В. М. Мануйлов

Курс лекций по линейной алгебре и геометрии

Механико-математический факультет
I-й курс, 2-й семестр
2007/2008 уч. год, поток механиков

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 $\cdot$.

В нашем курсе поле $\mathbb{K}$ всегда будет предполагаться полем вещественных или комплексных чисел.

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 Линейные подмногообразия

Определение 1.2.1 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 1.2.2 *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 1.2.3 $a_1 + V = a_2 + V \Longleftrightarrow a_1 - a_2 \in V$.

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 $v + b' \in V$. □

1.3 Аффинное пространство

Определение 1.3.1 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.

1.4 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:

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

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

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

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

Определение 1.4.7 Пусть $V \subset W$ — линейное подпространство линейного пространства W. Система векторов $a_1, \ldots, a_n$ называется линейно независимой относительно подпространства V, если из равенства $\lambda_1 a_1 + \ldots + \lambda_n a_n \in V$ следует, что все $\lambda_1, \ldots, \lambda_n$ равны нулю.

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. $\square$

Описанная в этой лемме система веторов называетсяотносительным базисом (относительно подпространства V). Точнее,

Определение 1.5.7 система векторов $e_{r+1}, \ldots, e_n$ называется относительным базисом относительно подпространства V, если, дополнив ее базисом подпространства V, мы получим базис пространства W.

В частности, относительный базис представляет собой линейно независимую относительно подпространства V систему векторов.

Лемма 1.5.8 *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. $\square$

Замечание: второе утверждение леммы неверно в бесконечномерном случае.

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 $_1+$ $_2$, 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 эквиваленты:*

(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. □

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. $\square$

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 $e_1, \dots, e_n$ — 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 $V \cong \mathbb{K}^n$. $\square$

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

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

Верно и обратное:

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

Доказательство. Допустим, что $\dim V < \dim W$, пусть $e_1, \dots, e_n$ — базис в W, тогда вектора $f(e_1), \dots, f(e_n) \in V$ 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 $\dim V < n$. $\square$

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 $e_1, \dots, e_n$ — базис в V. Тогда f является изоморфизмом тогда и только тогда, когда $f(e_1), \dots, f(e_n)$ — 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 $f(e_1, \dots, f(e_n))$ — 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 $g = f^{-1}$. $\square$

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. $\delta^i_j = \begin{cases} 1 & \text{при } i = j \\ 0 & \text{при } i \neq j \end{cases}$. 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 $\varepsilon^1, \ldots, \varepsilon^n$. Если $f = \lambda_1\varepsilon^1 + \ldots + \lambda_n\varepsilon^n = 0$ (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 $\forall f \in V'$, $\exists \lambda_1, \ldots, \lambda_n \in \mathbb{K}$ такие что $f = \lambda_i\varepsilon^i$. Возьмем произвольный функционал $f \in V'$, тогда, если $x = x^i e_i$, то $f(x) = x^i f(e_i)$. Возьмем $\lambda_i = f(e_i)$, тогда получим, что

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

что и требовалось доказать. □

Пусть нам даны базисы $e_1, \ldots, e_n$ и $\widetilde{e}_1, \ldots, \widetilde{e}_n$ в пространстве V и двойственные к ним базисы $\varepsilon^1, \ldots, \varepsilon^n$ и $\widetilde{\varepsilon}^1, \ldots, \widetilde{\varepsilon}^n$ в пространстве V'. Пусть C — матрица перехода от базиса $e_1, \ldots, e_n$ к $\widetilde{e}_1, \ldots, \widetilde{e}_n$, найдем матрицу перехода от базиса $\varepsilon^1, \ldots, \varepsilon^n$ к $\widetilde{\varepsilon}^1, \ldots, \widetilde{\varepsilon}^n$. 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 $\varepsilon^1, \ldots, \varepsilon^n$ и $\widetilde{\varepsilon}^1, \ldots, \widetilde{\varepsilon}^n$ 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, $f(\widetilde{e}_k) = f_i\varepsilon^i(c^j_k e_j) = f_i c^i_k$, так как $(\widetilde{e}_1 \ldots \widetilde{e}_n) = (e_1 \ldots e_n)C$. Отсюда получаем, что $\widetilde{f}_k = f_i c^i_k$. Следовательно $(\widetilde{f}_1 \ldots \widetilde{f}_n) = (f_1 \ldots f_n)C$, 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: $(\widetilde{\varepsilon}^1 \ldots \widetilde{\varepsilon}^n) = (\varepsilon^1 \ldots \varepsilon^n)(C^{-1})^t$, 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!

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 1.10.2 *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 Евклидовы и унитарные пространства

2.1 Евклидовы и унитарные пространства

Определение 2.1.1 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 2.1.2 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 2.1.3 Длиной вектора a в евклидовом или эрмитовом пространстве называется число $|a| = \sqrt{(a,a)}$.

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

Лемма 2.1.4 (Неравенство Коши–Буняковского) *Для любых двух векторов a, b евклидова или эрмитова 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. $\square$

Лемма 2.1.5 (Неравенство треугольника) *Для любых двух векторов a, b евклидова или эрмитова пространства имеет место неравенство $|a + b| \leqslant |a| + |b|$.*

Доказательство. 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. □

2.2 Процесс ортогонализации

Определение 2.2.1 Два вектора a, b называются ортогональными ($a \perp b$), если $(a, b) = 0$. Система векторов (или базис) $e_1, \ldots, e_n$ называется ортонормированной, если $(e_i, e_j) = \delta_{ij}$, т.е. если векторы попарно ортогональны и длина каждого из них равна 1.

Лемма 2.2.2 *Ортонормированная система векторов является линейно независимой.*

Доказательство. Пусть $\lambda_1 e_1 + \ldots + \lambda_n e_n = 0$. Умножим скалярно обе части этого равенства (слева) на вектор e_i: $((e_i, \lambda_1 e_1 + \ldots + \lambda_n e_n)) = \lambda_1(e_i, e_1) + \ldots + \lambda_n(e_i, e_n) = \lambda_i = 0$. □

Пусть $e_1, \ldots, e_n$ — некоторая линейно независимая система векторов. 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 2.2.3 *Существует такой набор попарно ортогональных векторов $a_1, \ldots, a_n$, что для каждого ноиера 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.

2.3 Ортогональное дополнение

Определение 2.3.1 Пусть $V \subset W$ — подпространство евклидова или эрмитова 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 2.3.2 $W = V \oplus V^\perp$.

Доказательство. Пусть $e_1, \dots, e_k$ — базис в V, дополним его до базиса всего пространства W векторами $e_{k+1}, \dots, e_n$. Применив процесс ортогонализации Грама–Шмидта, получим ортогональный базис $a_1, \dots, a_k, a_{k+1}, \dots, a_n$ в W, причем его первая часть будет базисом в V, т.к. $\langle a_1, \dots, a_k\rangle = \langle e_1, \dots, e_k\rangle = V$. Покажем, что векторы $a_{k+1}, \dots, a_n$ 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 $\nu_1, \dots, \nu_k$ должны быть нулевыми, так как иначе вектор v не был бы ортогонален всем векторам $a_1, \dots, a_k$. Верно и обратное: если первые k координат какого-то вектора в базисе $a_1, \dots, a_n$ 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 2.3.3 $|a_2| \geqslant |a_\perp|$.

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

Видно, что это определение корректно, т.к. $-1 \leqslant \frac{(a,b)}{|a|\cdot|b|} \leqslant 1$, и 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.

В многомерном случае геометрия аналогична обычной геометрии. Так, например, в прямоугольных треугольниках с одинаковой гипотенузой, чем больше угол, тем больше противолежащий катет.

Утверждение 2.3.5 *Пусть* $c = a_1 + b_1 = a_2 + b_2$, $a_1 \perp b_1$, $a_2 \perp b_2$ *и* $|b_1| < |b_2|$. *Тогда* $(\widehat{a_1, c}) < (\widehat{a_2, c})$.

Доказательство. Поскольку $(a_i, c) = (a_i, a_i + b_i) = |a_i|^2$, $i = 1, 2$, то $\cos(\widehat{a_i, c}) = \frac{|a_i|}{|c|}$, а из теоремы Пифагора $\sin(\widehat{a_i, c}) = \frac{|b_i|}{|c|}$, что и доказывает утверждение.

Утверждение 2.3.6 *(обозначения те же, что и ранее)* $(\widehat{a, a_0}) \leqslant (\widehat{a, a_1})$.

Доказательство. Найдем такое число λ, чтобы вектор λa_1 был бы ортогонален вектору $b = a - \lambda a_1$:

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

Если $(a_1, a) \leqslant 0$, то угол $(\widehat{a, a_1}) \geqslant \pi/2$ и утверждение очевидно. Если $(a_1, a) > 0$, то $\lambda > 0$ и угол между a и a_1 равен углу между a и λa_1. Применив предыдущую лемму, получаем требуемое утверждение. □

Определение 2.3.7 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.

2.4 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 $b = a_1 x^1 + \ldots + a_n x^n$, т.е. узнать неизвестные коэффициенты $x^1, \ldots, x^n$. Мы делаем m измерений (для точности берем $m > n$) и решаем систему уравнений $\left\{ \begin{array}{rcl} a_1^1 x^1 + \ldots + a_n^1 x^n & = & b^1 \\ \ldots \quad \ldots \quad \ldots & & \\ a_1^m x^1 + \ldots + a_n^m x^n & = & b^m \end{array} \right.$. 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 $c^j = a_1^j x^1 + \ldots + a_n^j x^n$ 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, $b_0 = x^1 a_1 + \ldots + x^n a_n$. Тогда, взяв скалярные произведения с векторами базиса, получаем следующие равенства:

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

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 координаату $x^1, \ldots, x^n$ 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 $a_1, \dots, a_n$ линейно независимы.

Параллелепипеды. Матрица Грама

Определение 2.4.1 Пусть $a_1, \dots, a_n$ — система векторов в векторном пространстве V. *Параллелепипедом*, натянутым на векторы $a_1, \dots, a_n$ называется множество векторов (точек) $\Pi(a_1, \dots, a_n) = \{c \in V : c = x^1 a_1 + \ldots + x^n a_n, \quad 0 \leqslant x^1, \dots, x^n \leqslant 1\}$.

Определение 2.4.2 Определим *n-мерный объем* Vol_n параллелепипеда $\Pi(a_1, \dots, a_n)$ индуктивно:

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

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

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

Теорема 2.4.3 $(\mathrm{Vol}_n \Pi(a_1, \dots, a_n))^2 = \det G(a_1, \dots, a_n)$.

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 $a_1, \dots, a_{n-1}$: $a_n = (a_n)_0 + (a_n)_\perp = \lambda_1 a_1 + \ldots + \lambda_{n-1} a_{n-1} + b$, где $b = (a_n)_\perp$ (т.е. b ортогонален векторам $a_1, \dots, a_{n-1}$) и $|b| = d(a_n, \langle a_1, \dots, a_{n-1} \rangle)$. Имеем:

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

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

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

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

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

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

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

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

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

Лемма 2.4.4 *Векторы $a_1,\ldots,a_n$ линейно зависимы тогда и только тогда, когда $\det G(a_1,\ldots,a_n) = 0$.*

Доказательство. Если векторы $a_1,\ldots,a_n$ линейно зависимы, то, без ограничения общности, можно считать, что $a_n = \lambda_1 a_1 + \ldots + \lambda_{n-1}a_{n-1}$. Тогда $\mathrm{Vol}_n \Pi(a_1,\ldots,a_n) = \mathrm{Vol}_{n-1}\Pi(a_1,\ldots,a_{n-1}) \cdot \underbrace{d(a_n, \langle a_1,\ldots,a_{n-1}\rangle)}_{=0} = 0$, следовательно, по предыдущей теореме, $\det G(a_1,\ldots,a_n) = 0$.

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

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

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

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

Пусть V — евклидово или эрмитово пространство, $a_1, \dots, a_n$ и $b_1, \dots, b_n$ — два базиса в пространстве V, и $C = \begin{pmatrix} c_1^1 & \dots & c_n^1 \\ \vdots & \ddots & \vdots \\ c_1^n & \dots & c_n^n \end{pmatrix}$ — матрица перехода от первого базиса к второму. Посмотрим, как связаны матрицы Грама $G = G(a_1, \dots, a_n)$ и $G' = G(b_1, \dots, b_n)$. Поскольку $b_k = c_k^i a_i$ и элементы матрицы G' равны $g'_{ij} = (b_i, b_j) = (c_i^k a_k, c_j^l a_l) = \overline{c_i^k}(a_k, a_l)c_j^l = \overline{c_i^k} g_{kl} c_j^l$, то $G' = \overline{C}^t G C$. Если базис $a_1, \dots, a_n$ был ортонормированным, то $G = E = \begin{pmatrix} 1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & 1 \end{pmatrix}$ и $G' = \overline{C}^t C$.

Утверждение 2.4.5 *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 $G = G(a_1, \dots, a_n)$, выберем в пространстве ортонормированный базис, пусть C — матрица перехода от этого ортонормированного базиса в базис $a_1, \dots, a_n$, тогда $G = \overline{C}^t C$.

$\Leftarrow$: пусть $G = \overline{C}^t C$, тогда C можно считать матрицей перехода от некоторого ортонормированного базиса $a_1, \dots, a_n$ к базису $b_1, \dots, b_n$, и тогда G — это матрица Грама для векторов $b_1, \dots, b_n$. $\square$

3 Линейные операторы

3.1 Линейные отображения

Определение 3.1.1 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 3.1.2 $L(V, W) \cong M_{m,n}(\mathbb{K})$.

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. $\square$

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 $A'_f = C_2^{-1} A_f C_1$.

Определение 3.1.3 *Ядром* $\operatorname{Ker} f$ линейного отображения $f : V \to W$ называется множество всех векторов, переходящих в ноль, $\operatorname{Ker} f = \{x \in V : f(x) = 0\}$.

Образом $\operatorname{Im} f$ линейного оператора $f : V \to W$ называется множество векторов $y \in W$, для которых существует прообраз, $\operatorname{Im} f = \{y \in W : \exists x \in V, f(x) = y\}$.

Лемма 3.1.4 *Ядро любого линейного оператора является линейным подпространством в V; образ любого линейного оператора является линейным подпространством в W.*

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 3.1.5 $\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim V$.

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. □

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 3.1.6 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 3.1.7 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 3.1.8 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 3.1.9 *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.

3.2 Инвариантное подпространство

Определение 3.2.1 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.

Определение 3.2.2 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. $A_f = \left(\begin{array}{c|c} A_{f_1} & \star \\ \hline 0 & \star \end{array}\right)$.

3.3 Невырожденные операторы. Собственные значения и собственные векторы

Определение 3.3.1 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 3.3.2 *Все эти пять свойств эквивалентны.*

Доказательство. 2) $\Longleftrightarrow$ 3), т.к. $\dim V = \dim \operatorname{Ker} f + \dim \operatorname{Im} f$.
1) $\Longleftrightarrow$ 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) $\Longleftrightarrow$ 4) — это мы знаем из курса высшей алгебры.
1) $\Longleftrightarrow$ 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 3.3.3 Пусть f — линейный оператор в линейном пространстве V. Если для некоторого числа $\lambda \in \mathbb{K}$ и для некоторого ненулевого вектора $v \in V$ выполняется равенство $f(v) = \lambda v$, то λ называется *собственным значением* оператора f, а v — *собственным вектором* оператора f, отвечающим собственному значению λ.

Лемма 3.3.4 *λ является собственным значением оператора f тогда и только тогда, когда оператор $f - \lambda \cdot id$ вырожден.*

Доказательство.
$\Longrightarrow$: Если $f(v) = \lambda v$, то $(f - \lambda \cdot id)(v) = 0$, значит, ядро оператора $(f - \lambda \cdot id)$ содержит ненулевой вектор v, откуда следует вырожденность этого оператора.
$\Longleftarrow$: Вырожденность $(f - \lambda \cdot id)$ означает наличие нетривиального ядра у этого оператора. Возьмем в качестве v любой ненулевой вектор из ядра $\mathrm{Ker}(f - \lambda \cdot id)$, тогда $f(v) = \lambda v$. □

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

Лемма 3.3.5 *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 3.3.6 *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. □

3.4 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 $V_2 \in \mathrm{Ker}\, f$. На самом деле $V_2 = \mathrm{Ker}\, f$, 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 3.4.1 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 3.4.2 *Если $f^2 = f$, то оператор $f : W \to W$ является оператором проектирования для некоторых V_1 и V_2.*

Доказательство. Возьмем $V_1 = \mathrm{Im}\, f$ и $V_2 = \mathrm{Ker}\, f$ 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 $a \in \mathrm{Ker}\, f$, т.е. $f(a) = 0$ и $a \in \mathrm{Im}\, f$, 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.

3.5 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 $p(f)$.

Аналогично, можно определить многочлен от матрицы. Для матрицы A определим $p(A)$ формулой $p(A) = a_0 E + a_1 A + a_2 A^2 + \ldots + a_n A^n$, где E — единичная матрица. Поскольку при фиксированном базисе имеется взаимно однозначное соотвествие между операторами и матрицами, все утверждения о многочленах от операторов допускают переформулировку для многочленов от матриц.

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 3.5.1 *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 3.5.2 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.

Аналогично можно определить минимальный многочлен для матриц вместо операторов.

Лемма 3.5.3 *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. $\square$

Лемма 3.5.4 *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

$$p(\lambda)x = a_0x + a_1\lambda x + \ldots + a_n\lambda^n x = a_0x + a_1fx + \ldots + a_nf^nx = p(f)x = 0,$$

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. $\square$

3.6 Характеристический многочлен

Определение 3.6.1 Многочлен $P_f(t) = \det(f - \lambda \cdot id)$ называется *характеристическим многочленом* оператора f.

Отметим, что характеристический многочлен можно определить и для в матриц (вместо операторов): $P_A(t) = \det(A - \lambda E)$, где A — матрица.

Отметим роль некоторых коэффициентов характеристического многочлена. Если его записать в виде $P_f(t) = a_0 + a_1t + \ldots + a_nt^n$, то тогда $a_n = (-1)^n$, $a_{n-1} = (-1)^{n-1}\operatorname{tr} f$, $a_0 = \det f$.

Лемма 3.6.2 *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 3.6.3 *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 $f_1x = \lambda x$, 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 $P_{f'}(t)$). $\Box$

Теорема 3.6.4 (Гамильтона-Кэли) *Характеристический многочлен $P_f(t)$ оператора $f : W \to W$ аннулирует этот оператор, т.е. $P_f(f) = 0$.*

Доказательство. В силу взаимно однозначного соответствия между матрицами и операторами, мы докажем "матричный вариант" этой теоремы: $P_A(A) = 0$ для произвольной матрицы A.

Запишем обратную матрицу (для тех значений λ, для которых она определена) в виде $(A - \lambda E)^{-1} = \frac{1}{P_A(\lambda)}C(\lambda)$, где $C(\lambda)$ — матрица, составленная из миноров матрицы $A - \lambda E$. Отсюда

$$(A - \lambda E)C(\lambda) = P_A(\lambda)E. \qquad (3)$$

Это равенство очевидно выполняется для всех λ, кроме корней характеристического многочлена. А поскольку обе части этого матричного равенства состоят из многочленов, из их непрерывности следует выполнение этого равенства для всех значений λ. Разложим матрицу $C(\lambda)$, состоящую из многочленов, по степеням λ: $C(\lambda) = C_0 + C_1\lambda + C_2\lambda^2 + \ldots + C_{n-1}\lambda^{n-1}$, где $C_0, C_1, \ldots, C_{n-1}$ — числовые матрицы. В таком же виде запишем характеристический многочлен $P_A(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \ldots + a_n\lambda^n$ и распишем матричное равенство (3) по степеням λ, т.е. для каждой степени напишем равенство коэффициентов левой и правой части матричного равенства:

$$\begin{array}{rcl} AC_0 & = & a_0E \\ AC_1 - C_0 & = & a_1E \\ AC_2 - C_1 & = & a_2E \\ \ldots & & \ldots \\ AC_{n-1} - C_{n-2} & = & a_{n-1}E \\ -C_{n-1} & = & a_nE. \end{array}$$

Умножим первое равенство на E, второе — на A, третье — на A^2, и т.д., и сложим. Тогда в правой части получится $a_0E + a_1A + a_2A^2 + \ldots + a_nA^n = P_A(A)$, а в левой — все слагаемые взаимно уничтожатся. Т.е. получится, что $P_A(A) = 0$. $\Box$

3.7 Диагонализируемые операторы

Определение 3.7.1 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 3.7.2 *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 векторы, 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). Применим индукцию по количеству линейно независимых векторов.

1) База индукции. Вектор a_1 отличен от нуля, поэтому система, состоящая из него одного линейно независима.

2) Индуктивный переход. Пусть первые $k-1$ векторов линейно независимы. Докажем, что тогда и первые k векторов тоже линейно независимы. Предположим обратное, т.е. что т.е. существуют скаляры $\alpha_1, \ldots, \alpha_k$, не все равные нулю, т.ч. $\alpha_1 a_1 + \ldots + \alpha_k a_k = 0$. Поскольку первые $k-1$ векторов по предположению линейно независимы, последний вектор есть линейная комбинация остальных, $a_k = -\frac{\alpha_1}{\alpha_k} a_1 - \ldots - \frac{\alpha_{k-1}}{\alpha_k} a_{k-1}$. Применив оператор f к обеим частям этого равенства, получим, что $f a_k = f(-\frac{\alpha_1}{\alpha_k} a_1 - \ldots - \frac{\alpha_{k-1}}{\alpha_k} a_{k-1})$, т.е. $\lambda_k a_k = -\frac{\alpha_1}{\alpha_k} \lambda_1 a_1 - \ldots - \frac{\alpha_{k-1}}{\alpha_k} \lambda_{k-1} a_{k-1}$, но, с другой стороны, $\lambda_k a_k = \lambda_k(-\frac{\alpha_1}{\alpha_k} a_1 - \ldots - \frac{\alpha_{k-1}}{\alpha_k} a_{k-1})$. Приравняв выражения в правых частях равенств, получим, что $(\lambda_k - \lambda_1)\frac{\alpha_1}{\alpha_k} a_1 + \ldots + (\lambda_k - \lambda_{k-1})\frac{\alpha_{k-1}}{\alpha_k} a_{k-1} = 0$. Т.к. все λ_i различны, то из линейной независимости векторов $a_1, \ldots, a_{k-1}$ следует, что все коэффициенты $\alpha_1, \ldots, \alpha_{k-1}$ равны нулю. Но тогда $\alpha_k = 0$, откуда следует линейная независимость системы из k векторов. $\square$

Лемма 3.7.3 *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 $A_f = \left(\begin{array}{c|ccc} \lambda & & \star & \\ \hline & \lambda_1 & & \star \\ \mathbf{0} & & \ddots & \\ & \mathbf{0} & & \lambda_{n-1} \end{array}\right)$ также будет верхнетреугольной. $\square$

Лемма 3.7.4 *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 3.7.5 *Пусть оператор f такой, что в алгебраически замкнутом поле характеристический многочлен $P_f(t)$ имеет единственный корень λ, тогда некоторая степень оператора $g = f - \lambda \cdot id$ равна нулю.*

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, $g^n = 0$. □

3.8 Жордановы клетки

Рассмотрим оператор f, заданный в базисе $e_1, \ldots, e_n$ матрицей

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}.$$

Такая матрица называется *жордановой клеткой* порядка n, с собственным значением λ. Действие такого оператора на базисных векторах устроено так: $f(e_1) = \lambda e_1$, $f(e_2) = \lambda e_2 + e_1$, $f(e_3) = \lambda e_3 + e_2, \ldots,$ $f(e_n) = \lambda e_n + e_{n-1}$.

Нормальной или жордановой формой называется блочно-диагональная матрица, все блоки которой являются жордановыми клетками.

3.9 Присоединенные векторы и корневое подпространство

Обозначим через $V_\lambda^{(1)}$ подпространство собственных векторов оператора f, отвечающих собственному значению λ (вместе с нулевым вектором). Иначе это можно записать как $V_\lambda^{(1)} = \operatorname{Ker}(f - \lambda \cdot id)$.

Аналогично определим $V_\lambda^{(2)} = \operatorname{Ker}(f - \lambda \cdot id)^2$, $V_\lambda^{(3)} = \operatorname{Ker}(f - \lambda \cdot id)^3$ и т.д.

Получается цепочка вложенных друг в друга подпространств $\{0\} \subset V_\lambda^{(1)} \subset V_\lambda^{(2)} \subset \ldots \subset V$.

Вектор v называется присоединенным вектором 1-го порядка оператора f, отвечающих собственному значению λ, если $v \in V_\lambda^{(2)}$, $v \notin V_\lambda^{(1)}$. Аналогично, присоединенным вектором порядка k, если $v \in V_\lambda^{(k+1)}$, $v \notin V_\lambda^{(k)}$.

Покажем, что рост подпространств $V_\lambda^{(k)}$ стабилизируется с ростом k.

Поскольку пространство, в котором действует наш оператор, конечномерно, то неубывающая числовая последовательность размерностей $\dim V_\lambda^{(1)} \leqslant \dim V_\lambda^{(2)} \leqslant \dim V_\lambda^{(3)} \leqslant \ldots$ не может неограниченно возрастать, значит, найдется какой-то номер k, для которого соседние члены этой последовательности совпадут.

Утверждение 3.9.1 *Если на каком-то шаге этой цепочки $V_\lambda^{(k-1)}$ и $V_\lambda^{(k)}$ совпали, то они будут совпадать и дальше, т.е. $V_\lambda^{(i)} = V_\lambda^{(i+1)}$ при $i > k$.*

Доказательство. Обозначим $f - \lambda \cdot id$ через g. 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. □

Обозначим через p тот номер, с которого происходит стабилизация цепочки подпространств $V_\lambda^{(k)}$. Тогда $\cup_{k=1}^{\infty} V_\lambda^{(k)} = V_\lambda^{(p)}$. Обозначим это пространство через V_λ и назовем его *корневым подпространством* оператора f, отвечающим собственному значению λ.

Утверждение 3.9.2 *Подпространство $V_\lambda \subset W$ инвариантно относительно оператора f.*

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

3.10 Теорема о разложении пространства в прямую сумму корневых подпространств

Лемма 3.10.1 *Пусть $f : V \to V$, λ — его собственное значение. Тогда имеет место разложение V в прямую сумму, $V = V_\lambda \oplus W$, где W — инвариантное подпространство, причем ограничение оператора f на W обратимо.*

Доказательство. Без ограничения общности можно считать, что $\lambda = 0$. Для этого достаточно перейти от оператора f к оператору $f - \lambda \cdot id$.

Положим $W = \operatorname{Im} f^p$ и докажем, что $V = V_\lambda \oplus W = \operatorname{Ker} f^n \oplus \operatorname{Im} f^p$. Инвариантность $\operatorname{Im} f^p$ очевидна: если $y \in \operatorname{Im} f^p$, то существует такой x, что $y = f^p(x)$; тогда $f(y) = f^{p+1}(x) \in \operatorname{Im} f^{p+1}$. Но $\operatorname{Im} f^{p+1} \subset \operatorname{Im} f^p$, а размерности этих образов совпадают, т.к. совпадают размерности соотвтествующих ядер, поэтому $\operatorname{Im} f^{p+1} = \operatorname{Im} f^p$.

Покажем, что сумма V_0 и W — прямая. Для этого достаточно доказать, что их пересечение равно нулю. Так как сумма размерностей ядра и образа равна размерности всего пространства, из того, что пересечение — нулевое, будет следовать, что $\dim V = \dim V_0 + \dim W$, т.е. $V = V_0 \oplus W$.

Пусть $v \in V_0 \cap W$, $v \neq 0$. Т.к. $v \in V_0$, то $f^p(v) = 0$, а т.к. $v \in W$, то $v = f^p(w)$ для некоторого вектора $w \in V$. Тогда $f^p(w) = v \neq 0$, $f^{2p}(w) = f^p(v) = 0$. Из того, что $f^p(w) \neq 0$, следует, что $w \notin V_0$, а из того, что $f^{2p}(w) = 0$, следует, что $w \in V_0^{(2p)} = V_0$. Полученное противоречие доказывает, что $V_0 \cap W = \{0\}$.

Докажем теперь, что ограничение f на W невырождено. Если бы это было не так, что 0 был бы собственным значением оператора $f|_W$, т.е. существовал бы собственный вектор $v \in W$, $v \neq 0$, $f(v) = 0$. Но это бы означало, что $v \in V_0$. Но $V_0 \cap W = \{0\}$. Т.о., 0 не может быть собственным значением для $f|_W$. □

Теорема 3.10.2 *Пусть дан оператор $f : V \to V$ (над алгебраически замкнутым полем). Тогда пространство V является прямой суммой всех корневых подпространств, т.е. $V = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_k}$, где $\lambda_1, \ldots, \lambda_k$ — все собственные значения оператора f.*

Доказательство. Применим нужное число раз предыдущую лемму. После первого применения получим $V = V_{\lambda_1} \oplus W$. Далее, из блочно-диагонального вида матрицы оператора f — она состоит из блоков, отвечающих ограничениям $f|_{V_{\lambda_1}}$ и $f|_W$ — следует, что собственные значения $f|_W$ — $\lambda_2, \lambda_3, \ldots, \lambda_k$. Продолжая этот процесс, мы исчерпаем все пространство. Действительно, если после k-кратного применения леммы мы получим $V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \ldots \oplus V_{\lambda_k} \oplus W'$, то $W' = \{0\}$ — иначе оператор ограничения $f|_{W'}$ будет иметь еще одно собственное значение (поле алгебраически замкнуто), которого не было у f — противоречие. □

Отметим, что ограничение оператора на корневое подпространство имеет единственное собственное значение.

3.11 Жорданова нормальная форма оператора

Теорема 3.11.1 (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. По теореме о корневом разложении, в подходящем базисе матрица оператора f имеет блочно-диагональный вид

$$A_f = \begin{pmatrix} A_{f|_{V_{\lambda_1}}} & & & \\ & A_{f|_{V_{\lambda_2}}} & & \\ & & \ddots & \\ & & & A_{f|_{V_{\lambda_k}}} \end{pmatrix},$$

поэтому достаточно доказать теорему для операторов с единственным собственным значением, каковыми являются ограничения оператора f на корневые подпространства.

Поэтому мы можем предположить без ограничения общности, что f имеет единственное собственное значение. Также без ограничения общности можно считать, что это собственное значение равно нулю.

Наша задача — построить базис в пространстве, в котором действует такой оператор, т.е. в корневом пространстве $V = V_0 = V_0^{(p)}$, которое является объединением подпространств $V_0^{(1)} \subset V_0^{(2)} \subset \ldots \subset V_0^{(p)}$.

Для подпространства $L \subset V$ система векторов $e_{r+1}, \ldots, e_n$ называется относительным базисом (относительно подпространства L), если, будучи дополненной базисом подпространства L, она становится базисом всего пространства V. В любом (конечномерном) пространстве можно выбрать относительный базис относительно любого подпространства.

На первом шаге рассмотрим подпространство $V_0^{(p-1)} \subset V_0^{(p)} = V$ и выберем относительный базис $e_1, \ldots, e_q$ относительно этого подпространства. Очевидно, он будет состоять из присоединенных векторов порядка $p-1$. Поскольку $f(V_0^{(p)}) = V_0^{(p)}$, $f(e_1), \ldots, f(e_q) \in V_0^{(p-1)}$. Покажем, что система векторов $f(e_1), \ldots, f(e_q)$ линейно независима относительно предыдущего подпространства $V_0^{(p-2)}$: если $\alpha_1 f(e_1) + \ldots + \alpha_q f(e_q) = f(\alpha_1 e_1 + \ldots + \alpha_q e_q) \in V_0^{(p-2)}$, то $\alpha_1 e_1 + \ldots + \alpha_q e_q \in V_0^{(p-1)}$, и из определения векторов $e_1, \ldots, e_q$ следует, что все $\alpha_1 = \ldots = \alpha_q = 0$. Дополним систему векторов $f(e_1), \ldots, f(e_q)$ до относительного базиса в $V_0^{(p-1)}$ относительно подпространства $V_0^{(p-2)}$ векторами $g_1, \ldots, g_s$. Т.о. относительный базис $V_0^{(p-1)}$ относительно подпространства $V_0^{(p-2)}$ состоит из векторов $f(e_1), \ldots, f(e_q), g_1, \ldots, g_s$. К этому относительному базису применим оператор f и вновь полученную систему векторов опять дополним до относительного базиса $V_0^{(p-2)}$ относительно подпространства $V_0^{(p-3)}$. Этот процесс можно продолжить до конца, т.е. до подпространства $V_0^{(1)}$ и его нулевого подпространства.

Запишем полученные векторы в виде таблицы:

$$\begin{array}{ccccccccccccc}
e_1 & \ldots & e_q & & & & & & & & & & \\
f(e_1) & \ldots & f(e_q) & g_1 & \ldots & g_s & & & & & & & \\
f^2(e_1) & \ldots & f^2(e_q) & f(g_1) & \ldots & f(g_s) & h_1 & \ldots & h_r & & & & \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
f^{p-1}(e_1) & \ldots & f^{p-1}(e_q) & f^{p-2}(g_1) & \ldots & f^{p-2}(g_s) & f^{p-3}(h_1) & \ldots & f^{p-3}(h_r) & \ldots & l_1 & \ldots & l_t
\end{array}$$

Векторы нижней строчки образуют базис подпространства $V_0^{(1)}$, векторы предпоследней строчки образуют относительный базис в $V_0^{(2)}$ относительно $V_0^{(1)}$, поэтому векторы двух последних строчек составляют базис подпространства $V_0^{(2)}$. Добавляя к ним векторы третьей с конца строчки, получаем базис подпространства $V_0^{(3)}$, и т.д. В итоге, все векторы таблицы составляют базис пространства V. Осталось проверить, что это — такой базис, в котором матрица оператора состоит из жордановых клеток.

Для этого рассмотрим вертикальные цепочки векторов таблицы. Обозначим $f^{p-1}(e_1) = \tilde{e}_1$, $f^{p-2}(e_1) = \tilde{e}_2$,... $e_1 = \tilde{e}_p$. Т.к. $\tilde{e}_1$ — собственный вектор, отвечающий нулевому собственному значению, то $f(\tilde{e}_1) = 0$. Далее, по определению, $f(\tilde{e}_2) = f^{p-1}(e_1) = \tilde{e}_1$, $f(\tilde{e}_3) = \tilde{e}_2$,.... Пусть $L_1 = \langle \tilde{e}_1, \ldots, \tilde{e}_p \rangle$ — линейная оболочка векторов первого стольбца. Тогда L_1 переходит в себя, т.е. является инвариантным подпространством, при этом матрица оператора f, ограниченного на L_1, имеет вид жордановой клетки с нулевым собственным значением. Аналогично, векторы второго столбца образуют инвариантное подпространство L_2, и матрица ограничения f на L_2 также является жордановой клеткой, и т.д. Таким образом, матрица оператора f состоит из стольких жордановых клеток, сколько столбцов в таблице.

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 и посчитаем количество клеток, ему отвечающих.

Введем числа $r_k(\lambda) = \mathrm{rk}(f - \lambda \cdot id)^k$ и $N_k(\lambda)$ — количество жордановых клеток размерности k, отвечающих собственному значению λ. Для блочно-диагональных матриц ранги можно считать по каждому блоку отдельно, и суммировать. 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 $J_m(\lambda_i - \lambda)$ не меняются при возведении таких клеток в любую степень (эти клетки невырождены).

Рассмотрим разность $r_0(\lambda) - r_1(\lambda)$. Для каждой отдельно взятой клетки с собственным значением λ 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 $r_0(\lambda) - r_1(\lambda)$ вклад, равный единице, т.е. эта разность равна общему количеству клеток, следовательно $r_0(\lambda) - r_1(\lambda) = N_1(\lambda) + N_2(\lambda) + N_3(\lambda) + \ldots$.

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

3.12 Функции от операторов и 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 ? Проверьте, что если матрица A приведена к жордановому виду $J(A)$ с помощью двух различных матриц перехода, C и D, т.е. если $A = CJ(A)C^{-1} = DJ(A)D^{-1}$, то матрицы $D^{-1}C$ и $f(J(A))$ перестановочны для любой функции f, для которой определено $f(J(A))$.

3.13 Овеществление и комплексификация

Овеществление

Определение 3.13.1 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 $e_1, \ldots, e_n$ — 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 $e_1, \ldots, e_n, ie_1, \ldots, ie_n$ (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 3.13.2 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 $e_1, \ldots, e_n$ 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 $e_1, \ldots, e_n, ie_1, \ldots, ie_n$ 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 3.13.3 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 3.13.4 *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 $e_1, \ldots, e_n$ — базис в $\widetilde{V}$, тогда $e_1, \ldots, e_n, e_{n+1} = j(e_1), \ldots, e_{2n} = j(e_n)$ 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 $i = 1, \ldots, n$, то в этом базисе матрица оператора имеет указанный вид. $\square$

Комплексификация

Определение 3.13.5 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 3.13.6 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 $(e_1, 0), \ldots, (e_n, 0)$.

В дальнейшем мы будем использовать обозначение $a + ib$ для пары (a, b) по аналгоии с комплексными числами.

3.14 Инвариантные подпространства в вещественном случае

В случае алгебраически замкнутого поля каждый оператор имеет собственные значения, и, следовательно, одномерные инвариантные подпространства. В вещественном случае это, вообще говоря, неверно, но имеется более слабое утверждение о существовании по крайней мере двумерных инвариантных подпространств.

Лемма 3.14.1 *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 $\beta \neq 0$). $\square$

4 Операторы в евклидовых и унитарных пространствах

4.1 Сопряженный оператор

Изучим теперь линейные операторы, действующие в евклидовых и эрмитовых пространствах, т.е. в пространствах со скалярным произведением.

Определение 4.1.1 Если для оператора $f : V \to V$ существует такой оператор g, что для любых векторов $a, b \in V$ выполняется равенство $(f(a), b) = (a, g(b))$, то g называется *сопряженным оператором* для f.

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

Сопряженный оператор обозначается $g = f^*$.

Лемма 4.1.2 *Если операторы f_1 и f_2 имеют сопряженные f_1^* и f_2^* соответственно, то операторы $h = f_1 + f_2$ и $g = f_1 f_2$ также имеют сопряженные h^* и g^*, причем $h^* = f_1^* + f_2^*$ b $g^* = f_2^* f_1^*$.*

Доказательство. приведем доказательство для второго утверждения (для композиции операторов), т.к. для первого оно очевидно. $(f_1 f_2(a), b) = (f_2(a), f_1^*(b)) = (a, f_2^* f_1^*(b))$. $\square$

Лемма 4.1.3 *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 эрмитово).*

Доказательство. Пусть матрица оператора f в ортонормированном базисе $e_1, \dots, e_n$ есть $A = \begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & \ddots & \vdots \\ a_1^n & \dots & a_n^n \end{pmatrix}$, а матрица оператора f^* (в том же базисе) есть $B = \begin{pmatrix} b_1^1 & \dots & b_n^1 \\ \vdots & \ddots & \vdots \\ b_1^n & \dots & b_n^n \end{pmatrix}$. 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 $B = \overline{A}^t$. $\square$

Отметим, что условие ортонормированности базиса в лемме является существенным.

Лемма 4.1.4 *Для любого оператора f существует ему сопряженный f^*.*

Доказательство. Выберем ортонормированный базис $e_1, \dots, e_n$, 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 $A_f = \begin{pmatrix} c_1^1 & \dots & c_n^1 \\ \vdots & \ddots & \vdots \\ c_1^n & \dots & c_n^n \end{pmatrix}$. Возьмем произвольные векторы $a = a^i e_i$, $b = b^i e_i$. Тогда

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

и легко видеть, что

$$(a, f(b)) = (\overline{a}^1 \dots \overline{a}^n) \begin{pmatrix} c_1^1 & \dots & c_n^1 \\ \vdots & \ddots & \vdots \\ c_1^n & \dots & c_n^n \end{pmatrix} \begin{pmatrix} b^1 \\ \vdots \\ b^n \end{pmatrix} = \overline{(b, g(a))} = (g(a), b),$$

т.е. $g = f^*$. $\square$

Следствие 4.1.5 *Если f^* — оператор, сопряженный к f, то $(f^*)^* = f$.*

Лемма 4.1.6 *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 $(f^*(v), u) = (v, f(u)) = 0$, т.к. $v \in V^\perp$, а $f(u) \in V$. Следовательно, $f^*(v) \in V^\perp$ для любого $v \in V^\perp$. $\square$

В случае евклидовых пространств операция перехода к сопряженному оператору является линейным оператором в пространстве $L(V)$ линейных операторов, квадрат которого равен единице. Далее мы увидим, что его собственными значениями являются числа 1 и -1. Вопрос: почему в случае эрмитовых пространств эта операция (перехода к сопряженному оператору) не является линейным оператором?

Определение 4.1.7 Оператор f называется *самосопряженным* (или *симметрическим*), если $f^* = f$. Оператор f называется *кососимметрическим*, если $f^* = -f$.

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.

Лемма 4.1.8 *Любой оператор единственным образом представляется в виде суммы симметрического и кососимметрического операторов.*

Доказательство. Разложение нужного вида дает формула $f = \frac{1}{2}(f + f^*) + \frac{1}{2}(f - f^*)$, первое слагаемое которой — симметрический оператор, а второе — кососимметрический. Единственность следует из того, что если оператор одновременно симметрический и кососимметрический, то он равен нулю. $\square$

Лемма 4.1.9 *Пусть λ — собственное значение самосопряженного оператора. Тогда $\lambda \in \mathbb{R}$.*

Доказательство. Ясно, что содержательной эта лемма является лишь в эрмитовом случае, в котором мы ее и будем доказывать. Пусть v — соответствующий собственный вектор, $f(v) = \lambda v$. Тогда $\overline{\lambda}(v, v) = (\lambda v, v) = (f(v), v) = (v, f^*(v)) = (v, f(v)) = (v, \lambda v) = \lambda(v, v)$. Поскольку $v \neq 0$, $\overline{\lambda} = \lambda$. $\square$

Лемма 4.1.10 *Пусть λ — собственное значение кососимметрического оператора. Тогда в евклидовом случае $\lambda = 0$, а в эрмитовом — λ чисто мнимое, $\lambda \in i\mathbb{R}$.*

Доказательство. Доказательство аналогично предыдущей лемме. В евклидовом случае $\lambda(v, v) = (\lambda v, v) = (f(v), v) = (v, f^*(v)) = -(v, f(v)) = -(v, \lambda v) = -\lambda(v, v)$, откуда $\lambda = -\lambda$. В эрмитовом случае $\overline{\lambda}(v, v) = (\lambda v, v) = (f(v), v) = (v, f^*(v)) = -(v, f(v)) = -(v, \lambda v) = -\lambda(v, v)$, откуда $\overline{\lambda} = -\lambda$. $\square$

Лемма 4.1.11 *Пусть λ_1, λ_2 — различные собственные значения самосопряженного или кососимметрического оператора, а v_1, v_2 — соответствующие собственные значения. Тогда $v_1 \perp v_2$.*

Доказательство. $\overline{\lambda}_1(v_1, v_2) = (\lambda_1 v_1, v_2) = (f(v_1), v_2) = (v_1, f^*(v_2)) = \pm(v_1, f(v_2)) = \pm(v_1, \lambda_2 v_2) = \pm\lambda_2(v_1, v_2)$, т.е. $(v_1, v_2)(\pm\lambda_2 - \overline{\lambda}_1) = 0$.
Рассмотрим второй сомножитель: $\pm\lambda_2 - \overline{\lambda}_1$. Если f самосопряжен, то собственные значения вещественны и знак — "плюс", т.е. $\lambda_1 - \lambda_2$; если f кососимметричен, то собственные значения число мнимы и знак — "минус", т.е. $-\lambda_1 + \lambda_2$. В обоих случаях это выражение отлично от нуля, следовательно, нулю равен первый сомножитель, $(v_1, v_2) = 0$. $\square$

Лемма 4.1.12 *Если L — инвариантное подпространство относительно самосопряженного или кососимметрического оператора f, то $L^\perp$ также будет инвариантно относительно f.*

Доказательство. Очевидно, т.к. $L^\perp$ инвариантно относительно $f^* = \pm f$. $\square$

4.2 Канонический вид матрицы самосопряженного оператора

Теорема 4.2.1 *Для любого самосопряженного оператора $f : V \to V$ 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. Проведем доказательство теоремы по индукции по размерности пространства.

1) Пусть $\dim V = 1$. 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 $\dim V \leqslant n$, докажем ее для $\dim V = n + 1$.

Сначала разберем случай эрмитова пространства. Пусть v — собственный вектор оператора f, $L = \langle v \rangle$, тогда $V = L \oplus L^{\perp}$, причем L и $L^{\perp}$ 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 $L^{\perp}$ 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 евклидова пространства. 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 $A_f = \left(\begin{array}{cc|c} a_{11} & a_{12} & \\ a_{21} & a_{22} & \mathbf{0} \\ \hline \multicolumn{2}{c|}{\mathbf{0}} & A_{f'} \end{array}\right)$. 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 4.2.2 *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 отвечающие м собственные векторы уже автоматически ортогональны друг другу.

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

4.3 Канонический вид матрицы кососимметрического оператора

Теорема 4.3.1 *У любого кососимметрического оператора $f : V \to V$ в эрмитовом пространстве V существует ортонормированный базис, в котором его матрица A_f имеет диагональный вид с чисто мнимыми числами на диагонали.*

Доказательство. Мы знаем, что если $L \subset V$ инвариантно относительно f, то и $L^{\perp}$ 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), $A_f = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$. □

Теорема 4.3.2 *У любого кососимметрического оператора $f : V \to V$ в евклидовом пространстве 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. Если у оператора f имеется одномерное инвариантное подпространство, то доказательство в точности совпадает с эрмитовым случаем, поэтому остается рассмотреть случай, когда можно выделить двумерное инвариантное подпространство $L \subset V$. При этом его ортогональное дополнение $L^{\perp}$ также инвариантно, и по предположению индукции можно считать, что матрица ограничения оператора f на $L^{\perp}$ уже имеет требуемый вид. Рассмотрим матрицу ограничения f на L в произвольном ортонормированном базисе двумерного подпространства L: это двумерная матрица, причем, в силу косой симметрии, на ее диагонали стоят нули, а вне диагонали — противоположные по знаку числа.

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

4.4 Изометрии

Определение 4.4.1 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 4.4.2 *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 эрмитово. Тогда $(a+b, a+b) = (a,a) + (b,b) + (a,b) + \overline{(a,b)} = (a,a) + (b,b) + 2\operatorname{Re}(a,b)$, откуда $\operatorname{Re}(a,b) = \frac{1}{2}((a+b,a+b)-(a,a)-(b,b))$, следовательно, вещественная часть скалярного произведения сохраняется. Взяв вектор ib вместо b, получаем

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

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

Лемма 4.4.3 *Следующие три утверждения эквивалентны:*
1) оператор f сохраняет скалярное произведение;
2) оператор f переводит некоторый ортонормированный базис в ортонормированный;
3) оператор f переводит произвольный ортонормированный базис в ортонормированный;
4) в любом ортонормированном базисе матрица A_f оператора f обладает свойством $\overline{A}_f^t A_f = E$.

Доказательство.
1) $\Rightarrow$ 3) 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.

Таким образом, первые три условия эквивалентны (3) $\Rightarrow$ 2) — очевидно). Покажем, что 3) и 4) эквивалентны.

3) $\Rightarrow$ 4). 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. $\overline{A}_f^t A_f = E$.

4) $\Rightarrow$ 3). 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. $\square$

4.5 Ортогональные и унитарные операторы

Определение 4.5.1 Оператор, сохраняющий скалярное произведение в евклидовых пространствах называется *ортогональным*, в эрмитовых пространствах — *унитарным.*

Лемма 4.5.2 *Пусть оператор f действует в евклидовом или эрмитовом пространстве V, а $L \subset V$ — инвариантное подпространство. Если f сохраняет скалярное произведение, то $L^\perp$ тоже инвариантно.*

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

Лемма 4.5.3 *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 $f(v) = \lambda v$. Т.к. оператор сохраняет скалярное произведение, то $(v, v) = (f(v), f(v)) = (\lambda v, \lambda v) = \overline{\lambda}\lambda(v, v) = |\lambda|^2(v, v)$, 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 $(v, v) = \lambda^2(v, v)$ с вещественным λ, следовательно, $\lambda = \pm 1$. $\square$

Лемма 4.5.4 *Если f — унитарный оператор, то его собственные вектора, отвечающие различным собственным значениям, взаимно ортогональны.*

Доказательство. Пусть $\lambda_1 \neq \lambda_2$, и $f(v_1) = \lambda_1 v_1$, $f(v_2) = \lambda_2 v_2$ (v_1 и v_2 — собственные вектора). Тогда $(v_1, v_2) = (f(v_1), f(v_2)) = (\lambda_1 v_1, \lambda_2 v_2) = \overline{\lambda}_1 \lambda_2 (v_1, v_2)$, 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. $\square$

Аналогичное утверждение верно и для ортогональных операторов, но у ортогонального оператора может быть не более двух различных собственных значений.

Теорема 4.5.5 *1) Если $f : V \to V$ — унитарный оператор, то существует ортонормированный базис, в котором его матрица A_f диагональна, причем на диагонали стоят числа, по модулю равные 1.*

2) Если $f : V \to V$ — 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 $L = \langle v \rangle$ — одномерное пространство, порожденное вектором v — будет инвариантным. Кроме того, его ортогональное дополнение $L^\perp$ также будет инвариантным по доказанной ранее лемме. Пользуясь этим замечанием, проведем теперь доказательство по индукции.

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

Пусть теорема верна для случая $\dim V = n$, докажем ее для $\dim V = n+1$. Возьмем одномерное инвариантное подпространство L, порожденное собственным вектором, тогда $V = L \oplus L^\perp$, и матрица A_f имеет вид $A_f = \left(\begin{array}{c|c} \lambda & \\ \hline & A' \end{array}\right)$, где A' — матрица оператора $f|_{L^\perp}$. Ограничение $f|_{L^\perp}$ оператора f на $L^\perp$ 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 собственные значения, то с ним можно поступить так же, как и в случае унитарного оператора. Если же их нет, то у оператора f найдется двумерное инвариантное подпространство L. По предположению индукции, для ограничения $f|_{L^\perp}$ существует ортонормированный базис в $L^\perp$, в котором матрица этого оператора имеет требуемый вид. Тогда матрица исходного оператора f будет блочно-диагональной, и все блоки, кроме первого (отвечающего подпространству L), имеют требуемый вид.

Ограничение оператора f на двумерное подпространство L также является ортогональным оператором. Выберем в L произвольный ортонормированный базис e_1, e_2. Поскольку длина вектора $f(e_1)$ должна быть равна 1, его координаты в базисе e_1, e_2 имеют вид $(\cos\varphi, \sin\varphi)$ для некоторого угла φ. Пусть (x, y) — координаты $f(e_2)$ в этом же базисе. Тогда $x^2 + y^2 = 1$ и $x\cos\varphi + y\sin\varphi = 0$, откуда получаются два решения: $x = -\sin\varphi$, $y = \cos\varphi$ и $x = \sin\varphi$, $y = -\cos\varphi$. Первое решение нам подходит — в этом случае двумерный блок — матрица ограничения f на L — имеет требуемый вид. Второе решение дает матрицу $\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$, котроая, как легко видеть, имеет вещественные собственные значения (ее харакеристический многочлен равен $\lambda^2 - 1$). Эту матрицу можно привести к диагональному виду с числами ± 1 на диагонали, что противоречит нашему предположению о том, что оператор f не имеет одномерных инвариантных подпространств.

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. $\square$

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.

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

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 взаимно-однозначное соответствие между квадратными матрицами и билинейными функциями, т.е. любая матрица задает какую-то функцию и разные матрицы задают разные функции.

Если в этом базисе векторы a, b имеют координаты $(a^1, \ldots, a^n)$ и $(b^1, \ldots, b^n)$ соответственно, то $g(a, b) = g_{ij} a^i b^j$, или, в матричной форме,

$$g(a,b) = \begin{pmatrix} a^1 & \ldots & a^n \end{pmatrix} \begin{pmatrix} g_{11} & \cdots & g_{1n} \\ \cdots & \cdots & \cdots \\ g_{n1} & \cdots & g_{nn} \end{pmatrix} \begin{pmatrix} b^1 \\ \ldots \\ b^n \end{pmatrix}.$$

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 $B(V) \cong Mat(n \times n)$.

Определение 5.1.2 *Рангом* билинейной 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, $\operatorname{rk} C^t G C = \operatorname{rk} G$.

Определение 5.1.3 *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 5.1.4 *Размерности левого и правого ядер совпадают и равны* $\dim G_L = \dim G_R = \dim V - \operatorname{rk} g$.

Доказательство. Если $a \in G_L$, то из равенства $g(a,b) = 0$ для любого вектора b следует, что

$$\begin{pmatrix} a^1 & \dots & a^n \end{pmatrix} \begin{pmatrix} g_{11} & \dots & g_{1n} \\ \dots & \dots & \dots \\ g_{n1} & \dots & g_{nn} \end{pmatrix} = 0,$$

т.е. координаты вектора a удовлетворяют однородной системе линейных уравнений с матрицей G^t. Поэтому размерность пространства решений этой системы совпадает с $\dim G_L$ и равна $n - \operatorname{rk} G$. Аналогично, правое ядро можно отождествить с множеством решений системы уравнений с матрицей G, и получить $\dim G_R = n - \operatorname{rk} G$. □

Определение 5.1.5 Билинейная функция g называется *невырожденной*, если $\dim G_L = \dim G_R = 0$ (это условие равносильно тому, что $\det G \neq 0$ или $\operatorname{rk} g = \dim V$, а также невырожденности матрицы G).

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.

5.2 Симметричные и кососимметричные 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 5.2.1 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 $g(b,a) = -g(a,b)$.

Утверждение 5.2.2 *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 5.2.3 Функция $f : V \to \mathbb{K}$ называется *квадратичной функцией (формой)*, если существует такая симметричная 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, $g(a,b) = \frac{1}{2}(f(a+b) - f(a) - f(b))$. Таким образом, симметричные билинейные функции и квадратичные функции находятся во взаимно однозначном соответствии.

Утверждение 5.2.4 *Любая билинейная функция допускает единственное разложение в сумму симметричной и кососимметричной билинейных функций.*

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

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 $g(b,a) = 0$ 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. $\square$

5.3 Ортогональное дополнение

Определение 5.3.1 Пусть $V \subset W$, а на W задана (косо)симметричная билинейная функция g, тогда $V^{\perp_g} = \{a \in W : g(a,b) = 0 \;\; \forall b \in V\}$.

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 $V_1^{\perp_g} = V_2$, $V_3^{\perp_g} = V_3$. 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 5.3.2 $\dim V^{\perp_g} \geqslant \dim W - \dim V$, *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 $a \in V^{\perp_g}$ 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

$$\left\{ \begin{array}{rcl} g(a, e_1) & = & 0 \\ \ldots\ldots & & \ldots \\ g(a, e_r) & = & 0. \end{array} \right.$$

Размерность пространства решений этой системы (т.е. пространства $V^{\perp_g}$) удовлетворяет неравенству $\dim V^{\perp_g} \geqslant n - r = \dim W - \dim V$ (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 $V \cap \operatorname{Ker} g = \{0\}$. $\square$

Лемма 5.3.3 $V \cap V^{\perp_g} = \operatorname{Ker} g_V$, *где* g_V *— ограничение* g *на* V.

Доказательство. Возьмем произвольный вектор $a \in V \cap V^{\perp_g}$, 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, $c \in V^{\perp_g}$, поэтому $c \in V \cap V^{\perp_g}$. □

Следствие 5.3.4 *Если* $\operatorname{Ker} g_V = \{0\}$ *(т.е. ограничение g на V невырождено), то* $W = V \oplus V^{\perp_g}$.

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

Лемма 5.3.5 *Если ограничения g на V и на $V^{\perp_g}$ невырождены, то* $(V^{\perp_g})^{\perp_g} = V$.

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

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

Лемма 5.3.6 *Пусть* $W = V \oplus V^{\perp_g}$, $e_1, \ldots, e_r$ *— базис в* V, $e_{r+1}, \ldots, e_n$ *— базис в* $V^{\perp_g}$. *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. □

5.4 Нормальный вид матрицы (косо)симметрической 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. 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 $g(a, b) = \alpha \neq 0$. При этом 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 5.4.1 *1) у любой симметричной билинейной функции на вещественном векторном пространстве существует базис, в котором ее матрица имеет диагональный вид с числами* 0, ± 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 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)-2).

Лемма 5.4.2 *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 $g(a,b) = 0$. Следовательно наша функция g тождественно нулевая, что противоречит предположению. $\square$

Далее будем действовать по индукции (база индукции $\dim W = 1$ уже проверена). Пусть утверждение теоремы верно для $\dim W < n$, 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 $W = V \oplus V^{\perp_g}$ и $\dim V^{\perp_g} = n-1$. По индукции в $V^{\perp_g}$ 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.

3) 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, $W = V \oplus V^{\perp_g}$, и $\dim V^{\perp_g} = n-2$. В $V^{\perp_g}$ по предположению индукции можно выбрать нужный базис. Добавив к нему векторы a и b, получим искомый базис. $\square$

Указанный вид матрицы называется нормальным. Для приведения матрицы симметричной билинейной 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 $g_{kl}x^k x^l$ и сделаем замену координат $\begin{cases} x^k &= \widetilde{x}^k + \widetilde{x}^l \\ x^l &= \widetilde{x}^k - \widetilde{x}^l \end{cases}$, тогда $x^k x^l = (\widetilde{x}^k)^2 - (\widetilde{x}^l)^2$.

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

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

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 $g(x, y) = x^1y^1 + \ldots + x^ky^k$.

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

5.5 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. Рассмотрим оставшийся случай — вещественной симметричной билинейной функции. Нормальный вид ее матрицы

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

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 5.5.1 Пара чисел (p, q) 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 5.5.2 (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: $e_1, \ldots, e_p, e_{p+1}, \ldots, e_{p+q}, e_{p+q+1}, \ldots, e_{p+q+s}$ и $\widetilde{e}_1, \ldots, \widetilde{e}_{p'}, \widetilde{e}_{p'+1}, \ldots, \widetilde{e}_{p'+q'}, \widetilde{e}_{p'+q'+1}, \ldots, \widetilde{e}_{p'+q'+s'}$, причем, как мы уже знаем, $s' = s$ и $p' + q' = p + q$. Рассмотрим в W подпространства

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

и

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

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 $x = x^1e_1 + \ldots + x^pe_p$, 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. $\square$

5.6 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 5.6.1 (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 $g(x,y) = |G_1|x^1y^1 + \frac{|G_2|}{|G_1|}x^2y^2 + \ldots + \frac{|G_r|}{|G_{r-1}|}x^ry^r$, где $|G_i|$ — i-й угловой минор.*

Доказательство. Будем искать новый базис в виде

$$\begin{array}{rcl} e_1' &=& e_1; \\ e_2' &=& e_2 + c_2^1 e_1'; \\ e_3' &=& e_3 + c_3^1 e_1' + c_3^2 e_2'; \\ \ldots && \ldots\ldots\ldots \\ e_n' &=& e_n + c_n^1 e_1' + \ldots + c_n^{n-1} e_{n-1}', \end{array}$$

где $e_1, \ldots, e_n$ — 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 $e_1, \ldots, e_k$ и векторов $e_1', \ldots, e_k'$ 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

$$\left\{ \begin{array}{rcl} c_k^1 g(e_1', e_1') + \ldots + c_k^{k-1} g(e_{k-1}', e_1') &=& -g(e_k, e_1') \\ \ldots\ldots\ldots\ldots\ldots\ldots && \ldots \\ c_k^1 g(e_1', e_{k-1}') + \ldots + c_k^{k-1} g(e_{k-1}', e_{k-1}') &=& -g(e_k, e_{k-1}') \end{array} \right. .$$

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

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

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 5.6.2 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 5.6.3 (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 $g(x,y) = \lambda_1 x^1 y^1 + \ldots + \lambda_n x^n y^n$ была невырожденной (т.е. чтобы все λ_i были положительными) необходимо, чтобы все $|G_i|$ были положительными. □

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

Определение 5.7.1 Вещественное пространство с заданной на нем невырожденной симметричной билинейной функцией называется *псевдоевклидовым* пространством.

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

Первый случай является естественным обобщением евклидового пространств. Скалярное (обобщенное) произведение задается симметричной билинейной функцией g, $(x,y) = g(x,y) = x^1y^1 + \ldots + x^py^p - x^{p+1}y^{p+1} - \ldots - x^ny^n$. Если в нормальном виде функции g все знаки — плюсы (т.е. если g положительно определена) то псевдоевклидово пространство является просто евклидовым. Матрица G функции g является обобщением матрицы Грама.

В псевдоевклидовом случае будем говорить, что базис $e_1,\ldots,e_n$ ортонормирован, если $g(e_i,e_j) = \begin{cases} 0 & \text{при } i \neq j \\ \pm 1 & \text{при } i = j \end{cases}$. Отметим, что в симплектическом случае такое невозможно, т.к. $g(x,x) = 0$ для любого вектора x.

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

Лемма 5.7.2 *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: $g(f_1f_2(x), f_1f_2(x)) = g(f_2(x), f_2(y)) = g(x,y)$. 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:

$$g(x,y) = g_{ij}x^iy^j = g(fx, fy) = g_{kl}a^k_ix^ia^l_jy^j = (A^tGA)_{ij}x^iy^j,$$

т.е. $G = A^tGA$. 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 $A^tA = E$, и это условие выделяет ортогональные (унитарные) матрицы.

Введем в рассмотрение следующие группы операторов:

- *псевдоортогональная группа* $O(p,q)$ — группа операторов, сохраняющих псевдоевклидовое скалярное произведение с сигнатурой (p,q) (если $q = 0$, то группу $O(p,0)$ называют ортогональной и обозначают $O(p)$).

- *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 $x^2 - y^2 = \pm 1$, 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.

5.8 Квадрики в аффинных пространствах

Напомним, что аффинное пространство это тройка, состоящая из множества точек $\mathcal{A}$, линейного пространства V и операции сложения точки и вектора (которая соответствует приложению начала вектора к данной точке и получению при этом новой точки — конца приложенного вектора). Для описания положения точки в аффинном пространстве нужно задать репер, который состоит из некоторой фиксированной точки O множества $\mathcal{A}$ и некоторого базиса пространства V. Произвольной точке $A \in \mathcal{A}$ можно сопоставить координаты (единственного) вектора v, удовлетворяющего равенству $\mathcal{O} + \sqsubseteq = \mathcal{A}$, в данном базисе. При переходе от одного репера к другому, координаты точки меняются по формуле

$$\begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \begin{pmatrix} c_1^1 & \dots & c_n^1 \\ \dots & \dots & \dots \\ c_1^n & \dots & c_n^n \end{pmatrix} \begin{pmatrix} \widetilde{x}^1 \\ \vdots \\ \widetilde{x}^n \end{pmatrix} + \begin{pmatrix} x_0^1 \\ \vdots \\ x_0^n \end{pmatrix}, \tag{4}$$

где $x^1, \dots, x^n$ и $\widetilde{x}^1, \dots, \widetilde{x}^n$ — координаты в старом и новом базисах, C — матрица перехода от старого к новому базису, а $(x_0^1, \dots, x_0^n)$ — координаты (в старом базисе) вектора w, удовлетворяющего равенству $O + w = \widetilde{O}$, где $\widetilde{O}$ — начальная точка нового репера.

Рассмотрим выражение $q(x) + l(x) + a$, где $q(x)$ — квадратичная функция, $q(x) = q_{ij}x^ix^j$, $l(x)$ — линейная функция, $l(x) = l_ix^i$, а a — число. Такое выражение называется *квадрикой*.

Мы знаем, что заменой базиса (приведением к нормальной форме) можно привести квадратичную функцию $q(x)$ к виду $q(x) = (x^1)^2 + \ldots + (x^p)^2 - (x^{p+1})^2 - \ldots - (x^{p+q})^2$. Если

$p+q=n$ (размерности пространства V), то с помощью умножения на скаляр можно привести квадрику к одному из видов:

$$(x^1)^2+\ldots+(x^p)^2-(x^{p+1})^2-\ldots-(x^{p+q})^2+1,$$

$$(x^1)^2+\ldots+(x^p)^2-(x^{p+1})^2-\ldots-(x^{p+q})^2-1,$$

$$(x^1)^2+\ldots+(x^p)^2-(x^{p+1})^2-\ldots-(x^{p+q})^2.$$

Если $p+q<n$, то линейную функцию можно привести к виду $l(x)=l_{p+q+1}x^{p+q+1}+\ldots+l_nx^n$. Если это выражение отлично от нуля, то преобразованием координат можно $l_{p+q+1}x^{p+q+1}+\ldots+l_nx^n+a$ преобразовать к виду x^{p+q+1}.

Подводя итог, получаем следующее утверждение:

Утверждение 5.8.1 *Квадрика приводится к одному из следующих видов:*

$$\begin{gathered}(x^1)^2+\ldots+(x^p)^2-(x^{p+1})^2-\ldots-(x^{p+q})^2+1,\\(x^1)^2+\ldots+(x^p)^2-(x^{p+1})^2-\ldots-(x^{p+q})^2-1,\\(x^1)^2+\ldots+(x^p)^2-(x^{p+1})^2-\ldots-(x^{p+q})^2,\\(x^1)^2+\ldots+(x^p)^2-(x^{p+1})^2-\ldots-(x^{p+q})^2+x^{p+q+1}.\end{gathered}$$

5.9 Симметрические билинейные функции на евклидовом пространстве

Приступим к рассмотрению симметрических билинейных функций в евклидовом пространстве. Пусть нам задано евклидово пространство V со скалярным произведением $(\cdot,\cdot)$.

Теорема 5.9.1 *Пусть g — симметричная билинейная функция на евклидовом пространстве V, тогда существует ортонормированный базис, в котором ее матрица диагональна.*

Доказательство. Пусть G — матрица билинейной функции g в каком-нибудь ортонормированном базисе. При переходе к другому базису матрица этой функции преобразуется по правилу $G'=C^tGC$, где C — матрица перехода. Если производится переход от одного ортонормированного базиса к другому ортонормированому, то матрица C ортогональна, следовательно, $C^t=C^{-1}$. Следовательно, матрица билинейной функции преобразуется по формуле $G'=C^{-1}GC$. Но по этой же формуле преобразуются матрицы линейных операторов, а для них уже известна теорема о приведении к диагональному виду. Именно, было доказано, что существует такая ортогональная матрица C, что матрица $C^{-1}GC$ диагональна. □

Лемма 5.9.2 *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: $\det(G'-\lambda E)=\det(C^tGC-\lambda C^tC)=\det(C^t(G-\lambda E)C)=\underbrace{(\det C)^2}_{=1}\det(G-\lambda E)=\det(G-\lambda E)$, и его корни, тем самым, не зависят от выбора ортонормированного базиса. □

Определение 5.9.3 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.

5.10 Пара симметричных билинейных функций

Теорема 5.10.1 *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 $(H' - \lambda E)CX = 0$, домножим его слева на C^t и получим $C^t(H' - \lambda E)CX = 0$, т.е. $(H - \lambda G)X = 0$.

Мы только что получили доказательство следующей леммы:

Лемма 5.10.2 *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.

6 Тензоры

6.1 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 $l \in V'$ Определим значение $\varphi_x(l) = l(x)$. 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 $l \in V'$ $\varphi_x(l) = 0$, т.е. $l(x) = 0$. Но единственный вектор в V, на котором любой функционал равен нулю, есть нулевой вектор, $x = 0$. Действительно, если это неверно, выберем базис $e_1 = x, e_2, \ldots, e_n$ 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 $e_1, \ldots, e_n$ в V векторы $\varphi_{e_1}, \ldots, \varphi_{e_n}$ линейно независимы. Но, поскольку $\dim V'' = n$, эти векторы составляют базис пространства V'', следовательно φ — изоморфизм.

При построении этого изоморфизма, мы 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.

Т.к. V и V'' изоморфны, и этот изоморфизм независим от выбора базиса, то мы можем эти два пространства отождествить, и смотреть на пространства V и V' как на двойственные друг к другу: элементы пространства V' есть линейные функции на V, а элементы пространства V можно считать линейными функциями на V'.

6.2 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

$$T : \underbrace{V \times \ldots \times V}_{p} \times \underbrace{V' \times \ldots \times V'}_{q} \to \mathbb{K},$$

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:

$$\begin{aligned}
T(v_1, \ldots, v_i' + v_i'', \ldots, v_p, l^1, \ldots, l^q) &= T(v_1, \ldots, v_i', \ldots, v_p, l^1, \ldots, l^q) + \\
&+ T(v_1, \ldots, v_i'', \ldots, v_p, l^1, \ldots, l^q), \\
T(v_1, \ldots, \lambda v_i, \ldots, v_p, l^1, \ldots, l^q) &= \lambda T(v_1, \ldots, v_i, \ldots, v_p, l^1, \ldots, l^q), \\
T(v_1, \ldots, v_p, l^1, \ldots, l'^j + l''^j, \ldots, l^q) &= T(v_1, \ldots, v_p, l^1, \ldots, l'^j, \ldots, l^q) + \\
&+ T(v_1, \ldots, v_p, l^1, \ldots, l''^j, \ldots, l^q), \\
T(v_1, \ldots, v_p, l^1, \ldots, \lambda l^j, \ldots, l^q) &= \lambda T(v_1, \ldots, v_p, l^1, \ldots, l^j, \ldots, l^q).
\end{aligned}$$

Определение 6.2.1 *Тензором* называется полилинейная функция. Тензор типа (p, q) — это полилинейная функция от p векторов из V и от q линейных функций из V'. Тензорами типа $(0, 0)$ называют скаляры.

Примеры:
1) скаляры — это тензоры типа $(0, 0)$ по определению;
2) векторы — это тензоры типа $(0, 1)$, т.к. любой вектор задает линейную функцию на V' переводящую элемент $l \in V'$ в $l(v) \in \mathbb{K}$ (ранее мы отождествили V'' и V);

3) линейные функции — это тензоры типа $(1,0)$, т.к. они задают отображение $l : V \to \mathbb{K}$, переводя вектор $v \in V$ в элемент $l(v) \in \mathbb{K}$;

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.

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:

$$\begin{aligned}
(T+S)(v_1,\ldots,v_p,l^1,\ldots,l^q) &:= T(v_1,\ldots,v_p,l^1,\ldots,l^q) + \\
&+ \ S(v_1,\ldots,v_p,l^1,\ldots,l^q); \\
(\lambda T)(v_1,\ldots,v_p,l^1,\ldots,l^q) &:= \lambda T(v_1,\ldots,v_p,l^1,\ldots,l^q).
\end{aligned}$$

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

$$\begin{aligned}
&(T\otimes S)(v_1,\ldots,v_p,v_{p+1},\ldots,v_{p+r},l^1,\ldots,l^q,l^{q+1},\ldots,l^{q+t}) := \\
&:= T(v_1,\ldots,v_p,l^1,\ldots,l^q)\cdot S(v_{p+1},\ldots,v_{p+r},l^{q+1},\ldots,l^{q+t}).
\end{aligned}$$

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 $l \in V'$ — тензор $v \otimes l \in \Theta_1^1$. Посмотрим, как этот тензор действует на своих аргументах. Возьмем произвольный вектор $a \in V$ и функцию $h \in V'$, получим, что $(v\otimes l)(a,h) = h(v)\cdot l(a)$.

2) возьмем две линейные функции $g, h \in V'$, тогда $g \otimes h$ будет тензором типа $(2,0)$ (т.е. билинейной функцией), $(g\otimes h)(v_1,v_2) = g(v_1)\cdot h(v_2)$.

6.3 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.

$$\begin{aligned}
&T(v_1,\ldots,v_p,l^1,\ldots,l^q) = T(v_1^{i_1}e_{i_1},\ldots,v_p^{i_p}e_{i_p},l^1_{j_1}\varepsilon^{j_1},\ldots,l^q_{j_q}\varepsilon^{j_q}) = \\
&= v_1^{i_1}\cdot\ldots\cdot v_p^{i_p}\cdot l^1_{j_1}\cdot\ldots\cdot l^q_{j_q}\cdot T(e_{i_1},\ldots,e_{i_p},\varepsilon^{j_1},\ldots,\varepsilon^{j_q}),
\end{aligned}$$

где $v_k^{i_k}$ — координаты вектора v_k, а $l^m_{j_j}$ — координаты линейной функции l^m.

Обозначим $T(e_{i_1},\ldots,e_{i_p},\varepsilon^{j_1},\ldots,\varepsilon^{j_q}) = T^{j_1,\ldots,j_q}_{i_1,\ldots,i_p} \in \mathbb{K}$. Зафиксировав базис, мы можем поставить тензору T в соответствие набор чисел $T^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}$ — 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

$$\begin{aligned}
\widetilde{T}^{l_1,\ldots,l_q}_{k_1,\ldots,k_p} &= T(\widetilde{e}_{k_1},\ldots,\widetilde{e}_{k_p},\widetilde{\varepsilon}^{l_1},\ldots,\widetilde{\varepsilon}^{l_q}) = \\
&= T(c^{i_1}_{k_1}e_{i_1},\ldots,c^{i_p}_{k_p}e_{i_p},d^{l_1}_{j_1}\varepsilon^{j_1},\ldots,d^{l_q}_{j_q}\varepsilon^{j_q}) = \\
&= T(e_{i_1},\ldots,e_{i_p},\varepsilon^{j_1},\ldots,\varepsilon^{j_q})\cdot c^{i_1}_{k_1}\cdot\ldots\cdot c^{i_p}_{k_p}\cdot d^{l_1}_{j_1}\cdot\ldots\cdot d^{l_q}_{j_q} = \\
&= T^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}\cdot c^{i_1}_{k_1}\cdot\ldots\cdot c^{i_p}_{k_p}\cdot d^{l_1}_{j_1}\cdot\ldots\cdot d^{l_q}_{j_q}.
\end{aligned}$$

Примеры:

1) если x — вектор, то $\widetilde{x}^k = d_i^k x^i$. Такой закон изменения координат называется *векторным* законом.

2) если l — линейная функция, то $\widetilde{l}_k = c_k^i l_i$. Такой закон изменения координат называется *ковекторным* законом. А величины, изменяющиеся по ковекторному закону называются *ковекторами.* 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 тензора: тензор — это сопоставление каждому базису пространства V набора 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 $T(e_{i_1},\ldots,e_{i_p},\varepsilon^{j_1},\ldots,\varepsilon^{j_q}) = T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}$.

Теперь мы можем отождествить линейные операторы и тензоры типа $(1,1)$. И те, и другие ведут себя одинаково при заменах координат, именно, матрица линейного оператора при переходе от одного базиса к другому ведет себя по тому же закону, что и набор чисел, задающий тензор типа $(1,1)$.

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 6.3.1 *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 $\widetilde{a}^{kl}\widetilde{a}_{lm} = \delta_m^k$. Подставим сюда $\widetilde{a}_{lm} = a_{ji}c_l^j c_m^i$, где $C = (c_m^i)$ — матрица перехода, тогда $\widetilde{a}^{kl}a_{ji}c_l^j c_m^i = \delta_m^k$. Умножим (и просуммируем по повторяющимся индексам) обе части этого равенства на элементы обратной матрицы к матрице перехода (напомним, что $c_i^k d_k^j = \delta_i^j$): $\widetilde{a}^{kl}a_{ji}c_l^j c_m^i d_p^m = \delta_m^k d_p^m$, т.е. $\widetilde{a}^{kl}a_{ji}c_l^j \delta_p^i = d_p^k$, или $\widetilde{a}^{kl}a_{jp}c_l^j = d_p^k$. Воспользуемся теперь обратимостью матрицы (a_{ij}), т.е. тем, что $a_{jp}a^{pr} = \delta_j^r$. Умножив обе части предпоследнего равенства на a^{pr} (и просуммировав), получим $\widetilde{a}^{kl}a_{jp}a^{pr}c_l^j = a^{pr}d_p^k$, или $\widetilde{a}^{kl}c_l^r = a^{pr}d_p^k$. Еще раз умножив обе части равенства на d_r^n просуммировав по r и воспользовавшись тем, что $c_l^r d_r^n = \delta_l^n$ и $\widetilde{a}^{kl}\delta_l^n = \widetilde{a}^{kn}$, получим $\widetilde{a}^{kn} = a^{pr}d_p^k d_r^n$. $\square$

6.4 Базис в пространстве тензоров

Построим базис в пространстве тензоров Θ_p^q. Напомним, что векторы и ковекторы отождествляются с тензорами 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_p^q. Прежде, чем доказывать это, вычислим значение такого тензора (произведения) на наборах аргументов из базисных векторов:

$$\begin{gathered}\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}(e_{k_1},\ldots,e_{k_p},\varepsilon^{l_1},\ldots,\varepsilon^{l_q}) = \\ = \varepsilon^{i_1}(e_{k_1})\cdot\ldots\cdot\varepsilon^{i_p}(e_{k_p})\cdot\varepsilon^{l_1}(e_{j_1})\cdot\ldots\cdot\varepsilon^{l_1}(e_{j_q}) = \\ = \delta_{k_1}^{i_1}\cdot\ldots\cdot\delta_{k_p}^{i_p}\cdot\delta_{j_1}^{l_1}\cdot\ldots\cdot\delta_{j_q}^{l_q} = \\ = \begin{cases} 1, & \text{если } i_1 = k_1,\ldots,i_p = k_p, j_1 = l_1,\ldots,j_q = l_q, \\ 0, & \text{в остальных случаях.}\end{cases}\end{gathered}$$

Лемма 6.4.1 *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_p^q.*

Доказательство. Сначала докажем линейную независимость этих произведений. Пусть существуют такие числа $\lambda_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}$, что линейная комбинация $\lambda_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}\varepsilon^{i_1} \otimes \ldots \otimes \varepsilon^{i_p} \otimes e_{j_1} \otimes \ldots \otimes e_{j_q}$. Применим этот тензор, как полилинейную функцию, к аргументам $e_{k_1}, \ldots, e_{k_p}, \varepsilon^{l_1}, \ldots, \varepsilon^{l_q}$ и получим

$$\lambda_{i_1,\ldots,i_p}^{j_1,\ldots,j_q} \underbrace{\varepsilon^{i_1} \otimes \ldots \otimes \varepsilon^{i_p} \otimes e_{j_1} \otimes \ldots \otimes e_{j_q}(e_{k_1}, \ldots, e_{k_p}, \varepsilon^{l_1}, \ldots, \varepsilon^{l_q})} = 0 \tag{5}$$

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 (5) может быть записано в виде

$$\lambda_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}\varepsilon^{i_1} \otimes \ldots \otimes \varepsilon^{i_p} \otimes e_{j_1} \otimes \ldots \otimes e_{j_q}(e_{k_1}, \ldots, e_{k_p}, \varepsilon^{l_1}, \ldots, \varepsilon^{l_q}) = \lambda_{k_1,\ldots,k_p}^{l_1,\ldots,l_q} = 0.$$

Но, т.к. это равенство имеет место для любого набора индексов $k_1, \ldots, k_p, l_1, \ldots, l_q$, то все $\lambda_{k_1,\ldots,k_p}^{l_1,\ldots,l_q}$ равны нулю. Линейная независимость доказана.

Удостоверимся теперь, что любой тензор можно представить в виде линейной комбинации этих базисных тензоров. Для этого достаточно доказать равенство

$$T = T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}\varepsilon^{i_1} \otimes \ldots \otimes \varepsilon^{i_p} \otimes e_{j_1} \otimes \ldots \otimes e_{j_q}. \tag{6}$$

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 $T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}$, а правая часть, как мы уже видели раньше, также равна $T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}\varepsilon^{i_1} \otimes \ldots \otimes \varepsilon^{i_p} \otimes e_{j_1} \otimes \ldots \otimes e_{j_q}(e_{k_1}, \ldots, e_{k_p}, \varepsilon^{l_1}, \ldots, \varepsilon^{l_q}) = T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}$. Итак, равенство проверено, т.е. для произвольного тензора T мы нашли его разложение в линейную комбинацию (6), причем числа $T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}$ являются координатами этого тензора в указанном базисе. □

6.5 Свертка тензоров

Пусть $T \in \Theta_p^q$ — 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

$$(sT)(v_2, \ldots, v_p, l^2, \ldots, l^q) := T(e_i, v_2, \ldots, v_p, \varepsilon^i, l^2, \ldots, l^q),$$

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 $(sT)_{i,i_2,\ldots,i_p}^{i,j_2,\ldots,j_q}$ (здесь опять производится суммирование по индексу i). При переходе к другому базису имеем

$$\widetilde{T}_{k_1,\ldots,k_p}^{l_1,\ldots,l_q} = T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q} \cdot c_{k_1}^{i_1} \cdot \ldots \cdot c_{k_p}^{i_p} \cdot d_{j_1}^{l_1} \cdot \ldots \cdot d_{j_q}^{l_q};$$

$$(\widetilde{sT})_{k_2,\ldots,k_p}^{l_2,\ldots,l_q} = \widetilde{T}_{k,k_2,\ldots,k_p}^{k,l_2,\ldots,l_q} = T_{i_1,i_2,\ldots,i_p}^{j_1,j_2,\ldots,j_q} \cdot \underbrace{c_k^{i_1}} \cdot c_{k_2}^{i_2} \cdot \ldots \cdot c_{k_p}^{i_p} \cdot \underbrace{d_{j_1}^k} \cdot d_{j_2}^{l_2} \cdot \ldots \cdot d_{j_q}^{l_q}.$$

Произведение подчеркнутых элементов равно $c_k^{i_1} d_{j_1}^k = \delta_{j_1}^{i_1}$, т.к. $CD = E$, поэтому ненулевыу слагаемые отвечают значениям $i_1 = j_1$. Обозначим $i_1 = j_1 = i$, тогда

$$\widetilde{T}_{k_1,\ldots,k_p}^{l_1,\ldots,l_q} = T_{i,i_2,\ldots,i_p}^{i,j_2,\ldots,j_q} \cdot c_{k_2}^{i_2} \cdot \ldots \cdot c_{k_p}^{i_p} \cdot d_{j_2}^{l_2} \cdot \ldots \cdot d_{j_q}^{l_q} = (sT)_{i_2,\ldots,i_p}^{j_2,\ldots,j_q} \cdot c_{k_2}^{i_2} \cdot \ldots \cdot c_{k_p}^{i_p} \cdot d_{j_2}^{l_2} \cdot \ldots \cdot d_{j_q}^{l_q}.$$

Т.е. координаты $(sT)_{i_2,\ldots,i_p}^{j_2,\ldots,j_q}$ 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_i^j), сверткой этого тензора будет число f_i^i (сумма диагональных элементов), т.е. в данном случае свертка — это след, $sf = \operatorname{tr} f$.

2) (частный случай предыдущего примера). Возьмем линейную функцию l, т.е. тензор типа $(1,0)$, и вектор a — тензор типа $(0,1)$. Тогда $l\otimes a$ будет тензором типа $(1,1)$ (контрольный вопрос: с каким линейным оператором отождествляется этот тензор?). Координаты этого тензора $(l\otimes a)_i^j = l_i a^j$, где l_i — набор координат линейной функции l. Тогда при полной свертке этого тензора получим $s(l\otimes a) = l_i a^i = l(a)$.

3) 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 имеется выделенный тензор типа $(2,0)$, отвечающий биоинейной функции, которая задает скалярное произведение. Это позволяет у любого тензора заменить векторный аргумент на ковекторный и наоборот. В координатах это выглядит следующим образом. Пусть $G = (g_{ij})$ — матрица Грама скалярного произведения $(\cdot,\cdot)$. Тензору T типа $(0,1)$ 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 $T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}$ ставит в соответствие тензор $g_{ij}\cdot T_{i_1,\ldots,i_p}^{j,j_2,\ldots,j_q}$. 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 $T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}$ операция поднятия индекса переводит в тензор $g^{ij}\cdot T_{i,i_2,\ldots,i_p}^{j_1,\ldots,j_q} \in \Theta_{p-1}^{q+1}$.

6.7 Оператор альтернирования. Кососимметрические тензоры

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. $T = T(v_1,\ldots,v_p)$; определим $f_\sigma(T) = \sigma T$, где $(\sigma T)(v_1,\ldots,v_p) := T(v_{\sigma(1)},\ldots,v_{\sigma(p)})$. 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 $(\sigma T)_{i_1,\ldots,i_p} = (\sigma T)(e_{i_1},\ldots,e_{i_p}) = T(e_{i_{\sigma(1)}},\ldots,e_{i_{\sigma(p)}}) = T_{i_{\sigma(1)}},\ldots,{i_{\sigma(p)}}$.

Тензор T называется *симметрическим*, если $\sigma T = T$ для любой перестановки $\sigma \in S_p$.

Тензор T называется *кососимметрическим*, если $\sigma T = (-1)^\sigma T$ для любой перестановки $\sigma \in S_p$.

Построим оператор *альтернирования* (приводящий к 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. 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 $T \in \Theta_p^0$, 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 $T \in \Theta_p^0$. 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 кососимметрических тензоров — *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 кососимметрических тензоров $T \in \Lambda_p$, $S \in \Lambda_q$, $R \in \Lambda_r$*

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

$$\begin{aligned} \sigma(S \otimes T)_{i_1,\ldots,i_{p+q}} &= (S \otimes T)_{i_{\sigma(1)},\ldots,i_{\sigma(p+q)}} = \\ &= S_{i_{\sigma(1)},\ldots,i_{\sigma(q)}} \cdot T_{i_{\sigma(q+1)},\ldots,i_{\sigma(p+q)}}; \qquad (7) \\ \sigma(T \otimes S)_{i_1,\ldots,i_{p+q}} &= (T \otimes S)_{i_{\sigma(1)},\ldots,i_{\sigma(p+q)}} = \\ &= T_{i_{\sigma(1)},\ldots,i_{\sigma(p)}} \cdot S_{i_{\sigma(p+1)},\ldots,i_{\sigma(p+q)}}. \qquad (8) \end{aligned}$$

Посмотрим, чем отличаются индексы у S и T в выражениях (7) и (8). Индексы в (7) — это $\sigma(1),\ldots,\sigma(q),\sigma(q+1),\ldots,\sigma(p+q)$, а индексы в (8) — это $\sigma(p+1),\ldots,\sigma(p+q),\sigma(1),\sigma(q)$. Пусть τ — перестановка

$$\begin{pmatrix} p+1 & \ldots & p+q & 1 & \ldots & p \\ 1 & \ldots & q & q+1 & \ldots & p+q \end{pmatrix}.$$

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. Легко видеть, что это число равно произведению pq, т.е. $(-1)^{\tau} = (-1)^{pq}$, что и требовалось показать.

3) Введем дополнительно обозначение $T_1 \wedge T_2 \wedge \ldots \wedge T_k := \operatorname{Alt}(T_1 \otimes T_2 \otimes \ldots \otimes T_k)$. Для доказательство ассоциативности нам также понадобится следующее равенство.

Утверждение 6.8.2 $\operatorname{Alt}((\operatorname{Alt} Q) \otimes R) = \operatorname{Alt}(Q \otimes R) = \operatorname{Alt}(Q \otimes (\operatorname{Alt} R))$ *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

$$\begin{aligned} \operatorname{Alt}((\operatorname{Alt} Q) \otimes R) &= \operatorname{Alt}\Big(\Big(\frac{1}{p!} \sum_{\sigma \in S_p} (-1)^{\sigma} \sigma Q\Big) \otimes R\Big) = \\ &= \frac{1}{p!} \sum_{\sigma \in S_p} (-1)^{\sigma} \operatorname{Alt}(\sigma Q \otimes R). \end{aligned}$$

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.

$$\widetilde{\sigma} = \begin{pmatrix} 1 & \ldots & p & p+1 & \ldots & p+q \\ \sigma(1) & \ldots & \sigma(p) & p+1 & \ldots & p+q \end{pmatrix}.$$

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. $(T \wedge S) \wedge R = T \wedge (S \wedge R)$. □

Содержание