Линейная алгебра и геометрия

Ольга Викторовна Сипачева
Кафедра общей топологии и геометрии
o-sipa@yandex.ru

Векторное, или линейное, пространство над полем K — это тройка $(V, +, \cdot)$, где $+\colon V \times V \to V$ — операция сложения элементов V (сумма $+(\boldsymbol{x}, \boldsymbol{y})$ обозначается $\boldsymbol{x} + \boldsymbol{y}$) и $\cdot\colon K \times V \to V$ — операция умножения элементов V на числа из K (произведение $\cdot, (\boldsymbol{x}, \lambda)$ обозначается $\lambda \boldsymbol{x}$), обладающие следующими свойствами:

1. $\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}$ для любых $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ (ассоциативность);
2. $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$ для любых $\boldsymbol{x}, \boldsymbol{y} \in V$ (коммутативность);
3. в V существует такой элемент $\mathbf{0}$ (нуль), что $\boldsymbol{x} + \mathbf{0} = \boldsymbol{x}$ для любого $\boldsymbol{x} \in V$;
4. для любого элемента $\boldsymbol{x} \in V$ существует такой элемент $-\boldsymbol{x} \in V$ (противоположный элемент), что $\boldsymbol{x} + (-\boldsymbol{x}) = \mathbf{0}$;
5. $\lambda(\boldsymbol{x} + \boldsymbol{y}) = \lambda \boldsymbol{x} + \lambda \boldsymbol{y}$ для любых $\lambda \in K$ и $\boldsymbol{x}, \boldsymbol{y} \in V$;
6. $(\lambda + \mu)\boldsymbol{x} = \lambda \boldsymbol{x} + \mu \boldsymbol{x}$ для любых $\lambda, \mu \in K$ и $\boldsymbol{x} \in V$;
7. $(\lambda\mu)\boldsymbol{x} = \lambda(\mu \boldsymbol{x})$ для любых $\lambda, \mu \in K$ и $\boldsymbol{x} \in V$;
8. $1\boldsymbol{x} = \boldsymbol{x}$ для любого $\boldsymbol{x} \in V$.

Предгильбертовым пространством называется пара $(V, (\cdot,\cdot))$, где V — векторное пространство над $\mathbb{R}$ с фиксированной положительно определенной симметрической билинейной функцией. Билинейная функция $(\cdot,\cdot)$ называется скалярным произведением.

Аксиомы скалярного произведения:

1. $(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{z}) = (\boldsymbol{x},\boldsymbol{z}) + (\boldsymbol{y},\boldsymbol{z})$ для любых $\boldsymbol{x},\boldsymbol{y} \in V$,
2. $(\lambda\boldsymbol{x},\boldsymbol{y}) = \lambda(\boldsymbol{x},\boldsymbol{y})$ для любых $\lambda \in \mathbb{R}$ и $\boldsymbol{x},\boldsymbol{y} \in V$,
3. $(\boldsymbol{x},\boldsymbol{y}) = (\boldsymbol{y},\boldsymbol{x})$ для любых $\boldsymbol{x},\boldsymbol{y} \in V$,
4. $(\boldsymbol{x},\boldsymbol{x}) > 0$ для любого $\boldsymbol{x} \in V$, $\boldsymbol{x} \neq \mathbf{0}$.

В случае конечномерного V такое пространство называется евклидовым.

Аффинным пространством, ассоциированным с векторным пространством V, называется пара $(\mathbb{A}, +)$, где $\mathbb{A}$ — множество (точек) и $+\colon \mathbb{A} \times V \to \mathbb{A}$ — операция прибавления вектора к точке, удовлетворяющая следующим условиям:

1. $A + (\boldsymbol{x} + \boldsymbol{y}) = (A + \boldsymbol{x}) + \boldsymbol{y}$ для любых $A \in \mathbb{A}$ и $\boldsymbol{x}, \boldsymbol{y} \in V$,
2. $A + \mathbf{0} = A$ для любой точки $A \in \mathbb{A}$,
3. для любых точек $A, B \in \mathbb{A}$ существует единственный вектор $\boldsymbol{x} \in V$, для которого $A + \boldsymbol{x} = B$ (его обозначают $\overrightarrow{AB}$).

Векторные пространства

Векторное, или линейное, пространство над полем K — это тройка $(V, +, \cdot)$, где $+ : V \times V \to V$ — операция сложения элементов V (сумма $+(\boldsymbol{x}, \boldsymbol{y})$ обозначается $\boldsymbol{x} + \boldsymbol{y}$) и $\cdot : K \times V \to V$ — операция умножения элементов V на числа из K (произведение $\cdot(\boldsymbol{x}, \lambda)$ обозначается $\lambda\boldsymbol{x}$), для которых выполнены следующие аксиомы:

1. $\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}$ для любых $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ (ассоциативность);
2. $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$ для любых $\boldsymbol{x}, \boldsymbol{y} \in V$ (коммутативность);
3. в V существует такой элемент $\mathbf{0}$ (нуль), что $\boldsymbol{x} + \mathbf{0} = \boldsymbol{x}$ для любого $\boldsymbol{x} \in V$;
4. для любого элемента $\boldsymbol{x} \in V$ существует такой элемент $-\boldsymbol{x} \in V$ (противоположный элемент), что $\boldsymbol{x} + (-\boldsymbol{x}) = \mathbf{0}$;
5. $\lambda(\boldsymbol{x} + \boldsymbol{y}) = \lambda\boldsymbol{x} + \lambda\boldsymbol{y}$ для любых $\lambda \in K$ и $\boldsymbol{x}, \boldsymbol{y} \in V$;
6. $(\lambda + \mu)\boldsymbol{x} = \lambda\boldsymbol{x} + \mu\boldsymbol{x}$ для любых $\lambda, \mu \in K$ и $\boldsymbol{x} \in V$;
7. $(\lambda\mu)\boldsymbol{x} = \lambda(\mu\boldsymbol{x})$ для любых $\lambda, \mu \in K$ и $\boldsymbol{x} \in V$;
8. $1\boldsymbol{x} = \boldsymbol{x}$ для любого $\boldsymbol{x} \in V$.

Элементами векторного пространства $(V, +, \cdot)$ считаются элементы множества V.

Вместо $(V, +, \cdot)$ пишут просто V.

Элементы множества V обычно (и именно в качестве элементов V как векторного пространства) называют векторами. При этом они могут быть одновременно многочленами, или функциями, или классами равных направленных отрезков и т.п. Операции сложения векторов и умножения вектора на число часто называют линейными операциями.

Следствия из аксиом векторного пространства

1. Вектор $\mathbf{0}$ единствен.
2. Противоположный вектор единствен.
3. Для любых $\boldsymbol{a}, \boldsymbol{b} \in V$ уравнение $\boldsymbol{x} + \boldsymbol{a} = \boldsymbol{b}$ имеет единственное решение, равное $\boldsymbol{b} + (-\boldsymbol{a})$. Это решение называется разностью векторов $\boldsymbol{b}$ и $\boldsymbol{a}$ и обозначается $\boldsymbol{b} - \boldsymbol{a}$.
4. Сумма произвольного числа векторов не зависит от расстановки скобок. ($\Longrightarrow$ мы будем часто опускать скобки)
5. $\lambda\mathbf{0} = \mathbf{0}$ для любого $\lambda \in K$.
6. $\lambda(-\boldsymbol{x}) = -\lambda\boldsymbol{x}$ для любых $\lambda \in K$, $\boldsymbol{x} \in V$.
7. $\lambda(\boldsymbol{x} - \boldsymbol{y}) = \lambda\boldsymbol{x} - \lambda\boldsymbol{y}$ для любых $\lambda \in K$, $\boldsymbol{x}, \boldsymbol{y} \in V$.
8. $0\boldsymbol{x} = \mathbf{0}$ для любого $\boldsymbol{x} \in V$.
9. $(-1)\boldsymbol{x} = -\boldsymbol{x}$ для любого $\boldsymbol{x} \in V$.
10. $(\lambda - \mu)\boldsymbol{x} = \lambda\boldsymbol{x} - \mu\boldsymbol{x}$ для любых $\lambda, \mu \in K$ и $\boldsymbol{x} \in V$.

Упражнения

1. Докажите следствия из аксиом векторного пространства.
2. Докажите, что аксиома 2 вытекает из остальных.
3. Выясните, какие ещё аксиомы вытекают из остальных.

Примеры

- Обычные векторы (классы направленных отрезков) в обычном пространстве, рассматриваемом в элементарной геометрии.
- Любое поле K является векторным пространством над полем K относительно операций поля. Любое поле K является также векторным пространством над любым своим подполем F.
- Множество действительных (комплексных, со значениями в поле K) функций с фиксированной областью определения является векторным пространством над полем действительных чисел (над полем комплексных чисел, над полем K). Операции определяются поточечно. Вместо всех функций можно рассматривать непрерывные функции, дифференцируемые функции ...
- Многочлены, многочлены степени, ограниченной некоторым числом ...

Примеры

- Частный случай — множество функций, определённых на $\{1, 2, \ldots, n\}$. Значения любой функции $f\colon \{1, 2, \ldots, n\} \to K$ можно записать в столбец:
$$f = \begin{pmatrix} f(1) \\ f(2) \\ \vdots \\ f(n) \end{pmatrix}.$$
- Арифметическое векторное пространство K^n над полем K — пространство строк или столбцов чисел (элементов K) фиксированного размера n. Операции:
$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot x_1 \\ \lambda \cdot x_2 \\ \ldots \\ \lambda \cdot x_n \end{pmatrix}.$$
- Матрицы.
- Функции от нескольких переменных, многочлены от нескольких переменных.

Примеры

- Пусть X — любое множество, и пусть $\mathscr{P}(X)$ — множество всех его подмножеств. Для $A, B \subset X$ положим $A + B = A \triangle B = (A \cup B) \smallsetminus (A \cap B)$, $0A = \varnothing$ и $1A = A$. Множество $\mathscr{P}(X)$ с так определёнными операциями — векторное пространство над полем $\mathbb{F}_2$. В этом легко убедиться, если вместо множеств $A \subset X$ рассмотреть их характеристические функции $\chi_A \colon X \to \mathbb{F}_2$.

  На $\mathscr{P}(X)$ можно определить также билинейное умножение, положив $A \cdot B = A \cap B$. В результате получится алгебра над полем $\mathbb{F}_2$. Её не следует путать с *булевой алгеброй* — дистрибутивной решёткой (относительно операций $\cup$ и $\cap$) с дополнениями. *Алгеброй множеств* называется именно булева алгебра.

Подпространства

Определение

Непустое подмножество U векторного пространства V над полем K называется подпространством векторного пространства V (обозначение: $U \underset{\text{lin}}{\subseteq} V$), если оно замкнуто относительно линейных операций, т.е.

1. для любых $\boldsymbol{x}, \boldsymbol{y} \in U$ имеем $\boldsymbol{x} + \boldsymbol{y} \in V$ и
2. для любых $\boldsymbol{x} \in U$ и $\lambda \in K$ имеем $\lambda \boldsymbol{x} \in V$.

Подпространство векторного пространства само является векторным пространством относительно тех же операций. Подпространство $U \underset{\text{lin}}{\subseteq} V$ собственное, если $U \neq V$, т.е. $U \underset{\text{lin}}{\subsetneq} V$.

Упражнения

1. Покажите, что в определении подпространства вместо непустоты множества U и выполнения условия 1 можно потребовать, чтобы множество U являлось подгруппой аддитивной группы V — получится равносильное определение.
2. Докажите, что отношение «быть подпространством» транзитивно.

Примеры

- V, $\{\mathbf{0}\}$
- Множество обычных векторов, параллельных данной прямой или плоскости.
- Множество непрерывных функций — подпространство пространства всех функций, множество дифференцируемых функций — подпространство пространства непрерывных функций и пространства всех функций, множество полиномиальных функций (не многочленов!) — подпространство пространств всех дифференцируемых функций, всех непрерывных функций и всех функций ...
- Множество (кольцо) $K[x]$ всех многочленов с коэффициентами из K, рассматриваемое как векторное пространство над K, — подпространство множества $K_{\leqslant n}[x]$ всех многочленов степени, не превосходящей k.
- Множество строк $(x_1, \ldots, x_n)$ чисел, компоненты которых удовлетворяют уравнению вида $\lambda_1 \cdot x_1 + \cdots + \lambda_n \cdot x_n = 0$, — подпространство арифметического векторного пространства K^n .

Линейная зависимость и линейные комбинации векторов

Система векторов — упорядоченный набор векторов, в котором элементы могут повторяться. По традиции записывается как множество с перенумерованными элементами.

Определение

Конечная система векторов $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\} \subset V$ называется линейно зависимой, если найдутся $\{\lambda_1, \ldots, \lambda_m\}$, не все равные нулю, для которых $\lambda \boldsymbol{x}_1 + \ldots \lambda \boldsymbol{x}_m = 0$.

Конечное множество векторов линейно зависимо, если при любой нумерации его элементов получается линейно зависимая система.

Бесконечное подмножество V называется линейно зависимым, если некоторое его непустое конечное подмножество линейно зависимо.

Подмножество, не являющееся линейно зависимым, называется линейно независимым.

Пусть $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m \in V$.

Определение

Всякое выражение вида $\lambda_1\boldsymbol{x}_1 + \dots + \lambda_m\boldsymbol{x}_m$, где $\lambda_1, \dots, \lambda_m \in K$, называется линейной комбинацией векторов $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$. Если не все $\lambda_1, \dots, \lambda_m$ равны нулю, то линейная комбинация $\lambda_1\boldsymbol{x}_1 + \dots + \lambda_m\boldsymbol{x}_m$ называется нетривиальной, в противном случае линейная комбинация называется тривиальной.

Говорят, что вектор $\boldsymbol{x} \in V$ линейно выражается через векторы $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$, если он равен некоторой их линейной комбинации.

Говорят, что вектор $\boldsymbol{x} \in V$ линейно выражается через подмножество $S \subset V$, если он линейно выражается через некоторые векторы $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m \in S$.

Соглашение

Значение пустой линейной комбинации принимается равным нулевому вектору.

Факты из курса алгебры в первом семестре

- Если система векторов содержит линейно зависимую подсистему, то она сама линейно зависима.
- Векторы линейно зависимы тогда и только тогда, когда хотя бы один из них линейно выражается через остальные.
- Пусть векторы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ линейно независимы. Вектор $\boldsymbol{x}$ линейно выражается через $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ тогда и только тогда, когда векторы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m, \boldsymbol{x}$ линейно зависимы.
- Пусть вектор $\boldsymbol{x}$ линейно выражается через векторы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$. Это выражение единственно тогда и только тогда, когда векторы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ линейно независимы.

Определение

Пусть $X \subset V$. Совокупность всевозможных (конечных) линейных комбинаций векторов из X называется линейной оболочкой множества X и обозначается $\langle X \rangle$. Говорят, что пространство V порождается множеством X, или натянуто на множество X, если $\langle X \rangle = V$.

Векторное пространство называется конечномерным, если оно порождается конечным числом векторов, и бесконечномерным в противном случае.

Линейная оболочка $\langle X \rangle$ множества $X \subset V$ — это наименьшее подпространство пространства V, содержащее X.

Замечание

Если $X \subset Y \subset V$, то $\langle X \rangle \subset \langle Y \rangle$. Для любого множества $X \subset V$ $\langle \langle X \rangle \rangle = \langle X \rangle$.

Основная лемма о линейной зависимости (из курса алгебры)

Если векторное пространство V порождается n векторами, то всякие $m > n$ векторов пространства V линейно зависимы.

Базис и размерность

Определение

Система векторов $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ в конечномерном векторном пространстве V называется базисом пространства V, если каждый вектор $\boldsymbol{x} \in V$ единственным образом выражается через $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Коэффициенты этого выражения называются координатами вектора $\boldsymbol{x}$ в базисе $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$.

Говорят, что система векторов $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ в векторном пространстве V полна, если любой вектор $\boldsymbol{x} \in V$ выражается через $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, т.е. $V = \langle\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}\rangle$.

Определение′

Базис — это полная линейно независимая система векторов.

Базис можно определить также как *максимальную* (по включению) *линейно независимую систему векторов* или как *минимальную полную систему векторов*.

Теорема

Все базисы конечномерного пространства V содержат одно и тоже число векторов.

Доказательство. Пусть $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ и $\{\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_m\}$ — базисы. Они линейно независимы. Основная лемма $\Longrightarrow$ $n \leqslant m$ и $m \leqslant n$. □

Теорема

Всякое конечномерное векторное пространство V обладает базисом. Из всякого конечного множества, порождающего V, можно выбрать базис.

Доказательство. Если $V = \{\mathbf{0}\}$, то доказывать нечего. Пусть $X \subset V$ — конечное множество, $V \neq \{\mathbf{0}\}$ и $V = \langle X \rangle$. Выберем любой ненулевой вектор $\boldsymbol{x}_1 \in X$. Предположим, что мы нашли линейно независимую систему векторов $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in X$. Если $\langle \boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \rangle \supset X$, то $V = \langle \boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \rangle$ согласно замечанию, а значит, $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ — базис. В противном случае возьмём $\boldsymbol{x}_{k+1} \in X \setminus \langle \boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \rangle$. Получим линейно независимую систему векторов $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+1} \in X$. X конечно $\Longrightarrow$ на некотором шаге процесс оборвётся. □

Определение

Число векторов в любом базисе конечномерного векторного пространства V называется размерностью пространства V и обозначается $\dim V$.

Примеры

- $\dim K^n = n$ для любого поля K. *Стандартный базис* состоит из векторов
$$(1,0,\dots,0),\ (0,1,\dots,0),\ \dots,\ (0,0,\dots,1).$$
- Размерность пространства K^K всех функций $K \to K$ бесконечна $\iff$ поле K бесконечно. Для конечного поля $\dim K^K = |K|$.
- Размерность $\dim K[x]$ бесконечна для любого поля. Базис: 1, x, x^2, x^3, $\dots$. $\dim K_{\leqslant n}[x] = n+1$ для любого поля.
- Размерность $\mathbb{C}$ как векторного пространства над $\mathbb{R}$ равна 2.
- Размерность $\mathbb{R}$ как векторного пространства над $\mathbb{Q}$ несчётна.

Лемма

Пусть $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k\}$ *— базис подпространства* U *векторного пространства* V *и* $\boldsymbol{x} \notin U$. *Тогда система векторов* $\boldsymbol{x}, \boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k$ *линейно независима.*

Доказательство. Допустим, что система векторов $\boldsymbol{x}, \boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k$ линейно зависима. Тогда $\boldsymbol{x}$ выражается через векторы $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k$ и потому принадлежит подпространству U, а это противоречит предположению. □

Следствие

Всякое подпространство U *конечномерного векторного пространства* V *конечномерно, причём* $\dim U \leqslant \dim V$.

Доказательство. Для $U = \{\mathbf{0}\}$ доказывать нечего. Пусть $U \neq \{\mathbf{0}\}$. Возьмём ненулевой вектор $\boldsymbol{x}_1 \in U$. Пусть мы выбрали линейно независимые векторы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in U$. Если $\langle\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}\rangle = U$, то U конечномерно. Если нет, выберем $\boldsymbol{x}_{k+1} \in U \setminus \langle\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}\rangle$. По лемме система $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+1}\}$ линейно независима. По основной лемме на шаге с номером $\leqslant \dim V$ мы получим систему, порождающую U. □

Следствие

Если V — конечномерное векторное пространство и U — его собственное подпространство, то $\dim U < \dim V$.

Доказательство. Пусть $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k\}$ — базис U. Возьмем любой вектор $\boldsymbol{x} \in V \setminus U$. По лемме система $\{\boldsymbol{x}, \boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k\}$ линейно независима. По основной лемме число элементов этой системы не превосходит $\dim V$. $\square$

Теорема

Пусть X — любое множество векторов в конечномерном векторном пространстве V. Всякую линейно независимую систему векторов из X можно дополнить до максимальной (по включению) линейно независимой системы векторов из X.

Доказательство. Пусть $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\} \subset X$ — линейно независимая система. Если $\langle\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}\rangle \supset X$, то она максимальна. Если нет, то возьмём $\boldsymbol{x}_{k+1} \in X \setminus \langle\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}\rangle \subset \langle X\rangle \setminus \langle\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}\rangle$. По лемме система $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+1}\}$ линейно независима. Пространство V конечномерно $\Longrightarrow$ $\langle X\rangle$ конечномерно $\Longrightarrow$ на некотором шаге получим $\langle\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}\rangle \supset X$, т.е. $\langle\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}\rangle = \langle X\rangle$. $\square$

Теорема

Во всяком конечномерном векторном пространстве V имеется базис. Более того, всякую линейно независимую систему векторов в V можно дополнить до базиса.

Определение

Множество векторов S в векторном пространстве V *полно*, если $V = \langle S \rangle$.

Базис (Гамеля) векторного пространства V — это полное линейно независимое множество векторов в V.

Другими словами, базис — это максимальное линейно независимое множество векторов, а также минимальное полное множество векторов.

Теорема

В любом векторном пространстве V имеется базис. Более того, всякое линейно независимое множество векторов в V можно дополнить до базиса.

Доказательство этой теоремы существенно использует лемму Цорна, которая равносильна аксиоме выбора (см. [Н.К. Верещагин, А. Шень, *Начала теории множеств*, Москва: МЦНМО, 2012]). Она сама равносильна аксиоме выбора.

Частичный порядок, или просто ***порядок***, на множестве X — это отношение $\leqslant$ на множестве X, обладающее следующими свойствами:

- $x \leqslant y$ и $y \leqslant z \implies x \leqslant z$ (транзитивность);
- $x \leqslant x$ для всякого $x \in X$ (рефлексивность);
- $x \leqslant y$ и $y \leqslant x \implies x = y$ (антисимметричность).

Пара $(X, \leqslant)$, т.е. множество X вместе с заданным на нём порядком, называется (*частично*) ***упорядоченным множеством***. Элементы $x, y \in X$ ***сравнимы***, если $x \leqslant y$ или $y \leqslant x$. Элемент $x \in X$ ***максимален***, если $x \leqslant y \implies x = y$. *Верхняя грань* множества $Y \subset X$ — любой элемент $x \in X$ со свойством $y \leqslant x$ для любого $y \in Y$. Если любые два элемента в X сравнимы относительно порядка $\leqslant$, то говорят, что множество X ***линейно упорядочено***.

На всяком подмножестве Y упорядоченного множества X естественно возникает индуцированный порядок (для $x, y \in Y$ полагаем $x \leqslant y$ в $Y \iff x \leqslant y$ в X).

Лемма Цорна

Пусть $(X, \leqslant)$ — упорядоченное множество и всякое подмножество X, на котором порядок $\leqslant$ линеен, имеет верхнюю грань. Тогда в X есть максимальный элемент.

Лемма Цорна (оригинальная формулировка)

Пусть $\mathscr{X}$ — семейство множеств со свойством: если $\mathscr{Y} \subset \mathscr{X}$ таково, что для любых $X, Y \in \mathscr{Y}$ либо $X \subset Y$, либо $Y \subset X$, то $\bigcup \mathscr{Y} \in \mathscr{X}$. Тогда в семействе $\mathscr{X}$ есть максимальный по включению элемент.

Доказательство теоремы. Возьмём в качестве $\mathscr{X}$ семейство всех линейно независимых множеств в V, содержащих данное множество $\mathcal{S}$. Пусть $\mathscr{Y} \subset \mathscr{X}$ удовлетворяет условию в лемме. Возьмём любые $n \in \mathbb{N}$ и $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \bigcup \mathscr{Y}$. Для $i \in \mathbb{N}$ найдём $\mathcal{S}_i \in \mathscr{Y}$, для которых $\boldsymbol{x}_i \in \mathcal{S}_i$. По условию на $\mathscr{Y}$ множества $\mathcal{S}_1, \ldots, \mathcal{S}_n$ можно упорядочить по включению. Пусть $\mathcal{S}_{k_1} \subset \mathcal{S}_{k_2} \subset \cdots \subset \mathcal{S}_{k_n}$. Тогда $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathcal{S}_{k_n}$. Множество $\mathcal{S}_{k_n}$ линейно независимо $\Longrightarrow$ $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ линейно независимы $\Longrightarrow$ $\bigcup \mathscr{Y}$ линейно независимо. Лемма Цорна $\Longrightarrow$ в $\mathscr{X}$ есть максимальное линейно независимое множество, содержащее $\mathcal{S}$. $\square$

Ранг системы векторов

Теорема

Число элементов во всех максимальных линейно независимых системах векторов произвольного подмножества X конечномерного векторного пространства одинаково и равно размерности $\dim\langle X\rangle$ линейной оболочки этого подмножества.

Доказательство. Очевидно, для всякой линейно независимой системы векторов $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in X$ имеем $\langle\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}\rangle \supset X$ — иначе в X найдётся вектор, который не выражается через $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, и его можно будет добавить к данной системе без нарушения линейной независимости, а это невозможно в силу её максимальности. Очевидно также, что $\langle\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}\rangle \subset \langle X\rangle$. Значит, $\langle\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}\rangle = \langle X\rangle$ и система $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ обязана быть базисом в $\langle X\rangle$. Число элементов во всех базисах одинаково. □

Это одинаковое число элементов в максимальных линейно независимых системах векторов из X называется рангом множества X и обозначается $\operatorname{rank} X$.

Координаты векторов

Пусть V — векторное пространство над полем K конечной размерности $\dim V = n$, и пусть $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ — какой-нибудь фиксированный базис. Тогда для любого вектора $\boldsymbol{x} \in V$ существуют однозначно определённые числа $x_1, \ldots, x_n \in K$ такие, что $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$. Набор чисел $(x_1, \ldots, x_n)$ — это набор координат вектора $\boldsymbol{x}$.

Если в базисе $\boldsymbol{E}$ вектор $\boldsymbol{x}$ имеет координаты $(x_1, \ldots, x_n)$, а вектор $\boldsymbol{y}$ — координаты $(y_1, \ldots, y_n)$, т.е. $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$ и $\boldsymbol{y} = y_1\boldsymbol{e}_1 + \cdots + y_n\boldsymbol{e}_n$, то

$$\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1)\boldsymbol{e}_1 + \cdots + (x_n + y_n)\boldsymbol{e}_n \quad \text{и} \quad \lambda\boldsymbol{x} = (\lambda \cdot x_1)\boldsymbol{e}_1 + \cdots + (\lambda \cdot x_n)\boldsymbol{e}_n,$$

т.е. вектор $\boldsymbol{x} + \boldsymbol{y}$ имеет координаты $(x_1 + y_1, \ldots, x_n + y_n)$, а вектор $\lambda\boldsymbol{x}$ — координаты $(\lambda \cdot x_1, \ldots, \lambda \cdot x_n)$.

Таким образом, при сложении векторов их координаты складываются, а при умножении вектора на число его координаты умножаются на это число. Поэтому *строка (столбец) координат линейной комбинации векторов есть линейная комбинация с теми же коэффициентами строк (столбцов) координат этих векторов.* Отсюда следует, что векторы линейно зависимы тогда и только тогда, когда линейно зависимы строки (столбцы) их координат.

Изоморфизм векторных пространств

Определение

Векторные пространства V и W над полем K изоморфны, если существует биекция $\varphi\colon V \to W$, сохраняющая операции сложения и умножения на число:

1. $\varphi(\boldsymbol{x}+\boldsymbol{y}) = \varphi(\boldsymbol{x}) + \varphi(\boldsymbol{y})$ для любых $\boldsymbol{x}, \boldsymbol{y} \in V$;
2. $\varphi(\lambda \boldsymbol{x}) = \lambda\varphi(\boldsymbol{x})$ для любых $\lambda \in K$ и $\boldsymbol{x} \in V$.

Обозначение: $V \cong W$. Биекция φ называется изоморфизмом пространств V и W.

Свойства изоморфизма

- $\varphi(\mathbf{0}_V) = \mathbf{0}_W$ и $\varphi^{-1}(\mathbf{0}_W) = \mathbf{0}_V$.
- Обратное отображение $\varphi^{-1}\colon W \to V$ существует и является изоморфизмом.
  Доказательство. Для $\boldsymbol{x}, \boldsymbol{y} \in W$ и $\lambda \in K$ имеем $\varphi(\varphi^{-1}(\boldsymbol{x}) + \varphi^{-1}(\boldsymbol{y})) = \varphi(\varphi^{-1}(\boldsymbol{x})) + \varphi(\varphi^{-1}(\boldsymbol{y})) = \boldsymbol{x}+\boldsymbol{y}$ и $\varphi(\lambda\varphi^{-1}(\boldsymbol{x})) = \lambda\boldsymbol{x}$. Применив φ^{-1} к обеим частям этих равенств, видим, что $\varphi^{-1}(\boldsymbol{x}+\boldsymbol{y}) = \varphi^{-1}(\boldsymbol{x}+\boldsymbol{y})$ и $\lambda\varphi^{-1}(\boldsymbol{x}) = \varphi^{-1}(\lambda\boldsymbol{x})$. $\square$
- Если V, U и W — векторные пространства над одним и тем же полем, $V \cong U$ и $U \cong W$, то $V \cong W$.

Теорема

Всякое векторное пространство V над полем K размерности n изоморфно арифметическому пространству K^n.

Доказательство. Зафиксируем базис $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ в V. Рассмотрим отображение $\varphi\colon V \to K^n$, которое каждому вектору $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$ ставит в соответствие строку его координат $(x_1, \ldots, x_n)$ в базисе $\boldsymbol{E}$. Разложение вектора по базису единственно $\implies$ это отображение корректно определено. Ясно, что это биекция, удовлетворяющая условиям 1 и 2 из определения изоморфизма. □

Пусть V — конечномерное векторное пространство и $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ — *фиксированный* базис в нём. Доказанная теорема позволяет отождествлять каждый вектор $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$ в V со строкой или столбцом его координат $x_1, \ldots, x_n$ и писать $\boldsymbol{x} = (x_1, \ldots, x_n)$ или $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. В частности, для $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in V$, где $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{in})$, имеем $\operatorname{rank}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} = \operatorname{rank}\begin{pmatrix} x_{11} & \ldots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{k1} & \ldots & x_{kn} \end{pmatrix}$.

Теорема

Конечномерные векторные пространства над одним и тем же полем изоморфны тогда и только тогда, когда они имеет одинаковую размерность.

Доказательство

Необходимость. Пусть $\varphi\colon V \to W$ — изоморфизм векторных пространств V и W и $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ — базис пространства V. Покажем, что $\{\varphi(\boldsymbol{e}_1), \dots, \varphi(\boldsymbol{e}_n)\}$ — базис пространства W. Эта система линейно независима, так как если $\lambda_1\varphi(\boldsymbol{e}_1) + \dots + \lambda_n\varphi(\boldsymbol{e}_n) = \mathbf{0}_W$, то

$$\varphi^{-1}(\lambda_1\varphi(\boldsymbol{e}_1) + \dots + \lambda_n\varphi(\boldsymbol{e}_n)) = \lambda_1\varphi^{-1}(\varphi(\boldsymbol{e}_1)) + \dots + \lambda_n\varphi^{-1}(\varphi(\boldsymbol{e}_n)) =$$
$$= \lambda_1\boldsymbol{e}_1 + \dots + \lambda_n\boldsymbol{e}_n = \mathbf{0}_V,$$

а значит, $\lambda_1 = \dots = \lambda_n = 0$. Любой вектор $\boldsymbol{x} \in W$ линейно выражается через $\varphi(\boldsymbol{e}_1), \dots, \varphi(\boldsymbol{e}_n)$, так $\varphi^{-1}(\boldsymbol{x}) = \lambda_1\boldsymbol{e}_1 + \dots + \lambda_n\boldsymbol{e}_n$ для некоторых $\lambda_1, \dots, \lambda_n \in K$, а значит, $\boldsymbol{x} = \lambda_1\varphi(\boldsymbol{e}_1) + \dots + \lambda_n\varphi(\boldsymbol{e}_n)$. Итак, $\{\varphi(\boldsymbol{e}_1), \dots, \varphi(\boldsymbol{e}_n)\}$ — базис пространства W. Следовательно, $\dim W = \dim V$.

Достаточность вытекает из предыдущей теоремы. $\square$

Упражнения

1. Покажите, что в любом векторном пространстве V все базисы имеют одну и ту же мощность (она называется размерностью пространства V).
2. Докажите, что векторные пространства над одним и тем же полем изоморфны тогда и только тогда, когда они имеют одинаковую размерность.
3. Покажите, что для бесконечного множества X и любого поля K пространство всех функций $f : X \to K$ с поточечными операциями сложения и умножения на число бесконечномерно. Что, если множество X конечно, но зато поле K бесконечно?
4. Вспомните, что множество $\mathscr{P}(X)$ всех подмножеств множества X с операциями, определёнными правилами $A + B = A \vartriangle B = (A \cup B) \smallsetminus (A \cap B)$, $0A = \varnothing$ и $1A = A$ для $A, B \subset X$, является векторным пространством над полем $\mathbb{F}_2$. Найдите размерность этого пространства.
5. Пусть V — n-мерное векторное пространство над произвольным полем K и $\boldsymbol{E}_1, \ldots, \boldsymbol{E}_n$ — базисы V. Перенумеруем векторы в базисах: $\boldsymbol{E}_i = \{\boldsymbol{e}_{i1}, \boldsymbol{e}_{i2}, \ldots, \boldsymbol{e}_{in}\}$, $i \leqslant n$. Верно ли, что перенумерацию всегда можно осуществить таким образом, что все системы $\{\boldsymbol{e}_{1i}, \boldsymbol{e}_{2i}, \ldots, \boldsymbol{e}_{ni}\}$, $i \leqslant n$, тоже окажутся базисами?

Матрица перехода

Пусть даны два базиса $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ и $\boldsymbol{E}' = \{\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n\}$ векторного пространства V. Выразим каждый вектор $\boldsymbol{e}'_j \in \boldsymbol{E}'$ через базис $\boldsymbol{E}$:

$$\boldsymbol{e}'_j = t_{1j}\boldsymbol{e}_1 + t_{2j}\boldsymbol{e}_2 + \cdots + t_{nj}\boldsymbol{e}_n.$$

Матрица

$$T = \begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} & \ldots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{pmatrix}$$

называется матрицей перехода от базиса $\boldsymbol{E}$ к базису $\boldsymbol{E}'$. Она всегда невырождена.

Очень нужная формула

Если вектор $\boldsymbol{x} \in V$ имеет координаты $(x_1, x_2, \ldots, x_n)$ в базисе $\boldsymbol{E}$ и координаты $(x'_1, x'_2, \ldots, x'_n)$ в базисе $\boldsymbol{E}'$, то

$$\boxed{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = T \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}}.$$

Имеем
$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \dots + x_n\boldsymbol{e}_n = x_1'\boldsymbol{e}_1' + x_2'\boldsymbol{e}_2' + \dots + x_n'\boldsymbol{e}_n'. \qquad (\star)$$

Подставим выражения для $\boldsymbol{e}_j'$:

$$\begin{aligned}\boldsymbol{x} = x_1'(t_{11}\boldsymbol{e}_1 + t_{21}\boldsymbol{e}_2 + \dots + t_{n1}\boldsymbol{e}_n) + x_2'(t_{12}\boldsymbol{e}_1 + t_{22}\boldsymbol{e}_2 + \dots + t_{n2}\boldsymbol{e}_n) + \\ + \dots + x_n'(t_{1n}\boldsymbol{e}_1 + t_{2n}\boldsymbol{e}_2 + \dots + t_{nn}\boldsymbol{e}_n).\end{aligned}$$

Соберём коэффициенты при каждом векторе $\boldsymbol{e}_i$:

$$\begin{aligned}\boldsymbol{x} = {} & (x_1' \cdot t_{11} + x_2' \cdot t_{12} + \dots + x_n' \cdot t_{1n})\boldsymbol{e}_1 + \\ & + (x_1' \cdot t_{21} + x_2' \cdot t_{22} + \dots + x_n' \cdot t_{2n})\boldsymbol{e}_2 + \\ & + \dots + (x_1' \cdot t_{n1} + x_2' \cdot t_{n2} + \dots + x_n' \cdot t_{nn})\boldsymbol{e}_n.\end{aligned} \qquad (\star\star)$$

$(\star)$ + $(\star\star)$ + единственность разложения вектора по базису $\Longrightarrow$

$$\begin{aligned} x_1 &= x_1' \cdot t_{11} + x_2' \cdot t_{12} + \dots + x_n' \cdot t_{1n}, \\ x_2 &= x_1' \cdot t_{21} + x_2' \cdot t_{22} + \dots + x_n' \cdot t_{2n}, \\ &\dots \\ x_n &= x_1' \cdot t_{n1} + x_2' \cdot t_{n2} + \dots + x_n' \cdot t_{nn}, \end{aligned} \qquad \text{т.е.} \qquad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ t_{21} & t_{22} & \dots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & \dots & t_{nn} \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix}.$$

Свойства матрицы перехода

Здесь для базисов $\boldsymbol{E}$ и $\boldsymbol{E}'$ будем обозначать матрицу перехода от $\boldsymbol{E}$ к $\boldsymbol{E}'$ через $T_{\boldsymbol{E}\to\boldsymbol{E}'}$.

1. Если $\boldsymbol{E}_1$ и $\boldsymbol{E}_2$ — два базиса конечномерного векторного пространства V, то
$$T_{\boldsymbol{E}_1\to\boldsymbol{E}_2} = T^{-1}_{\boldsymbol{E}_2\to\boldsymbol{E}_1}.$$
2. Если $\boldsymbol{E}_1$, $\boldsymbol{E}_2$ и $\boldsymbol{E}_3$ — три базиса конечномерного векторного пространства V, то
$$T_{\boldsymbol{E}_1\to\boldsymbol{E}_3} = T_{\boldsymbol{E}_1\to\boldsymbol{E}_2} T_{\boldsymbol{E}_2\to\boldsymbol{E}_3}.$$
3. Если V — векторное пространство размерности n и $\boldsymbol{E}$ — любой базис этого пространства, то всякая невырожденная квадратная матрица размера $n \times n$ является матрицей перехода от базиса $\boldsymbol{E}$ к некоторому другому базису.

Упражнение

Докажите эти свойства.

Параметрические уравнения подпространства

Пусть V — конечномерное векторное пространство и $U \underset{\text{lin}}{\subseteq} V$. Предположим, что $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ — базис в V и $\{\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_k\}$ — базис в U $(k \leqslant n)$. Пусть каждый вектор $\boldsymbol{\varepsilon}_i$ имеет в базисе $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ координаты $(a_{1i}, \ldots, a_{ni})$, т.е.

$$\boldsymbol{\varepsilon}_i = a_{1i}\boldsymbol{e}_1 + \cdots + a_{ni}\boldsymbol{e}_n.$$

Для любого вектора $\boldsymbol{x} \in U$ $\exists\, x_1, \ldots, x_n, x'_1, \ldots, x'_k \in K$ такие, что

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n = x'_1\boldsymbol{\varepsilon}_1 + \cdots + x'_k\boldsymbol{\varepsilon}_k. \qquad (\star)$$

Снова подставим выражения для $\boldsymbol{\varepsilon}_j$ и соберём коэффициенты при каждом $\boldsymbol{e}_i$:

$$\boldsymbol{x} = (x'_1 \cdot a_{11} + x'_2 \cdot a_{12} + \cdots + x'_k \cdot a_{1k})\boldsymbol{e}_1 + \cdots + (x'_1 \cdot a_{n1} + x'_2 \cdot a_{n2} + \cdots + x'_k \cdot a_{nk})\boldsymbol{e}_n. \quad (\star\star)$$

$(\star)$ + $(\star\star)$ + единственность разложения вектора по базису $\Longrightarrow$

$$\begin{aligned}
x_1 &= x'_1 \cdot a_{11} + x'_2 \cdot a_{12} + \cdots + x'_k \cdot a_{1k},\\
x_2 &= x'_1 \cdot a_{21} + x'_2 \cdot a_{22} + \cdots + x'_k \cdot a_{2k},\\
&\ldots\\
x_n &= x'_1 \cdot a_{n1} + x'_2 \cdot a_{n2} + \cdots + x'_k \cdot a_{nk},
\end{aligned}$$

Параметрические уравнения подпространства

Пусть V — конечномерное векторное пространство и $U \underset{\text{lin}}{\subseteq} V$. Предположим, что $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ — базис в V и $\{\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_k\}$ — базис в U ($k \leqslant n$). Пусть каждый вектор $\boldsymbol{\varepsilon}_i$ имеет в базисе $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ координаты $(a_{1i}, \ldots, a_{ni})$, т.е.

$$\boldsymbol{\varepsilon}_i = a_{1i}\boldsymbol{e}_1 + \cdots + a_{ni}\boldsymbol{e}_n.$$

Для любого вектора $\boldsymbol{x} \in U$ $\exists\, x_1, \ldots, x_n, x_1', \ldots, x_k' \in K$ такие, что

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n = x_1'\boldsymbol{\varepsilon}_1 + \cdots + x_k'\boldsymbol{\varepsilon}_k. \qquad (\star)$$

Снова подставим выражения для $\boldsymbol{\varepsilon}_j$ и соберём коэффициенты при каждом $\boldsymbol{e}_i$:

$$\boldsymbol{x} = (x_1' \cdot a_{11} + x_2' \cdot a_{12} + \cdots + x_k' \cdot a_{1k})\boldsymbol{e}_1 + \cdots + (x_1' \cdot a_{n1} + x_2' \cdot a_{n2} + \cdots + x_k' \cdot a_{nk})\boldsymbol{e}_n. \quad (\star\star)$$

$(\star)$ + $(\star\star)$ + единственность разложения вектора по базису $\Longrightarrow$

$$\begin{cases} x_1 = t_1 \cdot a_{11} + t_2 \cdot a_{12} + \cdots + t_k \cdot a_{1k}, \\ x_2 = t_1 \cdot a_{21} + t_2 \cdot a_{22} + \cdots + t_k \cdot a_{2k}, \\ \quad \ldots \\ x_n = t_1 \cdot a_{n1} + t_2 \cdot a_{n2} + \cdots + t_k \cdot a_{nk}, \end{cases}$$

— система уравнений, определяющая U.

Сумма и пересечение подпространств

Пусть V — любое векторное пространство над полем K, а U и W — его подпространства. Легко проверить, $U \cap W$ — тоже подпространство V. Однако $U \cup W$ не обязано быть подпространством.

Определение

Сумма $U + W$ подпространств U и W векторного пространства V — это множество

$$U + W = \{\boldsymbol{u} + \boldsymbol{w} : \boldsymbol{u} \in U, \ \boldsymbol{w} \in W\}.$$

Сложение векторов ассоциативно $\Longrightarrow$ сложение подпространств ассоциативно.

Упражнения

1. Докажите, что $U + W$ — наименьшее подпространство пространства V, содержащее $U \cup W$.
2. Докажите, что если поле K содержит $\geqslant n$ элементов, $V_1, \ldots, V_n$ — подпространства векторного пространства V над полем K и $V_1 \cup \cdots \cup V_n = V_1 + \cdots + V_n$, то для некоторого $i \leqslant n$ имеем $V_i = V$.

Определение

Сумма $U + W$ подпространств U и W называется прямой суммой и обозначается $U \oplus W$, если каждый вектор $\boldsymbol{v} \in U + W$ единственным образом представляется в виде $\boldsymbol{u} + \boldsymbol{w}$, где $\boldsymbol{u} \in U$ и $\boldsymbol{w} \in W$. При этом $\boldsymbol{u}$ называется проекцией вектора $\boldsymbol{v}$ на подпространство U (параллельно подпространству W), а $\boldsymbol{w}$ — проекцией вектора $\boldsymbol{v}$ на W (параллельно подпространству U).

Теорема

Сумма подпространств U и W произвольного векторного пространства является прямой $\iff$ $U \cap W = \{\mathbf{0}\}$.

Доказательство. $\Longrightarrow$: Пусть $U + W = U \oplus W$. Любой вектор $\boldsymbol{v} \in U \cap W$ можно представить как $\boldsymbol{v} + \mathbf{0}$ и как $\mathbf{0} + \boldsymbol{v}$. Представление единственно $\Longrightarrow$ $U \cap W = \{\mathbf{0}\}$.
$\Longleftarrow$: Пусть $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{w}_1 = \boldsymbol{v}_2 + \boldsymbol{w}_2$. Тогда $\boldsymbol{v}_1 - \boldsymbol{v}_2 = \boldsymbol{w}_2 - \boldsymbol{w}_1 \in U \cap W$. $U \cap W = \{\mathbf{0}\}$ $\Longrightarrow$ $\boldsymbol{v}_1 = \boldsymbol{v}_2$ и $\boldsymbol{w}_1 = \boldsymbol{w}_2$. $\square$

Упражнение

1. Пусть V — векторное пространство. Говорят, что его подпространства $V_1, \ldots, V_n$ линейно независимы, если любая система векторов $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$, где $\boldsymbol{v}_i \in V_i$ для $i \leqslant n$, линейно независима. Докажите, что следующие условия равносильны:
   - сумма $V_1 + \cdots + V_n$ прямая;
   - подпространства $V_1, \ldots, V_n$ линейно независимы;
   - если векторы $\boldsymbol{v}_i$, $i \leqslant n$, удовлетворяют условиям $\boldsymbol{v}_i \in V_i$ для $i \leqslant n$ и $\boldsymbol{v}_1 + \cdots + \boldsymbol{v}_n = \mathbf{0}$, то $\boldsymbol{v}_1 = \cdots = \boldsymbol{v}_n = \mathbf{0}$;
   - объединение любых базисов подпространств V_i, $i \leqslant n$, является базисом подпространства $V_1 + \cdots + V_n$;
   - $V_i \cap (V_{i+1} + \cdots + V_n) = \{\mathbf{0}\}$ для всех $i < n$.
2. Подумайте, как распространить утверждения пункта 1 на случай бесконечного (не обязательно счётного) числа слагаемых.
3. Пусть V — конечномерное векторное пространство и $V_1, \ldots, V_n$ — его подпространства. Докажите, что $V = V_1 \oplus \cdots \oplus V_n$ тогда и только тогда, когда $V = V_1 + \cdots + V_n$ и $\dim V = \dim V_1 + \cdots + \dim V_n$.

Примеры

- Если $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ — базис векторного пространства V, то
$$V = \langle \boldsymbol{e}_1 \rangle \oplus \cdots \oplus \langle \boldsymbol{e}_n \rangle.$$
- Пусть π_1 и π_2 — непараллельные плоскости в обычном пространстве. Тогда пространство является суммой двух подпространств — векторов, параллельных π_1, и векторов, параллельных π_2. Это не прямая сумма.
- Пусть π — плоскость в обычном пространстве и l — не параллельная ей прямая. Тогда пространство является суммой двух подпространств — векторов, параллельных π, и векторов, параллельных l. Это прямая сумма.
- Любая квадратная вещественная матрица A порядка n является суммой симметрической и кососимметрической матриц: $A = \frac{A+A^T}{2} + \frac{A-A^T}{2}$. Поэтому пространство $M_n(\mathbb{R})$ всех квадратных вещественных матриц порядка n является (прямой) суммой подпространства всех симметрических матриц и подпространства всех кососимметрических матриц.
- Пространство всех функций $\mathbb{R} \to \mathbb{R}$ является прямой суммой подпространства всех чётных функций и подпространства всех нечётных функций.

Теорема

Для любых двух подпространств U и W конечномерного векторного пространства V над полем K справедлива формула Грассмана

$$\boxed{\dim(U+W) = \dim U + \dim W - \dim(U \cap W)}\,.$$

Доказательство. Выберем базис $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k\}$ в подпространстве $U \cap W$ и дополним его векторами $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_l$ до базиса подпространства U и векторами $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_m$ до базиса подпространства W. Положим $\boldsymbol{F} = \{\boldsymbol{f}_1, \ldots, \boldsymbol{f}_l\}$ и $\boldsymbol{G} = \{\boldsymbol{g}_1, \ldots, \boldsymbol{g}_m\}$. Покажем, что $\boldsymbol{E} \cup \boldsymbol{F} \cup \boldsymbol{G}$ — базис суммы $U + W$.

Для $\boldsymbol{v} \in U + W$ имеем $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$, где $\boldsymbol{u} \in U$, $\boldsymbol{w} \in W$. Вектор $\boldsymbol{u}$ выражается через $\boldsymbol{E} \cup \boldsymbol{F}$ и $\boldsymbol{w}$ выражается через $\boldsymbol{E} \cup \boldsymbol{G}$ $\implies$ $\boldsymbol{v}$ выражается через $\boldsymbol{E} \cup \boldsymbol{F} \cup \boldsymbol{G}$.

Осталось показать, что система $\boldsymbol{E} \cup \boldsymbol{F} \cup \boldsymbol{G}$ линейно независима. Пусть

$$\underbrace{\alpha_1 \boldsymbol{e}_1 + \cdots + \alpha_k \boldsymbol{e}_k}_{\boldsymbol{a}} + \underbrace{\beta_1 \boldsymbol{f}_1 + \cdots + \beta_l \boldsymbol{f}_l}_{\boldsymbol{b}} + \underbrace{\gamma_1 \boldsymbol{g}_1 + \cdots + \gamma_m \boldsymbol{g}_m}_{\boldsymbol{c}} = \boldsymbol{0}. \qquad (\star)$$

Имеем $\boldsymbol{b} \in \langle \boldsymbol{F} \rangle = U$ и $\boldsymbol{b} = -\boldsymbol{a} - \boldsymbol{c} \in W$ $\implies$ $\boldsymbol{b} \in U \cap W = \langle \boldsymbol{E} \rangle$. Система $\boldsymbol{E} \cup \boldsymbol{F}$ линейно независима $\implies$ $\boldsymbol{b} = \boldsymbol{0}$ $\implies$ $\boldsymbol{a} + \boldsymbol{c} = \boldsymbol{0}$. Система $\boldsymbol{E} \cup \boldsymbol{G}$ линейно независима $\implies$ $\boldsymbol{a} = \boldsymbol{c} = \boldsymbol{0}$. Системы $\boldsymbol{E}$, $\boldsymbol{F}$ и $\boldsymbol{G}$ линейно независимы $\implies$ линейная комбинация $(\star)$ тривиальна. $\square$

Линейные функции

Определение

Линейная функция (или линейный функционал, или линейная форма) на векторном пространстве V над полем K — это любое отображение $\boldsymbol{f}: V \to K$ со свойствами

1. $\boldsymbol{f}(\boldsymbol{x}+\boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{y})$ для любых $\boldsymbol{x}, \boldsymbol{y} \in V$,
2. $\boldsymbol{f}(\lambda \boldsymbol{x}) = \lambda \boldsymbol{f}(\boldsymbol{x})$ для любых $\lambda \in K$, $\boldsymbol{x} \in V$.

Примеры

- На обычном пространстве: $\boldsymbol{x} \mapsto (\boldsymbol{x}, \boldsymbol{a})$ для фиксированного вектора $\boldsymbol{a}$.
- На арифметическом пространстве K^n: $(x_1, \dots, x_n) \mapsto x_i$ для $i \leqslant n$.
- На пространстве функций $K \to K$: $f \mapsto f(a)$ для фиксированного $a \in K$.
- На пространстве непрерывных вещественных функций: $f \mapsto \int\limits_a^b f(x)\,dx$.
- На пространстве дифференцируемых функций: $f \mapsto f'(a)$ для $a \in K$.
- На пространстве матриц: $A \mapsto \operatorname{tr} A$. Функция $A \mapsto \det A$ не линейна.

Множество всех линейных функций на векторном пространстве V над полем K образует векторное пространство над K относительно операций поточечного сложения функций и умножения функции на число:

$$(\boldsymbol{f}+\boldsymbol{g})(\boldsymbol{x})=\boldsymbol{f}(\boldsymbol{x})+\boldsymbol{g}(\boldsymbol{x}), \quad (\lambda \boldsymbol{f})(\boldsymbol{x})=\lambda\cdot\boldsymbol{f}(\boldsymbol{x}) \quad \text{для } \boldsymbol{x}\in V,\ \lambda\in K.$$

Определение

Пространство всех линейных функций (функционалов, форм) на V называется **сопряжённым пространством** по отношению к V и обозначается V^*.

Пусть V — n-мерное пространство с базисом $\mathscr{E}=\{\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n\}$. Значение линейной функции $\boldsymbol{f}\in V^*$ на векторе $\boldsymbol{x}\in V$ может быть выражено через координаты $(x_1,\ldots,x_n)$ этого вектора:

$$\boldsymbol{f}(\boldsymbol{x})=\boldsymbol{f}(x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_n)=x_1\cdot\boldsymbol{f}(\boldsymbol{e}_1)+\cdots+x_n\cdot\boldsymbol{f}(\boldsymbol{e}_n)=f_1\cdot x_1+\cdots+f_n\cdot x_n,$$

где $f_i=\boldsymbol{f}(\boldsymbol{e}_i)$ для $i\leqslant n$. Числа $f_1,\ldots,f_n$ не зависят от вектора $\boldsymbol{x}$, они определяются только функцией $\boldsymbol{f}$ и базисом $\boldsymbol{E}$. Эти числа называются **коэффициентами** функции $\boldsymbol{f}$ в базисе $\boldsymbol{E}$.

Всякая линейная функция $\boldsymbol{f}: V \to K$ однозначно определяется своими значениями на базисных векторах — коэффициентами $f_1 = \boldsymbol{f}(\boldsymbol{e}_1), \ldots, f_n = \boldsymbol{f}(\boldsymbol{e}_n)$. Коэффициентами могут быть любые числа.

Если $\boldsymbol{f}, \boldsymbol{g} \in V^*$ и $\lambda \in K$, то

$$(f+g)_i = f_i + g_i \quad \text{и} \quad (\lambda f)_i = \lambda \cdot f_i \qquad \text{для } i \leqslant n.$$

Значит, отображение $\boldsymbol{f} \mapsto (f_1, \ldots, f_n)$ — изоморфизм между V^* и арифметическим пространством K^n.

Теорема

Любое конечномерное векторное пространство V изоморфно своему сопряжённому пространству V^*.

Упражнение

Покажите, что если векторное пространство V бесконечномерно, то пространства V и V^* не изоморфны.

Взаимный базис

Пусть $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ — базис векторного пространства V. Определим систему функционалов $\boldsymbol{\mathcal{E}} = \{\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n\}$ в V^*, полагая

$$\boldsymbol{\varepsilon}_i(\boldsymbol{e}_j) = \delta_{ij} = \begin{cases} 1, & \text{если } i = j, \\ 0, & \text{если } i \neq j \end{cases} \qquad \text{для } i, j \leqslant n$$

(δ_{ij} — *символ Кронекера*). Заметим, что строка коэффициентов функционала $\boldsymbol{\varepsilon}_i$ — это i-я строка единичной матрицы. Значит, функционалы $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ линейно независимы. Поскольку пространство V^* n-мерно, эти функционалы составляют в нём базис.

Определение

Базис $\{\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n\}$ пространства V^*, определённый формулой $\boldsymbol{\varepsilon}_i(\boldsymbol{e}_j) = \delta_{ij}$, называется взаимным (а также двойственным или сопряжённым) базису $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ пространства V.

$$\boxed{\boldsymbol{\varepsilon}_i(\boldsymbol{e}_j) = \delta_{ij}}$$

Для любого $\boldsymbol{x} \in V$ с координатами $(x_1, \dots, x_n)$ в базисе $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ и всякого $i \leqslant n$ имеем

$$\boldsymbol{\varepsilon}_i(\boldsymbol{x}) = \boldsymbol{\varepsilon}_i(x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n) = x_1\boldsymbol{\varepsilon}_i(\boldsymbol{e}_1) + \dots + x_n\boldsymbol{\varepsilon}_i(\boldsymbol{e}_n) = x_i,$$

и для любого функционала $\boldsymbol{f} \in V^*$ с координатами $(f_1, \dots, f_n)$ в базисе $\{\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_n\}$ и всякого $i \leqslant n$ имеем

$$\boldsymbol{f}(\boldsymbol{e}_i) = f_1\boldsymbol{\varepsilon}_1(\boldsymbol{e}_i) + \dots + f_n\boldsymbol{\varepsilon}_n(\boldsymbol{e}_i) = f_i\boldsymbol{\varepsilon}_i(\boldsymbol{e}_i) = f_i.$$

Таким образом,

$$\boxed{x_i = \boldsymbol{\varepsilon}_i(\boldsymbol{x})} \qquad \text{и} \qquad \boxed{f_i = \boldsymbol{f}(\boldsymbol{e}_i)}.$$

Кроме того,

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n) = x_1\boldsymbol{f}(\boldsymbol{e}_1) + \dots + x_n\boldsymbol{f}(\boldsymbol{e}_n) = x_1 \cdot f_1 + \dots + x_n \cdot f_n.$$

Итак,

$$\boxed{\boldsymbol{f}(\boldsymbol{x}) = f_1 \cdot x_1 + \dots + f_n \cdot x_n} = (f_1, \dots, f_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Преобразование координат в сопряжённом пространстве

Пусть даны два базиса $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ и $\boldsymbol{E}' = \{\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_n\}$ векторного пространства V, и пусть $T = (t_{ij})$ — матрица перехода от $\boldsymbol{E}$ к $\boldsymbol{E}'$. Рассмотрим базисы $\boldsymbol{\mathcal{E}} = \{\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_n\}$ и $\boldsymbol{\mathcal{E}}' = \{\boldsymbol{\varepsilon}'_1, \dots, \boldsymbol{\varepsilon}'_n\}$ пространства V^*, сопряжённые к $\boldsymbol{E}$ и $\boldsymbol{E}'$, и матрицу перехода от $\boldsymbol{\mathcal{E}}$ к $\boldsymbol{\mathcal{E}}'$. Обозначим эту матрицу $S = (s_{ij})$.

Для $k, l \leqslant n$ имеем

$$\boldsymbol{\varepsilon}'_k = s_{1k}\boldsymbol{\varepsilon}_1 + \dots + s_{nk}\boldsymbol{\varepsilon}_n \quad \text{и} \quad \boldsymbol{e}'_k = t_{1k}\boldsymbol{e}_1 + \dots + t_{nk}\boldsymbol{e}_n.$$

Из равенств $\boldsymbol{\varepsilon}'_k(\boldsymbol{e}'_l) = \delta_{kl}$ и $\boldsymbol{\varepsilon}_i(\boldsymbol{e}_j) = \delta_{ij}$ вытекает, что

$$\delta_{kl} = \boldsymbol{\varepsilon}'_k(\boldsymbol{e}'_l) = \sum_{i=1}^{n} s_{ik}\boldsymbol{\varepsilon}_i\Big(\sum_{j=1}^{n} t_{jl}\boldsymbol{e}_j\Big) = \sum_{i,j=1}^{n} s_{ik}t_{jl}\boldsymbol{\varepsilon}_i(\boldsymbol{e}_j) = \sum_{i=1}^{n} s_{ik}t_{il}.$$

Следовательно, $S^T T = E$ и

$$\boxed{S = (T^{-1})^T}.$$

Естественный изоморфизм между V и V^{**}

Пусть V — конечномерное векторное пространство. Тогда между V и $V^{**} = (V^*)^*$ имеется *естественный* изоморфизм, который не зависит от выбора базиса или каких-либо других объектов.

Для каждого вектора $\boldsymbol{x} \in V$ определим функцию $e_{\boldsymbol{x}} : V^* \to K$ правилом

$$e_{\boldsymbol{x}}(\boldsymbol{f}) = \boldsymbol{f}(\boldsymbol{x}) \qquad \text{для } \boldsymbol{x} \in V.$$

Она линейна, т.е. $e_{\boldsymbol{x}} \in V^{**}$.

Функцию $e_{\boldsymbol{x}}$ называют функцией вычисления.

Теорема

Для всякого конечномерного векторного пространства V отображение $\Phi\colon V \to V^{**}$, определённое правилом $\boldsymbol{x} \mapsto e_{\boldsymbol{x}}$, является изоморфизмом.

Доказательство. Проверим, что отображение Φ линейно. Для $\boldsymbol{x}, \boldsymbol{y} \in V$ и $\boldsymbol{f} \in V^*$ имеем

$$e_{\boldsymbol{x}+\boldsymbol{y}}(\boldsymbol{f}) = \boldsymbol{f}(\boldsymbol{x}+\boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{y}) = e_{\boldsymbol{x}}(\boldsymbol{f}) + e_{\boldsymbol{y}}(\boldsymbol{f}) = (e_{\boldsymbol{x}} + e_{\boldsymbol{y}})(\boldsymbol{f}),$$

т.е. $\Phi(\boldsymbol{x}+\boldsymbol{y}) = \Phi(\boldsymbol{x}) + \Phi(\boldsymbol{y})$. Аналогично доказывается, что $\Phi(\lambda\boldsymbol{x}) = \lambda\Phi(\boldsymbol{x})$.

Проверим, что Φ биективно. Возьмём любой базис $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ в V и взаимный базис $\{\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_n\}$ в V^*.
Для $i, j \leqslant n$ имеем $e_{\boldsymbol{e}_i}(\boldsymbol{\varepsilon}_j) = \boldsymbol{\varepsilon}_j(\boldsymbol{e}_j) = \delta_{ij}$. Значит, $\{e_{\boldsymbol{e}_1}, \dots, e_{\boldsymbol{e}_n}\}$ — базис пространства V^{**}, взаимный с базисом $\{\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_n\}$. Отображение Φ переводит вектор $\boldsymbol{x} \in V$ с координатами $(x_1, \dots, x_n)$ в базисе $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ в вектор $\Phi(\boldsymbol{x}) \in V^{**}$ с теми же координатами в базисе $\{e_{\boldsymbol{e}_1}, \dots, e_{\boldsymbol{e}_n}\}$:

$$\Phi(x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n) = x_1\Phi(\boldsymbol{e}_1) + \dots + x_n\Phi(\boldsymbol{e}_n) = x_1 e_{\boldsymbol{e}_1} + \dots + x_n e_{\boldsymbol{e}_n}.$$

Значит, Φ — биекция. $\square$

Соответствие $\boldsymbol{x} \mapsto e_{\boldsymbol{x}}$ позволяет отождествлять пространства V и V^{**}, причём для этого отождествления не требуется фиксировать базис или что бы то ни было ещё. Так что каждый вектор $\boldsymbol{x} \in V$ можно рассматривать как линейную функцию от функционалов из V^*, определённую правилом $\boldsymbol{x}(\boldsymbol{f}) = \boldsymbol{f}(\boldsymbol{x})$. Векторы и функционалы ведут себя по отношению друг к другу совершенно одинаково. Вместо $\boldsymbol{f}(\boldsymbol{x})$ часто пишут $(\boldsymbol{f}, \boldsymbol{x})$.

Упражнение

Пусть V — конечномерное векторное пространство. Зафиксируем вектор $\boldsymbol{x}_0 \in V$. Сопоставив каждому функционалу $\boldsymbol{f} \in V^*$ число $(\boldsymbol{f}, \boldsymbol{x}_0)$, мы получим линейную функцию $\eta_{\boldsymbol{x}_0} \colon V^* \to K$ $(\eta_{\boldsymbol{x}_0} \in V^{**} = V)$. Аналогично, зафиксировав функционал $\boldsymbol{f}_0 \in V^*$ и сопоставив каждому вектору $\boldsymbol{x} \in V$ число $(\boldsymbol{f}_0, \boldsymbol{x})$, мы получим линейную функцию $\theta_{\boldsymbol{f}_0} \colon V \to K$ $(\theta_{\boldsymbol{f}_0} \in V^*)$.

Покажите, что отображения

$$\mathrm{H} \colon V \to V \quad \text{и} \quad \Theta \colon V^* \to V^*,$$

определённые правилами $\mathrm{H}(\boldsymbol{x}) = \Phi^{-1}(\eta_{\boldsymbol{x}})$ и $\Theta(\boldsymbol{f}) = \theta_{\boldsymbol{f}}$ (здесь Φ — естественный изоморфизм $V \to V^{**}$, $\boldsymbol{x} \mapsto e_{\boldsymbol{x}}$), являются изоморфизмами.

Свёртка

Пусть U и V — векторные пространства над полем K. Операция (функция от двух аргументов)

$$c\colon U \times V \to K$$

называется свёрткой, или спариванием, если она билинейна, т.е. для $\boldsymbol{u}_1, \boldsymbol{u}_2 \in U$, $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$ и $\lambda \in K$

$$\begin{aligned} c(\boldsymbol{u}_1 + \boldsymbol{u}_2, \boldsymbol{v}_1) &= c(\boldsymbol{u}_1, \boldsymbol{v}_1) + c(\boldsymbol{u}_2, \boldsymbol{v}_1), \qquad & c(\lambda \boldsymbol{u}_1, \boldsymbol{v}_1) &= \lambda c(\boldsymbol{u}_1, \boldsymbol{v}_1), \\ c(\boldsymbol{u}_1, \boldsymbol{v}_1 + \boldsymbol{v}_2) &= c(\boldsymbol{u}_1, \boldsymbol{v}_1) + c(\boldsymbol{u}_2, \boldsymbol{v}_2), \qquad & c(\boldsymbol{u}_1, \lambda \boldsymbol{v}_1) &= \lambda c(\boldsymbol{u}_1, \boldsymbol{v}_1). \end{aligned}$$

В частном случае, когда $U = V^*$ (и $U^* = V$), возникает естественная свёртка $c(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}(\boldsymbol{v}) (= \boldsymbol{v}(\boldsymbol{u}))$. Вместо $\boldsymbol{u}(\boldsymbol{v})$ часто пишут $(\boldsymbol{u}, \boldsymbol{v})$.

В случае, когда пространства U и V имеют одинаковую конечную размерность n, скажем, что базисы $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ и $\{\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n\}$ пространств U и V взаимны, если

$$c(\boldsymbol{e}_i, \boldsymbol{\varepsilon}_j) = \delta_{ij} \qquad \text{для } i, j \leqslant n$$

(взаимные базисы существуют не для всякой свёртки).

Аннулятор

Пусть U и V — векторные пространства, $c\colon U \times V \to K$ — свёртка и $X \subset V$. Множество

$$\operatorname{Ann} X = \{\boldsymbol{u} \in U : c(\boldsymbol{u}, \boldsymbol{v}) = 0 \ \forall \boldsymbol{v} \in X\}$$

называется аннулятором множества X. В частном случае, когда $U = V^*$ и $c(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v})$, для $\operatorname{Ann} X$ используют обозначение X^0:

$$X^0 = \operatorname{Ann} X = \{\boldsymbol{f} \in V^* : (\boldsymbol{f}, \boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) = 0 \ \forall \boldsymbol{x} \in X\}.$$

В этом случае имеем также $V = U^*$. Для $F \subset V = U^*$ аннулятор $\operatorname{Ann} F$ обозначается F_0 и называется нуль-пространством множества F:

$$F_0 = \operatorname{Ann} F = \{\boldsymbol{x} \in U : (\boldsymbol{x}, \boldsymbol{f}) = \boldsymbol{x}(\boldsymbol{f}) = \boldsymbol{f}(\boldsymbol{x}) = 0 \ \forall \boldsymbol{f} \in F\}.$$

Упражнения

1. Покажите, что $\operatorname{Ann} X$ — подпространство пространства U для любого $X \subset V$.
2. Покажите, что $\operatorname{Ann} X = \operatorname{Ann}\langle X\rangle$ для любого $X \subset V$.
3. Покажите, что $\operatorname{Ann}\operatorname{Ann} X = \langle X\rangle$ для любого $X \subset V$.

Теорема

Пусть U и V — взаимно сопряжённые конечномерные векторные пространства с естественной свёрткой $c(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}(\boldsymbol{v}) = \boldsymbol{v}(\boldsymbol{u})$. Тогда для всякого множества $X \subset V$

$$\boxed{\operatorname{rank} X + \dim \operatorname{Ann} X = \dim V}.$$

Доказательство. Пусть $\operatorname{rank} X = r$ и $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r\}$ — максимальная линейно независимая система векторов в X. Дополним это систему до базиса $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ в пространстве V. Пусть $\{\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n\}$ — взаимный базис в V^*.

Вектор $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n \in V$ принадлежит $\langle X\rangle \iff x_{r+1} = x_{r+2} = \cdots = x_n = 0$ $\iff \boldsymbol{\varepsilon}_{r+1}(\boldsymbol{x}) = \boldsymbol{\varepsilon}_{r+2}(\boldsymbol{x}) = \cdots = \boldsymbol{\varepsilon}_n(\boldsymbol{x}) = 0$. Значит, $\boldsymbol{\varepsilon}_{r+1}, \boldsymbol{\varepsilon}_{r+2}, \ldots, \boldsymbol{\varepsilon}_n \in \operatorname{Ann}\langle X\rangle = \operatorname{Ann} X$, а поскольку $\operatorname{Ann} X = \langle \operatorname{Ann} X\rangle$, имеем $\langle\{\boldsymbol{\varepsilon}_{r+1}, \boldsymbol{\varepsilon}_{r+2}, \ldots, \boldsymbol{\varepsilon}_n\}\rangle \subset \operatorname{Ann} X$.

Пусть теперь $\boldsymbol{f} = f_1\boldsymbol{\varepsilon}_1 + \cdots + f_n\boldsymbol{\varepsilon}_n \in V^*$ и $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n \in \langle X\rangle$. Имеем $\boldsymbol{f}(\boldsymbol{x}) = f_1 \cdot x_1 + \cdots + f_n \cdot x_n$. Поскольку $\boldsymbol{e}_i \in X$ для $i \leqslant r$, видим, что если $\boldsymbol{f} \in \operatorname{Ann} X$, т.е. $\boldsymbol{f}(\boldsymbol{x}) = 0$ для всех $\boldsymbol{x} \in X$, то $\boldsymbol{f}(\boldsymbol{e}_i) = f_i = 0$ для $i \leqslant r$. Значит, всякий функционал из $\operatorname{Ann} X$ является линейной комбинацией функционалов $\boldsymbol{\varepsilon}_{r+1}, \ldots, \boldsymbol{\varepsilon}_n$, т.е. $\operatorname{Ann} X \subset \langle\{\boldsymbol{\varepsilon}_{r+1}, \boldsymbol{\varepsilon}_{r+2}, \ldots, \boldsymbol{\varepsilon}_n\}\rangle$.

$\operatorname{Ann} X = \langle \boldsymbol{\varepsilon}_{r+1}, \ldots, \boldsymbol{\varepsilon}_n\rangle$, функции $\boldsymbol{\varepsilon}_{r+1}, \ldots, \boldsymbol{\varepsilon}_n$ линейно независимы $\Rightarrow \dim \operatorname{Ann} X = n - r$. $\square$

Функционалы и подпространства

Определение

Ядром линейного функционала $\boldsymbol{f}$ на произвольном векторном пространстве V над любым полем K называется множество

$$\operatorname{Ker}\boldsymbol{f} = \{\boldsymbol{x} \in V : \boldsymbol{f}(\boldsymbol{x}) = 0\}.$$

Замечания

1. Ядро произвольного линейного функционала на векторном пространстве является подпространством этого пространства.
2. Для произвольного множества $F \subset V^*$ $\quad F_0 = \bigcap\limits_{\boldsymbol{f} \in F} \operatorname{Ker}\boldsymbol{f}$.
3. Пусть V — векторное пространство с базисом $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$. Если $\boldsymbol{f}$ — линейная функция на V с координатами $(f_1, \ldots, f_n)$ во взаимном базисе, то

$$\boldsymbol{x} = (x_1, \ldots, x_n) \in \operatorname{Ker}\boldsymbol{f} \iff \boldsymbol{f}(\boldsymbol{x}) = 0 \iff f_1 \cdot x_1 + \cdots + f_n \cdot x_n = 0.$$

Теорема

Пусть V — конечномерное векторное пространство над полем K. Тогда

1. *множество $U \subset V$ является подпространством пространства $V \iff U$ есть пересечение ядер некоторого конечного множества линейных функций;*
2. *если в V зафиксирован базис, то $U \subset V$ является подпространством пространства V (и $\dim U = k$) $\iff$ U есть множество векторов, координаты которых в этом базисе являются решениями некоторой конечной однородной системы линейных уравнений (и ранг матрицы этой системы равен $\dim V - k$).*

Доказательство. 1. Импликация $\Longleftarrow$ вытекает из того, что пересечение подпространств является подпространством. Докажем $\Longrightarrow$.

Возьмём любой базис $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k\}$ в U. Дополним его до базиса $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ в пространстве V. Пусть $\{\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n\}$ — взаимный базис в V^*. Вектор $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n \in V$ принадлежит $U \iff x_{k+1} = \cdots = x_n = 0 \iff$

$$\boldsymbol{\varepsilon}_{k+1}(\boldsymbol{x}) = \boldsymbol{\varepsilon}_{k+2}(\boldsymbol{x}) = \cdots = \boldsymbol{\varepsilon}_n(\boldsymbol{x}) = 0 \iff \boldsymbol{x} \in \operatorname{Ker} \boldsymbol{\varepsilon}_{k+2} \cap \cdots \cap \operatorname{Ker} \boldsymbol{\varepsilon}_n.$$

❷ Импликация $\Longleftarrow$ была доказана в курсе алгебры. Докажем $\Longrightarrow$.

Пусть $\boldsymbol{E}'$ — фиксированный базис в V и $\boldsymbol{\mathcal{E}}'$ — взаимный базис в V^*. Пользуясь доказанным пунктом ❶, найдём конечное множество $F = \{\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m\} \subset V^*$, для которого $U = F_0 = \operatorname{Ker} \boldsymbol{f}_1 \cap \cdots \cap \operatorname{Ker} \boldsymbol{f}_m$. Для каждого $i \leqslant n$ пусть $(f'_{i1}, \ldots, f'_{im})$ — координаты функционала $\boldsymbol{f}_i$ в базисе $\boldsymbol{\mathcal{E}}'$, и для произвольного вектора $\boldsymbol{x} \in V$ пусть $(x'_1, \ldots, x'_n)$ — его координаты в базисе $\boldsymbol{E}'$. Имеем

$$\boldsymbol{x} = (x'_1, \ldots, x'_n) \in U \iff \begin{cases} \boldsymbol{f}_1(\boldsymbol{x}) = f'_{11} \cdot x'_1 + \cdots + f'_{1n} \cdot x'_n = 0, \\ \boldsymbol{f}_2(\boldsymbol{x}) = f'_{21} \cdot x'_1 + \cdots + f'_{2n} \cdot x'_n = 0, \\ \qquad \ldots, \\ \boldsymbol{f}_m(\boldsymbol{x}) = f'_{m1} \cdot x'_1 + \cdots + f'_{mn} \cdot x'_n = 0. \end{cases}$$

По предыдущей теореме

$$\operatorname{rank}\{\boldsymbol{f}'_1, \ldots, \boldsymbol{f}'_m\} = \operatorname{rank} \begin{pmatrix} f'_{11} & \ldots & f'_{1n} \\ \vdots & \ddots & \vdots \\ f'_{m1} & \ldots & f'_{mn} \end{pmatrix} = \dim V - \dim F_0 = \dim V - \dim U.$$

□

Линейные отображения

Пусть V и W — векторные пространства над полем K.

Определение

Отображение $f: V \to W$ называется линейным, если f сохраняет операции сложения и умножения на число:

1. $f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$ для всех $\boldsymbol{x}, \boldsymbol{y} \in V$;
2. $f(\lambda \boldsymbol{x}) = \lambda f(\boldsymbol{x})$ для всех $\lambda \in K$ и $\boldsymbol{x} \in V$.

Простейшие свойства линейных отображений

- $f(\mathbf{0}) = \mathbf{0}$;
- $f(-\boldsymbol{x}) = -f(\boldsymbol{x})$ для любого $\boldsymbol{x} \in V$;
- $f(\boldsymbol{x}-\boldsymbol{y}) = f(\boldsymbol{x}) - f(\boldsymbol{y})$ для любых $\boldsymbol{x}, \boldsymbol{y} \in V$.

Упражнение

Докажите эти свойства.

Примеры

- Линейные функции являются линейными отображениями из векторного пространства V над полем K в поле K, рассматриваемое как векторное пространство над самим собой.
- Поворот обычной плоскости.
- Если $V = U \oplus W$, то проектирование на подпространство U параллельно W можно рассматривать как отображение $V \to V$ и как отображение $V \to U$. В обоих случаях это линейное отображение.
- Дифференцирование — линейное отображение пространства (непрерывно) дифференцируемых функций на заданном промежутке числовой прямой в пространство всех (непрерывных) функций на этом промежутке.

Замечание

Пусть V и U — любые векторные пространства над одним и тем же полем K. На множестве $\mathscr{L}(V, U)$ линейных отображений $V \to U$ определены поточечные операции сложения и умножения на число: $(f+g)(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x})$, $\lambda f(\boldsymbol{x}) = f(\lambda \boldsymbol{x})$ для $\boldsymbol{x} \in V$, $\lambda \in K$. Относительно этих операций $\mathscr{L}(V, U)$ является векторным пространством.

Матрица линейного отображения

Всякое линейное отображение $f: V \to U$ однозначно определяется образами базисных векторов пространства V. В самом деле, если $\boldsymbol{E} = \{\boldsymbol{e}_\alpha : \alpha \in A\}$ (A — любое индексное множество) — базис пространства V, то для любого целого $n \leqslant |A|$, любых $\alpha_1, \ldots, \alpha_n \in A$ и любого вектора $\boldsymbol{x} = x_1 \boldsymbol{e}_{\alpha_1} + \cdots + x_n \boldsymbol{e}_{\alpha_n} \in V$ имеем

$$f(\boldsymbol{x}) = f(x_1 \boldsymbol{e}_{\alpha_1} + \cdots + x_n \boldsymbol{e}_{\alpha_n}) = x_1 f(\boldsymbol{e}_{\alpha_1}) + \cdots + x_n f(\boldsymbol{e}_{\alpha_n}).$$

Образами базисных векторов $\boldsymbol{e}_\alpha$ могут быть любые векторы $\boldsymbol{u}_\alpha \in U$, $\alpha \in A$. Действительно, для любых $\boldsymbol{u}_\alpha \in U$, $\alpha \in A$, отображение, определённое правилом

$$f(x_1 \boldsymbol{e}_{\alpha_1} + \cdots + x_n \boldsymbol{e}_{\alpha_n}) = x_1 \boldsymbol{u}_{\alpha_1} + \cdots + x_n \boldsymbol{u}_{\alpha_n}$$

для каждого $x_1 \boldsymbol{e}_{\alpha_1} + \cdots + x_n \boldsymbol{e}_{\alpha_n} \in V$, линейно, и $f(\boldsymbol{e}_\alpha) = \boldsymbol{u}_\alpha$.

То же рассуждение в конечномерном случае выглядит так:
Если $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ — базис пространства V, то для любого вектора $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n \in V$ имеем

$$f(\boldsymbol{x}) = f(x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n) = x_1 f(\boldsymbol{e}_1) + \cdots + x_n f(\boldsymbol{e}_n).$$

Образами базисных векторов могут быть любые векторы $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in U$: для поизвольных $\boldsymbol{u}_i \in U$ отображение, определённое правилом

$$f(x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n) = x_1\boldsymbol{u}_1 + \cdots + x_n\boldsymbol{u}_n$$

для каждого $x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n \in V$, линейно, и $f(\boldsymbol{e}_i) = \boldsymbol{u}_i$.

Пусть V и U конечномерны, $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ — базис V и $\boldsymbol{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m\}$ — базис U. Для каждого $j = 1, \ldots, n$ образ $f(\boldsymbol{e}_j)$ вектора $\boldsymbol{e}_j$ принадлежит пространству U, а значит, его можно разложить по базису $\boldsymbol{B}$ этого пространства:

$$f(\boldsymbol{e}_j) = a_{1j}\boldsymbol{b}_1 + \cdots + a_{mj}\boldsymbol{b}_m.$$

Матрица $A = (a_{ij})$ размера $m \times n$ называется матрицей линейного отображения f относительно базисов $\boldsymbol{E}$ и $\boldsymbol{B}$. Столбцы матрицы линейного отображения — это столбцы координат векторов $f(\boldsymbol{e}_1), \ldots, f(\boldsymbol{e}_n)$ в базисе $\boldsymbol{B}$.

Упражнения

Пусть V и U — конечномерные векторные пространства размерностей n и m соответственно. Зафиксируем в каждом из них произвольный базис.

1. Заметьте, что любая матрица размера $m \times n$ является матрицей некоторого линейного отображения $V \to U$ относительно зафиксированных базисов.
2. Покажите, что отображение, которое каждому линейному отображению $V \to U$ ставит в соответствие его матрицу относительно зафиксированных базисов, является изоморфизмом между векторными пространствами всех линейных отображений $V \to U$ и всех матриц размера $m \times n$.

Для произвольного вектора $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n$ из пространства V разложим его образ $\boldsymbol{y} = f(\boldsymbol{x})$ по базису $\boldsymbol{B}$ пространства U:

$$\boldsymbol{y} = y_1\boldsymbol{b}_1 + \dots + y_m\boldsymbol{b}_m.$$

С другой стороны, имеем

$$\boldsymbol{y} = f(\boldsymbol{x}) = f\Big(\sum_{j=1}^{n} x_j\boldsymbol{e}_j\Big) = \sum_{j=1}^{n} f(x_j\boldsymbol{e}_j) = \sum_{j=1}^{n} x_j f(\boldsymbol{e}_j) = \sum_{j=1}^{n} x_j\Big(\sum_{i=n}^{m} a_{ij}\boldsymbol{b}_i\Big) = \sum_{i=1}^{m}\Big(\sum_{j=n}^{m} a_{ij}x_j\Big)\boldsymbol{b}_i.$$

Единственность разложения вектора по базису $\Longrightarrow$ $y_i = \sum_{j=n}^{m} a_{ij}x_j$ для каждого $i = 1, \dots, m$. В матричной форме:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Когда базисы в пространствах V и U фиксированы и векторы отождествляются со столбцами их координат в этих базисах, пишут просто

$$\boldsymbol{y} = \boxed{f(\boldsymbol{x}) = A\boldsymbol{x}}.$$

Пример

Пусть $V = U \oplus W$, $\dim V = 7$, $\dim U = 4$. Возьмём базис $\boldsymbol{E}' = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4\}$ в U и дополним его до базиса $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_7\}$ в V. Для проектирования $\pi\colon V \to U$ имеем $\pi(\boldsymbol{e}_i) = \boldsymbol{e}_i$ для $i \leqslant 4$ и $\pi(\boldsymbol{e}_i) = \boldsymbol{0}$ для $i = 5, 6, 7$. Значит, матрица π относительно базисов $\boldsymbol{E}$ и $\boldsymbol{E}'$ такова:

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

Преобразование матрицы линейного отображения при замене базисов

Пусть V и U — конечномерные векторные пространства с базисами $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ и $\boldsymbol{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m\}$, и пусть $f\colon V \to U$ — линейное отображение с матрицей A относительно базисов $\boldsymbol{E}$ и $\boldsymbol{B}$. Предположим, что $\boldsymbol{E}' = \{\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n\} \subset V$ и $\boldsymbol{B}' = \{\boldsymbol{b}'_1, \ldots, \boldsymbol{b}'_m\} \subset U$ — другие базисы и A' — матрица f относительно базисов $\boldsymbol{E}'$ и $\boldsymbol{B}'$. Тогда для любого вектора $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n = x'_1\boldsymbol{e}'_1 + \cdots + x'_n\boldsymbol{e}'_n \in V$ и его образа $\boldsymbol{y} = f(\boldsymbol{x}) = y_1\boldsymbol{b}_1 + \cdots + y_m\boldsymbol{b}_m = y'_1\boldsymbol{b}'_1 + \cdots + y'_m\boldsymbol{b}'_m$ имеем

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} y'_1 \\ \vdots \\ y'_m \end{pmatrix} = A' \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}.$$

Если T — матрица перехода от $\boldsymbol{E}$ к $\boldsymbol{E}'$ и S — матрица перехода от $\boldsymbol{B}$ к $\boldsymbol{B}'$, то

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = T \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = S \begin{pmatrix} y'_1 \\ \vdots \\ y'_m \end{pmatrix}.$$

Значит,

$$S \begin{pmatrix} y'_1 \\ \vdots \\ y'_m \end{pmatrix} = AT \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} \quad \Longrightarrow \quad \begin{pmatrix} y'_1 \\ \vdots \\ y'_m \end{pmatrix} = S^{-1}AT \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} \quad \Longrightarrow \quad A' \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = S^{-1}AT \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}.$$

Подставляя вместо $\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ столбцы $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$, видим, что $\boxed{A' = S^{-1}AT}$.

В курсе алгебры была доказана лемма:

Лемма

При умножении матрицы на невырожденную матрицу её ранг не изменяется.

Из неё и формулы преобразования матрицы линейного отображения при замене базиса немедленно следует

Теорема

Ранг матрицы линейного отображения конечномерных векторных пространств не зависит от выбора базисов в этих пространствах.

Ядро и образ линейного отображения

Определение

Ядром линейного отображения $f\colon V \to U$ векторных пространств называется множество

$$\operatorname{Ker} f = \{\boldsymbol{x} \in V : f(\boldsymbol{x}) = \mathbf{0}_U\}.$$

Образом линейного отображения $f\colon V \to U$ векторных пространств называется множество

$$\operatorname{Im} f = \{\boldsymbol{y} \in U : \exists \boldsymbol{x} \in V,\ \text{для которого} f(\boldsymbol{x}) = \boldsymbol{y}\}.$$

Предложение

Ядро $\operatorname{Ker} f$ *любого линейного отображения* $f\colon V \to U$ *является подпространством пространства* V, *а образ* $\operatorname{Im} f$ *— подпространством пространства* U.

Доказательство. 1. $\operatorname{Ker} f \neq \varnothing$, так как $\mathbf{0}_V \in \operatorname{Ker} f$. Если $\boldsymbol{u}, \boldsymbol{v} \in \operatorname{Ker} f$ и λ — число, то $f(\boldsymbol{u})+f(\boldsymbol{v}) = \mathbf{0}_U + \mathbf{0}_U = \mathbf{0}_U$ и $f(\lambda\boldsymbol{u}) = \lambda f(\boldsymbol{u}) = \lambda\mathbf{0}_U = \mathbf{0}_U$, т.е. $\boldsymbol{u}+\boldsymbol{v} \in \operatorname{Ker} f$ и $\lambda\boldsymbol{u} \in \operatorname{Ker} f$.
2. $\operatorname{Im} f \neq \varnothing$, так как $\mathbf{0}_U = f(\mathbf{0}_V) \in \operatorname{Im} f$. Если $\boldsymbol{u}, \boldsymbol{v} \in \operatorname{Im} f$ (т.е. $\boldsymbol{u} = f(\boldsymbol{u}')$ и $\boldsymbol{v} = f(\boldsymbol{v}')$ для некоторых $\boldsymbol{u}', \boldsymbol{v}' \in V$) и λ — число, то $\boldsymbol{u}+\boldsymbol{v} = f(\boldsymbol{u}'+\boldsymbol{v}')$ и $f(\lambda\boldsymbol{u}) = f(\lambda\boldsymbol{u}')$, откуда $\boldsymbol{u}+\boldsymbol{v} \in \operatorname{Im} f$ и $\lambda\boldsymbol{u} \in \operatorname{Im} f$. $\square$

Упражнения

1. Опишите всевозможные ядра и образы линейных отображений из векторного пространства V над полем K в поле K, рассматриваемое как векторное пространство над самим собой.
2. Найдите ядро и образ поворота обычной плоскости.
3. Пусть $V = U \oplus W$. Найдите ядро и образ проектирования на подпространство U параллельно W, рассматриваемое а) как отображение $V \to V$ и б) как отображение $V \to U$.
4. Найдите ядро и образ линейного отображения дифференцирования пространства непрерывно дифференцируемых функций на заданном промежутке числовой прямой в пространство всех непрерывных функций на этом промежутке.

Предложение

Линейное отображение $f\colon V \to U$ инъективно тогда и только тогда, когда $\operatorname{Ker} f = \{\mathbf{0}\}$.

Доказательство. $\Longrightarrow$: $f(\mathbf{0}_V) = \mathbf{0}_U$ и f инъективно $\Longrightarrow$ $f(\boldsymbol{x}) \ne \mathbf{0}_U$ для $\boldsymbol{x} \ne \mathbf{0}_V$.
$\Longleftarrow$: Пусть $\operatorname{Ker} f = \{\mathbf{0}_V\}$. Для любых $\boldsymbol{x}, \boldsymbol{y} \in V$ имеем $f(\boldsymbol{x}) - f(\boldsymbol{y}) = f(\boldsymbol{x} - \boldsymbol{y})$, поэтому если $f(\boldsymbol{x}) = f(\boldsymbol{y})$, то $f(\boldsymbol{x} - \boldsymbol{y}) = \mathbf{0}_U$, а значит, $\boldsymbol{x} - \boldsymbol{y} \in \operatorname{Ker} f$, так что если $\operatorname{Ker} f = \{\mathbf{0}_V\}$, то $\boldsymbol{x} - \boldsymbol{y} = \mathbf{0}_V$, откуда $\boldsymbol{x} = \boldsymbol{y}$. □

Предложение

Если $f\colon V \to U$ — линейное отображение, $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n \in V$ и векторы $f(\boldsymbol{v}_1), \dots, f(\boldsymbol{v}_n)$ линейно независимы (в пространстве U), то и векторы $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ линейно независимы (в пространстве V).

Доказательство. Пусть $\lambda_1 \boldsymbol{v}_1 + \dots + \lambda_n \boldsymbol{v}_n = \mathbf{0}_V$. Тогда $f(\lambda_1 \boldsymbol{v}_1 + \dots + \lambda_n \boldsymbol{v}_n) = \lambda_1 f(\boldsymbol{v}_1) + \dots + \lambda_n f(\boldsymbol{v}_n) = f(\mathbf{0}_V) = \mathbf{0}_U$. Поскольку векторы $f(\boldsymbol{v}_1), \dots, f(\boldsymbol{v}_n)$ линейно независимы, имеем $\lambda_1 = \dots = \lambda_n = 0$. □

Лемма

Если V — конечномерное векторное пространство над полем K с базисом $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$, U — векторное пространство над K и $f\colon V \to U$ — линейное отображение, то

$$\operatorname{Im} f = \langle f(\boldsymbol{e}_1), \ldots, f(\boldsymbol{e}_n)\rangle.$$

Доказательство. Для любых чисел $\lambda_1, \ldots, \lambda_n \in K$ имеем

$$\lambda_1 f(\boldsymbol{e}_1) + \cdots + \lambda_n f(\boldsymbol{e}_n) = f(\lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_n \boldsymbol{e}_n).$$

□

Лемма

Пусть V и U — конечномерные векторные пространства с базисами $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ и $\boldsymbol{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m\}$, и пусть $f\colon V \to U$ — линейное отображение с матрицей A относительно базисов $\boldsymbol{E}$ и $\boldsymbol{B}$. Тогда

$$\dim \operatorname{Im} f = \operatorname{rank} A.$$

Доказательство. Достаточно вспомнить, что столбцы матрицы A — это координаты векторов $f(\boldsymbol{e}_1), \ldots, f(\boldsymbol{e}_n)$ в базисе $\boldsymbol{B}$, и применить предыдущую лемму.

□

Теорема

Для любого линейного отображения $f\colon V \to U$ конечномерного векторного пространства V

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

Доказательство. Выберем базис $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k\}$ в $\operatorname{Ker} f$ и дополним его до базиса $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ во всём пространстве V. Имеем $f(\boldsymbol{e}_1) = \cdots = f(\boldsymbol{e}_k) = \mathbf{0}_U$. По лемме

$$\operatorname{Im} f = \langle\{f(\boldsymbol{e}_{k+1}), \ldots, f(\boldsymbol{e}_n)\}\rangle.$$

Надо доказать, что векторы $f(\boldsymbol{e}_{k+1}), \ldots, f(\boldsymbol{e}_n)$ линейно независимы.

Пусть $\lambda_{k+1} f(\boldsymbol{e}_{k+1}) + \cdots + \lambda_n f(\boldsymbol{e}_n) = \mathbf{0}_U$. Рассмотрим вектор $\boldsymbol{x} = \lambda_{k+1}\boldsymbol{e}_{k+1} + \cdots + \lambda_n \boldsymbol{e}_n$. Имеем $f(\boldsymbol{x}) = \mathbf{0}_U$, т.е. $\boldsymbol{x} \in \operatorname{Ker} f = \langle\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k\}\rangle$. Значит, для некоторых $\lambda_1, \ldots, \lambda_k$ имеем $\boldsymbol{x} = \lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_k \boldsymbol{e}_k = \lambda_{k+1}\boldsymbol{e}_{k+1} + \cdots + \lambda_n \boldsymbol{e}_n$, и из линейной независимости базисных векторов $\boldsymbol{e}_i$ вытекает, что $\lambda_1 = \cdots = \lambda_n = 0$. □

Следствие

Линейное отображение $f\colon V \to U$ сюръективно тогда и только тогда, когда $\dim U = \dim V - \dim \operatorname{Ker} f$.

Факторпространство

Пусть V — векторное пространство над полем K и U — его подпространство. Будем говорить, что векторы $\boldsymbol{x}, \boldsymbol{y} \in V$ *сравнимы по подпространству* U, и писать $\boldsymbol{x} \equiv \boldsymbol{y}(U)$, если $\boldsymbol{x} - \boldsymbol{y} \in U$. Ясно, что отношение сравнимости по подпространству является отношением эквивалентности на V. Класс эквивалентности $[\boldsymbol{v}]$ вектора $\boldsymbol{v} \in V$ — это множество $\boldsymbol{v} + U$. Если $\boldsymbol{x} \equiv \boldsymbol{x}'(U)$ и $\boldsymbol{y} \equiv \boldsymbol{y}'(U)$, то для любых $\lambda, \mu \in K$ имеем

$$\lambda\boldsymbol{x} + \mu\boldsymbol{y} - (\lambda\boldsymbol{x}' + \mu\boldsymbol{y}') = \lambda(\boldsymbol{x} - \boldsymbol{x}') + \mu(\boldsymbol{y} - \boldsymbol{y}') \in U,$$

т.е. $\lambda\boldsymbol{x} + \mu\boldsymbol{y} \equiv \lambda\boldsymbol{x}' + \mu\boldsymbol{y}'(U)$. Следовательно, на классах эквивалентности корректно определены операции сложения и умножения на скаляр правилами

$$[\boldsymbol{x}] + [\boldsymbol{y}] = [\boldsymbol{x} + \boldsymbol{y}] \qquad \text{и} \qquad \lambda[\boldsymbol{x}] = [\lambda\boldsymbol{x}]$$

(«корректно» означает, что это определение не зависит от выбора представителей классов эквивалентности).

Относительно этих операций факторамножество V/U всех классов эквивалентности — векторное пространством над полем K. Оно называется факторпространством пространства V по подпространству U. Отображение $V \to V/U$, которое каждому вектору $\boldsymbol{x}$ ставит в соответствие его класс эквивалентности, называется естественным отображением. Ясно, что оно линейно.

Ядром естественного отображения $V \to V/U$ является подпространство U, поэтому в случае, когда V конечномерно, по доказанной раньше теореме имеем

$$\boxed{\dim V/U = \dim V - \dim U}\,.$$

Всякое сюръективное линейное отображение $f: V \to U$ можно представить (с точностью до изоморфизма) как естественное отображение при факторизации по ядру этого отображения. Точнее, имеет место следующая теорема.

Теорема

Пусть $f: V \to U$ — линейное отображение. Тогда отображение $\varphi: V/\operatorname{Ker} f \to \operatorname{Im} f$, определённое правилом $[\boldsymbol{v}] = \boldsymbol{v} + \operatorname{Ker} f \mapsto f(\boldsymbol{v})$, является изоморфизмом.

Доказательство. Проверим, что отображение φ определено корректно, т.е. если $[\boldsymbol{v}] = [\boldsymbol{u}]$, то $f(\boldsymbol{v}) = f(\boldsymbol{v})$. Это так: если $[\boldsymbol{v}] = [\boldsymbol{u}]$, то $\boldsymbol{v} - \boldsymbol{u} \in \operatorname{Ker} f$, а значит, $f(\boldsymbol{v}) - f(\boldsymbol{u}) = f(\boldsymbol{v} - \boldsymbol{u}) = \mathbf{0}$.

Линейность и сюръективность φ очевидны. Проверим инъективность. Если $\varphi([\boldsymbol{v}]) = \varphi([\boldsymbol{u}])$, то $f(\boldsymbol{v}) = f(\boldsymbol{u})$, откуда $f(\boldsymbol{v}) - f(\boldsymbol{u}) = f(\boldsymbol{v} - \boldsymbol{u}) = \mathbf{0}$ и $\boldsymbol{v} - \boldsymbol{u} \in \operatorname{Ker} f$, т.е. $\boldsymbol{v} - \boldsymbol{u} \in \operatorname{Ker} f$. Значит, $[\boldsymbol{v}] = \boldsymbol{v} + \operatorname{Ker} f = \boldsymbol{u} + \operatorname{Ker} f = [\boldsymbol{u}]$. □

Замечание

Векторы $[\boldsymbol{x}_1], \ldots, [\boldsymbol{x}_n] \in V/U$ линейно независимы тогда и только тогда, когда из $\lambda_1 \boldsymbol{x}_1 + \cdots + \lambda_n \boldsymbol{x}_n \in U$ вытекает, что $\lambda_1 = \cdots = \lambda_n = 0$. Следовательно, если $[\boldsymbol{x}_1], \ldots, [\boldsymbol{x}_n]$ линейно независимы в V/U, то $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ линейно независимы в V.

Упражнение

Покажите, что если $\boldsymbol{E}$ — базис в U и $\boldsymbol{E} \cup \boldsymbol{E}'$ — базис в V, причём $\boldsymbol{E} \cap \boldsymbol{E}' = \varnothing$, то $\{[\boldsymbol{e}'] : \boldsymbol{e}' \in \boldsymbol{E}'\}$ — базис в V/U.

Сопряжённые линейные отображения

Пусть V и U — линейные пространства и V^* и U^* — сопряжённые к ним пространства. Тогда для любого линейного отображения $f : V \to U$ и любого линейного функционала $\varphi \in U^*$ композиция φ и f является линейным функционалом на V: $(\varphi \circ f)(\boldsymbol{x}) = \varphi(f(\boldsymbol{x})) \in K$ для всякого $\boldsymbol{x} \in V$. Соответствие $\varphi \mapsto \varphi \circ f$ линейно по φ.

Определение

Для любого линейного отображения $f : V \to U$ линейное отображение $U^* \to V^*$, определённое правилом $\varphi \mapsto \varphi \circ f$, называется сопряжённым отображением по отношению к f и обозначается f^*.

Действие сопряжённого отображения можно записать как $(f^*(\varphi), \boldsymbol{x}) = (\varphi, f(\boldsymbol{x}))$ для $\boldsymbol{x} \in V$ и $\varphi \in U^*$, где $(\varphi, \boldsymbol{x}) = \varphi(\boldsymbol{x})$.
Правило $f \mapsto f^*$ определяет отображение $* : \mathscr{L}(V, U) \to \mathscr{L}(U^*, V^*)$ пространств линейных отображений. Оно, очевидно, линейно.

Упражнение

Покажите, что отображение $*$ всегда инъективно, но не всегда является изоморфизмом.

Пусть V и U конечномерны, $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ — базис V и $\boldsymbol{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m\}$ — базис U, и пусть $\boldsymbol{\mathcal{E}} = \{\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n\}$ и $\boldsymbol{\mathcal{B}} = \{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m\}$ — взаимные с ними базисы сопряжённых пространств V^* и U^*. Рассмотрим линейное отображение $f\colon V \to U$ с матрицей $A = (a_{ij})$ относительно базисов $\boldsymbol{E}$ и $\boldsymbol{B}$. Для любых $i \leqslant n$ и $j = 1, \ldots, m$ имеем

$$f^*(\boldsymbol{\beta}_j)(\boldsymbol{e}_i) = \boldsymbol{\beta}_j(f(\boldsymbol{e}_i)) = a_{ji}$$

(надо вспомнить, что координаты вектора $f(\boldsymbol{e}_i)$ в базисе $\boldsymbol{B}$ составляют i-й столбец матрицы A и элемент $\boldsymbol{\beta}_j$ взаимного базиса $\boldsymbol{\mathcal{B}}$ ставит в соответствие каждому вектору из U его j-ю координату в базисе $\boldsymbol{B}$). Значит, матрица A^* отображения f^*, сопряжённого к отображению f с матрицей A, — это транспонированная матрица A:

$$\boxed{A^* = A^T}.$$

Следовательно, в конечномерном случае $*$ — изоморфизм.

▸ к сопряжённым операторам

Комплексификация

Пусть $(V, +, \cdot)$ — векторное пространство над $\mathbb{R}$. Векторное пространство $V_{\mathbb{C}} = (V \times V, +_{\mathbb{C}}, \cdot_{\mathbb{C}})$ над $\mathbb{C}$, элементами которого являются пары $(\boldsymbol{x}, \boldsymbol{y}) \in V \times V$, а операции определены правилами

$$(\boldsymbol{x}, \boldsymbol{y}) + (\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{x} + \boldsymbol{u}, \boldsymbol{y} + \boldsymbol{v}) \quad \text{и} \quad (\lambda + i\mu)(\boldsymbol{x}, \boldsymbol{y}) = (\lambda\boldsymbol{x} - \mu\boldsymbol{y}, \mu\boldsymbol{x} + \lambda\boldsymbol{y})$$

(для краткости мы пишем $+$ вместо $+_{\mathbb{C}}$), называется комплексификацией пространства V. При этом векторы пространства V отождествляются с векторами вида $(\boldsymbol{x}, \boldsymbol{0})$ в пространстве $V_{\mathbb{C}}$.

Для любых $\boldsymbol{x}, \boldsymbol{y} \in V$ имеем $(\boldsymbol{x}, \boldsymbol{0}) + (\boldsymbol{0}, \boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{0}) + i(\boldsymbol{y}, \boldsymbol{0})$. Поэтому каждый вектор $(\boldsymbol{x}, \boldsymbol{y}) \in V_{\mathbb{C}}$ однозначно представляется в виде $\boldsymbol{x} + i\boldsymbol{y}$:

$$\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} \rightleftarrows \boldsymbol{x} + i\boldsymbol{y}.$$

Теорема

Любой базис $\boldsymbol{E}$ *пространства* V *является одновременно и базисом пространства* $V_{\mathbb{C}}$*.*

Доказательство. То, всякий вектор $\boldsymbol{x} + i\boldsymbol{y} \in V_{\mathbb{C}}$ выражается через $\boldsymbol{E}$ (с комплексными коэффициентами), очевидно. Проверим, что множество $\boldsymbol{E}$ линейно независимо над $\mathbb{C}$. Любая линейная комбинация элементов $\boldsymbol{E}$ с коэффициентами из $\mathbb{C}$ имеет вид $\sum_{i\leqslant n}(\lambda_i + i\mu_i)\boldsymbol{e}_{r_i}$, где $\boldsymbol{e}_{r_i}$ — попарно различные векторы из $\boldsymbol{E}$ и $\lambda_i, \mu_i \in \mathbb{R}$. Если $\sum_{i\leqslant n}(\lambda_i + i\mu_i)\boldsymbol{e}_{r_i} = \boldsymbol{0}$, то $\sum_{i\leqslant n}\lambda_i\boldsymbol{e}_{r_i} + i\sum_{i\leqslant n}\mu_i\boldsymbol{e}_{r_i} = \boldsymbol{0}$, т.е. $(\sum_{i\leqslant n}\lambda_i\boldsymbol{e}_{r_i}, \sum_{i\leqslant n}\mu_i\boldsymbol{e}_{r_i}) = (\boldsymbol{0}, \boldsymbol{0})$, и из линейной независимости векторов $\boldsymbol{e}_{r_i}$ над $\mathbb{R}$ следует, что $\lambda_i = \mu_i = 0$ для всех $i \leqslant n$. □

Следствие

Размерность комплексификации $V_{\mathbb{C}}$ *любого векторного пространства* V *над* $\mathbb{R}$ *равна размерности пространства* V*.*

Не всякий базис пространства $V_{\mathbb{C}}$ является одновременно базисом пространства V. Однако в любом базисе, который является таковым, координаты любого вектора $\boldsymbol{x} \in V$ как элемента $V_{\mathbb{C}}$ совпадают с его координатами как элемента V.

Пусть V и U — два векторных пространства над $\mathbb{R}$. Произвольное линейное отображение $f : V \to U$ однозначно продолжается до линейного отображения $f_{\mathbb{C}} : V_{\mathbb{C}} \to U_{\mathbb{C}}$ по формуле

$$f_{\mathbb{C}}(\boldsymbol{x} + i\boldsymbol{y}) = f(\boldsymbol{x}) + if(\boldsymbol{y}).$$

Отображение $f_{\mathbb{C}}$ называется комплексификацией отображения f.

Предположим, что пространства V и U конечномерны. Вспомним, что всякий базис $\boldsymbol{E}$ в V (базис $\boldsymbol{B}$ в U) является одновременно базисом в $V_{\mathbb{C}}$ (в $U_{\mathbb{C}}$), а матрица линейного отображения состоит из столбцов координат образов базисных векторов. Поскольку $f(\boldsymbol{e}) \in U$ для любого $\boldsymbol{e} \in \boldsymbol{E}$, а координаты в базисе $\boldsymbol{B}$ любого вектора $\boldsymbol{x} \in U$ как элемента U сопадают с координатами в том же базисе этого вектора как элемента $V_{\mathbb{C}}$, видим, что матрица отображения $f_{\mathbb{C}}$ совпадает с матрицей отображения f.

Овеществление

Пусть $(V, +, \cdot)$ — векторное пространство над $\mathbb{C}$. Забудем про возможность умножать векторы из V на комплексные числа и оставим только умножение на числа из $\mathbb{R}$, т.е. заменим операцию $\cdot : V \times \mathbb{C} \to V$ на её ограничение $\cdot_{\mathbb{R}} : V \times \mathbb{R} \to V$. Множество V и операцию $+$ оставим без изменений. В результате мы получим векторное пространство $(V, +, \cdot_{\mathbb{R}})$ над $\mathbb{R}$, которое будем обозначать $V_{\mathbb{R}}$. Оно называется овеществлением пространства V.

Теорема

Если $\boldsymbol{E}$ — базис векторного пространства V над полем $\mathbb{C}$, то $\boldsymbol{E} \cup i\boldsymbol{E}$ — базис $V_{\mathbb{R}}$.

(Здесь $i\boldsymbol{E} = \{i\boldsymbol{e} : \boldsymbol{e} \in \boldsymbol{E}\}$.)

Доказательство. Для каждого вектора $\boldsymbol{x} \in V$ (однозначно) определены векторы $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n \in \boldsymbol{E}$ и числа $x_1 = y_1 + iz_1, \dots, x_n = y_n + iz_n \in \mathbb{C}$ такие, что
$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n = (y_1 + iz_1)\boldsymbol{e}_1 + \dots + (y_n + iz_n)\boldsymbol{e}_1 = y_1\boldsymbol{e}_1 + \dots + y_n\boldsymbol{e}_n + z_1(i\boldsymbol{e}_1) + \dots + z_n(i\boldsymbol{e}_1).$$
Значит, $\langle \boldsymbol{E} \cup i\boldsymbol{E} \rangle = V$. Множество $\boldsymbol{E} \cup i\boldsymbol{E}$ линейно независимо над $\mathbb{R}$. Действительно, любая линейная комбинация векторов из $\boldsymbol{E} \cup i\boldsymbol{E}$ представима как $\sum_{i \leqslant n} \alpha_{r_i}\boldsymbol{e}_{r_i} + \sum_{i \leqslant m} \beta_{s_i}(i\boldsymbol{e}_{s_i})$, где $\boldsymbol{e}_{r_1}, \dots, \boldsymbol{e}_{r_n}, \boldsymbol{e}_{s_1}, \dots, \boldsymbol{e}_{s_m} \in \boldsymbol{E}$ (причём $\boldsymbol{e}_{r_i} \neq \boldsymbol{e}_{r_j}$ для $i \neq j$ и $\boldsymbol{e}_{s_i} \neq \boldsymbol{e}_{s_j}$ для $i \neq j$) и $\alpha_{r_1}, \dots, \alpha_{r_n}, \beta_{s_1}, \dots, \beta_{s_m} \in \mathbb{R}$. Для каждого индекса r_i, который не встречается среди индексов s_j, положим $\beta_{r_i} = 0$, и для каждого индекса s_j, который не встречается среди индексов r_i, положим $\alpha_{r_i} = 0$. В результате мы сможем записать нашу линейную комбинацию в виде $(\alpha_{t_1} + i\beta_{t_1})\boldsymbol{e}_{t_1} + \dots + (\alpha_{t_k} + i\beta_{t_k})\boldsymbol{e}_{t_k}$, где $\{t_1, \dots, t_k\} = \{r_1, \dots, r_n\} \cup \{s_1, \dots, s_m\}$ и все $\boldsymbol{e}_{t_i}$ разные. Из линейной независимости векторов $\boldsymbol{e}_{t_1}, \dots, \boldsymbol{e}_{t_k}$ над $\mathbb{C}$ следует, что если наша линейная комбинация равна $\boldsymbol{0}$, то $\alpha_{t_i} = \beta_{t_i} = 0$ для всех $i \leqslant k$, а значит, $\alpha_{r_i} = 0$ для $i \leqslant n$ и $\beta_{s_i} = 0$ для $i \leqslant m$. □

Упражнение

Покажите, что если V — векторное пространство над $\mathbb{C}$, $\boldsymbol{E} = \{\boldsymbol{e}_\iota : \iota \in I\}$ (I — любое множество индексов) — его базис и вектор $\boldsymbol{x} \in V$ имеет в этом базисе координаты $x_{\iota_1} = y_{\iota_1} + iz_{\iota_1}, \ldots, x_{\iota_n} = y_{\iota_n} + iz_{\iota_n}$, то в качестве элемента пространства $V_{\mathbb{R}}$ над $\mathbb{R}$ тот же вектор имеет в базисе $\boldsymbol{E} \cup i\boldsymbol{E}$ координаты $y_{\iota_1}, \ldots, y_{\iota_n}, z_{\iota_1}, \ldots, z_{\iota_n}$.

Следствие

Если V — конечномерное векторное пространство над полем $\mathbb{C}$ и $\dim V = n$, то $\dim V_{\mathbb{R}} = 2n$.

Пусть теперь V и U — два векторных пространства над $\mathbb{C}$ и $f : V \to U$ — линейное отображение. Это же самое отображение можно рассматривать также и как отображение $f : V_{\mathbb{R}} \to U_{\mathbb{R}}$ (потому что несущие множества у пространств V и $V_{\mathbb{R}}$ и у пространств U и $U_{\mathbb{R}}$ одни и те же). Относительно операций $+$ и $\cdot_{\mathbb{R}}$ оно остаётся линейным. Поэтому в случае, когда U и V конечномерны, определена не только его матрица A как отображения $V \to U$, но и матрица $A_{\mathbb{R}}$ как отображения $V_{\mathbb{R}} \to U_{\mathbb{R}}$.

Теорема

Пусть V и U — два конечномерных векторных пространства над $\mathbb{C}$ с базисами $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ и $\boldsymbol{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m\}$ соответственно, и пусть $f : V \to U$ — линейное отображение с матрицей $A = B + iC$ относительно этих базисов. Тогда матрица f как отображения $V_{\mathbb{R}} \to U_{\mathbb{R}}$ относительно базисов $\boldsymbol{E}_{\mathbb{R}} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n, i\boldsymbol{e}_1, \ldots, i\boldsymbol{e}_n\}$ и $\boldsymbol{B}_{\mathbb{R}} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m, i\boldsymbol{b}_1, \ldots, i\boldsymbol{b}_m\}$ равна $\begin{pmatrix} B & -C \\ C & B \end{pmatrix}$.

Доказательство. Пусть $A = (a_{ij})$, $B = (b_{ij})$ и $C = (c_{ij})$. По определению матрицы линейного отображения каждый вектор $f(\boldsymbol{e}_j)$, $j \leqslant n$, имеет координаты $a_{1j} = b_{1j} + ic_{1j}, \ldots, a_{mj} = b_{mj} + ic_{mj}$ в базисе $\boldsymbol{B}$ (когда он рассматривается как элемент U), а значит, координаты $b_{1j}, \ldots, b_{mj}, c_{1j}, \ldots, c_{mj}$ в базисе $\boldsymbol{E}_{\mathbb{R}}$ (когда он рассматривается как элемент $U_{\mathbb{R}}$). В силу линейности f над $\mathbb{C}$ каждый вектор $f(i\boldsymbol{e}_j)$, $j \leqslant n$, имеет координаты $ia_{1j} = -c_{1j} + ib_{1j}, \ldots, ia_{mj} = -c_{mj} + ib_{mj}$ в базисе $\boldsymbol{B}$ (когда он рассматривается как элемент U), а значит, координаты $-c_{1j}, \ldots, -c_{mj}, b_{1j}, \ldots, b_{mj}$ в базисе $\boldsymbol{E}_{\mathbb{R}}$ (когда он рассматривается как элемент $U_{\mathbb{R}}$). Из определения матрицы линейного отображения вытекает требуемое утверждение. □

Примеры

1. Комплексификация одномерного арифметического пространства $\mathbb{R}^1$ — арифметическое пространство $\mathbb{C}^1$. Каждое комплексное число $x + iy$, где $x, y \in \mathbb{R}$, можно трактовать как пару (x, y), причём пары умножаются на числа именно по тому закону, который по определению действует в комплексификации: $(x + iy)(u + iv) = xu - yv + i(xv + yu) \implies (x + iy)(u, v) = (xu - yv, xv + yu)$. Базис $\{\boldsymbol{e}\} = \{1\}$ пространства $\mathbb{R}^1$ является одновременно и базисом в $\mathbb{C}^1$ — любой элемент $\mathbb{C}^1$ есть некоторое комплексное число, умноженное на 1.
2. Овеществление одномерного пространства $\mathbb{C}^1$ изоморфно двумерному пространству $\mathbb{R}^2$. Изоморфизм задаётся правилом $\varphi\colon x + iy \mapsto (x, y)$. Базису $\{\boldsymbol{e}\} = \{1\}$ в пространстве $\mathbb{C}^1$ соответствует базис $\{1, i\}$ овеществления, который при изоморфизме φ переходит в стандартный базис $\{(1, 0), (0, 1)\}$.

   Рассмотрим поворот $f\colon \mathbb{C}^1 \to \mathbb{C}^1$ комплексной плоскости $\mathbb{C}^1$ на угол α. В области определения и области значений отображения f возьмём один и тот же базис $\{\boldsymbol{e}\} = \{1\}$. При повороте базисный вектор 1 переходит в вектор $\cos\alpha + i\sin\alpha$. Значит, поворот имеет матрицу $(\cos\alpha + i\sin\alpha)$ (размера 1×1). Матрица f как отображения $\mathbb{C}^1_{\mathbb{R}} \to \mathbb{C}^1_{\mathbb{R}}$ относительно базиса $\{\boldsymbol{e}, i\boldsymbol{e}\} = \{1, i\} = \{(1, 0), (0, 1)\}$ имеет вид $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$.

Линейные операторы

Определение

Линейный оператор в векторном пространстве V — это любое линейное отображение $f\colon V \to V$.

При определении матрицы линейного оператора в области определения и образе, как правило, берётся один и тот же базис.

Пусть V — конечномерное векторное пространство с базисом $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$. Матрица линейного оператора $\mathcal{A}\colon V \to V$ в базисе $\boldsymbol{E}$ — это квадратная матрица $A = (a_{ij})$ размера $n \times n$, которая определяется равенством

$$A\boldsymbol{e}_j = \sum_{i=1}^{n} a_{ij}\boldsymbol{e}_i.$$

В j-м столбце матрицы A стоят координаты вектора $\mathcal{A}\boldsymbol{e}_j$ в базисе $\boldsymbol{E}$.

Как и любое линейное отображение, всякий линейный оператор $A\colon V \to V$ однозначно определяется образами базисных векторов пространства V, которые могут быть любыми. Всякая квадратная матрица порядка n является матрицей некоторого линейного оператора в данном базисе.

Поскольку оператор $\mathcal{A}$ — частный случай линейного отображения, его матрица A в базисе $\boldsymbol{E}$ связывает координаты векторов и их образов в этом базисе точно так же, как и в общем случае: для любого вектора $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n$ из пространства V и его образа $\mathcal{A}\boldsymbol{x} = \boldsymbol{y} = y_1\boldsymbol{e}_1 + \dots + y_n\boldsymbol{e}_n$ имеем

$$\boxed{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}\,.$$

Когда базис в пространстве V фиксирован и векторы отождествляются со столбцами их координат в этом базисе, пишут просто $\boldsymbol{y} = A\boldsymbol{x}$.

В случае оператора формула преобразования матрицы при замене базиса имеет тот же вид, что и в случае общего линейного отображения: если V — конечномерное пространство, $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ и $\boldsymbol{E}' = \{\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_n\}$ — два базиса в нём, $\mathcal{A}\colon V \to V$ — линейный оператор, который имеет матрицу A в базисе $\boldsymbol{E}$ и матрицу A' в базисе $\boldsymbol{E}'$, а T — матрица перехода от $\boldsymbol{E}$ к $\boldsymbol{E}'$, то

$$\boxed{A' = T^{-1}AT}\,.$$

Определение

Говорят, что квадратные матрицы A и B с коэффициентами из поля K подобны (над полем K), если существует невырожденная матрица C с коэффициентами из K, для которой $\boxed{A = C^{-1}BC}$.

Теорема

Пусть V — векторное пространство конечной размерности n.

Квадратные матрицы A и B порядка n подобны тогда и только тогда, когда они являются матрицами одного и того же линейного оператора $V \to V$ в некоторых базисах пространства V.

Доказательство. *Достаточность* немедленно следует из формулы преобразования матрицы оператора при замене базиса: $A' = T^{-1}AT$.

Необходимость вытекает из той же формулы преобразования матрицы, а также того, что любая квадратная матрица порядка n является матрицей некоторого оператора $V \to V$ в любом данном базисе, и что любая невырожденная матрица является матрицей перехода от некоторого базиса пространства V к некоторому другому базису. □

Мы будем называть любую характеристику матриц, которая сохраняется преобразованием подобия, т.е. отображением матриц $A \mapsto C^{-1}AC$ для любой невырожденной матрицы C, матричным инвариантом (примеры — определитель и ранг). Любую характеристику операторов в векторном пространстве, не зависящую от выбора базиса в этом пространстве, будем называть инвариантом оператора. Любой инвариант матрицы оператора (в частности, её определитель и ранг) является инвариантом оператора.

Размерность подпространства $\operatorname{Im}\mathcal{A}$ пространства V называется рангом оператора $\mathcal{A}$ и обозначается $\operatorname{rank}\mathcal{A}$. В силу доказанного в разделе о линейных отображениях равенства $\dim\operatorname{Im}\mathcal{A} = \operatorname{rank} A$ (где A — матрица оператора $\mathcal{A}$) имеем

$$\boxed{\operatorname{rank}\mathcal{A} = \operatorname{rank} A}.$$

Таким образом, ранг оператора — это просто ранг его матрицы (в любом базисе).

Определителем оператора $\mathcal{A}$ называется определитель его матрицы в любом базисе.

Ранг и определитель — инварианты оператора.

Алгебра линейных операторов

Алгебра над полем K — это четвёрка $(A, +, \times, \cdot)$, где $+ : A \times A \to A$ — операция сложения элементов A (сумма $+(\boldsymbol{x}, \boldsymbol{y})$ обозначается $\boldsymbol{x} + \boldsymbol{y}$), $\times : A \times A \to A$ — операция умножения элементов A (произведение $\times(\boldsymbol{x}, \boldsymbol{y})$ обозначается $\boldsymbol{x} \times \boldsymbol{y}$) и $\cdot : K \times A \to A$ — операция умножения элементов A на числа из K (произведение $\cdot(\boldsymbol{x}, \lambda)$ обозначается $\lambda \boldsymbol{x}$), обладающие следующими свойствами:

1. $(V, +, \cdot)$ — векторное пространство над полем K;
2. для любых $\lambda \in K$ и $\boldsymbol{x}, \boldsymbol{y} \in V$ умножение $\times : A \times A \to A$ *билинейно*, т.е. для любых $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in A$ и $\lambda, \mu \in K$ выполняются условия
   - $(\boldsymbol{x} + \boldsymbol{y}) \times \boldsymbol{z} = \boldsymbol{x} \times \boldsymbol{z} + \boldsymbol{y} \times \boldsymbol{z}$,
   - $\boldsymbol{x} \times (\boldsymbol{y} + \boldsymbol{z}) = \boldsymbol{x} \times \boldsymbol{y} + \boldsymbol{x} \times \boldsymbol{z}$,
   - $(\lambda \boldsymbol{x}) \times (\mu \boldsymbol{y}) = (\lambda \cdot \mu)(\boldsymbol{x} \times \boldsymbol{y})$.

Элементами алгебры $(A, +, \times, \cdot)$ считаются элементы множества A. Вместо $(A, +, \times, \cdot)$ пишут просто A.

Если в алгебре A имеется *единица*, т.е. элемент $\boldsymbol{e} \in A$ такой, что $\boldsymbol{e} \times \boldsymbol{x} = \boldsymbol{x} \times \boldsymbol{e}$ для любого $\boldsymbol{x} \in A$, то A называется алгеброй с единицей.

Упражнение

1. Покажите, что условие ② в определении алгебры равносильно условию
   ②′ для любых $\lambda \in K$ и $\boldsymbol{x}, \boldsymbol{y} \in V$ $\quad (\lambda \boldsymbol{x}) \times \boldsymbol{y} = \lambda(\boldsymbol{x} \times \boldsymbol{y}) = \boldsymbol{x} \times (\lambda \boldsymbol{y})$.
2. Пусть K — поле и $(R, +, \times)$ — кольцо с единицей. Предположим, что существует сохраняющий единицу гомоморфизм колец $f\colon K \to R$, образ которого $f(K)$ содержится в центре кольца R (центр — это множество элементов, коммутирующих по умножению со всеми остальными элементами). Для $\lambda \in K$ и $\boldsymbol{x} \in R$ положим $\lambda \boldsymbol{x} = f(\lambda) \times \boldsymbol{x}$. Тем самым определена операция $\cdot\colon R \times K \to R$: для $\lambda \in K$ и $\boldsymbol{x} \in R$ полагаем $\cdot(\boldsymbol{x}, \lambda) = \lambda \boldsymbol{x}$.
   - Проверьте, что $(R, +, \times, \cdot)$ — алгебра с единицей над полем K.
   - Заметьте, что любая алгебра с единицей $(A, +, \times, \cdot)$ получается таким способом из кольца $(A, +, \times)$.

Обозначение

Множество всех линейных операторов в векторном пространстве V обозначается через $\operatorname{End} V$.

На множестве $\operatorname{End} V$ всех линейных операторов в векторном пространстве V над полем K имеются алгебраические операции: *произведение* операторов — композиция $\circ$:
$$(\mathcal{A} \circ \mathcal{B})\boldsymbol{v} = \mathcal{A}(\mathcal{B}\boldsymbol{v}) \quad \text{для } \mathcal{A}, \mathcal{B} \in \operatorname{End} V \text{ и } \boldsymbol{v} \in V,$$
сумма и умножение на число поточечные:
$$(\mathcal{A} + \mathcal{B})\boldsymbol{v} = \mathcal{A}\boldsymbol{v} + \mathcal{B}\boldsymbol{v} \quad \text{и} \quad (\lambda\mathcal{A})\boldsymbol{v} = \mathcal{A}(\lambda\boldsymbol{v}) \quad \text{для } \mathcal{A}, \mathcal{B} \in \operatorname{End} V,\ \boldsymbol{v} \in V \text{ и } \lambda \in K.$$
Множество $\operatorname{End} V$ с этими операциями является (ассоциативной) алгеброй с единицей. Единица — *тождественный* оператор $\mathcal{E}$, ноль — *нулевой* оператор Θ ($\mathcal{E}\boldsymbol{x} = \boldsymbol{x}$, $\Theta\boldsymbol{x} = \boldsymbol{0}$ для всех $\boldsymbol{x} \in V$).

Упражнение

Докажите, что оператор $\mathcal{A} \in \operatorname{End} V$ *идемпотентный*, т.е. $\mathcal{A} \circ \mathcal{A} = \mathcal{A}$, если и только если $\mathcal{A}$ является проектированием на некоторое подпространство.

Множество $M_n(K)$ всех квадратных матриц порядка n с коэффициентами из поля K тоже является алгеброй с единицей.

Две алгебры A и A' над одним и тем же полем K изоморфны, если существует биекция $\varphi\colon A \rightleftarrows A'$, сохраняющая все операции, т.е. такая, что
$$\varphi(x + y) = \varphi(x) + \varphi(y),\ \varphi(x \times y) = \varphi(x) \times \varphi(y),\ \varphi(\lambda x) = \lambda\varphi(x) \quad \forall\, x, y \in A,\ \lambda \in K.$$

Теорема

Для любого n-мерного векторного пространства V над полем K алгебры $\operatorname{End} V$ *и* $M_n(K)$ *изоморфны.*

Доказательство. Зафиксируем базис $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ в пространстве V. Пусть $\varphi\colon \operatorname{End} V \to M_n(K)$ — отображение, которое каждому линейному оператору $\mathcal{A}$ ставит в соответствие его матрицу A в базисе $\boldsymbol{E}$. Это биекция.

Пусть $\mathcal{A}, \mathcal{B} \in \operatorname{End} V$ и $\mathcal{C} = \mathcal{A} \circ \mathcal{B}$, и пусть $\varphi(\mathcal{A}) = A = (a_{ij})$, $\varphi(\mathcal{B}) = B = (b_{ij})$ и $\varphi(\mathcal{C}) = C = (c_{ij})$. Тогда для любого $k \leqslant n$ имеем

$$\mathcal{C}\boldsymbol{e}_k = (\mathcal{A}\mathcal{B})\boldsymbol{e}_k = \mathcal{A}\Big(\sum_{j=1}^{n} b_{jk}\boldsymbol{e}_j\Big) = \sum_{j=1}^{n} b_{jk}\mathcal{A}\boldsymbol{e}_j = \sum_{j=1}^{n} b_{jk}\Big(\sum_{i=1}^{n} a_{ij}\boldsymbol{e}_i\Big) = \sum_{i=1}^{n}\Big(\sum_{j=1}^{n} a_{ij}b_{jk}\Big)\boldsymbol{e}_i.$$

С другой стороны,

$$\mathcal{C}\boldsymbol{e}_k = \sum_{i=1}^{n} c_{ik}\boldsymbol{e}_i.$$

В силу единственности разложения вектора по базису получаем $c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$, т.е. $C = AB$. Таким образом, $\varphi(\mathcal{A} \circ \mathcal{B}) = \varphi(\mathcal{B})\varphi(\mathcal{B})$. Аналогично проверяется, что $\varphi(\mathcal{A} + \mathcal{B}) = \varphi(\mathcal{B}) + \varphi(\mathcal{B})$ и $\varphi(\lambda\mathcal{A}) = \lambda\varphi(\mathcal{A})$. □

Следствие

Для любого конечномерного векторного пространства V над полем K

$$\dim \operatorname{End} V = (\dim V)^2.$$

Замечание

Линейный оператор $\mathcal{B}$ является обратным к оператору $\mathcal{A}$ относительно умножения в алгебре $\operatorname{End} V$ тогда и только тогда, когда операторы $\mathcal{A}$ и B взаимно обратны как отображения. Обозначение: $\mathcal{B} = \mathcal{A}^{-1}$.

Не для всякого линейного оператора существует обратный: например, у проектирования на собственное подпространство обратного нет. Оператор, у которого есть обратный (т.е. биективный), называется невырожденным.

Замечание

Оператор $\mathcal{A}$ невырожден $\iff$ выполнено любое из равносильных условий

- $\det \mathcal{A} \neq 0$,
- $\mathcal{A}$ инъективен и сюръективен,
- $\operatorname{Ker} \mathcal{A} = \{\mathbf{0}\}$ и $\operatorname{Im} \mathcal{A} = V$,
- $\operatorname{Ker} \mathcal{A} = \{\mathbf{0}\}$,
- $\operatorname{Im} \mathcal{A} = V$.

Многочлены от операторов

Пусть $f(x) = a_0 + a_1 x + \cdots + a_m x^m \in K[x]$ — многочлен с коэффициентами из поля K и $\mathcal{A}$ — линейный оператор в векторном пространстве V. Тогда определён оператор

$$f(\mathcal{A}) = a_0 \mathcal{E} + a_1 \mathcal{A} + a_2 \mathcal{A}^2 + \cdots + a_m \mathcal{A}^m.$$

Если пространство V конечномерно и A — матрица оператора $\mathcal{A}$ в некотором базисе, то оператор $f(\mathcal{A})$ имеет в этом базисе матрицу

$$f(A) = a_0 E + a_1 A + a_2 A^2 + \cdots + a_m A^m.$$

Замечание

Для любого оператора $\mathcal{A}$ и любых многочленов $f(x)$ и $g(x)$ операторы $f(\mathcal{A})$ и $g(\mathcal{A})$ (так же как и их матрицы) коммутируют: $f(\mathcal{A})g(\mathcal{A}) = g(\mathcal{A})f(\mathcal{A})$.

Определение

Ненулевой многочлен $f(x) \in K[x]$ называется аннулирующим многочленом для линейного оператора $\mathcal{A}$, если $f(\mathcal{A}) = \Theta$. Аналогично определяется аннулирующий многочлен для матрицы.

Предложение

Для всякого линейного оператора $\mathcal{A}$ в конечномерном векторном пространстве V существует аннулирующий многочлен.

Доказательство. Пусть $\dim V = n$. Рассмотрим операторы $\mathcal{E}, \mathcal{A}, \mathcal{A}^2, \ldots, \mathcal{A}^{n^2}$. Некоторая их нетривиальная линейная комбинация $a_0\mathcal{E} + a_1\mathcal{A} + a_2\mathcal{A}^2 + \cdots + a_{n^2}\mathcal{A}^{n^2}$ равна нулевому оператору Θ, потому что $\dim \operatorname{End} V = n^2$. Многочлен $a_0 + a_1x + a_2x^2 + \cdots + a_{n^2}x^{n^2}$ является аннулирующим для $\mathcal{A}$. □

Упражнение

Покажите, что если A — верхнетреугольная (диагональная) матрица порядка n с диагональными элементами $a_{11}, \ldots, a_{nn}$, то матрица $f(A)$ тоже верхнетреугольная (диагональная) и имеет диагональные элементы $f(a_{11}), f(a_{22}), \ldots, f(a_{nn})$:

$$f\begin{pmatrix} a_{11} & & & & \\ 0 & a_{22} & & \ast & \\ \vdots & \vdots & \ddots & & \\ 0 & 0 & \ldots & a_{n-1\,n-1} & \\ 0 & 0 & \ldots & 0 & a_{nn} \end{pmatrix} = \begin{pmatrix} f(a_{11}) & & & & \\ 0 & f(a_{22}) & & \ast & \\ \vdots & \vdots & \ddots & & \\ 0 & 0 & \ldots & f(a_{n-1\,n-1}) & \\ 0 & 0 & \ldots & 0 & f(a_{nn}) \end{pmatrix}.$$

Инвариантные подпространства и собственные векторы

Определение

Подпространство U векторного пространства V называется инвариантным подпространством оператора $\mathcal{A}\colon V \to V$, если $\mathcal{A}U \subset U$, т.е. $\mathcal{A}\boldsymbol{u} \in U$ для любого вектора $\boldsymbol{u} \in U$.

Ограничение $\mathcal{A}|_U$ линейного оператора $\mathcal{A}$ на инвариантное подпространство U является линейным оператором в U.

Подпространства $\{\mathbf{0}\}$ и V инвариантны всегда.

Замечание

Для любого оператора $\mathcal{A}$ подпространства $\operatorname{Ker}\mathcal{A}$ и $\operatorname{Im}\mathcal{A}$ инвариантны.

Если U — инвариантное подпространство оператора $\mathcal{A}\colon V \to V$, то можно определить фактор-оператор $\widetilde{\mathcal{A}}\colon V/U \to V/U$ правилом $\widetilde{\mathcal{A}}([\boldsymbol{v}]) = [\mathcal{A}\boldsymbol{v}]$ для $\boldsymbol{v} \in V$.

Упражнение

Покажите, что это определение корректно.

Пусть $\mathcal{A}$ — линейный оператор в конечномерном векторном пространстве V и $U \subset V$ — инвариантное подпространство. Предположим, что $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k, \boldsymbol{e}_{k+1}, \ldots, \boldsymbol{e}_n\}$ — базис пространства V, в котором первые k векторов $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$ образуют базис подпространства U, так что $V = U \oplus \langle\{\boldsymbol{e}_{k+1}, \ldots, \boldsymbol{e}_n\}\rangle$. Пусть $A = (a_{ij})$ — матрица $\mathcal{A}$ в этом базисе.

Для каждого $i \leqslant k$ имеем $\mathcal{A}\boldsymbol{e}_i = a_{1i}\boldsymbol{e}_1 + \cdots + a_{ki}\boldsymbol{e}_k$ (так как $\mathcal{A}U \subset U$), причём числа $a_{1i}, \ldots, a_{ki}$ образуют i-й столбец матрицы A_U ограничения $\mathcal{A}|_U$ в базисе $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k\}$ подпространства U.

Теперь возьмём $i > k$. Имеем $\mathcal{A}\boldsymbol{e}_i = a_{1i}\boldsymbol{e}_1 + a_{ki}\boldsymbol{e}_k + a_{k+1\,i}\boldsymbol{e}_{k+1} + \cdots + a_{ni}\boldsymbol{e}_n \in a_{k+1\,i}\boldsymbol{e}_{k+1} + \cdots + a_{ni}\boldsymbol{e}_n + U$. Значит, $\mathcal{A}\boldsymbol{e}_i + U = (a_{k+1\,i}\boldsymbol{e}_{k+1} + U) + \cdots + (a_{ni}\boldsymbol{e}_n + U)$, т.е. числа $a_{k+1\,i}, \ldots, a_{ni}$ составляют $(i-k)$-й столбец матрицы $\widetilde{A}$ фактор-оператора $\widetilde{\mathcal{A}}\colon V/U \to V/U$.

Вывод: Матрица A оператора $\mathcal{A}$ в базисе $\boldsymbol{E}$ имеет вид

$$A = \begin{pmatrix} A_U & * \\ 0 & \widetilde{A} \end{pmatrix},$$

где A_U — матрица ограничения $\mathcal{A}|_U$, а $\widetilde{A}$ — матрица фактор-оператора $V/U \to V/U$ в базисе факторпространства $\{[\boldsymbol{e}_{k+1}], \ldots, [\boldsymbol{e}_n]\}$.

Если, сверх того, подпространство $W = \langle \boldsymbol{e}_{k+1}, \ldots, \boldsymbol{e}_n \rangle$ тоже инвариантно, то

$$A = \begin{pmatrix} A_U & 0 \\ 0 & A_W \end{pmatrix},$$

где A_W — матрица ограничения $\mathcal{A}|_W$. В этом случае говорят, что оператор $\mathcal{A}$ является *прямой суммой операторов* $\mathcal{A}|_U$ и $\mathcal{A}|_W$ и пишут $\mathcal{A} = \mathcal{A}|_U \oplus \mathcal{A}|_W$ (ясно, что при этом $V = U \oplus W$).

Аналогично определяется прямая сумма произвольного (конечного) числа операторов на произвольных векторных пространствах: если $U_1, \ldots, U_m$ — векторные пространства и $V = U_1 \oplus \cdots \oplus U_m$ (это означает, что любой вектор $\boldsymbol{v} \in V$ однозначным образом представляется в виде $\boldsymbol{v} = \boldsymbol{u}_1 + \cdots + \boldsymbol{u}_n$, где $\boldsymbol{u}_i \in U_i$, $i \leqslant m$), и если для каждого $i \leqslant m$ $\mathcal{A}_i$ — оператор в U_i, то оператор $\mathcal{A}\colon V \to V$, определённый правилом $\mathcal{A}(\boldsymbol{v}) = \mathcal{A}_1(\boldsymbol{u}_1) + \cdots + \mathcal{A}_m(\boldsymbol{u}_m)$ для $\boldsymbol{v} = \boldsymbol{u}_1 + \cdots + \boldsymbol{u}_m$, называется прямой суммой операторов $\mathcal{A}_1, \ldots, \mathcal{A}_m$ и обозначается $\mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_m$. По-другому его можно определить как тот единственный оператор $\mathcal{A}$ в $U_1 \oplus \cdots \oplus U_m$, сужение которого на каждое подпространство U_i, $i \leqslant m$, совпадает с $\mathcal{A}_i$.

Предложение

Пусть $\mathcal{A}$ — линейный оператор в векторном пространстве V, $f(x)$ — аннулирующий многочлен для $\mathcal{A}$ и $f(x) = f_1(x) \cdot f_2(x)$ для взаимно простых многочленов $f_1(x)$ и $f_2(x)$. Тогда $\mathcal{A}$ раскладывается в прямую сумму операторов $\mathcal{A}_1$ и $\mathcal{A}_2$ таких, что $f_1(\mathcal{A}_1) = \Theta$ и $f_2(\mathcal{A}_2) = \Theta$.

▸ К жордановой форме

Доказательство. Поскольку $f_1(x)$ и $f_2(x)$ взаимно просты, существуют многочлены $g_1(x)$ и $g_2(x)$, для которых $g_1(x) \cdot f_1(x) + g_2(x) \cdot f_2(x) = 1$, т.е. $g_1(\mathcal{A}) \cdot f_1(\mathcal{A}) + g_2(\mathcal{A}) \cdot f_2(\mathcal{A}) = \mathcal{E}$. Для любого $\boldsymbol{x} \in V$ имеем

$$\boldsymbol{x} = \mathcal{E}(\boldsymbol{x}) = f_1(\mathcal{A})(g_1(\mathcal{A})\boldsymbol{x}) + f_2(\mathcal{A})(g_2(\mathcal{A})\boldsymbol{x}) \in \operatorname{Im} f_1(\mathcal{A}) + \operatorname{Im} f_2(\mathcal{A}).$$

Значит, $V = \operatorname{Im} f_1(\mathcal{A}) + \operatorname{Im} f_2(\mathcal{A})$.

Поскольку $f(\mathcal{A}) = f_1(\mathcal{A})f_2(\mathcal{A}) = f_2(\mathcal{A})f_1(\mathcal{A}) = \Theta$, имеем $\operatorname{Im} f_1(\mathcal{A}) \subset \operatorname{Ker} f_2(\mathcal{A})$ и $\operatorname{Im} f_2(\mathcal{A}) \subset \operatorname{Ker} f_1(\mathcal{A})$. Значит, $V = \operatorname{Ker} f_1(\mathcal{A}) + \operatorname{Ker} f_2(\mathcal{A})$. Эта сумма прямая: если $\boldsymbol{x} \in \operatorname{Ker} f_1(\mathcal{A}) \cap \operatorname{Ker} f_2(\mathcal{A})$, то

$$\boldsymbol{x} = \mathcal{E}\boldsymbol{x} = g_1(\mathcal{A})(f_1(\mathcal{A})\boldsymbol{x}) + g_2(\mathcal{A})(f_2(\mathcal{A})\boldsymbol{x}) = g_1(\mathcal{A})\mathbf{0} + g_2(\mathcal{A})\mathbf{0} = \mathbf{0}.$$

Итак, $V = \operatorname{Ker} f_1(\mathcal{A}) \oplus \operatorname{Ker} f_2(\mathcal{A})$. Подпространства $\operatorname{Ker} f_i(\mathcal{A})$ инвариантны: если $\boldsymbol{x} \in \operatorname{Ker} f_i(\mathcal{A})$, то $\mathcal{A}\boldsymbol{x} \in \operatorname{Ker} f_i(\mathcal{A})$, так как $f_i(\mathcal{A})\mathcal{A}\boldsymbol{x} = \mathcal{A}f_i(\mathcal{A})\boldsymbol{x} = \mathcal{A}\mathbf{0} = \mathbf{0}$. □

Если пространство V конечномерно, $V = U_1 \oplus \dots \oplus U_m$, все U_i инвариантны и базис V составлен из базисов подпространств U_i, то матрица A оператора $\mathcal{A} = \mathcal{A}_1 \oplus \dots \oplus \mathcal{A}_m$ имеет блочно-диагональный вид

$$A = \begin{pmatrix} A_1 & & & & \\ & A_2 & & \text{\huge 0} & \\ & & \ddots & & \\ & \text{\huge 0} & & A_{m-1} & \\ & & & & A_m \end{pmatrix}.$$

Обратно, если в некотором базисе $\boldsymbol{E}$ матрица оператора $\mathcal{A} \colon V \to V$ имеет блочно-диагональный вид с диагональными блоками $A_1, \dots, A_m$, то V раскладывается в прямую сумму инвариантных подпространств $U_1, \dots, U_m$ таких, что для каждого $i \leqslant m$ A_i — матрица ограничения $\mathcal{A}|_{U_i}$ и базис $\boldsymbol{E}$ составлен из базисов подпространств U_i.

В частности, матрица A оператора $\mathcal{A}$ в базисе $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ диагональна тогда и только тогда, когда все подпространства $\langle \boldsymbol{e}_i \rangle$ инвариантны.

Замечание

Для $\boldsymbol{x} \in V$ подпространство $\langle \boldsymbol{x} \rangle$ инвариантно относительно линейного оператора $\mathcal{A} \colon V \to V$ тогда и только тогда, когда существует $\lambda \in K$, для которого $\mathcal{A}\boldsymbol{x} = \lambda \boldsymbol{x}$.

Определение

Пусть V — векторное пространство над полем K. Вектор $\boldsymbol{v} \in V$ называется собственным вектором линейного оператора $\mathcal{A}\colon V \to V$, если $\boldsymbol{v} \neq \mathbf{0}$ и $\mathcal{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ для некоторого числа $\lambda \in K$. При этом λ называется собственным значением, отвечающим собственному вектору $\boldsymbol{v}$.

Множество собственных значений линейного оператора называется спектром этого оператора.

Говорят, что оператор в конечномерном векторном пространстве диагонализируем, если в некотором базисе его матрица диагональна. Из предыдущих рассуждений вытекает следующее утверждение.

Предложение

Оператор $\mathcal{A}$ в конечномерном пространстве диагонализируем тогда и только тогда, когда в пространстве V имеется базис, состоящий из собственных векторов оператора $\mathcal{A}$.

Пример

Пусть $V = U \oplus W$ и $\mathcal{A}$ — проектирование на подпространство U параллельно W. Собственными векторами оператора $\mathcal{A}$ являются все ненулевые векторы $\boldsymbol{u} \in U$ (им всем отвечает собственное значение 1) и все ненулевые векторы $\boldsymbol{w} \in W$ (им всем отвечает собственное значение 0). Если оба подпространства U и W ненулевые, то спектр $\mathcal{A}$ равен $\{0, 1\}$. Если $U \neq \{\mathbf{0}\}$ и $W = \{\mathbf{0}\}$, то спектр равен $\{1\}$. Если $U = \{\mathbf{0}\}$ и $W \neq \{\mathbf{0}\}$, то спектр равен $\{0\}$. Если оба подпространства U и W нулевые, то спектр $\mathcal{A}$ пуст.

С этого момента мы будем предполагать, что все рассматриваемые операторы действуют в конечномерных пространствах, если только явно не оговорено противное.

Пусть $\mathcal{A}$ — линейный оператор в конечномерном пространстве V. Ненулевой вектор $\boldsymbol{x} \in V$ является собственным вектором оператора $\mathcal{A}$, которому отвечает собственное значение λ, тогда и только тогда, когда координаты $x_1, \ldots, x_n$ этого вектора в любом базисе пространства V удовлетворяют условию

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \text{т.е.} \quad (A - \lambda E) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

(здесь A — матрица оператора $\mathcal{A}$ в том же базисе, а E — единичная матрица). Другими словами, $(x_1, \ldots, x_n)$ — решение однородной системы уравнений с матрицей $A - \lambda E$. Эта система имеет ненулевое решение $\iff$ $\det(A - \lambda E) = 0$. Приходим к следующему выводу:

Предложение

Пусть $\boldsymbol{E}$ — произвольный базис конечномерного пространства V, и $\mathcal{A} \colon V \to V$ — линейный оператор с матрицей A в этом базисе. Число λ является собственным значением, отвечающим некоторому собственному вектору оператора $\mathcal{A}$, тогда и только тогда, когда $\det(A - \lambda E) = 0$.

Определение

Пусть A — произвольная квадратная матрица. Многочлен

$$\chi_A(\lambda) = \det(A - \lambda E)$$

называется характеристическим многочленом матрицы A, а уравнение $\chi_A(\lambda) = 0$ — характеристическим уравнением матрицы A.

Предложение

Характеристические многочлены подобных матриц совпадают.

Доказательство. Если $A = C^{-1}BC$, то

$$\det(A - \lambda E) = \det(C^{-1}BC - \lambda E) = \det(C^{-1}BC - \lambda C^{-1}EC) =$$
$$= \det(C^{-1}(A - \lambda E)C) = \det C^{-1} \cdot \det(A - \lambda E) \cdot \det C = \det(A - \lambda E).$$

$\square$

Определение

Пусть V — конечномерное векторное пространство, в котором задан некоторый базис, и $\mathcal{A}\colon V \to V$ — любой линейный оператор. Пусть A — матрица оператора $\mathcal{A}$ в данном базисе. Многочлен

$$\chi_{\mathcal{A}}(\lambda) = \det(A - \lambda E)$$

называется характеристическим многочленом оператора $\mathcal{A}$, а уравнение $\chi_{\mathcal{A}}(\lambda) = 0$ — характеристическим уравнением оператора $\mathcal{A}$.

Из независимости характеристического многочлена от базиса вытекает такое утверждение:

Следствие

Коэффициенты характеристического многочлена линейного оператора не зависят от выбора базиса, т.е. являются инвариантами линейного оператора.

Следствие

Если U — инвариантное подпространство оператора $\mathcal{A}\colon V \to V$, то $\chi_{\mathcal{A}}(\lambda) = \chi_{\mathcal{A}|_U}(\lambda) \cdot \chi_{\widetilde{\mathcal{A}}}(\lambda)$ (здесь $\mathcal{A}|_U$ — сужение $\mathcal{A}$ на U и $\widetilde{\mathcal{A}}$ — фактор-оператор $V/U \to V/U$). Если $V = U \oplus W$, то $\chi_{\mathcal{A}}(\lambda) = \chi_{\mathcal{A}|_U}(\lambda) \cdot \chi_{\mathcal{A}|_W}(\lambda)$.

Упражнение

Покажите, что коэффициент характеристического многочлена матрицы A (оператора с матрицей A) при λ^{n-1} равен $(-1)^{n-1}\operatorname{tr} A$.

Из инвариантности коэффициентов характеристического многочлена вытекает, что

- след матрицы является матричным инвариантом;
- след матрицы линейного оператора (в любом базисе) является инвариантом оператора.

След матрицы оператора (в любом базисе) называется **следом** линейного оператора.

Замечание

Спектр оператора совпадает с множеством корней его характеристического многочлена.

Лемма

Для любого линейного оператора $\mathcal{A}$ в векторном пространстве V над полем K и любого $\lambda \in K$ операторы $\mathcal{A}$ и $\mathcal{A} - \lambda\mathcal{E}$ имеют одни и те же инвариантные подпространства.

Доказательство. Если U — инвариантное подпространство оператора $\mathcal{A}$ и $\boldsymbol{x} \in U$, то $(\mathcal{A} - \lambda\mathcal{E})\boldsymbol{x} = \mathcal{A}\boldsymbol{x} - \lambda\boldsymbol{x} \in U + U = U$. Отсюда следует, что U инвариантно для $\mathcal{A} - \lambda\mathcal{E}$. То, что всякое подпространство, инвариантное для $\mathcal{A} - \lambda\mathcal{E}$, инвариантно и для $\mathcal{A}$, проверяется аналогично. $\square$

Предложение

Пусть V — векторное пространство над полем K, $\mathcal{A}: V \to V$ — линейный оператор и $\lambda \in K$. Тогда множество $V_\lambda = \{\boldsymbol{x} \in V : \mathcal{A}\boldsymbol{x} = \lambda\boldsymbol{x}\}$ является инвариантным подпространством оператора $\mathcal{A}$.

Доказательство. Ясно, что $V_\lambda = \operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E})$. Осталось применить лемму. $\square$

Определение

Для любого линейного оператора $\mathcal{A}$ в векторном пространстве V над полем K и любого $\lambda \in K$ подпространство

$$V_\lambda = \{\boldsymbol{x} \in V : \mathcal{A}\boldsymbol{x} = \lambda\boldsymbol{x}\} \subset V$$

называется собственным подпространством оператора $\mathcal{A}$, отвечающим собственному значению λ. Оно состоит из нулевого вектора и всех собственных векторов оператора $\mathcal{A}$, отвечающих собственному значению λ.

Предложение

Для любого корня λ характеристического многочлена линейного оператора $\mathcal{A}: V \to V$ для размерности собственного подпространства V_λ имеет место равенство

$$\boxed{\dim V_\lambda = \dim V - \operatorname{rank}(\mathcal{A} - \lambda\mathcal{E})}\,.$$

Доказательство. Требуемое равенство немедленно вытекает из определения собственного вектора с собственным значением λ как решения уравнения $\mathcal{A}\boldsymbol{x} = \lambda\boldsymbol{x}$. $\square$

Теорема

Собственные векторы любого линейного оператора, отвечающие разным собственным значениям, линейно независимы.

Доказательство. Пусть $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ — собственные векторы линейного оператора $\mathcal{A}\colon V \to V$, отвечающие разным собственным значениям $\lambda_1, \ldots, \lambda_n$, и пусть дана их линейная комбинация $\mu_1 \boldsymbol{u}_1 + \cdots + \mu_n \boldsymbol{u}_n$, равная нулевому вектору $\mathbf{0}$. Нам надо доказать, что $\mu_1 = \cdots = \mu_n = 0$.

Применим индукцию по n. Для $n = 1$ утверждение верно. Пусть $n > 1$ и для меньших n всё доказано. Имеем

$$\mathcal{A}(\mu_1 \boldsymbol{u}_1 + \cdots + \mu_n \boldsymbol{u}_n) = \mu_1 \mathcal{A} \boldsymbol{u}_1 + \cdots + \mu_n \mathcal{A} \boldsymbol{u}_n = \mu_1 \lambda_1 \boldsymbol{u}_1 + \cdots + \mu_n \lambda_n \boldsymbol{u}_n = \mathcal{A}\mathbf{0} = \mathbf{0}.$$

Вычитая данную линейную комбинацию, умноженную на λ_n, получаем

$$\mu_1(\lambda_1 - \lambda_n)\boldsymbol{u}_1 + \cdots + \mu_{n-1}(\lambda_{n-1} - \lambda_n)\boldsymbol{u}_{n-1} + \mu_n(\lambda_n - \lambda_n)\boldsymbol{u}_n =$$
$$= \mu_1(\lambda_1 - \lambda_n)\boldsymbol{u}_1 + \cdots + \mu_{n-1}(\lambda_{n-1} - \lambda_n)\boldsymbol{u}_{n-1} = \mathbf{0}.$$

Эта сумма — линейная комбинация собственных векторов с разными собственными значениями с числом слагаемых $< n$. Поскольку $\lambda_i - \lambda_n \neq 0$ для $i < n$, из индуктивного предположения следует, что $\mu_i = 0$ для $i \leqslant n - 1$. Из того, что данная линейная комбинация равна нулевому вектору, получаем $\mu_n = 0$. □

Следствие

Сумма собственных подпространств любого линейного оператора является прямой суммой.

Следствие

Если характеристический многочлен линейного оператора в n-мерном векторном пространстве имеет n различных корней, то этот оператор диагонализируем (равносильные условия: его матрица подобна диагональной; в пространстве имеется базис из собственных векторов оператора).

Теорема

Линейный оператор в ненулевом конечномерном векторном пространстве имеет хотя бы один собственный вектор тогда и только тогда, когда у его характеристического многочлена есть хотя бы один корень.

Поле K алгебраически замкнуто, если всякий многочлен ненулевой степени над K имеет хотя бы один корень. Поле комплексных чисел $\mathbb{C}$ алгебраически замкнуто. Любое поле изоморфно подполю алгебраически замкнутого поля.

Следствие

Любой линейный оператор в ненулевом конечномерном векторном пространстве над алгебраически замкнутым полем (в частности, над полем комплексных чисел) имеет собственный вектор.

Линейный оператор в вещественном векторном пространстве может не иметь собственных векторов (пример — поворот плоскости).

Предложение

Пусть $\mathcal{A}\colon V \to V$ — линейный оператор в векторном пространстве V над полем $\mathbb{R}$, $\mathcal{A}_{\mathbb{C}}$ — его комплексификация и $\boldsymbol{x} + i\boldsymbol{y} \in V_{\mathbb{C}}$ — собственный вектор оператора $\mathcal{A}_{\mathbb{C}}$ с собственным значением $\lambda + i\mu$, где $\lambda, \mu \in \mathbb{R}$, $\mu \neq 0$. Тогда линейная оболочка $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ в V является двумерным инвариантным подпространством оператора $\mathcal{A}$, причём $\mathcal{A}\boldsymbol{x} = \lambda\boldsymbol{x} - \mu\boldsymbol{y}$ и $\mathcal{A}\boldsymbol{y} = \mu\boldsymbol{x} + \lambda\boldsymbol{y}$.

Доказательство. Имеем $\mathcal{A}_{\mathbb{C}}(\boldsymbol{x} + i\boldsymbol{y}) = (\lambda + i\mu)(\boldsymbol{x} + i\boldsymbol{y})$. Значит, $\mathcal{A}\boldsymbol{x} + i\mathcal{A}\boldsymbol{y} = (\lambda\boldsymbol{x} - \mu\boldsymbol{y}) + i(\mu\boldsymbol{x} + \lambda\boldsymbol{y})$, откуда

$$\mathcal{A}\boldsymbol{x} = \lambda\boldsymbol{x} - \mu\boldsymbol{y} \quad \text{и} \quad \mathcal{A}\boldsymbol{y} = \mu\boldsymbol{x} + \lambda\boldsymbol{y}.$$

Мы показали, что подпространство $\langle \boldsymbol{x}, \boldsymbol{y} \rangle \subset V$ инвариантно для $\mathcal{A}$. Векторы $\boldsymbol{x}$ и $\boldsymbol{y}$ линейно независимы, так как если $\boldsymbol{y} = \alpha\boldsymbol{x}$, то $\boldsymbol{x} + i\boldsymbol{y} = (1 + i\alpha)\boldsymbol{x}$, и $\mathcal{A}_{\mathbb{C}}(\boldsymbol{x} + i\boldsymbol{y}) = \mathcal{A}_{\mathbb{C}}(1 + i\alpha)\boldsymbol{x} = (1 + i\alpha)\mathcal{A}_{\mathbb{C}}\boldsymbol{x} = (1 + i\alpha)\mathcal{A}\boldsymbol{x}$. С другой стороны, по предположению $\mathcal{A}_{\mathbb{C}}(\boldsymbol{x} + i\boldsymbol{y}) = (\lambda + i\mu)(\boldsymbol{x} + i\boldsymbol{y}) = (\lambda + i\mu)(1 + i\alpha)\boldsymbol{x}$. Значит, $\mathcal{A}\boldsymbol{x} = (\lambda + i\mu)\boldsymbol{x}$, а это невозможно, так как $\mathcal{A}$ переводит V в V и по предположению $\mu \neq 0$. $\square$

В качестве следствия получаем теорему:

Теорема

Любой линейный оператор в векторном пространстве над полем $\mathbb{R}$ имеет одномерное или двумерное инвариантное подпространство.

Пример

Оператор $\mathcal{A}$ поворота обычной плоскости на угол φ в ортонормированном базисе $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ имеет матрицу

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

У этой матрицы над полем $\mathbb{C}$ два собственных значения: $\cos\varphi - i\sin\varphi$ и $\cos\varphi + i\sin\varphi$. Им отвечают собственные векторы $\boldsymbol{e}_1 + i\boldsymbol{e}_2$ и $\boldsymbol{e}_1 - i\boldsymbol{e}_2$. Таким образом, над полем $\mathbb{C}$ матрица поворота подобна диагональной матрице

$$\begin{pmatrix} \cos\varphi - i\sin\varphi & 0 \\ 0 & \cos\varphi + i\sin\varphi \end{pmatrix}.$$

Корневые векторы и подпространства

Пусть V — векторное пространство размерности n над полем K и $\mathcal{A}\colon V \to V$ — линейный оператор.

Определение

Ненулевой вектор $\boldsymbol{v} \in V$ называется корневым вектором оператора $\mathcal{A}$, отвечающим числу $\lambda \in K$, если существует натуральное m, для которого $(\mathcal{A} - \lambda\mathcal{E})^m \boldsymbol{v} = \mathbf{0}$. Наименьшее такое m называется высотой корневого вектора $\boldsymbol{v}$.

Пусть $\boldsymbol{v}$ — корневой вектор высоты m, отвечающий числу λ. Тогда вектор $\boldsymbol{u} = (\mathcal{A} - \lambda\mathcal{E})^{m-1}\boldsymbol{v}$ — собственный вектор оператора $\mathcal{A}$, отвечающий собственному значению λ:

$$\mathcal{A}\boldsymbol{u} = (\mathcal{A} - \lambda\mathcal{E})\boldsymbol{u} + \lambda\boldsymbol{u} = (\mathcal{A} - \lambda\mathcal{E})^m\boldsymbol{v} + \lambda\boldsymbol{u} = \lambda\boldsymbol{u}.$$

Значит, λ — собственное значение оператора $\mathcal{A}$, т.е. корень его характеристического многочлена.

Легко видеть, что все корневые векторы, отвечающие корню λ, вместе с $\mathbf{0}$ образуют подпространство в V.

Определение

Множество всех корневых векторов, отвечающих корню λ, называется **корневым подпространством** и обозначается V^λ.

Замечание

Для любого собственного значения λ корневое подпространство V^λ содержит собственное подпространство V_λ.

Теорема

Для любого собственного значения λ корневое подпространство V^λ инвариантно относительно оператора $\mathcal{A}$.

Доказательство. Если $\boldsymbol{v} \in V^\lambda$ — корневой вектор высоты m, то $(\mathcal{A} - \lambda\mathcal{E})^m(\mathcal{A}\boldsymbol{v}) = \mathcal{A}(\mathcal{A} - \lambda\mathcal{E})^m\boldsymbol{v} = \mathcal{A}\mathbf{0} = \mathbf{0}$. $\square$

Пусть $f(x)$ — аннулирующий многочлен для оператора $\mathcal{A}$ со старшим коэффициентом 1. Предположим, что $f(x)$ раскладывается в произведение линейных множителей: $f(x) = (x - x_1)^{m_1} \ldots (x - x_k)^{m_k}$, где $x_i \neq x_j$ для $i \neq j$. Из предложения о взаимно простых множителях аннулирующего многочлена получаем $V = \mathrm{Ker}(\mathcal{A} - x_1\mathcal{E})^{m_1} \oplus \cdots \oplus \mathrm{Ker}(\mathcal{A} - x_k\mathcal{E})^{m_k}$. По определению корневого вектора всякий ненулевой вектор из $\mathrm{Ker}(\mathcal{A} - x_i\mathcal{E})^{m_i}$ является корневым вектором для $\mathcal{A}$, отвечающим числу x_i. Значит, если $\mathrm{Ker}(\mathcal{A} - x_i\mathcal{E})^{m_i} \neq \{\mathbf{0}\}$, то x_i — собственное значение оператора $\mathcal{A}$.

Вывод: Пусть $\mathcal{A}$ — линейный оператор в конечномерном пространстве V, у которого некоторый аннулирующий многочлен раскладывается в произведение линейных множителей. Тогда существуют инвариантные подпространства $V_1, \ldots, V_k \subset V$ такие, что

- $V = V_1 \oplus \cdots \oplus V_k$,
- $V_i \subset V^{\lambda_i}$ для некоторого собственного значения λ_i оператора $\mathcal{A}$, причём $\lambda_i \neq \lambda_j$ для $i \neq j$ (потом мы увидим, что V_i, $i \leqslant k$, — это в точности все корневые подпространства V^{λ_i} для $\mathcal{A}$),
- для каждого $i \leqslant k$ существует натуральное число m_i, для которого $(\mathcal{A}|_{V_i} - \lambda_i\mathcal{E})^{m_i} = \Theta$ (здесь $\mathcal{E}$ и Θ — тождественный и нулевой операторы в V_i).

Нильпотентные операторы

Определение

Оператор $\mathcal{A}$ называется **нильпотентным**, если $\mathcal{A}^h = \Theta$ для некоторого натурального h. Минимальное натуральное число h, для которого $\mathcal{A}^h = \Theta$, называется **высотой** (или **ступенью нильпотентности**) оператора $\mathcal{A}$.

Предложение

Всякий линейный оператор в конечномерном пространстве раскладывается в прямую сумму нильпотентного и невырожденного операторов.

Доказательство. Рассмотрим последовательность $V \underset{\text{lin}}{\supseteq} \operatorname{Im} \mathcal{A} \underset{\text{lin}}{\supseteq} \operatorname{Im} \mathcal{A}^2 \underset{\text{lin}}{\supseteq} \ldots$. Если $\operatorname{Im} \mathcal{A}^{i+1} \neq \operatorname{Im} \mathcal{A}^i$, то $\dim \operatorname{Im} \mathcal{A}^{i+1} < \dim \operatorname{Im} \mathcal{A}^i$. Значит, для некоторого k имеем $\operatorname{Im} \mathcal{A}^{k+1} = \operatorname{Im} \mathcal{A}^k$. Положим $U = \operatorname{Im} \mathcal{A}^k$ и $W = \operatorname{Ker} \mathcal{A}^k$. Подпространства U и V инвариантны, и $U \cap V = \{\mathbf{0}\}$. Поскольку $\dim V = \dim U + \dim W$, имеем $V = U \oplus W$, и $\mathcal{A} = \mathcal{A}|_U \oplus \mathcal{A}|_W$. При этом $\mathcal{A}|_U$ невырожден и $\mathcal{A}|_W$ нильпотентен. □

Всюду в этом разделе V — конечномерное векторное пространство и $\mathcal{A}$ — нильпотентный оператор в V высоты $h > 0$.

Жорданов базис для нильпотентного оператора

Назовём серией конечную последовательность ненулевых векторов $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k \in V$ такую, что $\mathcal{A}\boldsymbol{u}_i = \boldsymbol{u}_{i-1}$ для $i > 1$ и $\mathcal{A}\boldsymbol{u}_1 = \mathbf{0}$:

$$\mathbf{0} \xleftarrow{\mathcal{A}} \boldsymbol{u}_1 \xleftarrow{\mathcal{A}} \boldsymbol{u}_2 \xleftarrow{\mathcal{A}} \ldots \xleftarrow{\mathcal{A}} \boldsymbol{u}_k.$$

Вектор $\boldsymbol{u}_1$ — *первый член* серии, $\boldsymbol{u}_k$ — *последний член*, и число k — *длина* серии. Ясно, что $k \leqslant p$. Базис пространства V называется жордановым для оператора $\mathcal{A}$, если он распадается в серии.

Теорема

Для любого нильпотентного оператора в конечномерном пространстве существует жорданов базис.

Лемма 1

Высота h оператора $\mathcal{A}$ равна наибольшей из длин серий.

Доказательство. Если имеется серия длины k, то $\mathcal{A}^{k-1} \neq \Theta \implies h > k - 1$.

С другой стороны, $\mathcal{A}^{h-1} \neq \Theta \implies \exists\ \boldsymbol{u}_h \in V$, для которого $\mathcal{A}^{h-1}\boldsymbol{u}_h \neq \mathbf{0}$. Серия $\boldsymbol{u}_1 = \mathcal{A}^{h-1}\boldsymbol{u}_h \longleftarrow \boldsymbol{u}_2 = \mathcal{A}^{h-2}\boldsymbol{u}_{h_1} \longleftarrow \ldots \longleftarrow \boldsymbol{u}_{h-1} = \mathcal{A}\boldsymbol{u}_h \longleftarrow \boldsymbol{u}_h$ имеет длину h. $\square$

Лемма 2

Векторы, составляющие серию, линейно независимы.

Доказательство. Индукция по длине серии k. Если $k = 1$, доказывать нечего. Пусть $k > 1$ и для меньших k утверждение верно.

Пусть $\boldsymbol{u}_1 \leftarrow \dots \leftarrow \boldsymbol{u}_k$ — серия, и пусть $\lambda_1 \boldsymbol{u}_1 + \dots + \lambda_k \boldsymbol{u}_k = \boldsymbol{0}$. Применим $\mathcal{A}$:

$$\lambda_1 \mathcal{A}\boldsymbol{u}_1 + \lambda_2 \mathcal{A}\boldsymbol{u}_2 + \dots + \lambda_{k-1} \mathcal{A}\boldsymbol{u}_{k-1} + \lambda_k \mathcal{A}\boldsymbol{u}_k = \mathcal{A}\boldsymbol{0},$$

откуда $\lambda_2 \boldsymbol{u}_1 + \dots + \lambda_k \boldsymbol{u}_{k-1} = \boldsymbol{0}$. По индуктивному предположению $\lambda_2 = \dots = \lambda_k = 0$. Поскольку $\boldsymbol{u}_1 \neq \boldsymbol{0}$ и $\lambda_1 \boldsymbol{u}_1 = \boldsymbol{0}$, имеем $\lambda_1 = 0$. □

Лемма 3

Система векторов, распадающаяся на непересекающиеся серии, линейно независима $\iff$ линейно независима система первых векторов этих серий.

Доказательство. Пусть данная система состоит из серий

$$\boldsymbol{u}_{11} \leftarrow \boldsymbol{u}_{12} \leftarrow \cdots \leftarrow \boldsymbol{u}_{1k_1}, \quad \boldsymbol{u}_{21} \leftarrow \boldsymbol{u}_{22} \leftarrow \cdots \leftarrow \boldsymbol{u}_{2k_2}, \quad \ldots, \quad \boldsymbol{u}_{m1} \leftarrow \boldsymbol{u}_{m2} \leftarrow \cdots \leftarrow \boldsymbol{u}_{mk_m}.$$

Если она линейно независима, то линейна независима и её часть, состоящая из первых векторов.

Обратную импликацию будем доказывать индукцией по максимальной длине k серий в системе (т.е. по $k = \max_{i \leqslant m} k_i$). Если $k = 1$, то система целиком состоит из первых векторов. Предположим, что $k > 1$ и для меньших k утверждение верно.

Пусть

$$\lambda_{11}\boldsymbol{u}_{11} + \lambda_{12}\boldsymbol{u}_{12} + \cdots + \lambda_{1k_1}\boldsymbol{u}_{1k_1} + \cdots + \lambda_{m1}\boldsymbol{u}_{m1} + \lambda_{m2}\boldsymbol{u}_{m2} + \cdots + \lambda_{mk_m}\boldsymbol{u}_{mk_m} = \mathbf{0}.$$

Применим оператор $\mathcal{A}$:

$$\lambda_{11}\mathcal{A}\boldsymbol{u}_{11} + \lambda_{12}\mathcal{A}\boldsymbol{u}_{12} \cdots + \lambda_{1k_1}\mathcal{A}\boldsymbol{u}_{1k_1} + \cdots + \lambda_{m1}\mathcal{A}\boldsymbol{u}_{m1} + \lambda_{m2}\mathcal{A}\boldsymbol{u}_{m2} + \cdots + \lambda_{mk_m}\mathcal{A}\boldsymbol{u}_{mk_m} =$$

$$= \lambda_{12}\boldsymbol{u}_{11} + \cdots + \lambda_{1k_1}\boldsymbol{u}_{1\,k_1-1} + \cdots + \lambda_{m2}\boldsymbol{u}_{m1} + \cdots + \lambda_{mk_m}\boldsymbol{u}_{m\,k_m-1} = \mathbf{0}.$$

Каждая серия стала короче. По индуктивному предположению $\lambda_{12} = \cdots = \lambda_{1k_1} = \cdots = \lambda_{m2} = \cdots = \lambda_{mk_m} = 0$. Исходная линейная комбинация принимает вид

$$\lambda_{11}\boldsymbol{u}_{11} + \cdots + \lambda_{m1}\boldsymbol{u}_{m1} = \mathbf{0}.$$

Первые векторы серий линейно независимы $\Longrightarrow$ $\lambda_{11} = \cdots = \lambda_{m1} = 0$. □

Доказательство теоремы. Индукция по высоте оператора h. Если $h = 1$, то $\mathcal{A} = \Theta$ и годится любой базис. Пусть $h > 1$ и для меньших h утверждение верно.

Подпространство $\operatorname{Im}\mathcal{A} \subset V$ инвариантно, и высота ограничения $\mathcal{A}|_{\operatorname{Im}\mathcal{A}}$ равна $h-1$. По индуктивному предположению для $\mathcal{A}|_{\operatorname{Im}\mathcal{A}}$ существует жорданов базис

$$\boldsymbol{E} = \{\boldsymbol{e}_{11} \leftarrow \boldsymbol{e}_{12} \leftarrow \cdots \leftarrow \boldsymbol{e}_{1k_1},\ \ \boldsymbol{e}_{21} \leftarrow \boldsymbol{e}_{22} \leftarrow \cdots \leftarrow \boldsymbol{e}_{2k_2},\ \ \ldots,\ \ \boldsymbol{e}_{m1} \leftarrow \boldsymbol{e}_{m2} \leftarrow \cdots \leftarrow \boldsymbol{e}_{mk_m}\}.$$

Первые векторы серий из базиса составляют базис пересечения $\operatorname{Im}\mathcal{A} \cap \operatorname{Ker}\mathcal{A}$. Действительно, ясно, что они линейно независимы. Все они принадлежат $\operatorname{Im}\mathcal{A} \cap \operatorname{Ker}\mathcal{A}$, а значит, $\operatorname{Im}\mathcal{A} \cap \operatorname{Ker}\mathcal{A}$ содержит их линейную оболочку. С другой стороны, если $\boldsymbol{u} \in \operatorname{Im}\mathcal{A} \setminus \langle\{\boldsymbol{e}_{11}, \boldsymbol{e}_{21}, \ldots, \boldsymbol{e}_{m1}\}\rangle$, то в разложении $\boldsymbol{u}$ по базису $\boldsymbol{E}$ коэффициент при хотя бы одном векторе $\boldsymbol{e}_{rs}$, где $r \leqslant m$ и $1 < s \leqslant k_s$, окажется ненулевым:

$$\boldsymbol{u} = u_{11}\boldsymbol{e}_{11} + u_{12}\boldsymbol{e}_{12} + \cdots + u_{1k_1}\boldsymbol{e}_{1k_1} + \cdots + u_{r\,s-1}\boldsymbol{e}_{r\,s-1} + u_{rs}\boldsymbol{e}_{rs} + \cdots + u_{mk_m}\boldsymbol{e}_{mk_m}.$$

Поскольку

$$\mathcal{A}\boldsymbol{u} = u_{12}\boldsymbol{e}_{11} + u_{13}\boldsymbol{e}_{12} + \cdots + u_{1k_1}\boldsymbol{e}_{1\,k_1-1} + \cdots + u_{rs}\boldsymbol{e}_{r\,s-1} + \cdots + u_{mk_m}\boldsymbol{e}_{m\,k_m-1},$$

$u_{rs}\boldsymbol{e}_{r\,s-1} \neq \mathbf{0}$ и вектор $\boldsymbol{e}_{r\,s-1}$ не выражается через остальные, видим, что $\mathcal{A}\boldsymbol{u} \neq \mathbf{0}$, т.е. $\boldsymbol{u} \notin \operatorname{Ker}\mathcal{A}$. Следовательно, $\operatorname{Im}\mathcal{A} \cap \operatorname{Ker}\mathcal{A}$ содержится в линейной оболочке первых векторов серий, а значит, совпадает с ней.

Дополним линейно независимую систему $\{\boldsymbol{e}_{11}, \dots, \boldsymbol{e}_{m1}\}$ до базиса $\boldsymbol{E}'$ в $\operatorname{Ker} \mathcal{A}$, добавив векторы $\boldsymbol{e}_1, \dots, \boldsymbol{e}_l$. Каждый из них образует серию длины 1.

Пусть $\boldsymbol{e}_{1k_1} = \mathcal{A}\boldsymbol{e}_{1\,k_1+1}, \dots, \boldsymbol{e}_{mk_m} = \mathcal{A}\boldsymbol{e}_{m\,k_m+1}$. Имеем набор серий

$$\boldsymbol{E}'' = \{\boldsymbol{e}_1, \quad \dots, \quad \boldsymbol{e}_l, \quad \boldsymbol{e}_{11} \leftarrow \dots \leftarrow \boldsymbol{e}_{1k_1} \leftarrow \boldsymbol{e}_{1\,k_1+1},$$
$$\boldsymbol{e}_{21} \leftarrow \dots \leftarrow \boldsymbol{e}_{2k_2} \leftarrow \boldsymbol{e}_{2\,k_2+1}, \quad \dots, \quad \boldsymbol{e}_{m1} \leftarrow \dots \leftarrow \boldsymbol{e}_{mk_m} \leftarrow \boldsymbol{e}_{m\,k_m+1}\}.$$

По лемме 3 векторы в этом наборе линейно независимы. Покажем, что любой вектор $\boldsymbol{w} \in V$ выражается через них.

Вектор $\mathcal{A}\boldsymbol{w}$ принадлежит $\operatorname{Im} \mathcal{A}$ и потому разлагается по базису $\boldsymbol{E}$:

$$\mathcal{A}\boldsymbol{w} = w_{11}\boldsymbol{e}_{11} + \dots + w_{1k_1}\boldsymbol{e}_{1k_1} + \dots + w_{m1}\boldsymbol{e}_{11} + \dots + w_{1k_1}\boldsymbol{e}_{1k_1}.$$

Рассмотрим вектор

$$\boldsymbol{w}_1 = w_{11}\boldsymbol{e}_{12} + \dots + w_{1k_1}\boldsymbol{e}_{1\,k_1+1} + \dots + w_{m1}\boldsymbol{e}_{12} + \dots + w_{mk_m}\boldsymbol{e}_{m\,k_m+1}.$$

Имеем $\mathcal{A}\boldsymbol{w}_1 = \mathcal{A}\boldsymbol{w}$. Следовательно, $\boldsymbol{w}_2 = \boldsymbol{w} - \boldsymbol{w}_1 \in \operatorname{Ker} \mathcal{A}$, а значит, $\boldsymbol{w}_2$ выражается через векторы из базиса ядра $\boldsymbol{E}'$. Вспомнив, что $\boldsymbol{w}_1$ выражается через векторы из $\boldsymbol{E}''$, видим, что $\boldsymbol{w} = \boldsymbol{w}_1 + \boldsymbol{w}_2$ выражается через векторы из $\boldsymbol{E}''$. $\square$

В жордановом базисе $\boldsymbol{E} = \{\boldsymbol{e}_{11} \leftarrow \boldsymbol{e}_{12} \leftarrow \dots \leftarrow \boldsymbol{e}_{1k_1}, \quad \dots, \ \boldsymbol{e}_{m1} \leftarrow \boldsymbol{e}_{m2} \leftarrow \dots \leftarrow \boldsymbol{e}_{mk_m}\}$ матрица нильпотентного оператора $\mathcal{A}$ имеет блочно-диагональный вид, который называется жордановой формой:

$$A = \begin{array}{c} \begin{array}{cccccccccccccccc} 1 & 2 & 3 & 4 & \dots & k_1 & 1 & 2 & 3 & \dots & k_2 & \dots\dots\dots & 1 & 2 & 3 & \dots \; k_m \end{array} \\ \begin{pmatrix} \begin{matrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & \ddots & \ddots \\ 0 & & & 0 & 1 \\ & & & & 0 \end{matrix} & & & 0 \\ & \begin{matrix} 0 & 1 & & 0 \\ & 0 & 1 & \\ & & \ddots & \ddots \\ 0 & & 0 & 1 \\ & & & 0 \end{matrix} & & \\ & & \ddots & \\ 0 & & & \begin{matrix} 0 & 1 & & 0 \\ & 0 & 1 & \\ & & \ddots & \ddots \\ 0 & & 0 & 1 \\ & & & 0 \end{matrix} \end{pmatrix} \end{array}.$$

Блоки называются жордановыми клетками. В каждой клетке первый столбец нулевой, а все остальные столбцы (если они есть) ненулевые. Число клеток (серий) равно числу нулевых столбцов. Остальные столбцы линейно независимы. Значит, число клеток $= \dim V - \operatorname{rank} A$.

При возведении A в квадрат каждая единица либо сдвигается вправо на одну позицию, либо исчезает (если она находится в последнем столбце своей клетки). При этом вторые столбцы всех клеток (размера $\geqslant 2$) становятся нулевыми. Все ненулевые столбцы остаются линейно независимыми. Поэтому $\operatorname{rank} A^2 = \operatorname{rank} A -$ число клеток размера $\geqslant 2$.

...

При умножении матрицы A^k на A каждая единица либо сдвигается вправо на одну позицию, либо исчезает (если она находится в последнем столбце своей клетки). При этом $(k+1)$-е столбцы всех клеток (размера $\geqslant k+1$) становятся нулевыми. Все ненулевые столбцы остаются линейно независимыми. Поэтому $\operatorname{rank} A^{k+1} = \operatorname{rank} A^k -$ число клеток размера $\geqslant k+1$.

Таким образом, полагая $A^0 = E$, можем написать общую формулу: для каждого натурального k

$$\text{число клеток размера} \geqslant k = \operatorname{rank} A^{k-1} - \operatorname{rank} A^k$$

(считаем, что $\operatorname{rank} A^0 = \dim V$), откуда

$$\boxed{\text{число клеток размера } k = \operatorname{rank} A^{k-1} - 2\operatorname{rank} A^k + \operatorname{rank} A^{k+1}}.$$

Вывод: Жорданова форма матрицы нильпотентного оператора определена однозначно с точностью до перестановки клеток на диагонали.

Жорданова форма матрицы

Для любого $\lambda \in K$ и любого натурального k матрица размера k

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

называется жордановой клеткой порядка k с собственным значением λ. Для $k = 1$ имеем $J_1(\lambda) = (\lambda)$. Блочно-диагональная матрица вида

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

называется жордановой матрицей.

Если матрица A линейного оператора $\mathcal{A}$ является жордановой в некотором базисе, то этот базис называется жордановым базисом, а сама матрица A называется жордановой нормальной формой (или просто жордановой формой) матрицы оператора $\mathcal{A}$. Про матрицу A при этом говорят, что она приводится к жордановой форме.

Предположим, что оператор $\mathcal{A}$ таков, что для некоторого $\lambda \in K$ оператор $\mathcal{A} - \lambda\mathcal{E}$ нильпотентен. Мы видели, что для всякого нильпотентного оператора существует жорданов базис (причём все собственные числа жордановых клеток в его жордановой форме равны нулю). В этом базисе матрица оператора $\mathcal{A}$ имеет вид

$$A = \begin{array}{c} \begin{matrix} 1 & 2 & 3 & 4 & \dots & k_1 & 1 & 2 & 3 & \dots & k_2 & \dots & \dots\dots & 1 & 2 & 3 & \dots & k_m \end{matrix} \\ \left(\begin{array}{cccccccccccccc} \lambda & 1 & & & & & 0 & & & & & & & 0 \\ & \lambda & 1 & & & & & & & & & & & \\ & & \lambda & 1 & & & & & & & & & & \\ & & & \ddots & \ddots & & & & & & & & & \\ 0 & & & & \lambda & 1 & & & & & & & & \\ & & & & & \lambda & & & & & & & & \\ & & & & & & \lambda & 1 & & 0 & & & & \\ & & & & & & & \lambda & 1 & & & & & \\ & & & & & & & & \ddots & \ddots & & & & \\ & & & & & & 0 & & & \lambda & 1 & & & \\ & & & & & & & & & & \lambda & & & \\ & & & & & & & & & & & \ddots & & \\ & & & & & & & & & & & & \ddots & \\ & & & & & & & & & & & & & \ddots \\ & & & & & & & & & \lambda & 1 & & 0 & \\ & & & & & & & & & & \lambda & 1 & & \\ & & & & & & & & & & & \ddots & \ddots & \\ 0 & & & & & & & & & 0 & & & \lambda & 1 \\ & & & & & & & & & & & & & \lambda \end{array}\right) \end{array} = \begin{pmatrix} J_{k_1}(\lambda) & 0 & \dots & 0 \\ 0 & J_{k_2}(\lambda) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{k_m}(\lambda) \end{pmatrix}.$$

Вспомним, что если у линейного оператора $\mathcal{A}$ в конечномерном пространстве V некоторый аннулирующий многочлен раскладывается в произведение линейных множителей, то найдутся инвариантные подпространства $V_1, \dots, V_k \subset V$ и попарно различные собственные значения $\lambda_1, \dots, \lambda_k$ оператора $\mathcal{A}$ такие, что $V = V_1 \oplus \dots \oplus V_k$ и все операторы $\mathcal{A}|_{V_i} - \lambda_i \mathcal{E}$ нильпотентны. Для каждого $i \leqslant k$ возьмём жорданов базис $\boldsymbol{E}_i$ оператора $\mathcal{A}|_{V_i}$ в пространстве V_i. Пусть $k_{i1}, \dots, k_{im_i}$ — размеры жордановых клеток в этом базисе. Тогда $\boldsymbol{E} = \boldsymbol{E}_1 \cup \dots \cup \boldsymbol{E}_k$ — базис пространства $\boldsymbol{E}$, и в этом базисе оператор $\mathcal{A}$ имеет жорданову матрицу. Итак,

если некоторый аннулирующий многочлен линейного оператора в конечномерном векторном пространстве над полем K раскладывается в произведение линейных множителей (в частности, если поле K алгебраически замкнуто), то существует жорданов базис для этого оператора.

Из вида жордановой матрицы понятно, что

- $\chi_{\mathcal{A}}(\lambda) = (\lambda_1 - \lambda)^{\dim V_1} \cdot (\lambda_2 - \lambda)^{\dim V_2} \cdot \ldots \cdot (\lambda_k - \lambda)^{\dim V_k}$, так что $\lambda_1, \dots, \lambda_k$ — это все собственные значения $\mathcal{A}$.
- Для любого $\lambda \in K$ оператор $\mathcal{A} - \lambda \mathcal{E}$ имеет тот же жорданов базис, и его жорданова матрица состоит из клеток $J_{k_{i1}}(\lambda_i - \lambda), \dots, J_{k_{im_i}}(\lambda_i - \lambda)$, $i \leqslant k$.

$$
A = \begin{pmatrix}
J_{k_{11}}(\lambda_1) & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
0 & J_{k_{12}}(\lambda_1) & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & J_{k_{1m_1}}(\lambda_1) & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & J_{k_{21}}(\lambda_2) & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & \cdots & J_{k_{2m_2}}(\lambda_2) & \cdots & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & J_{k_{k1}}(\lambda_k) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & J_{k_{km_k}}(\lambda_k)
\end{pmatrix} =
$$

$$
= \begin{pmatrix}
\lambda_1 & 1 & & 0 & & & & & & & & & \text{\Huge 0} \\
& \lambda_1 & 1 & & & & & & & & & & \\
& & \ddots & \ddots & & & & & & & & & \\
0 & & & \lambda_1 & 1 & & & & & & & & \\
& & & & \lambda_1 & & & & & & & & \\
& & & & & \lambda_i & 1 & 0 & & & & & \\
& & & & & & \ddots & \ddots & & & & & \\
& & & & 0 & & & \lambda_i & 1 & & & & \\
& & & & & & & & \lambda_i & & & & \\
& & & & & & & & & \ddots & & & \\
& & & & & & & & & & \ddots & & \\
& & & & & & & & & & & \ddots & \\
& & & & & & & & & & \lambda_k & 1 & 0 \\
& & & & & & & & & & & \ddots & \ddots \\
& & & & & & & & & & 0 & \lambda_k & 1 \\
\text{\Huge 0} & & & & & & & & & & & & \lambda_k
\end{pmatrix}
$$

Замечание

Для любых натуральных r и s и любых различных $\mu, \lambda \in K$ $(J_r(\lambda) - \mu E)^s$ — верхнетреугольная матрица с диагональными элементами $(\lambda - \mu)^s$ и $(J_r(\lambda) - \lambda E)^r = \Theta$.

Для любого r и любого $i \leqslant k$ матрица $(A - \lambda_i E)^r$ имеет блочно-диагональный вид с блоками $J_{k_{js}}(\lambda_j - \lambda_i)^r$. Все такие блоки с $j \neq i$ невырождены, а блоки $J_{k_{is}}(\lambda_i - \lambda_i)^r$ вырождены и при $k_{is} \leqslant r$ равны Θ.

$$(A - \lambda_i E)^r =$$

$$= \begin{pmatrix}
J_{k_{11}}(\lambda_1-\lambda_i)^r & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & J_{k_{1m_1}}(\lambda_1-\lambda_i)^r & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \cdots & J_{k_{i1}}(0)^r & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \cdots & 0 & \cdots & J_{k_{im_2}}(0)^r & \cdots & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & J_{k_{k1}}(\lambda_k-\lambda_i)^r & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & J_{k_{km_k}}(\lambda_k-\lambda_i)^r
\end{pmatrix}$$

Отсюда и из формулы для числа жордановых клеток данного размера у нильпотентного оператора вытекает такое утверждение:

Теорема

Если матрица линейного оператора $\mathcal{A}$ приводится к жордановой форме, то эта форма определена однозначно с точностью до перестановки жордановых клеток на диагонали. А именно, для любого $\lambda \in K$ и любого натурального k

$$\boxed{\text{число клеток } J_k(\lambda) = \operatorname{rank}(\mathcal{A}-\lambda\mathcal{E})^{k-1} - 2\operatorname{rank}(\mathcal{A}-\lambda\mathcal{E})^{k} + \operatorname{rank}(\mathcal{A}-\lambda\mathcal{E})^{k+1}}$$

(ранг оператора $(\mathcal{A}-\lambda)^0$ полагается равным нулю).

В частности, число жордановых клеток, диагональные элементы которых не являются собственными значениями оператора, равно нулю.

Предложение

Если матрица линейного оператора $\mathcal{A}\colon V \to V$ приводится к жордановой форме, то пространство V раскладывается в прямую сумму всех корневых подпространств. Для каждого собственного значения λ оператора $\mathcal{A}$ размерность корневого подпространства V^{λ} равна кратности λ как корня характеристического многочлена.

Доказательство. Пусть $\boldsymbol{E} = \boldsymbol{E}_1 \cup \dots \cup \boldsymbol{E}_k$ — жорданов базис, в котором векторы из $\boldsymbol{E}_i$, $i \leqslant k$, соответствуют жордановым клеткам с λ_i на диагонали. Если в разложении вектора $\boldsymbol{x} \in V$ по базису $\boldsymbol{E}$ коэффициент хотя бы при одном базисном векторе из хотя бы одного набора $\boldsymbol{E}_j$, $j \neq i$, ненулевой, то $(\mathcal{A} - \lambda_i \mathcal{E})^s \boldsymbol{x} \neq \boldsymbol{0}$ для любого натурального s. Значит, все корневые векторы, отвечающие собственному значению λ_i, содержатся в $\langle \boldsymbol{E}_i \rangle$. С другой стороны, $\langle \boldsymbol{E}_i \rangle \subset V^{\lambda_i}$, поскольку $\langle \boldsymbol{E}_i \rangle \subset \operatorname{Ker}(\mathcal{A} - \lambda_i \mathcal{E})^{s_i}$, где s_i — максимальный размер жордановой клетки с λ_i на диагонали. Следовательно, $\langle \boldsymbol{E}_i \rangle$ совпадает с корневым подпространством V^{λ_i} для каждого i. При этом число векторов в $\boldsymbol{E}_i$ равно сумме размеров жордановых клеток с λ_i на диагонали, т.е. степени одночлена $\lambda_i - \lambda$ в разложении $\chi_{\mathcal{A}}(\lambda)$ на линейные множители. $\square$

Определение

Каждое собственное значение λ_0 линейного оператора $\mathcal{A}$ является корнем характеристического многочлена $\chi_{\mathcal{A}}(\lambda)$. Кратность корня λ_0 называется алгебраической кратностью характеристического значения λ_0. Размерность собственного подпространства V_{λ_0} называется геометрической кратностью собственного значения λ_0.

Предложение

Если у линейного оператора $\mathcal{A}$ существует жорданова форма, то геометрическая кратность всякого его собственного значения λ равно числу жордановых клеток с λ на диагонали в жордановой матрице оператора $\mathcal{A}$. Следовательно, геометрическая кратность не превосходит алгебраической кратности.

Доказательство. Собственное подпространство V_λ оператора $\mathcal{A}$ — это линейная оболочка векторов жорданова базиса, соответствующих первым столбцам жордановых клеток вида $J_k(\lambda)$. (Действительно, эти столбцы линейно независимы, и мы знаем, что $\dim V_\lambda = \dim V - \operatorname{rank}(\mathcal{A} - \lambda\mathcal{E})$, а это как раз число клеток вида $J_k(\lambda)$.) Поскольку размерность корневого подпространства V^λ равна сумме размеров жордановых клеток вида $J_k(\lambda)$, имеем $\dim V_\lambda \leqslant \dim V^\lambda$. $\square$

Поскольку любая диагональная матрица жорданова, из последнего предложения и единственности жордановой формы вытекает следующая теорема.

Теорема

Для диагонализируемости линейного оператора $\mathcal{A}$ в n-мерном векторном пространстве V (т.е. существования в пространстве V базиса из собственных векторов оператора $\mathcal{A}$) необходимо и достаточно выполнение двух условий:

1. *характеристический многочлен оператора $\mathcal{A}$ раскладывается на линейные множители;*
2. *геометрическая кратность каждого собственного значения оператора $\mathcal{A}$ равна его алгебраической кратности.*

Предложение

У линейного оператора $\mathcal{A}$ имеется жорданова форма $\iff$ характеристический многочлен $\chi_{\mathcal{A}}(\lambda)$ раскладывается в произведение линейных множителей и является аннулирующим.

Доказательство. *Достаточность.* Мы умеем строить жорданов базис, имея любой аннулирующий многочлен, разложенный на линейные множители.

Необходимость. Имеем $\chi_{\mathcal{A}}(\lambda) = (\lambda_1 - \lambda)^{s_1} \cdot (\lambda_2 - \lambda)^{s_2} \cdot \ldots \cdot (\lambda_k - \lambda)^{s_k}$, где $s_i = \dim V^{\lambda_i}$, так что в жордановом базисе $\mathcal{A}$ имеет матрицу

$$A = (-1)^{\dim V} \cdot (A - \lambda_1 E)^{s_1} \cdot (A - \lambda_2 E)^{s_2} \cdot \ldots \cdot (A - \lambda_k E)^{s_k}$$

и

$$\chi_A(A) = (-1)^{\dim V} \begin{pmatrix} J_{k_{11}}(\lambda_1-\lambda_1)^{s_1} \cdots J_{k_{11}}(\lambda_1-\lambda_k)^{s_k} & & 0 \\ & \ddots & \\ & J_{k_{ij}}(\lambda_i-\lambda_1)^{s_1} \cdots J_{k_{ij}}(\lambda_i-\lambda_k)^{s_k} & \\ & & \ddots \\ 0 & & J_{k_{km_k}}(\lambda_k-\lambda_1)^{s_1} \cdots J_{k_{km_k}}(\lambda_k-\lambda_k)^{s_k} \end{pmatrix},$$

причём $k_{ij} \leqslant s_i$ для $j \leqslant m_j$, а значит, $J_{k_{ij}}(\lambda_i - \lambda_i)^{s_i} = \Theta$.

Следовательно, $\chi_A(A) = \Theta$, откуда $\chi_{\mathcal{A}}(\mathcal{A}) = \Theta$. □

Теорема Гамильтона–Кэли

Теорема Гамильтона–Кэли

Для всякого линейного оператора в конечномерном векторном пространстве характеристический многочлен является аннулирующим.

Доказательство. Пусть $\mathcal{A}$ — линейный оператор в конечномерном векторном пространстве V над полем K. Предположим, что K алгебраически замкнуто. Тогда любой аннулирующий многочлен оператора $\mathcal{A}$ раскладывается на линейные множители. Следовательно, у $\mathcal{A}$ имеется жорданова форма. По предыдущему предложению характеристический многочлен является аннулирующим.

Пусть теперь K — произвольное поле, и пусть $\widetilde{K}$ — его алгебраически замкнутое расширение. Определим оператор $\widetilde{\mathcal{A}}\colon \widetilde{K}^n \to \widetilde{K}^n$ правилом $\widetilde{\mathcal{A}}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ для $x_1, \ldots, x_n \in \widetilde{K}$, где A — матрица $\mathcal{A}$ в любом базисе. Матрица оператора $\widetilde{\mathcal{A}}$ — это A, поэтому $\chi_{\widetilde{\mathcal{A}}}(\lambda) = \chi_A(\lambda)$. По доказанному $\chi_A(A) = \Theta$, а значит, $\chi_{\widetilde{\mathcal{A}}}(\widetilde{\mathcal{A}}) = \mathit{\Theta}$ ($\mathit{\Theta}$ — нулевой оператор). □

Следствия теоремы Гамильтона–Кэли

Теорема

Для линейного оператора A в n-мерном векторном пространстве V следующие условия равносильны:

1. *характеристический многочлен оператора $\mathcal{A}$ раскладывается на линейные множители;*
2. *существует базис пространства V, в котором $\mathcal{A}$ имеет жорданову матрицу;*
3. *пространство V является прямой суммой корневых подпространств оператора $\mathcal{A}$.*

Следствие

Для любого линейного оператора в n-мерном векторном пространстве над алгебраически замкнутым полем существует жорданов базис.

Следствие

Любая квадратная матрица с коэффициентами из алгебраически замкнутого поля K подобна жордановой матрице.

Теорема Гамильтона–Кэли сводит вычисление любой полиномиальной функции от линейного оператора в n-мерном пространстве к вычислению полиномиальной функции степени меньше n от этого оператора.

Минимальный многочлен

Пусть $\mathcal{A}\colon V \to V$ — линейный оператор в конечномерном векторном пространстве V.
По теореме Гамильтона–Кэли любой оператор и соответствующая ему матрица A (в любом базисе) аннулируется своим характеристическим многочленом.

Определение

Аннулирующий многочлен для линейного оператора $\mathcal{A}$ (матрицы A) наименьшей положительной степени со старшим коэффициентом 1 называется **минимальным многочленом** для оператора $\mathcal{A}$ (соответственно, для матрицы A) и обозначается $\mu_{\mathcal{A}}(\lambda)$ (соответственно, $\mu_A(\lambda)$).

Теорема

1. *Минимальный многочлен линейного оператора равен минимальному многочлену его матрицы в любом базисе;*
2. *минимальный многочлен является делителем любого аннулирующего многочлена;*
3. *минимальный многочлен единствен;*
4. *наборы корней (без учёта кратности) у минимального многочлена и характеристического многочлена совпадают.*

Доказательство. Утверждение 1 вытекает из изоморфизма алгебр операторов и матриц.

2: Пусть $\mathcal{A}$ — оператор и $f(\lambda)$ — любой аннулирующий его многочлен. Поделим $f(\lambda)$ на $\mu_{\mathcal{A}}(\lambda)$ с остатком:

$$f(\lambda) = \mu_{\mathcal{A}}(\lambda) \cdot g(\lambda) + r(\lambda), \quad \text{где } \deg r(\lambda) < \deg \mu_{\mathcal{A}}(\lambda).$$

Подставляя $\mathcal{A}$ вместо λ, получаем $r(\mathcal{A}) = f(\mathcal{A}) - \mu_{\mathcal{A}}(\mathcal{A})$, так что $r(\lambda)$ — аннулирующий многочлен для $\mathcal{A}$. Из минимальности $\mu_{\mathcal{A}}$ следует, что $r(\lambda) \equiv 0$.

Утверждение ③ вытекает из ②: любой из двух минимальных многочленов линейного оператора делит другой, а их старшие коэффициенты совпадают.

④: Теорема Гамильтона–Кэли + ② $\Longrightarrow$ $\mu_{\mathcal{A}}(\lambda)$ делит $\chi_{\mathcal{A}}(\lambda)$ $\Longrightarrow$ все корни минимального многочлена являются корнями характеристического многочлена.

Обратно, пусть λ_0 — корень характеристического многочлена. Тогда у $\mathcal{A}$ есть собственный вектор $\boldsymbol{v}$ с собственным значением λ_0. Применим к вектору $\boldsymbol{v}$ оператор $\mu_{\mathcal{A}}(\mathcal{A})$. Пусть $\mu_{\mathcal{A}}(\lambda) = \lambda^k + a_{k-1}\lambda^{k-1} + \cdots + a_1\lambda + a_0$. Подставив $\mathcal{A}$ вместо переменной λ, получаем

$$\begin{aligned}\mu_{\mathcal{A}}(A)\boldsymbol{v} = (\mathcal{A}^k + a_{k-1}\mathcal{A}^{k-1} + \cdots + a_1\mathcal{A} + a_0\mathcal{E})\boldsymbol{v} = \\ = \mathcal{A}^k\boldsymbol{v} + a_{k-1}\mathcal{A}^{k-1}\boldsymbol{v} + \cdots + a_1\mathcal{A}\boldsymbol{v} + a_0\mathcal{E}\boldsymbol{v} = \\ = \lambda_0^k\boldsymbol{v} + a_{k-1}\lambda_0^{k-1}\boldsymbol{v} + \cdots + a_1\lambda_0\boldsymbol{v} + a_0\boldsymbol{v} = \\ (\lambda_0^k + a_{k-1}\lambda_0^{k-1} + \cdots + a_1\lambda_0 + a_0)\boldsymbol{v} = \mu_{\mathcal{A}}(\lambda_0)\boldsymbol{v}.\end{aligned}$$

Поскольку многочлен $\mu_{\mathcal{A}}(\lambda)$ аннулирует оператор $\mathcal{A}$, имеем $\mu_{\mathcal{A}}(\mathcal{A})\boldsymbol{v} = \mathbf{0}$. Вектор $\boldsymbol{v}$ собственный $\Longrightarrow$ $\boldsymbol{v} \neq \mathbf{0}$ $\Longrightarrow$ $\mu_{\mathcal{A}}(\lambda_0) = 0$. $\square$

Теорема Гамильтона–Кэли $\Longrightarrow$ степень минимального многочлена оператора $\mathcal{A}\colon V \to V$ не превосходит $\dim V$. Она может быть любым числом от 1 до $\dim V$.

Пример

Если оператор имеет диагональную матрицу, то степень его минимального многочлена равна числу различных диагональных элементов.

Предложение

Для линейного оператора существует жорданов базис $\Longleftrightarrow$ *минимальный многочлен этого оператора раскладывается на линейные множители. При этом кратность каждого корня* λ *минимального многочлена равна максимальному размеру жордановой клетки с* λ *на диагонали в жордановой матрице оператора.*

Следствие

Линейный оператор диагонализируем тогда и только тогда, когда его минимальный многочлен раскладывается на линейные множители и не имеет кратных корней.

Упражнения

1. Докажите, что для любого натурального n, любого поля K и любых матриц $A, B \in M_n(K)$ характеристические многочлены матриц AB и BA совпадают.
2. Докажите, что для любого натурального n и любого бесконечного поля K всякая матрица $A \in M_n(K)$ подобна матрице A^T.
3. Докажите, что для любого натурального n и любого бесконечного поля K всякая матрица $A \in M_n(K)$ раскладывается в произведение двух симметричных матриц, одна из которых невырождена.
4. Докажите, что всякая матрица $A \in M_n(\mathbb{C})$ подобна симметричной матрице.

Сопряжённые операторы

Пусть теперь V — произвольное (не обязательно конечномерное) векторное пространство над произвольным полем K, и пусть $\mathcal{A}\colon V \to V$ — линейный оператор. Линейный оператор $\mathcal{A}^*\colon V^* \to V^*$, сопряжённый к оператору $\mathcal{A}$, определяется так же, как и в случае общих линейных отображений (см. раздел после факторпространств):

$$\boxed{\mathcal{A}^*(f) = f \circ \mathcal{A}} \qquad \text{для всякого } f \in V^*.$$

Если пространство V конечномерно и A — матрица оператора $\mathcal{A}$ в некотором его базисе, то, как мы знаем из теории общих линейных отображений, матрица A^* оператора $\mathcal{A}^*$ во взаимном базисе — это транспонированная матрица A:

$$\boxed{A^* = A^T}.$$

Отсюда следует, что сопряжённые операторы имеют одинаковые характеристические многочлены, и если матрица одного из них приводится к жордановой форме, то для другого это тоже верно, причём жордановы формы их матриц одинаковы.

Упражнения

1. Покажите, что оператор $\mathcal{A}$ в конечномерном пространстве V и оператор $\mathcal{B}$ в пространстве V^* сопряжены $\iff$ матрица оператора $\mathcal{A}$ в некотором базисе равна транспонированной матрице оператора $\mathcal{B}$ во взаимном базисе $\iff$ матрица оператора $\mathcal{A}$ в любом базисе равна транспонированной матрице оператора $\mathcal{B}$ во взаимном базисе.
2. Вспомните, что в случае, когда V конечномерно, существует естественный изоморфизм между пространствами V и V^{**}, который позволяет отождествить эти пространства. Покажите, что при таком отождествлении $\mathcal{A}^{**} = \mathcal{A}$.
3. Заметьте, что для конечномерного пространства V отображение $*$ пространства $\mathscr{L}(V)$ линейных операторов в V в пространство $\mathscr{L}(V)$ линейных операторов в V^*, определённое правилом $\mathcal{A} \mapsto \mathcal{A}^*$, является изоморфизмом. Верно ли это для произвольного пространства V?

Билинейные функции

Пусть V — векторное пространство над полем K.

Определение

Билинейной функцией на V называется функция $b\colon V \times V \to K$, линейная по каждому аргументу.

Линейность по каждому аргументу означает, что для любых $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ и любого $\lambda \in K$

$$
\begin{aligned}
b(x+y,z) &= b(x,z)+b(y,z), \qquad & b(x,y+z) &= b(x,y)+b(x,z),\\
b(\lambda x,y) &= \lambda b(x,y), \qquad & b(x,\lambda y) &= \lambda b(x,y).
\end{aligned}
$$

Билинейные функции складываются и умножаются на числа поточечно. Относительно этих операций все билинейные функции на V образуют векторное пространство над полем K. Будем обозначать его $\mathscr{B}(V)$.

Примеры

- Скалярное произведение в обычном пространстве.
- Для любых $f, g \in V^*$ $\;b(\boldsymbol{x}, \boldsymbol{y}) = f(\boldsymbol{x}) \cdot g(\boldsymbol{y})$ — билинейная функция на V.
- Для любых $\boldsymbol{x}, \boldsymbol{y} \in V$ $\;b(f, g) = f(\boldsymbol{x}) \cdot g(\boldsymbol{y})$ — билинейная функция на V^*.
- $\operatorname{tr} AB$ для $A, B \in M_n(K)$.
- В пространстве непрерывных функций на $[0, 1]$

$$\int\limits_0^1 f(x) \cdot g(x)\, dx, \qquad \int\limits_0^1 \int\limits_0^1 \varphi(x, y) \cdot f(x) \cdot g(y)\, dx dy,$$

где φ — фиксированная непрерывная функция на $[0, 1]$.

Корреляции и ядра

Задание билинейной функции $b\colon V \times V \to K$ равносильно заданию линейного отображения

$$b_{\text{лев}}\colon V \to V^*, \qquad \boldsymbol{v} \mapsto b(\boldsymbol{v}, \boldsymbol{x}).$$

Это отображение называется левой корреляцией билинейной функции b. Аналогично определяется правая корреляция

$$b_{\text{пр}}\colon V \to V^*, \qquad \boldsymbol{v} \mapsto b(\boldsymbol{x}, \boldsymbol{v}).$$

Отображения $b \mapsto b_{\text{лев}}$ и $b \mapsto b_{\text{пр}}$ — изоморфизмы между пространством $\mathscr{B}(V)$ и пространством $\mathscr{L}(V, V^*)$ линейных отображений $V \to V^*$. ▸ к матрицам ▸ к тензорам

Ядра левой и правой корреляций называются, соответственно, левым и правым ядрами билинейной функции:

$$\operatorname{Ker}_{\text{лев}} b = \{\boldsymbol{v} \in V : b(\boldsymbol{v}, \boldsymbol{x}) = 0 \ \forall \boldsymbol{x} \in V\}, \quad \operatorname{Ker}_{\text{пр}} b = \{\boldsymbol{v} \in V : b(\boldsymbol{x}, \boldsymbol{v}) = 0 \ \forall \boldsymbol{x} \in V\}.$$

Упражнение

Приведите пример векторного пространства V и билинейной функции $b \in \mathscr{B}(V)$, для которой одно из ядер тривиально, а другое нет.

Вектор $\boldsymbol{u} \in V$ ортогонален слева вектору $\boldsymbol{v}$, а вектор $\boldsymbol{v}$ ортогонален справа вектору $\boldsymbol{u}$, если $b(\boldsymbol{u}, \boldsymbol{v}) = 0$. Левое ортогональное дополнение ${}^{\perp}U$ подпространства $U \subset V$ — это множество векторов, ортогональных слева каждому вектору $\boldsymbol{u} \in U$:

$$ {}^{\perp}U = \{\boldsymbol{x} \in V : b(\boldsymbol{x}, \boldsymbol{u}) = 0 \ \forall \boldsymbol{u} \in U\}. $$

Аналогичным образом определяется правое ортогональное дополнение $U^{\perp}$:

$$ U^{\perp} = \{\boldsymbol{x} \in V : b(\boldsymbol{u}, \boldsymbol{x}) = 0 \ \forall \boldsymbol{u} \in U\}. $$

Левое ортогональное дополнение ${}^{\perp}U$ — не что иное как аннулятор $\operatorname{Ann} U$ относительно свёртки $c(\boldsymbol{x}, \boldsymbol{y}) = b(\boldsymbol{x}, \boldsymbol{y})$, а правое ортогональное дополнение $U^{\perp}$ — это $\operatorname{Ann} U$ относительно свёртки $c(\boldsymbol{x}, \boldsymbol{y}) = b(\boldsymbol{y}, \boldsymbol{x})$.

Свойства правого ортогонального дополнения

- $U^{\perp}$ — подпространство пространства V, и если $W \underset{\text{lin}}{\subset} U \underset{\text{lin}}{\subset} V$, то $W^{\perp} \underset{\text{lin}}{\supset} U^{\perp}$.
- $U \subset {}^{\perp}(U^{\perp})$.
- $\operatorname{Ker}_{\text{пр}} b \underset{\text{lin}}{\subset} U^{\perp}$ для любого $U \underset{\text{lin}}{\subset} V$, $\operatorname{Ker}_{\text{пр}} b = V^{\perp}$.

Свойства левого ортогонального дополнения аналогичны.

Определение

Ненулевой вектор $\boldsymbol{v} \in V$ называется изотропным, если он ортогонален (слева = справа) сам себе. Множество всех изотропных векторов — изотропный конус. Подпространство $U \underset{\text{lin}}{\subseteq} V$ изотропно (слева/справа), если существует ненулевой вектор $\boldsymbol{u} \in U$, которому ортогонален (слева/справа) каждый $\boldsymbol{x} \in U$. Подпространство $U \underset{\text{lin}}{\subseteq} V$ вполне изотропно, если $U = U^{\perp}$ $(= {}^{\perp}U)$.

Определение

Билинейная функция невырождена, если оба её ядра тривиальны.

Упражнения

1. Заметьте, что билинейная функция b на векторном пространстве V невырождена тогда и только тогда, когда $V^{\perp} = {}^{\perp}V = \{\mathbf{0}\}$.
2. Покажите, что $U \underset{\text{lin}}{\subseteq} V$ изотропно слева или справа тогда и только тогда, когда ограничение $b|_{U \times U} \colon U \times U \to K$ вырождено.
3. Приведите пример невырожденной билинейной функции на пространстве V, для которой всякое одномерное подпространство $U \subset V$ вполне изотропно.

Предложение

Пусть b — билинейная функция на векторном пространстве V и $U \underset{\text{lin}}{\subset} V$.

1. *Левое ортогональное дополнение ${}^{\perp}U$ подпространства U — это прообраз при левой корреляции $b_{\text{лев}}\colon V \to V^*$ аннулятора*
$$U^0 = \operatorname{Ann} U = \{\boldsymbol{f} \in V^* : (\boldsymbol{f}, \boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) = 0 \ \forall \boldsymbol{x} \in U\}$$
подпространства U (относительно естественного спаривания $c(\boldsymbol{f}, \boldsymbol{u}) = \boldsymbol{f}(\boldsymbol{u})$).
2. $U^{\perp} = b_{\text{пр}}^{-1}(U^0)$.
3. *Подпространство U вполне изотропно тогда и только тогда, когда его образы $b_{\text{лев}}(U)$ и $b_{\text{пр}}(U)$ при левой и правой корреляциях содержатся в аннуляторе U^0.*

Упражнение

Докажите, что если подпространство U векторного пространства V конечномерно и ограничение билинейной функции $b\colon V \times V \to K$ на это подпространство невырождено, то $V = U \oplus U^{\perp}$.

Билинейные функции на конечномерных пространствах

Всюду ниже V — конечномерное векторное пространство (если не указано иное).

Матрица билинейной функции

Пусть V — n-мерное векторное пространство, $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ — базис в V и b — любая билинейная функция на V.

Матрица

$$B = \begin{pmatrix} b(\boldsymbol{e}_1, \boldsymbol{e}_1) & b(\boldsymbol{e}_1, \boldsymbol{e}_2) & \ldots & b(\boldsymbol{e}_1, \boldsymbol{e}_n) \\ b(\boldsymbol{e}_2, \boldsymbol{e}_1) & b(\boldsymbol{e}_2, \boldsymbol{e}_2) & \ldots & b(\boldsymbol{e}_2, \boldsymbol{e}_n) \\ \vdots & \vdots & \ddots & \vdots \\ b(\boldsymbol{e}_n, \boldsymbol{e}_1) & b(\boldsymbol{e}_n, \boldsymbol{e}_2) & \ldots & b(\boldsymbol{e}_n, \boldsymbol{e}_n) \end{pmatrix}$$

называется матрицей Грама, или просто матрицей, а её элементы $b_{ij} = b(\boldsymbol{e}_i, \boldsymbol{e}_j)$ — коэффициентами билинейной функции b в базисе $\boldsymbol{E}$.

Для любых векторов $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$ и $\boldsymbol{y} = y_1\boldsymbol{e}_1 + \cdots + y_n\boldsymbol{e}_n$

$$b(\boldsymbol{x}, \boldsymbol{y}) = (x_1, \ldots, x_n)\, B \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Пусть даны два базиса $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ и $\boldsymbol{E}' = \{\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_n\}$ векторного пространства V, C — матрица перехода от $\boldsymbol{E}$ к $\boldsymbol{E}'$ и B — матрица билинейной функции $b\colon V \times V \to K$ в базисе $\boldsymbol{E}$. Тогда её матрица B' в базисе $\boldsymbol{E}'$ вычисляется по формуле

$$\boxed{B' = C^T BC}.$$

Следовательно, *ранг матрицы билинейной функции не зависит от базиса*. Ранг матрицы билинейной функции b в любом базисе называется рангом билинейной функции b и обозначается $\operatorname{rank} b$.

Предложение

Пусть b — билинейная функция на V, $\boldsymbol{E}$ — любой базис пространства V, B — матрица b в этом базисе и $\mathcal{E}$ — взаимный с $\boldsymbol{E}$ базис пространства V^*. Тогда матрица правой корреляции $b_{пр}$ относительно базисов $\boldsymbol{E}$ и $\mathcal{E}$ равна B, а матрица левой корреляции $b_{лев}$ относительно тех же базисов равна B^T.

к скалярному произведению

Предложение

$\dim \operatorname{Ker}_{лев} b = \dim \operatorname{Ker}_{пр} b$.

(В бесконечномерном случае это не так.)

Упражнение

Приведите пример билинейной функции b, для которой $\operatorname{Ker}_{\text{лев}} b \neq \operatorname{Ker}_{\text{пр}} b$.

Теорема

Следующие свойства билинейной функции b на n-мерном векторном пространстве V равносильны:

1. *функция b невырождена, т.е.* $\operatorname{Ker}_{\textit{лев}} b = \operatorname{Ker}_{\textit{пр}} b = \{\mathbf{0}\}$;
2. *матрица Грама функции b в некотором базисе невырождена;*
3. *матрица Грама функции b в любом базисе невырождена;*
4. *левая корреляция $b_{\textit{лев}} \colon V \to V^*$ — изоморфизм;*
5. *правая корреляция $b_{\textit{пр}} \colon V \to V^*$ — изоморфизм;*
6. *для любого $\boldsymbol{v} \in V \setminus \{\mathbf{0}\}$ существует $\boldsymbol{u} \in V$ такой, что $b(\boldsymbol{v}, \boldsymbol{u}) \neq 0$;*
7. *для любого $\boldsymbol{v} \in V \setminus \{\mathbf{0}\}$ существует $\boldsymbol{u} \in V$ такой, что $b(\boldsymbol{u}, \boldsymbol{v}) \neq 0$;*
8. *для любой линейной функции $f \colon V \to K$ существует $\boldsymbol{v} \in V$ такой, что $f(\boldsymbol{x}) = b(\boldsymbol{v}, \boldsymbol{x})$ для всех $\boldsymbol{x} \in V$;*
9. *для любой линейной функции $f \colon V \to K$ существует $\boldsymbol{v} \in V$ такой, что $f(\boldsymbol{x}) = b(\boldsymbol{x}, \boldsymbol{v})$ для всех $\boldsymbol{x} \in V$.*

При выполнении этих условий вектор $\boldsymbol{v}$ в 8 и 9 определён однозначно.

Предложение

Если билинейная функция b на конечномерном пространстве V невырождена, то для всякого подпространства $U \subset V$ выполняются равенства

$$\dim {}^{\perp}U = \dim V - \dim U = \dim U^{\perp} \quad \textit{и} \quad ({}^{\perp}U)^{\perp} = U = {}^{\perp}(U)^{\perp}.$$

Доказательство. Первые два равенства следуют из того, что ${}^{\perp}U$ и $U^{\perp}$ — прообразы $U^0 \subset V^*$ при изоморфизмах $b_{\text{лев}}$ и $b_{\text{пр}}$ и $\dim \operatorname{Ann} U = \dim V - \dim U$. Вторые два равенства вытекают из первых, поскольку оба подпространства $({}^{\perp}U)^{\perp}$ и ${}^{\perp}(U)^{\perp}$ содержат U и имеют размерность $\dim U$. □

Упражнение

Докажите, что размерность вполне изотропного подпространства невырожденной билинейной функции на пространстве V не превосходит $\frac{1}{2}\dim V$.

Канонический оператор

Билинейная функция b на V невырождена $\iff$ её корреляции $b_{\text{лев}}\colon V \to V^*$ и $b_{\text{пр}}\colon V \to V^*$ — изоморфизмы. Линейный оператор $\kappa_b = b_{\text{пр}}^{-1} \circ b_{\text{лев}}\colon V \to V$ называется **каноническим оператором** невырожденной билинейной функции b.

Теорема

Канонический оператор $\kappa_b\colon V \to V$ невырожденной билинейной функции $b\colon V \times V \to K$ является изоморфизмом. Это единственный линейный оператор $V \to V$ с тем свойством, что

$$\boxed{b(\boldsymbol{x}, \boldsymbol{y}) = b(\boldsymbol{y}, \kappa_b \boldsymbol{x})} \quad \text{для любых } \boldsymbol{x}, \boldsymbol{y} \in V.$$

Доказательство. Из определения следует, что если B — матрица b в некотором базисе $\boldsymbol{E}$, то матрица канонического оператора κ_b в $\boldsymbol{E}$ равна $B^{-1}B^T$. Любой оператор с указанным свойством должен иметь ту же матрицу. □

Имеем изоморфизмы $\mathscr{B}(V) \simeq \mathscr{L}(V, V^*)$: $b \mapsto b_{\text{лев}}$ и $\mathscr{B}(V) \simeq \mathscr{L}(V, V^*)$: $b \mapsto b_{\text{пр}}$, а также изоморфизм $b \mapsto \kappa_b$ между пространством невырожденных билинейных функций на V и пространством изоморфизмов $V \to V$.

Симметричные и кососимметричные билинейные функции

Ниже V — произвольное векторное пространство над полем K.

Определение

Билинейная функция b на V называется симметричной (или симметрической), если $b(\boldsymbol{x}, \boldsymbol{y}) = b(\boldsymbol{y}, \boldsymbol{x})$ для любых векторов $\boldsymbol{x}, \boldsymbol{y} \in V$.

Билинейная функция b на V называется кососимметричной (кососимметрической, антисимметричной), если $b(\boldsymbol{x}, \boldsymbol{y}) = -b(\boldsymbol{y}, \boldsymbol{x})$ для любых векторов $\boldsymbol{x}, \boldsymbol{y} \in V$.

Предложение

Если характеристика поля K отлична от 2, то билинейная функция b на V кососимметрична $\iff$ $b(\boldsymbol{x}, \boldsymbol{x}) = 0$ для всех $\boldsymbol{x} \in V$.

(Достаточно рассмотреть $b(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y})$.)

Множества симметричных и кососимметричных билинейных функций являются векторными пространствами относительно поточечных операций. Будем обозначать их $\mathscr{B}_{\text{сим}}(V)$ и $\mathscr{B}_{\text{кос}}(V)$.

Предложение

Множества $\mathscr{B}_{сим}(V)$ и $\mathscr{B}_{кос}(V)$ всех симметричных и всех кососимметричных билинейных функций на V являются линейными подпространствами векторного пространства $\mathscr{B}(V)$. Более того, $\mathscr{B}(V) = \mathscr{B}_{сим}(V) \oplus \mathscr{B}_{кос}(V)$.

Доказательство. Симметричное и кососимметричное слагаемые определяются так:

$$b_{сим}(\boldsymbol{x}, \boldsymbol{y}) = \frac{b(\boldsymbol{x},\boldsymbol{y})+b(\boldsymbol{y},\boldsymbol{x})}{2}, \qquad b_{кос}(\boldsymbol{x}, \boldsymbol{y}) = \frac{b(\boldsymbol{x},\boldsymbol{y})-b(\boldsymbol{y},\boldsymbol{x})}{2}. \qquad \square$$

Отображение $\mathscr{B}(V) \to \mathscr{B}(V)$, определённое правилом $b \mapsto b_{сим}$, является линейным оператором. Это оператор проектирования пространства $\mathscr{B}(V)$ на подпространство $\mathscr{B}_{сим}(V)$ параллельно $\mathscr{B}_{кос}(V)$. Он называется симметрированием.

Отображение $\mathscr{B}(V) \to \mathscr{B}(V)$, определённое правилом $b \mapsto b_{кос}$, является линейным оператором. Это оператор проектирования пространства $\mathscr{B}(V)$ на подпространство $\mathscr{B}_{кос}(V)$ параллельно $\mathscr{B}_{сим}(V)$. Он называется альтернированием.

Левая и правая ортогональность относительно (косо)симметричной билинейной функции b совпадают: для любого $U \underset{\text{lin}}{\subseteq} V$ имеем $U^{\perp} = {}^{\perp}U$. Следовательно, левое ядро совпадает с правым и называется просто ядром: $V^{\perp} = {}^{\perp}V = \operatorname{Ker} b$.

Предложение

Если b — (косо)симметричная билинейная функция и U — подпространство пространства V с тем свойством, что $\operatorname{Ker} b \oplus U = V$, то ограничение функции b на U невырождено.

Доказательство. Пусть $\boldsymbol{u} \in \operatorname{Ker} b|_{U\times U}$. Для любого вектора $\boldsymbol{x} \in V$ имеем $\boldsymbol{x} = \boldsymbol{v} + \boldsymbol{w}$, где $\boldsymbol{v} \in \operatorname{Ker} b$ и $\boldsymbol{w} \in U$, $\implies b(\boldsymbol{u}, \boldsymbol{x}) = b(\boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{u}, \boldsymbol{w}) = 0 \implies \boldsymbol{u} \in \operatorname{Ker} V \cap U$. □

Упражнение

Покажите, что для произвольных билинейных функций это, вообще говоря, не так.

Всюду ниже пространство V предполагается n-мерным.

Матрица $B = (b_{ij})$ любой симметричной билинейной функции на V симметрична, т.е. удовлетворяет условию $B^T = B$.

Матрица $B = (b_{ij})$ любой кососимметричной билинейной функции на V кососимметрична, т.е. удовлетворяет условию $B^T = -B$.

Теорема (Лагранжа)

Для симметричной билинейной функции b на конечномерном векторном пространстве V над любым полем характеристики $\neq 2$ существует базис пространства V, в котором матрица b диагональна.

Доказательство. Если $\dim V = 1$, то доказывать нечего. Предположим, что $\dim V = n > 1$ и для пространств меньшей размерности теорема верна.
Если $b \equiv 0$, то доказывать нечего. Пусть $b \not\equiv 0$. Тогда существует $\boldsymbol{e}_1$, для которого $b(\boldsymbol{e}_1, \boldsymbol{e}_1) \neq 0$ (надо рассмотреть $b(\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y})$). Подпространство $U = \langle \boldsymbol{e}_1 \rangle^{\perp}$ имеет размерность $n-1$ (так как это пространство решений уравнения $b(\boldsymbol{e}_1, \boldsymbol{x}) = 0$), и $U \cap \langle \boldsymbol{e}_1 \rangle = \{\boldsymbol{0}\}$. Значит, $V = U \oplus \langle \boldsymbol{e}_1 \rangle$. В U существует базис $\boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$, в котором матрица ограничения $b|_{U \times U}$ диагональна. $\square$

к эрмитовым функциям

Определение

Говорят, что симметричная билинейная функция $b\colon V \times V \to K$ имеет в базисе $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ пространства V канонический (или диагональный) вид, если матрица b в этом базисе диагональна, т.е. для всех векторов $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$ и $\boldsymbol{y} = y_1\boldsymbol{e}_1 + \cdots + y_n\boldsymbol{e}_n$ значение $b(\boldsymbol{x}, \boldsymbol{y})$ вычисляется по формуле вида $\sum_{i=1}^{n} \lambda_i x_i y_i$. При этом базис $\boldsymbol{E}$ называется каноническим базисом для b.

Определение

Базис векторного пространства V, в котором матрица кососимметричной билинейной функции b имеет блочно-диагональный вид

$$\begin{pmatrix} \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} & & & & & & \text{\Large 0} \\ & \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} & & & & & \\ & & \ddots & & & & \\ & & & \begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ \text{\Large 0} & & & & & & 0 \end{pmatrix},$$

называется симплектическим базисом относительно b.

Теорема

Для любой кососимметричной билинейной функции b на конечномерном векторном пространстве V над любым полем характеристики $\neq 2$ существует симплектический базис.

Доказательство. Индукция по $\dim V$. Если $\dim V = 1$, то $b \equiv 0$. Пусть $\dim V = n > 1$ и для меньших n теорема верна. Если $b \equiv 0$, то доказывать нечего. Пусть $\boldsymbol{e}_1$, $\boldsymbol{u}$ — любые два вектора, для которых $b(\boldsymbol{e}_1, \boldsymbol{u}) \neq 0$ (они существуют и не пропорциональны друг другу, так как $b \not\equiv 0$ и $b(\boldsymbol{e}_1, \boldsymbol{e}_1) = 0$). Положим $\boldsymbol{e}_2 = \frac{1}{b(\boldsymbol{e}_1, \boldsymbol{u})} \boldsymbol{u}$. Пусть $U = \langle \{\boldsymbol{e}_1, \boldsymbol{e}_2\} \rangle^{\perp}$. Имеем $\dim U = n - 2$. Действительно, $\dim U \geqslant n - 2$, поскольку U — это пространство решений системы двух линейных уравнений $b(\boldsymbol{e}_1, \boldsymbol{x}) = 0$ и $b(\boldsymbol{e}_2, \boldsymbol{x}) = 0$, и $U \cap \langle \boldsymbol{e}_1, \boldsymbol{e}_2 \rangle = \{\boldsymbol{0}\}$, поскольку для любого вектора $\boldsymbol{v} = \lambda \boldsymbol{e}_1 + \mu \boldsymbol{e} - 2 \in \langle \boldsymbol{e}_1, \boldsymbol{e}_2 \rangle$ имеем $b(\boldsymbol{e}_1, \boldsymbol{v}) \neq \boldsymbol{0}$ и $b(\boldsymbol{e}_2, \boldsymbol{v}) \neq \boldsymbol{0}$. Значит, $V = U \oplus \langle \{\boldsymbol{e}_1, \boldsymbol{e}_2\} \rangle$. По индуктивному предположению в U есть нужный базис $\{\boldsymbol{e}_3, \ldots, \boldsymbol{e}_n\}$ для ограничения b. $\square$

Следствие

На векторном пространстве V над полем характеристики $\neq 2$ существует невырожденная кососимметричная функция $\iff$ размерность V чётна.

Иногда кососимметричной называют билинейную функцию b на V с тем свойством, что $b(\boldsymbol{x}, \boldsymbol{x}) = 0$ для всех $\boldsymbol{x} \in V$. При таком определении доказанная теорема становится верной для любого поля.

Квадратичные формы

Всюду ниже V — n-мерное векторное пространство над произвольным полем K.

Определение

Пусть $b\colon V \times V \to K$ — симметричная билинейная функция. Функция $q\colon V \to K$, определённая правилом

$$q(\boldsymbol{x}) = b(\boldsymbol{x}, \boldsymbol{x}) \qquad \text{для каждого } \boldsymbol{x} \in V,$$

называется квадратичной формой (или квадратичной функцией), ассоциированной с b.

Если $B = (b_{ij})$ — матрица симметричной билинейной функции b в некотором базисе, то

$$q(\boldsymbol{x}) = (x_1, \ldots, x_n)\, B \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

где $x_1, \ldots, x_n$ — координаты вектора $\boldsymbol{x}$ в том же базисе. В координатах:

$$q(\boldsymbol{x}) = \sum_{i,j=1}^{n} b_{ij} x_i x_j.$$

Определение

Матрицей квадратичной формы q, ассоциированной с симметричной билинейной функцией b, в данном базисе называется матрица b в этом базисе. Ранг этой матрицы называется **рангом квадратичной формы** q.

Поляризация

Если характеристика поля K не равна 2, то по данной квадратичной форме q можно однозначно восстановить симметричную билинейную функцию b, с которой она ассоциирована:

$$b(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2}(q(\boldsymbol{x} + \boldsymbol{y}) - q(\boldsymbol{x}) - q(\boldsymbol{y})).$$

Процедура восстановления b по q называется **процедурой поляризации**. Таким образом, если характеристика поля K не равна 2, то имеется взаимно однозначное соответствие между симметричными билинейными функциями и квадратичными формами.

Согласно теореме Лагранжа для любой квадратичной формы $q\colon V \to K$ существует базис, в котором она записывается в виде суммы квадратов с коэффициентами:

$$q(\boldsymbol{x}) = \lambda_1 x_1^2 + \dots + \lambda_n x_n^2.$$

Такое выражение называется каноническим (или диагональным) видом квадратичной формы q.

Здесь и всюду дальше обозначения вида $x_1, \dots, x_n$, $\tilde{x}_1, \dots, \tilde{x}_n$, $y_1, \dots, y_n$ используются для координат вектора, к которому применяется рассматриваемая квадратичная форма, в рассматриваемом базисе (из контекста всегда ясно, каком). Иногда мы пишем $q(x_1, \dots, x_n)$ вместо $q(\boldsymbol{x})$, подразумевая, что каждый вектор может быть отождествлён с набором его координат.

Существует несколько алгоритмов приведения квадратичной формы к каноническому виду. Мы рассмотрим два.

Метод Лагранжа: выделение полных квадратов

Пусть характеристика поля $\neq 2$ и в некотором базисе

$$q(\boldsymbol{x}) = \sum_{i,j=1}^{n} b_{ij}x_ix_j = \sum_{i=1}^{n} b_{ii}x_i^2 + 2\sum_{1\leqslant i<j\leqslant n} b_{ij}x_ix_j.$$

Возможны два случая:

1. $b_{ii} \neq 0$ *хотя бы для одного* $i \leqslant n$

Пусть $b_{11} \neq 0$. Собираем вместе все слагаемые с x_1 и выносим b_{11} за скобки:

$$q(\boldsymbol{x}) = b_{11}\Big(x_1^2 + 2\frac{b_{12}}{b_{11}}x_1x_2 + \dots + 2\frac{b_{1n}}{b_{11}}x_1x_n\Big) + \dots .$$

Выделяем полный квадрат:

$$q(\boldsymbol{x}) = b_{11}\Big(x_1 + \frac{b_{12}}{b_{11}}x_2 + \dots + \frac{b_{1n}}{b_{11}}x_n\Big)^2 - \frac{b_{12}^2}{b_{11}^2}x_2^2 - \dots - \frac{b_{1n}^2}{b_{11}^2}x_n^2 + \dots .$$

Делаем замену переменных:

$$y_1 = x_1 + \frac{b_{12}}{b_{11}}x_2 + \dots + \frac{b_{1n}}{b_{11}}x_n, \quad y_2 = x_2, \ \dots, \ y_n = x_n.$$

Получаем $q(\boldsymbol{x}) = b_{11}y_1^2 + \tilde{q}(y_2, \dots, y_n)$. Дальше по рекурсии.

2. $b_{11} = b_{22} = \dots = b_{nn} = 0$. Ищем $i < j$, для которых $b_{ij} \neq 0$. Пусть $b_{12} \neq 0$. Делаем подстановку:

$$x_1 = y_1 - y_2, \quad x_2 = y_1 + y_2, \quad x_3 = y_3, \dots, x_n = y_n.$$

В выражении $q(y_1, \dots, y_n)$ той же квадратичной формы в новых координатах коэффициент при y_1^2 равен $2b_{12} \neq 0$. Применяем алгоритм 1.

Метод Якоби: ортогонализация

Теорема (формула Якоби)

Пусть $b\colon V \times V \to K$ *— симметричная билинейная функция на* n*-мерном векторном пространстве над любым полем* K, B *— её матрица в некотором базисе* $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ *и* $\Delta_1, \ldots, \Delta_n$ *— угловые миноры матрицы* B. *Положим* $\Delta_0 = 1$.

Если все Δ_i, $i \leqslant n$, *отличны от нуля, то существует базис* $\tilde{\boldsymbol{E}} = \{\tilde{\boldsymbol{e}}_1, \ldots, \tilde{\boldsymbol{e}}_n\}$ *векторного пространства* V, *в котором квадратичная функция* q, *ассоциированная с* b, *имеет вид*

$$q(y_1\tilde{\boldsymbol{e}}_1 + \cdots + y_n\tilde{\boldsymbol{e}}_n) = \frac{\Delta_1}{\Delta_0} y_1^2 + \cdots + \frac{\Delta_n}{\Delta_{n-1}} y_n^2,$$

причём матрица перехода от $\boldsymbol{E}$ *к* $\tilde{\boldsymbol{E}}$ *верхне-унитреугольна* (*верхнетреугольна и все диагональные элементы равны* 1).

▸ к эрмитовым функциям

Доказательство. Индукция по n. Пусть $n > 1$, для меньших n доказано.

Положим $U = \langle\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{n-1}\}\rangle$. По индуктивному предположению для билинейной функции $b|_{U\times U}$ в U есть базис $\{\tilde{\boldsymbol{e}}_1, \ldots, \tilde{\boldsymbol{e}}_{n-1}\}$ с нужными свойствами. Для $i, j < n$ имеем $b(\tilde{\boldsymbol{e}}_i, \tilde{\boldsymbol{e}}_j) = 0$, если $i \neq j$, и $b(\tilde{\boldsymbol{e}}_i, \tilde{\boldsymbol{e}}_i) = \frac{\Delta_i}{\Delta_{i-1}}$.

Положим
$$\tilde{\boldsymbol{e}}_n = \boldsymbol{e}_n - \frac{b(\boldsymbol{e}_n, \tilde{\boldsymbol{e}}_1)}{b(\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_1)}\tilde{\boldsymbol{e}}_1 - \cdots - \frac{b(\boldsymbol{e}_n, \tilde{\boldsymbol{e}}_{n-1})}{b(\tilde{\boldsymbol{e}}_{n-1}, \tilde{\boldsymbol{e}}_{n-1})}\tilde{\boldsymbol{e}}_{n-1}.$$

Покажем, что базис $\tilde{\boldsymbol{E}} = \{\tilde{\boldsymbol{e}}_1, \ldots, \tilde{\boldsymbol{e}}_n\}$ обладает нужными свойствами.

Матрица C перехода от $\boldsymbol{E}$ к $\tilde{\boldsymbol{E}}$ верхне-унитреугольна.

Для $k < n$ имеем
$$b(\tilde{\boldsymbol{e}}_n, \tilde{\boldsymbol{e}}_k) = b\left(\tilde{\boldsymbol{e}}_n - \frac{b(\boldsymbol{e}_n, \tilde{\boldsymbol{e}}_1)}{b(\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_1)}\tilde{\boldsymbol{e}}_1 - \cdots - \frac{b(\tilde{\boldsymbol{e}}_n, \tilde{\boldsymbol{e}}_{n-1})}{b(\tilde{\boldsymbol{e}}_{n-1}, \tilde{\boldsymbol{e}}_{n-1})}\tilde{\boldsymbol{e}}_{n-1},\ \tilde{\boldsymbol{e}}_k\right) =$$
$$= b(\tilde{\boldsymbol{e}}_n, \tilde{\boldsymbol{e}}_k) - \frac{b(\tilde{\boldsymbol{e}}_n, \tilde{\boldsymbol{e}}_1)}{b(\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_1)} b(\tilde{\boldsymbol{e}}_1, \boldsymbol{e}_k) - \cdots - \frac{b(\tilde{\boldsymbol{e}}_n, \tilde{\boldsymbol{e}}_{n-1})}{b(\tilde{\boldsymbol{e}}_{n-1}, \tilde{\boldsymbol{e}}_{n-1})} b(\tilde{\boldsymbol{e}}_{n-1}, \tilde{\boldsymbol{e}}_k) =$$
$$= b(\tilde{\boldsymbol{e}}_n, \tilde{\boldsymbol{e}}_k) - \frac{b(\tilde{\boldsymbol{e}}_n, \tilde{\boldsymbol{e}}_k)}{b(\tilde{\boldsymbol{e}}_k, \tilde{\boldsymbol{e}}_k)} b(\tilde{\boldsymbol{e}}_k, \tilde{\boldsymbol{e}}_k) = 0.$$

$\Longrightarrow$ матрица $\tilde{B} = (\tilde{b}_{ij})$ функции b в базисе $\tilde{\boldsymbol{E}}$ диагональна, диагональные элементы суть $\tilde{b}_{ii} = \frac{\Delta_i}{\Delta_{i-1}}$ для $i < n$ и $\tilde{b}_{nn} = b(\tilde{\boldsymbol{e}}_n, \tilde{\boldsymbol{e}}_n)$. Имеем $\det \tilde{B} = \det C^T \cdot \det B \cdot \det C = \det B = \Delta_n$. Значит, $\tilde{b}_{nn} = \frac{\Delta_n}{\Delta_{n-1}}$.

▸ к Граму–Шмидту

□

Квадратичные формы над $\mathbb{R}$ и над $\mathbb{C}$

Предложение

Для любой квадратичной формы q на n-мерном векторном пространстве над полем комплексных чисел существует базис, в котором

$$q(\boldsymbol{x}) = x_1^2 + \cdots + x_r^2$$

для некоторого $r \leqslant n$. При этом число r является инвариантом квадратичной функции q; более того, оно совпадает с рангом q.

Определение

Выражение $q(\boldsymbol{x}) = x_1^2 + \cdots + x_r^2$ для квадратичной формы q над полем комплексных чисел называется её нормальным видом.

Предложение

Для любой квадратичной формы q на n-мерном векторном пространстве над полем вещественных чисел существует базис, в котором

$$q(\boldsymbol{x}) = x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_{k+l}^2$$

для некоторых k и l, удовлетворяющих условию $r = k + l \leqslant n$. При этом число r является инвариантом квадратичной функции q; более того, оно совпадает с рангом q.

Определение

Выражение $q(\boldsymbol{x}) = x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_{k+l}^2$ для квадратичной формы q над полем вещественных чисел называется её **нормальным видом**. Числа k и l называются, соответственно, её **положительным** и **отрицательным индексами инерции**.

Теорема (закон инерции)

Положительный и отрицательный индексы инерции являются инвариантами квадратичной функции (т.е. не зависят от выбора базиса).

Доказательство. Пусть $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ и $\{\tilde{\boldsymbol{e}}_1, \ldots, \tilde{\boldsymbol{e}}_n\}$ — два базиса векторного пространства V над $\mathbb{R}$, в которых

$$q(\boldsymbol{x}) = x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_{k+l}^2 = \tilde{x}_1^2 + \cdots + \tilde{x}_s^2 - \tilde{x}_{s+1}^2 - \cdots - \tilde{x}_{s+t}^2. \qquad (\star)$$

Предположим, что $k > s$. Положим $U = \langle\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k\}\rangle$ и $W = \langle\{\tilde{\boldsymbol{e}}_{s+1}, \ldots, \tilde{\boldsymbol{e}}_n\}\rangle$. Формула Грассмана $\implies$ $\dim(U \cap W) = k + n - s - \dim(U + W) \geqslant k - s > 0$. Пусть $\boldsymbol{u} \neq \boldsymbol{0}$, $\boldsymbol{u} \in U \cap W$. Имеем

$$\boldsymbol{u} = \lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_k \boldsymbol{e}_k = \mu_{s+1} \tilde{\boldsymbol{e}}_{s+1} + \cdots + \mu_n \tilde{\boldsymbol{e}}_n.$$

Подставляем в $(\star)$: $\lambda_1^2 + \cdots + \lambda_k^2 = -\mu_{s+1}^2 - \cdots - \mu_n^2$. □

▸ к эрмитовым формам

Теорема (теорема Якоби)

Если все угловые миноры Δ_i матрицы Q квадратичной формы $q\colon V \to \mathbb{R}$ в некотором базисе $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ пространства V не равны нулю, то отрицательный индекс инерции формы q равен числу тех $i \leqslant n$, для которых $\Delta_i \cdot \Delta_{i-1} < 0$.

Доказательство. Прямое следствие метода Якоби и закона инерции. $\square$

Упражнение

Докажите, что теорема Якоби остаётся верной, если среди миноров Δ_i есть нулевые, но все миноры, соседние с нулевыми, ненулевые.

Положительно определённые билинейные функции

Определение

Квадратичная форма $q\colon V \to \mathbb{R}$ называется положительно определенной, если $q(\boldsymbol{x}) > 0$ для любого ненулевого вектора $\boldsymbol{x} \in V$. Симметричная билинейная функция $b\colon V \times V \to \mathbb{R}$ называется положительно определенной, если соответствующая ей квадратичная форма положительно определена.

Квадратичная форма $q\colon V \to \mathbb{R}$ называется отрицательно определенной, если $q(\boldsymbol{x}) > 0$ для любого ненулевого вектора $\boldsymbol{x} \in V$. Симметричная билинейная функция $b\colon V \times V \to \mathbb{R}$ называется отрицательно определенной, если соответствующая ей квадратичная форма отрицательно определена.

Аналогично определяются неотрицательно и неположительно определённые квадратичные формы и симметричные билинейные функции.

Замечание

Нормальный вид положительно определённой квадратичной формы в n-мерном пространстве есть $x_1^2 + \dots + x_n^2$.

Теорема (критерий Сильвестра)

Пусть V — n-мерное векторное пространство над $\mathbb{R}$ и $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ — любой базис в нём. Симметричная билинейная функция $b \colon V \times V \to \mathbb{R}$ положительно определена $\iff$ все угловые миноры Δ_i её матрицы B в базисе $\boldsymbol{E}$ положительны.

Доказательство

$\Longrightarrow$: Индукция по n. Пусть $n > 1$, для меньших верно. Положим $U = \langle\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_{n-1}\}\rangle$. Ограничение b на U положительно определено + индуктивное предположение $\Longrightarrow$ $\Delta_i > 0$ для $i < n$. Пусть $\tilde{\boldsymbol{E}}$ — базис, в котором матрица $\tilde{B}$ диагональна, и пусть C — матрица перехода от $\tilde{\boldsymbol{E}}$ к $\boldsymbol{E}$. Положительная определённость + формула изменения матрицы $\Longrightarrow$

$$\Delta_n = \det B = \det C^T \cdot \det \tilde{B} \cdot \det C = (\det C)^2 \cdot \det \tilde{B} > 0.$$

$\Longleftarrow$: по формуле Якоби. $\square$

▸ к эрмитовым функциям

Упражнения

1. Докажите, что симметричная билинейная функция $b\colon V \times V \to \mathbb{R}$ отрицательно определена $\iff$ знаки угловых миноров Δ_i её матрицы B в произвольном базисе $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ строго чередуются, причём $\Delta_1 < 0$.
2. Докажите, что симметричная билинейная функция $b\colon V \times V \to \mathbb{R}$ неотрицательно определена $\iff$ все главные миноры её матрицы B в произвольном базисе $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ неотрицательны.

Евклидовы пространства

Определение

Скалярным произведением на векторном пространстве над $\mathbb{R}$ называется фиксированная на этом пространстве положительно определенная симметричная билинейная функция. Скалярное произведение обозначается $(\cdot\,,\cdot)$.

Евклидовым (векторным) пространством называется конечномерное векторное пространство над $\mathbb{R}$ с фиксированным скалярным произведением.

Примеры (скалярного произведения)

- $|\vec{x}| \cdot |\vec{y}| \cdot \cos \widehat{\vec{x}, \vec{y}}$ в обычном пространстве.
- $(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^T \boldsymbol{y} = x_1 y_1 + \ldots + x_n y_n$ в арифметическом пространстве $\mathbb{R}^n$.
- $(f, g) = \int\limits_0^1 f(t) g(t)\, dt$ в $C([0,1])$.

Всюду ниже V — n-мерное евклидово пространство.

Определение

Длиной вектора $\boldsymbol{x}$ в евклидовом пространстве V называется число
$$|\boldsymbol{x}| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})}.$$
Угол $\widehat{\boldsymbol{x}, \boldsymbol{y}}$ между ненулевыми векторами $\boldsymbol{x}$ и $\boldsymbol{y}$ — это число $\alpha \in [0, \pi]$ такое, что
$$\cos\alpha = \frac{(\boldsymbol{x}, \boldsymbol{y})}{|\boldsymbol{x}||\boldsymbol{y}|}.$$

Угол определён для любых $\boldsymbol{x}, \boldsymbol{y} \neq \boldsymbol{0}$ в силу следующего утверждения:

Неравенство Коши–Буняковского

Для любых векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ евклидова пространства
$$(\boldsymbol{x}, \boldsymbol{y})^2 \leqslant (\boldsymbol{x}, \boldsymbol{x}) \cdot (\boldsymbol{y}, \boldsymbol{y}),$$
причем равенство имеет место тогда и только тогда, когда векторы $\boldsymbol{x}$ и $\boldsymbol{y}$ пропорциональны.

Доказательство. Если $\boldsymbol{y} \neq \boldsymbol{0}$, имеем $(\boldsymbol{x} - t\boldsymbol{y}, \boldsymbol{x} - t\boldsymbol{y}) \geqslant \boldsymbol{0}$ для всех $t \in \mathbb{R}$, т.е.
$t^2(\boldsymbol{y}, \boldsymbol{y}) - 2t(\boldsymbol{x}, \boldsymbol{y}) + (\boldsymbol{x}, \boldsymbol{x}) \geqslant 0 \implies$ дискриминант неположительный.
Дискриминант равен 0 $\iff$ $\boldsymbol{x} = t\boldsymbol{y}$ для некоторого $t \in \mathbb{R}$. $\square$

Неравенство треугольника

Для любых векторов $\boldsymbol{x}$ *и* $\boldsymbol{y}$ *евклидова пространства*

$$|\boldsymbol{x}+\boldsymbol{y}| \leqslant |\boldsymbol{x}| + |\boldsymbol{y}|.$$

Доказательство. Имеем $|\boldsymbol{x}+\boldsymbol{y}|^2 = (\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y}) = (\boldsymbol{x},\boldsymbol{x}) + 2(\boldsymbol{x},\boldsymbol{y}) + (\boldsymbol{y},\boldsymbol{y})$.
Неравенство Коши–Буняковского $\Longrightarrow$ $|\boldsymbol{x}+\boldsymbol{y}|^2 \leqslant (|\boldsymbol{x}| + |\boldsymbol{y}|)^2$. □

Определение

Для векторов $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ евклидова пространства матрица

$$G(\boldsymbol{v}_1, \dots, \boldsymbol{v}_k) = \begin{pmatrix} (\boldsymbol{v}_1,\boldsymbol{v}_1) & (\boldsymbol{v}_1,\boldsymbol{v}_2) & \dots & (\boldsymbol{v}_1,\boldsymbol{v}_k) \\ (\boldsymbol{v}_2,\boldsymbol{v}_1) & (\boldsymbol{v}_2,\boldsymbol{v}_2) & \dots & (\boldsymbol{v}_2,\boldsymbol{v}_k) \\ \vdots & \vdots & \ddots & \vdots \\ (\boldsymbol{v}_k,\boldsymbol{v}_1) & (\boldsymbol{v}_k,\boldsymbol{v}_2) & \dots & (\boldsymbol{v}_k,\boldsymbol{v}_k) \end{pmatrix}$$

называется матрицей Грама.

Теорема

Для произвольных векторов $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ евклидова пространства V

$$\det G(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k) \geqslant 0,$$

причём $\det G(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k) = 0 \iff$ векторы $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ линейно зависимы.

Доказательство. Векторы $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ линейно зависимы $\iff \lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_k \boldsymbol{v}_k = \boldsymbol{0}$ для некоторых $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$, причём не все λ_i равны нулю, откуда для каждого $\boldsymbol{v}_i$

$$0 = (\lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_k \boldsymbol{v}_k, \boldsymbol{v}_i) = \lambda_1(\boldsymbol{v}_1, \boldsymbol{v}_i) + \cdots + \lambda_k(\boldsymbol{v}_k, \boldsymbol{v}_i) = \lambda_1 g_{1i} + \cdots + \lambda_k g_{ki},$$

где g_{ji} — элементы матрицы Грама, т.е. строки матрицы Грама линейно зависимы.

Пусть система $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ линейно независима (тогда $k \leqslant n = \dim V$). Дополним её до базиса $\boldsymbol{E} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{e}_{k+1}, \ldots, \boldsymbol{e}_n\}$. $\det G(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$ — это угловой минор Δ_k матрицы положительно определённой симметричной билинейной функции $(\cdot, \cdot)$. Критерий Сильвестра $\implies \Delta_k > 0$. □

Предложение

Пусть $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ *— произвольный базис евклидова пространства* V, $G = (g_{ij}) = G(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ *и* $\boldsymbol{x} = \sum_{i=1}^n x_i \boldsymbol{e}_i$, $\boldsymbol{y} = \sum_{i=1}^n y_i \boldsymbol{e}_i$ *— любые векторы в* V. *Тогда*

$$(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i,j=1}^{n} x_i y_j g_{ij} = (x_1, \dots, x_n)\, G \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Векторы $\boldsymbol{x}$ и $\boldsymbol{y}$ называются ортогональными (запись $\boldsymbol{x} \perp \boldsymbol{y}$), если они ортогональны относительно $(\cdot, \cdot)$, т.е. если $(\boldsymbol{x}, \boldsymbol{y}) = 0$. Имеем $\boldsymbol{x} \perp \boldsymbol{y} \iff$ либо $|\boldsymbol{x}| = 0$ (т.е. $\boldsymbol{x} = \boldsymbol{0}$), либо $|\boldsymbol{y}| = 0$ (т.е. $\boldsymbol{y} = \boldsymbol{0}$), либо $\cos \widehat{\boldsymbol{x}, \boldsymbol{y}} = 0$.

Теорема Пифагора

Если $\boldsymbol{x} \perp \boldsymbol{y}$, *то* $|\boldsymbol{x} + \boldsymbol{y}|^2 = |\boldsymbol{x}|^2 + |\boldsymbol{y}|^2$.

Система векторов ортогональна, когда все векторы в ней попарно ортогональны.

Предложение

Любая ортогональная система ненулевых векторов линейно независима.

Определение

Система векторов $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ в евклидовом пространстве называется ортонормированной, если она ортогональна и каждый вектор $\boldsymbol{v}_i$ базиса имеет единичную длину, т.е. $(\boldsymbol{v}_i, \boldsymbol{v}_j) = \delta_{ij}$ для $i, j \leqslant k$.

Предложение

Базис $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ евклидова пространства ортонормирован тогда и только тогда, когда выполнено любое из эквивалентных условий:

1. $G(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) = E$;
2. *для любых* $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$ *и* $\boldsymbol{y} = y_1\boldsymbol{e}_1 + \cdots + y_n\boldsymbol{e}_n$ (*или для любых* $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ *и* $\boldsymbol{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$)
$$(\boldsymbol{x}, \boldsymbol{y}) = x_1 y_1 + \cdots + x_n y_n \qquad (\text{или} \quad (\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^T \boldsymbol{y});$$
3. *координаты любого вектора* $\boldsymbol{x}$ *в базисе* $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ *вычисляются по формуле* $x_i = (\boldsymbol{x}, \boldsymbol{e}_i)$ *для* $i \leqslant n$, *т.е.*
$$\boldsymbol{x} = (\boldsymbol{x}, \boldsymbol{e}_1)\boldsymbol{e}_1 + \cdots + (\boldsymbol{x}, \boldsymbol{e}_n)\boldsymbol{e}_n.$$

Процесс ортогонализации Грама–Шмидта

Если $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ — любой базис пространства V, то, применив к нему алгоритм Якоби, мы получим ортогональный базис $\tilde{\boldsymbol{E}} = \{\tilde{\boldsymbol{e}}_1, \dots, \tilde{\boldsymbol{e}}_n\}$:

$$\tilde{\boldsymbol{e}}_1 = \boldsymbol{e}_1, \qquad \tilde{\boldsymbol{e}}_i = \boldsymbol{e}_i - \frac{(\boldsymbol{e}_i, \tilde{\boldsymbol{e}}_1)}{(\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_1)}\tilde{\boldsymbol{e}}_1 - \dots - \frac{(\boldsymbol{e}_i, \tilde{\boldsymbol{e}}_{i-1})}{(\tilde{\boldsymbol{e}}_{i-1}, \tilde{\boldsymbol{e}}_{i-1})}\tilde{\boldsymbol{e}}_{i-1}.$$

При этом матрица перехода от $\boldsymbol{E}$ к $\tilde{\boldsymbol{E}}$ верхне-унитреугольная. Положив $\hat{\boldsymbol{e}}_i = \frac{\tilde{\boldsymbol{e}}_i}{|\tilde{\boldsymbol{e}}_i|}$ для $i \leqslant n$, мы получим ортонормированный базис $\hat{\boldsymbol{E}} = \{\hat{\boldsymbol{e}}_1, \dots, \hat{\boldsymbol{e}}_n\}$. Матрица перехода от $\boldsymbol{E}$ к $\hat{\boldsymbol{E}}$ верхнетреугольная, на диагонали стоя́т $\frac{1}{|\tilde{\boldsymbol{e}}_i|} = \sqrt{\frac{\Delta_{i-1}}{\Delta_i}}$.

Можно нормировать базисные векторы сразу:

$$\hat{\boldsymbol{e}}_1 = \frac{\boldsymbol{e}_1}{|\boldsymbol{e}_1|}, \qquad \hat{\boldsymbol{e}}_i = \frac{\boldsymbol{e}_i - (\boldsymbol{e}_i, \hat{\boldsymbol{e}}_1)\hat{\boldsymbol{e}}_1 - \dots - (\boldsymbol{e}_i, \hat{\boldsymbol{e}}_{i-1})\hat{\boldsymbol{e}}_{i-1}}{|\boldsymbol{e}_i - (\boldsymbol{e}_i, \hat{\boldsymbol{e}}_1)\hat{\boldsymbol{e}}_1 - \dots - (\boldsymbol{e}_i, \hat{\boldsymbol{e}}_{i-1})\hat{\boldsymbol{e}}_{i-1}|}.$$

Теорема

Во всяком евклидовом пространстве существует ортонормированный базис. Более того, ортонормированный базис любого подпространства евклидова пространства можно дополнить до ортонормированного базиса всего пространства.

Упражнения

1. Докажите, что во всяком счётномерном векторном пространстве над $\mathbb{R}$ со скалярным произведением существует ортонормированный базис.
2. Верно ли, что любую ортонормированную систему векторов в таком пространстве можно дополнить до ортонормированного базиса?
3. Пусть V — векторное пространство над $\mathbb{R}$ и $\boldsymbol{E}$ — произвольный базис в нём. Покажите, что на V существует единственное скалярное произведение, относительно которого базис $\boldsymbol{E}$ является ортонормированным.

Ортогональное проектирование

Пусть $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k\}$ — ортонормированный базис подпространства $U \subset V$ и $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ — ортонормированный базис V. Тогда $U^\perp = \langle\{\boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n\}\rangle$.

Предложение

Для каждого $U \underset{\mathrm{lin}}{\subseteq} V \quad V = U \oplus U^\perp$.

Таким образом, любой вектор $\boldsymbol{x} \in V$ единственным образом раскладывается в сумму $\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{w}$, где $\boldsymbol{u} \in U$ и $\boldsymbol{w} \in U^\perp$. Вектор $\boldsymbol{u}$ называется ортогональной проекцией вектора $\boldsymbol{x}$ на U (обозначение: $\mathrm{pr}_U\, \boldsymbol{x}$), а вектор $\boldsymbol{w}$ — ортогональной составляющей вектора $\boldsymbol{x}$ относительно U (обозначение: $\mathrm{ort}_U\, \boldsymbol{x}$). Отображение проектирования пространства V на U параллельно $U^\perp$ называется ортогональным проектированием на U и обозначается pr_U.

В указанном вверху ортонормированном базисе для любого $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n$ имеем

$$\mathrm{pr}_U\, \boldsymbol{x} = (\boldsymbol{x}, \boldsymbol{e}_1)\boldsymbol{e}_1 + \dots + (\boldsymbol{x}, \boldsymbol{e}_k)\boldsymbol{e}_k \quad \text{и} \quad \mathrm{ort}_U\, \boldsymbol{x} = (\boldsymbol{x}, \boldsymbol{e}_{k+1})\boldsymbol{e}_{k+1} + \dots + (\boldsymbol{x}, \boldsymbol{e}_n)\boldsymbol{e}_n.$$

Таким образом, процесс ортонормализации сводится к замене каждого базисного вектора его ортогональной составляющей относительно линейной оболочки уже построенных векторов и нормированию.

Свойства ортогонального дополнения

Пусть U и W — подпространства евклидова пространства V.

- $(U^{\perp})^{\perp} = U$ (из соображений размерности);
- $(U + W)^{\perp} = U^{\perp} \cap W^{\perp}$ (доказывается прямой проверкой);
- $(U \cap W)^{\perp} = U^{\perp} + W^{\perp}$ (из первых двух свойств).

Расстояние в евклидовом пространстве

Определение

Расстояние $\rho(\boldsymbol{x}, \boldsymbol{y})$ между векторами $\boldsymbol{x}$ и $\boldsymbol{y}$ в евклидовом пространстве V определяется формулой

$$\rho(\boldsymbol{x}, \boldsymbol{y}) = |\boldsymbol{x} - \boldsymbol{y}|.$$

Функция $\rho\colon V \times V \to \mathbb{R}$ — метрика (она рефлексивна, симметрична и удовлетворяет неравенству треугольника).

Определение

Расстояние $\rho(\boldsymbol{x}, U)$ между вектором $\boldsymbol{x}$ и подпространством U в евклидовом пространстве V определяется формулой

$$\rho(\boldsymbol{x}, U) = \inf_{\boldsymbol{u} \in U} \rho(\boldsymbol{x}, \boldsymbol{u}).$$

Углом между ненулевым вектором $\boldsymbol{x}$ и ненулевым подпространством U называется наименьший из углов между $\boldsymbol{x}$ и ненулевыми векторами из U.

Теорема

1. *Расстояние от вектора* $\boldsymbol{x}$ *евклидова пространства* V *до подпространства* U *равно* $|\operatorname{ort}_U \boldsymbol{x}|$, *причём единственным ближайшим к* $\boldsymbol{x}$ *вектором подпространства* U *является* $\operatorname{pr}_U \boldsymbol{x}$.
2. *Угол между ненулевым вектором* $\boldsymbol{x}$ *и ненулевым подпространством* U *равен углу между* $\boldsymbol{x}$ *и* $\operatorname{pr}_U \boldsymbol{x}$ (*если* $\operatorname{pr}_U \boldsymbol{x} \neq \mathbf{0}$) *или* $\frac{\pi}{2}$ (*если* $\operatorname{pr}_U \boldsymbol{x} = \mathbf{0}$).

Доказательство. 1. По теореме Пифагора

$$|(\operatorname{pr}_U \boldsymbol{x} - \boldsymbol{u}) + \operatorname{ort}_U \boldsymbol{x}|^2 = |\operatorname{pr}_U \boldsymbol{x} - \boldsymbol{u}|^2 + |\operatorname{ort}_U \boldsymbol{x}|^2 \geqslant |\operatorname{ort}_U \boldsymbol{x}|^2 = \rho^2(\boldsymbol{x}, \operatorname{pr}_U \boldsymbol{x}),$$

так что наименьшее значение $\rho(\boldsymbol{x}, U)$ достигается при $\boldsymbol{u} = \operatorname{pr}_U \boldsymbol{x}$ и равно $|\operatorname{ort}_U \boldsymbol{x}|$.

2. Пусть $\operatorname{pr}_U \boldsymbol{x} \neq \mathbf{0}$ и $\boldsymbol{u} \in U$, $\boldsymbol{u} \neq \mathbf{0}$. Надо доказать, что

$$\cos(\boldsymbol{x}, \boldsymbol{u}) = \frac{(\boldsymbol{x}, \boldsymbol{u})}{|\boldsymbol{x}| \cdot |\boldsymbol{u}|} \leqslant \frac{(\boldsymbol{x}, \operatorname{pr}_U \boldsymbol{x})}{|\boldsymbol{x}| \cdot |\operatorname{pr}_U \boldsymbol{x}|} = \cos(\boldsymbol{x}, \operatorname{pr}_U \boldsymbol{x}),$$

т.е. $(\boldsymbol{x}, \boldsymbol{u})|\operatorname{pr}_U \boldsymbol{x}| \leqslant (\boldsymbol{x}, \operatorname{pr}_U \boldsymbol{x})|\boldsymbol{u}| \Leftrightarrow (\operatorname{pr}_U \boldsymbol{x} + \operatorname{ort}_U \boldsymbol{x}, \boldsymbol{u})|\operatorname{pr}_U \boldsymbol{x}| \leqslant (\operatorname{pr}_U \boldsymbol{x} + \operatorname{ort}_U \boldsymbol{x}, \operatorname{pr}_U \boldsymbol{x})|\boldsymbol{u}|$ $\Leftrightarrow (\operatorname{pr}_U \boldsymbol{x}, \boldsymbol{u})|\operatorname{pr}_U \boldsymbol{x}| \leqslant (\operatorname{pr}_U \boldsymbol{x}, \operatorname{pr}_U \boldsymbol{x})|\boldsymbol{u}| \Leftrightarrow (\operatorname{pr}_U \boldsymbol{x}, \boldsymbol{u}) \leqslant |\operatorname{pr}_U \boldsymbol{x}| \cdot |\boldsymbol{u}|$. Это неравенство Коши–Буняковского. □

Теорема

Пусть U — ненулевое подпространство евклидова пространства V и $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k\}$ — любой базис в U. Тогда для произвольного вектора $\boldsymbol{x} \in V$

$$\rho(\boldsymbol{x}, U)^2 = \frac{\det G(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k, \boldsymbol{x})}{\det G(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k)}.$$

Доказательство. Пусть $\boldsymbol{x} \notin U$. Тогда $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k, \boldsymbol{x}\}$ — базис подпространства $U \oplus \langle \boldsymbol{x} \rangle$. Матрица $G(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k, \boldsymbol{x})$ — матрица симметричной билинейной функции в базисе $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k, \boldsymbol{x}\}$, и все её угловые миноры отличны от нуля. Применяя ортогонализацию Якоби, получаем ортогональный базис $\{\tilde{\boldsymbol{e}}_1, \ldots, \tilde{\boldsymbol{e}}_{k+1}\}$, причём $\tilde{\boldsymbol{e}}_{k+1} = \mathrm{ort}_U \, \boldsymbol{x}$ и $|\tilde{\boldsymbol{e}}_{k+1}|^2 = \frac{\Delta_{k+1}}{\Delta_k}$. □

Определение

Для векторов $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ в n-мерном евклидовом пространстве множество

$$P(\boldsymbol{a}_1, \dots, \boldsymbol{a}_k) = \left\{ \sum_{i=1}^{k} x_i \boldsymbol{a}_i : 0 \leqslant x_i \leqslant 1 \right\}$$

называется k-мерным параллелепипедом, натянутым на векторы $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$. Основанием этого k-мерного параллелепипеда называется $(k-1)$-мерный параллелепипед $P(\boldsymbol{a}_1, \dots, \boldsymbol{a}_{k-1})$, а высотой — длина вектора $\operatorname{ort}_{\langle\{\boldsymbol{a}_1, \dots, \boldsymbol{a}_{k-1}\}\rangle} \boldsymbol{a}_k$. Объём $\operatorname{vol} P(\boldsymbol{a}_1, \dots, \boldsymbol{a}_k)$ k-мерного параллелепипеда определяется рекурсивно: для $k = 1$ это длина вектора $\boldsymbol{a}_1$, для $k > 1$ — произведение объёма основания на высоту.

Теорема

$\operatorname{vol} P(\boldsymbol{a}_1, \dots, \boldsymbol{a}_k)^2 = \det G(\boldsymbol{a}_1, \dots, \boldsymbol{a}_k)$.

Доказательство. Индукция по k + предыдущая теорема. □

Ортогональные матрицы

Пусть $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ и $\boldsymbol{E}' = \{\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n\}$ — два ортонормированных базиса в евклидовом пространстве и $T = (t_{ij})$ — матрица перехода от $\boldsymbol{E}$ к $\boldsymbol{E}'$, т.е. $\boldsymbol{e}'_i = \sum_{j \leqslant n} t_{i,j} \boldsymbol{e}_j$. Вычисляя скалярное произведение $\boldsymbol{e}'_i$ и $\boldsymbol{e}'_j$ в базисах $\boldsymbol{E}$ и $\boldsymbol{E}'$, получаем

$$(\boldsymbol{e}'_i, \boldsymbol{e}'_j) = \sum_{k=1}^{n} t_{ki} t_{kj} \qquad \text{и} \qquad (\boldsymbol{e}'_i, \boldsymbol{e}'_j) = \delta_{ij}.$$

Значит,

- столбцы T образуют ортонормированный базис в пространстве $\mathbb{R}^n$ n-ок вещественных чисел относительно стандартного скалярного произведения (сумма произведений соответственных компонент) и
- $T^T T = E$, так что $T^{-1} = T^T$ и строки тоже образуют ортонормированный базис в $\mathbb{R}^n$.

Определение

Матрица A, удовлетворяющая условию $A^T A = E$, называется *ортогональной*.

Свойства ортогональных матриц

- Матрица ортогональна $\iff$ это матрица перехода между ортонормированными базисами.
- Если даны два базиса в евклидовом пространстве, один из них ортонормирован и матрица перехода ортогональна, то и другой базис ортонормирован.
- A ортогональна $\implies$ A^{-1} ортогональна.
- A и B ортогональны $\implies$ AB ортогональна.
- Множество O_n всех ортогональных матриц порядка n образует подгруппу в полной линейной группе $\mathrm{GL}_n(\mathbb{R})$ всех невырожденных квадратных матриц порядка n над полем $\mathbb{R}$. Это ортогональная группа O_n.

Теорема (о QR-разложении)

Всякая невырожденная матрица A может быть представлена в виде произведения QR, где Q — ортогональная, а R — верхнетреугольная матрица, причем диагональные элементы матрицы R положительны.

Доказательство. Рассмотрим столбцы матрицы A как столбцы координат векторов $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ в некотором ортонормированном базисе $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$. Так как матрица A невырождена, векторы $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ составляют базис $\boldsymbol{A}$. При этом A является матрицей перехода от $\boldsymbol{E}$ к $\boldsymbol{A}$. Пусть $\tilde{\boldsymbol{A}}$ — ортонормированный базис, полученный из базиса $\boldsymbol{A}$ с помощью процесса ортонормализации Грама–Шмидта. Матрица перехода от $\boldsymbol{A}$ к $\tilde{\boldsymbol{A}}$, а значит, и матрица R перехода от $\tilde{\boldsymbol{A}}$ к $\boldsymbol{A}$ — верхнетреугольная матрица с положительными элементами на диагонали. Поскольку базисы $\tilde{\boldsymbol{A}}$ и $\boldsymbol{E}$ ортонормированы, матрица Q перехода от $\boldsymbol{E}$ к $\tilde{\boldsymbol{A}}$ ортогональна. $\square$

Геометрический смысл определителя

Пусть $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ — произвольные векторы n-мерного евклидова пространства и A — квадратная матрица, столбцами которой являются столбцы координат векторов $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ в каком-нибудь ортонормированном базисе $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$. Тогда

$$\operatorname{vol} P(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) = |\det A|.$$

Действительно, для линейно независимых $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ имеем

$$G(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) = A^T E A = A^T A,$$

откуда

$$\det G(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) = (\det A)^2.$$

Знак $\det A$ можно принять за определение ориентации линейно независимой системы $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ по отношению к базису $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$.

Изоморфизм евклидовых пространств

Определение

Евклидовы пространства V и U называются изоморфными, если существует отображение $\Phi\colon V \to U$, которое является изоморфизмом векторных пространств и удовлетворяет условию
$$(\Phi(\boldsymbol{x}), \Phi(\boldsymbol{y})) = (\boldsymbol{x}, \boldsymbol{y})$$
для любых $\boldsymbol{x}, \boldsymbol{y} \in V$. Само отображение Φ называется при этом изоморфизмом евклидовых пространств V и U.

Замечание

Линейный оператор между евклидовыми пространствами — изоморфизм $\iff$ его матрица ортогональна.

Теорема

Любые два евклидовых пространства одинаковой размерности изоморфны.

Доказательство. Возьмём ортонормированные базисы $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ и $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ в n-мерных евклидовых пространствах U и V, положим $\Phi(\boldsymbol{v}_i) = \boldsymbol{u}_i$ и продолжим по линейности. $\square$

Линейные функции на евклидовом пространстве

Скалярное произведение невырождено $\Longrightarrow$ его корреляции $(\cdot\,,\cdot)_{\text{лев}} = (\cdot\,,\cdot)_{\text{пр}}$ — изоморфизмы. Скалярное произведение симметрично $\Longrightarrow$ они совпадают. Будем обозначать их $\mathscr{C}$.

Для каждого линейного функционала $f : V \to \mathbb{R}$ на евклидовом пространстве V найдётся вектор $\boldsymbol{v}_f \in V$ такой, что $f(\boldsymbol{x}) = (\boldsymbol{v}_f, \boldsymbol{x})$ для всех $\boldsymbol{x} \in V$.

Соответствие $\mathscr{C} : \boldsymbol{v} \mapsto f_{\boldsymbol{v}}$ — изоморфизм $V \to V^*$, и если $\boldsymbol{E}$ — любой ортонормированный базис в V и $\boldsymbol{\mathcal{E}}$ — взаимный с ним базис в V^*, то относительно базисов $\boldsymbol{E}$, $\boldsymbol{\mathcal{E}}$ этот изоморфизм имеет матрицу E.

Линейные операторы в евклидовых пространствах

Как и выше, через V везде обозначаем n-мерное евклидово пространство. Каждому линейному оператору $\mathcal{A}\colon V \to V$ соответствует билинейная функция $b_{\mathcal{A}}\colon V \times V \to \mathbb{R}$, определённая правилом $b_{\mathcal{A}}(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, \mathcal{A}\boldsymbol{y})$. Обозначим соответствие $\mathcal{A} \mapsto b_{\mathcal{A}}$ через β.

Теорема

1. *Матрица любого линейного оператора $\mathcal{A}\colon V \to V$ в любом ортонормированном базисе совпадает с матрицей билинейной функции $\beta(\mathcal{A})$ в том же базисе.*
2. *Отображение $\beta\colon \mathscr{L}(V) \to \mathscr{B}(V)$ — изоморфизм.*

Доказательство. 1. Матрица A оператора $\mathcal{A}$ в любом базисе характеризуется тем, что если вектор $\boldsymbol{y} \in V$ имеет в этом базисе координаты $y_1, \ldots, y_n$, а его образ $\mathcal{A}\boldsymbol{y}$ — координаты $x_1, \ldots, x_n$, то $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. Матрица $B_{\mathcal{A}}$ билинейной функции $\beta(\mathcal{A})$ характеризуется тем, что $b_{\mathcal{A}}(\boldsymbol{x}, \boldsymbol{y}) = (x_1, \ldots, x_n) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. В ортонормированном базисе скалярное произведение есть сумма произведений соответственных координат. $\square$

Сопряжённые операторы

Изоморфизм $\mathscr{C}: V \to V^*$ (корреляция) не зависит от базиса и позволяет отождествлять операторы $\mathcal{A}: V^* \to V^*$ с операторами $\mathscr{C}^{-1} \circ \mathcal{A} \circ \mathscr{C}: V \to V$.

Определение

Оператор $\mathcal{A}^*: V \to V$ в евклидовом пространстве V называется сопряжённым к линейному оператору $\mathcal{A}: V \to V$, если для любых $\boldsymbol{x}, \boldsymbol{y} \in V$

$$(\mathcal{A}\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, \mathcal{A}^*\boldsymbol{y}).$$

Такое определение (и обозначение) законно, так как сопряжённый оператор в новом («евклидовом») смысле — это $\mathscr{C}^{-1} \circ \mathcal{A}_{\text{сопр}} \circ \mathscr{C}$, где $\mathcal{A}_{\text{сопр}}$ — сопряжённый оператор в общем смысле.

Действительно, для $\boldsymbol{y} \in V$ $\mathscr{C}(\boldsymbol{y})$ — функционал $f_{\boldsymbol{y}}$, который каждому $\boldsymbol{x} \in V$ ставит в соответствие $(\boldsymbol{x}, \boldsymbol{y})$, $\mathcal{A}_{\text{сопр}}(f_{\boldsymbol{y}}) = f_{\boldsymbol{y}} \circ \mathcal{A}$ — функционал f, который каждому $\boldsymbol{x} \in V$ ставит в соответствие $(\mathcal{A}\boldsymbol{x}, \boldsymbol{y})$, и $\mathscr{C}^{-1}(f)$ — вектор $\tilde{\boldsymbol{y}}$ с тем свойством, что $(\boldsymbol{x}, \tilde{\boldsymbol{y}}) = f(\boldsymbol{x}) = (\mathcal{A}\boldsymbol{x}, \boldsymbol{y})$ для всех $\boldsymbol{x} \in V$.

Теорема

Для каждого линейного оператора $\mathcal{A}\colon V \to V$ сопряжённый оператор $\mathcal{A}^*$ существует и единствен, причём в любом ортонормированном базисе его матрица равна транспонированной матрице самогó оператора $\mathcal{A}$.

Доказательство. Существование уже доказали. Утверждение о матрицах, а вместе с ним и единственность, вытекает из того, что в ортонормированном базисе для операторов $\mathcal{A}$ и $\mathcal{B}$ с матрицами A и B и векторов $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ и $\boldsymbol{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ имеем

$$(A\boldsymbol{x}, \boldsymbol{y}) = \left(A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right)^T \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (x_1, \dots, x_n)\, A^T \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \qquad (\boldsymbol{x}, B\boldsymbol{y}) = (x_1, \dots, x_n)\, B \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}. \quad \square$$

Свойства операции сопряжения

- $(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*$,
- $(\mathcal{A}^*)^* = \mathcal{A}$,
- $(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*$,
- $(\lambda\mathcal{A})^* = \lambda\mathcal{A}^*$,
- $\mathcal{E}^* = \mathcal{E}$.

Самосопряжённые операторы

Определение

Линейный оператор $\mathcal{A}: V \to V$ в евклидовом пространстве называется самосопряженным (или симметричным, или симметрическим), если $\mathcal{A} = \mathcal{A}^*$, т.е. для любых $\boldsymbol{x}, \boldsymbol{y} \in V$
$$(\mathcal{A}\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, \mathcal{A}\boldsymbol{y}).$$

Теорема

Линейный оператор $\mathcal{A}$ в евклидовом пространстве V самосопряжен $\iff$ его матрица в некотором ортонормированном базисе симметрична $\iff$ его матрица в произвольном ортонормированном базисе симметрична.

Доказательство. Из общей теоремы о матрицах сопряжённых операторов. □

Предложение

Пусть $\mathcal{A}: V \to V$ — самосопряженный линейный оператор в евклидовом пространстве V и U — подпространство, инвариантное относительно $\mathcal{A}$. Тогда ортогональное дополнение $U^\perp$ тоже инвариантно относительно $\mathcal{A}$.

Теорема

Для любого самосопряжённого линейного оператора $\mathcal{A}$ в n-мерном евклидовом пространстве V найдётся ортонормированный базис, в котором матрица этого оператора диагональна.

Доказательство. Индукция по $n = \dim V$. Для $n = 1$ доказывать нечего. Если $n > 1$, то в V есть одно- или двумерное инвариантное подпространство U. Пусть $\dim U = 2$. Ограничение $\mathcal{A}$ на U — самосопряжённый оператор, значит, его матрица симметрична в любом ортонормированном базисе подпространства U. Характеристический многочлен симметричной матрицы 2×2 имеет корень (прямое вычисление) $\implies$ у $\mathcal{A}|_U$ есть собственный вектор $\boldsymbol{u}$. Он собственный и для всего оператора $\mathcal{A} \implies \langle\{\boldsymbol{u}\}\rangle$ — одномерное инвариантное подпространство.

Итак, всегда найдётся одномерное инвариантное подпространство U. Пусть $U = \langle\{\boldsymbol{u}\}\rangle$ для некоторого ненулевого $\boldsymbol{u} \in V$. По индуктивному предположению в $U^\perp$ есть ортонормированный базис $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{n-1}\}$, в котором матрица ограничения $\mathcal{A}|_U$ диагональна. Система $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{n_1}, \frac{\boldsymbol{u}}{|\boldsymbol{u}|}\}$ — ортонормированный базис пространства V, и в нём матрица оператора $\mathcal{A}$ диагональна, так как $U^\perp$ и U инвариантны. $\square$

Следствие

Если $\mathcal{A}$ *— самосопряженный линейный оператор евклидова пространства* V*, то*

1. *характеристический многочлен* $\chi_{\mathcal{A}}(\lambda)$ *раскладывается на линейные множители (над* $\mathbb{R}$*);*
2. *размерность каждого собственного подпространства* V_λ *равна кратности корня* λ *характеристического многочлена;*
3. *собственные векторы, отвечающие различным собственным значениям, ортогональны друг другу.*

Следствие

Для оператора $\mathcal{A}$ *следующие условия равносильны:*

1. *оператор* $\mathcal{A}$ *самосопряжён;*
2. *существует ортонормированный базис, в котором матрица* $\mathcal{A}$ *диагональна;*
3. *если* A *— матрица* $\mathcal{A}$ *в любом ортонормированном базисе, то существуют ортогональная матрица* T *и диагональная матрица* D*, для которых* $A = T^{-1}DT$.

Упражнение

Докажите, что оператор $\mathcal{A}$ в евклидовом пространстве удовлетворяет условию $\mathcal{A}^* = -\mathcal{A}$ тогда и только тогда, когда в некотором ортонормированном базисе его матрица имеет блочно-диагональный вид

$$\begin{pmatrix} \begin{matrix} 0 & a_1 \\ -a_1 & 0 \end{matrix} & & & & & & \text{\Huge 0} \\ & \begin{matrix} 0 & a_2 \\ -a_2 & 0 \end{matrix} & & & & & \\ & & \ddots & & & & \\ & & & \begin{matrix} 0 & a_k \\ -a_k & 0 \end{matrix} & & & \\ & & & & 0 & & \\ \text{\Huge 0} & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$$

для некоторых $k \leqslant n$ и $a_1, \ldots, a_k \in \mathbb{R}$.

Теорема

1. *Сумма самосопряженных линейных операторов и произведение самосопряженного оператора на вещественное число являются самосопряженными операторами.*
2. *Композиция двух самосопряженных операторов является самосопряженным оператором тогда и только тогда, когда эти операторы перестановочны.*

Приведение квадратичной функции к главным осям

Напомним, что между линейными операторами $\mathcal{A}\colon V \to V$ и билинейными функциями $b\colon V \times V \to \mathbb{R}$ имеется изоморфизм β, определённый правилом $\beta(\mathcal{A})(\boldsymbol{x}, \boldsymbol{y}) = (x, \mathcal{A}\boldsymbol{y})$, причём матрица $\mathcal{A}$ в любом ортонормированном базисе совпадает с матрицей $\beta(\mathcal{A})$ в том же базисе. Ясно, что оператор $\mathcal{A}$ самосопряжён тогда и только тогда, когда билинейная функция $\beta(\mathcal{A})$ симметрична.

Теорема

Пусть V — n-мерное евклидово пространство и $b\colon V \times V \to \mathbb{R}$ — симметричная билинейная функция. Тогда в V существует ортонормированный базис $\boldsymbol{E}$, в котором квадратичная функция $q\colon V \to \mathbb{R}$, соответствующая b, записывается в виде суммы квадратов с коэффициентами:

$$q(\boldsymbol{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2,$$

где $\lambda_i \in \mathbb{R}$ и x_i — координаты вектора $\boldsymbol{x}$ в базисе $\boldsymbol{E}$.

Числа $\lambda_1, \ldots, \lambda_n$ — это собственные значения самосопряжённого оператора $\mathcal{A}$, для которого $b = \beta(\mathcal{A})$. Нахождение ортонормированного базиса $\boldsymbol{E}$ из теоремы называется приведением квадратичной формы q к главным осям.

Приведение пары квадратичных форм к сумме квадратов

Теорема

Пусть V — произвольное n-мерное векторное пространство над $\mathbb{R}$, и пусть $q\colon V \to \mathbb{R}$ и $r\colon V \to \mathbb{R}$ — две квадратичные формы на V, причем r положительно определена. Тогда существует базис, в котором обе эти формы записываются в виде суммы квадратов с коэффициентами.

Доказательство. Пусть $b\colon V \to V$ — симметричная билинейная функция, соответствующая форме r. Рассмотрим евклидово пространство V со скалярным произведением $(\cdot,\cdot) = b(\cdot,\cdot)$. В V есть ортонормированный базис $\boldsymbol{E}$, в котором

$$q(\boldsymbol{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2.$$

В этом (и любом другом) ортонормированном базисе матрица скалярного произведения b равна E, т.е.

$$r(\boldsymbol{x}) = x_1^2 + x_2^2 + \cdots + x_n^2. \qquad \square$$

Матрицы Q и R форм q и r в базисе $\boldsymbol{E}$ имеют вид $Q = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ и $R = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$, поэтому

$$\det(Q - \lambda R) = (\lambda_1 - \lambda) \cdot (\lambda_2 - \lambda) \cdot \ldots \cdot (\lambda_n - \lambda).$$

В любом другом базисе матрицы имеют вид $Q' = C^T QC$ и $R' = C^T EC$ для (невырожденной) матрицы перехода C, и

$$\det(Q' - \lambda R') = \det C^2 \cdot (\lambda_1 - \lambda) \cdot (\lambda_2 - \lambda) \cdot \ldots \cdot (\lambda_n - \lambda).$$

Таким образом, $\lambda_1, \ldots, \lambda_n$ — корни многочлена $\det(Q' - \lambda R')$, где Q' и R' — матрицы квадратичных форм q и r в любом базисе.

Требование положительной определённости одной из форм существенно, но не необходимо.

Ортогональные операторы

Определение

Линейный оператор $\mathcal{A}\colon V \to V$ в евклидовом пространстве V называется ортогональным, если он сохраняет скалярное произведение, т.е.

$$(\mathcal{A}\boldsymbol{x}, \mathcal{A}\boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{y}) \qquad \text{для любых } \boldsymbol{x}, \boldsymbol{y} \in V.$$

Замечание

$(\mathcal{A}\boldsymbol{x}, \mathcal{A}\boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{y})$ для любых $\boldsymbol{x}, \boldsymbol{y} \in V \iff \mathcal{A}^* \circ \mathcal{A} = \mathcal{E} \iff \mathcal{A}^* = \mathcal{A}^{-1}$.

Действительно, тогда $(\mathcal{A}\boldsymbol{x}, \mathcal{A}\boldsymbol{y}) = (\boldsymbol{x}, \mathcal{A}^* \circ \mathcal{A}\boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{y})$ для любых $\boldsymbol{x}, \boldsymbol{y} \in V \iff$ $(\boldsymbol{x}, \mathcal{A}^* \circ \mathcal{A}\boldsymbol{y} - \boldsymbol{y}) = 0$ для любых $\boldsymbol{x}, \boldsymbol{y} \in V$.

Теорема

Линейный оператор $\mathcal{A}\colon V \to V$ в евклидовом пространстве V является ортогональным $\iff$ его матрица в некотором ортонормированном базисе ортогональна $\iff$ его матрица в произвольном ортонормированном базисе ортогональна.

Следствие

Линейный оператор в евклидовом пространстве ортогонален $\iff$ он переводит некоторый ортонормированный базис в ортонормированный базис $\iff$ он переводит любой ортонормированный базис в ортонормированный базис.

Теорема

Линейный оператор ортогонален тогда и только тогда, когда он сохраняет длины векторов. В частности, ортогональный оператор сохраняет углы между векторами.

Доказательство. $(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2}\big((\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y}) - (\boldsymbol{x}, \boldsymbol{x}) - (\boldsymbol{y}, \boldsymbol{y})\big)$. $\square$

Следствие

Собственные значения ортогонального линейного оператора равны ± 1.

Предложение

Пусть $\mathcal{A}\colon V \to V$ — ортогональный линейный оператор в евклидовом пространстве V и U — подпространство, инвариантное относительно $\mathcal{A}$. Тогда ортогональное дополнение $U^\perp$ тоже инвариантно относительно $\mathcal{A}$.

Теорема

Оператор в евклидовом пространстве ортогонален тогда и только тогда, когда в некотором ортонормированном базисе его матрица имеет блочно-диагональный вид

$$\begin{pmatrix} \cos\alpha_1 & -\sin\alpha_1 & & & & & & & & & & \\ \sin\alpha_1 & \cos\alpha_1 & & & & & & & & & & \\ & & \cos\alpha_2 & -\sin\alpha_2 & & & & & & & & \\ & & \sin\alpha_2 & \cos\alpha_2 & & & & & & & & \\ & & & & \ddots & & & & & & & \\ & & & & & \cos\alpha_k & -\sin\alpha_k & & & & & \\ & & & & & \sin\alpha_k & \cos\alpha_k & & & & & \\ & & & & & & & -1 & & & & \\ & & & & & & & & \ddots & & & \\ & & & & & & & & & -1 & & \\ & & & & & & & & & & 1 & \\ & & & & & & & & & & & \ddots \\ & & & & & & & & & & & & 1 \end{pmatrix}$$

для некоторых $k \leqslant n$ и $\alpha_1, \ldots, \alpha_k \in (-\pi, \pi) \setminus \{0\}$. Блоки (-1) и/или (1) могут отсутствовать.

Если заменить условие $\alpha_1, \ldots, \alpha_k \in (-\pi, \pi) \setminus \{0\}$ на $\alpha_1, \ldots, \alpha_k \in [-\pi, \pi]$, то можно утверждать, что в выписанной матрице имеется не больше одного блока (-1) и не больше одного блока (1).

Доказательство. Пусть $\mathcal{A}$ — ортогональный оператор в n-мерном евклидовом пространстве V. Индукция по n. Для $n = 1$ утверждение верно. Пусть $n = 2$ и $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ — ортонормированный базис. Положим $\alpha = \widehat{\boldsymbol{e}_1, \mathcal{A}\boldsymbol{e}_1}$. Для вектора $\mathcal{A}\boldsymbol{e}_2$ возможны два случая: либо он совпадает с вектором, который получается из $\boldsymbol{e}_2$ поворотом на угол α, либо он противоположен этому вектору. В первом случае оператор $\mathcal{A}$ — это поворот на угол α, во втором случае $\mathcal{A}$ — это отражение относительно биссектрисы ℓ между $\boldsymbol{e}_1$ и $\mathcal{A}\boldsymbol{e}_1$, и его матрица в ортонормированном базисе $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2\}$, где $\boldsymbol{e}'_2$ параллелен биссектрисе ℓ, имеет вид $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

Предположим, что $n > 2$ и для меньших n всё доказано. Пусть U — одно- или двумерное инвариантное подпространство пространства V для $\mathcal{A}$. В U есть подходящий базис, в $U^{\perp}$ тоже (по индуктивному предположению). Нужный базис в V получается расстановкой векторов из этих базисов в подходящем порядке. $\square$

Такое представление матрицы ортогонального оператора можно трактовать как разложение оператора на плоские вращения и отражения, так как блоки матрицы соответствуют сужениям оператора на инвариантные подпространства, которые можно осуществлять последовательно в любом порядке.

Для трехмерного евклидова пространства доказанная теорема означает, что матрица любого ортогонального оператора $\mathcal{A}$ в подходящем ортонормированном базисе имеет один из двух видов

$$\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

В первом случае оператор $\mathcal{A}$ представляет собой поворот на угол α вокруг третьей оси базиса, во втором — зеркальный поворот, т.е. поворот, совмещенный с отражением относительно плоскости, ортогональной оси поворота. Зеркальный поворот не может быть результатом непрерывного движения, так как он изменяет ориентацию пространства. Следовательно, *сколь угодно сложное реальное движение твёрдого тела с закрепленной точкой сводится к повороту вокруг подходящей оси на подходящий угол* (это теорема Эйлера).

Множество невырожденных линейных операторов n-мерного векторного пространства V образует группу, называемую полной линейной группой пространства V и обозначаемую $\mathrm{GL}(V)$. В случае n-мерного пространства V над $\mathbb{R}$ $\mathrm{GL}(V)$ изоморфна полной линейной группе $\mathrm{GL}_n(\mathbb{R})$ невырожденных матриц порядка n над $\mathbb{R}$.

Теорема

Ортогональные операторы в евклидовом пространстве V образуют подгруппу группы $\mathrm{GL}(n)$.

Доказательство. Из критерия ортогональности $\mathcal{A}^* \circ \mathcal{A} = \mathcal{A} \circ \mathcal{A}^* = \mathcal{E}$.

Группа ортогональных операторов на n-мерном евклидовом пространстве изоморфна ортогональной группе O_n ортогональных матриц порядка n.

Полярное разложение над $\mathbb{R}$

Определение

Самосопряженный линейный оператор $\mathcal{A}$ в евклидовом пространстве V называется положительно (неотрицательно) определённым, если соответствующая ему симметрическая билинейная функция $\beta(\mathcal{A})$ положительно (неотрицательно) определена, т.е. $(\boldsymbol{x}, \mathcal{A}\boldsymbol{x}) > 0$ ($(\boldsymbol{x}, \mathcal{A}\boldsymbol{x}) \geqslant 0$) для любого ненулевого вектора $\boldsymbol{x} \in V$.

Лемма

Самосопряженный оператор в евклидовом пространстве положительно (неотрицательно) определён тогда и только тогда, когда все его собственные значения положительны (неотрицательны).

Доказательство. Необходимость очевидна, достаточность вытекает из теоремы о приведении квадратичной функции к главным осям. $\square$

Лемма

Для всякого положительно (неотрицательно) определённого самосопряжённого оператора $\mathcal{A}$ в евклидовом пространстве существует единственный положительно (неотрицательно) определённый самосопряжённый оператор $\mathcal{B}$, для которого $\mathcal{A} = \mathcal{B}^2$ — квадратный корень из $\mathcal{A}$.

Доказательство. С у щ е с т в о в а н и е: надо рассмотреть ортонормированный базис $\boldsymbol{E}$ из собственных векторов (в нём матрица A оператора $\mathcal{A}$ диагональна).
Е д и н с т в е н н о с т ь: запишем матрицу B любого самосопряжённого оператора $\mathcal{B}$, для которого $\mathcal{B}^2 = \mathcal{A}$, в базисе $\boldsymbol{E}$. Имеем $AB = B^3 = BA$. Значит, собственные подпространства V_λ оператора $\mathcal{A}$ инвариантны относительно $\mathcal{B}$ (если $\boldsymbol{x} \in V_\lambda$, т.е. $A\boldsymbol{x} = \lambda\boldsymbol{x}$, то $A(B\boldsymbol{x}) = BA\boldsymbol{x} = \lambda B\boldsymbol{x}$, так что $B\boldsymbol{x} \in V_\lambda$) и достаточно рассмотреть ограничения на эти подпространства. Задача свелась к единственности симметричной неотрицательно определённой матрицы, квадрат которой равен λE, где $\lambda \geqslant 0$. Если $\lambda = 0$, такая матрица нулевая. Пусть $\lambda > 0$, и пусть C — симметричная матрица, для которой $C^2 = \lambda E$. Существует ортогональная матрица O, для которой O^TCO диагональна (и $O^T = O^{-1}$). Имеем $(O^{-1}CO)^2 = O^{-1}C^2O = O^{-1}\lambda EO = \lambda E$. Значит, C — скалярная матрица $\sqrt{\lambda}E$. □

Лемма

Пусть $\mathcal{A}$ *и* $\mathcal{B}$ *— линейные операторы в евклидовом пространстве и* $(\mathcal{A}\boldsymbol{x}, \mathcal{A}\boldsymbol{y}) = (\mathcal{B}\boldsymbol{x}, \mathcal{B}\boldsymbol{y})$ *для всех* $\boldsymbol{x}, \boldsymbol{y} \in V$. *Тогда существует ортогональный оператор* $\mathcal{O}$, *для которого* $\mathcal{A} = \mathcal{O} \circ \mathcal{B}$.

Доказательство. Имеем $\operatorname{Ker} \mathcal{A} = \operatorname{Ker} \mathcal{B}$. Значит, $\dim \operatorname{Im} \mathcal{A} = \dim \operatorname{Im} \mathcal{B} = r$. Пусть $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r$ — ортонормированный базис в $\operatorname{Im} \mathcal{A}$ и $\boldsymbol{v}_i$ таковы, что $\mathcal{A}\boldsymbol{v}_i = \boldsymbol{e}_i$ для $i \leqslant r$ ($\boldsymbol{v}_i$ линейно независимы). Положим $\tilde{\boldsymbol{e}}_i = \mathcal{B}\boldsymbol{v}_i$. Имеем $(\tilde{\boldsymbol{e}}_i, \tilde{\boldsymbol{e}}_j) = \delta_{ij}$ для $i, j \leqslant r$. Дополним $\{\tilde{\boldsymbol{e}}_1, \ldots, \tilde{\boldsymbol{e}}_r\}$ до ортонормированного базиса $\tilde{\boldsymbol{E}} = \{\tilde{\boldsymbol{e}}_1, \ldots, \tilde{\boldsymbol{e}}_n\}$ в V, а $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r\}$ — до ортонормированного базиса $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ в V. Оператор $\mathcal{O}$, который переводит $\tilde{\boldsymbol{E}}$ в $\boldsymbol{E}$, ортогонален. Дополним $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\}$ до базиса векторами из $\operatorname{Ker} \mathcal{A} = \operatorname{Ker} \mathcal{B}$. Операторы $\mathcal{A}$ и $\mathcal{O} \circ \mathcal{B}$ совпадают на базисе $\Longrightarrow$ они равны. □

Теорема (о полярном разложении над $\mathbb{R}$)

Для всякого линейного оператора $\mathcal{A}$ в евклидовом пространстве существуют полярные разложения $\mathcal{A} = \mathcal{O} \circ \mathcal{S}_1$ и $\mathcal{A} = \mathcal{S}_2 \circ \mathcal{O}$, где $\mathcal{S}_1$ и $\mathcal{S}_2$ — неотрицательно определённые самосопряжённые операторы и $\mathcal{O}$ — ортогональный оператор. Если $\mathcal{A}$ невырожден, то $\mathcal{O}$, $\mathcal{S}_1$ и $\mathcal{S}_2$ определены однозначно.

Доказательство. С у щ е с т в о в а н и е . Оператор $\mathcal{A}^* \circ \mathcal{A}$ положительно (неотрицательно) определён и самосопряжён $\implies$ существует положительно (неотрицательно) определённый самосопряжённый оператор $\mathcal{B} = \mathcal{B}^*$, для которого $\mathcal{B}^2 = \mathcal{A}^* \circ \mathcal{A}$. Положим $\mathcal{S}_1 = \mathcal{B}$. Имеем $(\mathcal{A}\boldsymbol{x}, \mathcal{A}\boldsymbol{y}) = (\mathcal{A}\boldsymbol{x}, \mathcal{A}^{**}\boldsymbol{y}) = (\mathcal{A}^* \circ \mathcal{A}\boldsymbol{x}, \boldsymbol{y}) = (\mathcal{S}_1^* \circ \mathcal{S}_1\boldsymbol{x}, \boldsymbol{y}) = (\mathcal{S}_1\boldsymbol{x}, \mathcal{S}_1\boldsymbol{y})$. По лемме $\mathcal{A} = \mathcal{O} \circ \mathcal{S}_1$ для ортогонального оператора $\mathcal{O}$. Положим $\mathcal{S}_2 = \mathcal{O} \circ \mathcal{S}_1 \circ \mathcal{O}^{-1}$. Матрица S_1 оператора $\mathcal{S}_1$ диагональна и имеет неотрицательные диагональные элементы в некотором ортонормированном базисе $\boldsymbol{E}$. Пусть O и S_2 — матрицы операторов $\mathcal{O}$ и $\mathcal{S}_2$ в том же базисе. Тогда O — матрица перехода от некоторого ортонормированного базиса к $\boldsymbol{E}$ и $S_1 = O^{-1}S_2O$ — матрица оператора $\mathcal{S}_2$ в этом базисе. Значит, оператор $\mathcal{S}_2$ самосопряжён и неотрицательно определён, и $\mathcal{A} = \mathcal{S}_2 \circ \mathcal{O}$.

Единственность. Пусть оператор $\mathcal{A}$ невырожден и $\mathcal{A} = \mathcal{O} \circ \mathcal{S}_1 = \mathcal{O}' \circ \mathcal{S}_1'$ — два полярных разложения. Имеем

$$\mathcal{A}^* \circ \mathcal{A} = \mathcal{S}_1'^* \circ \mathcal{O}'^* \circ \mathcal{O}' \circ \mathcal{S}_1' = \mathcal{S}_1'^2.$$

По лемме о квадратном корне $\mathcal{S}_1 = \mathcal{S}_1'$. Значит, $\mathcal{O}' = \mathcal{A} \circ \mathcal{S}_1^{-1} = \mathcal{O}$.

Единственность разложения вида $\mathcal{A} = \mathcal{S}_2 \circ \mathcal{O}$ доказывается так же. □

Полуторалинейные функции

Всюду в этом разделе V — n-мерное векторное пространство над $\mathbb{C}$.

На векторных пространствах над $\mathbb{C}$ не бывает положительно определённых билинейных функций: $b(i\boldsymbol{x}, i\boldsymbol{x}) = -b(\boldsymbol{x}, \boldsymbol{x})$.

Определение

Функция $b\colon V \times V \to \mathbb{C}$ называется полуторалинейной, если

1. b линейна по первому (или второму) аргументу, т.е.
$$b(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}) = b(\boldsymbol{x}, \boldsymbol{z}) + b(\boldsymbol{y}, \boldsymbol{z}) \quad \text{и} \quad b(\lambda\boldsymbol{x}, \boldsymbol{y}) = \lambda b(\boldsymbol{x}, \boldsymbol{y})$$
для всех $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ и $\lambda \in \mathbb{C}$;
2. b антилинейна по второму (или первому) аргументу, т.е.
$$b(\boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z}) = b(\boldsymbol{x}, \boldsymbol{y}) + b(\boldsymbol{x}, \boldsymbol{z}) \quad \text{и} \quad b(\boldsymbol{x}, \lambda\boldsymbol{y}) = \overline{\lambda} b(\boldsymbol{x}, \boldsymbol{y})$$
для всех $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ и $\lambda \in \mathbb{C}$ (черта — комплексное сопряжение).

Пусть $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ — базис в V. Полуторалинейная функция, как и билинейная, полностью определяется своими значениями $b_{ij} = b(\boldsymbol{e}_i, \boldsymbol{e}_j)$ на парах базисных векторов: если x_i , y_i — координаты векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ в $\boldsymbol{E}$, то

$$b(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i,j=1}^{n} b_{ij} x_i \overline{y}_j = (x_1, \ldots, x_n)\, B \begin{pmatrix} \overline{y}_1 \\ \vdots \\ \overline{y}_n \end{pmatrix},$$

где

$$B = (b_{ij}) = \begin{pmatrix} b(\boldsymbol{e}_1, \boldsymbol{e}_1) & b(\boldsymbol{e}_1, \boldsymbol{e}_2) & \ldots & b(\boldsymbol{e}_1, \boldsymbol{e}_n) \\ b(\boldsymbol{e}_2, \boldsymbol{e}_1) & b(\boldsymbol{e}_2, \boldsymbol{e}_2) & \ldots & b(\boldsymbol{e}_2, \boldsymbol{e}_n) \\ \vdots & \vdots & \ddots & \vdots \\ b(\boldsymbol{e}_n, \boldsymbol{e}_1) & b(\boldsymbol{e}_n, \boldsymbol{e}_2) & \ldots & b(\boldsymbol{e}_n, \boldsymbol{e}_n) \end{pmatrix}$$

— матрица полуторалинейной функции b.

В другом базисе $\boldsymbol{E}'$ та же функция b имеет матрицу

$$\boxed{B' = C^T B\, \overline{C}},$$

где $C = (c_{ij})$ — матрица перехода от $\boldsymbol{E}$ к $\boldsymbol{E}'$ и $\overline{C} = (\overline{c}_{ij})$.

Эрмитовы полуторалинейные функции

Определение

Полуторалинейная функция b называется эрмитовой, если для любых $\boldsymbol{x}, \boldsymbol{y} \in V$

$$b(\boldsymbol{x}, \boldsymbol{y}) = \overline{b(\boldsymbol{y}, \boldsymbol{x})}\,.$$

Полуторалинейная функция b называется косоэрмитовой, если для любых $\boldsymbol{x}, \boldsymbol{y} \in V$

$$b(\boldsymbol{x}, \boldsymbol{y}) = -\overline{b(\boldsymbol{y}, \boldsymbol{x})}\,.$$

Квадратная матрица $B \in M_n(\mathbb{C})$ называется эрмитовой (косоэрмитовой), если $B^T = \overline{B}$, т.е. $b_{ij} = \overline{b}_{ji}$ (если $B^T = -\overline{B}$, т.е. $b_{ij} = -\overline{b}_{ji}$).

Полуторалинейная функция (косо)эрмитова $\iff$ её матрица в нектором (любом) базисе (косо)эрмитова.

Если B — эрмитова матрица, то

- Все диагональные элементы B вещественные
- Определитель $\det B$ вещественный: $\overline{\det B} = \det \overline{B} = \det B^T = \det B$.

Если B — косоэрмитова матрица, то все диагональные элементы B чисто мнимые.

Всякая полуторалинейная функция есть сумма эрмитовой и косоэрмитовой:

$$b(\mathbf{x},\mathbf{y}) = \tfrac{1}{2}\left(b(\mathbf{x},\mathbf{y}) + \overline{b(\mathbf{y},\mathbf{x})}\right) + \tfrac{1}{2}\left(b(\mathbf{x},\mathbf{y}) - \overline{b(\mathbf{y},\mathbf{x})}\right).$$

Каждой эрмитовой полуторалинейной функции b соответствует эрмитова квадратичная форма

$$q(\mathbf{x}) = b(\mathbf{x},\mathbf{x}).$$

Замечание

Всякая эрмитова квадратичная функция принимает только вещественные значения.

Предложение

Всякая эрмитова полуторалинейная функция b однозначно определяется своей эрмитовой квадратичной функцией q.

$$q(\mathbf{x} \pm \mathbf{y}) = q(\mathbf{x}) + q(\mathbf{y}) \pm b(\mathbf{x},\mathbf{y}) \pm b(\mathbf{y},\mathbf{x})$$
$$q(\mathbf{x} \pm i\mathbf{y}) = q(\mathbf{x}) + q(\mathbf{y}) \mp ib(\mathbf{x},\mathbf{y}) \pm ib(\mathbf{y},\mathbf{x})$$
$$b(\mathbf{x},\mathbf{y}) = \tfrac{1}{4}\left(q(\mathbf{x}+\mathbf{y}) + iq(\mathbf{x}+i\mathbf{y}) - q(\mathbf{x}-\mathbf{y}) - iq(\mathbf{x}-i\mathbf{y})\right)$$

Теорема

В n-мерном векторном пространстве V над $\mathbb{C}$ для всякой эрмитовой полуторалинейной функции b существует базис $\boldsymbol{E}$, в котором функция b и соответствующая ей эрмитова квадратичная функция q имеют нормальный вид

$$b(\boldsymbol{x},\boldsymbol{y})=x_1\overline{y}_1+\ldots+x_k\overline{y}_k-x_{k+1}\overline{y}_{k+1}-\ldots-x_{k+\ell}\overline{y}_{k+\ell},\quad q(\boldsymbol{x})=|x_1|^2+\ldots+|x_k|^2-|x_{k+1}|^2-\ldots-|x_{k+\ell}|^2$$

(здесь x_i и y_i — координаты векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ в $\boldsymbol{E}$).

(Доказательство дословно повторяет доказательство теоремы Лагранжа.)

Числа k и ℓ — положительный и отрицательный индексы инерции формы q.

Теорема (закон инерции)

Положительный и отрицательный индексы инерции являются инвариантами эрмитовой квадратичной формы (т.е. не зависят от выбора базиса).

(Доказательство дословно повторяет доказательство в вещественном случае.)

Определение

Эрмитова квадратичная форма q (и соответствующая эрмитова полуторалинейная функция) положительно определена, если $q(\boldsymbol{x})>0$ для любого вектора $\boldsymbol{x}\neq\mathbf{0}$.

Эрмитова квадратичная форма положительно определена $\iff$ $k=n$ и $\ell=0$.

Теорема (формула Якоби)

Пусть $b\colon V \times V \to \mathbb{C}$ — эрмитова полуторалинейная функция на n-мерном пространстве V над $\mathbb{C}$, B — её матрица в некотором базисе $\boldsymbol{E}$ и $\Delta_1, \ldots, \Delta_n$ — угловые миноры B. Положим $\Delta_0 = 1$.

Если все Δ_i, $i \leqslant n$, отличны от нуля, то существует базис $\tilde{\boldsymbol{E}} = \{\tilde{\boldsymbol{e}}_1, \ldots, \tilde{\boldsymbol{e}}_n\}$ пространства V, в котором квадратичная функция q, ассоциированная с b, имеет вид

$$q(y_1\tilde{\boldsymbol{e}}_1 + \cdots + y_n\tilde{\boldsymbol{e}}_n) = \frac{\Delta_1}{\Delta_0}y_1^2 + \cdots + \frac{\Delta_n}{\Delta_{n-1}}y_n^2,$$

причём матрица перехода от $\boldsymbol{E}$ к $\tilde{\boldsymbol{E}}$ верхне-унитреугольна.

Доказательство дословно повторяет доказательство в случае билинейной функции.

Теорема (критерий Сильвестра)

Пусть V — конечномерное векторное пространство над $\mathbb{C}$ и $\boldsymbol{E}$ — любой базис в нём. Эрмитова полуторалинейная функция на V положительно определена $\iff$ все угловые миноры её матрицы в базисе $\boldsymbol{E}$ положительны.

Доказательство дословно повторяет доказательство в вещественном случае.

Унитарные пространства

Определение

Векторное пространство V над полем комплексных чисел $\mathbb{C}$ называется унитарным (или эрмитовым), если на V зафиксирована некоторая положительно определенная эрмитова полуторалинейная функция, называемая (эрмитовым) скалярным произведением и обозначаемая $(\cdot, \cdot)$.

Свойства эрмитова скалярного произведения

- $(\boldsymbol{x}, \boldsymbol{y}) = \overline{(\boldsymbol{y}, \boldsymbol{x})}$ для всех $\boldsymbol{x}, \boldsymbol{y} \in V$;
- $(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}) = (\boldsymbol{x}, \boldsymbol{z}) + (\boldsymbol{y}, \boldsymbol{z})$ для всех $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$;
- $(\lambda \boldsymbol{x}, \boldsymbol{y}) = \lambda(\boldsymbol{x}, \boldsymbol{y})$ для всех $\lambda \in \mathbb{C}$ и $\boldsymbol{x}, \boldsymbol{y} \in V$;
- $(\boldsymbol{x}, \boldsymbol{x}) > 0$ для всех $\boldsymbol{x} \in V \setminus \{\boldsymbol{0}\}$.

Из этих свойств следует антилинейность по второму аргументу:

- $(\boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x}, \boldsymbol{y}) + (\boldsymbol{x}, \boldsymbol{z})$ для всех $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$;
- $(\boldsymbol{x}, \lambda \boldsymbol{y}) = \overline{\lambda}(\boldsymbol{x}, \boldsymbol{y})$ для всех $\lambda \in \mathbb{C}$ и $\boldsymbol{x}, \boldsymbol{y} \in V$.

Примеры

1. В пространстве $\mathbb{C}^n$ строк длины n с комплексными элементами
$$\big((x_1, \dots, x_n), (y_1, \dots, y_n)\big) = x_1 \cdot \overline{y}_1 + \dots + x_n \cdot \overline{y}_n.$$
2. В пространстве квадратных матриц $M_n(\mathbb{C})$
$$(A, B) = \operatorname{tr}(A\overline{B}^T).$$
3. В пространстве непрерывных комплекснозначных функций на отрезке $[0, 1]$
$$(f, g) = \int\limits_0^1 f(t)\overline{g(t)}\,dt.$$

Определение

Длиной вектора $\boldsymbol{x}$ в унитарном пространстве V называется число $|\boldsymbol{x}| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})}$.

Неравенство Коши–Буняковского–Шварца

Для любых векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ унитарного пространства $|(\boldsymbol{x}, \boldsymbol{y})|^2 \leqslant (\boldsymbol{x}, \boldsymbol{x}) \cdot (\boldsymbol{y}, \boldsymbol{y})$, причем равенство имеет место тогда и только тогда, когда векторы $\boldsymbol{x}$ и $\boldsymbol{y}$ пропорциональны.

Доказательство. Положим $\alpha = \frac{(\boldsymbol{x}, \boldsymbol{y})}{|(\boldsymbol{x}, \boldsymbol{y})|}$. Имеем $\alpha \cdot \overline{\alpha} = 1$ и

$$(t\boldsymbol{x} + \alpha\boldsymbol{y}, t\boldsymbol{x} + \alpha\boldsymbol{y}) = t^2(\boldsymbol{x}, \boldsymbol{x}) + t\overline{\alpha}(\boldsymbol{x}, \boldsymbol{y}) + t\alpha\overline{(\boldsymbol{x}, \boldsymbol{y})} + (\boldsymbol{y}, \boldsymbol{y}) = t^2(\boldsymbol{x}, \boldsymbol{x}) + 2t|(\boldsymbol{x}, \boldsymbol{y})| + (\boldsymbol{y}, \boldsymbol{y}).$$

Дискриминант этого квадратного трёхчлена относительно t равен $4|(\boldsymbol{x}, \boldsymbol{y})|^2 - 4(\boldsymbol{x}, \boldsymbol{x})(\boldsymbol{y}, \boldsymbol{y})$. Он должен быть неположительным и может обращаться в ноль только когда $\boldsymbol{x}$ и $\boldsymbol{y}$ пропорциональны. □

Следствие (неравенство треугольника)

$$|\boldsymbol{x} + \boldsymbol{y}| \leqslant |\boldsymbol{x}| + |\boldsymbol{y}|.$$

Векторы $\boldsymbol{x}$ и $\boldsymbol{y}$ называются ортогональными (запись $\boldsymbol{x} \perp \boldsymbol{y}$), если они ортогональны относительно $(\cdot\,,\cdot)$, т.е. если $(\boldsymbol{x},\boldsymbol{y}) = (\boldsymbol{y},\boldsymbol{x}) = 0$. Ортогональное дополнение $U^{\perp}$ подпространства U унитарного пространства V определяется как в евклидовом случае:

$$U^{\perp} = \{\boldsymbol{x} \in V : (\boldsymbol{u},\boldsymbol{x}) = 0 \ \forall \boldsymbol{u} \in U\}.$$

Теорема Пифагора

Если $\boldsymbol{x} \perp \boldsymbol{y}$, то $|\boldsymbol{x}+\boldsymbol{y}|^2 = |\boldsymbol{x}|^2 + |\boldsymbol{y}|^2$.

Система векторов ортогональна, когда все векторы в ней попарно ортогональны.

Предложение

Любая ортогональная система ненулевых векторов линейно независима.

Определение

Система векторов $\{\boldsymbol{v}_1,\ldots,\boldsymbol{v}_k\}$ в унитарном пространстве называется ортонормированной, если она ортогональна и каждый вектор $\boldsymbol{v}_i$ базиса имеет единичную длину, т.е. $(\boldsymbol{v}_i,\boldsymbol{v}_j) = \delta_{ij}$ для $i,j \leqslant k$.

Предложение

Базис $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ *унитарного пространства ортонормирован тогда и только тогда, когда выполнено любое из эквивалентных условий:*

1. *для любых* $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$ *и* $\boldsymbol{y} = y_1\boldsymbol{e}_1 + \cdots + y_n\boldsymbol{e}_n$ $(\boldsymbol{x}, \boldsymbol{y}) = x_1\overline{y}_1 + \cdots + x_n\overline{y}_n$;
2. *координаты любого вектора* $\boldsymbol{x}$ *в базисе* $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ *вычисляются по формуле* $x_i = (\boldsymbol{x}, \boldsymbol{e}_i)$ *для* $i \leqslant n$, *т.е.* $\boldsymbol{x} = (\boldsymbol{x}, \boldsymbol{e}_1)\boldsymbol{e}_1 + \cdots + (\boldsymbol{x}, \boldsymbol{e}_n)\boldsymbol{e}_n$.

Процесс ортогонализации Грама–Шмидта

Если $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ — любой базис пространства V, то из него строится ортонормированный базис точно так же, как в вещественном случае: для $i \leqslant n$

$$\tilde{\boldsymbol{e}}_1 = \boldsymbol{e}_1, \qquad \tilde{\boldsymbol{e}}_i = \boldsymbol{e}_i - \frac{(\boldsymbol{e}_i, \tilde{\boldsymbol{e}}_1)}{(\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_1)}\tilde{\boldsymbol{e}}_1 - \cdots - \frac{(\boldsymbol{e}_i, \tilde{\boldsymbol{e}}_{i-1})}{(\tilde{\boldsymbol{e}}_{i-1}, \tilde{\boldsymbol{e}}_{i-1})}\tilde{\boldsymbol{e}}_{i-1}, \qquad \hat{\boldsymbol{e}}_i = \frac{\tilde{\boldsymbol{e}}_i}{|\tilde{\boldsymbol{e}}_i|}$$

или

$$\hat{\boldsymbol{e}}_1 = \frac{\boldsymbol{e}_1}{|\boldsymbol{e}_1|}, \qquad \hat{\boldsymbol{e}}_i = \frac{\boldsymbol{e}_i - (\boldsymbol{e}_i, \hat{\boldsymbol{e}}_1)\hat{\boldsymbol{e}}_1 - \cdots - (\boldsymbol{e}_i, \hat{\boldsymbol{e}}_{i-1})\hat{\boldsymbol{e}}_{i-1}}{|\boldsymbol{e}_i - (\boldsymbol{e}_i, \hat{\boldsymbol{e}}_1)\hat{\boldsymbol{e}}_1 - \cdots - (\boldsymbol{e}_i, \hat{\boldsymbol{e}}_{i-1})\hat{\boldsymbol{e}}_{i-1}|}.$$

Теорема

Во всяком конечномерном унитарном пространстве имеется ортонормированный базис. Более того, ортонормированный базис любого его подпространства можно дополнить до ортонормированного базиса всего пространства.

Упражнения

1. Докажите, что во всяком счётномерном унитарном пространстве существует ортонормированный базис.
2. Верно ли, что любую ортонормированную систему векторов в таком пространстве можно дополнить до ортонормированного базиса?
3. Пусть V — векторное пространство над $\mathbb{C}$ и $\boldsymbol{E}$ — произвольный базис в нём. Покажите, что на V существует единственное скалярное произведение, относительно которого базис $\boldsymbol{E}$ является ортонормированным. (Скажем, что это скалярное произведение порождено базисом $\boldsymbol{E}$.)

Ортогональное проектирование в унитарном пространстве

Определение и свойства ортогонального дополнения в конечномерном унитарном пространстве V совершенно аналогичны евклидову случаю. В частности, для любого подпространства U пространства V имеем $V = U \oplus U^{\perp}$. Ортогональная проекция $\mathrm{pr}_U\,\boldsymbol{x}$ вектора $\boldsymbol{x}$ на подпространство и его ортогональная составляющая $\mathrm{ort}_U\,\boldsymbol{x}$ тоже определяются точно так же.

Расстояние в унитарном пространстве

Определение

Расстояние $\rho(\boldsymbol{x}, \boldsymbol{y})$ между векторами $\boldsymbol{x}$ и $\boldsymbol{y}$ в унитарном пространстве V определяется формулой
$$\rho(\boldsymbol{x}, \boldsymbol{y}) = |\boldsymbol{x} - \boldsymbol{y}|.$$

Функция $\rho\colon V \times V \to \mathbb{R}$ — метрика (она рефлексивна, симметрична и удовлетворяет неравенству треугольника).

Определение

Расстояние $\rho(\boldsymbol{x}, U)$ между вектором $\boldsymbol{x}$ и подпространством U в унитарном пространстве V определяется формулой
$$\rho(\boldsymbol{x}, U) = \inf_{\boldsymbol{u} \in U} \rho(\boldsymbol{x}, \boldsymbol{u}).$$

Теорема

Расстояние от вектора $\boldsymbol{x}$ конечномерного унитарного пространства V до подпространства U равно $|\operatorname{ort}_U \boldsymbol{x}|$, причём единственным ближайшим к $\boldsymbol{x}$ вектором подпространства U является $\operatorname{pr}_U \boldsymbol{x}$.

Унитарные матрицы

Пусть $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ и $\boldsymbol{E}' = \{\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n\}$ — два ортонормированных базиса в n-мерном унитарном пространстве и $T = (t_{ij})$ — матрица перехода от $\boldsymbol{E}$ к $\boldsymbol{E}'$, т.е. $\boldsymbol{e}'_i = \sum_{j \leqslant n} t_{i,j} \boldsymbol{e}_j$. Вычисляя скалярное произведение $\boldsymbol{e}'_i$ и $\boldsymbol{e}'_j$ в базисах $\boldsymbol{E}$ и $\boldsymbol{E}'$, получаем

$$(\boldsymbol{e}'_i, \boldsymbol{e}'_j) = \sum_{k=1}^{n} t_{ki} \overline{t}_{kj} \qquad \text{и} \qquad (\boldsymbol{e}'_i, \boldsymbol{e}'_j) = \delta_{ij}.$$

Значит,

- столбцы T образуют ортонормированный базис в пространстве $\mathbb{C}^n$ n-ок комплексных чисел и
- $T^T \overline{T} = E$, так что $T^{-1} = \overline{T}^T$ и строки тоже образуют ортонормированный базис в $\mathbb{C}^n$;
- $|\det T| = \sqrt{\det T \cdot \overline{\det T}} = 1$.

Определение

Матрица U, удовлетворяющая условию $U^T \overline{U} = E$, называется унитарной.

Свойства унитарных матриц

- Матрица унитарна $\iff$ это матрица перехода между ортонормированными базисами.
- Если даны два базиса в унитарном пространстве, один из них ортонормирован и матрица перехода унитарна, то и другой базис ортонормирован.
- A унитарна $\implies$ A^{-1} унитарна.
- A и B унитарны $\implies$ AB унитарна.
- Множество U_n всех унитарных матриц порядка n образует подгруппу в полной линейной группе $\mathrm{GL}_n(\mathbb{C})$ всех невырожденных квадратных матриц порядка n над полем $\mathbb{C}$. Это унитарная группа U_n.

Изоморфизм унитарных пространств

Определение

Унитарные пространства V и U называются изоморфными, если существует отображение $\Phi\colon V \to U$, которое является изоморфизмом векторных пространств и удовлетворяет условию
$$(\Phi(\boldsymbol{x}), \Phi(\boldsymbol{y})) = (\boldsymbol{x}, \boldsymbol{y})$$
для любых $\boldsymbol{x}, \boldsymbol{y} \in V$. Само отображение Φ называется при этом изоморфизмом унитарных пространств V и U.

Замечание

Линейный оператор между конечномерными унитарными пространствами — изоморфизм $\iff$ его матрица унитарна.

Теорема

Любые два n-мерных унитарных пространства изоморфны.

Доказательство. Возьмём ортонормированные базисы $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ и $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ в n-мерных унитарных пространствах U и V, положим $\Phi(\boldsymbol{v}_i) = \boldsymbol{u}_i$ и продолжим по линейности. □

Линейные функции на унитарном пространстве

Для каждого линейного функционала $f : V \to \mathbb{C}$ на унитарном пространстве V найдётся единственный вектор $\boldsymbol{v}_f \in V$ такой, что $f(\boldsymbol{x}) = (\boldsymbol{v}_f, \boldsymbol{x})$ для всех $\boldsymbol{x} \in V$. Однако соответствие $\boldsymbol{x} \mapsto \boldsymbol{v}_f$ не линейно, а антилинейно. Тем не менее все общие понятия — сопряжённого пространства, взаимного базиса и пр. — определены и для унитарного пространства, и любое конечномерное унитарное пространство изоморфно своему сопряжённому V^* относительно скалярного произведения на V^*, порожденного любым базисом пространства V^*, взаимным с любым ортонормированным базисом пространства V.

Линейные операторы в унитарных пространствах

Через V всюду ниже обозначаем n-мерное унитарное пространство.

Операторы $V \to V$ имеют ту отличительную особенность, что их характеристические многочлены всегда имеют n корней $\Longrightarrow$ у любого такого оператора имеются собственные векторы и жорданов базис. В остальном теория таких операторов аналогична теории операторов в евклидовом пространстве.

Каждому линейному оператору $\mathcal{A}\colon V \to V$ соответствует полуторалинейная функция $b_{\mathcal{A}}\colon V \times V \to \mathbb{C}$, определённая правилом $b_{\mathcal{A}}(\boldsymbol{x}, \boldsymbol{y}) = (\mathcal{A}\boldsymbol{x}, \boldsymbol{y})$. Обозначим соответствие $\mathcal{A} \mapsto b_{\mathcal{A}}$ через α.

Теорема

1. *Матрица любого линейного оператора $\mathcal{A}\colon V \to V$ в любом ортонормированном базисе совпадает с транспонированной матрицей полуторалинейной функции $\alpha(\mathcal{A})$ в том же базисе.*
2. *Отображение α из пространства $\mathscr{L}(V)$ линейных операторов в V в пространство $\mathscr{B}(V)$ полуторалинейных функций на V — изоморфизм.*

Сопряжённые операторы в унитарном пространстве

Определение

Оператор $\mathcal{A}^*: V \to V$ в унитарном пространстве V называется сопряжённым к линейному оператору $\mathcal{A}: V \to V$, если для любых $\boldsymbol{x}, \boldsymbol{y} \in V$

$$(\mathcal{A}\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, \mathcal{A}^*\boldsymbol{y}).$$

Теорема

Для каждого линейного оператора $\mathcal{A}: V \to V$ с матрицей A в некотором (любом) ортонормированном базисе сопряжённый оператор $\mathcal{A}^*$ существует и единствен, его матрица в том же базисе есть эрмитово сопряжённая к A матрица $\overline{A}^T$.

Свойства операции сопряжения

- $(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*$,
- $(\mathcal{A}^*)^* = \mathcal{A}$,
- $(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*$,
- $(\lambda\mathcal{A})^* = \overline{\lambda}\mathcal{A}^*$,
- $\mathcal{E}^* = \mathcal{E}$.

Самосопряжённые операторы в унитарном пространстве

Определение

Линейный оператор $\mathcal{A}\colon V \to V$ в унитарном пространстве называется самосопряженным (или эрмитовым), если $\mathcal{A} = \mathcal{A}^*$, т.е. для любых $\boldsymbol{x}, \boldsymbol{y} \in V$
$$(\mathcal{A}\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, \mathcal{A}\boldsymbol{y}).$$

Теорема

Линейный оператор $\mathcal{A}$ в унитарном пространстве V самосопряжён $\iff$ его матрица в некотором ортонормированном базисе эрмитова $\iff$ его матрица в произвольном ортонормированном базисе эрмитова.

Доказательство. Из общей теоремы о матрицах сопряжённых операторов. □

Предложение

Пусть $\mathcal{A}\colon V \to V$ — самосопряжённый линейный оператор в конечномерном унитарном пространстве V и U — подпространство, инвариантное относительно $\mathcal{A}$. Тогда ортогональное дополнение $U^\perp$ тоже инвариантно относительно $\mathcal{A}$.

Замечания

1. Все собственные значения любого самосопряжённого оператора в унитарном пространстве вещественны.
2. Собственные векторы, соответствующие различным собственным значениям самосопряжённого оператора в унитарном пространстве, ортогональны.

Теорема

Для любого самосопряжённого линейного оператора $\mathcal{A}$ в конечномерном унитарном пространстве V найдётся ортонормированный базис, в котором матрица этого оператора диагональна и вещественна.

Следствие

Для оператора $\mathcal{A}$ в конечномерном унитарном пространстве следующие условия равносильны:

1. *оператор $\mathcal{A}$ самосопряжён;*
2. *существует ортонормированный базис, в котором матрица $\mathcal{A}$ диагональна и вещественна;*
3. *если A — матрица $\mathcal{A}$ в любом ортонормированном базисе, то существуют унитарная матрица T и вещественная диагональная матрица D, для которых $A = T^{-1}DT$.*

Упражнение

Докажите, что оператор $\mathcal{A}$ в унитарном пространстве удовлетворяет условию $\overline{\mathcal{A}}^* = -\mathcal{A}$ тогда и только тогда, когда в некотором ортонормированном базисе его матрица имеет диагональный вид с чисто мнимыми диагональными элементами.

Теорема

1. *Сумма самосопряжённых линейных операторов в унитарном пространстве и произведение самосопряжённого оператора на вещественное число являются самосопряжёнными операторами.*
2. *Композиция двух самосопряжённых операторов в унитарном пространстве является самосопряжённым оператором тогда и только тогда, когда эти операторы перестановочны.*

Приведение квадратичной функции к главным осям над $\mathbb{C}$

Напомним, что между линейными операторами $\mathcal{A}: V \to V$ и полуторалинейными функциями $b: V \times V \to \mathbb{C}$ на конечномерном унитарном пространстве имеется изоморфизм α, определённый правилом $\alpha(\mathcal{A})(\boldsymbol{x}, \boldsymbol{y}) = (\mathcal{A}\boldsymbol{x}, \boldsymbol{y})$, причём матрица $\mathcal{A}$ в любом ортонормированном базисе совпадает с транспонированной матрицей $\alpha(\mathcal{A})$ в том же базисе. Ясно, что оператор $\mathcal{A}$ самосопряжён тогда и только тогда, когда полуторалинейная функция $\alpha(\mathcal{A})$ эрмитова.

Теорема

Пусть V — n-мерное унитарное пространство и $b: V \times V \to \mathbb{C}$ — эрмитова полуторалинейная функция. Тогда в V существует ортонормированный базис $\boldsymbol{E}$, в котором квадратичная функция $q: V \to \mathbb{C}$, соответствующая b, записывается в виде суммы квадратов с вещественными коэффициентами:

$$q(\boldsymbol{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2,$$

где $\lambda_i \in \mathbb{R}$ и x_i — координаты вектора $\boldsymbol{x}$ в базисе $\boldsymbol{E}$.

Числа $\lambda_1, \ldots, \lambda_n$ — это собственные значения самосопряжённого оператора $\mathcal{A}$, для которого $b = \alpha(\mathcal{A})$.

Унитарные операторы

Определение

Линейный оператор $\mathcal{A}\colon V \to V$ в унитарном пространстве V называется унитарным, если он сохраняет скалярное произведение, т.е.

$$(\mathcal{A}\boldsymbol{x}, \mathcal{A}\boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{y}) \qquad \text{для любых } \boldsymbol{x}, \boldsymbol{y} \in V.$$

Замечание

$(\mathcal{A}\boldsymbol{x}, \mathcal{A}\boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{y})$ для любых $\boldsymbol{x}, \boldsymbol{y} \in V \iff \mathcal{A}^* \circ \mathcal{A} = \mathcal{E} \iff \mathcal{A}^* = \mathcal{A}^{-1}$.

Теорема

Линейный оператор $\mathcal{A}\colon V \to V$ в конечномерном унитарном пространстве V является унитарным $\iff$ его матрица в некотором ортонормированном базисе унитарна $\iff$ его матрица в произвольном ортонормированном базисе унитарна.

Следствие

Линейный оператор в конечномерном унитарном пространстве унитарен $\Longleftrightarrow$ *он переводит некоторый ортонормированный базис в ортонормированный базис* $\Longleftrightarrow$ *он переводит любой ортонормированный базис в ортонормированный базис.*

Теорема

Линейный оператор в унитарном пространстве унитарен тогда и только тогда, когда он сохраняет длины векторов.

Следствие

Все собственные значения унитарного линейного оператора по модулю равны 1.

Предложение

Пусть $\mathcal{A}\colon V \to V$ *— унитарный линейный оператор в унитарном пространстве* V *и* U *— подпространство, инвариантное относительно* $\mathcal{A}$. *Тогда ортогональное дополнение* $U^{\perp}$ *тоже инвариантно относительно* $\mathcal{A}$.

Теорема

Линейный оператор в конечномерном унитарном пространстве унитарен тогда и только тогда, когда в некотором ортонормированном базисе его матрица диагональна, причём все её диагональные элементы по модулю равны 1.

Таким образом, для унитарного оператора, в отличие от ортогонального, всегда существует базис, состоящий из собственных векторов.

Комплексификация евклидова пространства

Комплексификация $V_{\mathbb{C}}$ евклидова пространства V (или, более общо, векторного пространства над $\mathbb{R}$ со скалярным произведением) каноническим образом превращается в унитарное пространство, если определить эрмитово скалярное умножение по формуле

$$(\boldsymbol{x} + i\boldsymbol{y}, \boldsymbol{u} + i\boldsymbol{v}) = [(\boldsymbol{x}, \boldsymbol{u}) + (\boldsymbol{y}, \boldsymbol{v})] + i[(\boldsymbol{x}, \boldsymbol{v}) - (\boldsymbol{u}, \boldsymbol{y})].$$

При этом комплексное продолжение $\mathcal{A}_{\mathbb{C}}$ самосопряжённого (ортогонального) оператора $\mathcal{A}$ будет самосопряжённым (соответственно, унитарным) оператором на $V_{\mathbb{C}}$.

Полярное разложение над $\mathbb{C}$

Определение

Самосопряжённый линейный оператор $\mathcal{A}$ в унитарном пространстве V называется положительно (неотрицательно) определённым, если соответствующая ему эрмитова полуторалинейная функция $\alpha(\mathcal{A})$ положительно (неотрицательно) определена, т.е. $(\mathcal{A}\boldsymbol{x}, \boldsymbol{x}) = (\boldsymbol{x}, \mathcal{A}\boldsymbol{x}) > 0$ $((\mathcal{A}\boldsymbol{x}, \boldsymbol{x}) = (\boldsymbol{x}, \mathcal{A}\boldsymbol{x}) \geqslant 0)$ для любого ненулевого вектора $\boldsymbol{x} \in V$.

Лемма

Самосопряжённый оператор в конечномерном унитарном пространстве положительно (неотрицательно) определён тогда и только тогда, когда все его собственные значения положительны (неотрицательны).

Лемма

Для всякого положительно (неотрицательно) определённого самосопряжённого оператора $\mathcal{A}$ в конечномерном унитарном пространстве существует единственный положительно (неотрицательно) определённый самосопряжённый оператор $\mathcal{B}$, для которого $\mathcal{A} = \mathcal{B}^2$ — квадратный корень из $\mathcal{A}$.

Лемма

Пусть $\mathcal{A}$ и $\mathcal{B}$ — линейные операторы в конечномерном унитарном пространстве и $(\mathcal{A}\mathbf{x}, \mathcal{A}\mathbf{y}) = (\mathcal{B}\mathbf{x}, \mathcal{B}\mathbf{y})$ для всех $\mathbf{x}, \mathbf{y} \in V$. Тогда существует унитарный оператор $\mathcal{U}$, для которого $\mathcal{A} = \mathcal{U} \circ \mathcal{B}$.

Теорема (о полярном разложении над $\mathbb{C}$)

Для всякого линейного оператора $\mathcal{A}$ в конечномерном унитарном пространстве существуют полярные разложения $\mathcal{A} = \mathcal{U} \circ \mathcal{S}_1$ и $\mathcal{A} = \mathcal{S}_2 \circ \mathcal{U}$, где $\mathcal{S}_1$ и $\mathcal{S}_2$ — неотрицательно определённые самосопряжённые операторы и $\mathcal{U}$ — ортогональный оператор. Если $\mathcal{A}$ невырожден, то $\mathcal{U}$, $\mathcal{S}_1$ и $\mathcal{S}_2$ определены однозначно.

Аффинные пространства

Аффинным пространством, ассоциированным с векторным пространством V, или аффинным пространством над V, называется пара $(\mathbb{A}, +)$, где $\mathbb{A}$ — множество (точек) и $+\colon \mathbb{A} \times V \to \mathbb{A}$ — операция прибавления вектора к точке, удовлетворяющая условиям

1. $A + (\boldsymbol{x} + \boldsymbol{y}) = (A + \boldsymbol{x}) + \boldsymbol{y}$ для любых $A \in \mathbb{A}$ и $\boldsymbol{x}, \boldsymbol{y} \in V$,
2. $A + \mathbf{0} = A$ для любой точки $A \in \mathbb{A}$,
3. для любых точек $A, B \in \mathbb{A}$ существует единственный вектор $\boldsymbol{x} \in V$, для которого $A + \boldsymbol{x} = B$ (его обозначают $\overrightarrow{AB}$).

Элементы множества $\mathbb{A}$ называются точками. Вектор $\overrightarrow{AB}$ из условия 3 называется вектором, соединяющим точки A и B.

Аксиома 3 $\implies$ для любой фиксированной точки $A \in \mathbb{A}$ соответствие $\boldsymbol{v} \mapsto A + \boldsymbol{v}$ — биекция между V и $\mathbb{A}$.

Аксиома 1 $\implies$ для любых (не обязательно различных) трёх точек $A, B, C \in \mathbb{A}$ имеем $(A + \overrightarrow{AB}) + \overrightarrow{BC} = A + \overrightarrow{AC}$ и $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$ (аксиома треугольника).

Аксиома треугольника $\implies$ $\overrightarrow{AA} = \mathbf{0}$ и $\overrightarrow{AB} = -\overrightarrow{BA}$ для любых точек $A, B \in \mathbb{A}$.

Пример

$\mathbb{A} = V$ (векторы рассматриваются как точки).
Аксиомы ❶ и ❷ выполнены.
Аксиома ❸: Упорядоченной паре точек $\boldsymbol{u}, \boldsymbol{v} \in V$ сопоставляется вектор $\overrightarrow{\boldsymbol{u}\boldsymbol{v}} = \boldsymbol{v} - \boldsymbol{u}$.

Размерность $\dim \mathbb{A}$ аффинного пространства $\mathbb{A}$ над векторным пространством V полагается равной размерности векторного пространства V. Мы будем рассматривать главным образом аффинные пространства конечной размерности.

Всюду ниже в этом разделе предполается, что $\mathbb{A}$ — аффинное пространство, ассоциированное с векторным пространством V над полем K.

Зафиксировав точку $O \in \mathbb{A}$ («начало отсчёта»), можно отождествить каждую точку $X \in \mathbb{A}$ с её радиус-вектором $\overrightarrow{OX}$. Такое отождествление называется векторизацией аффинного пространства $\mathbb{A}$.

Аффинная система координат

Определение

Репером в n-мерном аффинном пространстве $\mathbb{A}$ над векторным пространством V называется пара $(O, \boldsymbol{E})$, где O — некоторая фиксированная точка в $\mathbb{A}$ (называемая началом координат) и $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ — базис векторного пространства V. С каждым репером связана аффинная система координат, в которой координатами точки $A \in \mathbb{A}$ являются координаты вектора $\overrightarrow{OA}$ в базисе $\boldsymbol{E}$.

Замечания

1. Если $(a_1, \ldots, a_n)$ и $(b_1, \ldots, b_n)$ — координаты точек A и B относительно репера $(O, \boldsymbol{E})$, то координаты вектора $\overrightarrow{AB}$ в базисе $\boldsymbol{E}$ равны $(b_1 - a_1, \ldots, b_n - a_n)$.
2. Если $B = A + \boldsymbol{v}$, $(a_1, \ldots, a_n)$ — координаты точки A относительно репера $(O, \boldsymbol{E})$ и $(v_1, \ldots, v_n)$ — координаты вектора $\boldsymbol{v}$ в базисе $\boldsymbol{E}$, то координаты точки B относительно репера $(O, \boldsymbol{E})$ равны $(a_1 + v_1, \ldots, a_n + v_n)$.

Барицентрические координаты

Определение

Пусть $k \geqslant 0$ и $A_0, \dots, A_k \in \mathbb{A}$. Выражение вида

$$\sum_{i=0}^{k} \lambda_i A_i, \qquad \text{где} \quad \lambda_i \in K \quad \text{и} \quad \sum_{i=0}^{k} \lambda_i = 1,$$

называется барицентрической линейной комбинацией точек $A_0, \dots, A_k \in \mathbb{A}$ с коэффициентами λ_i. Она считается равной точке A, определяемой равенством

$$\overrightarrow{OA} = \sum_{i=0}^{k} \lambda_i \overrightarrow{OA_i},$$

где O — произвольная точка пространства $\mathbb{A}$.

Барицентрическая комбинация точек $A_0, \dots, A_k$ с неотрицательными коэффициентами λ_i, $\sum \lambda_i = 1$, называется выпуклой комбинацией этих точек. Выпуклая комбинация $\frac{1}{k+1}(A_0 + \dots + A_k)$ называется центром тяжести системы точек $\{A_0, \dots, A_k\}$.

Барицентрическая линейная комбинация не зависит от выбора точки O:
$\overrightarrow{O'A} = \overrightarrow{O'O} + \overrightarrow{OA} = (\sum \lambda_i)\overrightarrow{O'O} + (\sum \lambda_i)\overrightarrow{OA} = \sum \lambda_i(\overrightarrow{O'O} + \overrightarrow{OA}) = \sum \lambda_i \overrightarrow{O'A}.$

Определение

Точки $A_0, A_1, \dots, A_k \in \mathbb{A}$ называются аффинно (не)зависимыми, если векторы $\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_n}$ линейно (не)зависимы в V.

Бесконечное множество $S \subset \mathbb{A}$ называется аффинно зависимым, если оно содержит конечное подмножество аффинно зависимых точек. Подмножество, не являющееся линейно зависимым, называется линейно независимым.

Определение

Пусть $\dim \mathbb{A} = n$ и $A_0, A_1, \dots, A_n \in \mathbb{A}$ — аффинно независимые точки. Тогда каждая точка $A \in \mathbb{A}$ единственным образом представляется в виде

$$A = \sum_{i=0}^{n} a_i A_i, \qquad \text{где} \quad \sum_{i=0}^{n} a_i = 1.$$

Числа $a_0, a_1, \dots, a_n$ называются барицентрическими координатами точки A относительно $A_0, A_1, \dots, A_n$.

Связь между барицентрическими и аффинными координатами: $a_0, a_1, \dots, a_n$ — барицентрические координаты произвольной точки A относительно аффинно независимых точек $A_0, A_1, \dots, A_n \iff a_1, \dots, a_n$ — её аффинные координаты относительно репера $(A_0, \boldsymbol{E})$, где $\boldsymbol{E} = \{\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_n}\}$, и $a_0 = 1 - \sum_{i=1}^{n} a_i$.

Теорема

Пусть $\mathbb{A}$ *—* n*-мерное аффинное пространство и* $X_0, \ldots, X_k \in \mathbb{A}$*. Тогда следующие условия равносильны:*

1. *точки* $X_0, \ldots, X_k \in \mathbb{A}$ *аффинно независимы;*
2. *ранг матрицы, составленной из их координат относительно некоторого (любого) репера с началом координат* X_0*, равен* k*;*
3. *ранг матрицы, составленной из их барицентрических координат относительно некоторых (любых)* $n+1$ *аффинно независимых точек равен* $k+1$*.*

Доказательство.

1 $\Leftrightarrow$ 2: из теории векторных пространств.

1 $\Leftrightarrow$ 3: Пусть $A_0, \ldots, A_n \in \mathbb{A}$ — любые аффинно независимые точки, и пусть $x_{0i}, \ldots, x_{ni}$ — соответствующие барицентрические координаты точки X_i для $i \leqslant k$. Тогда $x_{1i}, \ldots, x_{ni}$ — координаты вектора $\overrightarrow{A_0X_i}$ относительно базиса $\{\overrightarrow{A_0A_1}, \ldots, \overrightarrow{A_0A_n}\}$ в V.

$$\operatorname{rank}\begin{pmatrix} x_{00} & x_{10} & \dots & x_{k0} \\ x_{01} & x_{11} & \dots & x_{k1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{0n} & x_{1n} & \dots & x_{kn} \end{pmatrix} = \operatorname{rank}\begin{pmatrix} 1 & 1 & \dots & 1 \\ x_{01} & x_{11} & \dots & x_{k1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{0n} & x_{1n} & \dots & x_{kn} \end{pmatrix} =$$

$$= \operatorname{rank}\begin{pmatrix} 1 & 0 & \dots & 0 \\ x_{01} & x_{11} - x_{01} & \dots & x_{k1} - x_{01} \\ \vdots & \vdots & \ddots & \vdots \\ x_{0n} & x_{1n} - x_{0n} & \dots & x_{kn} - x_{0n} \end{pmatrix}$$

(к первой строке прибавили сумму всех остальных и вычли из каждого стобца первый). Ранг последней матрицы на 1 больше ранга матрицы

$$\begin{pmatrix} x_{11} - x_{01} & \dots & x_{k1} - x_{01} \\ \vdots & \ddots & \vdots \\ x_{1n} - x_{0n} & \dots & x_{kn} - x_{0n} \end{pmatrix},$$

составленной из координат векторов $\overrightarrow{X_0X_1}, \dots, \overrightarrow{X_0X_k}$. $\square$

Пусть в аффинном пространстве $\mathbb{A}$ даны два репера $(O, \boldsymbol{E})$ и $(O', \boldsymbol{E}')$. Рассмотрим произвольную точку $X \in \mathbb{A}$. Из равенства $\overrightarrow{OX} = \overrightarrow{O'X} + \overrightarrow{OO'}$ получаем формулу преобразования координат точки при переходе к другой системе координат:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = T \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} + \begin{pmatrix} x_{o'1} \\ x_{o'2} \\ \vdots \\ x_{o'n} \end{pmatrix},$$

где $x_1, x_2, \ldots, x_n$ и $x'_1, x'_2, \ldots, x'_n$ — координаты точки X относительно реперов $(O, \boldsymbol{E})$ и $(O', \boldsymbol{E}')$, $x_{o'1}, x_{o'2}, \ldots, x_{o'n}$ — координаты точки O' относительно репера $(O, \boldsymbol{E})$ и T — матрица перехода от базиса $\boldsymbol{E}$ к базису $\boldsymbol{E}'$.

Матрицей перехода от $(O, \boldsymbol{E})$ к $(O', \boldsymbol{E}')$ называется матрица

$$\widehat{T} = \begin{pmatrix} & T & & x_{o'1} \\ & & & \vdots \\ & & & x_{o'n} \\ 0 & \ldots & 0 & 1 \end{pmatrix}, \quad \text{для которой} \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix} = \widehat{T} \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \\ 1 \end{pmatrix}.$$

к аффинно-квадратичным функциям

Изоморфизм аффинных пространств

Скажем, что аффинные пространства $\mathbb{A}$ и $\mathbb{B}$ над векторным пространством V изоморфны, если существует биекция $\Phi\colon \mathbb{A} \to \mathbb{B}$ с тем свойством, что $\Phi(A + \boldsymbol{v}) = \Phi(A) + \boldsymbol{v}$ для любых $A \in \mathbb{A}$ и $\boldsymbol{v} \in V$.

Пусть V и W — векторные пространства над одним и тем же полем.

Определение

Аффинное пространство $\mathbb{A}$ над V и аффинное пространство $\mathbb{B}$ над W изоморфны, если существуют изоморфизм $\varphi\colon V \to W$ и биекция $\Phi\colon \mathbb{A} \to \mathbb{B}$ с тем свойством, что

$$\Phi(A + \boldsymbol{v}) = \Phi(A) + \varphi(\boldsymbol{v}) \qquad \text{для любых } A \in \mathbb{A} \text{ и } \boldsymbol{v} \in V.$$

При этом отображение $\varphi\colon V \to W$ называется дифференциалом, или линейной частью, отображения Φ.

Теорема

Два конечномерных аффинных пространства изоморфны тогда и только тогда, когда их размерности равны.

Доказательство. Если аффинные пространства $\mathbb{A}$ и $\mathbb{B}$ изоморфны, то векторные пространства V и W, ассоциированные с ними, изоморфны, а значит, $\dim V = \dim W$. Следовательно, $\dim \mathbb{A} = \dim \mathbb{B}$.

Обратно, если $\dim \mathbb{A} = \dim \mathbb{B}$, т.е. $\dim V = \dim W$, то существует изоморфизм $\varphi\colon V \to W$. Аффинный изоморфизм $f\colon \mathbb{A} \to \mathbb{B}$ строится по формуле $f(A + \boldsymbol{v}) = B + \varphi(\boldsymbol{v})$ для любых фиксированных $A \in \mathbb{A}$ и $B \in \mathbb{B}$. □

Вывод: Каждое аффинное пространство размерности n изоморфно пространству строк длины n над соответствующим полем. Вектор, идущий от одной строки к другой, равен разности этих строк.

Аффинные подпространства

Определение

Пусть U — подпространство векторного пространства V и $A \in \mathbb{A}$. Множество

$$\pi = \{A + \boldsymbol{x} \in \mathbb{A} : \boldsymbol{x} \in U\}$$

называется аффинным подпространством, или плоскостью, проходящей через точку A и имеющей направляющее подпространство U. Обозначение: $\pi = A + U$.

Размерностью плоскости π называется размерность ее направляющего подпространства U: $\dim \pi = \dim U$.

Нульмерная плоскость — это точка. Одномерная плоскость называется прямой, а $(n-1)$-мерная плоскость в n-мерном аффинном пространстве — гиперплоскостью.

Всякая плоскость $\pi = A + U$ является аффинным пространством, ассоциированным с U: выполнение аксиом 1 и 2 для π очевидно; для любых точек $A, B \in \pi$ вектор $\boldsymbol{u} = \overrightarrow{AB}$ принадлежит U и $B = A + \boldsymbol{u}$, причём вектор $\boldsymbol{u}$ определен однозначно в V, а значит, и в U $\Longrightarrow$ аксиома 3 тоже выполняется.

Замечания

1. Если $B \in \pi = A + U$, то $\pi = B + U$.
2. Для $\pi = A + U$ направляющее подпространство U есть $\{\overrightarrow{XY} : X, Y \in \pi\}$.
3. Пересечение плоскостей либо пусто, либо является плоскостью.
4. Для любого множества $S \subset \mathbb{A}$ и любой точки $X \in S$ плоскость
$$X + \langle\{\overrightarrow{XY} : Y \in S\}\rangle$$
является наименьшей плоскостью, содержащей S (т.е. пересечением всех плоскостей, содержащих S).

Определение

Наименьшая плоскость, содержащая множество $S \subset \mathbb{A}$, называется **аффинной оболочкой** множества S и обозначается $\operatorname{Aff} S$.

Для $X \in S$
$$\overrightarrow{XX} + \langle\{\overrightarrow{XY} : Y \in S\}\rangle = \left\{\sum_{i=1}^{k} \lambda_i \overrightarrow{XY_i} : k \in \mathbb{N},\ \lambda_i \in K,\ Y_i \in S\right\} =$$
$$= \left\{\sum_{i=0}^{k} \lambda_i \overrightarrow{XY_i} : k \in \mathbb{N},\ \lambda_i \in K,\ \sum \lambda_i = 1,\ Y_0 = X,\ Y_i \in S\right\}$$

$\Longrightarrow$ $\operatorname{Aff} S$ = множество всех барицентрических линейных комбинаций точек из S.

Предложение

Точки $A_0, \dots, A_k \in \mathbb{A}$ аффинно независимы $\iff$ они не содержатся ни в какой плоскости размерности $< k$.

Теорема

1. *Через любые $k+1$ точек аффинного пространства проходит плоскость размерности $\leqslant k$.*
2. *Если эти точки аффинно независимы, то через них проходит единственная плоскость размерности k.*

Доказательство. Пусть $A_0, \dots, A_k \in \mathbb{A}$. Тогда $\pi = A_0 + \langle\{\overrightarrow{A_0, A_1}, \dots, \overrightarrow{A_0, A_k}\}\rangle$ — плоскость размерности $\leqslant k$, проходящая через $A_0, \dots, A_k \in \mathbb{A}$. Если точки $A_0, \dots, A_k$ аффинно независимы, то π — единственная содержащая их $\leqslant k$-мерная плоскость, и $\dim \pi = k$. $\square$

Предложение

Пусть $S \subset \mathbb{A}$ и $M \subset S$ аффинно независимо. Множество M является максимальным аффинно независимым множеством в S $\iff$ $S \subset \operatorname{Aff} M$.

Доказательство. ⇒: Пусть $A_0 \in S$. Если $A_0 \in M$, то $A_0 \in \operatorname{Aff} M$. Предположим, что $A_0 \notin M$. M максимально $\implies$ множество $\{A_0\} \cup M$ аффинно зависимо. Пусть $A_1, \ldots, A_k \in M$ и точки $A_0, \ldots, A_k$ аффинно зависимы. Тогда векторы $\overrightarrow{A_0A_1}, \ldots, \overrightarrow{A_0A_k}$ линейно зависимы, причём $\overrightarrow{A_0A_1} \neq \mathbf{0}$ и, очевидно, $\overrightarrow{A_1A_0}$ выражается через $\overrightarrow{A_1A_2}, \overrightarrow{A_1A_3}, \ldots, \overrightarrow{A_1A_k}$ (это следует из того, что $\overrightarrow{A_0A_i} = \overrightarrow{A_0A_1} + \overrightarrow{A_1A_i}$ и точки $A_1, \ldots, A_k$ аффинно независимы). Пусть $\overrightarrow{A_1A_0} = \sum_{i=1}^k \lambda_i \overrightarrow{A_1A_i}$, причём $\lambda_1 = 1 - \sum_{i=2}^k \lambda_i$ (λ_1 может быть любым, так как это коэффициент при $\overrightarrow{A_1A_1} = \mathbf{0}$). Имеем $\sum_{i=1}^k \lambda_i = 1$ и $\overrightarrow{A_1A_0} = \sum_{i=1}^k \lambda_i \overrightarrow{A_1A_i}$, т.е. $A_0 = \sum_{i=1}^k \lambda_i A_i$, так что $A_0 \in \operatorname{Aff}(\{A_1, \ldots, A_k\}) \subset \operatorname{Aff} M$.

⇐: Предположим, что M не максимально. Пусть $X \in S$ и множество $\{X\} \cup M$ аффинно независимо. $S \subset \operatorname{Aff} M \implies X = \sum_{i=0}^k \lambda_i A_i$ для некоторых $k \geqslant 0$, $A_0, \ldots, A_k \in M$ и $\lambda_0, \ldots, \lambda_k \in K$ таких, что $\sum \lambda_i = 1$. По определению барицентрической линейной комбинации имеем $\overrightarrow{A_0X} = \sum_{i=0}^k \lambda_i \overrightarrow{A_0A_i} = \sum_{i=1}^k \lambda_i \overrightarrow{A_0A_i}$. Значит, векторы $\overrightarrow{A_0X}, \overrightarrow{A_0A_1}, \ldots, \overrightarrow{A_0A_k}$ линейно зависимы, т.е. точки $X, A_0, A_1, \ldots, A_k$ аффинно зависимы. Противоречие. □

Теорема (геометрическая характеризация аффинных подпространств)

Пусть $K \neq \mathbb{F}_2$. Непустое множество $S \subset \mathbb{A}$ является плоскостью $\iff$ вместе с любыми двумя различными точками оно содержит проходящую через них прямую.

Доказательство. $\Rightarrow$ очевидно, докажем $\Leftarrow$. Пусть $S \subset \mathbb{A}$ непусто и обладает указанным свойством. Покажем, что $S \supset \operatorname{Aff} S$. Пусть $X = \sum_{i=0}^{k} \lambda_i X_i$, где $k \in \mathbb{N}$, $\lambda_i \neq 0$, — барицентрическая линейная комбинация точек $X_0, X_1, \dots, X_k \in S$. Индукцией по k докажем, что $X \in S$. Для $k = 0, 1$ доказывать нечего. Пусть $k > 1$ и для меньших k всё доказано. Если найдётся $i \leqslant k$, для которого $\lambda_i \neq 1$ (пусть для определённости $\lambda_0 \neq 1$), то $X = \lambda_0 X_0 + (1-\lambda_0)\left(\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_0} X_i\right)$, причем по предположению индукции $Y = \left(\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_0} X_i\right) \in S$. Поскольку X лежит на прямой, проходящей через точки $Y \in S$ и X_0, имеем $X \in S$.

Предположим теперь, что $\lambda_i = 1$ для всех i, т.е. $k \cdot 1 = 0$ в поле K и $X = \sum_{i=0}^{k} X_i$. Возьмём $\alpha \in K \setminus \{0, 1\}$. Тогда $\sum_{i=0}^{k-1} 1/\alpha = 0$ и $1/\alpha \neq 1$. Имеем $X = \alpha\left(\sum_{i=0}^{k-1}(1/\alpha)X_i + X_k\right) + (1-\alpha)X_k$. В барицентрической линейной комбинации $Y = \sum_{i=0}^{k-1}(1/\alpha)X_i + X_k$ коэффициент при X_0 не равен 1 (и все коэффициенты ненулевые), так что по доказанному $Y \in S$. Поскольку X лежит на прямой, проходящей через точки Y и X_k, имеем $X \in S$. $\square$

Упражнение

Заметьте, что если $K = \mathbb{F}_2$, то для любых $A, B \in \mathbb{A}$ прямая, проходящая через точки A и B, состоит в точности из этих точек, т.е. $\mathrm{Aff}\{A, B\} = \{A, B\}$.

Докажите, что в случае $K = \mathbb{F}_2$ непустое множество $S \subset \mathbb{A}$ является плоскостью $\Longleftrightarrow$ вместе с любыми тремя различными точками оно содержит проходящую через них двумерную плоскость.

Взаимное расположение плоскостей

Объединение двух плоскостей, вообще говоря, не является плоскостью.

Теорема

Пусть $\pi = A + U$ и $\rho = B + W$ — две плоскости. Тогда

$$\operatorname{Aff}(\pi \cup \rho) = A + \langle \{\overrightarrow{AB}\} \cup (U + W) \rangle.$$

Пусть $\mathbb{A}$ конечномерно. Если $\pi \cap \rho \neq \varnothing$, то справедлива формула Грассмана

$$\dim \operatorname{Aff}(\pi \cup \rho) = \dim \pi + \dim \rho - \dim(\pi \cap \rho).$$

Если $\pi \cap \rho = \varnothing$, то

$$\dim \operatorname{Aff}(\pi \cup \rho) = \dim \pi + \dim \rho + 1 - \dim(U \cap W).$$

Доказательство. Положим $\sigma = A + \langle \{\overrightarrow{AB}\} \cup (U + W) \rangle$. Тогда $\pi, \rho \subset \sigma \implies \operatorname{Aff}(\pi \cup \rho) \subset \sigma$.

Обратно, $A \in \operatorname{Aff}(\pi \cup \rho) \implies \operatorname{Aff}(\pi \cup \rho) = A + V'$ для некоторого $V' \underset{\text{lin}}{\subset} V$. Из того, что $B \in \operatorname{Aff}(\pi \cup \rho)$, следует, что $\overrightarrow{AB} \in V'$. Для любого $\boldsymbol{u} \in U$ имеем $A + \boldsymbol{u} \in \pi \subset \operatorname{Aff}(\pi \cup \rho) \implies \boldsymbol{u} \in V'$. Для любого $\boldsymbol{w} \in W$ имеем $B + \boldsymbol{w} \in \rho \subset \operatorname{Aff}(\pi \cup \rho)$ и $B + \boldsymbol{w} = (A + \overrightarrow{AB}) + \boldsymbol{w} = A + (\overrightarrow{AB} + \boldsymbol{w}) \implies \overrightarrow{AB} + \boldsymbol{w} \in V'$. Поскольку $\overrightarrow{AB} \in V'$, имеем $\boldsymbol{w} \in V'$. Следовательно, $\langle \{\overrightarrow{AB}\} \cup (U + W) \rangle \subset V'$, т.е. $\sigma \subset \operatorname{Aff}(\pi \cup \rho)$.

Пусть $\pi \cap \rho \neq \varnothing$ и $X \in \pi \cap \rho$. Тогда $\overrightarrow{AB} = \overrightarrow{XB} - \overrightarrow{XA} \in U + W$. Следовательно, $\dim\langle\{\overrightarrow{AB}\} \cup (U + W)\rangle = \dim(U + W)$. Формула Грассмана $\Longrightarrow$

$$\begin{aligned}\dim \operatorname{Aff}(\pi \cup \rho) = \dim(U + W) = \dim U + \dim W - \dim(U \cap W) = \\ = \dim \pi + \dim \rho - \dim(\pi \cap \rho).\end{aligned}$$

Пусть $\pi \cap \rho = \varnothing$.

Если $\overrightarrow{AB} = \boldsymbol{u} + \boldsymbol{w}$ для $\boldsymbol{u} \in U$ и $\boldsymbol{w} \in W$, то по аксиоме 1 аффинного пространства $B = A + \overrightarrow{AB} = A + (\boldsymbol{u} + \boldsymbol{w}) = (A + \boldsymbol{u}) + \boldsymbol{w}$, т.е. $A + \boldsymbol{u} = B - \boldsymbol{w} \in \pi \cap \rho$. По предположению $\pi \cap \rho = \varnothing \Longrightarrow \overrightarrow{AB} \notin U + W$. Следовательно,

$$\dim\langle\{\overrightarrow{AB}\} \cup (U + W)\rangle = \dim(U + W) + 1$$

и

$$\begin{aligned}\dim \operatorname{Aff}(\pi \cup \rho) = \dim(U + W) + 1 = \dim U + \dim W - \dim(U \cap W) + 1 = \\ = \dim \pi + \dim \rho - \dim(U \cap W) + 1.\end{aligned}$$ $\square$

Замечание

Из доказательства видно, что плоскости $\pi = A + U$ и $\rho = B + W$ пересекаются тогда и только тогда, когда вектор $\overrightarrow{AB}$ лежит в $U + W$.

Определение

Плоскости $\pi = A + U$ и $\rho = B + W$ в аффинном пространстве $\mathbb{A}$ называются **параллельными**, если $U \subset W$ или $W \subset U$. Обозначение: $\pi \parallel \rho$.

Замечание

Плоскости π и ρ параллельны $\iff$ либо они не пересекаются, либо одна из них содержится в другой.

Определение

Плоскости π и ρ называются **скрещивающимися**, если они не параллельны и не пересекаются.

Для любых двух плоскостей есть три возможности: они могут

- пересекаться,
- не пересекаться и быть параллельными,
- скрещиваться.

Упражнение

Докажите, что если π — гиперплоскость в конечномерном аффинном пространстве $\mathbb{A}$, то любая не пересекающая π плоскость параллельна π.

Аффинно-линейные функции

Любое поле K можно рассматривать как одномерное векторное пространство над полем K, а также как аффинное пространство, ассоциированное с K. Пусть $\mathbb{A}$ — аффинное пространство, ассоциированное с векторным пространством V над K.

Определение

Отображение $f\colon \mathbb{A} \to K$ называется аффинно-линейной, или аффинной, функцией, если существует линейное отображение $\varphi\colon V \to K$ такое, что

$$f(X + \boldsymbol{x}) = f(X) + \varphi(\boldsymbol{x}) \qquad \text{для любых } X \in \mathbb{A} \text{ и } \boldsymbol{x} \in V.$$

Линейная функция φ называется дифференциалом функции f и обозначается df.

Замечание

Функция $f\colon \mathbb{A} \to K$ аффинно-линейна $\iff$ если $\exists O \in \mathbb{A}$ и $\exists$ линейное отображение (дифференциал) $df\colon V \to K$ такие, что $f(O + \boldsymbol{v}) = f(O) + df(\boldsymbol{v})$ для всех $\boldsymbol{v} \in V$.

Действительно, $\Rightarrow$ очевидно. $\Leftarrow$: $\forall X \in \mathbb{A}\ \exists \boldsymbol{x} \in V$ такой, что $X = O + \boldsymbol{x}$. Для любого $\boldsymbol{v} \in V$ имеем $f(X + \boldsymbol{v}) = f(O + \boldsymbol{x} + \boldsymbol{v}) = f(O) + df(\boldsymbol{x}) + df(\boldsymbol{v}) = f(X) + df(\boldsymbol{v})$.

Частный случай аффинно-линейных функций — постоянные функции (с $df \equiv 0$).

Пусть $\dim \mathbb{A} = n$ и в $\mathbb{A}$ зафиксирован репер $(O, \boldsymbol{E})$, где $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$, и пусть $df = f_1\boldsymbol{\varepsilon}_1 + \dots + f_n\boldsymbol{\varepsilon}_n$ — разложение дифференциала df по взаимному базису $\{\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_n\}$ в сопряженном пространстве V^*.

Если $x_1, \dots, x_n$ — координаты точки $X \in \mathbb{A}$ в $(O, \boldsymbol{E})$ и $f(O) = c \in K$, то

$$f(X) = f(O + \overrightarrow{OX}) = f(O) + df(\overrightarrow{OX}) = f(O) + df(x_1, \dots, x_n),$$

откуда

$$\boxed{f(X) = f_1x_1 + \dots + f_nx_n + c}. \qquad (*)$$

Обратно, любая функция $f\colon \mathbb{A} \to K$ вида $(*)$ аффинно-линейна: для любого вектора $\boldsymbol{v} = v_1\boldsymbol{e}_1 + \dots + v_n\boldsymbol{e}_n \in V$ точка $X + \boldsymbol{v} = O + (\overrightarrow{OX} + \boldsymbol{v})$ имеет координаты $(x_1 + v_1, \dots, x_n + v_n)$ относительно репера $(O, \boldsymbol{E})$, и $(*) \implies$

$$f(X + \boldsymbol{v}) = f_1(x_1 + v_1) + \dots + f_n(x_n + v_n) + c.$$

Следовательно,

$$f(X + \boldsymbol{v}) = df(f_1x_1 + \dots + f_nx_n + c) + (f_1v_1 + \dots + f_nv_n) = f(X) + \varphi(\boldsymbol{v}).$$

Выражение $f(X) = f_1x_1 + \dots + f_nx_n + c$ будем обозначать $f(x_1, \dots, x_n)$. Его называют аффинно-линейной формой.

Предложение

Барицентрические координаты относительно любых аффинно независимых точек в n-мерном аффинном пространстве — аффинно-линейные функции.

Доказательство. Пусть $x_0, x_1, \ldots, x_n$ — барицентрические координаты точки X относительно аффинно независимых точек $A_0, \ldots, A_n$. Тогда $x_1, \ldots, x_n$ — координаты вектора $\overrightarrow{A_0X}$ в базисе $\{\overrightarrow{A_0A_1}, \ldots, \overrightarrow{A_0A_n}\}$ векторного пространства V. Значит, $f_i(X) = x_i$, $i = 1, \ldots, n$, — аффинно-линейные функции. Поскольку $x_0 = 1 - \sum\limits_{i=1}^{n} x_i$, функция $f_0(X) = x_0$ тоже аффинно-линейна. $\square$

Упражнения

1. Докажите, что функция $f : \mathbb{A} \to K$ аффинно-линейна тогда и только тогда, когда для любой барицентрической линейной комбинации $\sum\limits_{i=0}^{k} \lambda_i X_i$ любых точек $X_0, \ldots, X_k \in \mathbb{A}$
$$f\Big(\sum_{i=0}^{k} \lambda_i X_i\Big) = \sum_{i=0}^{k} \lambda_i f(X_i)$$
(ниже доказана аналогичная теорема для общих аффинных отображений).
2. **Выпуклым многогранником** называется выпуклая оболочка конечного числа точек $A_0, \ldots, A_k$ в аффинном пространстве. Точки $A_0, \ldots, A_k$ называются **вершинами** многогранника.

   Выведите из упражнения 1, что ограничение любой аффинно-линейной функции на любой выпуклый многогранник достигает своего максимума (а также минимума) в одной из вершин многогранника.

Важнейшие приложения выпуклых многогранников связаны с линейным программированием, основная задача которого состоит в максимизации или минимизации данной аффинно-линейной функции на выпуклом многограннике.

Пример

Транспортная задача. Имеются пункты производства $\Pi_1, \ldots, \Pi_m$ некоторого продукта и пункты его реализации $\mathrm{P}_1, \ldots, \mathrm{P}_m$. Для всех $i \leqslant m$ и $j \leqslant n$ заданы следующие величины: объём производства a_i в Π_i, объём реализации b_j в P_j и затраты c_{ij} на перевозку единицы продукта из Π_i в P_j. Предполагается, что суммарный объём производства равен объёму реализуемого продукта: $\sum_{i=1}^m a_i = \sum_{j=1}^n b_j$. Требуется составить план перевозок (т.е. определить количество продукта x_{ij}, который предполагается доставить из Π_i в P_j), позволяющий полностью вывезти продукты всех производителей, обеспечивающий все пункты реализации и дающий минимум суммарных затрат на перевозку. Иными словами, на выпуклом многограннике

$$\sum_{j=1}^n x_{ij} = a_i, \qquad \sum_{i=1}^m x_{ij} = b_j, \qquad x_{ij} \geqslant 0$$

в mn-мерном пространстве с координатами x_{ij} $(i \leqslant m,\ j \leqslant n)$ требуется минимизировать функцию $\sum_{i=1}^m \sum_{j=1}^n c_{ij}x_{ij}$ — общую стоимость перевозки.

Определение

Множество точек, в которых аффинно-линейная функция $f : \mathbb{A} \to K$ обращается в ноль, называется **ядром** функции f и обозначается $\operatorname{Ker} f$.

Предложение

Ядро $\operatorname{Ker} f$ аффинно-линейной функции $f : \mathbb{A} \to K$ с дифференциалом $df : V \to K$ либо пусто, либо является плоскостью и равно $A + \operatorname{Ker} df$, где A — произвольная точка, удовлетворяющая условию $f(A) = 0$.

Доказательство. Для любой точки $B = A + \boldsymbol{v}$, где $\boldsymbol{v} \in \operatorname{Ker} df$, имеем $f(B) = f(A + \boldsymbol{v}) = f(A) + df(\boldsymbol{v}) = 0$. Значит, $A + \operatorname{Ker} df \subset \operatorname{Ker} f$.

Если $B \in \operatorname{Ker} f$, то $f(B) = f(A) + df(\overrightarrow{AB}) \implies df(\overrightarrow{AB}) = 0$, т.е. $\overrightarrow{AB} \in \operatorname{Ker} df$, а значит, $\operatorname{Ker} f \subset A + \operatorname{Ker} df$. $\square$

Из предложения следует, что пересечение ядер конечного числа аффинно-линейных функций либо пусто, либо является плоскостью.

Теорема

Пусть $\mathbb{A}$ — аффинное пространство размерности n. Тогда

1. *Множество точек из $\mathbb{A}$, координаты которых удовлетворяют совместной системе линейных уравнений ранга r, образуют $(n-r)$-мерную плоскость $\pi \subset \mathbb{A}$.*
2. *Произвольная k-мерная плоскость в аффинном пространстве $\mathbb{A}$ состоит из всех точек, координаты которых относительно произвольного репера составляют множество решений некоторой системы линейных уравнений ранга $n-k$.*

Доказательство. 1. Пусть дана совместная система линейных уравнений

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1, \\ \qquad \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m. \end{cases}$$

Рассмотрим аффинные функции $f_i \colon \mathbb{A} \to K$, $i = 1, \dots, m$, которые имеют координатную запись $f_i(X) = a_{i1}x_1 + \dots + a_{in}x_n - b_i$ относительно некоторого репера $(O, \boldsymbol{E})$.

Данную систему можно переписать в виде

$$\begin{cases} f_1(X) = 0, \\ \qquad \dots \\ f_m(X) = 0. \end{cases}$$

Видим, что числа $x_1, \dots, x_n$ являются решением данной системы $\iff$ они являются координатами относительно репера $(O, \boldsymbol{E})$ точки X, лежащей в плоскости $\pi = \operatorname{Ker} f_1 \cap \dots \cap \operatorname{Ker} f_m$.

Пусть $(x_1^0, \dots, x_n^0)$ — любое решение данной системы и X_0 — точка в $\mathbb{A}$ с координатами $x_1^0, \dots, x_n^0$ относительно репера $(O, \boldsymbol{E})$. Из курса алгебры знаем, что решения системы линейных уравнений суть суммы любого фиксированного частного решения неоднородной системы и решений однородной системы. Значит, $(x_1, \dots, x_n)$ — решение данной системы $\iff$ это набор координат точки $X = X_0 + \boldsymbol{x}$, где $\boldsymbol{x} \in V$ удовлетворяет системе

$$\begin{cases} df_1(X) = 0, \\ \qquad \dots \\ df_m(X) = 0 \end{cases}$$

(напомним, что df_i — дифференциал функции f_i).

Из курса алгебры знаем, что множество решений — подпространство U размерности $n - r$, где r — ранг матрицы коэффициентов однородной системы (см. также подраздел «Функционалы и подпространства» в начале этого курса, в конце раздела «Сопряжённое пространство»). Значит, множество решений данной системы — плоскость $\pi = X_0 + U$ размерности $n - r$.

2. Пусть $\pi = A + U$ — плоскость размерности k. В разделе о подпространствах векторных пространств была доказана теорема, что U — множество решений однородной системы линейных уравнений ранга $n - k$

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0, \\ \qquad \dots \\ a_{k1}x_1 + \dots + a_{kn}x_n = 0. \end{cases}$$

Точка X принадлежит плоскости $\pi = A + U \iff \overrightarrow{AX} \in U$. Если $a_1, \dots, a_n$ — координаты точки A, а $x_1, \dots, x_n$ — координаты точки X относительно репера $(O, \boldsymbol{E})$, то $x_1 - a_1, \dots, x_n - a_n$ — координаты вектора $\overrightarrow{AX}$ в базисе $\boldsymbol{E}$, которые удовлетворяют системе

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1, \\ \qquad \dots \\ a_{k1}x_1 + \dots + a_{kn}x_n = b_m, \end{cases}$$

где $b_i = \sum_{j=1}^{n} a_{ij}a_j$ для $i = 1, \dots, k$. $\square$

Аффинные отображения

Всюду ниже V и W — векторные пространства над одним и тем же полем K.

Определение

Пусть даны два аффинных пространства, $\mathbb{A}$ над V и $\mathbb{B}$ над W. Отображение $f : \mathbb{A} \to \mathbb{B}$ называется аффинным, если существует линейное отображение $\varphi : V \to W$ такое, что $f(A + \boldsymbol{v}) = f(A) + \varphi(\boldsymbol{v})$ для любых $A \in \mathbb{A}$ и $\boldsymbol{v} \in V$.

Отображение $\varphi : V \to W$ называется дифференциалом, или линейной частью, отображения f и обозначается df.

Замечания

1. $f(B) = f(A) + df(\overrightarrow{AB})$ для любых $A, B \in \mathbb{A}$ $\implies$ любое аффинное отображение однозначно определяется своим дифференциалом и значением в любой данной точке.
2. $df(\overrightarrow{AB}) = \overrightarrow{f(A)f(B)}$ для любых $A, B \in \mathbb{A}$.

Теорема

Пусть $\mathbb{A}$ и $\mathbb{B}$ — аффинные пространства над векторными пространствами V и W. Отображение $f : \mathbb{A} \to \mathbb{B}$ аффинно $\iff$ существуют точка $O \in \mathbb{A}$ и линейное отображение $\varphi : V \to W$ такие, что $f(O + \boldsymbol{v}) = f(O) + \varphi(\boldsymbol{v})$ для любого $\boldsymbol{v} \in V$. При этом $\varphi = df$.

Доказательство. $\Rightarrow$: очевидно.

$\Rightarrow$: Для любой точки $A \in \mathbb{A}$ существует вектор $\boldsymbol{w} \in V$ такой, что $A = O + \boldsymbol{w}$. Для всякого $\boldsymbol{v} \in V$ имеем

$$f(A + \boldsymbol{v}) = f(O + \boldsymbol{w} + \boldsymbol{v}) = f(O) + \varphi(\boldsymbol{w} + \boldsymbol{v}) = f(O) + \varphi(\boldsymbol{w}) + \varphi(\boldsymbol{v}) =$$
$$= f(O + \boldsymbol{w}) + \varphi(\boldsymbol{v}) = f(A) + \varphi(\boldsymbol{v}).$$

□

Предложение

Пусть $\mathbb{A}$ и $\mathbb{B}$ — аффинные пространства над векторными пространствами V и W и $f : \mathbb{A} \to \mathbb{B}$ — отображение. Тогда следующие условия равносильны:

1. *отображение f — изоморфизм аффинных пространств;*
2. *отображение f аффинно и биективно;*
3. *отображение f аффинно и его дифференциал $df : V \to W$ биективен, т.е. является изоморфизмом векторных пространств V и W.*

Доказательство. ❶ $\Leftrightarrow$ ❸: по определению изоморфизма.

❷ $\Rightarrow$ ❸: Пусть f — биекция. Для $A \in \mathbb{A}$ и $\boldsymbol{v} \in V$ имеем $f(A + \boldsymbol{v}) = f(A) + df(\boldsymbol{v})$. Если $\boldsymbol{v} \neq \mathbf{0}$, то $A + \boldsymbol{v} \neq A \implies f(A + \boldsymbol{v}) \neq f(A) \implies df(\boldsymbol{v}) \neq \mathbf{0}$, т.е. df инъективно.

Пусть $\boldsymbol{w} \in W$. Возьмём $A \in \mathbb{A}$. Отображение f сюръективно $\implies \exists$ точка $A' \in \mathbb{A}$ такая, что $f(A') = f(A) + \boldsymbol{w}$. Имеем $df(\overrightarrow{AA'}) = \overrightarrow{f(A)f(A')} = \boldsymbol{w}$, т.е. df сюръективно.

❸ $\Leftarrow$ ❷: Если $df : V \to W$ — изоморфизм и $A, A' \in \mathbb{A}$ различны, то $\overrightarrow{AA'} \neq \mathbf{0} \implies df(\overrightarrow{AA'}) \neq \mathbf{0} \implies f(A') = f(A) + df(\overrightarrow{AA'}) \neq f(A)$, т.е. f инъективно.

Возьмём $A \in \mathbb{A}$. df сюръективно $\implies \forall B \in \mathbb{B}$ существует $\boldsymbol{v} \in V$, для которого $df(\boldsymbol{v}) = \overrightarrow{f(A)B}$. Имеем $f(A + \boldsymbol{v}) = B$, т.е. f сюръективно. □

Теорема

Пусть $\mathbb{A}$ *и* $\mathbb{B}$ *— аффинные пространства над векторными пространствами* V *и* W*. Отображение* $f\colon \mathbb{A} \to \mathbb{B}$ *аффинно* $\iff$ *для любой барицентрической линейной комбинации* $\sum\limits_{i=0}^{k} \lambda_i X_i$ *любых точек* $X_0, \dots, X_k \in \mathbb{A}$ $f\Big(\sum\limits_{i=0}^{k} \lambda_i X_i\Big) = \sum\limits_{i=0}^{k} \lambda_i f(X_i)$.

Доказательство. $\Rightarrow$: Для $i \leqslant k$ положим $\boldsymbol{x}_i = \overrightarrow{X_0X_i}$. Имеем

$$f\Big(\sum_{i=0}^{k} \lambda_i X_i\Big) = f\Big(X_0 + \Big(\sum_{i=0}^{k} \lambda_i \boldsymbol{x}_i\Big)\Big) = f(X_0) + df\Big(\sum_{i=0}^{k} \lambda_i \boldsymbol{x}_i\Big) =$$

$$= \Big(\sum_{i=0}^{k} \lambda_i\Big) f(X_0) + df\Big(\sum_{i=0}^{k} \lambda_i \boldsymbol{x}_i\Big) = \sum_{i=0}^{k} \lambda_i (f(X_0) + df(\boldsymbol{x}_i)) = \sum_{i=0}^{k} \lambda_i f(X_i).$$

$\Leftarrow$: Пусть $\boldsymbol{E}$ — базис в V и $O \in \mathbb{A}$. Для любого $\boldsymbol{x} \in V$ имеем $\boldsymbol{x} = \sum\limits_{i=1}^{k} \lambda_i \boldsymbol{e}_i \in V$, где $k \in \mathbb{N}$, $\lambda_i \in K$ и $\boldsymbol{e}_i \in \boldsymbol{E}$. Положим $\varphi(\boldsymbol{x}) = \sum\limits_{i=1}^{k} \lambda_i \overrightarrow{f(O) f(O + \boldsymbol{e}_i)}$. Легко показать, что φ — линейное отображение $V \to W$.

Каждая точка $X \in \mathbb{A}$ имеет вид $X = O + \boldsymbol{x}$ для некоторого $\boldsymbol{x} = \sum\limits_{i=1}^{k} \lambda_i \boldsymbol{e}_i$. Положим $\lambda_0 = 1 - \sum\limits_{i=1}^{k} \lambda_i$, $X_0 = O$ и $X_i = O + \boldsymbol{e}_i$ для $i = 1, \ldots, k$. Имеем $\overrightarrow{OX} = \boldsymbol{x}$, $\overrightarrow{OX_0} = \boldsymbol{0}$, $\overrightarrow{OX_i} = \boldsymbol{e}_i$ для $i = 1, \ldots, k$ и

$$\overrightarrow{OX} = \sum_{i=0}^{k} \lambda_i \overrightarrow{OX_i}, \quad \text{т.е.} \quad X = \sum_{i=0}^{k} \lambda_i X_i.$$

По предположению

$$f(X) = \sum_{i=0}^{k} \lambda_i f(X_i).$$

По определению барицентрической линейной комбинации это означает, что

$$\overrightarrow{f(O)f(X)} = \sum_{i=0}^{k} \lambda_i \overrightarrow{f(O)f(X_i)} = \lambda_0 \overrightarrow{f(O)f(X_0)} + \sum_{i=1}^{k} \lambda_i \overrightarrow{f(O)f(X_i)} =$$
$$= \boldsymbol{0} + \sum_{i=1}^{k} \lambda_i \overrightarrow{f(O)f(O + \boldsymbol{e}_i)} = \varphi(\boldsymbol{x}),$$

т.е. $f(X) = f(O) + \varphi(\boldsymbol{x})$. Таким образом, для любого $\boldsymbol{x} \in V$ имеем $f(O + \boldsymbol{x}) = f(O) + \varphi(\boldsymbol{x})$, т.е. f — аффинное отображение с дифференциалом $df = \varphi$. $\square$

Следствие

Пусть $\mathbb{A}$ *и* $\mathbb{B}$ *— аффинные пространства над векторными пространствами* V *и* W *и* $\dim \mathbb{A} = n$. *тогда для любых аффинно независимых точек* $A_0, A_1, \ldots, A_n \in \mathbb{A}$ *и любых точек* $B_0, B_1, \ldots, B_n \in \mathbb{B}$ *существует единственное аффинное отображение* $f : \mathbb{A} \to \mathbb{B}$ *с тем свойством, что* $f(A_i) = B_i$ *для* $i = 0, 1, \ldots, n$.

Замечания

1. Для аффинных отображений $f_0, \ldots, f_k : \mathbb{A} \to \mathbb{B}$ и любых $\lambda_0, \ldots, \lambda_k \in K$, $\sum \lambda_i = 1$, можно определить поточечную барицентрическую линейную комбинацию $f = \sum_{i=0}^{k} \lambda_i f_i$ отображений f_i с коэффициентами λ_i, полагая
$$f(X) = \sum_{i=0}^{k} \lambda_i f_i(X) \qquad \text{для всех} \quad X \in \mathbb{A}.$$
По предыдущей теореме эта комбинация тоже является аффинным отображением $\mathbb{A} \to \mathbb{B}$.

2. Композиция аффинного отображения $f : \mathbb{A}_1 \to \mathbb{A}_2$ с дифференциалом $df : V_1 \to V_2$ и аффинного отображения $g : \mathbb{A}_2 \to \mathbb{A}_3$ с дифференциалом $dg : V_2 \to V_3$ является аффинным отображением $\mathbb{A}_1 \to \mathbb{A}_3$ с дифференциалом $dg \circ df : V_1 \to V_3$.

Матрица аффинного отображения

Пусть $\mathbb{A}$ — n-мерное аффинное пространство над V, $\mathbb{B}$ — m-мерное аффинное пространство над W, $(O_A, \boldsymbol{E}_A)$ — репер в $\mathbb{A}$, $(O_B, \boldsymbol{E}_B)$ — репер в $\mathbb{B}$ и $f\colon \mathbb{A} \to \mathbb{B}$ — аффинное отображение с дифференциалом $df\colon V \to W$. Матрицей аффинного отображения f в системах координат $(O_A, \boldsymbol{E}_A)$ и $(O_B, \boldsymbol{E}_B)$ называется матрица

$$\widehat{A} = \begin{pmatrix} & A & & \begin{matrix} o'_1 \\ \vdots \\ o'_m \end{matrix} \\ 0 & \dots & 0 & 1 \end{pmatrix},$$

где A — матрица дифференциала df в базисах $\boldsymbol{E}_A$ и $\boldsymbol{E}_B$ и $o'_1, \dots, o'_m$ — координаты точки $f(O_A)$ относительно репера $(O_B, \boldsymbol{E}_B)$.

Пусть $x_1, \dots, x_n$ — координаты точки $X \in \mathbb{A}$ относительно репера $(O_A, \boldsymbol{E}_A)$ (т.е. координаты вектора $\overrightarrow{O_A X}$ в базисе $\boldsymbol{E}_A$) и $x'_1, \dots, x'_m$ — координаты $f(X) \in \mathbb{B}$ относительно репера $(O_B, \boldsymbol{E}_B)$ (т.е. координаты вектора $\overrightarrow{O_B f(X)}$ в базисе $\boldsymbol{E}_B$). Тогда $f(X) = f(O_A) + df(\overrightarrow{O_A X}) = O_B + \overrightarrow{O_B f(O_A)} + df(\overrightarrow{O_A X})$. В координатах:

$$\boxed{\begin{pmatrix} x'_1 \\ \vdots \\ x'_m \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} o'_1 \\ \vdots \\ o'_m \end{pmatrix}}, \quad \text{или} \quad \boxed{\begin{pmatrix} x'_1 \\ \vdots \\ x'_m \\ 1 \end{pmatrix} = \widehat{A} \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix}}, \quad \text{или} \quad \boxed{\begin{cases} x'_1 = a_{11}x_1 + \dots + a_{1n}x_n + o'_1, \\ \quad \dots \\ x'_m = a_{m1}x_1 + \dots + a_{mn}x_n + o'_m \end{cases}}.$$

Пусть $(O_A, \boldsymbol{E}_A)$, $(\tilde{O}_A, \tilde{\boldsymbol{E}}_A)$ — два репера в аффинном пространстве $\mathbb{A}$ и $(O_B, \boldsymbol{E}_B)$, $(\tilde{O}_B, \tilde{\boldsymbol{E}}_B)$ — два репера в аффинном пространстве $\mathbb{A}$, и пусть T_A $(\widehat{T}_A)$ — матрица перехода от базиса $\boldsymbol{E}_A$ к базису $\tilde{\boldsymbol{E}}_A$ (от репера $(O_A, \boldsymbol{E}_A)$ к реперу $(\tilde{O}_A, \tilde{\boldsymbol{E}}_A)$), а T_B $(\widehat{T}_B)$ — соответствующие матрицы для $\mathbb{B}$. Тогда $\begin{pmatrix} x'_1 \\ \vdots \\ x'_m \\ 1 \end{pmatrix} = \widehat{T}_B \begin{pmatrix} \tilde{x}'_1 \\ \vdots \\ \tilde{x}'_m \\ 1 \end{pmatrix}$ и

$\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix} = \widehat{T}_A \begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \\ 1 \end{pmatrix}$, так что в новых координатах $\begin{pmatrix} \tilde{x}'_1 \\ \vdots \\ \tilde{x}'_m \\ 1 \end{pmatrix} = \widehat{T}_B^{-1} \widehat{A}\, \widehat{T}_A \begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \\ 1 \end{pmatrix}$, откуда

$$\boxed{\widehat{\tilde{A}} = \widehat{T}_B^{-1} \widehat{A}\, \widehat{T}_A} \qquad \text{и} \qquad \boxed{\tilde{A} = T_B^{-1} A\, T_A}$$

(второе равенство следует из того, что у матриц $\widehat{A}$, $\widehat{T}_A$ и $\widehat{T}_B$ последняя строка равна $(0 \ldots 01)$).

Рангом аффинного отображения f называется ранг его дифференциала. В координатах: $\operatorname{rank} f = \operatorname{rank} A = \operatorname{rank} \widehat{A} - 1$.

Аффинные операторы

Определение

Аффинным оператором называется аффинное отображение аффинного пространства в себя. Взаимно однозначный аффинный оператор называется аффинным преобразованием.

Примеры

1. Сдвиг, или параллельный перенос, на фиксированный вектор $\boldsymbol{v} \in V$:
$$t_{\boldsymbol{v}}(X) = X + \boldsymbol{v} \qquad \text{для всех } X \in \mathbb{A}.$$
2. Гомотетия с центром в фиксированной точке $O \in \mathbb{A}$ и коэффициентом $\lambda \in K \setminus \{0\}$:
$$f(X) = O + \lambda\overrightarrow{OX} \qquad \text{для всех } X \in \mathbb{A}.$$
Гомотетия с коэффициентом $\lambda = -1$ называется центральной симметрией.

Упражнения

1. Докажите, что всякое аффинное преобразование с дифференциалом $\lambda\mathcal{E}$, где $\lambda \neq 0, 1$, является гомотетией.
2. Докажите, что композиция гомотетий с разными центрами и коэффициентами λ и μ является гомотетией, если $\lambda \cdot \mu \neq 1$, и параллельным переносом, если $\lambda \cdot m = 1$.

Матрица аффинного оператора

Пусть $\mathbb{A}$ — n-мерное аффинное пространство над V, $(O, \boldsymbol{E})$ — репер в $\mathbb{A}$ и $f : \mathbb{A} \to \mathbb{A}$ — аффинный оператор с дифференциалом $df : V \to V$. Матрицей аффинного оператора f относительно репера $(O, \boldsymbol{E})$ называется матрица

$$\widehat{A} = \begin{pmatrix} & A & & \begin{matrix} o'_1 \\ \vdots \\ o'_n \end{matrix} \\ 0 & \dots & 0 & 1 \end{pmatrix},$$

где A — матрица дифференциала df в базисе $\boldsymbol{E}$ и $o'_1, \dots, o'_n$ — координаты точки $f(O)$ относительно репера $(O, \boldsymbol{E})$.

Пусть $x_1, \dots, x_n$ — координаты точки $X \in \mathbb{A}$ относительно репера $(O, \boldsymbol{E})$ (т.е. координаты вектора $\overrightarrow{OX}$ в базисе $\boldsymbol{E}$) и $x'_1, \dots, x'_n$ — координаты $f(X)$ (т.е. координаты вектора $\overrightarrow{Of(X)}$ в базисе $\boldsymbol{E}$). Тогда $f(X) = f(O) + df(\overrightarrow{OX}) = O + \overrightarrow{Of(O)} + df(\overrightarrow{OX})$. В координатах:

$$\boxed{\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} o'_1 \\ \vdots \\ o'_n \end{pmatrix}}, \quad \text{или} \quad \boxed{\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \\ 1 \end{pmatrix} = \widehat{A} \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix}}, \quad \text{или} \quad \boxed{\begin{cases} x'_1 = a_{11}x_1 + \dots + a_{1n}x_n + o'_1, \\ \quad \dots \\ x'_n = a_{n1}x_1 + \dots + a_{nn}x_n + o'_n \end{cases}}.$$

Пусть $(O, \boldsymbol{E})$ и $(\tilde{O}, \tilde{\boldsymbol{E}})$ — два репера в аффинном пространстве $\mathbb{A}$, и пусть T $(\widehat{T})$ — матрица перехода от базиса $\boldsymbol{E}$ к базису $\tilde{\boldsymbol{E}}$ (от репера $(O, \boldsymbol{E})$ к реперу $(\tilde{O}, \tilde{\boldsymbol{E}})$).

Тогда $\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \\ 1 \end{pmatrix} = \widehat{T} \begin{pmatrix} \tilde{x}'_1 \\ \vdots \\ \tilde{x}'_n \\ 1 \end{pmatrix}$ и $\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix} = \widehat{T} \begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \\ 1 \end{pmatrix}$, так что в новых координатах

$\begin{pmatrix} \tilde{x}'_1 \\ \vdots \\ \tilde{x}'_n \\ 1 \end{pmatrix} = \widehat{T}^{-1} \widehat{A}\, \widehat{T} \begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \\ 1 \end{pmatrix}$, откуда

$$\boxed{\widehat{\tilde{A}} = \widehat{T}^{-1} \widehat{A}\, \widehat{T}} \qquad \text{и} \qquad \boxed{\tilde{A} = T^{-1} A\, T}$$

(второе равенство следует из того, что у матриц $\widehat{A}$ и $\widehat{T}$ строка равна $(0 \ldots 01)$).

Аффинный оператор $\mathbb{A} \to \mathbb{A}$ называется невырожденным, если его ранг равен $\dim \mathbb{A}$. Аффинный оператор невырожден тогда и только тогда, когда он обратим, так что

$$\text{невырожденные аффинные операторы} = \text{аффинные преобразования}.$$

Если $f : \mathbb{A} \to \mathbb{A}$ — невырожденный аффинный оператор, то для любой точки $O \in \mathbb{A}$ и любых векторов $\boldsymbol{x}, \boldsymbol{y} \in V$ имеем $f(f^{-1}(O) + \boldsymbol{x}) = O + df(\boldsymbol{x})$ и

$$f^{-1}(O + \boldsymbol{y}) = f^{-1}(O) + (df)^{-1}(\boldsymbol{y}),$$

так что $d(f^{-1}) = (df)^{-1}$.

Множество всех невырожденных аффинных операторов в аффинном пространстве $\mathbb{A}$ (= аффинных преобразований пространства $\mathbb{A}$) является группой, которая называется полной аффинной группой пространства $\mathbb{A}$ и обозначается $\mathrm{GA}(\mathbb{A})$. Группа $\mathrm{GA}(\mathbb{A})$ содержит подгруппы сдвигов и операторов с неподвижной точкой.

Предложение

1. *Любой сдвиг является аффинным оператором с тождественным дифференциалом.*
2. *Любой аффинный оператор с тождественным дифференциалом — сдвиг.*
3. *Множество* $T(\mathbb{A})$ *всех сдвигов аффинного пространства* $\mathbb{A}$ *является подгруппой группы* $\mathrm{GA}(\mathbb{A})$, *изоморфная аддитивной группе векторного пространства* V.

Доказательство

1. $t_{\boldsymbol{v}}(X + \boldsymbol{x}) = X + \boldsymbol{x} + \boldsymbol{v} = (X + \boldsymbol{v}) + \boldsymbol{x} = t_{\boldsymbol{v}}(X) + \boldsymbol{x}$.
2. Если $f(X + \boldsymbol{x}) = f(X) + \boldsymbol{x}$ для всех $X \in \mathbb{A}$ и $\boldsymbol{x} \in V$, то для любой фиксированной точки $O \in \mathbb{A}$ и любой точки $X \in \mathbb{A}$ имеем
$f(X) = f(O + \overrightarrow{OX}) = f(O) + \overrightarrow{OX} = X + \overrightarrow{Of(O)}$.
3. $t_{\boldsymbol{v}} \circ t_{\boldsymbol{u}} = t_{\boldsymbol{v}+\boldsymbol{u}}$, $t_{\boldsymbol{v}}^{-1} = t_{\boldsymbol{v}}$, $t_{\boldsymbol{0}} = \mathrm{id}\,|_{\mathbb{A}}$. □

Ясно, что множество G_A всех аффинных преобразований с неподвижной точкой $A \in \mathbb{A}$, — тоже подгруппа, причём $\mathrm{G}_A \cap \mathrm{T}(\mathbb{A}) = \{\mathrm{id}\,|_{\mathbb{A}}\}$.

Теорема

Пусть $A \in \mathbb{A}$. Тогда каждое аффинное преобразование $f \in \mathrm{GA}(\mathbb{A})$ однозначно представляется в виде композиций

$$f = t_{\boldsymbol{v}} \circ g_A \qquad \text{и} \qquad f = g_A \circ t_{\boldsymbol{w}},$$

где $\boldsymbol{v}, \boldsymbol{w} \in V$ и $g_A \in \mathrm{G}_A$. Таким образом, $\mathrm{GA}(\mathbb{A}) = \mathrm{G}_A\, \mathrm{T}(\mathbb{A}) = \mathrm{T}(\mathbb{A})\, \mathrm{G}_A$.

Доказательство. Положим $\boldsymbol{v} = \overrightarrow{Af(A)}$ и $g = t_{-\boldsymbol{v}} f$. Имеем $g \in \mathrm{G}_A$ и $dg = df$, причём $f = t_{\boldsymbol{v}} g$ (больше не будем писать $\circ$).

Чтобы получить второе разложение, запишем $f = g(g^{-1} t_{\boldsymbol{v}} g)$. Дифференциал оператора $g^{-1} t_{\boldsymbol{v}} g$ тождественный (это композиция дифференциалов), поэтому $g^{-1} t_{\boldsymbol{v}} g = t_{\boldsymbol{w}}$ для некоторого $\boldsymbol{w} \in V$. В явном виде:

$$\begin{aligned} A + \boldsymbol{w} = t_{\boldsymbol{w}}(A) = (g^{-1} t_{\boldsymbol{v}} g)(A) = g^{-1}(t_{\boldsymbol{v}}(g(A))) = g^{-1}(t_{\boldsymbol{v}}(A)) = \\ = g^{-1}(A + \boldsymbol{v}) = g^{-1}(A) + (dg)^{-1}(\boldsymbol{v}) = A + (df)^{-1}(\boldsymbol{v}). \end{aligned}$$

Единственность: пусть $f = t_{\boldsymbol{v}} g$ и $f = t_{\boldsymbol{v}'} g'$. Умножая равенство $t_{\boldsymbol{v}} g = t_{\boldsymbol{v}'} g'$ слева на $t_{\boldsymbol{v}'}^{-1}$, а справа на g^{-1}, получим $t_{\boldsymbol{v} - \boldsymbol{v}'} = g' g^{-1}$. Сдвиг с неподвижной точкой тождествен $\implies$ $t_{\boldsymbol{v}} = t_{\boldsymbol{v}'}$ и $g = g'$. Единственность второго разложения доказывается аналогично. $\square$

Основная теорема аффинной геометрии

Основная теорема аффинной геометрии

Пусть $\mathbb{A}$ — аффинное пространство, ассоциированное с векторным пространством V над полем $\mathbb{R}$, причём $\dim \mathbb{A} \geqslant 2$, и пусть $f : \mathbb{A} \to \mathbb{A}$ — биекция, при которой образ любой прямой — прямая. Тогда f — аффинное преобразование.

Доказательство. Биективность $\implies$

- образ плоскости — плоскость (по теореме о геометрической характеризации аффинных подпространств),
- прообраз прямой — прямая,
- образ любой двумерной плоскости π — двумерная плоскость (если $\dim f(\pi) > 2$, то в $f(\pi)$ через точку P, не лежащую на прямой ℓ, проходит много прямых, не пересекающих $\ell \implies$ в плоскости π через $f^{-1}P$ проходит много прямых, параллельных $f^{-1}(\pi)$ — противоречие);
- параллельные прямые переходят в параллельные прямые.

Правило параллелограмма $\implies$ если $\overrightarrow{AB} = \overrightarrow{CD}$, то $\overrightarrow{f(A)f(B)} = \overrightarrow{f(C)f(D)}$. Значит, правило $\overrightarrow{AB} \mapsto \overrightarrow{f(A)f(B)}$ корректно определяет отображение $\varphi\colon V \to V$, причём $f(X + \boldsymbol{x}) = f(X) + \varphi(\boldsymbol{x})$. Надо доказать, что φ линейно.

Правило параллелограмма $\implies \varphi(\boldsymbol{x}+\boldsymbol{y}) = \varphi(\boldsymbol{x}) + \varphi(\boldsymbol{y})$.

Осталось доказать, что для любых $\lambda \in \mathbb{R}$ и $\boldsymbol{x} \in V$ имеем $\varphi(\lambda\boldsymbol{x}) = \lambda\varphi(\boldsymbol{x})$. Параллельные прямые переходят в параллельные $\implies \varphi(\lambda\boldsymbol{x}) = \lambda\varphi(\boldsymbol{x})$ для $\lambda \in \mathbb{Z}$ и для λ вида $\frac{1}{n}$ и для $\lambda \in \mathbb{Q}$.

Пусть O, A и B — любые аффинно независимые точки в $\mathbb{A}$. Возьмём на прямых $\mathrm{Aff}\{O, A\}$ и $\mathrm{Aff}\{O, B\}$ точки X и Y. Точки $f(X)$ и $f(Y)$ лежат на прямых $\mathrm{Aff}\{f(O), f(A)\}$ и $\mathrm{Aff}\{f(O), f(B)\}$. Это означает, что определены функции $\Phi\colon \mathbb{R} \to \mathbb{R}$ и $\Psi\colon \mathbb{R} \to \mathbb{R}$ такие, что для $t \in \mathbb{R}$ имеем $\varphi(t\overrightarrow{OA}) = \Phi(t)\overrightarrow{f(O)f(A)}$ и $\varphi(t\overrightarrow{OB}) = \Psi(t)\overrightarrow{f(O)f(B)}$. Параллельные прямые переходят в параллельные прямые $\implies \Phi \equiv \Psi$ и $\Phi(x+y) = \Phi(x) + \Phi(y)$.

Покажем, что $\Phi(t) = t$ для всех $t \in \mathbb{R}$. Это так для $t \in \mathbb{Q} \implies$ достаточно проверить, что если $x < y$, то $\Phi(x) < \Phi(y)$.

Для любых $x, y \in \mathbb{R}$ имеем $\Phi(x \cdot y) = \Phi(x) \cdot \Phi(y)$. Действительно, если $y = 0$, то $\Phi(y) = 0$, и доказывать нечего. Пусть $y \neq 0$. Векторы $x\overrightarrow{OA} + y\overrightarrow{OB}$ и $\frac{x}{y}\overrightarrow{OA} + \overrightarrow{OB}$ пропорциональны $\implies$ их образы $\Phi(x)\overrightarrow{f(O)f(A)} + \Phi(y)\overrightarrow{f(O)f(B)}$ и $\Phi\left(\frac{x}{y}\right)\overrightarrow{f(O)f(A)} + \overrightarrow{f(O)f(B)}$ тоже $\implies \Phi\left(\frac{x}{y}\right) = \frac{\Phi(x)}{\Phi(y)}$.

Достаточно проверить, что $\Phi(t) > 0$ для $t > 0$, так если $x < y$, то $\Phi(y) = \Phi(y - x + x) = \Phi(y - x) + \Phi(x)$. Для $t > 0$ имеем $\Phi(t) = (\Phi(\sqrt{s}))^2 > 0$. $\square$

Основная теорема аффинной геометрии (конечномерная версия)

Пусть $\mathbb{A}$ *— конечномерное аффинное пространство, ассоциированное с векторным пространством над полем* $\mathbb{R}$*, причём* $\dim\mathbb{A} \geqslant 2$*, и пусть* $f\colon \mathbb{A}\to\mathbb{A}$ *— биекция, при которой образы любых трёх точек, лежащих на одной прямой, лежат на одной прямой. Тогда* f *— аффинное преобразование.*

Доказательство. Ввиду предыдущей теоремы достаточно показать, что образ каждой прямой — прямая. Для этого достаточно проверить, что образы любых трёх аффинно независимых точек аффинно независимы.

Для любых $A_0,\dots,A_m\in\mathbb{A}$ имеем $f(\mathrm{Aff}\{A_0,\dots,A_m\})\subset\mathrm{Aff}\{f(A_0),\dots,f(A_m)\}$. Это легко доказать индукцией по m, заметив, что $X\in\mathrm{Aff}\{A_0,\dots,A_m\} \iff X=\sum_{i=0}^m\lambda_iA_i$, где $\sum_{i=0}^m\lambda_i=1 \iff X$ лежит на прямой, соединяющей любую точку A_i с $\lambda_i\neq1$ с точкой $\sum_{j\neq i}\frac{\lambda_j}{1-\lambda_i}A_j\in\mathrm{Aff}\{A_0,\dots,A_{i-1},A_{i+1},\dots,A_m\}$.

Пусть $\dim\mathbb{A}=n$ и $A_0,A_1,A_2\in\mathbb{A}$ — аффинно независимые точки. Предположим, что $f(A_0),f(A_1),f(A_2)$ аффинно зависимы. Возьмём точки $A_3,\dots,A_n$, для которых система $A_0,\dots,A_n$ аффинно независима (т.е. дополним систему линейно независимых векторов $\{\overrightarrow{A_0A_1},\overrightarrow{A_0A_2}\}$ до базиса). Ясно, что точки $f(A_0),\dots,f(A_n)$ аффинно зависимы. Значит, $\mathrm{Aff}\{f(A_0),\dots,f(A_n)\}\neq\mathbb{A}$ в противоречие с биективностью f и тем, что $\mathrm{Aff}\{A_0,\dots,A_n\}=\mathbb{A}$. □

Упражнение

Приведите пример аффинного пространства $\mathbb{A}$ над полем $\mathbb{R}$ и не являющейся аффинным преобразованием биекции $f\colon \mathbb{A} \to \mathbb{A}$, при которой образы любых трёх точек, лежащих на одной прямой, лежат на одной прямой.

Евклидовы аффинные пространства

Определение

Аффинное пространство $\mathbb{A}$ над евклидовым пространством V называется аффинно-евклидовым пространством (или просто евклидовым пространством, если ясно, что речь идет о точечно-векторном пространстве). Расстоянием $\rho(A, B)$ между двумя точками $A, B \in \mathbb{A}$ называется длина вектора $\overrightarrow{AB}$.

Среди всех аффинных систем координат в аффинно-евклидовом пространстве выделяются системы, связанные с ортонормированными реперами, т.е. реперами, базис которых ортонормирован. Такие системы координат называются прямоугольными.

Если $x_1, \ldots, x_n$ и $y_1, \ldots, y_n$ — координаты точек $X, Y \in \mathbb{A}$ относительно прямоугольной системы координат, то

$$\boxed{\rho(X, Y) = \sqrt{(y_1 - x_1)^2 + \cdots + (y_n - x_n)^2}}.$$

Определение

Два аффинных евклидовых пространства $\mathbb{A}$ и $\tilde{\mathbb{A}}$ называются изоморфными, если между ними существует изоморфизм аффинных пространств $\varphi\colon \mathbb{A} \to \tilde{\mathbb{A}}$, сохраняющий расстояние между точками, т.е. такой, что $\rho(A, B) = \tilde{\rho}(\varphi(A), \varphi(B))$ для любых $A, B \in \mathbb{A}$; здесь ρ — функция расстояния на $\mathbb{A}$ и $\tilde{\rho}$ — функция расстояния на $\tilde{\mathbb{A}}$.

Теорема

Любые аффинно-евклидовы пространства $\mathbb{A}$ (над евклидовым пространством V) и $\tilde{\mathbb{A}}$ (над $\tilde{V}$) одинаковой размерности изоморфны.

Доказательство. Выберем ортонормированные реперы $(O, \boldsymbol{E})$ в $\mathbb{A}$ и $(\tilde{O}, \tilde{\boldsymbol{E}})$ в $\tilde{\mathbb{A}}$. Для точки $X \in \mathbb{A}$ с координатами $x_1, \dots, x_n$ относительно репера $(O, \boldsymbol{E})$ положим образ $\varphi(X)$ равным точке $\tilde{X} \in \tilde{\mathbb{A}}$ с координатами $\tilde{x}_1, \dots, \tilde{x}_n$ относительно репера $(\tilde{O}, \tilde{\boldsymbol{E}})$. Отображение φ — изоморфизм аффинных пространств, и он сохраняет расстояния. □

Будем говорить, что угол между ненулевыми векторными подпространствами U и W евклидова пространства V равен нулю, если одно из них лежит в другом. В противном случае углом между U и W будем называть точную нижнюю грань углов между ненулевыми векторами $\boldsymbol{u} \in U_0$ и $\boldsymbol{w} \in W_0$, где U_0 и W_0 — ортогональные дополнения подпространства $U \cap W$ соответственно в U и W. Углом между плоскостями в аффинно-евклидовом пространстве называется угол между их направляющими подпространствами. Плоскости называются перпендикулярными, если угол между ними равен $\frac{\pi}{2}$. Расстояние $\rho(\pi, \tau)$ между плоскостями π и τ определяется формулой

$$\rho(\pi, \tau) = \inf\{\rho(X, Y) : X \in \pi, Y \in \tau\}.$$

Теорема

Для любых двух плоскостей $\pi = A + U$ и $\tau = B + W$ расстояние $\rho(\pi, \tau)$ равно длине ортогональной составляющей вектора $\overrightarrow{AB}$ в V относительно подпространства $U + W$.

Доказательство. Разложим V в прямую сумму

$$V = (U + W) \oplus (U + W)^{\perp}$$

и представим $\overrightarrow{AB}$ в виде

$$\overrightarrow{AB} = \mathrm{pr}_{U+W}\,\overrightarrow{AB} + \mathrm{ort}_{U+W}\,\overrightarrow{AB}.$$

Выберем любые две точки $X = A + \boldsymbol{u} \in \pi$ и $Y = B + \boldsymbol{w} \in \tau$. По теореме Пифагора

$$\rho(X, Y)^2 = |\overrightarrow{AB} + \boldsymbol{w} - \boldsymbol{u}|^2 = |\,\mathrm{pr}_{U+W}(\overrightarrow{AB} + \boldsymbol{w} - \boldsymbol{u})|^2 + |\,\mathrm{ort}_{U+W}\,\overrightarrow{AB}|^2 \geqslant |\,\mathrm{ort}_{U+W}\,\overrightarrow{AB}|^2.$$

Равенство достигается для точек X и Y таких, что $\mathrm{pr}_{U+W}\,\overrightarrow{XY} = \boldsymbol{u} - \boldsymbol{w}$. □

Всюду ниже $\mathbb{A}$ — евклидово аффинное пространство, ассоциированное с евклидовым векторным пространством V, размерности n.

Определение

Движение, или изометрия, пространства $\mathbb{A}$ — это любое отображение $f\colon \mathbb{A} \to \mathbb{A}$ с тем свойством, что $\rho(f(A), f(B)) = \rho(A, B)$ для любых точек $A, B \in \mathbb{A}$ (т.е. сохраняющее расстояния между точками).

Теорема

Отображение $f\colon \mathbb{A} \to \mathbb{A}$ — движение $\Longleftrightarrow$ f — аффинное отображение с ортогональным дифференциалом.

Доказательство. $\Leftarrow$ очевидно. Докажем $\Rightarrow$. Возьмём любую точку $O \in \mathbb{A}$ и для каждого вектора $\boldsymbol{x} \in V$ положим $\varphi(\boldsymbol{x}) = \overrightarrow{f(O)f(O+\boldsymbol{x})}$. Получили отображение $\varphi\colon V \to V$, сохраняющее длины векторов, причём $f(O+\boldsymbol{x}) = f(O) + \varphi(\boldsymbol{x})$.

Пусть $\boldsymbol{x}, \boldsymbol{y} \in V$. Длина вектора $\boldsymbol{x} - \boldsymbol{y}$ равна расстоянию между точками $O + \boldsymbol{x}$ и $O + \boldsymbol{y} \implies |\varphi(\boldsymbol{x}) - \varphi(\boldsymbol{y})| = |\boldsymbol{x} - \boldsymbol{y}|$.

$2(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{x}) + (\boldsymbol{y}, \boldsymbol{y}) - (\boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y}) \implies$ имеем $(\varphi(\boldsymbol{x}), \varphi(\boldsymbol{y})) = (\boldsymbol{x}, \boldsymbol{y})$.

Отображение φ линейно:

Если $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$, то

$0 = |\boldsymbol{z} - \boldsymbol{x} - \boldsymbol{y}|^2 = |\boldsymbol{z}|^2 + |\boldsymbol{y}|^2 + |\boldsymbol{x}|^2 - 2(\boldsymbol{z}, \boldsymbol{x}) - 2(\boldsymbol{z}, \boldsymbol{y}) + 2(\boldsymbol{x}, \boldsymbol{y}) = |\varphi(\boldsymbol{z}) - \varphi(\boldsymbol{x}) - \varphi(\boldsymbol{y})|^2$
$\implies \varphi(\boldsymbol{z}) = \varphi(\boldsymbol{x}) + \varphi(\boldsymbol{y})$.

Если $\boldsymbol{y} = \lambda \boldsymbol{x}$, то $0 = |\boldsymbol{y} - \lambda \boldsymbol{x}|^2 = |\boldsymbol{y}|^2 + \lambda^2 |\boldsymbol{x}|^2 - 2\lambda(\boldsymbol{y}, \boldsymbol{x}) = |\varphi(\boldsymbol{y}) - \lambda\varphi(\boldsymbol{x})|^2 \implies \varphi(\boldsymbol{y}) = \lambda\varphi(\boldsymbol{x})$. □

Упражнение

Докажите, что если отображение обычой двумерной аффинно-евклидовой плоскости в себя переводит любые две точки, расстояние между которыми равно 1, в две точки, расстояние между которыми равно 1, то это отображение — движение.

Совокупность всех движений образует группу, называемую группой изометрий пространства $\mathbb{A}$ и обозначаемую $\operatorname{Isom}\mathbb{A}$. Движение собственное (сохраняет ориентацию), если определитель его дифференциала равен 1. В противном случае (если определитель дифференциала равен -1) движение несобственное (меняет ориентацию). Собственные движения образуют подгруппу в $\operatorname{Isom}\mathbb{A}$.

Теорема

Для любого движения $f\colon \mathbb{A}\to\mathbb{A}$ найдётся вектор $\boldsymbol{u}\in V$ такой, что $df(\boldsymbol{u})=\boldsymbol{u}$ и $f=t_{\boldsymbol{u}}\circ\Phi$, где Φ — движение с неподвижной точкой.

Доказательство. Если у df есть собственный вектор с собственным значением 1, то положим $U=V_1$ (собственное подпространство), в противном случае положим $U=\{\mathbf{0}\}$. Пусть $W=U^{\perp}$. df ортогонален $\implies$ W инвариантно, и $(df-\mathcal{E})|_W$ невырождено.

Возьмём $O\in\mathbb{A}$ и положим $\boldsymbol{v}=\overrightarrow{Of(O)}$. Поскольку $V=U\oplus W$, имеем $\boldsymbol{v}=\boldsymbol{u}+\boldsymbol{w}$, где $\boldsymbol{u}\in U$ и $\boldsymbol{w}\in W$, причём $df(\boldsymbol{u})=\boldsymbol{u}$. Положим $g=t_{\boldsymbol{u}}^{-1}\circ f$.

Найдём неподвижную точку отображения g. Пусть $X=O+\boldsymbol{x}$ и $\boldsymbol{x}=\boldsymbol{y}+\boldsymbol{z}$, где $\boldsymbol{y}\in U$ и $\boldsymbol{z}\in W$. Тогда $g(X)=f(O)+df(\boldsymbol{x})-\boldsymbol{u}=f(O)+df(\boldsymbol{x})-\boldsymbol{u}=$ $O+\boldsymbol{v}+df(\boldsymbol{x})-\boldsymbol{u}=O+df(\boldsymbol{x})+\boldsymbol{w}=O+\boldsymbol{x}+(df-\mathcal{E})\boldsymbol{x}+\boldsymbol{w}$. Значит, $X=O+\boldsymbol{x}$ неподвижна $\iff$ $(df-\mathcal{E})\boldsymbol{x}+\boldsymbol{w}=\mathbf{0}$ $\iff$ $\boldsymbol{x}=\left((df-\mathcal{E})|_W\right)^{-1}(-\boldsymbol{w})$. □

Следствие

1. *Если* $f\colon \mathbb{A} \to \mathbb{A}$ *— движение и* df *не имеет собственных векторов с собственным значением* 1*, то* f *имеет неподвижную точку.*
2. *Если* $f\colon \mathbb{A} \to \mathbb{A}$ *— движение и* df *имеет собственный вектор* $\boldsymbol{u}$ *с собственным значением* 1*, то прямая* $f(O) + \langle \boldsymbol{u} \rangle$ *состоит из неподвижных точек отображения* $t_{-\boldsymbol{u}} \circ f$.

Замечания

1. Любое собственное движение прямой — сдвиг, несобственное движение прямой — композиция сдвига и отражения относительно неподвижной точки дифференциала.
2. Собственное движение двумерной плоскости — либо сдвиг, либо поворот относительно неподвижной точки. Несобственное движение плоскости — композиция сдвига вдоль прямой, состоящей из неподвижных точек дифференциала, и отражения относительно этой прямой.

3. Собственное движение трёхмерного пространства — композиция сдвига вдоль прямой, состоящей из неподвижных точек дифференциала, и вращения вокруг этой прямой (винтовое движение). Несобственное движение пространства — либо композиция симметрии относительно плоскости, состоящей из неподвижных точек дифференциала, и сдвига вдоль прямой, параллельного этой плоскости, либо является композицией симметрии относительно плоскости и вращения вокруг оси, перпендикулярной к этой плоскости.

Аффинно-квадратичные функции

В этом разделе $\mathbb{A}$ — n-мерное аффинное пространство, ассоциированное с векторным пространством V над полем K характеристики $\neq 2$.

Определение

Аффинно-квадратичной функцией на аффинном пространстве $\mathbb{A}$ называется отображение $Q\colon \mathbb{A} \to K$ с тем свойством, что для любой точки $O \in \mathbb{A}$ существуют квадратичная функция $q\colon V \to K$ и линейная функция $\ell\colon V \to K$ такие, что

$$Q(X) = Q(O) + 2\ell(\overrightarrow{OX}) + q(\overrightarrow{OX}) \qquad \forall X \in \mathbb{A}.$$

Если $\tilde{O}$ — любая другая точка, то

$$\begin{aligned}
Q(X) &= Q(\tilde{O} + \overrightarrow{\tilde{O}X}) = Q(O + \overrightarrow{O\tilde{O}} + \overrightarrow{\tilde{O}X}) = Q(O) + 2\ell(\overrightarrow{O\tilde{O}} + \overrightarrow{OX}) + q(\overrightarrow{O\tilde{O}} + \overrightarrow{OX}) = \\
&= Q(O) + 2\ell(\overrightarrow{O\tilde{O}}) + 2\ell(\overrightarrow{OX}) + q(\overrightarrow{O\tilde{O}}) + q(\overrightarrow{OX}) + 2b(\overrightarrow{OX}, \overrightarrow{O\tilde{O}}) = \\
&= (Q(O) + 2\ell(\overrightarrow{O\tilde{O}}) + q(\overrightarrow{O\tilde{O}})) + 2\ell(\overrightarrow{OX}) + 2b(\overrightarrow{OX}, \overrightarrow{O\tilde{O}}) + q(\overrightarrow{OX}) = \\
&= Q(\tilde{O}) + 2(\ell(\overrightarrow{OX}) + b(\overrightarrow{OX}, \overrightarrow{O\tilde{O}})) + q(\overrightarrow{OX}),
\end{aligned}$$

где b — полярная к q билинейная функция, $\implies$ q не зависит от O.

Если $(O, \boldsymbol{E})$ — репер, $B = (b_{ij})$ — матрица квадратичной функции q (билинейной функции b) в базисе $\boldsymbol{E}$ и $l_1, \ldots, l_n$ — координаты линейной функции ℓ во взаимном с $\boldsymbol{E}$ базисе и $x_1, \ldots, x_n$ — координаты точки X в $(O, \boldsymbol{E})$, то

$$Q(X) = \sum_{i,j} b_{ij} x_i x_j + 2 \sum_i l_i x_i + a_0 = (x_1, \ldots, x_n)\, B \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + 2(l_1, \ldots, l_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + a_0,$$

где $a_0 = Q(O)$.

Матрица аффинно-квадратичной функции Q:

$$\widehat{B} = \begin{pmatrix} & B & & l_1 \\ & & & \vdots \\ & & & l_n \\ l_1 & \ldots & l_n & a_0 \end{pmatrix},$$

для неё

$$Q\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = (x_1, \ldots, x_n, 1)\widehat{B} \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix}.$$

Для другого репера $(O', \boldsymbol{E}')$ с матрицей перехода $\widehat{C}$:

$$\widehat{B}' = \widehat{C}^T \widehat{B} \widehat{C}, \quad \begin{pmatrix} l'_1 \\ \vdots \\ l'_n \end{pmatrix} = C^T \left(B \begin{pmatrix} \tilde{o}_1 \\ \vdots \\ \tilde{o}_n \end{pmatrix} + \begin{pmatrix} l_1 \\ \vdots \\ l_n \end{pmatrix} \right), \quad a'_0 = Q(O')$$

(здесь $\tilde{o}_1, \ldots, \tilde{o}_n$ — координаты точки O' относительно репера $(O, \boldsymbol{E})$).

Если меняется только начало координат, а базис не меняется, то $B' = B$ и $\begin{pmatrix} l'_1 \\ \vdots \\ l'_n \end{pmatrix} = B \begin{pmatrix} \tilde{o}_1 \\ \vdots \\ \tilde{o}_n \end{pmatrix} + \begin{pmatrix} l_1 \\ \vdots \\ l_n \end{pmatrix}$. Значит, если для точки $\tilde{O} = \begin{pmatrix} \tilde{o}_1 \\ \vdots \\ \tilde{o}_n \end{pmatrix}$ имеем $B \begin{pmatrix} \tilde{o}_1 \\ \vdots \\ \tilde{o}_n \end{pmatrix} = -\begin{pmatrix} l_1 \\ \vdots \\ l_n \end{pmatrix}$, то в записи аффинно-квадратичной функции Q относительно репера $(\tilde{O}, \boldsymbol{E})$ нет линейных членов. Другими словами, для любого $\boldsymbol{x} \in V$ имеем $Q(\tilde{O} + \boldsymbol{x}) = Q(\tilde{O}) + q(\boldsymbol{x})$, т.е. $Q(\tilde{O} + \boldsymbol{x}) = Q(\tilde{O} - \boldsymbol{x})$.

Определение

Точка $\tilde{O}$ называется **центром** аффинно-квадратичной функции Q, если для любого вектора $\boldsymbol{x} \in V$

$$Q(\tilde{O} + \boldsymbol{x}) = Q(\tilde{O} - \boldsymbol{x}).$$

Множество центров $=$ множество решений уравнения $B \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = -\begin{pmatrix} l_1 \\ \vdots \\ l_n \end{pmatrix}$.

Приведение аффинно-квадратичной функции к каноническому виду

Теорема

Для любой аффинно-квадратичной функции $Q\colon \mathbb{A} \to K$, *не являющейся аффинно-линейной, существует репер* $(O, \boldsymbol{E})$, *в котором* Q *имеет вид*

$$Q(X) = \lambda_1 x_1^2 + \cdots + \lambda_r x_r^2 + a_0, \qquad r \leqslant n, \tag{$*$}$$

или

$$Q(X) = \lambda_1 x_1^2 + \cdots + \lambda_r x_r^2 + 2x_{r+1}, \qquad r < n, \tag{$**$}$$

где $\lambda_1, \ldots, \lambda_r \in K \setminus \{0\}$.

Замечание

Функция, которая имеет вид $(*)$ относительно некоторого репера, не приводится к виду $(**)$, и наоборот, так как в первом случае разность $\operatorname{rank} \widehat{B} - \operatorname{rank} B$ равна 1 или 0, а во втором — 2, и ранги не меняются при замене координат.

Доказательство. Теорема Лагранжа $\Longrightarrow$ существует базис $\boldsymbol{E}$, в котором $q(\boldsymbol{x}) = \lambda_1 x_1^2 + \dots + \lambda_r x_r^2$, причем $\lambda_1, \dots, \lambda_r$ ненулевые. Возьмём любую точку O'. Запишем Q относительно репера $(O', \boldsymbol{E})$:

$$Q(O' + \boldsymbol{x}) = \lambda_1 x_1'^2 + \dots + \lambda_r x_r'^2 + 2l_1' x_1' + \dots + 2l_n' x_n' + a_0'.$$

Перенесём начало координат в точку O'' с координатами $-\frac{l_1'}{\lambda_1}, \dots, -\frac{l_r'}{\lambda_r}, 0, \dots, 0$, т.е. сделаем замену координат $x_i'' = x_i' + \frac{l_i'}{\lambda_i}$ для $i \leqslant r$, $x_i'' = x_i'$ для $i > r$:

$$Q(O'' + \boldsymbol{x}) = \lambda_1 x_1''^2 + \dots + \lambda_r x_r''^2 + 2l_{r+1}' x_{r+1}'' + \dots + 2l_n' x_n'' + a_0''.$$

Если $l_{r+1}' = \dots = l_n' = 0$, то Q имеет вид $(*)$ (и O'' — центр функции Q).
Пусть $l_i' \neq 0$ для некоторого $i > r$. Пусть для определённости $l_{r+1}' \neq 0$. Тогда замена

$$x_{r+1} = l_{r+1}' x_{r+1}'' + \dots + l_n' x_n'' + \frac{a_0''}{2}, \qquad x_i = x_i'' \text{ для } i \neq r+1$$

приводит Q к виду $(**)$ (и Q не имеет центра). □

Замечание

В случае $K = \mathbb{C}$ ($K = \mathbb{R}$) функция Q приводится к виду $(*)$ или $(**)$, в котором $\lambda_i = 1$ ($\lambda_i = \pm 1$) для $i \leqslant r$, причём этот вид определён однозначно (для поля $\mathbb{R}$ число $+1$ и -1 не зависит от системы координат в силу закона инерции).

Аффинно-квадратичные функции в аффинно-евклидовых пространствах

Ниже $\mathbb{A}$ — n-мерное аффинное пространство, ассоциированное с евклидовым векторным пространством V.

Теорема

Для любой аффинно-квадратичной функции $Q\colon \mathbb{A} \to \mathbb{R}$, не являющейся аффинно-линейной, существует прямоугольная система координат $(O, \boldsymbol{E})$, в которой Q имеет вид

$$Q(X) = \lambda_1 x_1^2 + \dots + \lambda_r x_r^2 + a_0, \qquad r \leqslant n, \tag{$*$}$$

или

$$Q(X) = \lambda_1 x_1^2 + \dots + \lambda_r x_r^2 + 2\lambda_{r+1} x_{r+1}, \qquad r < n, \tag{$**$}$$

где $\lambda_1, \dots, \lambda_r \in \mathbb{R} \setminus \{0\}$ и $\lambda_{r+1} > 0$. Этот вид определён однозначно с точностью до нумерации переменных.

Доказательство. Начало доказательства прежнее, только надо ссылаться не на теорему Лагранжа, а на теорему о приведении квадратичной функции к главным осям. Если Q имеет центр, то и конец доказательства прежний.

Предположим, что центра нет и функция уже приведена к виду

$$Q(O'' + \boldsymbol{x}) = \lambda_1 x_1''^2 + \dots + \lambda_r x_r''^2 + 2l'_{r+1}x''_{r+1} + \dots + 2l'_n x''_n + a''_0,$$

где не все l'_j, $j > r$, равны 0, относительно ортонормированного репера $(O'', \boldsymbol{E}')$. Положим

$$\boldsymbol{e}_{r+1} = \frac{1}{\sqrt{{l'_{r+1}}^2 + \dots + {l'_n}^2}}(l'_{r+1}\boldsymbol{e}'_{r+1} + \dots + l'_n\boldsymbol{e}'_n) \in \langle \boldsymbol{e}'_{r+1}, \dots, \boldsymbol{e}'_n \rangle.$$

Это единичный вектор, ортогональный всем $\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_r$. Дополним его до базиса в $\langle \boldsymbol{e}'_{r+1}, \dots, \boldsymbol{e}'_n \rangle$ векторами $\boldsymbol{e}_{r+2}, \dots, \boldsymbol{e}_n$. Положим $\boldsymbol{E} = \{\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_r, \boldsymbol{e}_{r+1}, \dots, \boldsymbol{e}_n\}$. Относительно репера $(O'', \boldsymbol{E})$ функция Q имеет вид

$$Q(O'' + \boldsymbol{x}) = \lambda_1 x_1'''^2 + \dots + \lambda_r x_r'''^2 + 2\frac{1}{\sqrt{{l'_{r+1}}^2 + \dots + {l'_n}^2}} x'''_{r+1} + a'''_0.$$

Осталось сдвинуть начало координат вдоль вектора $\boldsymbol{e}_{r+1}$, чтобы обнулить a'''_0.

Единственность:

Тип $(*)$ или $(**)$ определён однозначно (из рассмотрения рангов).
Числа r и $\lambda_1, \dots, \lambda_r$ определены однозначно по теореме о приведении квадратичной функции в евклидовом пространстве к главным осям.
a_0 равно значению Q в центре и не зависит от выбора центра: если O_1 и O_2 — центры, то $Q(O_2) = Q(O_1) + q(\overrightarrow{O_1O_2}) = Q(O_2) + q(\overrightarrow{O_2O_1}) + q(\overrightarrow{O_1O_2})$ и $q(\overrightarrow{O_1O_2}) = 0$.
Осталось доказать единственность λ_{r+1}. Предположим, что относительно ортонормированных реперов $(O, \boldsymbol{E})$ и $(O', \boldsymbol{E}')$ Q записывается как

$$Q(X) = \lambda_1 x_1^2 + \dots + \lambda_r x_r^2 + 2\lambda_{r+1} x_{r+1} \quad \text{и} \quad Q(X) = \lambda_1 {x_1'}^2 + \dots + \lambda_r {x_r'}^2 + 2\tilde{\lambda}_{r+1} x_{r+1}',$$

и пусть B — (диагональная) матрица q относительно этих реперов. Рассмотрим оператор $\mathcal{B}\colon V \to V$ с матрицей B. Его матрица совпадает с матрицей q во всех ортонормированных базисах, так как при ортогональной замене координат матрицы операторов и квадратичных функций меняются одинаково ($C^T = C^{-1}$).
Значит, оператор $\mathcal{B}$ имеет одну и ту же матрицу в базисах $\boldsymbol{E}$ и $\boldsymbol{E}'$, причём эта матрица диагональна и диагональные элементы с номерами $> r$ нулевые. Из вида матрицы $\mathcal{B}$ в базисе $\boldsymbol{E}$ следует, что $\operatorname{Im} \mathcal{B} = \mathcal{B}(\langle\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_r\}\rangle) = \langle\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_r\}\rangle$. Из (того же самого) вида матрицы $\mathcal{B}$ в базисе $\boldsymbol{E}'$ следует, что
$\operatorname{Im} \mathcal{B} = \mathcal{B}(\langle\{\boldsymbol{e}_1', \dots, \boldsymbol{e}_r'\}\rangle) = \langle\{\boldsymbol{e}_1', \dots, \boldsymbol{e}_r'\}\rangle$.

Таким образом, $\langle\{\boldsymbol{e}_1,\ldots,\boldsymbol{e}_r\}\rangle = \langle\{\boldsymbol{e}'_1,\ldots,\boldsymbol{e}'_r\}\rangle$. Базисы ортонормированные $\Longrightarrow$ $\langle\{\boldsymbol{e}_{r+1},\ldots,\boldsymbol{e}_n\}\rangle = \langle\{\boldsymbol{e}_1,\ldots,\boldsymbol{e}_r\}\rangle^\perp = \langle\{\boldsymbol{e}'_1,\ldots,\boldsymbol{e}'_r\}\rangle^\perp = \langle\{\boldsymbol{e}'_{r+1},\ldots,\boldsymbol{e}'_n\}\rangle$. Значит, матрица C перехода от $\boldsymbol{E}$ к $\boldsymbol{E}'$ блочно-диагональная (по определению этой матрицы) и состоит из блока C_1 размера $r\times r$ и блока C_2 размера $(n-r)\times(n-r)$, каждый из которых, очевидно, является ортогональной матрицей (поскольку сама матрица C ортогональна). По формуле преобразования выражения для аффинно-квадратичной функции (точнее, его линейной части) при переходе к новой системе координат имеем

$$\begin{pmatrix} 0\\ \vdots\\ 0\\ \tilde\lambda_{r+1}\\ 0\\ \vdots\\ 0 \end{pmatrix} = C^T\left(B\begin{pmatrix} \tilde o_1\\ \vdots\\ \tilde o_r\\ \tilde o_{r+1}\\ \tilde o_{r+2}\\ \vdots\\ \tilde o_n \end{pmatrix} + \begin{pmatrix} 0\\ \vdots\\ 0\\ \lambda_{r+1}\\ 0\\ \vdots\\ 0 \end{pmatrix}\right) = \begin{pmatrix} C_1 & 0\\ 0 & C_2 \end{pmatrix}^T\left(\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 & 0 & \ldots & 0\\ 0 & \lambda_2 & \ldots & 0 & 0 & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \ldots & \lambda_r & 0 & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}\begin{pmatrix} \tilde o_1\\ \vdots\\ \tilde o_r\\ \tilde o_{r+1}\\ \tilde o_{r+2}\\ \vdots\\ \tilde o_n \end{pmatrix} + \begin{pmatrix} 0\\ \vdots\\ 0\\ \lambda_{r+1}\\ 0\\ \vdots\\ 0 \end{pmatrix}\right) =$$

$$= \begin{pmatrix} C_1 & 0\\ 0 & C_2 \end{pmatrix}^T\left(\begin{pmatrix} \lambda_1\cdot\tilde o_1\\ \vdots\\ \lambda_2\cdot\tilde o_r\\ 0\\ 0\\ \vdots\\ 0 \end{pmatrix} + \begin{pmatrix} 0\\ \vdots\\ 0\\ \lambda_{r+1}\\ 0\\ \vdots\\ 0 \end{pmatrix}\right),$$

где $\tilde o_1,\ldots,\tilde o_n$ — координаты точки O' относительно репера $(O,\boldsymbol{E})$. Матрица C_2 ортогональна, $\begin{pmatrix} \tilde\lambda_{r+1}\\ 0\\ \vdots\\ 0 \end{pmatrix} = C_2^T\begin{pmatrix} \lambda_{r+1}\\ 0\\ \vdots\\ 0 \end{pmatrix} \Longrightarrow \lambda_{r+1} = \tilde\lambda_{r+1}$. □

Квадрики

Здесь опять $\mathbb{A}$ — n-мерное аффинное пространство, ассоциированное с векторным пространством V над полем K характеристики $\neq 2$.

Определение

Множество вида $\Gamma(Q) = \{X \in \mathbb{A} : Q(X) = 0\}$, где $Q \colon \mathbb{A} \to K$ — аффинно-квадратичная функция, называется квадрикой, или гиперповерхностью второго порядка (если только оно не пусто и не является плоскостью). При $n = 2$ квадрики называются кониками, или кривыми второго порядка. При $n = 3$ квадрики называются также поверхностями второго порядка.

Предложение

Любая прямая либо целиком лежит на квадрике, либо пересекается с ней в не более чем двух точках.

Доказательство. Рассмотрим любую прямую $\ell = A + \langle \boldsymbol{v} \rangle = \{A + t\boldsymbol{v} : t \in K\}$, проходящую через точку A квадрики. Точки пересечения прямой с квадрикой соответствуют значениям t, удовлетворяющим квадратному уравнению. □

Определение

Точка O называется **центром** квадрики, если эта квадрика симметрична относительно O, т.е. вместе со всякой точкой $O + \boldsymbol{x}$ $(\boldsymbol{x} \in V)$ содержит точку $O - \boldsymbol{x}$. Центр квадрики, лежащий на ней самой, называется её **вершиной**.

Предложение

Если O — вершина квадрики $\Gamma(Q)$ и $A \in \Gamma(Q)$, $A \neq O$, то вся прямая, проходящая через точки O и A, содержится в квадрике $\Gamma(Q)$.

Доказательство. Из предыдущего предложения. □

Предложение

Любая квадрика содержит точку, не являющуюся ее вершиной.

Доказательство. По теореме о геометрической характеризации плоскостей: если все вершины принадлежат квадрике, то это плоскость. □

Теорема

Если K — бесконечное поле характеристики $\neq 2$ и Q_1 и Q_2 — две аффинно-квадратичные функции $\mathbb{A} \to K$, для которых $\Gamma(Q_1)$ и $\Gamma(Q_2)$ — квадрики и $\Gamma(Q_1) = \Gamma(Q_2)$, то эти функции пропорциональны.

Доказательство. Положим $\Gamma = \Gamma(Q_1) = \Gamma(Q_2)$. Пусть $O \in \Gamma$ — не вершина. Тогда $Q_1(O) = Q_2(O) = 0$ и $\ell_1, \ell_2 \not\equiv 0$ в разложениях

$$Q_1(O + \boldsymbol{x}) = q_1(\boldsymbol{x}) + 2\ell_1(\boldsymbol{x}) \quad \text{и} \quad Q_2(O + \boldsymbol{x}) = q_2(\boldsymbol{x}) + 2\ell_2(\boldsymbol{x}), \qquad \boldsymbol{x} \in V.$$

Рассмотрим произвольную прямую $\pi = O + \langle \boldsymbol{v} \rangle$. Точка $O + t\boldsymbol{v}$ этой прямой принадлежит $\Gamma \iff t$ — решение уравнений $t^2 q_i(\boldsymbol{v}) + 2t l_i(\boldsymbol{v}) = 0$, $i = 1, 2$. Следовательно, эти уравнения должны иметь одинаковые решения (относительно t). Поэтому для всех векторов $\boldsymbol{v}$ таких, что $\ell_1(\boldsymbol{v}), \ell_2(\boldsymbol{v}) \neq 0$, имеем $q_1(\boldsymbol{v}) \cdot \ell_2(\boldsymbol{v}) = q_2(\boldsymbol{v}) \cdot \ell_1(\boldsymbol{v})$. Умножим на $\ell_1(\boldsymbol{v}) \cdot \ell_2(\boldsymbol{v})$:

$$q_1(\boldsymbol{v}) \cdot \ell_2(\boldsymbol{v}) \cdot \ell_1(\boldsymbol{v}) \cdot \ell_2(\boldsymbol{v}) = q_2(\boldsymbol{v}) \cdot \ell_1(\boldsymbol{v}) \cdot \ell_1(\boldsymbol{v}) \cdot \ell_2(\boldsymbol{v})$$

для всех $\boldsymbol{v} \in V$. Поле K бесконечно $\Longrightarrow$ $q_1 \cdot \ell_2 \cdot \ell_1 \cdot \ell_2 = q_2 \cdot \ell_1 \cdot \ell_1 \cdot \ell_2 \Longrightarrow$ $q_1 \cdot \ell_2 = q_2 \cdot \ell_1$ (в кольце многочленов нет делителей нуля).

Предположим, что линейные функции ℓ_1 и ℓ_2 не пропорциональны, а значит, линейно независимы. Дополним их до базиса $\boldsymbol{\mathcal{E}}^*$ пространства V^* и рассмотрим взаимный базис $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ пространства V. В этом базисе $\ell_1(\boldsymbol{x}) = x_1$ и $\ell_2(\boldsymbol{x}) = x_2$, где $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$, а равенство $q_1 \cdot \ell_2 = q_2 \cdot \ell_1$ записывается в виде

$$\sum_{i,j} b_{ij}^{(1)} x_i x_j x_2 = \sum_{i,j} b_{ij}^{(2)} x_i x_j x_1.$$

Видим, что q_1 делится на x_1, а q_2 делится на x_2, причём частные от деления равны, т.е.

$$q_1(\boldsymbol{x}) = \ell(\boldsymbol{x}) \cdot x_1 \quad \text{и} \quad q_2(\boldsymbol{x}) = \ell(\boldsymbol{x}) \cdot x_2,$$

где $\ell(\boldsymbol{x})$ — некоторая линейная функция. Следовательно,

$$Q_1(O + \boldsymbol{v}) = (\ell(\boldsymbol{v}) + 2) \cdot x_1 \quad \text{и} \quad Q_2(O + \boldsymbol{v}) = (\ell(\boldsymbol{v}) + 2) \cdot x_2.$$

$\Gamma = \Gamma(Q_1) \implies$ плоскость $\pi : x_1 = 0$ содержится в $\Gamma \implies Q_2 \equiv 0$ на $\pi \implies$ $\ell(\boldsymbol{x}) + 2 = 0$ для любого $\boldsymbol{x} \in V$ с $x_1 = 0$ и $x_2 \neq 0$. Противоречие (достаточно рассмотреть $\boldsymbol{x} = \boldsymbol{e}_2$ и $\boldsymbol{x} = 2\boldsymbol{e}_2$). $\square$

Следствие

В аффинном пространстве над бесконечным полем K характеристики $\neq 2$ центр любой квадрики $\Gamma(Q)$ является также и центром аффинно-квадратичной функции Q.

Доказательство. Если O — центр квадрики $\Gamma(Q)$, то $\Gamma(Q) = \Gamma(\tilde{Q})$, где $Q(O + \boldsymbol{x}) = \tilde{Q}(O - \boldsymbol{x})$. Следовательно, $\tilde{Q} = \lambda Q$ для некоторого $\lambda \in K$. Сравнивая члены второй степени в выражениях для Q и $\tilde{Q}$, видим, что $\lambda = 1$, т.е. $Q = \tilde{Q}$, а значит, $Q(O + \boldsymbol{x}) = Q(O - \boldsymbol{x})$. □

Классификация квадрик

Аффинная классификация квадрик

Теорема

В n-мерном аффинном пространстве $\mathbb{A}$ над бесконечным полем K характеристики $\neq 2$ уравнение любой квадрики выбором системы координат приводится к одному из видов

$$\begin{aligned}
\lambda_1 x_1^2 + \dots + \lambda_r x_r^2 &= 0, \qquad r \leqslant n,\\
\lambda_1 x_1^2 + \dots + \lambda_r x_r^2 &= 1, \qquad r \leqslant n,\\
\lambda_1 x_1^2 + \dots + \lambda_r x_r^2 &= 2x_{r+1}, \qquad r < n,
\end{aligned}$$

где $\lambda_1, \dots, \lambda_r \in K \setminus \{0\}$.

В случае $K = \mathbb{C}$ ($K = \mathbb{R}$) уравнение квадрики можно привести к виду, в котором $\lambda_i = 1$ для всех $i \leqslant r$ ($\lambda_i = \pm 1$ для всех $i \leqslant r$), причём этот вид определён однозначно.

Определение

Две квадрики Γ_1 и Γ_2 называются аффинно эквивалентными, если найдется невырожденный аффинный оператор (т.е. аффинное преобразование) $\Phi\colon \mathbb{A} \to \mathbb{A}$ такой, что $\Phi(\Gamma_1) = \Gamma_2$.

Теорема

В аффинном пространстве $\mathbb{A}$ над бесконечным полем K характеристики $\neq 2$ две аффинно-квадратичные функции $Q_1\colon \mathbb{A} \to K$ и $Q_2\colon \mathbb{A} \to K$ задают аффинно эквивалентные квадрики тогда и только тогда, когда выражения для этих функций относительно некоторых реперов $(O, \boldsymbol{E})$ и $(O', \boldsymbol{E}')$ совпадают с точностью до ненулевого скалярного множителя.

Доказательство. Пусть $\widehat{C}$ — матрица перехода от $(O, \boldsymbol{E})$ к $(O', \boldsymbol{E}')$, и пусть $\widehat{B}_1$ — матрица Q_1 в $(O, \boldsymbol{E})$ и $\widehat{B}_2$ — матрица Q_2 в $(O', \boldsymbol{E}')$, причём $\widehat{B}_1 = \alpha \widehat{B}_2$. Тогда матрица Q_1 относительно $(O', \boldsymbol{E}')$ равна $\widehat{C}^T\,\widehat{B}_1\,\widehat{C}$. Рассмотрим преобразование $\Phi\colon \mathbb{A} \to \mathbb{A}$ с матрицей C относительно $(O, \boldsymbol{E})$. Для точки X с координатами $x_1, \ldots, x_n$ относительно $(O', \boldsymbol{E}')$ точка $\Phi(X)$ имеет координаты $C\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, так что

$$X \in \Gamma_1 \iff Q_1(X) = 0 \iff (x_1, \ldots, x_n, 1)\,\widehat{C}^T\,\widehat{B}_1\,\widehat{C}\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix} = 0 \iff \frac{1}{\alpha} Q_2(\Phi(X)) = 0.$$

Обратно, пусть Φ — аффинное преобразование с матрицей $\widehat{C}$ относительно репера $(O, \boldsymbol{E})$, $\Phi(\Gamma_1) = \Gamma_2$ и $\widehat{B}_1$ и $\widehat{B}_2$ — матрицы Q_1 и Q_2 относительно $(O, \boldsymbol{E})$. Пусть X — точка с координатами $x_1, \ldots, x_n$ относительно $(O, \boldsymbol{E})$ и $y_1, \ldots, y_n$ — координаты точки $\Phi(X) = Y$. Тогда

$$X \in \Gamma_1 \iff (x_1, \ldots, x_n, 1)\, \widehat{B}_1 \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix} = 0$$

и

$$\Phi(X) = Y \in \Gamma_2 \iff (y_1, \ldots, y_n, 1)\, \widehat{B}_2 \begin{pmatrix} y_1 \\ \vdots \\ y_n \\ 1 \end{pmatrix} = 0 \iff (x_1, \ldots, x_n, 1)\, C^T\, \widehat{B}_1\, C \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix}.$$

Из равносильности условий $X \in \Gamma_1$ и $\Phi(X) \in \Gamma_2$ следует, что $\widehat{B}_1 = \alpha \widehat{C}^T\, \widehat{B}_2\, \widehat{C}$. □

Следствие

В n-мерном вещественном аффинном пространстве $\mathbb{A}$ любая квадрика аффинно эквивалентна ровно одной из квадрик

$$\begin{aligned}
x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_{k+l}^2 &= 1, \qquad k \geqslant 0,\ \ l > 0,\ \ k + l \leqslant n,\\
x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_{k+l}^2 &= 0, \qquad 0 \leqslant l \leqslant k,\ k + l \leqslant n,\\
x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_{k+l}^2 &= 2x_{k+l+1}, \qquad 0 \leqslant l \leqslant k,\ k + l \leqslant n.
\end{aligned}$$

Метрическая классификация квадрик

Теорема

В n-мерном аффинно-евклидовом пространстве уравнение любой квадрики выбором прямоугольной системы координат приводится к одному из видов

$$\begin{aligned} \lambda_1 x_1^2 + \dots + \lambda_r x_r^2 &= 1, && r \leqslant n, \\ \lambda_1 x_1^2 + \dots + \lambda_r x_r^2 &= 0, && r \leqslant n, \\ \lambda_1 x_1^2 + \dots + \lambda_r x_r^2 &= 2x_{r+1}, && r < n, \end{aligned}$$

где $\lambda_1, \dots, \lambda_r \in K \setminus \{0\}$.

Для первого вида числа $\lambda_1, \dots, \lambda_r$ определены однозначно с точностью до перестановки. Для второго вида числа $\lambda_1, \dots, \lambda_r$ определены однозначно с точностью до перестановки и одновременного умножения на ненулевое число. Для третьего вида числа $\lambda_1, \dots, \lambda_r$ определены однозначно с точностью до перестановки и одновременного умножения на -1.

Определение

Две квадрики Γ_1 и Γ_2 в аффинно-евклидовом пространстве называются метрически эквивалентными, если найдется ортогональный аффинный оператор (т.е. изометрическое преобразование) $\Phi\colon \mathbb{A} \to \mathbb{A}$ такой, что $\Phi(\Gamma_1) = \Gamma_2$.

Теорема

Если в аффинно-евклидовом пространстве $\mathbb{A}$ две аффинно-квадратичные функции $Q_1\colon \mathbb{A} \to K$ и $Q_2\colon \mathbb{A} \to K$ задают квадрики, то эти квадрики метрически эквивалентны тогда и только тогда, когда выражения для функций Q_1 и Q_2 относительно некоторых ортонормированных реперов $(O, \boldsymbol{E})$ и $(O', \boldsymbol{E}')$ совпадают с точностью до ненулевого скалярного множителя.

Определение

Пусть $V_1, \ldots, V_k, V$ — векторные пространства над одним и тем же полем. Отображение $V_1 \times \cdots \times V_k \to V$ называется полилинейной функцией, если оно линейно по каждому из k аргументов при фиксированных значениях всех прочих аргументов.

Всюду ниже V — n-мерное векторное пространство над полем K, а p и q — неотрицательные целые числа.

Определение

Всякая полилинейная функция $\boldsymbol{t}\colon \underbrace{V \times \cdots \times V}_{p \text{ раз}} \times \underbrace{V^* \times \cdots \times V^*}_{q \text{ раз}} \to K$

называется тензором типа (p, q) и валентности (или ранга) $p + q$ на V.

Про тензор типа (p, q) говорят, что он p раз ковариантный и q раз контравариантный. Тензор типа $(p, 0)$ (типа $(0, q)$) называется просто ковариантным (контравариантным).

Предупреждение: В литературе p и q в обозначении типа часто указываются в обратном порядке.

Примеры

1. Тензор типа $(0,0)$ — скаляр.
2. Тензор типа $(1,0)$ — линейная функция $V \to K$, т.е. элемент V^* (ковектор).
3. Тензор типа $(0,1)$ — линейная функция $V^* \to K$, т.е. элемент $V^{**} = V$ (вектор).
4. Тензор типа $(2,0)$ (типа $(0,2)$) — билинейная функция на V (на V^*).
5. Тензор типа $(1,1)$ — билинейная функция $b \colon V \times V^* \to K$. Её левая корреляция $b_{\text{лев}} \colon V \to V^{**} = V$ каждому вектору $\boldsymbol{v}$ ставит в соответствие функционал $\boldsymbol{\varphi_v}$ на V^*, определённый правилом $\boldsymbol{\varphi_v}(\boldsymbol{f}) = b(\boldsymbol{v}, \boldsymbol{f})$ для $\boldsymbol{f} \in V^*$. В силу изоморфизма $V^{**} \rightleftarrows V$ функционал $\boldsymbol{\varphi_v} \in V^{**}$ есть функция вычисления $e_{\boldsymbol{x_v}}$ для некоторого вектора $\boldsymbol{x_v} \in V$. Соответствие $\boldsymbol{v} \mapsto \boldsymbol{x_v}$ — линейный оператор $V \to V$.

   Обратно, пусть имеется линейный оператор $\mathcal{A} \colon V \to V$. Определим тензор $\boldsymbol{t}$ типа $(1,1)$ правилом $\boldsymbol{t}(\boldsymbol{x}, \boldsymbol{f}) = \boldsymbol{f}(\mathcal{A}\boldsymbol{x}) = e_{\mathcal{A}\boldsymbol{x}}(\boldsymbol{f})$. Его левая корреляция $\boldsymbol{x} \mapsto e_{\mathcal{A}\boldsymbol{x}} = \mathcal{A}\boldsymbol{x}$ — оператор $\mathcal{A}$.

   Таким образом, между линейным пространством всех билинейных функций $V \times V^* \to K$ (тензоров типа $(1,1)$) и линейным пространством всех линейных операторов $V \to V$ имеется естественный изоморфизм, так что тензоры типа $(1,1)$ можно идентифицировать с операторами.

Координаты тензора

Соглашения об индексации:

- векторы нумеруются нижними индексами, ковекторы (функционалы) — верхними;
- координаты векторов нумеруются верхними индексами, ковекторов — нижними;
- если один и тот же индекс встречается в формуле снизу и сверху, то по нему производится суммирование от 1 до n.

Пусть $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ — базис в V и $\boldsymbol{\mathcal{E}} = \{\boldsymbol{\varepsilon}^1, \dots, \boldsymbol{\varepsilon}^n\}$ — взаимный базис в V^*. Любая полилинейная функция полностью определяется своими координатами — значениями на наборах базисных векторов.

Примеры

1. Пусть $\boldsymbol{t}(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{f})$ — тензор типа $(2, 1)$, и пусть $\boldsymbol{x} = x^i \boldsymbol{e}_i$, $\boldsymbol{y} = y^j \boldsymbol{e}_j$, $\boldsymbol{f} = f_k \boldsymbol{\varepsilon}^k$. Тогда

$$\boldsymbol{t}(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{f}) = \boldsymbol{t}(x^i \boldsymbol{e}_i, y^j \boldsymbol{e}_j, f_k \boldsymbol{\varepsilon}^k) = x^i \cdot y^j \cdot f_k \cdot \boldsymbol{t}(\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{\varepsilon}^k) = x^i \cdot y^j \cdot f_k \cdot t_{ij}^k,$$

   где $t_{ij}^k = \boldsymbol{t}(\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{\varepsilon}^k)$ для $i, j, k = 1, \ldots, n$ — *координаты* тензора $\boldsymbol{t}$ в базисе $\boldsymbol{E}$.

2. Пусть $\mathcal{A} \colon V \to V$ — линейный оператор с матрицей $A = (a_i^j)$, т.е. $\mathcal{A}\boldsymbol{e}_i = a_i^j \boldsymbol{e}_j$. Тогда $\mathcal{A}$ можно трактовать как тензор $\boldsymbol{t} \colon V \times V^* \to K$ (типа (1,1)), определённый правилом $\boldsymbol{t}(\boldsymbol{x}, \boldsymbol{f}) = \boldsymbol{f}(\mathcal{A}\boldsymbol{x})$. На паре базисных векторов $(\boldsymbol{e}_k, \boldsymbol{\varepsilon}^m)$ он принимает значение

$$\boldsymbol{\varepsilon}^m(\mathcal{A}\boldsymbol{e}_k) = \boldsymbol{\varepsilon}^m(a_k^j \boldsymbol{e}_j) = a_k^m.$$

   Таким образом, координаты тензора $\boldsymbol{t}$ — это в точности элементы матрицы A.

Для полилинейной функции общего вида

$$\boldsymbol{t}(\boldsymbol{x}_{i_1},\dots,\boldsymbol{x}_{i_p};\boldsymbol{f}^{j_1},\dots,\boldsymbol{f}^{j_q}) = \boldsymbol{t}(x^{k_1}_{i_1}\boldsymbol{e}_{k_1},\dots,x^{k_p}_{i_p}\boldsymbol{e}_{k_p};f^{j_1}_{m_1}\boldsymbol{\varepsilon}^{m_1},\dots,f^{j_q}_{m_q}\boldsymbol{\varepsilon}^{m_q}) = \\ = x^{k_1}_{i_1}\cdot\ldots\cdot x^{k_p}_{i_p}\cdot f^{j_1}_{m_1}\cdot\ldots\cdot f^{j_q}_{m_q}\cdot t^{m_1\dots m_q}_{k_1\dots k_p},$$

где $t^{m_1\dots m_q}_{k_1\dots k_p} = \boldsymbol{t}(\boldsymbol{e}_{k_1},\dots,\boldsymbol{e}_{k_p};\boldsymbol{\varepsilon}^{m_1},\dots,\boldsymbol{\varepsilon}^{m_q})$. Числа $t^{m_1\dots m_q}_{k_1\dots k_p}$ называются координатами (а также коэффициентами или компонентами) тензора $\boldsymbol{t}$ в базисе $\boldsymbol{E}$.

Пусть $\tilde{\boldsymbol{E}} = \{\tilde{\boldsymbol{e}}_1,\dots,\tilde{\boldsymbol{e}}_n\}$ и $\tilde{\boldsymbol{\mathcal{E}}} = \{\tilde{\boldsymbol{\varepsilon}}^1,\dots,\tilde{\boldsymbol{\varepsilon}}^n\}$ — другой базис в V и взаимный с ним базис в V^*, и пусть $C = (c^j_i)$ — матрица перехода от $\boldsymbol{E}$ к $\tilde{\boldsymbol{E}}$, т.е. $\tilde{\boldsymbol{e}}_i = c^j_i\boldsymbol{e}_j$, и $S = (s^j_i) = (C^{-1})^T$ — матрица перехода от $\boldsymbol{\mathcal{E}}$ к $\tilde{\boldsymbol{\mathcal{E}}}$, т.е. $\tilde{\boldsymbol{\varepsilon}}^i = s^i_j\boldsymbol{\varepsilon}^j$ (заметьте, что в первом случае номерам строк соответствуют верхние индексы, а во втором — нижние). Тогда

$$\tilde{t}^{m_1\dots m_q}_{k_1\dots k_p} = \boldsymbol{t}(\tilde{\boldsymbol{e}}_{k_1},\dots,\tilde{\boldsymbol{e}}_{k_p};\tilde{\boldsymbol{\varepsilon}}^{m_1},\dots,\tilde{\boldsymbol{\varepsilon}}^{m_q}) = \boldsymbol{t}(c^{i_1}_{k_1}\boldsymbol{e}_{i_1},\dots,c^{i_p}_{k_p}\boldsymbol{e}_{i_p};s^{m_1}_{j_1}\boldsymbol{\varepsilon}^{j_1},\dots,s^{m_q}_{j_q}\boldsymbol{\varepsilon}^{j_q}) = \\ = c^{i_1}_{k_1}\cdot\ldots\cdot c^{i_p}_{k_p}\cdot s^{m_1}_{j_1}\cdot\ldots\cdot s^{m_q}_{j_q}\cdot t^{j_1\dots j_q}_{i_1\dots i_p}. \qquad (*)$$

Говорят, что *на пространстве V задан тензор типа (p,q), если каждому базису пространства V поставлена в соответствие система n^{p+q} чисел $t^{m_1\dots m_q}_{k_1\dots k_p}$ таким образом, что системы, отвечающие разным базисам, связаны соотношением* $(*)$.

В дальнейшем мы считаем, что в пространстве V зафиксирован базис $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$, а в пространстве V^* — взаимный базис $\boldsymbol{\mathcal{E}} = \{\boldsymbol{\varepsilon}^1, \ldots, \boldsymbol{\varepsilon}^n\}$. Элементы векторных, сопряжённых и т.п. пространств обозначаются жирным шрифтом (возможно, с индексами, тильдами и пр.) и нумеруются в соответствии с соглашениями об индексации (когда оно применимо) — см., например, нумерацию элементов базисов выше. Их координаты обозначаются теми же буквами с теми же пометками, но светлым шрифтом, и индексируются в соответствии с теми же соглашениями. Например, рассматривая вектор $\boldsymbol{x} \in V$, мы будем подразумевать без пояснений, что $x^1, \ldots, x^n$ — его координаты в базисе $\boldsymbol{E}$.

Кроме того, как и выше, всегда подразумевается суммирование по одинаковым верхнему и нижнему индексам.

Арифметические операции над тензорами

- *Сложение* определяется поточечно для тензоров одного типа:

  $$(\boldsymbol{t}+\boldsymbol{s})(\boldsymbol{x}_1,\dots,\boldsymbol{x}_p;\boldsymbol{f}^1,\dots,\boldsymbol{f}^q)=\boldsymbol{t}(\boldsymbol{x}_1,\dots,\boldsymbol{x}_p;\boldsymbol{f}^1,\dots,\boldsymbol{f}^q)+\boldsymbol{s}(\boldsymbol{x}_1,\dots,\boldsymbol{x}_p;\boldsymbol{f}^1,\dots,\boldsymbol{f}^q).$$

  При сложении координаты тензоров складываются.
- *Умножение на скаляр* определяется поточечно. При умножении на скаляр координаты умножаются на тот же скаляр.

Таким образом, множество $T_p^q(V)$ всех тензоров типа (p,q) на V образует векторное пространство над полем K.

- *Умножение тензоров*: тензорное произведение тензоров $\boldsymbol{t}$ типа (p,q) и $\boldsymbol{t}'$ типа (p',q') — это тензор $\boldsymbol{t}\otimes\boldsymbol{s}$ типа $(p+p',q+q')$, определённый правилом

  $$\begin{gathered}(\boldsymbol{t}\otimes\boldsymbol{t}')(\boldsymbol{x}_1,\dots,\boldsymbol{x}_p,\boldsymbol{x}_{p+1},\dots,\boldsymbol{x}_{p+p'};\boldsymbol{f}^1,\dots,\boldsymbol{f}^q,\boldsymbol{f}^{q+1},\dots,\boldsymbol{f}^{q+q'})=\\ =\boldsymbol{t}(\boldsymbol{x}_1,\dots,\boldsymbol{x}_p;\boldsymbol{f}^1,\dots,\boldsymbol{f}^q)\cdot\boldsymbol{t}'(\boldsymbol{x}_{p+1},\dots,\boldsymbol{x}_{p+p'};\boldsymbol{f}^{q+1},\dots,\boldsymbol{f}^{q+q'}).\end{gathered}$$

  Его координаты — произведения соответствующих координат сомножителей. Умножение тензоров не коммутативно, но ассоциативно и дистрибутивно.

Примеры

1. Пусть $\boldsymbol{f}$ и $\boldsymbol{g}$ — тензоры типа $(1,0)$ на V (т.е. линейные функции из V^*) с координатами f_i и g_i в базисе $\boldsymbol{E}$ (это означает, что $\boldsymbol{f} = f_i\boldsymbol{\varepsilon}^i$ и $\boldsymbol{g} = g_i\boldsymbol{\varepsilon}^i$, т.е. $f_i = \boldsymbol{f}(\boldsymbol{e}_i)$ и $g_i = \boldsymbol{g}(\boldsymbol{e}_i)$). Тогда $\boldsymbol{f} \otimes \boldsymbol{g}$ — тензор типа $(2,0)$ (т.е. билинейная функция на V), определённый правилом
$$(\boldsymbol{f} \otimes \boldsymbol{g})(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{x}) \cdot \boldsymbol{g}(\boldsymbol{y})$$
(не все билинейные функции можно так разложить). Его координаты суть $t_{ij} = \boldsymbol{f}(\boldsymbol{e}_i) \cdot \boldsymbol{g}(\boldsymbol{e}_j) = f_i \cdot g_j$.
2. Пусть $\boldsymbol{f}$ — тензор типа $(1,0)$ и $\boldsymbol{v} = e_{\boldsymbol{v}}$ — тензор типа $(0,1)$. Тогда
$$(\boldsymbol{f} \otimes \boldsymbol{v})(\boldsymbol{x}; \boldsymbol{g}) = \boldsymbol{f}(\boldsymbol{x}) \cdot e_{\boldsymbol{v}}(\boldsymbol{g}) = e_{\boldsymbol{f}(\boldsymbol{x})\boldsymbol{v}}(\boldsymbol{g}).$$
Координаты этого тензора суть $(\boldsymbol{f} \otimes \boldsymbol{v})(\boldsymbol{e}_i; \boldsymbol{\varepsilon}^j) = f_i \cdot v^j$. Ему соответствует оператор $\boldsymbol{x} \mapsto \boldsymbol{f}(\boldsymbol{x})\boldsymbol{v}$ с матрицей $(f_i \cdot v^j) = \begin{pmatrix} f_1 \cdot v^1 & \dots & f_n \cdot v^1 \\ \vdots & \ddots & \vdots \\ f_1 \cdot v^n & \dots & f_n \cdot v^n \end{pmatrix}$.

Тензор, являющийся тензорным произведением тензоров ранга 1, называется **разложимым**. Всякий тензор можно представить как сумму разложимых.

Тензорное произведение векторных пространств

Теорема

Пусть V — векторное пространство над полем K с базисом $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$, и пусть $\boldsymbol{\mathcal{E}} = \{\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n\}$ — взаимный базис сопряжённого пространства.

Тензоры типа (p, q) на V составляют векторное пространство $T_p^q = T_p^q(V)$ над K размерности n^{p+q}. Векторы

$$\boldsymbol{\varepsilon}^{i_1} \otimes \cdots \otimes \boldsymbol{\varepsilon}^{i_p} \otimes \boldsymbol{e}_{j_1} \otimes \cdots \otimes \boldsymbol{e}_{j_q},$$

где $(i_1, \ldots, i_p)$ и $(j_1, \ldots, j_q)$ — всевозможные наборы натуральных чисел, не превосходящих p и q соответственно, образуют базис этого пространства.

Доказательство. Рассмотрим произвольный тензор $\boldsymbol{t}$ типа (p, q) с компонентами $t_{k_1 \ldots k_p}^{m_1 \ldots m_q} = \boldsymbol{t}(\boldsymbol{e}_{k_1}, \ldots, \boldsymbol{e}_{k_p}; \boldsymbol{\varepsilon}^{m_1}, \ldots, \boldsymbol{\varepsilon}^{m_q})$. Положим

$$\boldsymbol{t}' = t_{k_1 \ldots k_p}^{m_1 \ldots m_q} \boldsymbol{\varepsilon}^{k_1} \otimes \cdots \otimes \boldsymbol{\varepsilon}^{k_p} \otimes \boldsymbol{e}_{m_1} \otimes \cdots \otimes \boldsymbol{e}_{m_q}.$$

Поскольку $\boldsymbol{\varepsilon}^i(\boldsymbol{e}_j) = \boldsymbol{e}_i(\boldsymbol{\varepsilon}^j) = \delta_{ij}$, имеем

$$t'^{m_1 \ldots m_q}_{k_1 \ldots k_p} = \boldsymbol{t}'(\boldsymbol{e}_{k_1}, \ldots, \boldsymbol{e}_{k_p}; \boldsymbol{\varepsilon}^{m_1}, \ldots, \boldsymbol{\varepsilon}^{m_q}) = t_{i_1 \ldots i_p}^{j_1 \ldots j_q} \delta_{i_1}^{k_1} \cdot \ldots \cdot \delta_{i_p}^{k_p} \cdot \delta_{m_1}^{j_1} \cdot \ldots \cdot \delta_{m_q}^{j_q} = t_{k_1 \ldots k_p}^{m_1 \ldots m_q}$$

(здесь $\delta_i^j = \delta_j^i = \delta_{ij}$). Значит, $\boldsymbol{t}' = \boldsymbol{t}$, так что указанная система векторов полна.

Лемма (о линейной независимости)

Пусть V — любое векторное пространство и $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k \in V$. Если для всякого $i \leqslant k$ существует линейная функция F_i на V такая, что $F_i(\boldsymbol{v}_i) = 1$ и $F_i(\boldsymbol{v}_j) = 0$ для $j \neq i$, то векторы $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ линейно независимы.

Продолжение доказательства теоремы. Мы видели, что каждый тензор $\boldsymbol{t} \in T^q_p$ определяется своими компонентами $t^{m_1 \dots m_q}_{k_1 \dots k_p}$ по формуле $\boldsymbol{t} = t^{m_1 \dots m_q}_{k_1 \dots k_p} \boldsymbol{\varepsilon}^{k_1} \otimes \dots \otimes \boldsymbol{\varepsilon}^{k_p} \otimes \boldsymbol{e}_{m_1} \otimes \dots \otimes \boldsymbol{e}_{m_q}$. При сложении тензоров их компоненты складываются, а при умножении на скаляр умножаются на тот же скаляр. Значит, для фиксированных наборов индексов $k_1, \dots, k_p$ и $m_1, \dots, m_q$ функция $F^{m_1 \dots m_q}_{k_1 \dots k_p} \colon T^q_p \to K$, которая каждому тензору $\boldsymbol{t}$ ставит в соответствие $t^{m_1 \dots m_q}_{k_1 \dots k_p}$, линейна. При этом $F^{m_1 \dots m_q}_{k_1 \dots k_p}$ принимает значение 1 на векторе $\boldsymbol{\varepsilon}^{k_1} \otimes \dots \otimes \boldsymbol{\varepsilon}^{k_p} \otimes \boldsymbol{e}_{m_1} \otimes \dots \otimes \boldsymbol{e}_{m_q}$ и 0 на всех векторах такого вида, соответствующих другим наборам индексов. По лемме эти векторы линейно независимы. $\square$

Пространство T^q_p называется тензорным произведением p пространств V^* и q пространств V и обозначается $\underbrace{V^* \otimes \dots \otimes V^*}_{p \text{ раз}} \otimes \underbrace{V \otimes \dots \otimes V}_{q \text{ раз}}$ или $(V^*)^{\otimes p} \otimes V^{\otimes q}$.

Универсальное свойство тензорного произведения $(V^*)^{\otimes p} \otimes V^{\otimes q}$

Отображение

$$\otimes\colon \underbrace{V^* \times \dots \times V^*}_{p \text{ раз}} \times \underbrace{V \times \dots \times V}_{q \text{ раз}} \to \underbrace{V^* \otimes \dots \otimes V^*}_{p \text{ раз}} \otimes \underbrace{V \otimes \dots \otimes V}_{q \text{ раз}},$$

$$\otimes(\boldsymbol{f}_1, \dots, \boldsymbol{f}_p; \boldsymbol{x}_1, \dots, \boldsymbol{x}_q) = \boldsymbol{f}_1 \otimes \dots \otimes \boldsymbol{f}_p \otimes \boldsymbol{x}_1 \otimes \dots \otimes \boldsymbol{x}_q,$$

полилинейно. Для любой полилинейной функции $\varphi\colon V^* \times \dots \times V^* \times V \times \dots \times V \to K$ существует единственный линейный функционал $\ell_\varphi\colon V^* \otimes \dots \otimes V^* \otimes V \otimes \dots \otimes V \to K$, для которого диаграмма

$$\begin{tikzcd}
V^* \times \dots \times V^* \times V \times \dots \times V \arrow[d, "\otimes"'] \arrow[dr, "\varphi"] & \\
V^* \otimes \dots \otimes V^* \otimes V \otimes \dots \otimes V \arrow[r, "\ell_\varphi"'] & K
\end{tikzcd}$$

коммутативна, т.е. $\varphi = \ell_\varphi \circ \otimes$. Он определяется на элементах базиса формулой

$$\ell_\varphi(\boldsymbol{\varepsilon}^{k_1} \otimes \dots \otimes \boldsymbol{\varepsilon}^{k_p} \otimes \boldsymbol{e}_{m_1} \otimes \dots \otimes \boldsymbol{e}_{m_q}) = \varphi(\boldsymbol{\varepsilon}^{k_1}, \dots, \boldsymbol{\varepsilon}^{k_p}, \boldsymbol{e}_{m_1}, \dots, \boldsymbol{e}_{m_q})$$

и продолжается на всё пространство $(V^*)^{\otimes p} \otimes V^{\otimes q}$ по линейности.

Следовательно, $(V^* \otimes \dots \otimes V^* \otimes V \otimes \dots \otimes V)^* = V \otimes \dots \otimes V \otimes V^* \otimes \dots \otimes V^*$. Таким образом, линейные функционалы на пространстве $T_p^q = (V^*)^{\otimes p} \otimes V^{\otimes q}$ тензоров типа (p, q) — это в точности тензоры типа (q, p).

Общее определение тензорного произведения

Тензорное произведение $(V^*)^{\otimes p} \otimes V^{\otimes q}$ обладает тем свойством, что на $(V^*)^p \times V^q$ определено полилинейное умножение $\otimes\colon (V^*)^p \times V^q \to (V^*)^{\otimes p} \otimes V^{\otimes q}$, которое переводит наборы базисных элементов в базис пространства $(V^*)^{\otimes p} \otimes V^{\otimes q}$. Это свойство можно принять за определение тензорного произведения.

Определение

Пусть $V_1, \dots, V_k$ — любые векторные пространства над K, и пусть $\boldsymbol{E}_1, \dots, \boldsymbol{E}_k$ — их базисы. Пространство T вместе с полилинейным отображением

$$\otimes\colon V_1, \dots, V_k \to T$$

с тем свойством, что $\otimes(\boldsymbol{E}_1 \times \dots \times \boldsymbol{E}_k)$ — базис в T, называется тензорным произведением пространств $V_1, \dots, V_k$ и обозначается $V_1 \otimes \dots \otimes V_k$. Образ $\otimes(\boldsymbol{x}_1, \dots, \boldsymbol{x}_k)$ произвольного набора векторов $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_k) \in V_1 \otimes \dots \otimes V_k$ обозначается $\boldsymbol{x}_1 \otimes \dots \otimes \boldsymbol{x}_k$.

Тензорное произведение существует: достаточно взять векторное пространство T с базисом $\boldsymbol{E}_1 \times \dots \times \boldsymbol{E}_k$ (рассматриваемым просто как абстрактное множество, так что элементы T не связаны с элементами $V_1 \times \dots \times V_k$) и определить полилинейное отображение $\otimes\colon V_1, \dots, V_k \to T$, задав его на наборах базисных векторов формулой $\otimes(\boldsymbol{e}_1, \dots, \boldsymbol{e}_k) = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_k) \in T$ и продолжив по полилинейности.

Определение тензорного произведения не зависит от выбора базисов $\boldsymbol{E}_1, \dots, \boldsymbol{E}_k$

Действительно, пусть $\boldsymbol{E}_i = \{\boldsymbol{e}_{i,\alpha} : \alpha \in A_i\}$, и пусть $\boldsymbol{E}'_1 = \{\boldsymbol{e}'_\beta : \beta \in B\}$ — любой базис пространства V_1. Каждый вектор из T единственным образом представляется как сумма векторов вида $\lambda \boldsymbol{e}_{1,\alpha_1} \otimes \boldsymbol{e}_{2,\alpha_2} \otimes \dots \otimes \boldsymbol{e}_{k,\alpha_k}$, потому что операция $\otimes$ полилинейна и $\boldsymbol{E}_1 \times \boldsymbol{E}_2 \times \dots \times \boldsymbol{E}_k$ — базис. Каждый вектор из $\boldsymbol{E}_1$ линейно выражается единственным образом через элементы $\boldsymbol{E}'_1$, и в силу линейности отображения $\otimes$ по первому аргументу каждый элемент вида $\boldsymbol{e}_{1,\alpha_1} \otimes \boldsymbol{e}_{2,\tilde{\alpha}_2} \otimes \dots \otimes \boldsymbol{e}_{k,\tilde{\alpha}_k}$, где $\tilde{\alpha}_i \in A_i$, линейно выражается единственным образом через элементы $\boldsymbol{e}'_\beta \otimes \boldsymbol{e}_{2,\tilde{\alpha}_2} \otimes \dots \otimes \boldsymbol{e}_{k,\tilde{\alpha}_k}$ (с теми же $\tilde{\alpha}_2, \dots, \tilde{\alpha}_k$). Значит, каждый элемент пространства T единственным образом представляется как сумма элементов вида $\lambda \boldsymbol{e}'_\beta \otimes \boldsymbol{e}_{2,\alpha_2} \otimes \dots \boldsymbol{e}_{k,\alpha_k} = \lambda(\boldsymbol{e}'_\beta \otimes \boldsymbol{e}_{2,\alpha_2} \otimes \dots \boldsymbol{e}_{k,\alpha_k})$, т.е. как линейная комбинация элементов вида $\boldsymbol{e}'_\beta \otimes \boldsymbol{e}_{2,\alpha_2} \otimes \dots \boldsymbol{e}_{k,\alpha_k} \in \otimes(\boldsymbol{E}'_1 \times \boldsymbol{E}_2 \times \dots \times \boldsymbol{E}_k)$.

Точно так же доказывается, что при замене базисов $\boldsymbol{E}_i$ других сомножителей любыми другими базисами $\boldsymbol{E}'_i$ образ $\otimes(\boldsymbol{E}'_1 \times \dots \times \boldsymbol{E}'_k)$ остаётся базисом в T. $\square$

Тензорное произведение единственно

Для любого другого тензорного произведения $(\tilde{T}, \tilde{\otimes})$ пространств $V_1, \dots, V_k$ существует (единственный) изоморфизм $\varphi \colon \tilde{T} \to T$, удовлетворяющий условию $\varphi(\boldsymbol{x}_1 \tilde{\otimes} \boldsymbol{x}_2 \tilde{\otimes} \dots \tilde{\otimes} \boldsymbol{x}_k) = \boldsymbol{x}_1 \otimes \boldsymbol{x}_2 \otimes \dots \otimes \boldsymbol{x}_k$ для любых $\boldsymbol{x}_i \in V_i$: на базисе он определён правилом $\varphi(\boldsymbol{e}_{1,\alpha_1} \tilde{\otimes} \dots \tilde{\otimes} \boldsymbol{e}_{k,\alpha_k}) = (\boldsymbol{e}_{1,\alpha_1} \otimes \dots \otimes \boldsymbol{e}_{k,\alpha_k})$.

Из единственности тензорного произведения следует, что пространство $T^q_p(V)$ для конечномерного пространства V, определённое в начале раздела, — не что иное как тензорное произведение

$$(\underbrace{V\otimes\cdots\otimes V}_{p \text{ раз}}\otimes\underbrace{V^*\otimes\cdots\otimes V^*}_{q \text{ раз}})^* = \underbrace{V^*\otimes\cdots\otimes V^*}_{p \text{ раз}}\times\underbrace{V\otimes\cdots\otimes V}_{q \text{ раз}}$$

в общем смысле, так что все введённые в начале раздела определения и обозначения согласуются с общей теорией.

Тензорное произведение как факторпространство

Пусть V и W — любые векторные пространства над одним и тем же полем K. Обозначим ноль и сложение в V символами $\mathbf{0}_V$ и $\underset{V}{+}$, а ноль и сложение в W — символами $\mathbf{0}_W$ и $\underset{W}{+}$.

Рассмотрим векторное пространство U над K с базисом $V \times W$. Его элементы — $\mathbf{0}$ и всевозможные формальные линейные комбинации

$$\lambda_1(\boldsymbol{v}_1, \boldsymbol{w}_1) + \lambda_2(\boldsymbol{v}_2, \boldsymbol{w}_2) + \ldots + \lambda_k(\boldsymbol{v}_k, \boldsymbol{w}_k),$$

т.е. просто неупорядоченные наборы $\{\lambda_1(\boldsymbol{v}_1, \boldsymbol{w}_1), \lambda_2(\boldsymbol{v}_2, \boldsymbol{w}_2), \ldots, \lambda_k(\boldsymbol{v}_k, \boldsymbol{w}_k)\}$, где $k \in \mathbb{N}$ и для $i \leqslant k$ $\lambda_i \in K \setminus \{0\}$, $\boldsymbol{v}_i \in V$ и $\boldsymbol{w}_i \in W$, причём $(\boldsymbol{v}_i, \boldsymbol{w}_i) \neq (\boldsymbol{v}_j, \boldsymbol{w}_j)$ для $i \neq j$. Сложение $+$ здесь не имеет никакого отношения к операциям в V и W; пары вида $(\boldsymbol{v}, \boldsymbol{w})$ — просто элементы базиса, никак не связанные друг другом в пространстве U. Пространство U можно представить как векторное пространство функций $V \times W \to K$, принимающих ненулевые значения λ лишь в конечном числе точек $(\boldsymbol{v}, \boldsymbol{w})$. Выписанная выше формальная линейная комбинация соответствует функции, которая принимает значения λ_i в точках $(\boldsymbol{v}_i, \boldsymbol{w}_i)$ для $i \leqslant k$, а во всех остальных точках равна 0. Функции складываются как обычно (поточечно), и сложение формальных линейных комбинаций определяется соответственно.

Пусть R — подпространство пространства U, порождённое левыми частями *определяющих соотношений* тензорного произведения:

1. $(\boldsymbol{v}' \underset{V}{+} \boldsymbol{v}'', \boldsymbol{w}) - \big((\boldsymbol{v}', \boldsymbol{w}) + (\boldsymbol{v}'', \boldsymbol{w})\big) = \mathbf{0}$ (здесь $\underset{V}{+}$ — суммирование в V)
2. $(\boldsymbol{v}, \boldsymbol{w}' \underset{W}{+} \boldsymbol{w}'') - \big((\boldsymbol{v}, \boldsymbol{w}') + (\boldsymbol{v}, \boldsymbol{w}'')\big) = \mathbf{0}$ (здесь $\underset{W}{+}$ — суммирование в W)
3. $(\lambda \boldsymbol{v}, \boldsymbol{w}) - (\boldsymbol{v}, \lambda \boldsymbol{w}) = \mathbf{0}$
4. $\lambda(\boldsymbol{v}, \boldsymbol{w}) - (\lambda \boldsymbol{v}, \boldsymbol{w}) = \mathbf{0}$

для всех $\boldsymbol{v}, \boldsymbol{v}' \in V$, $\boldsymbol{w}, \boldsymbol{w}' \in W$ и $\lambda \in K$.

▸ унив. св-во

Тензорное произведение $V \otimes W$ — это факторпространство U/R. Образ элемента $(\boldsymbol{v}, \boldsymbol{w})$ при естественном отображении $U \to U/R = V \otimes W$ обозначается $\boldsymbol{v} \otimes \boldsymbol{w}$.

В силу определяющих соотношений отображение $V \times W \to V \otimes W$, определённое правилом $(\boldsymbol{v}, \boldsymbol{w}) \mapsto \boldsymbol{v} \otimes \boldsymbol{w}$, билинейно. Если $\{\boldsymbol{e}_\alpha : \alpha \in A\}$ и $\{\mathbf{e}_\beta : \beta \in B\}$ — базисы пространств V и W, то система $\{\boldsymbol{e}_\alpha \otimes \mathbf{e}_\beta : \alpha \in A, \beta \in B\}$ является — базисом пространства $V \otimes W$ (её полнота вытекает из билинейности операции $\otimes$, а линейная независимость доказывается точно так же, как в конечномерном случае — надо для каждой пары индексов $\alpha \in A$, $\beta \in B$ рассмотреть линейную функцию $T^{ab} : V \otimes W \to K$, которая каждому элементу вида $\boldsymbol{v} \otimes \boldsymbol{w} \in V \otimes W$ ставит в соответствие произведение α-й и β-й координат $v^\alpha \cdot w^\beta$ (а на остальные элементы пространства $V \otimes W$ продолжается по линейности) и применить лемму о линейной независимости).

Тензорное произведение произвольного конечного числа векторных пространств определяется по аналогии.

Универсальное свойство тензорного произведения

Отображение

$$\otimes\colon V_1 \times \cdots \times V_k \to V_1 \otimes \cdots \otimes V_k, \quad (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) \mapsto \boldsymbol{x}_1 \otimes \cdots \otimes \boldsymbol{x}_k,$$

полилинейно, и для любого векторного пространства U над K и любого полилинейного отображения $\varphi\colon V_1 \times \cdots \times V_k \to U$ существует единственное линейное отображение $\ell\colon V_1 \otimes \cdots \otimes V_k \to U$ со свойством $\varphi = \ell \circ \otimes$, т.е.

$$\varphi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) = \ell(\boldsymbol{x}_1 \otimes \cdots \otimes \boldsymbol{x}_k) \qquad \text{для любых } \boldsymbol{x}_i \in V_i: \qquad (*)$$

$$\begin{tikzcd} V_1 \times \cdots \times V_k \arrow[d, "\otimes"'] \arrow[dr, "\varphi"] & \\ V_1 \otimes \cdots \otimes V_k \arrow[r, "\ell"'] & U \end{tikzcd}$$

Действительно, пусть $\boldsymbol{E}_i$ — базисы в V_i. Для любого набора базисных векторов $(\boldsymbol{e}_{\alpha_1}, \ldots, \boldsymbol{e}_{\alpha_k})$, где $\boldsymbol{e}_{\alpha_i} \in \boldsymbol{E}_i$, положим $\ell(\boldsymbol{e}_{\alpha_1} \otimes \cdots \otimes \boldsymbol{e}_{\alpha_k}) = \varphi(\boldsymbol{e}_{\alpha_1}, \ldots, \boldsymbol{e}_{\alpha_k})$ и продолжим ℓ на остальные элементы тензорного произведения по линейности. Любое другое отображение со свойством $(*)$ должно совпадать с ℓ на базисе, а значит, и везде.

Элементы тензорного произведения $V_1 \otimes \dots \otimes V_n$ вида $\boldsymbol{v}_1 \otimes \dots \otimes \boldsymbol{v}_n$ называются разложимыми. Любое линейное отображение тензорного произведения полностью определяется образами разложимых элементов, поскольку в тензорном произведении есть базис, все элементы которого разложимы.

Если $\ell \colon V_1 \otimes \dots \otimes V_n \to U$ — любое линейное отображение в любое векторное пространство над K, то, полагая

$$\boldsymbol{t}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = \ell(\boldsymbol{x}_1 \otimes \dots \otimes \boldsymbol{x}_n) \quad \text{для } \boldsymbol{x}_i \in V_i,$$

мы получим полилинейное отображение $\boldsymbol{t} \colon V_1 \times \dots \times V_n \to U$. Значит, указанное в формулировке универсального свойства соответствие между полилинейными отображениями $V_1 \times \dots \times V_n \to U$ и линейными отображениями $V_1 \otimes \dots \otimes V_n \to U$ биективно. Ясно, что оно линейно. Рассматривая $U = K$, приходим к выводу:

Вывод: *Имеется естественный изоморфизм между векторным пространством всех полилинейных функций на $V_1 \times \dots \times V_n$ и пространством $(V_1 \otimes \dots \otimes V_n)^*$.*

Универсальное свойство является характеристическим — любое другое пространство с этим свойством совпадает с $V \otimes W$ (с точностью до сохраняющего операцию тензорного произведения изоморфизма). Более того:

Предложение

Если T — векторное пространство над K, для которого существует билинейное отображение $\tilde{\oplus} \colon V \times W \to T$, $(\boldsymbol{x}, \boldsymbol{y}) \mapsto \boldsymbol{x} \tilde{\otimes} \boldsymbol{y}$, с тем свойством, что для любой билинейной функции $b \colon V \times W \to K$ существует единственный линейный функционал $\ell \colon T \to K$, удовлетворяющий условию $b = \ell \circ \tilde{\oplus}$, то существует изоморфизм $\varphi \colon V \otimes W \to T$ такой, что

$$\varphi(\boldsymbol{x} \otimes \boldsymbol{y}) = \boldsymbol{x} \tilde{\oplus} \boldsymbol{y} \quad \text{для всех } \boldsymbol{x} \in V \text{ и } \boldsymbol{y} \in W.$$

Доказательство. В пространстве $V \otimes W$ имеется базис $\{\boldsymbol{e} \otimes \mathbf{e} : \boldsymbol{e} \in \boldsymbol{E}, \mathbf{e} \in \mathbf{E}\}$. На этом базисе φ определяется правилом

$$\varphi(\boldsymbol{e} \otimes \mathbf{e}) = \Psi(\boldsymbol{e}, \mathbf{e}) \qquad \text{для всех } \boldsymbol{e} \in \boldsymbol{E} \text{ и } \mathbf{e} \in \mathbf{E}.$$

Векторы $\varphi(\boldsymbol{e} \otimes \mathbf{e})$ линейно независимы в T в силу леммы о линейной независимости и составляют базис пространства T в силу условия о единственности. $\square$

На тензорное произведение произвольного конечного числа векторных пространств это утверждение распространяется по аналогии.

Примеры

1. $K[x,y] = K[x] \otimes K[y]$
   $\otimes\colon K[x] \times K[y] \to K[x,y]$ определяется правилом $(f \otimes g)(x,y) = f(x) \cdot g(y)$.
2. Пусть $\dim V = n$, и $\dim W = m$ и $\{\boldsymbol{\varepsilon}^1, \dots, \boldsymbol{\varepsilon}^n\}$, $\{\boldsymbol{\epsilon}^1, \dots, \boldsymbol{\epsilon}^m\}$ — взаимные с $\boldsymbol{E}$ и $\mathbf{E}$ базисы в V^* и W^*. Для $\boldsymbol{f} \in V^*$ и $\boldsymbol{g} \in W^*$ положим $(\boldsymbol{f} \otimes \boldsymbol{g})(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{x}) \cdot \boldsymbol{g}(\boldsymbol{y})$. Тогда $\otimes$ — билинейное отображение из $V^* \otimes W^*$ в пространство $\mathscr{B}(V,W)$ билинейных функций $V \times W \to K$, и $(\boldsymbol{\varepsilon}^i \otimes \boldsymbol{\epsilon}^j)(\boldsymbol{x}, \boldsymbol{y}) = x^i \cdot y^j$ для $\boldsymbol{x} \in V$, $\boldsymbol{y} \in W$. Всякая билинейная функция b на $V \times W$ однозначно представляется в виде $b(\boldsymbol{x}, \boldsymbol{y}) = \sum b_{ij} x^i y^j \implies V^* \otimes W^* = \mathscr{B}(V,W)$.
3. Пусть $\dim V = n$, $\dim W = m$ и $\boldsymbol{\mathcal{E}} = \{\boldsymbol{\varepsilon}^1, \dots, \boldsymbol{\varepsilon}^n\}$ — взаимный с $\boldsymbol{E}$ базис в V^*. Для $\boldsymbol{f} \in V^*$ и $\boldsymbol{w} \in W$ пусть $\boldsymbol{f} \otimes \boldsymbol{w}$ — линейное отображение $V \to W$, определённое правилом $(\boldsymbol{f} \otimes \boldsymbol{w})(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})\boldsymbol{w}$ для $\boldsymbol{x} \in V$. Отображение $\otimes\colon V^* \times W \to \mathscr{L}(V,W)$ билинейно ($\mathscr{L}(V,W)$ — векторное пространство линейных отображений $V \to W$). Имеем $\boldsymbol{\varepsilon}^i \otimes \mathbf{e}_j = E_{ji}$ (матричная единица). Матричные единицы образуют базис в $\mathscr{L}(V,W) \implies \mathscr{L}(V,W) = V^* \otimes W$.

4. Пусть L — расширение поля K. Поле L можно рассматривать как векторное пространство над K. Положим $V(L) = L \otimes V$. Это векторное пространство над K. Его можно превратить в векторное пространство над L, определив умножение на элементы L правилом

$$\lambda(\mu \otimes \boldsymbol{v}) = \lambda\mu \otimes \boldsymbol{v} \qquad \text{для} \quad \lambda, \mu \in L, \ \boldsymbol{v} \in V.$$

Пространство V можно вложить в $V(L)$, отождествив каждый вектор $\boldsymbol{v} \in V$ с $1 \otimes \boldsymbol{v} \in V(L)$. Тогда $\lambda\boldsymbol{v}$ отождествляется с $\lambda \otimes \boldsymbol{v}$, и всякий базис пространства V над K является базисом $V(L)$ над L (хотя есть и другие базисы).

С другой стороны, если $\{\lambda_\alpha : \alpha \in A\}$ — базис L над K, то всякий вектор пространства $V(L)$ однозначно представляется в виде $\sum_{i=1}^{n} \lambda_{\alpha_i} \boldsymbol{v}_i$, где λ_{α_i} — разные элементы базиса и $\boldsymbol{v}_i \in V$. Например, любой вектор из комплексификации вещественного пространства V однозначно представляется в виде $\boldsymbol{x} + i\boldsymbol{y}$, где $\boldsymbol{x}, \boldsymbol{y} \in V$.

Тензорное произведение линейных отображений

Пусть $\mathcal{A}\colon V \to V'$ и $\mathcal{B}\colon W \to W'$ — линейные отображения.

Определение

Тензорное произведение $\mathcal{A} \otimes \mathcal{B}$ отображений $\mathcal{A}$ и $\mathcal{B}$ — это отображение $V \otimes W \to V' \otimes W'$, определённое правилом

$$(\mathcal{A} \otimes \mathcal{B})(\boldsymbol{v}_1 \otimes \boldsymbol{w}_1 + \ldots + \boldsymbol{v}_k \otimes \boldsymbol{w}_k) = \mathcal{A}\boldsymbol{v}_1 \otimes \mathcal{B}\boldsymbol{w}_1 + \cdots + \mathcal{A}\boldsymbol{v}_k \otimes \mathcal{B}\boldsymbol{w}_k.$$

Упражнения

1. Проверьте, что это определение корректно (равные элементы переходят в равные).
2. Покажите, что тензорное произведение линейных отображений — линейное отображение.
3. Заметьте, что если $\mathcal{A}$ и $\mathcal{B}$ — операторы (т.е. $V' = V$ и $W' = W$), то $\mathcal{A} \otimes \mathcal{B} = (\mathcal{A} \otimes \boldsymbol{E}_W) \circ (\boldsymbol{E}_V \otimes \mathcal{B})$ (здесь $\boldsymbol{E}_W$ и $\boldsymbol{E}_V$ — тождественные операторы).

! 4. Докажите, что для конечномерных V и W пространство $\mathscr{L}(V \otimes W)$ линейных операторов на $V \otimes W$ является тензорным произведением пространств $\mathscr{L}(V)$ и $\mathscr{L}(W)$ относительно операции тензорного произведения операторов.

Пусть V, V', W и W' — конечномерные пространства, $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ и $\boldsymbol{E}' = \{\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_m\}$ — базисы V и V', $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ и $\mathbf{E}' = \{\mathbf{e}'_1, \dots, \mathbf{e}'_l\}$ — базисы W и W', и пусть $\mathcal{A}\colon V \to V'$ и $\mathcal{B}\colon W \to W'$ — линейные отображения с матрицами $A = (a_{ij})$ и $B = (b_{ij})$ относительно базисов $\boldsymbol{E}$, $\boldsymbol{E}'$, $\mathbf{E}$ и $\mathbf{E}'$. Тогда

$$(\mathcal{A} \otimes \mathcal{B})(\boldsymbol{e}_r \otimes \mathbf{e}_s) = (\mathcal{A}\boldsymbol{e}_r) \otimes (\mathcal{B}\mathbf{e}_s) = \Big(\sum_{i=1}^{n} a_{ir}\boldsymbol{e}'_i\Big) \otimes \Big(\sum_{j=1}^{l} b_{js}\mathbf{e}'_j\Big) = \sum_{i=1}^{m}\sum_{j=1}^{l} a_{ir}b_{js}(\boldsymbol{e}'_i \otimes \mathbf{e}'_j).$$

Таким образом, матрица оператора $\mathcal{A} \otimes \mathcal{B}$ есть

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & \cdots & a_{11}b_{1l} & \cdots & \cdots & a_{1n}b_{11} & a_{1n}b_{12} & \cdots & a_{1n}b_{1l} \\ a_{11}b_{21} & a_{11}b_{22} & \cdots & a_{11}b_{2l} & \cdots & \cdots & a_{1n}b_{21} & a_{1n}b_{22} & \cdots & a_{1n}b_{2l} \\ \vdots & \vdots & \ddots & \vdots & & & \vdots & \vdots & \ddots & \vdots \\ a_{11}b_{k1} & a_{11}b_{k2} & \cdots & a_{11}b_{kl} & \cdots & \cdots & a_{1n}b_{k1} & a_{1n}b_{k2} & \cdots & a_{1n}b_{kl} \\ \vdots & \vdots & & \vdots & \ddots & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \ddots & \vdots & \vdots & & \vdots \\ a_{m1}b_{11} & a_{m1}b_{12} & \cdots & a_{m1}b_{1l} & \cdots & \cdots & a_{mn}b_{11} & a_{mn}b_{12} & \cdots & a_{mn}b_{1l} \\ a_{m1}b_{21} & a_{m1}b_{22} & \cdots & a_{m1}b_{2l} & \cdots & \cdots & a_{mn}b_{21} & a_{mn}b_{22} & \cdots & a_{mn}b_{2l} \\ \vdots & \vdots & \ddots & \vdots & & & \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{k1} & a_{m1}b_{k2} & \cdots & a_{m1}b_{kl} & \cdots & \cdots & a_{mn}b_{k1} & a_{mn}b_{k2} & \cdots & a_{mn}b_{kl} \end{pmatrix}.$$

Она называется тензорным, или кронекеровым, произведением (или произведением Кронекера) матриц A и B.

Упражнение

Найдите определитель матрицы $A \otimes B$ для линейных операторов $\mathcal{A}, \mathcal{B}\colon V \to V$.

Свёртка

Пусть $p, q > 0$, $\boldsymbol{t}$ — тензор типа (p, q) и $\alpha \in \{1, \dots, p\}$, $\beta \in \{1, \dots, q\}$. Тензор

$$\bar{\boldsymbol{t}}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{p-1}; \boldsymbol{f}^1, \dots, \boldsymbol{f}^{q-1}) =$$
$$= \boldsymbol{t}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{\alpha-1}, \boldsymbol{e}_k, \boldsymbol{x}_\alpha, \dots, \boldsymbol{x}_{p-1}; \boldsymbol{f}^1, \dots, \boldsymbol{f}^{\beta-1}, \boldsymbol{\varepsilon}^k, \boldsymbol{f}^\beta, \dots, \boldsymbol{f}^{q-1})$$

называется свёрткой тензора $\boldsymbol{t}$ по α-му ковариантному индексу и β-му контравариантному индексу. Это тензор типа $(p-1, q-1)$.

Примеры

1. Свёртка тензора $\boldsymbol{t}(\boldsymbol{x}, \boldsymbol{f}) = (\boldsymbol{g} \otimes \boldsymbol{v})(\boldsymbol{x}, \boldsymbol{f}) = \boldsymbol{g}(\boldsymbol{x}) \cdot \boldsymbol{f}(\boldsymbol{v})$ (где $\boldsymbol{g} \in V^*$ и $\boldsymbol{v} \in V$) по единственным возможным индексам есть скаляр $\boldsymbol{g}(\boldsymbol{e}_i) \cdot \boldsymbol{\varepsilon}^i(\boldsymbol{v}) = g_i \cdot v^i = \boldsymbol{g}(\boldsymbol{v})$.
2. Свёртка тензора $\boldsymbol{t}(\boldsymbol{x}, \boldsymbol{f}) = \boldsymbol{f}(\mathcal{A}\boldsymbol{x})$ (где $\mathcal{A}$ — оператор) по единственным возможным индексам есть скаляр $\boldsymbol{\varepsilon}^i(\mathcal{A}\boldsymbol{e}_i)$ — след оператора $\mathcal{A}$.
3. Свёртка тензора $\boldsymbol{t} \otimes \boldsymbol{v}$, где $\boldsymbol{t}(\boldsymbol{x}, \boldsymbol{f}) = \boldsymbol{f}(\mathcal{A}\boldsymbol{x})$ для оператора $\mathcal{A}$ и $\boldsymbol{v} \in V$, по единственному ковариантному индексу и второму контравариантному индексу есть вектор $\boldsymbol{u}$ такой, что для всякого $\boldsymbol{f} \in V^*$
   $\boldsymbol{u}(\boldsymbol{f}) = \boldsymbol{f}(\mathcal{A}\boldsymbol{e}_i) \cdot \boldsymbol{v}(\boldsymbol{\varepsilon}^i) = \boldsymbol{f}(\mathcal{A}\boldsymbol{e}_i) \cdot v^i = \boldsymbol{f}(\mathcal{A}(v^i\boldsymbol{e}_i)) = \boldsymbol{f}(\mathcal{A}\boldsymbol{v})$, т.е. $\boldsymbol{u} = \mathcal{A}\boldsymbol{v}$.

4. Тензорное произведение двух операторов $\mathcal{A}\colon V \to V$ и $\mathcal{B}\colon V \to V$ с матрицами $A = (a_i^j)$ и $B = (b_i^j)$ (номера столбцов снизу) — это оператор $V \otimes V \to V \otimes V$, заданный на базисных векторах правилом $(\mathcal{A} \otimes \mathcal{B})(\boldsymbol{e}_i \otimes \boldsymbol{e}_j) = a_i^r b_j^s (\boldsymbol{e}_r \otimes \boldsymbol{e}_s)$. Его можно трактовать как тензор $\boldsymbol{\tau}$ типа $(1,1)$ на $V \otimes V$, заданный на базисных векторах правилом $\boldsymbol{\tau}(\boldsymbol{e}_i \otimes \boldsymbol{e}_j; \boldsymbol{\varepsilon}^r \otimes \boldsymbol{\varepsilon}^s) = a_i^r b_j^s$, или как тензор $\boldsymbol{t}$ типа $(2,2)$ на V, заданный на базисных векторах правилом $\boldsymbol{t}(\boldsymbol{e}_i, \boldsymbol{e}_j; \boldsymbol{\varepsilon}^r, \boldsymbol{\varepsilon}^s) = a_i^r b_j^s$. Его свёртки:
   - по первому ковариантному индексу и первому контравариантному индексу: $\boldsymbol{s}(\boldsymbol{e}_j, \boldsymbol{\varepsilon}^s) = a_k^k b_j^s$ — это оператор с матрицей $(\operatorname{tr} A) \cdot B$;
   - по первому ковариантному индексу и второму контравариантному индексу: $\boldsymbol{s}(\boldsymbol{e}_j, \boldsymbol{\varepsilon}^r) = a_k^r b_j^k$ — это оператор с матрицей AB (т.е. оператор $\mathcal{A} \circ \mathcal{B}$);
   - по второму ковариантному индексу и первому контравариантному индексу: $\boldsymbol{s}(\boldsymbol{e}_i, \boldsymbol{\varepsilon}^s) = a_i^k b_k^s$ — это оператор с матрицей BA;
   - по второму ковариантному индексу и второму контравариантному индексу: $\boldsymbol{s}(\boldsymbol{e}_i, \boldsymbol{\varepsilon}^r) = a_i^r b_k^k$ — это оператор с матрицей $(\operatorname{tr} B) \cdot A$.

Свёртка тензора не зависит от базиса.

Доказательство. У нас есть базис $\boldsymbol{E}$. Пусть $\tilde{\boldsymbol{E}} = \{\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_n\}$ — другой базис пространства V и $\tilde{\boldsymbol{\mathcal{E}}} = \{\tilde{\boldsymbol{\varepsilon}}_1, \dots, \tilde{\boldsymbol{\varepsilon}}_n\}$ — взаимный с ним базис пространства V^*. Свёртка тензора $\boldsymbol{t}$ типа (p, q) по α-му ковариантному индексу и β-му контравариантному индексу относительно базисов $\tilde{\boldsymbol{E}}$ и $\tilde{\boldsymbol{\mathcal{E}}}$ равна

$$\bar{\boldsymbol{t}}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{p-1}; \boldsymbol{f}^1, \dots, \boldsymbol{f}^{q-1}) = \\ = \boldsymbol{t}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{\alpha-1}, \tilde{\mathbf{e}}_k, \boldsymbol{x}_\alpha, \dots, \boldsymbol{x}_{p-1}; \boldsymbol{f}^1, \dots, \boldsymbol{f}^{\beta-1}, \tilde{\boldsymbol{\varepsilon}}_k, \boldsymbol{f}^\beta, \dots, \boldsymbol{f}^{q-1}).$$

Зафиксировав все аргументы тензора $\boldsymbol{t}$, кроме α-го и $(\alpha+\beta)$-го, мы получим билинейную функцию $b(\boldsymbol{x}, \boldsymbol{f})$. Нам надо доказать, что $b(\tilde{\mathbf{e}}_i, \tilde{\boldsymbol{\varepsilon}}^i) = b(\mathbf{e}_i, \boldsymbol{\varepsilon}^i)$. Пусть $C = (c_i^j)$ — матрица перехода от $\boldsymbol{E}$ к $\tilde{\boldsymbol{E}}$, т.е. $\tilde{\mathbf{e}}_i = c_i^j \mathbf{e}_j$. Тогда C^T — матрица перехода от $\tilde{\boldsymbol{\mathcal{E}}}$ к $\boldsymbol{\mathcal{E}}$, т.е. $\boldsymbol{\varepsilon}^i = c_j^i \tilde{\boldsymbol{\varepsilon}}^j$. Имеем

$$b(\tilde{\mathbf{e}}_i, \tilde{\boldsymbol{\varepsilon}}^i) = b(c_i^j \mathbf{e}_j, \tilde{\boldsymbol{\varepsilon}}^i) = b(\mathbf{e}_j, c_i^j \tilde{\boldsymbol{\varepsilon}}^i) = b(\mathbf{e}_j, \boldsymbol{\varepsilon}^j).$$

$\square$

Симметрирование и альтернирование

Будем рассматривать только тензоры типа $(p,0)$ (и $(0,q)$ — по аналогии) и поле характеристики 0.

Пусть $\boldsymbol{t}$ — произвольный тензор типа $(p,0)$ и $\pi \in S_p$ — подстановка. Положим

$$\boldsymbol{t}_\pi(\boldsymbol{x}_1,\dots,\boldsymbol{x}_p) = \boldsymbol{t}(\boldsymbol{x}_{\pi(1)},\dots,\boldsymbol{x}_{\pi(p)}).$$

Получили отображение

$$\vartheta_\pi\colon T_p^0(V)\to T_p^0(V), \qquad \boldsymbol{t}\mapsto \boldsymbol{t}_\pi.$$

Это невырожденный линейный оператор на $T_p^0(V)$. Имеем $\vartheta_\pi\circ\vartheta_\sigma = \vartheta_{\pi\sigma}$, так что $\pi\mapsto\vartheta_\pi$ — действие группы S_p на $T_p^0(V)$.

Определение

Тензор $\boldsymbol{t}$ типа $(p,0)$ называется симметричным (или симметрическим), если $\vartheta_\pi(\boldsymbol{t}) = \boldsymbol{t}$ для всех $\pi\in S_p$.

Множество симметричных тензоров в пространстве $T_p^0(V)$ образует подпространство. Обозначим его $ST_p(V)$.

Если $\boldsymbol{t}$ — симметричный тензор типа $(p,0)$, то его координаты в любом базисе *симметричны*, т.е. не меняются при любой перестановке индексов: $t_{i_1\ldots i_p} = t_{\pi(i_1)\ldots\pi(i_p)}$ для всех $\pi \in S_p$.

Замечание

Тензор симметричен $\iff$ его координаты в некотором (любом) базисе симметричны.

Определение

Симметрирование (или **симметризация**) тензоров из $T_p^0(V)$ — это линейный оператор Sym на $T_p^0(V)$, определённый правилом

$$\mathrm{Sym}\colon \boldsymbol{t} \mapsto \frac{1}{p!}\sum_{\pi\in S_p} \vartheta_\pi(\boldsymbol{t}).$$

Другими словами, $\mathrm{Sym}\,\boldsymbol{t}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p) = \frac{1}{p!}\sum_{\pi\in S_p} \boldsymbol{t}(\boldsymbol{x}_{\pi(1)},\ldots,\boldsymbol{x}_{\pi(p)})$. Координаты в любом базисе вычисляются по тому же правилу: $(\mathrm{Sym}\,t)_{i_1\ldots i_p} = \frac{1}{p!}\sum\limits_{\pi\in S_p} t_{\pi(1)\ldots\pi(p)}$.

Предложение

1. $\operatorname{Im} \operatorname{Sym} = ST_p(V)$, 2. $\operatorname{Sym}^2 = \operatorname{Sym}$.

Доказательство. Для $\sigma \in S_p$ и $\boldsymbol{t} \in T_p^0(V)$ имеем

$$\vartheta_\sigma(\operatorname{Sym} \boldsymbol{t}) = \vartheta_\sigma\Big(\frac{1}{p!}\sum_{\pi \in S_p} \vartheta_\pi(\boldsymbol{t})\Big) = \frac{1}{p!}\sum_{\pi \in S_p} \vartheta_{\sigma\pi}(\boldsymbol{t}) = \frac{1}{p!}\sum_{\tau \in S_p} \vartheta_\tau(\boldsymbol{t}) = \operatorname{Sym} \boldsymbol{t}.$$

Для любого $\boldsymbol{t} \in ST_p(V)$ имеем $\operatorname{Sym} \boldsymbol{t} = \frac{1}{p!}\sum_{\pi \in S_p} \vartheta_\pi(\boldsymbol{t}) = \frac{1}{p!}\sum_{\pi \in S_p} \boldsymbol{t} = \boldsymbol{t}$. □

Предложение

$\dim ST_p(V) = \frac{(n+p-1)!}{p!(n-1)!}$ *(число сочетаний с повторениями из n по p).*

Доказательство. Базис в $ST_p(V)$ составляют элементы $\operatorname{Sym}(\boldsymbol{\varepsilon}_{i_1} \otimes \cdots \otimes \boldsymbol{\varepsilon}_{i_p})$, $1 \leqslant i_1 \leqslant \cdots \leqslant i_p \leqslant n$, поскольку система всех таких элементов полна (так как она состоит из образов элементов базиса в $T_p^0(V)$) и линейно независима — это доказывается с помощью леммы о линейной независимости точно так же, как линейная независимость всех элементов вида $\boldsymbol{\varepsilon}_{i_1} \otimes \cdots \otimes \boldsymbol{\varepsilon}_{i_p}$ в $T_p^0(V)$. □

Определение

Тензор $\boldsymbol{t}$ типа $(p, 0)$ называется кососимметричным (кососимметрическим, антисимметричным, знакопеременным), если при перестановке любых двух аргументов функция $\boldsymbol{t}$ меняет знак: $\boldsymbol{t}(\ldots, \boldsymbol{v}, \ldots, \boldsymbol{u}, \ldots) = -\boldsymbol{t}(\ldots, \boldsymbol{u}, \ldots, \boldsymbol{v}, \ldots)$.

Равносильное условие: $\vartheta_\pi(\boldsymbol{t}) = \operatorname{sgn} \pi \boldsymbol{t}$ для всех $\pi \in S_p$.

Поскольку поле K имеет характеристику $0 \neq 2$, тензор $\boldsymbol{t}$ типа $(p, 0)$ кососимметричен $\iff$ значение $\boldsymbol{t}$ в наборе аргументов, среди которых есть хотя бы два совпадающих, равно нулю: $\boldsymbol{t}(\ldots, \boldsymbol{v}, \ldots, \boldsymbol{v}, \ldots) = 0$.

Множество кососимметричных тензоров в пространстве $T_p^0(V)$ образует подпространство. Обозначим его $\bigwedge T_p(V)$.

Если $\boldsymbol{t}$ — кососимметричный тензор типа $(p, 0)$, то при перестановке индексов его координаты в любом базисе умножаются на ± 1 (в зависимости от знака перестановки): $t_{i_1 \ldots i_p} = \operatorname{sgn} \pi \cdot t_{\pi(i_1) \ldots \pi(i_p)}$ для всех $\pi \in S_p$. Всякая координата с двумя одинаковыми индексами равна нулю.

Замечание

Тензор $\boldsymbol{t}$ кососимметричен $\iff$ его координаты в некотором (любом) базисе подчиняются правилу $t_{i_1 \ldots i_p} = \operatorname{sgn} \pi \cdot t_{\pi(i_1) \ldots \pi(i_p)}$ для всех $\pi \in S_p$.

Определение

Альтернирование тензоров из $T_p^0(V)$ — это линейный оператор Alt на $T_p^0(V)$, определённый правилом

$$\operatorname{Alt}\colon \boldsymbol{t} \mapsto \frac{1}{p!}\sum_{\pi\in S_p} \operatorname{sgn}\pi\, \vartheta_\pi(\boldsymbol{t}).$$

Другими словами, $\operatorname{Alt}\boldsymbol{t}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p) = \frac{1}{p!}\sum_{\pi\in S_p} \operatorname{sgn}\pi\, \boldsymbol{t}(\boldsymbol{x}_{\pi(1)},\ldots,\boldsymbol{x}_{\pi(p)})$. Координаты в любом базисе вычисляются по тому же правилу:

$$(\operatorname{Alt} t)_{i_1\ldots i_p} = \frac{1}{p!}\sum_{\pi\in S_p} \operatorname{sgn}\pi\, t_{\pi(1)\ldots\pi(p)}.$$

Предложение

1. $\operatorname{Im}\operatorname{Alt} = \bigwedge T_p(V)$,
2. $\operatorname{Alt}^2 = \operatorname{Alt}$,
3. $\operatorname{Alt}\vartheta_\pi(\boldsymbol{t}) = \operatorname{sgn}\pi\operatorname{Alt}\boldsymbol{t}$ *для всех* $\pi \in S_p$ *и* $\boldsymbol{t} \in T_p^0(V)$.

Доказательство. 1 и 2: Для $\sigma \in S_p$ и $\boldsymbol{t} \in T_p^0(V)$ имеем

$$\vartheta_\sigma(\operatorname{Alt}\boldsymbol{t}) = \vartheta_\sigma\left(\frac{1}{p!}\sum_{\pi\in S_p}\operatorname{sgn}\pi\vartheta_\pi(\boldsymbol{t})\right) = \frac{1}{p!}\sum_{\pi\in S_p}\operatorname{sgn}\pi\vartheta_\sigma(\vartheta_\pi(\boldsymbol{t})) =$$
$$= \frac{1}{p!}\sum_{\pi\in S_p}(\operatorname{sgn}\sigma)^2\operatorname{sgn}\pi\vartheta_{\sigma\pi}(\boldsymbol{t}) = \operatorname{sgn}\sigma\frac{1}{p!}\sum_{\pi\in S_p}\operatorname{sgn}\sigma\operatorname{sgn}\pi\vartheta_{\sigma\pi}(\boldsymbol{t}) =$$
$$= \operatorname{sgn}\sigma\frac{1}{p!}\sum_{\tau\in S_p}\operatorname{sgn}\tau\vartheta_\tau(\boldsymbol{t}) = \operatorname{sgn}\sigma\operatorname{Alt}\boldsymbol{t}.$$

Для любого $\boldsymbol{t} \in \bigwedge T_p(V)$ имеем $\operatorname{Alt}\boldsymbol{t} = \frac{1}{p!}\sum_{\pi\in S_p}\operatorname{sgn}\pi\vartheta_\pi(\boldsymbol{t}) = \frac{1}{p!}\sum_{\pi\in S_p}(\operatorname{sgn}\pi)^2\boldsymbol{t} = \boldsymbol{t}$.

3:
$$\operatorname{Alt}\vartheta_\sigma(\boldsymbol{t}) = \frac{1}{p!}\sum_{\pi\in S_p}\operatorname{sgn}\pi\vartheta_\pi(\vartheta_\sigma(\boldsymbol{t})) = \frac{1}{p!}\sum_{\pi\in S_p}\operatorname{sgn}\pi(\operatorname{sgn}\sigma)^2\vartheta_\pi(\vartheta_\sigma(\boldsymbol{t})) =$$
$$= \operatorname{sgn}\sigma\frac{1}{p!}\sum_{\pi\in S_p}\operatorname{sgn}\pi\operatorname{sgn}\sigma\vartheta_\pi(\vartheta_\sigma(\boldsymbol{t})) = \frac{1}{p!}\sum_{\tau\in S_p}\operatorname{sgn}\tau\vartheta_\tau(\boldsymbol{t}) = \operatorname{sgn}\sigma\operatorname{Alt}\boldsymbol{t}.$$

□

Предложение

$\dim \bigwedge T_p(V) = \frac{n!}{p!(n-p)!}$ *(число сочетаний из n по p).*

Доказательство. Базис в $\bigwedge T_p(V)$ составляют элементы $\mathrm{Alt}(\boldsymbol{\varepsilon}_{i_1} \otimes \cdots \otimes \boldsymbol{\varepsilon}_{i_p})$, $1 \leqslant i_1 < \cdots < i_p \leqslant n$. □

Замечание

Операции симметрирования и альтернирования являются проектированиями пространства $T_p^0(V)$ на подпространства $ST_p(V)$ и $\bigwedge T_p(V)$ соответственно.

Упражнения

1. Докажите, что $T_2(V) = ST_2(V) \oplus \bigwedge T_2(V)$.
2. Покажите, что для $p > 2$ $T_p(V) \neq ST_p(V) + \bigwedge T_p(V)$.

Предложение

Операции Sym *и* Alt *ассоциативны и дистрибутивны.*

Лемма

Для любых тензоров $\boldsymbol{f} \in T^0_k(V)$ *и* $\boldsymbol{g} \in T^0_m(V)$

1. $\operatorname{Sym}((\operatorname{Sym}\boldsymbol{f}) \otimes \boldsymbol{g}) = \operatorname{Sym}(\boldsymbol{f} \otimes (\operatorname{Sym}\boldsymbol{g})) = \operatorname{Sym}(\boldsymbol{f} \otimes \boldsymbol{g})$,
2. $\operatorname{Alt}((\operatorname{Alt}\boldsymbol{f}) \otimes \boldsymbol{g}) = \operatorname{Alt}(\boldsymbol{f} \otimes (\operatorname{Alt}\boldsymbol{g})) = \operatorname{Alt}(\boldsymbol{f} \otimes \boldsymbol{g})$.

Доказательство. Докажем первую половину 2 (остальное аналогично). По определению $\operatorname{Alt}\boldsymbol{f} = \frac{1}{k!}\sum_{\pi \in S_k} \operatorname{sgn}\pi\,\vartheta_\pi(\boldsymbol{f})$. Линейность альтернирования $\Longrightarrow$

$$\operatorname{Alt}((\operatorname{Alt}\boldsymbol{f}) \otimes \boldsymbol{g}) = \frac{1}{k!}\sum_{\pi \in S_k} \operatorname{sgn}\pi\,\operatorname{Alt}(\vartheta_\pi(\boldsymbol{f}) \otimes \boldsymbol{g}).$$

Пусть $\pi \in S_k$. Положим $\tilde{\pi}(i) = \pi(i)$ для $i \leqslant k$, $\tilde{\pi}(i) = i$ для $k < i \leqslant k + m$. Имеем $\vartheta_\pi(\boldsymbol{f}) \otimes \boldsymbol{g} = \vartheta_{\tilde{\pi}}(\boldsymbol{f} \otimes \boldsymbol{g})$. Получаем

$$\operatorname{Alt}(\vartheta_\pi(\boldsymbol{f}) \otimes \boldsymbol{g}) = \operatorname{Alt}(\vartheta_{\tilde{\pi}}(\boldsymbol{f} \otimes \boldsymbol{g})) = \operatorname{sgn}\tilde{\pi}\,\operatorname{Alt}(\boldsymbol{f} \otimes \boldsymbol{g}) \qquad \Longrightarrow$$

$$\operatorname{Alt}((\operatorname{Alt}\boldsymbol{f}) \otimes \boldsymbol{g}) = \frac{1}{k!}\sum_{\pi \in S_k} \operatorname{sgn}\pi \cdot \operatorname{sgn}\tilde{\pi}\,\operatorname{Alt}(\boldsymbol{f} \otimes \boldsymbol{g}) = \frac{1}{k!}\sum_{\pi \in S_k} \operatorname{Alt}(\boldsymbol{f} \otimes \boldsymbol{g}) = \operatorname{Alt}(\boldsymbol{f} \otimes \boldsymbol{g}). \qquad \square$$

* * *

Для тензоров типа $(0, q)$ теория совершенно аналогична.

Опускание и поднятие индексов

Если V — евклидово пространство, то между V и V^* имеется канонический изоморфизм $\varphi\colon V \to V^*$, действующий по правилу $\boldsymbol{v} \mapsto \boldsymbol{f}_{\boldsymbol{v}}$, $\boldsymbol{f}_{\boldsymbol{v}}(\boldsymbol{x}) = (\boldsymbol{v}, \boldsymbol{x})$ для всякого $\boldsymbol{x} \in V$. Обратный изоморфизм φ^{-1} действует по правилу $\boldsymbol{f} \mapsto \boldsymbol{v}_{\boldsymbol{f}}$, где $(\boldsymbol{v}_{\boldsymbol{f}}, \boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})$.

Пусть
$$\boldsymbol{t}\colon \underbrace{V \times \dots \times V}_{p \text{ раз}} \times \underbrace{V^* \times \dots \times V^*}_{q \text{ раз}} \to K$$
— полилинейная функция (тензор типа (p, q)). Заменим ковекторы $\boldsymbol{f} \in V^*$ соответствующими векторами $\boldsymbol{v}_{\boldsymbol{f}} \in V$. Получим полилинейную функцию
$$\boldsymbol{t}'\colon \underbrace{V \times \dots \times V}_{p+q \text{ раз}} \to K,$$
зависящую от $p + q$ векторов.

Пусть $t_{k_1\dots k_p}^{m_1\dots m_q} = \boldsymbol{t}(\boldsymbol{e}_{k_1}, \dots, \boldsymbol{e}_{k_p}; \boldsymbol{\varepsilon}^{m_1}, \dots, \boldsymbol{\varepsilon}^{m_q})$ — координаты тензора $\boldsymbol{t}$ и $t'_{k_1\dots k_p k_{p+1}\dots k_{p+q}} = \boldsymbol{t}'(\boldsymbol{e}_{k_1}, \dots, \boldsymbol{e}_{k_p}, \boldsymbol{e}_{k_{p+1}}, \dots, \boldsymbol{e}_{k_{p+q}})$ — координаты тензора $\boldsymbol{t}'$. Если базис $\boldsymbol{E}$ ортонормированный, то $\boldsymbol{v}_{\boldsymbol{\varepsilon}^i} = \boldsymbol{e}_i$, поэтому в этом случае $t'_{k_1\dots k_p k_{p+1}\dots k_{p+q}} = t_{k_1\dots k_p}^{k_{p+1}\dots k_{p+q}}$.

В случае произвольного базиса скалярное произведение определяется формулой
$$(\boldsymbol{x},\boldsymbol{y}) = (x_1,\dots,x_n)\,G\begin{pmatrix} y_1\\ \vdots\\ y_n\end{pmatrix},$$
где G — матрица Грама скалярного произведения: $g_{ij} = (\boldsymbol{e}_i,\boldsymbol{e}_j)$.

Скалярное произведение — тензор типа $(2,0)$, который называется метрическим тензором, и g_{ij} — его компоненты.

Поскольку канонический изоморфизм $\varphi\colon V \to V^*$ определяется правилом $\boldsymbol{v} \mapsto \boldsymbol{f}_{\boldsymbol{v}}$, $\boldsymbol{f}_{\boldsymbol{v}}(\boldsymbol{x}) = (\boldsymbol{v},\boldsymbol{x})$, имеем $\varphi(\boldsymbol{e}_i)(\boldsymbol{x}) = (\boldsymbol{e}_i,\boldsymbol{x}) = g_{ij}x^j$. Значит, координаты функционала $\varphi(\boldsymbol{e}_i)$ во взаимном базисе суть $g_{i1},\dots,g_{in}$, т.е. $G = G^T$ — матрица изоморфизма φ, а G^{-1} — матрица обратного изоморфизма φ^{-1}:
$$\varphi(\boldsymbol{e}_i) = g_{ij}\boldsymbol{\varepsilon}^j \qquad \text{и} \qquad \boldsymbol{e}_i = g_{ij}\varphi^{-1}(\boldsymbol{\varepsilon}^j) = g_{ij}\boldsymbol{v}_{\boldsymbol{\varepsilon}^j}.$$
Таким образом, скалярное произведение на сопряжённом пространстве, относительно которого φ — изоморфизм евклидовых пространств (т.е. $(\varphi(\boldsymbol{x}),\varphi(\boldsymbol{y})) = (\boldsymbol{x},\boldsymbol{y})$), имеет матрицу G^{-1}:
$$(\boldsymbol{f},\boldsymbol{g}) = (\varphi^{-1}(\boldsymbol{f}),\varphi^{-1}(\boldsymbol{g})) = (f_1,\dots,f_n)(G^{-1})^T G G^{-1}\begin{pmatrix} g_1\\ \vdots\\ g_n\end{pmatrix} = (f_1,\dots,f_n)G^{-1}\begin{pmatrix} g_1\\ \vdots\\ g_n\end{pmatrix},$$
так что координаты g^{ij} соответствующего метрического тензора равны координатам матрицы G^{-1}.

Отождествляя векторы $\boldsymbol{v_f}$ с ковекторами $\boldsymbol{f}$, подставим вместо последних q векторов $\boldsymbol{e}_i$ их выражения $\boldsymbol{e}_i = g_{ij}\boldsymbol{v}_{\boldsymbol{\varepsilon}^j}$ и заменим $\boldsymbol{v}_{\boldsymbol{\varepsilon}^j}$ на $\boldsymbol{\varepsilon}^j = \varphi(\boldsymbol{v}_{\boldsymbol{\varepsilon}^j})$:

$$\begin{aligned} t'_{k_1\dots k_p k_{p+1}\dots k_{p+q}} = \boldsymbol{t}'(\boldsymbol{e}_{k_1},\dots,\boldsymbol{e}_{k_p},\boldsymbol{e}_{k_{p+1}},\dots,\boldsymbol{e}_{k_{p+q}}) = \\ = \boldsymbol{t}(\boldsymbol{e}_{k_1},\dots,\boldsymbol{e}_{k_p}; g_{k_{p+1}i_1}\boldsymbol{\varepsilon}^{i_1},\dots,g_{k_{p+q}i_q}\boldsymbol{\varepsilon}^{i_q}) = g_{k_{p+1}i_1}\cdot\ldots\cdot g_{k_{p+q}i_q}\cdot t^{i_1\dots i_q}_{k_1\dots k_p}. \end{aligned}$$

Правило $\boldsymbol{t} \mapsto \boldsymbol{t}'$ определяет изоморфизм

$$T^q_p \rightleftarrows T_{p+q}, \qquad \text{т.е.} \qquad \underbrace{V^*\otimes\dots\otimes V^*}_{p\ \text{раз}}\otimes\underbrace{V\otimes\dots\otimes V}_{q\ \text{раз}} \rightleftarrows \underbrace{V^*\otimes\dots\otimes V^*}_{p+q\ \text{раз}},$$

при котором каждому тензору с координатами $t^{k_{p+1}\dots k_{p+1}}_{k_1\dots k_p}$ соответствует тензор с координатами $t'_{k_1\dots k_p k_{p+1}\dots k_{p+q}} = g_{i_1 k_{p+1}}\cdot\ldots\cdot g_{i_q k_{p+q}}\cdot t^{i_1\dots i_q}_{k_1\dots k_p}$. Он называется опусканием индексов.

Аналогично определяется операция поднятия индексов $T^q_p \to T^{p+q}$. Координаты меняются по формуле $t'^{k_1\dots k_p k_{p+1}\dots k_{p+1}} = g^{k_1 i_1}\cdot\ldots\cdot g^{k_p i_p}\cdot t^{k_{p+1}\dots k_{p+q}}_{i_1\dots i_p}$, где $g^{ij} = (\boldsymbol{\varepsilon}^i,\boldsymbol{\varepsilon}^j)$ — компоненты матрицы G^{-1}.

Тензорное произведение евклидовых пространств

Пусть E_1 и E_2 — два евклидовых пространства размерностей n_1 и n_2. На $E_1 \otimes E_2$ определим функцию от двух аргументов. Положим

$$(\boldsymbol{x} \otimes \boldsymbol{y}, \boldsymbol{u} \otimes \boldsymbol{v}) = (\boldsymbol{x}, \boldsymbol{u})_1 \cdot (\boldsymbol{y}, \boldsymbol{v})_2.$$

Для $\boldsymbol{x} = \boldsymbol{x}_1 \otimes \boldsymbol{y}_1 + \dots + \boldsymbol{x}_k \otimes \boldsymbol{y}_k$ и $\boldsymbol{u} = \boldsymbol{u}_1 \otimes \boldsymbol{v}_1 + \dots + \boldsymbol{u}_m \otimes \boldsymbol{v}_m$ положим

$$(\boldsymbol{x}, \boldsymbol{u}) = \sum_{i=1}^{k} \sum_{j=1}^{m} (\boldsymbol{x}_i \otimes \boldsymbol{y}_i, \boldsymbol{u}_j \otimes \boldsymbol{v}_j).$$

Эта функция определена корректно (принимает равные значения на равных векторах) и является скалярным произведением.

Пространство $E_1 \otimes E_2$ с таким скалярным произведением называется тензорным произведением евклидовых пространств E_1 и E_2.

Если $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_{n_1}\}$ и $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_{n_2}\}$ — ортонормированные базисы в E_1 и E_2, то $\{\boldsymbol{e}_i \otimes \mathbf{e}_j : i \leqslant n_1, j \leqslant n_2\}$ — ортонормированный базис в $E_1 \otimes E_2$.

Дальше всегда предполагается, что V — n-мерное векторное пространство над полем K, хотя почти все конструкции и утверждения без изменений переносятся на бесконечномерные пространства. Предположение конечномерности действительно нужно только в тех местах, где упоминаются сопряжённые пространства.

Прямая сумма векторных пространств

Определение

Прямой суммой конечного числа векторных пространств $V_1, \ldots, V_k$ называется векторное пространство $V_1 \oplus \cdots \oplus V_k = V_1 \times \cdots \times V_k$, составленное из всех последовательностей $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k)$, где $\boldsymbol{x}_i \in V_i$, с покомпонентными операциями $\lambda(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) = (\lambda\boldsymbol{x}_1, \ldots, \lambda\boldsymbol{x}_k)$ и $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) + (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k) = (\boldsymbol{x}_1 + \boldsymbol{y}_1, \ldots, \boldsymbol{x}_k + \boldsymbol{y}_k)$.

Прямой суммой $\bigoplus\limits_{i=0}^{\infty} V_i$ бесконечного числа векторных пространств $V_1, V_2, \ldots$ называется векторное пространство, составленное из всех последовательностей $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k, \boldsymbol{0}_{k+1}, \boldsymbol{0}_{k+2}, \ldots)$, где $k \in \mathbb{N}$, $\boldsymbol{x}_i \in V_i$ и $\boldsymbol{0}_j$ — нули пространств V_j, с покомпонентными операциями.

Каждое пространство V_i изоморфно подпространству прямой суммы $\bigoplus\limits_{i=0}^{\infty} V_i$, состоящему из всех наборов $(\boldsymbol{0}_1, \ldots, \boldsymbol{0}_{i-1}, \boldsymbol{x}, \boldsymbol{0}_{i+1}, \boldsymbol{0}_{i+2}, \ldots)$, где $\boldsymbol{x} \in V_i$. Если в пространствах V_i зафиксированы базисы $\boldsymbol{E}_i$, то базис пространства $\bigoplus\limits_{i=0}^{\infty} V_i$ составляют всевозможные наборы $(\boldsymbol{0}_1, \ldots, \boldsymbol{0}_{i-1}, \boldsymbol{e}, \boldsymbol{0}_{i+1}, \boldsymbol{0}_{i+2}, \ldots)$, где $i \in \mathbb{N}$ и $\boldsymbol{e} \in \boldsymbol{E}_i$.

Определение

Алгебра A с умножением $*$ называется градуированной алгеброй, если как векторное пространство она является прямой суммой своих подпространств: $A = \bigoplus\limits_{i\in\mathbb{Z}} A_i$, причём $A_i * A_j \subset A_{i+j}$ для любых $i, j \in \mathbb{Z}$.

В этом определении некоторые прямые слагаемые A_i могут быть тривиальными. Вместо $\mathbb{Z}$ можно рассматривать $\mathbb{N}$ или любую другую полугруппу.

Положим $T(V) = \bigoplus\limits_{q=0}^{\infty} T^p(V)$, где $T^0(V) = K$ и $T^q(V) = T_0^q(V) = V^{\otimes q}$ для $q \in \mathbb{N}$.

Элементы $T(V)$ записывают не в виде последовательностей, а в виде сумм

$$\boldsymbol{v}_1 \otimes \cdots \otimes \boldsymbol{v}_k + \boldsymbol{w}_1 \otimes \cdots \otimes \boldsymbol{w}_l + \cdots + \boldsymbol{u}_1 \otimes \cdots \otimes \boldsymbol{u}_m$$

— между этими формами записи имеется взаимно однозначное соответствие.

Для любых неотрицательных целых p и q и любых $\boldsymbol{t} \in T^p(V)$, $\boldsymbol{s} \in T^q(V)$ определено тензорное произведение $\boldsymbol{t} \otimes \boldsymbol{s}$ (см. подраздел «Арифметические операции над тензорами» в начале раздела о тензорах). Операция тензорного умножения билинейна, и для разложимых элементов пространств $T^p(V)$ и $T^q(V)$ она определяется правилом

$$(\boldsymbol{x}_1 \otimes \cdots \otimes \boldsymbol{x}_p) \otimes (\boldsymbol{y}_1 \otimes \cdots \otimes \boldsymbol{y}_q) = \boldsymbol{x}_1 \otimes \cdots \otimes \boldsymbol{x}_p \otimes \boldsymbol{y}_1 \otimes \cdots \otimes \boldsymbol{y}_q$$

(а на остальные элементы продолжается по билинейности).

Таким образом, на всей прямой сумме $T(V)$ определена билинейная операция умножения $\otimes$ (между элементами), т.е. $T(V)$ — алгебра. В ней есть единица — это $1 \in K = T^0(V)$. Поскольку для любых $\boldsymbol{t} \in T^p(V)$ и $\boldsymbol{s} \in T^q(V)$ имеем $\boldsymbol{t} \otimes \boldsymbol{s} \in T^{p+q}(V)$, $T(V)$ является градуированной алгеброй с единицей. Она называется тензорной алгеброй пространства V.

Если в V зафиксирован базис $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$, то базис векторного пространства $T(V)$ составляют всевозможные элементы вида $e_{i_1} \otimes \cdots \otimes e_{i_k}$, где $k \in \mathbb{N}$ и $i_j \leqslant n$.

Точно так же определяется градуированная алгебра с единицей $T_*(V) = \bigoplus\limits_{p=0}^{\infty} T_p(V)$, где $T_p(V) = T_p^0(V) = (V^*)^{\otimes p} = T^p(V^*)$. Она является и называется алгеброй полилинейных функций на V.

Универсальное свойство тензорной алгебры

Имеем $V \underset{\text{lin}}{\subseteq} T(V)$. Пусть $i\colon V \hookrightarrow T(V)$ — тождественное вложение.

Универсальное свойство тензорной алгебры пространства V

V — подпространство пространства $T(V)$ и любое линейное отображение $f\colon V \to A$ в любую алгебру A над полем K единственным образом продолжается до гомоморфизма алгебр $\bar{f}\colon T(V) \to A$ (отображение $\varphi\colon A_1 \to A_2$ из алгебры A_1 с умножением $*_1$ в алгебру A_2 с умножением $*_2$ называется гомоморфизмом алгебр, если оно линейно и $\varphi(a *_1 b) = \varphi(a) *_2 \varphi(b)$ для любых $a, b \in A_1$):

$$\begin{tikzcd} V \arrow[r, hook, "i"] \arrow[dr, "f"'] & T(V) \arrow[d, dashed, "\bar{f}"] \\ & A \end{tikzcd}$$

Действительно, в любой алгебре A с произведением $*$ выполнены определяющие соотношения 1–4 тензорного произведения (с заменой $\otimes$ на $*$). Полагая $\bar{f}\colon \boldsymbol{x}_1 \otimes \cdots \otimes \boldsymbol{x}_k \mapsto f(\boldsymbol{x}_1) * \cdots * f(\boldsymbol{x}_k)$ для $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in V$ и продолжая $\bar{f}$ на $T(V)$ по линейности, мы получим корректно определенный гомоморфизм алгебр.

Замечание

Для каждого векторного пространства V алгебра с универсальным свойством единственна (с точностью до изоморфизма): если $T'(V)$ — другая обладающая этим свойством алгебра с умножением $*$, то вложение $i\colon V \to T(V)$ должно продолжаться единственным образом до гомоморфизма алгебр $\bar{i}\colon T'(V) \to T(V)$. Поскольку $\bar{i}$ — гомоморфизм, для любых $k \in \mathbb{N}$ и $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k \in V$ должно выполняться условие $\bar{i}(\boldsymbol{x}_1 * \dots * \boldsymbol{x}_k) = \boldsymbol{x}_1 \otimes \dots \otimes \boldsymbol{x}_k$. Из единственности $\bar{i}$ следует, что V порождает алгебру $T'(V)$ (если $\boldsymbol{t} \in T'(V)$ не представляется в виде линейной комбинации произведений элементов V, то $\bar{i}$ может отображать этот элемент куда угодно). Наконец, поскольку элементы вида $\boldsymbol{x}_1 \otimes \dots \otimes \boldsymbol{x}_k$ образуют базис пространства $T(V)$ и $\bar{i}$ линейно, заключаем, что $\bar{i}$ — изоморфизм векторных пространств, а значит, и алгебр.

Замечание

Из универсального свойства следует, что любое линейное отображение векторных пространств $V \to W$ продолжается единственным образом до гомоморфизма алгебр $T(V) \to T(W)$.

Алгебраические объекты с подобными универсальными свойствами называются *свободными*. В частности, $T(V)$ — *свободная алгебра* векторного пространства V.

Симметрическая алгебра

Определение

Пусть V — векторное пространство над K с базисом $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$. Векторное пространство S вместе с симметричным полилинейным отображением

$$\vee\colon \underbrace{V \times \dots \times V}_{k\ \text{раз}} \to T$$

с тем свойством, что $\{\vee(\boldsymbol{e}_{i_1}, \dots, \boldsymbol{e}_{i_k}) : 1 \leqslant i_1 \leqslant \dots \leqslant i_k \leqslant n\}$ — базис в S, называется *k-й симметрической степенью* пространства V и обозначается $S^k(V)$. Образ $\vee(\boldsymbol{x}_1, \dots, \boldsymbol{x}_k)$ произвольного набора векторов $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_k) \in V \times \dots \times V$ обозначается $\boldsymbol{x}_1 \vee \dots \vee \boldsymbol{x}_k$.

Симметрическая степень существует: достаточно взять векторное пространство S с базисом $\{\boldsymbol{t}_{i_1,\dots,i_k} : 1 \leqslant i_1 \leqslant \dots \leqslant i_k \leqslant n\}$ и определить полилинейное отображение $\vee\colon V_1, \dots, V_k \to S$, задав его на наборах базисных векторов формулой $\vee(\boldsymbol{e}_{j_1}, \dots, \boldsymbol{e}_{j_k}) = \boldsymbol{t}_{\tilde{j}_1,\dots,\tilde{j}_k}$, где $\tilde{j}_1, \dots, \tilde{j}_k$ — числа $j_1, \dots, j_k$, упорядоченные по неубыванию (и продолжив по полилинейности).

Определение симметрической степени не зависит от выбора базиса $\boldsymbol{E}$

Действительно, если $\boldsymbol{E}' = \{\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n\}$ — любой другой базис пространства V, то система векторов $\{\boldsymbol{e}'_{i_1} \vee \cdots \vee \boldsymbol{e}'_{i_k} : 1 \leqslant i_1 \leqslant \cdots \leqslant i_k \leqslant n\}$ полна в силу полилинейности отображения $\vee$, и число элементов в этой системе совпадает с числом элементов в базисе $\{\vee(\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_k}) : 1 \leqslant i_1 \leqslant \cdots \leqslant i_k \leqslant n\}$. Значит, она является базисом.

Симметрическая степень единственна

Для любой другой k-й симметрической степени $(\tilde{S}, \tilde{\vee})$ пространства V существует (единственный) изоморфизм $\varphi\colon \tilde{S} \to S$, удовлетворяющий условию $\varphi(\boldsymbol{x}_1 \tilde{\vee} \boldsymbol{x}_2 \tilde{\vee} \ldots \tilde{\vee} \boldsymbol{x}_k) = \boldsymbol{x}_1 \vee \boldsymbol{x}_2 \vee \cdots \vee \boldsymbol{x}_k$ для любых $\boldsymbol{x}_i \in V$: на базисе он определён правилом $\varphi(\boldsymbol{e}_{i_1} \tilde{\vee} \ldots \tilde{\vee} \boldsymbol{e}_{i_k}) = (\boldsymbol{e}_{i_1} \vee \cdots \vee \boldsymbol{e}_{i_k})$.

Упражнение

Распространите определение симметрической степени на пространства произвольной размерности.

Универсальное свойство симметрической степени

Отображение $\quad \vee\colon \underbrace{V\times\dots\times V}_{k\ \text{раз}} \to S^k(V), \quad (\boldsymbol{x}_1,\dots,\boldsymbol{x}_k) \mapsto \boldsymbol{x}_1\vee\dots\vee\boldsymbol{x}_k,$

полилинейно и симметрично, и для любого векторного пространства U над K и любого симметричного полилинейного отображения $\varphi\colon V\times\dots\times V \to U$ существует единственное линейное отображение $\ell\colon S^k(V)\to U$ со свойством $\varphi = \ell\circ\vee$, т.е. $\varphi(\boldsymbol{x}_1,\dots,\boldsymbol{x}_k) = \ell(\boldsymbol{x}_1\vee\dots\vee\boldsymbol{x}_k)$ для любых $\boldsymbol{x}_i\in V_i$:

$$\begin{tikzcd}
V\times\dots\times V \arrow[d, "\vee"'] \arrow[dr, "\varphi"] & \\
S^k(V) \arrow[r, "\ell_\varphi"'] & K
\end{tikzcd}$$

Универсальное свойство характеризует симметрическую степень: любое другое пространство с этим свойством совпадает с $S^k(V)$ с точностью до сохраняющего операцию $\vee$ изоморфизма.

Симметрическая степень $S^k(V)$ является факторпространством тензорного произведения $V^{\otimes k}$ по подпространству, порождённому всеми элементами вида

$$\boldsymbol{v}_1\otimes\dots\otimes\boldsymbol{v}_k - \boldsymbol{v}_{\pi(1)}\otimes\dots\otimes\boldsymbol{v}_{\pi(k)}, \quad \text{где } \pi\in S_k \text{ (перестановка).}$$

Симметрическая алгебра

Положим $S(V) = \bigoplus\limits_{q=0}^{\infty} S^q(V)$. Элементы $S(V)$ записывают не в виде бесконечных последовательностей с нулями, а в виде сумм

$$\boldsymbol{v}_1 \vee \dots \vee \boldsymbol{v}_k + \boldsymbol{w}_1 \vee \dots \vee \boldsymbol{w}_l + \dots + \boldsymbol{u}_1 \vee \dots \vee \boldsymbol{u}_m.$$

Элементы вида $\boldsymbol{v}_1 \vee \dots \vee \boldsymbol{v}_k$ называются разложимыми.

Для любых неотрицательных целых p и q определим билинейную операцию $\vee \colon S^p(V) \times S^q(V) \to S^{p+q}(V)$, задав её на разложимых элементах правилом

$$(\boldsymbol{v}_1 \vee \dots \vee \boldsymbol{v}_p) \vee (\boldsymbol{w}_1 \vee \dots \vee \boldsymbol{w}_q) = \boldsymbol{v}_1 \vee \dots \vee \boldsymbol{v}_p \vee \boldsymbol{w}_1 \vee \dots \vee \boldsymbol{w}_q$$

и продолжив на все прочие элементы по билинейности. Тем самым мы задали билинейную операцию $\vee$ на пространстве $S(V)$. Она превращает пространство $S(V)$ в градуированную коммутативную алгебру с единицей, которая называется симметрической алгеброй пространства V.

Так же определяется градуированная коммутативная алгебра $S_*(V) = \bigoplus\limits_{p=0}^{\infty} S_p(V)$, где $S_p(V) = S^p(V^*)$, симметричных полилинейных функций на V.

Универсальное свойство симметрической алгебры

V — подпространство пространства $S(V)$ и любое линейное отображение $f\colon V \to A$ в любую коммутативную алгебру A над полем K единственным образом продолжается до гомоморфизма алгебр $\bar{f}\colon T(V) \to A$:

$$\begin{tikzcd} V \arrow[r, hook, "i"] \arrow[dr, "f"'] & S(V) \arrow[d, dashed, "\bar{f}"] \\ & A \end{tikzcd}$$

(здесь i — тождественное вложение).

Другими словами, $S(V)$ — *свободная коммутативная алгебра* векторного пространства V.

Единственность алгебры с указанным универсальным свойством для каждого векторного пространства V доказывается аналогично некоммутативному случаю.

Замечание

Конструкция симметрической алгебры (вместе с универсальным свойством) без изменений переносится на пространства произвольной размерности.

Если пространство V имеет базис $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$, то $\{\boldsymbol{e}_{i_1} \vee \dots \vee \boldsymbol{e}_{i_p} : p \in \mathbb{N}, 1 \leqslant i_1 \leqslant \dots \leqslant i_p \leqslant n\}$ — базис пространства $S(V)$. Соответствие

$$\boldsymbol{e}_{i_1} \vee \dots \vee \boldsymbol{e}_{i_p} \mapsto t_{i_1} \cdot \ldots \cdot t_{i_p}, \qquad 1 \leqslant i_1 \leqslant \dots \leqslant i_p \leqslant n,$$

— изоморфизм между $S(V)$ и алгеброй многочленов $K[t_1, \dots, t_n]$.

Вспомнив, что каждый элемент $\boldsymbol{\varepsilon}^i$ взаимного базиса на V^* ставит в соответствие каждому вектору $\boldsymbol{x} \in V$ его i-ю координату x^i и отождествляя каждый $\boldsymbol{\varepsilon}^i$ с переменной x^i, мы тем самым отождествляем $S(V^*)$ с $K[x^1, \dots, x^n]$. Каждый элемент $\lambda_1 \boldsymbol{\varepsilon}^{i_{11}} \vee \dots \vee \boldsymbol{\varepsilon}^{i_{1m_1}} + \dots + \lambda_k \boldsymbol{\varepsilon}^{i_{1k}} \vee \dots \vee \boldsymbol{\varepsilon}^{i_{1m_k}}$ отождествляется с многочленом $\lambda_1 x^{i_{11}} \cdot \ldots \cdot x^{i_{1m_1}} + \dots + \lambda_k x^{i_{1k}} \cdot \ldots \cdot x^{i_{1m_k}}$.

Если K — поле характеристики 0, то каждое пространство $S^q(V)$ можно отождествить с подпространством $ST^q(V)$ симметричных тензоров в $T^q(V)$.

Предложение

Если K — поле характеристики 0, то для каждого $q \in \mathbb{N}$ отображение

$$\mu \colon S^q(V) \to ST^q(V), \qquad (\boldsymbol{x}_1 \vee \dots \vee \boldsymbol{x}_q) \mapsto \mathrm{Sym}(\boldsymbol{x}_1 \otimes \dots \otimes \boldsymbol{x}_q),$$

является изоморфизмом.

Каждой симметричной полилинейной функции $\boldsymbol{s} \in ST_p(V)$ поставим в соответствие однородный многочлен $p_{\boldsymbol{s}}(x^1, \dots, x^n) \in S_p(V)$:

$$p_{\boldsymbol{s}}(x^1, \dots, x^n) = \boldsymbol{s}(\boldsymbol{x}, \dots, \boldsymbol{x}) \qquad \text{для вектора } \boldsymbol{x} \in V \text{ с координатами } x^1, \dots, x^n.$$

Предложение

Если K — поле характеристики 0, то отображение

$$ST_p(V) \to S_p(V), \qquad \boldsymbol{s} \mapsto p_{\boldsymbol{s}},$$

является изоморфизмом векторных пространств, обратным отображению $\mu\colon S_p(V) \to ST_p(V)$.

Доказательство. Для любой симметричной полилинейной функции вида

$$\boldsymbol{s} = \mathrm{Sym}(\boldsymbol{f}^1 \otimes \dots \otimes \boldsymbol{f}^p) = \mu(\boldsymbol{f}^1 \vee \dots \vee \boldsymbol{f}^p), \qquad \text{где } \boldsymbol{f}^i \in V^*,\ \boldsymbol{f}^i = f^i_j \boldsymbol{\varepsilon}^j,$$

имеем
$$p_{\boldsymbol{s}}(x^1, \dots, x^n) = \boldsymbol{f}^1(\boldsymbol{x}) \cdot \ldots \cdot \boldsymbol{f}^p(\boldsymbol{x}) = (f^1_j x^j) \cdot \ldots \cdot (f^p_i x^i).$$

В $S_p(V) \subset S(V^*)$ этот многочлен отождествляется с элементом $(f^1_j \boldsymbol{\varepsilon}^j) \vee \ldots \vee (f^p_i \boldsymbol{\varepsilon}^i) = \boldsymbol{f}^1 \vee \dots \vee \boldsymbol{f}^p = \mu^{-1}(\boldsymbol{s})$. □

Симметричная полилинейная функция $\boldsymbol{s}$ называется **поляризацией** однородного многочлена $p_{\boldsymbol{s}}$.

Умножение в алгебре симметричных полилинейных функций $ST_*(V) = \bigoplus_{p=0}^{\infty} ST_p(V)$, соответствующее умножению в алгебре $S_*(V) = \bigoplus_{p=0}^{\infty} S_p(V)$, определяется формулой

$$(\boldsymbol{t} \vee \boldsymbol{s})(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+m}) = \frac{p!q!}{(p+q)!} \sum_{\pi \in S_{k+m}} \boldsymbol{t}(\boldsymbol{x}_{\pi(1)}, \ldots, \boldsymbol{x}_{\pi(k)}) \cdot \boldsymbol{s}(\boldsymbol{x}_{\pi(k+1)}, \ldots, \boldsymbol{x}_{\pi(k+m)})$$

для $\boldsymbol{t} \in ST_k(V)$ и $\boldsymbol{s} \in ST_m(V)$. Произведение $\boldsymbol{t} \vee \boldsymbol{s}$ называется симметрическим произведением функций $\boldsymbol{t}$ и $\boldsymbol{s}$.

Упражнение

Проверьте, что симметрическое произведение p линейных функций $\boldsymbol{f}^1, \ldots, \boldsymbol{f}^p \in V^*$ задаётся формулой

$$(\boldsymbol{f}^1 \vee \cdots \vee \boldsymbol{f}^p)(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p) = \frac{1}{p!} \operatorname{per}(\boldsymbol{f}^i(\boldsymbol{x}_j)),$$

где $\operatorname{per} A$ — перманент квадратной матрицы A, который определяется аналогично определителю с той разницей, что все слагаемые берутся со знаком плюс независимо от чётности перестановки.

Для полей ненулевой характеристики эти формулы не имеют смысла.

Определение

Пусть V — векторное пространство над K с базисом $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$. Векторное пространство Λ вместе с кососимметричным полилинейным отображением

$$\wedge\colon \underbrace{V \times \cdots \times V}_{k \text{ раз}} \to \Lambda$$

с тем свойством, что $\{\wedge(\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_k} : 1 \leqslant i_1 < \cdots < i_k \leqslant n\}$ — базис в Λ, называется *k-й внешней степенью* пространства V и обозначается $\bigwedge^k(V)$. Образ $\wedge(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k)$ произвольного набора векторов $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) \in V \times \cdots \times V$ обозначается $\boldsymbol{x}_1 \wedge \ldots \wedge \boldsymbol{x}_k$.

Внешняя степень существует, не зависит от базиса и единственна по тем же причинам, что и симметрическая степень. Для $k > p$ внешняя степень $\bigwedge^k(V)$ тривиальна.

Упражнение

Распространите определение внешней степени на пространства произвольной размерности.

Универсальное свойство внешней степени

Отображение $\wedge\colon \underbrace{V\times\cdots\times V}_{k\ \text{раз}} \to \bigwedge^k(V)$, $(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_k)\mapsto \boldsymbol{x}_1\wedge\ldots\wedge\boldsymbol{x}_k$,

полилинейно и кососимметрично, и для любого векторного пространства U над K и любого симметричного полилинейного отображения $\varphi\colon V\times\cdots\times V\to U$ существует единственное линейное отображение $\ell\colon \bigwedge^k(V)\to U$ со свойством $\varphi=\ell\circ\wedge$, т.е. $\varphi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_k)=\ell(\boldsymbol{x}_1\wedge\ldots\wedge\boldsymbol{x}_k)$ для любых $\boldsymbol{x}_i\in V_i$:

$$\begin{tikzcd}
V\times\cdots\times V \arrow[d, "\wedge"'] \arrow[dr, "\varphi"] & \\
S^k(V) \arrow[r, "\ell_\varphi"'] & K
\end{tikzcd}$$

Универсальное свойство характеризует внешнюю степень: любое другое пространство с этим свойством совпадает с $\bigwedge^k(V)$ с точностью до сохраняющего операцию $\wedge$ изоморфизма.

Внешняя степень $\bigwedge^k(V)$ является факторпространством тензорного произведения $V^{\otimes k}$ по подпространству, порождённому всеми элементами вида

$$\cdots\otimes\boldsymbol{v}\otimes\cdots\otimes\boldsymbol{v}\otimes\cdots\quad\text{(хотя бы два сомножителя совпадают).}$$

Алгебра Грассмана

Положим $\bigwedge(V) = \bigoplus\limits_{q=0}^{\infty} \bigwedge^q(V)$. Элементы $\bigwedge(V)$ записывают не в виде бесконечных последовательностей с нулями, а в виде сумм

$$\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_k + \boldsymbol{w}_1 \wedge \ldots \wedge \boldsymbol{w}_l + \cdots + \boldsymbol{u}_1 \wedge \ldots \wedge \boldsymbol{u}_m.$$

Элементы вида $\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_k$ называются поливекторами, или k-векторами.

Для любых неотрицательных целых p и q определим билинейную операцию $\wedge\colon \bigwedge^p(V) \times \bigwedge^q(V) \to \bigwedge^{p+q}(V)$, задав её на поливекторах правилом

$$(\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_p) \wedge (\boldsymbol{w}_1 \wedge \ldots \wedge \boldsymbol{w}_q) = \boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_p \wedge \boldsymbol{w}_1 \wedge \ldots \wedge \boldsymbol{w}_q$$

и продолжив на все прочие элементы по билинейности. Тем самым мы задали билинейную операцию $\wedge$ на пространстве $\bigwedge(V)$. Она превращает пространство $\bigwedge(V)$ в градуированную алгебру с единицей, которая называется алгеброй Грассмана, или внешней алгеброй, пространства V. Вместо коммутативности в ней имеет место градуированная антикоммутативность

$$\boldsymbol{u} \wedge \boldsymbol{v} = (-1)^{pq} \boldsymbol{v} \wedge \boldsymbol{u} \quad \text{для } \boldsymbol{u} \in \textstyle\bigwedge^p(V) \text{ и } \boldsymbol{v} \in \bigwedge^q(V).$$

Градуированные алгебры с этим свойством называются суперкоммутативными.

Так же определяется градуированная внешняя алгебра $\bigwedge_*(V) = \bigoplus \bigwedge_p(V)$, где $\bigwedge_p(V) = \bigwedge^p(V^*)$, кососимметричных полилинейных функций на V.

Универсальное свойство алгебры Грассмана

V — подпространство пространства $\bigwedge(V)$ и любое линейное отображение $f: V \to A$ в любую алгебру с единицей A над полем K с умножением $*$, обладающее свойством $f(\boldsymbol{v}) * f(\boldsymbol{v}) = 0$, единственным образом продолжается до гомоморфизма алгебр с единицей $\bar{f}: \bigwedge(V) \to A$:

$$\begin{tikzcd} V \arrow[r, hook, "i"] \arrow[dr, "f"'] & \bigwedge(V) \arrow[d, dashed, "\bar{f}"] \\ & A \end{tikzcd}$$

(здесь i — тождественное вложение).

Как и в случаях тензорной и симметрической алгебр, алгебра с указанным универсальным свойством для каждого векторного пространства V единственна.

Если $\boldsymbol{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ — базис V, то $\{\boldsymbol{e}_{i_1} \wedge \dots \wedge \boldsymbol{e}_{i_p} : 1 \leqslant i_1 < \dots < i_p \leqslant n\}$ — базис пространства $\bigwedge^p(V)$ и $\{\boldsymbol{e}_{i_1} \wedge \dots \wedge \boldsymbol{e}_{i_p} : p \in \mathbb{N},\ 1 \leqslant i_1 < \dots < i_p \leqslant n\}$ — базис пространства $\bigwedge(V)$. Следовательно,

$$\dim \textstyle\bigwedge^p(V) = \frac{n!}{p!}(n-p)! \quad \text{и} \quad \dim \textstyle\bigwedge(V) = 2^n.$$

Таким образом, алгебра Грассмана конечномерного пространства конечномерна.

Пусть K — поле характеристики 0.

Предложение

Для каждого $q \in \mathbb{N}$ отображение $\mu\colon \bigwedge^q(V) \to \bigwedge T^q(V)$, определённое правилом $(\boldsymbol{x}_1 \wedge \dots \wedge \boldsymbol{x}_q) \mapsto \operatorname{Alt}(\boldsymbol{x}_1 \otimes \dots \otimes \boldsymbol{x}_q)$, является изоморфизмом.

Умножение в алгебре симметричных полилинейных функций $ST_*(V) = \bigoplus ST_p(V)$, соответствующее умножению в алгебре $S_*(V) = \bigoplus S_p(V)$, определяется формулой

$$(\boldsymbol{t} \wedge \boldsymbol{s})(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{k+m}) = \frac{p!q!}{(p+q)!} \sum_{\pi \in S_{k+m}} \boldsymbol{t}(\boldsymbol{x}_{\pi(1)}, \dots, \boldsymbol{x}_{\pi(k)}) \cdot \boldsymbol{s}(\boldsymbol{x}_{\pi(k+1)}, \dots, \boldsymbol{x}_{\pi(k+m)}).$$

Произведение $\boldsymbol{t} \wedge \boldsymbol{s}$ называется внешним произведением функций $\boldsymbol{t}$ и $\boldsymbol{s}$.

Упражнение

Покажите, что для $\boldsymbol{f}^1, \dots, \boldsymbol{f}^p \in V^*$ $\quad (\boldsymbol{f}^1 \wedge \dots \wedge \boldsymbol{f}^p)(\boldsymbol{x}_1, \dots, \boldsymbol{x}_p) = \frac{1}{p!}\det(\boldsymbol{f}^i(\boldsymbol{x}_j))$.

Плюккеровы координаты

Пусть теперь K — поле характеристики $\neq 2$.

Теорема

1. *Система векторов* $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ *в пространстве* V *линейно зависима* $\iff$ $\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_p = \mathbf{0}$.
2. *Если системы векторов* $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ *и* $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_p\}$ *в* V *линейно независимы, то* $\langle\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}\rangle = \langle\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_p\}\rangle \iff$ p*-векторы* $\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_p$ *и* $\boldsymbol{w}_1 \wedge \ldots \wedge \boldsymbol{w}_p$ *пропорциональны.*

Доказательство. 1: Предположим, что $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ линейно зависимы. Пусть для определённости $\boldsymbol{v}_p = \lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_{p-1}\boldsymbol{v}_{p-1}$. Тогда

$$\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_{p-1} \wedge \boldsymbol{v}_p = \sum_{i=1}^{p-1} \lambda_i \boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_{p-1} \wedge \boldsymbol{v}_i = \mathbf{0}.$$

Если система $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ линейно независима, то её можно продолжить до базиса. Тогда p-вектор $\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_p$ станет одним из базисным. Значит, он не может быть нулевым.

Если $\langle\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}\rangle = \langle\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_p\}\rangle$, то векторы $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_p$ линейно выражаются через векторы $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$. Значит, p-вектор $\boldsymbol{w}_1 \wedge \ldots \wedge \boldsymbol{w}_p$ линейно выражается через p-векторы вида $\boldsymbol{v}_{i_1} \wedge \ldots \wedge \boldsymbol{v}_{i_p}$. Имеем

$$\boldsymbol{v}_{i_1} \wedge \ldots \wedge \boldsymbol{v}_{i_p} = \begin{cases} \pm \boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_p, & \text{если } i_1, \ldots, i_p \text{ различны,} \\ \boldsymbol{0} & \text{в противном случае.} \end{cases}$$

Следовательно, p-вектор $\boldsymbol{w}_1 \wedge \ldots \wedge \boldsymbol{w}_p$ пропорционален p-вектору $\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_p$.

Если $\langle\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}\rangle \neq \langle\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_p\}\rangle$, то в пространстве V существует базис $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ такой, что

$$\langle\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}\rangle = \langle\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_p\}\rangle \quad \text{и} \quad \langle\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_p\}\rangle = \langle\{\boldsymbol{e}_{d+1}, \ldots, \boldsymbol{e}_{d+p}\}\rangle,$$

где $1 \leqslant d \leqslant p$. По доказанному p-вектор $\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_p$ пропорционален p-вектору $\boldsymbol{e}_1 \wedge \ldots \wedge \boldsymbol{e}_p$ и $\boldsymbol{w}_1 \wedge \ldots \wedge \boldsymbol{w}_p$ пропорционален p-вектору $\boldsymbol{e}_{d+1} \wedge \ldots \wedge \boldsymbol{e}_{d+p}$. Однако p-векторы $\boldsymbol{e}_1 \wedge \ldots \wedge \boldsymbol{e}_p$ и $\boldsymbol{e}_{d+1} \wedge \ldots \wedge \boldsymbol{e}_{d+p}$ базисные и потому не пропорциональны.

□

Плюккеровы координаты подпространств

Пусть $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ — базис пространства V и $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ — базис его подпространства U. Подпространству U соответствует p-вектор $\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_p$, причём по теореме соответствие между подпространствами размерности p и классами пропорциональных p-векторов взаимно однозначно.

Найдём координаты p-вектора $\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_p$ в базисе $\{\boldsymbol{e}_{i_1} \wedge \ldots \wedge \boldsymbol{e}_{i_p} : 1 \leqslant i_1 < \cdots < i_p \leqslant n\}$ пространства $\bigwedge^p(V)$.

Пусть $A = (a_{ij})$ — матрица размера $p \times n$, образованная записанными по строкам координатами векторов $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ в базисе $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$:

$$\boldsymbol{v}_i = \sum_{j=1}^{n} a_{ij}\boldsymbol{e}_j, \qquad i = 1, \ldots, p.$$

Имеем

$$\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_p = \sum_{i_1, \ldots, i_p \leqslant n} a_{1i_1} \cdot \ldots \cdot a_{pi_p} \boldsymbol{e}_{i_1} \wedge \ldots \wedge \boldsymbol{e}_{i_p}.$$

Если среди индексов $i_1, \ldots, i_p$ есть одинаковые, то $\boldsymbol{e}_{i_1} \wedge \ldots \wedge \boldsymbol{e}_{i_p} = \mathbf{0}$. Если нет, то множители в p-векторе $\boldsymbol{e}_{i_1} \wedge \ldots \wedge \boldsymbol{e}_{i_p}$ можно переставить в порядке возрастания. При этом сам p-вектор умножится на $(-1)^s$, где s — число инверсий в последовательности $i_1, \ldots, i_p$. Значит,

$$\boldsymbol{v}_1 \wedge \ldots \wedge \boldsymbol{v}_p = \sum_{i_1 < \cdots < i_p} M_{i_1 \ldots i_p} \boldsymbol{e}_{i_1} \wedge \ldots \wedge \boldsymbol{e}_{i_p},$$

где $M_{i_1 \ldots i_p}$ — минор порядка p матрицы A, образованный столбцами $i_1, \ldots, i_p$.

По теореме числа $M_{i_1\dots i_p}$ однозначно определяют подпространство U. Они называются плюккеровыми координатами подпространства U (соответствующими базису $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$). Поскольку p-векторы определяют одно и то же подпространство $\iff$ они пропорциональны, плюккеровы координаты определены с точностью до пропорциональности.

Плюккеровы координаты не могут быть произвольными наборами чисел, так как они соответствуют p-векторам, которые составляют лишь часть пространства $\bigwedge^p(V)$. Числа $\mu_{i_1\dots i_p}$, $1 \leqslant i_1 < \dots < i_p \leqslant n$, являются плюккеровыми координатами некоторого p-мерного подпространства $\iff$ существует матрица размера $p \times n$, для которой эти числа являются минорами максимального порядка.

Упражнение

Докажите, что числа $\mu_{i_1\dots i_p}$, $1 \leqslant i_1 < \dots < i_p \leqslant n$, являются плюккеровыми координатами некоторого p-мерного подпространства $\iff$ они удовлетворяют соотношениям Плюккера

$$\sum_{r=1}^{p+1} (-1)^k \mu_{i_1\dots i_{p-1} j_k} \cdot \mu_{j_1\dots \check{j}_k \dots j_{p+1}} = 0 \quad \text{для любых } i_1, \dots, i_{p-1}, j_1, \dots, j_{p+1} \leqslant n$$

(знак $\check{}$ обозначает пропуск индекса).