Линейное (векторное) пространство

Определение 1. Множество V называется *линейным (векторным) пространством* над некоторым полем $\mathbb{K}$, если заданы операция $+$ сложения двух элементов множества V и операция умножения $\cdot$ элементов множества V на элементы поля $\mathbb{K}$, которые удовлетворяют следующим условиям (аксиомам):

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

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

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

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

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

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

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

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

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

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

Свойства линейного пространства:

(i) нулевой элемент в множестве V определен однозначно,

(ii) для любого элемента обратный элемент определен однозначно,

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

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

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

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

Доказательство.

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

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

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

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

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

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

$\square$

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 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 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 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. $\square$

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.

Аффинное пространство

Определение 5. 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 соответсвующей 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.

Определение 6. 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 7. *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 8. $a_1 + V = a_2 + V \Longleftrightarrow a_1 - a_2 \in V$.

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 $v + b' \in V$. □

Определение 9. *Фактор-пространством* W/V линейного пространства W по, подпространству V называется множество $\{a+V : a \in W\}$ — множество всех линейных подмногообразий пространства W, заданных подпространством V, с определенными на нем операциями сложения, $(a+V)+(b+V) := (a+b)+V$, и умножения на скаляры, $\lambda(a+V) := \lambda a + V$.

Лемма 10. *Факторпространство линейного пространства само является линейным пространством.*

Доказательство.
1) Сначала докажем, что введенные нами в факторпространстве операции корректны, т.е. какие бы мы ни брали элементы подмногообразий в качестве a и b, мы получим одно и то же подмногообразие. Докажем это для сложения (для умножения доказывается аналогично): если

$$a_1 + V = a_2 + V, \quad b_1 + V = b_2 + V,$$

то

$$a_1 - a_2 = v \in V, \quad b_1 - b_2 = w \in V,$$

поэтому

$$(a_1+b_1)+V = (a_2+b_2+v+w)+V = (a_2+b_2)+(v+w)+V = (a_2+b_2)+V,$$

т.е. сложение определено корректно.
2) Проверим свойства линейного пространства. Ввиду простоты ограничимся проверкой одного из восьми свойств, например 5).

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

□

Линейная зависимость векторов

Определение 11. 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 ($I = \{1, \dots, n\}$), 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 12. *Линейной оболочкой* системы векторов линейного пространства называется множество всех векторов, являющихся их линейной комбинацией.

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

Определение 13. 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 14. *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 15. *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 16. *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). □

Размерность

Определение 17. 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 16 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 16 дает противоречие.

Определение 18. *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 19. Линейно независимая система векторов в пространстве V называется *максимальной*, если при добавлении любого другого вектора система векторов становится линейно зависимой.

Следствие 20. *В любом конечномерном пространстве существует максимальная (линейно независимая) система векторов.*

Определение 21. Максимальная система векторов называется *базисом* пространства.

Следствие 22. *Если векторы $v_1, \ldots, v_n$ составляют базис линейного пространства V, то любой вектор $v \in V$ можно представить в виде линейной комбинации этих векторов единственным образом.*

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 23. *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. $\square$

Лемма 24. *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.

Теорема 25. *Пусть V — подпространство векторного пространства W, тогда*

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

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

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

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

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

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

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

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

Пересечение и сумма подпространств

Лемма 26. *Пусть даны два линейных подпространства V_1 и V_2 пространства W, тогда $V_1 \cap V_2$ также является линейным подпространством.*

Доказательство. Для доказательства 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 $x + y, \lambda x \in V_2$, следовательно, $x + y, \lambda x \in V_1 \cap V_2$. $\square$

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 27. *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. $V_1 + V_2 = \langle V_1 \cup V_2 \rangle$ (линейная оболочка объединения множеств V_1 и V_2).

Лемма 28. *Для любых двух подпространств V_1 и V_2 их сумма $V_1 + V_2$ также будет линейным пространством.*

Доказательство. Возьмем произвольные векторы $a, b \in V_1 + V_2$,

$$a = a_1 + a_2, \quad b = b_1 + b_2, \qquad a_1, b_1 \in V_1, \quad a_2, b_2 \in V_2.$$

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 λa лежит в $V_1 + V_2$. Очевидно, что все условия определения линейного пространства будут выполнены, следовательно $V_1 + V_2$ является линейным пространством. $\square$

Теорема 29. $\dim V_1 + \dim V_2 = \dim(V_1 + V_2) + \dim(V_1 \cap V_2)$.

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 каждом из двух подпространств, V_1 и V_2, этот базис можно дополнить до базиса.

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, $\dim V_2 = r+q$. 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

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

следовательно, $v \in V_1 \cap V_2$, и его можно разложить по базису подпространства $V_1 \cap V_2$, $v = \alpha_1 e_1 + \ldots + \alpha_r e_r$, тогда

$$0 = v - v = \alpha_1 e_1 + \ldots + \alpha_r e_r + \lambda_{r+p+1} e_{r+p+1} + \ldots + \lambda_{r+p+q} e_{r+p+q},$$

следовательно

$$\lambda_{r+p+1} = \ldots = \lambda_{r+p+q} = 0 \quad \text{и} \quad \alpha_1 = \ldots = \alpha_r = 0,$$

т.к. $e_1, \ldots, e_r, e_{r+p+1}, \ldots, e_{r+p+q}$ линейно независимы (это базис V_2). 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 $e_1, \ldots, e_r, e_{r+1}, \ldots, e_{r+p}$ (это базис V_1) 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 в соответствующих подпространствах,

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

тогда

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

т.е. любой вектор $a \in V_1 + V_2$ есть линейная комбинация системы векторов $e_1, \ldots, e_r, e_{r+1}, \ldots, e_{r+p}, e_{r+p+1}, \ldots, e_{r+p+q}$, следовательно, эта система векторов является базисом в $V_1 + V_2$, значит, $\dim(V_1 + V_2) = r + p + q$, откуда следует утверждение теоремы. □

Прямая сумма подпространств. Внешняя прямая сумма

Определение 30. Сумма подпространств $V_1 + V_2$ называется *прямой суммой* (обозначение $V_1 \oplus V_2$), если $V_1 \cap V_2 = \{0\}$.

Следствие 31. *$\dim(V_1 \oplus V_2) = \dim V_1 + \dim V_2$.*

Доказательство. Утверждение следствия очевидно вытекает из предыдущей теоремы. □

Лемма 32. *Следующие утверждения эквиваленты:*

(1) сумма $V_1 + V_2$ прямая;

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

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

(4) 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 теоремы о размерностях суммы и пересечения.

(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): 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 $0 = v + (-v)$ — противоречие с 4).

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. (4) — частный случай (3), что верно для любого вектора, верно и для нулевого вектора. □

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 33. *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 $V_1 \oplus V_2$. Такое отождествление позволяет отождествить $\mathbb{K} \oplus \cdots \oplus \mathbb{K}$ с $\mathbb{K}^n$.

Координаты

Определение 34. 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.

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

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 суммы $x = x^1 e_1 + \ldots + x^n e_n$, или чуть более короткой $\sum_{i=1}^n x^i e_i$, мы часто будем писать $x^i e_i$, на самом деле подразумевая сумму (но не записывая знак суммирования). Координаты векторов мы часто будем записывать в виде столбцов, т.е. $x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$. Также обычно верхний индекс будет соответствовать столбцам, а нижний — строкам.

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 35. *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). $\square$

Изоморфизмы векторных пространств

Определение 36. 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 (хотя бы один) изоморфизм.

Лемма 37. *Если $f : V \to W$ — изоморфизм, то обратное отображение $f^{-1} : W \to V$ также будет изоморфизмом.*

Доказательство. Отметим сначала, что отображение f^{-1} существует, 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. □

Изоморфность является отношением эквивалентности, т.е. обладает свойствами
• симметричности (если V изоморфно W, то W изоморфно V),
• рефлексивности (любое пространство изоморфно самому себе) и
• транзитивности (если V изоморфно W и W изоморфно U, то V изоморфно U).
Первые два свойства мы проверили, третье легко проверяется.

Лемма 38. *Если $\dim V = n$, то V изоморфно пространству $\mathbb{K}^n$ столбцов (строк) из n элементов.*

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

Следствие 39. *Если $\dim V = \dim W$, то $V \cong W$.*

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

Верно и обратное:

Лемма 40. *Если $V \cong W$, то $\dim V = \dim W$.*

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

Лемма 41. *Пусть $\dim V = \dim W$, а отображение $f : V \to W$ удовлетворяет условиям линейности:*
1) $f(v_1 + v_2) = f(v_1) + f(v_2)\ \forall v_1, v_2 \in V$,
2) $f(\lambda v) = \lambda f(v)\ \forall v \in V, \lambda \in \mathbb{K}$.
Пусть $e_1, \ldots, e_n$ — базис в V. Тогда f является изоморфизмом тогда и только тогда, когда $f(e_1), \ldots, f(e_n)$ — базис в W.

Доказательство. Если f — изоморфизм, то из равенства $\lambda_1 f(e_1) + \ldots + \lambda_n f(e_n) = 0$ следует $f(\lambda_1 e_1 + \ldots + \lambda_n e_n) = 0$, откуда заключаем, что $\lambda_1 e_1 + \ldots + \lambda_n e_n = 0$, значит, все $\lambda_i = 0$.

Обратно, пусть $f(e_1), \ldots, f(e_n)$ — базис в W. Для проверки взаимной однозначности отображения f достаточно проверить, что отображение f^{-1} корректно определено. Пусть $w \in W$ имеет разложение по базису $w = w^1 f(e_1) + \ldots + w^n f(e_n)$. Тогда определим отображение $g : W \to V$ равенством $g(w) = w^1 e_1 + \ldots + w^n e_n$. Очевидная проверка показывает, что $g = f^{-1}$. $\square$

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 42. $V \cong V'$.

Доказательство. Пусть $e_1, \ldots, e_n$ — базис в V. Определим функционалы $\varepsilon^i \in V'$, $i = 1, \ldots, n$, равенствами $\varepsilon^i(e_j) = \delta^i_j$ (δ^i_j — символ Кронекера, т.е. $\delta^i_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$. Поскольку $f(x^i e_i) = x^i f(e_i)$, то значение функционала на произвольном векторе полностью определяется значениями функционала на базисных векторах и координатами этого вектора, т.е. функционалы ε^i полностью заданы нашими условиями.

Докажем, что $\varepsilon^1, \ldots, \varepsilon^n$ будет базисом в V'. Для этого нужно доказать линейную независимость и максимальность.

1) линейная независимость:

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:

надо показать, 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. $\square$

Базис $\varepsilon^1, \ldots, \varepsilon^n$ называется *двойственным базисом* к базису $e_1, \ldots, e_n$.

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 $(\widetilde{e}_1 \dots \widetilde{e}_n) = (e_1 \dots e_n)C$. Отсюда получаем, что $\widetilde{f}_k = f_i c^i_k$. Следовательно $(\widetilde{f}_1 \dots \widetilde{f}_n) = (f_1 \dots f_n)C$, или (после транспонирования, обозначенного индексом t) $\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} = (C^{-1})^t \begin{pmatrix} \widetilde{f}_1 \\ \vdots \\ \widetilde{f}_n \end{pmatrix}$.

Сравнивая эту формулу с формулой изменения координат при замене базиса в линейном пространстве, получаем, что матрицей перехода от базиса $\varepsilon^1, \dots, \varepsilon^n$ к базису $\widetilde{\varepsilon}^1, \dots, \widetilde{\varepsilon}^n$ является матрица $(C^{-1})^t$, т.е. эти базисы связаны равенством $(\widetilde{\varepsilon}^1 \dots \widetilde{\varepsilon}^n) = (\varepsilon^1 \dots \varepsilon^n)(C^{-1})^t$.

Итак, мы доказали следующее утверждение:

Лемма 43. *Пусть C — матрица перехода от базиса $e_1, \dots, e_n$ к базису $\widetilde{e}_1, \dots, \widetilde{e}_n$ в линейном пространстве V. Тогда двойственные базисы $\varepsilon^1, \dots, \varepsilon^n$ и $\widetilde{\varepsilon}^1, \dots, \widetilde{\varepsilon}^n$ к этим базисам связаны равенством $(\widetilde{\varepsilon}^1 \dots \widetilde{\varepsilon}^n) = (\varepsilon^1 \dots \varepsilon^n)(C^{-1})^t$; а координаты функционала f в этих двойственных базисах связаны равенством $(\widetilde{f}_1 \dots \widetilde{f}_n) = (f_1 \dots f_n)C$.*

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 $\widetilde{\varepsilon} = \lambda^{-1}\varepsilon$. 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!

Канонический изоморфизм между пространством и его вторым двойственным

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

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

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

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

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

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

Пример: двойственное пространство пространства многочленов

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

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

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

Лемма 44. *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 $\mathbb{K}_n[x]'$. $\square$

Евклидовы и унитарные пространства

Определение 45. Линейное пространство V над полем $\mathbb{R}$ называется *евклидовым*, если на нем определена функция $f : V \times V \to \mathbb{R}$ двух аргументов (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 46. Линейное пространство V над полем $\mathbb{C}$ называется *эрмитовым*, если на нем определена функция $f: V \times V \to \mathbb{C}$ двух аргументов (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 $V = \mathbb{R}[t]$ — пространство многочленов над полем $\mathbb{R}$ (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 47. Длиной вектора a в евклидовом или эрмитовом пространстве называется число $|a| = \sqrt{(a,a)}$.

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

Лемма 48 (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. □

Лемма 49 (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

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

откуда следует неравенство треугольника. □

Процесс ортогонализации

Определение 50. Два вектора a, b называются ортогональными ($a \perp b$), если $(a, b) = 0$. Система векторов $e_1, \ldots, e_n$ называется ортонормированной, если $(e_i, e_j) = \delta_{ij}$, т.е. если векторы попарно ортогональны и длина каждого из них равна 1.

Лемма 51. *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 $e_1, \ldots, e_j$ — 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 $V_1 \subset V_2 \subset \ldots \subset V_j$.

Лемма 52. *Существует такой набор попарно ортогональных векторов $a_1, \ldots, a_j$, что для каждого номера i линейная оболочка $\langle a_1, \ldots, a_i \rangle$ совпадает с V_i.*

Доказательство. (индукция по количеству векторов)

1) При $j = 1$ утверждение очевидно.

2) Пусть это утверждение выполнено для количества векторов, равного j, докажем его для $j + 1$. Т.к. утверждение верно для j векторов, то мы можем считать, что векторы $a_1, \ldots, a_j$ с указанными свойствами уже построены. Поскольку, по предположению индукции, $\langle a_1, \ldots, a_j \rangle = \langle e_1, \ldots, e_j \rangle$, среди векторов $a_1, \ldots, a_j$ не может быть нулевого вектора (иначе $\dim\langle a_1, \ldots, a_j \rangle < j = \dim\langle e_1, \ldots, e_j \rangle$).

Построим вектор a_{j+1} в виде

$$a_{j+1} = e_{j+1} + \lambda_1 a_1 + \ldots + \lambda_j a_j. \tag{3}$$

Линейная оболочка векторов $a_1, \ldots, a_{j+1}$ совпадает с $\langle e_1, \ldots, e_{j+1} \rangle$ при любых λ_i, поэтому мы будем подбирать коэффициенты λ_i так, чтобы выполнялось условие $(a_{j+1}, a_i) = 0$ для всех $i = 1, \ldots, j$. Рассмотрим скалярное произведение обеих частей (3) на a_i слева:

$$0 = (a_i, a_{j+1}) = (a_i, e_{j+1}) + \lambda_1(a_i, a_1) + \ldots + \lambda_j(a_i, a_j).$$

Поскольку $(a_i, a_k) = 0$ при $k \neq i$ по предположению индукции, то $0 = (a_i, e_{j+1}) + \lambda_i(a_i, a_i)$, следовательно $\lambda_i = -\frac{(a_i, e_{j+1})}{(a_i, a_i)}$ (знаменатель отличен от нуля, т.к. $a_i \neq 0$). □

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

Пусть $e_1, \ldots, e_n$ — базис в линейном пространстве V. Применим к нему процесс ортогонализации Грама-Шмидта и получим другой базис V, $a_1, \ldots, a_n$. Это, действительно, базис, т.к. их количество равно размерности V, и, поскольку $\langle a_1, \ldots, a_n \rangle = V$, они линейно независимы.

Лемма 53. *Матрица C перехода от базиса $e_1, \ldots, e_n$ к базису $a_1, \ldots, a_n$ верхнетреугольная, с единицами на диагонали.*

Доказательство. Поскольку $\langle e_1, \ldots, e_j \rangle = \langle a_1, \ldots, a_j \rangle$, для любых $\lambda_1, \ldots, \lambda_j \in \mathbb{K}$ существуют такие $\mu^1, \ldots, \mu^j \in \mathbb{K}$, что $\lambda_1 a_1 + \cdots + \lambda_j a_j = \mu^1 e_1 + \cdots + \mu^j e_j$. Тогда формула (3) примет вид

$$a_{j+1} = e_{j+1} + \mu^1 e_1 + \cdots + \mu^j e_j,$$

т.е.

$$\begin{array}{lcl} a_1 & = & e_1; \\ a_2 & = & e_2 + \mu_2^1 e_1; \\ a_3 & = & e_3 + \mu_3^1 e_1 + \mu_3^2 e_2; \\ \cdots & \cdots & \cdots\ \cdots\ \cdots\ \cdots\ \cdots \\ a_n & = & e_n + \mu_n^1 e_1 + \cdots + \mu_n^{n-1} e_{n-1}. \end{array}$$

Значит, матрица перехода имеет вид

$$C = \begin{pmatrix} 1 & \mu_2^1 & \mu_3^1 & \mu_4^1 & \vdots & \mu_n^1 \\ 0 & 1 & \mu_3^2 & \mu_4^2 & \vdots & \mu_n^2 \\ 0 & 0 & 1 & \mu_4^3 & \vdots & \mu_n^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

□

Ортогональные матрицы. QR-разложение

Определение 54. Базис $e_1, \dots, e_n$ евклидова или эрмитова пространства V называется *ортогональным*, если $(e_i, e_j) = 0$ при $i \neq j$. Он называется *ортонормированным*, если он ортогональный и $(e_i, e_i) = 1$, $i = 1, \dots, n$.

Если $e_1, \dots, e_n$ — ортогональный базис, то базис $\widetilde{e}_1, \dots, \widetilde{e}_n$, где $\widetilde{e}_i = \frac{1}{\sqrt{(e_i, e_i)}} e_i$, $i = 1, \dots, n$, — ортонормированный.

Определение 55. Матрица C с коэффициентами из $\mathbb{R}$ называется *ортогональной*, если $C^t C = E$. Матрица C с коэффициентами из $\mathbb{C}$ называется *унитарной*, если $\overline{C}^t C = E$ (здесь через $\overline{C}$ обозначает матрицу, полученную из матрицы C заменой всех элементов на их комплексно сопряженные).

Лемма 56. *Пусть $e_1 \dots, e_n$ — ортонормированный базис евклидова (эрмитова) пространства V, матрица C — матрица перехода к базису $\widetilde{e}_1, \dots, \widetilde{e}_n$. Базис $\widetilde{e}_1, \dots, \widetilde{e}_n$ ортонормированный тогда и только тогда, когда матрица C ортогональная (унитарная).*

Доказательство. Пусть $C = \begin{pmatrix} c_1^1 & \cdots & c_n^1 \\ \vdots & \ddots & \vdots \\ c_1^n & \cdots & c_n^n \end{pmatrix}$. Тогда $\widetilde{e}_i = c_i^j e_j$, $\widetilde{e}_k = c_k^l e_l$, поэтому $(\widetilde{e}_i, \widetilde{e}_k) = (c_i^j e_j, c_k^l e_l) = \overline{c}_i^j c_k^l (e_j, e_l) = \sum_{j=1}^n \overline{c}_i^j c_k^j$. Ортонормированность базиса $\widetilde{e}_1, \dots, \widetilde{e}_n$ равносильна тому, что $\sum_{j=1}^n \overline{c}_i^j c_k^j = 0$ при $i \neq k$, и $\sum_{j=1}^n \overline{c}_i^j c_i^j = 1$, $i = 1, \dots, n$. Но $\sum_{j=1}^n \overline{c}_i^j c_k^j$ — это элемент матрицы $\overline{C}^t C$ с индексами i и k.

□

Таким образом, ортогональные (унитарные) матрицы — это матрицы перехода от ортонормированного базиса к ортонормированному.

Если рассматривать столбцы матрицы как координаты векторов (записанные в ортонормированном базисе), то эта матрица ортогональна (унитарна), если эти вектора-столбцы ортонормированы.

Теорема 57. *Любая обратимая матрица C допускает единственное представление в виде произведения $C = QR$, где Q — ортогональная (унитарная) матрица, а R — верхнетреугольная, с положительными числами на диагонали.*

Доказательство. Пусть $e_1, \ldots, e_n$ — ортонормированный базис. Обратимость C позволяет ее рассматривать как матрицу перехода к некоторому другому базису $\widetilde{e}_1, \ldots, \widetilde{e}_n$. Применим к базису $\widetilde{e}_1, \ldots, \widetilde{e}_n$ процесс ортогонализации Грама–Шмидта, в результате которого получим ортогональный базис $\widetilde{a}_1, \ldots, \widetilde{a}_n$. Нормализуем его, и получим уже ортонормированный базис $a_1, \ldots, a_n$. Поскольку матрица перехода от базиса $\widetilde{e}_1, \ldots, \widetilde{e}_n$ к базису $\widetilde{a}_1, \ldots, \widetilde{a}_n$ верхнетреугольная с единицами на диагонали, матрица перехода от базиса $\widetilde{e}_1, \ldots, \widetilde{e}_n$ к базису $a_1, \ldots, a_n$ тоже верхнетреугольная, но на диагонали могут быть произвольные положительные числа (т.к. мы делили вектора $\widetilde{a}_i$ на их длины)). Обозначим эту матрицу через T.

Итак, $(\widetilde{e}_1, \ldots, \widetilde{e}_n) = (e_1, \ldots, e_n)C$, $(a_1, \ldots, a_n) = (\widetilde{e}_1, \ldots, \widetilde{e}_n)T$. Отсюда $(a_1, \ldots, a_n) = (e_1, \ldots, e_n)CT$. Поскольку базис $a_1, \ldots, a_n$ ортонормированный, матрица CT, служащая матрицей перехода от одного ортонормированного базиса к другому ортонормированному, является ортогональной (унитарной). Обозначим ее Q. Тогда $CT = Q$. Обозначим также $R = T^{-1}$. Тогда $C = QR$. Матрица, обратная к верхнетреугольной, также верхнетреугольна, и на ее диагонали также положительные числа.

Мы доказали существование требуемого представления в виде произведения. Остается доказать единственность. Пусть $C = Q_1R_1 = Q_2R_2$, где Q_1, Q_2 ортогональные (унитарные), а R_1, R_2 верхнетреугольные с положительными числами на диагонали. Тогда $Q_2^{-1}Q_1 = R_2R_1^{-1}$. Простая проверка показывает, что как ортогональные (унитарные) матрицы, так и верхнетреугольные с положительными числами на диагонали, образуют группу, т.е. обратные элементы и произведения принадлежат этой же группе. Таким образом, ортогональная матрица $Q = Q_2^{-1}Q_1$ равна верхнетреугольной с положительными числами на диагонали $R = R_2R_1^{-1}$. Остается проверить, что единственная матрица, которая является одновременно ортогональной (унитарной) и верхнетреугольной с положительными числами на диагонали, является единичной матрицей. Рассмторим такую матрицу $Q = \begin{pmatrix} q_1^1 & q_2^1 & q_3^1 & \cdots & q_n^1 \\ 0 & q_2^2 & q_3^2 & \cdots & q_n^2 \\ 0 & 0 & q_3^3 & \cdots & q_n^3 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & q_n^n \end{pmatrix}$. Ее ортогональность (унитарность) означает, что вектор-столбцы образуют ортонормированный базис. Единичная длина первого столбца означает, что $q_1^1 = 1$ (напомним, что на диагонали должны быть положительные числа). Ортогональность первого и второго столбцов дает $q_2^1 = 0$. Ортогональность третьего столбца с первым и вторым дает $q_3^1 = q_3^2 = 0$ и т.д.

□

Ортогональное дополнение

Определение 58. 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. $V^\perp = \{w \in W : (v, w) = 0 \ \ \forall v \in V\}$.

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

Лемма 59. $W = V \oplus V^\perp$.

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

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

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

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

Лемма 60. $|a_2| \geqslant |a_\perp|$.

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 61. 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.

Лемма 62. *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.

Лемма 63. *(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 64. 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.

Метод наименьших квадратов

Допустим, что мы исследуем какое-нибудь природное явление и хотим описать его линейной формулой, т.е. мы предполагаем, что какая-то величина b линейно зависит от других — $a_1, \ldots, a_n$, и хотим получить эту зависимость $b = a_1x^1 + \ldots + a_nx^n$, т.е. узнать неизвестные коэффициенты $x^1, \ldots, x^n$. Мы делаем m измерений (для точности берем $m > n$) и решаем систему уравнений $\left\{ \begin{array}{lcl} a_1^1x^1 + \ldots + a_n^1x^n & = & b^1 \\ \ldots \quad \ldots \quad \ldots & & \\ a_1^mx^1 + \ldots + a_n^mx^n & = & b^m \end{array} \right.$. 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 $c^j = a_1^jx^1 + \ldots + a_n^jx^n$ будет наименьшим. В качестве отклонения удобно рассмотреть корень из суммы квадратов отклонений координат $\sqrt{(b^1 - c^1)^2 + \ldots + (b^m - c^m)^2} = |b - c|$, что равно длине вектора $b - c$. Будем искать такое псевдо-решение.

Рассмотрим векторы

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

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

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

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

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

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

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

Матрица этой системы уравнений

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

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

Параллелепипеды. Матрица Грама

Определение 65. Пусть $a_1, \ldots, a_n$ — система векторов в векторном пространстве V. *Параллелепипедом*, натянутым на векторы $a_1, \ldots, a_n$ называется множество векторов (точек) $\Pi(a_1, \ldots, a_n) = \{c \in V : c = x^1 a_1 + \ldots + x^n a_n, \quad 0 \leqslant x^1, \ldots, x^n \leqslant 1\}$.

Определение 66. Определим *n-мерный объем* Vol_n параллелепипеда $\Pi(a_1, \ldots, a_n)$ индуктивно:
1) одномерный объем $\mathrm{Vol}_1 \Pi(a_1) := |a_1|$ — это длина вектора;
2)
$$\mathrm{Vol}_n \Pi(a_1, \ldots, a_n) := \mathrm{Vol}_{n-1} \Pi(a_1, \ldots, a_{n-1}) \cdot d(a_n, \langle a_1, \ldots, a_{n-1} \rangle).$$

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

Теорема 67. $(\mathrm{Vol}_n \Pi(a_1, \ldots, a_n))^2 = \det G(a_1, \ldots, a_n)$.

Доказательство. (по индукции)
1) При $n = 1$, очевидно, $|a_1|^2 = (a_1, a_1)$.
2) Пусть утверждение верно для размерности $n-1$, докажем его для размерности n. Спроектируем вектор a_n на линейную оболочку векторов $a_1, \ldots, a_{n-1}$: $a_n = (a_n)_0 + (a_n)_\perp = \lambda_1 a_1 + \ldots + \lambda_{n-1} a_{n-1} + b$, где $b = (a_n)_\perp$ (т.е. b ортогонален векторам $a_1, \ldots, a_{n-1}$) и $|b| = d(a_n, \langle a_1, \ldots, a_{n-1} \rangle)$. Имеем:

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

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

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

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

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

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

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

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

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

Лемма 68. *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 $a_n=\lambda_1a_1+\ldots+\lambda_{n-1}a_{n-1}$. 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 $a_1\neq0$, 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 $\mathrm{Vol}_{k-1}\Pi(a_1,\dots,a_{k-1})\neq 0$, а $\mathrm{Vol}_k\Pi(a_1,\dots,a_k)=0$. Т.к. $\mathrm{Vol}_k\Pi(a_1,\dots,a_k)=\mathrm{Vol}_{k-1}\Pi(a_1,\dots,a_{k-1})\cdot d(a_k,\langle a_1,\dots,a_{k-1}\rangle)$, то $d(a_k,\langle a_1,\dots,a_{k-1}\rangle)=0$, т.е. $a_k\in\langle a_1,\dots,a_{k-1}\rangle$. Следовательно векторы $a_1,\dots,a_k$, а ,значит, и $a_1,\dots,a_n$, линейно зависимы. □

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 $b_k=c_k^ia_i$ и элементы матрицы G' равны $g'_{ij}=(b_i,b_j)=(c_i^ka_k,c_j^la_l)=\overline{c_i^k}(a_k,a_l)c_j^l=\overline{c_i^k}g_{kl}c_j^l$, то $G'=\overline{C}^tGC$. Если базис $a_1,\dots,a_n$ был ортонормированным, то $G=E=\begin{pmatrix}1 & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & 1\end{pmatrix}$ и $G'=\overline{C}^tC$.

Лемма 69. *Произвольная квадратная матрица G является матрицей Грама для некоторого набора линейно независимых векторов тогда и только тогда, когда существует такая невырожденная матрица C, что $G=\overline{C}^tC$.*

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 $G=\overline{C}^tC$.

$\Leftarrow$: пусть $G=\overline{C}^tC$, 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 $b_1,\dots,b_n$. □

Билинейные функции (формы)

Определение 70. 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.

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

Если выбрать базис $e_1,\dots,e_n$ 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.

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 $g(x,y)=g(x^ie_i,y^je_j)=x^ig(e_i,e_j)y^j=x^ig_{ij}y^j$ или, в матричной форме

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

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 71. *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 72. *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 73. *Размерности левого и правого ядер совпадают и равны* $\dim G_L = \dim G_R = \dim V - \operatorname{rk} g$.

Доказательство. Запишем в матричной форме условие принадлежности вектора $x = (x^1 \dots x^n)$ левому ядру G_L:

$$x = (x^1 \dots x^n) \in G_L \Leftrightarrow (x^1 \dots x^n) G \begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix} = 0, \ \forall \begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix}.$$

Рассмотрим строку $(x^1 \dots x^n) G = (u^1 \dots u^n)$. Для нее должно выполняться $\sum_{i=1}^n u^i y^i = 0$ для любых значений $(y^1 \dots y^n)$, в частности и для набора $y^1{=}0, \dots, y^{k-1}{=}0, y^k{=}1, y^{k+1}{=}0, \dots, y^n{=}0$. Отсюда следует, что $u^k = 0$, а в силу произвольности выбора номера k получаем, что вся строка $(u^1 \dots u^n) = (x^1 \dots x^n) G$ является нулевой. Транспонируем равенство $(x^1 \dots x^n) G = 0$:

$$x = (x^1 \dots x^n) \in G_L \Leftrightarrow G^t \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = 0.$$

Находим размерность пространства решений последней системы:

$$\dim G_L = \dim V - \operatorname{rk} G^t = \dim V - \operatorname{rk} G = \dim V - \operatorname{rk} g.$$

Аналогичным образом сопоставляя правому ядру G_R систему однородных уравнений и вычисляя размерность пространства ее решений, завершаем доказательство. $\square$

Определение 74. Билинейная функция g называется *невырожденной*, если $\dim G_L = \dim G_R = 0$ (это условие равносильно тому, что $\det G \neq 0$ или $\operatorname{rk} g = \dim V$).

Примеры:

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

2) билинейная функция может быть невырождена на всем пространстве, но быть вырожденной на подпространстве! Например, пусть $G = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, рассмотрим вектор $a = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ и подпространство $V = \langle a \rangle$. Т.к. $g(a,a) = 0$, то для любых двух векторов $b, c \in V$ (т.е. коллинеарных вектору a) $g(b,c) = 0$, 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.

Симметричные и кососимметричные билинейные 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, $g^t(a,b) = g(b,a)$.

Определение 75. 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 76. *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 77. Функция $f : V \to \mathbb{K}$ называется *квадратичной функцией (формой)*, если существует такая билинейная функция g, что $f(a) = g(a,a)$ для любого $a \in V$.

Заметим, что если f — квадратичная функция, то

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

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

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

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

Утверждение 78. *Любая билинейная функция допускает единственное разложение в сумму симметричной и кососимметричной билинейных функций.*

Доказательство.

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

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

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

Метод Лагранжа выделения полных квадратов и нормальный вид симметричной функции

Теорема 79. *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.*

Доказательство.

Для приведения матрицы симметричной билинейной функции к нормальному виду воспользуемся методом Лагранжа выделения полных квадратов.

Мы установили ранее, что любая симметричная билинейная функция $g(x,y)$ однозначно восстанавливается по своей квадратичной форме $Q(x) = g(x,x)$. Будем приводить к нормальному виду квадратичную форму $Q(x) = g_{11}(x^1)^2 + 2g_{12}x^1x^2 + \ldots$.

Возможны три случая:

а) Коэффициент g_{11} при квадрате $(x^1)^2$ не равен нулю. Выделим полный квадрат, содержащий $(x^1)^2$:

$$\begin{aligned} Q(x) &= g_{11}\Big((x^1)^2 + \frac{2g_{12}}{g_{11}}x^1x^2 + \ldots + \frac{2g_{1n}}{g_{11}}x^1x^n\Big) + \sum_{i,j>1} g_{ij}x^ix^j = \\ &= g_{11}\Big(x^1 + \frac{g_{12}}{g_{11}}x^2 + \ldots + \frac{g_{1n}}{g_{11}}x^n\Big)^2 - \sum_{i,j>1} \frac{g_{1i}g_{1j}}{g_{11}}x^ix^j + \sum_{i,j>1} g_{ij}x^ix^j. \end{aligned}$$

Сделав замену координат

$$\begin{cases} \widetilde{x}^1 = x^1 + \frac{g_{12}}{g_{11}}x^2 + \ldots + \frac{g_{1n}}{g_{11}}x^n, \\ \widetilde{x}^2 = x^2, \\ \ldots \\ \widetilde{x}^n = x^n. \end{cases}$$

получим $Q(x) = g_{11}(\widetilde{x}^1)^2 + \ldots$, где многоточие стоит вместо слагаемых, не содержащих переменную $\widetilde{x}^1$.

б) $g_{11} = 0$, но имеется ненулевое слагаемое вида $g_{ii}(x^i)^2$ с $i > 1$. В этом случае, мы можем сделать другую замену координат:

$$\begin{cases} x^i &= \widetilde{x}^1, \\ x^1 &= \widetilde{x}^i, \\ x^k &= \widetilde{x}^k,\ k \neq 1, i. \end{cases}$$

После этой замены коэффициент $\widetilde{g}_{11} = g_{ii}$ перед $(\widetilde{x}^1)^2$ уже отличен от нуля, что сводит задачу к предыдущему случаю.

в) все коэффициенты g_{ii} при квадратах координат $(x^i)^2$ равны нулю. Этот случай можно свести к б) линейной заменой координат: сначала мы найдем ненулевое слагаемое в $Q(x)$ вида $2g_{kl}x^k x^l$ (этого нельзя сделать только у формы тождественно равной нулю), а затем сделаем такую замену координат:

$$\left\{\begin{array}{ccc} x^k & = & \widetilde{x}^k + \widetilde{x}^l, \\ x^l & = & \widetilde{x}^k - \widetilde{x}^l, \\ x^m & = & \widetilde{x}^m,\ m \neq k, l \end{array}\right.$$

После этой замены (только уже в новых координатах) $x^k x^l = (\widetilde{x}^k)^2 - (\widetilde{x}^l)^2$ и у нас есть на выбор два квадрата $2g_{kl}(\widetilde{x}^k)^2$ и $2g_{kl}(\widetilde{x}^l)^2$.

После выделения полного квадрата $(\widetilde{x}^1)^2$ оставшиеся в $Q(x)$ слагаемые, не содержащие $\widetilde{x}^1$, представляют собой квадратичную форму уже от меньшего числа неизвестных. Применив очевидное индуктивное предположение получим, что $Q(x) = \widetilde{g}_{11}(\widetilde{x}^1)^2 + \widetilde{g}_{22}(\widetilde{x}^2)^2 + \ldots + \widetilde{g}_{nn}(\widetilde{x}^n)^2$. Говоря другими словами, в новом базисе матрица $\widetilde{G}$ квадратичной функции $Q(x)$ будет диагональной с элементами $\widetilde{g}_{11}, \widetilde{g}_{22}, \ldots, \widetilde{g}_{nn}$ на диагонали.

Далее изучение вещественного случая будет отличаться от комплексного. "Меняя местами" переменные, т.е. делая преобразования координат как в случае б), мы можем добиться того, что положительные коэффициенты $\widetilde{g}_{ii} > 0$ идут первыми (если они есть конечно), затем отрицательные $\widetilde{g}_{ii} < 0$, и наконец нулевые $\widetilde{g}_{ii} = 0$:

$$\widetilde{g}_{11} > 0, \ldots, \widetilde{g}_{kk} > 0, \widetilde{g}_{k+1,k+1} < 0, \ldots, \widetilde{g}_{k+l,k+l} < 0, \widetilde{g}_{k+l+1,k+l+1} = 0, \ldots, \widetilde{g}_{k+l+s,k+l+s} = 0.$$

Эту диагональную матрицу можно упростить еще больше, сделав следующее преобразование "гомотетии":

$$\left\{\begin{array}{ccc} x'^1 & = & \sqrt{|\widetilde{g}_{11}|}\widetilde{x}^1, \\ \ldots & \ldots & \ldots\ \ldots\ \ldots\ \ldots \\ x'^{k+l} & = & \sqrt{|\widetilde{g}_{k+1,k+1}|}\widetilde{x}^{k+l}, \\ x'^m & = & \widetilde{x}^m,\ k+l \leq m \leq n \end{array}\right.$$

Окончательно (в системе координат $(x'^1, \ldots, x'^n)$) вид нашей квадратичной формы будет таков:

$$Q(x) = (x'^1)^2 + \ldots + (x'^k)^2 - (x'^{k+1})^2 - \ldots - (x'^{k+l})^2.$$

Данный вид называется **нормальным видом вещественной квадратичной формы** $Q(x)$.

В комплексном же случае нет нужды "сортировать" квадраты координат на положительные и отрицательные, достаточно записать подряд сначала ненулевые квадраты, а корень можно извлечь непосредственно из $\widetilde{g}_{ii}$:

$$\left\{\begin{array}{ccc} x'^1 & = & \sqrt{\widetilde{g}_{11}}\widetilde{x}^1, \\ \ldots, & & \\ x'^r & = & \sqrt{\widetilde{g}_{r,r}}\widetilde{x}^r, \\ x'^m & = & \widetilde{x}^m,\ r \leq m \leq n \end{array}\right.$$

И в системе координат $(x'^1, \ldots, x'^n)$) вид нашей комплексной квадратичной формы (называемый **нормальным видом комплексной квадратичной формы**) будет выглядеть так:

$$Q(x) = (x'^1)^2 + \ldots + (x'^r)^2.$$

Переходя к соответствующей вещественной симметричной билинейной форме, мы заключаем, что для нее существует базис, в котором она имеет нормальный вид

$$g(x, y) = x^1 y^1 + \ldots + x^k y^k - x^{k+1} y^{k+1} - \ldots - x^{k+l} y^{k+l}.$$

Аналогично, комплексная билинейная симметрическая функция имеет нормальный вид

$$g(x,y) = x^1y^1 + \ldots + x^ry^r.$$

Очевидно, что число r ненулевых квадратов равняется рангу данной формы. $\square$

Закон инерции квадратичной формы

Теперь обсудим вопрос о единственности нормального вида квадратичной формы (симметричной билинейной функции), заданной на векторном пространстве W. Для комплексной билинейной симметричной функции нормальный вид определяется единственным числом — рангом r квадратичной (билинейной) функции, поэтому не зависит от способа приведения к нормальному виду. Рассмотрим 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: k — количество единиц l — количество минус единиц, s — количество нулей. Видно, что $k+l = \operatorname{rk} g$, $s = \dim W - \operatorname{rk} g$ следовательно, $k+l$ и s не зависят от выбора базиса.

Определение 80. Набор чисел (k,l,s) называется *сигнатурой* вещественной квадратичной функции (симметричной билинейной функции).

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

Теорема 81 (Закон инерции квадратичной формы)**.** *Число k положительных квадратов, также как и число l отрицательных квадратов в нормальном виде вещественной квадратичной формы $Q(x)$ не зависят от способа приведения формы $Q(x)$ к нормальному виду.*

Доказательство. Пусть функция $Q(x)$ имеет нормальный вид в двух разных базисах: $e_1,\ldots,e_k,e_{k+1},\ldots,e_{k+l},e_{k+l+1},\ldots,e_{k+l+s}$ и $\widetilde{e}_1,\ldots,\widetilde{e}_{k'},\widetilde{e}_{k'+1},\ldots,\widetilde{e}_{k'+l'},\widetilde{e}_{k'+l'+1},\ldots,\widetilde{e}_{k'+l'+s}$, причем, как мы уже знаем, $s'=s$ и $k'+l'=k+l$. Рассмотрим в W подпространства

$$V_+ = \langle e_1,\ldots,e_k\rangle,\ V_- = \langle e_{k+1},\ldots,e_{k+l}\rangle,\ V_0 = \langle e_{k+l+1},\ldots,e_{k+l+s}\rangle$$

и

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

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

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

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

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 $Q(x) > 0$, но, с другой стороны, т.к. $x \in \widetilde{V}_- \oplus \widetilde{V}_0$, то $Q(x) \leqslant 0$. Получили противоречие. Случай $k < k'$ рассматривается аналогично. □

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 Δ_k минором порядка k квадратной матрицы G называется минор, составленный из первых k строк и k столбцов.

Теорема 82 (Якоби). *Пусть G — матрица симметричной билинейной функции g в некотором базисе. Пусть все угловые миноры Δ_k до порядка $r = \operatorname{rk} G$ отличны от 0. Тогда существует базис, в котором функция g имеет вид $g(x,y) = \Delta_1 x^1 y^1 + \frac{\Delta_2}{\Delta_1} x^2 y^2 + \ldots + \frac{\Delta_r}{\Delta_{r-1}} x^r y^r$, где Δ_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 $\Delta'_k = |C_k^t||G_k||C_k| = \Delta_k$).

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 $g(e_1, e_1) = \Delta_1 \neq 0$. Аналогично можно найти c_3^1 и c_3^2 из условия $g(e'_3, e'_1) = g(e'_3, e'_2) = 0$, затем найдем 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 Δ_{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 угловые миноры не изменяются). Таким образом, мы можем найти векторы $e'_1, \dots, e'_r$ 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 $g'_{ii} = g(e'_i, e'_i)$ отличны от нуля при $1 \leqslant i \leqslant r$.

Оставшиеся векторы $e'_{r+1}, \dots, e'_n$ найдем из условия $g'_{ij} = g(e'_i, e'_j) = 0$, если $i \leqslant r$ а $j > r$. Для этого можно 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

$$\left(\begin{array}{ccc|c} g'_{11} & & & \\ & \ddots & & \mathbf{0} \\ & & g'_{rr} & \\ \hline & \mathbf{0} & & \star \end{array}\right),$$

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

$$G' = \left(\begin{array}{ccc|ccc} g'_{11} & & & & & \\ & \ddots & & & \mathbf{0} & \\ & & g'_{rr} & & & \\ \hline & & & 0 & & \\ & \mathbf{0} & & & \ddots & \\ & & & & & 0 \end{array}\right).$$

Поскольку $\Delta'_k = \Delta_k$ для всех $k = 1, \dots, n$, имеем равенства

$$\begin{aligned} \Delta_1 = \Delta'_1 &= g(e'_1, e'_1); \\ \Delta'_2 = \Delta_2 = g(e'_1, e'_1) g(e'_2, e'_2) = \Delta_1 g(e'_2, e'_2) &\;\Rightarrow\; g(e'_2, e'_2) = \frac{\Delta_2}{\Delta_1}; \\ \Delta'_3 = \Delta_3 = \Delta'_2 g(e'_3, e'_3) = \Delta_2 g(e'_3, e'_3) &\;\Rightarrow\; g(e'_3, e'_3) = \frac{\Delta_3}{\Delta_2}; \\ \dots\dots\dots\dots\dots\dots\dots\dots & \\ \Delta'_r = \Delta_r = \Delta'_{r-1} g(e'_r, e'_r) = \Delta_{r-1} g(e'_r, e'_r) &\;\Rightarrow\; g(e'_r, e'_r) = \frac{\Delta_r}{\Delta_{r-1}}. \end{aligned}$$

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

$$\begin{aligned} g(x, y) &= g(e'_1, e'_1) x^1 y^1 + \ldots + g(e'_r, e'_r) x^r y^r = \\ &= \Delta_1 x^1 y^1 + \frac{\Delta_2}{\Delta_1} x^2 y^2 + \ldots + \frac{\Delta_r}{\Delta_{r-1}} x^r y^r, \end{aligned}$$

что и требовалось доказать. □

Определение 83. Квадратичная форма $Q(x)$ в вещественном пространстве V называется *положительно определенной*, если $Q(x) > 0$ для любого ненулевого $x \in V$ (эквивалентное определение: если ее сигнатура равна $(\dim V, 0, 0)$).

Теорема 84 (Критерий Сильвестра). *Функция $Q(x)$ 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 $Q(x)$ имеет вид $Q(x) = \lambda_1(x^1)^2 + \ldots + \lambda_n(x^n)^2$, где n — размерность пространства, а $\lambda_i > 0$ при всех $i = 1, \ldots, n$, следовательно, $Q(x)$ положительно определена.

Обратно, предположим, что функция $Q(x)$ положительно определена. Докажем сначала, что все угловые миноры будут ненулевыми. Отметим сразу, что $\det G = \Delta_n \neq 0$, т.к. функция $Q(x)$ невырождена и ее ранг равен n. Поскольку функция $Q(x)$ положительно определена, то и ее ограничения на любые подпространства также будут положительно определенными, а, следовательно, и невырожденными. Т.к. G_k — это матрица ограничения $Q(x)$ на $\langle e_1, \ldots, e_k \rangle$, то $\Delta_k \neq 0$ 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 Δ_i были положительными. □

Обобщенное ортогональное дополнение

Определение 85. Пусть $V \subset W$, а на W задана (косо)симметричная билинейная функция g, тогда $V^\perp = \{a \in W : g(a, b) = 0 \;\; \forall b \in V\}$.

Иногда пишут $\perp_g$ вместо $\perp$, чтобы явно указать, что это определение зависит от (косо)симметричной функции g.

Пример.

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

Напомним, что если билинейная функция g симметрична или кососимметрична, ее правое и левое ядро совпадают. Мы его обозначаем через $\operatorname{Ker} g$.

Лемма 86. $\dim V^\perp \geqslant \dim W - \dim V$, *причем равенство достигается тогда и только тогда, когда* $\operatorname{Ker} g \cap V = \{0\}$.

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

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

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

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

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

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

Лемма 87. *$V \cap V^\perp = \operatorname{Ker} g_V$, где g_V — ограничение g на V.*

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

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

Следствие 88. *Если $\operatorname{Ker} g_V = \{0\}$ (т.е. ограничение g на V не вырождено), то $W = V \oplus V^\perp$.*

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

Лемма 89. *Если ограничения g на V и на $V^\perp$ невырождены, то $(V^\perp)^\perp = V$.*

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

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

Лемма 90. *Пусть $W = V \oplus V^\perp$, $e_1, \ldots, e_r$ — базис в V, $e_{r+1}, \ldots, e_n$ — базис в $V^\perp$. Тогда в базисе $e_1, \ldots, e_n$ матрица билинейной функции имеет вид $G = \left(\begin{array}{c|c} \star & 0 \\ \hline 0 & \star \end{array}\right)$.*

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

Кососимметрические билинейные функции

Пусть g — кососимметрическая билинейная функция на пространстве W.

Рассмотрим сначала случай, когда размерность W мала.

1. Т.к. $g(e,e) = -g(e,e)$, то любая кососимметричная функция на одномерном пространстве тождественно равна нулю.

2. В двумерном случае, если g не тождественно равна нулю, то найдутся такие векторы $a, b \in W$, что $g(a,b) = \alpha \neq 0$. Приэтом эти два вектора обязаны быть линейно независимыми: если $b = \lambda a$, то $g(a,b) = \lambda g(a,a) = 0$. Следовательно, они образуют базис в двумерном пространстве W. 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 91. *Для любой кососимметричной билинейной функции g 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. Этот вид единственен, с точностью до перестановки блоков.*

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

Итак, любая кососимметрическая билинейная функция имеет нормальный вид $g(x, y) = x^1y^2 - x^2y^1 + x^3y^4 - x^4y^3 + \ldots + x^{2k-1}y^{2k} - x^{2k}y^{2k-1}$. Единственность следует из того, что число $2k = \operatorname{rk} G$ не зависит от выбора базиса.

□

Симметрические билинейные функции на евклидовом пространстве

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

Рассмотрим симметричную билинейную функцию g и соответствующую ей квадратичную функцию $Q(x) = g(x, x)$.

Пусть $t \in \mathbb{R}$. Вектор-функция $a(t)$ называется *дифференцируемой*, если все координаты $a^1(t), \ldots, a^n(t)$ векторнозначной функции $a(t)$ дифференцируемы. Легко проверить, что это определение не зависит от выбора базиса, в котором эти координаты записаны (хотя сами координаты от базиса зависят). Производную вектор-функции a обозначим a' или $\frac{da}{dt}$. Рассмотрим числовую функцию $Q(a(t)) = g(a(t), a(t))$.

Лемма 92. $\frac{dQ(a(t))}{dt} = 2g(a(t), \frac{da(t)}{dt})$.

Доказательство. Выберем базис; пусть g_{ij} — коэффициенты матрицы симметричной билинейной функции (= матрицы квадратичной функции), $a^i(t)$ — координаты вектор-функции в этом базисе. Тогда

$$\begin{aligned}
\frac{dQ(a(t))}{dt} &= \frac{d}{dt}\sum_{i,j=1}^{n} g_{ij}a^i(t)a^j(t) = \\
&= \sum_{i,j=1}^{n}\left(g_{ij}\frac{da^i(t)}{dt}a^j(t) + g_{ij}a^i(t)\frac{da^j(t)}{dt}\right) = \\
&= g\left(a(t), \frac{da(t)}{dt}\right) + g\left(\frac{da(t)}{dt}, a(t)\right) = \\
&= 2g\left(a(t), \frac{da(t)}{dt}\right).
\end{aligned}$$

□

Нам понадобится следующее утверждение из курса математического анализа.

Теорема 93. *Непрерывная функция f, заданная на замкнутом ограниченном множестве $S \subset \mathbb{R}^n$ достигает своего максимального значения в некоторой точке $x \in S$.*

Применим это утверждение к функции $f(x) = Q(x)$ на множестве векторов x единичной длины, т.е. $S = \{x \in V : |x| = 1\}$. Итак, существует вектор $a \in S$, на котором функция $Q(x)$ достигает своего максимума на множестве S.

Возьмем вектор $b \in V$, длины 1, и рассмотрим вектор-функцию $x(t) = \cos ta + \sin tb$.

Лемма 94. *Если $(a, b) = 0$, то $g(a, b) = 0$.*

Доказательство. Сначала заметим, что если $(a, b) = 0$, то $|x(t)| = 1$ для любого t, поэтому $x(t) \in S$. Действительно, $|x(t)|^2 = (\cos ta + \sin tb, \cos ta + \sin tb) = \cos^2 t(a, a) + 2\cos t \sin t(a, b) + \sin^2 t(b, b) = \cos^2 t + \sin^2 t = 1$.

Так как функция $Q(x)$ достигает максимума на S при $x = a$, функция $h(t) = Q(x(t))$ достигает максимума при $t = 0$ (т.к. $x(0) = a$). Эта функция дифференцируемая, поэтому $\frac{d}{dt}h(t)|_{t=0} = 0$. По Лемме 92, $0 = g(x(0), \frac{dx}{dt}|_{t=0}) = g(a, b)$ (т.к. $\frac{dx}{dt}|_{t=0} = b$).

□

Теорема 95. *Пусть Q — квадратичная функция на евклидовом пространстве V. Тогда существует ортонормированный базис $e_1, \ldots, e_n$, в котором матрица G функции Q имеет диагональный вид.*

Доказательство. Доказываем по индукции по размерности. Если $n = \dim V = 1$, то утверджение очевидно — одномерная матрица есть число. Докажем индукционный переход. Пусть a — вектор единичной длины,на котором функция $Q(x)$ достигает максимума на S. Возьмем его в качестве первого базисного вектора, $e_1 = a$, и дополним его до ортонормированного базиса всего V векторами $e_2, \ldots, e_n$. Поскольку в этом базисе коэффициенты квадратичной функции равны $g_{ij} = g(e_i, e_j)$, по Лемме 94 получаем, что $g_{1i} = 0 = g_{i,1}$ для $i = 2, \ldots, n$. Обозначим $g_{11} = \lambda_1$. Тогда матрица квадратичной функции Q в указанном базисе имеет вид $G = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & G' & \\ 0 & & & \end{pmatrix}$. Матрица G' — матрица квадратичной функции на подпространстве $L = \langle e_2, \ldots, e_n \rangle$, $\dim L = n - 1$. По предположению индукции, в L существует ортонормированный базис, в котором матрица этой функции имеет диагональный вид, $\widetilde{G}' = \begin{pmatrix} \lambda_2 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$.

□

Лемма 96. *Указанный в теореме диагональный вид единственен с точностью до перестановки диагональных элементов.*

Доказательство. Пусть G — матрица квадратичной функции Q в некотором ортонормированном базисе, а $\widetilde{G}$ — ее же матрица в другом ортонормированном базисе. Тогда $\widetilde{G} = C^t G C$, где C — ортогональная матрица перехода, и

$$\det(\widetilde{G} - \lambda E) = \det(C^t G C - \lambda E) = \det(C^t(G - \lambda E)C) = \det(G - \lambda E),$$

поэтому корни многочлена $\det(G - \lambda E)$ не зависят от выбора ортонормрованного базиса. Но если $\widetilde{G} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$, то эти корни равны $\lambda_1, \ldots, \lambda_n$.

□

Определение 97. Этот диагональный вид называется *каноническим видом* билинейной функции. Векторы соответствующего базиса (он тоже называется каноническим базисом) называются главными осями функции Q, и иногда приведение к каноническому виду называют приведением к главным осям.

Например, в трехмерном пространстве уравнение $Q(x) = 1$ задает поверхность второго порядка, а ее уравнение в главных осях имеет вид $\lambda_1(x^1)^2 + \lambda_2(x^2)^2 + \lambda_3(x^3)^2 = 1$.

Пара симметричных билинейных функций

Теорема 98. *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 ортонормированном базисе $e_1, \ldots, e_n$, а X' — столбец координат этого же вектора в первоначальном 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. $(H - \lambda G)X = 0$.

Таким образом, мы доказали следующее утверджение:

Лемма 99. *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.

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

Определение 100. 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 пространства. 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 $g(e_i, e_j) = \begin{cases} 0 & i \neq j \\ \pm 1 & i = j \end{cases}$. В симплектическом случае такое невозможно, т.к. $g(x,x) = 0$ для любого вектора x. Напомним, что матрица невырожденной кососимметричной функции имеет вид $\begin{pmatrix} \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} \end{pmatrix}$ (т.е. без клеток размера 1, т.к. наличие нулевых одномерных клеток противоречит невырожденности функции g). Следовательно, размерность четна. Пусть матрица функции g имеет указанный вид в базисе $e_1, \ldots, e_{2m}$. Рассмотрим базис $e_1, e_3, \ldots, e_{2m-1}, e_2, e_4, \ldots, e_{2m}$, в этом базисе матрица функции g имеет вид $\left(\begin{array}{c|c} \mathbf{0} & \begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix} \\ \hline \begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix} & \mathbf{0} \end{array}\right)$. Такой базис называется *гамильтоновым*. В симплектическом пространстве значение функции g на паре векторов также называют их скалярным произведением (еще более обобщенным, чем псевдоевклидово).

Определение 101. Вектор $x \neq 0$ называется *изотропным*, если $g(x,x) = 0$.

Отметим, что в евклидовых пространствах изотропных векторов не существует. В псевдоевклидовом пространстве с обобщенным скалярным произведением с сигнатурой $(p, n-p, 0)$ вектор с координатами $(0, \ldots, 0, 1, 0 \ldots, 0, 1, 0, \ldots, 0)$ (единицы на местах i, j, где $i \leqslant p$, $j > p$) является изотропным.

Рассмотрим в псевдоевклидовом пространстве множество всех матриц перехода от ортонормированного базиса к ортонормированному.

Аналогично, в симплектическом пространстве рассмотрим множество всех матриц перехода от гамильтонова базиса к гамильтонову.

Лемма 102. *Каждое из этих двух множеств образует группу (по умножению).*

Доказательство. Так как оба базиса — одного типа (ортонормированные или гамильтоновы), матрица соответствующей билинейной функции имеет в них один и тот же вид G. Если C — матрица перехода, то $G = C^tGC$. Тогда $(C^t)^{-1}GC^{-1} = G$, поэтому матрица C^{-1} лежит в том же множестве. Если C_1 и C_2 удовлетворяют равенству $G = C^tGC$, то $(C_1C_2)^tG(C_1C_2) = C_2^t(C_1^tGC_1)C_2 = C_2^tGC_2 = G$, и C_1C_2 также удовлетовряет этому равенству. Наконец, единичная матрица, являясь матрицей перехода от некторого базиса к нему же, также удовлетворяет этому равенству.

□

Эти группы имеют специальные названия и обозначения:

- *псевдоортогональная группа* $O(p,q)$ — группа матриц перехода в псевдоевклидовом пространстве с сигнатурой $(p,q,0)$ (если $q=0$, то группу $O(p,0)$ называют ортогональной и обозначают $O(p)$).

- *симплектическая группа* $Sp(2m)$ — группа матриц перехода в симплектическом пространстве размерности $2m$.

Дадим описание этих групп в малых размерностях.

Группа $O(2)$ — группа матриц перехода от ортонормированного базиса к ортонормированному на плоскости. Такие замены базиса могут собственными или несобственными (т.е. сохранять ориентацию плоскости или менять ее). Если ориентация сохраняется, 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 ориентация меняется, то это еще и композиция с симметрией, и матрица имеет вид $\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$.

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

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

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 перехода должна удовлетворять условию $C^t\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, что равносильно условию $\det = 1$ (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.

Симплектические пространства. Лагранжевы подпространства

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

Определение 103. Подпространство $V \subset W$ называется *изотропным*, если $V \subset V^{\perp_g}$.

Лемма 104. *V изотропно тогда и только тогда, когда $g(x, y) = 0$ для всех $x, y \in V$.*

Доказательство. Напомним, что $V^{\perp_g} = \{x \in W : g(x, y) = 0 \ \forall y \in V\}$. Если $g(x, y) = 0$ для всех $x, y \in V$ и если $x \in V$, то $x \in V^{\perp_g}$, т.е. $V \subset V^{\perp_g}$. Обратно, если $V \subset V^{\perp_g}$, то для любого $x \in V$ имеем $x \in V^{\perp_g}$, т.е. $g(x, y) = 0$ для любого $y \in V$ (и, значит, $g(x, y) = 0$ для любых $x, y \in V$). □

Лемма 105. *Если подпространство $V \subset W$ изотропно, то $\dim V \leqslant \frac{1}{2}\dim W$, причем изотропное подпространство максимальной размерности $\frac{\dim W}{2}$ существует.*

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

Возьмем гамильтонов базис $e_1, \ldots, e_m, e_{m+1}, \ldots, e_{2m}$, где $m = \frac{\dim W}{2}$, и пусть $V = \langle e_1, \ldots, e_m\rangle$. Т.к. этот базис гамильтонов, то в этом базисе матрица g имеет вид $\left(\begin{array}{c|c} 0 & E \\ \hline -E & 0 \end{array}\right)$, а матрица ограничения $g|_V$ функции g на V является просто нулевой. Следовательно, $g(x, y) = 0$ для любых $x, y \in V$, значит, V изотропно и имеет размерность $\dim V = \frac{\dim W}{2}$. □

Определение 106. Изотропное подпространство максимальной размерности называется *лагранжевым.*

Следующее утверждение показывает, что любое изотропное подпространство можно вложить в лагранжево.

Лемма 107. *Для любого изотропного подпространства $V \subset W$ симплектического пространства существует такое лагранжево подпространство $\widetilde{V}$, что $V \subset \widetilde{V}$.*

Отметим, что подпространство $\widetilde{V}$, указанное в лемме, не единственно.

Доказательство. Пусть $\dim V = r \leqslant m = \frac{\dim W}{2}$, тогда, в силу невырожденности функции g на W, $\dim V^{\perp_g} = 2m - r$ и $V \subset V^{\perp_g}$. Обозначим через g' ограничение функции g на $V^{\perp_g}$. Очевидно, функция g' вырождена. Действительно, т.к. g невырождена, то $(V^{\perp_g})^{\perp_g} = V$ и $\operatorname{Ker} g' = V^{\perp_g} \cap (V^{\perp_g})^{\perp_g} = V$. Приведем g' к нормальному виду и получим, что в некотором базисе $\underbrace{e_1, \ldots, e_{m-r}}_{m-r}, \underbrace{e_{m-r+1}, \ldots, e_{2m-2r}}_{m-r}, \underbrace{e_{2m-2r+1}, \ldots, e_{2m-r}}_{r}$ ее матрица будет иметь вид $\left(\begin{array}{c|c|c} 0 & E & 0 \\ \hline -E & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$. Т.к. $V = \operatorname{Ker} g$, то ему отвечает последняя группа векторов, т.е. $V = \langle e_{2m-2r+1}, \ldots, e_{2m-r}\rangle$. Добавив к V векторы $e_{m-r+1}, \ldots, e_{2m-2r}$, получим лагранжево подпространство $\widetilde{V} = \langle e_{m-r+1}, \ldots, e_{2m-2r}, e_{2m-2r+1}, \ldots, e_{2m-r}\rangle$. Оно, очевидно, содержит V и действительно является лагранжевым, т.к. функция g тождественно равна нулю на $\widetilde{V}$ (в матрице ограничению g на $\widetilde{V}$ отвечает правый нижний квадрат, состоящий из нулей) и $\dim \widetilde{V} = r + (m - r) = m$. □

Аффинная классификация квадрик

Пусть в векторном пространстве V размерности n заданы квадратичная функция Q, линейная функция f и число a. Квадрикой называется уравнение $Q(x) + 2f(x) + a = 0$. Обозначим через r ранг квадратичной функции Q.

Лемма 108. *Любая квадрика в подходящем базисе имеет один из следующих видов:*

I $x_1^2 + \dots + x_k^2 - x_{k+1}^2 - \dots - x_r^2 = 1$, *где* $0 \leqslant k \leqslant r$;

II $x_1^2 + \dots + x_k^2 - x_{k+1}^2 - \dots - x_r^2 = 0$, *где* $k \geqslant \frac{r}{2}$;

III $x_1^2 + \dots + x_k^2 - x_{k+1}^2 - \dots - x_r^2 = 2x_{r+1}$, *где* $k \geqslant \frac{r}{2}$.

Доказательство. Квадратичная функция приводится заменой базиса к виду $x_1^2 + \dots + x_k^2 - x_{k+1}^2 - \dots - x_r^2$. Если после этого линейная функция содержит координату с номером больше r, то всю линейную часть можно заменой базиса привести к виду $2x_{r+1}$ и получить тип III. Если такой координаты нет, то группируя полные квадраты, можно от линейной части избавиться, и в зависимости от оставшейся константы получаем тип I или II. $\square$

Линейные операторы

Линейные отображения

Определение 109. 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 $e_1, \dots, e_n$ — базис в V, а $e_1', \dots, e_m'$ — 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. $f(e_i) = a_i^k e_k'$ (a_i^k — коэффициенты разложения вектора $f(e_i)$ по базису e'), тогда $f(x) = x^i a_i^k e_k' = y^k e_k'$ — разложение значения по базису $e_1', \dots, e_k'$. Координаты x^i вектора x в базисе пространства V и координаты y^k значения отображения $f(x)$ в базисе пространства W связаны следующим соотношением:

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

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

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 $A_f = E$;

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

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 ($X = C_1\widetilde{X}$, $Y = C_2\widetilde{Y}$) имеем $C_2\widetilde{Y} = A_f C_1 \widetilde{X}$, т.е. $\widetilde{Y} = C_2^{-1} A_f C_1 \widetilde{X} = \widetilde{A}_f \widetilde{X}$. Окончательно получаем формулу для матрицы оператора в новых базисах $\widetilde{A}_f = C_2^{-1} A_f C_1$.

Определение 111. *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 отображения $f : V \to W$ называется множество векторов $y \in W$, для которых существует прообраз, $\operatorname{Im} f = \{y \in W : \exists x \in V, f(x) = y\}$.

Лемма 112. *Ядро любого линейного отображения является линейным подпространством в V; образ любого линейного отображения 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 отображения аналогична. □

Лемма 113. $\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim V$.

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) максимальность. 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: $\widetilde{A}_f = C^{-1} A_f C$, где C — матрица перехода, а A_f — матрица оператора в старом базисе.

Определение 114. 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

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

Определение 115. 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:

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

Определение 116. 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 117. *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.

Инвариантное подпространство

Определение 118. 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 119. 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. $A_f = \left(\begin{array}{c|c} A_{f_1} & \star \\ \hline 0 & \star \end{array}\right)$.

Невырожденные операторы. Собственные значения и собственные векторы

Определение 120. 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 121. *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 122. 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 123. *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 $\operatorname{Ker}(f - \lambda \cdot id)$, тогда $f(v) = \lambda v$. $\square$

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

Лемма 124. *Пространство $V(\lambda)$ инвариантно относительно оператора f.*

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

Лемма 125. *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. $\square$

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 $V_2 \subset \operatorname{Ker} f$. На самом деле $V_2 = \operatorname{Ker} f$, т.к. если $f(w) = 0$, то в разложении $w = v_1 + v_2$ имеем $v_1 = 0$, т.е. $w \in V_2$.

Определение 126. 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 127. *Если $f^2 = f$, то оператор $f : W \to W$ является оператором проектирования для некоторых V_1 и V_2.*

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

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

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

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, $f(w) = f(v_1) = v_1$, и оператор f действительно является оператором проектирования вдоль V_2 на V_1. $\square$

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.

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 128. *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 129. 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 130. *Для любого оператора f существует, и притом единственный, минимальный многочлен.*

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

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

Лемма 131. *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 $fx = \lambda x$. Заметим, что тогда $f^2 x = f(f(x)) = f(\lambda x) =$

$\lambda f x = \lambda^2 x$, и, аналогично, $f^k x = \lambda^k x$ для всех k. Пусть $p(t)$ — минимальный многочлен, т.е. $p(f) \equiv 0$, тогда $p(f)x = 0$, следовательно

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

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

Обратно, пусть λ — корень минимального многочлена, тогда $p(t) = (t - \lambda)q(t)$, $\deg q < \deg p$ (и тогда $p(f) = (f - \lambda \cdot id)q(f)$), 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. $\square$

Характеристический многочлен

Определение 132. 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 $P_f(t) = a_0 + a_1 t + \ldots + a_n t^n$, то тогда $a_n = (-1)^n$, $a_{n-1} = (-1)^{n-1} \operatorname{tr} f$, $a_0 = \det f$.

Лемма 133. *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. Тогда, т.к. $\det(A_f - \lambda E) = \det(A_{f_1} - \lambda E)\det(A_{f'} - \lambda E)$, то $P_f(t) = P_{f_1}(t) P_{f'}(t)$.

Лемма 134. *Пусть $V(\lambda)$ — инвариантное подпространство, образованное собственными векторами, отвечающими собственному значению λ. Тогда кратность корня характеристического многочлена не меньше размерности подпространства $V(\lambda)$.*

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

Теорема 135 (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

$$(A - \lambda E)C(\lambda) = P_A(\lambda)E. \tag{4}$$

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 (4) 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 $P_A(A) = 0$. $\square$

Диагонализируемые операторы

Определение 136. Оператор f называется *диагонализируемым*, если существует такой базис, что матрица этого оператора в этом базисе диагональна.

Примеры:

1) операторы проектирования диагонализируемы,

2) Оператор с матрицей $A = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$ не диагонализируемый. Действительно, все его собственные значения равны 0, поэтому если бы он был диагонализируемым, его диагональный вид $\widetilde{A}$ был бы нулевой матрицей. Но $\widetilde{A} = C^{-1}AC$ дает противоречие. Это рассуждение годится для любых операторов, у которых нет ненулевых собственных значений.

Лемма 137. *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 $fa_k = f(-\frac{\alpha_1}{\alpha_k}a_1 - \ldots - \frac{\alpha_{k-1}}{\alpha_k}a_{k-1})$, 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 $a_k = 0$, откуда следует линейная независимость системы из k векторов. $\square$

Лемма 138. *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 139. *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 140. *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 (по теореме Гамильтона–Кэли), то $P_g(g) = (-1)^n g^n = 0$, значит, $g^n = 0$. $\square$

Операторы, которые в некоторой степени равны 0, называются *нильпотентными*. Лемма утверждает, что если λ — единственное собственное значение оператора f, то оператор $g = f - \lambda \cdot id$ является нильпотентным.

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.

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 $V_\lambda^{(1)} = \mathrm{Ker}(f - \lambda \cdot id)$.

Аналогично определим $V_\lambda^{(2)} = \mathrm{Ker}(f - \lambda \cdot id)^2$, $V_\lambda^{(3)} = \mathrm{Ker}(f - \lambda \cdot id)^3$ 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 $\dim V_\lambda^{(1)} \leqslant \dim V_\lambda^{(2)} \leqslant \dim V_\lambda^{(3)} \leqslant \ldots \leqslant \dim V$ не может неограниченно возрастать, значит, найдется какой-то номер k, для которого соседние члены этой последовательности совпадут.

Утверждение 141. *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. $\square$

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 142. *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, $f(x) \in V_\lambda$. $\square$

Теорема о разложении пространства в прямую сумму корневых подпространств

Лемма 143. *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 $f - \lambda \cdot id$ 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 $V = V_\lambda \oplus W = \operatorname{Ker} f^p \oplus \operatorname{Im} f^p$. 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 $f|_W$. $\square$

Теорема 144 (о корневом разложении). *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. $\square$

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

Жорданова нормальная форма оператора

Теорема 145 (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 $f(V_0^{(p)}) = V_0^{(p-1)}$, $f(e_1), \ldots, f(e_q) \in V_0^{(p-1)}$. 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:

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

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 $L_1 = \langle \tilde{e}_1, \dots, \tilde{e}_p \rangle$ — 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 $r_0(\lambda) - r_1(\lambda) = N_1(\lambda) + N_2(\lambda) + N_3(\lambda) + \dots$.

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. $r_1(\lambda) - r_2(\lambda) = N_2(\lambda) + N_3(\lambda) + \dots$. Аналогично получаем, что $r_2(\lambda) - r_3(\lambda) = N_3(\lambda) + N_4(\lambda) + \dots$ и т.д., $r_{i-1}(\lambda) - r_i(\lambda) = N_i(\lambda) + N_{i+1}(\lambda) + \dots$. 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. $\square$

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

Овеществление и комплексификация

Овеществление

Определение 146. 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 $\dim$ напоминает, над каким полем мы рассматриваем размерность пространства).

Определение 147. 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 148. 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 149. *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. □

Комплексификация

Определение 150. 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 151. 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.

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 152. *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 $\beta \neq 0$). $\square$

Операторы в евклидовых и эрмитовых пространствах

Сопряженный оператор

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

Определение 153. 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 154. *Если операторы f_1 и f_2 имеют сопряженные f_1^* и f_2^* соответственно, то операторы $h = f_1+f_2$ и $g = f_1f_2$ 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. $(f_1f_2(a), b) = (f_2(a), f_1^*(b)) = (a, f_2^*f_1^*(b))$. $\square$

Лемма 155. *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 $e_1, \ldots, e_n$ есть $A = \begin{pmatrix} a_1^1 & \ldots & a_n^1 \\ \vdots & \ddots & \vdots \\ a_1^n & \ldots & a_n^n \end{pmatrix}$, а матрица оператора f^* (в том же базисе) есть $B = \begin{pmatrix} b_1^1 & \ldots & b_n^1 \\ \vdots & \ddots & \vdots \\ b_1^n & \ldots & b_n^n \end{pmatrix}$. Тогда $a_i^j = (e_j, a_i^k e_k) = (e_j, f(e_i)) = (f^*(e_j), e_i) = (b_j^l e_l, e_i) = \overline{b}_j^i$, откуда получаем $A = \overline{B}^t$, или $B = \overline{A}^t$. □

Отметим, что условие ортонормированности базиса в лемме является существенным.

Лемма 156. *Для любого оператора f существует ему сопряженный f^*.*

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

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

и легко видеть, что

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

т.е. $g = f^*$. □

Следствие 157. *Если f^* — оператор, сопряженный к f, то $(f^*)^* = f$.*

Лемма 158. *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 $v \in V^\perp$. □

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

Заметим, что для матриц операторов, *записанных в ортонормированном базисе*, 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 160. *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 161. *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 162. *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 163. *Пусть λ_1, λ_2 — различные собственные значения самосопряженного или кососимметрического оператора, а v_1, v_2 — соответствующие собственные векторы. 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. $\lambda_2 - \lambda_1$; если f кососимметричен, то собственные значения число мнимы и знак "минус", т.е. $-\lambda_2 + \lambda_1$. В обоих случаях это выражение отлично от нуля, следовательно, нулю равен первый сомножитель, $(v_1, v_2) = 0$. $\square$

Лемма 164. *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 $f^* = \pm f$. $\square$

Канонический вид матрицы самосопряженного оператора

Теорема 165. *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. Приведем два доказательства существования.

1. Доказательство индукцией 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 длины 1, $L = \langle v \rangle$, тогда $V = L \oplus L^{\perp}$, причем L и $L^{\perp}$ оба инвариантны относительно f. Выберем базис $e_2, \ldots, e_{n+1}$ в пространстве $L^{\perp}$, тогда в ортонормированном базисе $v, e_2, \ldots, e_{n+1}$ 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 пусть L — одно из них. Выберем ортонормированные базисы e_1, e_2 в L и $e_3, \ldots, e_{n+1}$ в $L^{\perp}$. Тогда в базисе $e_1, \ldots, e_{n+1}$ матрица оператора имеет вид $A_f = \left(\begin{array}{cc|c} a_{11} & a_{12} & \multirow{2}{*}{$\mathbf{0}$} \\ a_{21} & a_{22} & \\ \hline \multicolumn{2}{c|}{\mathbf{0}} & A_{f'} \end{array}\right)$. По предположению индукции матрица ограничения f' на ортогональное дополнение к инвариантному подпространству имеет искомый вид. Осталось лишь разобраться с двумерной клеткой $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Т.к. $A^t = A$, то $a_{21} = a_{12}$. Для дальнейшего доказательства нам потребуется

Лемма 166. *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 $D = (a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}^2) = (a_{11} - a_{22})^2 + 4a_{12}^2 \geqslant 0$. $\square$

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.

Приведем теперь второе доказательство.

2. Доказательство сведением к квадратичным функциям (только для вещественных линейных пространств). Пусть A — симметричная матрица оператора f в некотором ортонормированном базисе. Рассмотрим квадратичную функцию Q с этой же матрицей в том же базисе. Тогда существует ортонормированный базис, в котором матрица функции Q имеет диагональный вид $\widetilde{A}$, т.е. существует такая ортогональная матрица C, что $\widetilde{A} = C^t A C$. Но ее ортогональность означает, что $C^t = C^{-1}$, поэтому можно записать

$\widetilde{A} = C^{-1}AC$. Но по этой формуле изменяется матрица оператора, поэтому в том базисе, в котором диагональна матрица квадратичной функции Q, матрица оператора f тоже будет диагональной.

Единственность следует из того, что на диагонали находятся корни характеристического многочлена с учетом их кратности. $\square$

Канонический вид матрицы кососимметрического оператора

Теорема 167. *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), $A_f = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$. $\square$

Теорема 168. *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: если канонический вид матрицы состоит из двумерных блоков $\begin{pmatrix} 0 & a_1 \\ -a_1 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & a_k \\ -a_k & 0 \end{pmatrix}$ и $n-2k$ нулевых одномерных блоков, то корни характеристического многочлена — это числа $\pm ia_1, \ldots, \pm ia_k$ и 0 кратности $n-2k$. Это показывает независимость чисел k и $a_1, \ldots, a_k$ от выбора базиса. $\square$

Изометрии

Определение 169. 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 170. *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 171. *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. □

Ортогональные и унитарные операторы

Определение 172. Оператор, сохраняющий скалярное произведение в евклидовых пространствах называется *ортогональным*, в эрмитовых пространствах — *унитарным.*

Лемма 173. *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 $v \in L$. 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. □

Лемма 174. *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 175. *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 176. *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 (в подходящем базисе) 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 $\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$, которая, 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. □

Неотрицательные и положительные операторы. Извлечение неотрицательного квадратного корня из оператора

Определение 177. Оператор $f : V \to V$ называется *неотрицательным*, если
1) $f^* = f$,
2) $(fv, v) \geqslant 0$ для любого $v \in V$

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

Лемма 178. *Оператор f является неотрицательным $\iff$ в некотором ортонормированном базисе его матрица диагональна с неотрицательными элементами на диагонали.*

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

Лемма 179 (извлечение квадратного корня)**.** *Если f — неотрицательный оператор, то существует (и притом единственный) такой неотрицательный оператор g, что $g^2 = f$.*

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

матрицей $A_g = \begin{pmatrix} \sqrt{\lambda_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{\lambda_n} \end{pmatrix}$ в этом же базисе, тогда $g^2 = f$.

Единственность. Пусть наряду с оператором g, построенным выше, существует еще один такой неотрицательный оператор h, что $h^2 = f$. В некотором базисе $a_1, \ldots, a_n$ матрица оператора g имеет вид $A_g = \begin{pmatrix} \sqrt{\lambda_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{\lambda_n} \end{pmatrix}$, а т.к. оператор h неотрицательный, то и его матрица в некотором (вообще говоря другом) базисе $b_1, \ldots, b_n$ имеет вид $A_h = \begin{pmatrix} \mu_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mu_n \end{pmatrix}$. Т.к. λ_i и μ_i^2 — это собственные значения оператора f (с учетом кратности), то μ_i совпадают с $\sqrt{\lambda_i}$ с точностью до перестановки. Без ограничения общности можно считать, что они уже переставлены так, чтобы μ_i в точности совпадали с $\sqrt{\lambda_i}$. Построим изометрию u, переводящую один базис в другой, $u(a_i) = b_i$, тогда

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

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

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

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

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

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

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

$$\lambda_i u_i^j = \lambda_j u_i^j \forall i, j. \tag{5}$$

Точно так же получаем, что

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

Из равенства (5) следует, что $\sqrt{\lambda_i} u_i^j = \sqrt{\lambda_j} u_i^j$ для всех i, j, т.е. $ug = gu$. □

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

Лемма 180. *Оператор f неотрицателен тогда и только тогда, когда $f = h^*h$ для некоторого оператора h.*

Доказательство.
Если f неотрицателен, то в качестве оператора h можно взять квадратный корень из f.
Если $f = h^*h$, то $f^* = (h^*h)^* = h^*(h^*)^* = h^*h = f$, т.е. f — самосопряженный. Для любого вектора a имеем $(fa, a) = (h^*ha, a) = (ha, ha) \geqslant 0$, следовательно, f — неотрицательный. $\square$

Определение 181. Оператор f называется *положительным*, если он неотрицателен и обратим.

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

Лемма 182. *Если оператор $f : V \to V$ положителен, то формула $\langle a, b\rangle := (fa, b)$, где $a, b \in V$, задает (новое) скалярное произведение $\langle\cdot,\cdot\rangle$ в евклидовом/эрмитовом пространстве V со скалярным произведением $(\cdot,\cdot)$.*

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

Полярное разложение операторов

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

Теорема 183 (о полярном разложении)**.** *Для любого оператора $f : W \to W$ существует разложение в произведение двух операторов $f = uh$, где u — унитарный (или ортогональный), а h — неотрицательный. Причем оператор h определен однозначно, а если f обратим, то и u определен однозначно.*

Доказательство.
Доказательство приведем только для случая, когда f обратим.
Существование. Т.к. f обратим, то и f^* обратим, а, следовательно, и f^*f обратим. Более того, f^*f — неотрицательный, следовательно, он положительный. Пусть $h = (f^*f)^{1/2}$, тогда h — неотрицательный и обратимый, т.к. $h^2 = f^*f$, т.е. существует самосопряженный (и даже положительный) оператор h^{-1}. Рассмотрим оператор $u = fh^{-1}$. Тогда $u^* = (h^{-1})^*f^* = h^{-1}f^*$ и $u^*u = h^{-1}f^*fh = h^{-1}h^2h^{-1} = id$, следовательно, u — унитарный (или ортогональный) оператор. Мы получили искомое разложение $f = uh$.
Единственность. Пусть $f = uh = vk$, где u, v — унитарные (ортогональные), а h, k — неотрицательные операторы. Тогда имеем $h^2 = f^*f = (vk)^*vk = k^*v^*vk = k^*k = k^2$, т.е. h и k — это неотрицательные квадратные корни из f^*f, следовательно, $k = h$ (здесь мы не пользуемся обратимостью f), и, значит, $u = v = fh^{-1}$.

Канонический вид матрицы нормального оператора

Определение 184. *Оператор f называется нормальным, если он коммутирует со своим сопряженным, т.е. $ff^* = f^*f$.*

Самосопряженные, кососимметрические, ортогональные и унитарные операторы являются нормальными. Жорданова клетка задает оператор, который не является нормальным.

Теорема 185. *Оператор f, действующий в эрмитовом пространстве W, является нормальным тогда и только тогда, когда в некотором ортонормированном базисе его матрица диагональна.*

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

Теперь перейдем к доказательству существования ортонормированного базиса, в котором матрица нормального оператора диагональна. Пусть λ — собственное значение оператора f, $V_\lambda = \{v \in W : fv = \lambda v\} \subset W$ - подпространство собственных векторов, отвечающих λ. Подпространство V_λ, очевидно, инвариантное относительно f, будет инвариантно также относительно f^*. Действительно, возьмем произвольный $v \in V_\lambda$. Тогда $f^*fv = f^*\lambda v = \lambda f^*v$, но с другой стороны, $f^*fv = ff^*v$, поэтому $f(f^*v) = \lambda f^*v$, следовательно, вектор $w = f^*v$ является собственным (т.к. $f(w) = \lambda w$), отвечающим тому же собственному значению λ, поэтому $w = f^*v \in V_\lambda$, что и доказывает инвариантность V_λ относительно f^*.

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

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

Отдельно надо рассмотреть случай, когда $V_\lambda = W$. В этом случае $f = \lambda \cdot id$, и его матрица в любом базисе диагональна.

$\square$

Лемма 186. *Пусть f — нормальный оператор, а оператор g коммутирует с f. Тогда g коммутирует и с f^*.*

Доказательство. Запишем матрицы всех трех операторов в ортонормированном базисе, в котором матрица оператора f диагональна:

$$A_f = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \; A_{f^*} = \begin{pmatrix} \overline{\lambda}_1 & & \\ & \ddots & \\ & & \overline{\lambda}_n \end{pmatrix}, \; A_g = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Коммутирование операторов равносильно коммутированию их матриц. Условие коммутирования A_f и A_g можно записать в виде $a_{ij}(\lambda_i - \lambda_j) = 0$ для всех $i, j = 1, \ldots, n$, а условие коммутирования A_{f^*} и A_g имеет вид $a_{ij}(\overline{\lambda}_i - \overline{\lambda}_j) = 0$ для всех $i, j = 1, \ldots, n$. Если $a_{ij} \neq 0$, то условие $\lambda_i - \lambda_j = 0$ равносильно условию $\overline{\lambda}_i - \overline{\lambda}_j = 0$.

□

Одновременное приведение двух коммутирующих нормальных операторов к каноническому виду

Теорема 187. *Пусть f и g — нормальные операторы в эрмитовом пространстве W. Ортонормированный базис, в котором матрицы обоих операторов диагональны, существует тогда и только тогда, когда f и g коммутируют.*

Доказательство. Если требуемый базис существует, то операторы коммутируют, т.к. коммутируют их матрицы в этом базисе. Обратно, предположим, что $fg = gf$. Тогда $fg^* = g^*f$. Пусть $V_\lambda = \{v \in W : fv = \lambda v\}$ — собственное подпространство. Мы уже знаем, что оно инвариантно для f, и его ортогональное дополнение $V_\lambda^\perp$ — тоже (т.к. V_λ инвариантно для f^*).

Пусть $v \in V_\lambda$. Поскольку $fg = gf$, $f(g(v)) = g(f(v)) = g(\lambda v) = \lambda g(v)$, т.е. $g(v) \in V_\lambda$, т.е. V_λ инвариантно для g. Аналогично, поскольку $fg^* = g^*f$, подпространство V_λ инвариантно также для g^*, откуда следует, что $V_\lambda^\perp$ инвариантно для g. Итак, V_λ и $V_\lambda^\perp$ инвариантны как для f, так и для g. Это позволяет доказать утверждение по индукции. База (в размерности 1) очевидна. Для индукционного перехода заметим, что если взять в качестве базиса W объединение ортонормированных базисов V_λ и $V_\lambda^\perp$, то матрицы операторов f и g будут иметь вид $A_f = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}$ и $A_g = \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}$, где матрицы B_1, C_1 — матрицы ограничения операторов на V_λ, а B_2, C_2 — матрицы ограничения операторов на $V_\lambda^\perp$. Ограничения нормальных операторов на инвариантные (для самого оператора и для его сопряженного) подпространства являются также нормальными, поэтому для них по предположению индукции теорема верна.

Отдельно надо рассмотреть случай, когда $V_\lambda = W$. В этом случае $f = \lambda \cdot id$, и его матрица в любом базисе диагональна. В качестве искомого базиса возьмем тот, в котором матрица оператора g диагональна.

□

Точно так же доказывается аналогичное утверждение о самосопряженных операторах в евклидовых пространствах:

Теорема 188. *Пусть f и g — самосопряженные операторы в евклидовом пространстве W. Ортонормированный базис, в котором матрицы обоих операторов диагональны, существует тогда и только тогда, когда f и g коммутируют.*

Доказательство. Самосопряженные операторы в евклидовом пространстве обладают собственными векторами (а больше от эритовых пространств ничего и не нужно для доказательства предыдущего утверждения).

□

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:

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

Определение 189. *Тензором* называется полилинейная функция. Тензор типа (p, q) — это полилинейная функция от p векторов из V и от q линейных функций из V'. Тензором типа $(0, 0)$ называют скаляры.

Примеры:

1) скаляры — это тензоры типа $(0, 0)$ по поределению;

2) векторы — это тензоры типа $(0, 1)$, т.к. любой вектор задает линейную функцию на V' переводящую элемент $f \in V'$ в $f(v) \in \mathbb{K}$;

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

4) билинейные функции — это тензоры типа $(2, 0)$, т.к. они задают отображение $V \times V \to \mathbb{K}$.

5) линейные операторы — это тензоры типа $(1, 1)$, этот пример мы обсудим позже.

Отметим, что в силу канонического изоморфизма $V \cong V''$ роль векторного пространства V и его двойственного пространства V' в определении тензора симметричны.

Множество тензоров типа (p, q) мы будем обозначать Θ_p^q. На этом множестве можно ввести структуру линейного пространства, для этого нужно определить операции сложения двух тензоров и их умножения на скаляры. Пусть $T, S \in \Theta_p^q$, тогда:

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

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

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

Так определенное умножение тензоров дистрибутивно (т.е. $(T + \lambda S) \otimes R = T \otimes R + \lambda S \otimes R$), ассоциативно (т.е. $(T \otimes S) \otimes R = T \otimes (S \otimes R)$), но не коммутативно.

Примеры:

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

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

Координатное определение тензоров

Перейдем теперь к координатному описанию тензоров. Зафиксируем базис $e_1, \dots, e_n$ в пространстве V, ему будет соответствовать двойственный базис $\varepsilon^1, \dots, \varepsilon^n$ в двойственном пространстве V'. Т.к. тензор — это полилинейная функция, то ее значение определяется значениями на базисных векторах, т.е.

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

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

Обозначим $T(e_{i_1}, \dots, e_{i_p}, \varepsilon^{j_1}, \dots, \varepsilon^{j_q}) = T^{j_1,\dots,j_q}_{i_1,\dots,i_p} \in \mathbb{K}$. Поэтому, зафиксировав базис, мы можем поставить тензору T в соответствие набор чисел $T^{j_1,\dots,j_q}_{i_1,\dots,i_p}$ — 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 $D = (d^i_k) = C^{-1}$, $d^j_k c^k_i = \delta^j_i$, т.е. $\widetilde{\varepsilon}^i = d^i_k \varepsilon^k$. Тогда

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

Примеры:

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

2) если f — линейная функция, то $\widetilde{f}_k = c^i_k f_i$. 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*:

$$\widetilde{T}^{l_1,\dots,l_q}_{k_1,\dots,k_p} = T^{j_1,\dots,j_q}_{i_1,\dots,i_p} \cdot c^{i_1}_{k_1} \cdot \ldots \cdot c^{i_p}_{k_p} \cdot d^{l_1}_{j_1} \cdot \ldots \cdot d^{l_q}_{j_q}.$$

Теперь мы можем дать другое (координатное) определение тензору: тензор - это набор чисел $T^{j_1,\dots,j_q}_{i_1,\dots,i_p}$, 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 $T(e_{i_1}, \dots, e_{i_p}, \varepsilon^{j_1}, \dots, \varepsilon^{j_q}) = T^{j_1,\dots,j_q}_{i_1,\dots,i_p}$.

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. $a^{ij}a_{jk} = \delta^i_k$.

Лемма 190. *a^{ij} является тензором типа $(0,2)$, т.е. этот набор чисел при переходе к другому базису изменяется по формуле $\widetilde{a}^{kl} = a^{ij} d^k_i d^l_j$.*

Доказательство. В новом базисе выполнено равенство $\widetilde{a}^{kl}\widetilde{a}_{lm} = \delta^k_m$. Подставим сюда $\widetilde{a}_{lm} = a_{ji}c^j_l c^i_m$, где $C = (c^i_m)$ — матрица перехода, тогда $\widetilde{a}^{kl}a_{ji}c^j_l c^i_m = \delta^k_m$. Умножим (и просуммируем по повторяющимся индексам) обе части этого равенства на элементы обратной матрицы к матрице перехода (напомним, что $c^k_i d^j_k = \delta^j_i$): $\widetilde{a}^{kl}a_{ji}c^j_l c^i_m d^m_p = \delta^k_m d^m_p$, т.е. $\widetilde{a}^{kl}a_{ji}c^j_l \delta^i_p = d^k_p$, или $\widetilde{a}^{kl}a_{jp}c^j_l = d^k_p$. Воспользуемся теперь обратимостью матрицы (a_{ij}), т.е. тем, что $a_{jp}a^{pr} = \delta^r_j$. Умножив обе части предпоследнего равенства на a^{pr} (и просуммировав), получим $\widetilde{a}^{kl}a_{jp}a^{pr}c^j_l = a^{pr}d^k_p$, или $\widetilde{a}^{kl}c^r_l = a^{pr}d^k_p$. Еще раз умножив обе части равенства на d^n_r просуммировав по r и воспользовавшись тем, что $c^r_l d^n_r = \delta^n_l$ и $\widetilde{a}^{kl}\delta^n_l = \widetilde{a}^{kn}$, получим $\widetilde{a}^{kn} = a^{pr}d^k_p d^n_r$. □

Базис в пространстве тензоров

Построим базис в пространстве тензоров Θ^q_p. Векторы и ковекторы — это тензоры типа $(0,1)$ и $(1,0)$ соответственно. Рассмотрим произведение $\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}$, оно состоит из тензоров типа $(0,1)$ и тензоров типа $(1,0)$, следовательно само является тензором типа (p,q). Всего таких различных произведений получится n^{p+q}, т.к. из n ковекторов надо выбрать p и из n векторов надо выбрать q. Докажем, что эти элементы (произведения такого вида) образуют базис в Θ^q_p. Прежде, чем доказывать это, вычислим значение такого тензора (произведения) на наборах аргументов из базисных векторов:

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

Лемма 191. *Множество произведений вида* $\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}$ *является базисом в* Θ^q_p.

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

$$
\lambda^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}\underbrace{\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}(e_{k_1},\ldots,e_{k_p},\varepsilon^{l_1},\ldots,\varepsilon^{l_q})} = 0 \tag{6}
$$

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 (6) может быть записано в виде

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

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

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

$$
T = T^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}. \tag{7}
$$

А из-за полилинейности это равенство достаточно проверять на базисных аргументах вида $e_{k_1},\ldots,e_{k_p},\varepsilon^{l_1},\ldots,\varepsilon^{l_q}$. Левая часть равенства по определению равна $T^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}$, а правая часть, как мы уже видели раньше, также равна $T^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}(e_{k_1},\ldots,e_{k_p},\varepsilon^{l_1},\ldots,\varepsilon^{l_q}) = T^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}$. Итак, равенство проверено, т.е. для произвольного тензора T мы нашли его разложение в линейную комбинацию (7), причем числа $T^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}$ являются координатами этого тензора в указанном базисе. $\square$

Линейные операторы как тензоры

Перейдем теперь к обещанному отождествлению линейных операторов и тензоров типа $(1,1)$.

Утверждение 192. *Имеет место канонический изоморфизм* $\Theta^1_1 \cong L(V)$.

Доказательство. Пусть дан линейный оператор $f: V\to V$, нам надо по нему построить билинейную функцию $T_f = T_f(v,\varphi)$ с одним векторным ($v\in V$) и одним ковекторным ($\varphi\in V'$) аргументами. Определим его равенством $T_f(v,\varphi) := \varphi(f(v))\in\mathbb{K}$. T_f действительно является полилинейным отображением от двух аргументов. Соответствие $f\mapsto T_f$ задает линейное отображение $L(V)\to\Theta^1_1$, конструкция которого не зависит от выбора базисов. Доказать, что отображение $L(V)\to\Theta^1_1$ будет изоморфизмом, можно двумя способами.

Первый — это просто записать все в координатах и сравнить.

Второй способ. Размерности пространств $L(V)$ и Θ^1_1 совпадают (они равны $(\dim V)^2$), поэтому надо проверить лишь то, что у этого отображения нулевое ядро, т.е. что если $T_f = 0$, то и $f=0$. Пусть $T_f(v,\varphi)=0$ для любых $v\in V$, $\varphi\in V'$, т.е. для любого φ $\varphi(f(v))=0$, следовательно, $f(v)=0$. Но, т.к. это верно для всех векторов $v\in V$, то имеем $f=0$. $\square$

Теперь мы можем отождествлять линейные операторы и тензоры типа $(1,1)$.

Например, тождественный оператор задается в любом базисе матрицей (δ^j_i). Проверим в явном виде, что символ Кронекера δ^j_i является тензором. Запишем тензорный закон изменения координат $\widetilde{\delta}^k_l = \delta^j_i c^i_l d^k_j = c^i_l d^k_i$, что, в свою очередь, равно $c^i_l d^k_i = \delta^k_l$, т.к. $D = C^{-1}$ и $CD = E$. Значит, символ Кронекера действительно является тензором типа $(1,1)$.

Свертка тензоров

Пусть $T\in\Theta^q_p$ — тензор с хотя бы одним нижним и одним верхним индексами, т.е. $p,q>0$. Пусть $e_1,\ldots,e_n$ — базис в V. Зафиксируем один векторный и один ковекторный аргумент (пусть это будут первые по порядку аргументы), на их место поставим базисные элементы e_i и ε^i и определим полилинейную функцию sT от $p-1$ векторых и $q-1$ ковекторных аргументов по формуле

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

в которой подразумевается суммирование по индексу i. Проверим, что определение тензора sT не зависит от выбора базиса, т.е. что его координаты будут преобразовываться по тензорному закону. Имеем $(sT)^{i,j_2,\ldots,j_q}_{i,i_2,\ldots,i_p}$ (здесь опять производится суммирование по индексу i). При переходе к другому базису имеем

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

$$(\widetilde{sT})^{l_2,\ldots,l_q}_{k_2,\ldots,k_p} = \widetilde{T}^{k,l_2,\ldots,l_q}_{k,k_2,\ldots,k_p} = T^{j_1,j_2,\ldots,j_q}_{i_1,i_2,\ldots,i_p}\cdot \underbrace{c^{i_1}_k}\cdot c^{i_2}_{k_2}\cdot\ldots\cdot c^{i_p}_{k_p}\cdot \underbrace{d^k_{j_1}}\cdot d^{l_2}_{j_2}\cdot\ldots\cdot d^{l_q}_{j_q}.$$

Произведение подчеркнутых элементов равно $c_k^{i_1} d_{j_1}^k = \delta_{j_1}^{i_1}$, 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) 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 $s(g \otimes a \otimes b) = g_{ij} a^i b^j = g(a,b)$.

Поднятие и опускание индексов

Пусть нам дано евклидово пространство V, т.е. векторное пространство над $\mathbb{R}$ со скалярным произведением $(\cdot,\cdot)$. Тогда существует канонический изоморфизм $V \cong V'$, поэтому у любого тензора можно заменить векторный аргумент на ковекторный и наоборот. Например, возьмем тензор типа $(0,1)$, т.е. вектор $T : V' \to \mathbb{R}$. Отождествим его с ковектором $(T, \cdot) : V \to \mathbb{R}$, тогда пполучим канонический изоморфизм пространств тензоров Θ_0^1 и Θ_1^0. В координатах это выглядит следующим образом. Пусть $G = (g_{ij})$ — матрица Грама скалярного произведения $(\cdot,\cdot)$. Каждому тензору T с координатами T^i поставим в соответствие $T_j = g_{ij} T^i$ (это произведение двух тензоров, g и T и свертка по индексу i). Таким образом мы у тензора T опустили индекс.

Обобщая эту операцию, дадим определение операции опускания индекса.

Опускание индекса — это отображение $\Theta_p^q \to \Theta_{p+1}^{q-1}$, которое тензору $T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}$ 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 $g^{ij} \cdot T_{i,i_2,\ldots,i_p}^{j_1,\ldots,j_q} \in \Theta_{p-1}^{q+1}$.

Оператор альтернирования. Внешние формы

Рассмотрим линейное пространство Θ_p^0 тензоров с одними нижними индексами, т.е. полилинейные функции от p векторов. Также рассмотрим группу перестановок S_p. Если взять какую-нибудь перестановку $\sigma \in S_p$, то можно определить линейный оператор $f_\sigma : \Theta_p^0 \to \Theta_p^0$ следующим образом. Пусть $T \in \Theta_p^0$, т.е. $T = T(v_1, \ldots, v_p)$; 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 $(\sigma T)_{i_1,\ldots,i_p} = (\sigma T)(e_{i_1}, \ldots, e_{i_p}) = T(e_{i_{\sigma(1)}}, \ldots, e_{i_{\sigma(p)}}) = T_{i_{\sigma(1)}}, \ldots, i_{\sigma(p)}$.

Поскольку нам вскоре понадобится делить на целые числа, с этого момента, говоря о тензорах, будем считать, что характеристика поля $\mathbb{K}$ нулевая. При желании для простоты можно считать, что $\mathbb{K} = \mathbb{R}$.

Построим оператор *альтернирования* (приводящий к аналогу свойства кососимметричности)

$$\mathrm{Alt} : \Theta^0_p \to \Theta^0_p; \quad \mathrm{Alt}\, T := \frac{1}{p!} \sum_{\sigma \in S_p} (-1)^\sigma \sigma T.$$

Этот оператор будет линейным, т.к. является суммой линейных операторов. Назовем тензор $T \in \Theta^0_p$ *кососимметрическим* (или *внешней формой*), если $\sigma T = (-1)^\sigma T$ для любой перестановки $\sigma \in S_p$. В пространстве всех тензоров с нижними индексами определим подпространство $\Lambda_p \subset \Theta^0_p$ 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 193. *Оператор* Alt *является оператором проектирования на подпространство внешних форм* Λ_p.

Доказательство. Нам потребуются следующие равенства:

Утверждение 194. $\sigma(\mathrm{Alt}\, T) = \mathrm{Alt}(\sigma T) = (-1)^\sigma \mathrm{Alt}\, T$.

Доказательство. Применим перестановку σ к тензору $\mathrm{Alt}\, T$:

$$\sigma(\mathrm{Alt}\, T) = \sigma\Big(\frac{1}{p!} \sum_{\rho \in S_p} (-1)^\rho \rho T\Big) = \frac{1}{p!} \sum_{\rho \in S_p} (-1)^\rho ((\sigma\rho) T),$$

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: $\sigma(\mathrm{Alt}\, T) = \frac{1}{p!} \sum_{\tau \in S_p} (-1)^\rho \tau T$. А, поскольку $(-1)^\tau = (-1)^\rho (-1)^\sigma$, то

$$\sigma(\mathrm{Alt}\, T) = \frac{1}{p!} \sum_{\tau \in S_p} (-1)^\sigma (-1)^\tau \tau T = (-1)^\sigma \Big(\frac{1}{p!} \sum_{\tau \in S_p} (-1)^\tau \tau T\Big) = (-1)^\sigma \mathrm{Alt}\, T,$$

т.е. мы доказали, что $\sigma(\mathrm{Alt}\, T) = (-1)^\sigma \mathrm{Alt}\, T$. Теперь докажем, что $\mathrm{Alt}(\sigma T) = (-1)^\sigma \mathrm{Alt}\, T$.

По определению $\mathrm{Alt}(\sigma T) = \frac{1}{p!} \sum_{\rho \in S_p} (-1)^\rho ((\rho\sigma) T)$. Обозначим $\tau = \rho\sigma$ и получим

$$\begin{aligned} \mathrm{Alt}(\sigma T) &= \frac{1}{p!} \sum_{\rho \in S_p} (-1)^\rho ((\rho\sigma) T) = \frac{1}{p!} \sum_{\tau \in S_p} (-1)^\rho (\tau T) = \frac{1}{p!} \sum_{\tau \in S_p} (-1)^\tau (-1)^\sigma (\tau T) = \\ &= (-1)^\sigma \Big(\frac{1}{p!} \sum_{\tau \in S_p} (-1)^\tau (\tau T)\Big) = (-1)^\sigma \mathrm{Alt}\, T. \end{aligned}$$

□

Перейдем теперь собственно к доказательству леммы.

1. Проверим, что $\mathrm{Im}\, \mathrm{Alt} \subset \Lambda_p$. Действительно, поскольку $\sigma(\mathrm{Alt}\, T) = (-1)^\sigma \mathrm{Alt}\, T$, то по определению $\mathrm{Alt}\, T \in \Lambda_p$ для любого $T \in \Theta^0_p$, поэтому $\mathrm{Im}\, \mathrm{Alt} \subset \Lambda_p$.

2. Докажем, что если $T \in \Lambda_p$, то $\mathrm{Alt}\, T = T$. Действительно, поскольку $T \in \Lambda_p$, то $\sigma T = (-1)^\sigma T$ и

$$\mathrm{Alt}\, T = \frac{1}{p!} \sum_{\sigma \in S_p} (-1)^\sigma \sigma T = \frac{1}{p!} \sum_{\sigma \in S_p} (-1)^\sigma (-1)^\sigma T = \frac{1}{p!} \sum_{\sigma \in S_p} T = \frac{1}{p!} p! T = T.$$

3. Проверим, что $\mathrm{Alt}^2 = \mathrm{Alt}$, т.е. что $\mathrm{Alt}(\mathrm{Alt}\, T) = \mathrm{Alt}\, T$ для любого $T \in \Theta^0_p$. Действительно, в п.1 мы доказали, что $S = \mathrm{Alt}\, T \in \Lambda_p$, в п.2 — что $\mathrm{Alt}\, S = S$, подставив $\mathrm{Alt}\, T$ вместо S, получим $\mathrm{Alt}(\mathrm{Alt}\, T) = \mathrm{Alt}\, T$. □

Внешнее умножение, его свойства

Определим аналог тензорного умножения для внешних форм — *внешнее тензорное умножение* (обозначается $\wedge$): для $T \in \Lambda_p$, $S \in \Lambda_q$ положим $T \wedge S := \mathrm{Alt}(T \otimes S)$.

Лемма 195. *Введенное нами внешнее тензорное умножение обладает следующими свойствами: для любых внешних форм $T \in \Lambda_p$, $S \in \Lambda_q$, $R \in \Lambda_r$*
1) $(T + \lambda S) \wedge R = T \wedge R + \lambda S \wedge R$ (дистрибутивность);
2) $S \wedge T = (-1)^{pq} T \wedge S$ (антикоммутативность) ;
3) $(T \wedge S) \wedge R = T \wedge (S \wedge R)$ (ассоциативность).

Доказательство. 1) Дистрибутивность следует из дистрибутивности операции $\otimes$ и линейности оператора Alt.

2) По определению

$$S \wedge T = \mathrm{Alt}(S \otimes T) = \frac{1}{(p+q)!} \sum_{\sigma \in S_{p+q}} \sigma(S \otimes T);$$
$$T \wedge S = \mathrm{Alt}(T \otimes S) = \frac{1}{(p+q)!} \sum_{\sigma \in S_{p+q}} \sigma(T \otimes S).$$

Рассмотрим координаты тензоров $\sigma(S \otimes T)$ и $\sigma(T \otimes S)$. Имеем

$$\begin{aligned} \sigma(S \otimes T)_{i_1,\ldots,i_{p+q}} &= (S \otimes T)_{i_{\sigma(1)},\ldots,i_{\sigma(p+q)}} = \\ &= S_{i_{\sigma(1)},\ldots,i_{\sigma(q)}} \cdot T_{i_{\sigma(q+1)},\ldots,i_{\sigma(p+q)}}; \qquad (8) \\ \sigma(T \otimes S)_{i_1,\ldots,i_{p+q}} &= (T \otimes S)_{i_{\sigma(1)},\ldots,i_{\sigma(p+q)}} = \\ &= T_{i_{\sigma(1)},\ldots,i_{\sigma(p)}} \cdot S_{i_{\sigma(p+1)},\ldots,i_{\sigma(p+q)}}. \qquad (9) \end{aligned}$$

Посмотрим, чем отличаются индексы у S и T в выражениях (8) и (9). Индексы в (8) — это $\sigma(1), \ldots, \sigma(q), \sigma(q+1), \ldots, \sigma(p+q)$, а индексы в (9) — это $\sigma(p+1), \ldots, \sigma(p+q), \sigma(1), \ldots, \sigma(q)$. 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

$$\begin{aligned} S \wedge T &= \frac{1}{(p+q)!} \sum_{\sigma \in S_{p+q}} (-1)^{\sigma} \sigma(S \otimes T) = \\ &= \frac{1}{(p+q)!} \sum_{\sigma \in S_{p+q}} (-1)^{\tau} (-1)^{\sigma\tau} (\sigma\tau)(T \otimes S) = \\ &= (-1)^{\tau} \frac{1}{(p+q)!} \sum_{\rho \in S_{p+q}} (-1)^{\rho} \rho(T \otimes S) = (-1)^{\tau} T \wedge S, \end{aligned}$$

и нам осталось определить $(-1)^{\tau}$. Для вычисления знака перестановки надо подсчитать количество элементарных перестановок, ее составляющих. Лекго видеть, что это число равно произведению pq, т.е. $(-1)^{\tau} = (-1)^{pq}$, что и требовалось показать.

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

Утверждение 196. $\mathrm{Alt}((\mathrm{Alt}\, Q) \otimes R) = \mathrm{Alt}(Q \otimes R) = \mathrm{Alt}(Q \otimes (\mathrm{Alt}\, R))$ *для любых тензоров $Q \in \Theta^0_p$, $R \in \Theta^0_q$.*

Доказательство. Ограничимся доказательством первого из равенств (второе доказывается аналогично). Поскольку операция $\otimes$ обладает свойством дистрибутивности, а оператор Alt линеен, имеем

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

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

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

$\square$

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

Базис в пространстве внешних форм

Построим базис в пространстве кососимметрических тензоров Λ_p. Как нам известно, базисом в пространстве Θ_p^0 являются всевозможные произведения вида $\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p}$. Однако при проектировании на подпространство Λ_p эти произведения переходят в $\varepsilon^{i_1}\wedge\ldots\wedge\varepsilon^{i_p} = \operatorname{Alt}(\varepsilon^{i_1}\otimes\ldots\otimes\varepsilon^{i_p})$ и становятся линейно зависимыми. Действительно, если $i_k = i_l$, то $\varepsilon^{i_k}\wedge\varepsilon^{i_l} = 0$, поэтому из базиса нужно выкинуть все произведения, в которых встречаются повторяющиеся индексы. Далее, т.к. $\varepsilon^{i_k}\wedge\varepsilon^{i_l} = -\varepsilon^{i_l}\wedge\varepsilon^{i_k}$, то нужно также выкинуть все произведения с неупорядоченными индексами и оставить только произведения с индексами упорядоченными, например, по возрастанию, т.е. такие, что $i_1 < \ldots < i_p$. Всего таких упорядоченных произведений будет C_n^p, где $n = \dim V$. Очевидно, что любой тензор является линейной комбинацией тензоров $\varepsilon^{i_1}\wedge\ldots\wedge\varepsilon^{i_p}$, $i_1 < \ldots < i_p$, т.к. из базиса мы выкинули только заведомо линейно зависимые вектора. Поэтому для доказательства того, что эта система векторов действительно является базисом Λ_p достаточно доказать ее линейную независимость.

Пусть линейная комбинация этих тензоров равна нулю,

$$\sum_{i_1<\ldots<i_p} \lambda_{i_1,\ldots,i_p} \varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p} = 0.$$

Перепишем эту сумму в виде

$$\sum_{i_1<\ldots<i_p} \lambda_{i_1,\ldots,i_p} \Big(\frac{1}{p!} \sum_{\sigma\in S_p} (-1)^\sigma \underbrace{\varepsilon^{i_{\sigma(1)}} \otimes \ldots \otimes \varepsilon^{i_{\sigma(p)}}} \Big) = 0. \tag{10}$$

Подчеркнутые элементы — это элементы базиса в пространстве Θ_p^0 и они линейно независимы. Отсюда будет следовать равенство нулю всех коэффициентов, если только все эти подчеркнутые элементы будут различны. Докажем, что все базисные элементы, встречающиеся в этой сумме, различны. Распишем эту сумму подробнее (без общего сомножителя $\frac{1}{p!}$, который на линейную зависимость не влияет)

$$\ldots \pm \lambda_{i_1,\ldots,i_p} \sum_{\sigma\in S_p} \varepsilon^{i_{\sigma(1)}} \otimes \ldots \otimes \varepsilon^{i_{\sigma(p)}} \pm \ldots \pm \lambda_{j_1,\ldots,j_p} \sum_{\rho\in S_p} \varepsilon^{j_{\rho(1)}} \otimes \ldots \otimes \varepsilon^{j_{\rho(p)}} \pm \ldots .$$

Допустим, что существуют такие $\sigma, \rho \in S_p$ и индексы $i_1, \ldots, i_p$, $j_1, \ldots, j_p$, что $\varepsilon^{i_{\sigma(1)}} \otimes \ldots \otimes \varepsilon^{i_{\sigma(p)}} = \varepsilon^{j_{\rho(1)}} \otimes \ldots \otimes \varepsilon^{j_{\rho(p)}}$, т.е. $i_{\sigma(1)} = j_{\rho(1)}, \ldots, i_{\sigma(p)} = j_{\rho(p)}$. Перейдем к перестановке $\tau = \rho\sigma^{-1}$, тогда $i_1 = j_{\tau(1)}, \ldots, i_p = j_{\tau(p)}$. Но индексы i и j у нас упорядочены по возрастанию, следовательно, τ — тождественная перестановка. Значит, $\sigma = \rho$ и тогда $i_1 = j_1, \ldots, i_p = j_p$. Следовательно, все слагаемые в (10) различны, значит все коэффициенты в (10) равны нулю, что и доказывает линейную независимость системы тензоров $\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p}$, $i_1 < \ldots < i_p$, и то, что она является базисом в Λ_p. $\square$