# Линейная алгебра и геометрия

**ПАНОВ Тарас Евгеньевич**

курс лекций

Механико-математический факультет МГУ

# Оглавление


Предисловие 3

Глава 1. Линейные пространства 5
1.1. Линейные пространства и подпространства. Примеры 5
1.2. Линейная зависимость. Базис. Размерность 7
1.3. Пересечение и сумма подпространств, их размерности 10
1.4. Прямая сумма подпространств. Внешняя прямая сумма 11
1.5. Факторпространство. Размерность факторпространства 12
1.6. Координаты вектора. Закон изменения координат при замене базиса 14
1.7. Линейные отображения и изоморфизмы. Ядро и образ 16
1.8. Матрица линейного отображения. Преобразование матрицы линейного отображения при заменах базисов 18
1.9. Двойственное пространство $V^*$, двойственный базис. Отсутствие изоморфизма $V \cong V^*$ в бесконечномерном случае (пример) 19
1.10. Второе двойственное пространство, канонический изоморфизм $V \cong V^{**}$ 21
1.11. Сопряжённое линейное отображение 22

Глава 2. Линейные операторы 23
2.1. Матрица линейного оператора. Определитель и след оператора. Невырожденные операторы. Группы $GL_n$ и $SL_n$ 23
2.2. Проекторы, их алгебраическая характеризация 25
2.3. Многочлены от оператора. Минимальный аннулирующий многочлен 26
2.4. Овеществление и комплексификация 27
2.5. Инвариантные подпространства. Ограничение оператора и фактор-оператор. Собственные значения и собственные векторы 29
2.6. Характеристический многочлен. Теорема Гамильтона–Кэли 31
2.7. Диагонализируемые операторы. Критерий диагонализируемости 33
2.8. Нильпотентные операторы. Нормальный вид 35
2.9. Корневые векторы. Теорема о разложении в прямую сумму корневых подпространств 38
2.10. Жорданова нормальная форма оператора 40
2.11. Вычисление многочленов и функций от матриц. Экспонента линейного оператора (без обоснования сходимости) 41

Глава 3. Геометрия евклидовых и эрмитовых пространств 45
3.1. Аффинные пространства, системы координат, подпространства 45
3.2. Евклидовы и эрмитовы пространства, примеры. Неравенство Коши–Буняковского, неравенство треугольника 46

3.3. Ортогональные системы векторов, ортонормированные базисы. Ортогонализация Грама-Шмидта 49
3.4. Ортогональные и унитарные матрицы. $QR$-разложение 50
3.5. Ортогональное дополнение. Проекция и ортогональная составляющая. Угол между вектором и подпространством 52
3.6. Аффинные евклидовы пространства. Расстояние от точки до подпространства. Расстояние между подпространствами 53
3.7. Определитель матрицы Грама и многомерный объём 55
3.8. Метод наименьших квадратов 57
3.9. Изоморфизмы евклидовых и эрмитовых пространств. Канонический изоморфизм евклидова пространства и его сопряжённого 58

Глава 4. Операторы в евклидовых и эрмитовых пространствах 61
4.1. Сопряжённые операторы в евклидовых и эрмитовых пространствах 61
4.2. Самосопряжённые операторы. Канонический вид 62
4.3. Самосопряжённые проекторы. Спектральное разложение самосопряжённого оператора 64
4.4. Кососимметрические и косоэрмитовы операторы. Канонический вид. Эрмитово разложение 65
4.5. Ортогональные и унитарные операторы. Канонический вид. Группы $O_n$ и $SO_n$, $U_n$ и $SU_n$ 66
4.6. Положительные самосопряжённые операторы. Полярное разложение 70
4.7. Нормальные операторы 71

Глава 5. Билинейные и полуторалинейные функции 73
5.1. Билинейные и полуторалинейные функции, их матрицы. Закон изменения матрицы при замене базиса. Канонический изоморфизм пространства билинейных функций и пространства $\mathrm{Hom}(V, V^*)$ 73
5.2. Симметрические, кососимметрические и эрмитовы функции. Квадратичные формы. Нормальный вид 75
5.3. Нормальный вид кососимметрических и эрмитовых функций 79
5.4. Закон инерции. Единственность нормального вида 81
5.5. Теорема Якоби. Критерий Сильвестра 82
5.6. Симметрические билинейные функции в евклидовых пространствах. Канонический вид 84
5.7. Приведение пары форм к диагональному виду. Собственные значения и собственные векторы пары форм 85

Глава 6. Тензоры 87
6.1. Полилинейные функции 87
6.2. Тензоры: координатное определение 88
6.3. Тензорное произведение, свёртка, опускание и поднятие индексов 89
6.4. Базис в пространстве тензоров 91
6.5. Симметрические и кососимметрические тензоры, симметризация и альтернирование 92
6.6. Внешнее произведение кососимметрических тензоров, внешние формы 93

## Предисловие

Курс лекций подготовлен Пановым Тарасом Евгеньевичем, профессором кафедры высшей геометрии и топологии механико-математического факультета МГУ, в 2012 году.

В лекциях встречаются четыре типа утверждений: «теоремы», «леммы», «предложения» и «следствия». Теоремы — это наиболее важные утверждения, основные «кирпичики» курса. Леммы — это более технические утверждения, которые как правило являются промежуточными шагами в доказательствах теорем. Предложения — это как правило простые утверждения, которые тем не менее представляют самостоятельный интерес. Следствия — утверждения, которые непосредственно вытекают из предыдущих; они формулируются отдельно для последующих ссылок на них.

Бесконечномерные линейные пространства рассматриваются лишь в главе 1. Начиная с главы 2 все пространства предполагаются конечномерными. Изучение бесконечномерных пространств будет продолжено в курсе функционального анализа.

При подготовке данного курса лекций использовались классические учебники Гельфанда [**Ге**], Кострикина–Манина [**КМ**] и Постникова [**По**], а также предыдущие курсы лекций кафедры высшей геометрии и топологии (в частности, курсы профессоров И. А. Дынникова и В. М. Мануйлова, которые также доступны на сайте кафедры).

Данные лекции (а также программа экзамена и список теоретических задач) доступны на странице Т. Е. Панова на сайте кафедры высшей геометрии и топологии: `http://higeom.math.msu.su/people/taras/`

## Литература


[**Ге**] И. М. Гельфанд. *Лекции по линейной алгебре.* Москва, «Наука», 1974.

[**КМ**] А. И. Кострикин, Ю. И. Манин. *Линейная алгебра и геометрия.* Москва, «Наука», 1986.

[**Ма**] В. М. Мануйлов. *Курс лекций по линейной алгебре и геометрии.* `http://mech.math.msu.su/~manuilov/linalg.html`

[**По**] М. М. Постников. *Лекциии по геометрии. Семестр II. Линейная алгебра.* Москва, «Наука», 1986.

Глава 1

# Линейные пространства

## 1.1. Линейные пространства и подпространства. Примеры

Определение 1.1.1. *Линейным* (или *векторным*) *пространством* над полем $\mathbf{k}$ называется множество $V$ с заданными на нём операциями *сложения* «+» двух элементов множества $V$,

$$+\colon V \times V \to V, \qquad (\boldsymbol{u}, \boldsymbol{v}) \mapsto \boldsymbol{u} + \boldsymbol{v}$$

и *умножения* «$\cdot$» элементов $V$ на элементы поля $\mathbf{k}$,

$$\cdot\colon \mathbf{k} \times V \to V, \qquad (\lambda, \boldsymbol{v}) \mapsto \lambda \cdot \boldsymbol{v},$$

которые удовлетворяют следующим условиям:

1) $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$ для любых $\boldsymbol{u}, \boldsymbol{v} \in V$;
2) $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$ для любых $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$;
3) существует такой элемент $\mathbf{0} \in V$, что $\boldsymbol{v} + 0 = \boldsymbol{v}$ для любого $\boldsymbol{v} \in V$;
4) для любого $\boldsymbol{v} \in V$ существует такой элемент $-\boldsymbol{v} \in V$, что $\boldsymbol{v} + (-\boldsymbol{v}) = \mathbf{0}$;
5) $\lambda \cdot (\boldsymbol{u} + \boldsymbol{v}) = \lambda \cdot \boldsymbol{u} + \lambda \cdot \boldsymbol{v}$ для любых $\boldsymbol{u}, \boldsymbol{v} \in V$ и $\lambda \in \mathbf{k}$;
6) $(\lambda + \mu) \cdot \boldsymbol{v} = \lambda \cdot \boldsymbol{v} + \mu \cdot \boldsymbol{v}$ для любых $\boldsymbol{v} \in V$ и $\lambda, \mu \in \mathbf{k}$;
7) $\lambda \cdot (\mu \cdot \boldsymbol{v}) = (\lambda\mu) \cdot \boldsymbol{v}$ для любых $\boldsymbol{v} \in V$ и $\lambda, \mu \in \mathbf{k}$;
8) $1 \cdot \boldsymbol{v} = \boldsymbol{v}$ для любого $\boldsymbol{v} \in V$.

Элементы линейного пространства называются *векторами*. Свойства 1)–4) означают, что $V$ является абелевой (коммутативной) группой относительно операции сложения. Элемент $\mathbf{0}$ называется *нулевым вектором*, а элемент $(-\boldsymbol{v})$ называется *противоположным вектором* к $\boldsymbol{v}$.

Элементы поля $\mathbf{k}$ иногда называют *скалярами*. Свойства 5)–8) означают, что поле $\mathbf{k}$ *линейно действует* на $V$. Обычно бы будем опускать знак умножения $\cdot$.

За исключением некоторых примеров, в качестве поля $\mathbf{k}$ у нас будет выступать поле вещественных чисел $\mathbb{R}$ или поле комплексных чисел $\mathbb{C}$.

Далее говоря о *пространстве* мы будем иметь ввиду линейное пространство.

Вот некоторые простейшие свойства линейных пространств.

Предложение 1.1.2.

а) $0\boldsymbol{v} = \lambda\mathbf{0} = \mathbf{0}$ *для любых* $\boldsymbol{v} \in V$, $\lambda \in \mathbf{k}$*;*
б) $(-1)\boldsymbol{v} = -\boldsymbol{v}$ *для любого* $\boldsymbol{v} \in V$*;*
в) *если* $\lambda\boldsymbol{v} = \mathbf{0}$*, то либо* $\lambda = 0$*, либо* $\boldsymbol{v} = \mathbf{0}$*.*

Доказательство. Докажем а). Действительно, $0\boldsymbol{v} + 0\boldsymbol{v} = (0+0)\boldsymbol{v} = 0\boldsymbol{v}$, откуда $0\boldsymbol{v} = \mathbf{0}$ по свойству сокращения в абелевой группе. Аналогично, $\lambda\mathbf{0} + \lambda\mathbf{0} = \lambda(\mathbf{0} + \mathbf{0}) = \lambda\mathbf{0}$, т.е. $\lambda\mathbf{0} = \mathbf{0}$.

Докажем б). Действительно, $\boldsymbol{v} + (-1)\boldsymbol{v} = 1\boldsymbol{v} + (-1)\boldsymbol{v} = (1 + (-1))\boldsymbol{v} = 0\boldsymbol{v} = \mathbf{0}$, т.е. вектор $(-1)\boldsymbol{v}$ противоположен к $\boldsymbol{v}$.

Наконец, докажем в). Если $\lambda \neq 0$, то $\mathbf{0} = \lambda^{-1}(\lambda \boldsymbol{v}) = (\lambda^{-1}\lambda)\boldsymbol{v} = 1\boldsymbol{v} = \boldsymbol{v}$. □

Пример 1.1.3.

1. Множество $\{\mathbf{0}\}$, состоящее из одного элемента $\mathbf{0}$, является линейным пространством над любым полем.

2. Множества векторов на прямой, на плоскости, в пространстве, являются линейными пространствами над полем $\mathbb{R}$.

3. Поле $\mathbf{k}$ является векторным пространством над самим собой.

4. Поле $\mathbb{C}$ является линейным пространством над полем $\mathbb{R}$, а поле $\mathbb{R}$ является линейным пространством над полем рациональных чисел $\mathbb{Q}$. Более общо, если $\mathbf{k}_1$ — подполе в $\mathbf{k}_2$ (т.е. $\mathbf{k}_2$ является *расширением* поля $\mathbf{k}_1$), то $\mathbf{k}_2$ является линейным пространством над $\mathbf{k}_1$.

5. Пусть
$$\mathbf{k}^n := \big\{(x_1, \dots, x_n)\colon x_i \in \mathbf{k}\big\}$$
— множество последовательностей (*строк*) фиксированной длины $n$ из элементов поля $\mathbf{k}$. Операции покомпонентного сложения и умножения на скаляры, т.е.
$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) := (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n),$$
$$\lambda \cdot (x_1, x_2, \dots, x_n) := (\lambda x_1, \lambda x_2, \dots, \lambda x_n),$$
задают на $\mathbf{k}^n$ структуру линейного пространства над $\mathbf{k}$. Оно называется *$n$-мерным координатным* (или *арифметическим*) *пространством над* $\mathbf{k}$. Мы в основном будем иметь дело с пространствами $\mathbb{R}^n$ и $\mathbb{C}^n$.

При $n = 1$ мы получаем пространство из примера 3.

6. Множество функций на произвольном множестве $X$ со значениями в поле $\mathbf{k}$, обозначаемое $\mathbf{k}^X$, является линейным пространством относительно операций поточечного сложения (т.е. значение функции $f + g$ в точке $x \in X$ полагается равным $f(x) + g(x)$ и поточечного умножения на скаляры (т.е. $(\lambda \cdot f)(x) = \lambda f(x)$). В случае, когда $X$ — конечное множество из $n$ элементов, мы получаем пространство $\mathbf{k}^n$ из предыдущего примера.

7. Множество $C(\mathbb{R})$ непрерывных функций на вещественной прямой и множество $C[a, b]$ непрерывных функций на отрезке являются линейными пространствами над $\mathbb{R}$. Также линейными пространствами являются множества дифференцируемых функций (на прямой или на отрезке).

8. Множество решений однородной системы линейных уравнений является линейным пространством.

9. Рассмотрим множество $\mathbb{R}^\infty$, состоящее из бесконечных последовательностей вещественных чисел, в которых лишь конечное число членов отлично от нуля (такие последовательности называются *финитными*). Тогда $\mathbb{R}^\infty$ — линейное пространство относительно операций поэлементного сложения и умножения на числа. Пространство $\mathbb{R}^\infty$ можно отождествить с бесконечным объединением $\bigcup_{n \geqslant 0} \mathbb{R}^n$.

Пространство $\widehat{\mathbb{R}}^\infty$ *всех* бесконечных последовательностей также является линейным пространством.

10. Множество $\mathbf{k}[x]$ многочленов от одной переменной с коэффициентами в $\mathbf{k}$ является линейным пространством. Также линейным пространством является множество $\mathbf{k}_n[x]$ многочленов степени не выше $n$.

11. Множество всех матриц размера $m \times n$ с элементами из $\mathbf{k}$ образует линейное пространство $\mathrm{Mat}_{\mathbf{k}}(m,n)$ относительно операций сложения матриц и поэлементного умножения матриц на числа. При $m=1$ мы получаем пространство строк $\mathbf{k}^n$.

В предыдущих примерах мы столкнулись с ситуацией, когда подмножество линейного пространства само является линейным пространством. Это приводит к следующему определению.

Определение 1.1.4. Подмножество $W \subset V$ линейного пространства $V$ называется *подпространством*, если для любых векторов $\boldsymbol{u}, \boldsymbol{v} \in W$ и скаляра $\lambda \in \mathbf{k}$ мы имеем $\boldsymbol{u}+\boldsymbol{v} \in W$ и $\lambda\boldsymbol{u} \in W$. Другими словами, $W$ — подпространство, если $W$ само является линейным пространством относительно операций, заданных в пространстве $V$.

Пример 1.1.5. Вот некоторые примеры подпространств.

1. $\{\mathbf{0}\}$ является подпространством в любом пространстве $V$.

2. Множество векторов, коллинеарных заданному вектору, является подпространством в пространстве всех векторов на плоскости или в пространстве.

3. Пространство $C(\mathbb{R})$ непрерывных функций является подпространством в пространстве $\mathbb{R}^{\mathbb{R}}$ всех функций на $\mathbb{R}$.

4. Пространство $\mathbb{R}^\infty$ финитных последовательностей является подпространством в пространстве $\widehat{\mathbb{R}}^\infty$ всех последовательностей.

5. $\mathbf{k}_n[x]$ является подпространством в $\mathbf{k}_m[x]$ при $m \geqslant n$, а также в $\mathbf{k}[x]$.

## 1.2. Линейная зависимость. Базис. Размерность

Пусть $V$ — линейное пространство над полем $\mathbf{k}$.

Определение 1.2.1. *Линейной комбинацией* системы векторов $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ пространства $V$ называется сумма вида $\lambda_1\boldsymbol{v}_1 + \ldots + \lambda_k\boldsymbol{v}_k$, где $\lambda_i \in \mathbf{k}$.

*Линейной комбинацией* бесконечной системы векторов $\{\boldsymbol{v}_i \colon i \in I\}$ называется сумма вида $\sum_{i\in I} \lambda_i\boldsymbol{v}_i$, в которой лишь *конечное* число скаляров $\lambda_i$ отлично от нуля.

Линейная комбинация $\sum_{i\in I} \lambda_i\boldsymbol{v}_i$ называется *тривиальной*, если в ней все коэффициенты $\lambda_i$ равны нулю.

*Линейной оболочкой* системы векторов $\{\boldsymbol{v}_i \colon i \in I\}$ называется множество всех линейных комбинаций этой системы. Для линейной оболочки используется обозначение $\langle \boldsymbol{v}_i \colon i \in I\rangle$ или $\langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\rangle$ в случае конечной системы.

Система векторов $\{\boldsymbol{v}_i \colon i \in I\}$ (конечная или бесконечная) называется *линейно зависимой*, если существуют числа $\lambda_i$, не все равные нулю, такие, что $\sum_{i\in I} \lambda_i\boldsymbol{v}_i = \mathbf{0}$ (т.е. существует нетривиальная линейная комбинация векторов системы, равная нулю). В противном случае система называется *линейно независимой*.

Предложение 1.2.2. *Линейная оболочка $\langle \boldsymbol{v}_i \colon i \in I\rangle$ является линейным подпространством в $V$. Более того, $\langle \boldsymbol{v}_i \colon i \in I\rangle$ является наименьшим по включению линейным подпространством, содержащим все векторы системы $\{\boldsymbol{v}_i \colon i \in I\}$.*

Доказательство. Сумма векторов системы и результат умножения вектора системы на скаляр являются линейными комбинациями и потому принадлежат линейной оболочке. Следовательно, $\langle \boldsymbol{v}_i \colon i \in I\rangle$ — подпространство. Если $W$ — произвольное подпространство, содержащее все векторы из $\{\boldsymbol{v}_i \colon i \in I\}$, то $W$ также содержит все их линейные комбинации, а значит $W$ содержит $\langle \boldsymbol{v}_i \colon i \in I\rangle$. $\square$

Лемма 1.2.3. *Если система векторов $\{\boldsymbol{v}_i\colon i\in I\}$ линейно зависима, то один из векторов системы является линейной комбинацией остальных.*

Доказательство. Пусть $\lambda_1\boldsymbol{v}_1+\ldots+\lambda_k\boldsymbol{v}_k=0$, причем существует $\lambda_i\neq 0$. Тогда

$$\lambda_i\boldsymbol{v}_i=-\lambda_1\boldsymbol{v}_1-\ldots-\lambda_{i-1}\boldsymbol{v}_{i-1}-\lambda_{i+1}\boldsymbol{v}_{i+1}-\ldots-\lambda_k\boldsymbol{v}_k.$$

Умножив обе части этого равенства на $\lambda_i^{-1}$, получим, что $\boldsymbol{v}_i$ является линейной комбинацией векторов $\boldsymbol{v}_1,\ldots,\boldsymbol{v}_{i-1},\boldsymbol{v}_{i+1},\ldots,\boldsymbol{v}_k$. □

Определение 1.2.4. Линейно независимая система векторов $\{\boldsymbol{v}_i\colon i\in I\}$ называется *базисом* пространства $V$, если каждый вектор $\boldsymbol{v}\in V$ представляется в виде линейной комбинации $\sum_{i\in I}\lambda_i\boldsymbol{v}_i$. Другими словами, базисом называется *максимальная* (по включению) линейно независимая система векторов в пространстве $V$.

Пространство $V$ называется *конечномерным*, если в нём существует базис, состоящий из конечного числа векторов. В противном случае пространство называется *бесконечномерным*.

Предложение 1.2.5. *Если $\{\boldsymbol{v}_i\colon i\in I\}$ — базис пространства $V$, то представление любого вектора $\boldsymbol{v}\in V$ в виде линейной комбинации $\sum_{i\in I}\lambda_i\boldsymbol{v}_i$ единственно.*

Доказательство. Действительно, если $\boldsymbol{v}=\sum_{i\in I}\lambda_i\boldsymbol{v}_i=\sum_{i\in I}\mu_i\boldsymbol{v}_i$, то получаем $\mathbf{0}=\sum_{i\in I}(\lambda_i-\mu_i)\boldsymbol{v}_i$. Так как система $\{\boldsymbol{v}_i\colon i\in I\}$ линейно независима, из последнего равенства вытекает, что $\lambda_i=\mu_i$, т.е. два представления $\boldsymbol{v}$ в виде линейных комбинаций совпадают. □

Теорема 1.2.6. *В конечномерном пространстве все базисы состоят из одного числа элементов.*

Доказательство этой теоремы будет опираться на следующую лемму.

Лемма 1.2.7. *Пусть $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_m$ и $\boldsymbol{f}_1,\ldots,\boldsymbol{f}_n$ — две (конечных) линейно независимых системы векторов, причём вторая система содержится в линейной оболочке первой системы. Тогда $n\leqslant m$.*

Доказательство. Пусть $\boldsymbol{f}_j=a_{1j}\boldsymbol{e}_1+\ldots+a_{mj}\boldsymbol{e}_m,\ \ a_{ij}\in\mathbf{k},\ \ j=1,\ldots,n$. Так как $\boldsymbol{f}_1,\ldots,\boldsymbol{f}_n$ — линейно независимая система, мы имеем

$$x_1\boldsymbol{f}_1+\ldots+x_n\boldsymbol{f}_n=\mathbf{0}\quad\Longleftrightarrow\quad x_1=\ldots=x_n=0. \tag{1.1}$$

Подставляя в линейную комбинацию (1.1) выражения $\boldsymbol{f}_i$ через $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_m$, получаем:

$$\begin{aligned}\mathbf{0}&=x_1(a_{11}\boldsymbol{e}_1+\ldots+a_{m1}\boldsymbol{e}_m)+\ldots+x_n(a_{1n}\boldsymbol{e}_1+\ldots+a_{mn}\boldsymbol{e}_m)=\\&=(a_{11}x_1+\ldots+a_{1n}x_n)\boldsymbol{e}_1+\ldots+(a_{m1}x_1+\ldots+a_{mn}x_n)\boldsymbol{e}_m.\end{aligned}$$

Так как $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_m$ — линейно независимая система, предыдущее равенство равносильно системе уравнений:

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

Если $n>m$, то эта система имеет ненулевое решение, что противоречит (1.1). □

Доказательство теоремы 1.2.6. Пусть $V$ — конечномерное пространство. По определению, в $V$ существует конечный базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$. Пусть $\{\boldsymbol{f}_i \colon i \in I\}$ — другой базис. Если это базис бесконечен, то в нём содержится конечная линейно независимая система $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$, где $n > m$. При этом, так как $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ — базис, мы имеем $\{\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n\} \subset \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_m \rangle$, что противоречит лемме 1.2.7. Следовательно базис $\{\boldsymbol{f}_i \colon i \in I\}$ конечен, т.е. имеет вид $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$. Тогда $\{\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n\} \subset \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_m \rangle$ и $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m\} \subset \langle \boldsymbol{f}_1, \ldots, \boldsymbol{f}_n \rangle$, и из леммы 1.2.7 вытекает, что $m = n$. □

Определение 1.2.8. *Размерностью* конечномерного линейного пространства $V$ (обозначение: $\dim V$) называется число элементов в любом базисе $V$. Если же $V$ бесконечномерно, то мы пишем $\dim V = \infty$.

Размерность линейной оболочки системы векторов $\{\boldsymbol{e}_i \colon i \in I\}$ называется *рангом* системы векторов.

Замечание. В пространстве $\{\mathbf{0}\}$ базисом естественно считать пустое множество $\varnothing$. Мы имеем $\dim\{\mathbf{0}\} = 0$, так как пустое множество состоит из 0 элементов.

Предложение 1.2.9. *Подпространство $W$ конечномерного пространства $V$ конечномерно, причём $\dim W \leqslant \dim V$, и равенство достигается только при $W = V$.*

Доказательство. Пусть $\dim V = m$ и $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ — базис пространства $V$. Если $\dim W > m$, то в $W$ найдётся линейно независимая система $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ с $n > m$. Тогда $\{\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n\} \subset \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_m \rangle = V$, что противоречит лемме 1.2.7. Следовательно, $\dim W \leqslant \dim V$.

Пусть $\dim W = \dim V = m$ и пусть $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m$ — базис в $W$. Тогда каждый вектор $\boldsymbol{e}_i$ линейно выражается через $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m$, так как иначе мы бы получили линейно независимую систему $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m, \boldsymbol{e}_i$ из $m+1$ векторов в $V$, что противоречит теореме 1.2.6. Следовательно, любой вектор из $V$ лежит в $\langle \boldsymbol{f}_1, \ldots, \boldsymbol{f}_m \rangle = W$, т.е. $V = W$. □

Теорема 1.2.10. *Любой базис подпространства $W$ конечномерного пространства $V$ можно дополнить до базиса всего пространства $V$.*

Доказательство. Согласно предложению 1.2.9, пространство $W$ конечномерно; пусть $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r$ — его базис. Если $W = V$, то $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r$ — базис в $V$ и доказывать нечего. В противном случае в $V$ найдётся вектор $\boldsymbol{e}_{r+1} \notin \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_r \rangle = W$. Рассмотрим подпространство $W_1 = \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_r, \boldsymbol{e}_{r+1} \rangle \subset V$. Если $W_1 = V$, то всё доказано. В противном случае аналогично строим подпространство $W_2 = \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_r, \boldsymbol{e}_{r+1}, \boldsymbol{e}_{r+2} \rangle \subset V$, и так далее. Пусть $k = \dim V - \dim W$. Тогда на $k$-м шаге мы получим подпространство $W_k \subset V$ с $\dim W_k = \dim V$. Согласно предложению 1.2.9, $W_k = V$, а значит мы дополнили базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r$ в $W$ до базиса в $V$ векторами $\boldsymbol{e}_{r+1}, \ldots, \boldsymbol{e}_{r+k}$. □

Замечание. На самом деле предыдущая теорема имеет место и в бесконечномерном случае. В частности, в любом пространстве (даже бесконечномерном) существует базис. Доказательство этого факта, хотя и не сложно, использует абстрактные теоретико-множественные построения (*лемму Цорна*), которые выходят за рамки данного курса. Подробности можно найти в [**КМ**, §1.2].

Пример 1.2.11.

1. В арифметическом пространстве $\mathbf{k}^n$ имеется *стандартный* базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$, где $\boldsymbol{e}_i = (0, \ldots, 1, \ldots, 0)$ — строка, в которой на $i$-м месте стоит 1, а на остальных местах — нули. Таким образом, $\dim \mathbf{k}^n = n$.

2. В пространстве $\mathbf{k}_n[x]$ многочленов степени $\leqslant n$ имеется базис из одночленов $1, x, x^2, \ldots, x^n$. Таким образом, $\dim \mathbf{k}_n[x] = n + 1$. В пространстве $\mathbf{k}[x]$ всех многочленов имеется бесконечный базис из одночленов $1, x, x^2, x^3, \ldots$ всех степеней. Таким образом, $\dim \mathbf{k}[x] = \infty$.

3. В пространстве финитных последовательностей $\mathbb{R}^\infty$ имеется бесконечный базис $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots$, где $\boldsymbol{e}_i = (0, \ldots, 0, 1, 0, \ldots)$ — последовательность, в которой на $i$-м месте стоит 1, а на остальных местах — нули. Заметим, что эта же система $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots$ *не является* базисом в пространстве $\widehat{\mathbb{R}}^\infty$ всех последовательностей. Действительно, например, последовательность, состоящая из одних единиц не представляется в виде (конечной) линейной комбинации последовательностей $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots$.

**Далее все пространства мы будем предполагать конечномерными, если явно не указано противное.** Бесконечномерным пространствам будет посвящён отдельный курс функционального анализа. В этом курсе линейной алгебры мы лишь иногда будем встречаться с ними в примерах.

### 1.3. Пересечение и сумма подпространств, их размерности

Предложение 1.3.1. *Пересечение $V_1 \cap V_2$ подпространств пространства $V$ также является подпространством.*

Доказательство. Для любых $\boldsymbol{u}, \boldsymbol{v} \in V_1 \cap V_2$ и $\lambda \in \mathbf{k}$ сумма $\boldsymbol{u} + \boldsymbol{v}$ и произведение $\lambda \boldsymbol{v}$ также лежат и в $V_1$, и в $V_2$, а значит и в пересечении $V_1 \cap V_2$. □

В отличие от пересечения, объединение подпространств $V_1 \cup V_2$ в общем случае не будет линейным подпространством.

Определение 1.3.2. *Суммой* $V_1 + V_2$ подпространств $V_1$ и $V_2$ пространства $V$ называется множество всех векторов $\boldsymbol{v} \in V$, которые можно представить в виде суммы $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$, где $\boldsymbol{v}_1 \in V_1$ и $\boldsymbol{v}_2 \in V_2$.

Предложение 1.3.3. *Сумма подпространств является линейной оболочкой их объединения: $V_1 + V_2 = \langle V_1 \cup V_2 \rangle$. Таким образом, $V_1 + V_2$ является линейным подпространством.*

Доказательство. Включение $V_1 + V_2 \subset \langle V_1 \cup V_2 \rangle$ следует из того, что вектор $\boldsymbol{v}_1 + \boldsymbol{v}_2$ является линейной комбинацией векторов $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V_1 \cup V_2$.

Докажем обратное включение $\langle V_1 \cup V_2 \rangle \subset V_1 + V_2$. Рассмотрим линейную комбинацию $\boldsymbol{v} = \lambda_1 \boldsymbol{u}_1 + \ldots + \lambda_n \boldsymbol{u}_n$ векторов $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in V_1 \cup V_2$. Можно считать, что $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ лежат в $V_1$, а $\boldsymbol{u}_{k+1}, \ldots, \boldsymbol{u}_n$ лежат в $V_2$. Тогда мы имеем $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$, где $\boldsymbol{v}_1 = \lambda_1 \boldsymbol{u}_1 + \ldots + \lambda_k \boldsymbol{u}_k \in V_1$ и $\boldsymbol{v}_2 = \lambda_{k+1} \boldsymbol{u}_{k+1} + \ldots + \lambda_n \boldsymbol{u}_n \in V_2$. Следовательно, $\boldsymbol{v} \in V_1 + V_2$. □

Теорема 1.3.4. *Для любых подпространств $V_1$ и $V_2$ линейного пространства $V$ имеет место равенство*

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

Доказательство. Выберем базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$ пространства $V_1 \cap V_2$. Воспользовавшись теоремой 1.2.10, дополним его до базиса $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_l$ пространства $V_1$ и до базиса $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k, \boldsymbol{g}_1, \ldots, \boldsymbol{g}_m$ пространства $V_2$. Тогда мы имеем

$$\dim(V_1 \cap V_2) = k, \quad \dim V_1 = k + l, \quad \dim V_2 = k + m. \tag{1.2}$$

Докажем, что $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_l, \boldsymbol{g}_1, \ldots, \boldsymbol{g}_m$ — базис пространства $V_1 + V_2$.

Прежде всего заметим, что так как $V_1 + V_2 = \langle V_1 \cup V_2 \rangle$, любой вектор из $V_1 + V_2$ линейно выражается через эту систему векторов. Остаётся проверить, что эта система линейно независима. Пусть имеет место равенство

$$\lambda_1 \boldsymbol{e}_1 + \ldots + \lambda_k \boldsymbol{e}_k + \mu_1 \boldsymbol{f}_1 + \ldots + \mu_l \boldsymbol{f}_l + \nu_1 \boldsymbol{g}_1 + \ldots + \nu_m \boldsymbol{g}_m = \mathbf{0}. \tag{1.3}$$

Перепишем его в виде

$$\lambda_1 \boldsymbol{e}_1 + \ldots + \lambda_k \boldsymbol{e}_k + \mu_1 \boldsymbol{f}_1 + \ldots + \mu_l \boldsymbol{f}_l = -\nu_1 \boldsymbol{g}_1 - \ldots - \nu_m \boldsymbol{g}_m.$$

Вектор, стоящий в обеих частях этого равенства, лежит как в $V_1$, так и в $V_2$. Следовательно, он лежит в $V_1 \cap V_2$ и линейно выражается через $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$. Так как векторы $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_l$ линейно независимы по построению, мы получаем, что $\mu_1 = \ldots = \mu_l = 0$. Аналогичным образом доказывается, что $\nu_1 = \ldots = \nu_m = 0$. Тогда из линейной независимости $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$ и (1.3) следует, что $\lambda_1 = \ldots = \lambda_k = 0$.

Итак, система $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_l, \boldsymbol{g}_1, \ldots, \boldsymbol{g}_m$ порождает пространство $V_1 + V_2$ и линейно независима. Следовательно, это — базис в $V_1 + V_2$ и $\dim(V_1 + V_2) = k + l + m$. Отсюда и из (1.2) вытекает требуемое равенство. □

## 1.4. Прямая сумма подпространств. Внешняя прямая сумма

Определение 1.4.1. Сумма $V_1 + V_2$ подпространств пространства $V$ называется *прямой* (обозначение: $V_1 \oplus V_2$), если для любого вектора $\boldsymbol{v} \in V_1 + V_2$ представление $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$, где $\boldsymbol{v}_1 \in V_1$ и $\boldsymbol{v}_2 \in V_2$, единственно.

Теорема 1.4.2. *Следующие условия эквивалентны для подпространств $V_1$, $V_2$:*

а) *сумма $V_1 + V_2$ прямая;*
б) $V_1 \cap V_2 = \{\mathbf{0}\}$.
в) *если $\mathbf{0} = \boldsymbol{v}_1 + \boldsymbol{v}_2$, где $\boldsymbol{v}_1 \in V_1$ и $\boldsymbol{v}_2 \in V_2$, то $\boldsymbol{v}_1 = \boldsymbol{v}_2 = \mathbf{0}$;*
г) $\dim V_1 + \dim V_2 = \dim(V_1 + V_2)$*;*

Доказательство. Мы докажем импликации а)⇒ б)⇒ в)⇒ а) и б)⇔ г).

а)⇒ б). Пусть существует $\boldsymbol{v} \in V_1 \cap V_2$, $\boldsymbol{v} \neq \mathbf{0}$. Тогда $\mathbf{0} = \mathbf{0} + \mathbf{0} = \boldsymbol{v} + (-\boldsymbol{v})$, где $\boldsymbol{v} \in V_1$ и $\boldsymbol{v} \in V_2$. Получаем, что представление вектора $\mathbf{0}$ в виде суммы векторов из $V_1$ и $V_2$ не единственно, т.е. сумма $V_1 + V_2$ не прямая.

б)⇒ в). Если существует представление $\mathbf{0} = \boldsymbol{v}_1 + \boldsymbol{v}_2$, где $\boldsymbol{v}_1 \in V_1$, $\boldsymbol{v}_2 \in V_2$ и $\boldsymbol{v}_1 \neq \mathbf{0}$, то $\boldsymbol{v}_1 \in V_1$ и $\boldsymbol{v}_1 = -\boldsymbol{v}_2 \in V_2$, т.е. $\boldsymbol{v}_1 \in V_1 \cap V_2$ и $V_1 \cap V_2 \neq \{\mathbf{0}\}$. Противоречие.

в)⇒ а). Пусть у вектора $\boldsymbol{v} \in V$ есть два разложения: $\boldsymbol{v} = \boldsymbol{u}_1 + \boldsymbol{u}_2 = \boldsymbol{v}_1 + \boldsymbol{v}_2$, где $\boldsymbol{u}_1, \boldsymbol{v}_1 \in V_1$ и $\boldsymbol{u}_2, \boldsymbol{v}_2 \in V_2$. Тогда $\mathbf{0} = (\boldsymbol{u}_1 - \boldsymbol{v}_1) + (\boldsymbol{u}_2 - \boldsymbol{v}_2)$, где $\boldsymbol{u}_1 - \boldsymbol{v}_1 \in V_1$ и $\boldsymbol{u}_2 - \boldsymbol{v}_2 \in V_2$. Следовательно, $\boldsymbol{u}_1 - \boldsymbol{v}_1 = \boldsymbol{u}_2 - \boldsymbol{v}_2 = \mathbf{0}$, т.е. два разложения вектора $\boldsymbol{v}$ совпадают.

б)⇔ г). Эта эквивалентность вытекает из теоремы 1.3.4, так как лишь тривиальное пространство может иметь нулевую размерность. □

Понятие прямой суммы обобщается на любое конечное число подпространств:

Определение 1.4.3. Сумма $V_1 + \ldots + V_n$ подпространств пространства $V$ называется *прямой*, если для любого вектора $\boldsymbol{v} \in V_1 + \ldots + V_n$ представление $\boldsymbol{v} = \boldsymbol{v}_1 + \ldots + \boldsymbol{v}_n$, где $\boldsymbol{v}_i \in V_i$, единственно.

Теорема 1.4.4. *Для подпространств $V_1, \ldots, V_n$ пространства $V$ следующие условия эквивалентны:*

а) *сумма* $V_1 + \ldots + V_n$ *прямая;*
б) $V_i \cap (V_1 + \ldots + V_{i-1} + V_{i+1} + \ldots + V_n) = \{\mathbf{0}\}$ *для любого* $i = 1, \ldots, n$*;*
в) $V_i \cap (V_{i+1} + \ldots + V_n) = \{\mathbf{0}\}$ *для любого* $i = 1, \ldots, n-1$*;*
г) *если* $\mathbf{0} = \boldsymbol{v}_1 + \ldots + \boldsymbol{v}_n$*, где* $\boldsymbol{v}_i \in V_i$*, то* $\boldsymbol{v}_1 = \ldots = \boldsymbol{v}_n = \mathbf{0}$*;*
д) $\dim V_1 + \ldots + \dim V_n = \dim(V_1 + \ldots + V_n)$.

Доказательство теоремы проводится по индукции по $n$; мы его опускаем.

Пример 1.4.5. При $n \geqslant 3$ условие б) в предыдущей теореме более сильное, чем условие $V_i \cap V_j = \{\mathbf{0}\}$ при $1 \leqslant i < j \leqslant n$. Это последнее условие не гарантирует, что сумма подпространств прямая. Действительно, рассмотрим следующие три подпространства в $\mathbb{R}^2$ со стандартным базисом $\boldsymbol{e}_1, \boldsymbol{e}_2$:

$$V_1 = \{\boldsymbol{e}_1\}, \quad V_2 = \{\boldsymbol{e}_2\}, \quad V_3 = \{\boldsymbol{e}_1 + \boldsymbol{e}_2\}$$

(можно также взять любые три различные прямые, пересекающиеся в нуле). Тогда $V_i \cap V_j = \{\mathbf{0}\}$ при $i \neq j$, но сумма $V_1 + V_2 + V_3$ не прямая, так как, например, вектор $\boldsymbol{e}_1 + \boldsymbol{e}_2$ допускает два различных представления в виде суммы $\boldsymbol{v}_1 + \boldsymbol{v}_2 + \boldsymbol{v}_3$ с $\boldsymbol{v}_i \in V_i$, а именно $\boldsymbol{e}_1 + \boldsymbol{e}_2 = \boldsymbol{e}_1 + \boldsymbol{e}_2 + \mathbf{0} = \mathbf{0} + \mathbf{0} + (\boldsymbol{e}_1 + \boldsymbol{e}_2)$. Заметим также, что $\dim V_i = 1$, а $\dim(V_1 + V_2 + V_3) = \dim \mathbb{R}^2 = 2$.

Определение 1.4.6. Пусть $V_1, \ldots, V_n$ — линейные пространства над одним полем $\mathbf{k}$. Их *внешней прямой суммой*, обозначаемой через $V_1 \oplus \ldots \oplus V_n$, называется линейное пространство, состоящее из всех упорядоченных наборов $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$, где $\boldsymbol{v}_i \in V_i$ с операциями, определёнными покомпонентно:

$$(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) + (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) := (\boldsymbol{u}_1 + \boldsymbol{v}_1, \ldots, \boldsymbol{u}_n + \boldsymbol{v}_n),$$
$$\lambda \cdot (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) := (\lambda \boldsymbol{v}_1, \ldots, \lambda \boldsymbol{v}_n).$$

## 1.5. Факторпространство. Размерность факторпространства

Пусть $V$ — линейное пространство, а $W \subset V$ — подпространство.

Определение 1.5.1. *Классом смежности* вектора $\boldsymbol{v} \in V$ по подпространству $W$ называется множество $\boldsymbol{v} + W$, состоящее из всех векторов вида $\boldsymbol{v} + \boldsymbol{w}$, где $\boldsymbol{w} \in W$.

Лемма 1.5.2. *Равенство* $\boldsymbol{v}_1 + W = \boldsymbol{v}_2 + W$ *имеет место тогда и только тогда, когда* $\boldsymbol{v}_1 - \boldsymbol{v}_2 \in W$.

Доказательство. Пусть $\boldsymbol{v}_1 + W = \boldsymbol{v}_2 + W$. Тогда $\boldsymbol{v}_1 \in \boldsymbol{v}_1 + W = \boldsymbol{v}_2 + W$, а значит найдётся такой вектор $\boldsymbol{w} \in W$, что $\boldsymbol{v}_1 = \boldsymbol{v}_2 + \boldsymbol{w}$. Следовательно, $\boldsymbol{v}_1 - \boldsymbol{v}_2 = \boldsymbol{w} \in W$.

Обратно, пусть $\boldsymbol{v} := \boldsymbol{v}_1 - \boldsymbol{v}_2 \in W$. Докажем, что $\boldsymbol{v}_1 + W \subset \boldsymbol{v}_2 + W$. Возьмём произвольный вектор $\boldsymbol{u} \in \boldsymbol{v}_1 + W$. Тогда $\boldsymbol{u} = \boldsymbol{v}_1 + \boldsymbol{w}$ для некоторого $\boldsymbol{w} \in W$. Мы имеем $\boldsymbol{u} = \boldsymbol{v}_1 + \boldsymbol{w} = \boldsymbol{v}_2 + (\boldsymbol{v} + \boldsymbol{w})$, где $\boldsymbol{v} + \boldsymbol{w} \in W$. Следовательно, $\boldsymbol{u} \in \boldsymbol{v}_2 + W$ и $\boldsymbol{v}_1 + W \subset \boldsymbol{v}_2 + W$. Противоположное включение $\boldsymbol{v}_2 + W \subset \boldsymbol{v}_1 + W$ доказывается аналогично. Итак, $\boldsymbol{v}_1 + W = \boldsymbol{v}_2 + W$. □

Определение 1.5.3. *Факторпространством* $V/W$ линейного пространства $V$ по подпространству $W$ называется множество $\{\boldsymbol{v} + W : \boldsymbol{v} \in V\}$ всех классов смежности по подпространству $W$ с операциями сложения и умножения на скаляры:

$$(\boldsymbol{u} + W) + (\boldsymbol{v} + W) := (\boldsymbol{u} + \boldsymbol{v}) + W,$$
$$\lambda \cdot (\boldsymbol{v} + W) := \lambda \boldsymbol{v} + W.$$

Предложение 1.5.4. *Приведённые выше операции определены на классах смежности корректно и задают на $V/W$ структуру линейного пространства.*

Доказательство. Вначале проверим корректность определения операций, т.е. независимость результата операции от выбора вектора $\boldsymbol{v}$ в смежном классе $\boldsymbol{v}+W$. Докажем корректность для сложения. Если $\boldsymbol{u}_1+W=\boldsymbol{u}_2+W$ и $\boldsymbol{v}_1+W=\boldsymbol{v}_2+W$, то $\boldsymbol{u}:=\boldsymbol{u}_1-\boldsymbol{u}_2\in W$ и $\boldsymbol{v}:=\boldsymbol{v}_1-\boldsymbol{v}_2\in W$ в силу леммы 1.5.2. Следовательно,

$$(\boldsymbol{u}_1+W)+(\boldsymbol{v}_1+W)=(\boldsymbol{u}_1+\boldsymbol{v}_1)+W=(\boldsymbol{u}_2+\boldsymbol{v}_2)+(\boldsymbol{u}+\boldsymbol{v})+W=\\=(\boldsymbol{u}_2+\boldsymbol{v}_2)+W=(\boldsymbol{u}_2+W)+(\boldsymbol{v}_2+W),$$

т.е. сложение определено корректно. Корректность определения умножения на скаляры проверяется аналогично.

Теперь докажем, что $V/W$ — линейное пространство. Свойства 1) и 2) из определения 1.1.1 (ассоциативность и коммутативность сложения для смежных классов) вытекают из соответствующих свойств сложения в $V$. Нулевым элементом в $V/W$ является смежный класс $\mathbf{0}+W=W$, а противоположным элементом для $\boldsymbol{v}+W$ является $(-\boldsymbol{v})+W$. Проверим свойство 5):

$$\lambda\big((\boldsymbol{u}+W)+(\boldsymbol{v}+W)\big)=\lambda\big((\boldsymbol{u}+\boldsymbol{v})+W\big)=\lambda(\boldsymbol{u}+\boldsymbol{v})+W=(\lambda\boldsymbol{u}+\lambda\boldsymbol{v})+W=\\=(\lambda\boldsymbol{u}+W)+(\lambda\boldsymbol{v}+W)=\lambda(\boldsymbol{u}+W)+\lambda(\boldsymbol{v}+W).$$

Оставшиеся свойства 6)–8) проверяются аналогично. □

Теорема 1.5.5. *Пусть $W$ — подпространство линейного пространства $V$. Тогда*

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

Доказательство. Пусть $\dim V=n$, $\dim W=k$ и пусть $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_k$ — базис в $W$. Дополним его до базиса $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_k,\boldsymbol{e}_{k+1},\ldots,\boldsymbol{e}_n$ в $V$. Докажем, что классы

$$\boldsymbol{e}_{k+1}+W,\ldots,\boldsymbol{e}_n+W \tag{1.4}$$

образуют базис в $V/W$.

Вначале докажем, что классы (1.4) линейно независимы. Пусть

$$\lambda_{k+1}(\boldsymbol{e}_{k+1}+W)+\ldots+\lambda_n(\boldsymbol{e}_n+W)=\mathbf{0}+W.$$

Тогда $(\lambda_{k+1}\boldsymbol{e}_{k+1}+\ldots+\lambda_n\boldsymbol{e}_n)+W=\mathbf{0}+W$, т.е. $\boldsymbol{v}:=\lambda_{k+1}\boldsymbol{e}_{k+1}+\ldots+\lambda_n\boldsymbol{e}_n\in W$ в силу леммы 1.5.2. Так как $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_k$ — базис в $W$, мы можем записать $\boldsymbol{v}=\lambda_1\boldsymbol{e}_1+\ldots+\lambda_k\boldsymbol{e}_k$. Тогда получаем

$$\lambda_1\boldsymbol{e}_1+\ldots+\lambda_k\boldsymbol{e}_k-\lambda_{k+1}\boldsymbol{e}_{k+1}-\ldots-\lambda_n\boldsymbol{e}_n=\mathbf{0}.$$

Так как $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_k,\boldsymbol{e}_{k+1},\ldots,\boldsymbol{e}_n$ — базис в $V$, отсюда вытекает, что $\lambda_{k+1}=\ldots=\lambda_n=0$, т.е. классы (1.4) линейно независимы.

Осталось доказать, что классы (1.4) порождают всё пространство $V/W$. Возьмём произвольный вектор $\boldsymbol{v}+W\in V/W$. Разложим вектор $\boldsymbol{v}$ по базису в $V$:

$$\boldsymbol{v}=\lambda_1\boldsymbol{e}_1+\ldots+\lambda_k\boldsymbol{e}_k+\lambda_{k+1}\boldsymbol{e}_{k+1}+\ldots+\lambda_n\boldsymbol{e}_n.$$

Тогда

$$\boldsymbol{v}+W=(\lambda_{k+1}\boldsymbol{e}_{k+1}+\ldots+\lambda_n\boldsymbol{e}_n)+(\lambda_1\boldsymbol{e}_1+\ldots+\lambda_k\boldsymbol{e}_k)+W=\\=(\lambda_{k+1}\boldsymbol{e}_{k+1}+\ldots+\lambda_n\boldsymbol{e}_n)+W=\lambda_{k+1}(\boldsymbol{e}_{k+1}+W)+\ldots+\lambda_n(\boldsymbol{e}_n+W).$$

Итак, (1.4) — базис в $V/W$, а значит $\dim V/W=n-k$. □

## 1.6. Координаты вектора. Закон изменения координат при замене базиса

Определение 1.6.1. Пусть $V$ — линейное пространство и $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ — базис в $V$. Любой вектор $\boldsymbol{x} \in V$ единственным образом представляется в виде линейной комбинации базисных векторов: $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n$. Числа $x_1, \ldots, x_n \in \mathbf{k}$ называются *координатами* вектора $\boldsymbol{x}$ в базисе $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$.

Далее мы приведём ряд соглашений об обозначениях, которые существенно упростят выкладки, связанные с координатами.

Соглашение (обозначения Эйнштейна для координат). Мы, как правило, будем писать индексы у координат сверху, а не снизу, т.е. $x^1, \ldots, x^n$ вместо $x_1, \ldots, x_n$. Вместо длинной записи разложения вектора по базису $\boldsymbol{x} = x^1\boldsymbol{e}_1 + \ldots + x^n\boldsymbol{e}_n$ мы будем использовать запись $\boldsymbol{x} = x^i\boldsymbol{e}_i$, подразумевая сумму $\sum_{i=1}^n x^i\boldsymbol{e}_i$ (т.е. суммирование всегда будет подразумеваться по повторяющимся верхним и нижним индексам).

При работе с матрицами координаты вектора $\boldsymbol{x}$ в базисе $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ мы будем записывать в виде столбца высоты $n$, обозначая его простой (нежирной) буквой $x$, т.е. $x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$. Часто для экономии места вместо $\begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$ будем писать $(x^1 \ldots x^n)^t$.

Пусть в пространстве $V$ заданы два базиса: «старый» $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ и «новый» $\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n$. Нам будет удобно обозначать векторы нового базиса через $\boldsymbol{e}_{1'}, \ldots, \boldsymbol{e}_{n'}$. Запишем формулы, выражающие векторы нового базиса через старый базис:

$$\begin{aligned} &\boldsymbol{e}_{1'} = c^1_{1'}\boldsymbol{e}_1 + \ldots + c^n_{1'}\boldsymbol{e}_n, \\ &\ldots \qquad\qquad \ldots \\ &\boldsymbol{e}_{n'} = c^1_{n'}\boldsymbol{e}_1 + \ldots + c^n_{n'}\boldsymbol{e}_n, \end{aligned}$$

или, используя обозначения Эйнштейна,

$$\boldsymbol{e}_{i'} = c^i_{i'}\boldsymbol{e}_i, \quad i' = 1, \ldots n.$$

Эти формулы равносильны одному матричному равенству

$$(\boldsymbol{e}_{1'} \ldots \boldsymbol{e}_{n'}) = (\boldsymbol{e}_1 \ldots \boldsymbol{e}_n) \begin{pmatrix} c^1_{1'} & \cdots & c^1_{n'} \\ \vdots & \ddots & \vdots \\ c^n_{1'} & \cdots & c^n_{n'} \end{pmatrix}.$$

Определение 1.6.2. Матрица

$$C = (c^i_{i'}) = \begin{pmatrix} c^1_{1'} & \cdots & c^1_{n'} \\ \vdots & \ddots & \vdots \\ c^n_{1'} & \cdots & c^n_{n'} \end{pmatrix}$$

называется *матрицей перехода* от базиса $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ к базису $\boldsymbol{e}_{1'}, \ldots, \boldsymbol{e}_{n'}$. Её столбцами являются координаты новых базисных векторов в старом базисе.

Теорема 1.6.3 (закон изменения координат). *Пусть $x^1, \ldots, x^n$ — координаты вектора $\boldsymbol{x}$ в базисе $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$, а $x^{1'}, \ldots, x^{n'}$ — координаты этого же вектора в базисе*

$\boldsymbol{e}_{1'}, \dots, \boldsymbol{e}_{n'}$. *Тогда два набора координат связаны следующими формулами.*
*В развёрнутом координатном виде:*

$$x^1 = c^1_{1'} x^{1'} + \dots + c^1_{n'} x^{n'},$$
$$\dots\dots$$
$$x^n = c^n_{1'} x^{1'} + \dots + c^n_{n'} x^{n'}.$$

*В матричном виде:*

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

*В обозначениях Эйнштейна:*

$$x^i = c^i_{i'} x^{i'}, \quad i = 1, \dots, n.$$

Доказательство. Хотя все три формулы закона преобразования координат эквивалентны, мы докажем их по отдельности, чтобы лучше освоиться с различными обозначениями. Мы имеем

$$\begin{aligned} x^1 \boldsymbol{e}_1 + \dots + x^n \boldsymbol{e}_n = \boldsymbol{x} = x^{1'} \boldsymbol{e}_{1'} + \dots + x^{n'} \boldsymbol{e}_{n'} = \\ = x^{1'} (c^1_{1'} \boldsymbol{e}_1 + \dots + c^n_{1'} \boldsymbol{e}_n) + \dots + x^{n'} (c^1_{n'} \boldsymbol{e}_1 + \dots + c^n_{n'} \boldsymbol{e}_n) = \\ = (c^1_{1'} x^{1'} + \dots + c^1_{n'} x^{n'}) \boldsymbol{e}_1 + \dots + (c^n_{1'} x^{1'} + \dots + c^n_{n'} x^{n'}) \boldsymbol{e}_n. \end{aligned}$$

Так как $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ — базис, мы получаем $x^i = c^i_{1'} x^{1'} + \dots + c^i_{n'} x^{n'}$ для $i = 1, \dots, n$.

Та же выкладка в матричных обозначениях имеет вид

$$(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n) \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \boldsymbol{x} = (\boldsymbol{e}_{1'}, \dots, \boldsymbol{e}_{n'}) \begin{pmatrix} x^{1'} \\ \vdots \\ x^{n'} \end{pmatrix} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n) \begin{pmatrix} c^1_{1'} & \cdots & c^1_{n'} \\ \vdots & \ddots & \vdots \\ c^n_{1'} & \cdots & c^n_{n'} \end{pmatrix} \begin{pmatrix} x^{1'} \\ \vdots \\ x^{n'} \end{pmatrix},$$

откуда

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

Наконец, используя обозначения Эйнштейна, получаем

$$x^i \boldsymbol{e}_i = \boldsymbol{x} = x^{i'} \boldsymbol{e}_{i'} = x^{i'} c^i_{i'} \boldsymbol{e}_i,$$

откуда $x^i = x^{i'} c^i_{i'} = c^i_{i'} x^{i'}$. □

Из доказательства видно, что выкладка, использующая обозначения Эйнштейна имеет наиболее простой вид. Далее в аналогичных выкладках мы как правило будем пользоваться обозначениями Эйнштейна.

Обратим внимание также, что в определении матрицы перехода мы выражаем *новые* векторы через *старые*, а в законе преобразования координат, наоборот, *старые* координаты выражаются через *новые*.

Соглашение (обозначения Эйнштейна для матриц). Пусть $A = (a^i_j)$ — матрица размера $l \times m$, а $B = (b^j_k)$ — матрица размера $m \times n$. Тогда закон умножения матриц в обозначениях Эйнштейна выглядит следующим образом: для компонент

$l \times n$-матрицы $C = (c^i_k)$, получаемой как произведение матриц $A$ и $B$, имеет место соотношение $c^i_k = a^i_j b^j_k$ (по повторяющемуся индексу $j$ производится суммирование).

Компоненты единичной (квадратной) матрицы $E$ задаются *символом Кронекера*: $E = (\delta^i_j)$, где

$$\delta^i_j = \begin{cases} 1, & \text{при } i = j, \\ 0, & \text{при } i \neq j. \end{cases}$$

Матрица $D = (d^j_k)$ является обратной к матрице $C = (c^i_j)$, т.е. $D = C^{-1}$, если выполнено соотношение $c^i_j d^j_k = \delta^i_k$.

Предложение 1.6.4.

а) *Матрица $C_{\boldsymbol{e}' \to \boldsymbol{e}} = (c^{i'}_i)$ перехода от базиса $\boldsymbol{e}_{1'}, \ldots, \boldsymbol{e}_{n'}$ к базису $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ является обратной к матрице $C_{\boldsymbol{e} \to \boldsymbol{e}'} = (c^i_{i'})$ перехода от $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ к $\boldsymbol{e}_{1'}, \ldots, \boldsymbol{e}_{n'}$, т.е.*

$$C_{\boldsymbol{e} \to \boldsymbol{e}'} C_{\boldsymbol{e}' \to \boldsymbol{e}} = E.$$

*В частности, матрица перехода всегда невырождена (обратима).*

б) *Если $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$, $\boldsymbol{e}_{1'}, \ldots, \boldsymbol{e}_{n'}$, $\boldsymbol{e}_{1''}, \ldots, \boldsymbol{e}_{n''}$ — три базиса, то для соответствующих матриц перехода имеет место соотношение*

$$C_{\boldsymbol{e} \to \boldsymbol{e}''} = C_{\boldsymbol{e} \to \boldsymbol{e}'} C_{\boldsymbol{e}' \to \boldsymbol{e}''}.$$

Доказательство. Первое утверждение следует из второго, если положить $\boldsymbol{e}_{i''} = \boldsymbol{e}_i$. Докажем второе утверждение. Пусть $C_{\boldsymbol{e} \to \boldsymbol{e}'} = (c^i_{i'})$, $C_{\boldsymbol{e}' \to \boldsymbol{e}''} = (c^{i'}_{i''})$, $C_{\boldsymbol{e} \to \boldsymbol{e}''} = (c^i_{i''})$. Тогда

$$c^i_{i''} \boldsymbol{e}_i = \boldsymbol{e}_{i''} = c^{i'}_{i''} \boldsymbol{e}_{i'} = c^{i'}_{i''} c^i_{i'} \boldsymbol{e}_i = c^i_{i'} c^{i'}_{i''} \boldsymbol{e}_i,$$

откуда $c^i_{i''} = c^i_{i'} c^{i'}_{i''}$, т.е. $C_{\boldsymbol{e} \to \boldsymbol{e}''} = C_{\boldsymbol{e} \to \boldsymbol{e}'} C_{\boldsymbol{e}' \to \boldsymbol{e}''}$. □

## 1.7. Линейные отображения и изоморфизмы. Ядро и образ

Определение 1.7.1. Пусть $V$ и $W$ — линейные пространства над полем $\mathbf{k}$. Отображение $\mathcal{A}\colon V \to W$ называется *линейным*, если для любых векторов $\boldsymbol{u}, \boldsymbol{v} \in V$ и скаляра $\lambda \in \mathbf{k}$ выполнены равенства

$$\mathcal{A}(\boldsymbol{u} + \boldsymbol{v}) = \mathcal{A}\boldsymbol{u} + \mathcal{A}\boldsymbol{v}, \quad \mathcal{A}(\lambda \boldsymbol{v}) = \lambda \mathcal{A}\boldsymbol{v}.$$

Биективное (т.е. взаимно однозначное) линейное отображение $\mathcal{A}\colon V \to W$ называется *изоморфизмом.* Пространства $V$ и $W$ называются изоморфными, если между ними существует изоморфизм.

Вот два важнейших примера линейных отображений.

Пример 1.7.2.

1. Линейное отображение $f\colon V \to \mathbf{k}$ пространства $V$ над полем $\mathbf{k}$ в поле $\mathbf{k}$ (рассматриваемое как 1-мерное линейное пространство) называется *линейной функцией* или *линейным функционалом.* Линейные функции мы вскоре отдельно рассмотрим более подробно.

2. Линейное отображение $\mathcal{A}\colon V \to V$ пространства $V$ в себя называется *линейным оператором.* Линейные операторы будут подробно изучены во второй главе.

Теорема 1.7.3. *Два пространства $V$ и $W$ над полем $\mathbf{k}$ изоморфны тогда и только тогда, когда они имеют одинаковые размерности.*

Доказательство. Из определения изоморфизма вытекает, что свойства системы векторов быть линейной независимой и порождать всё пространство сохраняются при изоморфизмах, т.е. при изоморфизме базис переходит в базис. Следовательно, если $\mathcal{A}\colon V \to W$ — изоморфизм, то $\dim V = \dim W$.

Пусть теперь $\dim V = \dim W = n$. Выберем базисы $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ и $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ в $V$ и $W$ соответственно. Тогда формула

$$\mathcal{A}(x^i \boldsymbol{e}_i) = x^i \boldsymbol{f}_i$$

(где по $i$ подразумевается суммирование) определяет линейное отображение $\mathcal{A}\colon V \to W$. Оно является биекцией, так как формула $\mathcal{A}^{-1}(x^i \boldsymbol{f}_i) = x^i \boldsymbol{e}_i$ определяет обратное отображение. Итак, $\mathcal{A}$ — изоморфизм. □

В качестве стандартного примера $n$-мерного пространства над $\mathbb{R}$ мы будем рассматривать координатное пространство $\mathbb{R}^n$; в силу предыдущей теоремы оно изоморфно любому другому $n$-мерному вещественному пространству.

Определение 1.7.4. Пусть $\mathcal{A}\colon V \to W$ — линейное отображение. Множество векторов $\boldsymbol{v} \in V$, для которых $\mathcal{A}\boldsymbol{v} = \boldsymbol{0}$, называется *ядром* отображения $\mathcal{A}$ и обозначается $\operatorname{Ker} \mathcal{A}$. *Образ* линейного отображения $\mathcal{A}\colon V \to W$, обозначаемый $\operatorname{Im} \mathcal{A}$, определяется так же, как и для любого отображения: $\operatorname{Im} \mathcal{A} = \{\mathcal{A}\boldsymbol{v}\colon \boldsymbol{v} \in V\}$.

Предложение 1.7.5. *Пусть $\mathcal{A}\colon V \to W$ — линейное отображение. Тогда $\operatorname{Ker} \mathcal{A}$ является подпространством в $V$, а $\operatorname{Im} \mathcal{A}$ является подпространством в $W$.*

Доказательство. Пусть $\boldsymbol{u}, \boldsymbol{v} \in \operatorname{Ker} \mathcal{A}$, т.е. $\mathcal{A}\boldsymbol{u} = \mathcal{A}\boldsymbol{v} = \boldsymbol{0}$. Тогда $\mathcal{A}(\boldsymbol{u} + \boldsymbol{v}) = \mathcal{A}\boldsymbol{u} + \mathcal{A}\boldsymbol{v} = \boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0}$ и $\mathcal{A}(\lambda \boldsymbol{v}) = \lambda \mathcal{A}\boldsymbol{v} = \boldsymbol{0}$. Следовательно, $\boldsymbol{u} + \boldsymbol{v} \in \operatorname{Ker} \mathcal{A}$ и $\lambda \boldsymbol{v} \in \operatorname{Ker} \mathcal{A}$, а значит $\operatorname{Ker} \mathcal{A}$ — подпространство в $V$.

Пусть теперь $\boldsymbol{x}, \boldsymbol{y} \in \operatorname{Im} \mathcal{A}$, т.е. существуют $\boldsymbol{u}, \boldsymbol{v} \in V$, такие, что $\mathcal{A}\boldsymbol{u} = \boldsymbol{x}$ и $\mathcal{A}\boldsymbol{v} = \boldsymbol{y}$. Тогда $\mathcal{A}(\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{x} + \boldsymbol{y}$ и $\mathcal{A}(\lambda \boldsymbol{u}) = \lambda \boldsymbol{x}$. Следовательно, $\boldsymbol{x} + \boldsymbol{y} \in \operatorname{Im} \mathcal{A}$ и $\lambda \boldsymbol{x} \in \operatorname{Im} \mathcal{A}$, а значит $\operatorname{Im} \mathcal{A}$ — подпространство в $W$. □

Теорема 1.7.6. *Пусть $\mathcal{A}\colon V \to W$ — линейное отображение. Тогда соответствие $\boldsymbol{v} + \operatorname{Ker} \mathcal{A} \mapsto \mathcal{A}\boldsymbol{v}$ задаёт изоморфизм между факторпространством $V/\operatorname{Ker} \mathcal{A}$ и подпространством $\operatorname{Im} \mathcal{A}$.*

Доказательство. Это доказательство полностью повторяет доказательство теоремы из курса алгебры о том, что «гомоморфный образ группы изоморфен факторгруппе по ядру гомоморфизма».

Сначала проверим, что $\boldsymbol{v} + \operatorname{Ker} \mathcal{A} \mapsto \mathcal{A}\boldsymbol{v}$ действительно корректно определяет отображение $\widetilde{\mathcal{A}}\colon V/\operatorname{Ker} \mathcal{A} \to \operatorname{Im} \mathcal{A}$. Для этого нужно проверить, что если $\boldsymbol{u} + \operatorname{Ker} \mathcal{A} = \boldsymbol{v} + \operatorname{Ker} \mathcal{A}$, то $\mathcal{A}\boldsymbol{u} = \mathcal{A}\boldsymbol{v}$. В силу леммы 1.5.2, из равенства классов смежности $\boldsymbol{u} + \operatorname{Ker} \mathcal{A} = \boldsymbol{v} + \operatorname{Ker} \mathcal{A}$ вытекает, что $\boldsymbol{u} - \boldsymbol{v} \in \operatorname{Ker} \mathcal{A}$, т.е. $\mathcal{A}\boldsymbol{u} = \mathcal{A}\boldsymbol{v} + \mathcal{A}(\boldsymbol{u} - \boldsymbol{v}) = \mathcal{A}\boldsymbol{v}$. Итак, отображение $\widetilde{\mathcal{A}}\colon V/\operatorname{Ker} \mathcal{A} \to \operatorname{Im} \mathcal{A}$ определено корректно.

Линейность и сюръективность отображения $\widetilde{\mathcal{A}}$ очевидны. Проверим, что оно также инъективно. Пусть $\widetilde{\mathcal{A}}(\boldsymbol{u} + \operatorname{Ker} \mathcal{A}) = \widetilde{\mathcal{A}}(\boldsymbol{v} + \operatorname{Ker} \mathcal{A})$. Это означает, что $\mathcal{A}\boldsymbol{u} = \mathcal{A}\boldsymbol{v}$, т.е. $\boldsymbol{u} - \boldsymbol{v} \in \operatorname{Ker} \mathcal{A}$. Тогда из леммы 1.5.2 следует, что $\boldsymbol{u} + \operatorname{Ker} \mathcal{A} = \boldsymbol{v} + \operatorname{Ker} \mathcal{A}$, т.е. $\widetilde{\mathcal{A}}$ инъективно. Так как линейное отображение $\widetilde{\mathcal{A}}\colon V/\operatorname{Ker} \mathcal{A} \to \operatorname{Im} \mathcal{A}$ сюръективно и инъективно, оно является изоморфизмом. □

Следствие 1.7.7. *Для любого линейного отображения $\mathcal{A}\colon V \to W$ мы имеем*

$$\dim V = \dim \operatorname{Ker} \mathcal{A} + \dim \operatorname{Im} \mathcal{A}.$$

Доказательство. Из предыдущей теоремы следует, что $\dim(V/\operatorname{Ker}\mathcal{A}) = \dim\operatorname{Im}\mathcal{A}$, а $\dim(V/\operatorname{Ker}\mathcal{A}) = \dim V - \dim\operatorname{Ker}\mathcal{A}$ по теореме 1.5.5 о размерности факторпространства. $\square$

## 1.8. Матрица линейного отображения. Преобразование матрицы линейного отображения при заменах базисов

Пусть $\mathcal{A}\colon V \to W$ — линейное отображение, $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ — базис в $V$, а $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ — базис в $W$.

Определение 1.8.1. *Матрицей отображения* $\mathcal{A}\colon V \to W$ *по отношению к базисам* $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ *и* $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ называется матрица

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \ldots & a_m^1 \\ a_1^2 & a_2^2 & \ldots & a_m^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \ldots & a_m^n \end{pmatrix}$$

размера $n \times m$, в которой $i$-й столбец составлен из координат вектора $\mathcal{A}(\boldsymbol{e}_i)$ относительно базиса $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$:

$$\mathcal{A}\boldsymbol{e}_i = a_i^j \boldsymbol{f}_j.$$

Зная матрицу линейного отображения $\mathcal{A}$, мы можем найти образ любого вектора $\boldsymbol{x} \in V$ при отображении $\mathcal{A}$ следующим образом.

Предложение 1.8.2. *Пусть* $\boldsymbol{x} = x^j \boldsymbol{e}_j$ *— произвольный вектор из* $V$, *а* $\boldsymbol{y} = y^i \boldsymbol{f}_i$ *— его образ в* $W$, *т.е.* $\boldsymbol{y} = \mathcal{A}\boldsymbol{x}$. *Тогда*

$$y^i = a_j^i x^j \qquad \text{или} \qquad \begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix} = \begin{pmatrix} a_1^1 & \ldots & a_m^1 \\ \vdots & \ddots & \vdots \\ a_1^n & \ldots & a_m^n \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^m \end{pmatrix}.$$

Доказательство. Действительно,

$$y^i \boldsymbol{f}_i = \boldsymbol{y} = \mathcal{A}\boldsymbol{x} = \mathcal{A}(x^j \boldsymbol{e}_j) = x^j \mathcal{A}\boldsymbol{e}_j = x^j a_j^i \boldsymbol{f}_i.$$

Так как $\{\boldsymbol{f}_i\}$ — базис, отсюда следует, что $y^i = a_j^i x^j$. $\square$

Определение 1.8.3. Множество всех линейных отображений $\mathcal{A}\colon V \to W$ из $V$ в $W$ с операциями сложения и умножения

$$(\mathcal{A}_1 + \mathcal{A}_2)(\boldsymbol{v}) := \mathcal{A}_1\boldsymbol{v} + \mathcal{A}_2\boldsymbol{v}, \quad (\lambda\mathcal{A})(\boldsymbol{v}) := \lambda(\mathcal{A}\boldsymbol{v})$$

является линейным пространством. Оно называется *пространством линейных отображений* из $V$ в $W$ и обозначается $\operatorname{Hom}_{\mathbf{k}}(V, W)$ или просто $\operatorname{Hom}(V, W)$.

Предложение 1.8.4. *Пусть* $\dim V = m$ *и* $\dim W = n$. *Тогда пространство линейных отображений* $\operatorname{Hom}_{\mathbf{k}}(V, W)$ *изоморфно пространству матриц* $\operatorname{Mat}_{\mathbf{k}}(n, m)$.

Доказательство. Выберем базисы $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ и $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ в $V$ и $W$ соответственно. Определим отображение $\operatorname{Hom}_{\mathbf{k}}(V, W) \to \operatorname{Mat}_{\mathbf{k}}(n, m)$, которое сопоставляет линейному отображению его матрицу в выбранных базисах. Непосредственно проверяется, что это отображение линейно. Кроме того, оно биективно: обратное отображение сопоставляет $n \times m$-матрице $A = (a_j^i)$ линейное отображение, определяемое в координатах формулой из предложения 1.8.2. Следовательно, наше отображение $\operatorname{Hom}_{\mathbf{k}}(V, W) \to \operatorname{Mat}_{\mathbf{k}}(n, m)$ является изоморфизмом. $\square$

Пусть в пространстве $V$ выбран новый базис $\boldsymbol{e}_{1'}, \dots, \boldsymbol{e}_{m'}$, а в пространстве $W$ — новый базис $\boldsymbol{f}_{1'}, \dots, \boldsymbol{f}_{n'}$.

Теорема 1.8.5 (закон изменения матрицы линейного отображения). *Имеет место соотношение*

$$A' = D^{-1}AC,$$

*где $A$ — матрица линейного отображения $\mathcal{A}\colon V \to W$ по отношению к базисам $\boldsymbol{e}_1, \dots, \boldsymbol{e}_m$ и $\boldsymbol{f}_1, \dots, \boldsymbol{f}_n$; $A'$ — матрица отображения $\mathcal{A}$ по отношению к базисам $\boldsymbol{e}_{1'}, \dots, \boldsymbol{e}_{m'}$ и $\boldsymbol{f}_{1'}, \dots, \boldsymbol{f}_{n'}$; $C = C_{\boldsymbol{e}\to\boldsymbol{e}'}$ — матрица перехода от базиса $\boldsymbol{e}_1, \dots, \boldsymbol{e}_m$ к базису $\boldsymbol{e}_{1'}, \dots, \boldsymbol{e}_{m'}$ и $D = D_{\boldsymbol{f}\to\boldsymbol{f}'}$ — матрица перехода от $\boldsymbol{f}_1, \dots, \boldsymbol{f}_n$ к $\boldsymbol{f}_{1'}, \dots, \boldsymbol{f}_{n'}$.*

Доказательство. Пусть $C = (c^i_{i'})$ и $A = (a^j_i)$, тогда

$$\mathcal{A}\boldsymbol{e}_{i'} = \mathcal{A}(c^i_{i'}\boldsymbol{e}_i) = c^i_{i'}\mathcal{A}\boldsymbol{e}_i = c^i_{i'}a^j_i\boldsymbol{f}_j.$$

С другой стороны, если $A' = (a^{j'}_{i'})$ и $D = (d^j_{j'})$, то

$$\mathcal{A}\boldsymbol{e}_{i'} = a^{j'}_{i'}\boldsymbol{f}_{j'} = a^{j'}_{i'}d^j_{j'}\boldsymbol{f}_j.$$

Сравнивая последние два соотношения, с учётом того, что $\{\boldsymbol{f}_j\}$ — базис, получаем $a^j_i c^i_{i'} = d^j_{j'}a^{j'}_{i'}$. Это эквивалентно соотношению $AC = DA'$, т.е. $A' = D^{-1}AC$. □

## 1.9. Двойственное пространство $V^*$, двойственный базис. Отсутствие изоморфизма $V \cong V^*$ в бесконечномерном случае (пример)

Напомним, что *линейной функцией* называется линейное отображение $f\colon V \to \mathbf{k}$. Как и всякое множество линейных отображений между двумя пространствами, множество линейных функций является линейным пространством.

Определение 1.9.1. Пространство $\mathrm{Hom}(V, \mathbf{k})$ линейных функций $f\colon V \to \mathbf{k}$ называется *двойственным* (или *сопряжённым*) *пространством* к $V$ и обозначается $V^*$.

Пусть $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ — базис в $V$. Значение линейной функции $\xi \in V^*$ на любом векторе $\boldsymbol{x} = x^i\boldsymbol{e}_i \in V$ определяется её значениями на базисных векторах, так как $\xi(\boldsymbol{x}) = x^i\xi(\boldsymbol{e}_i)$. Определим линейные функции $\varepsilon^1, \dots, \varepsilon^n \in V^*$ по правилу

$$\varepsilon^i(\boldsymbol{e}_j) = \delta^i_j.$$

Тогда для любого вектора $\boldsymbol{x} = x^j\boldsymbol{e}_j$ мы имеем

$$\varepsilon^i(\boldsymbol{x}) = \varepsilon^i(x^j\boldsymbol{e}_j) = x^j\varepsilon^i(\boldsymbol{e}_j) = x^j\delta^i_j = x^i.$$

В связи с этим функции $\varepsilon^i$ часто называют *координатными функциями.*

Предложение 1.9.2. *Линейные функции $\varepsilon^1, \dots, \varepsilon^n$ образуют базис в $V^*$.*

Доказательство. Линейная независимость. Пусть $x_1\varepsilon^1 + \ldots + x_n\varepsilon^n = o$. Это равенство означает, что линейная функция $\xi := x_i\varepsilon^i$ равна нулю на любом векторе из $V$. Вычислим её на векторе $\boldsymbol{e}_j$: $0 = \xi(\boldsymbol{e}_j) = x_i\varepsilon^i(\boldsymbol{e}_j) = x_i\delta^i_j = x_j$. Итак, все коэффициенты $x_j$ равны нулю, а значит $\varepsilon^1, \dots, \varepsilon^n \in V^*$ линейно независимы.

Теперь проверим, что $\varepsilon^1, \dots, \varepsilon^n$ порождают всё пространство $V^*$. Мы утверждаем, что любая линейная функции $\xi$ представляется в виде линейной комбинации $\xi = \xi_i\varepsilon^i$, где $\xi_i = \xi(\boldsymbol{e}_i)$. Действительно, для любого вектора $\boldsymbol{x} = x^j\boldsymbol{e}_j \in V$ мы имеем

$$\xi_i\varepsilon^i(\boldsymbol{x}) = \xi_i x^i = \xi(\boldsymbol{e}_i)x^i = \xi(x^i\boldsymbol{e}_i) = \xi(\boldsymbol{x}).$$

Следовательно, $\xi = \xi_i \varepsilon^i$, т.е. $\varepsilon^1, \ldots, \varepsilon^n$ — базис в $V^*$. □

Определение 1.9.3. Базис $\varepsilon^1, \ldots, \varepsilon^n$ пространства $V^*$ называется *двойственными* (или *сопряжённым*) *базисом* к $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$.

Следствие 1.9.4. $\dim V = \dim V^*$.

Пусть теперь $\boldsymbol{e}_{1'}, \ldots, \boldsymbol{e}_{n'}$ — другой базис пространства $V$ и $C = (c^i_{i'})$ — матрица перехода, $\boldsymbol{e}_{i'} = c^i_{i'} \boldsymbol{e}_i$. Рассмотрим двойственные базисы $\varepsilon^1, \ldots, \varepsilon^n$ и $\varepsilon^{1'}, \ldots, \varepsilon^{n'}$.

Предложение 1.9.5. *Матрица перехода от* $\varepsilon^1, \ldots, \varepsilon^n$ *к* $\varepsilon^{1'}, \ldots, \varepsilon^{n'}$ *есть* $(C^{-1})^t$.

Доказательство. Для любого вектора $\boldsymbol{x} = x^i \boldsymbol{e}_i = x^{i'} \boldsymbol{e}_{i'}$ мы имеем $\varepsilon^i(\boldsymbol{x}) = x^i = c^i_{i'} x^{i'} = c^i_{i'} \varepsilon^{i'}(\boldsymbol{x})$. Следовательно,
$$\varepsilon^i = c^i_{i'} \varepsilon^{i'}.$$
Это эквивалентно матричному соотношению
$$\begin{pmatrix} \varepsilon^1 \\ \vdots \\ \varepsilon^n \end{pmatrix} = C \begin{pmatrix} \varepsilon^{1'} \\ \vdots \\ \varepsilon^{n'} \end{pmatrix}$$
или
$$(\varepsilon^{1'} \ldots \varepsilon^{n'}) = (\varepsilon^1 \ldots \varepsilon^n)(C^{-1})^t.$$
Это означает, что $(C^{-1})^t$ — это матрица перехода от $\varepsilon^1, \ldots, \varepsilon^n$ к $\varepsilon^{1'}, \ldots, \varepsilon^{n'}$. □

Из следствия 1.9.4 также вытекает, что пространства $V$ и $V^*$ изоморфны. Однако для построения изоморфизма между ними нам необходимо выбрать базис в $V$ (и двойственный базис в $V^*$); изоморфизм между $V$ и $V^*$ «неканоничен» в том смысле, что он зависит от выбора базиса. Разные базисы дают разные изоморфизмы.

Для бесконечномерных пространств ситуация иная: пространства $V$ и $V^*$ *никогда* не изоморфны, пространство $V^*$ всегда «больше». Мы не будем приводить доказательства этого факта (и даже не будем приводить его точной формулировки), а лишь проиллюстрируем его на примере. Это последний раз, когда у нас появляются бесконечномерные пространства; здесь нам также понадобится поле, отличное от $\mathbb{R}$ и $\mathbb{C}$.

Напомним, что $\mathbf{k}^\infty$ — это пространство финитных последовательностей, т.е. бесконечных последовательностей элементов поля $\mathbf{k}$, в которых лишь конечное число элементов отличны от нуля. Через $\widehat{\mathbf{k}}^\infty$ мы обозначали пространство всех бесконечных последовательностей.

Предложение 1.9.6. *Двойственное пространство к* $\mathbf{k}^\infty$ *изоморфно* $\widehat{\mathbf{k}}^\infty$.

Доказательство. Мы воспользуемся тем фактом, что в пространстве $\mathbf{k}^\infty$ имеется стандартный базис $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots$, где $\boldsymbol{e}_i = (0, \ldots, 0, 1, 0, \ldots)$ — последовательность, в которой на $i$-м месте стоит 1, а на остальных местах — нули.

Рассмотрим отображение
$$\mathcal{A}\colon (\mathbf{k}^\infty)^* \to \widehat{\mathbf{k}}^\infty, \quad f \mapsto \bigl(f(\boldsymbol{e}_1), f(\boldsymbol{e}_2), \ldots\bigr),$$
которое линейной функции $f \in (\mathbf{k}^\infty)^*$ ставит в соответствие последовательность её значений на базисных векторах $\boldsymbol{e}_i$. Это отображение очевидно линейно. Кроме того, отображение $\mathcal{A}$ биективно: обратное отображение $\mathcal{A}^{-1}$ задаётся формулой
$$\mathcal{A}^{-1}\bigl((x_1, x_2, \ldots)\bigr) = f \in (\mathbf{k}^\infty)^*, \quad \text{где } f(\boldsymbol{e}_i) = x_i.$$

Так как любой элемент $\boldsymbol{y} \in \mathbf{k}^\infty$ есть (конечная) линейная комбинация элементов $\boldsymbol{e}_i$, значение линейной функции $f$ на $\boldsymbol{y}$ однозначно восстанавливается по её значениям $f(\boldsymbol{e}_i)$. Итак, $\mathcal{A}$ — изоморфизм. □

Пусть $\mathbf{k} = \mathbb{Z}_2$ — поле из двух элементов.

Предложение 1.9.7. *Пространства $\mathbb{Z}_2^\infty$ и $\widehat{\mathbb{Z}}_2^\infty$ неизоморфны. Таким образом, пространство $\mathbb{Z}_2^\infty$ не изоморфно своему двойственному пространству.*

Доказательство. Дело в том, что $\mathbb{Z}_2^\infty$ как множество счётно, а $\widehat{\mathbb{Z}}_2^\infty$ не является счётным, так что между $\mathbb{Z}_2^\infty$ и $\widehat{\mathbb{Z}}_2^\infty$ нельзя установить биекцию. (Множество $X$ называется *счётным*, если имеется биекция между $X$ и множеством натуральных чисел $\mathbb{N}$).

Действительно, элементы множества $\mathbb{Z}_2^\infty$ соответствуют конечным последовательностям из нулей и единиц. Количество таких последовательностей длины $n$ конечно (равно $2^n$), а поэтому $\mathbb{Z}_2^\infty$ счётно как счётное объединение конечных множеств. (Это также можно увидеть, отождествив $\mathbb{Z}_2^\infty$ с множеством рациональных чисел на отрезке $[0,1]$ в двоичной записи).

С другой стороны, множество $\widehat{\mathbb{Z}}_2^\infty$ всех последовательностей из нулей и единиц не является счётным. Действительно, предположим, что нам удалось перенумеровать все такие последовательности: $a_1, a_2, \ldots,$. Рассмотрим последовательность $b$, в которой $k$-й элемент отличается от $k$-го элемента последовательности $a_k$. Тогда последовательность $b$ не может присутствовать в списке $a_1, a_2, \ldots$, так как она отличается от $k$-ой последовательности из списка по крайней мере в $k$-м члене. Полученное противоречие показывает, что $\widehat{\mathbb{Z}}_2^\infty$ не является счётным. (Множество $\widehat{\mathbb{Z}}_2^\infty$ можно также отождествить с множеством всех чисел на отрезке $[0,1]$ в двоичной записи.) □

В качестве задачи полезно доказать, что пространство $\mathbb{R}^\infty$ также не изоморфно своему двойственному пространству $\widehat{\mathbb{R}}^\infty$ (указание: в $\mathbb{R}^\infty$ имеется счётный базис, а в $\widehat{\mathbb{R}}^\infty$ не существует счётного базиса).

## 1.10. Второе двойственное пространство, канонический изоморфизм $V \cong V^{**}$

Мы видели, что пространства $V$ и $V^*$ изоморфны (в конечномерном случае), однако для построения изоморфизма нам требовалось выбрать базис в $V$. Сейчас мы построим изоморфизм между пространством $V$ и его *вторым двойственным пространством* $V^{**}$, который не требует выбора базиса.

По определению, элементами пространства $V^{**}$ являются линейные функции на пространстве линейных функций $V^*$.

Теорема 1.10.1. *Пусть $V$ — конечномерное линейное пространство. Отображение $\varphi\colon V \to V^{**}$, сопоставляющее вектору $\boldsymbol{x} \in V$ линейную функцию $\varphi_{\boldsymbol{x}}$ на $V^*$, задаваемую формулой*

$$\varphi_{\boldsymbol{x}}(\xi) := \xi(\boldsymbol{x}), \qquad \text{для } \xi \in V^*,$$

*является изоморфизмом.*

Доказательство. Очевидно, что $\varphi_{\boldsymbol{x}}$ — линейная функция на $V^*$. Кроме того,

$$\varphi(\boldsymbol{x}+\boldsymbol{y}) = \varphi_{\boldsymbol{x}+\boldsymbol{y}} = \varphi_{\boldsymbol{x}} + \varphi_{\boldsymbol{y}} = \varphi(\boldsymbol{x}) + \varphi(\boldsymbol{y})$$

и $\varphi(\lambda \boldsymbol{x}) = \lambda\varphi(\boldsymbol{x})$, т.е. отображение $\varphi$ линейно.

Докажем, что $\varphi\colon V \to V^{**}$ инъективно, т.е. $\operatorname{Ker}\varphi = \{\boldsymbol{0}\}$. Пусть $\varphi(\boldsymbol{x}) = \varphi_{\boldsymbol{x}} = o$. Последнее равенство означает, что $\varphi_{\boldsymbol{x}}(\xi) = \xi(\boldsymbol{x}) = 0$ для любой линейной функции $\xi \in V^*$. В частности, это верно для всех линейных функций $\varepsilon^1, \ldots, \varepsilon^n$ двойственного базиса к произвольному базису $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ в $V$. Следовательно, $\varepsilon^i(\boldsymbol{x}) = x^i = 0$, т.е. все координаты вектора $\boldsymbol{x} \in V$ в базисе $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ равны нулю. Это означает, что $\boldsymbol{x} = \boldsymbol{0}$, т.е. $\operatorname{Ker}\varphi = \{\boldsymbol{0}\}$.

Так как $\dim V = \dim V^{**} = \dim\operatorname{Ker}\varphi + \dim\operatorname{Im}\varphi$ (последнее равенство вытекает из следствия 1.7.7) и $\dim\operatorname{Ker}\varphi = 0$, мы получаем $\dim\operatorname{Im}\varphi = \dim V^{**}$. Следовательно, $\operatorname{Im}\varphi = V^{**}$ и $\varphi$ сюръективно.

Итак, $\varphi$ линейно и биективно, а значит это — изоморфизм. $\square$

При построении изоморфизма $\varphi\colon V \to V^{**}$ мы ни разу не использовали базис (базис использовался только при доказательстве). Изоморфизм, который не зависит от выбора базиса, называется *каноническим*.

## 1.11. Сопряжённое линейное отображение

Определение 1.11.1. Пусть $\mathcal{A}\colon V \to W$ — линейное отображение. Отображение $\mathcal{A}^*\colon W^* \to V^*$, заданное формулой

$$(\mathcal{A}^*\xi)(\boldsymbol{v}) := \xi(\mathcal{A}\boldsymbol{v}) \qquad \text{для } \xi \in W^*, \boldsymbol{v} \in V,$$

называется *сопряжённым* к $\mathcal{A}$.

Непосредственно проверяется, что $\mathcal{A}^*$ — линейное отображение.

Предложение 1.11.2. *Пусть $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ и $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ — базисы пространств $V$ и $W$ соответственно. Тогда матрица отображения $\mathcal{A}\colon V \to W$ в этих базисах и матрица сопряжённого отображения $\mathcal{A}\colon W^* \to V^*$ в двойственных базисах $\varphi^1, \ldots, \varphi^n$ и $\varepsilon^1, \ldots, \varepsilon^m$ получаются друг из друга траспонированием.*

Доказательство. Пусть $A = (a^i_j)$ — матрица отображения $\mathcal{A}$ в базисах $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ и $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$. Тогда для любого вектора $\boldsymbol{x} = x^i\boldsymbol{e}_i$ мы имеем

$$(\mathcal{A}^*\varphi^i)(\boldsymbol{x}) = \varphi^i(\mathcal{A}\boldsymbol{x}) = \varphi^i(a^j_k x^k \boldsymbol{f}_j) = a^j_k x^k \varphi^i(\boldsymbol{f}_j) = a^j_k x^k \delta^i_j = a^i_k x^k = a^i_k \varepsilon^k(\boldsymbol{x}).$$

Следовательно, $\mathcal{A}^*\varphi^i = a^i_k\varepsilon^k$. Это равенство означает, что в $i$-й строке матрицы $A = (a^i_k)$ стоят координаты образа $\varphi^i$ при отображении $\mathcal{A}^*$ по отношению к базису $\varepsilon^1, \ldots, \varepsilon^m$. По определению это означает, что матрица линейного отображения $\mathcal{A}^*$ получается из матрицы $A$ транспонированием. $\square$

Глава 2

# Линейные операторы

## 2.1. Матрица линейного оператора. Определитель и след оператора. Невырожденные операторы. Группы $GL_n$ и $SL_n$

Начиная с этой главы, все пространства предполагаются конечномерными. Напомним, что *линейным оператором* называется линейное отображение $\mathcal{A}\colon V \to V$ пространства $V$ в себя. Далее мы будем называть линейные операторы просто «операторами».

В определении матрицы $A$ линейного отображения $\mathcal{A}\colon V \to V$ естественно в качестве базисов в обоих экземплярах пространства $V$ брать один и тот же базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Получаемая квадратная матрица $A = (a_i^j)$ называется *матрицей линейного оператора* $\mathcal{A}\colon V \to V$ в базисе $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Таким образом, $i$-й столбец матрицы $A$ составлен из координат вектора $\mathcal{A}\boldsymbol{e}_i$ относительно базиса $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$:

$$\mathcal{A}\boldsymbol{e}_i = a_i^j \boldsymbol{e}_j.$$

Пример 2.1.1.

1. *Тождественный* оператор id переводит каждый вектор $\boldsymbol{v} \in V$ в себя: $\mathrm{id}\,\boldsymbol{v} = \boldsymbol{v}$. Матрицей оператора id в любом базисе является единичная матрица $E$. Обратно, если матрица оператора $\mathcal{A}$ в каком-то базисе есть $E$, то $\mathcal{A} = \mathrm{id}$.

2. Рассмотрим оператор дифференцирования $\frac{d}{dx}$ в пространстве $\mathbf{k}_2[x]$ многочленов степени не выше 2. Тогда $\frac{d}{dx}1 = 0$, $\frac{d}{dx}x = 1$ и $\frac{d}{dx}x^2 = 2x$. Таким образом, матрицей оператора $\frac{d}{dx}$ в базисе $1, x, x^2$ является матрица

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

3. В пространстве $\mathbb{R}^3$ рассмотрим оператор $\mathrm{pr}_{\boldsymbol{v}}$ ортогонального проектирования на направление вектора $\boldsymbol{v} = (1,1,1)$. Найдём матрицу этого оператора в стандартном базисе $\boldsymbol{e}_1 = (1,0,0)$, $\boldsymbol{e}_2 = (0,1,0)$ и $\boldsymbol{e}_3 = (0,0,1)$. По формуле из аналитической геометрии, для любого вектора $\boldsymbol{u} \in \mathbb{R}^3$ мы имеем $\mathrm{pr}_{\boldsymbol{v}}\boldsymbol{u} = \frac{(\boldsymbol{u},\boldsymbol{v})}{(\boldsymbol{v},\boldsymbol{v})}\boldsymbol{v}$. Следовательно,

$$\mathrm{pr}_{\boldsymbol{v}}\boldsymbol{e}_1 = \mathrm{pr}_{\boldsymbol{v}}\boldsymbol{e}_2 = \mathrm{pr}_{\boldsymbol{v}}\boldsymbol{e}_3 = \frac{1}{3}(1,1,1) = \frac{1}{3}\boldsymbol{e}_1 + \frac{1}{3}\boldsymbol{e}_2 + \frac{1}{3}\boldsymbol{e}_3.$$

Таким образом, матрица оператора $\mathrm{pr}_{\boldsymbol{v}}$ имеет вид

$$\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

Теорема 2.1.2 (закон изменения матрицы линейного оператора). *Имеет место соотношение*

$$A' = C^{-1}AC,$$

*где $A$ — матрица оператора $\mathcal{A}\colon V \to V$ в базисе $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$, $A'$ — матрица в базисе $\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_n$ и $C$ — матрица перехода от базиса $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ к базису $\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_n$.*

Доказательство. Так как мы выбираем один и тот же базис в обоих экземплярах пространства $V$, мы должны подставить $C = D$ в формулу $A' = D^{-1}AC$ преобразования матрицы линейного отображения из теоремы 1.8.5. $\square$

Матрицы $A$ и $A'$, удовлетворяющие соотношению $A' = C^{-1}AC$, где $C$ — невырожденная матрица, называются *подобными*. Таким образом, матрицы одного оператора в разных базисах подобны.

*След* $\operatorname{tr} A$ матрицы $A = (a^i_j)$ — это сумма её диагональных элементов, $\operatorname{tr} A = a^i_i$.

Лемма 2.1.3. *Определитель и след подобных матриц равны.*

Доказательство. Пусть $A' = C^{-1}AC$. Тогда для определителя имеем

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

Для вычисления следа используем обозначения Эйнштейна: $a^{i'}_{j'} = c^{i'}_i a^i_j c^j_{j'}$, откуда

$$\operatorname{tr} A' = a^{i'}_{i'} = c^{i'}_i a^i_j c^j_{i'} = c^j_{i'} c^{i'}_i a^i_j = \delta^j_i a^i_j = a^i_i = \operatorname{tr} A.$$

$\square$

Определение 2.1.4. *Определитель* (соответственно, *след*) линейного оператора $\mathcal{A}\colon V \to V$ — это определитель (соответственно, след) матрицы оператора $\mathcal{A}$ в любом базисе; обозначается $\det \mathcal{A}$ (соответственно, $\operatorname{tr} \mathcal{A}$).

Оператор $\mathcal{A}$ называется *невырожденным*, если $\det \mathcal{A} \neq 0$.

*Композицией* операторов $\mathcal{A}\colon V \to V$ и $\mathcal{B}\colon V \to V$ называется линейный оператор $\mathcal{A} \circ \mathcal{B}\colon V \to V$, определяемый формулой $(\mathcal{A} \circ \mathcal{B})(\boldsymbol{v}) = \mathcal{A}(\mathcal{B}\boldsymbol{v})$.

Предложение 2.1.5. *Матрица композиции операторов $\mathcal{A} \circ \mathcal{B}$ в любом базисе есть произведение матриц операторов $\mathcal{A}$ и $\mathcal{B}$ в этом базисе.*

Доказательство. Пусть $A = (a^i_j)$ и $B = (b^i_j)$ — матрицы операторов $\mathcal{A}$ и $\mathcal{B}$ в базисе $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ и пусть $C = AB = (c^i_j)$. Тогда

$$(\mathcal{A} \circ \mathcal{B})(\boldsymbol{e}_i) = \mathcal{A}(\mathcal{B}\boldsymbol{e}_i) = \mathcal{A}(b^j_i \boldsymbol{e}_j) = b^j_i \mathcal{A}\boldsymbol{e}_j = b^j_i a^k_j \boldsymbol{e}_k = a^k_j b^j_i \boldsymbol{e}_k = c^k_i \boldsymbol{e}_k,$$

т.е. $C = AB$ есть матрица оператора $\mathcal{A} \circ \mathcal{B}$. $\square$

Теорема 2.1.6. *Следующие условия эквивалентны для оператора $\mathcal{A}\colon V \to V$:*

а) *оператор $\mathcal{A}$ невырожден, т.е.* $\det \mathcal{A} \neq 0$*;*
б) *оператор $\mathcal{A}$ обратим, т.е. существует $\mathcal{A}^{-1}\colon V \to V$, $\mathcal{A} \circ \mathcal{A}^{-1} = \mathrm{id}$;*
в) $\operatorname{Im} \mathcal{A} = V$*;*
г) $\operatorname{Ker} \mathcal{A} = \{\mathbf{0}\}$*.*

Доказательство. Мы докажем импликации а)⇒ б)⇒ в)⇒ г⇒ а).

а)⇒ б). Пусть $\det \mathcal{A} \neq 0$ и $A$ — матрица оператора $\mathcal{A}$ в любом базисе. Тогда $\det A \neq 0$. Следовательно, существует обратная матрица $A^{-1}$. Рассмотрим оператор $\mathcal{A}^{-1}$, который в выбранном базисе задаётся матрицей $A^{-1}$. Тогда по предыдущему предложению, матрица оператора $\mathcal{A} \circ \mathcal{A}^{-1}$ есть $AA^{-1} = E$, а значит $\mathcal{A} \circ \mathcal{A}^{-1} = \mathrm{id}$.

б)⇒ в). Пусть $\mathcal{A}$ обратим. Предположим, что $\operatorname{Im}\mathcal{A} \neq V$. Выберем такой $\boldsymbol{v} \in V$, что $\boldsymbol{v} \notin \operatorname{Im}\mathcal{A}$. Пусть $\mathcal{A}^{-1}(\boldsymbol{v}) = \boldsymbol{u}$. Тогда $\mathcal{A}\boldsymbol{u} \in \operatorname{Im}\mathcal{A}$. С другой стороны, $\mathcal{A}\boldsymbol{u} = \mathcal{A} \circ \mathcal{A}^{-1}(\boldsymbol{v}) = \boldsymbol{v}$. Противоречие.

в)⇒ г). Пусть $\operatorname{Im}\mathcal{A} = V$, т.е. $\dim V = \dim\operatorname{Im}\mathcal{A}$. Согласно следствию 1.7.7, $\dim V = \dim\operatorname{Ker}\mathcal{A} + \dim\operatorname{Im}\mathcal{A}$. Следовательно, $\dim\operatorname{Ker}\mathcal{A} = 0$, т.е. $\operatorname{Ker}\mathcal{A} = \{\mathbf{0}\}$.

г)⇒ а). Пусть $\operatorname{Ker}\mathcal{A} = \{\mathbf{0}\}$. Предположим, что $\det\mathcal{A} = 0$. Тогда $\det A = 0$, где $A$ — матрица оператора $\mathcal{A}$ в любом базисе $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$. Это означает, что система линейных уравнений $Ax = 0$ имеет ненулевое решение $x^1, \dots, x^n$. Согласно предложению 1.8.2, это означает, что $\mathcal{A}\boldsymbol{x} = \mathbf{0}$, где $\boldsymbol{x} = x^i\boldsymbol{e}_i \neq \mathbf{0}$. Таким образом, $\operatorname{Ker}\mathcal{A}$ содержит ненулевой вектор $\boldsymbol{x}$. Противоречие. □

Множество $\operatorname{Hom}(V, V)$ всех линейных операторов $\mathcal{A}\colon V \to V$ в фиксированном пространстве $V$ образует кольцо относительно операций сложения и композиции (если включить в рассмотрение и умножение операторов на элементы поля $\mathbf{k}$, то получаемый объект называется *алгеброй* над полем $\mathbf{k}$). Наряду с $\operatorname{Hom}(V, V)$ для этого кольца (или алгебры) используется обозначение $\operatorname{End}(V)$.

Невырожденные операторы в $V$ образуют (неабелеву) группу относительно композиции. Эта группа называется *общей линейной группой* пространства $V$ и обозначается $GL(V)$. Если $\dim V = n$, то группа $GL(V)$ изоморфна группе невырожденных квадратных матриц размера $n$ с элементами из поля $\mathbf{k}$ по умножению; эта группа обозначается $GL_n(\mathbf{k})$.

Матрицы (или операторы) с определителем 1 образуют подгруппу в $GL_n(\mathbf{k})$; эта подгруппа называется *специальной линейной группой* и обозначается $SL_n(\mathbf{k})$.

## 2.2. Проекторы, их алгебраическая характеризация

Определение 2.2.1. Пусть пространство $V$ представлено в виде прямой суммы двух подпространств: $V = V_1 \oplus V_2$. Тогда для любого вектора $\boldsymbol{v} \in V$ имеется единственное разложение $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$, $\boldsymbol{v}_1 \in V_1$, $\boldsymbol{v}_2 \in V_2$. Оператор $\mathcal{P}\colon V \to V$, переводящий вектор $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$ в вектор $\boldsymbol{v}_1$, называется *проектором* на $V_1$ вдоль $V_2$.

Для такого проектора $\mathcal{P}$ мы очевидно имеем $\operatorname{Im}\mathcal{P} = V_1$ и $\operatorname{Ker}\mathcal{P} = V_2$.

Теорема 2.2.2. *Следующие условия эквивалентны для оператора $\mathcal{A}\colon V \to V$:*

а) *$\mathcal{A}$ является проектором;*

б) $\mathcal{A}^2 = \mathcal{A}$.

Доказательство. Если $\mathcal{A}$ — проектор на $V_1$ вдоль $V_2$, то для $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$ имеем

$$\mathcal{A}^2\boldsymbol{v} = \mathcal{A}\big(\mathcal{A}(\boldsymbol{v}_1 + \boldsymbol{v}_2)\big) = \mathcal{A}\boldsymbol{v}_1 = \boldsymbol{v}_1 = \mathcal{A}\boldsymbol{v},$$

т.е. $\mathcal{A}^2 = \mathcal{A}$.

Пусть теперь $\mathcal{A}^2 = \mathcal{A}$. Положим $V_1 := \operatorname{Im}\mathcal{A}$ и $V_2 := \operatorname{Ker}\mathcal{A}$. Мы покажем, что $\mathcal{A}$ — проектор на $V_1$ вдоль $V_2$.

Сначала докажем, что $V = V_1 \oplus V_2$, т.е., что $V = V_1 + V_2$ и $V_1 \cap V_2 = \{\mathbf{0}\}$. Пусть $\boldsymbol{v} \in V_1 \cap V_2$. Тогда $\boldsymbol{v} \in V_1 = \operatorname{Im}\mathcal{A}$, т.е. существует такой вектор $\boldsymbol{u} \in V$, что $\mathcal{A}\boldsymbol{u} = \boldsymbol{v}$, и $\boldsymbol{v} \in V_2 = \operatorname{Ker}\mathcal{A}$, т.е. $\mathcal{A}\boldsymbol{v} = \mathbf{0}$. Тогда

$$\boldsymbol{v} = \mathcal{A}\boldsymbol{u} = \mathcal{A}^2\boldsymbol{u} = \mathcal{A}(\mathcal{A}\boldsymbol{u}) = \mathcal{A}\boldsymbol{v} = \mathbf{0},$$

т.е. $\boldsymbol{v} = \mathbf{0}$. Итак, $V_1$ и $V_2$ действительно образуют прямую сумму. Кроме того,

$$\dim(V_1 \oplus V_2) = \dim V_1 + \dim V_2 = \dim\operatorname{Im}\mathcal{A} + \dim\operatorname{Ker}\mathcal{A} = \dim V,$$

т.е. $V = V_1 \oplus V_2$.

Рассмотрим теперь произвольный вектор $\boldsymbol{v} \in V$ и представим его в виде $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$, $\boldsymbol{v}_1 \in V_1 = \operatorname{Im} \mathcal{A}$, $\boldsymbol{v}_2 \in V_2 = \operatorname{Ker} \mathcal{A}$. Тогда существует $\boldsymbol{u} \in V$ такой, что $\boldsymbol{v}_1 = \mathcal{A}\boldsymbol{u}$, а $\mathcal{A}\boldsymbol{v}_2 = \mathbf{0}$. Мы имеем

$$\mathcal{A}\boldsymbol{v} = \mathcal{A}\boldsymbol{v}_1 + \mathcal{A}\boldsymbol{v}_2 = \mathcal{A}(\mathcal{A}\boldsymbol{u}) = \mathcal{A}\boldsymbol{u} = \boldsymbol{v}_1.$$

Итак, $\mathcal{A}$ — действительно проектор на $V_1$ вдоль $V_2$. □

Матрица проектора на $V_1$ вдоль $V_2$ в базисе, составленном из базисов пространств $V_1$ и $V_2$ имеет вид $\left(\begin{array}{c|c} E & 0 \\ \hline 0 & 0 \end{array}\right)$, где $E$ — единичная матрица размера $k = \dim V_1$, а 0 обозначает матрицу из нулей соответствующего размера.

## 2.3. Многочлены от оператора. Минимальный аннулирующий многочлен

Пусть $\mathcal{A}\colon V \to V$ — оператор. Мы уже рассматривали его квадрат $\mathcal{A}^2$. На самом деле каждому многочлену $P(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n \in \mathbf{k}[t]$ можно сопоставить оператор

$$P(\mathcal{A}) := a_0 \operatorname{id} + a_1 \mathcal{A} + a_2 \mathcal{A}^2 + \ldots + a_n \mathcal{A}^n,$$

который называется *многочленом от оператора* $\mathcal{A}$.

Определение 2.3.1. Многочлен $P(t)$ называется *аннулирующим* оператор $\mathcal{A}$, если $P(\mathcal{A}) = \mathcal{O}$ (нулевой оператор).

Пример 2.3.2. Многочлен $t - 1$ аннулирует тождественный оператор id.

Предложение 2.3.3. *У любого оператора существует ненулевой аннулирующий многочлен.*

Доказательство. Пусть $\dim V = n$. Рассмотрим $n^2 + 1$ операторов $\mathcal{A}^0 = \operatorname{id}$, $\mathcal{A}^1 = \mathcal{A}, \mathcal{A}^2 \ldots, \mathcal{A}^{n^2}$. Так как размерность пространства операторов равна $n^2$, эти операторы линейно зависимы, т.е. существуют числа $a_0, a_1, \ldots, a_{n^2}$, не все равные нулю, такие, что $a_0 \operatorname{id} + a_1 \mathcal{A} + \ldots + a_{n^2} \mathcal{A}^{n^2} = \mathcal{O}$. Тогда ненулевой многочлен $P(t) = a_0 + a_1 t + \ldots + a_{n^2} t^{n^2}$ аннулирует оператор $\mathcal{A}$. □

Определение 2.3.4. Ненулевой многочлен $P(t)$ называется *минимальным аннулирующим многочленом* (или просто *минимальным многочленом*) для оператора $\mathcal{A}$, если $P(\mathcal{A}) = 0$, многочлен $P(t)$ имеет наименьшую степень среди всех ненулевых аннулирующих многочленов и его старший коэффициент равен 1.

Предложение 2.3.5. *Для любого оператора существует единственный минимальный аннулирующий многочлен.*

Доказательство. Мы уже доказали существование ненулевого аннулирующего многочлена. Среди всех аннулирующих многочленов выберем многочлен наименьшей степени и разделим его на старший коэффициент. В результате мы по определению получим минимальный многочлен.

Докажем единственность. Пусть $P_1(t)$ и $P_2(t)$ — два минимальных многочлена для оператора $\mathcal{A}$. Степени и старшие коэффициенты многочленов $P_1(t)$ и $P_2(t)$ равны. Тогда $P_1(t) - P_2(t)$ будет аннулирующим многочленом меньшей степени, а значит $P_1(t) - P_2(t) = 0$. □

## 2.4. Овеществление и комплексификация

При работе с линейными пространствами и операторами часто бывает удобно изменить поле скаляров. Здесь мы рассмотрим две такие операции: переход от пространств над полем вещественных чисел $\mathbb{R}$ к пространствам над полем комплексных чисел $\mathbb{C}$ (*комплексификация* вещественного пространства) и переход от пространств над $\mathbb{C}$ к пространствам над $\mathbb{R}$ (*овеществление* комплексного пространства).

**Овеществление.**

Определение 2.4.1. Пусть $V$ — комплексное пространство (пространство над полем $\mathbb{C}$). Рассмотрим пространство $V_{\mathbb{R}}$, состоящее из тех же векторов, что и $V$, но вместо операции умножения на все комплексные числа мы оставим лишь умножение не вещественные числа. Тогда $V_{\mathbb{R}}$ — вещественное пространство (пространство над полем $\mathbb{R}$), которое называется *овеществлением* пространства $V$.

Предложение 2.4.2. *Пусть* $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ *— базис пространства* $V$. *Тогда* $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$, $i\boldsymbol{e}_1, \dots, i\boldsymbol{e}_n$ *— базис пространства* $V_{\mathbb{R}}$. *Таким образом,* $\dim V_{\mathbb{R}} = 2 \dim V$.

Доказательство. Проверим, что векторы $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n, i\boldsymbol{e}_1, \dots, i\boldsymbol{e}_n$ линейно независимы в $V_{\mathbb{R}}$. Пусть

$$\lambda_1 \boldsymbol{e}_1 + \ldots + \lambda_n \boldsymbol{e}_n + \mu_1 i\boldsymbol{e}_1 + \ldots + \mu_n i\boldsymbol{e}_n = \boldsymbol{0},$$

в пространстве $V_{\mathbb{R}}$, где $\lambda_k, \mu_k \in \mathbb{R}$. Тогда в пространстве $V$ мы имеем

$$(\lambda_1 + i\mu_1)\boldsymbol{e}_1 + \ldots + (\lambda_n + i\mu_n)\boldsymbol{e}_n = \boldsymbol{0}.$$

Так как векторы $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ линейно независимы в $V$, мы имеем $\lambda_k + i\mu_k = 0$, т.е. $\lambda_k = \mu_k = 0$. Следовательно, $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n, i\boldsymbol{e}_1, \dots, i\boldsymbol{e}_n$ линейно независимы в $V_{\mathbb{R}}$.

Теперь проверим, что любой вектор $\boldsymbol{v} \in V_{\mathbb{R}}$ представляется в виде линейной комбинации (с вещественными коэффициентами) векторов $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n, i\boldsymbol{e}_1, \dots, i\boldsymbol{e}_n$. Рассмотрим $\boldsymbol{v}$ как вектор из $V$. Так как $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ — базис в $V$, мы имеем

$$\boldsymbol{v} = \alpha_1 \boldsymbol{e}_1 + \ldots + \alpha_n \boldsymbol{e}_n$$

для некоторых $\alpha_k \in \mathbb{C}$. Запишем $\alpha_k = \lambda_k + i\mu_k$, где $\lambda_k, \mu_k \in \mathbb{R}$. Тогда

$$\boldsymbol{v} = \lambda_1 \boldsymbol{e}_1 + \ldots + \lambda_n \boldsymbol{e}_n + \mu_1 i\boldsymbol{e}_1 + \ldots + \mu_n i\boldsymbol{e}_n.$$

Итак, $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n, i\boldsymbol{e}_1, \dots, i\boldsymbol{e}_n$ — базис в $V_{\mathbb{R}}$. □

Определение 2.4.3. Пусть $V$ — комплексное пространство и $\mathcal{A}\colon V \to V$ — оператор. Тогда тот же оператор, рассматриваемый в пространстве $\mathcal{A}_{\mathbb{R}}$, называется *овеществлением* оператора $\mathcal{A}$ и обозначается $\mathcal{A}_{\mathbb{R}}$.

Предложение 2.4.4. *Запишем матрицу оператора* $\mathcal{A}$ *в базисе* $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ *пространства* $V$ *в виде* $A + iB$, *где* $A$ *и* $B$ *— вещественные матрицы. Тогда*

а) *матрица оператора* $\mathcal{A}_{\mathbb{R}}$ *в базисе* $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n, i\boldsymbol{e}_1, \dots, i\boldsymbol{e}_n$ *есть* $\left(\begin{array}{c|c} A & -B \\ \hline B & A \end{array}\right)$;

б) $\det \mathcal{A}_{\mathbb{R}} = |\det \mathcal{A}|^2$.

Доказательство. Пусть $A = (a^l_k)$ и $B = (b^l_k)$. Тогда

$$\mathcal{A}_{\mathbb{R}}(\boldsymbol{e}_k) = \mathcal{A}(\boldsymbol{e}_k) = (a^l_k + ib^l_k)\boldsymbol{e}_l = a^l_k \boldsymbol{e}_l + b^l_k i\boldsymbol{e}_l,$$

$$\mathcal{A}_{\mathbb{R}}(i\boldsymbol{e}_k) = \mathcal{A}(i\boldsymbol{e}_k) = i\mathcal{A}(\boldsymbol{e}_k) = i(a^l_k + ib^l_k)\boldsymbol{e}_l = -b^l_k \boldsymbol{e}_l + a^l_k i\boldsymbol{e}_l,$$

и утверждение а) вытекает из определения матрицы оператора.

Для доказательства утверждения б) произведём следующие элементарные преобразования над строками и столбцами матрицы оператора $\mathcal{A}_{\mathbb{R}}$:

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

Отсюда получаем $\det \mathcal{A}_{\mathbb{R}} = \det(A-iB)\det(A+iB) = \overline{\det \mathcal{A}} \cdot \det \mathcal{A} = |\det \mathcal{A}|^2$. □

**Комплексная структура.**

Определение 2.4.5. Пусть $V$ — вещественное пространство. *Комплексной структурой* на $V$ называется такой оператор $\mathcal{J}\colon V \to V$, что $\mathcal{J}^2 = -\operatorname{id}$.

Предложение 2.4.6. *Пусть $V$ — вещественное пространство с комплексной структурой $\mathcal{J}$. Введём на $V$ операцию умножения на комплексные числа по правилу*

$$(\lambda + i\mu)\cdot \boldsymbol{v} = \lambda \boldsymbol{v} + \mu \mathcal{J}(\boldsymbol{v}).$$

*Тогда $V$ превращается в комплексное пространство $\widetilde{V}$, для которого $\widetilde{V}_{\mathbb{R}} = V$, а овеществление оператора умножения на $i$ есть $\mathcal{J}$.*

Доказательство. Необходимо проверить свойства 5)–8) из определения линейного пространства над $\mathbb{C}$. Эта проверка осуществляется непосредственно. □

Предложение 2.4.7. *Пусть $\mathcal{J}$ — комплексная структура на $V$. Тогда*

а) *размерность вещественного пространства $V$ чётна;*

б) *в подходящем базисе матрица оператора $\mathcal{J}$ имеет вид* $\left(\begin{array}{c|c} 0 & -E \\ \hline E & 0 \end{array}\right)$.

Доказательство. Рассмотрим комплексное пространство $\widetilde{V}$ из предыдущего предложения. Так как любой базис в $V$ порождает $\widetilde{V}$, это пространство конечномерно. Так как $V = \widetilde{V}_{\mathbb{R}}$, из предложения 2.4.2 следует, что $\dim V = 2\dim \widetilde{V}$ — чётно.

Далее, если $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n$ — базис комплексного пространства $\widetilde{V}$, то $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n$, $i\boldsymbol{e}_1,\ldots,i\boldsymbol{e}_n$ — базис пространства $\widetilde{V}_{\mathbb{R}} = V$. В этом базисе оператор $\mathcal{J}$ (овеществление оператора умножения на $i$) имеет указанный вид согласно предложению 2.4.4 а). □

**Комплексификация.**

Определение 2.4.8. Пусть $V$ — пространство над $\mathbb{R}$. Рассмотрим пространство $V \oplus V$, состоящее из пар $(\boldsymbol{u},\boldsymbol{v})$, где $\boldsymbol{u},\boldsymbol{v} \in V$, и введём на нём комплексную структуру следующим образом: $\mathcal{J}(\boldsymbol{u},\boldsymbol{v}) := (-\boldsymbol{v},\boldsymbol{u})$. Получаемое пространство $\widetilde{V \oplus V}$ над полем $\mathbb{C}$ называется *комплексификацией* пространства $V$ и обозначается $V_{\mathbb{C}}$.

Предложение 2.4.9. *Пусть $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n$ — базис пространства $V$. Тогда векторы $(\boldsymbol{e}_1,\boldsymbol{0}),\ldots,(\boldsymbol{e}_n,\boldsymbol{0})$ образуют базис пространства $V_{\mathbb{C}}$. Таким образом,* $\dim V_{\mathbb{C}} = \dim V$.

Доказательство. Как всегда, нужно проверить, что $(\boldsymbol{e}_1,\boldsymbol{0}),\ldots,(\boldsymbol{e}_n,\boldsymbol{0})$ линейно независимы и порождают всё пространство $V_{\mathbb{C}}$. Пусть

$$\alpha_1(\boldsymbol{e}_1,\boldsymbol{0}) + \ldots + \alpha_n(\boldsymbol{e}_n,\boldsymbol{0}) = (\boldsymbol{0},\boldsymbol{0}) \tag{2.1}$$

для некоторых $\alpha_k = \lambda_k + i\mu_k \in \mathbb{C}$. Так как $(\lambda_k + i\mu_k)\cdot(\boldsymbol{e}_k, \mathbf{0}) = \lambda_k(\boldsymbol{e}_k, \mathbf{0}) + \mu_k \mathcal{J}(\boldsymbol{e}_k, \mathbf{0}) = (\lambda_k \boldsymbol{e}_k, \mu_k \boldsymbol{e}_k)$, из равенства (2.1) мы получаем

$$(\lambda_1 \boldsymbol{e}_1 + \ldots + \lambda_n \boldsymbol{e}_n, \mu_1 \boldsymbol{e}_1 + \ldots + \mu_n \boldsymbol{e}_n) = (\mathbf{0}, \mathbf{0}),$$

откуда все $\lambda_k, \mu_k$, а значит и $\alpha_k$, равны нулю. Итак, $(\boldsymbol{e}_1, \mathbf{0}), \ldots, (\boldsymbol{e}_n, \mathbf{0})$ линейно независимы. То, что они порождают пространство $V_{\mathbb{C}}$, проверяется аналогично. □

Определение 2.4.10. Пусть $V$ — вещественное пространство и $\mathcal{A}\colon V \to V$ — оператор. Оператор $\mathcal{A}_{\mathbb{C}}\colon V_{\mathbb{C}} \to V_{\mathbb{C}}$, заданный формулой $\mathcal{A}_{\mathbb{C}}(\boldsymbol{u}, \boldsymbol{v}) := (\mathcal{A}\boldsymbol{u}, \mathcal{A}\boldsymbol{v})$, называется *комплексификацией* оператора $\mathcal{A}$.

Предложение 2.4.11. *Пусть $A$ — матрица оператора $\mathcal{A}$ в базисе $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Тогда оператор $\mathcal{A}_{\mathbb{C}}$ в базисе $(\boldsymbol{e}_1, \mathbf{0}), \ldots, (\boldsymbol{e}_n, \mathbf{0})$ имеет ту же матрицу $A$.*

Доказательство. $\mathcal{A}_{\mathbb{C}}(\boldsymbol{e}_k, \mathbf{0}) = (\mathcal{A}\boldsymbol{e}_k, \mathbf{0}) = (a^l_k \boldsymbol{e}_l, \mathbf{0}) = a^l_k(\boldsymbol{e}_l, \mathbf{0})$. □

При работе с комплексифицированным пространством $V_{\mathbb{C}}$ удобно записывать векторы $(\boldsymbol{u}, \boldsymbol{v}) \in V_{\mathbb{C}}$ в виде $\boldsymbol{u} + i\boldsymbol{v}$. Тогда действие комплексифицированного оператора $\mathcal{A}_{\mathbb{C}}$ записывается в виде $\mathcal{A}_{\mathbb{C}}(\boldsymbol{u} + i\boldsymbol{v}) = \mathcal{A}\boldsymbol{u} + i\mathcal{A}\boldsymbol{v}$.

Предложение 2.4.12. *Пространство $(V_{\mathbb{C}})_{\mathbb{R}}$ канонически изоморфно $V \oplus V$.*

Доказательство. Действительно, $V_{\mathbb{C}} = \widetilde{V \oplus V}$, а $(\widetilde{V \oplus V})_{\mathbb{R}} = V \oplus V$ согласно предложению 2.4.6 (мы просто сначала добавили, а потом убрали умножение на комплексные скаляры). □

Можно доказать (задача), что $(V_{\mathbb{R}})_{\mathbb{C}}$ канонически изоморфно $V \oplus \overline{V}$, где $\overline{V}$ — *комплексно сопряжённое пространство*, в котором сложение то же, что и в $V$, а умножение на комплексные числа определено по формуле $\lambda \cdot \boldsymbol{v} := \overline{\lambda}\boldsymbol{v}$.

Как мы вскоре убедимся, комплексификация предоставляет весьма полезный инструмент для работы с операторами в вещественных пространствах.

## 2.5. Инвариантные подпространства. Ограничение оператора и фактор-оператор. Собственные значения и собственные векторы

Определение 2.5.1. Подпространство $W \subset V$ называется *инвариантным* относительно оператора $\mathcal{A}\colon V \to V$, если $\mathcal{A}(W) \subset W$.

Пример 2.5.2. Ядро $\operatorname{Ker}\mathcal{A}$ и образ $\operatorname{Im}\mathcal{A}$ оператора $\mathcal{A}$ являются инвариантными подпространствами.

Пусть $W \subset V$ — инвариантное подпространство для оператора $\mathcal{A}\colon V \to V$. Выберем базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$ в $W$ и дополним его до базиса $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k, \boldsymbol{e}_{k+1}, \ldots, \boldsymbol{e}_n$ в $V$. Пусть $A = (a^i_j)$ — матрица оператора $\mathcal{A}$ в этом базисе. Тогда $\mathcal{A}\boldsymbol{e}_j = a^1_j \boldsymbol{e}_1 + \ldots + a^k_j \boldsymbol{e}_k$ при $j = 1, \ldots, k$. Это означает, что матрица $A$ имеет вид

$$A = \left(\begin{array}{c|c} * & * \\ \hline 0 & * \end{array}\right),$$

где в левом нижнем углу стоит матрица размера $(n-k) \times k$ из нулей.

Аналогично, если имеет место разложение $V = W_1 \oplus W_2$ в прямую сумму инвариантных подпространств, $\mathcal{A}(W_1) \subset W_1$ и $\mathcal{A}(W_2) \subset W_2$, то в подходящем базисе матрица оператора $\mathcal{A}$ будет иметь *блочно-диагональный вид*

$$A = \left(\begin{array}{c|c} * & 0 \\ \hline 0 & * \end{array}\right),$$

Определение 2.5.3. Пусть $W \subset V$ — инвариантное подпространство для оператора $\mathcal{A}\colon V \to V$. Тогда оператор $\widehat{\mathcal{A}}\colon W \to W$, определённый равенством $\widehat{\mathcal{A}}\boldsymbol{w} := \mathcal{A}\boldsymbol{w}$ для $\boldsymbol{w} \in W$, называется *ограничением* оператора $\mathcal{A}$ на подпространство $W$ и часто обозначается $\mathcal{A}|_W$.

Линейный оператор $\widetilde{\mathcal{A}}\colon V/W \to V/W$, определённый на классах смежности по правилу $\widetilde{\mathcal{A}}(\boldsymbol{v} + W) = \mathcal{A}\boldsymbol{v} + W$, называется *фактор-оператором.*

Определение фактор-оператора корректно. Действительно, если $\boldsymbol{v} + W = \boldsymbol{u} + W$, то $\boldsymbol{v} - \boldsymbol{u} \in W$, $\mathcal{A}(\boldsymbol{v} - \boldsymbol{u}) \in W$, и мы имеем

$$\widetilde{\mathcal{A}}(\boldsymbol{v} + W) = \mathcal{A}(\boldsymbol{u} + \boldsymbol{v} - \boldsymbol{u}) + W = \mathcal{A}\boldsymbol{u} + \mathcal{A}(\boldsymbol{v} - \boldsymbol{u}) + W = \mathcal{A}\boldsymbol{u} + W = \widetilde{\mathcal{A}}(\boldsymbol{u} + W).$$

Предложение 2.5.4. *Пусть $W \subset V$ — инвариантное подпространство для оператора $\mathcal{A}\colon V \to V$. Пусть $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$ — базис в $W$ и $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k, \boldsymbol{e}_{k+1}, \ldots, \boldsymbol{e}_n$ — базис в $V$. Тогда матрица оператора $\mathcal{A}$ в этом базисе имеет вид*

$$A = \left(\begin{array}{c|c} \widehat{A} & * \\ \hline 0 & \widetilde{A} \end{array}\right),$$

*где $\widehat{A}$ — матрица ограничения $\mathcal{A}|_W$ в базисе $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$, а $\widetilde{A}$ — матрица фактор-оператора в базисе $\boldsymbol{e}_{k+1} + W, \ldots, \boldsymbol{e}_n + W$ факторпространства $V/W$.*

Доказательство. Это вытекает из предыдущих рассуждений. □

Определение 2.5.5. Ненулевой вектор $\boldsymbol{v} \in V$ называется *собственным* для оператора $\mathcal{A}$, если $\mathcal{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ для некоторого $\lambda \in \mathbf{k}$.

Число $\lambda \in \mathbf{k}$ называется *собственным значением*, если существует собственный вектор $\boldsymbol{v}$, для которого $\mathcal{A}\boldsymbol{v} = \lambda\boldsymbol{v}$.

Предложение 2.5.6. *Все собственные векторы, отвечающие собственному значению $\lambda$, и вектор $\mathbf{0}$ образуют подпространство, которое совпадает с ядром оператора $\mathcal{A} - \lambda \cdot \mathrm{id}$.*

Доказательство. Равенство $\mathcal{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ имеет место тогда и только, когда $\boldsymbol{v} \in \operatorname{Ker}(\mathcal{A} - \lambda \cdot \mathrm{id})$. □

Определение 2.5.7. Пусть $\lambda$ — собственное значение для оператора $\mathcal{A}$. Подпространство $V_\lambda = \operatorname{Ker}(\mathcal{A} - \lambda \cdot \mathrm{id})$ называется *собственным подпространством*, соответствующим $\lambda$.

Предложение 2.5.8. *Собственное подпространство $V_\lambda$ инвариантно.*

Доказательство. Действительно, если $\boldsymbol{v} \in V_\lambda$, то $\mathcal{A}\boldsymbol{v} = \lambda\boldsymbol{v} \in V_\lambda$. □

Пример 2.5.9.

1. Для тождественного оператора $\mathrm{id}\colon V \to V$ все ненулевые векторы являются собственными с собственным значением 1.

2. Ядро любого оператора $\mathcal{A}\colon V \to V$ состоит из собственных векторов с собственным значением 0 и нулевого вектора.

3. Все собственные значения проектора $\mathcal{P}\colon V \to V$ суть 0 или 1. Причём если $\mathcal{P}$ — проектор на $U$ вдоль $W$, то $U$ — это собственное подпространство, соответствующее $\lambda = 1$, а $W$ — собственное подпространство, соответствующее $\lambda = 0$.

## 2.6. Характеристический многочлен. Теорема Гамильтона–Кэли

Определение 2.6.1. Многочлен $P_{\mathcal{A}}(t) := \det(\mathcal{A} - t \cdot \mathrm{id})$ называется *характеристическим многочленом* оператора $\mathcal{A}$.

Как и всякий определитель оператора, характеристический многочлен $P_{\mathcal{A}}(t)$ можно вычислять как $\det(A - tE)$, где $A$ — матрица оператора $\mathcal{A}$ в любом базисе, а $E$ — единичная матрица.

Некоторые из коэффициентов характеристического многочлена $P_{\mathcal{A}}(t) = a_n t^n + a_{n-1}t^{n-1} + \ldots + a_0$ нам уже знакомы: $a_n = (-1)^n$, $a_{n-1} = (-1)^{n-1} \operatorname{tr} \mathcal{A}$, $a_0 = \det \mathcal{A}$.

Предложение 2.6.2. *Собственные значения оператора $\mathcal{A}$ — это в точности корни его характеристического многочлена.*

Доказательство. Если $\lambda$ — собственное значение, то оператор $\mathcal{A} - \lambda \cdot \mathrm{id}$ вырожден, т.е. $\det(\mathcal{A} - \lambda \cdot \mathrm{id}) = 0$, а значит $\lambda$ — корень многочлена $P_{\mathcal{A}}(t)$. Обратно, если $\lambda$ — корень многочлена $P_{\mathcal{A}}(t)$, то $\det(\mathcal{A} - \lambda \cdot \mathrm{id}) = 0$, а значит $\lambda$ — собственное значение. $\square$

Теорема 2.6.3.

а) *Оператор $\mathcal{A}\colon V \to V$ в нетривиальном пространстве над полем $\mathbb{C}$ имеет инвариантное подпространство размерности 1.*

б) *Оператор $\mathcal{A}\colon V \to V$ в нетривиальном пространстве над полем $\mathbb{R}$ имеет инвариантное подпространство размерности 1 или 2.*

Доказательство. Для доказательства а) заметим, что так как поле $\mathbb{C}$ алгебраически замкнуто, характеристический многочлен $P_{\mathcal{A}}(t)$ имеет корень $\lambda$. Значит оператор $\mathcal{A}$ имеет собственный вектор $\boldsymbol{v}$, т.е. $\mathcal{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ и $\langle \boldsymbol{v} \rangle$ — одномерное инвариантное подпространство.

Докажем б). Если характеристический многочлен имеет вещественный корень, то мы получаем одномерное инвариантное подпространство. Предположим, что все корни многочлена $P_{\mathcal{A}}(t)$ комплексны. Пусть $\lambda + i\mu$ — корень. Тогда $\lambda + i\mu$ — собственное значение комплексифицированного оператора $\mathcal{A}_{\mathbb{C}}$ (напомним, что в подходящих базисах матрицы операторов $\mathcal{A}$ и $\mathcal{A}_{\mathbb{C}}$ совпадают). Возьмём соответствующий собственный вектор $\boldsymbol{u} + i\boldsymbol{v} \in V_{\mathbb{C}}$. Тогда

$$\mathcal{A}\boldsymbol{u} + i\mathcal{A}\boldsymbol{v} = \mathcal{A}_{\mathbb{C}}(\boldsymbol{u} + i\boldsymbol{v}) = (\lambda + i\mu)(\boldsymbol{u} + i\boldsymbol{v}) = (\lambda\boldsymbol{u} - \mu\boldsymbol{v}) + i(\mu\boldsymbol{u} + \lambda\boldsymbol{v}).$$

Следовательно, $\mathcal{A}\boldsymbol{u} = \lambda\boldsymbol{u} - \mu\boldsymbol{v}$ и $\mathcal{A}\boldsymbol{v} = \mu\boldsymbol{u} + \lambda\boldsymbol{v}$, и линейная оболочка $\langle \boldsymbol{u}, \boldsymbol{v} \rangle \subset V$ является инвариантным подпространством для $\mathcal{A}$. $\square$

Пусть $V_\lambda = \operatorname{Ker}(\mathcal{A} - \lambda \cdot \mathrm{id})$ — собственное подпространство.

Предложение 2.6.4. *Размерность собственного подпространства $V_\lambda$ не превосходит кратности $\lambda$ как корня характеристического многочлена.*

Доказательство. Пусть $\dim V_\lambda = k$. Выберем базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$ в пространстве $V_\lambda$ и дополним его до базиса в $V$. Так как $\mathcal{A}\boldsymbol{e}_i = \lambda \boldsymbol{e}_i$ при $i = 1, \ldots, k$, матрица оператора $\mathcal{A}$ в выбранном базисе имеет вид

$$A = \left(\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & * \\ 0 & & \lambda & \\ \hline & 0 & & \widetilde{A} \end{array}\right).$$

Тогда $P_\mathcal{A}(t) = \det(A - tE) = (\lambda - t)^k \det(\widetilde{A} - tE) = (\lambda - t)^k P_{\widetilde{\mathcal{A}}}(t)$, где $\widetilde{\mathcal{A}}$ — фактороператор. Отсюда вытекает, что кратность корня $\lambda$ не меньше $k = \dim V_\lambda$. □

Теорема 2.6.5 (Гамильтона–Кэли). *Характеристический многочлен $P_\mathcal{A}(t)$ оператора $\mathcal{A}\colon V \to V$ аннулирует этот оператор, т.е. $P_\mathcal{A}(\mathcal{A}) = \mathcal{O}$.*

Мы приведём два доказательства этого фундаментального факта. Первое доказательство более элементарное, но использует специальный трюк. Второе доказательство идейно проще, но использует понятие факторпространства.

Первое доказательство. Для квадратной матрицы $M$ обозначим через $\widehat{M}$ матрицу того же размера, на $ij$-м месте которой стоит алгебраическое дополнение $ji$-го элемента матрицы $M$. Как известно из курса алгебры, имеет место тождество

$$M\widehat{M} = \det M \cdot E.$$

Теперь возьмём в качестве $M$ матрицу $A - tE$, где $A$ — матрица оператора $\mathcal{A}$ в произвольном базисе. Тогда

$$(A - tE)(\widehat{A - tE}) = \det(A - tE) \cdot E = P_\mathcal{A}(t)E. \tag{2.2}$$

По определению элементы матрицы $\widehat{A - tE}$ являются многочленами от $t$ степени не выше $n - 1$, где $n = \dim V$. Следовательно, эту матрицу можно записать в виде

$$\widehat{A - tE} = B_0 + tB_1 + t^2 B_2 + \ldots + t^{n-1} B_{n-1},$$

где $B_i$ — числовые матрицы. Подставив это разложение вместе с разложением характеристического многочлена $P_\mathcal{A}(t) = a_0 + a_1 t + \ldots + a_n t^n$ в формулу (2.2), получим

$$(A - tE)(B_0 + tB_1 + t^2 B_2 + \ldots + t^{n-1} B_{n-1}) = (a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n)E.$$

Приравнивая коэффициенты при различных степенях $t$, получим

$$\begin{aligned} AB_0 &= a_0 E, \\ -B_0 + AB_1 &= a_1 E, \\ -B_1 + AB_2 &= a_2 E, \\ &\ldots \\ -B_{n-2} + AB_{n-1} &= a_{n-1} E \\ -B_{n-1} &= a_n E \end{aligned}$$

Умножив слева обе части второго равенства на $A$, третьего — на $A^2$, и т.д., и сложив все полученные равенства, получим

$$0 = a_0 E + a_1 A + a_2 A^2 + \ldots + a_n A^n.$$

В правой части стоит результат подстановки $A$ в характеристический многочлен. $\square$

Второе доказательство. Вначале докажем теорему над алгебраически замкнутым полем, например, над полем $\mathbb{C}$. Проведём индукцию по размерности $V$. Если $V$ одномерно, то $\mathcal{A} = \lambda \cdot \mathrm{id}$ — умножение на скаляр $\lambda$. Тогда $P_{\mathcal{A}}(t) = \lambda - t$, а значит $P_{\mathcal{A}}(\mathcal{A}) = \lambda \cdot \mathrm{id} - \mathcal{A} = \mathcal{O}$.

Пусть теперь $\dim V = n$, и предположим, что теорема доказана для пространств размерности $n-1$. Выберем инвариантное одномерное подпространство $U$ для $\mathcal{A}$; это подпространство порождено собственным вектором $\boldsymbol{u}$ с собственным значением $\lambda$. Дополним вектор $\boldsymbol{u}$ до базиса. В этом базисе матрица оператора $\mathcal{A}$ имеет вид

$$A = \left(\begin{array}{c|c} \lambda & * \\ \hline 0 & \widetilde{A} \end{array}\right),$$

где $\widetilde{A}$ — матрица фактор-оператора $\widetilde{\mathcal{A}}$, действующего в фактор-пространстве $V/U$. Отсюда вытекает, что

$$P_{\mathcal{A}}(t) = \det(A - tE) = (\lambda - t)\det(\widetilde{A} - tE) = (\lambda - t)P_{\widetilde{\mathcal{A}}}(t),$$

где $P_{\widetilde{\mathcal{A}}}(t)$ — характеристический многочлен фактор-оператора. Так как $\dim V/U = n-1$, по предположению индукции $P_{\widetilde{\mathcal{A}}}(\widetilde{\mathcal{A}}) = \mathcal{O}$. По определению фактор-оператора это означает, что для любого вектора $\boldsymbol{v} \in V$ имеем $P_{\widetilde{\mathcal{A}}}(\mathcal{A})\boldsymbol{v} \in U$, т.е. $P_{\widetilde{\mathcal{A}}}(\mathcal{A})\boldsymbol{v} = \mu\boldsymbol{u}$ для некоторого $\mu$. Следовательно,

$$P_{\mathcal{A}}(\mathcal{A})\boldsymbol{v} = (\lambda \cdot \mathrm{id} - \mathcal{A})P_{\widetilde{\mathcal{A}}}(\mathcal{A})\boldsymbol{v} = (\lambda \cdot \mathrm{id} - \mathcal{A})(\mu\boldsymbol{u}) = 0,$$

так как $\mathcal{A}\boldsymbol{u} = \lambda\boldsymbol{u}$. Итак, $P_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$ для операторов в комплексных пространствах.

Для доказательства теоремы над полем $\mathbb{R}$ воспользуемся комплексификацией $\mathcal{A}_{\mathbb{C}}$ оператора $\mathcal{A}$. Так как в соответствующих базисах пространств $V$ и $V_{\mathbb{C}}$ матрицы операторов $\mathcal{A}$ и $\mathcal{A}_{\mathbb{C}}$ совпадают, мы имеем $P_{\mathcal{A}}(t) = P_{\mathcal{A}_{\mathbb{C}}}(t)$ и $P_{\mathcal{A}}(A) = P_{\mathcal{A}_{\mathbb{C}}}(A) = 0$. $\square$

## 2.7. Диагонализируемые операторы. Критерий диагонализируемости

Определение 2.7.1. Оператор $\mathcal{A}$ называется *диагонализируемым*, если существует базис, в котором матрица этого оператора диагональна.

По определению матрицы оператора, базис, в котором матрица оператора диагональна, состоит из собственных векторов. Поэтому оператор диагонализируем тогда и только тогда, когда для него существует базис из собственных векторов.

Теорема 2.7.2 (критерий диагонализируемости). *Оператор $\mathcal{A}$ в $n$-мерном пространстве $V$ диагонализируем тогда и только тогда, когда его характеристический многочлен имеет в точности $n$ корней (с учётом кратностей), и размерность каждого собственного подпространства $V_\lambda$ равна кратности корня $\lambda$.*

Для доказательства теоремы нам понадобится лемма.

Лемма 2.7.3. *Собственные подпространства $V_{\lambda_1}, \ldots, V_{\lambda_k}$, соответствующие попарно различным собственным значениям $\lambda_1, \ldots, \lambda_k$ оператора $\mathcal{A}$, образуют прямую сумму $V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_k}$.*

Доказательство. Проведём индукцию по $k$. При $k = 1$ утверждение очевидно.

По определению прямой суммы мы должны проверить, что соотношение

$$v_1 + \ldots + v_k = \mathbf{0}, \tag{2.3}$$

где $v_i \in V_{\lambda_i}$, влечёт $v_1 = \ldots = v_k = \mathbf{0}$. Применив к (2.3) оператор $\mathcal{A}$, получим

$$\lambda_1 v_1 + \ldots + \lambda_k v_k = \mathbf{0}$$

Умножим (2.3) на $\lambda_k$ и вычтем из предыдущего соотношения:

$$(\lambda_1 - \lambda_k)v_1 + \ldots + (\lambda_{k-1} - \lambda_k)v_{k-1} = \mathbf{0}.$$

По предположению индукции получаем $(\lambda_1 - \lambda_k)v_1 = \ldots = (\lambda_{k-1} - \lambda_k)v_{k-1} = \mathbf{0}$. Так как по условию $\lambda_1 - \lambda_k \neq 0, \ldots, \lambda_{k-1} - \lambda_k \neq 0$, получаем $v_1 = \ldots = v_{k-1} = \mathbf{0}$. Тогда из (2.3) следует, что и $v_k = \mathbf{0}$. □

Сформулируем одно полезное следствие этой леммы.

Следствие 2.7.4. *Собственные векторы оператора, отвечающие различным собственным значениям, линейно независимы.*

Доказательство теоремы 2.7.2. Предположим, что оператор $\mathcal{A}$ диагонализируем. Пусть на диагонали матрицы $D$ оператора $\mathcal{A}$ стоят числа $\lambda_1, \ldots, \lambda_k$, причём число $\lambda_i$ присутствует $r_i$ раз. Тогда мы имеем $P_{\mathcal{A}}(t) = \det(D - tE) = \prod_{i=1}^{k}(\lambda_i - t)^{r_i}$. Следовательно, многочлен $P_{\mathcal{A}}(t)$ имеет $\sum_{i=1}^{k} r_i = n$ корней, и каждому корню $\lambda_i$ соответствует $r_i$ линейно независимых собственных векторов, т.е. $\dim V_{\lambda_i} = r_i$.

Предположим теперь, что многочлен $P_{\mathcal{A}}(t)$ имеет различные корни $\lambda_1, \ldots, \lambda_k$, причём кратность корня $\lambda_i$ равна $r_i$, $\sum_{i=1}^{k} r_i = n$ и $\dim V_{\lambda_i} = r_i$. Согласно лемме 2.7.3, пространства $V_{\lambda_1}, \ldots, V_{\lambda_k}$ образуют прямую сумму, а по условию сумма их размерностей равна $n = \dim V$. Следовательно, $V = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_k}$. Выбрав базис в каждом из подпространств $V_{\lambda_i}$ и взяв объединение этих базисов, мы получим базис пространства $V$, состоящий из собственных векторов. Итак, оператор $\mathcal{A}$ диагонализируем. □

Следствие 2.7.5. *Пусть характеристический многочлен $P_{\mathcal{A}}(t)$ имеет $n = \dim V$ различных корней. Тогда оператор $\mathcal{A}$ диагонализируем.*

Набор собственных значений оператора $\mathcal{A}$ часто называют его *спектром* (эта терминология будет прояснена в курсе функционального анализа, когда будут рассматриваться операторы с непрерывным спектром в бесконечномерных пространствах). Если все собственные значения имеют кратность 1 как корни характеристического многочлена, то говорят о *простом спектре*. Таким образом, операторы с простым спектром диагонализируемы. Появление кратных корней является «особенностью», которая устраняется произвольно малым возмущением коэффициентов матрицы оператора. Таким образом, над полем $\mathbb{C}$ «почти все» операторы диагонализируемы.

Пример 2.7.6.

1. Оператор, заданный матрицей $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ в стандартном базисе $\mathbb{R}^2$, не диагонализируем, так как его характеристический многочлен $t^2 + 1$ не имеет вещественных корней. Однако тот же оператор в $\mathbb{C}^2$ диагонализируем: в базисе $\boldsymbol{f}_1 = (1\, i)$, $\boldsymbol{f}_2 = (1\, {-i})$ его матрица $\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$ диагональна.

2. Оператор, заданный матрицей $\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$, не диагонализируем ни над каким полем по другой причине: его характеристический многочлен $(t-1)^2$ имеет корень 1 кратности 2, но при этом размерность соответствующего собственного подпространства равна 1 (вектор $\boldsymbol{e}_2 = (0\,1)$ не является собственным).

## 2.8. Нильпотентные операторы. Нормальный вид

Определение 2.8.1. Оператор $\mathcal{A}$ называется *нильпотентным*, если $\mathcal{A}^k = \mathcal{O}$ для некоторого $k$. Минимальное число $k$, для которого $\mathcal{A}^k = \mathcal{O}$, называется *степенью нильпотентности* оператора $\mathcal{A}$.

Пример 2.8.2. Рассмотрим оператор $\mathcal{A}$, заданный в базисе $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ матрицей

$$\begin{pmatrix}0 & 1 & & 0\\ & 0 & \ddots & \\ & & \ddots & 1\\ 0 & & & 0\end{pmatrix} \tag{2.4}$$

(над диагональю стоят единицы, а на остальных местах — нули). Действие этого оператора на базисные векторы описывается схемой $\boldsymbol{e}_n \mapsto \boldsymbol{e}_{n-1} \mapsto \ldots \mapsto \boldsymbol{e}_1 \mapsto \mathbf{0}$. Отсюда видно, что $\mathcal{A}^n = \mathcal{O}$, т.е. оператор $\mathcal{A}$ нильпотентен и имеет степень $n$.

Сформулируем несколько простых свойств нильпотентных операторов.

Предложение 2.8.3. *Пусть $\mathcal{A}\colon V \to V$ — нильпотентный оператор, причём $\dim V = n$. Тогда*

а) *единственным собственным значением оператора $\mathcal{A}$ является* 0*;*
б) *оператор $\mathcal{A}$ диагонализируем тогда и только тогда, когда $\mathcal{A} = \mathcal{O}$;*
в) *$\mathcal{A}^n = \mathcal{O}$, т.е. степень нильпотентности $\mathcal{A}$ не превосходит $n = \dim V$.*

Доказательство. Докажем а). Пусть $\mathcal{A}^k = \mathcal{O}$ и $\mathcal{A}^{k-1} \neq \mathcal{O}$. Значит существует такой вектор $\boldsymbol{v}$, что $\boldsymbol{u} := \mathcal{A}^{k-1}\boldsymbol{v} \neq \mathbf{0}$. Тогда $\mathcal{A}\boldsymbol{u} = \mathcal{A}^k\boldsymbol{v} = \mathbf{0}$, т.е. $\boldsymbol{u}$ — собственный вектор с собственным значением 0. Если теперь $\lambda \neq 0$ — другое собственное значение, то по определению найдётся $\boldsymbol{w} \neq \mathbf{0}$, такой, что $\mathcal{A}\boldsymbol{w} = \lambda\boldsymbol{w}$. Тогда $\mathbf{0} = \mathcal{A}^k\boldsymbol{w} = \lambda^k\boldsymbol{w}$. Отсюда $0 = \lambda^k$, т.е. $\lambda = 0$ — противоречие.

Докажем б). Если $\mathcal{A}$ диагонализируем, то на диагонали его диагональной матрицы стоят собственные значения, которые все равны нулю в силу а). Следовательно, матрица нулевая и $\mathcal{A} = \mathcal{O}$.

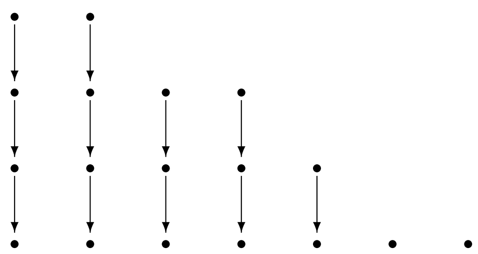
Докажем в). Из утверждения а) вытекает, что характеристический многочлен оператора $\mathcal{A}$ есть $(-t)^n$. Тогда $\mathcal{A}^n = \mathcal{O}$ по теореме Гамильтона–Кэли. □

Следующая теорема показывает, что любой нильпотентный оператор является прямой суммой операторов из примера 2.8.2.

Теорема 2.8.4. *Пусть $\mathcal{A}\colon V \to V$ — нильпотентный оператор. Тогда в пространстве $V$ существует базис, в котором матрица оператора $\mathcal{A}$ имеет блочно-диагональный вид с блоками из матриц* (2.4) *произвольных размеров. Такой вид матрицы оператора единствен с точностью до перестановки блоков.*

Базис, существование которого утверждается в этой теореме, называется *нормальным*, а матрица оператора в таком базисе называется *нормальным видом* (или *нормальной формой*) нильпотентного оператора.

Доказательство теоремы 2.8.4. Базис, в котором матрица оператора состоит из блоков вида (2.4), удобно изображать в виде диаграммы



В этой диаграмме точки изображают элементы нормального базиса, а стрелки описывают действие оператора $\mathcal{A}$. Элементы нижней строки оператор переводит в нуль, т.е. в ней стоят собственные векторы оператора (с собственным значением 0), входящие в базис. Каждый столбец соответствует одному блоку вида (2.4), причём размер блока равен высоте соответствующего столбца (количеству точек в столбце).

Итак, нам нужно доказать существование базиса, действие оператора $\mathcal{A}$ на элементы которого описывается диаграммой указанного вида. Проведём индукцию по размерности пространства $V$. Если $\dim V = 1$, то нильпотентный оператор $\mathcal{A}$ является нулевым, и любой ненулевой вектор в $V$ образует нормальный базис. Пусть теперь $\dim V = n > 1$, и пусть для размерностей, меньших $n$, существование нормального базиса уже доказано. Пусть $V_0 = \operatorname{Ker}\mathcal{A}$ — подпространство собственных векторов для $\mathcal{A}$. Так как $\dim V_0 > 0$, имеем $\dim V/V_0 < n$.

Рассмотрим фактор-оператор $\widetilde{\mathcal{A}}\colon V/V_0 \to V/V_0$, $\widetilde{\mathcal{A}}(\boldsymbol{v} + V_0) = \mathcal{A}\boldsymbol{v} + V_0$. По индуктивному предположению $\widetilde{\mathcal{A}}$ имеет нормальный базис. Можно считать его непустым: иначе $V = V_0$ и любой базис в $V_0$ будет нормальным для $\mathcal{A}$. Построим диаграмму $\widetilde{D}$ для элементов нормального базиса оператора $\widetilde{\mathcal{A}}$, в каждом её столбце возьмём самый верхний вектор $\widetilde{\boldsymbol{e}}_i$, $i = 1, \ldots, m$ (здесь $m$ — количество столбцов в $\widetilde{D}$), и положим $\widetilde{\boldsymbol{e}}_i = \boldsymbol{e}_i + V_0$, $\boldsymbol{e}_i \in V$. Теперь построим диаграмму $D$ из векторов пространства $V$ следующим образом. Для $i = 1, \ldots, m$ столбец с номером $i$ диаграммы $D$ будет состоять (сверху вниз) из векторов $\boldsymbol{e}_i, \mathcal{A}\boldsymbol{e}_i, \ldots, \mathcal{A}^{h_i-1}\boldsymbol{e}_i, \mathcal{A}^{h_i}\boldsymbol{e}_i$, где $h_i$ — высота $i$-го столбца в диаграмме $\widetilde{D}$. Так как $\widetilde{\mathcal{A}}^{h_i}\widetilde{\boldsymbol{e}}_i = \mathbf{0}$, мы имеем $\mathcal{A}^{h_i}\boldsymbol{e}_i \in V_0$ и $\mathcal{A}^{h_i+1}\boldsymbol{e}_i = \mathbf{0}$. Выберем базис в линейной оболочке $\langle \mathcal{A}^{h_1}\boldsymbol{e}_1, \ldots, \mathcal{A}^{h_m}\boldsymbol{e}_m\rangle \subset V_0$, дополним его до базиса $V_0$ и поставим дополняющие векторы в качестве новых столбцов (высоты один) в нижней строке диаграммы $D$; оператор $\mathcal{A}$ переводит их нуль.

Таким образом, построенная диаграмма $D$ из векторов пространства $V$ имеет в точности такой вид, как требуется для нормального базиса. Нужно лишь проверить, что векторы, составляющие диаграмму, действительно образуют базис в $V$.

Сначала покажем, что векторы из $D$ порождают всё $V$. Пусть $\boldsymbol{v} \in V$. Положим $\widetilde{\boldsymbol{v}} = \boldsymbol{v} + V_0$. По предположению $\widetilde{\boldsymbol{v}} = \sum_{i=1}^{m}\sum_{j=0}^{h_i-1}\lambda_{ij}\widetilde{\mathcal{A}}^j\widetilde{\boldsymbol{e}}_i$. Тогда

$$\boldsymbol{v} - \sum_{i=1}^{m}\sum_{j=0}^{h_i-1}\lambda_{ij}\mathcal{A}^j\boldsymbol{e}_i \in V_0.$$

Но все векторы $\mathcal{A}^j\boldsymbol{e}_i$, $j \leqslant h_i - 1$, лежат в строках диаграммы $D$, начиная со второй снизу, а подпространство $V_0$ порождено векторами из нижней строки $D$ по построению. Поэтому $\boldsymbol{v}$ можно представить в виде линейной комбинации векторов из $D$.

Остаётся проверить линейную независимость векторов из $D$. Сначала докажем, что векторы нижней строки линейно независимы. Действительно, если некоторая их нетривиальная линейная комбинация равна нулю, то она должна иметь вид $\sum_{i=1}^{m}\lambda_i\mathcal{A}^{h_i}\boldsymbol{e}_i=\mathbf{0}$, ибо остальные элементы нижней строки дополняют базис линейной оболочки $\langle\mathcal{A}^{h_1}\boldsymbol{e}_1,\ldots,\mathcal{A}^{h_m}\boldsymbol{e}_m\rangle$ до базиса $V_0$. Но все $h_i\geqslant 1$, поэтому

$$\mathcal{A}\Big(\sum_{i=1}^{m}\lambda_i\mathcal{A}^{h_i-1}\boldsymbol{e}_i\Big)=\mathbf{0},$$

так что

$$\sum_{i=1}^{m}\lambda_i\mathcal{A}^{h_i-1}\boldsymbol{e}_i\in V_0\qquad\text{и}\qquad\sum_{i=1}^{m}\lambda_i\widetilde{\mathcal{A}}^{h_i-1}\widetilde{\boldsymbol{e}}_i=\mathbf{0}.$$

Из последнего соотношения следует, что все $\lambda_i=0$, так как векторы $\widetilde{\mathcal{A}}^{h_i-1}\widetilde{\boldsymbol{e}}_i$ составляют нижнюю строку диаграммы $\widetilde{D}$ и являются частью базиса пространства $V/V_0$.

Наконец, покажем, что если имеется любая нетривиальная линейная комбинация векторов $D$, равная нулю, то из неё можно получить нетривиальную линейную зависимость между векторами нижней строки $D$. Отметим самую верхнюю строку $D$, в которой имеются ненулевые коэффициенты этой воображаемой линейной комбинации. Пусть номер этой строки (считая снизу) равен $h$. Применим к этой комбинации оператор $\mathcal{A}^{h-1}$. При этом её часть, лежащая в $h$-й строке, перейдёт в нетривиальную линейную комбинацию элементов нижней строки, а остальные слагаемые обратятся в нуль. Это завершает доказательство существования нормального базиса.

Теперь докажем единственность. Размеры блоков — это высоты столбцов диаграммы. Если расположить столбцы, как на рисунке, в порядке убывания, то их высоты однозначно определяются, если известны длины строк в диаграмме, начиная с нижней, в порядке убывания. Из предыдущего рассуждения следует, что длина нижней строки равна $\dim V_0=\dim\operatorname{Ker}\mathcal{A}$ и не зависит от выбора базиса. Длина второй снизу строки равна размерности ядра фактор-оператора $\widetilde{\mathcal{A}}$ в пространстве $V/\operatorname{Ker}\mathcal{A}$, т.е. $\dim\operatorname{Ker}\mathcal{A}^2-\dim\operatorname{Ker}\mathcal{A}$, что также не зависит от выбора базиса. Продолжая далее, мы видим, что длина $k$-й снизу строки равна размерности ядра фактор-оператора в пространстве $V/\operatorname{Ker}\mathcal{A}^{k-1}$, т.е. $\dim\operatorname{Ker}\mathcal{A}^k-\dim\operatorname{Ker}\mathcal{A}^{k-1}$. Это завершает доказательство единственности. □

Замечание. На практике для нахождения нормального базиса нильпотентного оператора используется следующая модификация процедуры, изложенной в доказательстве теоремы. Сначала находим степень нильпотентности $k$ оператора $\mathcal{A}$, возводя матрицу в степень, пока не получим 0. Векторы из первой сверху строки диаграммы $D$ соответствуют элементам базиса факторпространства $V/\operatorname{Ker}\mathcal{A}^{k-1}$. Для их нахождения мы выбираем максимальную линейно независимую систему векторов $V$, не лежащих в $\operatorname{Ker}\mathcal{A}^{k-1}$. Затем мы «спускаем» найденные векторы на одну строку вниз, применяя к ним оператор $\mathcal{A}$. Полученная система векторов лежит в $\operatorname{Ker}\mathcal{A}^{k-1}$, но не лежит в $\operatorname{Ker}\mathcal{A}^{k-2}$, и мы заполняем вторую строку, дополняя эти векторы до базиса в $\operatorname{Ker}\mathcal{A}^{k-1}/\operatorname{Ker}\mathcal{A}^{k-2}$. Затем мы спускаем все векторы из второй строки ещё на одну строку и заполняем третью строку, дополняя векторы, пришедшие из второй строки, до базиса в $\operatorname{Ker}\mathcal{A}^{k-2}/\operatorname{Ker}\mathcal{A}^{k-3}$. И так далее. На последнем шаге мы заполняем нижнюю строку, дополняя векторы, пришедшие сверху, до базиса в $\operatorname{Ker}\mathcal{A}$.

## 2.9. Корневые векторы. Теорема о разложении в прямую сумму корневых подпространств

Если оператор $\mathcal{A}$ имеет всего одно собственное значение $\lambda$, то его характеристический многочлен имеет вид $(-1)^n(t-\lambda)^n$, а значит $(\mathcal{A}-\lambda\cdot\mathrm{id})^n=0$ по теореме Гамильтона–Кэли. Следовательно, оператор $\mathcal{A}-\lambda\cdot\mathrm{id}$ нильпотентен и к нему можно применить теорему из предыдущего раздела.

В случае, когда имеется более одного различного собственного значения $\lambda$, соответствующие операторы $\mathcal{A}-\lambda\cdot\mathrm{id}$ не будут нильпотентными.

Определение 2.9.1. Вектор $\boldsymbol{v}\in V$ называется *корневым вектором* оператора $\mathcal{A}$, отвечающим числу $\lambda\in\mathbf{k}$, если существует такое $m$, что $(\mathcal{A}-\lambda\cdot\mathrm{id})^m\boldsymbol{v}=\mathbf{0}$.

Обозначим через $R_\lambda$ множество всех корневых векторов, отвечающих $\lambda$.

Предложение 2.9.2. *$R_\lambda$ является подпространством в $V$.*

Доказательство. Пусть $\boldsymbol{u},\boldsymbol{v}\in R_\lambda$, т.е. $(\mathcal{A}-\lambda\cdot\mathrm{id})^l\boldsymbol{u}=(\mathcal{A}-\lambda\cdot\mathrm{id})^m\boldsymbol{v}=\mathbf{0}$ для некоторых $l,m$. Тогда $(\mathcal{A}-\lambda\cdot\mathrm{id})^m(\mu\boldsymbol{u})=\mathbf{0}$ для любого $\mu\in\mathbf{k}$. Положим $p=\max\{l,m\}$. Тогда $(\mathcal{A}-\lambda\cdot\mathrm{id})^p(\boldsymbol{u}+\boldsymbol{v})=\mathbf{0}$, т.е. $R_\lambda$ — действительно подпространство. □

Подпространство $R_\lambda\subset V$ называется *корневым подпространством* для оператора $\mathcal{A}$, отвечающим $\lambda$.

Предложение 2.9.3. *Подпространство $R_\lambda$ нетривиально тогда и только тогда, когда $\lambda$ — собственное значение оператора $\mathcal{A}$. При этом $V_\lambda\subset R_\lambda$, т.е. корневое подпространство содержит собственное подпространство.*

Доказательство. Действительно, если $\lambda$ — собственное значение, то существует ненулевой вектор $\boldsymbol{v}$, для которого $(\mathcal{A}-\lambda\cdot\mathrm{id})\boldsymbol{v}=\mathbf{0}$, т.е. $\boldsymbol{v}\in R_\lambda$ и $R_\lambda$ нетривиально. Отсюда также следует, что $V_\lambda\subset R_\lambda$.

Обратно, пусть $R_\lambda$ содержит ненулевой вектор $\boldsymbol{u}$, для которого $(\mathcal{A}-\lambda\cdot\mathrm{id})^m\boldsymbol{u}=\mathbf{0}$, причём $m$ минимально, т.е. $\boldsymbol{v}:=(\mathcal{A}-\lambda\cdot\mathrm{id})^{m-1}\boldsymbol{u}\neq\mathbf{0}$. Тогда имеем $(\mathcal{A}-\lambda\cdot\mathrm{id})\boldsymbol{v}=(\mathcal{A}-\lambda\cdot\mathrm{id})^m\boldsymbol{u}=\mathbf{0}$, т.е. $\boldsymbol{v}$ — собственный вектор, отвечающий $\lambda$. □

Далее будем рассматривать только нетривиальные корневые подпространства $R_\lambda$.

Теорема 2.9.4. *Пусть $\mathcal{A}$ — оператор в пространстве $V$ над алгебраически замкнутым полем, и пусть $\lambda_1,\ldots,\lambda_k$ — все собственные значения оператора $\mathcal{A}$. Тогда $V$ является прямой суммой всех корневых подпространств, т.е.*

$$V=R_{\lambda_1}\oplus\ldots\oplus R_{\lambda_k}.$$

Доказательство теоремы будет опираться на три леммы.

Лемма 2.9.5. *Подпространство $R_\lambda$ инвариантно относительно любого оператора $\mathcal{A}-\mu\cdot\mathrm{id}$, $\mu\in\mathbf{k}$ (в частности, $R_\lambda$ инвариантно относительно $\mathcal{A}$). Ограничение*

$$(\mathcal{A}-\mu\cdot\mathrm{id})|_{R_\lambda}\colon R_\lambda\to R_\lambda$$

*при $\lambda\neq\mu$ является обратимым, а при $\lambda=\mu$ — нильпотентным оператором.*

Доказательство. Пусть $\boldsymbol{v}\in R_\lambda$, т.е. $(\mathcal{A}-\lambda\cdot\mathrm{id})^m\boldsymbol{v}=\mathbf{0}$. Тогда

$$(\mathcal{A}-\lambda\cdot\mathrm{id})^m(\mathcal{A}-\mu\cdot\mathrm{id})\boldsymbol{v}=(\mathcal{A}-\mu\cdot\mathrm{id})(\mathcal{A}-\lambda\cdot\mathrm{id})^m\boldsymbol{v}=\mathbf{0},$$

так как $(\mathcal{A}-\mu\cdot\mathrm{id})$ и $(\mathcal{A}-\lambda\cdot\mathrm{id})^m$ являются многочленами от оператора $\mathcal{A}$, а любые два многочлена от оператора коммутируют. Итак, $R_\lambda$ является $(\mathcal{A}-\mu\cdot\mathrm{id})$-инвариантным подпространством, и мы можем рассмотреть ограничение $(\mathcal{A}-\mu\cdot\mathrm{id})|_{R_\lambda}$.

Пусть $\boldsymbol{v}\in\mathrm{Ker}(\mathcal{A}-\mu\cdot\mathrm{id})|_{R_\lambda}$, т.е. $\boldsymbol{v}\in R_\lambda$ и $\mathcal{A}\boldsymbol{v}=\mu\boldsymbol{v}$. Тогда $(\mathcal{A}-\lambda\cdot\mathrm{id})\boldsymbol{v}=(\mu-\lambda)\boldsymbol{v}$, а значит

$$(\mu-\lambda)^m\boldsymbol{v}=(\mathcal{A}-\lambda\cdot\mathrm{id})^m\boldsymbol{v}=\boldsymbol{0}.$$

Следовательно, если $\lambda\neq\mu$, то $\boldsymbol{v}=0$. Таким образом, при $\lambda\neq\mu$ мы получаем, что ядро оператора $(\mathcal{A}-\mu\cdot\mathrm{id})|_{R_\lambda}$ тривиально, а значит этот оператор обратим.

Наконец, если $\boldsymbol{e}_1,\dots,\boldsymbol{e}_r$ — базис в $R_\lambda$ и $(\mathcal{A}-\lambda\cdot\mathrm{id})^{m_i}\boldsymbol{e}_i=\boldsymbol{0}$, то $(\mathcal{A}-\lambda\cdot\mathrm{id})^m\boldsymbol{v}=\boldsymbol{0}$ для любого вектора $\boldsymbol{v}\in R_\lambda$, где $m$ — наибольшее из чисел $m_1,\dots,m_r$. Это означает, что оператор $(\mathcal{A}-\lambda\cdot\mathrm{id})|_{R_\lambda}$ нильпотентен. $\square$

Лемма 2.9.6. *Корневые подпространства $R_{\lambda_1},\dots,R_{\lambda_k}$, соответствующие различным собственным значениям $\lambda_1,\dots,\lambda_k$, образуют прямую сумму $R_{\lambda_1}\oplus\dots\oplus R_{\lambda_k}$.*

Доказательство. Проведём индукцию по $k$. При $k=1$ утверждение очевидно. Предположим, что утверждение доказано для $k-1$ подпространств.

Докажем, что соотношение

$$\boldsymbol{v}_1+\dots+\boldsymbol{v}_k=\boldsymbol{0}, \tag{2.5}$$

где $\boldsymbol{v}_i\in R_{\lambda_i}$, влечёт $\boldsymbol{v}_1=\dots=\boldsymbol{v}_k=\boldsymbol{0}$. Имеем $(\mathcal{A}-\lambda_k\cdot\mathrm{id})^p\boldsymbol{v}_k=\boldsymbol{0}$ для некоторого $p$. Применив к (2.5) оператор $(\mathcal{A}-\lambda_k\cdot\mathrm{id})^p$, получим

$$(\mathcal{A}-\lambda_k\cdot\mathrm{id})^p\boldsymbol{v}_1+\dots+(\mathcal{A}-\lambda_k\cdot\mathrm{id})^p\boldsymbol{v}_{k-1}=\boldsymbol{0}. \tag{2.6}$$

Так как подпространства $R_{\lambda_1},\dots,R_{\lambda_{k-1}}$ инвариантны относительно $\mathcal{A}-\lambda_k\cdot\mathrm{id}$, мы имеем $(\mathcal{A}-\lambda_k\cdot\mathrm{id})^p\boldsymbol{v}_i\in R_{\lambda_i}$, $i=1,\dots,k-1$. Тогда по предположению индукции из (2.6) вытекает, что $(\mathcal{A}-\lambda_k\cdot\mathrm{id})^p\boldsymbol{v}_i=\boldsymbol{0}$, $i=1,\dots,k-1$. Так как по предыдущей лемме оператор $\mathcal{A}-\lambda_k\cdot\mathrm{id}$ в пространствах $R_{\lambda_1},\dots,R_{\lambda_{k-1}}$ обратим, отсюда следует, что $\boldsymbol{v}_1=\dots=\boldsymbol{v}_{k-1}=\boldsymbol{0}$. Тогда из (2.5) получаем, что и $\boldsymbol{v}_k=\boldsymbol{0}$. $\square$

Лемма 2.9.7. *Размерность корневого подпространства $R_\lambda$ равна кратности $\lambda$ как корня характеристического многочлена оператора $\mathcal{A}$.*

Доказательство. Обозначим через $r_\lambda$ кратность корня $\lambda$. Пусть $\widehat{\mathcal{A}}=\mathcal{A}|_{R_\lambda}$ — ограничение оператора $\mathcal{A}$ на $R_\lambda$ и $\widetilde{\mathcal{A}}\colon V/R_\lambda\to V/R_\lambda$ — фактор-оператор. Тогда для характеристических многочленов мы имеем

$$P_{\mathcal{A}}(t)=P_{\widehat{\mathcal{A}}}(t)P_{\widetilde{\mathcal{A}}}(t)=(\lambda-t)^{\dim R_\lambda}P_{\widetilde{\mathcal{A}}}(t) \tag{2.7}$$

(см. предложение 2.5.4), откуда $\dim R_\lambda\leqslant r_\lambda$.

Предположим, что $\dim R_\lambda<r_\lambda$. Тогда из (2.7) следует, что $\lambda$ является корнем многочлена $P_{\widetilde{\mathcal{A}}}(t)$, т.е. собственным значением оператора $\widetilde{\mathcal{A}}$. Пусть $\boldsymbol{v}+R_\lambda$ — соответствующий (ненулевой) собственный вектор, т.е.

$$\widetilde{\mathcal{A}}(\boldsymbol{v}+R_\lambda)=\lambda(\boldsymbol{v}+R_\lambda)\qquad\text{или}\qquad\mathcal{A}\boldsymbol{v}+R_\lambda=\lambda\boldsymbol{v}+R_\lambda.$$

Отсюда вытекает, что $\mathcal{A}\boldsymbol{v}-\lambda\boldsymbol{v}=(\mathcal{A}-\lambda\cdot\mathrm{id})\boldsymbol{v}\in R_\lambda$. По определению $R_\lambda$ это означает, что $\boldsymbol{0}=(\mathcal{A}-\lambda\cdot\mathrm{id})^m(\mathcal{A}-\lambda\cdot\mathrm{id})\boldsymbol{v}=(\mathcal{A}-\lambda\cdot\mathrm{id})^{m+1}\boldsymbol{v}$, т.е. $\boldsymbol{v}\in R_\lambda$. Но тогда $\boldsymbol{v}+R_\lambda$ — нулевой вектор пространства $V/R_\lambda$. Противоречие. $\square$

Доказательство теоремы 2.9.4. Пусть $\dim V = n$ и пусть $r_i$ — кратность корня $\lambda_i$, $i = 1\ldots,k$. Тогда $\sum_{i=1}^k r_i = n$ (здесь мы пользуемся алгебраической замкнутостью поля) и из двух предыдущих лемм вытекает, что $\dim(R_{\lambda_1}\oplus\ldots\oplus R_{\lambda_k}) = \sum_{i=1}^k r_i = \dim V$. Следовательно, $V = R_{\lambda_1}\oplus\ldots\oplus R_{\lambda_k}$. □

Разложение $V = R_{\lambda_1}\oplus\ldots\oplus R_{\lambda_k}$ называется *корневым разложением* для $V$.

## 2.10. Жорданова нормальная форма оператора

Давайте посмотрим, как выглядит матрица оператора $\mathcal{A}|_{R_\lambda}$ (ограничения оператора $\mathcal{A}$ на корневое подпространство $R_\lambda$). Так как $(\mathcal{A}-\lambda\cdot\mathrm{id})|_{R_\lambda}$ является нильпотентным оператором (лемма 2.9.5), в пространстве $R_\lambda$ можно выбрать нормальный базис для этого оператора. Тогда матрица оператора $(\mathcal{A}-\lambda\cdot\mathrm{id})|_{R_\lambda}$ в этом базисе будет состоять из блоков вида (2.4), а значит матрица оператора $\mathcal{A}|_{R_\lambda}$ в том же базисе будет состоять из блоков вида

$$J_\lambda = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} \tag{2.8}$$

(на диагонали стоит $\lambda$, над диагональю — единицы, а на остальных местах нули).

Определение 2.10.1. Матрица (2.8) называется *жордановой клеткой*. Если матрица оператора $\mathcal{A}$ в некотором базисе является блочно-диагональной с блоками вида (2.8) (возможно, соответствующими различным $\lambda$), то такая матрица называется *жордановой нормальной формой* оператора $\mathcal{A}$. Базис, в котором оператор имеет жорданову нормальную форму, называется *жордановым*.

Теорема 2.10.2. *Для любого оператора $\mathcal{A}$ в пространстве $V$ над алгебраически замкнутым полем существует жорданов базис (в котором оператор имеет жорданову нормальную форму). Жорданова нормальная форма оператора единственна с точностью до перестановки блоков (клеток).*

Доказательство. Существование жордановой формы является прямым следствием теорем о разложении в сумму корневых подпространств и существования нормального вида для нильпотентных операторов. Действительно, пусть $\lambda_1,\ldots,\lambda_k$ — все собственные значения $\mathcal{A}$. Выберем в каждом корневом пространстве $R_{\lambda_i}$ нормальный базис для нильпотентного оператора $(\mathcal{A}-\lambda_i\cdot\mathrm{id})|_{R_{\lambda_i}}$. Тогда объединение этих базисов даст жорданов базис для оператора $\mathcal{A}$ в силу наличия корневого разложения $V = R_{\lambda_1}\oplus\ldots\oplus R_{\lambda_k}$ (здесь мы пользуемся алгебраической замкнутостью поля).

Докажем единственность жордановой формы. Надо показать, что количество жордановых клеток фиксированного размера с одним и тем же $\lambda$ не зависит от способа приведения к жордановой форме (т.е. от выбора жорданова базиса). Выберем произвольный жорданов базис. Пусть $W_{\lambda_i}$ — линейная оболочка части этого базиса, отвечающей всем клеткам с $\lambda_i$ на диагонали. Тогда ограничение оператора $(\mathcal{A}-\lambda_i\cdot\mathrm{id})$ на $W_{\lambda_i}$ — нильпотентный оператор, а значит $W_{\lambda_i}\subset R_{\lambda_i}$. Кроме того, $V = W_{\lambda_1}\oplus\ldots\oplus W_{\lambda_k}$ по определению жорданова базиса и $V = R_{\lambda_1}\oplus\ldots\oplus R_{\lambda_k}$ (корневое разложение). Следовательно, $\dim W_{\lambda_i} = \dim R_{\lambda_i}$ и $W_{\lambda_i} = R_{\lambda_i}$. Итак, подпространства, отвечающие клеткам с собственным значением $\lambda_i$, не зависят от способа приведения к жордановой форме и равны $R_{\lambda_i}$.

Таким образом, мы свели доказательство единственности жордановой формы к случаю, когда оператор $\mathcal{A}$ имеет одно собственное значение $\lambda$. Любой жорданов базис для такого оператора будет также нормальным базисом для нильпотентного оператора $\mathcal{A} - \lambda \cdot \mathrm{id}$. Для нильпотентных операторов мы уже доказали единственность нормального вида (т.е. жордановой формы) в теореме 2.8.4. □

На языке матриц данная теорема означает, что любая квадратная комплексная матрица $A \in \mathrm{Mat}_n(\mathbb{C})$ подобна матрице $J$, состоящей из жордановых клеток, т.е. $J = C^{-1}AC$ для некоторой невырожденной матрицы $C$.

Замечания. В отличие от жордановой формы, жорданов базис оператора далеко не единствен. Например, для тождественного оператора id любой базис будет жордановым.

Теорема 2.10.2 не имеет места над полем $\mathbb{R}$. Например, оператор, заданный матрицей $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ в $\mathbb{R}^2$, не приводится к жордановой форме (докажите).

На практике для нахождения жордановой формы и жорданова базиса оператора, заданного матрицей $A$, можно использовать следующий алгоритм. Сначала находим характеристический многочлен и его корни $\lambda_1, \ldots, \lambda_k$ (собственные значения) с кратностями. Затем находим корневые подпространства $R_{\lambda_i}$. Для этого возводим матрицу $A - \lambda_i E$ в степень до тех пор, пока не наступит стабилизация ранга: $\mathrm{rk}(A - \lambda_i E)^{m_i} = \mathrm{rk}(A - \lambda_i E)^{m_{i+1}}$. Тогда $R_{\lambda_i} = \mathrm{Ker}(A - \lambda_i E)^{m_i}$, а число $m_i$ будет размером максимальной жордановой клетки, отвечающей $\lambda_i$. Далее в каждом пространстве $R_{\lambda_i}$ находим нормальный базис для нильпотентного оператора $(\mathcal{A} - \lambda_i \cdot \mathrm{id})|_{R_{\lambda_i}}$ (как описано в конце раздела 2.8); объединение этих базисов и будет жордановым базисом для $\mathcal{A}$.

Зная жорданову форму, легко вычислить минимальный многочлен оператора.

Следствие 2.10.3. *Минимальный аннулирующий многочлен оператора $\mathcal{A}$ над полем $\mathbb{C}$ есть $P(t) = \prod_{i=1}^{k}(t - \lambda_i)^{m_i}$, где $\lambda_1, \ldots, \lambda_k$ — все собственные значения $\mathcal{A}$, а $m_i$ — размер максимальной жордановой клетки, отвечающей $\lambda_i$.*

## 2.11. Вычисление многочленов и функций от матриц. Экспонента линейного оператора (без обоснования сходимости)

Одним из важных применений жордановой формы является эффективное вычисление многочленов и функций от операторов (матриц).

Прежде всего получим формулу для многочлена от жордановой клетки.

Предложение 2.11.1. *Пусть $f(t)$ — многочлен. Тогда его значение на жордановой клетке* (2.8) *размера $n$ вычисляется по формуле*

$$f(J_\lambda) = \begin{pmatrix} f(\lambda) & \frac{f'(\lambda)}{1!} & \ldots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ & \ddots & \ddots & \vdots \\ & & f(\lambda) & \frac{f'(\lambda)}{1!} \\ 0 & & & f(\lambda) \end{pmatrix}. \tag{2.9}$$

Доказательство. Вначале по индукции проверим эту формулу для многочлена $f(t) = t^m$, т.е. для $m$-й степени жордановой клетки. Пусть $J_\lambda^m = (a_{ij}^m)$, т.е. $(ij)$-й элемент матрицы $J_\lambda^m$ есть $a_{ij}^m$. По предположению индукции для $f(t) = t^{m-1}$ имеем

$$a_{ij}^{m-1} = \frac{f^{(j-i)}(\lambda)}{(j-i)!} = C_{m-1}^{j-i}\lambda^{m-1-j+i},$$

где $C_k^i = \frac{k!}{i!(k-i)!}$ — биномиальный коэффициент; мы считаем $C_k^i = 0$ при $i < 0$ или $i > k$. Тогда из соотношения $J_\lambda^m = J_\lambda^{m-1}J_\lambda$ по правилу умножения матриц вычисляем

$$a_{ik}^m = \sum_{j=1}^n a_{ij}^{m-1}a_{jk}^1 = \sum_{j=1}^n C_{m-1}^{j-i}\lambda^{m-1-j+i}C_1^{k-j}\lambda^{1-k+j} =$$
$$= C_{m-1}^{k-1-i}\lambda^{m+i-k} + C_{m-1}^{k-i}\lambda^{m+i-k} = C_m^{k-i}\lambda^{m-k+i},$$

что и требовалось доказать. Осталось заметить, что формула (2.9) линейна по $f$, а значит она верна для любого многочлена $f(t)$. □

На основе формулы (2.9) мы можем также вычислять многочлены от матриц в жордановой форме $J$, так как многочлен применяется к такой матрице поблочно.

Теперь если $A$ — произвольная матрица, то мы можем привести её к жордановой форме, т.е. найти жорданову матрицу $J$, для которой $A = CJC^{-1}$. При возведении матрицы $A$ в степень мы получаем

$$A^m = (CJC^{-1})^m = CJC^{-1}CJC^{-1}\dots CJC^{-1} = CJ^mC^{-1}.$$

Тогда аналогичная формула верна и для произвольного многочлена $f(t)$:

$$f(A) = Cf(J)C^{-1}. \tag{2.10}$$

При помощи этой формулы мы можем вычислить любой многочлен от матрицы $A$, зная её жорданову форму и жорданов базис.

Определение 2.11.2. Пусть $\mathcal{A}$ — оператор в комплексном пространстве с собственными значениями $\lambda_1, \dots, \lambda_k$ кратностей $r_{\lambda_1}, \dots, r_{\lambda_k}$ соответственно. Говорят, что два многочлена $f(t)$ и $g(t)$ *совпадают на спектре оператора* $\mathcal{A}$, если

$$f^{(j)}(\lambda_i) = g^{(j)}(\lambda_i) \quad \text{при } i = 1, \dots, k,\ j = 0, \dots, r_{\lambda_i} - 1. \tag{2.11}$$

Предложение 2.11.3. *Если два многочлена $f(t)$ и $g(t)$ совпадают на спектре оператора $\mathcal{A}$, то $f(\mathcal{A}) = g(\mathcal{A})$.*

Доказательство. Действительно, при вычислении многочлена $f(t)$ от матрицы $A$ по формулам (2.9) и (2.10) используются производные многочлена $f(t)$ в собственных значениях $\lambda_i$ порядка не выше $r_{\lambda_i} - 1$. □

Это утверждение позволяет находить многочлен $f(t)$ большой степени от матрицы $A$ (например, возводить матрицу в большую степень), вовсе не вычисляя её жордановой формы, следующим образом. Если матрица $A$ размера $n$, то можно найти многочлен $g(t)$ степени не выше $n-1$, удовлетворяющий соотношениям (2.11), при помощи формул интерполяции или методом неопределённых коэффициентов. Тогда мы имеем $f(A) = g(A)$. Если же нам известна жорданова форма матрицы $A$, то можно ещё снизить степень многочлена $g$, заменив в соотношениях (2.11) числа

$r_{\lambda_i}$ на числа $m_{\lambda_i}$ — размеры максимальных жордановых клеток с $\lambda_i$ на диагонали (жорданов базис для этого знать не обязательно).

На самом деле формулы (2.9) и (2.10) можно использовать также для вычисления более общих функций $f$ от операторов или матриц (а не только многочленов). Если функция $f$ гладкая и хорошо приближается многочленами (такие функции называются *аналитическими*) в окрестности собственных значений матрицы $A$, то можно определить $f(A)$ как предел последовательности многочленов, получаемых обрезанием ряда Тейлора для $f$. При этом, однако, необходимо обосновывать сходимость получаемых последовательностей (или рядов) из матриц. Мы этого делать не будем, а скажем лишь, что таким образом можно определить $e^A$, $\sin A$, $\cos A$, а также $\sqrt{A}$ и $\ln A$ для матриц с положительными собственными значениями. Можно также использовать формулы (2.9) и (2.10) в качестве *определения* $f(A)$ для функций $f$, которые определены на спектре $A$.

Мы рассмотрим экспоненту более подробно.

Определение 2.11.4. *Экспонентой* оператора $\mathcal{A}$ называется оператор

$$e^{\mathcal{A}} = \sum_{k=0}^{\infty} \frac{1}{k!}\mathcal{A}^k = \mathrm{id} + \mathcal{A} + \frac{\mathcal{A}^2}{2} + \frac{\mathcal{A}^3}{6} + \ldots.$$

Вот простое, но важное свойство экспоненты.

Предложение 2.11.5. *Имеет место соотношение* $\det e^{\mathcal{A}} = e^{\operatorname{tr}\mathcal{A}}$.

Доказательство. Приведём матрицу оператора к жордановой форме: $A = CJC^{-1}$. Используя формулы (2.9) и (2.10) для $f(t) = e^t$, вычисляем

$$\det e^A = \det(Ce^JC^{-1}) = \det e^J = \prod_{i=1}^{k} e^{r_{\lambda_i}\lambda_i} = e^{\sum_{i=1}^{k} r_{\lambda_i}\lambda_i} = e^{\operatorname{tr} J} = e^{\operatorname{tr} A}.$$

□

Основное свойство числовой экспоненты $e^a e^b = e^{a+b}$, вообще говоря, *нарушается* для экспоненты операторов. Однако есть важный случай, когда оно выполнено.

Предложение 2.11.6. *Если операторы* $\mathcal{A}, \mathcal{B}\colon V \to V$ *коммутируют, т.е.* $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$, *то* $e^{\mathcal{A}}e^{\mathcal{B}} = e^{\mathcal{A}+\mathcal{B}}$.

Доказательство. Мы имеем:

$$e^{\mathcal{A}}e^{\mathcal{B}} = \Bigl(\sum_{i=0}^{\infty}\frac{1}{i!}\mathcal{A}^i\Bigr)\Bigl(\sum_{j=0}^{\infty}\frac{1}{j!}\mathcal{B}^j\Bigr) = \sum_{i,j\geqslant 0}\frac{1}{i!j!}\mathcal{A}^i\mathcal{B}^j = \sum_{k\geqslant 0}\sum_{i=0}^{k}\frac{1}{i!(k-i)!}\mathcal{A}^i\mathcal{B}^{k-i} =$$

$$= \sum_{k\geqslant 0}\frac{1}{k!}\sum_{i=0}^{k}\frac{k!}{i!(k-i!)}\mathcal{A}^i\mathcal{B}^{k-i} = \sum_{k\geqslant 0}\frac{1}{k!}(\mathcal{A}+\mathcal{B})^k = e^{\mathcal{A}+\mathcal{B}}.$$

Коммутативность $\mathcal{A}$ и $\mathcal{B}$ используется в том месте, где $(\mathcal{A}+\mathcal{B})^k$ раскладывается по формуле бинома Ньютона. □

Глава 3

# Геометрия евклидовых и эрмитовых пространств

## 3.1. Аффинные пространства, системы координат, подпространства

В геометрии на плоскости или в пространстве рассматриваются точки и векторы. Для формализацией этих понятий и взаимосвязей между ними служит понятие аффинного пространства.

Определение 3.1.1. *Аффинным пространством* называется пара $(\mathfrak{A}, V)$, состоящая из множества $\mathfrak{A}$, элементы которого называются *точками*, и векторного пространства $V$ над полем $\mathbf{k}$, с дополнительной операцией сложения

$$+\colon \mathfrak{A} \times V \to \mathfrak{A}, \quad (P, \boldsymbol{v}) \mapsto P + \boldsymbol{v}$$

для $P \in \mathfrak{A}$ и $\boldsymbol{v} \in V$. (Говоря неформально к точке $P$ можно «приложить» вектор $\boldsymbol{v}$ и тогда его «конец» — это точка $P+\boldsymbol{v}$.) При этом требуется, чтобы операция сложения точек и векторов удовлетворяла следующим условиям:

1) $P + \mathbf{0} = P$ для любой точки $P \in \mathfrak{A}$;
2) $(P + \boldsymbol{u}) + \boldsymbol{v} = P + (\boldsymbol{u} + \boldsymbol{v})$ для любых $P \in \mathfrak{A}$, $\boldsymbol{u}, \boldsymbol{v} \in V$;
3) для любых $P, Q \in \mathfrak{A}$ существует единственный вектор $\boldsymbol{v} \in V$, такой, что $P + \boldsymbol{v} = Q$.

Часто аффинным пространством называют просто множество точек $\mathfrak{A}$ из определения выше (особенно когда из контекста понятно, какое векторное пространство $V$ имеется ввиду). Вектор $\boldsymbol{v}$, однозначно сопоставляемый паре точек $P, Q \in \mathfrak{A}$ в силу свойства 3), обозначается $\overline{PQ}$. Тогда из свойства 2) вытекает, что $\overline{PQ} + \overline{QR} = \overline{PR}$.

Замечание. Свойства 1)–2) из определения аффинного пространства означают, что на множестве $\mathfrak{A}$ задано *действие* абелевой группы векторов пространства $V$. Свойство 3) по определению означает, что это действие *свободно и транзитивно*. Множество, на котором задано свободное и транзитивное действие группы $G$ называется *главным однородным пространством* группы $G$. Таким образом, аффинное пространство $\mathfrak{A}$ — это главное однородное пространство абелевой группы $V$.

Пример 3.1.2.

1. Точки плоскости и (трёхмерного) пространства образуют аффинные пространства. Заметим, что точки «не помнят» начала координат: все точки на плоскости или в пространстве равноправны, пока мы не ввели там систему координат.

2. Рассмотрим совместную неоднородную систему линейных уравнений $Ax = b$, где $A$ — матрица, а $x$ и $b$ — столбцы. Пусть $\mathfrak{A}$ — множество решений $x$ этой системы, а $V$ — векторное пространство решений $y$ однородной системы $Ay = 0$. Тогда $(\mathfrak{A}, V)$ — аффинное пространство. Действительно, если $x \in \mathfrak{A}$ и $y \in V$, то $A(x+y) = Ax+Ay = b$ и поэтому $x + y \in \mathfrak{A}$.

3. Из всякого векторного пространства $V$ можно получить аффинное пространство, взяв в качестве $\mathfrak{A}$ множество векторов $V$; при этом сложение точек и векторов —

это просто сложение векторов в исходном пространстве $V$. Аффинное пространство, получаемое при помощи этой процедуры из $\mathbb{R}^n$, мы будем обозначать через $\mathbb{A}^n$.

Определение 3.1.3. *Аффинной системой координат* (или *репером*) в аффинном пространстве $(\mathfrak{A}, V)$ называется набор $(O, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$, состоящий из точки $O \in \mathfrak{A}$, называемой *началом координат*, и базиса $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ в векторном пространстве $V$.

*Координатами* точки $P \in \mathfrak{A}$ в системе координат $(O, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ называются координаты $x^1, \ldots, x^n$ вектора $\overline{OP}$ в базисе $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$, т.е.

$$\overline{OP} = x^1\boldsymbol{e}_1 + \ldots + x^n\boldsymbol{e}_n.$$

Определение 3.1.4. *Аффинным подпространством* в аффинном пространстве $(\mathfrak{A}, V)$ называется пара $(\mathfrak{B}, W)$, состоящая из подмножества $\mathfrak{B} \subset \mathfrak{A}$ и векторного подпространства $W \subset V$, такая, что $(\mathfrak{B}, W)$ является аффинным пространством относительно операций в $(\mathfrak{A}, V)$.

Эквивалентно, $(\mathfrak{B}, W)$ называется аффинным подпространством в $(\mathfrak{A}, V)$, если

1) для любых $P \in \mathfrak{B}$ и $\boldsymbol{w} \in W$ точка $P + \boldsymbol{w}$ лежит в $\mathfrak{B}$;
2) для любых $P, Q \in \mathfrak{B}$ вектор $\overline{PQ}$ лежит в $W$.

Пример 3.1.5. Аффинные подпространства $(\mathfrak{B}, W)$ в $\mathbb{A}^n$ можно задавать двумя способами:

а) репером, т.е. точкой $P \in \mathfrak{B}$ и базисом $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_k$ векторного пространства $W \subset \mathbb{R}^n$. При этом любая другая точка $P' \in \mathfrak{B}$ представляется в виде $P' = P + x^1\boldsymbol{f}_1 + \ldots + x^k\boldsymbol{f}_k$. Для такого способа задания используется обозначение
$$\mathfrak{B} = P + \langle \boldsymbol{f}_1, \ldots, \boldsymbol{f}_k \rangle.$$
б) как множество решений совместной неоднородной системы уравнений:
$$\mathfrak{B} = \{x \in \mathbb{R}^n \colon Ax = b\}.$$
При этом $W$ — это пространство решений однородной системы $Ax = 0$.

Первый способ обобщает параметрическое задание прямых и плоскостей. Второй способ обобщает задание прямых и плоскостей уравнениями.

Переход от второго способа задания подпространства к первому заключается в решении неоднородной системы $Ax = b$: точка $P$ — это частное решение, а набор $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_k$ — это фундаментальная система решений однородной системы $Ax = 0$.

Для перехода от первого способа задания подпространства ко второму необходимо задать линейную оболочку $\langle \boldsymbol{f}_1, \ldots, \boldsymbol{f}_k \rangle$ однородной системой $Ax = 0$; тогда столбец $b$ правых частей неоднородной системы $Ax = b$ получается при подстановке координат точки $P$ в уравнения системы $Ax = 0$.

## 3.2. Евклидовы и эрмитовы пространства, примеры. Неравенство Коши–Буняковского, неравенство треугольника

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

1) *билинейность*, т.е.

$$(\lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2, \boldsymbol{v}) = \lambda_1(\boldsymbol{u}_1, \boldsymbol{v}) + \lambda_2(\boldsymbol{u}_2, \boldsymbol{v}) \quad \text{и}$$
$$(\boldsymbol{u}, \mu_1 \boldsymbol{v}_1 + \mu_2 \boldsymbol{v}_2) = \mu_1(\boldsymbol{u}, \boldsymbol{v}_1) + \mu_2(\boldsymbol{u}, \boldsymbol{v}_2)$$

для любых $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{R}$ и $\boldsymbol{u}, \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{v}, \boldsymbol{v}_1, \boldsymbol{v}_2 \in V$;

2) *симметричность*: $(\boldsymbol{v}, \boldsymbol{u}) = (\boldsymbol{u}, \boldsymbol{v})$ для любых $\boldsymbol{u}, \boldsymbol{v} \in V$;

3) *положительная определённость*: $(\boldsymbol{v}, \boldsymbol{v}) \geqslant 0$ для любого $\boldsymbol{v} \in V$, причём $(\boldsymbol{v}, \boldsymbol{v}) = 0$ только при $\boldsymbol{v} = \boldsymbol{0}$.

Свойство билинейности выражает линейность скалярного произведения по каждому из аргументов. Ввиду наличия свойства симметричности, билинейность очевидно вытекает из линейности по любому из двух аргументов.

Определение 3.2.2. Линейное пространство над полем $\mathbb{C}$ называется *эрмитовым*, (или *унитарным*) если на парах его векторов определена функция $f\colon V \times V \to \mathbb{C}$ (обозначаемая $(a, b) := f(a, b)$ и называемая *скалярным произведением*), удовлетворяющая следующим свойствам:

1) *полуторалинейность*, т.е.

$$(\lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2, \boldsymbol{v}) = \overline{\lambda}_1(\boldsymbol{u}_1, \boldsymbol{v}) + \overline{\lambda}_2(\boldsymbol{u}_2, \boldsymbol{v}) \quad \text{и}$$
$$(\boldsymbol{u}, \mu_1 \boldsymbol{v}_1 + \mu_2 \boldsymbol{v}_2) = \mu_1(\boldsymbol{u}, \boldsymbol{v}_1) + \mu_2(\boldsymbol{u}, \boldsymbol{v}_2)$$

для любых $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{C}$ и $\boldsymbol{u}, \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{v}, \boldsymbol{v}_1, \boldsymbol{v}_2 \in V$;

2) *эрмитовость*: $(\boldsymbol{v}, \boldsymbol{u}) = \overline{(\boldsymbol{u}, \boldsymbol{v})}$ для любых $\boldsymbol{u}, \boldsymbol{v} \in V$; в частности, $(\boldsymbol{v}, \boldsymbol{v})$ вещественно для любого $\boldsymbol{v} \in V$.

3) *положительная определённость*: $(\boldsymbol{v}, \boldsymbol{v}) \geqslant 0$ для любого $\boldsymbol{v} \in V$, причём $(\boldsymbol{v}, \boldsymbol{v}) = 0$ только при $\boldsymbol{v} = \boldsymbol{0}$.

Свойство полуторалинейности выражает линейность скалярного произведения по второму аргументу и *антилинейность* по первому. Ввиду наличия свойства эрмитовости, полуторалинейность очевидно вытекает из линейности по второму аргументу.

Пример 3.2.3. 1. Скалярное произведение векторов $\boldsymbol{u} = (u^1, \ldots, u^n)$ и $\boldsymbol{v} = (v^1, \ldots, v^n)$ в пространстве $\mathbb{R}^n$ определяется по формуле

$$(\boldsymbol{u}, \boldsymbol{v}) := u^1 v^1 + u^2 v^2 + \ldots + u^n v^n. \tag{3.1}$$

Скалярное произведение векторов в $\mathbb{C}^n$ определяется по формуле

$$(\boldsymbol{u}, \boldsymbol{v}) := \overline{u^1} v^1 + \overline{u^2} v^2 + \ldots + \overline{u^n} v^n. \tag{3.2}$$

2. Скалярное произведение в пространстве $\mathrm{Mat}_{\mathbb{C}}(n, n)$ квадратных комплексных матриц размера $n$ задаётся с помощью формулы

$$(A, B) := \mathrm{Tr}(\overline{A}^t B) = \sum_{i,j=1}^{n} \overline{a_{ij}} b_{ij}.$$

При отождествлении пространства $\mathrm{Mat}_{\mathbb{C}}(n, n)$ с $\mathbb{C}^{n^2}$ это скалярное произведение переходит в стандартное скалярное произведение из предыдущего примера.

3. Рассмотрим пространство $C[a, b]$ вещественнозначных функций, непрерывных на отрезке $[a, b]$. Зададим скалярное произведение функций $f$ и $g$ по формуле

$$(f, g) := \int_a^b f(x) g(x) dx.$$

Тогда свойства 1) и 2) скалярного произведения очевидны, а 3) вытекает из того, что интеграл $\int_a^b f^2(x)dx$ от неотрицательной функции $f^2(x)$ неотрицателен и обращается в нуль только при $f(x) \equiv 0$.

Аналогично, скалярное произведение в пространстве комплекснозначных функций можно определить по формуле $(f, g) := \int_a^b \overline{f(x)} g(x)dx$.

Определение 3.2.4. Пусть $V$ — евклидово или эрмитово пространство. Для $\boldsymbol{v} \in V$ величина $\sqrt{(\boldsymbol{v}, \boldsymbol{v})}$ называется *длиной* вектора $\boldsymbol{v}$ и обозначается $|\boldsymbol{v}|$.

Векторы $\boldsymbol{u}, \boldsymbol{v} \in V$, скалярное произведение которых равно нулю, называются *перпендикулярными* или *ортогональными.* В этом случае пишут $\boldsymbol{u} \perp \boldsymbol{v}$.

Предложение 3.2.5. *Пусть $\boldsymbol{u}$ — ненулевой вектор евклидова или эрмитова пространства $V$. Тогда для любого вектора $\boldsymbol{v} \in V$ существует единственное разложение $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$, где вектор $\boldsymbol{v}_1$ коллинеарен вектору $\boldsymbol{u}$, а вектор $\boldsymbol{v}_2$ ортогонален $\boldsymbol{u}$.*

Доказательство. Сначала докажем единственность. Пусть $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$ — такое разложение. Тогда для $\lambda \in \mathbb{R}$ имеем $\boldsymbol{v}_1 = \lambda \boldsymbol{u}$, $\boldsymbol{v}_2 = \boldsymbol{v} - \lambda \boldsymbol{u}$. Условие $\boldsymbol{u} \perp \boldsymbol{v}_2$ влечёт

$$\mathbf{0} = (\boldsymbol{u}, \boldsymbol{v}_2) = (\boldsymbol{u}, \boldsymbol{v} - \lambda \boldsymbol{u}) = (\boldsymbol{u}, \boldsymbol{v}) - \lambda(\boldsymbol{u}, \boldsymbol{u}).$$

Отсюда $\lambda = (\boldsymbol{u}, \boldsymbol{v})/(\boldsymbol{u}, \boldsymbol{u})$ и

$$\boldsymbol{v}_1 = \frac{(\boldsymbol{u}, \boldsymbol{v})}{(\boldsymbol{u}, \boldsymbol{u})} \boldsymbol{u}. \tag{3.3}$$

Тем самым векторы $\boldsymbol{v}_1$ и $\boldsymbol{v}_2 = \boldsymbol{v} - \boldsymbol{v}_1$ определены однозначно. С другой стороны, определив $\boldsymbol{v}_1$ по этой формуле, мы получим $\boldsymbol{v}_2 = (\boldsymbol{v} - \boldsymbol{v}_1) \perp \boldsymbol{u}$. □

Определение 3.2.6. Вектор (3.3) называется *ортогональной проекцией* вектора $\boldsymbol{v}$ на направление вектора $\boldsymbol{u}$ и обозначается $\operatorname{pr}_{\boldsymbol{u}} \boldsymbol{v}$, а вектор $\boldsymbol{v} - \operatorname{pr}_{\boldsymbol{u}} \boldsymbol{v}$ называется *ортогональной составляющей* вектора $\boldsymbol{v}$ относительно $\boldsymbol{u}$ и обозначается $\operatorname{ort}_{\boldsymbol{u}} \boldsymbol{v}$.

Длина ортогональной проекции вычисляется по формуле

$$|\operatorname{pr}_{\boldsymbol{u}} \boldsymbol{v}| = \sqrt{(\operatorname{pr}_{\boldsymbol{u}} \boldsymbol{v}, \operatorname{pr}_{\boldsymbol{u}} \boldsymbol{v})} = \sqrt{\Bigl(\frac{(\boldsymbol{u}, \boldsymbol{v})}{(\boldsymbol{u}, \boldsymbol{u})} \boldsymbol{u}, \frac{(\boldsymbol{u}, \boldsymbol{v})}{(\boldsymbol{u}, \boldsymbol{u})} \boldsymbol{u}\Bigr)} = \sqrt{\frac{\overline{(\boldsymbol{u}, \boldsymbol{v})}(\boldsymbol{u}, \boldsymbol{v})}{(\boldsymbol{u}, \boldsymbol{u})}} = \frac{|(\boldsymbol{u}, \boldsymbol{v})|}{|\boldsymbol{u}|}.$$

Теорема 3.2.7 (неравенство Коши–Буняковского). *Для любых двух векторов $\boldsymbol{u}, \boldsymbol{v}$ евклидова или эрмитова пространства имеет место неравенство*

$$|(\boldsymbol{u}, \boldsymbol{v})| \leqslant |\boldsymbol{u}| \cdot |\boldsymbol{v}|,$$

*причем равенство имеет место только в случае, когда векторы $\boldsymbol{u}, \boldsymbol{v}$ коллинеарны.*

Доказательство. Если $\boldsymbol{u} = \mathbf{0}$, утверждение очевидно. Пусть $\boldsymbol{u} \neq \mathbf{0}$. Запишем $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$, где $\boldsymbol{v}_1 = \operatorname{pr}_{\boldsymbol{u}} \boldsymbol{v}$ и $\boldsymbol{v}_2 = \operatorname{ort}_{\boldsymbol{u}} \boldsymbol{v}$. Тогда $(\boldsymbol{v}_1, \boldsymbol{v}_2) = 0$, и мы имеем

$$|\boldsymbol{v}|^2 = (\boldsymbol{v}, \boldsymbol{v}) = (\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{v}_1 + \boldsymbol{v}_2) = (\boldsymbol{v}_1, \boldsymbol{v}_1) + (\boldsymbol{v}_1, \boldsymbol{v}_2) + (\boldsymbol{v}_2, \boldsymbol{v}_1) + (\boldsymbol{v}_2, \boldsymbol{v}_2) = |\boldsymbol{v}_1|^2 + |\boldsymbol{v}_2|^2.$$

Отсюда $|\boldsymbol{v}_1| \leqslant |\boldsymbol{v}|$, причём равенство достигается только при $\boldsymbol{v}_2 = \mathbf{0}$, т.е. когда вектор $\boldsymbol{v}$ коллинеарен вектору $\boldsymbol{u}$. Осталось заметить, что $|\boldsymbol{v}_1| = |\operatorname{pr}_{\boldsymbol{u}} \boldsymbol{v}| = \frac{|(\boldsymbol{u}, \boldsymbol{v})|}{|\boldsymbol{u}|}$, так что неравенство $|\boldsymbol{v}_1| \leqslant |\boldsymbol{v}|$ эквивалентно требуемому. □

Определение 3.2.8. *Углом* между двумя ненулевыми векторами $\boldsymbol{u}, \boldsymbol{v}$ евклидова пространства называется величина

$$\angle(\boldsymbol{u}, \boldsymbol{v}) := \arccos \frac{(\boldsymbol{u}, \boldsymbol{v})}{|\boldsymbol{u}|\,|\boldsymbol{v}|} \in [0, \pi].$$

Неравенство Коши–Буняковского гарантирует, что угол между ненулевыми векторами всегда определен.

Следствие 3.2.9 (неравенство треугольника). *Для любых двух векторов $\boldsymbol{u}, \boldsymbol{v}$ евклидова или эрмитова пространства выполнено неравенство*

$$|\boldsymbol{u} + \boldsymbol{v}| \leqslant |\boldsymbol{u}| + |\boldsymbol{v}|.$$

Доказательство. В обеих частях неравенства стоят неотрицательные величины, поэтому при возведении в квадрат получается равносильное неравенство

$$(\boldsymbol{u} + \boldsymbol{v}, \boldsymbol{u} + \boldsymbol{v}) \leqslant (\boldsymbol{u}, \boldsymbol{u}) + (\boldsymbol{v}, \boldsymbol{v}) + 2|\boldsymbol{u}|\,|\boldsymbol{v}|.$$

После раскрытия скобок в левой части и сокращения подобных членов мы получаем следующее неравенство:

$$(\boldsymbol{u}, \boldsymbol{v}) + (\boldsymbol{v}, \boldsymbol{u}) \leqslant 2|\boldsymbol{u}|\,|\boldsymbol{v}|,$$

которое следует из неравенства Коши–Буняковского. □

## 3.3. Ортогональные системы векторов, ортонормированные базисы. Ортогонализация Грама-Шмидта

Предложение 3.3.1. *Пусть $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ — набор попарно ортогональных ненулевых векторов евклидова или эрмитова пространства. Тогда эти векторы линейно независимы.*

Доказательство. Пусть некоторая линейная комбинация данных векторов равна нулю:

$$\sum_{i=1}^{k} \lambda_i \boldsymbol{v}_i = \boldsymbol{0}.$$

Умножим обе части этого равенства скалярно на $\boldsymbol{v}_j$ и воспользуемся линейностью скалярного произведения по второму аргументу:

$$0 = \Big(\boldsymbol{v}_j, \sum_{i=1}^{k} \lambda_i \boldsymbol{v}_i\Big) = \sum_{i=1}^{k} \lambda_i (\boldsymbol{v}_j, \boldsymbol{v}_i) = \lambda_j (\boldsymbol{v}_j, \boldsymbol{v}_j),$$

так как по условию остальные слагаемые в этой сумме равны нулю. Поскольку по условию $\boldsymbol{v}_j \neq \boldsymbol{0}$, из положительной определенности скалярного произведения следует, что $(\boldsymbol{v}_j, \boldsymbol{v}_j) \neq 0$, а значит, $\lambda_j = 0$. Это выполнено для любого $j = 1, \ldots, k$, следовательно, линейная комбинация $\sum_{i=1}^{k} \lambda_i \boldsymbol{v}_i$ тривиальна. □

Определение 3.3.2. Базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ евклидова или эрмитова пространства называется *ортогональным*, если его векторы попарно ортогональны. Если при этом длина каждого вектора равна 1, то базис называется *ортонормированным.*

Теорема 3.3.3. *Пусть $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ — набор линейно независимых векторов пространства $V$. Тогда существует такой набор попарно ортогональных векторов $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k$, что для каждого $i = 1, \ldots, k$ линейная оболочка $\langle \boldsymbol{b}_1, \ldots, \boldsymbol{b}_i \rangle$ совпадает с $\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_i \rangle$.*

Доказательство. При $k = 1$ утверждение очевидно: можно взять $\boldsymbol{b}_1 = \boldsymbol{a}_1$. Предположим, что утверждение верно для наборов из $i$ векторов, и докажем его для наборов из $i+1$ вектора. Пусть $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_i$ — ортогональный набор, построенный по набору $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_i$. Мы хотим, чтобы для нового вектора $\boldsymbol{b}_{i+1}$ линейная оболочка $\langle \boldsymbol{b}_1, \ldots, \boldsymbol{b}_i, \boldsymbol{b}_{i+1}\rangle$ совпадала с $\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_i, \boldsymbol{a}_{i+1}\rangle = \langle \boldsymbol{b}_1, \ldots, \boldsymbol{b}_i, \boldsymbol{a}_{i+1}\rangle$, и поэтому будем искать $\boldsymbol{b}_{i+1}$ в виде

$$\boldsymbol{b}_{i+1} = \boldsymbol{a}_{i+1} + \lambda_1 \boldsymbol{b}_1 + \ldots + \lambda_i \boldsymbol{b}_i.$$

Коэффициенты $\lambda_1, \ldots, \lambda_i$ будем подбирать так, чтобы вектор $\boldsymbol{b}_{i+1}$ был ортогонален всем предыдущим векторам $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_i$. Умножив скалярно предыдущее равенство слева на $\boldsymbol{b}_j$ и использовав то, что $(\boldsymbol{b}_j, \boldsymbol{b}_\ell) = 0$ при $j \neq \ell$, получаем

$$0 = (\boldsymbol{b}_j, \boldsymbol{b}_{i+1}) = (\boldsymbol{b}_j, \boldsymbol{a}_{i+1}) + \lambda_j(\boldsymbol{b}_j, \boldsymbol{b}_j),$$

откуда $\lambda_j = -\frac{(\boldsymbol{b}_j, \boldsymbol{a}_{i+1})}{(\boldsymbol{b}_j, \boldsymbol{b}_j)}$ для $j = 1, \ldots, i$. Окончательно для вектора $\boldsymbol{b}_{i+1}$ получаем

$$\begin{aligned}\boldsymbol{b}_{i+1} &= \boldsymbol{a}_{i+1} - \frac{(\boldsymbol{b}_1, \boldsymbol{a}_{i+1})}{(\boldsymbol{b}_1, \boldsymbol{b}_1)}\boldsymbol{b}_1 - \frac{(\boldsymbol{b}_2, \boldsymbol{a}_{i+1})}{(\boldsymbol{b}_2, \boldsymbol{b}_2)}\boldsymbol{b}_2 - \ldots - \frac{(\boldsymbol{b}_i, \boldsymbol{a}_{i+1})}{(\boldsymbol{b}_i, \boldsymbol{b}_i)}\boldsymbol{b}_i = \\ &= \boldsymbol{a}_{i+1} - \mathrm{pr}_{\boldsymbol{b}_1} \boldsymbol{a}_{i+1} - \mathrm{pr}_{\boldsymbol{b}_2} \boldsymbol{a}_{i+1} - \ldots - \mathrm{pr}_{\boldsymbol{b}_i} \boldsymbol{a}_{i+1}.\end{aligned} \tag{3.4}$$

При этом $\boldsymbol{b}_{i+1} \neq \boldsymbol{0}$ (так как $\boldsymbol{b}_1 \ldots, \boldsymbol{b}_i, \boldsymbol{a}_{i+1}$ линейно независимы), $\boldsymbol{b}_{i+1}$ ортогонален векторам $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_i$, а $\langle \boldsymbol{b}_1, \ldots, \boldsymbol{b}_i, \boldsymbol{b}_{i+1}\rangle = \langle \boldsymbol{b}_1, \ldots, \boldsymbol{b}_i, \boldsymbol{a}_{i+1}\rangle = \langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_i, \boldsymbol{a}_{i+1}\rangle$. □

Индуктивная процедура перехода от набора $\boldsymbol{a}_1 \ldots, \boldsymbol{a}_k$ к ортогональному набору $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k$ называется *процессом ортогонализации Грама–Шмидта*. Условие $\langle \boldsymbol{b}_1, \ldots, \boldsymbol{b}_i\rangle = \langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_i\rangle$ при $i = 1, \ldots, k$ означает, что матрица перехода от $\boldsymbol{a}_1 \ldots, \boldsymbol{a}_k$ к $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k$ является *верхнетреугольной.*

Следствие 3.3.4. *В евклидовом или эрмитовом пространстве $V$ существуют ортонормированные базисы.*

Доказательство. Действительно, возьмём произвольный базис пространства $V$ и применим к нему ортогонализацию Грама–Шмидта. В результате получим ортогональный базис $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$. Тогда базис, состоящий из векторов $\frac{\boldsymbol{b}_1}{|\boldsymbol{b}_1|}, \ldots, \frac{\boldsymbol{b}_n}{|\boldsymbol{b}_n|}$ будет ортонормированным. □

Предложение 3.3.5. *Пусть векторы $\boldsymbol{u}$ и $\boldsymbol{v}$ имеют координаты $u^1, \ldots, u^n$ и $v^1, \ldots, v^n$ в некотором ортонормированном базисе пространства $V$. Тогда их скалярное произведение вычисляется по формуле*

$$(\boldsymbol{u}, \boldsymbol{v}) = u^1 v^1 + u^2 v^2 + \ldots + u^n v^n.$$

Доказательство. Пусть $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ — ортонормированный базис. Тогда $(\boldsymbol{u}, \boldsymbol{v}) = (u^i \boldsymbol{e}_i, v^j \boldsymbol{e}_j) = u^i v^j (\boldsymbol{e}_i, \boldsymbol{e}_j) = u^i v^j \delta_{ij} = u^1 v^1 + u^2 v^2 + \ldots + u^n v^n$. □

## 3.4. Ортогональные и унитарные матрицы. $QR$-разложение

Определение 3.4.1. Матрица перехода от одного ортонормированного базиса евклидова (соответственно, эрмитова) пространства к другому ортонормированному базису называется *ортогональной* (соответственно, *унитарной*).

Предложение 3.4.2. *Следующие условия эквивалентны:*

а) *матрица $C$ ортогональна (соответственно, унитарна);*

б) *$C^t C = E$ (соответственно, $\overline{C}^t C = E$);*

в) *столбцы матрицы $C$ образуют ортонормированный базис пространства $\mathbb{R}^n$ (соответственно, пространства $\mathbb{C}^n$);*

г) $CC^t = E$ *(соответственно,* $\overline{C}C^t = E$*);*

д) *строки матрицы $C$ образуют ортонормированный базис пространства $\mathbb{R}^n$ (соответственно, пространства $\mathbb{C}^n$).*

Доказательство. Условия б) и г) эквивалентны, так как каждое из них эквивалентно равенству $C^t = C^{-1}$ (соответственно, $\overline{C}^t = C^{-1}$). Эквивалентности б)$\Leftrightarrow$в) и г)$\Leftrightarrow$д) вытекают из правила умножения матриц и формул (3.1) и (3.2) для скалярного произведения в $\mathbb{R}^n$ и $\mathbb{C}^n$.

Докажем импликацию а)$\Rightarrow$б). Пусть $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ и $\boldsymbol{e}_{1'}, \ldots, \boldsymbol{e}_{n'}$ — два ортонормированных базиса в эрмитовом пространстве и $C = (c^i_{i'})$ — матрица перехода, т.е. $\boldsymbol{e}_{i'} = c^i_{i'}\boldsymbol{e}_i$. Тогда $(\boldsymbol{e}_i, \boldsymbol{e}_j) = \delta_{ij}$ и $(\boldsymbol{e}_{i'}, \boldsymbol{e}_{j'}) = \delta_{i'j'}$, откуда

$$\delta_{i'j'} = (\boldsymbol{e}_{i'}, \boldsymbol{e}_{j'}) = (c^i_{i'}\boldsymbol{e}_i, c^j_{j'}\boldsymbol{e}_j) = \overline{c}^i_{i'}c^j_{j'}(\boldsymbol{e}_i, \boldsymbol{e}_j) = \overline{c}^i_{i'}c^j_{j'}\delta_{ij} = \overline{c}^i_{i'}\delta_{ij}c^j_{j'}. \tag{3.5}$$

Согласно правилу умножения матриц, это эквивалентно матричному соотношению $E = \overline{C}^t EC$ или $\overline{C}^t C = E$. Случай евклидова пространства рассматривается аналогично.

Осталось доказать импликацию б)$\Rightarrow$а). Пусть имеет место тождество $\overline{C}^t C = E$ или, в обозначениях Эйнштейна, $\overline{c}^i_{i'}\delta_{ij}c^j_{j'} = \delta_{i'j'}$. Возьмём произвольный ортонормированный базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Из соотношения $\overline{C}^t C = E$ вытекает, что матрица $C$ невырождена, и поэтому можно рассмотреть новый базис $\boldsymbol{e}_{1'}, \ldots, \boldsymbol{e}_{n'}$, где $\boldsymbol{e}_{i'} = c^i_{i'}\boldsymbol{e}_i$. Тогда аналогично выкладке (3.5) мы получаем $(\boldsymbol{e}_{i'}, \boldsymbol{e}_{j'}) = \overline{c}^i_{i'}\delta_{ij}c^j_{j'} = \delta_{i'j'}$, т.е. базис $\boldsymbol{e}_{1'}, \ldots, \boldsymbol{e}_{n'}$ также ортонормирован, и $C$ — матрица перехода от одного ортонормированного базиса к другому. □

Теорема 3.4.3 ($QR$-разложение). *Для любой невырожденной вещественной (соответственно, комплексной) матрицы $A$ имеет место разложение*

$$A = QR,$$

*где $Q$ — ортогональная (соответственно, унитарная) матрица, а $R$ — верхнетреугольная матрица с положительными числами на диагонали.*

Доказательство. Пусть $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ — базис пространства $\mathbb{R}^n$ (или $\mathbb{C}^n$), состоящий из столбцов матрицы $A$. Применив к нему ортогонализацию Грама–Шмидта, получим ортогональный базис $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$. Пусть $C$ — матрица перехода, т.е.

$$(\boldsymbol{b}_1 \, \ldots \, \boldsymbol{b}_n) = (\boldsymbol{a}_1 \, \ldots \, \boldsymbol{a}_n)\, C$$

или $B = AC$, где $B$ — матрица, столбцы которой суть $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$. При этом матрица $C$ — верхнетреугольная с единицами на диагонали (это следует из соотношений (3.4)). Если мы видоизменим процедуру ортогонализации Грама–Шмидта и на каждом шаге вместо вектора $\boldsymbol{b}_i$ будем брать $\boldsymbol{b}'_i = \frac{\boldsymbol{b}_i}{|\boldsymbol{b}_i|}$, то полученная матрица $B' = (\boldsymbol{b}'_1, \ldots, \boldsymbol{b}'_n)$ будет ортогональной. При этом мы будем иметь $B' = AC'$, где $C'$ — верхнетреугольная матрица, на диагонали которой стоят положительные числа $\frac{1}{|\boldsymbol{b}_i|}$. Тогда положив $Q = B'$ и $R = C'^{-1}$, мы получим требуемое разложение $A = QR$. □

Замечание. $QR$-разложение имеет место также и для вырожденных матриц $A$ (при этом на диагонали $R$ могут стоять нули), а также для прямоугольных матриц $A$ произвольного размера (задача).

### 3.5. Ортогональное дополнение. Проекция и ортогональная составляющая. Угол между вектором и подпространством

Определение 3.5.1. Пусть $W \subset V$ — подпространство евклидова или эрмитова пространства. *Ортогональным дополнением* к $W$ называется множество $W^\perp$, состоящее из векторов, ортогональных всем векторам из $W$, т.е.

$$W^\perp = \{\boldsymbol{v} \in V : (\boldsymbol{v}, \boldsymbol{w}) = 0 \quad \text{для всех } \boldsymbol{w} \in W\}.$$

Легко видеть, что ортогональное дополнение $W^\perp$ является подпространством.

Предложение 3.5.2. *Для любого подпространства $W \subset V$ имеет место разложение $V = W \oplus W^\perp$.*

Доказательство. Пусть $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ — базис в $W$, дополним его до базиса всего пространства $V$ векторами $\boldsymbol{a}_{k+1}, \ldots, \boldsymbol{a}_n$. Применив ортогонализацию Грама–Шмидта, получим ортогональный базис $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k, \boldsymbol{b}_{k+1}, \ldots, \boldsymbol{b}_n$ в $V$, причём его первые $k$ векторов будут базисом в $W$, так как $\langle \boldsymbol{b}_1, \ldots, \boldsymbol{b}_k \rangle = \langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_k \rangle = W$. В то же время $\boldsymbol{b}_{k+1}, \ldots, \boldsymbol{b}_n$ лежат в $W^\perp$ по определению ортогонального дополнения. Итак, для любого вектора $\boldsymbol{v} \in V$ мы имеем разложение по базису

$$\boldsymbol{v} = \underbrace{\lambda_1 \boldsymbol{b}_1 + \ldots + \lambda_k \boldsymbol{b}_k}_{\in W} + \underbrace{\lambda_{k+1} \boldsymbol{b}_{k+1} + \ldots + \lambda_n \boldsymbol{b}_n}_{\in W^\perp},$$

т.е. $V = W + W^\perp$.

Осталось доказать, что эта сумма прямая. Пусть $\boldsymbol{v} \in W \cap W^\perp$. Так как $\boldsymbol{v} \in W^\perp$, мы имеем $(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{0}$ для всех $\boldsymbol{w} \in W$. Так как $\boldsymbol{v} \in W$, в качестве $\boldsymbol{w}$ мы можем взять сам вектор $\boldsymbol{v}$. Тогда $(\boldsymbol{v}, \boldsymbol{v}) = 0$, т.е. $\boldsymbol{v} = \boldsymbol{0}$ и сумма — прямая. □

Определение 3.5.3. Пусть $W \subset V$ — подпространство евклидова или эрмитова пространства. Для произвольного вектора $\boldsymbol{v} \in V$ запишем разложение $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$, где $\boldsymbol{v}_1 \in W$, а $\boldsymbol{v}_2 \in W^\perp$. Тогда вектор $\boldsymbol{v}_1$ называется *ортогональной проекцией* вектора $\boldsymbol{v}$ на подпространство $W$ и обозначается $\mathrm{pr}_W \boldsymbol{v}$, а вектор $\boldsymbol{v}_2 = \boldsymbol{v} - \mathrm{pr}_W \boldsymbol{v}$ называется *ортогональной составляющей* вектора $\boldsymbol{v}$ относительно подпространства $W$ и обозначается $\mathrm{ort}_W \boldsymbol{v}$.

Ясно, что $\mathrm{ort}_W \boldsymbol{v} = \mathrm{pr}_{W^\perp} \boldsymbol{v}$.

Предложение 3.5.4. *Пусть подпространство $W \subset V$ задано как линейная оболочка системы векторов: $W = \langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_k \rangle$. Тогда проекция вектора $\boldsymbol{v} \in V$ на $W$ есть линейная комбинация*

$$\mathrm{pr}_W \boldsymbol{v} = \lambda_1 \boldsymbol{a}_1 + \ldots + \lambda_k \boldsymbol{a}_k,$$

*коэффициенты которой находятся из системы линейных уравнений*

$$\begin{cases} (\boldsymbol{a}_1, \boldsymbol{a}_1)\lambda_1 + (\boldsymbol{a}_1, \boldsymbol{a}_2)\lambda_2 + \ldots + (\boldsymbol{a}_1, \boldsymbol{a}_k)\lambda_k = (\boldsymbol{a}_1, \boldsymbol{v}), \\ (\boldsymbol{a}_2, \boldsymbol{a}_1)\lambda_1 + (\boldsymbol{a}_2, \boldsymbol{a}_2)\lambda_2 + \ldots + (\boldsymbol{a}_2, \boldsymbol{a}_k)\lambda_k = (\boldsymbol{a}_2, \boldsymbol{v}), \\ \qquad \qquad \qquad \qquad \ldots\ldots \\ (\boldsymbol{a}_k, \boldsymbol{a}_1)\lambda_1 + (\boldsymbol{a}_k, \boldsymbol{a}_2)\lambda_2 + \ldots + (\boldsymbol{a}_k, \boldsymbol{a}_k)\lambda_k = (\boldsymbol{a}_k, \boldsymbol{v}). \end{cases}$$

Доказательство. Запишем $\boldsymbol{v} = \mathrm{pr}_W \boldsymbol{v} + \mathrm{ort}_W \boldsymbol{v}$. Тогда вектор $\mathrm{ort}_W \boldsymbol{v} = \boldsymbol{v} - \lambda_1 \boldsymbol{a}_1 - \ldots - \lambda_k \boldsymbol{a}_k$ ортогонален каждому из векторов $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$. Взяв скалярное произведение $\boldsymbol{a}_i$ с $\mathrm{ort}_W \boldsymbol{v}$, мы получаем

$$(\boldsymbol{a}_i, \boldsymbol{v} - \lambda_1 \boldsymbol{a}_1 - \ldots - \lambda_k \boldsymbol{a}_k) = 0,$$

что эквивалентно $i$-му уравнению системы. $\square$

Определение 3.5.5. Пусть $V$ — евклидово пространство. *Углом* между вектором $\boldsymbol{v} \in V$ и подпространством $W \subset V$ называется наименьший из углов между $\boldsymbol{v}$ и произвольным вектором $\boldsymbol{w} \in W$:

$$\angle(\boldsymbol{v}, W) := \min_{\boldsymbol{w} \in W} \angle(\boldsymbol{v}, \boldsymbol{w}).$$

Формально угол $\angle(\boldsymbol{v}, W)$ нужно было определить как *точную нижнюю грань* углов между $\boldsymbol{v}$ и $\boldsymbol{w} \in W$, но легко видеть, что множество таких углов замкнуто и поэтому точная нижняя грань достигается на некотором векторе $\boldsymbol{w} \in W$. На самом деле этот вектор есть $\mathrm{pr}_W \boldsymbol{v}$, как показано в следующем утверждении.

Предложение 3.5.6. *Угол между вектором и подпространством равен углу между вектором и его проекцией на это подпространство:*

$$\angle(\boldsymbol{v}, W) = \angle(\boldsymbol{v}, \mathrm{pr}_W \boldsymbol{v}).$$

Доказательство. Пусть $\boldsymbol{w} \in W$ — произвольный вектор. Обозначим $\alpha = \angle(\boldsymbol{v}, \mathrm{pr}_W \boldsymbol{v})$, $\beta = \angle(\boldsymbol{v}, \boldsymbol{w})$ и $\boldsymbol{v}_1 = \mathrm{pr}_W \boldsymbol{v}$. Необходимо показать, что $\alpha \leqslant \beta$. Так как $0 \leqslant \alpha, \beta \leqslant \pi$, неравенство $\alpha \leqslant \beta$ эквивалентно неравенству $\cos\alpha \geqslant \cos\beta$, т.е.

$$\frac{(\boldsymbol{v}, \boldsymbol{v}_1)}{|\boldsymbol{v}|\,|\boldsymbol{v}_1|} \geqslant \frac{(\boldsymbol{v}, \boldsymbol{w})}{|\boldsymbol{v}|\,|\boldsymbol{w}|}. \tag{3.6}$$

Запишем $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$, где $\boldsymbol{v}_2 = \mathrm{ort}_W \boldsymbol{v} \in W^\perp$. Тогда $(\boldsymbol{v}, \boldsymbol{v}_1) = (\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{v}_1) = |\boldsymbol{v}_1|^2$ и $(\boldsymbol{v}, \boldsymbol{w}) = (\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}) = (\boldsymbol{v}_1, \boldsymbol{w})$. Подставив это в (3.6), получим $|\boldsymbol{v}_1| \geqslant \frac{(\boldsymbol{v}_1, \boldsymbol{w})}{|\boldsymbol{w}|}$, что вытекает из неравенства Коши–Буняковского. $\square$

## 3.6. Аффинные евклидовы пространства. Расстояние от точки до подпространства. Расстояние между подпространствами

Определение 3.6.1. Аффинное пространство $(\mathfrak{A}, V)$ называется *евклидовым*, если линейное пространство $V$ является евклидовым.

*Расстоянием* между точками $P$ и $Q$ аффинного евклидова пространства $(\mathfrak{A}, V)$ называется длина вектора $\overline{PQ}$:

$$d(P, Q) := |\overline{PQ}|.$$

*Расстоянием* между точкой $P$ и аффинным подпространством $(\mathfrak{B}, W)$ называется наименьшее из расстояний между $P$ и произвольной точкой $Q \in \mathfrak{B}$:

$$d(P, \mathfrak{B}) := \min_{Q \in \mathfrak{B}} d(P, Q).$$

*Расстоянием* между аффинными подпространствами $(\mathfrak{C}, U)$ и $(\mathfrak{B}, W)$ называется наименьшее из расстояний между точками $P \in \mathfrak{C}$ и $Q \in \mathfrak{B}$:

$$d(\mathfrak{C}, \mathfrak{B}) := \min_{P \in \mathfrak{C},\, Q \in \mathfrak{B}} d(P, Q).$$

Следующее утверждение обобщает утверждения о расстоянии между точкой и прямой или плоскостью из аналитической геометрии.

Теорема 3.6.2. *Расстояние между точкой $P$ и аффинным подпространством $(\mathfrak{B}, W)$ равно длине ортогональной составляющей вектора $\overline{PQ}$, соединяющего $P$ с произвольной точкой $Q \in \mathfrak{B}$, относительно пространства $W$:*

$$d(P,\mathfrak{B}) = |\operatorname{ort}_W \overline{PQ}| \qquad \text{для любой точки } Q \in \mathfrak{B}.$$

Доказательство. Вначале мы докажем, что ортогональная составляющая $\operatorname{ort}_W \overline{PQ}$ не зависит от выбора точки $Q \in \mathfrak{B}$. Пусть $Q' \in \mathfrak{B}$ — другая точка. Мы имеем $\overline{PQ} = \operatorname{pr}_W \overline{PQ} + \operatorname{ort}_W \overline{PQ}$ и $\overline{PQ'} = \operatorname{pr}_W \overline{PQ'} + \operatorname{ort}_W \overline{PQ'}$. С другой стороны, $\overline{PQ'} = \overline{PQ} + \overline{QQ'}$, где $\overline{QQ'} \in W$. Тогда

$$\overline{PQ'} = \underbrace{\operatorname{pr}_W \overline{PQ'}}_{\in W} + \underbrace{\operatorname{ort}_W \overline{PQ'}}_{\in W^\perp} = \overline{QQ'} + \overline{PQ} = \underbrace{\overline{QQ'} + \operatorname{pr}_W \overline{PQ}}_{\in W} + \underbrace{\operatorname{ort}_W \overline{PQ}}_{\in W^\perp}.$$

Отсюда в силу единственности разложения вектора в прямой сумме $V = W \oplus W^\perp$ получаем $\operatorname{ort}_W \overline{PQ'} = \operatorname{ort}_W \overline{PQ}$, что и требовалось.

Теперь докажем, что для любой точки $Q \in \mathfrak{B}$ мы имеем $|\overline{PQ}| \geqslant |\operatorname{ort}_W \overline{PQ}|$. Действительно,

$$|\overline{PQ}|^2 = (\overline{PQ}, \overline{PQ}) = (\operatorname{pr}_W \overline{PQ} + \operatorname{ort}_W \overline{PQ}, \operatorname{pr}_W \overline{PQ} + \operatorname{ort}_W \overline{PQ}) =$$
$$= (\operatorname{pr}_W \overline{PQ}, \operatorname{pr}_W \overline{PQ}) + (\operatorname{ort}_W \overline{PQ}, \operatorname{ort}_W \overline{PQ}) = |\operatorname{pr}_W \overline{PQ}|^2 + |\operatorname{ort}_W \overline{PQ}|^2 \geqslant |\operatorname{ort}_W \overline{PQ}|^2,$$

где в предпоследнем равенстве мы воспользовались тем, что $(\operatorname{pr}_W \overline{PQ}, \operatorname{ort}_W \overline{PQ}) = 0$. Следовательно, $d(P,\mathfrak{B}) = \min_{Q\in\mathfrak{B}} |\overline{PQ}| \geqslant |\operatorname{ort}_W \overline{PQ}|$.

Осталось доказать, что значение $|\operatorname{ort}_W \overline{PQ}|$ достигается, т.е. $|\overline{PQ'}| = |\operatorname{ort}_W \overline{PQ}|$ для некоторой точки $Q' \in \mathfrak{B}$. Для этого возьмём в качестве $Q'$ точку $P + \operatorname{ort}_W \overline{PQ}$. Тогда, по определению расстояния, $|\overline{PQ'}| = |\operatorname{ort}_W \overline{PQ}|$. С другой стороны,

$$Q' = P + \operatorname{ort}_W \overline{PQ} = P + \overline{PQ} - \operatorname{pr}_W \overline{PQ} = Q - \operatorname{pr}_W \overline{PQ},$$

где $Q \in \mathfrak{B}$ и $\operatorname{pr}_W \overline{PQ} \in W$, т.е. $Q' \in \mathfrak{B}$. □

В аналитической геометрии расстояние между скрещивающимися прямыми в пространстве вычислялось как длина их общего перпендикуляра. Аналогичным образом вычисляется расстояние между аффинными подпространствами в общем случае:

Теорема 3.6.3. *Расстояние между двумя аффинными подпространствами $(\mathfrak{C}, U)$ и $(\mathfrak{B}, W)$ равно длине ортогональной составляющей вектора $\overline{PQ}$, соединяющего произвольную точку $P \in \mathfrak{C}$ с произвольной точкой $Q \in \mathfrak{B}$, относительно пространства $U + W$:*

$$d(\mathfrak{C},\mathfrak{B}) = |\operatorname{ort}_{U+W} \overline{PQ}| \qquad \text{для любых точек } P \in \mathfrak{C}, Q \in \mathfrak{B}.$$

Доказательство. Доказательство аналогично предыдущему. Вначале докажем, что ортогональная составляющая $\operatorname{ort}_{U+W} \overline{PQ}$ не зависит от выбора точек $P \in \mathfrak{C}$, $Q \in \mathfrak{B}$. Пусть $P' \in \mathfrak{C}$, $Q' \in \mathfrak{B}$ — другие точки. Мы имеем $\overline{PQ} = \operatorname{pr}_{U+W} \overline{PQ} + \operatorname{ort}_{U+W} \overline{PQ}$ и $\overline{P'Q'} = \operatorname{pr}_{U+W} \overline{P'Q'} + \operatorname{ort}_{U+W} \overline{P'Q'}$. С другой стороны, $\overline{P'Q'} = \overline{P'P} + \overline{PQ} + \overline{QQ'}$, где $\overline{P'P} \in U$ и $\overline{QQ'} \in W$. Тогда

$$\overline{P'Q'} = \underbrace{\operatorname{pr}_{U+W} \overline{P'Q'}}_{\in U+W} + \underbrace{\operatorname{ort}_{U+W} \overline{P'Q'}}_{\in (U+W)^\perp} = \underbrace{\overline{P'P} + \overline{QQ'} + \operatorname{pr}_{U+W} \overline{PQ}}_{\in U+W} + \underbrace{\operatorname{ort}_{U+W} \overline{PQ}}_{\in (U+W)^\perp}.$$

Отсюда в силу единственности разложения вектора в прямой сумме получаем $\operatorname{ort}_{U+W} \overline{P'Q'} = \operatorname{ort}_{U+W} \overline{PQ}$, что и требовалось.

Так же, как в предыдущем предложении, мы доказываем, что для $P \in \mathfrak{C}$ и $Q \in \mathfrak{B}$ мы имеем $|\overline{PQ}| \geqslant |\operatorname{ort}_{U+W} \overline{PQ}|$. Следовательно,

$$d(\mathfrak{C}, \mathfrak{B}) = \min_{P \in \mathfrak{C},\, Q \in \mathfrak{B}} |\overline{PQ}| \geqslant |\operatorname{ort}_{U+W} \overline{PQ}|.$$

Осталось доказать, что значение $|\operatorname{ort}_{U+W} \overline{PQ}|$ достигается в некоторой паре точек, т.е. $|\overline{P'Q'}| = |\operatorname{ort}_{U+W} \overline{PQ}|$ для некоторых точек $P' \in \mathfrak{C}$, $Q' \in \mathfrak{B}$. Запишем вектор $\operatorname{pr}_{U+W} \overline{PQ} \in U + W$ в виде суммы:

$$\operatorname{pr}_{U+W} \overline{PQ} = \boldsymbol{u} + \boldsymbol{w}, \tag{3.7}$$

где $\boldsymbol{u} \in U$ и $\boldsymbol{w} \in W$. Теперь возьмём $P' = P + \boldsymbol{u}$, тогда очевидно $P' \in (\mathfrak{C}, U)$. Далее, возьмём $Q' = P' + \operatorname{ort}_{U+W} \overline{PQ}$. Тогда, по определению расстояния, $|\overline{P'Q'}| = |\operatorname{ort}_{U+W} \overline{PQ}|$. С другой стороны,

$$Q' = P' + \operatorname{ort}_{U+W} \overline{PQ} = P + \boldsymbol{u} + \overline{PQ} - \operatorname{pr}_{U+W} \overline{PQ} = Q - \boldsymbol{w},$$

где $Q \in \mathfrak{B}$ и $\boldsymbol{w} \in W$, т.е. $Q' \in (\mathfrak{B}, W)$. □

Если $d(\mathfrak{C}, \mathfrak{B}) \neq 0$, то точки $P'$ и $Q'$, найденные в предыдущем доказательстве, различны. Прямая, содержащая $P'$ и $Q'$, перпендикулярна каждому из подпространств $(\mathfrak{C}, U)$ и $(\mathfrak{B}, W)$ и называется их *общим перпендикуляром*. Такая прямая единственна тогда и только тогда, когда векторы $\boldsymbol{u}$ и $\boldsymbol{w}$ в разложении (3.7) определены однозначно, т.е. когда $U \cap W = \{\mathbf{0}\}$. Например, это так в случае скрещивающихся прямых в 3-мерном пространстве.

## 3.7. Определитель матрицы Грама и многомерный объём

Определение 3.7.1. *Матрицей Грама* системы векторов $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ евклидова пространства называется матрица

$$G = G(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k) = \begin{pmatrix} (\boldsymbol{a}_1, \boldsymbol{a}_1) & (\boldsymbol{a}_1, \boldsymbol{a}_2) & \ldots & (\boldsymbol{a}_1, \boldsymbol{a}_k) \\ (\boldsymbol{a}_2, \boldsymbol{a}_1) & (\boldsymbol{a}_2, \boldsymbol{a}_2) & \ldots & (\boldsymbol{a}_2, \boldsymbol{a}_k) \\ \vdots & \vdots & \ddots & \vdots \\ (\boldsymbol{a}_k, \boldsymbol{a}_1) & (\boldsymbol{a}_k, \boldsymbol{a}_2) & \ldots & (\boldsymbol{a}_k, \boldsymbol{a}_k) \end{pmatrix}.$$

Матрица $G$ симметрична: $G^t = G$ (в эрмитовом пространстве матрица $G$ эрмитова, т.е. $\overline{G}^t = G$.)

Матрица Грама уже появлялась как матрица системы для нахождения коэффициентов проекции вектора на подпространство $\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_k \rangle$ (см. предложение 3.5.4).

Предложение 3.7.2. *Пусть $G$ — матрица Грама системы векторов $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$, а $A = (a^i_j)$ — матрица, в столбцы которой записаны координаты векторов $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ в некотором ортонормированном базисе. Тогда имеет место соотношение*

$$G = A^t A.$$

Доказательство. Это следует из закона умножения матриц и формулы для скалярного произведения в ортонормированном базисе (предложение 3.3.5). □

Определение 3.7.3. Пусть $P \in (\mathfrak{A}, V)$ — точка в аффинном пространстве, а $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ — набор векторов в линейном пространстве $V$. *Параллелепипедом* с вершиной в точке $P$, натянутым на векторы $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$, называется следующее множество точек аффинного пространства $\mathfrak{A}$:

$$\Pi(P; \boldsymbol{a}_1, \dots, \boldsymbol{a}_k) := \{Q \in \mathfrak{A} \colon Q = P + x^1 \boldsymbol{a}_1 + \ldots + x^n \boldsymbol{a}_k, \quad 0 \leqslant x^i \leqslant 1\}.$$

Определение 3.7.4. Определим *$k$-мерный объём* $\operatorname{vol}_k$ параллелепипеда $\Pi(P; \boldsymbol{a}_1, \dots, \boldsymbol{a}_k)$ в аффинном евклидовом пространстве индуктивно:

1) одномерный объём $\operatorname{vol}_1 \Pi(P; \boldsymbol{a}_1) := |\boldsymbol{a}_1|$ — это длина вектора;
2) $\operatorname{vol}_k \Pi(P; \boldsymbol{a}_1, \dots, \boldsymbol{a}_k) := \operatorname{vol}_{k-1} \Pi(P; \boldsymbol{a}_1, \dots, \boldsymbol{a}_{k-1}) \cdot |\operatorname{ort}_{\langle \boldsymbol{a}_1, \dots, \boldsymbol{a}_{k-1} \rangle} \boldsymbol{a}_k|$.

Это определение обобщает определение площади параллелограмма (или объёма 3-мерного параллелепипеда) как произведение длины (или площади) основания на высоту. Из определения объёма $\operatorname{vol}_k \Pi(P; \boldsymbol{a}_1, \dots, \boldsymbol{a}_k)$ видно, что он не зависит от вершины $P$; поэтому далее мы будем использовать обозначение $\operatorname{vol}_k \Pi(\boldsymbol{a}_1, \dots, \boldsymbol{a}_k)$.

Теорема 3.7.5. *Квадрат объёма равен определителю матрицы Грама:*

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

*В частности, объём* $\operatorname{vol}_k \Pi(\boldsymbol{a}_1, \dots, \boldsymbol{a}_k)$ *не зависит от порядка векторов.*

Доказательство. Индукция по $k$. При $k = 1$, очевидно, $|\boldsymbol{a}_1|^2 = (\boldsymbol{a}_1, \boldsymbol{a}_1)$.

Пусть утверждение доказано для $\operatorname{vol}_{k-1}$, докажем его для $\operatorname{vol}_k$. Рассмотрим разложение $\boldsymbol{a}_k = \operatorname{pr}_{\langle \boldsymbol{a}_1, \dots, \boldsymbol{a}_{k-1} \rangle} \boldsymbol{a}_k + \operatorname{ort}_{\langle \boldsymbol{a}_1, \dots, \boldsymbol{a}_{k-1} \rangle} \boldsymbol{a}_k$, где $\operatorname{pr}_{\langle \boldsymbol{a}_1, \dots, \boldsymbol{a}_{k-1} \rangle} \boldsymbol{a}_k = \lambda_1 \boldsymbol{a}_1 + \ldots + \lambda_{k-1} \boldsymbol{a}_{k-1}$, и обозначим $\boldsymbol{b} = \operatorname{ort}_{\langle \boldsymbol{a}_1, \dots, \boldsymbol{a}_{k-1} \rangle} \boldsymbol{a}_k$. Тогда $(\boldsymbol{a}_i, \boldsymbol{b}) = 0$ при $i = 1, \dots, k-1$ и $(\boldsymbol{a}_k, \boldsymbol{b}) = (\boldsymbol{b}, \boldsymbol{b})$. Мы имеем

$$\det G = \begin{vmatrix} (\boldsymbol{a}_1, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_1, \boldsymbol{a}_k) \\ \vdots & \ddots & \vdots \\ (\boldsymbol{a}_k, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_k, \boldsymbol{a}_k) \end{vmatrix} = \begin{vmatrix} (\boldsymbol{a}_1, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_1, \boldsymbol{a}_{k-1}) & (\boldsymbol{a}_1, \lambda_1 \boldsymbol{a}_1 + \ldots + \lambda_{k-1} \boldsymbol{a}_{k-1} + \boldsymbol{b}) \\ \vdots & \ddots & \vdots & \vdots \\ (\boldsymbol{a}_k, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_k, \boldsymbol{a}_{k-1}) & (\boldsymbol{a}_k, \lambda_1 \boldsymbol{a}_1 + \ldots + \lambda_{k-1} \boldsymbol{a}_{k-1} + \boldsymbol{b}) \end{vmatrix} =$$

$$= \lambda_1 \begin{vmatrix} (\boldsymbol{a}_1, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_1, \boldsymbol{a}_{k-1}) & (\boldsymbol{a}_1, \boldsymbol{a}_1) \\ \vdots & \ddots & \vdots & \vdots \\ (\boldsymbol{a}_k, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_k, \boldsymbol{a}_{k-1}) & (\boldsymbol{a}_k, \boldsymbol{a}_1) \end{vmatrix} + \ldots + \lambda_{k-1} \begin{vmatrix} (\boldsymbol{a}_1, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_1, \boldsymbol{a}_{k-1}) & (\boldsymbol{a}_1, \boldsymbol{a}_{k-1}) \\ \vdots & \ddots & \vdots & \vdots \\ (\boldsymbol{a}_k, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_k, \boldsymbol{a}_{k-1}) & (\boldsymbol{a}_k, \boldsymbol{a}_{k-1}) \end{vmatrix} +$$

$$+ \begin{vmatrix} (\boldsymbol{a}_1, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_1, \boldsymbol{a}_{k-1}) & (\boldsymbol{a}_1, \boldsymbol{b}) \\ \vdots & \ddots & \vdots & \vdots \\ (\boldsymbol{a}_{k-1}, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_{k-1}, \boldsymbol{a}_{k-1}) & (\boldsymbol{a}_{k-1}, \boldsymbol{b}) \\ (\boldsymbol{a}_k, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_k, \boldsymbol{a}_{k-1}) & (\boldsymbol{a}_k, \boldsymbol{b}) \end{vmatrix} = \begin{vmatrix} (\boldsymbol{a}_1, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_1, \boldsymbol{a}_{k-1}) & 0 \\ \vdots & \ddots & \vdots & \vdots \\ (\boldsymbol{a}_{k-1}, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_{k-1}, \boldsymbol{a}_{k-1}) & 0 \\ (\boldsymbol{a}_k, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_k, \boldsymbol{a}_{k-1}) & (\boldsymbol{b}, \boldsymbol{b}) \end{vmatrix} =$$

$$= \begin{vmatrix} (\boldsymbol{a}_1, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_1, \boldsymbol{a}_{k-1}) \\ \vdots & \ddots & \vdots \\ (\boldsymbol{a}_{k-1}, \boldsymbol{a}_1) & \dots & (\boldsymbol{a}_{k-1}, \boldsymbol{a}_{k-1}) \end{vmatrix} \cdot (\boldsymbol{b}, \boldsymbol{b}) = \big(\operatorname{vol}_{k-1} \Pi(\boldsymbol{a}_1, \dots, \boldsymbol{a}_{k-1})\big)^2 |\boldsymbol{b}|^2 = \big(\operatorname{vol}_k \Pi(\boldsymbol{a}_1, \dots, \boldsymbol{a}_k)\big)^2.$$

$\square$

Следствие 3.7.6. *Векторы* $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ *линейно зависимы тогда и только тогда, когда* $\det G(\boldsymbol{a}_1, \dots, \boldsymbol{a}_k) = 0$.

Доказательство. Действительно, предположим, что векторы $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ линейно зависимы. Можно считать, что $\boldsymbol{a}_k$ линейно выражается через $\boldsymbol{a}_1, \dots, \boldsymbol{a}_{k-1}$. Тогда

$\operatorname{ort}_{\langle \boldsymbol{a}_1,\dots,\boldsymbol{a}_{k-1}\rangle} \boldsymbol{a}_k = \mathbf{0}$ и, следовательно,

$$\det G = \bigl(\operatorname{vol}_k \Pi(\boldsymbol{a}_1,\dots,\boldsymbol{a}_k)\bigr)^2 = \bigl(\operatorname{vol}_{k-1} \Pi(\boldsymbol{a}_1,\dots,\boldsymbol{a}_{k-1})\bigr)^2 |\operatorname{ort}_{\langle \boldsymbol{a}_1,\dots,\boldsymbol{a}_{k-1}\rangle} \boldsymbol{a}_k|^2 = 0.$$

Обратно, пусть $\det G = (\operatorname{vol}_k \Pi(\boldsymbol{a}_1,\dots,\boldsymbol{a}_k))^2 = 0$. Тогда, в силу индуктивного определения объема, мы имеем $\operatorname{vol}_i \Pi(\boldsymbol{a}_1,\dots,\boldsymbol{a}_i) = 0$, а $\operatorname{vol}_{i-1} \Pi(\boldsymbol{a}_1,\dots,\boldsymbol{a}_{i-1}) \neq 0$ для некоторого $i$. Так как $\operatorname{vol}_i \Pi(\boldsymbol{a}_1,\dots,\boldsymbol{a}_i) = \operatorname{vol}_{i-1} \Pi(\boldsymbol{a}_1,\dots,\boldsymbol{a}_{i-1}) \operatorname{ort}_{\langle \boldsymbol{a}_1,\dots,\boldsymbol{a}_{i-1}\rangle} \boldsymbol{a}_i$, это означает, что $\operatorname{ort}_{\langle \boldsymbol{a}_1,\dots,\boldsymbol{a}_{i-1}\rangle} \boldsymbol{a}_i = \mathbf{0}$, т.е. $\boldsymbol{a}_i$ линейно выражается через $\boldsymbol{a}_1,\dots,\boldsymbol{a}_{i-1}$. □

Следствие 3.7.7. *Пусть* $\dim V = n$ *и* $A = (a^i_j)$ *— квадратная матрица из координат векторов* $\boldsymbol{a}_1,\dots,\boldsymbol{a}_n$ *в некотором ортонормированном базисе. Тогда*

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

Доказательство. Из предложения 3.7.2 и предыдущей теоремы получаем

$$\bigl(\operatorname{vol}_n \Pi(\boldsymbol{a}_1,\dots,\boldsymbol{a}_n)\bigr)^2 = \det G = \det(A^t A) = (\det A)^2.$$

□

## 3.8. Метод наименьших квадратов

Часто на практике при исследовании какого-нибудь природного или социального явления делается допущение, что это явление описывается линейной формулой. Точнее, мы предполагаем, что некоторая величина $b$ линейно зависит от других величин $a_1,\dots,a_n$, и хотим найти эту зависимость

$$b = a_1x^1 + \ldots + a_nx^n,$$

т.е. найти неизвестные коэффициенты $x^1,\dots,x^n$ (это называется *моделью линейной регрессии*). Для нахождения зависимости $b$ от $a_1,\dots,a_n$ делается большое число $m$ измерений (как правило $m \gg n$), и по таблице измеренных значений записывается система линейных уравнений

$$\begin{cases} a^1_1x^1 + \ldots + a^1_nx^n = b^1 \\ \quad\cdots\quad\cdots\quad\cdots \\ a^m_1x^1 + \ldots + a^m_nx^n = b^m \end{cases} \tag{3.8}$$

в которой число неизвестных меньше числа уравнений. Такая система, как правило, несовместна. Поэтому находится «наилучшее приближённое» решение $x^1,\dots,x^n$, для которого отклонение значений $b^i$ от $a^i_jx^j = a^i_1x^1 + \ldots + a^i_nx^n$ будет наименьшим.

Метод наименьших квадратов решает задачу нахождения наилучшего приближённого решения в предположении, что в качестве меры отклонения берётся сумма квадратов разностей величин $a^i_1x^1 + \ldots + a^i_nx^n$ и $b^i$.

Определение 3.8.1. *Псевдорешением* системы (3.8) называется набор $\tilde{x}^1,\dots,\tilde{x}^n$, который минимизирует сумму квадратов разностей левых и правых частей уравнений системы, т.е. минимизирует величину

$$(a^1_jx^j - b^1)^2 + (a^2_jx^j - b^2)^2 + \ldots + (a^m_jx^j - b^m)^2 \tag{3.9}$$

по всем $(x^1,\dots,x^n) \in \mathbb{R}^n$. Эта величина называется *квадратичным отклонением.*

Пусть $A = (a^i_j)$ — матрица системы (3.8), $\boldsymbol{a}_1,\dots,\boldsymbol{a}_n \in \mathbb{R}^m$ — векторы-столбцы этой матрицы, а $\boldsymbol{b} \in \mathbb{R}^m$ — вектор правых частей.

Теорема 3.8.2. *Псевдорешение системы* (3.8) *находится как решение системы*

$$\begin{cases} (\boldsymbol{a}_1, \boldsymbol{a}_1)x^1 + \ldots + (\boldsymbol{a}_1, \boldsymbol{a}_n)x^n = (\boldsymbol{a}_1, \boldsymbol{b}), \\ \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\ (\boldsymbol{a}_n, \boldsymbol{a}_1)x^1 + \ldots + (\boldsymbol{a}_n, \boldsymbol{a}_n)x^n = (\boldsymbol{a}_n, \boldsymbol{b}). \end{cases}$$

*Другими словами, псевдорешение — это набор коэффициентов в разложении проекции* $\mathrm{pr}_{\langle \boldsymbol{a}_1,\ldots,\boldsymbol{a}_n\rangle}\,\boldsymbol{b}$ *по векторам* $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$, *а квадратичное отклонение псевдорешения — это квадрат длины вектора* $\mathrm{ort}_{\langle \boldsymbol{a}_1,\ldots,\boldsymbol{a}_n\rangle}\,\boldsymbol{b}$.

Доказательство. Квадратичное отклонение (3.9) набора $x^1, \ldots, x^n$ — это по определению квадрат длины вектора $\boldsymbol{a}_j x^j - \boldsymbol{b}$, т.е. квадрат расстояния между точками $\boldsymbol{b}$ и $\boldsymbol{a}_j x^j \in \langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\rangle$ аффинного пространства $\mathbb{A}^m$. Мы знаем из теоремы 3.6.2, что расстояние между $\boldsymbol{b}$ и точкой $\boldsymbol{a}_j x^j$ подпространства $\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\rangle$ минимально, когда $\boldsymbol{a}_j x^j$ — это проекция вектора $\boldsymbol{b}$ на $\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\rangle$. Коэффициенты в разложении проекции по векторам подпространства находятся из указанной системы (предложение 3.5.4), а минимальное расстояние равно $|\,\mathrm{ort}_{\langle \boldsymbol{a}_1,\ldots,\boldsymbol{a}_n\rangle}\,\boldsymbol{b}|$. □

Аналогично, методом наименьших квадратов можно находить более сложные зависимости величины $b$ от $a_1, \ldots, a_n$. Например, в случае неоднородной линейной зависимости $b = x^0 + a_1 x^1 + \ldots + a_n x^n$ можно находить коэффициенты $x^0, x^1, \ldots, x^n$. В случае, когда предполагаемая зависимость $b$ от одной величины $a$ выражается многочленом $n$-й степени $b = x^0 + a x^1 + a^2 x^2 + \ldots + a^n x^n$ с неизвестными коэффициентами $x^0, x^1, \ldots, x^n$, метод наименьших квадратов позволяет находить наилучшее приближение для этих коэффициентов.

## 3.9. Изоморфизмы евклидовых и эрмитовых пространств. Канонический изоморфизм евклидова пространства и его сопряжённого

Определение 3.9.1. Два евклидовых или эрмитовых пространства $V$ и $W$ называются *изоморфными*, если существует изоморфизм линейных пространств $\mathcal{A}\colon V \to W$, сохраняющий скалярное произведение, т.е.

$$\bigl(\mathcal{A}\boldsymbol{u}, \mathcal{A}\boldsymbol{v}\bigr) = (\boldsymbol{u}, \boldsymbol{v}) \qquad \text{для любых } \boldsymbol{u}, \boldsymbol{v} \in V.$$

Предложение 3.9.2. *Два евклидовых или эрмитовых пространства* $V$ *и* $W$ *изоморфны тогда и только тогда, когда их размерности совпадают.*

Доказательство. Если $V$ и $W$ изоморфны как евклидовы (эрмитовы) пространства, то они изоморфны как линейные пространства, а потому $\dim V = \dim W$.

Доказательство обратного утверждения аналогично доказательству соответствующего утверждения для линейных пространств (теорема 1.7.3): изоморфизм между евклидовыми (эрмитовыми) пространствами $n$ устанавливается при помощи взаимно однозначного соответствия между базисами $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ в $V$ и $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ в $W$. Для того, чтобы получаемый изоморфизм линейных пространств $\mathcal{A}\colon V \to W$ сохранял скалярное произведение, базисы необходимо выбрать ортонормированными. В этом случае мы имеем $\boldsymbol{f}_i = \mathcal{A}\boldsymbol{e}_i$ и $(\boldsymbol{e}_i, \boldsymbol{e}_j) = (\boldsymbol{f}_i, \boldsymbol{f}_j) = \delta_{ij}$. Поэтому для любых векторов $\boldsymbol{u} = u^i \boldsymbol{e}_i$ и $\boldsymbol{v} = v^j \boldsymbol{e}_j$ из $V$ мы имеем

$$(\mathcal{A}\boldsymbol{u}, \mathcal{A}\boldsymbol{v}) = \overline{u}^i v^j (\mathcal{A}\boldsymbol{e}_i, \mathcal{A}\boldsymbol{e}_j) = \overline{u}^i v^j (\boldsymbol{f}_i, \boldsymbol{f}_j) = \overline{u}^i v^j \delta_{ij} = \sum_{i=1}^n \overline{u}^i v^i = (\boldsymbol{u}, \boldsymbol{v}),$$

т.е. изоморфизм $\mathcal{A}$ сохраняет скалярное произведение. $\square$

Так как пространства $V$ и $V^*$ имеют одну размерность (в конечномерном случае), они изоморфны. Однако построение изоморфизма между ними требует выбора базисов и в этом смысле неканонично. Оказывается, что в присутствии скалярного произведения можно установить изоморфизм $V \to V^*$ каноническим образом, т.е. не прибегая к выбору базисов.

Пусть $V$ — евклидово пространство. Каждому вектору $\boldsymbol{u} \in V$ сопоставим линейную функцию $\xi_{\boldsymbol{u}} = (\boldsymbol{u}, \cdot\,)$, заданную по формуле $\xi_{\boldsymbol{u}}(\boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v})$.

Теорема 3.9.3. *Пусть $V$ — евклидово пространство. Отображение $\boldsymbol{u} \mapsto \xi_{\boldsymbol{u}} = (\boldsymbol{u}, \cdot\,)$ устанавливает канонический изоморфизм $\mathcal{A} \colon V \to V^*$.*

Доказательство. Линейность отображения $\boldsymbol{u} \mapsto \xi_{\boldsymbol{u}}$ вытекает из линейности скалярного произведения по первому аргументу. Так как $\dim V = \dim V^*$, чтобы проверить, что $\mathcal{A} \colon V \to V^*$ — изоморфизм, достаточно проверить, что $\ker \mathcal{A} = \{\mathbf{0}\}$. Пусть $\boldsymbol{v} \in \ker \mathcal{A}$, т.е. $\mathcal{A}\boldsymbol{v} = \xi_{\boldsymbol{v}} = o$. Следовательно, $\xi_{\boldsymbol{v}}(\boldsymbol{w}) = (\boldsymbol{v}, \boldsymbol{w}) = 0$ для любого $\boldsymbol{w} \in V$. Но тогда и $(\boldsymbol{v}, \boldsymbol{v}) = 0$, значит $\boldsymbol{v} = \mathbf{0}$ и ядро отображения $\mathcal{A}$ нулевое. $\square$

Аналогичным образом для эрмитова пространства $V$ устанавливается канонический изоморфизм $\overline{V} \to V^*$, $\boldsymbol{u} \mapsto (\boldsymbol{u}, \cdot\,)$, где $\overline{V}$ — комплексно сопряжённое пространство (с умножением на скаляры, определённым по формуле $\lambda \cdot \boldsymbol{v} := \overline{\lambda}\boldsymbol{v}$). Это позволяет отождествить два понятия «сопряжённого» пространства для эрмитова пространства $V$.

Глава 4

# Операторы в евклидовых и эрмитовых пространствах

## 4.1. Сопряжённые операторы в евклидовых и эрмитовых пространствах

Пусть $\mathcal{A}\colon V \to V$ — линейный оператор в евклидовом пространстве $V$. В разделе 1.11 мы определили сопряжённое линейное отображение $\mathcal{A}^*\colon V^* \to V^*$ по формуле

$$(\mathcal{A}^*\xi)(\boldsymbol{v}) := \xi(\mathcal{A}\boldsymbol{v}) \qquad \text{для } \xi \in V^*, \boldsymbol{v} \in V.$$

При каноническом отождествлении $V \leftrightarrow V^*$, $\boldsymbol{u} \leftrightarrow \xi_{\boldsymbol{u}} = (\boldsymbol{u},\,\cdot\,)$ оператор $\mathcal{A}^*\colon V^* \to V^*$ переходит в оператор $\mathcal{A}^*\colon V \to V$, (который мы для простоты будем обозначать тем же символом $\mathcal{A}^*$), удовлетворяющий соотношению $\xi_{\mathcal{A}^*\boldsymbol{u}} = \mathcal{A}^*\xi_{\boldsymbol{u}}$ для любого вектора $\boldsymbol{u} \in V$. Чтобы установить связь между $\mathcal{A}\colon V \to V$ и $\mathcal{A}^*\colon V \to V$ непосредственно, вычислим значение линейных функций $\xi_{\mathcal{A}^*\boldsymbol{u}}$ и $\mathcal{A}^*\xi_{\boldsymbol{u}}$ на векторе $\boldsymbol{v} \in V$:

$$\xi_{\mathcal{A}^*\boldsymbol{u}}(\boldsymbol{v}) = (\mathcal{A}^*\boldsymbol{u}, \boldsymbol{v}), \quad (\mathcal{A}^*\xi_{\boldsymbol{u}})(\boldsymbol{v}) = \xi_{\boldsymbol{u}}(\mathcal{A}\boldsymbol{v}) = (\boldsymbol{u}, \mathcal{A}\boldsymbol{v}).$$

Так как $\xi_{\mathcal{A}^*\boldsymbol{u}} = \mathcal{A}^*\xi_{\boldsymbol{u}}$, мы получаем $(\mathcal{A}^*\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \mathcal{A}\boldsymbol{v})$ для любых $\boldsymbol{u}, \boldsymbol{v} \in V$.

Определение 4.1.1. Пусть $\mathcal{A}\colon V \to V$ — оператор в евклидовом или эрмитовом пространстве $V$. Линейный оператор $\mathcal{A}^*\colon V \to V$, удовлетворяющий соотношению

$$(\mathcal{A}^*\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \mathcal{A}\boldsymbol{v})$$

для любых $\boldsymbol{u}, \boldsymbol{v} \in V$, называется *сопряжённым* к $\mathcal{A}$.

Соотношение $(\mathcal{A}^*\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \mathcal{A}\boldsymbol{v})$ определяет оператор $\mathcal{A}^*$ однозначно. Действительно, если $(\mathcal{A}'\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \mathcal{A}\boldsymbol{v})$ для другого оператора $\mathcal{A}'\colon V \to V$, то мы получаем $((\mathcal{A}^* - \mathcal{A}')\boldsymbol{u}, \boldsymbol{v}) = 0$ для любых $\boldsymbol{u}, \boldsymbol{v} \in V$. В частности, $((\mathcal{A}^* - \mathcal{A}')\boldsymbol{u}, (\mathcal{A}^* - \mathcal{A}')\boldsymbol{u}) = 0$, т.е. $(\mathcal{A}^* - \mathcal{A}')\boldsymbol{u} = \boldsymbol{0}$ для любого $\boldsymbol{u} \in V$. Следовательно, $\mathcal{A}^* = \mathcal{A}'$.

Предложение 4.1.2. *Пусть $A$ — матрица оператора $\mathcal{A}\colon V \to V$ в ортонормированном базисе евклидова (эрмитова) пространства $V$. Тогда матрица сопряжённого оператора $\mathcal{A}^*\colon V \to V$ в том же базисе есть $A^t$ (соответственно, $\overline{A}^t$).*

Доказательство. Это следует из предложения 1.11.2 (о матрице сопряжённого отображения), но мы также дадим прямое доказательство. Пусть $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ — ортонормированный базис, $A = (a^i_j)$ — матрица оператора $\mathcal{A}\colon V \to V$ в этом базисе, а $\widetilde{A} = (\widetilde{a}^i_j)$ — матрица оператора $\mathcal{A}^*\colon V \to V$. Тогда мы имеем

$$(\mathcal{A}\boldsymbol{e}_j, \boldsymbol{e}_k) = (a^i_j\boldsymbol{e}_i, \boldsymbol{e}_k) = \overline{a}^i_j(\boldsymbol{e}_i, \boldsymbol{e}_k) = \overline{a}^i_j\delta_{ik} = \overline{a}^k_j,$$
$$(\boldsymbol{e}_j, \mathcal{A}^*\boldsymbol{e}_k) = (\boldsymbol{e}_j, \widetilde{a}^i_k\boldsymbol{e}_i) = \widetilde{a}^i_k(\boldsymbol{e}_j, \boldsymbol{e}_i) = \widetilde{a}^i_k\delta_{ji} = \widetilde{a}^j_k.$$

Так как $(\mathcal{A}\boldsymbol{e}_j, \boldsymbol{e}_k) = (\boldsymbol{e}_j, \mathcal{A}^*\boldsymbol{e}_k)$, мы получаем $\overline{a}^k_j = \widetilde{a}^j_k$ или $\overline{A}^t = \widetilde{A}$. □

Предложение 4.1.3. *Имеют место следующие равенства:*

а) $(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*$, $\quad (\lambda\mathcal{A})^* = \overline{\lambda}\mathcal{A}^*$;

б) $(\mathcal{AB})^* = \mathcal{B}^*\mathcal{A}^*$.

Доказательство. Это следует из свойств операции транспонирования матриц, благодаря предложению 4.1.2. □

Докажем ключевую лемму, которая нам не раз понадобится в дальнейшем.

Лемма 4.1.4. *Если $W \subset V$ — инвариантное подпространство относительно $\mathcal{A}$, то $W^\perp$ — инвариантное подпространство относительно $\mathcal{A}^*$.*

Доказательство. Пусть $\boldsymbol{u} \in W^\perp$. Тогда для любого $\boldsymbol{w} \in W$ имеем $(\mathcal{A}^*\boldsymbol{u}, \boldsymbol{w}) = (\boldsymbol{u}, \mathcal{A}\boldsymbol{w}) = 0$ так как $\mathcal{A}\boldsymbol{w} \in W$, а $\boldsymbol{u} \in W^\perp$. Следовательно, $\mathcal{A}^*\boldsymbol{u} \in W^\perp$. □

## 4.2. Самосопряжённые операторы. Канонический вид

Определение 4.2.1. Оператор $\mathcal{A}\colon V \to V$ в евклидовом или эрмитовом пространстве называется *самосопряжённым*, если $\mathcal{A}^* = \mathcal{A}$, т.е. выполнено соотношение

$$(\mathcal{A}\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \mathcal{A}\boldsymbol{v})$$

для любых $\boldsymbol{u}, \boldsymbol{v} \in V$.

Предложение 4.2.2. *Матрица $A$ самосопряжённого оператора $\mathcal{A}$ в ортонормированном базисе евклидова (эрмитова) пространства симметрична (эрмитова), т.е. $A^t = A$ (соответственно, $\overline{A}^t = A$).*

*Если матрица оператора $\mathcal{A}$ в некотором ортонормированном базисе симметрична (эрмитова), то оператор $\mathcal{A}$ самосопряжён.*

Доказательство. Первое утверждение вытекает из предложения 4.1.2: так как матрица оператора $\mathcal{A}^*$ есть $\overline{A}^t$ и $\mathcal{A}^* = \mathcal{A}$, мы получаем $\overline{A}^t = A$.

Докажем второе утверждение. Пусть $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ — ортонормированный базис, в котором матрица $A$ оператора $\mathcal{A}$ эрмитова, т.е. $\overline{A}^t = A$. Тогда из предложения 4.1.2 следует, что матрица $\overline{A}^t$ оператора $\mathcal{A}^*$ в том же базисе совпадает с $A$. Следовательно, $\mathcal{A}^* = \mathcal{A}$ и оператор $\mathcal{A}$ самосопряжён. □

В связи с этим самосопряжённые операторы в евклидовом пространстве также называют *симметрическими*, а в эрмитовом пространстве — *эрмитовыми*.

Вот основное свойство самосопряжённых операторов.

Теорема 4.2.3. *Самосопряжённый оператор диагонализируем в ортонормированном базисе. Другими словами, для самосопряжённого оператора существует ортонормированный базис из собственных векторов.*

Доказательство будет опираться на лемму, которая важна сама по себе.

Лемма 4.2.4. *Все корни характеристического многочлена самосопряжённого оператора $\mathcal{A}$ вещественны.*

Доказательство. Вначале докажем лемму для эрмитова пространства. В этом случае корни характеристического многочлена суть собственные значения оператора $\mathcal{A}$. Пусть $\lambda \in \mathbb{C}$ такой корень и $\boldsymbol{v} \neq \boldsymbol{0}$ — соответствующий собственный вектор, т.е. $\mathcal{A}\boldsymbol{v} = \lambda\boldsymbol{v}$. Тогда

$$\overline{\lambda}(\boldsymbol{v}, \boldsymbol{v}) = (\lambda\boldsymbol{v}, \boldsymbol{v}) = (\mathcal{A}\boldsymbol{v}, \boldsymbol{v}) = (\boldsymbol{v}, \mathcal{A}\boldsymbol{v}) = (\boldsymbol{v}, \lambda\boldsymbol{v}) = \lambda(\boldsymbol{v}, \boldsymbol{v}).$$

Так как $(\boldsymbol{v}, \boldsymbol{v}) \neq 0$, получаем $\overline{\lambda} = \lambda$, т.е. $\lambda \in \mathbb{R}$.

Случай евклидова пространства сводится к эрмитовому случаю при помощи комплексификации. Пусть $A$ — матрица самосопряжённого оператора $\mathcal{A}$ в ортонормированном базисе евклидова пространства $V$. Тогда матрица $A$ вещественна и симметрична. Та же матрица $A$ будет матрицей комплексифицированного оператора $\mathcal{A}_{\mathbb{C}}$ в соответствующем базисе эрмитова пространства $V_{\mathbb{C}}$. Этот базис также ортонормирован, а матрица $A$, будучи вещественной и симметричной, является эрмитовой. Следовательно, оператор $\mathcal{A}_{\mathbb{C}}$ самосопряжён, а корни его характеристического многочлена вещественны и совпадают с корнями многочлена оператора $\mathcal{A}$. □

Доказательство теоремы 4.2.3. Используем индукцию по размерности пространства $V$. При $\dim V = 1$ доказывать нечего. Предположим, что утверждение доказано для операторов в пространствах размерности $n-1$ и докажем его для пространства $V$ размерности $n$.

В силу предыдущей леммы у самосопряжённого оператора $\mathcal{A}$ имеется собственный вектор $\boldsymbol{v}$, т.е. одномерное инвариантное подпространство $W = \langle \boldsymbol{v} \rangle$. В силу леммы 4.1.4 ортогональное дополнение $W^{\perp}$ инвариантно относительно оператора $\mathcal{A}^* = \mathcal{A}$. Так как $\dim W^{\perp} = n-1$, в пространстве $W^{\perp}$ имеется ортонормированный базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{n-1}$ из собственных векторов оператора $\mathcal{A}|_{W^{\perp}}$. Тогда $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{n-1}, \frac{\boldsymbol{v}}{|\boldsymbol{v}|}$ — ортонормированный базис из собственных векторов оператора $\mathcal{A}$. □

Ортонормированный базис, в котором матрица самосопряжённого оператора $\mathcal{A}$ диагональна, называется *каноническим*, а сама диагональная матрица называется *каноническим видом* самосопряжённого оператора.

На практике для нахождения канонического базиса используется лемма.

Лемма 4.2.5. *Собственные векторы, отвечающие различным собственным значениям самосопряжённого оператора $\mathcal{A}$, взаимно ортогональны.*

Доказательство. Пусть $\mathcal{A}\boldsymbol{u} = \lambda\boldsymbol{u}$ и $\mathcal{A}\boldsymbol{v} = \mu\boldsymbol{v}$, где $\lambda \neq \mu$ — вещественные собственные значения. Тогда

$$\lambda(\boldsymbol{u}, \boldsymbol{v}) = (\lambda\boldsymbol{u}, \boldsymbol{v}) = (\mathcal{A}\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \mathcal{A}\boldsymbol{v}) = (\boldsymbol{u}, \mu\boldsymbol{v}) = \mu(\boldsymbol{u}, \boldsymbol{v}),$$

откуда $(\boldsymbol{u}, \boldsymbol{v}) = 0$, так как $\lambda \neq \mu$. □

Для нахождения канонического базиса самосопряжённого оператора $\mathcal{A}$ находятся все его собственные подпространства, а затем в каждом из них выбирается ортонормированный базис. Объединение этих базисов и будет каноническим базисом для $\mathcal{A}$.

В евклидовом пространстве верно утверждение, обратное к теореме 4.2.3:

Предложение 4.2.6. *Если оператор $\mathcal{A}$ в евклидовом пространстве диагонализируем в ортонормированном базисе, то $\mathcal{A}$ самосопряжён.*

Доказательство. Действительно, диагональная матрица симметрична, а оператор, имеющий симметричную матрицу в ортонормированном базисе евклидова пространства самосопряжён согласно предложению 4.2.2. □

В эрмитовом пространстве класс операторов, диагонализируемых в ортонормированном базисе, шире, чем самосопряжённые операторы (так как на диагонали могут стоять не вещественные числа). Такие операторы мы изучим позже.

## 4.3. Самосопряжённые проекторы. Спектральное разложение самосопряжённого оператора

Пусть пространство $V$ представлено в виде прямой суммы $V = U \oplus W$. Напомним (см. раздел 2.2), что оператор $\mathcal{P}\colon V \to V$, переводящий $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$ в $\boldsymbol{u}$ (где $\boldsymbol{u} \in U$, $\boldsymbol{w} \in W$), называется проектором на $U$ вдоль $W$. При этом $\operatorname{Im}\mathcal{P} = U$ и $\operatorname{Ker}\mathcal{P} = W$.

Определение 4.3.1. Проектор $\mathcal{P}\colon V \to V$ на $U$ вдоль $W$ называется *ортогональным*, если $W = U^{\perp}$. Такой проектор будем обозначать через $\operatorname{pr}_U$.

Это обозначение вполне согласуется с предыдущими: если $V = U \oplus U^{\perp}$, то для любого $\boldsymbol{v} \in V$ мы имеем $\boldsymbol{v} = \operatorname{pr}_U \boldsymbol{v} + \operatorname{ort}_U \boldsymbol{v}$.

Предложение 4.3.2. *Проектор $\mathcal{P}\colon V \to V$ является самосопряжённым оператором тогда и только тогда, когда он ортогонален.*

Доказательство. Пусть $\mathcal{P} = \operatorname{pr}_U$ — ортогональный проектор. Выберем ортонормированный базис в $U$ и дополним его до ортонормированного базиса в $V$. Тогда в этом ортонормированном базисе матрица оператора $\operatorname{pr}_U$ диагональна (с единицами и нулями на диагонали), а значит оператор $\operatorname{pr}_U$ самосопряжён.

Обратно пусть $\mathcal{P}$ — самосопряжённый проектор на $U$ вдоль $W$. Возьмём произвольные векторы $\boldsymbol{u} \in U = \operatorname{Im}\mathcal{P}$ и $\boldsymbol{w} \in W = \operatorname{Ker}\mathcal{P}$. Тогда $\boldsymbol{u} = \mathcal{P}\boldsymbol{v}$ для некоторого $\boldsymbol{v} \in V$ и $\mathcal{P}\boldsymbol{w} = \boldsymbol{0}$. Мы имеем

$$(\boldsymbol{u}, \boldsymbol{w}) = (\mathcal{P}\boldsymbol{v}, \boldsymbol{w}) = (\boldsymbol{v}, \mathcal{P}\boldsymbol{w}) = 0,$$

откуда получаем $W = U^{\perp}$ и $\mathcal{P} = \operatorname{pr}_U$. □

Пример 4.3.3. Пусть $\boldsymbol{u} = (u^1, \dots, u^n)^t \in \mathbb{R}^n$ — ненулевой вектор-столбец и $U = \langle \boldsymbol{u} \rangle$ — одномерное подпространство. Тогда матрица оператора $\operatorname{pr}_{\boldsymbol{u}} = \operatorname{pr}_U$ в стандартном базисе $\mathbb{R}^n$ есть

$$\frac{1}{|\boldsymbol{u}|^2} u u^t = \frac{1}{|\boldsymbol{u}|^2} \begin{pmatrix} u^1 \\ u^2 \\ \vdots \\ u^n \end{pmatrix} \begin{pmatrix} u^1 & u^2 & \cdots & u^n \end{pmatrix} = \frac{1}{|\boldsymbol{u}|^2} \begin{pmatrix} u^1u^1 & u^1u^2 & \cdots & u^1u^n \\ u^2u^1 & u^2u^2 & \cdots & u^2u^n \\ \vdots & \vdots & \ddots & \vdots \\ u^nu^1 & u^nu^2 & \cdots & u^nu^n \end{pmatrix}$$

Напомним, что спектром оператора $\mathcal{A}$ называется множество его собственных значений. Для каждого собственного значения $\lambda$ рассмотрим ортогональный проектор $\operatorname{pr}_{V_\lambda}$ на соответствующее собственное подпространство $V_\lambda$.

Теорема 4.3.4 (спектральное разложение). *Пусть $\mathcal{A}$ — самосопряжённый оператор. Тогда имеет место разложение*

$$\mathcal{A} = \sum_{\lambda} \lambda \operatorname{pr}_{V_\lambda},$$

*где сумма берётся по всем собственным значениями. При этом проекторы $\operatorname{pr}_{V_\lambda}$ удовлетворяют соотношениям $\operatorname{pr}_{V_\lambda} \mathcal{A} = \mathcal{A} \operatorname{pr}_{V_\lambda} = \lambda \operatorname{pr}_{V_\lambda}$ и $\operatorname{pr}_{V_\lambda} \operatorname{pr}_{V_\mu} = \mathcal{O}$ при $\lambda \neq \mu$.*

Доказательство. Соотношения $\operatorname{pr}_{V_\lambda} \mathcal{A} = \mathcal{A} \operatorname{pr}_{V_\lambda} = \lambda \operatorname{pr}_{V_\lambda}$ и $\operatorname{pr}_{V_\lambda} \operatorname{pr}_{V_\mu} = \mathcal{O}$ при $\lambda \neq \mu$ выполнены, так как они выполнены для матриц входящих в них операторов в каноническом базисе для оператора $\mathcal{A}$ (в этом базисе матрицы всех входящих в соотношения операторов диагональны). Разложению $V = \bigoplus_\lambda V_\lambda$ в прямую сумму собственных подпространств соответствует разложение тождественного оператора в сумму ортогональных проекторов

$$\mathrm{id} = \sum_{\lambda} \mathrm{pr}_{V_\lambda}.$$

Умножив это соотношение слева на $\mathcal{A}$ и использовав соотношение $\mathcal{A}\,\mathrm{pr}_{V_\lambda} = \lambda\,\mathrm{pr}_{V_\lambda}$, получим требуемое. □

## 4.4. Кососимметрические и косоэрмитовы операторы. Канонический вид. Эрмитово разложение

Определение 4.4.1. Оператор $\mathcal{A}\colon V \to V$ в евклидовом (эрмитовом) пространстве называется *кососимметрическим* (соответственно, *косоэрмитовым*), если $\mathcal{A}^* = -\mathcal{A}$, т.е. выполнено соотношение

$$(\mathcal{A}\boldsymbol{u}, \boldsymbol{v}) = -(\boldsymbol{u}, \mathcal{A}\boldsymbol{v})$$

для любых $\boldsymbol{u}, \boldsymbol{v} \in V$.

Предложение 4.4.2. *Матрица $A$ кососимметрического (косоэрмитова) оператора $\mathcal{A}$ в ортонормированном базисе евклидова (эрмитова) пространства кососимметрична (косоэрмитова), т.е. $A^t = -A$ (соответственно, $\overline{A}^t = -A$).*

*Если матрица оператора $\mathcal{A}$ в некотором ортонормированном базисе кососимметрична (косоэрмитова), то оператор $\mathcal{A}$ кососимметричен (косоэрмитов).*

Доказательство. То же, что и для самосопряжённых операторов (предложение 4.2.2). □

Теорема 4.4.3. *Для косоэрмитова оператора $\mathcal{A}$ существует ортонормированный базис, в котором его матрица диагональна с чисто мнимыми числами на диагонали. Другими словами, для косоэрмитова оператора существует ортонормированный базис из собственных векторов, а все собственные значения — чисто мнимые.*

Доказательство. Доказательство, как и в случае самосопряжённого оператора, основано на лемме 4.1.4. Будем вести индукцию по размерности пространства $V$. При $\dim V = 1$ доказывать нечего. Предположим, что утверждение доказано для операторов в пространствах размерности $n-1$ и докажем его для размерности $n$.

Выберем собственный вектор $\boldsymbol{v}$ для $\mathcal{A}$, т.е. одномерное инвариантное подпространство $W = \langle \boldsymbol{v} \rangle$. В силу леммы 4.1.4 ортогональное дополнение $W^\perp$ инвариантно относительно оператора $\mathcal{A}^*$, а значит оно инвариантно и относительно $\mathcal{A} = -\mathcal{A}^*$. Так как $\dim W^\perp = n-1$, в пространстве $W^\perp$ имеется ортонормированный базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{n-1}$ из собственных векторов оператора $\mathcal{A}|_{W^\perp}$. Тогда $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{n-1}, \frac{\boldsymbol{v}}{|\boldsymbol{v}|}$ — ортонормированный базис из собственных векторов оператора $\mathcal{A}$.

Пусть $D$ — диагональная матрица косоэрмитова оператора $\mathcal{A}$ в ортонормированном базисе из собственных векторов. Так как $\mathcal{A}^* = -\mathcal{A}$, получаем $\overline{D}^t = -D$. Следовательно, диагональные элементы матрицы $D$ (собственные числа) удовлетворяют соотношению $\overline{\lambda} = -\lambda$, т.е. являются чисто мнимыми. □

Кососимметрические операторы в евклидовом пространстве, как правило, не диагонализируемы. Например, как мы видели ранее, оператор с матрицей $\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}$ с ненулевым $a$ не диагонализируем в вещественном пространстве.

Теорема 4.4.4. *Для кососимметрического оператора $\mathcal{A}$ существует ортонормированный базис, в котором его матрица блочно-диагональная с блоками размера 1 или 2, причём блоки размера 1 нулевые, а блоки размера 2 имеют вид $\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}$ с ненулевыми $a \in \mathbb{R}$.*

Доказательство. В пространстве размерности 1 или 2 доказывать нечего, так как кососимметрическая матрица и так имеет там требуемый вид. Предположим, что утверждение доказано для операторов в пространствах размерности не больше $n-1$ и докажем его для пространства $V$ размерности $n$ (где $n \geqslant 3$).

В силу теоремы 2.6.3 б) для оператора $\mathcal{A}$ существует 1-мерное или 2-мерное инвариантное подпространство $W \subset V$. Как и в случае косоэрмитовых операторов, из леммы 4.1.4 следует, что ортогональное дополнение $W^{\perp}$ также инвариантно.

По предположению индукции, в пространстве $W^{\perp}$ имеется требуемый базис для оператора $\mathcal{A}|_{W^{\perp}}$. Выбрав произвольный ортонормированный базис в $W$ и взяв объединение базисов в $W^{\perp}$ и $W$, мы получим ортнормированный базис пространства $V$, в котором матрица оператора $\mathcal{A}$ состоит из блоков требуемого вида и ещё одного блока размера 1 или 2 — матрицы оператора $\mathcal{A}|_W$. Этот последний блок — кососимметрическая матрица размера 1 или 2, т.е. она тоже имеет требуемый вид. □

Ортонормированный базис, в котором матрица кососимметрического или косоэрмитова оператора $\mathcal{A}$ имеет вид, описанный в предыдущих теоремах, называется *каноническим*, а сама матрица называется *каноническим видом* оператора.

Заметим, что если $\mathcal{A}$ — эрмитов оператор, то оператор $i\mathcal{A}$ косоэрмитов и наоборот. Поэтому теорему 4.4.3 можно было свести к теореме 4.2.3.

Теорема 4.4.5 (эрмитово разложение). *Для любого оператора $\mathcal{A}$ в эрмитовом пространстве существует единственное представление в виде*

$$\mathcal{A} = \mathcal{R} + i\mathcal{I},$$

*где $\mathcal{R}$ и $\mathcal{I}$ — эрмитовы операторы.*

Доказательство. Сначала докажем единственность. Если $\mathcal{A} = \mathcal{R} + i\mathcal{I}$ — эрмитово разложение, то $\mathcal{A}^* = \mathcal{R}^* - i\mathcal{I}^* = \mathcal{R} - i\mathcal{I}$. Из этих двух соотношений получаем

$$\mathcal{R} = \frac{1}{2}(\mathcal{A}^* + \mathcal{A}), \qquad \mathcal{I} = \frac{i}{2}(\mathcal{A}^* - \mathcal{A}),$$

т.е. операторы $\mathcal{R}$ и $\mathcal{I}$ определены однозначно и эрмитово разложение единственно.

С другой стороны, операторы $\mathcal{R}$ и $\mathcal{I}$, задаваемые предыдущими формулами, очевидно эрмитовы (самосопряжены), так что эрмитово разложение существует. □

В одномерном эрмитовом пространстве $\mathbb{C}$ эрмитовы операторы — это вещественные числа, а операторы $\mathcal{R}$ и $\mathcal{I}$ в эрмитовом разложении — это вещественная и мнимая части комплексного числа.

## 4.5. Ортогональные и унитарные операторы. Канонический вид. Группы $O_n$ и $SO_n$, $U_n$ и $SU_n$

Предложение 4.5.1. *Следующие условия для оператора $\mathcal{A}\colon V \to V$ в евклидовом или эрмитовом пространстве эквивалентны:*

а) *оператор $\mathcal{A}$ сохраняет длины векторов, т.е. $|\mathcal{A}\boldsymbol{v}| = |\boldsymbol{v}|$ для любого $\boldsymbol{v} \in V$;*

б) *оператор* $\mathcal{A}$ *сохраняет скалярное произведение, т.е.* $(\mathcal{A}\boldsymbol{u}, \mathcal{A}\boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v})$ *для любых* $\boldsymbol{u}, \boldsymbol{v} \in V$*;*

в) *оператор* $\mathcal{A}$ *переводит ортонормированные базисы в ортонормированные, т.е. если* $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ *— ортонормированный базис, то* $\mathcal{A}\boldsymbol{e}_1, \ldots, \mathcal{A}\boldsymbol{e}_n$ *— также ортонормированный базис;*

г) *матрица* $A$ *оператора* $\mathcal{A}$ *в ортонормированном базисе ортогональна (унитарна), т.е.* $A^t A = E$ *(соответственно,* $\overline{A}^t A = E$*);*

д) $\mathcal{A}^*\mathcal{A} = \mathrm{id}$*, т.е. сопряжённый оператор к* $\mathcal{A}$ *является его обратным.*

Доказательство. Мы докажем импликации а)$\Leftrightarrow$б) и б)$\Rightarrow$в)$\Rightarrow$г)$\Rightarrow$д)$\Rightarrow$б).

а)$\Rightarrow$б). В евклидовом пространстве имеем $(\boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{u}) + 2(\boldsymbol{u}, \boldsymbol{v}) + (\boldsymbol{v}, \boldsymbol{v})$, откуда

$$(\boldsymbol{u}, \boldsymbol{v}) = \frac{1}{2}\big((\boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}) - (\boldsymbol{u}, \boldsymbol{u}) - (\boldsymbol{v}, \boldsymbol{v})\big).$$

Поэтому, если оператор $\mathcal{A}$ сохраняет длины, т.е. скалярные произведения вида $(\boldsymbol{v}, \boldsymbol{v})$, то он сохраняет и все скалярные произведения.

В эрмитовом пространстве имеем $(\boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{u}) + (\boldsymbol{u}, \boldsymbol{v}) + \overline{(\boldsymbol{u}, \boldsymbol{v})} + (\boldsymbol{v}, \boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{u}) + 2\,\mathrm{Re}(\boldsymbol{u}, \boldsymbol{v}) + (\boldsymbol{v}, \boldsymbol{v})$, откуда

$$\mathrm{Re}(\boldsymbol{u}, \boldsymbol{v}) = \frac{1}{2}\big((\boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}) - (\boldsymbol{u}, \boldsymbol{u}) - (\boldsymbol{v}, \boldsymbol{v})\big)$$

и аналогично

$$\mathrm{Im}(\boldsymbol{u}, \boldsymbol{v}) = -\frac{1}{2}\big((\boldsymbol{u}+i\boldsymbol{v}, \boldsymbol{u}+i\boldsymbol{v}) - (\boldsymbol{u}, \boldsymbol{u}) - (\boldsymbol{v}, \boldsymbol{v})\big).$$

Поэтому, если оператор $\mathcal{A}$ сохраняет длины, то он сохраняет и скалярное произведение $(\boldsymbol{u}, \boldsymbol{v})$, так как сохраняет его вещественную и мнимую части.

б)$\Rightarrow$а). Очевидно.

б)$\Rightarrow$в). Пусть $(\mathcal{A}\boldsymbol{u}, \mathcal{A}\boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v})$. Тогда если $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ — ортонормированный базис, то $(\mathcal{A}\boldsymbol{e}_i, \mathcal{A}\boldsymbol{e}_j) = (\boldsymbol{e}_i, \boldsymbol{e}_j) = \delta_{ij}$, т.е. базис $\mathcal{A}\boldsymbol{e}_1, \ldots, \mathcal{A}\boldsymbol{e}_n$ также ортонормирован.

в)$\Rightarrow$г). Пусть $\mathcal{A}$ переводит ортонормированный базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ в ортонормированный базис $\mathcal{A}\boldsymbol{e}_1, \ldots, \mathcal{A}\boldsymbol{e}_n$ и $A = (a^i_j)$ — матрица оператора в базисе $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Тогда

$$\delta_{ij} = (\mathcal{A}\boldsymbol{e}_i, \mathcal{A}\boldsymbol{e}_j) = (a^k_i \boldsymbol{e}_k, a^\ell_j \boldsymbol{e}_\ell) = \overline{a}^k_i a^\ell_j (\boldsymbol{e}_k, \boldsymbol{e}_\ell) = \overline{a}^k_i a^\ell_j \delta_{k\ell} = \overline{a}^k_i \delta_{k\ell} a^\ell_j.$$

Это эквивалентно матричному соотношению $E = \overline{A}^t E A$ или $\overline{A}^t A = E$.

г)$\Rightarrow$д). Это следует из предложения 4.1.2.

д)$\Rightarrow$б). Пусть $\mathcal{A}^*\mathcal{A} = \mathrm{id}$. Тогда $(\mathcal{A}\boldsymbol{u}, \mathcal{A}\boldsymbol{v}) = (\mathcal{A}^*\mathcal{A}\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v})$, т.е. $\mathcal{A}$ сохраняет скалярное произведение. □

Определение 4.5.2. Оператор $\mathcal{A}\colon V \to V$ в евклидовом (эрмитовом) пространстве, удовлетворяющий одному из эквивалентных условий из предложения 4.5.1, называется *ортогональным* (соответственно, *унитарным*).

Иногда ортогональные и унитарные операторы называют *изометрическими.*

Как и в случае самосопряжённых операторов, для приведения ортогонального или унитарного оператора к каноническому виду нам понадобится утверждение об инвариантности ортогонального дополнения.

Лемма 4.5.3. *Пусть* $\mathcal{A}\colon V \to V$ *— ортогональный или унитарный оператор, а* $W \subset V$ *— инвариантное подпространство. Тогда ортогональное дополнение* $W^\perp$ *также инвариантно относительно* $\mathcal{A}$*.*

Доказательство. Пусть $\boldsymbol{u} \in W^{\perp}$. Нам надо доказать, что $\boldsymbol{\mathcal{A}u} \in W^{\perp}$, т.е., что $(\boldsymbol{\mathcal{A}u}, \boldsymbol{w}) = 0$ для любого $\boldsymbol{w} \in W$. Мы знаем, что $\mathcal{A}(W) \subset W$. Поскольку оператор $\mathcal{A}$ обратим, $\mathcal{A}(W) = W$. Тогда найдётся такой вектор $\boldsymbol{v} \in W$, что $\boldsymbol{w} = \boldsymbol{\mathcal{A}v}$, а значит $(\boldsymbol{\mathcal{A}u}, \boldsymbol{w}) = (\boldsymbol{\mathcal{A}u}, \boldsymbol{\mathcal{A}v}) = (\boldsymbol{u}, \boldsymbol{v}) = 0$. □

Лемма 4.5.4. *Собственные значения ортогонального (унитарного) оператора $\mathcal{A}$ по модулю равны единице.*

Доказательство. Действительно, пусть $\boldsymbol{\mathcal{A}v} = \lambda \boldsymbol{v}$ для $\boldsymbol{v} \neq \mathbf{0}$. Тогда

$$(\boldsymbol{v}, \boldsymbol{v}) = (\boldsymbol{\mathcal{A}v}, \boldsymbol{\mathcal{A}v}) = (\lambda \boldsymbol{v}, \lambda \boldsymbol{v}) = \overline{\lambda}\lambda(\boldsymbol{v}, \boldsymbol{v}) = |\lambda|^2(\boldsymbol{v}, \boldsymbol{v}),$$

откуда $|\lambda|^2 = 1$. □

Теорема 4.5.5. *Для унитарного оператора $\mathcal{A}$ существует ортонормированный базис, в котором его матрица диагональна, а все диагональные элементы по модулю равны единице.*

Доказательство. Полностью аналогично доказательству теорем 4.2.3 и 4.4.3 для самосопряжённых или косоэрмитовых операторов. Шаг индукции проводим, выбирая собственный вектор $\boldsymbol{v}$ и устанавливая инвариантность подпространства $\langle \boldsymbol{v} \rangle^{\perp}$ при помощи леммы 4.5.3. □

Теорема 4.5.6. *Для ортогонального оператора $\mathcal{A}$ существует ортонормированный базис, в котором его матрица блочно-диагональная с блоками размера 1 или 2, причём блоки размера 1 имеют вид* (1) *или* (−1)*, а блоки размера 2 имеют вид* $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$*, где $\varphi \neq \pi k$ с целым $k$.*

Доказательство. В пространстве размерности 1 ортогональная матрица и так имеет вид (1) или (−1). В пространстве размерности 2 любая ортогональная матрица имеет вид $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$ (если определитель равен 1) или $\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$ (если определитель равен −1). В первом случае мы уже имеем требуемый вид (а оператор представляет собой поворот на угол $\varphi$ в положительном направлении). Во втором случае оператор представляет собой симметрию относительно прямой под углом $\frac{\varphi}{2}$ к оси абсцисс. Такой оператор имеет два ортогональных собственных вектора: $(\cos\frac{\varphi}{2}, \sin\frac{\varphi}{2})$ (вектор вдоль оси симметрии) и $(-\sin\frac{\varphi}{2}, \cos\frac{\varphi}{2})$ (вектор, перпендикулярный оси симметрии). В ортонормированном базисе из этих собственных векторов оператор $\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$ принимает требуемый вид $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Далее действуем по индукции, как и при доказательстве теоремы 4.4.4 для кососимметрических операторов. Предположим, что утверждение доказано для операторов в пространствах размерности не больше $n-1$ и докажем его для пространства $V$ размерности $n$ (где $n \geqslant 3$).

В силу теоремы 2.6.3 б) для оператора $\mathcal{A}$ существует 1-мерное или 2-мерное инвариантное подпространство $W \subset V$. Из леммы 4.5.3 следует, что ортогональное дополнение $W^{\perp}$ также инвариантно.

По предположению индукции, в пространстве $W^{\perp}$ имеется требуемый базис для ортогонального оператора $\mathcal{A}|_{W^{\perp}}$. Выбрав ортонормированный базис в пространстве $W$, как описано в начале доказательства, и взяв объединение базисов в $W^{\perp}$ и $W$, мы получим требуемый ортнормированный базис пространства $V$. □

Базисы, описанные в теоремах 4.5.5 и 4.5.6, называются *каноническими*, а соответствующие им матрицы — *каноническим видом* унитарного или ортогонального оператора.

Пример 4.5.7. 1. Как указано в доказательстве предыдущей теоремы, ортогональный оператор с матрицей $\begin{pmatrix}\cos\varphi & \sin\varphi\\ \sin\varphi & -\cos\varphi\end{pmatrix}$ имеет канонический вид $\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$. Тот же канонический вид будет, если оператор рассматривать как унитарный.

2. Канонический вид ортогонального оператора с матрицей $\begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}$ — это та же самая матрица. В то же время канонический вид этого оператора, рассматриваемого как унитарный оператор, есть $\begin{pmatrix}e^{i\varphi} & 0\\ 0 & e^{-i\varphi}\end{pmatrix}$.

3. В трёхмерном пространстве канонический вид ортогонального оператора есть

$$\begin{pmatrix}\cos\varphi & -\sin\varphi & 0\\ \sin\varphi & \cos\varphi & 0\\ 0 & 0 & \pm1\end{pmatrix}$$

где в левом нижнем углу стоит 1 или $-1$ в зависимости от знака определителя оператора. (Операторы, канонический вид которых имеет три блока $(1)$ или $(-1)$, получаются при $\varphi = \pi k$.) Если определитель положителен, то такой оператор представляет собой поворот (вокруг оси третьего вектора канонического базиса). Если же определитель отрицателен, то оператор — это «поворот с переворотом», т.е. композиция поворота и симметрии относительно плоскости, перпендикулярной оси поворота.

Отсюда, в частности, следует, что композиция двух поворотов — это снова поворот вокруг некоторой оси (так как в каноническом виде всегда происходит всего один поворот).

4. В четырёхмерном пространстве уже бывают «независимые повороты». А именно, канонический вид ортогонального оператора общего вида с положительным определителем представляет собой матрицу из двух блоков размера 2:

$$\begin{pmatrix}\cos\varphi & -\sin\varphi & 0 & 0\\ \sin\varphi & \cos\varphi & 0 & 0\\ 0 & 0 & \cos\psi & -\sin\psi\\ 0 & 0 & \sin\psi & \cos\psi\end{pmatrix}$$

Это — композиция двух независимых поворотов: на угол $\varphi$ в плоскости первого и второго базисных векторов и на угол $\psi$ в плоскости третьего и четвёртого базисных векторов. Такой оператор не сводится к одному повороту.

Произведение ортогональных операторов очевидно является ортогональным оператором, и поэтому ортогональные операторы в евклидовом пространстве $V$ образуют подгруппу в общей линейной группе $GL(V)$. Эта подгруппа называется *ортогональной группой* и обозначается $O(V)$. Если $\dim V = n$, то группа $O(V)$ изоморфна группе ортогональных матриц размера $n$; эта группа обозначается $O_n$.

Аналогично, унитарные операторы в эрмитовом пространстве $V$ образуют *унитарную группу*, которая обозначается $U(V)$. Если $\dim V = n$, то группа $U(V)$ изоморфна группе унитарных матриц размера $n$; эта группа обозначается $U_n$.

Наконец, ортогональные (унитарные) матрицы с определителем 1 образуют подгруппу в $O_n$ (соответственно, в $U_n$). Эта подгруппа называется *специальной ортогональной группой* (соответственно, *специальной унитарной группой*) и обозначается $SO_n$ (соответственно, $SU_n$).

## 4.6. Положительные самосопряжённые операторы. Полярное разложение

Определение 4.6.1. Самосопряжённый оператор $\mathcal{A}$ называется *положительным*, если $(\mathcal{A}\boldsymbol{v}, \boldsymbol{v}) > 0$ для любого ненулевого вектора $\boldsymbol{v} \in V$.

Лемма 4.6.2. *Самосопряжённый оператора $\mathcal{A}$ положителен тогда и только тогда, когда все его собственные значения положительны.*

Доказательство. Пусть $(\mathcal{A}\boldsymbol{v}, \boldsymbol{v}) > 0$ при $\boldsymbol{v} \neq \boldsymbol{0}$. Рассмотрим собственный вектор $\boldsymbol{v}$ с собственным значением $\lambda$. Тогда $\lambda(\boldsymbol{v}, \boldsymbol{v}) = (\mathcal{A}\boldsymbol{v}, \boldsymbol{v}) > 0$, откуда $\lambda > 0$.

Обратно, пусть все собственные значения $\lambda_i$ положительны. Выберем канонический ортонормированный базис из собственных векторов $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ с собственными значениями $\lambda_1, \ldots, \lambda_n$. Тогда для ненулевого вектора $\boldsymbol{v} = v^i \boldsymbol{e}_i$ мы имеем

$$(\mathcal{A}\boldsymbol{v}, \boldsymbol{v}) = \big(\mathcal{A}(v^i \boldsymbol{e}_i), v^j \boldsymbol{e}_j\big) = \overline{v}^i v^j (\mathcal{A}\boldsymbol{e}_i, \boldsymbol{e}_j) = \sum_{i=1}^{n} \lambda_i |v^i|^2 > 0. \qquad \square$$

Теорема 4.6.3. *Для положительного оператора $\mathcal{A}$ существует положительный оператор $\mathcal{P}$, удовлетворяющий соотношению $\mathcal{P}^2 = \mathcal{A}$.*

Доказательство. Выберем канонический базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ для $\mathcal{A}$, в котором его матрица диагональна с положительными числами $\lambda_1, \ldots, \lambda_n$ на диагонали. Рассмотрим оператор $\mathcal{P}$ задаваемый в том же базисе диагональной матрицей с числами $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}$ на диагонали. Тогда, очевидно, $\mathcal{P}^2 = \mathcal{A}$. Кроме того, оператор $\mathcal{P}$ самосопряжён (так как он имеет симметричную матрицу в ортонормированном базисе) и положителен в силу леммы 4.6.2. $\square$

Оператор $\mathcal{P}$, построенный в предыдущей теореме, называется, *положительным корнем* из положительного оператора $\mathcal{A}$ и обозначается $\sqrt{\mathcal{A}}$.

Самосопряжённый оператор $\mathcal{A}$ называется *неотрицательным*, если $(\mathcal{A}\boldsymbol{v}, \boldsymbol{v}) \geqslant 0$ для любого вектора $\boldsymbol{v} \in V$. Все утверждения выше переносятся без изменений на неотрицательные операторы.

Теорема 4.6.4 (полярное разложение). *Для любого невырожденного оператора $\mathcal{A}$ в евклидовом или эрмитовом пространстве существует представление в виде*

$$\mathcal{A} = \mathcal{P}\mathcal{U},$$

*где $\mathcal{P}$ — положительный, а $\mathcal{U}$ — ортогональный (унитарный) оператор.*

Доказательство. Если $\mathcal{A} = \mathcal{P}\mathcal{U}$, то $\mathcal{A}^* = \mathcal{U}^*\mathcal{P}$ и $\mathcal{A}\mathcal{A}^* = \mathcal{P}\mathcal{U}\mathcal{U}^*\mathcal{P} = \mathcal{P}^2$. Поэтому, если бы оператор $\mathcal{A}\mathcal{A}^*$ был положительным, можно было бы взять $\mathcal{P} = \sqrt{\mathcal{A}\mathcal{A}^*}$.

Оператор $\mathcal{A}\mathcal{A}^*$ действительно является самосопряжённым и положительным:

$$(\mathcal{A}\mathcal{A}^*\boldsymbol{v}, \boldsymbol{v}) = (\mathcal{A}^*\boldsymbol{v}, \mathcal{A}^*\boldsymbol{v}) > 0,$$

при $\boldsymbol{v} \neq \boldsymbol{0}$, так как $\mathcal{A}^*\boldsymbol{v} \neq \boldsymbol{0}$ в силу невырожденности $\mathcal{A}$. Теперь положим $\mathcal{P} = \sqrt{\mathcal{A}\mathcal{A}^*}$ и $\mathcal{U} = \mathcal{P}^{-1}\mathcal{A}$. Тогда $\mathcal{U}\mathcal{U}^* = \mathcal{P}^{-1}\mathcal{A}\mathcal{A}^*\mathcal{P}^{-1} = \mathcal{P}^{-1}\mathcal{P}^2\mathcal{P}^{-1} = \mathrm{id}$. Следовательно, $\mathcal{U}$ — ортогональный (унитарный) оператор и $\mathcal{A} = \mathcal{P}\mathcal{U}$. $\square$

Аналогично, рассмотрев положительный оператор $\mathcal{A}^*\mathcal{A}$, можно доказать существование второго полярного разложения $\mathcal{A} = \mathcal{U}'\mathcal{P}'$ (где $\mathcal{P}' = \sqrt{\mathcal{A}^*\mathcal{A}}$).

Замечание. На самом деле положительный корень из положительного оператора единствен (задача), поэтому полярное разложение также единственно.

Для произвольного оператора $\mathcal{A}$ существует полярное разложение $\mathcal{A} = \mathcal{P}\mathcal{U}$, где $\mathcal{P}$ — неотрицательный оператор, а $\mathcal{U}$ — ортогональный или унитарный оператор (задача). Однако такое разложение не единственно.

В одномерном эрмитовом пространстве $\mathbb{C}$ положительные операторы — это положительные вещественные числа, а унитарные операторы — это комплексные числа, по модулю равные 1, т.е. вида $e^{i\varphi}$. Поэтому полярное разложение — это представление комплексного числа $z$ в полярных координатах: $z = \rho e^{i\varphi}$, что объясняет название.

## 4.7. Нормальные операторы

Определение 4.7.1. Оператор $\mathcal{A}$ в евклидовом или эрмитовом пространстве называется *нормальным*, если он коммутирует с сопряжённым, т.е. $\mathcal{A}^*\mathcal{A} = \mathcal{A}\mathcal{A}^*$.

Все специальные классы операторов, рассмотренные выше в этой главе (от самосопряжённых до унитарных) являются нормальными.

Лемма 4.7.2. *Пусть $\boldsymbol{v}$ — собственный вектор нормального оператора $\mathcal{A}$ с собственным значением $\lambda$. Тогда $\boldsymbol{v}$ также является собственным вектором для сопряжённого оператора $\mathcal{A}^*$, с собственным значением $\overline{\lambda}$.*

Доказательство. Если оператор $\mathcal{A}$ нормален, то

$$(\mathcal{A}\boldsymbol{v}, \mathcal{A}\boldsymbol{v}) = (\mathcal{A}^*\mathcal{A}\boldsymbol{v}, \boldsymbol{v}) = (\mathcal{A}\mathcal{A}^*\boldsymbol{v}, \boldsymbol{v}) = (\mathcal{A}^*\boldsymbol{v}, \mathcal{A}^*\boldsymbol{v}),$$

т.е. $|\mathcal{A}\boldsymbol{v}| = |\mathcal{A}^*\boldsymbol{v}|$ для любого вектора $\boldsymbol{v}$. Поскольку вместе с оператором $\mathcal{A}$ нормален и каждый оператор вида $\mathcal{A} - \lambda\,\mathrm{id}$, отсюда следует, что

$$|(\mathcal{A} - \lambda\,\mathrm{id})\boldsymbol{v}| = |(\mathcal{A}^* - \overline{\lambda}\,\mathrm{id})\boldsymbol{v}|$$

для любого $\lambda$. Поэтому, если $(\mathcal{A} - \lambda\,\mathrm{id})\boldsymbol{v} = \boldsymbol{0}$, то $(\mathcal{A}^* - \overline{\lambda}\,\mathrm{id})\boldsymbol{v} = \boldsymbol{0}$. □

В эрмитовом пространстве класс нормальных операторов — это в точности операторы, диагонализируемые в ортонормированном базисе:

Теорема 4.7.3. *Оператор в эрмитовом пространстве диагонализируем в ортонормированном базисе тогда и только тогда, когда он нормален.*

Доказательство. Пусть матрица оператора $\mathcal{A}$ в некотором ортонормированном базисе диагональна с числами $\lambda_1, \ldots, \lambda_n$ на диагонали. Тогда сопряжённый оператор $\mathcal{A}^*$ в том же базисе имеет диагональную матрицу с числами $\overline{\lambda}_1, \ldots, \overline{\lambda}_n$. Так как диагональные матрицы коммутируют, мы получаем $\mathcal{A}^*\mathcal{A} = \mathcal{A}\mathcal{A}^*$, т.е. $\mathcal{A}$ нормален.

Обратно, пусть $\mathcal{A}^*\mathcal{A} = \mathcal{A}\mathcal{A}^*$. Доказательство диагонализируемости в ортнормированном базисе будем вести по индукции по размерности пространства $V$. При $\dim V = 1$ доказывать нечего. Предположим, что утверждение доказано для пространств размерности не больше $n-1$ и докажем его для размерности $n$.

Выберем собственный вектор $\boldsymbol{v}$ с собственным значением $\lambda$ для $\mathcal{A}$, т.е. одномерное инвариантное подпространство $W = \langle \boldsymbol{v} \rangle$. Докажем, что ортогональное дополнение $W^\perp$ также инвариантно относительно $\mathcal{A}$. Пусть $\boldsymbol{u} \in W^\perp$, т.е. $(\boldsymbol{u}, \boldsymbol{v}) = 0$. Тогда

$$(\mathcal{A}\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \mathcal{A}^*\boldsymbol{v}) = (\boldsymbol{u}, \overline{\lambda}\boldsymbol{v}) = \overline{\lambda}(\boldsymbol{u}, \boldsymbol{v}) = 0$$

(где мы воспользовались леммой 4.7.2). Следовательно, $\mathcal{A}\boldsymbol{u} \in W^\perp$ и пространство $W^\perp$ инвариантно относительно оператора $\mathcal{A}$.

С другой стороны, пространство $W^\perp$ инвариантно относительно оператора $\mathcal{A}^*$ в силу леммы 4.1.4. Так как $W^\perp$ инвариантно и относительно $\mathcal{A}$ и относительно $\mathcal{A}^*$, ограничение $\mathcal{A}|_{W^\perp}$ является нормальным оператором. Так как $\dim W^\perp = n-1$, по предположению индукции в пространстве $W^\perp$ имеется ортонормированный базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{n-1}$ из собственных векторов оператора $\mathcal{A}|_{W^\perp}$. Тогда $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{n-1}, \frac{\boldsymbol{v}}{|\boldsymbol{v}|}$ — ортонормированный базис из собственных векторов оператора $\mathcal{A}$. $\square$

В евклидовом пространстве аналог этой теоремы не имеет места: оператор $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ нормален, но не диагонализируем при $b \neq 0$. На самом деле мы знаем из раздела 4.2, что в евклидовом пространстве класс операторов, диагонализируемых в ортонормированном базисе, — это в точности самосопряжённые операторы.

Замечание. Можно доказать, что для нормального оператора в евклидовом пространстве существует ортонормированный базис, в котором его матрица состоит из блоков размера 1 или 2, причём блоки размера 2 имеют вид $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ (задача).

Соберём в одну таблицу всю информацию о специальных классах операторов из этой главы (отдельно для евклидова и эрмитова пространства):

| Название | Определение | Свойство матрицы (в ортонормиров. базисе) | Канонический вид |
|---|---|---|---|
| Самосопряжённый (симметрический) | $\mathcal{A}^* = \mathcal{A}$ | $A^t = A$ (симметричная) | диагональный |
| Кососимметрический | $\mathcal{A}^* = -\mathcal{A}$ | $A^t = -A$ (кососимметр.) | блоки $(0)$ и $\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}$ |
| Ортогональный | $\mathcal{A}^*\mathcal{A} = \mathrm{id}$ | $A^tA = E$ (ортогональная) | блоки $(\pm 1)$ и $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$ |
| Нормальный | $\mathcal{A}^*\mathcal{A} = \mathcal{A}\mathcal{A}^*$ | $A^tA = AA^t$ | блоки $(a)$ и $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ |
| Самосопряжённый (эрмитов) | $\mathcal{A}^* = \mathcal{A}$ | $\overline{A}^t = A$ (эрмитова) | диагональный с вещественными числами на диагонали |
| Косоэрмитов | $\mathcal{A}^* = -\mathcal{A}$ | $\overline{A}^t = -A$ (косоэрмитова) | диагональный с чисто мнимыми числами на диагонали |
| Унитарный | $\mathcal{A}^*\mathcal{A} = \mathrm{id}$ | $\overline{A}^tA = E$ (унитарная) | на диагонали — числа, равные по модулю 1 |
| Нормальный | $\mathcal{A}^*\mathcal{A} = \mathcal{A}\mathcal{A}^*$ | $\overline{A}^tA = A\overline{A}^t$ | диагональный |

Глава 5

# Билинейные и полуторалинейные функции

## 5.1. Билинейные и полуторалинейные функции, их матрицы. Закон изменения матрицы при замене базиса. Канонический изоморфизм пространства билинейных функций и пространства $\mathrm{Hom}(V, V^*)$

Определение 5.1.1. Пусть $V$ — линейное пространство над полем $\mathbf{k}$. Функция $\mathcal{B}\colon V \times V \to \mathbf{k}$ называется *билинейной функцией*, если она линейна по каждому аргументу:

$$\mathcal{B}(\lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2, \boldsymbol{y}) = \lambda_1\mathcal{B}(\boldsymbol{x}_1, \boldsymbol{y}) + \lambda_2\mathcal{B}(\boldsymbol{x}_2, \boldsymbol{y}) \quad \text{и}$$
$$\mathcal{B}(\boldsymbol{x}, \mu_1\boldsymbol{y}_1 + \mu_2\boldsymbol{y}_2) = \mu_1\mathcal{B}(\boldsymbol{x}, \boldsymbol{y}_1) + \mu_2\mathcal{B}(\boldsymbol{x}, \boldsymbol{y}_2)$$

для любых $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbf{k}$ и $\boldsymbol{x}, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}, \boldsymbol{y}_1, \boldsymbol{y}_2 \in V$.

*Матрицей билинейной функции* $\mathcal{B}$ в базисе $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ пространства $V$ называется квадратная матрица $B = (b_{ij})$ размера $n$, где $b_{ij} = \mathcal{B}(\boldsymbol{e}_i, \boldsymbol{e}_j)$.

Пример 5.1.2. Скалярное произведение в евклидовом пространстве является билинейной функцией. Таким образом, понятие билинейной функции обобщает понятие скалярного произведения (вместо трёх условий на функцию $V \times V \to \mathbb{R}$ требуется выполнение лишь первого, т.е. билинейности). Матрица билинейной функции является обобщением матрицы Грама скалярного произведения.

Зная матрицу $B = (b_{ij})$ билинейной функции, можно восстановить значение $\mathcal{B}(\boldsymbol{x}, \boldsymbol{y})$ на любой паре векторов $\boldsymbol{x} = x^i\boldsymbol{e}_i$ и $\boldsymbol{y} = y^j\boldsymbol{e}_j$:

$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{y}) = \mathcal{B}(x^i\boldsymbol{e}_i, y^j\boldsymbol{e}_j) = x^iy^j\mathcal{B}(\boldsymbol{e}_i, \boldsymbol{e}_j) = b_{ij}x^iy^j = x^tBy.$$

Выражение $B(x, y) = b_{ij}x^iy^j = x^tBy$ называется *билинейной формой* (здесь, как обычно, мы предполагаем, что $x = (x^1, \ldots, x^n)^t$ — это столбец высоты $n$, так что $x^t$ — это строка длины $n$). Билинейная форма представляет собой однородный многочлен степени 2 от двух наборов переменных $x^1, \ldots, x^n$ и $y^1, \ldots, y^n$, который линеен по $x$ при фиксированных $y$ и линеен по $y$ при фиксированных $x$.

Теорема 5.1.3 (закон изменения матрицы билинейной функции). *Имеет место соотношение*

$$B' = C^tBC,$$

*где $B$ — матрица билинейной функции $\mathcal{B}\colon V \times V \to \mathbf{k}$ в базисе $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$, $B'$ — матрица в базисе $\boldsymbol{e}_{1'}, \ldots, \boldsymbol{e}_{n'}$ и $C$ — матрица перехода от базиса $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ к базису $\boldsymbol{e}_{1'}, \ldots, \boldsymbol{e}_{n'}$.*

Доказательство. Пусть $B = (b_{ij})$, $B' = (b'_{ij})$, $C = (c^i_{i'})$. Мы имеем

$$b'_{ij} = \mathcal{B}(\boldsymbol{e}_{i'}, \boldsymbol{e}_{j'}) = \mathcal{B}(c^i_{i'}\boldsymbol{e}_i, c^j_{j'}\boldsymbol{e}_j) = c^i_{i'}c^j_{j'}\mathcal{B}(\boldsymbol{e}_i, \boldsymbol{e}_j) = c^i_{i'}b_{ij}c^j_{j'},$$

что эквивалентно матричному соотношению $B' = C^tBC$.

То же рассуждение можно провести, используя матричную запись билинейных форм. Значение $\mathcal{B}(\boldsymbol{x},\boldsymbol{y})$ можно записать в виде билинейных форм от новых или старых координат:

$$\mathcal{B}(\boldsymbol{x},\boldsymbol{y}) = B(x,y) = B(x',y').$$

Мы имеем $x = Cx'$ и $y = Cy'$. Тогда

$$(x')^t B' y' = B(x',y') = B(x,y) = x^t B y = (x')^t C^t B C y'.$$

Так как это верно для любых наборов $x', y'$, получаем $B' = C^t B C$. $\square$

Следствие 5.1.4. *Ранг матрицы билинейной функции не зависит от базиса.*

Доказательство. Так как матрица $C$ обратима, $\operatorname{rk} B' = \operatorname{rk}(C^t B C) = \operatorname{rk} B$. $\square$

Определение 5.1.5. *Рангом* билинейной функции $\mathcal{B}$ (обозначается $\operatorname{rk}\mathcal{B}$) называется ранг её матрицы в произвольном базисе. Билинейная функция $\mathcal{B}$ в пространстве $V$ называется *невырожденной*, если $\operatorname{rk}\mathcal{B} = \dim V$.

Множество $\mathrm{B}(V)$ всех билинейных функций в пространстве $V$ образует линейное пространство относительно операций сложения функций и умножения функций на скаляры. Сопоставление билинейной функции $\mathcal{B}$ её матрицы $B$ в фиксированном базисе $\boldsymbol{e}_1,\dots,\boldsymbol{e}_n$ устанавливает изоморфизм между пространством $\mathrm{B}(V)$ и пространством квадратных матриц $\mathrm{Mat}_{\mathbf{k}}(n)$. Как и в случае пространства линейных операторов $\mathrm{Hom}(V,V)$, этот изоморфизм неканоничен, так он зависит от выбора базиса.

Наряду с $\mathrm{B}(V)$ рассмотрим пространство $\mathrm{Hom}(V,V^*)$ линейных отображений из $V$ в двойственное пространство $V^*$.

Теорема 5.1.6. *Отображение $\varphi\colon \mathrm{B}(V)\to\mathrm{Hom}(V,V^*)$, сопоставляющее билинейной функции $\mathcal{B}$ линейное отображение $\widetilde{\mathcal{B}}\colon V\to V^*$, задаваемое формулой*

$$\widetilde{\mathcal{B}}(\boldsymbol{x}) = \mathcal{B}(\boldsymbol{x},\,\cdot\,) \quad \text{для } \boldsymbol{x}\in V,$$

*является каноническим изоморфизмом. (Здесь $\mathcal{B}(\boldsymbol{x},\,\cdot\,)\in V^*$ — линейная функция, значение которой на векторе $\boldsymbol{y}\in V$ есть $\mathcal{B}(\boldsymbol{x},\boldsymbol{y})$.)*

Доказательство. Отображение $\varphi$ линейно, так как билинейная функция линейна по первому аргументу $\boldsymbol{x}$. Кроме того, отображение $\varphi$ биективно: обратное отображение $\varphi^{-1}$ ставит в соответствие линейному отображению $\widetilde{\mathcal{B}}\colon V\to V^*$ билинейную функцию $\mathcal{B}$, заданную по формуле $\mathcal{B}(\boldsymbol{x},\boldsymbol{y}) = \widetilde{\mathcal{B}}(\boldsymbol{x})(\boldsymbol{y})$. Следовательно, $\varphi$ — изоморфизм. Этот изоморфизм каноничен, так как в его конструкции не использовался базис. $\square$

В комплексном пространстве $V$ наряду с билинейными функциями рассматриваются полуторалинейные:

Определение 5.1.7. Пусть $V$ — линейное пространство над полем $\mathbb{C}$. Функция $\mathcal{S}\colon V\times V\to\mathbb{C}$ называется *полуторалинейной функцией*, если она линейна по второму аргументу и антилинейна (или *полулинейна*) по первому каждому аргументу:

$$\mathcal{S}(\lambda_1\boldsymbol{x}_1+\lambda_2\boldsymbol{x}_2,\boldsymbol{y}) = \overline{\lambda}_1\mathcal{S}(\boldsymbol{x}_1,\boldsymbol{y})+\overline{\lambda}_2\mathcal{S}(\boldsymbol{x}_2,\boldsymbol{y}) \quad \text{и}$$
$$\mathcal{S}(\boldsymbol{x},\mu_1\boldsymbol{y}_1+\mu_2\boldsymbol{y}_2) = \mu_1\mathcal{S}(\boldsymbol{x},\boldsymbol{y}_1)+\mu_2\mathcal{S}(\boldsymbol{x},\boldsymbol{y}_2)$$

для любых $\lambda_1,\lambda_2,\mu_1,\mu_2\in\mathbb{C}$ и $\boldsymbol{x},\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y},\boldsymbol{y}_1,\boldsymbol{y}_2\in V$.

*Матрица полуторалинейной функции $\mathcal{S}$* в базисе $\boldsymbol{e}_1,\dots,\boldsymbol{e}_n$ определяется как $S = (s_{ij})$, где $s_{ij} = \mathcal{S}(\boldsymbol{e}_i,\boldsymbol{e}_j)$.

Значение $\mathcal{S}(\boldsymbol{x}, \boldsymbol{y})$ выражается через матрицу $S = (s_{ij})$ и координаты векторов $\boldsymbol{x} = x^i \boldsymbol{e}_i$ и $\boldsymbol{y} = y^j \boldsymbol{e}_j$ следующим образом:

$$\mathcal{S}(\boldsymbol{x}, \boldsymbol{y}) = \mathcal{S}(x^i \boldsymbol{e}_i, y^j \boldsymbol{e}_j) = \overline{x}^i y^j \mathcal{S}(\boldsymbol{e}_i, \boldsymbol{e}_j) = s_{ij} \overline{x}^i y^j = \overline{x}^t S y.$$

Выражение $S(x, y) = s_{ij} \overline{x}^i y^j = \overline{x}^t S y$ называется *полуторалинейной формой.*

Пример полуторалинейной функции — эрмитово скалярное произведение.

Матрицы $S$ и $S'$ полуторалинейной функции $\mathcal{S}$ в разных базисах связаны соотношением $S' = \overline{C}^t S C$, которое доказывается аналогично соотношению для билинейных функций. Отсюда следует, что ранг матрицы полуторалинейной функции не зависит от выбора базиса.

## 5.2. Симметрические, кососимметрические и эрмитовы функции. Квадратичные формы. Нормальный вид

Определение 5.2.1. Билинейная функция $\mathcal{B}\colon V \times V \to \mathbf{k}$ называется *симметрической*, если $\mathcal{B}(\boldsymbol{y}, \boldsymbol{x}) = \mathcal{B}(\boldsymbol{x}, \boldsymbol{y})$, и *кососимметрической*, если $\mathcal{B}(\boldsymbol{y}, \boldsymbol{x}) = -\mathcal{B}(\boldsymbol{x}, \boldsymbol{y})$, для любых $\boldsymbol{x}, \boldsymbol{y} \in V$.

Полуторалинейная функция $\mathcal{S}\colon V \times V \to \mathbb{C}$ в комплексном пространстве называется *эрмитовой*, если $\mathcal{S}(\boldsymbol{y}, \boldsymbol{x}) = \overline{\mathcal{S}(\boldsymbol{x}, \boldsymbol{y})}$, и *косоэрмитовой*, если $\mathcal{S}(\boldsymbol{y}, \boldsymbol{x}) = -\overline{\mathcal{S}(\boldsymbol{x}, \boldsymbol{y})}$.

Матрица симметрической (кососимметрической) билинейной функции в любом базисе симметрична (соответственно, кососимметрична). Матрица эрмитовой (косоэрмитовой) полуторалинейной функции в любом базисе эрмитова (соответственно, косоэрмитова). Кроме того, полуторалинейная функция $\mathcal{S}$ является эрмитовой тогда и только тогда, когда функция $i\mathcal{S}$ является косоэрмитовой.

Далее мы будем предполагать, что характеристика поля $\mathbf{k}$ отлична от 2.

Определение 5.2.2. *Квадратичной формой* над $\mathbf{k}$ называется однородный многочлен второй степени от $n$ переменных $x = (x^1, \ldots, x^n)$, т.е. многочлен вида

$$Q(x) = Q(x^1, \ldots, x^n) = q_{ij} x^i x^j = \sum_{i=1}^{n} q_{ii} (x^i)^2 + \sum_{i<j} 2 q_{ij} x^i x^j,$$

где $q_{ji} = q_{ij} \in \mathbf{k}$. Симметричная матрица $Q = (q_{ij})$ размера $n \times n$ называется *матрицей квадратичной формы.*

Если $B(x, y) = b_{ij} x^i y^j$ — симметрическая билинейная форма, то $B(x, x) = b_{ij} x^i x^j$ является квадратичной формой с матрицей $B$. Таким образом, квадратичная форма $B(x, x)$ полностью определяет симметрическую билинейную форму $B(x, y)$, а значит и симметрическую билинейную функцию $\mathcal{B}(\boldsymbol{x}, \boldsymbol{y})$. Это можно увидеть и не прибегая к выбору базиса: для симметрической билинейной функции имеет место соотношение

$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2}\bigl(\mathcal{B}(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y}) - \mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) - \mathcal{B}(\boldsymbol{y}, \boldsymbol{y})\bigr),$$

т.е. значение $\mathcal{B}$ на произвольной паре векторов можно восстановить, зная лишь значения $\mathcal{B}$ на парах совпадающих векторов. Функцию $V \to \mathbf{k}$, $\boldsymbol{x} \mapsto \mathcal{B}(\boldsymbol{x}, \boldsymbol{x})$ называют *квадратичной функцией.*

Теорема 5.2.3. *Для симметрической билинейной функции $\mathcal{B}$ над полем характеристики, отличной от 2, существует базис, в котором её матрица диагональна.*

*Другими словами, любую квадратичную форму $Q(x)$ линейной заменой координат $x = Cy$ можно привести к виду*

$$Q(y) = r_{11}(y^1)^2 + \ldots + r_{nn}(y^n)^2.$$

Мы приведём два доказательства этого факта. В первом случае будем работать с квадратичными формами и координатами, а во втором — с симметрическими билинейными функциями и базисами. Каждое из доказательств будет проведено таким образом, что его можно будет использовать как алгоритм, а не просто как доказательство существования нужного преобразования.

Первое доказательство (метод Лагранжа). Пусть $Q(x) = q_{ij}x^i x^j$ — квадратичная форма. Доказательство заключается в последовательном упрощении $Q(x)$, использующем основное и два вспомогательных преобразования.

*Основное преобразование* производится, если в квадратичной форме $Q(x) = q_{ij}x^i x^j$ первый коэффициент $q_{11}$ не равен нулю. Тогда имеем

$$Q(x^1, \ldots, x^n) = q_{11}(x^1)^2 + 2q_{12}x^1x^2 + \ldots + 2q_{1n}x^1x^n + \sum_{i,j>1} q_{ij}x^ix^j =$$
$$= q_{11}\Bigl(x^1 + \frac{q_{12}}{q_{11}}x^2 + \ldots + \frac{q_{1n}}{q_{11}}x^n\Bigr)^2 - q_{11}\Bigl(\frac{q_{12}}{q_{11}}x^2 + \ldots + \frac{q_{1n}}{q_{11}}x^n\Bigr)^2 + \sum_{i,j>1} q_{ij}x^ix^j =$$
$$= q_{11}\Bigl(x^1 + \frac{q_{12}}{q_{11}}x^2 + \ldots + \frac{q_{1n}}{q_{11}}x^n\Bigr)^2 + Q'(x^2, \ldots, x^n),$$

где $Q'(x^2, \ldots, x^n)$ — некоторая квадратичная форма от $n-1$ переменных. Теперь сделаем замену координат

$$u^1 = x^1 + \frac{q_{12}}{q_{11}}x^2 + \ldots + \frac{q_{1n}}{q_{11}}x^n,$$
$$u^2 = x^2, \quad \ldots, \quad u^n = x^n.$$

В результате форма $Q(x)$ преобразуется к виду

$$Q(u^1, \ldots, u^n) = q_{11}(u^1)^2 + Q'(u^2, \ldots, u^n).$$

Если в форме $Q'(u^2, \ldots, u^n)$ первый коэффициент (т.е. $q'_{22}$) не равен нулю, то мы снова можем применить основное преобразование, и т.д.

*Первое вспомогательное преобразование* производится, если $q_{11} = 0$, но существует $q_{ii} \neq 0$. В этом случае мы делаем замену $u^1 = x^i$, $u^i = x^1$, а остальные координаты без изменений. В результате получаем $q'_{11} \neq 0$.

*Второе вспомогательное преобразование* производится, если все коэффициенты $q_{ii}$ при квадратах равны нулю, но при этом есть хотя бы один ненулевой коэффициент (в противном случае $Q(x) \equiv 0$ уже имеет нужный вид). Пусть $q_{ij} \neq 0$, где $i < j$. Произведём замену координат

$$x^i = u^i, \quad x^j = u^i + u^j, \quad x^k = u^k, \quad \text{при } k \neq i, j.$$

В результате форма $Q(x)$ преобразуется к виду

$$Q(x) = 2q_{ij}x^ix^j + \ldots = 2q_{ij}u^i(u^i + u^j) + \ldots = 2q_{ij}(u^i)^2 + \ldots,$$

где $\ldots$ означает члены, не содержащие квадратов. Далее мы можем применить предыдущие преобразования.

Последовательно применяя основное преобразование и (если нужно) вспомогательные преобразования, мы приводим форму $Q(x)$ к диагональному виду. □

Второе доказательство (метод поиска базиса). Этот метод можно рассматривать как обобщение метода ортогонализации Грама–Шмидта. Базис, в котором матрица билинейной функции $\mathcal{B}$ диагональна — это «ортогональный» базис в смысле «скалярного произведения», задаваемого симметрической билинейной функцией $\mathcal{B}$. Здесь также имеется основное и вспомогательные преобразования.

Пусть $B = (b_{ij})$ — матрица билинейной функции $\mathcal{B}$ в исходном базисе $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$.

Основное преобразование производится, если $b_{11} = \mathcal{B}(\boldsymbol{e}_1, \boldsymbol{e}_1) \neq 0$ (это всегда так, если симметрическая билинейная функция $\mathcal{B}$ задаёт скалярное произведение, т.е. является положительно определённой). Выберем новый базис следующим образом:

$$
\begin{aligned}
&\boldsymbol{e}_{1'} = \boldsymbol{e}_1,\\
&\boldsymbol{e}_{2'} = \boldsymbol{e}_2 - \frac{\mathcal{B}(\boldsymbol{e}_1, \boldsymbol{e}_2)}{\mathcal{B}(\boldsymbol{e}_1, \boldsymbol{e}_1)}\boldsymbol{e}_1 = \boldsymbol{e}_2 - \frac{b_{12}}{b_{11}}\boldsymbol{e}_1,\\
&\quad\dots\\
&\boldsymbol{e}_{n'} = \boldsymbol{e}_n - \frac{\mathcal{B}(\boldsymbol{e}_1, \boldsymbol{e}_n)}{\mathcal{B}(\boldsymbol{e}_1, \boldsymbol{e}_1)}\boldsymbol{e}_1 = \boldsymbol{e}_n - \frac{b_{1n}}{b_{11}}\boldsymbol{e}_1.
\end{aligned}
\tag{5.1}
$$

В результате мы получаем $\mathcal{B}(\boldsymbol{e}_{1'}, \boldsymbol{e}_{i'}) = 0$ при $i > 1$. Таким образом, матрица $B'$ билинейной функции $\mathcal{B}$ в новом базисе принимает вид

$$
B' = \left(\begin{array}{c|ccc}
b_{11} & 0 & \cdots & 0 \\
\hline
0 & & & \\
\vdots & & \widetilde{B}' & \\
0 & & &
\end{array}\right)
$$

где $\widetilde{B}'$ — матрица размера $(n-1)\times(n-1)$ билинейной функции $\mathcal{B}$ на подпространстве $\langle \boldsymbol{e}_{2'}, \dots, \boldsymbol{e}_{n'} \rangle$. Далее мы работаем уже с этой матрицей $\widetilde{B}'$.

Первое вспомогательное преобразование производится, если $b_{11} = 0$, но имеется $b_{ii} \neq 0$. Тогда делаем замену, меняющую местами 1-й и $i$-й базисный векторы.

Второе вспомогательное преобразование производится, если все $b_{ii}$ равны нулю, но при этом билинейная функция $\mathcal{B}$ не является тождественно нулевой, т.е. $b_{ij} = \mathcal{B}(\boldsymbol{e}_i, \boldsymbol{e}_j) \neq 0$ для некоторых $i < j$. Произведём замену базиса

$$
\boldsymbol{e}_{i'} = \boldsymbol{e}_i + \boldsymbol{e}_j, \quad \boldsymbol{e}_{j'} = \boldsymbol{e}_j, \quad \boldsymbol{e}_{k'} = \boldsymbol{e}_k \quad \text{при } k \neq i, j.
$$

Тогда в новом базисе мы имеем

$$
b'_{ii} = \mathcal{B}(\boldsymbol{e}_{i'}, \boldsymbol{e}_{i'}) = \mathcal{B}(\boldsymbol{e}_i + \boldsymbol{e}_j, \boldsymbol{e}_i + \boldsymbol{e}_j) = 2\mathcal{B}(\boldsymbol{e}_i, \boldsymbol{e}_j) = 2b_{ij} \neq 0.
$$

Далее мы можем применить предыдущие преобразования.

Последовательно применяя основное преобразование и дополняя его в необходимых случаях вспомогательными преобразованиями, мы получаем базис $\boldsymbol{f}_1, \dots, \boldsymbol{f}_n$, в котором матрица билинейной функции $\mathcal{B}$ имеет диагональный вид. □

Обратим внимание, что основное и вспомогательное преобразование в обоих доказательствах — это одно и то же преобразование, просто в первом случае оно записано через координаты, а во втором — через базисы. Так что диагональные матрицы, получаемые первым и вторым методом, совпадают, как и все промежуточные матрицы.

Замечание. Если при приведении матрицы билинейной функции (квадратичной формы) к диагональному виду использовалось лишь основное преобразование, то матрица перехода от исходного базиса к базису, в котором матрица имеет диагональный вид, является верхнетреугольной (как и в случае процесса ортогонализации Грама–Шмидта). Если же хоть раз применялось вспомогательное преобразование, то матрица перехода может не быть верхнетреугольной.

Пример 5.2.4. Над полем $\mathbb{Z}_2$ симметрическая билинейная функция с матрицей $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ не приводится к диагональному виду заменой базиса (задача).

Над полем $\mathbb{R}$ квадратичную форму можно далее упростить:

Предложение 5.2.5. *Для любой симметрической билинейной функции $\mathcal{B}$ в пространстве над полем $\mathbb{R}$ существует базис, в котором её матрица имеет диагональный вид с $1$, $-1$ и $0$ на диагонали. Другими словами, вещественную квадратичную форму $Q(x)$ линейной заменой координат $x = Cy$ можно привести к виду*

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

Доказательство. Сначала мы с помощью теоремы 5.2.3 приведём квадратичную форму к виду

$$Q(u) = r_{11}(u^1)^2 + \ldots + r_{nn}(u^n)^2.$$

Если $r_{ii} > 0$, то замена $y^i = \sqrt{r_{ii}}u^i$ приводит слагаемое $r_{ii}(u^i)^2$ к виду $(y^i)^2$. Если же $r_{ii} < 0$, то замена $y^i = \sqrt{|r_{ii}|}u^i$ приводит слагаемое $r_{ii}(u^i)^2$ к виду $-(y^i)^2$. В результате получаем требуемый вид квадратичной формы с коэффициентами $1$, $-1$ и $0$. $\square$

Вид, описанный в предложении 5.2.5, называется *нормальным видом* вещественной симметрической билинейной формы (вещественной квадратичной формы). Как мы увидим ниже, это — наиболее простой вид, к которому можно привести квадратичную форму над полем $\mathbb{R}$.

Над полем $\mathbb{C}$ квадратичную форму можно ещё больше упростить:

Предложение 5.2.6. *Для любой симметрической билинейной функции $\mathcal{B}$ над полем $\mathbb{C}$ существует базис, в котором её матрица имеет диагональный вид с $1$ и $0$ на диагонали. Другими словами, комплексную квадратичную форму $Q(x)$ линейной заменой координат $x = Cz$ можно привести к виду*

$$(z^1)^2 + \ldots + (z^r)^2.$$

Доказательство. Сначала мы с помощью предложения 5.2.5 приведём квадратичную форму к виду $(y^1)^2 + \ldots + (y^p)^2 - (y^{p+1})^2 - \ldots - (y^{p+q})^2$. Затем сделаем замену координат $y^k = z^k$ при $k \leqslant p$ и $y^k = iz^k$ при $k > p$. В результате получим требуемый вид, где $r = p + q = \operatorname{rk} Q$. $\square$

Вид, описанный в предложении 5.2.6, называется *нормальным видом* комплексной симметрической билинейной формы (комплексной квадратичной формы).

## 5.3. Нормальный вид кососимметрических и эрмитовых функций

Пусть $\mathcal{B}$ — кососимметрическая билинейная функция (над полем характеристики, отличной от 2). Соответствующую ей кососимметрическую билинейную форму $B(x,y) = b_{ij}x^iy^j$, где $b_{ji} = -b_{ij}$, можно представить в виде

$$B(x,y) = \sum_{i<j} b_{ij}(x^iy^j - x^jy^i).$$

Теорема 5.3.1. *Для любой кососимметрической билинейной функции $\mathcal{B}$ существует базис, в котором её матрица блочно-диагональная с блоками размера 1 или 2, причём блоки размера 1 нулевые, а блоки размера 2 имеют вид* $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

*Другими словами, любую кососимметрическую билинейную форму $B(x,y)$ линейной заменой координат $x = Cy$ можно привести к виду*

$$(x^1y^2 - x^2y^1) + (x^3y^4 - x^4y^3) + \ldots + (x^{2k-1}y^{2k} - x^{2k}y^{2k-1}).$$

Доказательство. Проведём индукцию по размерности пространства $V$.

При $\dim V = 1$ доказывать нечего, так как кососимметрическая функция нулевая.

Пусть $\dim V = 2$. Тогда матрица кососимметрической функции в произвольном базисе $\boldsymbol{e}_1, \boldsymbol{e}_2$ имеет вид $\begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}$, где $b = b_{12} = \mathcal{B}(\boldsymbol{e}_1, \boldsymbol{e}_2)$. Пусть $b \neq 0$ (иначе мы уже имеем два блока из нулей). Тогда в новом базисе $\boldsymbol{e}_{1'} = \boldsymbol{e}_1$ и $\boldsymbol{e}_{2'} = \frac{1}{b}\boldsymbol{e}_2$ мы имеем

$$b'_{12} = \mathcal{B}(\boldsymbol{e}_{1'}, \boldsymbol{e}_{2'}) = \mathcal{B}(\boldsymbol{e}_1, \frac{1}{b}\boldsymbol{e}_2) = \frac{1}{b}\mathcal{B}(\boldsymbol{e}_1, \boldsymbol{e}_2) = 1,$$

т.е. матрица кососимметрической формы имеет требуемый вид $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Теперь предположим, что утверждение уже доказано для пространств размерности меньше $n$, и докажем его для размерности $n$. Можно считать, что функция $\mathcal{B}$ не является тождественно нулевой (иначе доказывать нечего). Пусть $b_{12} = \mathcal{B}(\boldsymbol{e}_1, \boldsymbol{e}_2) \neq 0$ для некоторого базиса $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Попытаемся заменить базисные векторы так, чтобы новые векторы $\boldsymbol{e}_{1'}, \ldots, \boldsymbol{e}_{n'}$ удовлетворяли соотношениям

$$\mathcal{B}(\boldsymbol{e}_{1'}, \boldsymbol{e}_{2'}) = 1, \quad \mathcal{B}(\boldsymbol{e}_{1'}, \boldsymbol{e}_{i'}) = \mathcal{B}(\boldsymbol{e}_{2'}, \boldsymbol{e}_{i'}) = 0 \quad \text{при } i \geqslant 3. \tag{5.2}$$

Новый базис будем искать в виде

$$\boldsymbol{e}_{1'} = \boldsymbol{e}_1, \quad \boldsymbol{e}_{2'} = c\boldsymbol{e}_2, \quad \boldsymbol{e}_{i'} = \boldsymbol{e}_i + c^1_{i'}\boldsymbol{e}_1 + c^2_{i'}\boldsymbol{e}_2 \quad \text{при } i \geqslant 3.$$

Подставив эти соотношения в (5.2), получим $c = \frac{1}{b_{12}}$ и

$$0 = \mathcal{B}(\boldsymbol{e}_{1'}, \boldsymbol{e}_{i'}) = \mathcal{B}(\boldsymbol{e}_1, \boldsymbol{e}_i + c^1_{i'}\boldsymbol{e}_1 + c^2_{i'}\boldsymbol{e}_2) = b_{1i} + c^2_{i'}b_{12},$$

откуда $c^2_{i'} = -\frac{b_{1i}}{b_{12}}$. Аналогично,

$$0 = \mathcal{B}(\boldsymbol{e}_{2'}, \boldsymbol{e}_{i'}) = \mathcal{B}(c\boldsymbol{e}_2, \boldsymbol{e}_i + c^1_{i'}\boldsymbol{e}_1 + c^2_{i'}\boldsymbol{e}_2) = c(b_{2i} - c^1_{i'}b_{12}),$$

откуда $c^1_{i'} = \frac{b_{2i}}{b_{12}}$. Окончательно, наша замена базиса имеет вид

$$\boldsymbol{e}_{1'} = \boldsymbol{e}_1, \quad \boldsymbol{e}_{2'} = \frac{1}{b_{12}}\boldsymbol{e}_2, \quad \boldsymbol{e}_{i'} = \boldsymbol{e}_i + \frac{b_{2i}}{b_{12}}\boldsymbol{e}_1 - \frac{b_{1i}}{b_{12}}\boldsymbol{e}_2 \quad \text{при } i \geqslant 3.$$

Ввиду соотношений (5.2), в новом базисе матрица билинейной функции $\mathcal{B}$ имеет вид

$$B' = \left(\begin{array}{cc|ccc} 0 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 0 & \cdots & 0 \\ \hline 0 & 0 & & & \\ \vdots & \vdots & & \widetilde{B}' & \\ 0 & 0 & & & \end{array}\right)$$

где $\widetilde{B}'$ — матрица билинейной функции $\mathcal{B}$ на подпространстве $\langle \boldsymbol{e}_{3'}, \ldots, \boldsymbol{e}_{n'} \rangle$. Так как это пространство имеет размерность $n-2$, по предположению индукции в нём существует требуемый базис $\boldsymbol{e}_{3''}, \ldots, \boldsymbol{e}_{n''}$. Тогда в базисе $\boldsymbol{e}_{1'}, \boldsymbol{e}_{2'}, \boldsymbol{e}_{3''}, \ldots, \boldsymbol{e}_{n''}$ исходного пространства $V$ матрица кососимметрической функции $\mathcal{B}$ имеет требуемый вид. □

Вид, описанный в теореме 5.3.1, называется *нормальным видом* кососимметрической билинейной формы.

Следствие 5.3.2.

а) *Ранг кососимметрической билинейной функции — чётное число;*
б) *Кососимметрическая билинейная функция в пространстве нечётной размерности всегда вырождена.*

Теперь рассмотрим полуторалинейные функции в комплексном пространстве.

Теорема 5.3.3. *Для любой эрмитовой полуторалинейной функции $\mathcal{S}$ существует базис, в котором её матрица имеет диагональный вид с $1$, $-1$ и $0$ на диагонали.*

Доказательство. Доказательство аналогично доказательству теоремы 5.2.3 для симметрических билинейных функций: отличие имеется лишь во вспомогательном преобразовании. Мы проведём доказательство методом поиска базиса.

Пусть $S = (s_{ij})$ — матрица полуторалинейной функции $\mathcal{S}$ в исходном базисе.

Основное преобразование производится, если хотя бы один диагональный элемент $s_{ii} = \mathcal{S}(\boldsymbol{e}_i, \boldsymbol{e}_i)$ отличен от нуля. Пусть $s_{11} \neq 0$. Замена базиса производится по тем же формулам (5.1), что и для симметрической билинейной функции (надо лишь заменить $b_{ij}$ на $s_{ij}$ и $\mathcal{B}$ на $\mathcal{S}$ в формулах).

Первое вспомогательное преобразование производится, если $s_{11} = 0$, но существует $s_{ii} \neq 0$, и заключается в перестановке 1-го и $i$-го базисных векторов. Второе вспомогательное преобразование производится, если все $s_{ii} = \mathcal{S}(\boldsymbol{e}_i, \boldsymbol{e}_i)$ равны нулю. Пусть $s_{12} = \mathcal{S}(\boldsymbol{e}_1, \boldsymbol{e}_2) \neq 0$. Здесь возможны два случая: $\operatorname{Re} s_{12} \neq 0$ и $\operatorname{Re} s_{12} = 0$. В первом случае производим ту же замену, что и для симметрических билинейных функций:

$$\boldsymbol{e}_{1'} = \boldsymbol{e}_1 + \boldsymbol{e}_2, \quad \boldsymbol{e}_{2'} = \boldsymbol{e}_2, \quad \ldots, \quad \boldsymbol{e}_{n'} = \boldsymbol{e}_n.$$

Тогда в новом базисе мы имеем

$$s'_{11} = \mathcal{S}(\boldsymbol{e}_{1'}, \boldsymbol{e}_{1'}) = \mathcal{S}(\boldsymbol{e}_1 + \boldsymbol{e}_2, \boldsymbol{e}_1 + \boldsymbol{e}_2) = \mathcal{S}(\boldsymbol{e}_1, \boldsymbol{e}_2) + \mathcal{S}(\boldsymbol{e}_2, \boldsymbol{e}_1) = 2 \operatorname{Re} s_{12} \neq 0.$$

Далее мы можем применить основное преобразование.

Если же $\operatorname{Re} s_{12} = 0$, то $\operatorname{Im} s_{12} \neq 0$ (так как $s_{12} \neq 0$ по предположению). В этом случае делаем замену

$$\boldsymbol{e}_{1'} = \boldsymbol{e}_1 + i\boldsymbol{e}_2, \quad \boldsymbol{e}_{2'} = \boldsymbol{e}_2, \quad \ldots, \quad \boldsymbol{e}_{n'} = \boldsymbol{e}_n.$$

Тогда в новом базисе мы имеем

$$s'_{11} = \mathcal{S}(\boldsymbol{e}_{1'}, \boldsymbol{e}_{1'}) = \mathcal{S}(\boldsymbol{e}_1 + i\boldsymbol{e}_2, \boldsymbol{e}_1 + i\boldsymbol{e}_2) = i\mathcal{S}(\boldsymbol{e}_1, \boldsymbol{e}_2) - i\mathcal{S}(\boldsymbol{e}_2, \boldsymbol{e}_1) = -2\operatorname{Im} s_{12} \neq 0,$$

и мы снова можем применить основное преобразование.

Последовательно применяя основное преобразование и дополняя его в необходимых случаях вспомогательными преобразованием, мы получаем базис $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$, в котором матрица эрмитовой функции $\mathcal{S}$ имеет диагональный вид. Пусть $\mathcal{S}(\boldsymbol{f}_i, \boldsymbol{f}_i) = r_{ii}$. Эти числа вещественны в силу эрмитовости. Далее доказательство завершается так же, как и для симметрических функций. □

Эрмитова полуторалинейная форма, соответствующая диагональной матрице из предыдущей теоремы, имеет вид

$$\overline{x}^1 y^1 + \ldots + \overline{x}^p y^p - \overline{x}^{p+1} y^{p+1} - \ldots - \overline{x}^{p+q} y^{p+q}.$$

Этот вид называется *нормальным видом* эрмитовой полуторалинейной формы.

## 5.4. Закон инерции. Единственность нормального вида

В случае симметрической билинейной формы над $\mathbb{C}$ нормальный вид зависит только от её ранга, и поэтому мы получаем:

ПРЕДЛОЖЕНИЕ 5.4.1. *Две комплексные симметрические билинейные формы (комплексные квадратичные формы) получаются друг из друга линейной заменой координат тогда и только тогда, когда их ранги совпадают.*

Аналогично, нормальный вид кососимметрической билинейной формы (над любым полем) зависит только от её ранга, и поэтому мы получаем:

ПРЕДЛОЖЕНИЕ 5.4.2. *Две кососимметрические билинейные формы получаются друг из друга линейной заменой координат тогда и только тогда, когда их ранги совпадают.*

В случае вещественных симметрических билинейных форм и в случае эрмитовых полуторалинейных форм ситуация сложнее: их нормальный вид не определяется одним лишь рангом, а зависит ещё от количества 1 и $-1$ на диагонали матрицы. Оказывается, что нормальный вид такой формы не зависит от способа приведения к нормальному виду:

ТЕОРЕМА 5.4.3 (закон инерции). *Количество* 1, $-1$ *и* 0 *на диагонали матрицы вещественной симметрической билинейной функции* $\mathcal{B}$ *не зависит от способа приведения к нормальному виду.*

*Другими словами, если квадратичная форма* $Q(x)$ *вещественной линейной заменой* $x = Cy$ *приводится к виду*

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

*а вещественной линейной заменой* $x = Cz$ *— к виду*

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

*то мы имеем* $p = p'$ *и* $q = q'$.

Доказательство. Пусть $(x^1, \dots, x^n)$ – координаты в исходном базисе $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ пространства $V$, $(y^1, \dots, y^n)$ — координаты в базисе $\boldsymbol{f}_1, \dots, \boldsymbol{f}_n$, а $(z^1, \dots, z^n)$ — координаты в базисе $\boldsymbol{g}_1, \dots, \boldsymbol{g}_n$. Рассмотрим подпространства

$$U_+ = \langle \boldsymbol{f}_1, \dots, \boldsymbol{f}_p \rangle, \qquad U_- = \langle \boldsymbol{f}_{p+1}, \dots, \boldsymbol{f}_{p+q} \rangle, \qquad U_0 = \langle \boldsymbol{f}_{p+q+1}, \dots, \boldsymbol{f}_n \rangle,$$
$$W_+ = \langle \boldsymbol{g}_1, \dots, \boldsymbol{g}_{p'} \rangle, \qquad W_- = \langle \boldsymbol{g}_{p'+1}, \dots, \boldsymbol{g}_{p'+q'} \rangle, \qquad W_0 = \langle \boldsymbol{g}_{p'+q'+1}, \dots, \boldsymbol{g}_n \rangle.$$

Для ненулевого вектора $\boldsymbol{x} \in U_+$ мы имеем $\boldsymbol{x} = y^1\boldsymbol{f}_1 + \ldots + y^p\boldsymbol{f}_p$ и поэтому $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = (y^1)^2 + \ldots + (y^p)^2 > 0$. Аналогично, если $\boldsymbol{x} \in U_- \oplus U_0$, то $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) \leqslant 0$. Для ненулевого вектора $\boldsymbol{x} \in W_+$ мы имеем $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) > 0$, а для $\boldsymbol{x} \in W_- \oplus W_0$ имеем $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) \leqslant 0$.

Предположим, что $p > p'$. Тогда

$$\dim U_+ + \dim(W_- \oplus W_0) = p + (n - p') > n = \dim V,$$

значит, $U_+ \cap (W_- \oplus W_0) \neq \{\mathbf{0}\}$. Возьмём ненулевой вектор $\boldsymbol{x}$ в этом пересечении. Так как $\boldsymbol{x} \in U_+$, имеем $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) > 0$. С другой стороны, из $\boldsymbol{x} \in W_- \oplus W_0$ следует, что $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) \leqslant 0$. Противоречие.

Следовательно, $p = p'$. Кроме того, $p + q = p' + q' = \operatorname{rk} \mathcal{B}$, а значит и $q = q'$. □

Имеет место также закон инерции для эрмитовых полуторалинейных функций, который доказывается полностью аналогично:

Теорема 5.4.4. *Количество $1$, $-1$ и $0$ на диагонали матрицы эрмитовой полуторалинейной функции не зависит от способа приведения к нормальному виду.*

Определение 5.4.5. Разность $p - q$ между числом положительных и отрицательных диагональных элементов в нормальном виде называется *сигнатурой* вещественной симметрической билинейной функции (эрмитовой полуторалинейной функции).

Из теоремы 5.4.3 следует, что сигнатура, как и ранг, является инвариантном вещественной симметрической билинейной функции (эрмитовой полуторалинейной функции), т.е. не зависит от базиса.

Следствие 5.4.6. *Две вещественные симметрические билинейные формы или две эрмитовы полуторалинейные функции получаются друг из друга линейной заменой координат тогда и только тогда, когда их ранги и сигнатуры совпадают.*

Все результаты этого раздела можно свести в одно утверждение: нормальный вид симметрической или кососимметрической билинейной функции или эрмитовой полуторалинейной функции единствен.

## 5.5. Теорема Якоби. Критерий Сильвестра

Теорема Якоби позволяет (при выполнении некоторого дополнительного условия) найти нормальный вид квадратичной формы без нахождения преобразования.

Напомним, что *угловым минором* порядка $k$ матрицы называется минор (определитель подматрицы), составленный из первых $k$ строк и первых $k$ столбцов.

Теорема 5.5.1 (Якоби). *Предположим, что все угловые миноры матрицы $Q$ квадратичной формы отличны от нуля до порядка $r = \operatorname{rk} Q$. Тогда существует замена координат, приводящая данную квадратичную форму к виду*

$$|Q_1|(x^1)^2 + \frac{|Q_2|}{|Q_1|}(x^2)^2 + \ldots + \frac{|Q_r|}{|Q_{r-1}|}(x^r)^2,$$

*где $|Q_i|$ — угловой минор порядка $i$.*

Сначала докажем лемму. Скажем, что для квадратичной формы имеет место *регулярный случай*, если она приводится к диагональному виду последовательным применением исключительно основного преобразования метода Лагранжа.

Лемма 5.5.2. *Для квадратичной формы $Q(x)$ имеет место регулярный случай тогда и только тогда, когда все угловые миноры её матрицы $Q$ отличны от нуля до порядка $r = \operatorname{rk} Q$.*

Доказательство. Пусть угловые миноры до порядка $r$ отличны от нуля. Тогда $q_{11} = |Q_1|$ — угловой минор порядка 1, который не равен нулю по предположению. Значит, на первом шаге применимо основное преобразование метода Лагранжа.

Предположим теперь, что после $k$-кратного применения основного преобразования метода Лагранжа матрица квадратичной формы принимает вид

$$Q' = \left(\begin{array}{ccc|cc} q'_{11} & & 0 & & \\ & \ddots & & & 0 \\ 0 & & q'_{kk} & & \\ \hline & & & q'_{k+1,k+1} & \cdots \\ & 0 & & \vdots & \ddots \end{array}\right) \tag{5.3}$$

Заметим, что матрица замены координат для основного преобразования есть

$$C = \begin{pmatrix} 1 & -\frac{q_{12}}{q_{11}} & \cdots & -\frac{q_{1n}}{q_{11}} \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

(см. формулы (5.1)). Для угловых подматриц $Q_k$ мы имеем $Q'_k = C_k^t Q_k C_k$, где $C_k$ — угловая подматрица матрицы $C$. Так как $\det C_k = 1$, мы получаем $|Q'_k| = |Q_k|$, т.е. угловые миноры матрицы квадратичной формы не меняются при основном преобразовании метода Лагранжа.

Возвращаясь к матрице (5.3), мы получаем $|Q_{k+1}| = |Q'_{k+1}| = q'_{11} \cdots q'_{kk} q'_{k+1,k+1} \neq 0$ при $k < r$ по предположению. Следовательно, $q'_{k+1,k+1} \neq 0$, и мы снова можем применить основное преобразование.

После $r$-кратного применения основного преобразования мы получаем матрицу (5.3), где $k = r$ и матрица в правом нижнем углу равна нулю. Следовательно, квадратичная форма приведена к диагональному виду последовательным применением основного преобразования метода Лагранжа, и мы имеем регулярный случай.

Пусть теперь имеет место регулярный случай, т.е. форма приведена к диагональному виду с ненулевыми числами $q'_{11}, \ldots, q'_{rr}$ на диагонали последовательным применением основного преобразования. Тогда, так как угловые миноры не меняются при основном преобразовании, мы имеем $|Q_k| = |Q'_k| = q'_{11} \cdots q'_{kk} \neq 0$ при $k \leqslant r$. □

Доказательство теоремы Якоби. В силу предыдущей леммы, мы можем привести квадратичную форму к диагональному виду

$$q'_{11}(u^1)^2 + \ldots + q'_{rr}(u^r)^2$$

используя лишь основное преобразование метода Лагранжа. Тогда $|Q_k| = |Q'_k| = q'_{11}\cdots q'_{kk}$ при $k \leqslant r$, т.е. $q'_{kk} = \frac{|Q_k|}{|Q_{k-1}|}$, что и требовалось. □

Определение 5.5.3. Симметрическая билинейная функция $\mathcal{B}$ называется *положительно определённой*, если $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) > 0$ при $\boldsymbol{x} \neq \mathbf{0}$. Соответствующая квадратичная форма $Q(x)$ удовлетворяет условию $Q(x) > 0$ при $x \neq 0$ и также называется положительно определённой.

Положительно определённая симметрическая билинейная функция задаёт в пространстве $V$ скалярное произведение, т.е. превращает $V$ в евклидово пространство.

Теорема 5.5.4 (критерий Сильвестра). *Симметрическая билинейная функция (квадратичная форма) положительно определена тогда и только тогда, когда все угловые миноры её матрицы в некотором базисе положительны.*

Доказательство. Пусть все угловые миноры $|Q_k|$ матрицы квадратичной формы $Q(x)$ положительны. Тогда в силу теоремы Якоби квадратичная форма приводится к виду $Q(u) = q'_{11}(u^1)^2 + \ldots + q'_{nn}(u^n)^2$, где $n = \operatorname{rk} Q = \dim V$, а $q'_{kk} = \frac{|Q_k|}{|Q_{k-1}|} > 0$. Такая квадратичная форма положительно определена, так как $Q(u) > 0$ при $u \neq 0$.

Обратно, пусть $Q(x)$ положительно определена. Так как в любом базисе мы имеем $q_{ii} = \mathcal{B}(\boldsymbol{e}_i, \boldsymbol{e}_i) > 0$, всегда применимо основное преобразование метода Лагранжа. Тогда последовательно применяя основное преобразование, мы приведём квадратичную форму к виду $Q(u) = q'_{11}(u^1)^2 + \ldots + q'_{nn}(u^n)^2$, где $q'_{ii} > 0$. Следовательно, $|Q_k| = |Q'_k| = q'_{11}\cdots q'_{kk} > 0$ для любого $k$. □

## 5.6. Симметрические билинейные функции в евклидовых пространствах. Канонический вид

Пусть $V$ — евклидово пространство. Мы знаем из теоремы 3.9.3, что отображение $\boldsymbol{x} \mapsto \xi_{\boldsymbol{x}} = (\boldsymbol{x}, \cdot\,)$ устанавливает канонический изоморфизм $V \to V^*$ между $V$ и его двойственным пространством $V^*$. Это позволяет нам отождествить пространства линейных отображений $\operatorname{Hom}(V, V)$ и $\operatorname{Hom}(V, V^*)$. С другой стороны, $\operatorname{Hom}(V, V)$ — это пространство $\operatorname{End}(V)$ линейных операторов, а в силу теоремы 5.1.6, $\operatorname{Hom}(V, V^*)$ — это пространство билинейных функций $\mathrm{B}(V)$. Если вникнуть в построение этих изоморфизмов, то мы увидим, что в явном виде канонический изоморфизм между пространством операторов и пространством билинейных функций в евклидовом пространстве описывается следующим утверждением, которое легко доказать и непосредственно:

Предложение 5.6.1. *Пусть $V$ — евклидово пространство. Отображение $\mathcal{A} \mapsto \mathcal{B}_{\mathcal{A}} = (\mathcal{A}\cdot\,, \cdot\,)$ устанавливает изоморфизм $\psi$: $\operatorname{End}(V) \to \mathrm{B}(V)$ между пространством операторов и пространством билинейных функций. Здесь $\mathcal{B}_{\mathcal{A}} = (\mathcal{A}\cdot\,, \cdot\,)$ — билинейная функция, задаваемая формулой $\mathcal{B}_{\mathcal{A}}(\boldsymbol{x}, \boldsymbol{y}) = (\mathcal{A}\boldsymbol{x}, \boldsymbol{y})$.*

Доказательство. Так как пространства $\operatorname{End}(V)$ и $\mathrm{B}(V)$ имеют одинаковую размерность $n^2$, достаточно проверить мономорфность отображения $\psi$. Пусть $\psi(\mathcal{A}) = 0$, т.е. $\mathcal{B}_{\mathcal{A}}$ — тождественно нулевая функция. Тогда $(\mathcal{A}\boldsymbol{x}, \boldsymbol{y}) = 0$ для любых $\boldsymbol{x}, \boldsymbol{y}$. В частности, $(\mathcal{A}\boldsymbol{x}, \mathcal{A}\boldsymbol{x}) = 0$ для любого $\boldsymbol{x}$, т.е. $\mathcal{A}\boldsymbol{x} = \mathbf{0}$ и $\mathcal{A} = \mathcal{O}$ — нулевой оператор. □

Это утверждение имеет важные следствия: оно позволяет переводить утверждения об операторах в утверждения о билинейных функциях и наоборот. Одно из основных приложений заключается в следующем:

Теорема 5.6.2. *Для билинейной симметрической функции в евклидовом пространстве существует ортонормированный базис, в котором её матрица диагональна. Другими словами, квадратичная форма приводится к диагональному виду ортогональным преобразованием.*

Первое доказательство. Пусть $\mathcal{B}$ — симметрическая билинейная функция и $\mathcal{A}$ — соответствующий ей оператор, т.е. $\mathcal{B} = \mathcal{B}_{\mathcal{A}}$. Тогда из $\mathcal{B}(\boldsymbol{x}, \boldsymbol{y}) = \mathcal{B}(\boldsymbol{y}, \boldsymbol{x})$ получаем $(\mathcal{A}\boldsymbol{x}, \boldsymbol{y}) = (\mathcal{A}\boldsymbol{y}, \boldsymbol{x}) = (\boldsymbol{x}, \mathcal{A}\boldsymbol{y})$, т.е. оператор $\mathcal{A}$ самосопряжён. Выберем канонический ортонормированный базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ для $\mathcal{A}$, т.е. $\mathcal{A}\boldsymbol{e}_i = \lambda_i \boldsymbol{e}_i$. Тогда для матрицы $B = (b_{ij})$ функции $\mathcal{B}$ в этом базисе имеем

$$b_{ij} = \mathcal{B}(\boldsymbol{e}_i, \boldsymbol{e}_j) = (\mathcal{A}\boldsymbol{e}_i, \boldsymbol{e}_j) = (\lambda_i \boldsymbol{e}_i, \boldsymbol{e}_j) = \lambda_i \delta_{ij},$$

т.е. матрица $B$ диагональна (и совпадает с матрицей оператора $\mathcal{A}$). □

Второе доказательство. Посмотрим, как преобразуются матрица билинейной функции и матрица оператора при ортогональном преобразовании. Пусть $B$ — матрица билинейной функции в некотором ортонормированном базисе. При ортогональном преобразовании с матрицей $C$ матрица $B$ переходит в матрицу $B' = C^t B C$. Так как матрица $C$ ортогональна, то же преобразование мы можем записать в виде $B' = C^{-1} B C$. Но это — закон преобразования для матрицы оператора. Так как оператор с симметричной матрицей $B$ в ортонормированном базисе самосопряжён, его можно привести к диагональному виду ортогональным преобразованием. □

Диагональный вид, к которому приводится симметрическая билинейная функция (квадратичная форма) ортогональным преобразованием, называется *каноническим.*

Предложение 5.6.3. *Канонический вид симметрической билинейной функции (квадратичной формы) единствен с точностью до перестановки диагональных элементов. Эти элементы представляют собой собственные значения матрицы квадратичной формы в любом ортонормированном базисе.*

Доказательство. Пусть $Q$ — матрица квадратичной формы в ортонормированном базисе. Тогда диагональные элементы канонического вида — это собственные значения самосопряжённого оператора с матрицей $Q$, т.е. корни уравнения $\det(Q - tE) = 0$. В другом ортонормированном базисе матрица квадратичной формы есть $Q' = C^t Q C$ и её собственные значения находятся из уравнения $\det(Q' - tE) = 0$. Так как $\det(Q' - tE) = \det(C^t Q C - t C^t C) = \det(C^t (Q - tE) C) = \det(C^t C) \det(Q - tE) = \det(Q - tE)$, собственные значения матриц $Q$ и $Q'$ совпадают. □

Собственные векторы матрицы квадратичной формы также называют её *главными осями*, а приведение к каноническому виду — *приведением к главным осям.*

## 5.7. Приведение пары форм к диагональному виду. Собственные значения и собственные векторы пары форм

Модификация теоремы 5.6.2 позволяет одновременно приводить к диагональному виду сразу две квадратичные формы, одна из которых положительно определена:

Теорема 5.7.1. *Пусть даны две квадратичные формы $Q(x)$ и $B(x)$, причём форма $Q(x)$ положительно определена. Тогда существует линейная замена координат $x = Cy$, приводящая форму $Q(x)$ к нормальному виду $(y^1)^2 + \ldots + (y^n)^2$, а форму $B(x)$ — к диагональному виду $\lambda_1 (y^1)^2 + \ldots + \lambda_n (y^n)^2$.*

Доказательство. Положительно определённая симметрическая билинейная функция, соответствующая квадратичной форме $Q(x)$, превращает $V$ в евклидово пространство. В исходных координатах матрица Грама скалярного произведения есть $Q$. В любом ортонормированном базисе матрица Грама (она же матрица квадратичной формы $Q(x)$) будет единичной. Согласно теореме 5.6.2, существует ортонормированный базис, в котором матрица формы $B(x)$ имеет диагональный вид. □

Обратим внимание, что матрица $C$ замены координат из предыдущей теоремы не является ортогональной: вместо соотношения $C^tC = E$ она удовлетворяет соотношению $C^tQC = E$. Другими словами, столбцы матрицы $C$ образуют ортонормированный базис относительно скалярного произведения в $\mathbb{R}^n$ с матрицей Грама $Q$.

Если мы попытаемся превратить доказательство теоремы 5.7.1 в практический алгоритм, то нам придётся действовать в два шага: сначала найти матрицу $C'$, приводящую $Q$ к единичному виду, т.е. $(C')^tQC' = E$, а матрицу $B$ к некоторому виду $B' = (C')^tBC'$; затем найти ортогональную матрицу $C''$, приводящую $B'$ к каноническому (диагональному) виду, т.е. $(C'')^tB'C'' = D$ — диагональная матрица; тогда для $C = C'C''$ мы имеем $C^tQC = E$ и $C^tBC = D$.

На практике, однако, это метод не очень эффективен. Более эффективный метод основан на следующем определении:

Определение 5.7.2. Пусть даны две квадратичные формы $Q(x)$ и $B(x)$, причём $Q(x)$ положительно определена. Корни уравнения $\det(B - tQ) = 0$ называются *собственными значениями пары форм* $Q(x)$ и $B(x)$.

Пусть $\lambda$ — собственное значение пары форм $Q(x)$ и $B(x)$. Ненулевой вектор $y$, удовлетворяющий системе уравнений $(B-\lambda Q)y = 0$, называется *собственным вектором пары форм*, соответствующим собственному значению $\lambda$.

Теорема 5.7.3. *Предположим, что линейная замена $x = Cy$ приводит положительно определённую форму $Q(x)$ к нормальному виду $(y^1)^2 + \ldots + (y^n)^2$, а форму $B(x)$ — диагональному виду $\lambda_1(y^1)^2 + \ldots + \lambda_n(y^n)^2$. Тогда числа $\lambda_1, \ldots, \lambda_n$ суть собственные векторы пары форм $Q(x)$ и $B(x)$, а столбцы матрицы $C$ образуют базис из собственных векторов пары форм, который является ортонормированным относительно скалярного произведения, задаваемого формой $Q(x)$.*

Доказательство. Мы имеем $C^tQC = E$, а $C^tBC = D$, где $D$ — диагональная матрица с числами $\lambda_1, \ldots, \lambda_n$ на диагонали. Тогда для любого $i$ матрица $D - \lambda_iE$ вырождена, следовательно,

$$0 = \det(D-\lambda_iE) = \det(C^tBC-\lambda_iC^tQC) = \det(C^t(B-\lambda_iQ)C) = \det(C)^2\det(B-\lambda_iQ).$$

Так как матрица $C$ невырождена, отсюда следует, что $\det(B - \lambda_iQ) = 0$, т.е. $\lambda_i$ — собственное значение пары форм.

Пусть $c_i$ — $i$-й столбец матрицы $C$. Из соотношения $C^t(B - \lambda_iQ)C = D - \lambda_iE$ мы получаем, что $i$-й столбец матрицы $C^t(B-\lambda_iQ)C$ нулевой, т.е. $C^t(B-\lambda_iQ)c_i = 0$. Так как матрица $C$ обратима, отсюда следует, что $(B - \lambda_iQ)c_i = 0$, т.е. $c_i$ — собственный вектор пары форм, отвечающий собственному значению $\lambda_i$.

Наконец, соотношение $C^tQC = E$ выражает тот факт, что столбцы матрицы $C$ образуют ортонормированный базис относительно скалярного произведения с матрицей Грама $Q$. □

Глава 6

# Тензоры

Понятие тензора обобщает объекты, рассматривавшиеся нами ранее: векторы, ковекторы (линейные функции), линейные операторы, билинейные функции. Имеется два подхода к определению тензора: бескоординатный (через полилинейные функции) и координатный (собственно тензоры). Начнём с первого.

## 6.1. Полилинейные функции

Пусть $V$ — линейное пространство над полем $\mathbf{k}$ нулевой характеристики (обычно $\mathbf{k} = \mathbb{R}$ или $\mathbb{C}$) и $V^*$ — двойственное пространство. Элементы $\boldsymbol{v} \in V$ — это, как обычно, векторы, а элементы $\xi \in V^*$ здесь мы будем называть *ковекторами*.

Определение 6.1.1. *Полилинейной функцией типа* $(p, q)$ называется функция

$$\mathcal{T}\colon \underbrace{V^* \times \ldots \times V^*}_{p} \times \underbrace{V \times \ldots \times V}_{q} \to \mathbf{k}$$

от $p$ ковекторных и $q$ векторных аргументов, которая линейна по каждому аргументу, т.е. удовлетворяет соотношениям

$$\mathcal{T}(\xi^1, \ldots, \lambda'\xi^{j'} + \lambda''\xi^{j''}, \ldots, \xi^p, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_q) = \\ = \lambda'\mathcal{T}(\xi^1, \ldots, \xi^{j'}, \ldots, \xi^p, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_q) + \lambda''\mathcal{T}(, \xi^1, \ldots, \xi^{j''}, \ldots, \xi^p, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_q).$$

$$\mathcal{T}(\xi^1, \ldots, \xi^p, \boldsymbol{v}_1, \ldots, \mu'\boldsymbol{v}'_i + \mu''\boldsymbol{v}''_i, \ldots, \boldsymbol{v}_q) = \\ = \mu'\mathcal{T}(\xi^1, \ldots, \xi^p, \boldsymbol{v}_1, \ldots, \boldsymbol{v}'_i, \ldots, \boldsymbol{v}_q) + \mu''\mathcal{T}(\xi^1, \ldots, \xi^p, \boldsymbol{v}_1, \ldots, \boldsymbol{v}''_i, \ldots, \boldsymbol{v}_q),$$

Полилинейные функции типа $(p, q)$ образуют линейное пространство, в котором сумма и умножение на скаляры определены по формуле

$$(\lambda\mathcal{T} + \mu\mathcal{S})(\xi^1, \ldots, \xi^p, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_q) = \lambda\mathcal{T}(\xi^1, \ldots, \xi^p, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_q) + \mu\mathcal{S}(\xi^1, \ldots, \xi^p, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_q).$$

Мы будем обозначать это пространство через $\mathrm{P}^p_q(V)$.

Пример 6.1.2. Линейные функции (ковекторы) являются полилинейными типа $(0, 1)$, а билинейные — типа $(0, 2)$. Ввиду наличия канонического изоморфизма $(V^*)^* \cong V$ векторы являются полилинейными функциями типа $(1, 0)$: значение вектора $\boldsymbol{v}$ на ковекторе $\xi$ определяется по формуле $\boldsymbol{v}(\xi) := \xi(\boldsymbol{v})$.

Следующее утверждение показывает, что операторы также можно рассматривать как полилинейные функции:

Предложение 6.1.3. *Сопоставление линейному оператору* $\mathcal{A}$ *полилинейной функции* $\mathcal{T}_{\mathcal{A}}$ *типа* $(1, 1)$*, задаваемой формулой* $\mathcal{T}_{\mathcal{A}}(\xi, \boldsymbol{v}) := \xi(\mathcal{A}(\boldsymbol{v}))$*, устанавливает канонический изоморфизм* $\mathrm{End}(V) \xrightarrow{\cong} \mathrm{P}^1_1(V)$.

Доказательство. Прежде всего заметим, что данное отображение $\mathrm{End}(V) \to \mathrm{P}^1_1(V)$ линейно. Пусть $\dim V = n$. Тогда размерность пространства $\mathrm{P}^1_1(V)$ равна $n^2$ — это доказывается при помощи выбора базиса, так же как и для билинейных функций (см. также раздел 6.4). Поэтому размерности пространств $\mathrm{End}(V)$ и $\mathrm{P}^1_1(V)$ равны, а значит достаточно доказать, что $\mathrm{End}(V) \to \mathrm{P}^1_1(V)$ — мономорфизм. Пусть $\mathcal{T}_\mathcal{A}$ — тождественно нулевая полилинейная функция, т.е. $\xi(\mathcal{A}(\boldsymbol{v})) = 0$ для любых $\boldsymbol{v} \in V$ и $\xi \in V^*$. Получаем, что любая линейная функция обращается в нуль на векторе $\mathcal{A}(\boldsymbol{v})$, т.е. $\mathcal{A}(\boldsymbol{v}) = \mathbf{0}$. Так как это верно для любого $\boldsymbol{v}$, получаем $\mathcal{A} = \mathcal{O}$. Итак отображение $\mathcal{A} \mapsto \mathcal{T}_\mathcal{A}$ мономорфно, а значит задаёт изоморфизм. □

## 6.2. Тензоры: координатное определение

Зафиксируем базис $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ в пространстве $V$. В пространстве $V^*$ имеется двойственный базис $\varepsilon^1, \dots, \varepsilon^n$, где $\varepsilon^i(\boldsymbol{e}_j) = \delta^i_j$. Тогда значение любой полилинейной функции $\mathcal{T} \in \mathrm{P}^p_q(V)$ определяется её значениями на базисных векторах и ковекторах:

$$\begin{aligned}(6.1)\quad \mathcal{T}(\xi^1, \dots, \xi^p, \boldsymbol{v}_1, \dots, \boldsymbol{v}_q) &= \mathcal{T}(\xi^1_{i_1}\varepsilon^{i_1}, \dots, \xi^p_{i_p}\varepsilon^{i_p}, v^{j_1}_1\boldsymbol{e}_{j_1}, \dots, v^{j_q}_q\boldsymbol{e}_{j_q}) \\ &= \xi^1_{i_1}\cdots\xi^p_{i_p}v^{j_1}_1\cdots v^{j_q}_q\mathcal{T}(\varepsilon^{i_1}, \dots, \varepsilon^{i_p}, \boldsymbol{e}_{j_1}, \dots, \boldsymbol{e}_{j_q}).\end{aligned}$$

Сопоставим полилинейной функции $\mathcal{T} \in \mathrm{P}^p_q(V)$ и базису $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ набор из $n^{p+q}$ чисел $T = \{T^{i_1,\dots,i_p}_{j_1,\dots,j_q}\}$, где

$$(6.2)\qquad T^{i_1,\dots,i_p}_{j_1,\dots,j_q} := \mathcal{T}(\varepsilon^{i_1}, \dots, \varepsilon^{i_p}, \boldsymbol{e}_{j_1}, \dots, \boldsymbol{e}_{j_q}).$$

Посмотрим, как преобразуется это набор при заменах базиса. Пусть $C = (c^i_{i'})$ — матрица перехода от базиса $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ к базису $\boldsymbol{e}_{1'}, \dots, \boldsymbol{e}_{n'}$. Тогда мы имеем $\boldsymbol{e}_{j'} = c^j_{j'}\boldsymbol{e}_j$, $\varepsilon^{i'} = c^{i'}_i\varepsilon^i$ и

$$\begin{aligned}T^{i'_1,\dots,i'_p}_{j'_1,\dots,j'_q} &= \mathcal{T}(\varepsilon^{i'_1}, \dots, \varepsilon^{i'_p}, \boldsymbol{e}_{j'_1}, \dots, \boldsymbol{e}_{j'_q}) = \mathcal{T}(c^{i'_1}_{i_1}\varepsilon^{i_1}, \dots, c^{i'_p}_{i_p}\varepsilon^{i_p}, c^{j_1}_{j'_1}\boldsymbol{e}_{j_1}, \dots, c^{j_q}_{j'_q}\boldsymbol{e}_{j_q}) = \\ &= c^{i'_1}_{i_1}\cdots c^{i'_p}_{i_p}c^{j_1}_{j'_1}\cdots c^{j_q}_{j'_q}\mathcal{T}(\varepsilon^{i_1}, \dots, \varepsilon^{i_p}, \boldsymbol{e}_{j_1}, \dots, \boldsymbol{e}_{j_q}) = c^{i'_1}_{i_1}\cdots c^{i'_p}_{i_p}c^{j_1}_{j'_1}\cdots c^{j_q}_{j'_q}T^{i_1,\dots,i_p}_{j_1,\dots,j_q}.\end{aligned}$$

Определение 6.2.1. *Тензором* типа $(p, q)$ называется соответствие

$$\text{базисы в } V \quad \longmapsto \quad \text{наборы из } n^{p+q} \text{ чисел } T = \{T^{i_1,\dots,i_p}_{j_1,\dots,j_q}\},$$

при котором наборы $T = \{T^{i_1,\dots,i_p}_{j_1,\dots,j_q}\}$ и $T' = \{T^{i'_1,\dots,i'_p}_{j'_1,\dots,j'_q}\}$, соответствующие различным базисам $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ и $\boldsymbol{e}_{1'}, \dots, \boldsymbol{e}_{n'}$, связаны соотношением

$$T^{i'_1,\dots,i'_p}_{j'_1,\dots,j'_q} = c^{i'_1}_{i_1}\cdots c^{i'_p}_{i_p}c^{j_1}_{j'_1}\cdots c^{j_q}_{j'_q}T^{i_1,\dots,i_p}_{j_1,\dots,j_q}.$$

Это соотношение называется *тензорным законом преобразования*. Числа $T^{i_1,\dots,i_p}_{j_1,\dots,j_q}$ называются *компонентами* тензора $T$.

Тензоры типа $(p, q)$ образуют линейное пространство $\mathrm{T}^p_q(V)$ относительно операций покомпонентного сложения и умножения на числа.

Полилинейная функция $\mathcal{T} \in \mathrm{P}^p_q(V)$ определяет тензор $T \in \mathrm{T}^p_q(V)$ по формуле (6.2). Обратно, тензор определяет полилинейную функцию по формуле (6.1). Таким образом, пространство полилинейных функций $\mathrm{P}^p_q(V)$ можно отождествить с пространством тензоров $\mathrm{T}^p_q(V)$. Это соответствие обобщает соответствие между билинейными функциями (или линейными операторами) и их матрицами.

Далее мы не будем различать полилинейные функции и тензоры. Мы будем говорить, например, что операторы — это тензоры типа $(1,1)$. Более точно, тензор, соответствующий оператору, — это сопоставление каждому базису матрицы оператора в этом базисе.

Пример 6.2.2.
1. Скаляры $\lambda \in \mathbf{k}$ естественно считать тензорами типа $(0,0)$: они не меняются при замене базиса.

2. Векторы — это тензоры типа $(1,0)$. Тензорный закон преобразования $v^{i'} = c_i^{i'} v^i$ описывает изменение координат вектора при замене базиса (теорема 1.6.3).

3. Ковекторы (линейные функции) — это тензоры типа $(0,1)$. Тензорный закон преобразования $\xi_{i'} = c_{i'}^i \xi_i$ — это закон преобразования координат линейной функции при замене базиса (см. раздел 1.9).

4. Билинейные функции — это тензоры типа $(0,2)$. Тензорный закон преобразования $T_{i'j'} = c_{i'}^i c_{j'}^j T_{ij}$ — это закон изменения матрицы билинейной функции (в матричном виде: $T' = C^t T C$, см. теорему 5.1.3).

5. Линейные операторы — это тензоры типа $(1,1)$. Тензорный закон преобразования $T_{j'}^{i'} = c_i^{i'} c_{j'}^j T_j^i$ — это закон изменения матрицы оператора (в матричном виде: $T' = C^{-1} T C$, см. теорему 2.1.2).

## 6.3. Тензорное произведение, свёртка, опускание и поднятие индексов

**Тензорное произведение.**

Определение 6.3.1. *Тензорным произведением* двух полилинейных функций $\mathcal{T} \in \mathrm{P}_q^p(V)$ и $\mathcal{S} \in \mathrm{P}_s^r(V)$ называется полилинейная функция $\mathcal{T} \otimes \mathcal{S} \in \mathrm{P}_{q+s}^{p+r}(V)$, заданная по формуле

$$\mathcal{T} \otimes \mathcal{S}(\xi^1, \dots, \xi^p, \xi^{p+1}, \dots, \xi^{p+r}, \boldsymbol{v}_1, \dots, \boldsymbol{v}_q, \boldsymbol{v}_{q+1}, \dots, \boldsymbol{v}_{q+s}) = \\ = \mathcal{T}(\xi^1, \dots, \xi^p, \boldsymbol{v}_1, \dots, \boldsymbol{v}_q) \cdot \mathcal{S}(\xi^{p+1}, \dots, \xi^{p+r}, \boldsymbol{v}_{q+1}, \dots, \boldsymbol{v}_{q+s}).$$

Полилинейной функции $\mathcal{T} \otimes \mathcal{S}$ соответствует тензор $T \otimes S \in \mathrm{T}_{q+s}^{p+r}(V)$, компоненты которого задаются формулой

$$(T \otimes S)_{j_1,\dots,j_q,j_{q+1},\dots,j_{q+s}}^{i_1,\dots,i_p,i_{p+1},\dots,i_{p+r}} = T_{j_1,\dots,j_q}^{i_1,\dots,i_p} \cdot S_{j_{q+1},\dots,j_{q+s}}^{i_{p+1},\dots,i_{p+r}}.$$

Тензор $T \otimes S$ называется *тензорным произведением* тензоров $T$ и $S$.

Операция тензорного произведения, очевидно, ассоциативна (т.е. $(T \otimes S) \otimes R = T \otimes (S \otimes R)$) и дистрибутивна (т.е. $(\lambda T + \mu S) \otimes R = \lambda T \otimes R + \mu S \otimes R$), но не коммутативна (вообще говоря, $T \otimes S \neq S \otimes T$).

Сумма и тензорное произведение превращают множество тензоров (всех типов) в пространстве $V$ в кольцо (точнее, алгебру над полем $\mathbf{k}$), которая называется *тензорной алгеброй* или *алгеброй Грассмана* и обозначается $\mathrm{T}(V)$.

Пример 6.3.2. Пусть $\xi, \eta \in V^*$ — линейные функции, т.е. тензоры типа $(0,1)$. Их тензорное произведение $\xi \otimes \eta$ является тензором типа $(0,2)$, т.е. билинейной функцией. По определению, значение этой билинейной функции на паре векторов задаётся формулой $(\xi \otimes \eta)(\boldsymbol{u}, \boldsymbol{v}) = \xi(\boldsymbol{u}) \cdot \eta(\boldsymbol{v})$.

**Свёртка.** Пусть теперь $T = \{T^{i_1,\dots,i_p}_{j_1,\dots,j_q}\}$ — тензор с хотя бы одним верхним и нижним индексом, т.е. $p > 0$ и $q > 0$. Зафиксируем один верхний и один нижний индекс (пусть для простоты это будут первые индексы) и сформируем следующий новый набор из $n^{p+q-2}$ чисел:

$$cT = \{T^{k,i_2,\dots,i_p}_{k,j_2,\dots,j_q}\},$$

где как обычно по повторяющемуся индексу $k$ подразумевается суммирование.

Предложение 6.3.3. $cT = \{T^{k,i_2,\dots,i_p}_{k,j_2,\dots,j_q}\}$ *является тензором типа* $(p-1, q-1)$.

Доказательство. Необходимо проверить тензорный закон. Мы имеем

$$\begin{aligned}(cT)^{i'_2,\dots,i'_p}_{j'_2,\dots,j'_q} = T^{k',i'_2,\dots,i'_p}_{k',j'_2,\dots,j'_q} = c^{k'}_{i_1} c^{i'_2}_{i_2} \cdots c^{i'_p}_{i_p} c^{j_1}_{k'} c^{j_2}_{j'_2} \cdots c^{j_q}_{j'_q} T^{i_1,i_2,\dots,i_p}_{j_1,j_2,\dots,j_q} = \\ = c^{i'_2}_{i_2} \cdots c^{i'_p}_{i_p} c^{j_2}_{j'_2} \cdots c^{j_q}_{j'_q} c^{k'}_{i_1} c^{j_1}_{k'} T^{i_1,i_2,\dots,i_p}_{j_1,j_2,\dots,j_q} = c^{i'_2}_{i_2} \cdots c^{i'_p}_{i_p} c^{j_2}_{j'_2} \cdots c^{j_q}_{j'_q} \delta^{j_1}_{i_1} T^{i_1,i_2,\dots,i_p}_{j_1,j_2,\dots,j_q} = \\ = c^{i'_2}_{i_2} \cdots c^{i'_p}_{i_p} c^{j_2}_{j'_2} \cdots c^{j_q}_{j'_q} T^{k,i_2,\dots,i_p}_{k,j_2,\dots,j_q} = c^{i'_2}_{i_2} \cdots c^{i'_p}_{i_p} c^{j_2}_{j'_2} \cdots c^{j_q}_{j'_q} (cT)^{i_2,\dots,i_p}_{j_2,\dots,j_q},\end{aligned}$$

что и требовалось. □

Определение 6.3.4. Тензор $cT = \{T^{k,i_2,\dots,i_p}_{k,j_2,\dots,j_q}\}$ называется *свёрткой* тензора $T = \{T^{i_1,i_2,\dots,i_p}_{j_1,j_2,\dots,j_q}\}$ по (первым) верхнему и нижнему индексам. Свёртка, очевидно, задаёт линейное отображение $\mathrm{T}^p_q(V) \to \mathrm{T}^{p-1}_{q-1}(V)$.

Операцию свёртки можно проводить несколько раз до исчерпания верхних или нижних индексов. Последняя возможная свёртка (после которой не остаётся либо верхних, либо нижних индексов) называется *полной свёрткой*.

Пример 6.3.5.

1. Пусть $\mathcal{A}$ — оператор, т.е. тензор типа $(1,1)$. Результатом его свёртки будет тензор типа $(0,0)$, т.е. скаляр. Этот скаляр — это сумма $a^i_i$ диагональных элементов матрицы оператора $\mathcal{A}$ в любом базисе, т.е. след оператора: $c\mathcal{A} = \operatorname{tr}\mathcal{A}$. Проверка тензорного закона для свёртки в данном случае сводится к проверке независимости следа от базиса (лемма 2.1.3).

2. Пусть $B = \{b_{j_1 j_2}\}$ — билинейная функция (тензор типа $(0,2)$), а $u = \{u^{i_1}\}$, $v = \{v^{i_2}\}$ — векторы (тензоры типа $(1,0)$). Рассмотрим тензор $B \otimes u \otimes v = \{b_{j_1 j_2} u^{i_1} v^{i_2}\}$ типа $(2,2)$. Его полная свёртка есть скаляр $b_{kl} u^k v^l = B(u,v)$ — значение билинейной функции на данной паре векторов.

**Опускание и поднятие индексов.** Пусть $V$ — евклидово пространство. Тогда соответствие $\boldsymbol{v} \mapsto (\boldsymbol{v}, \cdot\,)$ задаёт канонический изоморфизм $V \to V^*$, т.е. позволяет отождествить тензоры типа $(1,0)$ с тензорами типа $(0,1)$. В координатах это выглядит следующим образом. Пусть $G = (g_{ij})$ — матрица Грама скалярного произведения. Так как скалярное произведение — это билинейная функция, $G$ является тензором типа $(0,2)$. Тогда при изоморфизме $V \to V^*$ вектор $T$ с координатами $T^i$ переходит в ковектор с координатами $T_j = g_{ij} T^i$. Таким образом, мы «опустили индекс» у тензора $T$ при помощи фиксированного тензора $G$ типа $(0,2)$. Эта операция обобщается следующим образом:

Определение 6.3.6. *Опускание индекса* — это линейное отображение $\mathrm{T}^p_q(V) \to \mathrm{T}^{p-1}_{q+1}(V)$ тензоров в евклидовом пространстве $V$, которое тензору $T = \{T^{i_1,\dots,i_p}_{j_1,\dots,j_q}\}$ типа $(p,q)$ ставит в соответствие тензор $\{g_{ij} T^{i,i_2,\dots,i_p}_{j_1,\dots,j_q}\}$ типа $(p-1, q+1)$.

Пример 6.3.7. Рассмотрим изоморфизм $\psi$: $\mathrm{End}(V) \to \mathrm{B}(V)$ между пространством операторов и пространством билинейных функций (см предложение 5.6.1), сопоставляющий оператору $\mathcal{A}$ билинейную функцию $\mathcal{B}_\mathcal{A} = (\mathcal{A}\,\cdot\,,\,\cdot\,)$. В координатах это выглядит так: оператору с матрицей $A = (a^i_k)$ сопоставляется билинейная функция с матрицей $B_A = (g_{ij}a^i_k)$. Таким образом, $B_A$ — это тензор типа $(0,2)$, получаемый в результате опускания индекса у тензора $A$ типа $(1,1)$.

Для того, чтобы определить операцию, обратную к опусканию индекса, рассмотрим набор $\{g^{kl}\}$, состоящий из элементов обратной матрицы к матрице Грама $G = (g_{ij})$, т.е. $g^{kl}g_{lj} = \delta^k_j$.

Предложение 6.3.8. *$\{g^{kl}\}$ является тензором типа $(2,0)$.*

Доказательство. Необходимо проверить тензорный закон. В новом базисе мы имеем $g^{k'i'}g_{i'j'} = \delta^{k'}_{j'}$. Так как $\{g_{ij}\}$ — тензор типа $(0,2)$, мы имеем $g_{i'j'} = c^i_{i'}c^j_{j'}g_{ij}$. Подставив это в предыдущую формулу, получим $g^{k'i'}c^i_{i'}c^j_{j'}g_{ij} = \delta^{k'}_{j'}$. Теперь умножим обе части этого равенства на компоненты обратной матрицы $c^{j'}_k$ (и просуммируем по $j'$): $g^{k'i'}c^i_{i'}c^j_{j'}c^{j'}_k g_{ij} = \delta^{k'}_{j'}c^{j'}_k$. Так как $c^j_{j'}c^{j'}_k = \delta^j_k$, отсюда получаем $g^{k'i'}c^i_{i'}g_{ik} = c^{k'}_k$. Далее умножим обе части на $g^{kl}$: $g^{k'i'}c^i_{i'}g_{ik}g^{kl} = c^{k'}_k g^{kl}$. Так как $g_{ik}g^{kl} = \delta^l_i$, получаем $g^{k'i'}c^l_{i'} = c^{k'}_k g^{kl}$. Наконец, умножим обе части на $c^{l'}_l$: $g^{k'i'}c^l_{i'}c^{l'}_l = c^{k'}_k c^{l'}_l g^{kl}$. Так как $c^l_{i'}c^{l'}_l = \delta^{l'}_{i'}$, окончательно получаем $g^{k'l'} = c^{k'}_k c^{l'}_l g^{kl}$. Это и есть тензорный закон преобразования для тензора типа $(2,0)$. □

Определение 6.3.9. *Поднятие индекса* — это линейное отображение $\mathrm{T}^p_q(V) \to \mathrm{T}^{p+1}_{q-1}(V)$ тензоров в евклидовом пространстве $V$, которое тензору $T = \{T^{i_1,\dots,i_p}_{j_1,\dots,j_q}\}$ типа $(p,q)$ ставит в соответствие тензор $\{g^{ij}T^{i_1,\dots,i_p}_{j,j_2,\dots,j_q}\}$ типа $(p+1,q-1)$.

Операция опускания (или поднятия) индекса является композицией тензорного произведения с тензором $\{g_{ij}\}$ (или $\{g^{ij}\}$) и свёртки.

## 6.4. Базис в пространстве тензоров

Базис в пространстве тензоров $\mathrm{T}^p_q(V)$ можно задать при помощи операции тензорного произведения.

Рассмотрим полилинейную функцию $\varepsilon^{j_1}\otimes\ldots\otimes\varepsilon^{j_q}\otimes\boldsymbol{e}_{i_1}\otimes\ldots\otimes\boldsymbol{e}_{i_p}$. По определению, её значение на $q$ векторах и $p$ ковекторах задаётся формулой

$$(\boldsymbol{e}_{i_1}\otimes\ldots\otimes\boldsymbol{e}_{i_p}\otimes\varepsilon^{j_1}\otimes\ldots\otimes\varepsilon^{j_q})(\xi^1,\ldots,\xi^p,\boldsymbol{v}_1,\ldots,\boldsymbol{v}_q) = \\ = \boldsymbol{e}_{i_1}(\xi^1)\cdots\boldsymbol{e}_{i_p}(\xi^p)\,\varepsilon^{j_1}(\boldsymbol{v}_1)\cdots\varepsilon^{j_q}(\boldsymbol{v}_q) = \xi^1_{i_1}\cdots\xi^p_{i_p}v^{j_1}_1\cdots v^{j_q}_q,$$

а соответствующий тензор типа $(p,q)$ имеет компоненты

$$T^{k_1,\dots,k_p}_{l_1,\dots,l_q} = (\boldsymbol{e}_{i_1}\otimes\ldots\otimes\boldsymbol{e}_{i_p}\otimes\varepsilon^{j_1}\otimes\ldots\otimes\varepsilon^{j_q})(\varepsilon^{k_1},\ldots,\varepsilon^{k_p},\boldsymbol{e}_{l_1},\ldots,\boldsymbol{e}_{l_q}) = \delta^{k_1}_{i_1}\cdots\delta^{k_p}_{i_p}\delta^{j_1}_{l_1}\cdots\delta^{j_q}_{l_p}$$

(другими словами, компонента этого тензора с верхними индексами $i_1,\ldots,i_p$ и нижними индексами $j_1,\ldots,j_q$ — единица, а остальные компоненты — нули).

Теорема 6.4.1. *Тензоры $\boldsymbol{e}_{i_1}\otimes\ldots\otimes\boldsymbol{e}_{i_p}\otimes\varepsilon^{j_1}\otimes\ldots\otimes\varepsilon^{j_q}$, отвечающие всевозможным значениями индексов $i_1,\ldots,i_p$ и $j_1,\ldots,j_q$, образуют базис в пространстве $\mathrm{T}^p_q(V)$.*

Доказательство. Сначала докажем линейную независимость данных тензоров. Предположим, существуют такие числа $\lambda^{i_1,\dots,i_p}_{j_1,\dots,j_q}$, что линейная комбинация $\lambda^{i_1,\dots,i_p}_{j_1,\dots,j_q}\boldsymbol{e}_{i_1}\otimes\ldots\otimes\boldsymbol{e}_{i_p}\otimes\varepsilon^{j_1}\otimes\ldots\otimes\varepsilon^{j_q}$ равна нулю. Применив эту полилинейную функцию к аргументам $\varepsilon^{k_1},\dots,\varepsilon^{k_p},\boldsymbol{e}_{l_1},\dots,\boldsymbol{e}_{l_q}$, получим

$$(6.3)\quad 0=(\lambda^{i_1,\dots,i_p}_{j_1,\dots,j_q}\boldsymbol{e}_{i_1}\otimes\ldots\otimes\boldsymbol{e}_{i_p}\otimes\varepsilon^{j_1}\otimes\ldots\otimes\varepsilon^{j_q})(\varepsilon^{k_1},\dots,\varepsilon^{k_p},\boldsymbol{e}_{l_1},\dots,\boldsymbol{e}_{l_q})=$$
$$=\lambda^{i_1,\dots,i_p}_{j_1,\dots,j_q}\delta^{k_1}_{i_1}\cdots\delta^{k_p}_{i_p}\delta^{j_1}_{l_1}\cdots\delta^{j_q}_{l_p}=\lambda^{k_1,\dots,k_p}_{l_1,\dots,l_q},$$

т.е. все коэффициенты линейной комбинации равны нулю.

Теперь докажем, что любой тензор $T=\{T^{i_1,\dots,i_p}_{j_1,\dots,j_q}\}$ представляется в виде линейной комбинации данных тензоров. А именно, докажем, что

$$T=T^{i_1,\dots,i_p}_{j_1,\dots,j_q}\boldsymbol{e}_{i_1}\otimes\ldots\otimes\boldsymbol{e}_{i_p}\otimes\varepsilon^{j_1}\otimes\ldots\otimes\varepsilon^{j_q},$$

т.е. координаты тензора в данном базисе суть его компоненты. В силу полилинейности это равенство достаточно проверить на наборах базисных векторов и ковекторов, т.е. на аргументах вида $\varepsilon^{k_1},\dots,\varepsilon^{k_p},\boldsymbol{e}_{l_1},\dots,\boldsymbol{e}_{l_q}$. При подстановке этих аргументов в левую часть мы по определению получаем $T^{k_1,\dots,k_p}_{l_1,\dots,l_q}$, а при подстановке в правую часть мы получаем то же самое в силу выкладки, аналогичной (6.3). □

Следствие 6.4.2. *Размерность пространства* $\mathrm{T}^p_q(V)$ *равна* $n^{p+q}$.

## 6.5. Симметрические и кососимметрические тензоры, симметризация и альтернирование

Здесь мы будем рассматривать только тензоры с нижними индексами. Мы будем обозначать через $\Sigma_q$ группу перестановок индексов $1,\dots,q$. Через $(-1)^\sigma$ будем обозначать знак перестановки $\sigma\in\Sigma_q$.

Определение 6.5.1. Полилинейная функция $\mathcal{T}\in\mathrm{P}^0_q(V)$ называется *симметрической*, если

$$\mathcal{T}(\boldsymbol{v}_{\sigma(1)},\dots,\boldsymbol{v}_{\sigma(q)})=\mathcal{T}(\boldsymbol{v}_1,\dots,\boldsymbol{v}_q),$$

и *кососимметрической*, если

$$\mathcal{T}(\boldsymbol{v}_{\sigma(1)},\dots,\boldsymbol{v}_{\sigma(q)})=(-1)^\sigma\mathcal{T}(\boldsymbol{v}_1,\dots,\boldsymbol{v}_q),$$

для любых векторов $\boldsymbol{v}_1,\dots,\boldsymbol{v}_q$ и перестановки $\sigma\in\Sigma_q$.

Компоненты тензора $T\in\mathrm{T}^0_q(V)$, соответствующего симметрической полилинейной функции, удовлетворяет соотношениям

$$T_{j_1,\dots,j_q}=T_{j_{\sigma(1)},\dots,j_{\sigma(q)}},$$

а компоненты тензора, соответствующего кососимметрической полилинейной функции, — соотношениям

$$T_{j_1,\dots,j_q}=(-1)^\sigma T_{j_{\sigma(1)},\dots,j_{\sigma(q)}}.$$

Такие тензоры называются, соответственно, *симметрическими* и *кососимметрическими*. В частности, у косимметрического тензора могут быть отличны от нуля лишь компоненты $T_{j_1,\dots,j_q}$, у которых все индексы различны.

Симметрические и кососимметрические тензоры образуют подпространства в пространстве тензоров $\mathrm{T}^0_q(V)$, которые обозначаются $\mathrm{S}_q(V)$ и $\Lambda_q(V)$ соответственно.

При $q = 2$ симметрические и кососимметрические тензоры — это симметрические и кососимметрические билинейные функции соответственно.

Замечание. Хотя подпространства $\mathrm{S}_q(V)$ и $\Lambda_q(V)$ и образуют прямую сумму в пространстве $\mathrm{T}^0_q(V)$, при $q > 2$ разложение $\mathrm{T}^0_q(V) = \mathrm{S}_q(V) \oplus \Lambda_q(V)$ не имеет места, в отличие от случая билинейных функций.

Определение 6.5.2.
*Симметризацией* называется линейный оператор $\mathrm{Sym}\colon \mathrm{T}^0_q(V) \to \mathrm{T}^0_q(V)$, который тензору $T \in \mathrm{T}^0_q(V)$ ставит в соответствие тензор $\mathrm{Sym}\, T$ с компонентами

$$(\mathrm{Sym}\, T)_{j_1,\dots,j_q} := \frac{1}{q!} \sum_{\sigma\in\Sigma_q} T_{j_{\sigma(1)},\dots,j_{\sigma(q)}}.$$

*Альтернированием* называется линейный оператор $\mathrm{Alt}\colon \mathrm{T}^0_q(V) \to \mathrm{T}^0_q(V)$, который тензору $T \in \mathrm{T}^0_q(V)$ ставит в соответствие тензор $\mathrm{Alt}\, T$ с компонентами

$$(\mathrm{Alt}\, T)_{j_1,\dots,j_q} := \frac{1}{q!} \sum_{\sigma\in\Sigma_q} (-1)^\sigma T_{j_{\sigma(1)},\dots,j_{\sigma(q)}}.$$

Легко видеть, что $\mathrm{Sym}\, T$ — симметрический тензор, а $\mathrm{Alt}\, T$ — кососимметрический тензор для любого $T \in \mathrm{T}^0_q(V)$.

Предложение 6.5.3. *Операторы* Sym *и* Alt *являются проекторами на подпространства* $\mathrm{S}_q(V)$ *и* $\Lambda_q(V)$ *соответственно.*

Доказательство. Оба утверждения доказываются аналогично. Докажем второе. В силу теоремы 2.2.2, достаточно показать, что $\mathrm{Alt}\, T = T$ для любого кососимметрического тензора $T$. Мы имеем

$$(\mathrm{Alt}\, T)_{j_1,\dots,j_q} := \frac{1}{q!} \sum_{\sigma\in\Sigma_q} (-1)^\sigma T_{j_{\sigma(1)},\dots,j_{\sigma(q)}} = \frac{1}{q!} \sum_{\sigma\in\Sigma_q} T_{j_1,\dots,j_q} = T_{j_1,\dots,j_q},$$

где второе равенство выполнено в силу кососимметричности $T$. □

## 6.6. Внешнее произведение кососимметрических тензоров, внешние формы

Для кососимметрических тензоров имеется аналог тензорного произведения:

Определение 6.6.1. *Внешним произведением* кососимметрических тензоров $P \in \Lambda_p(V)$ и $Q \in \Lambda_q(V)$ называется кососимметрический тензор

$$P \wedge Q := \frac{(p+q)!}{p!\, q!} \mathrm{Alt}(P \otimes Q)$$

(роль коэффициента будет объяснена ниже).

Теорема 6.6.2. *Внешнее произведение кососимметрических тензоров обладает следующими свойствами: для любых* $P \in \Lambda_p(V)$, $Q \in \Lambda_q(V)$, $R \in \Lambda_r(V)$ *и* $\lambda, \mu \in \mathbf{k}$

а) $(\lambda P + \mu Q) \wedge R = \lambda P \wedge R + \mu Q \wedge R$ *(дистрибутивность, при* $p = q$*);*
б) $Q \wedge P = (-1)^{pq} P \wedge Q$ *(антикоммутативность);*
в) $(P \wedge Q) \wedge R = P \wedge (Q \wedge R)$ *(ассоциативность).*

Доказательство. а) Дистрибутивность вытекает из дистрибутивности операции $\otimes$ и линейности оператора Alt.

б) Для доказательства антикоммутативности достаточно проверить, что имеет место соотношение $\mathrm{Alt}(Q \otimes P) = (-1)^{pq} \mathrm{Alt}(P \otimes Q)$. Введём перестановку

$$\tau = \begin{pmatrix} 1 & \dots & p & p+1 & \dots & p+q \\ q+1 & \dots & q+p & 1 & \dots & q \end{pmatrix}.$$

Тогда $(-1)^\tau = (-1)^{pq}$, так как $\tau$ — результат композиции $pq$ элементарных подстановок. Мы имеем

$$\mathrm{Alt}(Q \otimes P)_{i_1,\dots,i_{p+q}} = \frac{1}{(p+q)!} \sum_{\sigma \in \Sigma_{p+q}} (-1)^\sigma Q_{i_{\sigma(1)},\dots,i_{\sigma(q)}} P_{i_{\sigma(q+1)},\dots,i_{\sigma(q+p)}} =$$
$$= \frac{1}{(p+q)!} \sum_{\sigma \in \Sigma_{p+q}} (-1)^\tau (-1)^{\sigma\tau} P_{i_{\sigma\tau(1)},\dots,i_{\sigma\tau(p)}} Q_{i_{\sigma\tau(p+1)},\dots,i_{\sigma\tau(p+q)}} =$$
$$= (-1)^\tau \frac{1}{(p+q)!} \sum_{\varphi \in \Sigma_{p+q}} (-1)^\varphi P_{i_{\varphi(1)},\dots,i_{\varphi(p)}} Q_{i_{\varphi(p+1)},\dots,i_{\varphi(p+q)}} = (-1)^{pq} \mathrm{Alt}(P \otimes Q)_{i_1,\dots,i_{p+q}}.$$

Тем самым б) доказано.

Далее мы будем использовать следующие обозначения. Для $\sigma \in \Sigma_p$ и $P \in \mathrm{T}^0_p(V)$ обозначим через $\sigma P$ тензор с компонентами $(\sigma P)_{i_1,\dots,i_p} := P_{i_{\sigma(1)},\dots,i_{\sigma(p)}}$. По определению альтернирования, $\mathrm{Alt}\, P = \frac{1}{p!} \sum_{\sigma \in \Sigma_p} (-1)^\sigma \sigma P$, и имеет место соотношение

$$\sigma(\mathrm{Alt}\, P) = \mathrm{Alt}(\sigma P) = (-1)^\sigma \mathrm{Alt}\, P.$$

Для доказательства в) нам понадобится лемма:

Лемма 6.6.3. *Для любых тензоров $P \in \mathrm{T}^0_p(V)$ и $Q \in \mathrm{T}^0_q(V)$ имеем*

$$\mathrm{Alt}((\mathrm{Alt}\, P) \otimes Q) = \mathrm{Alt}(P \otimes Q) = \mathrm{Alt}(P \otimes \mathrm{Alt}(Q)).$$

Доказательство. Докажем лишь первое равенство (второе доказывается аналогично). Поскольку операция $\otimes$ дистрибутивна, а оператор Alt линеен, имеем

$$\mathrm{Alt}((\mathrm{Alt}\, P) \otimes Q) = \mathrm{Alt}\Big(\Big(\frac{1}{p!} \sum_{\sigma \in \Sigma_p} (-1)^\sigma \sigma P\Big) \otimes Q\Big) = \frac{1}{p!} \sum_{\sigma \in \Sigma_p} (-1)^\sigma \mathrm{Alt}(\sigma P \otimes Q).$$

Каждой перестановке $\sigma \in \Sigma_p$ сопоставим перестановку $\widetilde{\sigma} \in \Sigma_{p+q}$, которая на первых $p$ индексах действует как $\sigma$, а остальные оставляет на месте, т.е.

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

При этом, очевидно, $(-1)^{\widetilde{\sigma}} = (-1)^\sigma$ и $\sigma P \otimes Q = \widetilde{\sigma}(P \otimes Q)$. Тогда имеем

$$\mathrm{Alt}((\mathrm{Alt}\, P) \otimes Q) = \frac{1}{p!} \sum_{\sigma \in \Sigma_p} (-1)^\sigma \mathrm{Alt}(\widetilde{\sigma}(P \otimes Q)) =$$
$$= \frac{1}{p!} \sum_{\sigma \in \Sigma_p} (-1)^\sigma (-1)^{\widetilde{\sigma}} \mathrm{Alt}(P \otimes Q) = \frac{1}{p!} \sum_{\sigma \in \Sigma_p} \mathrm{Alt}(P \otimes Q) = \mathrm{Alt}(P \otimes Q).$$

$\square$

Теперь вернёмся к доказательству теоремы 6.6.2 в). Мы имеем

$$(P \wedge Q) \wedge R = \frac{(p+q+r)!}{(p+q)!\,r!} \operatorname{Alt}\bigl((P \wedge Q) \otimes R\bigr) =$$
$$= \frac{(p+q+r)!}{(p+q)!\,r!} \operatorname{Alt}\Bigl(\frac{(p+q)!}{p!\,q!} \operatorname{Alt}(P \otimes Q) \otimes R\Bigr) = \frac{(p+q+r)!}{p!\,q!\,r!} \operatorname{Alt}(P \otimes Q \otimes R),$$

где в последнем равенстве мы воспользовались предыдущей леммой. Аналогично проверяется, что $P \wedge (Q \wedge R) = \frac{(p+q+r)!}{p!\,q!\,r!} \operatorname{Alt}(P \otimes Q \otimes R)$. □

Выберем теперь базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ в $V$, и пусть $\varepsilon^1, \ldots, \varepsilon^n$ — двойственный базис в $V^* = \mathrm{T}^0_1(V) = \Lambda_1(V)$. Рассмотрим кососимметрические тензоры

$$\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p} = p! \operatorname{Alt}(\varepsilon^{i_1} \otimes \ldots \otimes \varepsilon^{i_p}) = \sum_{\sigma \in \Sigma_p} (-1)^\sigma \varepsilon^{i_{\sigma(1)}} \otimes \ldots \otimes \varepsilon^{i_{\sigma(p)}}. \tag{6.4}$$

В силу теоремы 6.6.2 б) имеем $\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p} = (-1)^\sigma \varepsilon^{i_{\sigma(1)}} \wedge \ldots \wedge \varepsilon^{i_{\sigma(p)}}$ для $\sigma \in \Sigma_p$.

Замечание. Обратим внимание, что ввиду выбора коэффициента $\frac{(p+q)!}{p!\,q!}$ в определении внешнего произведения, выражение $\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p}$ оказалось линейной комбинацией выражений $\varepsilon^{i_{\sigma(1)}} \otimes \ldots \otimes \varepsilon^{i_{\sigma(p)}}$ *с целыми коэффициентами.* Таким образом, кососимметрические тензоры $\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p}$ определены над конечными полями (характеристики, отличной от двух) и даже над целыми числами, что удобно с алгебраической точки зрения.

Теорема 6.6.4. *Кососимметрические тензоры* $\{\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p},\ i_1 < \ldots < i_p\}$ *образуют базис в пространстве* $\Lambda_p(V)$. *В частности,* $\dim \Lambda_p(V) = C^p_n$.

Доказательство. Сначала докажем, что любой кососимметрический тензор $T \in \mathrm{T}^0_p(V)$ представляется в виде линейной комбинации данных тензоров. Разложим $T$ по базису $\varepsilon^{i_1} \otimes \ldots \otimes \varepsilon^{i_p}$:

$$T = T_{i_1,\ldots,i_p} \varepsilon^{i_1} \otimes \ldots \otimes \varepsilon^{i_p}.$$

Теперь применим к обеим частям оператор Alt. Так как $T$ — кососимметрический тензор, $\operatorname{Alt} T = T$. С другой стороны, $\operatorname{Alt}(\varepsilon^{i_1} \otimes \ldots \otimes \varepsilon^{i_p}) = \frac{1}{p!} \varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p}$ по определению $\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p}$. Итак, получаем

$$T = \frac{1}{p!} \sum_{i_1,\ldots,i_p} T_{i_1,\ldots,i_p} \varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p} = \sum_{i_1 < \ldots < i_p} T_{i_1,\ldots,i_p} \varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p},$$

что и даёт требуемое представление в виде линейной комбинации.

Предположим, что тензоры $\{\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p},\ i_1 < \ldots < i_p\}$ линейно зависимы, т.е.

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

Тогда из (6.4) получаем

$$\sum_{i_1 < \ldots < i_p} \lambda_{i_1,\ldots,i_p} \sum_{\sigma \in \Sigma_p} (-1)^\sigma \varepsilon^{i_{\sigma(1)}} \otimes \ldots \otimes \varepsilon^{i_{\sigma(p)}} = 0.$$

В этой сумме все тензоры $\varepsilon^{i_{\sigma(1)}} \otimes \ldots \otimes \varepsilon^{i_{\sigma(p)}}$ различны, следовательно они линейно независимы. Отсюда получаем, что $\lambda_{i_1,\ldots,i_p} = 0$. □

Определение 6.6.5. Выражение $T = \sum_{i_1 < \ldots < i_p} T_{i_1, \ldots, i_p} \varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p}$ (представляющее собой разложение кососимметрического тензора $T \in \Lambda_p(V)$ по базису $\{\varepsilon^{i_1} \wedge \ldots \wedge \varepsilon^{i_p},\, i_1 < \ldots < i_p\}$) называется *внешней p-формой*.

Внешние формы складываются и умножаются аналогично многочленам от $n$ переменных. Различие заключается в том, что «переменные» $\varepsilon^1, \ldots, \varepsilon^n$ антикоммутируют, т.е. удовлетворяют соотношениям $\varepsilon^i \wedge \varepsilon^j = -\varepsilon^j \wedge \varepsilon^i$ (в частности, $\varepsilon^i \wedge \varepsilon^i = 0$). Это предоставляет очень удобный формализм для работы с кососимметрическими тензорами.