В. М. Мануйлов

Курс лекций по линейной алгебре и геометрии

Механико-математический факультет
I-й курс, 2-й семестр

1 Линейные пространства

1.1 Линейное (векторное) пространство

Определение 1.1.1 Множество V называется *линейным (векторным) пространством* над некоторым полем $\mathbb{K}$, если заданы операция $+$ сложения двух элементов множества V и операция умножения $\cdot$ элементов множества V на элементы поля $\mathbb{K}$, которые удовлетворяют следующим условиям (аксиомам):

(i) $a + b = b + a \ \forall a, b \in V$,

(ii) $(a + b) + c = a + (b + c) \ \forall a, b, c \in V$,

(iii) $\exists 0 \in V : a + 0 = a \ \forall a \in V$,

(iv) $\forall a \in V \ \exists(-a) : a + (-a) = 0$,

(v) $\lambda \cdot (a + b) = \lambda \cdot a + \lambda \cdot b \ \forall a, b \in V, \forall \lambda \in \mathbb{K}$,

(vi) $(\lambda + \mu) \cdot a = \lambda \cdot a + \mu \cdot a \ \forall a \in V, \forall \lambda, \mu \in \mathbb{K}$,

(vii) $\lambda \cdot (\mu \cdot a) = (\lambda\mu) \cdot a \ \forall a \in V, \forall \lambda, \mu \in \mathbb{K}$,

(viii) $\mathbf{1} \cdot a = a \ \forall a \in V$.

Первые 4 свойства определяют на V структуру абелевой группы, а последние 4 свойства — структуру алгебры над полем $\mathbb{K}$.

Элементы множества V обычно называются векторами, а элементы поля $\mathbb{K}$ — скалярами или числами. Обычно мы будем опускать знак умножения $\cdot$.

Свойства линейного пространства:

(i) нулевой элемент в множестве V определен однозначно,

(ii) для любого элемента обратный элемент определен однозначно,

(iii) $\mathbf{0} \cdot a = 0 \ \forall a \in V$,

(iv) $\lambda \cdot 0 = 0 \ \forall \lambda \in \mathbb{K}$,

(v) $(-a) = (-\mathbf{1}) \cdot a \ \forall a \in V$,

(vi) если $\lambda \cdot a = 0$, то либо $\lambda = \mathbf{0}$, либо $a = 0$.

Доказательство.

(i) если $0'$ — другой нулевой элемент, то $0 = 0 + 0' = 0'$.

(ii) если $b + a = 0$, то $(-a) = (-a) + 0 = (-a) + b + a = (-a) + a + b = 0 + b = b$.

(iii) $0 = a + (-a) = \mathbf{1}a + (-a) = (\mathbf{1} + \mathbf{0})a + (-a) = \mathbf{1}a + \mathbf{0}a + (-a) = a + \mathbf{0}a + (-a) = a + (-a) + \mathbf{0}a = 0 + \mathbf{0}a = \mathbf{0}a$.

(iv) если $\lambda = \mathbf{0}$, то равенство $\mathbf{0} \cdot 0 = 0$ доказано в предыдущем пункте; если $\lambda \neq \mathbf{0}$, то $a + \lambda a = \lambda\lambda^{-1}a + \lambda 0 = \lambda(\lambda^{-1}a + 0) = \lambda(\lambda^{-1}a) = a$, и $\lambda 0 = 0$ следует из единственности нулевого элемента.

(v) $a + (-\mathbf{1})a = \mathbf{1}a + (-\mathbf{1})a = (\mathbf{1} + (-\mathbf{1}))a = \mathbf{0}a = 0$, поэтому $(-\mathbf{1})a = (-a)$ в силу единственности обратного элемента.

(vi) пусть $\lambda \cdot a = 0$; если $\lambda \neq 0$, то $a = \lambda^{-1}\lambda a = \lambda^{-1}0 = 0$.

□

Примеры линейных пространств:

(i) множество, состоящее из одного элемента $\{0\}$ является линейным пространством над любым полем,

(ii) множество векторов на прямой, на плоскости, в пространстве,

(iii) наборы из n чисел, $V = \{a_1, \ldots, a_n : a_i \in \mathbb{K}\}$, где сложение и умножение на скаляры определяется покомпонетно,

(iv) множество $\mathbb{K}_n[t]$ — множество многочленов степени не выше n с коэффициентами из поля $\mathbb{K}$ от переменной t,

(v) множество функций $F(x)$, определенных на некотором произвольном множестве X, со значениям в множестве $\mathbb{K}$,

(vi) множество решений однородной системы линейных уравнений,

(vii) $\mathbb{R}$ является линейным пространством над полем $\mathbb{Q}$,

(viii) $\mathbb{C}$ является линейным пространством над полем $\mathbb{R}$.

Определение 1.1.2 Пусть дано линейное пространство V. Линейной функцией (линейным функционалом) называют отображение $f : V \to \mathbb{K}$, обладающее свойствами: $f(a+b) = f(a)+f(b)$ и $f(\lambda a) = \lambda f(a)$ $\forall a, b \in V, \lambda \in \mathbb{K}$.

(ix) множество линейных функционалов является линейным пространством (оно называется двойственным пространством к V).

Определение 1.1.3 Пусть дано линейное пространство W, его непустое подмножество $V \subset W$ называется *подпространством*, если оно замкнуто относительно операций, определенных в пространстве W, т.е., если выполнены следующие свойства:
1) $a + b \in V$ $\forall a, b \in V$,
2) $\lambda a \in V$ $\forall a \in V, \lambda \in \mathbb{K}$.

Лемма 1.1.4 *Подпространство V линейного пространства W само является линейным пространством над тем же полем и с теми же операциями, что и W.*

Доказательство. Все условия определения линейного пространства выполнены, т.к. все элементы V являются элементами W, а для элементов W они выполнены по определению. □

Примеры подпространств. Пусть пространство W — это множество векторов на плоскости, тогда следующие множества будут подпространствами:
1) $\{0\}$,
2) множество всех векторов, коллинеарных некоторому заданному вектору,
3) само пространство W.

1.2 Аффинное пространство

Определение 1.2.1 Аффинным пространством называется тройка $(\mathcal{A}, V, +)$, состоящая из множества $\mathcal{A}$, векторного пространства V и операции сложения $+ : \mathcal{A} \times V \to \mathcal{A}$, (т.е. складывать можно элемент множества $\mathcal{A}$ с элементом векторного пространства, при этом в результате получается элемент множества $\mathcal{A}$), которая удовлетворяет следующим свойствам:
1) для любых $A, B \in \mathcal{A}$ существует единственный вектор $v \in V$, такой что $B = A + v$;
2) $A + 0 = A$ для любого $A \in \mathcal{A}$, где 0 — нулевой вектор;
3) $(A + v) + w = A + (v + w)$ для любых $A \in \mathcal{A}, v, w \in V$.

В обозначении аффинного пространства часто опускают знак плюс и пишут просто $(\mathcal{A}, V)$. Также, если из контекста понятно, какое пространство V имеется в виду, то и его не указывают и говорят об аффинном пространстве $\mathcal{A}$. Элементы аффинного пространства (или множества $\mathcal{A}$) называют точками. Любая пара точек $A, B \in \mathcal{A}$ однозначно определяет вектор v равенством $B = A + v$ (свойство 1) и такой вектор обозначается AB.

Примеры:
1) $\mathcal{A}$ — это обычная плоскость, V — двумерное векторное пространство векторов плоскости, $+$ — приложение вектора к точке.
2) Рассмотрим систему линейных уравнений $AX = B$, где A — матрица, X, B — столбцы. Пусть $\mathcal{A}$ — множество решений этой системы, V — множество решений соотвтетсвующей однородной системы $AX = 0$, $+$ — суммирование столбцов. Если $X_B \in \mathcal{A}$ и $X_0 \in V$, то $X_B + X_0 \in \mathcal{A}$.
3) Возьмем какое-нибудь векторное пространство V, в качестве $\mathcal{A}$ возьмем его же, $+$ — операция сложения в этом векторном пространстве.
Последний пример показывает, что имеется естественное соответствие между векторными и аффинными пространствами, и теория аффинных пространств полностью параллельна теории векторных пространств, поэтому в дальнейшем мы ограничимся только случаем векторных пространств, постоянно помня, что все результаты могут быть сформулированы в терминах точек и векторов аффинного пространства. Отождествляя $\mathcal{A}$ и V, мы будем иногда называть элементы векторного пространства точками.

1.3 Факторпространство

Определение 1.3.1 Пусть дано линейное пространство W, его элемент $a \in W$ и его подпространство $V \subset W$. *Линейным подмногообразием* называется множество всех векторов вида $a + v$, где $v \in V$.

Для линейных подмногообразий удобно пользоваться обозначением $a + V$.

Лемма 1.3.2 *Линейное подмногообразие $a+V$ является линейным подпространством пространства W тогда и только тогда, когда $a \in V$.*

Доказательство.
1) если $a \in V$, то $a + V = 0 + V = V$ (совпадают как множества).
2) пусть $a + V$ является подпространством. Т.к. $a \in a + V$, то $2a \in a + V$, что равносильно существованию некоторого $b \in V$, такого что $2a = a + b$, но тогда получаем, что $a = b$, т.е. $a \in V$. □

Лемма 1.3.3 $a_1 + V = a_2 + V \iff a_1 - a_2 \in V$.

Доказательство.
$\Rightarrow$: пусть $a_1 + V = a_2 + V$, тогда $a_1 \in a_1 + V = a_2 + V$, значит, найдется такой вектор $b \in V$, что $a_1 = a_2 + b$, т.е. $a_1 - a_2 \in V$.
$\Leftarrow$: пусть $a_1 - a_2 \in V$, т.е. $a_1 - a_2 = v \in V$. Возьмем произвольный элемент $b \in a_1 + V$. Тогда $b = a_1 + b'$ для какого-то вектора $b' \in V$. Покажем, что $b \in a_2 + V$. Действительно, $b = a_1 + b' = a_2 + (v + b')$, причем $v + b' \in V$. □

Определение 1.3.4 *Фактор-пространством* W/V линейного пространства W по, подпространству V называется множество $\{a + V : a \in W\}$ — множество всех линейных подмногообразий пространства W, заданных подпространством V, с определенными на нем операциями сложения, $(a + V) + (b + V) := (a + b) + V$, и умножения на скаляры, $\lambda(a + V) := \lambda a + V$.

Лемма 1.3.5 *Факторпространство линейного пространства само является линейным пространством.*

Доказательство.
1) Сначала докажем, что введенные нами в факторпространстве операции корректны, т.е. какие бы мы ни брали элементы подмногообразий в качестве a и b, мы получим одно и то же подмногообразие. Докажем это для сложения (для умножения доказывается аналогично): если $a_1 + V = a_2 + V$, $b_1 + V = b_2 + V$ то $a_1 - a_2 = v \in V$, $b_1 - b_2 = w \in V$, поэтому $(a_1 + b_1) + V = (a_2 + b_2 + v + w) + V = (a_2 + b_2) + (v + w) + V = (a_2 + b_2) + V$, т.е. сложение определено корректно.

2) Проверим свойства линейного пространства. Ввиду простоты ограничимся проверкой одного из восьми свойств, например 5).

$$\begin{aligned}\lambda((a+V)+(b+V)) &= \lambda((a+b)+V) = \lambda(a+b)+V = \\ &= (\lambda a+\lambda b)+V = (\lambda a+V)+(\lambda b+V) = \\ &= \lambda(a+V)+\lambda(b+V).\end{aligned}$$

□

1.4 Линейная зависимость векторов

Определение 1.4.1 Пусть дано линейное пространство V и некоторая система (множество) векторов $\{v_i : i \in I\} \subset V$ этого пространства. Если множество индексов I (а, значит, и система векторов) конечно ($I = \{1, \ldots, n\}$), их *линейной комбинацией* называется выражение вида $\lambda_1 v_1 + \ldots + \lambda_n v_n$, где λ_i — это числа (скаляры) из поля $\mathbb{K}$. Если множество I бесконечно, линейной комбинацией бесконечной системы векторов называется выражение аналогичного вида, $\sum_{i\in I} \lambda_i v_i$, в котором лишь *конечное* число скаляров λ_i отлично от нуля.

Определение 1.4.2 *Линейной оболочкой* системы векторов линейного пространства называется множество всех векторов, являющихся их линейной комбинацией.

Линейная оболочка системы векторов $e_1, \ldots, e_n$ часто обозначается $\langle e_1, \ldots, e_n\rangle$.

Определение 1.4.3 Система векторов $\{a_i : i \in I\}$ называется линейно зависимой, если существуют числа λ_i, не все равные нулю, такие, что $\sum_{k\in I} \lambda_i a_i = 0$, в противном случае система векторов называется линейно независимой.

Лемма 1.4.4 *Если система векторов $\{a_i : i \in I\}$ линейно зависима, то один из них является линейной комбинацией остальных.*

Доказательство. Пусть $\lambda_1 a_1 + \ldots + \lambda_k a_k = 0$, причем существует $\lambda_i \neq 0$, тогда имеем $\lambda_i a_i = -\lambda_1 a_1 - \ldots - \lambda_{i-1} a_{i-1} - \lambda_{i+1} a_{i+1} - \ldots \lambda_k a_k$, умножив обе части этого равенства на λ_i^{-1}, получим, что a_i есть линейная комбинация остальных векторов. □

Лемма 1.4.5 *Если система векторов $a_1, \ldots, a_n$ линейно независима, а система векторов $a_1, \ldots, a_n, a_{n+1}$ линейно зависима, то a_{n+1} является линейной комбинацией векторов $a_1, \ldots, a_n$.*

Доказательство. аналогично доказательству предыдущей леммы, с тем лишь замечанием, что если $\lambda_{n+1} = 0$, то ненулевой коэффициент λ_i находится среди первых n скаляров, но тогда первые n векторов линейно зависимы, что противоречит предположению. □

Лемма 1.4.6 *Пусть дана линейно независимая система векторов $e_1, \ldots, e_n$ и пусть существует линейно независимая система векторов $f_1, \ldots, f_m \in \langle e_1, \ldots, e_n\rangle$, тогда $m \leqslant n$.*

Доказательство. Пусть $f_i = a_{i1}e_1 + \ldots + a_{in}e_n$, $a_{ij} \in \mathbb{K}$, $i = 1, \ldots, m$. Т.к. $f_1, \ldots, f_m$ — линейно независимая система векторов, то

$$x_1 f_1 + \ldots + x_m f_m = 0 \Longleftrightarrow x_1 = \ldots = x_m = 0. \tag{1}$$

Подставляя в линейную комбинацию из (1) выражение f_i через $e_1, \ldots, e_n$, получаем:

$$\begin{aligned}0 &= x_1(a_{11}e_1 + \ldots a_{1n}e_n) + \ldots + x_m(a_{m1}e_1 + \ldots + a_{mn}e_n) = \\ &= (x_1a_{11} + \ldots + x_m a_{m1})e_1 + \ldots + (x_1a_{1n} + \ldots + x_m a_{mn})e_n,\end{aligned}$$

что равносильно (т.к. $e_1, \ldots, e_n$ — линейно независимая система векторов) системе уравнений:

$$\begin{cases} x_1a_{11} + \ldots + x_m a_{m1} = 0 \\ \qquad\qquad \ldots \\ x_1a_{1n} + \ldots + x_m a_{mn} = 0. \end{cases}$$

Если $m > n$, то эта система имеет ненулевое решение, что противоречит (1). □

1.5 Размерность

Определение 1.5.1 Определим *ранг* системы векторов: пусть S - непустая система векторов в некотором линейном пространстве V, тогда:
1) если S состоит только из $0 \in V$, то ранг $r(S) := 0$;
2) пусть e_1 — произвольный ненулевой вектор из системы S; если существует такой вектор e_2, что система $\{e_1, e_2\}$ будет линейно независимой, то рассмотрим эту систему векторов; если, далее, существует такой вектор e_3, что система $\{e_1, e_2, e_3\}$ будет линейно независимой, то будем рассматривать эту систему векторов, и т.д.;
3а) если процедура в п.2) закончится на конечном шаге, т.е. мы дойдем до линейно независимой системы векторов $\{e_1, \ldots, e_n\}$ и далее уже нельзя будет найти вектор e_{n+1}, чтобы расширить эту систему, то определим ранг как $r(S) := n$;
3б) если процедура в п.2) не закончится на конечном шаге, то ранг $r(S) := \infty$.

Докажем, что наше определение корректно. Сначала предположим, что, действуя, как в п.2), двумя способами, мы получили две конечные системы векторов $e_1, \ldots, e_n$ и $f_1, \ldots, f_m$, и пусть $m \neq n$. Тогда без ограничения общности можно считать, что $m > n$. Но, т.к. по определению к системе векторов $e_1, \ldots, e_n$ больше нельзя добавить ни одного вектора, то все $f_i \in \langle e_1, \ldots, e_n \rangle$, $i = 1, \ldots, m$, и по лемме 1.4.6 имеем $m \leqslant n$. Получили противоречие.
Теперь предположим, что один способ нам дал конечную систему векторов $e_1, \ldots, e_n$, а второй способ выбора векторов $f_1, f_2, \ldots$ не заканчивается ни на каком конечном шаге. Но тогда система векторов $f_1, \ldots, f_{n+1}$ линейно независима, и еще одно применение леммы 1.4.6 дает противоречие.

Определение 1.5.2 *Размерность* линейного пространства V равна $\dim V := r(V)$. Пространство V называется *конечномерным*, если $\dim V < \infty$. В противном случае пространство называется *бесконечномерным*.

Примеры бесконечномерных векторных пространств: множество $\mathbb{R}$ над полем $\mathbb{Q}$, пространство непрерывных функций на отрезке (докажите).

Определение 1.5.3 Система линейно независимых векторов в пространстве V называется *максимальной*, если при добавлении любого другого вектора система векторов становится линейно зависимой.

Следствие 1.5.4 *В любом конечномерном пространстве существует максимальная система векторов.*

Определение 1.5.5 Максимальная система векторов называется *базисом* пространства.

Бесконечномерные пространства мы почти не будем рассматривать в нашем курсе и все следующие определения, леммы и теоремы относятся к случаю конечномерных пространств. Полезным упражнением является проверка истинности таких утверждений в бесконечномерном случае (иногда сложно даже переформулировать конечномерные утверждения)

Лемма 1.5.6 *Пусть дано подпространство V некоторого векторного пространства W, и пусть $e_1, \ldots, e_r$ — базис в V. Тогда его можно дополнить до базиса всего пространства.*

Доказательство. Т.к. $e_1, \ldots, e_r$ — базис, то эти векторы линейно независимы; тогда, просто проделав процедуру п.2) в определении ранга системы векторов, мы получим базис всего пространства. □

Лемма 1.5.7 *Если V — подпространство векторного пространства W, то $\dim V \leqslant \dim W$. Если же $\dim V = \dim W$, то $V = W$.*

Доказательство. Из предыдущей леммы следует, что количество векторов в базисе подпространства не превышает количества векторов в базисе всего пространства, отсюда вытекает первое утверждение леммы. Докажем второе утверждение. Пусть $V \neq W$, т.е. существует вектор $w \in W$, $w \notin V$. Выберем базис $e_1, \ldots, e_r$ в V. Тогда система векторов $e_1, \ldots, e_r, w$ будет линейно независимой в W, что невозможно, т.к. $\dim W = r$. Действительно, если $\lambda_1 e_1 + \ldots + \lambda_r e_r + \lambda v = 0$ и хотя бы один из коэффициентов не равен нулю, то $\lambda \neq 0$ (противоречие с тем, что $e_1, \ldots, e_r$ — базис в V), но тогда вектор w есть линейная комбинация векторов $e_1, \ldots, e_r$, что противоречит предположению. □

Замечание: второе утверждение леммы неверно в бесконечномерном случае.

Теорема 1.5.8 *Пусть V — подпространство векторного пространства W, тогда*

$$\dim V + \dim W/V = \dim W.$$

Доказательство. Пусть $\dim V = r$, $\dim W = n$. Выберем произвольный базис $e_1, \ldots, e_r$ в V и дополним его до базиса $e_1, \ldots, e_r, e_{r+1}, \ldots, e_n$ в W. Докажем, что $e_{r+1} + V, \ldots, e_n + V$ будет базисом в W/V.

1) докажем линейную независимость этой системы векторов в W/V. Пусть они линейно зависимы, тогда существуют не все равные нулю $\lambda_{r+1}, \ldots, \lambda_n \in \mathbb{K}$, т.ч.

$$\lambda_{r+1}(e_{r+1} + V) + \ldots + \lambda_n(e_n + V) = 0 + V.$$

Раскрыв скобки, получим $(\lambda_{r+1}e_{r+1} + \ldots + \lambda_n e_n) + V = 0 + V$, поэтому существует такой вектор $v \in V$, что $v = \lambda_{r+1}e_{r+1} + \ldots + \lambda_n e_n$. Но, поскольку $v \in V$, его можно представить в виде $v = \lambda_1 e_1 + \ldots + \lambda_r e_r$. Тогда получаем, что $\lambda_1 e_1 + \ldots + \lambda_r e_r - \lambda_{r+1}e_{r+1} - \ldots - \lambda_n e_n = 0$, то есть, система векторов $e_1, \ldots, e_n$ линейно зависима, что неверно, так как это базис в W.

2) Докажем теперь, что любой вектор из W/V является линейной комбинацией векторов $e_{r+1} + V, \ldots, e_n + V$. Возьмем произвольный вектор $a + V$ из фактор-пространства W/V. Т.к. $a \in W$, то $a = \lambda_1 e_1 + \ldots + \lambda_r e_r + \lambda_{r+1}e_{r+1} + \ldots + \lambda_n e_n$. Тогда

$$\begin{aligned} a + V &= \underbrace{\lambda_1 e_1 + \ldots + \lambda_r e_r}_{\in V} + \lambda_{r+1}e_{r+1} + \ldots + \lambda_n e_n + V = \\ &= \lambda_{r+1}e_{r+1} + \ldots + \lambda_n e_n + V = \lambda_{r+1}(e_{r+1} + V) + \ldots + \lambda_n(e_n + V), \end{aligned}$$

Таким образом, $e_{r+1} + V, \ldots, e_n + V$ является базисом в W/V и, следовательно, $\dim W/V = n - r$. Отсюда вытекает утверждение теоремы. □

1.6 Пересечение и сумма подпространств

Лемма 1.6.1 *Пусть даны два линейных подпространства V_1 и V_2 пространства W, тогда $V_1 \cap V_2$ также является линейным подпространством.*

Доказательство. Для доказательство необходимо проверить, что множество $V_1 \cap V_2$ замкнуто относительно операций сложения и умножения на скаляры. Т.к. множества V_1 и V_2 замкнуты относительно этих операций, то $\forall x, y \in V_1 \cap V_2, \forall \lambda \in \mathbb{K}$ получаем, что $x + y, \lambda x \in V_1$ и $x + y, \lambda x \in V_1$, следовательно, $x + y, \lambda x \in V_1 \cap V_2$. □

Замечание. В отличие от пересечения, объединение подпространств $V_1 \cup V_2$ в общем случае не будет линейным подпространством. Например, если $V_1 = \langle\sqrt{2}\rangle$, а $V_2 = \langle\sqrt{3}\rangle$ над полем $\mathbb{Q}$, то вектор $\sqrt{2} + \sqrt{3}$ не будет принадлежать $V_1 \cup V_2$.

Определение 1.6.2 *Суммой* $V_1 + V_2$ подпространств V_1 и V_2 называется множество всех векторов $v \in W$, которые можно представить в виде суммы $v = v_1 + v_2$, где $v_1 \in V_1$ и $v_2 \in V_2$, т.е. $V_1 + V_2 = \langle V_1 \cup V_2 \rangle$.

Лемма 1.6.3 *Для любых двух подпространств V_1 и V_2 их сумма $V_1 + V_2$ также будет линейным пространством.*

Доказательство. Возьмем произвольные векторы $a, b \in V_1 + V_2$, $a = a_1 + a_2$, $b = b_1 + b_2$, $a_1, b_1 \in V_1, a_2, b_2 \in V_2$. Тогда $a+b = (a_1+b_1)+(a_2+b_2) \in V_1+V_2$. Аналогично доказывается, что для $\lambda \in \mathbb{K}$, $a \in V_1+V_2$, их произведение $\lambda a \in V_1+V_2$. Очевидно, что все условия определения линейного пространства будут выполнены, следовательно $V_1 + V_2$ является линейным пространством. □

Теорема 1.6.4 $\dim V_1 + \dim V_2 = \dim(V_1 + V_2) + \dim(V_1 \cap V_2)$.

Доказательство. Пусть $e_1, \ldots, e_r$ — базис в $V_1 \cap V_2$, $\dim(V_1 \cap V_2) = r$. Т.к. $V_1 \cap V_2 \subset V_1$ и $V_1 \cap V_2 \subset V_2$, то этот базис можно дополнить до базисов в V_1 и V_2.

Пусть $e_1, \ldots, e_r, e_{r+1}, \ldots, e_{r+p}$ — базис в V_1, $\dim V_1 = r + p$; $e_1, \ldots, e_r, e_{r+p+1}, \ldots, e_{r+p+q}$ — базис в V_2, $\dim V_1 = r + q$. Докажем, что $e_1, \ldots, e_r, e_{r+1}, \ldots, e_{r+p}, e_{r+p+1}, \ldots, e_{r+p+q}$ — базис в $V_1 + V_2$:

1) (линейная независимость). Пусть $\lambda_1 e_1 + \ldots + \lambda_{r+p+q} e_{r+p+q} = 0$, тогда

$$\underbrace{\lambda_1 e_1 + \ldots + \lambda_r e_r + \lambda_{r+1} e_{r+1} + \ldots + \lambda_{r+p} e_{r+p}}_{\in V_1} =$$
$$= \underbrace{-(\lambda_{r+p+1} e_{r+p+1} + \ldots + \lambda_{r+p+q} e_{r+p+q})}_{\in V_2} = v,$$

следовательно, $v \in V_1 \cap V_2$, и его можно разложить по базису, $v = \alpha_1 + \ldots + \alpha_r e_r$, тогда

$$0 = v - v = \alpha_1 + \ldots + \alpha_r e_r + \lambda_{r+p+1} e_{r+p+1} + \ldots + \lambda_{r+p+q} e_{r+p+q},$$

следовательно $\lambda_{r+p+1} = \ldots = \lambda_{r+p+q} = 0$ и $\alpha_1 = \ldots = \alpha_r = 0$, т.к. $e_1, \ldots, e_r, e_{r_p+1}, \ldots, e_{r+p+q}$ линейно независимы. Поэтому $v = 0$. Но тогда $v = \lambda_1 e_1 + \ldots + \lambda_r e_r + \lambda_{r+1} e_{r+1} + \ldots + \lambda_{r+p} e_{r+p} = 0$, и из линейной независимости системы векторов $e_1, \ldots, e_r, e_{r+1}, \ldots, e_{r+p}$ заключаем, что $\lambda_1 = \ldots = \lambda_r = \lambda_{r+1} = \ldots = \lambda_{r+p} = 0$. Итак, все $\lambda_i = 0$, $i = 1, \ldots, r + p + q$, следовательно $e_1, \ldots, e_r, e_{r+1}, \ldots, e_{r+p}, e_{r+p+1}, \ldots, e_{r+p+q}$ линейно независимы.

2) (максимальность). Возьмем произвольный вектор $a \in V_1 + V_2$, $a = a_1 + a_2$, где $a_1 \in V_1$, $a_2 \in V_2$. Разложим векторы a_1 и a_2 по базисам,

$$\begin{aligned} a_1 &= \lambda_1 e_1 + \ldots + \lambda_r e_r + \lambda_{r+1} e_{r+1} + \ldots + \lambda_{r+p} e_{r+p}, \\ a_2 &= \mu_1 e_1 + \ldots + \mu_r e_r + \mu_{r+p+1} e_{r+p+1} + \ldots + \mu_{r+p+q} e_{r+p+q}, \end{aligned}$$

тогда

$$\begin{aligned} a &= (\lambda_1 + \mu_1) e_1 + \ldots + (\lambda_r + \mu_r) e_r + \lambda_{r+1} e_{r+1} + \ldots + \lambda_{r+p} e_{r+p} + \\ &+ \mu_{r+p+1} e_{r+p+1} + \ldots + \mu_{r+p+q} e_{r+p+q}, \end{aligned}$$

следовательно система векторов $e_1, \ldots, e_r, e_{r+1}, \ldots, e_{r+p}, e_{r+p+1}, \ldots, e_{r+p+q}$ является базисом в $M_1 + M_2$, значит, $\dim(V_1 + V_2) = r + p + q$, откуда следует утверждение теоремы. □

1.7 Прямая сумма подпространств. Внешняя прямая сумма

Определение 1.7.1 Сумма подпространств V_1+V_2 называется *прямой суммой* (обозначение $V_1 \oplus V_2$), если $V_1 \cap V_2 = \{0\}$.

Следствие 1.7.2 $\dim(V_1 \oplus V_2) = \dim V_1 + \dim V_2$.

Доказательство. Утверждение следствия очевидно вытекает из предыдущей теоремы. □

Лемма 1.7.3 *Следующие утверждения эквивалентны:*

(i) *сумма $V_1 + V_2$ прямая;*

(ii) $\dim V_1 + \dim V_2 = \dim(V_1 + V_2)$;

(iii) *разложение любого вектора a вида $a = v_1 + v_2$, где $v_1 \in V_1$, $v_2 \in V_2$, единственно;*

(iv) *если $0 = v_1 + v_2$, где $v_1 \in V_1$, $v_2 \in V_2$, то $v_1 = v_2 = 0$.*

Доказательство. То, что 1) $\iff$ 2), вытекает из последнего следствия предыдущей лекции.

Докажем, что 1) $\Rightarrow$ 4). Пусть $0 = v_1 + v_2$, $v_1 \in V_1$, $v_2 \in V_2$, но $v_1 \neq 0$, а следовательно и $v_2 \neq 0$, тогда получаем, что $v_2 = -v_1$, т.е. $v_2 \in V_1$ и , следовательно $V_1 \cap V_2 \ni v_2 \neq 0$, т.е. сумма не прямая.

1) $\Leftarrow$ 4) доказывается аналогично. Если сумма не прямая, то $\exists v \in V_1 \cap V_2$, $v \neq 0$, тогда $v \in V_1$, $-v \in V_2$ и $0 = v + (-v)$.

Докажем 4) $\Rightarrow$ 3). Пусть у некоторого вектора a есть два разложения, $a = v_1 + v_2 = v_1' + v_2'$, $v_1, v_1' \in V_1$, $v_2, v_2' \in V_2$, тогда $0 = \underbrace{(v_1 = v_1')}_{\in V_1} + \underbrace{(v_2 - v_2')}_{\in V_2}$. Но тогда $v_1 - v_1' = v_2 - v_2' = 0$.

То, что 4) $\Leftarrow$ 3) очевидно, т.к. разложение любого вектора единственно, то и разложение нулевого вектора единственно. $\square$

Понятие прямой суммы можно обобщить на любое конечное число подпространств: сумма $V_1 + \ldots + V_n$ будет прямой, если

$$\forall i = 1, \ldots, n \quad V_i \cap (V_1 + \ldots + V_{i-1} + V_{i+1} + \ldots + V_n) = \{0\}. \tag{2}$$

Если сумма $V_1 + \ldots + V_n$ прямая, то для любого вектора a из этой суммы разложение вида $a = v_1 + \ldots + v_n$, где $v_i \in V_i$, $i = 1, \ldots, n$, единственно.

Замечание. Условие (2) более сильное, чем условие $V_i \cap V_j = \{0\}$ $\forall i, j = 1, \ldots, n$. Например, если взять три прямые (вектора, коллинеарные этим прямым), пересекающиеся в одной точке, то сумма любых двух из них будет прямой суммой, но сумма всех трех — нет, т.к. любой вектор третьей прямой можно представить в виде суммы векторов первых двух прямых, следовательно его разложение не будет единственно.

Внешняя прямая сумма

Определение 1.7.4 *Внешней прямой суммой* двух линейных пространств V_1, V_2 над одним полем $\mathbb{K}$ (не обязательно являющихся подпространствами одного пространства) называется новое линейное пространство $V_1 \oplus V_2$ над полем $\mathbb{K}$, состоящее из всех пар (v_1, v_2), где $v_i \in V_i$, $i = 1, 2$, с операциями сложения и умножения на скаляры:
1) $(v_1, v_2) + (v_1', v_2') = (v_1 + v_1', v_2 + v_2')$,
2) $\lambda(v_1, v_2) = (\lambda v_1, \lambda v_2)$,
где $v_i, v_i' \in V_i$, $\lambda \in \mathbb{K}$.

Если при этом отождествить сами пространства V_1 и V_2 с подмножествами внешней прямой суммы следующим образом: $V_1 \leftrightarrow (V_1, 0)$ и $V_2 \leftrightarrow (0, V_2)$, то их можно рассматривать как подпространства пространства $V_1 \oplus V_2$.

1.8 Координаты

Определение 1.8.1 Пусть дано линейное пространство V и базис $e_1, \ldots, e_n$ этого пространства, тогда любой вектор $x \in V$ можно представить в виде $x = \lambda_1 e_1 + \ldots + \lambda_n e_n$. Числа $\lambda_1, \ldots, \lambda_n \in K$ называются *координатами* вектора x в этом базисе.

Введем некоторые соглашения для записи координат. Индексы у координат мы обычно будем писать не снизу, а сверху, т.е. не x_i, а x^i. Вместо длинной записи $x = x^1 e_1 + \ldots + x^n e_n$ мы часто будем писать $x^i e_i$, на самом деле подразумевая сумму $\sum_{i=1}^n x^i e_i$. Координаты векторов мы часто будем записывать в виде столбцов, т.е. $x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$.

Корректность определения координат следует из свойств базиса (линейная независимость и максимальность).

Замена координат

Пусть нам даны два базиса $e_1, \ldots, e_n$ и $\widetilde{e}_1, \ldots, \widetilde{e}_n$ одного векторного пространства, тогда можно записать следующие равенства:

$$\begin{array}{rcl} \widetilde{e}_1 & = & c_1^1 e_1 + \ldots + c_1^n e_n, \\ \ldots & & \ldots \quad \ldots \\ \widetilde{e}_n & = & c_n^1 e_1 + \ldots + c_n^n e_n, \end{array}$$

которые равносильны одному матричному равенству

$$(\widetilde{e}_1 \ldots \widetilde{e}_n) = (e_1 \ldots e_n) \begin{pmatrix} c_1^1 & \ldots & c_n^1 \\ \vdots & \ddots & \vdots \\ c_1^n & \ldots & c_n^n \end{pmatrix}.$$

Матрица

$$C = \begin{pmatrix} c_1^1 & \ldots & c_n^1 \\ \vdots & \ddots & \vdots \\ c_1^n & \ldots & c_n^n \end{pmatrix}$$

называется *матрицей перехода* от базиса $e_1, \ldots, e_n$ к базису $\widetilde{e}_1, \ldots, \widetilde{e}_n$.

Лемма 1.8.2 *Пусть $x^1, \ldots, x^n$ — координаты вектора x в базисе $e_1, \ldots, e_n$, а $\widetilde{x}^1, \ldots, \widetilde{x}^n$ — координаты этого же вектора в базисе $\widetilde{e}_1, \ldots, \widetilde{e}_n$. Тогда*

$$\begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = C \begin{pmatrix} \widetilde{x}^1 \\ \vdots \\ \widetilde{x}^n \end{pmatrix}.$$

Доказательство. Так как $x^j e_j = x = \widetilde{x}^i \widetilde{e}_i = \widetilde{x}^i e_j c_i^j = (\widetilde{x}^i c_i^j) e_j$, из линейной независимости векторов $e_1, \ldots, e_n$ следует равенство координат: $x^j = \widetilde{x}^i c_i^j \ \forall j$ (подразумевается суммирование по индексу i). □

1.9 Изоморфизмы векторных пространств

Определение 1.9.1 Пусть даны два линейных пространства V и W над одним полем $\mathbb{K}$. Тогда биективное (т.е. взаимно однозначное) отображение $f : V \to W$ называется *изоморфизмом*, если выполнены следующие условия (условия линейности):
1) $f(v_1 + v_2) = f(v_1) + f(v_2) \ \forall v_1, v_2 \in V$,
2) $f(\lambda v) = \lambda f(v) \ \forall v \in V, \lambda \in \mathbb{K}$.
Два линейных пространства V и W называются изоморфными ($V \cong W$), если между ними существует изоморфизм.

Лемма 1.9.2 *Если $f : V \to W$ — изоморфизм, то обратное отображение $f^{-1} : W \to V$ также будет изоморфизмом.*

Доказательство. поскольку отображение f взаимно однозначно. Докажем только первый пункт, т.е., что $f^{-1}(w_1 + w_2) = f^{-1}(w_1) + f^{-1}(w_2) \ \forall w_1, w_2 \in W$ (второй пункт доказывается аналогично):

$$f(f^{-1}(w_1) + f^{-1}(w_2)) = f(f^{-1}(w_1)) + f(f^{-1}(w_2)) = w_1 + w_2 = f(f^{-1}(w_1 + w_2)),$$

следовательно $f^{-1}(w_1) + f^{-1}(w_2) = f^{-1}(w_1 + w_2)$, т.к. отображение f взаимно однозначно. □

Таким образом, изоморфность является отношением эквивалентности, т.е. обладает свойствами симметричности (любое пространство изоморфно самому себе), рефлексивности (если V изоморфно W, то W изоморфно V) и транзитивности (если V изоморфно W и W изоморфно U, то V изоморфно U).

Лемма 1.9.3 *Если* $\dim V = n$, *то* V *изоморфно пространству* $\mathbb{K}^n$ *столбцов (строк) из* n *элементов.*

Доказательство. Пусть $e_1, \ldots, e_n$ — базис в V, тогда построим отображение $f : V \to \mathbb{K}^n$ следующим образом: если $x = x^i e_i$, то $f(x) = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$. Легко проверить, что это отображение будет изоморфизмом, а следовательно $V \cong \mathbb{K}^n$. □

Следствие 1.9.4 *Если* $\dim V = \dim W$, *то* $V \cong W$.

Доказательство. Пусть $\dim V = \dim W = n$, тогда $V \cong \mathbb{K}^n \cong W$. □

Верно и обратное:

Лемма 1.9.5 *Если* $V \cong W$, *то* $\dim V = \dim W$.

Доказательство. Допустим, что $\dim V < \dim W$, пусть $e_1, \ldots, e_n$ — базис в W, тогда вектора $f(e_1), \ldots, f(e_n) \in V$ должны быть линейно независимыми. Действительно, если $\lambda_1 f(e_1) + \ldots + \lambda_n f(e_n) = 0$, то, применив к обеим частям этого равенства отображение f^{-1}, получим $\lambda_1 e_1 + \ldots + \lambda_n e_n = 0$, откуда $\lambda_1 = \ldots = \lambda_n = 0$. Но их линейная независимость противоречит предположению $\dim V < n$. □

Лемма 1.9.6 *Пусть* $\dim V = \dim W$, *а отображение* $f : V \to W$ *удовлетворяет условиям линейности:*
1) $f(v_1 + v_2) = f(v_1) + f(v_2)\ \forall v_1, v_2 \in V$,
2) $f(\lambda v) = \lambda f(v)\ \forall v \in V, \lambda \in \mathbb{K}$.
Пусть $e_1, \ldots, e_n$ — *базис в* V. *Тогда* f *является изоморфизмом тогда и только тогда, когда* $f(e_1), \ldots, f(e_n)$ — *базис в* W.

Доказательство. Если f — изоморфизм, то из равенства $\lambda_1 f(e_1) + \ldots + \lambda_n f(e_n) = 0$ следует $f(\lambda_1 e_1 + \ldots + \lambda_n e_n) = 0$, откуда заключаем, что $\lambda_1 e_1 + \ldots + \lambda_n e_n = 0$, значит, все $\lambda_i = 0$.
Обратно, пусть $f(e_1, \ldots, f(e_n))$ — базис в W. Для проверки взаимной однозначности отображения f достаточно проверить, что отображение f^{-1} корректно определено. Пусть $w \in W$ имеет разложение по базису $w = w^1 f(e_1) + \ldots + w^n f(e_n)$. Тогда определим отображение $g : W \to V$ равенством $g(w) = w^1 e_1 + \ldots + w^n e_n$. Очевидная проверка показывает, что $g = f^{-1}$. □

1.10 Двойственное векторное пространство

Напомним, что линейным функционалом на векторном пространстве V (над полем $\mathbb{K}$) называется отображение $f : V \to \mathbb{K}$, удовлетворяющее условиям линейности
1) $f(v_1 + v_2) = f(v_1) + f(v_2)\ \forall v_1, v_2 \in V$;
2) $f(\lambda v) = \lambda f(v)\ \forall v \in V, \lambda \in \mathbb{K}$.
Зададим на множестве V' всех линейных функционалов $f : V \to \mathbb{K}$, операции:
1) $(f_1 + f_2)(v) = f_1(v) + f_2(v)$, $f_1, f_2 \in V'$;
2) $(\lambda f)(v) = \lambda f(v)$, $f \in V'$, $\lambda \in \mathbb{K}$.
Эти операции превращают V' в линейное пространство. Это пространство называется *двойственным пространством* к V.

Лемма 1.10.1 $V \cong V'$.

Доказательство. Пусть $e_1, \ldots, e_n$ — базис в V, тогда докажем, что $\varepsilon^1, \ldots, \varepsilon^n$ будет базисом в V', где функционалы $\varepsilon^i \in V'$, $i = 1, \ldots, n$, определяются равенствами $\varepsilon^i(e_j) = \delta^i_j$ (δ^i_j — символ Кронекера, т.е. $\delta^i_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$. Поскольку $f(x^i e_i) = x^i f(e_i)$, то значение функционала на произвольном векторе полностью определяется значениями функционала на базисных векторах

и координатами этого вектора, т.е. функционалы ε^i полностью заданы нашими условиями. Нам нужно доказать два пункта:

1) линейная независимость векторов $\varepsilon^1, \dots, \varepsilon^n$. Если $f = \lambda_1\varepsilon^1 + \dots + \lambda_n\varepsilon^n = 0$ (равенство нулю в V' означает, что $f(v) = 0$ для любого вектора $v \in V$), то $f(e_i) = \lambda_i = 0$, т.е. все $\lambda_i = 0$. Следовательно эти функционалы линейно независимы.

2) максимальность, т.е. что $\forall f \in V', \exists \lambda_1, \dots, \lambda_n \in \mathbb{K}$ такие что $f = \lambda_i\varepsilon^i$. Возьмем произвольный функционал $f \in V'$, тогда, если $x = x^ie_i$, то $f(x) = x^if(e_i)$. Возьмем $\lambda_i = f(e_i)$, тогда получим, что

$$f(x) = x^if(e_i) = x^i\lambda_i = \lambda_ix^j\varepsilon^i(e_j) = \lambda_i\varepsilon^i(x^je_j) = \lambda_i\varepsilon^i(x),$$

что и требовалось доказать. □

Пусть нам даны базисы $e_1, \dots, e_n$ и $\widetilde{e}_1, \dots, \widetilde{e}_n$ в пространстве V и двойственные к ним базисы $\varepsilon^1, \dots, \varepsilon^n$ и $\widetilde{\varepsilon}^1, \dots, \widetilde{\varepsilon}^n$ в пространстве V'. Пусть C — матрица перехода от базиса $e_1, \dots, e_n$ к $\widetilde{e}_1, \dots, \widetilde{e}_n$, найдем матрицу перехода от базиса $\varepsilon^1, \dots, \varepsilon^n$ к $\widetilde{\varepsilon}^1, \dots, \widetilde{\varepsilon}^n$. Возьмем произвольный функционал $f \in V'$, тогда $f = f_i\varepsilon^i = \widetilde{f}_j\widetilde{\varepsilon}^j$, где f_i и $\widetilde{f}_j$ — это координаты функционала f в базисах $\varepsilon^1, \dots, \varepsilon^n$ и $\widetilde{\varepsilon}^1, \dots, \widetilde{\varepsilon}^n$ соответственно. Вычислим значение функционала f на векторе $\widetilde{e}_k$ двумя способами. С одной стороны, $f(\widetilde{e}_k) = \widetilde{f}_j\widetilde{\varepsilon}^j(\widetilde{e}_k) = \widetilde{f}_k$, а с другой стороны, $f(\widetilde{e}_k) = f_i\varepsilon^i(c^j_ke_j) = f_ic^i_k$, так как $(\widetilde{e}_1 \dots \widetilde{e}_n) = (e_1 \dots e_n)C$. Отсюда получаем, что $\widetilde{f}_k = f_ic^i_k$. Следовательно $(\widetilde{f}_1 \dots \widetilde{f}_n) = (f_1 \dots f_n)C$, или (после транспонирования, обозначенного индексом t) $\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} = (C^{-1})^t \begin{pmatrix} \widetilde{f}_1 \\ \vdots \\ \widetilde{f}_n \end{pmatrix}$, и, наконец, отсюда получаем связь между базисами: $(\widetilde{\varepsilon}^1 \dots \widetilde{\varepsilon}^n) = (\varepsilon^1 \dots \varepsilon^n)(C^{-1})^t$, т.е. матрица перехода от базиса ε к $\widetilde{\varepsilon}$ равна $(C^{-1})^t$.

Мы доказали, что $V \cong V'$, однако выбор изоморфизма $f : V \to V'$ зависит от выбора базиса в пространстве V. Действительно, пусть $V = \mathbb{R}$, тогда базисом является любое ненулевое число $e \in \mathbb{R}$. Выберем также еще один базис $\widetilde{e} = \lambda e$, $\lambda \neq 0, 1$. Пусть ε — базис в V', двойственный к e, т.е. $\varepsilon(e) = 1$, а $\widetilde{\varepsilon}$ — двойственный базис к $\widetilde{e}$, т.к. $\widetilde{\varepsilon}(\widetilde{e}) = 1$, то $\widetilde{\varepsilon} = \lambda^{-1}e$. Изоморфизм $f : V \to V'$, отвечающий базису e, задан следующим образом: $\forall x = \alpha e$, $f(x) = \alpha\varepsilon$. Перейдем к базису $\widetilde{e}$, тогда изоморфизм $\widetilde{f} : V \to V'$ будет задаваться следующим образом: $\widetilde{f}(x) = \widetilde{f}(\alpha\lambda^{-1}\widetilde{e}) = \alpha\lambda^{-2}\varepsilon$. Т.е. разные выборы базиса в пространстве V дают разные изоморфизмы!

1.11 Канонический изоморфизм между пространством и его вторым двойственным

Рассмотрим пространство, двойственное к двойственному. Оно называется *вторым двойственным* пространством: $V'' = (V')'$. Элементы пространства V'' — это линейные функционалы на пространстве V', т.е. функции, аргументами которых являются элементы множества V'. Очевидно, что $\dim V'' = \dim V' = \dim V$.

Определим отображение $\varphi : V \to V''$. Для каждого вектора $x \in V$ функционал $\varphi(x) = \varphi_x$ должен отображать функционалы (элементы множества V') в поле скаляров. Пусть $f \in V'$ Определим значение $\varphi_x(f) = f(x)$. Очевидно, что φ_x — это линейное отображение $V' \to \mathbb{K}$. Кроме того, $\varphi(x_1 + x_2) = \varphi(x_1) + \varphi(x_2)$ и $\varphi(\lambda x) = \lambda\varphi(x)$. Докажем, что φ есть изоморфизм. Для этого нам нужно доказать, что из условия $\varphi_x = 0$ следует, что $x = 0$. Условие $\varphi_x = 0$ означает, что для любого функцинала $f \in V'$ $\varphi_x(f) = 0$, т.е. $f(x) = 0$. Но единственный вектор в V, на котором любой функционал равен нулю, есть нулевой вектор, $x = 0$. Действительно, если это неверно, выберем базис $e_1 = x, e_2, \dots, e_n$ в V, тогда для функционала ε^1 имеем $\varepsilon^1(x) = 1 \neq 0$.

Так как $\varphi_x = 0 \iff x = 0$, то для произвольного базиса $e_1, \dots, e_n$ в V векторы $\varphi_{e_1}, \dots, \varphi_{e_n}$ линейно независимы. Но, поскольку $\dim V'' = n$, эти векторы составляют базис пространства V'', следовательно φ — изоморфизм.

При построении этого изоморфизма, мы ни разу не использовали базис (базис мы использовали только при доказательстве), поэтому этот изоморфизм не зависит от выбора базиса в пространстве V, а его конструкция универсальна и годится для любого пространства V! Такие изоморфизмы называются *каноническими*.

Т.к. V и V'' изоморфны *канонически*, то мы можем эти два пространства просто отождествить, и смотреть на пространства V и V' как на двойственные друг к другу (V' — двойственное к V, а $V = V''$ — двойственное к V').

Легко видеть, что базис $\varphi_{e_1}, \ldots, \varphi_{e_n}$ в V'' — двойственный к базису $\varepsilon^1, \ldots, \varepsilon^n$, если $\varepsilon^1, \ldots, \varepsilon^n$ — двойственный базис к $e_1, \ldots, e_n$. При отождествлении пространств V и V'', базисы $\varphi_{e_1}, \ldots, \varphi_{e_n}$ и $e_1, \ldots, e_n$ отождествятся, и тогда базисы $e_1, \ldots, e_n$ и $\varepsilon^1, \ldots, \varepsilon^n$ будут взаимно двойственными.

Пример: двойственное пространство пространства многочленов

Рассмотрим пространство $\mathbb{K}_n[x]$ многочленов степени не выше n с коэффициентами из поля $\mathbb{K}$ от переменной $x \in \mathbb{K}$. Зафиксируем произвольное $x = x_0$, и каждому многочлену $p(x)$ поставим в соответствие число $p(x) \mapsto p(x_0) \in \mathbb{K}$. Каждое x_0 задает свое отображение $ev_{x_0} : \mathbb{K}_n[x] \to \mathbb{K}$. Т.к.

$$ev_{x_0}(p(x) + q(x)) = p(x_0) + q(x_0) = ev_{x_0}(p(x)) + ev_{x_0}(q(x))$$

и $ev_{x_0}(\lambda p(x)) = \lambda ev_{x_0}(p)x$, то отображение ev_{x_0} линейно для каждого x_0. Таким образом, каждое значение x_0 задает элемент ev_{x_0} двойственного пространства $\mathbb{K}_n[x]'$.

Лемма 1.11.1 *Если $x_0, x_1, \ldots, x_n$ — попарно различные значения, то $ev_{x_0}, ev_{x_1}, \ldots, ev_{x_n}$ будет базисом в двойственном пространстве $\mathbb{K}_n[x]'$.*

Доказательство. Если нам удастся построить базис в пространстве $\mathbb{K}_n[x]$, который будет двойственным к $ev_{x_0}, ev_{x_1}, \ldots, ev_{x_n}$, то отсюда будет следовать, что $ev_{x_0}, ev_{x_1}, \ldots, ev_{x_n}$ будет двойственным к базису в $\mathbb{K}_n[x]$, т.е. будет базисом в $\mathbb{K}_n[x]'$. Построим такой базис:

Нам нужно найти такие многочлены $p^0(x), p^1(x), \ldots, p^n(x)$, что $ev_{x_i}(p^j(x)) = \delta_i^j$, т.е. значение i-й функции ev_{x_i} на всех базисных многочленах, кроме $p^i(x)$, равно 0, а на $p^i(x)$ равно 1. Эти многочлены можно построить, используя, например, интерполяционную формулу Лагранжа:

$$p^i(x) = \frac{(x - x_0)\ldots(x - x_{i-1})(x - x_{i+1})\ldots(x - x_n)}{(x_i - x_0)\ldots(x_i - x_{i-1})(x_i - x_{i+1})\ldots(x_i - x_n)}.$$

Докажем, что эти многочлены образуют базис.

1) линейная независимость: $p(x) = \lambda_i p^i(x) = 0$ только, если все $\lambda_i = 0$, т.к. $p(x_i) = \lambda_i \ \forall i$.

2) максимальность: возьмем произвольный многочлен $p(x)$, тогда $p(x) = p(x_i)p^i(x)$, т.е. является линейной комбинацией многочленов $p^i(x)$.

(здесь, согласно тензорным обозначениям, подразумевается суммирование по индексу i).

Таким образом, мы доказали, что $p^0(x), p^1(x), \ldots, p^n(x)$ — базис в $\mathbb{K}_n[x]$, а значит двойственный к нему базис $ev_{x_0}, ev_{x_1}, \ldots, ev_{x_n}$ будет базисом в $\mathbb{K}_n[x]'$. □

2 Линейные операторы

2.1 Линейные отображения

Определение 2.1.1 Пусть V, W — два векторных пространства над одним полем $\mathbb{K}$. Отображение $f : V \to W$ называется линейным, если $\forall x, y \in V$, $\lambda \in \mathbb{K}$ выполняются равенства $f(x+y) = f(x) + f(y)$ и $f(\lambda x) = \lambda f(x)$.

Пример: множество V' — это множество линейных отображений при $W = \mathbb{K}$.

Пусть $e_1, \ldots, e_n$ — базис в V, а $\widetilde{e}_1, \ldots, \widetilde{e}_m$ — базис в W. Если $x = x^i e_i \in V$, то $f(x) = f(x^i e_i) = x^i f(e_i)$, т.е., для вычисления значения функции в любой точке, достаточно знать ее значения на базисных векторах, т.е. $f(e_i) = a_i^k \widetilde{e}_k$ (a_i^k — коэффициенты разложения вектора $f(e_i)$ по базису $\widetilde{e}$), тогда $f(x) = x^i a_i^k \widetilde{e}_k = y^k \widetilde{e}_k$ — разложение значения по базисным векторам $\widetilde{e}_1, \ldots, \widetilde{e}_k$. Координаты x^i вектора x в базисе пространства V и координаты y^k значения отображения $f(x)$ в базисе пространства W связаны следующим соотношением:

$$\begin{pmatrix} y^1 \\ \vdots \\ y^m \end{pmatrix} = \begin{pmatrix} a_1^1 & \ldots & a_n^1 \\ \vdots & \ddots & \vdots \\ a_1^m & \ldots & a_n^m \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix},$$

или, в матричной форме, $Y = AX$, где Y и X — столбцы координат векторов $f(x)$ и x соответственно, а матрица $A_f = A = (a_i^k)$ является матрицей, определяемой линейным отображением f (и определяющей его).

Мы видим, что задание базисов в V и W позволяет сопоставить каждому линейному отображению f его матрицу A_f, причем это сопоставление взаимо однозначно. Поэтому существует биективное отображение между множеством линейных отображений $L(V, W)$ из V в W и множеством матриц $M_{m,n}(\mathbb{K})$ с коэффициентами из поля $\mathbb{K}$ размера $m \times n$.

Лемма 2.1.2 $L(V, W) \cong M_{m,n}(\mathbb{K})$.

Доказательство. Достаточно проверить, что построенное выше биективное отображение $L(V, W) \to M_{m,n}(\mathbb{K})$ будет линейным. Но это следует из того, что все отображения из $L(V, W)$ линейны. □

Примеры:

1) Рассмотрим отображение $f(x) \equiv 0$, ему будет соответствовать нулевая матрица $A_f = 0$;

2) Если $W = V$, а отображение тождественно, $f = id : V \to V$, т.е. $f(x) = x$ $\forall x \in V$, то ему соответствует единичная матрица $A_f = E_n$;

3) Отображению $f(x) = \lambda x$ соответствует матрица $A_f = \lambda E_n$.

Еще раз отметим, что соответствие $f \mapsto A_f$ зависит от выбора базисов в пространствах V и W.

Изменим базисы в V (матрица перехода C_1) и в W (матрица перехода C_2), тогда, естественно, изменится и матрица данного линейного отображения. Если в первоначальных базисах координаты были связаны матричным соотношением $Y = A_f X$, то в новых базисах ($X = C_1 X'$, $Y = C_2 Y'$) имеем $C_2 Y' = A_f C_1 X'$, т.е. $Y' = C_2^{-1} A_f C_1 X' = A'_f X'$. Окончательно получаем формулу для матрицы оператора в новых базисах $A'_f = C_2^{-1} A_f C_1$.

Если $W = V$, то мы получим отображение пространства в себя. Такие отображения называются *линейными операторами*. Матрица линейного оператора всегда квадратная, при этом в обоих экземплярах пространства V берется один и тот же базис. Тогда при переходе к другому базису матрица линейного оператора изменяется следующим образом: $A'_f = C^{-1} A_f C$, где C — матрица перехода, а A_f — матрица оператора в старом базисе.

Определение 2.1.3 Определим $\det f$ равенством $\det f = \det A_f$.

Чтобы определение было корректным, надо, чтобы эта величина не зависела от выбора базиса в пространстве, т.е. возьмем два разным базиса с матрицей перехода C, тогда

$$\det A'_f = \det(C^{-1} A_f C) = \det C^{-1} \det A_f \det C = \det A_f.$$

Определение 2.1.4 Определим след $\operatorname{tr} f$ линейного оператора равенством $\operatorname{tr} f = \operatorname{tr} A_f$ (сумма диагональных элементов матрицы A_f).

Аналогично проверяем, что определение корректно:

$$\operatorname{tr} A'_f = \operatorname{tr}(C^{-1}A_f C) = \operatorname{tr}(A_f C C^{-1}) = \operatorname{tr} A_f.$$

Определение 2.1.5 Определим ранг $\operatorname{rk} f$ линейного оператора равенством $\operatorname{rk} f = \operatorname{rk} A_f$.

Он тоже, очевидно, не будет зависеть от выбора базиса.

Определение 2.1.6 *Ядро* линейного оператора $\operatorname{Ker} f = \{x \in V : f(x) = 0\}$ — множество всех векторов, переходящих в ноль.
Образ линейного оператора $\operatorname{Im} f = \{y \in V : \exists x \in V, f(x) = y\}$ — множество векторов $y \in V$, для которых существует прообраз.

Аналогичные определения можно дать и для более общего случая — для линейных отображений одного векторного пространства в другое.

Лемма 2.1.7 *Ядро и образ любого линейного оператора являются линейными подпространствами в V.*

Доказательство. Доказательство очевидно, надо просто проверить, что эти множества замкнуты относительно операций сложения и умножения на скаляры. Например, в случае ядра, если $x, y \in \operatorname{Ker} f$, $\lambda \in \mathbb{K}$, то $f(x) = f(y) = 0$, поэтому $f(x+y) = 0$, $f(\lambda x) = 0$ и $x + y, \lambda x \in \operatorname{Ker} f$. Проверка для образа оператора аналогична. □

Лемма 2.1.8 $\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim V$.

Доказательство. Пусть $e_1, \ldots, e_r$ — базис в $\operatorname{Ker} f$, дополним его до базиса $e_1, \ldots, e_r, e_{r+1}, \ldots, e_n$ всего пространства V. Докажем, что $\dim \operatorname{Im} f = n - r$. Для этого рассмотрим набор векторов $f(e_{r+1}), \ldots, f(e_n)$ и докажем, что он является базисом в $\operatorname{Im} f$.
1) линейная независимость. Пусть $\lambda_{r+1} f(e_{r+1}) + \ldots + \lambda_n f(e_n) = f(\lambda_{r+1} e_{r+1} + \ldots + \lambda_n e_n) = 0$, следовательно $\lambda_{r+1} e_{r+1} + \ldots + \lambda_n e_n \in \operatorname{Ker} f$, но тогда $\lambda_{r+1} e_{r+1} + \ldots + \lambda_n e_n = \mu_1 e_1 + \ldots + \mu_r e_r$ для некоторых $\mu_1, \ldots, \mu_r$. Т.к. векторы $e_1, \ldots, e_n$ линейно независимы, то все $\lambda_i = 0$ (и μ_j тоже), следовательно векторы $f(e_{r+1}), \ldots, f(e_n)$ линейно независимы.
2) полнота. Возьмем произвольный $y \in \operatorname{Im} f$, следовательно существует такой $x \in V$, что $f(x) = y$. Если $x = x^i e_i$ (суммирование по индексу i, пробегающему от 1 до n), то $y = f(x) = f(x^i e_i) = x^i f(e_i)$, что является линейной комбинацией векторов $f(e_{r+1}), \ldots, f(e_n)$, т.к. при $i = 1, \ldots, r$ $e_i \in \operatorname{Ker} f$ и $f(e_i) = 0$.
Следовательно $f(e_{r+1}), \ldots, f(e_n)$ — базис в $\operatorname{Im} f$, отсюда уже вытекает утверждение леммы. □

Определение 2.1.9 *Композицией* двух линейных операторов $f, g : V \to V$ называются линейные операторы $f \circ g, g \circ f : V \to V$, где $(f \circ g)(x) = f(g(x))$ и $(g \circ f)(x) = g(f(x))$.

Можно легко показать, что в фиксированном базисе $A_{f \circ g} = A_f \cdot A_g$, также легко проверить, что для множества операторов выполнены все аксиомы кольца (если умножение — композиция), т.е. множество линейных операторов имеет структуру кольца с единицей, роль которой играет тождественный оператор.

2.2 Инвариантное подпространство

Определение 2.2.1 Пусть дан линейный оператор $f : W \to W$ и $V \subset W$ — подпространство в W. Оно называется *инвариантным* подпространством относительно f, если его образ лежит в нем самом, т.е. $f(V) \subset V$.

Примеры:
1) $V = \operatorname{Ker} f$ будет инвариантным подпространством, т.к. $\forall x \in V \quad f(x) = 0 \in V$,
2) $V = \operatorname{Im} f$ будет инвариантным подпространством, т.к. по определению $\operatorname{Im} f$ образ любого элемента ему принадлежит.

Рассмотрим подробнее матрицы операторов. Пусть V — инвариантное относительно f подпространство в W. Пусть $e_1, \ldots, e_r$ — базис в V, дополним его до базиса $e_1, \ldots, e_n$ в W. Пусть A_f — матрица оператора в этом базисе, тогда она имеет следующий вид: $A_f = \left(\begin{array}{c|c} \star & \star \\ \hline 0 & \star \end{array}\right)$, т.е. ее можно разбить по ширине и высоте на две части, отвечающие векторам $e_1, \ldots, e_r$ и $e_{r+1}, \ldots, e_n$, причем в нижнем левом углу будут стоять одни нули. Действительно, т.к. V инвариантно, то $f(e_i) \in V$ при $1 \leqslant i \leqslant r$, следовательно, $f(e_i) = \alpha_i^1 e_1 + \ldots + \alpha_i^r e_r$. Коэффициенты в этом разложении по базису — это i-й столбец матрицы A_f, а здесь на $r+1, \ldots, n$-ых местах стоят нули.

Если $W = V_1 \oplus V_2$, где V_1, V_2 — инвариантные подпространства, то и правый верхний угол матрицы A_f будет нулевой, и эта матрица будет иметь следующий вид: $A_f = \left(\begin{array}{c|c} \star & 0 \\ \hline 0 & \star \end{array}\right)$. Доказательство этого аналогично предыдущему.

Ограничение оператора и фактор-оператор

Определение 2.2.2 Пусть V — инвариантное относительно f подпространство, тогда оператор $f_1 : V \to V$, определенный равенством $f_1(v) = f(v)$, $v \in V$, называется *ограничением* оператора f на подпространство V и часто обозначается $f|_V$.

Матрицей оператора $f|_V$ будет левый верхний угол матрицы оператора f, т.е. $A_f = \left(\begin{array}{c|c} A_{f_1} & \star \\ \hline 0 & \star \end{array}\right)$.

Если V инвариантно, то можно рассмотреть оператор $f' : W/V \to W/V$ (*фактор-оператор*), определенный равенством $f'(a+V) = f(a)+V$. Для проверки корректности такого определения предположим, что $a_1 + V = a_2 + V$, тогда

$$\begin{aligned} f'(a_1+V) &= f(a_1) + V = f(a_2 + \underbrace{(a_1 - a_2)}_{\in V}) + V = \\ &= f(a_2) + \underbrace{f(a_1 - a_2)}_{\in V} + V = f(a_2) + V = \\ &= f'(a_2 + V). \end{aligned}$$

Если $e_1 \ldots, e_r$ — базис в V, а $e_{r+1}, \ldots, e_n$ — его дополнение до базиса в W, то $e_{r+1}+V, \ldots, e_n+V$ будет базисом в фактор-пространстве W/V. Матрицей оператора в этом базисе будет правый нижний угол матрицы оператора A_f, т.е. $A_f = \left(\begin{array}{c|c} A_{f_1} & \star \\ \hline 0 & A_{f'} \end{array}\right)$.

2.3 Невырожденные операторы. Собственные значения и собственные векторы

Определение 2.3.1 Линейный оператор $f : V \to V$ называется *невырожденным*, если выполнено одно из следующих условий:
1) $\det f \neq 0$;
2) $\operatorname{Ker} f = \{0\}$;
3) $\operatorname{Im} f = V$;

4) $\operatorname{rk} f = \dim V$;
5) $\exists g : V \to V$, такой что $g \circ f = f \circ g = id$, т.е. существует обратный оператор.

Лемма 2.3.2 *Все эти пять свойств эквивалентны.*

Доказательство. 2) $\iff$ 3), т.к. $\dim V = \dim \operatorname{Ker} f + \dim \operatorname{Im} f$.
1) $\iff$ 2): Пусть существует ненулевой вектор $x \in \operatorname{Ker} f$. Выберем такой базис в V, чтобы x был первым вектором базиса, тогда в матрице оператора A_f первый столбец будет нулевым, тогда $\det f = 0$. Обратно, если $\det f = 0$, то у системы уравнений $A_f X = 0$ существует ненулевое решение, т.е. под действием оператора f некоторый ненулевой вектор переходит в 0. Но тогда $\operatorname{Ker} f \neq \{0\}$.
1) $\iff$ 4) — это мы знаем из курса высшей алгебры.
1) $\iff$ 5). Если $\det f \neq 0$ и A_f — матрица оператора f, то $\det A_f \neq 0$, следовательно существует обратная матрица A_f^{-1}, ей соответствует некоторый оператор g. Т.к. $A_f A_f^{-1} = A_f^{-1} A_f = E$, то $f \circ g = g \circ f = id$. Обратно, если существует обратный оператор, то его матрица будет обратной к матрице оператора f, следовательно $\det f = \det A_f \neq 0$.
Замечание. Обратный оператор (если он существует) единственен.
Оператор, для которого ни одно из этих свойств не выполняется называется *вырожденным*.

Собственные значения и собственные векторы

Определение 2.3.3 Число $\lambda \in \mathbb{K}$ называется *собственным значением* оператора f, если оператор $f - \lambda \cdot id$ вырожденный. Тогда у этого оператора — ненулевое ядро. Любой ненулевой вектор, принадлежащий ядру оператора $f - \lambda \cdot id$, называется *собственным вектором*, отвечающим собственному значению λ.

Рассмотрим пространство $V(\lambda) = \operatorname{Ker}(f - \lambda \cdot id)$ — подпространство собственных векторов, отвечающих одному и тому же собственному значению λ.

Лемма 2.3.4 *Пространство $V(\lambda)$ инвариантно относительно оператора f.*

Доказательство. Если $x \in V(\lambda)$, т.е. $(f - \lambda \cdot id)(x) = 0$, тогда $f(x) = \lambda x \in V(\lambda)$. □

Лемма 2.3.5 *Пусть $\mathbb{K}$ — алгебраически замкнутое поле (т.е. любой многочлен $f \in \mathbb{K}_n[x]$, $\deg f > 0$, имеет корень), например, поле комплексных чисел. Тогда у любого оператора $f : W \to W$, где $\dim W > 1$, существует нетривиальное инвариантное подпространство (отличное от нуля и от всего пространства).*

Доказательство. Рассмотрим уравнение $\det(f - \lambda \cdot id) = 0$. В силу алгебраической замкнутости поля, это уравнение имеет корень λ_0, тогда λ_0 будет собственным значением f и тогда $\dim V(\lambda_0) > 0$ и $V(\lambda_0)$ инвариантно. Если $V(\lambda_0) \neq W$, то оно нетривиально. Если же случайно получилось, что $V(\lambda_0) = W$, то f имеет вид $f = \lambda_0 \cdot id$, т.е. является просто оператором умножения на число, и тогда любое подпространство будет инвариантным. □

2.4 Проекторы

Если $W = V_1 \oplus V_2$, то для любого вектора w имеет место единственное разложение вида $w = v_1 + v_2$, где $v_1 \in V_1$, $v_2 \in V_2$. Рассмотрим линейный оператор $f : W \to W$, определенный формулой $f(w) = v_1$. Т.к. $v_1 = v_1 + 0$, то $f(V_1) \subset V_1$, т.е. V_1 инвариантно относительно f, более того на подпространстве имеем $f|_{V_1} = id_{V_1}$. Т.к. все вектора из V_2 переходят в 0, то $V_2 \in \operatorname{Ker} f$. На самом деле $V_2 = \operatorname{Ker} f$, т.к. если $f(w) = 0$, то в разложении $w = v_1 + v_2$ имеем $v_2 = 0$, т.е. $w \in V_2$.

Определение 2.4.1 Операторы указанного вида называются операторами проектирования или просто *проекторами* вдоль V_2 на V_1.

Проекторы обладают замечательным свойством: если f — проектор, то $f^2 = f$. Докажем обратное утверждение.

Теорема 2.4.2 *Если $f^2 = f$, то оператор $f : W \to W$ является оператором проектирования для некоторых V_1 и V_2.*

Доказательство. Возьмем $V_1 = \operatorname{Im} f$ и $V_2 = \operatorname{Ker} f$ и докажем, что f — проектор вдоль V_2 на V_1.

Сначала докажем, что $W = V_1 \oplus V_2$, т.е., что $W = V_1 + V_2$ и $V_1 \cap V_2 = \{0\}$. Допустим, что существует ненулевой вектор $a \in V_1 \cap V_2$, тогда $a \in \operatorname{Ker} f$, т.е. $f(a) = 0$ и $a \in \operatorname{Im} f$, т.е. существует такой вектор $b \in W$, что $f(b) = a$. Тогда

$$a = f(b) = f^2(b) = f(a) = 0,$$

следовательно, $a = 0$. Мы получили, что V_1 и V_2 действительно образуют прямую сумму и $V_1 \oplus V_2 \subset W$. Но, т.к.

$$\dim(V_1 \oplus V_2) = \dim V_1 + \dim V_2 = \dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim W,$$

то $V_1 \oplus V_2 = W$.

Возьмем теперь произвольный вектор $w \in W$, тогда $w = v_1 + v_2$, где $v_1 \in V_1$, $v_2 \in V_2$, следовательно, $f(w) = f(v_1 + v_2) = f(v_1) + f(v_2) = f(v_1) + 0 = f(v_1)$, т.к. $v_2 \in \operatorname{Ker} f$. Нам осталось доказать, что, если $v_1 \in \operatorname{Im} f$, то $f(v_1) = v_1$. Пусть $b \in W$ — прообраз v_1, т.е. $f(b) = v_1$, тогда $v_1 = f(b) = f^2(b) = f(v_1)$, следовательно оператор f действительно является оператором проектирования вдоль V_2 на V_1. □

Матрица оператора проектирования в базисе, составленном из базисов подпространств V_1 и V_2 имеет следующий вид: $\begin{pmatrix} 1 & & & & & \\ & \ddots & & & \mathbf{0} & \\ & & 1 & & & \\ & & & 0 & & \\ & \mathbf{0} & & & \ddots & \\ & & & & & 0 \end{pmatrix}$, где количество единиц равно размерности подпространства V_1.

2.5 Нильпотентные операторы

Определение 2.5.1 Оператор f называется *нильпотентным*, если существует такое натуральное число k, что $f^k = 0$.

Рассмотрим пример нильпотентного оператора. Пусть матрица оператора в некотором базисе $e_1, \ldots, e_n$ имеет вид $A_f = \begin{pmatrix} 0 & 1 & & \mathbf{0} \\ & 0 & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & 0 \end{pmatrix}$, т.е. базисные вектора он сдвигает по циклу: $e_n \mapsto e_{n-1} \mapsto \ldots \mapsto e_1 \mapsto 0$. При возведении этой матрицы в степени, диагональ с единицами будет все дальше и дальше смещаться в правый верхний угол и через некоторое время вообще исчезнет, т.е. будучи возведенным в степень n, матрица станет нулевой, а значит и сам оператор станет нулевым в этой степени, т.е. будет нильпотентным.

Теорема 2.5.2 *Если оператор $f : W \to W$ нильпотентный, то найдется базис, в котором матрица оператора будет иметь блочно-диагональный вид $A_f = \begin{pmatrix} A_{n_1} & & & \mathbf{0} \\ & A_{n_2} & & \\ & & \ddots & \\ \mathbf{0} & & & A_{n_k} \end{pmatrix}$, где каждый блок (клетка) A_{n_i} размерности n_i имеет вид $\begin{pmatrix} 0 & 1 & & \mathbf{0} \\ & 0 & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & 0 \end{pmatrix}$. Такой вид матрицы оператора единственен с точностью до перестановки блоков.*

Такая форма матрицы оператора называется *нормальной*.

Доказательство.

1) Единственность. Пусть f — нильпотентный оператор, рассмотрим последовательность операторов $f^0 = id, f^1 = f, f^2, \ldots, f^n = 0$. Обозначим $r_i = \operatorname{rk} f^i$, в частности $r_0 = \dim W$, $r_n = 0$, и мы имеем цепочку неравенств $r_0 \geqslant r_1 \geqslant \ldots \geqslant r_n$. Обозначим через N_k количество клеток размера $k \times k$. Свяжем между собой числа r_i и N_k. Для этого рассмотрим разность $r_0 - r_1$. Для каждой отдельно взятой клетки такая разность равна 1, т.к. ранг клетки размера $k \times k$ равен $k-1$, т.е. на единицу меньше, чем размерность. Поэтому каждая клетка вносит в разность $r_0 - r_1$ вклад, равный единице, т.е. эта разность равна общему количеству клеток, следовательно $r_0 - r_1 = N_1 + N_2 + N_3 + \ldots$.

Рассмотрим разность $r_1 - r_2$. Для клеток размера 1×1 такая разность равна нулю, а для клеток размера $n \times n$ при $n \geqslant 2$ ранг клетки равен $n-1$, а ранг ее квадрата равен $n-2$, и их разность равна единице. Поэтому разность $r_1 - r_2$ равна количеству клеток размера $n \times n$ при $n \geqslant 2$, т.е. $r_1 - r_2 = N_2 + N_3 + \ldots$. Аналогично получаем, что $r_2 - r_3 = N_3 + N_4 + \ldots$ и т.д., $r_{i-1} - r_i = N_i + N_{i+1} + \ldots$. Вычитая из предыдущего равенства последующее, получаем $N_i = (r_{i-1} - r_i) - (r_i - r_{r+1}) = r_{i-1} - 2r_i + r_{r+1}$. Т.к. ранги от выбора базиса не зависят, то и числа N_i от базиса не зависят, следовательно количество клеток каждого размера будет одно и то же, поэтому нормальная форма оператора единственна с точность до перестановки клеток, из которых она состоит.

2) Существование. Проведем индукцию по размерности пространства W.

Если $\dim W = 1$, тогда, т.к. f нильпотентен, то $f = 0$ и в любом базисе имеет нужный (нулевой) вид.

В предположении, что утверждение теоремы верно для $\dim W < n$, докажем его для $\dim W = n$. Рассмотрим образ оператора $V = \operatorname{Im} f$ — инвариантное подпространство, и ядро $U = \operatorname{Ker} f$. Рассмотрим ограничение f на V — оператор $f_1 : V \to V$. Покажем, что $\dim V < \dim W$. Если мы предположим противное, $\dim V = \dim W$, это будет означать, что $V = W$, следовательно, оператор f будет обратимым, следовательно, $\det f \neq 0$. Но тогда $\det f^n = (\det f)^n \neq 0$ для любой степени n, что приводит к противоречию с нильпотентностью.

Итак, $\dim V < \dim W$, и в подходящем базисе матрица оператора f имеет блочный вид $A_f = \begin{pmatrix} A_{f_1} & \star \\ \mathbf{0} & \star \end{pmatrix}$. Тогда его степени имеют такой же вид: $A_f^n = \begin{pmatrix} A_{f_1}^n & \star \\ \mathbf{0} & \star \end{pmatrix}$, следовательно, если $A_f^n = 0$, то и $A_{f_1}^n = 0$, и оператор f_1 также нильпотентный (причем $\dim V < n$). Значит, по предположению индукции оператор f_1 имеет базис, в котором он представим в нужном нам виде $A_{f_1} = \begin{pmatrix} A_{k_1} & & & \mathbf{0} \\ & A_{k_2} & & \\ & & \ddots & \\ \mathbf{0} & & & A_{k_t} \end{pmatrix}$, где t — количество клеток. Этому разбиению на блоки будет соответствовать базис

$$\underbrace{a_1^{(1)}, \ldots, a_{k_1}^{(1)}}_{\text{отвечает } A_{k_1}}, \underbrace{a_1^{(2)}, \ldots, a_{k_2}^{(2)}}_{\text{отвечает } A_{k_2}}, \ldots, \underbrace{a_1^{(t)}, \ldots, a_{k_t}^{(t)}}_{\text{отвечает } A_{k_t}}.$$

Все эти вектора вместе образуют базис в V, обозначим эту группу векторов через (I).

Дополним ее до базиса всего W. Если $\operatorname{rk} f = r$, то $\dim V = r$, следовательно, $\dim U = n - r$. Под действием оператора f_1 вектор $a_{k_i}^{(i)}$ переходит в $a_{k_i-1}^{(i)}$, $a_{k_i-1}^{(i)} \mapsto a_{k_i-2}^{(i)}$, $\ldots$, $a_2^{(i)} \mapsto a_1^{(i)}$, $a_1^{(i)} \mapsto 0$. Следовательно вектора $a_1^{(1)}, \ldots, a_1^{(t)} \in \operatorname{Ker} f_1$, более того, $\operatorname{Ker} f_1 = \langle a_1^{(1)}, \ldots, a_1^{(t)} \rangle$, т.к. все пространство разбивается в прямую сумму подпространств, отвечающих клеткам $A_{k_1}, \ldots, A_{k_t}$. Т.к. $U \cap V = \operatorname{Ker} f_1 \subset \operatorname{Ker} f = U$, то базис $a_1^{(1)}, \ldots, a_1^{(t)}$ подпространства $U \cap V$ можно дополнить до базиса пространства U. Дополним его векторами $b_1, \ldots, b_{n-r-t}$ и обозначим эту группу векторов через (II) (количество векторов в ней равно $\dim U - \dim(U \cap V) = (n-r) - t$). Для того, чтобы дополнить группы векторов (I) и (II) до базиса всего пространства W, нам нужно еще t векторов. Рассмотрим группу векторов (I). Под действием оператора f_1 в векторы $a_{k_1}^{(1)}, \ldots, a_{k_t}^{(1)}$ ничего не переходит, однако во всем пространстве W есть вектора, в них переходящие, т.к. $a_{k_1}^{(1)}, \ldots, a_{k_t}^{(1)} \in$

$\operatorname{Im} f = V$. Пусть это вектора $c_1, \ldots, c_t$, обозначим эту группу векторов через (III). Докажем, что объединение групп векторов (I), (II) и (III) образует базис в W.

Линейная независимость. Пусть линейная комбинация этих векторов равна нулю:

$$\alpha_1^{(1)} a_1^{(1)} + \ldots + \alpha_{k_1}^{(1)} a_{k_1}^{(1)} + \ldots + \alpha_1^{(t)} a_1^{(t)} + \ldots + \alpha_{k_t}^{(t)} a_{k_t}^{(t)} + \\ + \beta_1 b_1 + \ldots + \beta_{n-r-t} b_{n-r-t} + \gamma_1 c_1 + \ldots + \gamma_t c_t = 0.$$

Применим к обоим частям равенства оператор f, тогда получим

$$\alpha_2^{(1)} a_1^{(1)} + \ldots + \alpha_{k_1}^{(1)} a_{k_1-1}^{(1)} + \ldots + \alpha_2^{(t)} a_1^{(t)} + \ldots + \alpha_{k_t}^{(t)} a_{k_t-1}^{(t)} + \\ + 0 + \ldots + 0 + \gamma_1 a_{k_1}^{(1)} + \ldots + \gamma_t a_{k_t}^{(t)} = 0.$$

Но вектора группы (I) образуют базис в V, следовательно, они линейно независимы, и все присутствующие здесь коэффициенты равны нулю, т.е.

$$\alpha_2^{(1)} = \ldots = \alpha_{k_1}^{(1)} = \ldots = \alpha_2^{(t)} = \ldots = \alpha_{k_t}^{(t)} = \gamma_1 = \ldots = \gamma_t = 0.$$

Тогда от первоначальной суммы остается лишь

$$\alpha_1^{(1)} a_1^{(1)} + \ldots + \alpha_1^{(t)} a_1^{(t)} + \beta_1 b_1 + \ldots + \beta_{n-r-t} b_{n-r-t} = 0,$$

а эти вектора как раз образуют базис в U, следовательно они линейно независимы и все оствшиеся коэффициенты также равны нулю, следовательно система векторов групп $(I) \cup (II) \cup (III)$ линейно независима.

Полнота этой системы векторов очевидна, т.к. количество векторов равно $k_1 + \ldots + k_t + (n - r - t) + t = r + (n - r - t) + t = n = \dim W$, значит это действительно базис.

Посмотрим теперь, как выглядит матрица оператора f в этом базисе. Перепишем этот базис в следующем порядке:

$$\underbrace{a_1^{(1)}, \ldots, a_{k_1}^{(1)}, c_1}, \ldots\ldots, \underbrace{a_1^{(t)}, \ldots, a_{k_t}^{(t)}, c_t}, b_1, \ldots, b_{n-r-t}.$$

По построению, $f(b_i) = 0$ для $i = 1, \ldots, n - r - t$, а в каждой выделенной группе вектора переходят друг в друга циклически: $c_i \mapsto a_{k_i}^{(i)} \mapsto \ldots \mapsto a_1^{(i)} \mapsto 0$, $i = 1, \ldots, t$, т.е. каждой группе отвечает клетка размера на единицу больше, чем A_{k_i} т.к. к каждой циклической цепочке векторов мы добавили еще по одному вектору. В этом базисе матрица оператора будет иметь вид: $A_f = \begin{pmatrix} A_{k_1+1} & & & & & & \\ & A_{k_2+1} & & & & \mathbf{0} & \\ & & \ddots & & & & \\ & & & A_{k_t+1} & & & \\ & & & & 0 & & \\ & \mathbf{0} & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$, где одномерные нулевые клетки отвечают векторам b_i. Теорема доказана. $\square$

2.6 Многочлены от операторов

Пусть $f : V \to V$ — линейный оператор. Тогда каждому многочлену $p(t) \in \mathbb{K}_n[t]$ можно поставить в соответствие оператор по следующему правилу:

$$a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n \mapsto a_0 id + a_1 f + a_2 f^2 + \ldots + a_n f^n$$

— этот многочлен от оператора, также являющийся оператором, мы будем обозначать через $p(f)$.

Может так получиться, что $p(f)$ — нулевой оператор, тогда многочлен $p(t)$ называется *аннулирующим многочленом* для оператора f.

Пример: Если $f = id$, то $p(t) = t - 1$ будет аннулирующим многочленом, т.к. $p(f) = f - id = 0$.

Лемма 2.6.1 *У любого оператора f существует аннулирующий многочлен.*

Доказательство. Пусть $\dim V = n$, рассмотрим операторы $\underbrace{f^0 = id, f^1 = f, f^2, \ldots, f^{n^2}}_{n^2+1}$. Размерность векторного пространства линейных операторов равна n^2, следовательно эти операторы (поскольку их количество больше размерности) линейно зависимы, тогда существуют такие числа $a_0, a_1, \ldots, a_{n^2}$, не все равные нулю, что $a_0 id + a_1 f + \ldots + a_{n^2} f^{n^2} = 0$, но тогда получаем, что многочлен $p(t) = a_0 + a_1 t + \ldots + a_{n^2} t^{n^2}$ аннулирует оператор f. □

Минимальный многочлен

Определение 2.6.2 Многочлен $p(t)$ называется *минимальным многочленом* для оператора f, если он аннулирует этот оператор, имеет наименьшую степень среди всех аннулирующих многочленов и его старший коэффициент равен 1.

Лемма 2.6.3 *Для любого оператора f существует, и притом единственный, минимальный многочлен.*

Доказательство. 1) Существование. Мы уже показали, что для любого оператора существует аннулирующий многочлен. Поэтому мы можем выбрать из всех аннулирующих многочленов многочлены с наименьшей степенью и поделить их на старший коэффициент. То, что получится, по определению будет минимальным многочленом.

2) Единственность. Допустим, что $p_1(t)$ и $p_2(t)$ — два минимальных многочлена для одногои того же оператора f, тогда $\deg p_1 = \deg p_2$ и их старшие коэффициенты равны 1, поэтому многочлен $p_1(t) - p_2(t)$ будет аннулирующим многочленом меньшей степени, что противоречит предположению. □

Лемма 2.6.4 *Число λ будет собственным значением оператора f тогда и только тогда, когда λ — корень минимального многочлена для f.*

Доказательство. Пусть λ — собственное значение оператора f, тогда существует такой ненулевой вектор x, что $fx = \lambda x$. Пусть $p(t)$ — минимальный многочлен, т.е. $p(f) \equiv 0$, тогда $p(f)x = 0$, следовательно

$$p(\lambda)x = a_0 x + a_1 \lambda x + \ldots + a_n \lambda^n x = a_0 x + a_1 f x + \ldots + a_n f^n x = p(f)x = 0,$$

поэтому $p(\lambda) = 0$ (так как $x \neq 0$).

Обратно, пусть λ — корень минимального многочлена, тогда $p(t) = (t - \lambda)q(t)$, $\deg q < \deg p$, поэтому $q(t)$ не аннулирует f, следовательно, существует такой ненулевой вектор x, что $q(f)x = y \neq 0$. Тогда

$$(f - \lambda \cdot id)y = (f - \lambda \cdot id)q(f)x = p(f)x = 0,$$

следовательно, y — это собственный вектор оператора f, а λ — его собственное значение. □

2.7 Характеристический многочлен

Определение 2.7.1 Многочлен $P_f(t) = \det(f - \lambda \cdot id)$ называется *характеристическим многочленом* оператора f.

Отметим роль некоторых коэффициентов характеристического многочлена. Если его записать в виде $P_f(t) = a_0 + a_1 t + \ldots + a_n t^n$, то тогда $a_n = (-1)^n$, $a_{n-1} = (-1)^{n-1} \operatorname{tr} f$, $a_0 = \det f$.

Лемма 2.7.2 *λ является собственным значением оператора f тогда и только тогда, когда λ — корень характеристического многочлена $P_f(t)$.*

Доказательство. Если λ — собственное значение, то оператор $g = f - \lambda \cdot id$ вырожденный, следовательно $\det(f - \lambda \cdot id) = 0$, т.е. λ — корень $P_f(t)$. Обратно, если λ — корень $P_f(t)$, то $\det(f - \lambda \cdot id) = 0$, следовательно оператор $f - \lambda \cdot id$ вырожденный, значит, λ является собственным значением оператора f. $\square$

Если $V \subset W$ — инвариантное подпространство, тогда, как мы знаем, матрица оператора $f : W \to W$ имеет вид: $A_f = \begin{pmatrix} A_{f_1} & \star \\ \mathbf{0} & A_{f'} \end{pmatrix}$, где f_1 — ограничение f на V, а f' — фактор-оператор. Тогда, т.к. $\det f = \det f_1 \det f'$, то $P_f(t) = P_{f_1}(t) P_{f'}(t)$.

Лемма 2.7.3 *Пусть $V(\lambda)$ — инвариантное подпространство, образованное собственными векторами, отвечающими собственному значению λ. Тогда кратность корня характеристического многочлена не меньше размерности подпространства $V(\lambda)$.*

Доказательство. Если f_1 — ограничение оператора f на $V(\lambda)$, то для любого $x \in V(\lambda)$ будем иметь, что $f_1 x = \lambda x$, поэтому матрица этого оператора имеет вид $A_{f_1} = \begin{pmatrix} \lambda & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda \end{pmatrix}$, а матрица оператора f — вид: $A_f = \left(\begin{array}{ccc|c} \lambda & & \mathbf{0} & \\ & \ddots & & \star \\ \mathbf{0} & & \lambda & \\ \hline & \mathbf{0} & & A_{f'} \end{array}\right)$. Следовательно $P_f(t) = P_{f_1}(t) P_{f'}(t) = (\lambda - t)^{\dim V(\lambda)} \cdot P_{f'}(t)$, поэтому кратность корня λ не меньше размерности подпространства $V(\lambda)$ (но может быть и больше, если λ является корнем многочлена $P_{f'}(t)$). $\square$

Теорема 2.7.4 (Гамильтона-Кэли) *Характеристический многочлен $P_f(t)$ оператора $f : W \to W$ аннулирует этот оператор, т.е. $P_f(f) = 0$.*

Доказательство. Мы ограничимся доказательством теоремы только случаем алгебраически замкнутых полей.

Проведем индукцию по размерности пространства.

1) База индукции. Пусть $\dim W = 1$. В этом случае матрица оператора f имеет вид $A_f = (\lambda)$ и оператор имеет вид $fx = \lambda x$. Характеристический многочлен имеет вид $P_f(t) = \lambda - t$, следовательно, $P_f(f)x = \lambda x - fx = 0$ для любого $x \in W$. База индукции верна.

2) Индуктивный переход. Пусть утверждение теоремы верно для $\dim W < n$, докажем его для $\dim W = n$. Т.к. поле алгебраически замкнуто, то многочлен $P_f(t)$ имеет корень λ, т.е. $P_f(t) = (\lambda - t) q(t)$ для некоторого многочлена $q(t)$. Тогда, как уже отмечалось, λ будет собственным значением оператора f. Значит, существует собственный вектор $x \in W$, отвечающий собственному значению λ. Рассмотрим одномерное инвариантное линейное подпространство V, порожденное вектором x. Тогда матрица оператора f в подходящем базисе будет иметь вид $A_f = \begin{pmatrix} \lambda & \star \\ \mathbf{0} & A_{f'} \end{pmatrix}$. Пусть $P_{f'}(t)$ — характеристический многочлен фактор-оператора f', тогда $P_f(t) = (\lambda - t) P_{f'}(t)$ и, более того, т.к. $\dim(W/V) = n - 1$, то, по предположению индукции, $P_{f'}(f') = 0$ (здесь 0 — это нулевой вектор фактор-пространства). Это значит, что для любого $a \in W$

$$P_{f'}(f')(a + V) = V. \tag{3}$$

Вспомним теперь определение фактор-оператора в фактор-пространстве: $f'(a + V) = f(a) + V$. Легко показать, что, применяя оператор несколько раз подряд, получим $(f')^k(a + V) = f^k(a) + V$ для любой степени k. Но тогда $P_{f'}(f')(a + V) = P_{f'}(f)a + V$. Сравнивая это равенство с (3), получаем $P_{f'}(f)a + V = V$, следовательно, $P_{f'}(f)a \in V$ для произвольного вектора $a \in W$. Поскольку $\dim V = 1$, $P_{f'}(f)a = \alpha x$ для некоторого скаляра α. Поэтому

$$P_f(f)a = (\lambda \cdot id - f) P_{f'}(f)a = (\lambda \cdot id - f)\alpha x = \alpha(\lambda x - fx) = 0,$$

т.к. x — собственный вектор, отвечающий собственному значению λ. Итак, мы получили, что $P_f(f)a = 0$ для любого вектора a, т.е. $P_f(t)$ аннулирует f. $\square$

2.8 Еще несколько утверждений об операторах

Определение 2.8.1 Оператор f называется *диагонализируемым*, если существует такой базис, что матрица этого оператора в этом базисе диагональна.

Примеры:
1) операторы проектирования диагонализируемы,
2) нильпотентные операторы не диагонализируемы (если они не нулевые), так как любая диагональная нильпотентная матрица равна нулю.

Лемма 2.8.2 *Пусть характеристический многочлен $P_f(t)$ имеет $n = \dim W$ различных корней, тогда оператор f диагонализируемый.*

Доказательство. Пусть $\lambda_1, \ldots, \lambda_n$ — корни $P_f(t)$, т.е. собственные значения оператора f, а $a_1, \ldots, a_n$ — отвечающие им собственные вектора, т.е. $fa_i = \lambda_i a_i$, $i = 1, \ldots, n$. Если бы мы знали, что $a_1, \ldots, a_n$ — базис, то матрица оператора в этом базисе имела бы вид: $A_f = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$.

Докажем, что $a_1, \ldots, a_n$ является базисом. Для этого нам достаточно доказать линейную независимость этих векторов (так как их количество совпадает с размерностью пространства). Проведем индукцию по размерности пространства.

1) База индукции. Если $\dim W = 1$, то вектор a_1 будет базисом. База индукции верна.

2) Индуктивный переход. Пусть для $\dim W < n$ утверждение верно, докажем его для $\dim W = n$. Допустим, что это не так, т.е. существуют скаляры $\alpha_1, \ldots, \alpha_n$, не все равные нулю, т.ч. $\alpha_1 a_1 + \ldots + \alpha_n a_n = 0$. Без ограничения общности можно считать, что $a_n \neq 0$, тогда вектор a_n есть линейная комбинация остальных векторов, $a_n = -\frac{\alpha_1}{\alpha_n} a_1 - \ldots - \frac{\alpha_{n-1}}{\alpha_n} a_{n-1}$. Применив оператор f к обеим частям этого равенства, получим, что $fa_n = f(-\frac{\alpha_1}{\alpha_n} a_1 - \ldots - \frac{\alpha_{n-1}}{\alpha_n} a_{n-1})$, т.е. $\lambda_n a_n = -\frac{\alpha_1}{\alpha_n}\lambda_1 a_1 - \ldots - \frac{\alpha_{n-1}}{\alpha_n}\lambda_{n-1} a_{n-1}$, но, с другой стороны, $\lambda_n a_n = \lambda_n(-\frac{\alpha_1}{\alpha_n} a_1 - \ldots - \frac{\alpha_{n-1}}{\alpha_n} a_{n-1})$. Приравняв выражения в правых частях равенств, получим, что $(\lambda_n - \lambda_1)\frac{\alpha_1}{\alpha_n} a_1 + \ldots + (\lambda_n - \lambda_{n-1})\frac{\alpha_{n-1}}{\alpha_n} a_{n-1} = 0$. Рассмотрим подпространство $V = \langle a_1, \ldots, a_{n-1} \rangle$, оно будет инвариантно относительно оператора f и $\dim V = n - 1$. Если f_1 — ограничение f на V, то $a_1, \ldots, a_{n-1}$ — собственные вектора оператора f_1, следовательно, по предположению индукции, они линейно независимы. Т.к. все λ_i различны, то $\alpha_1 = \ldots = \alpha_{n-1} = 0$, но тогда $a_n = 0$, что невозможно, т.к. это собственный вектор. Полученное противоречие доказывает линейную независимость векторов $a_1, \ldots, a_n$. □

Лемма 2.8.3 *Над алгебраически замкнутым полем матрицу любого оператора можно привести к верхнетреугольному виду* $\begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$ *заменой базиса.*

Доказательство. Индукция по размерности пространства.

1) Если $\dim W = 1$, то утверждение очевидно, т.к. любая матрица размера 1 является верхнетреугольной.

2) Пусть утверждение верно для $\dim W < n$, докажем его для $\dim W = n$. Т.к. поле алгебраически замкнуто, характеристический многочлен имеет корень λ, он будет собственным значением. Ему соответствует собственный вектор, который порождает одномерное инвариантное подпространство V, тогда матрица оператора f имеет вид $A_f = \begin{pmatrix} \lambda & \star \\ \mathbf{0} & A_{f'} \end{pmatrix}$. По предположению индукции матрицу $A_{f'}$ можно привести к верхнетреугольному виду $A_{f'} = \begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ \mathbf{0} & & \lambda_{n-1} \end{pmatrix}$,

тогда матрица $A_f = \left(\begin{array}{c|ccc} \lambda & \multicolumn{3}{c}{\star} \\ \hline & \lambda_1 & & \star \\ \mathbf{0} & & \ddots & \\ & \mathbf{0} & & \lambda_{n-1} \end{array}\right)$ также будет верхнетреугольной. $\square$

Лемма 2.8.4 *Если матрица оператора верхнетреугольная, то на главной диагонали стоят собственные значения этого оператора.*

Доказательство. Пусть $A_f = \begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$, тогда $A_{f-t\cdot id} = \begin{pmatrix} \lambda_1 - t & & \star \\ & \ddots & \\ \mathbf{0} & & \lambda_n - t \end{pmatrix}$, и $\det(f - t\cdot id) = (\lambda_1 - t)\cdot\ldots\cdot(\lambda_n - t)$, т.е. $\lambda_1, \ldots, \lambda_n$ — это корни характеристического многочлена, а значит собственные значения. $\square$

Лемма 2.8.5 *Пусть оператор f такой, что в алгебраически замкнутом поле характеристический многочлен $P_f(t)$ имеет единственный корень λ, тогда оператор $g = f - \lambda\cdot id$ является нильпотентным.*

Доказательство. Т.к. у оператора f только одно собственное значение λ, то в некотором базисе его матрица будет иметь вид $A_f = \begin{pmatrix} \lambda & & \star \\ & \ddots & \\ \mathbf{0} & & \lambda \end{pmatrix}$. Матрица оператора g в этом базисе имеет вид $A_g = \begin{pmatrix} 0 & & \star \\ & \ddots & \\ \mathbf{0} & & 0 \end{pmatrix}$. Характеристический многочлен оператора g имеет вид $P_g(t) = (-1)^n t^n$, и, поскольку он является аннулирующим, то $P_g(g) = (-1)^n g^n = 0$, значит, $g^n = 0$ и, следовательно, оператор g является нильпотентным. $\square$

2.9 Корневое подпространство

Определение 2.9.1 Пусть дан оператор $f : W \to W$ и λ — произвольный скаляр. *Корнєвым подпространством*, отвечающим числу λ называется множество V_λ, состоящее из всех векторов $x \in W$, для которых существует такое натуральное число n, что $(f - \lambda\cdot id)^n x = 0$.

Утверждение 2.9.2 *Множество V_λ действительно является подпространством.*

Доказательство. Если $x \in V_\lambda$, то $(f - \lambda\cdot id)(\alpha x) = 0$, следовательно, $\alpha x \in V_\lambda$ для любого $\alpha \in \mathbb{K}$.

Если $x, y \in V_\lambda$, то существуют числа $n_x, n_y \in \mathbb{N}$, для которых $(f-\lambda\cdot id)^{n_x} x = 0$, $(f-\lambda\cdot id)^{n_y} y = 0$. Пусть $n = \max(n_x, n_y)$. Тогда $(f - \lambda\cdot id)^n (x + y) = (f - \lambda\cdot id)^n x + (f - \lambda\cdot id)^n y = 0$, значит, V_λ является подпространством. $\square$

Утверждение 2.9.3 *Подпространство $V_\lambda \subset W$ инвариантно относительно оператора f.*

Доказательство. Возьмем произвольный вектор $x \in V_\lambda$, т.е. $(f - \lambda\cdot id)^n x = 0$, тогда $(f - \lambda\cdot id)^n fx = f(f - \lambda\cdot id)^n x = f(0) = 0$, значит, $fx \in V_\lambda$. $\square$

Утверждение 2.9.4 *Подпространство V_λ нетривиально (т.е. отлично от нулевого вектора) тогда и только тогда, когда λ — собственное значение оператора f.*

Доказательство. Если λ — собственное значение, то существует собственный вектор $x \neq 0$, т.ч. $(f - \lambda \cdot id)x = 0$, т.е. $x \in V_\lambda$.

Обратно, если подпространство V_λ содержит ненулевой вектор x, то пусть n — наименьшее натуральное число, при котором $(f - \lambda \cdot id)^n x = 0$. Тогда $y = (f - \lambda \cdot id)^{n-1} x \neq 0$. Но тогда $(f - \lambda \cdot id)y = (f - \lambda \cdot id)^n x = 0$, следовательно y — собственный вектор оператора f, а λ — собственное значение. □

Обозначим $g = f - \lambda \cdot id$. Очевидно, что если $g^n x = 0$ для какого-то $n \in \mathbb{N}$ и для какого-то вектора x, то и $g^m x = 0$ при всех $m > n$. Т.е. получаем цепочку вложений: $\operatorname{Ker} g \subseteq \operatorname{Ker} g^2 \subseteq \ldots \subseteq \operatorname{Ker} g^n \subseteq \ldots \subseteq V_\lambda$.

Утверждение 2.9.5 *Если на каком-то шаге этой цепочки ядра* $\operatorname{Ker} g^{k-1}$ *и* $\operatorname{Ker} g^k$ *совпали, то они будут совпадать и дальше, т.е. если* $\operatorname{Ker} g^{k-1} = \operatorname{Ker} g^k$, *то* $\operatorname{Ker} g^i = \operatorname{Ker} g^{k-1}$ *при* $i > k$.

Доказательство. По условию мы имеем $g^k x = 0 \iff x \in \operatorname{Ker} g^k \iff x \in \operatorname{Ker} g^{k-1} \iff g^{k-1} x = 0$. Докажем, что $g^{k+1} x = 0 \iff g^k x = 0$, т.е. что $\operatorname{Ker} g^{k+1} = \operatorname{Ker} g^k$.

Если $g^k x = 0$, то, очевидно, $g^{k+1} x = g(g^k x) = 0$. Обратно, если $g^{k+1} x = 0$, то $g^k(gx) = 0 \iff g^{k-1}(gx) = 0$ (по условию), следовательно, $g^k x = 0$. Мы доказали, что $\operatorname{Ker} g^{k+1} = \operatorname{Ker} g^k$. Проделав эту операцию нужное число раз, получим что $\operatorname{Ker} g^{k+2} = \operatorname{Ker} g^{k+1}$ и т.д. □

Из этого утверждения следует, что при возведении оператора g в степени, его ядро не может бесконечно увеличиваться потому, что если в этой цепочке встретилось хотя бы одно равенство, то и дальше будут идти равенства. А равенство заведомо встретится не позже степени $n = \dim W$, т.к. $\dim \operatorname{Ker} g^n \leqslant \dim W$, а с каждым шагом она увеличивается (если увеличивается) не менее, чем на единицу. При возведении оператора g в очередную степень рост его ядра заведомо прекратится при $n = \dim W$, поэтому нам достаточно ограничиться этой степенью: для любого $x \in V_\lambda$ $g^{\dim W} x = 0$.

В последующих наших рассуждениях о корневых подпространствах мы ограничимся подпространствами, отвечающими собственным значениям, т.к. все остальные подпространства будут состоять только из нулевого вектора.

2.10 Теорема о разложении пространства в прямую сумму корневых подпространств

Теорема 2.10.1 *Пусть дан оператор* $f : W \to W$ *(над алгебраически замкнутым полем). Тогда пространство* W *является прямой суммой всех корневых подпространств, т.е.* $W = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_k}$, *где* $\lambda_1, \ldots, \lambda_k$ — *собственные значения оператора* f.

Доказательство. Проведем доказательство этой теоремы по следующему плану:

I) введем в рассмотрение новые подпространства $V_1, \ldots, V_k$ и докажем, что $V_i \subseteq V_{\lambda_i}$, $i = 1, \ldots, k$;

II) докажем, что все пространство является их суммой, т.е. $W = V_1 + \ldots + V_k$;

III) докажем, что эта сумма прямая, т.е. $W = V_1 \oplus \ldots \oplus V_k$;

IV) докажем, что $V_i = V_{\lambda_i}$, откуда уже получим утверждение теоремы.

I. Т.к. поле алгебраически замкнуто, то характеристический многочлен можно разложить на линейные множители: $P_f(t) = (t - \lambda_1)^{r_1} \cdot \ldots \cdot (t - \lambda_k)^{r_k}$, причем $r_1 + \ldots + r_k = \deg P_f(t) = \dim W$. Рассмотрим многочлены

$$p_i(t) = (t - \lambda_1)^{r_1} \cdot \ldots \cdot (t - \lambda_{i-1})^{r_{i-1}} (t - \lambda_{i+1})^{r_{i+1}} \cdot \ldots \cdot (t - \lambda_k)^{r_k},$$

т.е. $p_i(t)$ — это $P_f(t)$, из которого выкинута i-ая скобка. Обозначим $f_i = p_i(f)$ и $V_i = \operatorname{Im} f_i$.

Докажем, что $V_i \subseteq V_{\lambda_i}$. Возьмем произвольный вектор $v \in V_i$, тогда у него существует прообраз $w \in W$, $v = f_i w = p_i(f)w$. Но

$$(f - \lambda_i \cdot id)^{r_i} v = (f - \lambda_i \cdot id)^{r_i} p_i(f) w = P_f(f) w = 0,$$

следовательно, $v \in V_{\lambda_i}$.

II. Здесь нам надо доказать, что любой вектор $w \in W$ можно представить в виде $w = v_1 + \ldots + v_k$, где $v_i \in V_i$. Т.к. многочлены $p_1(t), \ldots, p_k(t)$ взаимно просты, то существуют такие многочлены $q_1(t), \ldots, q_k(t)$, что $p_1(t)q_1(t) + \ldots + p_k(t)q_k(t) \equiv 1$. Тогда $p_1(f)q_1(f) + \ldots + p_k(f)q_k(f) = id$, и для любого $w \in W$ имеем $w = p_1(f)\underbrace{q_1(f)w}_{=u_1} + \ldots + p_k(f)\underbrace{q_k(f)w}_{=u_k}$, т.е. $w = p_1(f)u_1 + \ldots + p_k(f)u_k$.

По определению, $v_i = p_i(f)u_i \in \operatorname{Im} f_i = V_i$. Мы получили искомое разложение $w = v_1 + \ldots + v_k$.

III. Докажем, что пересечение каждого из подпространств V_i с суммой остальных состоит только из нулевого вектора, т.е. что $V_i \cap (+_{j \neq i} V_j) = \{0\}$, откуда будет следовать, что сумма прямая. Возьмем произвольный вектор $a \in V_i \cap (+_{j \neq i} V_j)$, т.е. $a \in V_i$ и $a \in +_{j \neq i} V_j$. Тогда $(f - \lambda_i \cdot id)^{r_i} a = 0$ и $a = a_1 + \ldots + a_{i-1} + a_{i+1} + \ldots + a_n$, где $a_j \in V_j$. По определению, $(f - \lambda_j \cdot id)^{r_j} a_j = 0$, а, так как при $j \neq i$ многочлен $p_i(t)$ содержит множитель $(t - \lambda_j)^{r_j}$, то $p_i(f)a_j = 0$. Поэтому $p_i(f)a = p_i(f)a_1 + \ldots + p_i(f)a_{i-1} + p_i(f)a_{i+1} + \ldots + p_i(f)a_k = 0$. Т.к. многочлены $(t - \lambda_i)^{r_i}$ и $p_i(t)$ взаимно просты, то найдутся такие многочлены $u(t)$, $v(t)$, что $u(t)(t - \lambda_i)^{r_i} + v(t)p_i(t) \equiv 1$, следовательно, $u(f)(f - \lambda_i \cdot id)^{r_i} + v(f)p_i(f) = id$. Тогда $a = u(f)\underbrace{(f - \lambda_i \cdot id)^{r_i} a}_{=0} + v(f)\underbrace{p_i(t)a}_{=0} = 0$, т.е. пересечение $V_i \cap (+_{j \neq i} V_j)$ состоит только из нулевого вектора, поэтому сумма прямая.

IV. Возьмем произвольный вектор $a \in V_{\lambda_i}$, тогда, т.к. $W = \oplus_{i=1}^{k} V_i$, то $a = a' + a''$, где $a' \in V_i$, $a'' \in \oplus_{j \neq i} V_j$. Докажем, что $a'' = 0$. Т.к. $a \in V_{\lambda_i}$, то существует такая степень m, что $(f - \lambda_i \cdot id)^m a = 0$. Пусть $n = \max(m, r_i)$, тогда $(f - \lambda_i \cdot id)^n a = 0$ и $(f - \lambda_i \cdot id)^n a' = 0$, следовательно $(f - \lambda_i \cdot id)^n a'' = 0$. Как было показано в предыдущем пункте, $p_i(f)a'' = 0$. Т.к. многочлены $(t - \lambda_i)^n$ и $p_i(t)$ взаимно просты, то существуют такие многочлены $u(t)$, $v(t)$, что $u(t)(t - \lambda_i)^n + v(t)p_i(t) \equiv 1$, тогда $a'' = u(f)(f - \lambda_i \cdot id)^n a'' + v(f)p_i(f)a'' = 0$. Мы доказали, что $a = a' \in V_i$, т.е., что $V_{\lambda_i} \subseteq V_i$. Обратное включение было доказано в первом пункте. Следовательно, $V_{\lambda_i} = V_i$ и $W = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_k}$. Теорема доказана. □

Лемма 2.10.2 *Пусть $p(t)$ — аннулирующий многочлен для оператора $f : W \to W$ и пусть его можно разложить в произведение двух взаимно простых многочленов $p(t) = p_1(t)p_2(t)$, то пространство W является прямой суммой двух подпространств, $W = V_1 \oplus V_2$, каждое из которых инвариантно и $p_i(t)$ аннулирует сужение f на V_i, т.е. $p_i(f)V_i = \{0\}$, $i = 1, 2$.*

Доказательство. Возьмем $V_1 = \operatorname{Im} p_2(f)$ и $V_2 = \operatorname{Im} p_1(t)$.

То, что V_i инвариантны и их прямая сумма равна W, доказывается так же, как предыдущая теорема о том, что пространство является прямой суммой всех его корневых подпространств. Докажем, что $p_1(f)V_1 = \{0\}$. Возьмем произвольный вектор $v_1 \in V_1$, тогда существует такой вектор $w \in W$, что $v_1 = p_1(f)w$, тогда $p_1(f)v_1 = p_1(f)p_2(f)w = p(f)w = 0$, т.к. $p(t)$ аннулирует f. Аналогично доказывается, что $p_2(f)V_2 = \{0\}$. □

Применяя эту лемму несколько раз, получаем аналогичное утверждение для произвольного числа сомножителей.

Следствие 2.10.3 *Пусть $p(t)$ — аннулирующий многочлен для оператора $f : W \to W$ и пусть его можно разложить в произведение m взаимно простых многочленов $p(t) = p_1(t) \cdot \ldots \cdot p_m(t)$, то пространство W является прямой суммой m подпространств, $W = V_1 \oplus \ldots \oplus V_m$, каждое из которых инвариантно и $p_i(t)$ аннулирует сужение f на V_i, т.е. $p_i(f)V_i = \{0\}$, $i = 1, \ldots, m$.*

2.11 Жорданова клетка. Жорданова нормальная форма оператора

Определение 2.11.1 Матрица $J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & \ldots & 0 & 1 \end{pmatrix}$ называется *жордановой клеткой* размера n с собственным значением λ. Блочно-диагональная матрица с жордановыми клетками на главной диагонали называется жордановой матрицей (жордановой формой).

Теорема 2.11.2 (Жордана о приведении матрицы оператора к нормальной форме) *Для любого оператора f существует базис, в котором его матрица A_f будет жордановой. Такая матрица единственна с точностью до перестановки блоков (клеток).*

Доказательство.

1) Существование. Если в каждом из корневых подпространств V_{λ_i} оператора f выбрать базис, а потом из них составить базис всего пространства (это будет просто объединение всех базисов корневых подпространств), то матрица оператора в этом базисе будет блочно-диагональной, $A_f = \begin{pmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_k \end{pmatrix}$, где матрица A_i, $i = 1, \ldots, k$, — это матрица ограничения $f|_{V_{\lambda_i}}$ оператора f на подпространство V_{λ_i}. Т.к. каждая из матриц A_i имеет только одно собственное значение λ_i, то оператор $f|_{V_{\lambda_i}} - \lambda_i \cdot id : V_{\lambda_i} \to V_{\lambda_i}$ (и его матрица A_i) будет нильпотентным, и, согласно теореме о нильпотентных операторах, в V_{λ_i} существует базис, в котором матрица оператора $f|_{V_{\lambda_i}} - \lambda_i \cdot id$ имеет вид $\begin{pmatrix} J_{n_1}(0) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & J_{n_p}(0) \end{pmatrix}$. Тогда матрица оператора $f|_{V_{\lambda_i}}$ имеет аналогичный вид $A_i = \begin{pmatrix} J_{n_1}(\lambda_i) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & J_{n_p}(\lambda_i) \end{pmatrix}$, но уже с числами λ_i вместо нулей на диагонали. Следовательно, матрица A_f состоит из жордановых клеток.

2) Единственность. Надо показать, что каким бы способом мы не привели бы матрицу оператора f к жордановой нормальной форме, количество жордановых клеток фиксированной размерности с собственным значением λ одно и то же. Для этого зафиксируем собственное значение λ_i и посчитаем количество клеток, ему отвечающих. Обозначим через $V^{(i)}$ подпространство, отвечающее блоку A_i, т.е. всем жордановым клеткам с собственным значением λ_i. Тогда $V^{(i)} \subseteq V_{\lambda_i}$ и, следовательно, $\dim V^{(i)} \leqslant \dim V_{\lambda_i}$, но, более того, $W = V^{(1)} \oplus \ldots \oplus V^{(k)}$. Но мы также знаем, что $W = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_k}$, значит $\dim V^{(1)} + \ldots + \dim V^{(k)} = \dim V_{\lambda_1} + \ldots + \dim V_{\lambda_k}$, поэтому $\dim V^{(i)} = \dim V_{\lambda_i}$, значит, $V^{(i)} = V_{\lambda_i}$ для $i = 1, \ldots, k$. Мы получили, что при произвольном разбиении на клетки, подпространства, отвечающие клеткам с собственным значением λ_i одинаковы и равны V_{λ_i}. Таким образом, разбиение матрицы оператора f на блоки, отвечающие различным собственным значениям, единственно.

Рассмотрим теперь блок A_i в матрице A_f, соответствующий собственному значению λ_i. Тогда матрица $A_i - \lambda_i E$ будет нильпотентной, а она, как мы уже знаем, имеет единственное разложение на жордановы клетки с собственным значением 0. Значит, матрица A_i имеет единственное разложение на жордановы клетки с собственным значением λ_i, поэтому вся матрица также допускает единственное (с точночтью до перестановок) разложение на жордановы клетки, т.к. разбиение на матрицы единственно, и разбиение каждой из них на жордановы клетки также единственно.

Еще одно доказательство единственности жордановой нормальной формы матрицы оператора можно получить по аналогии с аналогичным утверждением для нильпотентного оператора, при этом заодно получатся формулы для количества клеток определенной размерности. Введем числа $r_k(\lambda) = \operatorname{rk}(f - \lambda \cdot id)^k$ и $N_k(\lambda)$ — количество жордановых клеток размерности k, отвечающих собственному значению λ. Тогда легко заметить, что при вычислении разности $r_{k+1}(\lambda) - r_k(\lambda)$ нужно учитывать только клетки, отвечающие собственному значению λ, поскольку ранги остальных клеток $J_{n_i}(\lambda_i - \lambda)$ не меняются при возведении таких клеток в любую степень (эти клетки невырождены). Поэтому $r_{k+1}(\lambda) - r_k(\lambda) = N_k(\lambda) + N_{k+1}(\lambda) + \ldots$, откуда (так же, как для нильпотентных операторов) следует, что $N_k(\lambda) = r_{k+1}(\lambda) + r_{r-1}(\lambda) - 2r_k(\lambda)$. □

2.12 Приложения теоремы Жордана

I. Задача о рекуррентных последовательностях. Пусть заданы первые k элементов $x_1, \ldots, x_k$ некоторой числовой последовательности, и пусть известно, что каждый элемент этой последовательности, начиная с номера $k+1$, линейно зависит от предыдущих k элементов по

формуле $x_{n+1} = a_1 x_{n-k+1} + a_2 x_{n-k+2} + \ldots + a_k x_n$, $n > k$. Тогда мы можем последовательно вычислять все элементы этой последовательности. Но как быть, если нам, например, нужно вычислить x_{2002}, не будем же мы вычислять по этой формуле все элементы подряд. Если посмотреть на правило, по которому вычисляется каждый следующий элемент, можно заметить, что $x_{n+1} = (a_1 \ldots a_k) \begin{pmatrix} x_{n-k+1} \\ \vdots \\ x_n \end{pmatrix}$, тогда видно, что столбцы $\begin{pmatrix} x_{n-k+1} \\ \vdots \\ x_n \end{pmatrix}$ и $\begin{pmatrix} x_{n-k+2} \\ \vdots \\ x_{n+1} \end{pmatrix}$ связаны между собой оператором,

$$\begin{pmatrix} x_{n-k+2} \\ \vdots \\ \vdots \\ x_{n+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & & \mathbf{0} \\ & \ddots & \ddots & \\ \mathbf{0} & & 0 & 1 \\ a_1 & a_2 & \ldots & a_k \end{pmatrix}}_{=A} \begin{pmatrix} x_{n-k+1} \\ \vdots \\ \vdots \\ x_n \end{pmatrix},$$

тогда $\begin{pmatrix} x_{n+1} \\ \vdots \\ x_{n+k} \end{pmatrix} = A \begin{pmatrix} x_n \\ \vdots \\ x_{n+k-1} \end{pmatrix} = \ldots = A^n \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$, поэтому для вычисления любого элемента последовательности нам достаточно уметь возводить матрицу A в степени. Рассмотрим жорданову форму A' матрицы A, пусть C — матрица перехода к жордановому базису, $A = CA'C^{-1}$, тогда $A^n = \underbrace{CA'C^{-1} \cdot CA'C^{-1} \cdot \ldots \cdot CA'C^{-1}}_{n \text{ раз}}$. А если жорданова форма диагональна, как, например, у рекуррентного соотношения $x_{n+1} = x_n + x_{n-1}$ (числа Фибоначчи), то она очень легко возводится в любую степень, и мы легко можем получить формулу общего члена этой последовательности.

II. Задача о вычислении функций от матриц. Если функция $f(x)$ достаточно гладкая, т.е. имеет достаточно много производных, то для нее можно написать формулу Тейлора, которая будет иметь достаточно много членов, $f(t) = f(\lambda) + \frac{f'(\lambda)}{1!}(t-\lambda) + \frac{f''(\lambda)}{2!}(t-\lambda)^2 + \ldots + \frac{f^{(m)}(\lambda)}{m!}(t-\lambda)^m + r_m$ (в качестве последнего слагаемого можно взять, например, остаточный член в форме Лагранжа). Если матрица A — жорданова клетка, $A = \begin{pmatrix} \lambda & 1 & & \mathbf{0} \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ \mathbf{0} & & & \lambda \end{pmatrix}$, то

$$f(A) = \begin{pmatrix} f(\lambda) & \frac{f'(\lambda)}{1!} & \ldots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ & \ddots & \ddots & \vdots \\ & & f(\lambda) & \frac{f'(\lambda)}{1!} \\ \mathbf{0} & & & f(\lambda) \end{pmatrix},$$

т.е. значение функции $f(A)$ определяется только значением функции $f(t)$ и ее $n-1$ производной в точке $t = \lambda$, а все производные более высоких порядков (т.е. все последующие слагаемые формулы Тейлора) дают нулевой вклад. То есть, мы можем взять формулу Тейлора для этой функции, обрубить ее на $n-1$-й производной, и мы получим многочлен $p(t)$, причем $p(A) = f(A)$, а вычислять значение многочлена от матрицы мы умеем. Если матрица произвольна, то ее нужно привести к жордановой форме, $A' = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}$, где $A_1, \ldots, A_m$ — жордановые клетки. Т.к. $f(A') = \begin{pmatrix} f(A_1) & & 0 \\ & \ddots & \\ 0 & & f(A_m) \end{pmatrix}$ и $f(A) = Cf(A')C^{-1}$, то формулу Тейлора нам достаточно обрубить на $k-1$-й производной, где k — максимальный размер жордановой клетки в жордановой форме матрицы A, тогда мы получим такой многочлен $p(t)$, что $p(A) = f(A)$. Этот многочлен называется интерполяционным.

Вопрос: Почему определение $f(A)$ не зависит от способа приведения к жордановой форме ?

2.13 Овеществление и компексификация

Овеществление

Определение 2.13.1 Пусть V — векторное пространство над полем комплексных чисел $\mathbb{C}$. Рассмотрим пространство $V_{\mathbb{R}}$, состоящее из тех же векторов, что и V, только вместо операции умножения на все комплексные числа мы ограничимся умножением только на вещественные числа. Тогда $V_{\mathbb{R}}$ будет линейным пространством над полем вещественных чисел $\mathbb{R}$, оно называется *овеществлением* пространства V.

Пусть $e_1, \dots, e_n$ — базис в пространстве V, тогда он не будет базисом пространства $V_{\mathbb{R}}$, так как не все вектора являются их линейными комбинациями с вещественными числами, а на комплексные числа мы больше не можем умножать. Базисом в $V_{\mathbb{R}}$ будут вектора $e_1, \dots, e_n, ie_1, \dots, ie_n$ (проверьте), следовательно $\dim_{\mathbb{R}} V_{\mathbb{R}} = 2 \dim_{\mathbb{C}} V$ (индекс у dim непоминает, над каким полем мы рассматриваем размерность пространства).

Определение 2.13.2 Пусть дан оператор $f : V \to V$, тогда этот оператор, рассматриваемый на пространстве $V_{\mathbb{R}}$, называется *овеществлением оператора* f и обозначается $f_{\mathbb{R}}$.

Посмотрим, как связаны матрицы операторов f и $f_{\mathbb{R}}$. Пусть в базисе $e_1, \dots, e_n$ пространства V $f(e_k) = c_k^j e_j$. Матрицу $A_f = (c_k^j)$ оператора f можно разложить на вещественную и чисто мнимую часть, т.к. ее элементы — это комплексные числа, т.е. $A_f = A + iB$, где $A = (\operatorname{Re} c_k^j)$, $B = (\operatorname{Im} c_k^j)$. Тогда матрица оператора $f_{\mathbb{R}}$ в базисе $e_1, \dots, e_n, ie_1, \dots, ie_n$ будет иметь вид $A_{f_{\mathbb{R}}} = \left(\begin{array}{c|c} A & -B \\ \hline B & A \end{array}\right)$. Посчитаем $\det A_{f_{\mathbb{R}}}$, для чего сделаем следующие элементарные преобразования над строками и столбцами матрицы $A_{f_{\mathbb{R}}}$:

$$\begin{pmatrix} A & -B \\ B & A \end{pmatrix} \to \begin{pmatrix} A - iB & -B - iA \\ B & A \end{pmatrix} \to$$

$$\to \begin{pmatrix} A - iB & -B - iA + i(A - iB) \\ B & A + iB \end{pmatrix} = \begin{pmatrix} A - iB & 0 \\ B & A + iB \end{pmatrix},$$

поэтому

$$\det A_{f_{\mathbb{R}}} = \det(A - iB) \det(A + iB) = \overline{\det(A_f)} \cdot \det A_f = |\det A_f|^2.$$

Комплексная структура

Определение 2.13.3 Пусть V — векторное пространство над $\mathbb{R}$. *Комплексной структурой* на V называется такой линейный оператор $j : V \to V$, что $j^2 = -id$.

Тогда пространство V можно рассматривать как векторное пространство над $\mathbb{C}$, так как на V можно ввести операцию умножения на комплексные числа: $(a + ib)v := av + bj(v)$. То, что это определение корректно (свойства v-viii определения векторного пространства), проверяется тривиально.

Лемма 2.13.4 *Пусть j — комплексная структура на вещественном векторном пространстве V. Тогда*

1) $\dim V$ четна;

2) в подходящем базисе матрица оператора j имеет вид $A_j = \left(\begin{array}{c|c} 0 & -E \\ \hline E & 0 \end{array}\right)$.

Доказательство.

1. Обозначим через $\widetilde{V}$ пространство V, рассматриваемое как комплексное (с помошью комплексной структуры j). Размерность пространства $\widetilde{V}$ конечна, т.к. по базису пространства V можно разложить любой вектор (возможно неоднозначно). Пусть $e_1, \dots, e_n$ — базис в $\widetilde{V}$, тогда $e_1, \dots, e_n, e_{n+1} = j(e_1), \dots, e_{2n} = j(e_n)$ будет базисом в V, следовательно, $\dim V = 2n$.

2. Т.к. $j(e_i) = e_{n+i}$ и $j(e_{n+i}) = -e_i$ для $i = 1, \dots, n$, то в этом базисе матрица оператора имеет указанный вид. □

Комплексификация

Определение 2.13.5 Пусть V — векторное пространство над полем $\mathbb{R}$. Рассмотрим пространство $V_{\mathbb{C}} = V \oplus V = \{(a,b) : a,b \in V\}$ и определим комплексную структуру следующим образом: $j(a,b) := (-b,a)$ (нетрудно убедиться, что $j^2 = -id$). Пространство с такой комплексной структурой называется *комплексификацией* пространства V.

Покажем, что $\dim_{\mathbb{C}} V_{\mathbb{C}} = \dim_{\mathbb{R}} V$. Если $e_1, \ldots, e_n$ — базис в пространстве V, то $(e_1,0), \ldots, (e_n,0)$ будет базисом в пространстве $V_{\mathbb{C}}$. Действительно, поскольку $j(e_i,0) = (0,e_i)$, то умножением на мнимую единицу мы можем получить вектора $(0,e_1), \ldots, (0,e_n)$ и, следовательно, любой вектор (a,b), где $a,b \in V$.

Определение 2.13.6 Если дан оператор $f : V \to V$, то оператор $f_{\mathbb{C}} : V_{\mathbb{C}} \to V_{\mathbb{C}}$, заданный формулой $f_{\mathbb{C}}(a,b) := (fa, fb)$, называется *комплексификацией* оператора f.

Легко убедиться, что так определенный $f_{\mathbb{C}}$ действительно будет линейным оператором. Если A_f — матрица оператора f в базисе $e_1, \ldots, e_n$, то эта же матрица будет матрицей оператора $f_{\mathbb{C}}$ в базисе $(e_1,0), \ldots, (e_n,0)$.

2.14 Два канонических изоморфизма

Посмотрим, что будет, если к векторному пространству применить комплексификацию, а потом овеществление, и наоборот.

Утверждение 2.14.1 *Векторные пространства $(V_{\mathbb{C}})_{\mathbb{R}}$ и $V \oplus V$ канонически изоморфны.*

Доказательство. По определению, $V_{\mathbb{C}}$ и $V \oplus V$ совпадают как множества, поэтому $(V_{\mathbb{C}})_{\mathbb{R}}$ и $V \oplus V$ совпадают как множества. Операция сложения при комплексификации и при овеществлении у нас не менялась, а к операции умножения на скаляры мы сначала добавили умножение на комплексные числа, а затем его запретили, следовательно операция умножения тоже не изменилась. Значит $(V_{\mathbb{C}})_{\mathbb{R}}$ и $V \oplus V$ совпадают как линейные пространства. □

Если применять сначала овеществление, а потом комплексификацию, результат будет более сложным и нам потребуется рассмотреть сопряженное комплексное пространство. Пусть V — векторное пространство над $\mathbb{C}$, тогда сопряженное комплексное пространство $\overline{V}$ равно V как множество, с такой же операцией сложения, а операция умножения определяется следующим образом: $\lambda \cdot v := \overline{\lambda} \cdot v \ \forall \lambda \in \mathbb{C}, v \in V$.

Утверждение 2.14.2 *Векторные пространства $(V_{\mathbb{R}})_{\mathbb{C}}$ и $V \oplus \overline{V}$ канонически изоморфны.*

Доказательство. У нас введены две комплексные структуры: первая определена умножением на мнимую единицу $i \in \mathbb{C}$ в V, а вторая появляется при комплексификации: $j : V_{\mathbb{R}} \oplus V_{\mathbb{R}} \to V_{\mathbb{R}} \oplus V_{\mathbb{R}}$, $j(a,b) = (-b,a)$. Легко убедиться, что оператор j перестановочен с умножением на i, следовательно, является линейным оператором над полем комплексных чисел. Т.к. $j^2 + id = 0$, то $t^2 + 1$ является аннулирующим многочленом для оператора j. Значит, этот оператор может иметь только два собственных значения, а именно корни $\pm i$. Рассмотрим два подпространства собственных векторов, отвечающих этим собственным значениям: V^+ и V^-. Докажем, что $(V_{\mathbb{R}})_{\mathbb{C}} = V^+ \oplus V^-$, а затем, что пространства V^+ и V, а также пространства V^- и $\overline{V}$ канонически изоморфны друг другу. Пусть $(a,b) \in V^+$. Тогда $j(a,b) = (-b,a) = i(a,b) = (ia, ib)$, следовательно, $a = ib$, $ia = -b$ и $V^+ = \{(a,-ia) : a \in V\}$. Аналогично получаем, что $V^- = \{(b,ib) : b \in V\}$. Очевидно, что $V^+ \cap V^- = \{0\}$. Возьмем произвольный вектор $(c,d) \in (V_{\mathbb{R}})_{\mathbb{C}}$. Тогда $(c,d) = \underbrace{\left(\frac{c+id}{2}, -i\frac{c+id}{2}\right)}_{\in V^+} + \underbrace{\left(\frac{c-id}{2}, i\frac{c-id}{2}\right)}_{\in V^-}$, следовательно, $(V_{\mathbb{R}})_{\mathbb{C}} = V^+ \oplus V^-$.

Равенства $V^+ = \{(a,-ia) : a \in V\}$ и $V^- = \{(b,ib) : b \in V\}$ позволяют отождествить эти пространства с V (и с $\overline{V}$) как множества, и даже как векторные пространства над полем $\mathbb{R}$. Такие отождествления задаются отображениями $\varphi_+ : V \to V^+$, $\varphi_+(a) = (a,-ia)$, и $\varphi_- : V \to V^-$, $\varphi_-(b) = (b,ib)$. Однако нам нужно построить канонический изоморфизм комплексных векторных

пространств. На V^+ комплексная структура (оператор j) совпадает с комплексной структурой на V (умножение на i), так как для любого вектора a из этого пространства имеем $j(\varphi_+(a)) = j(a,-ia) = i(a,-ia) = (ia,a) = \varphi_+(ia)$. На V^- комплексная структура (оператор j) не совпадает с комплексной структурой на V, но совпадает с комплексной структурой на $\overline{V}$ (умножение на $-i$), так как для любого вектора b из этого пространства имеем $j(\varphi_-(b)) = j(b,ib) = -i(b,ib) = (-ib,b) = \varphi_-(-ib)$. Таким образом, отображения $\varphi_\pm$ задают канонические изоморфизмы $V^+ \cong V$, $V^- \cong \overline{V}$. □

Также, как при действиях с комплексными числами, часто вместо пар $(a,b) \in V_{\mathbb{C}}$ мы будем писать $a+ib$.

2.15 Вещественная Жорданова нормальная форма

Лемма 2.15.1 *Если $f : V \to V$ — оператор в вещественном векторном пространстве и $\dim V \geqslant 1$, то в пространстве V существует либо одномерное, либо двумерное инвариантное подпространство.*

Доказательство. Если $\dim V = 1$, то утверждение леммы очевидно. Если $\dim V > 1$, то пусть $\lambda = \alpha + i\beta$ — собственное значение оператора $f_{\mathbb{C}}$. Тогда в $V_{\mathbb{C}}$ есть собственный вектор $a+ib$, $a,b \in V$, отвечающий собственному значению λ. Тогда

$$f_{\mathbb{C}}(a+ib) = (\alpha+i\beta)(a+ib) = \alpha a - \beta b + i(\alpha b + \beta a),$$

но с другой стороны $f_{\mathbb{C}}(a+ib) = f(a) + if(b)$, следовательно, $f(a) = \alpha a - \beta b$ и $f(b) = \alpha b + \beta a$. Т.к. α и β — это вещественные числа, то $f(a), f(b) \in \langle a,b\rangle$, значит, $\langle a,b\rangle$ — инвариантное подпространство. Очевидно, что оно либо одномерное, либо двумерное (на самом деле оно всегда будет получаться двумерным, если $\beta \neq 0$). □

Наряду с обычными жордановыми клетками $J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & \mathbf{0} \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & \lambda \end{pmatrix}$, $\lambda \in \mathbb{R}$, рассмотрим клетки вида $J_{2k}(\alpha,\beta) = \begin{pmatrix} \alpha & \beta & 1 & 0 & & & & \\ -\beta & \alpha & 0 & 1 & & & \mathbf{0} & \\ & & \alpha & \beta & \ddots & & & \\ & & -\beta & \alpha & & \ddots & & \\ & & & & \ddots & & 1 & 0 \\ & & & & & \ddots & 0 & 1 \\ & \mathbf{0} & & & & & \alpha & \beta \\ & & & & & & -\beta & \alpha \end{pmatrix}$, $\alpha,\beta \in \mathbb{R}$, т.е. клетки, в которых вместо λ стоят матрицы $\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$, а вместо единиц стоят единичные матрицы $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Теорема 2.15.2 (о вещественной жордановой нормальной форме) *Для любого оператора $f : W \to W$, где W — пространство над полем $\mathbb{R}$, существует базис, в котором матрица оператора состоит из клеток вида $J_k(\lambda)$ и $J_{2k}(\alpha,\beta)$.*

Отметим, что, в отличие от Теоремы Жордана, числа λ должны быть вещественными.

Доказательство. Пусть $p(t) = (t-\lambda_1)^{r_1} \cdot \ldots \cdot (t-\lambda_j)^{r_j} \cdot (t^2+a_1t+b_1)^{s_1} \cdot \ldots \cdot (t^2+a_lt+b_l)^{s_l}$ — минимальный многочлен для оператора f. Поскольку все сомножители взаимно просты, W есть прямая сумма инвариантных подпространств, $W = V_1^{(1)} \oplus \ldots \oplus V_j^{(1)} \oplus V_1^{(2)} \oplus \ldots \oplus V_l^{(2)}$, в которой подпространства первой группы отвечают степеням линейных сомножителей, а второй группы — квадратичных. Тогда матрица оператора f имеет вид $A_f =$

$$\left(\begin{array}{c|c} A_1^{(1)} & \\ \ddots & \mathbf{0} \\ A_j^{(1)} & \\ \hline & A_1^{(2)} \\ \mathbf{0} & \ddots \\ & A_l^{(2)} \end{array}\right),$$ где каждая клетка $A_i^{(*)}$ есть матрица ограничения оператора f на подпространство $V_i^{(*)}$. Рассмотрим клетку первого типа $A_i^{(1)}$. Ограничение оператора на это подпространство имеет единственное (и, притом, вещественное) собственное значение — λ_i, следовательно, такая клетка приводятся к жордановому виду (алебраическая замкнутость поля здесь не нужна). Теперь посмотрим на клетку второго типа $A_i^{(2)}$. Она отвечает сомножителю вида $(t^2+at+b)^s$ и является матрицей ограничения оператора f на подпространство $V = V_i^{(2)}$. Обозначим через $g : V \to V$ ограничение f на V и перейдем к пространству $\widetilde{V} = V_{\mathbb{C}}$ — комплексификации пространства V, и к комплексификации $g_{\mathbb{C}}$ оператора g. Т.к. аннулирующий многочлен $(t^2+at+b)^s = (t-\lambda)^s(t-\overline{\lambda})^s$ пространства $\widetilde{V}$ имеет два комплексных корня $\lambda, \overline{\lambda}$ (они будут комплексно сопряжены, поскольку коэффициенты a, b вещественны), то $\widetilde{V}$ является прямой суммой корневых подпространств, $\widetilde{V} = V_\lambda \oplus V_{\overline{\lambda}}$. Отметим, что в комплексифицированном пространстве $\widetilde{V}$, аналогично случаю поля комплексных чисел, существует вещественно линейный, но не линейный, оператор комплексного сопряжения, $(a,b) \mapsto (a,-b)$. Мы для краткости будем его обозначать чертой, так же, как и в поле $\mathbb{C}$. Тогда, по определению корневого пространства имеем

$$\begin{array}{rcl} v \in V_\lambda & \iff & \exists m \in \mathbb{N} : (g_{\mathbb{C}} - \lambda \cdot id)^m v = 0; \\ v \in V_{\overline{\lambda}} & \iff & \exists m \in \mathbb{N} : (g_{\mathbb{C}} - \overline{\lambda} \cdot id)^m v = 0. \end{array}$$

Т.к. $\overline{(g_{\mathbb{C}} - \lambda \cdot id)^m v} = (g_{\mathbb{C}} - \overline{\lambda} \cdot id)^m \overline{v}$, то $v \in V_\lambda \iff \overline{v} \in V_{\overline{\lambda}}$. Поэтому, если выбрать в V_λ жорданов базис $c_1, \ldots, c_n$, то $\overline{c}_1, \ldots, \overline{c}_n$ будет базисом в $V_{\overline{\lambda}}$, более того будем иметь, что при $i > 1$ вектор c_i (соотв. $\overline{c}_i$) под действием $g_{\mathbb{C}}$ переходит в вектор $\lambda c_i + c_{i-1}$ (соотв. в $\overline{\lambda}\overline{c}_i + \overline{c}_{i-1}$) и c_1 (соотв. $\overline{c}_1$) переходит в λc_1 (соотв. в $\overline{\lambda}\overline{c}_1$). Рассмотрим вектора $a_i = \frac{1}{2}(c_i + \overline{c}_i)$ и $b_i = \frac{1}{2i}(c_i - \overline{c}_i)$. легко видеть, что эти векторы вещественны (т.е. $\overline{a}_i = a_i$, $\overline{b}_i = b_i$), поэтому они лежат в $V \subset \widetilde{V}$. Т.к. $c_1, \ldots, c_n, \overline{c}_1, \ldots, \overline{c}_n$ — базис в комплексном пространстве $\widetilde{V}$, то вектора $a_1, \ldots, a_n, b_1, \ldots, b_n$ образуют базис в вещественном пространстве V. Пусть $\lambda = \alpha + i\beta$, где $\alpha, \beta \in \mathbb{R}$, тогда при $i > 1$

$$\begin{array}{rcl} g(a_i) & = & g_{\mathbb{C}}(\frac{1}{2}(c_i + \overline{c}_i)) = \frac{1}{2}(g_{\mathbb{C}}(c_i) + g_{\mathbb{C}}(\overline{c}_i)) = \\ & = & \frac{1}{2}(\alpha c_i + \alpha\overline{c}_i + i\beta c_i - i\beta\overline{c}_i + c_{i-1} + \overline{c}_{i-1}) = \\ & = & \alpha\frac{1}{2}(c_i + \overline{c}_i) + i\beta\frac{1}{2}(c_i - \overline{c}_i) + \frac{1}{2}(c_{i-1} + \overline{c}_{i-1}) = \\ & = & \alpha a_i - \beta b_i + a_{i-1}. \end{array}$$

Аналогично $g(b_i) = \beta a_i + \alpha b_i + b_{i-1}$, $g(a_1) = \alpha a_1 - \beta b_i$, $g(b_1) = \beta a_1 + \alpha b_1$. Следовательно в базисе $a_1, \ldots, a_n, b_1, \ldots, b_n$ матрица оператора g является клеткой второго типа. $\square$

3 Евклидовы и унитарные пространства

3.1 Нормированное векторное пространство

Определение 3.1.1 Пусть V — векторное пространство над полем $\mathbb{R}$ или $\mathbb{C}$. *Нормой* называется функция $p : V \to [0, \infty)$ со свойствами:
1) $p(\lambda v) = |\lambda| p(v)$ для любых $\lambda \in \mathbb{K}$, $v \in V$;
2) из равенства $p(v) = 0$ следует, что $v = 0$;
3) $p(v + w) \leqslant p(v) + p(w)$ для любых векторов $v, w \in V$ (неравенство треугольника).

Обозначается норма как $p(v) = \|v\|$.

Пример.
На плоскости можно ввести много норм. Если x, y — координаты на плоскости, и $v = (x, y)$, то норму можно ввести, например, одной из следующих формул: $\|v\| = \begin{cases} \max(|x|, |y|) \\ \sqrt{x^2 + y^2} \\ |x| + |y| \end{cases}$.

Пример.
На пространстве вещественных квадратных матриц порядка n можно ввести норму по формуле $\|(a_{ij})\| = \sqrt{\sum_{i,j=1}^{n} a_{ij}^2}$.

Определение 3.1.2 *Единичным шаром* называется множество векторов, норма которых не превосходит единицу, т.е. $\{v \in V : \|v\| \leqslant 1\}$.

Пример:
В нашем примере норм на плоскости единичными шарами будут: квадрат, ограниченный прямыми $x = \pm 1$, $y = \pm 1$ в случае первой нормы; круг радиуса 1 с центром в начале координат в случае второй нормы; квадрат, ограниченный прямыми $y = \pm x \pm 1$ в случае третей нормы.

Определение 3.1.3 Множество $M \subset V$ называется *выпуклым*, если для любых двух векторов (точек) $a, b \in M$ и для любого числа $t \in [0, 1]$ вектор (точка) $t \cdot a + (1 - t) \cdot b \in M$.

Другими словами, если какие-то две точки принадлежат множеству M, то и весь отрезок, их соединяющий, должен принадлежать M.

Лемма 3.1.4 *Единичный шар является выпуклым множеством.*

Доказательство. Для любых двух точек v, w единичного шара, имеем $\|v\| \leqslant 1$, $\|w\| \leqslant 1$, тогда

$$\|tv + (1 - t)w\| \leqslant \|tv\| + \|(1 - t)w\| = |t| \, \|v\| + |1 - t| \, \|w\| \leqslant t + (1 - t) = 1,$$

т.е. точка $tv + (1 - t)w$ тоже принадлежит единичному шару для всех $0 \leqslant t \leqslant 1$. □

Разберем отдельно нормы на одномерных вещественных и комплексных пространствах.
Пусть V — одномерное вещественное векторное пространство. Возьмем произвольный вектор $v \neq 0$ в этом пространстве, тогда найдет такое число α, что $\|\alpha v\| = 1$, возьмем вектор $e = \alpha v$ в качестве базиса. Тогда единичный шар — это множество $\{w \in V : w = \lambda e, |\lambda| \leqslant 1\}$, т.е. симметричный относительно начала координат отрезок на прямой. Аналогично, если V — одномерное комплексное пространство, то единичным шаром будет круг на комплексной плоскости с центром в начале координат.

Определение 3.1.5 Множество $M \subset V$ называется центрально-симметричным, если его пересечение с любым одномерным векторным подпространством имеет один из вышеописанных видов (центрально-симметричный отрезок для вещественного случая и круг для комплексного).

Пример:
Очевидно, что единичный шар является центрально-симметричным множеством.

Теорема 3.1.6 *Пусть M — центрально-симметричное выпуклое множество в линейном пространстве V, тогда на пространстве V существует такая норма, для которой множество M является единичным шаром.*

Доказательство. Пусть $L \subset V$ — произвольное одномерное подпространство, тогда $L \cap M$ устроено так же, как и единичные шары в одномерных пространствах, т.е. $L \cap M = \{w = \lambda v : |\lambda| \leqslant 1\}$ для некоторого вектора $v \in L$. Определим норму на L следующим образом: произвольный вектор $a \in L$ запишем в виде $a = \alpha v$ и положим $\|a\| := |\alpha|$. Покажем, что так определить норму можно лишь единственным образом. Действительно, если множество $L \cap M$ задается каким-нибудь другим вектором v', т.е. $L \cap M = \{w = \mu v' : |\mu| \leqslant 1\}$. Тогда легко видеть, что $v' = \beta v$, где $|\beta| = 1$, поэтому замена вектора v на v' не меняет норму.

Возьмем произвольный вектор $a \in V$, он задает одномерное пространство $L = \langle a \rangle \subset V$, определим норму вектора a как описаано выше. Проверим, будут ли выполнены необходимые условия нормы. Первые два условия выполнены очевидным образом, а третье (неравенство треугольника) надо проверять. Для этого возьмем произвольные векторы $a, b \in V$. Векторы $\frac{a}{\|a\|}$, $\frac{b}{\|b\|}$ лежат в M, так как их норма равна 1. Поскольку множество M выпукло, то для любого $t \in [0,1]$ вектор $t\frac{a}{\|a\|}+(1-t)\frac{b}{\|b\|}$ также лежит в M. Возьмем $t = \frac{\|a\|}{\|a\|+\|b\|}$, тогда вектор $c = \frac{a}{\|a\|+\|b\|}+\frac{b}{\|a\|+\|b\|} = \frac{a+b}{\|a\|+\|b\|}$ лежит в M, следовательно $\|c\| \leqslant 1$. Таким образом, $\|c\| = \frac{a+b}{\|a\|+\|b\|} = \frac{\|a+b\|}{\|a\|+\|b\|} \leqslant 1$, отсюда уже следует неравенство треугольника $\|a+b\| \leqslant \|a\| + \|b\|$. □

3.2 Евклидовы и унитарные пространства

Определение 3.2.1 Линейное пространство V над полем $\mathbb{R}$ называется *евклидовым*, если на нем определена функция $f : V \times V \to \mathbb{R}$ (обозначается $f(a,b) = (a,b)$ и называется скалярным произведением), удовлетворяющая следующим свойствам:

1) линейность по второму аргументу: $(a, b+\lambda c) = (a,b) + \lambda(a,c)$ для любых $a, b, c \in V$, $\lambda \in \mathbb{R}$;

2) симметричность: $(b,a) = (a,b)$ для любых $a, b \in V$;

3) положительная определенность: $(a,a) \geqslant 0$ для любого $a \in V$, причем, если $(a,a) = 0$, то $a = 0$.

Видно, что благодаря второму свойству эта функция также будет линейной и по первому аргументу, т.е. она билинейна.

Определение 3.2.2 Линейное пространство V над полем $\mathbb{C}$ называется *унитарным* (или *эрмитовым*), если на нем определена функция $f : V \times V \to \mathbb{C}$ (обозначается $f(a,b) = (a,b)$, называется скалярным произведением), удовлетворяющая следующим свойствам:

1) линейность по второму аргументу: $(a, b+\lambda c) = (a,b) + \lambda(a,c)$ для любых $a, b, c \in V$, $\lambda \in \mathbb{C}$;

2) эрмитовость: $(b,a) = \overline{(a,b)}$ для любых $a, b \in V$;

3) положительная определенность: $(a,a) \geqslant 0$, причем, если $(a,a) = 0$, то $a = 0$. Т.к. $(a,a) = \overline{(a,a)}$ (свойство 2), то число (a,a) вещественно, и неравенство $(a,a) \geqslant 0$ имеет смысл.

Используя второе свойство можно получить, что $(a+\lambda b, c) = (a,c) + \overline{\lambda}(b,c)$, т.е. она полу(анти)линейна по первому аргументу, такая функция называется полуторалинейной.

Пример:

Пусть $V = \mathbb{K}[t]$ — пространство многочленов над полем $\mathbb{K}$ (это один из немногих случаев, когда конечномерность пространства не играет существенной роли и не обязательно ограничиваться многочленами фиксированной степени), возьмем два произвольных вещественных числа a, b, $a < b$. Определим скалярное произведение двух многочленов $p(t), q(t)$ по следующей формуле: $(p(t), q(t)) = \int_a^b p(t)q(t)\,dt$. То, что выполнены первые два условия скалярного произведения, сразу вытекает из свойств интеграла, проверим положительную определенность. Действительно, $(p(t), p(t)) = \int_a^b p^2(t)\,dt \geqslant 0$, интеграл от неотрицательной функции неотрицателен, причем, если $\int_a^b p^2(t)\,dt = 0$, то $p(t) \equiv 0$. Аналогично, для пространства многочленов над полем $\mathbb{C}$ скалярное произведение можно задать формулой $(p(t), q(t)) = \int_a^b \overline{p(t)}q(t)\,dt$.

Определение 3.2.3 Евклидовы и унитарные пространства называются *гильбертовыми.*

Определение 3.2.4 Длиной вектора a в гильбертовом пространстве называется число $|a| = \sqrt{(a,a)}$.

Это определение корректно, т.к. $(a,a) \geqslant 0$.

Утверждение 3.2.5 *Длина является нормой в гильбертовом пространстве.*

Доказательство. То, что выполнены первые два свойства нормы вытекает из свойств скалярного произведения, проверим третье свойство (неравенство треугольника), для этого нам понадобится неравенство Коши–Буняковского, $|(a,b)| \leqslant |a| \cdot |b|$, которое мы докажем после завершения доказательства утверждения.

Из неравенства Коши–Буняковского следует, что $(a,b) + (b,a) \leqslant 2|a| \cdot |b|$. Добавив к обеим частям неравенства $(a,a) + (b,b)$, получим

$$(a+b, a+b) = (a,a) + (b,b) + (a,b) + (b,a) \leqslant (a,a) + (b,b) + 2|a| \cdot |b| = (|a| + |b|)^2,$$

откуда следует неравенство треугольника. □

Лемма 3.2.6 (Неравенство Коши–Буняковского) *Для любых двух векторов a, b гильбертова пространства имеет место неравенство $|(a,b)| \leqslant |a| \cdot |b|$.*

Доказательство. Начнем с более простого — вещественного — случая. Рассмотрим скалярный квадрат $(a - \lambda b, a - \lambda b) \geqslant 0$, следовательно, для всех $\lambda \in \mathbb{R}$ имеем квадратичное неравенство $(a,a) - 2\lambda(a,b) + \lambda^2(b,b) \geqslant 0$, следовательно дискриминант этого квадратного трехчлена неположительный, т.е. $(a,b)^2 - (a,a)(b,b) \leqslant 0$. Перенося $(a,a)(b,b)$ в правую часть и извлекая корень, получаем искомое неравенство.

Перейдем теперь к комплексному случаю. Возьмем произвольное $\lambda \in \mathbb{C}$. Тогда $(a - \lambda b, a - \lambda b) = (a,a) - \overline{\lambda}(b,a) - \lambda(a,b) + |\lambda|^2(b,b) \geqslant 0$. Т.к. (a,b) — комплексное число, то для некоторого угла φ выполнено равенство $(a,b) = |(a,b)|e^{i\varphi}$. Ограничимся только теми λ, для которых $\lambda e^{i\varphi} = \mu \in \mathbb{R}$, тогда наш скалярный квадрат можно переписать в виде $(a - \lambda b, a - \lambda b) = (a,a) - 2\mu|(a,b)| + \mu^2(b,b)$. Далее, действуя, как в вещественном случае, получаем нужный результат. □

3.3 Процесс ортогонализации

Определение 3.3.1 Два вектора a, b называются ортогональными ($a \perp b$), если $(a,b) = 0$. Базис $e_1, \ldots, e_n$ называется ортонормированным, если $(e_i, e_j) = \delta_{ij}$, т.е. если векторы базиса попарно ортогональны и длина каждого из них равна 1.

Пусть $e_1, \ldots, e_n$ — некоторая линейно независимая система векторов гильбертова пространства. Обозначим через V_i линейную оболочку первых i векторов, $V_i = \langle e_1, \ldots, e_i \rangle$, и получим расширяющуюся цепочку подпространств $V_1 \subset V_2 \subset \ldots \subset V_n$.

Утверждение 3.3.2 *Существует такой набор попарно ортогональных векторов $a_1, \ldots, a_n$, что для каждого номера i линейная оболочка $\langle a_1, \ldots, a_i \rangle$ совпадает с V_i.*

Доказательство. (индукция по количеству векторов)

1) При $n = 1$ утверждение очевидно.

2) Пусть это утверждение выполнено для количества векторов, равного n, докажем его для $n+1$. Т.к. утверждение верно для n векторов, то мы можем считать, что векторы $a_1, \ldots, a_n$ с указанными свойствами уже построены. Построим вектор a_{n+1} в виде $a_{n+1} = e_{n+1} + \lambda_1 a_1 + \ldots + \lambda_n a_n$. Линейная оболочка векторов $a_1, \ldots, a_{n+1}$ совпадает с $\langle e_1, \ldots, e_{n+1} \rangle$ при любых λ_i, поэтому мы будем подбирать коэффициенты λ_i так, чтобы выполнялось условие $(a_{n+1}, a_i) = 0$ для всех $i = 1, \ldots, n$. Рассмотрим скалярное произведение $0 = (a_{n+1}, a_i) = (e_{n+1}, a_i) + \lambda_1(a_1, a_i) + \ldots + \lambda_n(a_n, a_i)$. Поскольку $(a_j, a_i) = 0$ при $j \neq i$ по предположению индукции, то $0 = (e_{n+1}, a_i) + \lambda_i(a_i, a_i)$, следовательно $\lambda_i = -\frac{(e_{n+1}, a_i)}{(a_i, a_i)}$ (знаменатель отличен от нуля). □

Таким образом, чтобы получить вектор a_{n+1}, надо из вектора e_{n+1} вычесть его ортогональные проекции на векторы $a_1, \ldots, a_n$. Этот метод ортогонализации называется методом Грама–Шмидта.

3.4 Ортогональное дополнение

Определение 3.4.1 Пусть $V \subset W$ — подпространство гильбертова пространства. *Ортогональным дополнением* $V^\perp$ к V в W называется множество, состоящее из векторов, ортогональных всем векторам из V, т.е. $V^\perp = \{v \in V : (v, w) = 0 \ \ \forall w \in W\}$.

Очевидным образом проверяется, что $V^\perp$ является подпространством, а не просто подмножеством.

Лемма 3.4.2 $W = V \oplus V^\perp$.

Доказательство. Пусть $e_1, \ldots, e_k$ — базис в V, дополним его до базиса всего пространства W векторами $e_{k+1}, \ldots, e_n$. Применив процесс ортогонализации Грама–Шмидта, получим ортогональный базис $a_1, \ldots, a_k, a_{k+1}, \ldots, a_n$ в W, причем его первая часть будет базисом в V, т.к. $\langle a_1, \ldots, a_k\rangle = \langle e_1, \ldots, e_k\rangle = V$. Покажем, что векторы $a_{k+1}, \ldots, a_n$ образуют базис в $V^\perp$. Пусть $v \in V^\perp$, $v = \nu_1 a_1 + \ldots + \nu_k a_k + \nu_{k+1} a_{k+1} + \ldots + \nu_n a_n$ — разложение вектора v по базису пространства W. Коэффициенты $\nu_1, \ldots, \nu_k$ должны быть нулевыми, так как иначе вектор v не был бы ортогонален всем векторам $a_1, \ldots, a_k$. Верно и обратное: если первые k координат какого-то вектора в базисе $a_1, \ldots, a_n$ равны нулю, то этот вектор принадлежит $V^\perp$. Следовательно, произвольный вектор $w \in W$ может быть представлен в виде суммы двух слагаемых — из V и из $V^\perp$: $w = \underbrace{\lambda_1 a_1 + \ldots + \lambda_k a_k}_{\in V} + \underbrace{\lambda_{k+1} a_{k+1} + \ldots + \lambda_n a_n}_{\in V^\perp}$, т.е. $W = V + V^\perp$.

Докажем, что эта сумма прямая. Возьмем произвольный вектор $v \in V \cap V^\perp$. Т.к. $v \in V^\perp$, то $(v, w) = 0$ для любого вектора $w \in V$. Поскольку $v \in V$, мы можем в качестве w взять сам вектор v, тогда $(v, v) = 0$, значит, $v = 0$. Следовательно, пересечение состоит только из нулевого вектора, и сумма — прямая. □

Из разложения в прямую сумму $W = V \oplus V^\perp$ следует, что любой вектор $a \in W$ можно единственным способом представить в виде $a = a_0 + a_\perp$, где $a_0 \in V$, $a_\perp \in V^\perp$. Вектор a_0 называется (ортогональной) проекцией вектора a на подпространство V, а вектор $a_\perp$ называется ортогональной составляющей вектора a.

Пусть имеется другое разложение вектора a: $a = a_1 + a_2$, где $a_1 \in V$, (а на a_2 дополнительных условий нет), тогда имеет место

Утверждение 3.4.3 $|a_2| \geqslant |a_\perp|$.

Доказательство. Обозначим $a_0 - a_1 = b \in V$, тогда $a = a_1 + a_2 = a_0 - (a_0 - a_1) + a_2 = a_0 + (a_2 - b)$, следовательно, $a_2 - b = a_\perp$, т.е. $a_2 = a_\perp + b$. Тогда $(a_2, a_2) = (a_\perp, a_\perp) + 2\underbrace{(a_\perp, b)}_{=0} + (b, b) \geqslant (a_\perp, a_\perp)$, откуда получаем $|a_2| \geqslant |a_\perp|$. □

Определение 3.4.4 Углом между двумя ненулевыми векторами в евклидовом пространстве называется величина $\widehat{(a, b)} := \arccos \frac{(a,b)}{|a| \cdot |b|}$.

Видно, что это определение не противоречит здравому смыслу, т.е. угол равен нулю тогда и только тогда, когда вектора коллинеарны, и угол — прямой тогда и только тогда, когда вектора ортогональны.

Утверждение 3.4.5 *(обозначения те же, что и в предыдущем утверждении)* $\widehat{(a, a_0)} \leqslant \widehat{(a, a_1)}$.

Доказательство. Очевидно, что если $\lambda > 0$, то $\widehat{(a, b)} = \widehat{(a, \lambda b)}$, т.е. вектор b можно заменить на любой другой, коллинеарный ему (и направленный в ту же сторону). Найдем такое число λ, чтобы вектор λa_1 был бы ортогонален вектору $b = a - \lambda a_1$:

$$(\lambda a_1, a - \lambda a_1) = 0 \Leftrightarrow (a_1, a - \lambda a_1) = 0 \Leftrightarrow (a_1, a) - \lambda \underbrace{(a_1, a_1)}_{\neq 0} = 0 \Leftrightarrow \lambda = \frac{(a_1, a)}{(a_1, a_1)}.$$

Тогда $(\widehat{a,a_1}) = (\widehat{a,\lambda a_1})$, $\cos(\widehat{a,a_0}) = \frac{(a,a_0)}{|a|\cdot|a_0|} = \frac{|a_0|}{|a|}$, $\cos(\widehat{a,\lambda a_1}) = \frac{(a,\lambda a_1)}{|a|\cdot|\lambda a_1|} = \frac{(a-\lambda a_1,\lambda a_1)+(\lambda a_1,\lambda a_1)}{|a|\cdot|\lambda a_1|} = \frac{|\lambda a_1|}{|a|}$. Следовательно,

$$\sin^2(\widehat{a,a_0}) = 1 - \frac{(a_0,a_0)}{(a,a)} = \frac{(a_\perp,a_\perp)}{(a,a)}, \quad \sin^2(\widehat{a,\lambda a_1}) = 1 - \frac{(\lambda a_1,\lambda a_1)}{(a,a)} = \frac{(b,b)}{(a,a)}.$$

Т.к. $(a_\perp,a_\perp) \leqslant (b,b)$, то $\sin^2(\widehat{a,a_0}) \leqslant \sin^2(\widehat{a,\lambda a_1})$, и $\sin(\widehat{a,a_0}) \leqslant \sin(\widehat{a,\lambda a_1})$. Т.к. оба угла принадлежат $[0,\pi/2]$ (их косинусы неотрицательны), получаем $(\widehat{a,a_0}) \leqslant (\widehat{a,\lambda a_1}) = (\widehat{a,a_1})$. □

Определение 3.4.6 Расстоянием $d(a,V)$ от вектора a до подпространства V называется наименьшее из всех возможных длин векторов, соединяющих векторы (точки) подпространства V с данным вектором, т.е. $d(a,V) := \min_{a_1\in V}|a-a_1|$. Углом $(\widehat{a,V})$ между вектором a и подпространством V называется наименьший из всех углов между вектором a и произвольным вектором $a_1 \in V$, т.е. $(\widehat{a,V}) := \min_{a_1\in V}(\widehat{a,a_1})$.

Очевидно, что $d(a,V) = |a_\perp|$ — расстояние от вектора до подпространства равно длине ортогональной составляющей при проекции вектора на подпространство, а $(\widehat{a,V}) = (\widehat{a,a_0})$ — угол между вектором и подпространством равен углу между вектором и его проекцией на данное подпространство.

3.5 Метод наименьших квадратов

Допустим, что мы исследуем какое-нибудь природное явление и хотим описать его линейной формулой, т.е. мы предполагаем, что какая-то величина b линейно зависит от других — $a_1,\ldots,a_n$, и хотим получить эту зависимость $b = a_1x^1+\ldots+a_nx^n$, т.е. узнать неизвестные коэффициенты $x^1,\ldots,x^n$. Мы делаем m измерений (для точности берем $m > n$) и решаем систему уравнений $\begin{cases} a_1^1x^1+\ldots+a_n^1x^n &= b^1 \\ \ldots\quad\ldots\quad\ldots & \\ a_1^mx^1+\ldots+a_n^mx^n &= b^m \end{cases}$. Вообще говоря, эта переопределенная система не имеет решения. Поэтому нам надо найти наиболее приближенное решение $x^1,\ldots,x^n$ в том смысле, что отклонение значений b^j от $c^j = a_1^jx^1+\ldots+a_n^jx^n$ будет наименьшим. В качестве отклонения удобно рассмотреть корень из суммы квадратов отклонений координат $\sqrt{(b^1-c^1)^2+\ldots+(b^m-c^m)^2} = |b-c|$, что равно длине вектора $b-c$. Будем искать такое псевдо-решение.

Рассмотрим векторы

$$a_1 = \begin{pmatrix} a_1^1 \\ \vdots \\ a_1^m \end{pmatrix}, \quad \ldots, \quad a_n = \begin{pmatrix} a_n^1 \\ \vdots \\ a_n^m \end{pmatrix}, \quad b = \begin{pmatrix} b^1 \\ \vdots \\ b^m \end{pmatrix}.$$

Пусть $V = \langle a_1,\ldots,a_n\rangle$. Обычно m намного больше n, и векторы $a_1,\ldots,a_n$ линейно независимы. Если они все-таки линейно зависимы, следует отбросить какое-то их количество, чтобы оставшиеся образовали базис подпространства V. Будем считать, что это уже сделано, и векторы $a_1,\ldots,a_n$ линейно независимы.

Спроектируем вектор b на подпространство V. Получим разложение $b = b_0 + b_\perp$, и, как мы доказали ранее, $|b_\perp| = |b-b_0|$ будет наименьшей длиной векторов, соединяющих b с V, т.е. b_0 определяет искомое максимально приближенное псевдо-решение. Чтобы найти его, разложим вектор b_0 по базису подпространства V, $b_0 = x^1a_1+\ldots+x^na_n$. Тогда, взяв скалярные произведения с векторами базиса, получаем следующие равенства:

$$\begin{array}{rcl} (a_1, x^1a_1+\ldots+x^na_n) &=& (a_1,b) \\ \ldots\quad\ldots\quad\ldots\quad\ldots & & \\ (a_n, x^1a_1+\ldots+x^na_n) &=& (a_n,b), \end{array}$$

т.е. надо решить систему уравнений (уже квадратную):

$$\left\{\begin{array}{rcl} (a_1,a_1)x^1+\ldots+(a_1,a_n)x^n & = & (a_1,b) \\ \ldots\quad\ldots\quad\ldots & & \\ (a_n,a_1)x^1+\ldots+(a_n,a_n)x^n & = & (a_n,b). \end{array}\right.$$

Решив ее, найдем координааты $x^1,\ldots,x^n$ вектора b_0, которые и есть наше псевдо-решение.

Матрица этой системы уравнений

$$G=G(a_1,\ldots,a_n)=\begin{pmatrix} (a_1,a_1) & \ldots & (a_1,a_n) \\ \vdots & \ddots & \vdots \\ (a_n,a_1) & \ldots & (a_n,a_n) \end{pmatrix}$$

называется *матрицей Грама*. Далее мы убедимся, что ее определитель отличен от нуля, когда векторы $a_1,\ldots,a_n$ линейно независимы.

Параллелепипеды. Матрица Грама

Определение 3.5.1 Пусть $a_1,\ldots,a_n$ — система векторов в векторном пространстве V. *Параллелепипедом*, натянутым на векторы $a_1,\ldots,a_n$ называется множество векторов (точек) $\Pi(a_1,\ldots,a_n)=\{c\in V : c=x^1a_1+\ldots+x^na_n,\ \ 0\leqslant x^1,\ldots,x^n\leqslant 1\}$.

Определение 3.5.2 Определим *n-мерный объем* Vol_n параллелепипеда $\Pi(a_1,\ldots,a_n)$ индуктивно:

1) одномерный объем $\mathrm{Vol}_1\,\Pi(a_1):=|a_1|$ — это длина вектора;

2)
$$\mathrm{Vol}_n\,\Pi(a_1,\ldots,a_n):=\mathrm{Vol}_{n-1}\,\Pi(a_1,\ldots,a_{n-1})\cdot d(a_n,\langle a_1,\ldots,a_{n-1}\rangle).$$

Очевидно, что объем есть неотрицательная величина. Корректность этого определения, т.е. независимость объема от порядка векторов при индуктивном переходе, вытекает из следующей теоремы.

Теорема 3.5.3 $(\mathrm{Vol}_n\,\Pi(a_1,\ldots,a_n))^2=\det G(a_1,\ldots,a_n)$.

Доказательство. (по индукции)

1) При $n=1$, очевидно, $|a_1|^2=(a_1,a_1)$.

2) Пусть утверждение верно для размерности $n-1$, докажем его для размерности n. Спроектируем вектор a_n на линейную оболочку векторов $a_1,\ldots,a_{n-1}$: $a_n=(a_n)_0+(a_n)_\perp=\lambda_1a_1+\ldots+\lambda_{n-1}a_{n-1}+b$, где $b=(a_n)_\perp$ (т.е. b ортогонален векторам $a_1,\ldots,a_{n-1}$) и $|b|=d(a_n,\langle a_1,\ldots,a_{n-1}\rangle)$. Имеем:

$$\det G(a_1,\dots,a_n)=\det\begin{pmatrix}(a_1,a_1)&\dots&(a_1,a_n)\\ \vdots&\ddots&\vdots\\ (a_n,a_1)&\dots&(a_n,a_n)\end{pmatrix}=$$

$$=\det\begin{pmatrix}(a_1,a_1)&\dots&(a_1,a_{n-1})&(a_1,\lambda_1a_1+\ldots+\lambda_{n-1}a_{n-1}+b)\\ \vdots&\ddots&\vdots&\vdots\\ (a_n,a_1)&\dots&(a_n,a_{n-1})&(a_n,\lambda_1a_1+\ldots+\lambda_{n-1}a_{n-1}+b)\end{pmatrix}=$$

$$=\det\begin{pmatrix}(a_1,a_1)&\dots&(a_1,a_{n-1})&\lambda_1(a_1,a_1)+\ldots+\lambda_{n-1}(a_1,a_{n-1})+(a_1,b)\\ \vdots&\ddots&\vdots&\vdots\\ (a_n,a_1)&\dots&(a_n,a_{n-1})&\lambda_1(a_n,a_1)+\ldots+\lambda_{n-1}(a_n,a_{n-1})+(a_n,b)\end{pmatrix}=$$

$$=\lambda_1\underbrace{\det\begin{pmatrix}(a_1,a_1)&\dots&(a_1,a_{n-1})&(a_1,a_1)\\ \vdots&\ddots&\vdots&\vdots\\ (a_n,a_1)&\dots&(a_n,a_{n-1})&(a_n,a_1)\end{pmatrix}}_{=0}+\dots$$

$$\dots+\lambda_{n-1}\underbrace{\det\begin{pmatrix}(a_1,a_1)&\dots&(a_1,a_{n-1})&(a_1,a_{n-1})\\ \vdots&\ddots&\vdots&\vdots\\ (a_n,a_1)&\dots&(a_n,a_{n-1})&(a_n,a_{n-1})\end{pmatrix}}_{=0}+$$

$$+\det\begin{pmatrix}(a_1,a_1)&\dots&(a_1,a_{n-1})&(a_1,b)\\ \vdots&\ddots&\vdots&\vdots\\ (a_n,a_1)&\dots&(a_n,a_{n-1})&(a_n,b)\end{pmatrix}=$$

$$=\det\begin{pmatrix}(a_1,a_1)&\dots&(a_1,a_{n-1})&\underbrace{(a_1,b)}_{=0}\\ \vdots&\ddots&\vdots&\vdots\\ (a_{n-1},a_1)&\dots&(a_{n-1},a_{n-1})&\underbrace{(a_{n-1},b)}_{=0}\\ (a_n,a_1)&\dots&(a_n,a_{n-1})&\underbrace{(a_n,b)}_{=(b,b)}\end{pmatrix}=$$

$$=\det\begin{pmatrix}(a_1,a_1)&\dots&(a_1,a_{n-1})\\ \vdots&\ddots&\vdots\\ (a_{n-1},a_1)&\dots&(a_{n-1},a_{n-1})\end{pmatrix}\cdot(b,b)=\det G(a_1,\dots,a_{n-1})\cdot|b|^2=$$

$$=(\mathrm{Vol}_{n-1}\Pi(a_1,\dots,a_{n-1}))^2\cdot|b|^2=(\mathrm{Vol}_n\Pi(a_1,\dots,a_n))^2.$$

Лемма 3.5.4 *Векторы $a_1,\dots,a_n$ линейно зависимы тогда и только тогда, когда* $\det G(a_1,\dots,a_n)=0$.

Доказательство. Если векторы $a_1,\dots,a_n$ линейно зависимы, то, без ограничения общности, можно считать, что $a_n=\lambda_1a_1+\ldots+\lambda_{n-1}a_{n-1}$. Тогда $\mathrm{Vol}_n\Pi(a_1,\dots,a_n)=\mathrm{Vol}_{n-1}\Pi(a_1,\dots,a_{n-1})\cdot\underbrace{d(a_n,\langle a_1,\dots,a_{n-1}\rangle)}_{=0}=0$, следовательно, по предыдущей теореме, $\det G(a_1,\dots,a_n)=0$.

Пусть теперь $\det G(a_1,\dots,a_n)=0$. Покажем линейную зависимость векторов $a_1,\dots,a_n$. Если $a_1=0$, то линейная зависимость очевидна. Предположим, что $a_1\neq0$, и рассмотрим последовательность чисел

$$0\neq\mathrm{Vol}_1\Pi(a_1),\mathrm{Vol}_2(a_1,a_2,),\dots,\mathrm{Vol}_n\Pi(a_1,\dots,a_n)=0.$$

Существует такой номер k, что $\mathrm{Vol}_{k-1}\Pi(a_1,\dots,a_{k-1})\neq0$, а $\mathrm{Vol}_k\Pi(a_1,\dots,a_k)$=0. Т.к. $\mathrm{Vol}_k\Pi(a_1,\dots,a_k)=\mathrm{Vol}_{k-1}\Pi(a_1,\dots,a_{k-1})\cdot d(a_k,\langle a_1,\dots,a_{k-1}\rangle)$, то $d(a_k,\langle a_1,\dots,a_{k-1}\rangle)=0$, т.е.

$a_k \in \langle a_1, \ldots, a_{k-1}\rangle$. Следовательно векторы $a_1, \ldots, a_k$, а ,значит, и $a_1, \ldots, a_n$, линейно зависимы. $\square$

Пусть нам дан линейный оператор $f : V \to V$, где V — гильбертово пространство, пусть также $a_1, \ldots, a_n$ — базис в пространстве V, а $b_i = f(a_i)$, $i = 1, \ldots, n$, — значения оператора на этих базисных векторах. Посмотрим, как связаны матрицы Грама $G = G(a_1, \ldots, a_n)$ и $G' = G(b_1, \ldots, b_n)$. Пусть матрица оператора f в базисе $a_1, \ldots, a_n$ имеет вид $C = A_f = \begin{pmatrix} c_1^1 & \ldots & c_n^1 \\ \vdots & \ddots & \vdots \\ c_1^n & \ldots & c_n^n \end{pmatrix}$, тогда $b_k = c_k^i a_i$ и элементы матрицы G' равны $g'_{ij} = (b_i, b_j) = (c_i^k a_k, c_j^l a_l) = \overline{c_i^k}(a_k, a_l)c_j^l = \overline{c_i^k} g_{kl} c_j^l$. Следовательно, $G' = \overline{C}^t G C$. Если базис ортонормирован, то $G = E = \begin{pmatrix} 1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & 1 \end{pmatrix}$ и $G' = \overline{C}^t C$.

Утверждение 3.5.5 *Произвольная квадратная матрица G является матрицей Грама для некоторого набора (линейно независимых) векторов тогда и только тогда, когда существует такая (невырожденная) матрица C, что $G = \overline{C}^t C$.*

Доказательство.
$\Rightarrow$: пусть $G = G(a_1, \ldots, a_n)$, выберем в пространстве ортонормированный базис, пусть f — линейный оператор, переводящий этот ортонормированный базис в векторы $a_1, \ldots, a_n$, тогда $G = \overline{C}^t C$, где $C = A_f$.

$\Leftarrow$: пусть $G = \overline{C}^t C$, тогда C можно считать матрицей некоторого оператора f в некотором ортонормированном базисе $a_1, \ldots, a_n$, и тогда G — это матрица Грама для векторов $f(a_1), \ldots, f(a_n)$. $\square$

3.6 Канонический изоморфизм $V \cong V'$

Лемма 3.6.1 *Пусть V' — двойственное пространство к пространству V. Если V — евклидово пространство, то существует канонический (т.е. не зависящий от базиса) изоморфизм между V и V'. Если V — эрмитово пространство, то существует канонический изоморфизм между сопряженным пространством $\overline{V}$ и V'.*

Доказательство.
Зафиксируем произвольный вектор $v \in V$ и рассмотрим функцию $l_v : V \to \mathbb{K}$, где $\mathbb{K}$ есть поле $\mathbb{R}$ или $\mathbb{C}$, определенную равенством $l_v(x) = (v, x)$. Линейность этой функции следует из линейности скалярного произведения.

Сначала рассмотрим случай, когда V — евклидово пространство. Рассмотрим отображение $f : V \to V'$, определенное формулой $f(v) := l_v$. Это — линейное отображение, т.к. $l_{v_1+\lambda v_2} = l_{v_1} + \lambda l_{v_2}$, более того, это отображение не зависит от выбора базиса. Проверим, будет ли f изоморфизмом, т.е. надо доказать, что $\operatorname{Ker} f = \{0\}$ (этого будет достаточно, так как размерности пространств V и V' совпадают). Возьмем произвольный вектор $w \in \operatorname{Ker} f$, тогда $f(w) = l_w = 0$, следовательно, $l_w(x) = (w, x) = 0$ для любого $x \in V$. Но тогда и $(w, w) = 0$, значит, $w = 0$, т.е. ядро отображения f нулевое.

Теперь рассмотрим случай, когда V — унитарное пространство. Напомним, что, по определению сопряженного пространства, умножение на комплексные числа устроено по правилу $\lambda \cdot v := \overline{\lambda} v$. Тогда отображение $f : \overline{V} \to V'$, $f(v) = l_v$, также будет линейным (над полем $\mathbb{C}$), потому что $l_{\lambda \cdot v} = l_{\overline{\lambda} v} = \lambda l_v$. Аналогично вещественному случаю получаем, что ядро f нулевое и f — изоморфизм. $\square$

Напомним, что если пространства канонически изоморфны, то их можно просто отождествить с помощью канонического изоморфизма, и в дальнейшем не различать.

4 Операторы в евклидовых и унитарных пространствах

4.1 Изоморфизмы гильбертовых пространств

Определение 4.1.1 Пусть V и W - гильбертовы пространства. Отображение $f : V \to W$ называется *изоморфизмом* гильбертовых пространств, если выполнены следующие свойства:
1) f — линейное отображение;
2) f сохраняет скалярное произведение, т.е. $(f(a), f(b)) = (a, b)$ для любых $a, b \in V$;
3) существует обратное отображение $f^{-1} : W \to V$, обладающее свойствами 1 и 2.

Утверждение 4.1.2 *Два гильбертовых пространства изоморфны (как гильбертовы пространства) тогда и только тогда, когда их размерности совпадают.*

Доказательство.
Если пространства V и W изоморфны как гильбертовы пространства, то они также изоморфны как векторные пространства, поэтому их размерности совпадают. Докажем обратное утверждение. Пусть $\dim V = \dim W = n$, $a_1, \ldots, a_n$ — ортонормированный базис в V, а $b_1, \ldots, b_n$ — ортонормированный базис в W. Отображение f достаточно определить на базисных векторах, а затем продолжить на все V по линейности. Пусть $f(a_i) = b_i$, $i = 1, \ldots, n$. Проверим условие 2). Возьмем два вектора $x, y \in V$, $x = x^i a_i$, $y = y^i a_i$. Тогда $f(x) = x^i b_i$, $f(y) = y^i b_i$ и $(f(x), f(y)) = \overline{x}^i (b_i, b_j) y^j = \overline{x}^i \delta_{ij} y^j = \overline{x}^i (a_i, a_j) y^j = (x, y)$. Обратное отображение, очевидно, существует и линейно, и также удовлетворяет свойству 2). □

Утверждение 4.1.3 *Пусть $f : V \to V$ — некоторое отображение (не обязательно линейное). Если f сохраняет скалярное произведение, то оно линейно.*

Доказательство. Возьмем ортонормированный базис $a_1, \ldots, a_n$ пространства V. Пусть $b_i = f(a_i)$, $i = 1, \ldots, n$. Проверим будут ли вектора образовывать базис (линейно независимы). Т.к. сохраняется скалярное произведение, то $(b_i, b_j) = (a_i, a_j) = \delta_{ij}$, и система векторов $b_1, \ldots, b_n$ ортонормирована. Проверим ее линейную независимость: если $\lambda_1 b_1 + \ldots + \lambda_n b_n = 0$, то, умножая скалярно на вектор b_k, получаем $(b_k, \lambda_1 b_1 + \ldots + \lambda_n b_n) = \lambda_k = 0$. Таким образом, векторы $b_1, \ldots, b_n$ образуют базис (вообще говоря, любая система попарно ортогональных векторов линейно независима).

Возьмем произвольный вектор $x = x^i a_i$, тогда $x^i = x^i (a_i, a_i) = (a_i, x)$, а, т.к. f сохраняет скалярное произведение, то $x^i = (a_i, x) = (f(a_i), f(x)) = (b_i, f(x))$, следовательно, $f(x) = x^i b_i$. Таким образом, $f(x)$ имеет те же координаты в базисе $b_1, \ldots, b_n$, что и x в базисе $a_1, \ldots, a_n$, следовательно, при отображении f сохраняются координаты. После этого замечания проверка линейности очевидна. □

Определение 4.1.4 Линейный оператор $f : V \to V$ называется *изометрией*, если $|f(a)| = |a|$ для любого вектора $a \in V$ (т.е. если он сохраняет длины векторов).

Утверждение 4.1.5 *Оператор f является изометрией тогда и только тогда, когда он сохраняет скалярное произведение.*

Доказательство. Если f сохраняет скалярное произведение, то он, в частности, сохраняет и длины векторов. Остается проверить обратное утверждение. Пусть f сохраняет длины векторов (т.е. изометрия).

1) В случае, когда пространство V евклидово, возьмем два произвольных вектора $a, b \in V$, тогда $(a + b, a + b) = (a, a) + (b, b) + 2(a, b)$, откуда $(a, b) = \frac{1}{2}((a + b, a + b) - (a, a) - (b, b))$, следовательно, оператор f сохраняет скалярное произведение.

2) Рассмотрим теперь случай, когда пространство V унитарно. Тогда будем иметь $(a + b, a + b) = (a, a) + (b, b) + (a, b) + \overline{(a, b)} = (a, a) + (b, b) + 2\,\mathrm{Re}(a, b)$, откуда $\mathrm{Re}(a, b) = \frac{1}{2}((a + b, a + b) - (a, a) - (b, b))$, следовательно, вещественная часть скалярного произведения сохраняется. Взяв вектор ib вместо b, получаем

$$\begin{aligned}(a + ib, a + ib) &= (a, a) + (ib, ib) + (a, ib) + (ib, a) = (a, a) + (b, b) + i(a, b) - i(b, a) = \\ &= (a, a) + (b, b) + i(a, b) - i\overline{(a, b)} = (a, a) + (b, b) - 2\,\mathrm{Im}(a, b),\end{aligned}$$

следовательно мнимая часть скалярного произведения также выражается через длины векторов, $\mathrm{Im}(a,b) = -\frac{1}{2}((a+ib, a+ib) - (a,a) - (b,b))$ и, значит, тоже сохраняется. □

Лемма 4.1.6 *Следующие три утверждения эквивалентны:*
1) оператор f сохраняет скалярное произведение;
2) оператор f переводит произвольный ортонормированный базис в ортонормированный;
3) в любом ортонормированном базисе матрица A_f оператора f обладает свойством $\overline{A}_f^t A_f = E$.

Доказательство.
1) $\Rightarrow$ 2) очевидно, т.к. $(f(a_i), f(a_j)) = (a_i, a_j) = \delta_{ij}$.
2) $\Rightarrow$ 1). Пусть $a_1, \ldots, a_n$ — ортонормированный базис, $b_i = f(a_i)$, и $b_1, \ldots, b_n$ — также ортонормированный базис. Возьмем произвольные векторы $x = x^i a_i$ и $y = y^i a_i$. Тогда $f(x) = x^i b_i$, $f(y) = y^i b_i$, и $(f(x), f(y)) = (x^i b_i, y^j b_j) = \overline{x}^i (b_i, b_j) y^j = \overline{x}^i (a_i, a_j) y^j = (x, y)$, значит, f сохраняет скалярное произведение.
2) $\Rightarrow$ 3). Возьмем ортонормированный базис $a_1, \ldots, a_n$, тогда $\underbrace{G(f(a_1), \ldots, f(a_n))}_{=E} = \overline{A}_f^t \underbrace{G(a_1, \ldots, a_n)}_{=E} A_f$, т.е. $\overline{A}_f^t A_f = E$.
3) $\Rightarrow$ 2). Для любого ортонормированного базиса $a_1, \ldots, a_n$ имеем $G(f(a_1), \ldots, f(a_n)) = \overline{A}_f^t \underbrace{G(a_1, \ldots, a_n)}_{=E} A_f$, значит, $\overline{A}_f^t A_f = E$, следовательно, векторы $f(a_1), \ldots, f(a_n)$ также ортонормированны. □

4.2 Ортогональные и унитарные операторы

Определение 4.2.1 Оператор, сохраняющий скалярное произведение в евклидовых пространствах называется *ортогональным*, в унитарных пространствах — *унитарным*.

Лемма 4.2.2 *Пусть оператор f действует в гильбертовом пространстве W, а $V \subset W$ — инвариантное подпространство. Если f сохраняет скалярное произведение, то $V^\perp$ тоже инвариантно.*

Доказательство. Возьмем произвольный вектор $a \in V^\perp$; надо доказать, что $f(a) \in V^\perp$, т.е., что $(f(a), v) = 0$ для любого $v \in V$. Мы знаем, что $f(V) \subseteq V$, но, поскольку ортонормированный базис переходит в ортонормированный базис, то $\dim f(V) = \dim V$, следовательно $f(V) = V$. Тогда найдется такой вектор $w \in V$, что $v = f(w)$, и тогда $(f(a), v) = (f(a), f(w)) = (a, w) = 0$, что и требовалось доказать. □

Лемма 4.2.3 *Если f — унитарный оператор, то все его собственные значения по модулю равны 1, если же f — ортогональный, то все его собственные значения равны ± 1.*

Доказательство. Пусть λ — собственное значение, тогда для собственного вектора v выполнено равенство $fv = \lambda v$. Т.к. оператор сохраняет скалярное произведение, то $(v, v) = (fv, fv) = (\lambda v, \lambda v) = \overline{\lambda}\lambda(v, v) = |\lambda|^2 (v, v)$, а т.к. $v \neq 0$, то $(v, v) \neq 0$, следовательно, $|\lambda|^2 = 1$, т.е. $|\lambda| = 1$. Если же оператор f ортогональный, то будем иметь $(v, v) = \lambda^2 (v, v)$, следовательно, $\lambda = \pm 1$. □

Лемма 4.2.4 *Если f — унитарный оператор, то его собственные вектора, отвечающие различным собственным значениям, взаимно ортогональны.*

Доказательство. Пусть $\lambda_1 \neq \lambda_2$, и $fv_1 = \lambda_1 v_1$, $fv_2 = \lambda_2 v_2$ (v_1 и v_2 — собственные вектора). Тогда $(v_1, v_2) = (fv_1, fv_2) = (\lambda_1 v_1, \lambda_2 v_2) = \overline{\lambda_1}\lambda_2 (v_1, v_2)$, следовательно, либо $(v_1, v_2) = 0$, т.е. $v_1 \perp v_2$, либо $\overline{\lambda_1}\lambda_2 = 1$. Но во втором случае, поскольку $|\lambda_1| = |\lambda_2| = 1$, имеем $\lambda_1^{-1} = \overline{\lambda_1}$, поэтому $\overline{\lambda_1}\lambda_2 = \lambda_1^{-1}\lambda_2 = 1$, откуда следует, что $\lambda_1 = \lambda_2$, что невозможно по предположению. Следовательно, векторы v_1 и v_2 ортогональны. □

Теорема 4.2.5 *1) Если $f : W \to W$ — унитарный оператор, то существует ортонормированный базис, в котором его матрица A_f диагональна, причем на диагонали стоят числа, по модулю равные 1.*

2) Если $f : W \to W$ — ортогональный оператор, то существует ортонормированный базис, в котором A_f имеет блочно-диагональный вид с блоками размера 1 и 2, причем одномерные блоки — это ± 1, а двумерные блоки имеют вид $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$ для некоторого угла φ.

3) Указанные канонические виды матриц унитарного и ортогонального оператора единственны с точностью до перестановки диагональных элементов и двумерных блоков.

Доказательство.

1) Пусть λ — собственное значение оператора f (оно существует, т.к. поле $\mathbb{C}$ алгебраически замкнуто) и v — собственный вектор, отвечающий этому значению. Тогда $V = \langle v \rangle$ — одномерное пространство, порожденное вектором v — будет инвариантным. Пользуясь этим замечанием, проведем теперь доказательство по индукции.

Если $\dim W = 1$, то утверждение теоремы очевидно.

Пусть теорема верна для случая $\dim W = n$, докажем ее для $\dim W = n+1$. Возьмем одномерное инвариантное подпространство V, порожденное собственным вектором, тогда $W = V \oplus V^\perp$, и матрица A_f имеет вид $A_f = \left(\begin{array}{c|c} \lambda & \\ \hline & A' \end{array}\right)$, где A' — матрица оператора $f|_{V^\perp}$. Ограничение $f|_{V^\perp}$ оператора f на $V^\perp$ также будет унитарным (так как f сохраняет скалярные произведения), следовательно, по предположению индукции матрицу A' можно представить в требуемом виде, но тогда и вся матрица будет представлена в таком виде. Поскольку на диагонали будут стоять собственные значения оператора f (и его ограничений), то они все по модулю равны 1.

2) Если у оператора f есть вещественные корни, то с ними можно поступить так же, как и в случае унитарного оператора. Мы знаем, что у оператора f есть либо одномерное, либо двумерное инвариантное подпространство V, из ограничения $f|_{V^\perp}$ мы также можем выделить либо одномерно, либо двумерное инвариантное подпространство и т.д. В итоге мы представим матрицу A_f в блочно-диагональном виде с блоками размера 1 или 2. Причем одномерные блоки — это не что иное, как собственные значения оператора, т.е. ± 1.

Рассмотрим теперь какой-нибудь двумерный блок и обозначим для краткости через g ограничение оператора f на соответствующее инвариантное двумерное подпространство V. У оператора g нет вещественных собственных значений, т.к. иначе V распадалось бы на два одномерных инвариантных подпространства. Пусть комплексные числа λ и $\overline{\lambda}$ — собственные значения оператора g (они сопряжены друг другу, поскольку коэффициенты характеристического уравнения вещественны). Комплексифицируем V, получим пространство $V_\mathbb{C}$, элементами которого являются векторы вида $a + ib$, где $a, b \in V$. В пространстве $V_\mathbb{C}$ есть собственный вектор оператора g, отвечающий значению λ, пусть это вектор $c \in V_\mathbb{C}$, тогда $g_\mathbb{C} c = \lambda c$. Пусть $\lambda = \alpha + i\beta$ и $c = a + ib$, где $\alpha, \beta \in \mathbb{R}, a, b \in V$. Тогда $g_\mathbb{C} c = ga + igb$, но, с другой стороны, $g_\mathbb{C} c = \lambda c = \alpha a - \beta b + i(\alpha b + \beta a)$, следовательно, имеем $ga = \alpha a - \beta b$ и $gb = \beta a + \alpha b$. Но это значит, что в базисе a, b матрица оператора g имеет вид $\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$. Т.к. $\alpha^2 + \beta^2 = |\lambda| = 1$, то существует угол φ, для которого $\alpha = \cos\varphi, \beta = -\cos\varphi$. Но тогда матрица оператора g будет иметь требуемый вид. Следовательно любую двумерную клетку можно привести к такому виду.

Осталось доказать, что требуемый вид матрица оператора g имеет в ортонормированном базисе. Для этого достаточно проверить, что $|a| = |b|$ и $(a, b) = 0$ (тогда базис $e_1 = \frac{a}{|a|}$, $e_2 = \frac{b}{|b|}$ ортонормированный, и матрица оператора g в нем имеет такой же вид, что и в базисе a, b). Имеем

$$(a, a) = (ga, ga) = (\alpha a - \beta b, \alpha a - \beta b) = \alpha^2 (a, a) + \beta^2 (b, b) - 2\alpha\beta (a, b),$$

поэтому

$$\underbrace{(1-\alpha^2)}_{=\beta^2}(a,a)-\beta^2(b,b)+2\alpha\beta(a,b)=0,$$

т.е.

$$\beta^2((a,a)-(b,b))+2\alpha\beta(a,b)=0. \tag{4}$$

Аналогично,

$$(a,b)=(ga,gb)=(\alpha a-\beta b,\beta a+\alpha b)=\alpha\beta(a,a)+\alpha^2(a,b)-\beta^2(a,b)-\alpha\beta(b,b),$$

поэтому

$$\alpha\beta((a,a)-(b,b))+\underbrace{(\alpha^2-\beta^2-1)}_{=-2\beta^2}(a,b)=0,$$

т.е.

$$\alpha\beta((a,a)-(b,b))-2\beta^2(a,b)=0. \tag{5}$$

Обозначим $(a,a)-(b,b)=x$, $(a,b)=y$. Равенства (4), (5) дают систему уравнений для определения x и y $\begin{cases}\beta^2x+2\alpha\beta y &= 0\\ \alpha\beta x-2\beta^2y &= 0\end{cases}$. Определитель этой системы $\begin{vmatrix}\beta^2 & 2\alpha\beta\\ \alpha\beta & -2\beta^2\end{vmatrix}=-2\beta^3-2\alpha^2\beta=-2\beta$. Поскольку $\beta\neq 0$ (иначе собственное значение λ было бы вещественным и нам не пришлось бы возиться с двумерными клетками), определитель системы не равен нулю, следовательно эта система имеет только одно решение, а именно нулевое, т.е. $x=(a,a)-(b,b)=0$ и $y=(a,b)=0$, откуда получаем, что длины векторов a и b совпадают, а сами эти векторы ортогональны.

3) В случае унитарного оператора единственность очевидна, т.к. на диагонали там стоят корни характеристического многочлена с учетом их кратности. В случае ортогонального оператора корни характеристического многочлена двумерной клетки являются комплексными корнями характеристического многочлена оператора, следовательно, двумерные клетки тоже определяются однозначно. □

4.3 Сопряженный оператор

Лемма 4.3.1 *Пусть даны два (не обязательно линейные) отображения f,g гильбертова пространства V в себя. Если для любых векторов $a,b\in V$ выполняется равенство $(fa,b)=(a,gb)$, то f и g — линейные операторы.*

Доказательство. Поскольку операторы f и g играют симметричную роль, достаточно доказать утверждение леммы для отображения f. Равенство $f(a_1+a_2)-f(a_1)-f(a_2)=0$ следует из того, что для любого вектора $b\in V$ верно равенство $(f(a_1+a_2)-f(a_1)-f(a_2),b)=(f(a_1+a_2),b)-(f(a_1),b)-(f(a_2),b)=(a_1+a_2,gb)-(a_1,gb)-(a_2,gb)=(0,gb)=0$. Аналогично проверяется, что $f(\lambda a)=\lambda f(a)$: $(f(\lambda a)-\lambda f(a),b)=(f(\lambda a),b)-(\lambda f(a),b)=(\lambda a,gb)-\overline{\lambda}(a,gb)=0$. □

Определение 4.3.2 Если для оператора $f:V\to V$ существует такой оператор g, что для любых векторов $a,b\in V$ выполняется равенство $(fa,b)=(a,gb)$, то g называется *сопряженным оператором* для f.

Докажем единственность сопряженного оператора. Допустим, что g_1 и g_2 — два сопряженных оператора для f. Тогда $(fa,b)=(a,g_1b)=(a,g_2b)$, т.е. $(a,g_1b-g_2b)=0$ для любого вектора a, следовательно при любом b имеем $g_1b-g_2b=0$, следовательно, $g_1=g_2$.

Сопряженный оператор обозначается $g=f^*$.

Лемма 4.3.3 *Если операторы f_1 и f_2 имеют сопряженные f_1^* и f_2^* соответственно, то оператор $g=f_1f_2$ также имеет сопряженный g^*, причем $g^*=f_2^*f_1^*$.*

Доказательство. $(f_1 f_2 a, b) = (f_2 a, f_1^* b) = (a, f_2^* f_1^* b)$. □

Лемма 4.3.4 *Если в ортонормированном базисе матрица оператора f равна A и существует сопряженный оператор f^*, то матрица этого оператора в том же базисе равна A^t (если пространство евклидово) или $\overline{A}^t$ (если пространство унитарно).*

Доказательство. Пусть матрица оператора f в ортонормированном базисе $e_1, \ldots, e_n$ есть $A = \begin{pmatrix} a_1^1 & \ldots & a_n^1 \\ \vdots & \ddots & \vdots \\ a_1^n & \ldots & a_n^n \end{pmatrix}$, а матрица оператора f^* (в том же базисе) есть $B = \begin{pmatrix} b_1^1 & \ldots & b_n^1 \\ \vdots & \ddots & \vdots \\ b_1^n & \ldots & b_n^n \end{pmatrix}$. Тогда $a_i^j = (e_j, a_i^k e_k) = (e_j, f e_i) = (f^* e_j, e_i) = (b_j^l e_l, e_i) = \overline{b}_j^i$, откуда получаем $A = \overline{B}^t$, или $B = \overline{A}^t$. □

Отметим, что условие ортонормированности базиса в лемме является существенным.

Лемма 4.3.5 *Для любого оператора f существует ему сопряженный f^*.*

Доказательство. Выберем ортонормированный базис $e_1, \ldots, e_n$, и пусть A_f — матрица оператора f в этом базисе. Тогда матрица $\overline{A}_f^t$ будет (в этом же базисе) матрицей некоторого оператора g. Пусть $A_f = \begin{pmatrix} c_1^1 & \ldots & c_n^1 \\ \vdots & \ddots & \vdots \\ c_1^n & \ldots & c_n^n \end{pmatrix}$. Возьмем произвольные векторы $a = a^i e_i$, $b = b^i e_i$. Тогда

$$fb = \begin{pmatrix} c_1^1 & \ldots & c_n^1 \\ \vdots & \ddots & \vdots \\ c_1^n & \ldots & c_n^n \end{pmatrix} \begin{pmatrix} b^1 \\ \vdots \\ b^n \end{pmatrix}; \qquad ga = (a^1 \ldots a^n) \begin{pmatrix} \overline{c}_1^1 & \ldots & \overline{c}_n^1 \\ \vdots & \ddots & \vdots \\ \overline{c}_1^n & \ldots & \overline{c}_n^n \end{pmatrix},$$

и легко видеть, что

$$(a, fb) = (\overline{a}^1 \ldots \overline{a}^n) \begin{pmatrix} c_1^1 & \ldots & c_n^1 \\ \vdots & \ddots & \vdots \\ c_1^n & \ldots & c_n^n \end{pmatrix} \begin{pmatrix} b^1 \\ \vdots \\ b^n \end{pmatrix} = \overline{(b, ga)} = (ga, b),$$

т.е. $g = f^*$. □

Следствие 4.3.6 *Если f^* — оператор, сопряженный к f, то $(f^*)^* = f$.*

Утверждение 4.3.7 *Если V — инвариантное подпространство относительно f, то $V^\perp$ — инвариантное подпространство относительно f^*.*

Доказательство. Возьмем произвольный вектор $v \in V^\perp$, тогда $\forall u \in V$ имеем $(f^* v, u) = (v, fu) = 0$, т.к. $v \in V^\perp$, а $fu \in V$. Следовательно, $f^* v \in V^\perp$ для любого $v \in V^\perp$. □

Определение 4.3.8 Оператор f называется *самосопряженным* (или *симметрическим*), если $f^* = f$. Оператор f называется *кососимметрическим*, если $f^* = -f$. Оператор f называется *нормальным*, если $f^* f = f f^*$.

Кроме того, в терминах сопряженного оператора можно дать другое определение ортогонального (унитарного) оператора: оператор f называется *ортогональным* (в евклидовом пространстве) или *унитарным* (в унитарном пространстве), если $f^* = f^{-1}$ (или, что то же самое, $f^* f = id$).

Заметим, что для матриц операторов будут выполнены эти же свойства, что и для операторов, т.е. матрица симметрического оператора является симметрической (в вещественном случае), матрица кососимметрического оператора является кососимметрической и т.д.

Легко показать, что все ортогональные, унитарные, симметрические, кососимметрические операторы являются нормальными операторами (т.е. свойство $f^* f = f f^*$, очевидно, будет выполнено).

Лемма 4.3.9 *Пусть $f : W \to W$ — оператор проектирования. Он является самосопряженным тогда и только тогда, когда проектирование является ортогональным.*

Доказательство. Пусть $W = V_1 \oplus V_2$, f — оператор проектирования на V_1 вдоль V_2, тогда $V_1 = \operatorname{Im} f$, $V_2 = \operatorname{Ker} f$. Условие ортогональности проектирования равносильно условию $V_1 \perp V_2$.

Пусть оператор f самосопряженный. Возьмем произвольные векторы $a \in \operatorname{Ker} f$ и $b \in \operatorname{Im} f$ (тогда $b = fc$ для некоторого $c \in W$), тогда $(a, b) = (a, fc) = (fa, c) = (0, c) = 0$, следовательно $a \perp b$, откуда получаем $\operatorname{Ker} f \perp \operatorname{Im} f$.

Предположим теперь, что $\operatorname{Ker} f \perp \operatorname{Im} f$. Возьмем произвольные векторы $u, v \in W$ и представим их в виде $u = u_1 + u_2$, $v = v_1 + v_2$, где $u_1, v_1 \in \operatorname{Im} f$, $u_2, v_2 \in \operatorname{Ker} f$. Тогда

$$(fu, v) = (\underbrace{fu_1}_{=u_1}, v) + (\underbrace{fu_2}_{=0}, v) = (u_1, v_1 + v_2) = (u_1, v_1) + \underbrace{(u_1, v_2)}_{=0} = (u_1, v_1).$$

Аналогично получаем, что $(u, fv) = (u_1, v_1)$. Поэтому $(fu, v) = (u, fv)$ для любых векторов $u, v \in W$, значит, оператор f самосопряжен. □

4.4 Канонический вид матрицы самосопряженного оператора

Теорема 4.4.1 *Для любого самосопряженного оператора $f : W \to W$ существует ортонормированный базис, в котором его матрица имеет диагональный вид с вещественными числами на диагонали. Указанный канонический вид матрицы самосопряженного оператора единственен с точностью до перестановки диагональных элементов.*

Доказательство.

1) Короткое и неправильное: приведем матрицу к жордановой форме, под диагональю будут нули, следовательно, т.к. $\overline{A}^t = A$, над диагональю тоже будут нули. Ошибка в том, что жорданов базис в общем случае не будет ортогональным.

2) Правильное. При доказательстве мы будем использовать следующую лемму.

Лемма 4.4.2 *Если V — инвариантное подпространство относительно самосопряженного оператора f, то $V^\perp$ также будет инвариантно относительно f.*

Доказательство. Очевидно, т.к. $V^\perp$ инвариантно относительно $f^* = f$. □

Проведем доказательство теоремы по индукции.

1) Пусть $\dim W = 1$. Очевидно, $A_f = (a)$ — диагональная матрица. В случае поля комплексных чисел, a будет вещественным числом, т.к. $\overline{a} = a$.

2) Допустим, что теорема доказана для $\dim W \leqslant n$, докажем ее для $\dim W = n + 1$.

Сначала разберем случай поля комплексных чисел. Пусть v — собственный вектор оператора f, $V = \langle v \rangle$, тогда $W = V \oplus V^\perp$, причем V и $V^\perp$ оба инвариантны относительно f, тогда матрица оператора f имеет вид $A_f = \left(\begin{array}{c|c} a & \mathbf{0} \\ \hline \mathbf{0} & A_{f'} \end{array}\right)$, где ограничение f' оператора f на $V^\perp$ тоже самосопряжено, следовательно, по предположению индукции, его можно привести к искомому виду. В итоге матрица A_f будет диагональной, причем, т.к. $\overline{a} = a$, на диагонали будут вещественные числа.

Теперь перейдем к случаю поля вещественных чисел. Мы знаем, что у любого оператора над полем вещественных чисел существует либо одномерное, либо двумерное инвариантное подпространство. Если у этого оператора есть хотя бы одно одномерное инвариантное подпространство, то можно действовать аналогично предыдущему случаю. Если же у этого оператора имеются лишь двумерные инвариантные подпространства, то его матрица имеет вид $A_f = \left(\begin{array}{cc|c} a_{11} & a_{12} & \mathbf{0} \\ a_{21} & a_{22} & \\ \hline \multicolumn{2}{c|}{\mathbf{0}} & A_{f'} \end{array}\right)$. По предположению индукции матрица ограничения f' на ортогональное дополнение к инвариантному подпространству имеет искомый вид. Осталось лишь разобраться с двумерной клеткой $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Т.к. $A^t = A$, то $a_{21} = a_{12}$. Для дальнейшего доказательства нам потребуется

Лемма 4.4.3 *У двумерной симметричной матрицы характеристический многочлен имеет вещественные корни.*

Доказательство.

$$\det \begin{pmatrix} a_{11}-t & a_{12} \\ a_{21} & a_{22}-t \end{pmatrix} = t^2 - (a_{11}+a_{22})t + a_{11}a_{22} - a_{12}^2.$$

Дискриминант этого многочлена равен $D = (a_{11}+a_{22})^2 - 4(a_{11}a_{22} - a_{12}^2) = (a_{11}-a_{22})^2 + 4a_{12}^2 \geqslant 0$. □

Пусть λ_1, λ_2 — собственные числа матрицы A. Тогда в базисе, составленным из собственных векторов, она имеет вид $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. Если $\lambda_1 = \lambda_2$, то мы всегда можем выбрать два ортонормированных собственных вектора в качестве базиса двумерного пространства. Если же $\lambda_1 \neq \lambda_2$, то применим следующее утверждение.

Лемма 4.4.4 *Собственые вектора, отвечающие различным собственным значениям самосопряженного оператора взаимно ортогональны.*

Доказательство. Пусть $fv_1 = \lambda v_1$, $fv_2 = \lambda_2 v_2$, тогда $(fv_1, v_2) = (\lambda_1 v_1, v_2) = \lambda_1(v_1, v_2)$. Но, с другой стороны, $(fv_1, v_2) = (v_1, fv_2) = (v_1, \lambda_2 v_2) = \lambda_2(v_1, v_2)$. Таким образом, $\lambda_1(v_1, v_2) = \lambda_2(v_1, v_2)$, следовательно, $(v_1, v_2) = 0$, т.к. $\lambda_1 \neq \lambda_2$. □

Доказательство существования канонического вида закончено. Единственность следует из того, что на диагонали там стоят корни характеристического многочлена с учетом их кратности. □

4.5 Канонический вид матрицы кососимметрического оператора

Теорема 4.5.1 *У любого кососимметрического оператора $f : W \to W$ над полем комплексных чисел существует ортонормированный базис, в котором его матрица A_f имеет диагональный вид с чисто мнимыми числами на диагонали.*

Доказательство. Мы знаем, что если $V \subset W$ инвариантно относительно f, то и $V^\perp$ будет инвариантно относительно f. Т.к. $f^* = -f$, то $f^*(V^\perp) = -f(V^\perp) = f(V^\perp) \subset V^\perp$, т.е. $V^\perp$ будет инвариантно относительно f. Следовательно матрицу оператора f можно привести к диагональному виду (таким же способом, каким мы делали это ранее — по индукции, соответствующий базис будет ортонормированным), $A_f = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$. Т.к. $f^* = -f$, то $\overline{A}^t = -A$, следовательно $\overline{\lambda}_i = -\lambda_i$, $i = 1, \ldots, n$, следовательно все λ_i — чисто мнимые числа. □

Теорема 4.5.2 *У любого кососимметрического оператора $f : W \to W$ над полем вещественных чисел существует ортонормированный базис, в котором матрица A_f имеет блочно-диагональный вид, причем все блоки либо одномерные (равные нулю), либо двумерные, имеющие вид $\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$, $a \in \mathbb{R}$.*

Доказательство. Разложив пространство W на инвариантные подпространства (одномерные или двумерные) оператора f, получим искомый блочно-диагональный вид. Т.к. $\overline{A}_f^t = -A_f$, то на диагонали этой матрицы стоят нули, следовательно все одномерные клетки нулевые, и по диагонали двумерных клеток стоят нули. Таким образом, двумерные клетки имеют вид $\begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$, но из того же свойства кососимметричности следует, что $b = -a$. □

Единственность канонического вида матриц кососимметрических операторов с точностью до перестановки блоков также следует из того, что эти блоки определяются характеристическим многочленом оператора.

4.6 Частичные изометрии

Определение 4.6.1 Оператор f называется *частичной изометрией*, если он изометричен на ортогональном дополнении к ядру, т.е. если $|fa| = |a|$ для любого вектора $a \in (\operatorname{Ker} f)^\perp$.

Операторы частичной изометрии можно рассматривать как некоторое обобщение ортогональных и унитарных операторов.

Пример.

Пусть оператор f задается в ортонормированном базисе e_1, e_2 на плоскости матрицей $A_f = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, т.е. $fe_2 = e_1$, $fe_1 = 0$. Очевидно, что $\operatorname{Im} f = \operatorname{Ker} f = \langle e_1 \rangle$, следовательно, $(\operatorname{Ker} f)^\perp = \langle e_2 \rangle$. Частичная изометричность оператора f означает, что он не меняет длин векторов, коллинеарных вектору e_2.

Лемма 4.6.2 *Если оператор $f : W \to W$ является частичной изометрией, то и оператор f^* будет частичной изометрией.*

Доказательство. Обозначим $V = (\operatorname{Ker} f)^\perp$. Т.к. $W = V \oplus V^\perp$, то мы можем выбрать ортонормированный базис $e_1, \dots, e_r$ в подпространстве V и дополнить его до ортонормированного базиса всего пространства векторами $e_{r+1}, \dots, e_n \in V^\perp$. Рассмотрим образ этого базиса, т.е. векторы $\widetilde{e}_i = f(e_i)$ (это имеет смысл делать только при $i = 1, \dots, r$, т.к. $f(e_i) = 0$ для $i > r$). Т.к. f сохраняет расстояния на V, то он сохраняет и скалярное произведение на V, следовательно векторы $\widetilde{e}_1, \dots, \widetilde{e}_r$ будут ортонормированными. Обозначим $\langle \widetilde{e}_1, \dots, \widetilde{e}_r \rangle = \widetilde{V}$ и дополним векторы $\widetilde{e}_1, \dots, \widetilde{e}_r$ до ортонормированного базиса всего пространства векторами $\widetilde{e}_{r+1}, \dots, \widetilde{e}_n$ (эти векторы будут базисом в $\widetilde{V}^\perp$).

Посмотрим как действует оператор f^* на векторах $\widetilde{e}_1, \dots, \widetilde{e}_r$. Координаты вектора $f^*(\widetilde{e}_i)$ $(i \leqslant r)$ в базисе $e_1, \dots, e_n$ будут равны

$$(f^*(\widetilde{e}_i), e_j) = (\widetilde{e}_i, fe_j) = (\widetilde{e}_i, \widetilde{e}_j) = \begin{cases} (e_i, e_j) = \delta_{ij} & \text{при } j \leqslant r \\ 0 & \text{при } j > r \end{cases},$$

следовательно, $f^*(\widetilde{e}_i) = \begin{cases} e_i & \text{при } i \leqslant r \\ 0 & \text{при } i > r \end{cases}$. Ясно, что $\operatorname{Ker} f^* = \widetilde{V}^\perp$, следовательно, $(\operatorname{Ker} f^*)^\perp = \widetilde{V}$, и для любого $a \in \widetilde{V}$, $a = a^i \widetilde{e}_i$ (суммаа берется по $i = 1, \dots, r$) выполнено равенство

$$(f^*a, f^*a) = (f^* a^i \widetilde{e}_i, f^* a^i \widetilde{e}_i) = (a^i e_i, a^i e_i) = \sum_{i=1}^{r} (a^i)^2 = (a, a),$$

следовательно, оператор f^* изометричен (сохраняет длины) на $\widetilde{V} = (\operatorname{Ker} f^*)^\perp$, т.е. является частичной изометрией. □

Поскольку $(f^*)^* = f$, то очевидно будет выполнено и обратное утверждение, а именно, если f^* — частичная изометрия, то и f — частичная изометрия.

Лемма 4.6.3 *Оператор $f^* f$ является оператором проектирования тогда и только тогда, когда f является частичной изометрией.*

Доказательство. Пусть f — частичная изометрия. Воспользуемся обозначениями из доказательства предыдущей леммы. Поскольку $f(e_i) = \begin{cases} \widetilde{e}_i & \text{при } i \leqslant r \\ 0 & \text{при } i > r \end{cases}$ и $f^*(\widetilde{e}_i) = \begin{cases} e_i & \text{при } i \leqslant r \\ 0 & \text{при } i > r \end{cases}$, получаем, что $f^* f(e_i) = \begin{cases} e_i & \text{при } i \leqslant r \\ 0 & \text{при } i > r \end{cases}$, следовательно, $f^* f$ является оператором проектирования (причем ортогонального) на $V = (\operatorname{Ker} f)^\perp$.

Обратно, пусть $p = f^* f$ является оператором проектирования на некоторое подпространство V, $W = V \oplus V^\perp$. Т.к. $(f^* f)^* = f^* (f^*)^* = f^* f$, то p является самосопряженным проектором, следовательно он является ортогональным проектором и $V^\perp = \operatorname{Ker} p$. Возьмем произвольный

вектор $a \in W$, $a = a_1 + a_2$, где $a_1 \in V$, $a_2 \in V^\perp$. Тогда $pa = pa_1 = a_1$ и $(fa, fa) = (f^*fa, a) = (pa, a) = (a_1, a) = (a_1, a_1)$, откуда следует, что для любого вектора $v \in V$ $(fv, fv) = (v, v)$, а для любого вектора $u \in V^\perp$ $(fu, fu) = 0$, т.е. $fu = 0$, таким образом, оператор f изометричен на $V = (\mathrm{Ker}\, f)^\perp$, т.е. является частичной изометрией. $\square$

Лемма 4.6.4 *Оператор f является частичной изометрией тогда и только тогда, когда $f = ff^*f$.*

Доказательство. Если f — частичная изометрия, то $p = f^*f$ — ортогональный проектор на $V = (\mathrm{Ker}\, f)^\perp$. Тогда для любого $a \in W$, $a = a_1 + a_2$, где $a_1 \in V$, $a_2 \in V^\perp = \mathrm{Ker}\, f$, будем иметь $fa = fa_1 + fa_2 = fa_1$, и, поскольку $pa = a_1$, то $ff^*fa = fpa = fa_1 = fa$.

Обратно, пусть $f = ff^*f$. Тогда $f^*f = f^*ff^*f = (f^*f)^2$, следовательно, f^*f является проектором, значит, f является частичной изометрией. $\square$

4.7 Неотрицательные и положительные операторы. Извлечение неотрицательного квадратного корня из оператора

Определение 4.7.1 Оператор $f : V \to V$ называется *неотрицательным*, если
1) $f^* = f$,
2) $(fv, v) \geqslant 0$ для любого $v \in V$

Упражнение: в случае поля комплексных чисел первое условие является следствием второго, а в вещественном — нет.

Лемма 4.7.2 *Оператор f является неотрицательным $\iff$ в некотором ортонормированном базисе его матрица диагональна с неотрицательными элементами на диагонали.*

Доказательство.
$\Rightarrow$: Т.к. $f^* = f$, то существует ортонормированный базис $e_1, \ldots, e_n$, в котором матрица оператора диагональна, $A_f = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$, $\lambda_i \in \mathbb{R}$. Тогда $(fe_i, e_i) = \lambda_i \geqslant 0$ для всех $i = 1, \ldots, n$.

$\Leftarrow$: очевидно. $\square$

Лемма 4.7.3 (извлечение квадратного корня) *Если f — неотрицательный оператор, то существует (и притом единственный) такой неотрицательный оператор g, что $g^2 = f$.*

Доказательство.
Существование. Т.к. f — неотрицательный, то в некотором ортонормированном базисе его матрица имеет вид $A_f = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$, причем все $\lambda_i \geqslant 0$. Возьмем оператор g с матрицей $A_g = \begin{pmatrix} \sqrt{\lambda_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{\lambda_n} \end{pmatrix}$ в этом же базисе, тогда $g^2 = f$.

Единственность. Пусть наряду с оператором g, построенным выше, существует еще один такой неотрицательный оператор h, что $h^2 = f$. В некотором базисе $a_1, \ldots, a_n$ матрица оператора g имеет вид $A_g = \begin{pmatrix} \sqrt{\lambda_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{\lambda_n} \end{pmatrix}$, а т.к. оператор h неотрицательный, то и его матрица в некотором (вообще говоря другом) базисе $b_1, \ldots, b_n$ имеет вид $A_h = \begin{pmatrix} \mu_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mu_n \end{pmatrix}$. Т.к. λ_i

и μ_i^2 — это собственные значения оператора f (с учетом кратности), то μ_i совпадают с $\sqrt{\lambda_i}$ с точностью до перестановки. Без ограничения общности можно считать, что они уже переставлены так, чтобы μ_i в точности совпадали с $\sqrt{\lambda_i}$. Построим изометрию u, переводящую один базис в другой, $u(a_i) = b_i$, тогда

$$(u^*hu)a_i = (u^*h)b_i = u^*\mu_i b_i = \mu_i u^* b_i = \sqrt{\lambda_i}u^* b_i = \sqrt{\lambda_i}a_i = ga_i,$$

следовательно, $u^*hu = g$ и $h = ugu^*$ (т.к. оператор u обратим и $u^{-1} = u^*$). Возведя обе части этого равенства в квадрат, получим

$$f = g^2 = (u^*hu)^2 = u^*huu^*hu = u^*h^2u = u^*fu,$$

значит, $uf = fu$, т.е. оператор u перестановочен с f. Если бы u был перестановочен и с g, т.е. если бы было верно равенство $ug = gu$, то тогда получилось бы, что $h = ugu^* = guu^* = g$, что нам и нужно доказать.

Докажем перестановочность u с g. Пусть матрица оператора u в базисе $a_1, \ldots, a_n$ имеет вид $A_u = \begin{pmatrix} u_1^1 & \ldots & u_n^1 \\ \vdots & \ddots & \vdots \\ u_1^n & \ldots & u_n^n \end{pmatrix}$ (в этом базисе матрицы операторов f и g имеют диагональный вид). Тогда

$$A_{uf} = A_u A_f = \begin{pmatrix} u_1^1 & \ldots & u_n^1 \\ \vdots & \ddots & \vdots \\ u_1^n & \ldots & u_n^n \end{pmatrix} \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 u_1^1 & \ldots & \lambda_n u_n^1 \\ \vdots & \ddots & \vdots \\ \lambda_1 u_1^n & \ldots & \lambda_n u_n^n \end{pmatrix},$$

аналогично, $A_{fu} = \begin{pmatrix} \lambda_1 u_1^1 & \ldots & \lambda_1 u_n^1 \\ \vdots & \ddots & \vdots \\ \lambda_n u_1^n & \ldots & \lambda_n u_n^n \end{pmatrix}$. Перестановочность u с f означает, что

$$\lambda_i u_i^j = \lambda_j u_i^j \forall i, j. \tag{6}$$

Точно так же получаем, что

$$A_{ug} = \begin{pmatrix} \sqrt{\lambda_1}u_1^1 & \ldots & \sqrt{\lambda_n}u_n^1 \\ \vdots & \ddots & \vdots \\ \sqrt{\lambda_1}u_1^n & \ldots & \sqrt{\lambda_n}u_n^n \end{pmatrix}, \qquad A_{gu} = \begin{pmatrix} \sqrt{\lambda_1}u_1^1 & \ldots & \sqrt{\lambda_1}u_n^1 \\ \vdots & \ddots & \vdots \\ \sqrt{\lambda_n}u_1^n & \ldots & \sqrt{\lambda_n}u_n^n \end{pmatrix}.$$

Из равенства (6) следует, что $\sqrt{\lambda_i}u_i^j = \sqrt{\lambda_j}u_i^j$ для всех i, j, т.е. $ug = gu$. □

Единственный неотрицательный квадратный корень из неотрицательного оператора f мы будем обозначать $f^{1/2}$.

Лемма 4.7.4 *Оператор f неотрицателен тогда и только тогда, когда $f = h^*h$ для некоторого оператора h.*

Доказательство.

Если f неотрицателен, то в качестве оператора h можно взять квадратный корень из f.

Если $f = h^*h$, то $f^* = (h^*h)^* = h^*(h^*)^* = h^*h = f$, т.е. f — самосопряженный. Для любого вектора a имеем $(fa, a) = (h^*ha, a) = (ha, ha) \geqslant 0$, следовательно, f — неотрицательный. □

Определение 4.7.5 Оператор f называется *положительным*, если он неотрицателен и обратим.

Для любого положительного оператора существует ортонормированный базис, в котором его матрица диагональна, причем на диагонали стоят положительные числа.

Лемма 4.7.6 *Если оператор $f : V \to V$ положителен, то формула $\langle a, b\rangle := (fa, b)$, где $a, b \in V$, задает (новое) скалярное произведение $\langle\cdot,\cdot\rangle$ в гильбертовом пространстве V со скалярным произведением $(\cdot,\cdot)$.*

Доказательство. Проверим необходимые условия скалярного произведения:
1) линейность очевидна, т.к. оператор f линейный и скалярное произведение $(\cdot,\cdot)$ также линейно;
2) симметричность следует из самосопряженности f;
3) положительная определенность: $\langle a, b\rangle \geqslant 0$, т.к. $\langle a, b\rangle = (fa, a) = (g^2a, a) = (ga, ga) \geqslant 0$, где g — неотрицательный квадратный корень из f. Причем оператор g положителен, поскольку обратим (иначе $f = g^2$ был бы необратим). Равенство $\langle a, a\rangle = 0$ эквивалентно $(ga, ga) = 0$, или $ga = 0$, но из обратимости g отсюда вытекает $a = 0$. □

4.8 Полярное разложение операторов

На примере комплексных чисел мы знаем, что они допускают так называемое полярное разложение — разложение в виде $z = r \cdot \mathrm{e}^{i\varphi}$, где $r \geqslant 0$ — вещественное число, и $|\mathrm{e}^{i\varphi}| = 1$. Оказывается, что подобным образом можно раскладывать и линейные операторы.

Теорема 4.8.1 (о полярном разложении) *Для любого оператора $f : W \to W$ существует разложение в произведение двух операторов $f = uh$, где u — унитарный (или ортогональный), а h — неотрицательный. Причем оператор h определен однозначно, а, если f обратим, то и u определен однозначно.*

Доказательство.
Разберем сначала более простой случай, когда f обратим.
Существование. Т.к. f обратим, то и f^* обратим, а, следовательно, и f^*f обратим. Более того, f^*f — неотрицательный, следовательно, он положительный. Пусть $h = (f^*f)^{1/2}$, тогда h — неотрицательный и обратимый, т.к. $h^2 = f^*f$, т.е. существует самосопряженный (и даже положительный) оператор h^{-1}. Рассмотрим оператор $u = fh^{-1}$. Тогда $u^* = (h^{-1})^*f^* = h^{-1}f^*$ и $u^*u = h^{-1}f^*fu = h^{-1}h^2h^{-1} = id$, следовательно, u — унитарный (или ортогональный) оператор. Мы получили искомое разложение $f = uh$.
Единственность. Пусть $f = uh = vk$, где u, v — унитарные (ортогональные), а h, k — неотрицательные операторы. Тогда имеем $h^2 = f^*f = (vk)^*vk = k^*v^*vk = k^*k = k^2$, т.е. h и k — это неотрицательные квадратные корни из f^*f, следовательно, $k = h$ (здесь мы не пользуемся обратимостью f), и, значит, $u = v = fh^{-1}$.
Рассмотрим теперь общий случай, когда обратимость f не предполагается. Единственность неотрицательного сомножителя в полярном разложении не связана с обратимостью и уже доказана. Докажем существование полярного разложения. Пусть $V = (\operatorname{Ker} f)^\perp$, $V^\perp = \operatorname{Ker} f$, т.е. $fv = 0\ \forall v \in V^\perp$ и $fv \neq 0\ \forall v \in V$, $v \neq 0$. Пусть $\widetilde{V} = \operatorname{Im} f$, тогда $\dim V = \dim \widetilde{V}$ и $\dim V^\perp = \dim \widetilde{V}^\perp$, т.к.

$$\dim V = \dim W - \dim V^\perp = \dim W - \dim \operatorname{Ker} f = \dim \operatorname{Im} f = \dim \widetilde{V}.$$

Возьмем такой оператор $g : W \to W$, что он равен нулю на всех векторах из V и изоморфно отображает $V^\perp$ в $\widetilde{V}^\perp$ (таких операторов много, возьмем какой-нибудь из них). Тогда оператор $f + g$ будет изоморфизмом W на себя, следовательно он обратим.
Действительно, пусть $\underbrace{e_1, \ldots, e_r}_{\text{базис } V}, \underbrace{e_{r+1}, \ldots, e_n}_{\text{базис } V^\perp}$ — базис пространства W. Обозначим $\widetilde{e}_i = (f + g)e_i = fe_i$ при $i \leqslant r$ и $\widetilde{e}_i = (f+g)e_i = ge_i$ при $i > r$, тогда $\widetilde{e}_1, \ldots, \widetilde{e}_r, \widetilde{e}_{r+1}, \ldots, \widetilde{e}_n$ будет базисом в W (вторая группа векторов образует базис в $\widetilde{V}^\perp$ по определению частичной изометрии g, а первая группа векторов образует базис в $\widetilde{V}$, поскольку ограничение оператора f на V осуществляет взаимно-однозначное соответствие между V и $\widetilde{V}$), следовательно, $f+g$ — изоморфизм W на себя. Тогда, по уже доказанному, существует полярное разложение $f + g = uh$, где u — унитарный (ортогональный), а h — неотрицательный оператор, причем $h = ((f + g)^*(f + g))^{1/2} = (f^*f + g^*g + g^*f + f^*g)^{1/2}$.
Докажем лемму.

Лемма 4.8.2 $f^*f(W) \subset V$, $g^*g(W) \subset V^\perp$, $g^*f = f^*g = 0$.

Доказательство. Пусть $a \in W$, $b \in V^\perp$. Тогда $(f^*fa, b) = (fa, \underbrace{fb}_{=0}) = 0$, поэтому $f^*fa \in V$.

Аналогично проверяется, что $g^*g(W) \subset V^\perp$ (на самом деле здесь равенство, а не вложение, но нам это не потребуется).

Возьмем $a = a_1 + a_2 \in W$, $a_1 \in V$, $a_2 \in V^\perp$. Тогда $f^*ga = f^*\underbrace{ga_1}_{=0} + f^*ga_2 = f^*ga_2$. Для любого $c \in W$ имеем $(f^*ga_2, c) = (\underbrace{ga_2}_{\in\widetilde{V}^\perp}, \underbrace{fc}_{\in\widetilde{V}}) = 0$, поэтому $f^*ga_2 = 0$. Таким образом, $f^*ga = 0$ для любого вектора $a \in W$. Аналогично доказывается, что $g^*f = 0$. □

Продлжим доказательство теоремы. Нам известно, что $f + g = uh$, где $h = (f^*f + g^*g)^{1/2}$. Пусть p — оператор ортогонального проектирования на V. Очевидно, что $fp = f$ и $gp = 0$. Умножая равенство $f + g = uh$ справа на p, получим $f = uhp$. Все, что осталось доказать, — это неотрицательность оператора hp. На самом деле мы докажем, что $hp = (f^*f)^{1/2}$, откуда и будет следовать его неотрицательность.

Пусть $e_1, \ldots, e_r$ — ортонормированный базис в V, а $e_{r+1}, \ldots, e_n$ — ортонормированный базис в $V^\perp$. Из леммы 4.8.2 видно, что в базисе $e_1, \ldots, e_r, e_{r+1}, \ldots, e_n$ матрицы операторов f^*f и g^*g имеют вид $A_{f^*f} = \left(\begin{array}{c|c} \star & 0 \\ \hline 0 & 0 \end{array}\right)$, $A_{g^*g} = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & \star \end{array}\right)$, поэтому матрицы их неотрицательных квадратных корней имеют такой же вид, $A_{\sqrt{f^*f}} = \left(\begin{array}{c|c} \star & 0 \\ \hline 0 & 0 \end{array}\right)$, $A_{\sqrt{g^*g}} = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & \star \end{array}\right)$, а матрица проектора p имеет вид $A_p = \left(\begin{array}{c|c} E & 0 \\ \hline 0 & 0 \end{array}\right)$, откуда получаем, что

$$(g^*g)^{1/2}(f^*f)^{1/2} = (f^*f)^{1/2}(g^*g)^{1/2} = (g^*g)^{1/2}p = 0$$

и $(f^*f)^{1/2}p = (f^*f)^{1/2}$. Но тогда

$$((f^*f)^{1/2} + (g^*g)^{1/2})^2 = f^*f + g^*g,$$

поэтому (единственность неотрицательного квадратного корня)

$$h = (f^*f + g^*g)^{1/2} = (f^*f)^{1/2} + (g^*g)^{1/2}.$$

Умножая это выражение справа на p, получаем $hp = (f^*f)^{1/2}p + (g^*g)^{1/2}p = (f^*f)^{1/2}$. □

4.9 Канонический вид матрицы нормального оператора

Теорема 4.9.1 *Оператор f, действующий в унитарном пространстве W, является нормальным (т.е. $f^*f = ff^*$) тогда и только тогда, когда в некотором ортонормированном базисе его матрица диагональна.*

Доказательство. Если в некотором ортонормированном базисе матрица оператора f диагональна, $A_f = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$, то матрица сопряженного оператора имеет в том же базисе вид $A_{f^*} = \begin{pmatrix} \overline{\lambda}_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \overline{\lambda}_n \end{pmatrix}$. Так как $A_fA_{f^*} = A_{f^*}A_f$, то $f^*f = ff^*$.

Теперь перейдем к доказательству существования ортонормированного базиса, в котором матрица нормального оператора диагональна. Пусть λ — собственное значение оператора f,

$V = \{v \in W : fv = \lambda v\} \subset W$ - подпространство собственных векторов, отвечающих λ. Подпространство V, очевидно, инвариантное относительно f, будет инвариантно также относительно f^*. Действительно, возьмем произвольный $v \in V$. Тогда $f^*fv = f^*\lambda v = \lambda f^*v$, но с другой стороны, $f^*fv = ff^*v$, поэтому $f(f^*v) = \lambda f^*v$, следовательно, вектор f^*v является собственным, отвечающим тому же собственному значению λ, поэтому $f^*v \in V$, что и доказывает инвариантность V относительно f^*.

Мы знаем, что, если подпространство инвариантно относительно f, то ортогональное дополнение инвариантно относительно f^*. Т.к. V инвариантно относительно обоих операторов f и f^*, то и $V^\perp$ будет инвариантно относительно обоих операторов.

Доказательство теоремы проведем по индукции (по размерности пространства) точно так же, как и в предыдущих теоремах о приведении к диагональному виду. Возьмем инвариантное подпространство V, тогда в подходящем ортонормированном базисе его матрица имеет блочный вид $A_f = \left(\begin{array}{c|c} \star & 0 \\ \hline 0 & \star \end{array}\right)$. Ограничения оператора f на V и $V^\perp$ тоже нормальны, следовательно (по предположению индукции) эти блоки диагональны в подходящих базисах, следовательно, и вся матрица диагональна. $\square$

5 Билинейные и полуторалинейные функции

5.1 Билинейные функции (формы)

Определение 5.1.1 Пусть V — векторное пространство над полем $\mathbb{K}$. Функция $g : V \times V \to \mathbb{K}$ называется *билинейной функцией*, если она линейна по каждому аргументу, т.е.

$$\begin{aligned} g(a_1 + a_2, b) &= g(a_1, b) + g(a_2, b) && \forall a_1, a_2, b \in V; \\ g(\lambda a, b) &= \lambda g(a, b) && \forall a, b \in V, \lambda \in \mathbb{K}; \\ g(a, b_1 + b_2) &= g(a, b_1) + g(a, b_2) && \forall a, b_1, b_2 \in V; \\ g(a, \lambda b) &= \lambda g(a, b) && \forall a, b \in V, \lambda \in \mathbb{K}. \end{aligned}$$

Если выбрать базис $e_1, \ldots, e_n$ в пространстве V, то билинейную функцию можно записать матрицей $G = (g_{ij})$, где $g_{ij} = g(e_i, e_j)$. Причем (если базис зафиксирован), то существует взаимнооднозначное соответствие между квадратными матрицами и билинейными функциями, т.е. любая матрица задает какую-то функцию и разные матрицы задают разные функции.

Пример.

Если $g(a, b) = (a, b)$ — обычное скалярное произведение в евклидовом пространстве, то g будет билинейной функцией, а ее матрица G будет просто матрицей Грама. Поэтому на матрицу билинейной функции можно смотреть как на обобщение матрицы Грама.

Если G — матрица билинейной функции g, то значение этой функции на двух любых векторах восстанавливается по формуле $g(x, y) = g(x^i e_i, y^j e_j) = x^i g(e_i, e_j) y^j = x^i g_{ij} y^j$ или, в матричной форме

$$g(x, y) = (x^1 \ldots x^n) G \begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix}.$$

При замене базиса $e'_k = c^i_k e_i$, где $C = (c^i_k)$ — матрица перехода, матрица билинейной функции изменится следующим образом:

$$g'_{kl} = g(e'_k, e'_l) = g(c^i_k e_i, c^j_l e_j) = c^i_k g(e_i, e_j) c^j_l = c^i_k g_{ij} c^j_l,$$

т.е. $G' = C^t G C$, где G' — матрица той же билинейной функции в новом базисе.

На множестве билинейных функций можно естественным образом определить структуру линейного пространства (над тем же полем $\mathbb{K}$), причем размерность этого пространства будет n^2, где $n = \dim V$. Обозначается оно $B(V)$. Очевидно, что $B(V) \cong Mat(n \times n)$.

В случае поля комплексных чисел рассматривают также полуторалинейные функции.

Определение 5.1.2 Пусть V — векторное пространство над полем $\mathbb{C}$. Функция $f : V \times V \to \mathbb{C}$ называется *полуторалинейной функцией*, если она линейна по второму аргументу и антилинейна по первому, т.е.

$$\begin{aligned} g(a_1 + a_2, b) &= g(a_1, b) + g(a_2, b) && \forall a_1, a_2, b \in V; \\ g(\lambda a, b) &= \overline{\lambda} g(a, b) && \forall a, b \in V, \lambda \in \mathbb{C}; \\ g(a, b_1 + b_2) &= g(a, b_1) + g(a, b_2) && \forall a, b_1, b_2 \in V; \\ g(a, \lambda b) &= \lambda g(a, b) && \forall a, b \in V, \lambda \in \mathbb{C}. \end{aligned}$$

И в этом случае можно, аналогичным образом, ввести понятие матрицы полуторалинейной функции, но при переходе к новому базису матрица полуторалинейной формы будет изменяться по формуле $G' = \overline{C}^t G C$.

Определение 5.1.3 *Рангом* билинейной (полутроалинейной) функции называется ранг ее матрицы в произвольном базисе, $\operatorname{rk} g = \operatorname{rk} G$.

Формулы перехода к другому базису показывают, что это определение корректно. Действительно, поскольку матрица перехода C обратима, $\operatorname{rk} C^t G C = \operatorname{rk} \overline{C}^t G C = \operatorname{rk} G$.

5.2 Канонический изоморфизм $B(V) \cong L(V, V')$. Невырожденность

Пусть $L(V, V')$ — пространство всех линейных отображений $f : V \to V'$, где V' — двойственное пространство к V. Пусть нам дано такое отображение $f : V \to V'$, и $e_1, \ldots, e_n$ — базис в V. Возьмем произвольный вектор $x \in V$, ему отвечает линейная функция $f(x) \in V'$, $f(x) : V \to \mathbb{K}$. Возьмем произвольный вектор $y \in V$, тогда функционал $f(x)$ можно применить к вектору y:

$$f(x)y = f(x^i e_i)y^j e_j = x^i f(e_i)e_j y^j = x^i f_{ij} y^j,$$

где $f_{ij} = f(e_i)e_j$. Таким образом, элементу $f \in L(V, V')$ ставится в соответствие матрица $F = (f_{ij})$. Обратно, каждая матрица $F = (f_{ij})$ задает (если зафиксирован базис) отображение $f : V \to V'$ по формуле $f(x)y = x^i f_{ij} y^j$. Другими словами, фиксация базиса устанавливает изоморфизм $L(V, V') \cong Mat(n \times n)$.

Лемма 5.2.1 *Пространства $B(V)$ и $L(V, V')$ канонически изоморфны.*

Доказательство. Возьмем произвольное билинейное отображение $g \in B(V)$ и зафиксируем его первый аргумент, тогда получится линейное отображение $V \to \mathbb{K}$, заданное формулой $x \mapsto g(a, x)$, следовательно, $g(a, x) \in V'$ при фиксированном a. Теперь мы можем построить отображение $B(V) \to L(V, V')$ по следующему правилу: билинейному отображению $g \in B(V)$ ставим в соответствие отображение $\widetilde{g} : V \to V'$, определенное равенством $\widetilde{g}(a) = g(a, x)$. Ясно, что $\widetilde{g} \in L(V, V')$. Можно построить также и обратное отображение: пусть $f : V \to V'$ — линейное отображение, тогда $f(a) \in V'$ при фиксированном $a \in V$, т.е. $f(a)b \in \mathbb{K}$ и $\widetilde{f}(x, y) = f(x)y$ — билинейная функция, т.е. $\widetilde{f} \in B(V)$. Эти два отображения — $g \mapsto \widetilde{g}$, $f \mapsto \widetilde{f}$ — взаимно обратны, следовательно они осуществляют изоморфизм $B(V) \cong L(V, V')$. Т.к. этот изоморфизм не зависит от выбора базиса, то это канонический изоморфизм. □

Докажем канонический изоморфизм $B(V) \cong L(V, V')$ еще одним способом (с помощью матриц). Выберем базис $e_1, \ldots, e_n$ в V, возьмем отображение $g \in B(V)$, пусть $G = (g_{ij})$ — его матрица, т.е. $g_{ij} = g(e_i, e_j)$. Возьмем двойственный базис $\varepsilon^1, \ldots, \varepsilon^n$ в V' и отображение $\widetilde{g} : V \to V'$ (то же, что и в предыдущем абзаце). Пусть $\widetilde{G} = (\widetilde{g}_{ij})$ — его матрица, т.е. $\widetilde{g}(e_i) = \widetilde{g}_{ij}\varepsilon^j$. Посмотрим, как связаны между собой матрицы G и $\widetilde{G}$. С одной стороны, $\widetilde{g}(e_i)e_k = (\widetilde{g}_{ij}\varepsilon^j)e_k = \widetilde{g}_{ik}$, а с другой стороны, $\widetilde{g}(e_i)e_k = g(e_i, e_k) = g_{ik}$, следовательно, матрицы G и $\widetilde{G}$ совпадают в произвольном (но одном и том же) базисе, откуда и следует канонический изоморфизм $B(V) \cong L(V, V')$. □

Как мы уже знаем, при фиксированном первом аргументе билинейная функция $x \mapsto g(a, x) = {}_a g(x)$ является линейным функционалом. Аналогично, при фиксированном правом аргументе имеем линейный функционал $x \mapsto g(x, a) = g_a(x)$. Фиксация первого (соотв. второго) аргумента определяет отображение $L_g : V \to V'$ равенством $a \mapsto {}_a g$ (соотв. отображение $R_g : V \to V'$ равенством $a \mapsto g_a$).

Определение 5.2.2 *Левым ядром* билинейной функции $g \in B(V)$ называется множество $G_L = \{a \in V : g(a, b) = 0 \ \forall b \in V\}$. *Правым ядром* билинейной функции называется множество $G_R = \{a \in V : g(b, a) = 0 \ \forall b \in V\}$.

Очевидно, что множества G_L и G_R являются подпространствами в V.

Лемма 5.2.3 *Размерности левого и правого ядер совпадают и равны* $\dim G_L = \dim G_R = \dim V - \operatorname{rk} g$.

Доказательство. Т.к. матрица G билинейной функции $g : V \times V \to \mathbb{K}$ и матрица A_{L_g} отображения $L_g : V \to V'$ (см. предыдущую лемму) совпадают, то $\operatorname{rk} G = \operatorname{rk} A_{L_g}$. Более того, $a \in G_L$ тогда и только тогда, когда $a \in \operatorname{Ker} L_g$, т.е. $G_L = \operatorname{Ker} L_g$. Т.к. $\dim V' = \dim V$, то $\dim G_L = \dim \operatorname{Ker} L_g = \dim V' - \operatorname{rk} A_{L_g} = \dim V - \operatorname{rk} g$. Аналогично (заменяя первый аргумент на второй), легко видеть, что матрица G^t совпадает с матрицей A_{R_g} отображения $R_g : V \to V'$ и $\dim G_R = \dim V' - \operatorname{rk} G^t = \dim V - \operatorname{rk} g$. □

Определение 5.2.4 Билинейная функция g называется *невырожденной*, если $\dim G_L = \dim G_R = 0$ (это условие равносильно тому, что $\det G \neq 0$ или $\operatorname{rk} g = \dim V$).

Примеры:

1) пусть $G = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Найдем левое и правое ядро: $g(a,b) = (a^1\ a_2) \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b^1 \\ b^2 \end{pmatrix} = a^1 b^2$. Следовательно, $G_L = \langle e_2 \rangle$ и $G_R = \langle e_1 \rangle$, где e_1, e_2 — базис.

2) билинейная функция может быть невырождена на всем пространстве, но быть вырожденной на подпространстве! Например, пусть $G = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, рассмотрим вектор $a = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ и подпространство $V = \langle a \rangle$. Т.к. $g(a,a) = 0$, то для любых двух векторов $b, c \in V$ (т.е. коллинеарных вектору a) $g(a,b) = 0$, т.е. ограничение билинейной функции g на подпространство V вырождено, в то время как сама функция g невырождена, т.к. $\det G \neq 0$. Отметим, что в этом примере левое и правое ядро совпали, потому (как мы увидим далее) что функция симметрична.

5.3 Симметричные, кососимметричные и эрмитовы функции

Если $g : V \times V \to \mathbb{K}$ — билинейная функция, то функция $(a,b) \mapsto g(b,a)$, полученная из функции g заменой первого и второго аргументов, также является билинейной функцией. Мы будем ее обозначать g^t, $g^t(a,b) = g(b,a)$. Это же обозначение мы будем использовать и для полуторалинейных функций над полем комплексных чисел.

Определение 5.3.1 Билинейная функция g называется *симметричной*, если $g^t = g$, т.е. если $g(b,a) = g(a,b)$; *кососимметричной*, если $g^t = -g$, т.е. если $g(b,a) = -g(a,b)$. Над полем комплексных чисел полуторалинейная функция g называется *эрмитовой*, если $g^t = \overline{g}$, т.е. если $g(b,a) = \overline{g(a,b)}$.

Формально можно также ввести понятие косоэрмитовой полуторалинейной функции, удовлетворяющей условию $g(b,a) = -\overline{g(a,b)}$, однако мы сейчас убедимся, что это понятие сводится к понятию эрмитовой функции.

Лемма 5.3.2 *Если g — косоэрмитова функция, то ig — эрмитова.*

Доказательство. Нам дано, что $g^t = -\overline{g}$. Умножим обе части этого равенства на мнимую единицу, $ig^t = -i\overline{g} = \overline{ig}$. □

Утверждение 5.3.3 *Если билинейная функция симметрична (или кососимметрична), то ее левое и правое ядра совпадают.*

Доказательство. Очевидно (см. определение левого и правого ядер). □

В случае (косо)симметричной билинейной функции g корректно определено ее ядро $\operatorname{Ker} g = G_L = G_R$.

Определение 5.3.4 Функция $f : V \to \mathbb{K}$ называется *квадратичной функцией (формой)*, если существует такая билинейная функция g, что $f(a) = g(a,a)$ для любого $a \in V$.

Заметим, что если f — квадратичная функция, то

$$f(a+b) = g(a+b, a+b) = g(a,a) + g(a,b) + g(b,a) + g(b,b),$$

следовательно, $g(b,a) + g(a,b) = f(a+b) - f(a) - f(b)$. Если поле $\mathbb{K}$ не характеристики 2 (поля характеристики 2 мы вообще далее не будем рассматривать, поскольку нам часто придется исползовать деление пополам), то определим билинейную функцию h равенством $h(a,b) = \frac{1}{2}(g(a,b) + g(b,a))$. Эта функция является симметричной и $h(a,a) = g(a,a) = f(a)$, т.е. ее можно использовать вместо произвольной функции g в определении квадратичной функции.

Таким образом, определение квадратичной функции можно сделать более жестким, а именно потребовать симметричность функции g.

Мы получили, что любая симметричная билинейная функция g задает квадратичную функцию f равенством $f(a) = g(a,a)$, и наоборот, по формуле $g(a,b) = \frac{1}{2}(f(a+b) - f(a) - f(b))$ любая квадратичная функция задает симметричную билинейную функцию. Ясно, что это соответствие взаимно однозначно, когда характеристика поля $\mathbb{K}$ отлична от 2.

Утверждение 5.3.5 *Любая билинейная функция допускает единственное разложение в сумму симметричной и кососимметричной билинейных функций.*

Доказательство.

Существование указанного разложения очевидно, т.к. $g = \frac{1}{2}(g + g^t) + \frac{1}{2}(g - g^t)$.

Единственность. Сначала убедимся, что если билинейная функция одновременно симметрична и кососимметрична, то она нулевая. Действительно, если $g(a,b) = g(b,a) = -g(b,a)$, то $2g(b,a) = 0$, поэтому (характеристика поля не равна 2) $g(b,a) = 0$ для всех a, b.

Пусть теперь функция g обладает двумя разложениями указанного вида, $g = g_1 + g_2 = h_1 + h_2$, где g_1, h_1 симметричны, а g_2, h_2 кососимметричны. Тогда $0 = (g_1 - h_1) + (g_2 - h_2)$, откуда следует, что как функция $g_2 - h_2$, так и функция $g_1 - h_1$, должна быть одновременно симметричной и кососимметричной, поэтому обе эти функции равны нулю. □

5.4 Ортогональное дополнение

Определение 5.4.1 Пусть $V \subset W$, а на W задана (косо)симметричная билинейная функция g, тогда $V^\perp = \{a \in W : g(a,b) = 0 \ \ \forall b \in V\}$.

Пример.

Пусть $G = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ в базисе e_1, e_2 и пусть $a = e_1 + e_2$. Обозначим $V_1 = \langle e_1 \rangle$, $V_2 = \langle e_2 \rangle$, $V_3 = \langle a \rangle$. Тогда $V_1^\perp = V_2$, $V_3^\perp = V_3$. Видно, что ортогональное дополнение в случае произвольной билинейной симметричной функции не очень похоже на обычное ортогональное дополнение. Случай кососимметричной билинейной функции будет выглядеть еще более экзотичным.

Лемма 5.4.2 $\dim V^\perp \geqslant \dim W - \dim V$, *причем равенство достигается тогда и только тогда, когда* $\operatorname{Ker} g \cap V = \{0\}$.

Доказательство. Выберем базис $e_1, \ldots, e_r$ в V. Тогда любой вектор $b \in V$ можно записать в виде $b = b^i e_i$, $1 \leqslant i \leqslant r$. Тогда условие $a \in V^\perp$ эквивалентно равенствам $g(a, e_i) = 0$ для всех $i = 1, \ldots, r$. Дополним выбранный базис до базиса всего пространства. Тогда на n координат вектора a мы будем иметь систему из r линейных уравнений

$$\begin{cases} g(a,e_1) &= 0 \\ \ldots\ldots & \ldots \\ g(a,e_r) &= 0. \end{cases}$$

Размерность пространства решений этой системы (т.е. пространства $V^\perp$) удовлетворяет неравенству $\dim V^\perp \geqslant n - r = \dim W - \dim V$ (здесь стоит неравенство, потому что некоторые уравнения системы могут быть линейно зависимыми).

Равенство будет достигаться тогда и только тогда, когда ранг этой системы равен в точности r, т.е. все строки линейно независимы. Пусть для некоторых коэффициентов λ_i

$$\sum_{i=1}^{r} \lambda_i g(a, e_i) = g(a, \sum_{i=1}^{r} \lambda_i e_i) = g(a,b) = 0$$

(здесь $b = \sum_{i=1}^r \lambda_i e_i \in V$). Если ранг системы уравнений меньше r, т.е. если строки линейно зависимы, то это равносильно тому, что для некоторого $b \in V$, $b \neq 0$, равенство $g(a,b) = 0$ выполнено для всех $a \in W$ (т.е. отображение $a \mapsto g(a,b)$ тождественно равно нулю). Но это значит, что $b \in \operatorname{Ker} g$, или, другими словами, $V \cap \operatorname{Ker} g \neq \{0\}$. Но тогда линейная независимость, наоборот, означает, что $V \cap \operatorname{Ker} g = \{0\}$. □

Лемма 5.4.3 *$V \cap V^\perp = \operatorname{Ker} g_V$, где g_V — ограничение g на V.*

Доказательство. Возьмем произвольный вектор $a \in V \cap V^\perp$, тогда $a \in V$ и $g(a,b) = 0$ для любого $b \in V$, поэтому, по определению, $a \in \operatorname{Ker} g_V$.

Возьмем произвольный вектор $c \in \operatorname{Ker} g_V$, тогда $c \in V$ (т.к. g_V определена только на V) и $g(c,b) = 0$ для любого $b \in V$, следовательно, $c \in V^\perp$, поэтому $c \in V \cap V^\perp$. □

Следствие 5.4.4 *Если $\operatorname{Ker} g_V = \{0\}$ (т.е. ограничение g на V не вырождено), то $W = V \oplus V^\perp$.*

Доказательство. Т.к. $V \cap V^\perp = \{0\}$, то $V + V^\perp = V \oplus V^\perp \subset W$ (т.е. сумма — прямая), причем $\dim(V \oplus V^\perp) = \dim V + \dim V^\perp \geqslant \dim V + (\dim W - \dim V) = \dim W$, следовательно, $V \oplus V^\perp = W$. □

Лемма 5.4.5 *Если ограничения g на V и на $V^\perp$ невырождены, то $(V^\perp)^\perp = V$.*

Доказательство. Вложение $V \subset (V^\perp)^\perp$ имеет место независимо от условий на g. Действительно, если $a \in V$, то $g(a,b) = 0$ для любого $b \in V^\perp$.

Докажем совпадение V и $(V^\perp)^\perp$. Поскольку ограничения g на V и $V^\perp$ навырождены, имеют место разложения $W = V \oplus V^\perp = V^\perp \oplus (V^\perp)^\perp$. Поэтому $\dim V = \dim(V^\perp)^\perp$, и из совпадения размерностей пространства $(V^\perp)^\perp$ и его подпространства V следует их совпадение. □

Лемма 5.4.6 *Пусть $W = V \oplus V^\perp$, $e_1, \ldots, e_r$ — базис в V, $e_{r+1}, \ldots, e_n$ — базис в $V^\perp$. Тогда в базисе $e_1, \ldots, e_n$ матрица билинейной функции имеет вид $G = \left(\begin{array}{c|c} \star & 0 \\ \hline 0 & \star \end{array}\right)$.*

Доказательство. Поскольку при $i \leqslant r$, $j > r$ (и наоборот, при $i > r$, $j \leqslant r$) $g(e_i, e_j) = 0$, так как векторы e_i, e_j принадлежат различным прямым слагаемым, то в левом нижнем и правом верхнем углах матрицы будут нули. □

5.5 Нормальный вид матрицы (косо)симметрической или эрмитовой функции

Посмотрим подробнее, как устроены линейные пространства с заданными на них (косо)симметричными билинейными функциями.

Начнем с одномерного случая.

I. Симметричная функция в одномерном вещественном пространстве. Пусть $e \in V$ — базис в V, $a = \alpha e$ и $b = \beta b$, тогда $g(a,b) = \alpha\beta g(e,e)$. Если $g(e,e) = 0$, то такая функция вырождена на V. Если $g(e,e) > 0$, то, изменив длину базисного вектора, можно получить такой вектор e', что $g(e',e') = 1$. Если $g(e,e) < 0$, то таким же способом можно получить $g(e',e') = -1$. Таким образом, существует базис, в котором матрица (одномерная) этой функции имеет один из трех видов, а именно (0), (1) или (-1).

II. Симметричная функция в одномерном комплексном пространстве. Этот случай аналогичен предыдущему. Если $g(e,e) = \lambda \in \mathbb{C}$, то, умножением e на $\lambda^{-1/2}$ можно получить такой вектор e', что $g(e',e') = 1$. Итак, в этом случае существует базис, в котором функция имеет один из двух видов — (0) или (1).

III. Эрмитова функция в одномерном комплексном пространстве. Поскольку $g(a,a) = \overline{g(a,a)}$, то $g(a,a)$ вещественна для любого a. Если $e' = \gamma e$, то $g(e',e') = |\gamma| g(e,e)$, т.е. $g(e',e')$ имеет тот же знак, что и $g(e,e)$. Поэтому в этом случае матрица функции опять имеет один из трех видов — (0), (1) или (-1).

IV. Кососимметричная функция. Т.к. $g(e,e) = -g(e,e)$, то любая кососимметричная функция на одномерном пространстве тождественно равна нулю. В двумерном случае, если g не тождественно равна нулю, то найдутся такие векторы a, b, что $g(a,b) = \alpha \neq 0$. Приэтом эти два

вектора обязаны быть линейно независимыми, (если $b = \lambda a$, то $g(a,b) = \lambda g(a,a) = 0$) следовательно они образуют базис в двумерном пространстве. В этом базисе матрица функции g имеет вид $G = \begin{pmatrix} g(a,a) & g(a,b) \\ g(b,a) & g(b,b) \end{pmatrix} = \begin{pmatrix} 0 & \alpha \\ -\alpha & 0 \end{pmatrix}$. Взяв базис $e_1 = a$, $e_2 = \frac{1}{\alpha}b$, получим в нем матрицу $G = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Таким образом, у любой кососимметрической функции на двумерном пространстве существует базис, в котором ее матрица имеет один из двух видов — $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ или $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Теорема 5.5.1 *1) у любой симметричной билинейной функции на вещественном векторном пространстве существует базис, в котором ее матрица имеет диагональный вид с числами* $0, \pm 1$ *на диагонали.*

2) у любой симметричной билинейной функции на комплексном векторном пространстве существует базис, в котором ее матрица имеет диагональный вид с числами $0, 1$ *на диагонали.*

3) у любой эрмитовой полуторалинейной функции на комплексном векторном пространстве существует базис, в котором ее матрица имеет диагональный вид с числами $0, \pm 1$ *на диагонали.*

4) у любой кососимметричной билинейной функции на вещественном или комплексном векторном пространстве существует базис, в котором его матрица имеет блочно-диагональный вид с блоками $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ *и* (0) *на диагонали.*

Доказательство. Пункты 1)-3).

Лемма 5.5.2 *Если* $g \neq 0$, *где* g *— функция из пунктов 1)-3) теоремы, то существует такой вектор* a, *что* $g(a,a) \neq 0$.

Доказательство. Предположим противное. Пусть $g(a,a) = 0$ для любого вектора a. Тогда, т.к. $g(a+b, a+b) - g(a,a) - g(b,b) = g(a,b) + g(b,a)$, то $g(a,b) + g(b,a) = 0$ для любых векторов a, b. Если g — симметричная, то сразу получаем $g(a,b) = 0$. Если g — эрмитова, то имеем $g(a,b) + \overline{g(a,b)} = 2\,\mathrm{Re}(g(a,b)) = 0$. Взяв вместо вектора b вектор ib, получим $\mathrm{Re}(g(a,ib)) = \mathrm{Im}(g(a,b)) = 0$ для любых векторов a, b. Следовательно наша функция g тождественно нулевая, что противоречит предположению. □

Далее будем действовать по индукции (база индукции $\dim W = 1$ уже проверена). Пусть утверждение теоремы верно для $\dim V < n$, докажем его для $\dim W = n$. Если $g = 0$, то ее матрица нулевая и всё очевидно. Если же $g \neq 0$, то возьмем такой вектор $a \in W$, что $g(a,a) \neq 0$ (по лемме он существует). Возьмем $V = \langle a \rangle \subset W$. Т.к. ограничение g_V функции g на V невырождено, то $W = V \oplus V^\perp$ и $\dim V^\perp = n - 1$. По индукции в $V^\perp$ уже существует нужный базис, добавив к нему вектор (точнее, некоторое его кратное $e = \lambda a$, такое что $g(e,e) = \pm 1$), получим искомый базис.

4) Если $g \neq 0$, то найдутся такие линейно независимые векторы $a, b \in W$, что $g(a,b) \neq 0$. Без ограничения общности можно считать, что $g(a,b) = 1$. Возьмем $V = \langle a, b \rangle \subset W$. Ограничение g_V функции g на V невырождено, т.к. матрица функции g_V равна $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, следовательно, $W = V \oplus V^\perp$, и $\dim V^\perp = n - 2$. В $V^\perp$ по предположению индукции можно выбрать нужный базис. Добавив к нему векторы a и b, получим искомый базис. □

Указанный вид матрицы называется нормальным. Для приведения матрицы симметричной билинейной или эрмитовой функции к нормальному виду удобно использовать метод Лагранжа выделения полных квадратов. Дадим его краткое описание.

Любая билинейная функция имеет вид $g(x,y) = g_{ij}x^i y^j$, где $x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$, $y = \begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix}$. Заменив y на x, получим квадратичную функцию $g(x,x) = g_{ij}x^i x^j$.

Возможны два случая:

а) Если все коэффициенты при $(x^i)^2$ равны нулю, это можно исправить: найдем ненулевое слагаемое вида $g_{kl}x^kx^l$ и сделаем замену координат $\left\{\begin{array}{rcl} x^k &=& \widetilde{x}^k+\widetilde{x}^l \\ x^l &=& \widetilde{x}^k-\widetilde{x}^l \end{array}\right.$, тогда $x^kx^l=(\widetilde{x}^k)^2-(\widetilde{x}^l)^2$.

б) Если имеется ненулевое слагаемое вида $g_{ii}(x^i)^2$ (без ограничения общности можно считать, что $i=1$), тогда выделим полный квадрат, содержащий $(x^1)^2$:

$$\begin{aligned} g(x,x) &= g_{11}\Big((x^1)^2+\frac{2g_{12}}{g_{11}}x^1x^2+\ldots+\frac{2g_{1n}}{g_{11}}x^1x^n\Big)+\text{слагаемые, не содержащие } x^1 = \\ &= g_{11}\Big(x^1+\frac{g_{12}}{g_{11}}x^2+\ldots+\frac{g_{1n}}{g_{11}}x^n\Big)^2+\text{слагаемые, не содержащие } x^1. \end{aligned}$$

Сделав замену координат $\widetilde{x}^1=\sqrt{|g_{11}|}\Big(x^1+\frac{g_{12}}{g_{11}}x^2+\ldots+\frac{g_{1n}}{g_{11}}x^n\Big)$, получим $g(x,x)=\pm(\widetilde{x}^1)^2+\ldots$. С оставшимися слагаемыми, не содержащими x^1, можно проделать то же самое, в итоге получим, что $g(x,x)=\pm(\widetilde{x}^1)^2\pm(\widetilde{x}^2)^2\pm\ldots\pm(\widetilde{x}^k)^2$, т.е. в новом базисе матрица функции g будет диагональной.

Итак, мы знаем, что для любой билинейной вещественной симметрической функции в векторном пространстве W существует базис, в котором она имеет нормальный вид $g(x,y)=x^1y^1+\ldots+x^py^p-x^{p+1}y^{p+1}-\ldots-x^{p+q}y^{p+q}$.

Аналогично, комплексная билинейная симметрическая функция имеет нормальный вид $g(x,y)=x^1y^1+\ldots+x^ky^k$.

Нормальный вид эрмитовой функции — $g(x,y)=\overline{x}^1y^1+\ldots+\overline{x}^py^p-\overline{x}^{p+1}y^{p+1}-\ldots-\overline{x}^{p+q}y^{p+q}$.

А кососимметрическая функция имеет нормальный вид $g(x,y)=x^1y^2-x^2y^1+x^3y^4-x^4y^3+\ldots+x^{2i-1}y^{2i}-x^{2i}y^{2i-1}$.

5.6 Единственность нормального вида

Теперь обсудим вопрос о единственности нормального вида. Для комплексной билинейной симметричной функции (соотв. для кососимметричной функции) нормальный вид определяется единственным числом — k (соотв. $2i$), которое равно рангу функции, поэтому не зависит от способа приведения к нормальному виду. Оставшиеся два случая похожи друг на друга, поэтому обсудим один из них — случай вещественной симметричной билинейной функции. Нормальный вид ее матрицы

$$G=\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & \mathbf{0} & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & \mathbf{0} & & & & & & & 0 \end{pmatrix}.$$

Введем в рассмотрение три числа: p — количество единиц q — количество минус единиц, s — количество нулей. Видно, что $p+q=\operatorname{rk}g$, $s=\dim W-\operatorname{rk}g$ следовательно, $p+q$ и s не зависят от выбора базиса.

Определение 5.6.1 Набор чисел (p,q,s) называется *сигнатурой* вещественной симметричной билинейной функции (или квадратичной функции).

Если функция g невырождена, то $s=0$, и для определения p и q достаточно знать хотя бы одно из них или их разность. Поэтому в случае невырожденной функции сигнатурой часто называют число $p-q$.

Теорема 5.6.2 (инерции) *Если вещественная симметричная билинейная функция в пространстве W приведена к нормальному виду двумя разными способами, то числа p, q и s одни и те же.*

Доказательство. Пусть функция g имеет нормальный вид в двух базисах: $e_1, \dots, e_p, e_{p+1}, \dots, e_{p+q}, e_{p+q+1}, \dots, e_{p+q+s}$ и $\widetilde{e}_1, \dots, \widetilde{e}_{p'}, \widetilde{e}_{p'+1}, \dots, \widetilde{e}_{p'+q'}, \widetilde{e}_{p'+q'+1}, \dots, \widetilde{e}_{p'+q'+s'}$, причем, как мы уже знаем, $s' = s$ и $p' + q' = p + q$. Рассмотрим в W подпространства

$$V_+ = \langle e_1, \dots, e_p \rangle,\ V_- = \langle e_{p+1}, \dots, e_{p+q} \rangle,\ V_0 = \langle e_{p+q+1}, \dots, e_{p+q+s} \rangle$$

и

$$\widetilde{V}_+ = \langle \widetilde{e}_1, \dots, \widetilde{e}_{p'} \rangle,\ \widetilde{V}_- = \langle \widetilde{e}_{p'+1}, \dots, \widetilde{e}_{p'+q'} \rangle,\ \widetilde{V}_0 = \langle \widetilde{e}_{p'+q'+1}, \dots, \widetilde{e}_{p'+q'+s} \rangle.$$

Если $x \in V_+$ и $x \neq 0$, то $g(x,x) > 0$ (неравенство строгое, так как ограничение функции g на V_+ невырождено). Действительно, если $x = x^1 e_1 + \ldots + x^p e_p$, то $g(x,x) = (x^1)^2 + \ldots + (x^p)^2 > 0$. Аналогично, если $x \in V_- \oplus V_0$, то $g(x,x) \leqslant 0$. Такая же ситуация имеет место с подпространствами $\widetilde{V}_+$, $\widetilde{V}_-$ и $\widetilde{V}_0$.

Пусть $p > p'$, тогда $\dim V_+ = p$, $\dim(\widetilde{V}_- \oplus \widetilde{V}_0) = \dim \widetilde{V}_- + \dim \widetilde{V}_0 = q' + s = \dim W - p'$. Следовательно,

$$\dim V_+ + \dim(\widetilde{V}_- \oplus \widetilde{V}_0) = p + \dim W - p' > \dim W,$$

значит, пересечение этих подпространств нетривиально, $V_+ \cap (\widetilde{V}_- \oplus \widetilde{V}_0) \neq \{0\}$. Возьмем произвольный ненулевой вектор их этого пересечения, $x \in V_+ \cap (\widetilde{V}_- \oplus \widetilde{V}_0)$. Т.к. $x \in V_+$, $x \neq 0$, то $g(x,x) > 0$, но, с другой стороны, т.к. $x \in \widetilde{V}_- \oplus \widetilde{V}_0$, то $g(x,x) \leqslant 0$. Получили противоречие. Случай $p < p'$ рассматривается аналогично. □

5.7 Теорема Якоби. Критерий Сильвестра

Сигнатура и нормальный вид симметричной билинейной (или квадратичной) функции могут быть определены и без нахождения в явном виде замены координат.

Напомним, что угловым минором порядка k квадратной матрицы называется минор, составленный из первых k строк и k столбцов.

Теорема 5.7.1 (Якоби) *Пусть G — матрица симметричной билинейной функции g в некотором базисе. Пусть все угловые миноры до порядка $r = \operatorname{rk} G$ отличны от 0. Тогда существует базис, в котором функция g имеет вид $g(x,y) = |G_1| x^1 y^1 + \frac{|G_2|}{|G_1|} x^2 y^2 + \ldots + \frac{|G_r|}{|G_{r-1}|} x^r y^r$, где $|G_i|$ — i-й угловой минор.*

Доказательство. Будем искать новый базис в виде

$$\begin{aligned} e'_1 &= e_1; \\ e'_2 &= e_2 + c_2^1 e'_1; \\ e'_3 &= e_3 + c_3^1 e'_1 + c_3^2 e'_2; \\ \dots &\quad \dots\dots\dots \\ e'_n &= e_n + c_n^1 e'_1 + \ldots + c_n^{n-1} e'_{n-1}, \end{aligned}$$

где $e_1, \dots, e_n$ — тот базис, в котором нам дана матрица функции g. Такое преобразование удобно тем, что (также как в обычном процессе ортогонализации) для любого номера k линейные оболочки векторов $e_1, \dots, e_k$ и векторов $e'_1, \dots, e'_k$ совпадают. Матрица перехода при таком преобразовании верхнетреугольна, т.е. имеет вид $C = \begin{pmatrix} 1 & & \star \\ & \ddots & \\ \mathbf{0} & & 1 \end{pmatrix}$, поэтому ее определитель равен 1. Кроме того, поскольку изменение матрицы G при такой замене координат сводится

к элементарным преобразованиям строк и столбцов, определители угловых миноров не изменяются (иначе это можно показать так: если мы ограничимся первыми k координатами и пространством $\langle e_1, \ldots, e_k \rangle$, то матрица перехода C_k также будет иметь единичный определитель, и матрица углового минора изменится по формуле $G'_k = C_k^t G_k C_k$, поэтому $|G'_k| = |C_k^t||G_k||C_k| = |G_k|$).

Коэффициенты c_i^j будем искать следующим образом. Сначала мы их найдем для индексов $j \leqslant r$, где $r = \operatorname{rk} G$. Коэффициент c_2^1 определим из равенства $g(e_1, e'_2) = g(e'_1, e'_2) = 0$. Это равенство дает уравнение $g(e_1, e_2) + c_2^1 g(e_1, e_1) = 0$, которое имеет решение, поскольку по предположению $g(e_1, e_1) = |G_1| \neq 0$. Аналогично можно найти c_3^1 и c_3^2 из условия $g(e'_3, e'_1) = g(e'_3, e'_2) = 0$, затем — c_4^1, c_4^2 и c_4^3 и т.п. пока нижний индекс не станет равным r. Действительно, если мы уже нашли векторы $e'_2, \ldots, e'_{k-1}$, то коэффициенты c_k^i для вектора e'_k ищутся из условия $g(e'_k, e'_1) = \ldots = g(e'_k, e'_{k-1}) = 0$, которое дает систему линейных уравнений

$$\begin{cases} c_k^1 g(e'_1, e'_1) + \ldots + c_k^{k-1} g(e'_{k-1}, e'_1) &= -g(e_k, e'_1) \\ \ldots\ldots\ldots\ldots\ldots\ldots & \ldots \\ c_k^1 g(e'_1, e'_{k-1}) + \ldots + c_k^{k-1} g(e'_{k-1}, e'_{k-1}) &= -g(e_k, e'_{k-1}) \end{cases}.$$

Эта система имеет решение, так как ее матрица невырождена — ее определитель равен $|G_{k-1}|$ (если бы мы не заменили векторы e_i на e'_i, то матрица системы просто совпадала бы с $k-1$-м угловым минором, а при заменах указанного вида определитель угловых миноров не изменяется). Таким образом, мы можем найти векторы $e'_1, \ldots, e'_r$ нового базиса. При этом матрица функции g примет вид

$$\left(\begin{array}{ccc|c} g(e'_1, e'_1) & & & \\ & \ddots & & \star \\ & & g(e'_r, e'_r) & \\ \hline & \star & & \star \end{array}\right),$$

где левый верхний угол — диагональная матрица. Отметим, что ранг этой диагональной матрицы равен r (ранг углового минора не меняется), поэтому все значения $g(e'_i, e'_i)$ отличны от нуля при $1 \leqslant i \leqslant r$.

Оставшиеся векторы $e'_{r+1}, \ldots, e'_n$ найдем из условия $g(e'_i, e'_j) = 0$, если $i \leqslant r$ а $j > r$. Для этого надо взять $e'_j = e_j - \frac{g(e_j, e'_1)}{g(e'_1, e'_1)} e'_1 - \ldots - \frac{g(e_j, e'_r)}{g(e'_r, e'_r)} e'_r$. Теперь матрица функции g примет вид

$$\left(\begin{array}{ccc|c} g(e'_1, e'_1) & & & \\ & \ddots & & \mathbf{0} \\ & & g(e'_r, e'_r) & \\ \hline & \mathbf{0} & & \star \end{array}\right),$$

а, поскольку ранг этой матрицы равен r, т.е. рангу углового минора порядка r, то правый нижний угол этой матрицы равен нулю. Окончательно, в построенном базисе матрица функции g примет вид

$$G' = \left(\begin{array}{ccc|ccc} g(e'_1, e'_1) & & & & & \\ & \ddots & & & \mathbf{0} & \\ & & g(e'_r, e'_r) & & & \\ \hline & & & 0 & & \\ & \mathbf{0} & & & \ddots & \\ & & & & & 0 \end{array}\right).$$

Поскольку $|G'_k| = |G_k|$ для всех $k = 1, \ldots, n$, имеем равенства

$$
\begin{aligned}
|G_1| = |G'_1| &= g(e'_1, e'_1); \\
|G'_2| = |G_2| = g(e'_1, e'_1)g(e'_2, e'_2) = |G_1|g(e'_2, e'_2) &\Leftarrow g(e'_2, e'_2) = \frac{|G_2|}{|G_1|}; \\
|G'_3| = |G_3| = |G'_2|g(e'_3, e'_3) = |G_2|g(e'_3, e'_3) &\Leftarrow g(e'_3, e'_3) = \frac{|G_3|}{|G_3|}; \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
|G'_r| = |G_r| = |G'_{r-1}|g(e'_r, e'_r) = |G_{r-1}|g(e'_r, e'_r) &\Leftarrow g(e'_r, e'_r) = \frac{|G_r|}{|G_{r-1}|}.
\end{aligned}
$$

В построенном базисе $e'_1, \ldots, e'_n$ функция g имеет вид

$$
\begin{aligned}
g(x, y) &= g(e'_1, e'_1)x^1y^1 + \ldots + g(e'_r, e'_r)x^ry^r = \\
&= |G_1|x^1y^1 + \frac{|G_2|}{|G_1|}x^2y^2 + \ldots + \frac{|G_r|}{|G_{r-1}|}x^ry^r,
\end{aligned}
$$

что и требовалось доказать. □

Определение 5.7.2 Симметричная билинейная функция g в вещественном пространстве V называется *положительно определенной*, если ее сигнатура равна $(\dim V, 0, 0)$ (эквивалентное определение — если $g(x, x) > 0$ для любого ненулевого $x \in V$).

Теорема 5.7.3 (Критерий Сильвестра) *Функция g положительно определена тогда и только тогда, когда в ее матрице G (в произвольном базисе) все угловые миноры положительны.*

Доказательство. Если все угловые миноры положительны, то по предыдущей теореме функция g имеет вид $g(x, x) = \lambda_1(x^1)^2 + \ldots + \lambda_n(x^n)^2$, где n — размерность пространства, а $\lambda_i > 0$ при всех $i = 1, \ldots, n$, следовательно, g положительно определена.

Обратно, предположим, что функция g положительно определена. Докажем сначала, что все угловые миноры будут ненулевыми. Отметим сразу, что $|G| = |G_n| \neq 0$, т.к. функция g невырождена и ее ранг равен n. Поскольку функция g положительно определена, то и ее ограничения на любые подпространства также будут положительно определенными, а, следовательно, и невырожденными. Т.к. G_k — это матрица ограничения g на $\langle e_1, \ldots, e_k \rangle$, то $|G_k| \neq 0$ для всех k. Следовательно, все угловые миноры ненулевые. Воспользуемся предыдущей теоремой: чтобы функция $g(x, y) = \lambda_1x^1y^1 + \ldots + \lambda_nx^ny^n$ была невырожденной (т.е. чтобы все λ_i были положительными) необходимо, чтобы все $|G_i|$ были положительными. □

5.8 Пространства с обобщенным скалярным произведением. Группы операторов, сохраняющих скалярное произведение

Определение 5.8.1 Вещественное пространство с заданной на нем невырожденной симметричной билинейной функцией называется *псевдоевклидовым* пространством.

Комплексное пространство с заданной на нем невырожденной эрмитовой полуторалинейной функцией называется *псевдоэрмитовым* пространством.

Вещественное пространство с заданной на нем невырожденной кососимметричной билинейной функцией называется *симплектическим* пространством.

Первые два случая являются естественными обобщениями евклидового и эрмитового пространств. Скалярное (обобщенное) произведение задается симметричной билинейной или эрмитовой полуторалинейной функцией g, $(x, y) = g(x, y) = x^1y^1 + \ldots + x^py^p - x^{p+1}y^{p+1} - \ldots - x^ny^n$ (в псевдоэрмитовом случае добавляется комплексное сопряжение). Если в нормальном виде функции g все знаки — плюсы (т.е. если g положительно определена) то псевдоевклидово пространство является просто евклидовым, а псевдоэрмитово пространство — эрмитовым (унитарным). Матрица G функции g является обобщением матрицы Грама.

В псевдоевклидовом и псевдоэрмитовом случае будем говорить, что базис $e_1, \ldots, e_n$ ортонормирован, если $g(e_i, e_j) = \begin{cases} 0 & \text{при } i \neq j \\ \pm 1 & \text{при } i = j \end{cases}$. В симплектическом случае такое невозможно, т.к. $g(x, x) = 0$ для любого вектора x. Однако невырожденная кососимметричная функция имеет вид $\begin{pmatrix} \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} \end{pmatrix}$ (т.е. без клеток размера 1, т.к. наличие нулевой одномерной клетки противоречит невырожденности функции g). Следовательно, размерность четна. Пусть матрица функции g имеет указанный вид в базисе $e_1, \ldots, e_{2m}$. Рассмотрим базис $e_1, e_3, \ldots, e_{2m-1}, e_2, e_4, \ldots, e_{2m}$, в этом базисе матрица функции g имеет вид $\left(\begin{array}{ccc|ccc} & & & 1 & & \\ & \mathbf{0} & & & \ddots & \\ & & & & & 1 \\ \hline -1 & & & & & \\ & \ddots & & & \mathbf{0} & \\ & & -1 & & & \end{array}\right)$. Такой базис называется *гамильтоновым.* В симплектическом пространстве значение функции g на паре векторов также называют их скалярным произведением (еще более обобщенным, чем псевдоевклидовое или псевдоэрмитово).

Рассмотрим теперь операторы, действующие в псевдоевклидовом, псевдоэрмитовом или симплектическом пространстве V. Пусть оператор $f : V \to V$ сохраняет скалярное произведение, т.е. $g(fx, fy) = g(x, y)$ для всех $x, y \in V$.

Лемма 5.8.2 *Все операторы, сохраняющие скалярное произведение, образуют группу.*

Доказательство. Очевидно, что если операторы f_1, f_2 сохраняют скалярное произведение, то и их композиция также его сохраняет: $g(f_1 f_2 x, f_1 f_2 x) = g(f_2 x, f_2 y) = g(x, y)$. Если обратный оператор существует, то он также сохраняет скалярное произведение: $g(x, y) = g(ff^{-1}x, ff^{-1}y) = g(f^{-1}x, f^{-1}y)$.

Нам осталось показать, почему обратный оператор существует для любого оператора f, сохраняющего скалярное произведение. Возьмем какой-нибудь базис $e_1, \ldots, e_n$, пусть $A = (a^i_j)$ — матрица оператора f, $G = g_{ij}$ — матрица функции g в выбранном базисе. Тогда условие сохранения скалярного произведения примет вид:

$$g(x, y) = g_{ij} x^i y^j = g(fx, fy) = g_{kl} a^k_i x^i a^l_j y^j = (A^t G A)_{ij} x^i y^j,$$

т.е. $G = A^t G A$ (в псевдоэрмитовом случае надо еще учесть комплексное сопряжение, $G = \overline{A}^t G A$). Т.к. матрица G невырождена, то и A обязана быть невырожденной, следовательно, существует и A^{-1}, которой соответствует (в том же базисе) оператор f^{-1}. □

Если G — единичная матрица, то $A^t A = E$ (в псевдоэрмитовом случае $\overline{A}^t A = E$), и эти условия выделяют ортогональные (унитарные) матрицы.

Введем в рассмотрение следующие группы операторов:

- *псевдоортогональная группа* $O(p, q)$ — группа операторов, сохраняющих псевдоевклидовое скалярное произведение с сигнатурой $(p, q, 0)$ (если $q = 0$, то группу $O(p, 0)$ называют ортогональной и обозначают $O(p)$).
- *псевдоунитарная группа* $U(p, q)$ — группа операторов, сохраняющих псевдоэрмитовое скалярное произведение с сигнатурой $(p, q, 0)$ (если $q = 0$, то группу $O(p, 0)$ называют унитарной и обозначают $U(p)$).
- *симплектическая группа* $Sp(2m)$ — группа операторов, сохраняющих симплектическое скалярное произведение в пространстве размерности $2m$.

Дадим описание этих групп в малых размерностях.

Группа $O(2)$ — группа операторов, сохраняющих скалярное произведение на плоскости. Такие преобразования могут собственными или несобственными (сохранять ориентацию плоскости или менять ее). Если оператор сохраняет ориентацию, то это просто поворот на некоторый угол φ и его матрица имеет вид $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$, а если оператор меняет ориентацию, то это еще и композиция с симметрией, и матрица имеет вид $\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$.

Группа $O(1,1)$. В этом случае скалярное произведение задано формулой $g(x,y) = x^1y^1 - x^2y^2$ и вектор с координатами x, y будет иметь длину ±1, если $x^2 - y^2 = \pm1$, т.е. если его конец лежит на одной из гипербол, $x^2 - y^2 = 1$ или $x^2 - y^2 = -1$. В этом случае стандартный базис $e_1 = (1,0)$, $e_2 = (0,1)$ также будет ортонормированным, но при повороте вектора e_1 против часовой стрелки, вектор e_2 будет поворачиваться по часовой стрелке и новый ортонормированный базис e_1', e_2' будет симметричен относительно прямой $y = x$. В этом случае матрица A оператора, сохраняющего скалярное произведение, должна удовлетворять условию $A^t \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Здесь будет уже не два, как в случае $O(2)$, а четыре различных класса операторов:

$$\begin{pmatrix} \operatorname{ch}\varphi & \operatorname{sh}\varphi \\ \operatorname{sh}\varphi & \operatorname{ch}\varphi \end{pmatrix}, \quad \begin{pmatrix} -\operatorname{ch}\varphi & -\operatorname{sh}\varphi \\ \operatorname{sh}\varphi & \operatorname{ch}\varphi \end{pmatrix}, \quad \begin{pmatrix} \operatorname{ch}\varphi & \operatorname{sh}\varphi \\ -\operatorname{sh}\varphi & -\operatorname{ch}\varphi \end{pmatrix}, \quad \begin{pmatrix} -\operatorname{ch}\varphi & -\operatorname{sh}\varphi \\ -\operatorname{sh}\varphi & -\operatorname{ch}\varphi \end{pmatrix}.$$

Поэтому иногда говорят, что псевдоортогональная группа состоит из четырех компонент.

Рассмотрим теперь симплектическую группу $Sp(2)$. Матрица скалярного произведения имеет вид $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Матрица A оператора, сохраняющего скалярное произведение, должна удовлетворять условию $A^t \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, что равносильно условию $\det A = 1$ (проверьте!). Таким образом, группа $Sp(2)$ совпадает с группой матриц, имеющих единичный определитель, $SL_2(\mathbb{R})$. Однако при больших размерностях пространства симплектические группы не совпадают ни с какими, уже известными нам.

5.9 Связь между псевдоортогональной, псевдоунитарной и симплектической группами

Теорема 5.9.1 *$U(p,q) = O(2p,2q) \cap Sp(2n)$, где $n = p + q$.*

Доказательство. Пусть V — псевдоэрмитово пространство, функцию g, задающую псевдоэрмитово скалярное произведение, можно разложить на вещественную и чисто мнимую часть, т.е. $g = g_{\mathrm{Re}} + ig_{\mathrm{Im}}$, где $g_{\mathrm{Re}}(x,y) = \operatorname{Re} g(x,y)$ и $g_{\mathrm{Im}}(x,y) = \operatorname{Im} g(x,y)$. Далее для доказательства нам потребуется

Теорема 5.9.2 *Функция g_{Re} определяет на $V_{\mathbb{R}}$ псевдоевклидово скалярное произведение с удвоенной сигнатурой, а функция g_{Im} определяет на $V_{\mathbb{R}}$ симплектическое скалярное произведение. Если $e_1, \ldots, e_n$ — ортонормированный базис в V, тогда $e_1, \ldots, e_n, ie_1, \ldots, ie_n$ является одновременно ортонормированным базисом для g_{Re} и (с точностью до перестановки) гамильтоновым базисом для g_{Im}.*

Доказательство. Функции g_{Re} и g_{Im}, очевидно, билинейны над полем $\mathbb{R}$. Поскольку

$$g_{\mathrm{Re}}(x,y) = \operatorname{Re} g(x,y) = \operatorname{Re} \overline{g(y,x)} = \operatorname{Re} g(y,x) = g_{\mathrm{Re}}(y,x),$$
$$g_{\mathrm{Im}}(x,y) = \operatorname{Im} g(x,y) = \operatorname{Im} \overline{g(y,x)} = -\operatorname{Im} g(y,x) = -g_{\mathrm{Im}}(y,x),$$

функция g_{Re} симметрична, а функция g_{Im} кососимметрична.

Проверим теперь, что базис $e_1, \ldots, e_n, ie_1, \ldots, ie_n$ является ортонормированным для g_{Re} и гамильтоновым для g_{Im}. Действительно, $g_{\mathrm{Re}}(e_j, e_k) = \operatorname{Re} g(e_j, e_k) = \pm\delta_{jk}$, $g_{\mathrm{Re}}(ie_j, ie_k) =$

$\operatorname{Re} g(ie_j, ie_k) = \operatorname{Re} g(e_j, e_k) = \pm\delta_{jk}$, $g_{\mathrm{Re}}(e_j, ie_k) = \operatorname{Re} g(e_j, ie_k) = \operatorname{Re} ig(e_j, e_k) = \operatorname{Re} \pm i = 0$. Таким образом, этот базис действительно будет ортонормированным, причем сигнатура равна $(2p, 2q)$, т.к. на каждый базисный вектор e_k псевдоэрмитова пространства с $g(e_k, e_k) = 1$ приходится два вектора, e_k и ie_k с $g_{\mathrm{Re}}(e_k, e_k) = g_{\mathrm{Re}}(ie_k, ie_k) = 1$.

Аналогично $g_{\mathrm{Im}}(e_j, e_k) = g_{\mathrm{Im}}(ie_j, ie_k) = \operatorname{Im} g(e_j, e_k) = 0$ и $g_{\mathrm{Im}}(ie_j, e_k) = \operatorname{Im} g(ie_j, e_k) = -g(e_j, e_k) = \pm\delta_{jk}$. Если $g(e_j, e_k = 1)$, то нам надо поменять местами векторы ie_j и e_k. После такой перестановки матрица функции g_{Im} будет иметь вид $\left(\begin{array}{c|c} 0 & E \\ \hline -E & 0 \end{array}\right)$. □

Равенство $U(p, q) = O(2p, 2q) \cap Sp(2n)$ является следствием этой теоремы. Докажем включения в обе стороны. Пусть $f : V \to V$, $f \in U(p, q)$, т.е. $g(fx, fy) = g(x, y)$ для всех $x, y \in V$. Тогда f сохраняет и вещественную часть функции g, т.е. g_{Re}. Следовательно, $f \in O(2p, 2q)$. Аналогично, f сохраняет мнимую часть g, g_{Im}, следовательно, $f \in Sp(2n)$. Поэтому $U(p, q) \subseteq O(2p, 2q) \cap Sp(2n)$.

Обратно, пусть $f : V_{\mathbb{R}} \to V_{\mathbb{R}}$, причем $f \in O(2p, 2q) \cap Sp(2n)$, т.е. f сохраняет g_{Re} и g_{Im}. Т.к. f сохраняет и вещественную, и мнимую часть функции g, то он сохраняет и саму g, т.е. $g(fx, fy) = g(x, y)$ для любых $x, y \in V$. Покажем, что f является линейным оператором над полем $\mathbb{C}$ (а не только над $\mathbb{R}$). Оператор f обратим, поэтому для любого вектора y существует прообраз — вектор z, для которого $fz = y$. Чтобы доказать, что $ifx = f(ix)$, вычислим $g(ifx - f(ix), y)$:

$$g(ifx - f(ix), y) = -ig(fx, fz) - g(f(ix), fz) = -ig(x, z) + ig(x, z) = 0,$$

т.е. мы получили, что $g(ifx - f(ix), y) = 0$ для любого вектора y. В силу невырожденности g, получаем $ifx = f(ix)$, т.е. линейность над полем $\mathbb{C}$. Следовательно $f \in U(p, q)$, значит, $O(2p, 2q) \cap Sp(2n) \subseteq U(p, q)$. □

5.10 Симплектические пространства. Лагранжевы подпространства

Приступим к более подробному рассмотрению симплектических скалярных произведений.

Определение 5.10.1 Вектор x называется *изотропным*, если $g(x, x) = 0$.

Пусть W — симплектическое пространство с невырожденной кососимметричной билинейной функцией g.

Определение 5.10.2 Подпространство $V \subset W$ называется *изотропным*, если $V \subset V^{\perp}$.

Очевидно, что если V изотропно, то $g(x, y) = 0$ для всех $x, y \in V$.

Лемма 5.10.3 *Если подпространство $V \subset W$ изотропно, то $\dim V \leqslant \frac{1}{2} \dim W$, причем изотропное подпространство максимальной размерности $\frac{\dim W}{2}$ существует.*

Доказательство. Так как функция g невырождена, то $\dim V^{\perp} = \dim W - \dim V$, причем, поскольку $V \subset V^{\perp}$, то $\dim V^{\perp} \geqslant \dim V$. Следовательно, $\dim W = \dim V + \dim V^{\perp} \geqslant 2 \dim V$, откуда $\dim V \leqslant \frac{\dim W}{2}$.

Возьмем гамильтонов базис $e_1, \ldots, e_m, e_{m+1}, \ldots, e_{2m}$, где $m = \frac{\dim W}{2}$, и пусть $V = \langle e_1, \ldots, e_m \rangle$. Т.к. этот базис гамильтонов, то в этом базисе матрица g имеет вид $\left(\begin{array}{c|c} 0 & E \\ \hline -E & 0 \end{array}\right)$, а матрица ограничения $g|_V$ функции g на V является просто нулевой. Следовательно, $g(x, y) = 0$ ля любых $x, y \in V$, значит, V изотропно и имеет размерность $\dim V = \frac{\dim W}{2}$. □

Определение 5.10.4 Изотропное подпространство максимальной размерности называется *лагранжевым.*

Следующее утверждение показывает, что любое изотропное подпространство можно вложить в лагранжево.

Лемма 5.10.5 *Для любого изотропного подпространства $V \subset W$ симплектического пространства существует такое лагранжево подпространство $\widetilde{V}$, что $V \subset \widetilde{V}$.*

Отметим, что подпространство $\widetilde{V}$, указанное в лемме, не единственно.

Доказательство. Пусть $\dim V = r \leqslant m = \frac{\dim W}{2}$, тогда, в силу невырожденности функции g на W, $\dim V^\perp = 2m - r$ и $V \subset V^\perp$. Обозначим через g' ограничение функции g на $V^\perp$. Очевидно, функция g' вырождена. Действительно, т.к. g невырождена, то $(V^\perp)^\perp$ и $\operatorname{Ker} g' = V^\perp \cap (V^\perp)^\perp = V$. Приведем g' к нормальному виду и получим, что в некотором базисе $\underbrace{e_1, \ldots, e_{m-r}}_{m-r}, \underbrace{e_{m-r+1}, \ldots, e_{2m-2r}}_{m-r}, \underbrace{e_{2m-2r+1}, \ldots, e_{2m-r}}_{r}$ ее матрица будет иметь вид $\left(\begin{array}{c|c|c} 0 & E & 0 \\ \hline -E & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$. Т.к. $V = \operatorname{Ker} g$, то ему отвечает последняя группа векторов, т.е. $V = \langle e_{2m-2r+1}, \ldots, e_{2m-r} \rangle$. Добавив к V векторы $e_{m-r+1}, \ldots, e_{2m-2r}$, получим лагранжево подпространство $\widetilde{V}$. Оно, очевидно, содержит V и действительно является лагранжевым, т.к. функция g тождественно равна нулю на $\widetilde{V}$ (в матрице ограничению g на $\widetilde{V}$ отвечает правый нижний квадрат, состоящий из нулей) и $\dim \widetilde{V} = r + (m - r) = m$. $\square$

Теорема 5.10.6 *Пусть $V \subset W$ — лагранжево подпространство, тогда существует другое лагранжевое подпространство $\widetilde{V}$, такое что $W = V \oplus \widetilde{V}$.*

Доказательство. Возьмем произвольный базис $e_1, \ldots, e_m$ в V, где $m = \frac{\dim W}{2}$, и дополним его до базиса всего пространства W векторами $e_{m+1}, \ldots, e_{2m}$. Вспомним про канонический изоморфизм $B(W) \cong L(W, W')$. Функции g ставится в соответствие линейное отображение $\widetilde{g} : W \to W'$, причем, т.к. g невырождена, то $\widetilde{g}$ обратим, т.е. существует обратное отображение $\widetilde{g}^{-1} : W' \to W$. Возьмем двойственный базис $\varepsilon^1, \ldots, \varepsilon^m, \varepsilon^{m+1}, \ldots, \varepsilon^{2m}$ в W', применим к нему отображение $\widetilde{g}^{-1}$ и обозначим $e_i^0 = \widetilde{g}^{-1}(\varepsilon^i)$ для $i = 1, \ldots, m$. Векторы $e_1^0, \ldots, e_m^0, e_1, \ldots, e_m$ образуют базис в W, для доказательства этого достаточно проверить их линейную независимость, так как их количество совпадает с размерностью.

Пусть $\lambda_1 e_1^0 + \ldots + \lambda_m e_m^0 + \mu_1 e_1 + \ldots + \mu_m e_m = 0$, следовательно, $g(\lambda_1 e_1^0 + \ldots + \mu_m e_m, e_i) = 0$ для всех $i = 1, \ldots, m$. Т.к. V изотропно, то $g(e_k, e_i) = 0$ для всех значений индексов, а $g(e_k^0, e_i) = \widetilde{g}(e_k^0)(e_j) = \widetilde{g}(\widetilde{g}^{-1}(\varepsilon^k))(e_i) = \varepsilon^k(e_i) = \delta_i^k$. Поэтому $g(\lambda_1 e_1^0 + \ldots + \mu_m e_m, e_i) = \lambda_i$, следовательно все $\lambda_i = 0$. Поэтому $\mu_1 e_1 + \ldots + \mu_m e_m = 0$, но $e_1, \ldots, e_m$ — это базис в V, поэтому они линейно независимы и все $\mu_i = 0$. Следовательно, набор векторов $e_1^0, \ldots, e_m^0, e_1, \ldots, e_m$ действительно будет базисом в W. В этом базисе матрица функции g имеет вид $\left(\begin{array}{c|c} \star & E \\ \hline -E & 0 \end{array}\right)$ (в правом верхнем углу будет E, а в нижем левом будет $-E$ потому, что $g(e_i^0, e_j) = -g(e_j, e_i^0) = \delta_{ij}$, а в правом нижнем углу будет нулевая матрица, т.к. $g(e_i, e_j) = 0$).

Чтобы избавиться от ненулевого верхнего левого угла, перейдем к новому базису. Векторы $e_1, \ldots, e_m$ оставим без изменений, а вместо векторов $e_1^0, \ldots, e_m^0$ рассмотрим векторы

$$\begin{aligned} \overline{e}_1 &= e_1^0, \\ \overline{e}_2 &= e_2^0 + c_2^1 e_1, \\ \overline{e}_3 &= e_3^0 + c_3^1 e_1 + c_3^2 e_2, \\ \ldots &\quad \ldots\ldots\ldots\ldots\ldots \\ \overline{e}_m &= e_m^0 + c_m^1 e_1 + \ldots + c_m^{m-1} e_{m-1}. \end{aligned}$$

При такой замене все углы матрицы, кроме левого верхнего, не изменятся, поскольку $g(\overline{e}_i, e_j) = g(e_j^0, e_i)$.

Подберем коэффициенты c_j^i так, чтобы получилось $g(\overline{e}_i, \overline{e}_i) = 0$ для всех индексов. Т.к. $g(\overline{e}_2, \overline{e}_1) = g(\overline{e}_2, e_1^0) = g(e_2^0, e_1^0) + c_2^1 g(e_1, e_1^0) = 0$, то коэффициент c_2^1 можно найти из этого равенства. Далее, условия $g(\overline{e}_3, \overline{e}_1) = 0$ и $g(\overline{e}_3, \overline{e}_2) = 0$ дают нам систему уравнений на неизвестные c_3^1 и c_3^2 с невырожденной (на самом деле, единичной) матрицей, следовательно, и эти коэффициенты можно найти. Прожолжая эту процедуру, находим все коэффициенты c_j^i. В базисе

$\overline{e}_1, \ldots, \overline{e}_m, e_1, \ldots, e_m$ матрица функции g имеет стандартный вид $\left(\begin{array}{c|c} 0 & E \\ \hline -E & 0 \end{array}\right)$, поэтому подпространство $\widetilde{V} = \langle \overline{e}_1, \ldots, \overline{e}_m \rangle$ будет лагранжевым (так как $g(\overline{e}_i, \overline{e}_j) = 0$ для всех индексов). Также очевидно, что $W = V \oplus \widetilde{V}$. □

5.11 Симметрические билинейные функции на евклидовом пространстве

Приступим к рассмотрению симметрических функций на евклидовом пространстве. Пусть нам задано такое пространство V со скалярным произведением $(\cdot, \cdot)$. Раньше мы установили следующие два канонических изоморфизма:

1) если V — евклидово пространство, то $V \cong V'$. Этот изоморфизм в сторону $V \to V'$ можно записать, например, следующим образом: $a \in V \mapsto (a, \cdot) \in V'$.

2) $B(V) \cong L(V, V')$ для любого пространства V (здесь евклидовость ни при чем). Этот изоморфизм можно записать так: $B(V) \ni g \mapsto \widetilde{g} \in L(V, V')$, где $\widetilde{g}(a) \in V'$ определяется соотношением $\widetilde{g}(a)(b) = g(a, b)$.

Из этих двух изоморфизмов видно, что для евклидова пространства имеем канонический изоморфизм $B(V) \cong L(V, V) = L(V)$, т.е. мы избавляемся от "лишнего" пространства V'. Проще всего этот изоморфизм будет записать в обратную сторону, т.е. $L(V) \to B(V)$. Пусть $f \in L(V)$, т.е. $f : V \to V$, тогда оператору f можно поставить в соответствие билинейную функцию $g_f(a, b) = (fa, b)$. Отображение $f \mapsto g_f$, $L(V) \to B(V)$, задает изоморфизм. Действительно, т.к. размерности пространств $L(V)$ и $B(V)$ совпадают, то достаточно проверить лишь то, что у этого отображения нулевое ядро, т.е. проверить, что если $g_f = 0$, то и $f = 0$. Если $(fa, b) = 0$ для всех $a, b \in V$, то, зафиксировав a, получим, что $(fa, b) = 0$ для любого b, следовательно, $fa = 0$. Но т.к. это равенство выполняется для всех a, то $f = 0$.

Отсюда следует существование обратного отображения $B(V) \to L(V)$, $B(V) \ni g \mapsto f_g \in L(V)$, где оператор f_g определяется равенством $g(a, b) = (f_g a, b)$.

Лемма 5.11.1 *1) Если функция g симметрична, то f_g — самосопряженный оператор.*
2) Если f — самосопряженный оператор, то g_f — симметричная функция.

Доказательство. 1) Пусть $g(a, b) = g(b, a)$ для всех $a, b \in V$, тогда $(f_g b, a) = (f_g a, b) = (b, f_g a)$, что означает самосопряженность.

2) Если $(fa, b) = (a, fb)$ для всех $a, b \in V$, то $g_f(a, b) = (fa, b) = (a, fb) = (fb, a) = g_f(b, a)$, т.е. g_f симметрична. □

Исходя из этой леммы можно сделать вывод, что, поскольку самосопряженный оператор можно привести к диагональному виду, то то же самое можно сделать и с симметричной билинейной функцией.

Теорема 5.11.2 *Пусть g — симметричная билинейная функция на евклидовом пространстве V, тогда существует ортонормированный базис, в котором ее матрица диагональна.*

Доказательство. Для самосопряженного оператора f_g существует ортонормированный базис $e_1, \ldots, e_n$, в котором его матрица диагональна, $A_{f_g} = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$, где λ_i — вещественные числа. Это значит, что $f_g e_i = \lambda_i e_i$, $i = 1, \ldots, n$. Рассмотрим матрицу G функции g: $G = (g_{ij})$. По определению $g_{ij} = g(e_i, e_j) = (f_g e_i, e_j) = (\lambda_i e_i, e_j) = \lambda_i \delta_{ij}$, т.е. матрица G диагональна и, более того, равна матрице A_{f_g}. □

Другое доказательство. Возьмем матрицу G билинейной функции в каком-нибудь ортонормированном базисе. При переходе к другому базису матрица этой функции преобразуется по правилу $G' = C^t G C$, где C — матрица перехода. Поскольку C — матрица перехода от одного ортонормированного базиса к другому ортонормированному, то матрица C ортогональна, следовательно, $C^t = C^{-1}$. Следовательно, матрица билинейной функции преобразуется по правилу $G' = C^{-1} G C$. Но по такому же правилу преобразуются линейные операторы, а для них уже известна теорема о приведении к диагональному виду. □

Лемма 5.11.3 *Указанный в теореме диагональный вид единственен с точностью до перестановки диагональных элементов.*

Доказательство. Диагональные элементы λ_i, $i = 1, \ldots, n$, удовлетворяют уравнению $\det(G - \lambda E) = 0$, однако при переходе к другому ортонормированному базису это уравнение имеет тот же вид: $\det(G' - \lambda E) = \det(C^t G C - \lambda C^t C) = \det(C^t (G - \lambda E) C) = \underbrace{(\det C)^2}_{=1} \det(G - \lambda E) =$ $\det(G - \lambda E)$, и его корни, тем самым, не зависят от выбора ортонормированного базиса. □

Определение 5.11.4 Этот диагональный вид называется *каноническим видом* билинейной функции. Собственные векторы оператора f_g называются главными осями функции g, и иногда приведение к каноническому виду называют приведением к главным осям.

Например, в трехмерном пространстве уравнение $g(x, x) = 1$ задает поверхность второго порядка, а ее уравнение в главных осях имеет вид $\lambda_1 (x^1)^2 + \lambda_2 (x^2)^2 + \lambda_3 (x^3)^3 = 1$.

5.12 Пара симметричных билинейных функций

Теорема 5.12.1 *Пусть на векторном пространстве V заданы две билинейные симметричные функции g и h, и пусть g положительно определена, т.е. $g(x, x) > 0$ $\forall x \neq 0$. Тогда существует базис в V, в котором одновременно матрица функции g имеет нормальный вид, а матрица функции h — канонический (т.е. матрица функции g единична, а матрица функции h — диагональна).*

Доказательство. Идея доказательства состоит в том, что, поскольку функция g положительно определена, то на V с ее помощью можно определить скалярное произведение по формуле $(a, b) := g(a, b)$ (все аксиомы скалярного произведения легко проверяются). Следовательно, на V можно ввести структуру евклидового пространства. При этом в любом ортонормированная базисе (относительно введенного только что скалярного произведения) матрица Грама (она же — матрица функции g) будет единичной. По предыдущей теореме существует ортонормированный базис, в котором матрица функции h имеет канонический вид. □

Покажем, как найти канонический вид функции h и канонический базис. Рассмотрим определитель $\det(H - \lambda G)$, где H — матрица функции h, а G — матрица функции g в некотором базисе $e_1, \ldots, e_n$. Введем скалярное произведение с помощью функции g. Пусть $e'_1, \ldots, e'_n$ — ортонормированный базис (по отношению к введенному скалярному произведению). Пусть H' и G' — матрицы этих функций в базисе $e'_1, \ldots, e'_n$. Пусть также C — матрица перехода от базиса $e'_1, \ldots, e'_n$ к $e_1, \ldots, e_n$. Тогда $G = C^t G' C$ и $H = C^t H' C$, причем, т.к. базис $e'_1, \ldots, e'_n$ ортонормирован, то $G' = E$. Получим

$$\det(H - \lambda G) = \det(C^t H' C - \lambda C^t E C) = \underbrace{\det C^t}_{\neq 0} \det(H' - \lambda E) \underbrace{\det C}_{\neq 0},$$

следовательно, многочлены $\det(H - \lambda G)$ и $\det(H' - \lambda E)$ отличаются лишь числовым множителем, значит их корни совпадают. Т.к. диагональные элемента канонического вида функции h — это корни многочлена $\det(H' - \lambda E)$, то они будут также корнями уравнения $\det(H - \lambda G)$. Последнее уравнение называется *обобщенным характеристическим уравнением.* для пары симметричных билинейных функций.

Нахождение канонического базиса. Пусть x — собственный вектор матрицы H', отвечающий собственному значению λ. Пусть X — столбец координат этого вектора в первоначальном базисе $e_1, \ldots, e_n$, а X' — столбец координат этого же вектора в ортонормированном базисе $e'_1, \ldots, e'_n$. Чтобы найти X', нужно решить уравнение $(H' - \lambda E) X' = 0$. Т.к. $X' = CX$, то это уравнение равносильно уравнению $(H' - \lambda E) C X = 0$, домножим его слева на C^t и получим $C^t (H' - \lambda E) C X = 0$, т.е. $(H - \lambda G) X = 0$.

Мы только что получили доказательство следующей леммы:

Лемма 5.12.2 *Корни многочлена* $\det(H-\lambda G)$ *не зависят от выбора базиса и являются диагональными элементами канонического вида матрицы функции* h*, а координаты векторов канонического базиса ищутся как решение системы уравнений* $(H-\lambda G)X=0$*.*

Покажем на примере, что требование положительной определенности хотя бы одной из двух функций существенно. Пусть билинейные симметричные функциии g и h заданы матрицами $G=\begin{pmatrix}0&1\\1&0\end{pmatrix}$ и $H=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$. Ни одна из них не является положительно определенной. Допустим, что существует базис, удовлетворяющий условию теоремы, тогда уравнение $\det(H-\lambda G)=0$ должно иметь вещественные корни (т.к. эти корни суть диагональные элементы матрицы канонического вида функции h). Но это уравнение не имеет вещественных корней, следовательно, такого базиса не существует.

5.13 Принцип минимакса

Мы знаем, что существует базис, в котором матрица симметричной билинейной функции g имеет диагональный вид. Упорядочим диагональные элементы по убыванию, $\lambda_1\geqslant\lambda_2\geqslant\ldots\geqslant\lambda_n$. Обозначим через q квадратичную функцию $q(x)=g(x,x)$.

Теорема 5.13.1 (принцип минимакса) *Для квадратичной функции* q *имеет место равенство*

$$\lambda_k=\min_{V,\dim V=k-1}\Big(\max_{|x|=1,x\perp V}q(x)\Big)$$

(здесь минимум берется по всем подпространствам V *размерности* $k-1$*, а максимум — по всем векторам единичной длины, ортогональным к* V*).*

Доказательство. Пусть $V_0=\langle e_1,\ldots,e_{k-1}\rangle$, где $e_1,\ldots,e_n$ — канонический ортонормированный базис, в котором функция q имеет диагональный вид, $q(x)=\lambda_1(x^1)^2+\ldots+\lambda_n(x^n)^2$, где $x=x^ie_i$. Пусть $x\perp V_0$. Тогда $q(x)=\sum_{i=k}^n\lambda_i(x^i)^2$. Если $x\neq0$, то $\frac{x}{|x|}$ — единичный вектор, следовательно, $q(\frac{x}{|x|})=\frac{q(x)}{|x|^2}$. Найдем $\max\limits_{x\perp V_0}\frac{q(x)}{|x|^2}$. С одной стороны, $\frac{q(x)}{|x|^2}=\frac{\sum_{i=k}^n\lambda_i(x^i)^2}{\sum_{i=k}^n(x^i)^2}\leqslant\lambda_k$. Но, если взять $x=e_k$, то получим, что $\frac{q(x)}{|x|^2}=\lambda_k$. Следовательно, $\max\limits_{x\perp V_0}\frac{q(x)}{|x|^2}=\lambda_k$.

Проверим, что для любого подпространства V размерности $k-1$ имеет место оценка $\max\limits_{x\perp V}\frac{q(x)}{|x|^2}\geqslant\lambda_k$. Т.к. $\dim V=k-1$, то $\dim V^\perp=n-k+1$. Пусть $V'=\langle e_1,\ldots,e_k\rangle$, $\dim V'=k$. Поскольку $\dim V^\perp+\dim V'=n+1$, то пересечение этих подпространств ненулевое, и существует ненулевой вектор $x_0\in V^\perp\cap V'$. Тогда очевидно, что $\max\limits_{x\perp V}\frac{q(x)}{|x|^2}\geqslant\frac{q(x_0)}{|x_0|^2}$. Пусть $x_0=x_0^ie_i$, тогда $x_0^{k+1}=\ldots=x_0^n=0$ (т.к. $x_0\in V'$) и

$$\frac{q(x_0)}{|x_0|^2}=\frac{\sum_{i=1}^k\lambda_i(x_0^i)^2}{\sum_{i=1}^k(x_0^i)^2}\geqslant\frac{\sum_{i=1}^k\lambda_k(x_0^i)^2}{\sum_{i=1}^k(x_0^i)^2}=\lambda_k$$

(т.к. $\lambda_i\geqslant\lambda_k$ при $i\leqslant k$). Таким образом, минимум по всем подпространствам размерности $k-1$ тоже $\geqslant\lambda_k$. Но мы знаем, что этот минимум достигается на подпространстве V_0. □

Пример.

Рассмотрим эллипсоид $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$, $0<a\leqslant b\leqslant c$. На левую часть его уравнения можно смотреть как на квадратичную функцию $q(x)=\lambda_1(x^1)^2+\lambda_2(x^2)^2+\lambda_3(x^3)^2$, где $\lambda_1\geqslant\lambda_2\geqslant\lambda_3$. При этом $\lambda_1=\max\limits_{|x|=1}q(x)$, $\lambda_2=\min\limits_V\Big(\max\limits_{|x|=1,x\perp V}q(x)\Big)$, где минимум берется по всевозможным прямым V, $\dim V=1$. Таким образом, для определения λ_2 надо взять всевозможные плоскости, проходящие через начало координат ($x\perp V$), в каждой плоскости найти $\max\limits_{|x|=1}q(x)$, и затем найти минимум по всем плоскостям. Для определения λ_3 можно заменить $q(x)$ на $-q(x)$, тогда $-\lambda_3=\max\limits_{|x|=1}(-q(x))$, или $\lambda_3=\min\limits_{|x|=1}q(x)$.

Следствие 5.13.2 *Пусть $q(x) = \lambda_1(x^1)^2 + \ldots + \lambda_n(x^n)^2$ — квадратичная функция на евклидовом пространстве W, и пусть задано подпространство $V \subset W$ коразмерности 1 (т.е. $\dim V = \dim W - 1$). Рассмотрим ограничение q' квадратичной функции q на V. Пусть $q'(y) = \mu_1(y^1)^2 + \ldots + \mu_{n-1}(y^{n-1})^2$ — канонический вид квадратичной функции q', причем числа μ_i упорядочены по убыванию, т.е. $\mu_1 \geqslant \ldots \geqslant \mu_{n-1}$. Тогда они находятся между соответствующими числами λ_i, т.е.*

$$\lambda_1 \geqslant \mu_1 \geqslant \lambda_2 \geqslant \mu_2 \geqslant \lambda_3 \geqslant \ldots \geqslant \lambda_{n-1} \geqslant \mu_{n-1} \geqslant \lambda_n.$$

Доказательство. Сравним числа

$$\lambda_{k+1} = \min_{L \subset W, \dim L = k} \Big(\max_{x \perp L, |x|=1} q(x) \Big) = \min_{K \subset W, \dim K = n-k} \Big(\max_{x \in K, |x|=1} q(x) \Big)$$

и

$$\mu_k = \min_{L \subset V, \dim L = k-1} \Big(\max_{x \perp L, |x|=1} q'(x) \Big) = \min_{K \subset V, \dim K = n-k} \Big(\max_{x \in K, |x|=1} q'(x) \Big)$$

(мы здесь заменили L на $K = L^\perp$). Поскольку в этих равенствах в скобках стоят одни и те же выражения, а минимум потом берется по разным множествам: для λ_{k+1} — по всем подпространствам $K \subset W$ размерности $n-k$, а для μ_k — по всем подпространствам $K \subset V$ той же размерности, причем для μ_k минимум берется по меньшему множеству, чем для λ_{k+1}. Следовательно, $\mu_k \geqslant \lambda_{k+1}$.

Теперь сравним λ_k и μ_k. Заменим функцию $q(x)$ на $\widetilde{q}(x) = -q(x)$. Это тоже будет квадратичная функция, диагональные элементы ее матрицы после упорядочения по убыванию будут равны $\widetilde{\lambda}_1 = -\lambda_n \geqslant \ldots \geqslant \widetilde{\lambda}_n = -\lambda_1$, т.е. $\widetilde{\lambda}_i = -\lambda_{n-i+1}$. Аналогично для функции $\widetilde{q}' = -q'$ получим, что $\widetilde{\mu}_1 = -\mu_{n-1} \geqslant \ldots \geqslant \widetilde{\mu}_{n-1} = -\mu_1$, т.е. $\widetilde{\mu}_i = -\mu_{n-i}$. Для функции $\widetilde{q}'$ напишем уже доказанное нами неравенство $\widetilde{\mu}_i \geqslant \widetilde{\lambda}_{i+1}$, из которого, если вернуться к λ и μ, получаем $-\mu_{n-i} \geqslant -\lambda_{n-i}$, или $\mu_k \leqslant \lambda_k$, где $k = n - i$. □

Пример.

Рассмотрим эллипсоид с полуосями a, b, c $(a \geqslant b \geqslant c > 0)$. В сечении его плоскостью, проходящей через центр, мы получим эллипс с некоторыми полуосями α, β $(\alpha \geqslant \beta > 0)$. Из следствия получаем, что $a \geqslant \alpha \geqslant b \geqslant \beta \geqslant c$. Отметим, что если в сечении получилась окружность, т.е. если $\alpha = \beta$, то радиус этой окружности может быть равен только средней полуоси b.

6 Тензоры

6.1 Тензоры. Пространство тензоров

Пусть нам задано векторное пространство V (над полем $\mathbb{K}$), и пусть V' — двойственное пространство. Возьмем произведение p экземпляров пространства V и q экземпляров пространства V' и рассмотрим функцию

$$T : \underbrace{V \times \ldots \times V}_{p \text{ раз}} \times \underbrace{V' \times \ldots \times V'}_{q \text{ раз}} \to \mathbb{K},$$

т.е. T — функция от p векторов и q линейных функций, принимающая значения в $\mathbb{K}$. Такая функция называется полилинейной, если она линейна по каждому аргументу при фиксированных остальных аргументах, т.е. если выполняются равенства:

$$\begin{aligned}
T(v_1, \ldots, v_i' + v_i'', \ldots, v_p, f^1, \ldots, f^q) &= T(v_1, \ldots, v_i', \ldots, v_p, f^1, \ldots, f^q) + \\
&+ T(v_1, \ldots, v_i'', \ldots, v_p, f^1, \ldots, f^q), \\
T(v_1, \ldots, \lambda v_i, \ldots, v_p, f^1, \ldots, f^q) &= \lambda T(v_1, \ldots, v_i, \ldots, v_p, f^1, \ldots, f^q), \\
T(v_1, \ldots, v_p, f^1, \ldots, f'^j + f''^j, \ldots, f^q) &= T(v_1, \ldots, v_p, f^1, \ldots, f'^j, \ldots, f^q) + \\
&+ T(v_1, \ldots, v_p, f^1, \ldots, f''^j, \ldots, f^q), \\
T(v_1, \ldots, v_p, f^1, \ldots, \lambda f^j, \ldots, f^q) &= \lambda T(v_1, \ldots, v_p, f^1, \ldots, f^j, \ldots, f^q).
\end{aligned}$$

Определение 6.1.1 *Тензором* называется полилинейная функция. Тензор типа (p, q) — это полилинейная функция от p векторов из V и от q линейных функций из V'. Тензором типа $(0, 0)$ называют скаляры.

Примеры:
1) скаляры — это тензоры типа $(0, 0)$ по пределению;
2) векторы — это тензоры типа $(0, 1)$, т.к. любой вектор задает линейную функцию на V' переводящую элемент $f \in V'$ в $f(v) \in \mathbb{K}$;
3) линейные функции — это тензоры типа $(1, 0)$, т.к. они задают отображение $f : V \to \mathbb{K}$, переводя вектор $v \in V$ в элемент $f(v) \in \mathbb{K}$;
4) билинейные функции — это тензоры типа $(2, 0)$, т.к. они задают отображение $V \times V \to \mathbb{K}$.
5) линейные операторы — это тензоры типа $(1, 1)$, этот пример мы обсудим позже.

Отметим, что в силу канонического изоморфизма $V \cong V''$ роль векторного пространства V и его двойственного пространства V' в определении тензора симметричны.

Множество тензоров типа (p, q) мы будем обозначать Θ_p^q. На этом множестве можно ввести структуру линейного пространства, для этого нужно определить операции сложения двух тензоров и их умножения на скаляры. Пусть $T, S \in \Theta_p^q$, тогда:

$$\begin{aligned}
(T + S)(v_1, \ldots, v_p, f^1, \ldots, f^q) &:= T(v_1, \ldots, v_p, f^1, \ldots, f^q) + \\
&+ S(v_1, \ldots, v_p, f^1, \ldots, f^q); \\
(\lambda T)(v_1, \ldots, v_p, f^1, \ldots, f^q) &:= \lambda T(v_1, \ldots, v_p, f^1, \ldots, f^q).
\end{aligned}$$

Легко проверить выполнение всех аксиом линейного пространства и убедиться, что множество Θ_p^q действительно будет линейным пространством.

Можно также определить операцию умножения тензоров (разных типов) друг на друга. Пусть $T \in \Theta_p^q$, $S \in \Theta_r^t$, тогда новый тензор $T \otimes S \in \Theta_{p+r}^{q+t}$ опреляется по формуле

$$\begin{aligned}
(T \otimes S)(v_1, \ldots, v_p, v_{p+1}, \ldots, v_{p+r}, f^1, \ldots, f^q, f^{q+1}, \ldots, f^{q+t}) := \\
:= T(v_1, \ldots, v_p, f^1, \ldots, f^q) \cdot S(v_{p+1}, \ldots, v_{p+r}, f^{q+1}, \ldots, f^{q+t}).
\end{aligned}$$

Так определенное умножение тензоров дистрибутивно (т.е. $(T + \lambda S) \otimes R = T \otimes R + \lambda S \otimes R$), ассоциативно (т.е. $(T \otimes S) \otimes R = T \otimes (S \otimes R)$), но не коммутативно.

Примеры:

1) произведение вектора $v \in V$ и линейной функции $f \in V'$ — тензор $v \otimes f \in \Theta_1^1$. Посмотрим, как этот тензор действует на своих аргументах. Возьмем произвольный вектор $a \in V$ и функцию $h \in V'$, получим, что $(v \otimes f)(a, h) = h(v) \cdot f(a)$.

2) возьмем две линейные функции $g, h \in V'$, тогда $g \otimes h$ будет тензором типа $(2,0)$ (т.е. билинейной функцией), $(g \otimes h)(v_1, v_2) = g(v_1) \cdot h(v_2)$.

6.2 Координатное определение тензоров

Перейдем теперь к координатному описанию тензоров. Зафиксируем базис $e_1, \ldots, e_n$ в пространстве V, ему будет соответствовать двойственный базис $\varepsilon^1, \ldots, \varepsilon^n$ в двойственном пространстве V'. Т.к. тензор — это полилинейная функция, то ее значение определяется значениями на базисных векторах, т.е.

$$T(v_1, \ldots, v_p, f^1, \ldots, f^q) = T(v_1^{i_1} e_{i_1}, \ldots, v_p^{i_p} e_{i_p}, f_{j_1}^1 \varepsilon^{j_1}, \ldots, f_{j_q}^q \varepsilon^{j_q}) =$$
$$= v_1^{i_1} \cdot \ldots \cdot v_p^{i_p} \cdot f_{j_1}^1 \cdot \ldots \cdot f_{j_q}^q \cdot T(e_{i_1}, \ldots, e_{i_p}, \varepsilon^{j_1}, \ldots, \varepsilon^{j_q}),$$

где $v_k^{i_k}$ — координаты вектора v_k, а $f_{j_j}^l$ — координаты линейной функции f^l.

Обозначим $T(e_{i_1}, \ldots, e_{i_p}, \varepsilon^{j_1}, \ldots, \varepsilon^{j_q}) = T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q} \in \mathbb{K}$. Поэтому, зафиксировав базис, мы можем поставить тензору T в соответствие набор чисел $T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}$ — его значения на базисных векторах. Естественно, этот набор чисел будет зависеть от выбора базиса. Посмотрим, как изменяются эти числа при переходе от одного базиса к другому. Пусть мы перешли от базиса e_i к базису $\widetilde{e}_i$, и $C = (c_i^k)$ — матрица перехода, т.е. $\widetilde{e}_i = c_i^k e_k$. Тогда двойственный базис ε^i тоже сменится на $\widetilde{\varepsilon}^i$, причем матрица перехода $D = (d_k^i) = C^{-1}$, $d_k^j c_i^k = \delta_i^j$, т.е. $\widetilde{\varepsilon}^i = d_k^i \varepsilon^k$. Тогда

$$\widetilde{T}_{k_1,\ldots,k_p}^{l_1,\ldots,l_q} = T(\widetilde{e}_{k_1}, \ldots, \widetilde{e}_{k_p}, \widetilde{\varepsilon}^{l_1}, \ldots, \widetilde{\varepsilon}^{l_q}) =$$
$$= T(c_{k_1}^{i_1} e_{i_1}, \ldots, c_{k_p}^{i_p} e_{i_p}, d_{j_1}^{l_1} \varepsilon^{j_1}, \ldots, d_{j_q}^{l_q} \varepsilon^{j_q}) =$$
$$= T(e_{i_1}, \ldots, e_{i_p}, \varepsilon^{j_1}, \ldots, \varepsilon^{j_q}) \cdot c_{k_1}^{i_1} \cdot \ldots \cdot c_{k_p}^{i_p} \cdot d_{j_1}^{l_1} \cdot \ldots \cdot d_{j_q}^{l_q} =$$
$$= T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q} \cdot c_{k_1}^{i_1} \cdot \ldots \cdot c_{k_p}^{i_p} \cdot d_{j_1}^{l_1} \cdot \ldots \cdot d_{j_q}^{l_q}.$$

Примеры:

1) если x — вектор, то $\widetilde{x}^k = d_i^k x^i$. Такой закон изменения координат называется *вєкторным* законом.

2) если f — линейная функция, то $\widetilde{f}_k = c_k^i f_i$. Такой закон изменения координат называется *ковєкторным* законом. А величины, изменяющиеся по ковекторному закону называются *ковєкторами.* Таким образом, линейные функции (элементы пространства V') — это ковекторы.

Мы получили *тензорный закон изменения координат*:

$$\widetilde{T}_{k_1,\ldots,k_p}^{l_1,\ldots,l_q} = T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q} \cdot c_{k_1}^{i_1} \cdot \ldots \cdot c_{k_p}^{i_p} \cdot d_{j_1}^{l_1} \cdot \ldots \cdot d_{j_q}^{l_q}.$$

Теперь мы можем дать другое (координатное) определение тензору: тензор - это набор чисел $T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}$, который при замене координат преобразуется по тензорному закону.

Это и предыдущее определения тензора эквивалентны, так как по такому набору чисел, пользуясь линейностью, можно восстановить полилинейную функцию T, для которой $T(e_{i_1}, \ldots, e_{i_p}, \varepsilon^{j_1}, \ldots, \varepsilon^{j_q}) = T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}$.

Рассмотрим еще один пример — пример тензоров типа $(0,2)$. Пусть (a_{ij}) — матрица некоторой билинейной функции, или, что то же самое, тензор типа $(2,0)$. Предположим, что матрица (a_{ij}) обратима и обозначим элементы обратной матрицы через a^{ij}, т.е. $a^{ij} a_{jk} = \delta_k^i$.

Лемма 6.2.1 *a^{ij} является тензором типа $(0,2)$, т.е. этот набор чисел при переходе к другому базису изменяется по формуле $\widetilde{a}^{kl} = a^{ij} d_i^k d_j^l$.*

Доказательство. В новом базисе выполнено равенство $\widetilde{a}^{kl}\widetilde{a}_{lm} = \delta^k_m$. Подставим сюда $\widetilde{a}_{lm} = a_{ji}c^j_l c^i_m$, где $C = (c^i_m)$ — матрица перехода, тогда $\widetilde{a}^{kl}a_{ji}c^j_l c^i_m = \delta^k_m$. Умножим (и просуммируем по повторяющимся индексам) обе части этого равенства на элементы обратной матрицы к матрице перехода (напомним, что $c^k_i d^j_k = \delta^j_i$): $\widetilde{a}^{kl}a_{ji}c^j_l c^i_m d^m_p = \delta^k_m d^m_p$, т.е. $\widetilde{a}^{kl}a_{ji}c^j_l\delta^i_p = d^k_p$, или $\widetilde{a}^{kl}a_{jp}c^j_l = d^k_p$. Воспользуемся теперь обратимостью матрицы (a_{ij}), т.е. тем, что $a_{jp}a^{pr} = \delta^r_j$. Умножив обе части предпоследнего равенства на a^{pr} (и просуммировав), получим $\widetilde{a}^{kl}a_{jp}a^{pr}c^j_l = a^{pr}d^k_p$, или $\widetilde{a}^{kl}c^r_l = a^{pr}d^k_p$. Еще раз умножив обе части равенства на d^n_r просуммировав по r и воспользовавшись тем, что $c^r_l d^n_r = \delta^n_l$ и $\widetilde{a}^{kl}\delta^n_l = \widetilde{a}^{kn}$, получим $\widetilde{a}^{kn} = a^{pr}d^k_p d^n_r$. □

6.3 Базис в пространстве тензоров

Построим базис в пространстве тензоров Θ^q_p. Векторы и ковекторы — это тензоры типа $(0,1)$ и $(1,0)$ соответственно. Рассмотрим произведение $\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}$, оно состоит из тензоров типа $(0,1)$ и тензоров типа $(1,0)$, следовательно само является тензором типа (p,q). Всего таких различных произведений получится n^{p+q}, т.к. из n ковекторов надо выбрать p и из n векторов надо выбрать q. Докажем, что эти элементы (произведения такого вида) образуют базис в Θ^q_p. Прежде, чем доказывать это, вычислим значение такого тензора (произведения) на наборах аргументов из базисных векторов:

$$\begin{aligned}\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}(e_{k_1},\ldots,e_{k_p},\varepsilon^{l_1},\ldots,\varepsilon^{l_q}) =\\ = \varepsilon^{i_1}(e_{k_1})\cdot\ldots\cdot\varepsilon^{i_p}(e_{k_p})\cdot\varepsilon^{l_1}(e_{j_1})\cdot\ldots\cdot\varepsilon^{l_1}(e_{j_q}) =\\ = \delta^{i_1}_{k_1}\cdot\ldots\cdot\delta^{i_p}_{k_p}\cdot\delta^{l_1}_{j_1}\cdot\ldots\cdot\delta^{l_q}_{j_q} =\\ = \begin{cases}1, & \text{если } i_1=k_1,\ldots,i_p=k_p, j_1=l_1,\ldots,j_q=l_q,\\ 0, & \text{в остальных случаях.}\end{cases}\end{aligned}$$

Лемма 6.3.1 *Множество произведений вида $\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}$ является базисом в Θ^q_p.*

Доказательство. Сначала докажем линейную независимость этих произведений. Пусть существуют такие числа $\lambda^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}$, что линейная комбинация $\lambda^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}$. Применим этот тензор, как полилинейную функцию, к аргументам $e_{k_1},\ldots,e_{k_p},\varepsilon^{l_1},\ldots,\varepsilon^{l_q}$ и получим

$$\lambda^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}\underbrace{\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}(e_{k_1},\ldots,e_{k_p},\varepsilon^{l_1},\ldots,\varepsilon^{l_q})}=0 \tag{7}$$

Подчеркнутое выражение равно 1, если $i_1=k_1,\ldots,i_p=k_p,j_1=l_1,\ldots,j_q=l_q$ и 0 в остальных случаях, следовательно, равенство (7) может быть записано в виде

$$\lambda^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}(e_{k_1},\ldots,e_{k_p},\varepsilon^{l_1},\ldots,\varepsilon^{l_q}) = \lambda^{l_1,\ldots,l_q}_{k_1,\ldots,k_p} = 0.$$

Но, т.к. это равенство имеет место для любого набора индексов $k_1,\ldots,k_p,l_1,\ldots,l_q$, то все $\lambda^{l_1,\ldots,l_q}_{k_1,\ldots,k_p}$ равны нулю. Линейная независимость доказана.

Удостоверимся теперь, что любой тензор можно представить в виде линейной комбинации этих базисных тензоров. Для этого достаточно доказать равенство

$$T = T^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}. \tag{8}$$

А из-за полилинейности это равенство достаточно проверять на базисных аргументах вида $e_{k_1},\ldots,e_{k_p},\varepsilon^{l_1},\ldots,\varepsilon^{l_q}$. Левая часть равенства по определению равна $T^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}$, а правая часть, как мы уже видели раньше, также равна $T^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}(e_{k_1},\ldots,e_{k_p},\varepsilon^{l_1},\ldots,\varepsilon^{l_q}) = T^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}$. Итак, равенство проверено, т.е. для произвольного тензора T мы нашли его разложение в линейную комбинацию (8), причем числа $T^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}$ являются координатами этого тензора в указанном базисе. □

6.4 Линейные операторы как тензоры

Перейдем теперь к обещанному отождествлению линейных операторов и тензоров типа $(1,1)$.

Утверждение 6.4.1 *Имеет место канонический изоморфизм* $\Theta^1_1 \cong L(V)$.

Доказательство. Пусть дан линейный оператор $f : V \to V$, нам надо по нему построить билинейную функцию $T_f = T_f(v, \varphi)$ с одним векторным ($v \in V$) и одним ковекторным ($\varphi \in V'$) аргументами. Определим его равенством $T_f(v, \varphi) := \varphi(f(v)) \in \mathbb{K}$. T_f действительно является полилинейным отображением от двух аргументов. Соответствие $f \mapsto T_f$ задает линейное отображение $L(V) \to \Theta^1_1$, конструкция которого не зависит от выбора базисов. Доказать, что отображение $L(V) \to \Theta^1_1$ будет изоморфизмом, можно двумя способами.

Первый — это просто записать все в координатах и сравнить.

Второй способ. Размерности пространств $L(V)$ и Θ^1_1 совпадают (они равны $(\dim V)^2$), поэтому надо проверить лишь то, что у этого отображения нулевое ядро, т.е. что если $T_f = 0$, то и $f = 0$. Пусть $T_f(v, \varphi) = 0$ для любых $v \in V$, $\varphi \in V'$, т.е. для любого φ $\varphi(f(v)) = 0$, следовательно, $f(v) = 0$. Но, т.к. это верно для всех векторов $v \in V$, то имеем $f = 0$. □

Теперь мы можем отождествлять линейные операторы и тензоры типа $(1,1)$.

Например, тождественный оператор задается в любом базисе матрицей (δ^j_i). Проверим в явном виде, что символ Кронекера δ^j_i является тензором. Запишем тензорный закон изменения координат $\widetilde{\delta}^k_l = \delta^j_i c^i_l d^k_j = c^i_l d^k_i$, что, в свою очередь, равно $c^i_l d^k_i = \delta^k_l$, т.к. $D = C^{-1}$ и $CD = E$. Значит, символ Кронекера действительно является тензором типа $(1,1)$.

6.5 Свертка тензоров

Пусть $T \in \Theta^q_p$ — тензор с хотя бы одним нижним и одним верхним индексами, т.е. $p, q > 0$. Пусть $e_1, \ldots, e_n$ — базис в V. Зафиксируем один векторный и один ковекторный аргумент (пусть это будут первые по порядку аргументы), на их место поставим базисные элементы e_i и ε^i и определим полилинейную функцию sT от $p-1$ векторых и $q-1$ ковекторных аргументов по формуле

$$(sT)(v_2, \ldots, v_p, f^2, \ldots, f^q) := T(e_i, v_2, \ldots, v_p, \varepsilon^i, f^2, \ldots, f^q),$$

в которой подразумевается суммирование по индексу i. Проверим, что определение тензора sT не зависит от выбора базиса, т.е. что его координаты будут преобразовываться по тензорному закону. Имеем $(sT)^{i, j_2, \ldots, j_q}_{i, i_2, \ldots, i_p}$ (здесь опять производится суммирование по индексу i). При переходе к другому базису имеем

$$\widetilde{T}^{l_1, \ldots, l_q}_{k_1, \ldots, k_p} = T^{j_1, \ldots, j_q}_{i_1, \ldots, i_p} \cdot c^{i_1}_{k_1} \cdot \ldots \cdot c^{i_p}_{k_p} \cdot d^{l_1}_{j_1} \cdot \ldots \cdot d^{l_q}_{j_q};$$

$$(\widetilde{sT})^{l_2, \ldots l_q}_{k_2, \ldots, k_p} = \widetilde{T}^{k, l_2, \ldots, l_q}_{k, k_2, \ldots, k_p} = T^{j_1, j_2, \ldots, j_q}_{i_1, i_2, \ldots, i_p} \cdot \underbrace{c^{i_1}_{k}} \cdot c^{i_2}_{k_2} \cdot \ldots \cdot c^{i_p}_{k_p} \cdot \underbrace{d^{k}_{j_1}} \cdot d^{l_2}_{j_2} \cdot \ldots \cdot d^{l_q}_{j_q}.$$

Произведение подчеркнутых элементов равно $c^{i_1}_k d^k_{j_1} = \delta^{i_1}_{j_1}$, т.к. $CD = E$, поэтому ненулевыу слагаемые отвечают значениям $i_1 = j_1$. Обозначим $i_1 = j_1 = i$, тогда

$$\widetilde{T}^{l_1, \ldots, l_q}_{k_1, \ldots, k_p} = T^{i, j_2, \ldots, j_q}_{i, i_2, \ldots, i_p} \cdot c^{i_2}_{k_2} \cdot \ldots \cdot c^{i_p}_{k_p} \cdot d^{l_2}_{j_2} \cdot \ldots \cdot d^{l_q}_{j_q} = (sT)^{j_2, \ldots, j_q}_{i_2, \ldots, i_p} \cdot c^{i_2}_{k_2} \cdot \ldots \cdot c^{i_p}_{k_p} \cdot d^{l_2}_{j_2} \cdot \ldots \cdot d^{l_q}_{j_q}.$$

Т.е. координаты $(sT)^{j_2, \ldots, j_q}_{i_2, \ldots, i_p}$ действительно преобразуются по тензорному закону, значит, действительно, sT — тензор типа $(p-1, q-1)$. Этот тензор называется *сверткой* тензора T. Такую операцию свертки можно проводить несколько раз до исчерпания верхних или нижних индексов. Последняя возможная свертка (после которой не остается либо нижних, либо верхних индексов) называется *полной* сверткой.

Примеры:

1) Возьмем линейный оператор — тензор типа $(1,1)$, результатом свертки будет скаляр. Пусть нам дан оператор $f : V \to V$, который задается матрицей (f^j_i), сверткой этого тензора будет число f^i_i (сумма диагональных элементов), т.е. в данном случае свертка — это след, $sf = \operatorname{tr} f$.

2) Рассмотрим полную свертку тензора типа $(2,2)$. Возьмем билинейную функцию g, т.е. тензор типа $(2,0)$, и два вектора a, b — тензоры типа $(0,1)$. Тогда $g \otimes a \otimes b$ будет тензором типа $(2,2)$. Координаты этого тензора $(g \otimes a \otimes b)_{ij}^{kl} = g_{ij}a^k b^l$, где g_{ij} — матрица билинейной функции g. Тогда при полной свертке этого тензора получим $s(g \otimes a \otimes b) = g_{ij}a^i b^j = g(a,b)$.

6.6 Поднятие и опускание индексов

Пусть нам дано евклидово пространство V, т.е. векторное пространство над $\mathbb{R}$ со скалярным произведением $(\cdot,\cdot)$. Тогда существует канонический изоморфизм $V \cong V'$, поэтому у любого тензора можно заменить векторный аргумент на ковекторный и наоборот. Например, возьмем тензор типа $(0,1)$, т.е. вектор $T : V' \to \mathbb{R}$. Отождествим его с ковектором $(T,\cdot) : V \to \mathbb{R}$, тогда пполучим канонический изоморфизм пространств тензоров Θ_0^1 и Θ_1^0. В координатах это выглядит следующим образом. Пусть $G = (g_{ij})$ — матрица Грама скалярного произведения $(\cdot,\cdot)$. Каждому тензору T с координатами T^i поставим в соответствие $T_j = g_{ij}T^i$ (это произведение двух тензоров, g и T и свертка по индексу i). Таким образом мы у тензора T опустили индекс.

Обобщая эту операцию, дадим определение операции опускания индекса.

Опускание индекса — это отображение $\Theta_p^q \to \Theta_{p+1}^{q-1}$, которое тензору $T_{i_1,\dots,i_p}^{j_1,\dots,j_q}$ ставит в соответствие тензор $g_{ij} \cdot T_{i_1,\dots,i_p}^{j,j_2,\dots,j_q}$. Здесь мы опустили первый индекс j_1. Аналогично можно опустить любой другой верхний индекс.

Поднятие индекса. Аналогичным образом можно и поднимать индексы. Для этого используется матрица $G^{-1} = (g^{ij})$ (вспомним, что матрица Грама обратима и что g^{ij} есть тензор типа $(0,2)$). Тензор $T_{i_1,\dots,i_p}^{j_1,\dots,j_q}$ операция поднятия индекса переводит в тензор $g^{ij} \cdot T_{i,i_2,\dots,i_p}^{j_1,\dots,j_q} \in \Theta_{p-1}^{q+1}$.

6.7 Оператор альтернирования. Внешние формы

Рассмотрим линейное пространство Θ_p^0 тензоров с одними нижними индексами, т.е. полилинейные функции от p векторов. Также рассмотрим группу перестановок S_p. Если взять какую-нибудь перестановку $\sigma \in S_p$, то можно определить линейный оператор $f_\sigma : \Theta_p^0 \to \Theta_p^0$ следующим образом. Пусть $T \in \Theta_p^0$, т.е. $T = T(v_1,\dots,v_p)$; определим $f_\sigma(T) = \sigma T$, где $(\sigma T)(v_1,\dots,v_p) := T(v_{\sigma(1)},\dots,v_{\sigma(p)})$. Эта операция перестановки аргументов сумму тензоров переводит в сумму, а умножение тензора на скаляр — в умножение на скаляр, следовательно, f_σ — линейный оператор. кроме того, $f_{\sigma_1}f_{\sigma_2} = f_{\sigma_1\sigma_2}$. Координаты тензоров T и σT связаны между собой равенством $(\sigma T)_{i_1,\dots,i_p} = (\sigma T)(e_{i_1},\dots,e_{i_p}) = T(e_{i_{\sigma(1)}},\dots,e_{i_{\sigma(p)}}) = T_{i_{\sigma(1)}},\dots,i_{\sigma(p)}$.

Поскольку нам вскоре понадобится делить на целые числа, с этого момента, говоря о тензорах, будем считать, что характеристика поля $\mathbb{K}$ нулевая. При желании для простоты можно считать, что $\mathbb{K} = \mathbb{R}$.

Построим оператор *альтернирования* (приводящий к аналогу свойства кососимметричности)

$$\operatorname{Alt} : \Theta_p^0 \to \Theta_p^0; \quad \operatorname{Alt} T := \frac{1}{p!}\sum_{\sigma \in S_p} (-1)^\sigma \sigma T.$$

Этот оператор будет линейным, т.к. является суммой линейных операторов. Назовем тензор $T \in \Theta_p^0$ *кососимметрическим* (или *внешней формой*), если $\sigma T = (-1)^\sigma T$ для любой перестановки $\sigma \in S_p$. В пространстве всех тензоров с нижними индексами определим подпространство $\Lambda_p \subset \Theta_p^0$ всех кососимметрических тензоров (проверка того, что множество кососимметрических тензоров в действительности есть подпространство, очевидна). Если $p = 2$, то условие кососимметричности эквивалентно условию $T_{ij} = -T_{ji}$.

Лемма 6.7.1 *Оператор* Alt *является оператором проектирования на подпространство внешних форм* Λ_p.

Доказательство. Нам потребуются следующие равенства:

Утверждение 6.7.2 $\sigma(\operatorname{Alt} T) = \operatorname{Alt}(\sigma T) = (-1)^\sigma \operatorname{Alt} T$.

Доказательство. Применим перестановку σ к тензору $\operatorname{Alt} T$:

$$\sigma(\operatorname{Alt} T) = \sigma\Big(\frac{1}{p!}\sum_{\rho\in S_p}(-1)^\rho \rho T\Big) = \frac{1}{p!}\sum_{\rho\in S_p}(-1)^\rho((\sigma\rho)T),$$

т.к. σ — это линейный оператор. Когда ρ пробегает всю группу S_p, перестановка $\tau = \sigma\rho$ тоже пробегает всю группу S_p, поэтому полученное выражение можно записать так: $\sigma(\operatorname{Alt} T) = \frac{1}{p!}\sum_{\tau\in S_p}(-1)^\rho \tau T$. А, поскольку $(-1)^\tau = (-1)^\rho(-1)^\sigma$, то

$$\sigma(\operatorname{Alt} T) = \frac{1}{p!}\sum_{\tau\in S_p}(-1)^\sigma(-1)^\tau\tau T = (-1)^\sigma\Big(\frac{1}{p!}\sum_{\tau\in S_p}(-1)^\tau\tau T\Big) = (-1)^\sigma \operatorname{Alt} T,$$

т.е. мы доказали, что $\sigma(\operatorname{Alt} T) = (-1)^\sigma \operatorname{Alt} T$. Теперь докажем, что $\operatorname{Alt}(\sigma T) = (-1)^\sigma \operatorname{Alt} T$.

По определению $\operatorname{Alt}(\sigma T) = \frac{1}{p!}\sum_{\rho\in S_p}(-1)^\rho((\rho\sigma)T)$. Обозначим $\tau = \rho\sigma$ и получим

$$\begin{aligned}\operatorname{Alt}(\sigma T) &= \frac{1}{p!}\sum_{\rho\in S_p}(-1)^\rho((\rho\sigma)T) = \frac{1}{p!}\sum_{\tau\in S_p}(-1)^\rho(\tau T) = \frac{1}{p!}\sum_{\tau\in S_p}(-1)^\tau(-1)^\sigma(\tau T) = \\ &= (-1)^\sigma\Big(\frac{1}{p!}\sum_{\tau\in S_p}(-1)^\tau(\tau T)\Big) = (-1)^\sigma \operatorname{Alt} T.\end{aligned}$$

□

Перейдем теперь собственно к доказательству леммы.

1. Проверим, что $\operatorname{Im}\operatorname{Alt} \subset \Lambda_p$. Действительно, поскольку $\sigma(\operatorname{Alt} T) = (-1)^\sigma \operatorname{Alt} T$, то по определению $\operatorname{Alt} T \in \Lambda_p$ для любого $T \in \Theta^0_p$, поэтому $\operatorname{Im}\operatorname{Alt} \subset \Lambda_p$.

2. Докажем, что если $T \in \Lambda_p$, то $\operatorname{Alt} T = T$. Действительно, поскольку $T \in \Lambda_p$, то $\sigma T = (-1)^\sigma T$ и

$$\operatorname{Alt} T = \frac{1}{p!}\sum_{\sigma\in S_p}(-1)^\sigma\sigma T = \frac{1}{p!}\sum_{\sigma\in S_p}(-1)^\sigma(-1)^\sigma T = \frac{1}{p!}\sum_{\sigma\in S_p} T = \frac{1}{p!}p!T = T.$$

3. Проверим, что $\operatorname{Alt}^2 = \operatorname{Alt}$, т.е. что $\operatorname{Alt}(\operatorname{Alt} T) = \operatorname{Alt} T$ для любого $T \in \Theta^0_p$. Действительно, в п.1 мы доказали, что $S = \operatorname{Alt} T \in \Lambda_p$, в п.2 — что $\operatorname{Alt} S = S$, подставив $\operatorname{Alt} T$ вместо S, получим $\operatorname{Alt}(\operatorname{Alt} T) = \operatorname{Alt} T$. □

6.8 Внешнее умножение, его свойства

Определим аналог тензорного умножения для внешних форм — *внешнее тензорное умножение* (обозначается $\wedge$): для $T \in \Lambda_p$, $S \in \Lambda_q$ положим $T \wedge S := \operatorname{Alt}(T \otimes S)$.

Лемма 6.8.1 *Введенное нами внешнее тензорное умножение обладает следующими свойствами: для любых внешних форм $T \in \Lambda_p$, $S \in \Lambda_q$, $R \in \Lambda_r$*

1) $(T + \lambda S) \wedge R = T \wedge R + \lambda S \wedge R$ (дистрибутивность);

2) $S \wedge T = (-1)^{pq} T \wedge S$ (антикоммутативность) ;

3) $(T \wedge S) \wedge R = T \wedge (S \wedge R)$ (ассоциативность).

Доказательство. 1) Дистрибутивность следует из дистрибутивности операции $\otimes$ и линейности оператора Alt.

2) По определению

$$S \wedge T = \operatorname{Alt}(S \otimes T) = \frac{1}{(p+q)!}\sum_{\sigma\in S_{p+q}}\sigma(S \otimes T);$$

$$T \wedge S = \operatorname{Alt}(T \otimes S) = \frac{1}{(p+q)!}\sum_{\sigma\in S_{p+q}}\sigma(T \otimes S).$$

Рассмотрим координаты тензоров $\sigma(S\otimes T)$ и $\sigma(T\otimes S)$. Имеем

$$\begin{aligned}\sigma(S\otimes T)_{i_1,\dots,i_{p+q}} &= (S\otimes T)_{i_{\sigma(1)},\dots,i_{\sigma(p+q)}} = \\ &= S_{i_{\sigma(1)},\dots,i_{\sigma(q)}}\cdot T_{i_{\sigma(q+1)},\dots,i_{\sigma(p+q)}}; \qquad (9)\\ \sigma(T\otimes S)_{i_1,\dots,i_{p+q}} &= (T\otimes S)_{i_{\sigma(1)},\dots,i_{\sigma(p+q)}} = \\ &= T_{i_{\sigma(1)},\dots,i_{\sigma(p)}}\cdot S_{i_{\sigma(p+1)},\dots,i_{\sigma(p+q)}}. \qquad (10)\end{aligned}$$

Посмотрим, чем отличаются индексы у S и T в выражениях (9) и (10). Индексы в (9) — это $\sigma(1),\dots,\sigma(q),\sigma(q+1),\dots,\sigma(p+q)$, а индексы в (10) — это $\sigma(p+1),\dots,\sigma(p+q),\sigma(1),\sigma(q)$. Пусть τ — перестановка

$$\begin{pmatrix} p+1 & \dots & p+q & 1 & \dots & p\\ 1 & \dots & q & q+1 & \dots & p+q\end{pmatrix}.$$

Тогда, как легко видеть, $\sigma(S\otimes T)=\sigma\tau(T\otimes S)$. Поэтому

$$\begin{aligned}S\wedge T &= \frac{1}{(p+q)!}\sum_{\sigma\in S_{p+q}}(-1)^\sigma\sigma(S\otimes T) = \\ &= \frac{1}{(p+q)!}\sum_{\sigma\in S_{p+q}}(-1)^\tau(-1)^{\sigma\tau}(\sigma\tau)(S\otimes T) = \\ &= (-1)^\tau\frac{1}{(p+q)!}\sum_{\rho\in S_{p+q}}(-1)^\rho\rho(S\otimes T) = (-1)^\tau T\wedge S,\end{aligned}$$

и нам осталось определить $(-1)^\tau$. Для вычисления знака перестановки надо подсчитать количество элементарных перестановок, ее составляющих. Лекго видеть, что это число равно произведению pq, т.е. $(-1)^\tau=(-1)^{pq}$, что и требовалось показать.

3) Введем дополнительно обозначение $T_1\wedge T_2\wedge\ldots\wedge T_k := \mathrm{Alt}(T_1\otimes T_2\otimes\ldots\otimes T_k)$. Для доказательство ассоциативности нам также понадобится следующее равенство.

Утверждение 6.8.2 $\mathrm{Alt}((\mathrm{Alt}\,Q)\otimes R)=\mathrm{Alt}(Q\otimes R)=\mathrm{Alt}(Q\otimes(\mathrm{Alt}\,R))$ *для любых тензоров* $Q\in\Theta^0_p$, $R\in\Theta^0_q$.

Доказательство. Ограничимся доказательством первого из равенств (второе доказывается аналогично). Поскольку операция $\otimes$ обладает свойством дистрибутивности, а оператор Alt линеен, имеем

$$\begin{aligned}\mathrm{Alt}((\mathrm{Alt}\,Q)\otimes R) &= \mathrm{Alt}\Big(\Big(\frac{1}{p!}\sum_{\sigma\in S_p}(-1)^\sigma\sigma Q\Big)\otimes R\Big) = \\ &= \frac{1}{p!}\sum_{\sigma\in S_p}(-1)^\sigma\,\mathrm{Alt}(\sigma Q\otimes R).\end{aligned}$$

Каждой перестановке $\sigma\in S_p$ поставим в соответствие такую перестановку $\widetilde{\sigma}\in S_{p+q}$, которая на первых p индексах действует как σ, а остальные оставляет на месте, т.е.

$$\widetilde{\sigma}=\begin{pmatrix}1 & \dots & p & p+1 & \dots & p+q\\ \sigma(1) & \dots & \sigma(p) & p+1 & \dots & p+q\end{pmatrix}.$$

При этом, очевидно, $(-1)^{\widetilde{\sigma}}=(-1)^\sigma$.

Тогда $\sigma Q \otimes R = \widetilde{\sigma}(Q \otimes R)$, и

$$\begin{aligned}
\operatorname{Alt}((\operatorname{Alt} Q) \otimes R) &= \frac{1}{p!} \sum_{\sigma \in S_p} (-1)^{\sigma} \operatorname{Alt}(\widetilde{\sigma}(Q \otimes R)) = \\
&= \frac{1}{p!} \sum_{\sigma \in S_p} (-1)^{\sigma} (-1)^{\widetilde{\sigma}} \operatorname{Alt}(Q \otimes R) = \\
&= \frac{1}{p!} \sum_{\sigma \in S_p} \operatorname{Alt}(Q \otimes R) = \frac{1}{p!} p! \operatorname{Alt}(Q \otimes R) = \\
&= \operatorname{Alt}(Q \otimes R).
\end{aligned}$$

□

Докажем теперь ассоциативность внешнего умножения. Обозначим $Q = T \otimes S$, тогда $\operatorname{Alt} Q = \operatorname{Alt}(T \otimes S)$ и

$$\begin{aligned}
(T \wedge S) \wedge R &= \operatorname{Alt}((T \wedge S) \otimes R) = \operatorname{Alt}(\operatorname{Alt}(T \otimes S) \otimes R) = \\
&= \operatorname{Alt}((\operatorname{Alt} Q) \otimes R) = \operatorname{Alt}(Q \otimes R) = \operatorname{Alt}(T \otimes S \otimes R) = \\
&= T \wedge S \wedge R.
\end{aligned}$$

Аналогично получим, что $T \wedge (S \wedge R) = T \wedge S \wedge R$, т.е. $(T \wedge S) \wedge R = T \wedge (S \wedge R)$. □

6.9 Базис в пространстве внешних форм

Построим базис в пространстве кососимметрических тензоров Λ_p. Как нам известно, базисом в пространстве Θ_p^0 являются всевозможные произведения вида $\varepsilon^{i_1} \otimes \ldots \otimes \varepsilon^{i_p}$. Однако при проектировании на подпространство Λ_p эти произведения переходят в $\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p} = \operatorname{Alt}(\varepsilon^{i_1} \otimes \ldots \otimes \varepsilon^{i_p})$ и становятся линейно зависимыми. Действительно, если $i_k = i_l$, то $\varepsilon^{i_k} \wedge \varepsilon^{i_l} = 0$, поэтому из базиса нужно выкинуть все произведения, в которых встречаются повторяющиеся индексы. Далее, т.к. $\varepsilon^{i_k} \wedge \varepsilon^{i_l} = -\varepsilon^{i_l} \wedge \varepsilon^{i_k}$, то нужно также выкинуть все произведения с неупорядоченными индексами и оставить только произведения с индексами упорядоченными, например, по возрастанию, т.е. такие, что $i_1 < \ldots < i_p$. Всего таких упорядоченных произведений будет C_n^p, где $n = \dim V$. Очевидно, что любой тензор является линейной комбинацией тензоров $\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p}$, $i_1 < \ldots < i_p$, т.к. из базиса мы выкинули только заведомо линейно зависимые вектора. Поэтому для доказательства того, что эта система векторов действительно является базисом Λ_p достаточно доказать ее линейную независимость.

Пусть линейная комбинация этих тензоров равна нулю,

$$\sum_{i_1 < \ldots < i_p} \lambda_{i_1, \ldots, i_p} \varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p} = 0.$$

Перепишем эту сумму в виде

$$\sum_{i_1 < \ldots < i_p} \lambda_{i_1, \ldots, i_p} \Big(\frac{1}{p!} \sum_{\sigma \in S_p} (-1)^{\sigma} \underbrace{\varepsilon^{i_{\sigma(1)}} \otimes \ldots \otimes \varepsilon^{i_{\sigma(p)}}} \Big) = 0. \tag{11}$$

Подчеркнутые элементы — это элементы базиса в пространстве Θ_p^0 и они линейно независимы. Отсюда будет следовать равенство нулю всех коэффициентов, если только все эти подчеркнутые элементы будут различны. Докажем, что все базисные элементы, встречающиеся в этой сумме, различны. Распишем эту сумму подробнее

$$\ldots + \lambda_{i_1, \ldots, i_p} \sum_{\sigma \in S_p} \varepsilon^{i_{\sigma(1)}} \otimes \ldots \otimes \varepsilon^{i_{\sigma(p)}} + \ldots + \lambda_{j_1, \ldots, j_p} \sum_{\rho \in S_p} \varepsilon^{j_{\rho(1)}} \otimes \ldots \otimes \varepsilon^{j_{\rho(p)}} + \ldots .$$

Допустим, что существуют такие $\sigma, \rho \in S_p$ и индексы $i_1, \ldots, i_p, j_1, \ldots, j_p$, что $\varepsilon^{i_{\sigma(1)}} \otimes \ldots \otimes \varepsilon^{i_{\sigma(p)}} = \varepsilon^{j_{\rho(1)}} \otimes \ldots \otimes \varepsilon^{j_{\rho(p)}}$, т.е. $i_{\sigma(1)} = j_{\rho(1)}, \ldots, i_{\sigma(p)} = j_{\rho(p)}$. Перейдем к перестановке $\tau = \rho\sigma^{-1}$, тогда $i_1 = j_{\tau(1)}, \ldots, i_p = j_{\tau(p)}$. Но индексы i и j у нас упорядочены по возрастанию, следовательно, τ — тождественная перестановка. Значит, $\sigma = \rho$ и тогда $i_1 = j_1, \ldots, i_p = j_p$. Следовательно, все слагаемые в (11) различны, значит все коэффициенты в (11) равны нулю, что и доказывает линейную независимость системы тензоров $\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p}$, $i_1 < \ldots < i_p$, и то, что она является базисом в Λ_p. □

6.10 Связь между линейной зависимостью и тривиальностью внешнего произведения

Рассмотрим пространство Λ_1. Оно совпадает с Θ_1^0 и с V', поэтому его элементами являются ковекторы или линейные функции на V.

Теорема 6.10.1 *Пусть $\varphi^1, \ldots, \varphi^p \in \Lambda_1 = V'$, тогда $\varphi^1 \wedge \ldots \wedge \varphi^p = 0$ тогда и только тогда, когда ковекторы $\varphi^1, \ldots, \varphi^p$ линейно зависимы.*

Доказательство. Если ковекторы линейно зависимы, то один из них можно выразить в виде линейной комбинации остальных. Пусть $\varphi^p = \lambda_1\varphi^1 + \ldots + \lambda_{p-1}\varphi^{p-1}$, тогда

$$\begin{aligned} \varphi^1 \wedge \ldots \wedge \varphi^p &= \varphi^1 \wedge \ldots \wedge \varphi^{p-1} \wedge (\lambda_1\varphi^1 + \ldots + \lambda_{p-1}\varphi^{p-1}) = \\ &= \lambda_1\varphi^1 \wedge \ldots \wedge \varphi^{p-1} \wedge \varphi^1 + \ldots + \lambda_{p-1}\varphi^1 \wedge \ldots \wedge \varphi^{p-1} \wedge \varphi^{p-1} = 0 \end{aligned}$$

поскольку в каждом слагаемом встречается пара совпадающих сомножителей.

Для доказательства обратного утверждения предположим, что $\varphi^1 \wedge \ldots \wedge \varphi^p = 0$ и разложим ковекторы $\varphi^1, \ldots, \varphi^p$ по некоторому базису $\varepsilon^1, \ldots, \varepsilon^n$ пространства V', $\varphi^k = \varphi^k_{i_k}\varepsilon^{i_k}$. Тогда

$$\begin{aligned} 0 &= \varphi^1_{i_1}\varepsilon^{i_1} \wedge \ldots \wedge \varphi^p_{i_p}\varepsilon^{i_p} = \varphi^1_{i_1} \cdot \ldots \cdot \varphi^p_{i_p} \cdot \varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p} = \\ &= \varphi^1_{i_1} \cdot \ldots \cdot \varphi^p_{i_p} \cdot \sum_{\sigma \in S_p} (-1)^\sigma \varepsilon^{i_{\sigma(1)}} \otimes \ldots \otimes \varepsilon^{i_{\sigma(p)}}. \end{aligned}$$

Эту полилинейную функцию можно применить к наборам векторов $e_{k_1}, \ldots, e_{k_p}$, причем ее значение на этих наборах равно нулю (т.к. она тождественно нулевая). Вычислим ее значение на таком наборе векторов:

$$0 = \varphi^1_{i_1} \cdot \ldots \cdot \varphi^p_{i_p} \cdot \underbrace{\sum_{\sigma \in S_p} (-1)^\sigma \varepsilon^{i_{\sigma(1)}} \otimes \ldots \otimes \varepsilon^{i_{\sigma(p)}}(e_{k_1}, \ldots, e_{k_p})}. \tag{12}$$

В подчеркнутой сумме имеется только одно ненулевое слагаемое, а именно, то слагаемое, в котором $i_{\sigma(1)} = k_1, \ldots, i_{\sigma(p)} = k_p$, т.е. когда $i_1 = k_{\sigma^{-1}(1)}, \ldots, i_p = k_{\sigma^{-1}(p)}$, причем это ненулевое слагаемое равно $(-1)^\sigma$. Поэтому наша сумма (12) преобразуется в

$$0 = \sum_{\sigma \in S_p} (-1)^\sigma \varphi^1_{k_{\sigma^{-1}(1)}} \cdot \ldots \cdot \varphi^p_{k_{\sigma^{-1}(p)}} = \sum_{\rho \in S_p} (-1)^\rho \varphi^1_{k_{\rho(1)}} \cdot \ldots \cdot \varphi^p_{k_{\rho(p)}} \tag{13}$$

(здесь мы переобозначили $\rho = \sigma^{-1}$). А это не что иное, как формула определителя некоторой матрицы.

Рассмотрим прямоугольную матрицу

$$\Phi = \begin{pmatrix} \varphi^1_1 & \ldots & \varphi^1_n \\ \vdots & \ddots & \vdots \\ \varphi^p_1 & \ldots & \varphi^p_n \end{pmatrix}$$

и пусть $\Phi_{k_1,\ldots,k_p}$ — это квадратная матрица (минор), составленная из столбцов $k_1, \ldots, k_p$ матрицы Φ, а остальные столбцы выкинуты. Тогда правая часть (13) будет формулой определителя этой матрицы, т.е. $0 = \det \Phi_{k_1,\ldots,k_p}$. Т.к. числа $k_1, \ldots, k_p$ мы можем брать любыми, то все максимальные миноры в матрице Φ равны нулю, следовательно ее строки линейно зависимы, следовательно ковекторы $\varphi^1, \ldots, \varphi^p$ линейно зависимы. □

Оглавление

1 Линейные пространства **2**
1.1 Линейное (векторное) пространство . . . 2
1.2 Аффинное пространство . . . 3
1.3 Факторпространство . . . 4
1.4 Линейная зависимость векторов . . . 5
1.5 Размерность . . . 6
1.6 Пересечение и сумма подпространств . . . 7
1.7 Прямая сумма подпространств. Внешняя прямая сумма . . . 8
1.8 Координаты . . . 9
1.9 Изоморфизмы векторных пространств . . . 10
1.10 Двойственное векторное пространство . . . 11
1.11 Канонический изоморфизм между пространством и его вторым двойственным . . 12

2 Линейные операторы **14**
2.1 Линейные отображения . . . 14
2.2 Инвариантное подпространство . . . 16
2.3 Невырожденные операторы. Собственные значения и собственные векторы . . . 16
2.4 Проекторы . . . 17
2.5 Нильпотентные операторы . . . 18
2.6 Многочлены от операторов . . . 20
2.7 Характеристический многочлен . . . 21
2.8 Еще несколько утверждений об операторах . . . 23
2.9 Корневое подпространство . . . 24
2.10 Теорема о разложении пространства в прямую сумму корневых подпространств . 25
2.11 Жорданова клетка. Жорданова нормальная форма оператора . . . 26
2.12 Приложения теоремы Жордана . . . 27
2.13 Овеществление и компексификация . . . 29
2.14 Два канонических изоморфизма . . . 30
2.15 Вещественная Жорданова нормальная форма . . . 31

3 Евклидовы и унитарные пространства **33**
3.1 Нормированное векторное пространство . . . 33
3.2 Евклидовы и унитарные пространства . . . 34
3.3 Процесс ортогонализации . . . 35
3.4 Ортогональное дополнение . . . 36
3.5 Метод наименьших квадратов . . . 37
3.6 Канонический изоморфизм $V \cong V'$ . . . 40

4 Операторы в евклидовых и унитарных пространствах **41**
4.1 Изоморфизмы гильбертовых пространств . . . 41
4.2 Ортогональные и унитарные операторы . . . 42
4.3 Сопряженный оператор . . . 44
4.4 Канонический вид матрицы самосопряженного оператора . . . 46
4.5 Канонический вид матрицы кососимметрического оператора . . . 47
4.6 Частичные изометрии . . . 48
4.7 Неотрицательные и положительные операторы. Извлечение неотрицательного квадратного корня из оператора . . . 49
4.8 Полярное разложение операторов . . . 51
4.9 Канонический вид матрицы нормального оператора . . . 52

5 Билинейные и полуторалинейные функции **54**
5.1 Билинейные функции (формы) . . . 54
5.2 Канонический изоморфизм $B(V) \cong L(V, V')$. Невырожденность . . . 55
5.3 Симметричные, кососимметричные и эрмитовы функции . . . 56
5.4 Ортогональное дополнение . . . 57
5.5 Нормальный вид матрицы (косо)симметрической или эрмитовой функции . . . 58
5.6 Единственность нормального вида . . . 60
5.7 Теорема Якоби. Критерий Сильвестра . . . 61
5.8 Пространства с обобщенным скалярным произведением. Группы операторов, сохраняющих скалярное произведение . . . 63
5.9 Связь между псевдоортогональной, псевдоунитарной и симплектической группами 65
5.10 Симплектические пространства. Лагранжевы подпространства . . . 66
5.11 Симметрические билинейные функции на евклидовом пространстве . . . 68
5.12 Пара симметричных билинейных функций . . . 69
5.13 Принцип минимакса . . . 70

6 Тензоры **72**
6.1 Тензоры. Пространство тензоров . . . 72
6.2 Координатное определение тензоров . . . 73
6.3 Базис в пространстве тензоров . . . 74
6.4 Линейные операторы как тензоры . . . 75
6.5 Свертка тензоров . . . 75
6.6 Поднятие и опускание индексов . . . 76
6.7 Оператор альтернирования. Внешние формы . . . 76
6.8 Внешнее умножение, его свойства . . . 77
6.9 Базис в пространстве внешних форм . . . 79
6.10 Связь между линейной зависимостью и тривиальностью внешнего произведения . 80