Введение в линейную алгебру

Кожевников Павел Александрович

14 июня 2020 г.

Глава 1

Векторные пространства

§ 1. Векторные пространства и подпространства

Аксиомы и их следствия

Бинарная операция на множестве Z — это некоторое отображение $\varphi : Z \times Z \to Z$, таким образом для каждой упорядоченной пары (a, b) элементов множества Z определен элемент $\varphi(a, b) \in Z$ — *результат* применения операции φ к a и b. Ниже будут встречаться бинарные операции в записи, более привычной для арифметических действий. Скажем, операцию можно обозначить «+» и тогда вместо $\varphi(a, b)$ писать $a + b$. *Унарная операция* на множестве Z — это просто отображение $Z \to Z$. Скажем, умножение на число 2 формально можно считать таким отображением $\psi : \mathbb{F} \to \mathbb{F}$, что $\forall x \in \mathbb{F}\ \psi(x) = 2x$.

Далее предполагаем, что V — произвольное множество, а $\mathbb{F}$ — некоторое поле (скажем, поле $\mathbb{R}$), при этом на V определена бинарная операция, которую называем *сложением* и обозначаем «+», а также для каждого $\lambda \in \mathbb{F}$ на V определена унарная операция, которую называем *умножением* на λ. Результат применения умножения на λ к $\mathbf{a} \in V$ обозначаем $\lambda \cdot \mathbf{a}$ или $\lambda\mathbf{a}$.

> **Определение.** V называется *векторным пространством (или линейным пространством) над полем* $\mathbb{F}$*)*, если выполнены следующие свойства V1—V8 (аксиомы векторного пространства):
> V1. $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}) \quad (\forall\ \mathbf{a}, \mathbf{b}, \mathbf{c} \in V)$;
> V2. $\exists\ \mathbf{0} \in V\ \forall\ \mathbf{a} \in V$: $\mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a}$;
> V3. $\forall\ \mathbf{a} \in V\ \exists\ \mathbf{x} \in V$: $\mathbf{x} + \mathbf{a} = \mathbf{a} + \mathbf{x} = \mathbf{0}$;
> V4. $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a} \quad (\forall\ \mathbf{a}, \mathbf{b} \in V)$;
> V5. $\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b} \quad (\forall\ \mathbf{a}, \mathbf{b} \in V,\ \forall\ \lambda \in \mathbb{F}\ )$;
> V6. $(\lambda + \mu)\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a} \quad (\forall\ \mathbf{a} \in V,\ \forall\ \lambda, \mu \in \mathbb{F}\ )$;
> V7. $1 \cdot \mathbf{a} = \mathbf{a} \quad (\forall\ \mathbf{a} \in V)$;
> V8. $(\lambda\mu)\mathbf{a} = \lambda(\mu\mathbf{a}) \quad (\forall\ \mathbf{a} \in V,\ \forall\ \lambda, \mu \in \mathbb{F}\ )$.

На протяжении главы V будет обозначать векторное пространство. Элементы векторного пространства V будем называть *векторами* (независимо от их природы). Элементы поля $\mathbb{F}$ будем иногда называть *константами* или *скалярами*.

Заметим, что аксиомы V1—V4 повторяют аксиомы абелевой группы. Аксиомы V1 и V4 называются аксиомами *ассоциативности* и *коммутативности* сложения. Вектор $\mathbf{0}$ из аксиомы V2 (нейтральный элемент относительно сложения) называется *нулевым* вектором. Вектор $\mathbf{x}$ из аксиомы V3 называется *противоположным* вектором для вектора $\mathbf{a}$ и обозначатся $-\mathbf{a}$. Используя пртивоположный вектор, можно определить операцию *вычитания* равенством $\mathbf{a} - \mathbf{b} := \mathbf{a} + (-\mathbf{b})$, в частности, V2 принимает вид $\mathbf{a} - \mathbf{a} = \mathbf{0}$.

Из аксиом следуют привычные правила, которыми мы пользуемся, скажем, при работе с векторами в геометрии. Например, аксиомы V1 и V4 обеспечивают то, что результат вычисления суммы векторов $\sum\limits_{i=1}^{k} \mathbf{a}_i$ не зависит от порядка выполнения операций. Отметим следствия аксиом V1—V4.

Предложение 1.1. *1). Нулевой вектор единственный;*
2). $\forall\ \mathbf{a} \in V$ *противоположный ему вектор единственный;*
3). (закон сокращения) $\mathbf{a}+\mathbf{b} = \mathbf{a}+\mathbf{c} \Leftrightarrow \mathbf{b} = \mathbf{c} \qquad (\forall\ \mathbf{a}, \mathbf{b}, \mathbf{c} \in V)$*.*

$\rhd$ 1). Пусть $\mathbf{0}$ и $\mathbf{0}'$ — два нулевых вектора. Тогда $\mathbf{0} = \mathbf{0} + \mathbf{0}' = \mathbf{0}'$, то есть $\mathbf{0}$ и $\mathbf{0}'$ совпадают.
2). Пусть векторы $\mathbf{x}$ и $\mathbf{y}$ оба являются противоположными для вектора $\mathbf{a}$. Тогда $\mathbf{x} = \mathbf{x}+\mathbf{0} = \mathbf{x}+(\mathbf{a}+\mathbf{y}) = (\mathbf{x}+\mathbf{a})+\mathbf{y} = \mathbf{0}+\mathbf{y} = \mathbf{y}$, то есть $\mathbf{x}$ и $\mathbf{y}$ совпадают.
3). $\mathbf{a}+\mathbf{b} = \mathbf{a}+\mathbf{c} \Rightarrow -\mathbf{a}+(\mathbf{a}+\mathbf{b}) = -\mathbf{a}+(\mathbf{a}+\mathbf{c}) \Rightarrow (-\mathbf{a}+\mathbf{a})+\mathbf{b} = (-\mathbf{a}+\mathbf{a})+\mathbf{c} \Rightarrow \mathbf{0}+\mathbf{b} = \mathbf{0}+\mathbf{c} \Rightarrow \mathbf{b} = \mathbf{c}$. Обратное следствие очевидно. $\square$

Аксиомы V5 и V6 можно назвать свойством линейности, или дистрибутивности (по векторам и по константам соответственно), V7 — «условие нормировки». Зафиксируем еще некоторые следствия аксиом.

Предложение 1.2. *1)* $0 \cdot \mathbf{a} = \lambda \cdot \mathbf{0} = \mathbf{0} \qquad (\forall\ \mathbf{a} \in V,\ \forall\ \lambda \in \mathbb{F})$*;*
2) $-(\lambda\mathbf{a}) = (-\lambda)\mathbf{a} = \lambda(-\mathbf{a}) \qquad (\forall\ \mathbf{a} \in V,\ \forall\ \lambda \in \mathbb{F})$*;*
3) $\lambda(\mathbf{a}-\mathbf{b}) = \lambda\mathbf{a} - \lambda\mathbf{b} \qquad (\forall\ \mathbf{a}, \mathbf{b} \in V,\ \forall\ \lambda \in \mathbb{F})$*;*
4) $(\lambda-\mu)\mathbf{a} = \lambda\mathbf{a} - \mu\mathbf{a} \qquad (\forall\ \mathbf{a} \in V,\ \forall\ \lambda, \mu \in \mathbb{F})$ *.*

$\rhd$ 1) $0 \cdot \mathbf{a} = (0+0) \cdot \mathbf{a} = 0 \cdot \mathbf{a} + 0 \cdot \mathbf{a}$. С другой стороны, $0 \cdot \mathbf{a} = \mathbf{0} + 0 \cdot \mathbf{a}$. По закону сокращения (см. предложение 1.1, 3) $0 \cdot \mathbf{a} = \mathbf{0}$. Аналогично $\lambda \cdot \mathbf{0} = \lambda(\mathbf{0}+\mathbf{0}) = \lambda \cdot \mathbf{0} + \lambda \cdot \mathbf{0}$, откуда $\lambda \cdot \mathbf{0} = \mathbf{0}$.
2). $(-\lambda)\mathbf{a} + \lambda\mathbf{a} = (-\lambda+\lambda)\mathbf{a} = 0 \cdot \mathbf{a} = \mathbf{0}$, поэтому $(-\lambda)\mathbf{a}$ является противоположным для $\lambda\mathbf{a}$, то есть равен $-(\lambda\mathbf{a})$. Аналогично $\lambda(-\mathbf{a}) + \lambda\mathbf{a} = \lambda(-\mathbf{a}+\mathbf{a}) = \lambda \cdot \mathbf{0} = \mathbf{0}$, поэтому $\lambda(-\mathbf{a}) = -(\lambda\mathbf{a})$.
3). Вытекает из утверждения 2) и определения вычитания векторов: $\lambda(\mathbf{a}-\mathbf{b}) = \lambda(\mathbf{a}+(-\mathbf{b})) = \lambda\mathbf{a} + \lambda(-\mathbf{b}) = \lambda\mathbf{a} - \lambda\mathbf{b}$.
4). Аналогично 3). $\square$

Подпространства

Определение. Непустое подмножество U векторного пространства V называется *подпространством*, если $\forall\ \mathbf{a}, \mathbf{b} \in U$, $\forall\ \lambda \in \mathbb{F}$ выполнено:
П1. $\mathbf{a}+\mathbf{b} \in U$;
П2. $\lambda\mathbf{a} \in U$.

Тот факт, что U является подпространством в векторном пространстве V, будем обозначать $U \leqslant V$.

Свойства П1 и П2 означают, что подпространство является подмножеством, замкнутым относительно операций сложения и умножения на число, тем самым, подпространство само является векторным пространством относительно операций в объемлющем векторном пространстве. Любое векторное пространство V содержит *тривиальные* подпространства V и $O = \{\mathbf{0}\}$ (*нулевое* подпространство).

Предложение 1.3. *Пусть* $U \leqslant V$*, тогда* $\mathbf{0} \in U$*.*

$\rhd$ Пусть $\mathbf{a} \in V$, тогда, согласно П2, $\mathbf{0} = 0 \cdot \mathbf{a} \in V$. $\square$

Предложение 1.4. *Пусть* $U \leqslant V$ *и* $\mathbf{a} \in U$ *тогда* $-\mathbf{a} \in U$*.*

$\rhd$ Достаточно в П2 подставить $\lambda = -1$. $\square$

Видим, что подпространтво — это подгруппа абелевой группы $(V, +)$, замкнутая относительно умножения на константы.

Предложение 1.5. *Пересечение подпространств является подпространством.*

$\rhd$ Пусть $U_i \leqslant V$ для $i \in I$, где I — некоторое множество индексов; $U = \bigcap\limits_{i \in I} U_i$. Проверим П1 для множества U (П2 проверяется аналогично).

Пусть $\mathbf{a}, \mathbf{b} \in U$, тогда $\mathbf{a}, \mathbf{b} \in U_i$ $(\forall\ i \in I)$. Так как U_i — подпространство, то $\mathbf{a}+\mathbf{b} \in U_i$ $(\forall\ i \in I)$, тем самым $\mathbf{a}+\mathbf{b} \in U$, что и требовалось. $\square$

Упражнение. Пусть $U_1 \leqslant V$, $U_2 \leqslant V$. Докажите, что объединение $U_1 \cup U_2$ является подпространством $\Leftrightarrow U_1 \subset U_2$ или $U_2 \subset U_1$.

Линейные комбинации

Пусть $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k \in V$, $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{F}$.

> **Определение.** Сумма $\sum\limits_{i=1}^{k} \lambda_i \mathbf{a}_i$ называется *линейной комбинацией* векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ с *коэффициентами* $\lambda_1, \lambda_2, \ldots, \lambda_k$.

Линейная комбинация $\sum\limits_{i=1}^{k} \lambda_i \mathbf{a_i}$ называется *тривиальной*, если все ее коэффициенты равны 0, то есть $\lambda_1 = \ldots = \lambda_k = 0$. В противном случае линейная комбинация называется *нетривиальной*. Ясно, что тривиальная линейная комбинация равна $\mathbf{0}$.

Количество слагаемых k иногда называют *длиной* линейной комбинации. Можно позволить себе работать и с пустой линейной комбинацией ($k = 0$), полагая ее значение равным $\mathbf{0}$.

Линейную комбинацию условимся записывать также в следующем компактном виде: $\sum\limits_{i=1}^{k} \lambda_i \mathbf{a}_i = \mathrm{a}\lambda$, где $\mathrm{a} = (\mathbf{a}_1\, \mathbf{a}_2\, \ldots\, \mathbf{a}_k)$ — строка векторов, а $\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{pmatrix}$ — столбец коэффициентов. Это согласуется с правилом умножения матриц (умножаем строку на столбец).

В том случае, когда $\mathbf{b} \in V$ равен некоторой линейной комбинации векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ говорят, что $\mathbf{b}$ *раскладывается* по векторам $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ или *линейно выражается* через векторы $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$.

Предложение 1.6. *Пусть $U \leqslant V$. Тогда $\forall\ \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k \in U$ и $\forall\ \lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{F}$: $\sum\limits_{i=1}^{k} \lambda_i \mathbf{a}_i \in U$.*

$\triangleright$ Достаточно многократно применить П1 и П2 $\square$

Таким образом, условие замкнутости относительно взятия любой линейной комбинации эквивалентно определению подпространства (в П1 и П2 мы видим частные случаи линейных комбинаций: $\mathbf{a} + \mathbf{b}$ и $\lambda\mathbf{a}$).

Линейная оболочка

Важное понятие линейной оболочки, которое определим ниже, позволяет, в частности, конструировать подпространства.

> **Определение.** *Линейной оболочкой* системы векторов $\mathbf{a}_1, \ldots, \mathbf{a}_k$ называется множество всех векторов, которые линейно выражаются через $\mathbf{a}_1, \ldots, \mathbf{a}_k$.

Линейная оболочка векторов $\mathbf{a}_1, \ldots, \mathbf{a}_k$ обозначается $\langle \mathbf{a}_1, \ldots, \mathbf{a}_k \rangle$. Формальная запись опеделения:
$\langle \mathbf{a}_1, \ldots, \mathbf{a}_k \rangle := \left\{ \sum\limits_{i=1}^{k} \lambda_i \mathbf{a}_i \mid \lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{F} \right\}$.

Несложно распространить определение на произвольные (в том числе бесконечные) системы векторов.

> **Определение.** *Линейной оболочкой* системы векторов $\mathcal{A}$ называется множество всех векторов, каждый из которых линейно выражается через несколько векторов из $\mathcal{A}$.

Обозначение линейной оболочки: $\langle \mathcal{A} \rangle$. Очевидно, для любого подмножества $\mathcal{A} \subset V$ выполнено $\mathcal{A} \subset \langle \mathcal{A} \rangle$. Условимся считать, что $\langle \varnothing \rangle = O$.

Отметим, что везде рассматриваются линейные комбинации из конечного количества слагаемых, хотя длина линейной комбинации не ограничивается. Формально,
$\langle \mathcal{A} \rangle = \left\{ \sum\limits_{i=1}^{k} \lambda_i \mathbf{a}_i \mid k \in \mathbb{Z}^+,\ \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k \in \mathcal{A},\ \lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{F} \right\}$.

Предложение 1.7. *$\langle\mathcal{A}\rangle$ является подпространством для любой системы векторов $\mathcal{A}$.*

$\triangleright$ Проверим для $\langle\mathcal{A}\rangle$ свойство П1 (П2 проверяется аналогично). Пусть $\mathbf{a},\mathbf{b}\in\langle\mathcal{A}\rangle$. Это означает, что $\mathbf{a}=\sum\limits_{i=1}^{k}\alpha_i\mathbf{a}_i$, $\mathbf{b}=\sum\limits_{j=1}^{m}\beta_j\mathbf{b}_j$, где $\mathbf{a}_i,\mathbf{b}_j\in\mathcal{A}$, $\alpha_i,\beta_j\in\mathbb{F}$. Тогда $\mathbf{a}+\mathbf{b}=\sum\limits_{i=1}^{k}\alpha_i\mathbf{a}_i+\sum\limits_{j=1}^{m}\beta_j\mathbf{b}_j$. В правой части линейная комбинация длины $k+m$ векторов из $\mathcal{A}$, значит $\mathbf{a}+\mathbf{b}\in\langle\mathcal{A}\rangle$. $\square$

Пусть подмножество $\mathcal{A}\subset V$ и подпространство $U\leqslant V$ таковы, что $\mathcal{A}\subset U$. Предложение 1.6 показывает, что тогда и $\langle\mathcal{A}\rangle\subset U$. Таким образом, предложение 1.7 по сути означает, что $\langle\mathcal{A}\rangle$ — это минимальное (по включению) подпространство, содержащее $\mathcal{A}$.

Упражнение. Подмножество $\mathcal{A}\subset V$ является подпространством $\Leftrightarrow$ $\mathcal{A}=\langle\mathcal{A}\rangle$.

Примеры

I. Пусть O — фиксированная точка геометрического пространства (начало отсчета). Множество всех радиус-векторов (с обычными операциями сложения векторов и умножения на число) — пример векторного пространства над $\mathbb{R}$. Радиус-вектор можно отождествить с точкой пространства — концом этого радиус-вектора. Тогда нетривиальные подпространства — это прямые и плоскости, проходящие через O.
Если векторы $\mathbf{a}_1,\dots,\mathbf{a}_k$ коллинеарны и среди них есть хотя бы один ненулевой, то $\langle\mathbf{a}_1,\dots,\mathbf{a}_k\rangle$ — это прямая.
Если же векторы $\mathbf{a}_1,\dots,\mathbf{a}_k$ компланарны, но не коллинеарны, то $\langle\mathbf{a}_1,\dots,\mathbf{a}_k\rangle$ — это плоскость.

II.1. Множество $\mathbf{M}_{m\times n}(\mathbb{F})$ матриц $m\times n$ (заполненных константами) — векторное пространство относительно операций сложения матриц и умножения на число.
$\mathbb{F}^n$ отождествляем с множеством столбцов $\mathbf{M}_{n\times 1}(\mathbb{F})$.
Пусть $\mathbf{M}^{+}_{n\times n}(\mathbb{F})=\{A\in\mathbf{M}_{n\times n}\,|\,A^T=A\}$ — множество симметричных матриц $n\times n$,
$\mathbf{M}^{-}_{n\times n}(\mathbb{F})=\{A\in\mathbf{M}_{n\times n}\,|\,A^T=-A\}$ — множество кососимметричных матриц $n\times n$.
Тогда $\mathbf{M}^{+}_{n\times n}(\mathbb{F})\leqslant\mathbf{M}_{n\times n}(\mathbb{F})$, $\mathbf{M}^{-}_{n\times n}(\mathbb{F})\leqslant\mathbf{M}_{n\times n}(\mathbb{F})$.

II.2. Для матрицы $A\in\mathbf{M}_{m\times n}$ рассмотрим однородную систему линейных уравнений (СЛУ) $AX=O$ (где $X\in\mathbf{M}_{n\times 1}$ — столбец неизвестных) и множество ее решений $U=\mathrm{Sol}(AX=O)$. Тогда $U\leqslant\mathbf{M}_{n\times 1}$.
Если $A_iX=O$, $i=1,2,\dots,k$ — несколько однородных СЛУ относительно $X\in\mathbf{M}_{n\times 1}$, то СЛУ, представляющая собой объединение этих систем, имеет в качестве множества решений пересечение $\bigcap\limits_{i=1}^{k}\mathrm{Sol}(A_iX=O)$.

III.1. Пусть S — некоторое множество. Множество $\mathbf{F}(S)$ всех функций $f:S\to\mathbb{R}$ — векторное пространство над $\mathbb{R}$ относительно операций сложения функций и умножения функции на число.
Пусть $\mathbf{F}^{+}(\mathbb{R})=\{f\in\mathbf{F}(\mathbb{R})\,|\,\forall\,x\in\mathbb{F}\;f(-x)=f(x)\}$ — множество всех четных функций, аналогично
$\mathbf{F}^{-}(\mathbb{R})=\{f\in\mathbf{F}(\mathbb{R})\,|\,\forall\,x\in\mathbb{R}\;f(-x)=-f(x)\}$ — множество всех нечетных функций.
Тогда $\mathbf{F}^{+}(\mathbb{R})\leqslant\mathbf{F}(\mathbb{R})$, $\mathbf{F}^{-}(\mathbb{R})\leqslant\mathbf{F}(\mathbb{R})$.

III.2. Если S — интервал числовой прямой, имеется цепочка вложенных подпространств $\mathbf{F}(S)\geqslant\mathbf{C}(S)\geqslant\mathbf{C}^1(S)\geqslant\ldots\geqslant\mathbf{C}^k(S)\geqslant\ldots\geqslant\mathbf{C}^\infty(S)\geqslant\mathbf{P}(S)$,
где $\mathbf{C}(S)$ — множество непрерывных функций $S\to\mathbb{R}$, $\mathbf{C}^k(S)$ — множество функций, имеющих непрерывную k-ю производную, $\mathbf{P}(S)$ — множество полиномиальных функций (многочленов).
Множество (формальных) многочленов $\mathbf{P}$ может быть задано линейной оболочкой: $\mathbf{P}=\langle 1,x,x^2,x^3,\ldots\rangle$.
$\mathbf{P}\geqslant\mathbf{P}_n=\langle 1,x,x^2,\dots,x^n\rangle$, где $\mathbf{P}_n$ — множество многочленов f степени $\deg f\leqslant n$.

III.3. В пространстве $\mathbf{F}(\mathbb{N}) = \{(f_1, f_2, \ldots) \mid f_i \in \mathbb{R}\}$ всех последовательностей есть подпространство ограниченных последовательностей, а в нем — подпространство сходящихся последовательностей.

«Бесконечная система линейных уравнений» $f_{i+1} = f_i + f_{i-1}$, $i = 2, 3, \ldots$, задает подпространство *фибоначчиевых* последовательностей.

III.4. Рассмотрим линейное однородное дифференциальное уравнение $x^{(n)} + a_{n-1}(t)x^{(n-1)} + \ldots + a_0(t)x = 0$ относительно неизвестной функции $x(t)$ (где $a_i(t)$ — данные непрерывные функции $\mathbb{R} \to \mathbb{R}$). Множество решений этого уравнения — подпространство в пространстве всех функций $\mathbb{R} \to \mathbb{R}$.

IV. Если $\mathbb{F}$ — подполе некоторого поля $\mathbb{K}$, то операции в $\mathbb{K}$ индуцируют на $\mathbb{K}$ структуру векторного пространства над $\mathbb{F}$.

§ 2. Линейная зависимость. Размерность и ранг. Базис

Линейная зависимость

Продолжаем работать в векторном пространстве V над полем $\mathbb{F}$.

Определение. Система векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ называется *линейно зависимой*, если некоторая их нетривиальная линейная комбинация равна $\mathbf{0}$, и *линейно независимой* в противном случае.

Полагают, что пустая система векторов линейно независима (формально это согласуется с определением).

Предложение 2.1. *Система векторов* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k \in V$ *(*$k \geqslant 2$*) линейно зависима* $\Leftrightarrow$ *среди векторов* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ **найдется** *вектор, который линейно выражается через остальные* $k-1$ *векторов этой системы.*

$\triangleright$ $\boxed{\Rightarrow}$ Пусть $\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \ldots + \lambda_k\mathbf{a}_k = \mathbf{0}$, и не все коэффициенты равны 0, скажем $\lambda_k \neq 0$. Тогда поделим равенство на $-\lambda_k$ и перенесем $\mathbf{a}_k$ в другую часть; получим $\mathbf{a}_k = \sum\limits_{i=1}^{k-1} \mu_i\mathbf{a}_i$, где $\mu_i = -\dfrac{\lambda_i}{\lambda_k}$, $i = 1, 2, \ldots, k-1$.

$\boxed{\Leftarrow}$ Пусть, скажем, вектор $\mathbf{a}_k$ раскладывается по векторам $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{k-1}$: $\mathbf{a}_k = \sum\limits_{i=1}^{k-1} \mu_i\mathbf{a}_i$. Тогда $\mu_1\mathbf{a}_1 + \mu_2\mathbf{a}_2 + \ldots + \mu_{k-1}\mathbf{a}_{k-1} - \mathbf{a}_k$ — нетривиальная линейная комбинация, равная $\mathbf{0}$. $\square$

Предложение 2.2. *1) Если в конечной системе векторов имеется некоторая линейно зависимая подсистема, то и вся система линейно зависима.*

2) Любая подсистема линейно независимой системы линейно независима.

$\triangleright$ 1) Пусть, скажем, для системы векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ ее подсистема $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ $(m \leqslant k)$ линейно зависима. Тогда существует нетривиальная линейная комбинация $\sum\limits_{i=1}^{m} \mu_i\mathbf{a}_i$, равная $\mathbf{0}$. Значит, $\mu_1\mathbf{a}_1 + \mu_2\mathbf{a}_2 + \ldots + \mu_m\mathbf{a}_m + 0 \cdot \mathbf{a}_{m+1} + \ldots + 0 \cdot \mathbf{a}_k$ — нетривиальная линейная комбинация, равная $\mathbf{0}$.

2) Это переформулировка утверждения 1). $\square$

Следствие. *Любая система векторов, содержащая* $\mathbf{0}$*, является линейно зависимой.*

Предложение 2.3. *Пусть векторы* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ *и* $\mathbf{b}$ *таковы, что* $\mathbf{b} \in \langle \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k \rangle$*. Тогда коэффициенты* $\lambda_1, \lambda_2, \ldots, \lambda_k$ *в разложении* $\mathbf{b} = \sum\limits_{i=1}^{k} \lambda_i\mathbf{a}_i$ *определяются однозначно* $\Leftrightarrow$ *система* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ *линейно независима.*

$\triangleright$ $\boxed{\Rightarrow}$ Предположим, что напротив, система $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ линейно зависима и существует нетривиальная линейная комбинация $\sum\limits_{i=1}^{k} \alpha_i \mathbf{a_i}$, равная $\mathbf{0}$. Тогда можно прибавить эту линейную комбинацию к имеющейся линейной комбинации, равной $\mathbf{b}$, и получить новое линейное выражение: $\mathbf{b} = \sum\limits_{i=1}^{k} (\lambda_i + \alpha_i)\mathbf{a_i}$ (оно действительно хотя бы в одном коэффициенте отличается от разложения $\mathbf{b} = \sum\limits_{i=1}^{k} \lambda_i \mathbf{a_i}$).

$\boxed{\Leftarrow}$ Пусть система $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ линейно независима, и предположим, что наряду с разложением $\mathbf{b} = \sum\limits_{i=1}^{k} \lambda_i \mathbf{a}_i$ имеется разложение $\mathbf{b} = \sum\limits_{i=1}^{k} \mu_i \mathbf{a}_i$. Вычитая из первого равенства второе, получаем $\sum\limits_{i=1}^{k} (\lambda_i - \mu_i)\mathbf{a}_i = \mathbf{0}$. Так как $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ — линейно независимая система, то левая часть полученного равенства — тривиальная линейная комбинация, откуда $\lambda_i = \mu_i$, $i = 1, 2, \ldots, k$. $\square$

Пусть $\mathrm{a} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)$ — строка векторов, где $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ — линейно независимая система. Тогда предыдущее предложение можно интерпретировать как закон сокращения: $\mathrm{a}\lambda = \mathrm{a}\mu \Rightarrow \lambda = \mu$. Этот закон выполнен как для столбцов $\lambda, \mu \in \mathbf{M}_{k\times 1}$, так и для матриц $\lambda, \mu \in \mathbf{M}_{k\times m}$. Обратим внимание на то, что если отказаться от условия линейной независимости системы $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$, то следствие $\mathrm{a}\lambda = \mathrm{a}\mu \Rightarrow \lambda = \mu$, вообще говоря, неверно.

Ранг и размерность

Определение. Целое неотрицательное число r называется *рангом* непустой системы (возможно, бесконечной) $\mathcal{A}$ векторов из V если в системе $\mathcal{A}$ найдется линейно независимая подсистема из r векторов, а любая подсистема из $r+1$ векторов является линейно зависимой.

Определение. Будем говорить, что $\mathcal{A}$ имеет бесконечный ранг, если для любого $r \in \mathbb{N}$ в $\mathcal{A}$ найдется линейно независимая подсистема из r векторов.

Обозначение для ранга: $\operatorname{rg}\mathcal{A}$. В частности, $\operatorname{rg}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)$ — ранг конечной системы векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$.

В том случае, когда $\mathcal{A}$ является подпространством в V, более употребительное название для ранга — *размерность*. Обозначение для размерности — $\dim\mathcal{A}$. Итак, если $U \leqslant V$, то $\boxed{\dim U = \operatorname{rg} U}$. Пространство размерности k называют *k-мерным*. Если $\dim V < \infty$, то пространство V называют *конечномерным*, иначе — *бесконечномерным*.

Очевидно, система из одного нулевого вектора имеет ранг 0, а ранг любой конечной системы из k векторов не превосходит k.

Предложение 2.4. *Конечная система векторов* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ *линейно независима* $\Leftrightarrow$ $\operatorname{rg}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k) = k$.

$\triangleright$ Сразу следует из определения. $\square$

Предложение 2.5. *Пусть* $\mathcal{A}_1, \mathcal{A}_2$ *— две системы векторов из* V*, причем* $\operatorname{rg}\mathcal{A}_1 = r_1$, $\operatorname{rg}\mathcal{A}_2 = r_2$. *Тогда* $\operatorname{rg}(\mathcal{A}_1 \bigcup \mathcal{A}_2) \leqslant r_1 + r_2$.

$\triangleright$ Пусть это не так, и в объединении наборов $\mathcal{A}_1$ и $\mathcal{A}_2$ нашлась линейно независимая подсистема из $r_1 + r_2 + 1$ векторов. Но (по определению ранга и предложению 2.2) среди этих векторов не более r_1 векторов из $\mathcal{A}_1$ и не более r_2 векторов из $\mathcal{A}_2$. Противоречие. $\square$

Следствие. *Пусть* $\mathcal{A}_i$, $i = 1, 2, \ldots, k$ *— системы векторов из* V*, причем* $\operatorname{rg}\mathcal{A}_i = r_i$. *Тогда* $\operatorname{rg}(\mathcal{A}_1 \bigcup \ldots \bigcup \mathcal{A}_k) \leqslant \sum\limits_{i=1}^{k} r_i$.

Предложение 2.6. *Пусть* $\mathcal{A}, \mathcal{B}$ *— две системы векторов из* V*, причем* $\mathcal{A} \subset \mathcal{B}$ *и* $\operatorname{rg}\mathcal{B} = r$. *Тогда* $\operatorname{rg}\mathcal{A} \leqslant r$.

$\triangleright$ Сразу следует из определения. $\square$

Предыдущее предложение почти очевидно: если к системе векторов $\mathcal{A}$ добавить некоторые векторы (расширить $\mathcal{A}$ до системы $\mathcal{B}$), то ранг не уменьшится. Оказывается, ранг не изменится, если к $\mathcal{A}$ добавлять линейные комбинации векторов из $\mathcal{A}$ (и наоборот, ранг не изменится, если из системы векторов удалить вектор, который раскладывается по оставшимся векторам). Обобщение этого факта составляет содержание следующей основной теоремы о рангах. Доказательству теоремы предпошлем две леммы (которые являются частными случаями теоремы).

Лемма 1. *Пусть $\mathcal{A}$ — система векторов из V, $\operatorname{rg}\mathcal{A} = r < \infty$ и $\mathbf{a}_1, \ldots, \mathbf{a}_r$ — линейно независимая подсистема векторов из $\mathcal{A}$. Тогда любой вектор из $\mathcal{A}$ раскладывается по $\mathbf{a}_1, \ldots, \mathbf{a}_r$.*

$\triangleright$ Пусть $\mathbf{a}$ — произвольный вектор из $\mathcal{A}$. Из определения ранга следует, что система из $r+1$ векторов $\mathbf{a}$, $\mathbf{a}_1, \ldots, \mathbf{a}_r$ линейно зависима. Тогда найдутся числа μ, $\lambda_1, \ldots,$ $\lambda_r \in \mathbb{R}$, не все равные нулю и такие, что $\mu\mathbf{a} + \sum\limits_{i=1}^{r} \lambda_i \mathbf{a}_i = \mathbf{0}$. При этом $\mu \neq 0$, иначе система $\mathbf{a}_1, \ldots, \mathbf{a}_r$ была бы линейно зависимой. Отсюда $\mathbf{a} = -\sum\limits_{i=1}^{r} \dfrac{\lambda_i}{\mu}\mathbf{a}_i$. $\square$

Лемма 2. *Пусть $\mathbf{a}_1, \ldots, \mathbf{a}_k$, $\mathbf{b}$ — такие векторы из V, что $\mathbf{b} \in \langle \mathbf{a}_1, \ldots, \mathbf{a}_k \rangle$. Тогда $\operatorname{rg}(\mathbf{a}_1, \ldots, \mathbf{a}_k) = \operatorname{rg}(\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{b})$.*

$\triangleright$ Положим $r = \operatorname{rg}(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ $(r \leqslant k)$. Предположим, что утверждение неверно, и в системе $\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{b}$ нашлась линейно независимая подсистема из $r+1$ векторов. Тогда один из этих $r+1$ векторов — это $\mathbf{b}$ (так как в системе $\mathbf{a}_1, \ldots, \mathbf{a}_k$ нет линейно независимой подсистемы из $r+1$ векторов). Итак, пусть для определенности $\mathbf{b}, \mathbf{a}_1, \ldots, \mathbf{a}_r$ — линейно независимая система. Из предложения 2.2 следует, что система $\mathbf{a}_1, \ldots, \mathbf{a}_r$ линейно независима, значит, по лемме 1 каждый из векторов $\mathbf{a}_1, \ldots, \mathbf{a}_k$ лежит в $\langle \mathbf{a}_1, \ldots, \mathbf{a}_r \rangle$. По условию $\mathbf{b}$ равен линейной комбинации векторов $\mathbf{a}_1, \ldots, \mathbf{a}_k$: $\mathbf{b} = \sum\limits_{i=1}^{k} \lambda_i \mathbf{a}_i$. Подставив в это выражение вместо векторов $\mathbf{a}_1, \ldots, \mathbf{a}_k$ их разложения по векторам $\mathbf{a}_1, \ldots, \mathbf{a}_r$, получим, что $\mathbf{b}$ раскладывается по векторам $\mathbf{a}_1, \ldots, \mathbf{a}_r$. Но это противоречит линейной независимости системы $\mathbf{b}, \mathbf{a}_1, \ldots, \mathbf{a}_r$ (см. предложение 2.1). $\square$

Теорема 2.1 (основная теорема о рангах)**.** *Пусть $\mathcal{A}$ и $\mathcal{B}$ — две такие системы векторов из V, что $\mathcal{A} \subset \mathcal{B}$. Пусть $\operatorname{rg}\mathcal{A} = r < \infty$. Тогда $\operatorname{rg}\mathcal{B} = r \Leftrightarrow$ любой вектор из $\mathcal{B}$ принадлежит $\langle \mathcal{A} \rangle$.*

$\triangleright$ $\boxed{\Rightarrow}$ В системе $\mathcal{A}$ зафиксируем некоторую линейно независимую подсистему $\mathbf{a}_1, \ldots, \mathbf{a}_r$ из r векторов. Так как $\mathbf{a}_1, \ldots, \mathbf{a}_r \in \mathcal{B}$ и $\operatorname{rg}\mathcal{B} = r$, то по лемме 1 (примененной к системе $\mathcal{B}$) любой вектор из $\mathcal{B}$ лежит в $\langle \mathbf{a}_1, \ldots, \mathbf{a}_r \rangle$ и, следовательно, в $\langle \mathcal{A} \rangle$.

$\boxed{\Leftarrow}$ Предположим, что утверждение неверно, и в системе $\mathcal{B}$ нашлась линейно независимая подсистема из $r+1$ векторов $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{r+1}$. Каждый из них линейно выражается через несколько (конечное число) векторов из $\mathcal{A}$, поэтому можно выбрать конечную систему векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_l$ из $\mathcal{A}$, через которые линейно выражается каждый из векторов $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{r+1}$. Имеем $\operatorname{rg}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_l, \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{r+1}) \geqslant \operatorname{rg}(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{r+1}) = r+1$. С другой стороны, применяя многократно лемму 2, имеем: $\operatorname{rg}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_l, \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{r+1}) = \operatorname{rg}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_l, \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_r) = \operatorname{rg}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_l, \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{r-1}) = \ldots = \operatorname{rg}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_l) \leqslant \operatorname{rg}\mathcal{A} = r$. Противоречие. $\square$

Следствие 1. *Для любой системы векторов $\mathcal{A}$ из V выполнено $\operatorname{rg}\mathcal{A} = \dim\langle \mathcal{A} \rangle$.*

Следствие 2. *Пусть даны подпространства $U \leqslant W \leqslant V$ такие, что $\dim U = \dim W < \infty$. Тогда $U = W$.*

> **Определение.** Подсистему $\mathbf{a}_1, \ldots, \mathbf{a}_k$ системы векторов $\mathcal{A}$ будем называть *базисной*, если
> 1. $\mathbf{a}_1, \ldots, \mathbf{a}_k$ — линейно независима,
> 2. любой вектор из $\mathcal{A}$ принадлежит линейной оболочке $\langle \mathbf{a}_1, \ldots, \mathbf{a}_k \rangle$.

Теорема 2.2. *Пусть* $\operatorname{rg}\mathcal{A} = r < \infty$. *Тогда подсистема* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ *системы* $\mathcal{A}$ *является базисной для* $\mathcal{A} \Leftrightarrow \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ *линейно независима и* $k = r$.

$\triangleright$ Сразу следует из основной теоремы 2.1. $\square$

Предложение 2.7. *Пусть* $\operatorname{rg}\mathcal{A} = r < \infty$. *Пусть* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ *— линейно независимая подсистема системы* $\mathcal{A}$. *Тогда систему* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ *можно дополнить* $r - k$ *векторами до базисной подсистемы.*

$\triangleright$ Если $k < r$, то $k = \operatorname{rg}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k) < \operatorname{rg}\mathcal{A}$, и по основной теореме 2.1 в $\mathcal{A}$ найдется вектор $\mathbf{a}_{k+1}$, который не выражается линейно через $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$. Тогда $\operatorname{rg}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{k+1}) > k$, следовательно $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{k+1}$ — линейно независимая подсистема. Продолжая процесс добавления векторов, в конце концов придем к базисной подсистеме. $\square$

Упражнение.* Теорию ранга и размерности можно развить, начиная с другого (эквивалентного) определения ранга: рангом подмножества $\mathcal{A} \subset V$ можем называть минимальное r такое, что существует подмножество $\mathcal{B}$ такое, что $|\mathcal{B}| = r$ и $\mathcal{A} \subset \langle \mathcal{B} \rangle$.

Базис

> **Определение.** Упорядоченный конечный набор векторов $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ называется *базисом* векторного пространства V, если
> B1. $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ — линейно независимая система;
> B2. $V = \langle \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n \rangle$.

Видим, что определение базиса согласуется с определением базисной подсистемы.

Базисы часто будем для краткости обозначать одной буквой, имея в виду упорядоченный набор векторов или строку из векторов, например: $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$.

Теорема 2.3 (Описание базисов)**.** *Пусть* $\dim V = n < \infty$. *Тогда упорялрченный набор векторов* $\mathbf{e}_1, \ldots, \mathbf{e}_k$ *является базисом пространства* $V \Leftrightarrow$ *система* $\mathbf{e}_1, \ldots, \mathbf{e}_k$ *линейно независима и* $k = n$.

$\triangleright$ Следует из теоремы 2.2. $\square$

Существование базисов в векторных пространствах конечной размерности следует из предыдущей теоремы. Если $\dim V = n < \infty$, то каждый базис пространства V содержит ровно n векторов. Если $\dim V = \infty$, то конечного базиса в V не существует (в этом курсе мы не будем заниматься понятием бесконечного базиса).

Предложение 2.8. *Пусть* $\dim V = n < \infty$, *и* $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k$ *— линейно независимая система векторов. Тогда эту систему можно дополнить до базиса* $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k, \mathbf{e}_{k+1}, \ldots, \mathbf{e}_n$ *пространства* V.

$\triangleright$ Следует из предложения 2.7. $\square$

Координаты

> **Определение.** Пусть в векторном пространстве V зафиксирован базис $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$. Коэффициенты $x_1, x_2, \ldots, x_n$ в разложении $\mathbf{a} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n$ вектора $\mathbf{a} \in V$ по этому базису называются *координатами* вектора $\mathbf{a}$ в базисе e.

Из предложения 2.3 следует, что упорядоченный набор координат данного вектора $\mathbf{a}$ в данном базисе однозначно определен. Упорядоченный набор координат $(x_1, x_2, \ldots, x_n)$ удобно записывать в виде столбца: $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. Этот столбец называется *координатным столбцом* вектора $\mathbf{a}$ в базисе e. Для любого упорядоченного набора координат имеется вектор из V именно с таким набором координат. Таким образом, если в векторном пространстве V зафиксирован базис e, то имеется взаимно-однозначное соответствие между множеством V и множеством $\mathbf{M}_{n\times 1}$ (вектору сопоставлятся координатный столбец). Запись $\mathbf{a} = \mathrm{e}X$ будет означать, что вектор $\mathbf{a}$ имеет координатный столбец X в базисе e.

Также разложения вида $\sum\limits_{i=1}^{n} x_i\mathbf{e}_i$ можно записывать в виде так называемого *тензорного суммирования* — следующей договоренности, предложенной Эйнштейном: наличие одинаковых верхнего и нижнего индекса означает суммирования по этому индексу (от 1 до $n = \dim V$). Так, если в координатах x_i вместо нижних индексов использовать верхние, то запись $\sum\limits_{i=1}^{n} x^i\mathbf{e}_i$ можно сократить до $x^i\mathbf{e}_i$.

Следующее предложение показывает, что записи вида $\mathbf{a} = \mathrm{e}X$ согласуются с дистрибутивностью матричного умножения.

Предложение 2.9 (линейность сопоставления координат)**.** *Пусть в V зафиксирован базис* $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$*. Тогда при сложении векторов соответствующие координаты складываются, а при умножении вектора на число $\lambda \in \mathbb{R}$ соответствующие координаты умножаются на λ.*

$\triangleright$ По условию $\mathbf{a} = \mathrm{e}X = \sum\limits_{i=1}^{n} x_i\mathbf{e}_i$, $\mathbf{b} = \mathrm{e}Y = \sum\limits_{i=1}^{n} y_i\mathbf{e}_i$.

Сложив равенства, имеем $\mathbf{a} + \mathbf{b} = \sum\limits_{i=1}^{n} (x_i + y_i)\mathbf{e}_i = \mathrm{e}(X + Y)$.

Умножив первое равенство на λ, имеем $\lambda\mathbf{a} = \sum\limits_{i=1}^{n} (\lambda x_i)\mathbf{e}_i = \mathrm{e}(\lambda X)$. $\square$

Итак, при фиксации базиса e равенство $\mathbf{a} = \mathrm{e}X$ возникает важное линейное взаимно-однозначное соответствие (ниже для таких соответствий вводится термин «изоморфизм») между V и $\mathbf{M}_{n\times 1}$. Эта «координатизация» дает возможность вопросы о линейных операциях в V перевести на матричный язык (вместо абстрактных элементов V можно заниматься столбцами). В частности, отметим такое следствие предложения 2.9 (оно согласуется с общими свойствами изоморфизма — см. главу 2).

Следствие. *Пусть в V зафиксирован базис* e*, и в этом базисе векторы $\mathbf{a}_1, \ldots, \mathbf{a}_k$ имеют координатные столбцы $X_1, \ldots, X_k$ соответственно: $\mathbf{a}_i = \mathrm{e}X_i$, $i = 1, \ldots, k$. Тогда система векторов $\mathbf{a}_1, \ldots, \mathbf{a}_k$ линейно зависима $\Leftrightarrow$ система столбцов $X_1, \ldots, X_k$ линейно зависима.*

$\triangleright$ Согласно предложению 2.9, $\sum\limits_{i=1}^{k} \lambda_i\mathbf{a}_i = \mathbf{0} \Leftrightarrow \sum\limits_{i=1}^{k} \lambda_i X_i = O$. $\square$

Замена базиса и координат

Пусть $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ и $\mathrm{e}' = (\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n)$ — два базиса в векторном пространстве V $(n = \dim V)$.

> **Определение.** Матрица S размера $n\times n$, j-ый столбец которой равен координатному столбцу вектора $\mathbf{e}'_j$ в базисе e $(j = 1, 2, \ldots, n)$, называется *матрицей перехода* от базиса e к базису e′.

Определение означает, что столбцы $s_{\bullet 1}, \ldots, s_{\bullet n}$ матрицы перехода от e к e′ удовлетворяют равенствам $\mathbf{e}'_1 = \mathrm{e}s_{\bullet 1}$, …, $\mathbf{e}'_n = \mathrm{e}s_{\bullet n}$. Эти равенства можно записать в виде $\boxed{\mathrm{e}' = \mathrm{e}S}$. По сути это компактная запись определения матрицы перехода.

Предложение 2.10. *Пусть в пространстве V выбран базис* e*. Матрица $S \in \mathbf{M}_{n\times n}$ является матрицей перехода от* e *к некоторому базису* e′ *тогда и только тогда, когда S невырожденная.*

$\triangleright$ Согласно следствию из предложения 2.9, S является матрицей перехода к некоторому базису e′ тогда и только тогда, когда столбцы S образуют базис в $\mathbf{M}_{n\times 1}$. $\square$

Предложение 2.11. *Пусть в пространстве V выбран базис* e*. Матрица перехода от* e *к* e *равна единичной матрице E_n.*

▷ Следует из определения матрицы перехода. □

Теорема 2.4. *Пусть* $\dim V = n < \infty$ *и* $\mathbf{a} \in V$ *имеет в базисах* e *и* e' *координатные столбцы* X *и* X'. *Тогда*

$$\boxed{X = SX'},$$

где S *— матрица перехода от базиса* e *к базису* e'.

▷ По условию $\mathbf{a} = \mathrm{e}X = \mathrm{e}'X'$. Так как $\mathrm{e}' = \mathrm{e}S$, имеем $\mathrm{e}X = (\mathrm{e}S)X' = \mathrm{e}(SX')$. В последнем равенстве используется ассоциативность умножения матриц, она верна и в случае «необычной» матрицы e, элементы которой — векторы. Далее из закона сокращения (здесь используем, что e — линейно независимая система) получаем $X = SX'$. □

Предложение 2.12. *Пусть* e, e', e'' *— три базиса в* V. *Пусть* S *— матрица перехода от* e *к* e', *а* R *— матрица перехода от* e' *к* e''. *Тогда матрица перехода от* e *к* e'' *равна* SR.

▷ $\mathrm{e}'' = \mathrm{e}'R = (\mathrm{e}S)R = \mathrm{e}(SR)$. □

Следствие. *Если матрица перехода от базиса* e *к базису* e' *равна* S, *то матрица перехода от базиса* e' *к базису* e *равна* S^{-1}.

▷ Положив в предложении 2.12 $\mathrm{e}'' = \mathrm{e}$, с учетом предложения 2.11 получим $SR = E$. □

Примеры

I. Понятие базиса на плоскости и в пространстве согласуется с определением из курса геометрии: базисы на плоскости — упорядоченные пары неколлинеарных векторов (т.е. плоскость является двумерным пространством); базисы в пространстве — упорядоченные тройки некомпланарных векторов (т.е. геометрическое пространство является трехмерным).

II.1. *Стандартный базис* в $\mathbf{M}_{m\times n}$ образуют матрицы, элементы которых — все нули, кроме одной единицы. Тогда числа, записанные в ячейках произвольной матрицы $A \in \mathbf{M}_{m\times n}$, — это коэффициенты в разложении A по стандартному базису. Отсюда $\dim \mathbf{M}_{m\times n} = mn$.

Упражнение. Найдите $\dim \mathbf{M}^{+}_{n\times n}(\mathbb{R})$ и $\dim \mathbf{M}^{-}_{n\times n}(\mathbb{R})$. Укажите некоторые базисы этих подпространств.

II.2. Пространство столбцов $V = \mathbf{M}_{n\times 1}(\mathbb{F})$ можно назвать *стандартным* n-мерным пространством над полем $\mathbb{F}$ (в согласии с замечаниями о «координатизации»).
Базисы в V — во взаимно-однозначном соответствии с множеством $GL_n(\mathbb{F})$ невырожденных матриц: каждая невырожденная матрица $A = (a_{\bullet 1}\ \ldots\ a_{\bullet n})$ определяет базис $(a_{\bullet 1}, \ldots, a_{\bullet n})$.
В частности, верхнетреугольная матрица с ненулевыми константами на главной диагонали определяет так называемый *треугольный базис* простанства V.
Та же матрица A совпадает с матрицей перехода от стандартного базиса к базису $(a_{\bullet 1}, \ldots, a_{\bullet n})$.

II.3. Подпространство $U \leqslant \mathbb{R}^n = \mathbf{M}_{n\times 1}(\mathbb{R})$ может быть задано как линейная оболочка нескольких столбцов: $U = \langle a_{\bullet 1}, \ldots, a_{\bullet k}\rangle$. Базис U в этом случае можно найти, выполнив элементарные преобразования столбцов (они не изменяют их линейную оболочку) до ступенчатого вида. Имеем $\dim U = \operatorname{rg} A$, где A — матрица, $(a_{\bullet 1}\ \ldots\ a_{\bullet k})$.

II.4. Подпространство $U \leqslant \mathbb{R}^n$ может быть задано как множество решений некоторой однородной СЛУ (системы линейных уравнений) $AX = O$, т.е. $U = \mathrm{Sol}(AX = O)$. При этом если $\operatorname{rg} A = r$, то $\dim U = n - r$. Определение базиса в $U = \mathrm{Sol}(AX = O)$ совпадает с определением ФСР (фундаментальной системы решений) СЛУ $AX = O$.
В указанном виде может быть задано любое подпространство $U \leqslant \mathbb{R}^n$ (имеется алгоритм получения СЛУ, для которой данная линейная оболочка столбцов являетcя множеством решений; см. также предложение 2.10 главы 5).

III.1. Набор степеней $1, x, x^2, \ldots, x^n$ — базис в пространстве $\mathbf{P}_n$ (формальных) многочленов степени не выше n. Так, $\dim \mathbf{P}_n = n + 1$. (Пространство $\mathbf{P}$ всех многочленов уже не является конечномерным.)

Также набор функций $1, x, x^2, \ldots, x^n$ является базисом в пространстве полиномиальных функций $\mathbb{R} \to \mathbb{R}$ степени не выше n. Действительно, если линейная комбинация $\sum\limits_{i=0}^{n} \lambda_i x^i$ равна нулевой функции, то все коэффициенты равны 0, иначе многочлен степени k будет иметь больше чем k корней.

III.2. Рассмотрим пространство фибоначчиевых последовательностей $V = \{(x^1, x^2, \ldots) \,|\, x^{i+1} = x^i + x^{i-1},\ i = 2, 3, \ldots\}$. Каждая последовательность из V однозначно задается первыми двумя членами x^1, x^2. Рассмотрим две последовательности $e_1 = (1, 0, 1, 1, 2, \ldots)$ и $e_2 = (0, 1, 1, 2, 3, \ldots)$. Они образуют *стандартный базис* в V: каждая последовательность $(x^1, x^2, \ldots) \in V$ совпадает с линейной комбинацией $x^1 e_1 + x^2 e_2$ (поскольку совпадают первые два члена). В частности, $\dim V = 2$.

В V можно обнаружить известные последовательности — геометрические прогрессии. Действительно $(1, \lambda, \lambda^2, \ldots) \in V \Leftrightarrow \lambda^{i+1} = \lambda^i + \lambda^{i-1}$, $i = 2, 3, \ldots$. Последнее множество равенств эквивалентно единственному равенству $\lambda^2 = \lambda + 1$, откуда $\lambda_{1,2} = \frac{1 \pm \sqrt{5}}{2}$. Найденные прогрессии $g_i = (1, \lambda_i, \lambda_i^2, \ldots)$, $i = 1, 2$, образуют базис в V.

Линейное выражение (обычной) последовательности Фибоначчи $f = (1, 1, 2, 3, 5, 8, \ldots)$ через g_i позволит найти явную формулу для чисел Фибоначчи.

Равенство $f = c_1 g_1 + c_2 g_2$ эквивалентно системе равенств для первых двух членов:
$1 = c_1 + c_2$, $1 = c_1 \lambda_1 + c_2 \lambda_2$. Отсюда $c_1 = \frac{\sqrt{5}+1}{2\sqrt{5}} = \frac{\lambda_1}{\sqrt{5}}$, $c_2 = \frac{\sqrt{5}-1}{2\sqrt{5}} = -\frac{\lambda_2}{\sqrt{5}}$. И окончательно, n-е число Фибоначчи равно $c_1 \lambda_1^{n-1} + c_2 \lambda_2^{n-1} = \frac{\lambda_1^n - \lambda_2^n}{\sqrt{5}}$.

IV. Пусть $\mathbb{F}$ — конечное поле из q элементов (например, $\mathbb{F} = \mathbb{Z}_p$ для некоторого простого p), а V — векторное пространство над $\mathbb{F}$, $\dim V = n$.

Фиксация любого базиса определяет биекцию $V \to \mathbf{M}_{n\times 1}(\mathbb{F}) = \mathbb{F}^n$, откуда $|V| = q^n$.

Далее, выбор (упорядоченной) линейно независимой системы $\mathbf{a}_1, \ldots, \mathbf{a}_k$ в V можно осуществить, последовательно выбирая $\mathbf{a}_1 \neq \mathbf{0}$, $\mathbf{a}_2 \notin \langle \mathbf{a}_1 \rangle$, $\mathbf{a}_3 \notin \langle \mathbf{a}_1, \mathbf{a}_2 \rangle$, и т.д. Для выбора $\mathbf{a}_{i+1}$ имеется $(q^n - q^i)$ возможностей (все векторы из V, исключая лежащие в i-мерном подпространстве $\langle \mathbf{a}_1, \ldots, \mathbf{a}_i \rangle$). Отсюда количество таких линейно независимых систем равно $\prod\limits_{i=1}^{k} (q^n - q^i)$.

В частности, количество базисов в V (равное также $|GL_n(\mathbb{F})|$) равно $\prod\limits_{i=1}^{n} (q^n - q^i)$.

Линейно независимая система $\mathbf{a}_1, \ldots, \mathbf{a}_k$ определяет подпространство $\langle \mathbf{a}_1, \ldots, \mathbf{a}_k \rangle$, и каждое k-мерное подпространство определяется таким образом столькими способами, сколько в нем базисов. Отсюда находим количество k-мерных подпространств в V как $\dfrac{\prod\limits_{i=1}^{k} (q^n - q^i)}{\prod\limits_{i=1}^{k} (q^k - q^i)}$.

§ 3. Сумма подпространств. Прямая сумма

Сумма подпространств

Определение. Пусть $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_k$ — подмножества в V. *Суммой* (по Минковскому) подмножеств $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_k$ называется множество $\{\mathbf{a}_1 + \mathbf{a}_2 + \ldots + \mathbf{a}_k \,|\, \mathbf{a}_i \in \mathcal{A}_i\}$, $i = 1, \ldots, k$.

Обозначение для суммы подмножеств: $\mathcal{A}_1 + \mathcal{A}_2 + \ldots + \mathcal{A}_k$ или $\sum\limits_{i=1}^{k} \mathcal{A}_i$. Из определения ясно, что $\mathcal{A}_1 + \mathcal{A}_2 = \mathcal{A}_2 + \mathcal{A}_1$, $(\mathcal{A}_1 + \mathcal{A}_2) + \mathcal{A}_3 = \mathcal{A}_1 + (\mathcal{A}_2 + \mathcal{A}_3)$.

Далее будем заниматься почти всегда только суммами подпространств.

Следующее предложение связывает понятия суммы и линейной оболочки.

Предложение 3.1. *Пусть* $U_i = \langle \mathcal{A}_i \rangle$, $i = 1, 2, \ldots, k$. *Тогда*

$$\sum_{i=1}^{k} U_i = \langle \mathcal{A}_1 \cup \mathcal{A}_2 \cup \ldots \cup \mathcal{A}_k \rangle.$$

$\triangleright$ По определению, $\sum\limits_{i=1}^{k} U_i$ — множество элементов вида $\sum\limits_{i=1}^{k} \mathbf{a}_i$, где $\mathbf{a}_i$ может быть значением произвольной линейной комбинации векторов из $\mathcal{A}_i$, а значит множество сумм $\sum\limits_{i=1}^{k} \mathbf{a}_i$ совпадает с множеством значений линейных комбинации векторов из $\mathcal{A}_1 \cup \mathcal{A}_2 \cup \ldots \cup \mathcal{A}_k$. $\square$

Следствие. *Сумма нескольких (конечного числа) подпространств является подпространством.*

Видим, что сумма $\sum\limits_{i=1}^{k} U_i$ подпространств $U_1, \ldots, U_k$ — это минимальное по включению подпространство, содержащее каждое из подпространств $U_1, \ldots, U_k$.

Предложение 3.2. *Пусть* $U_i \leqslant V$, $i = 1, 2, \ldots, k$. *Тогда*

$$\dim \left(\sum_{i=1}^{k} U_i \right) \leqslant \sum_{i=1}^{k} \dim U_i.$$

$\triangleright$ По предыдущему предложению, $\sum\limits_{i=1}^{k} U_i = \langle U_1 \cup U_2 \cup \ldots \cup U_k \rangle$. Но по основной теореме 2.1 и предложению 2.5 имеем $\dim \langle U_1 \cup U_2 \cup \ldots \cup U_k \rangle = \mathrm{rg}(U_1 \cup U_2 \cup \ldots \cup U_k) \leqslant \sum\limits_{i=1}^{k} \dim U_i$. $\square$

Для операции сложения подмножеств выполнены далеко не все свойства, которыми обладает операция сложения векторов. Например, вообще говоря не выполнен закон сокращения, скажем, для любого подпространства $U \leqslant V$ справедливо равенство $U + U = U$. Также не работает аналогия между суммой подпространств и теоретико-множественным объединением: скажем не всегда $(U_1 + U_2) \cap U_3 = (U_1 \cap U_3) + (U_2 \cap U_3)$ в отличие от теоретико-множественного тождества $(U_1 \cup U_2) \cap U_3 = (U_1 \cap U_3) \cup (U_2 \cap U_3)$.

Упражнение. Приведите пример подпространств U_1, U_2, U_3 некоторого векторного пространства V таких, что $(U_1 + U_2) \cap U_3 \neq (U_1 \cap U_3) + (U_2 \cap U_3)$.

Прямая сумма

Важен следующий специальный случай суммы подпространств $U_i \leqslant V$.

Определение. Сумма U подпространств $U_1, U_2, \ldots, U_k$ называется (внутренней) *прямой суммой*, если $\forall \mathbf{a} \in U$ имеется единственный набор $\mathbf{a}_i \in U_i$ $(i = 1, 2, \ldots, k)$ такой, что $\mathbf{a} = \sum\limits_{i=1}^{k} \mathbf{a}_i$.

Обозначение для прямой суммы: $U_1 \bigoplus U_2 \bigoplus \ldots \bigoplus U_k$ или $\bigoplus\limits_{i=1}^{k} U_i$. Иногда говорят, что U *разложено в прямую сумму* подпространств $U_1, U_2, \ldots, U_k$. Пример разложения в прямую сумму: $V = \langle \mathbf{e}_1 \rangle \bigoplus \langle \mathbf{e}_2 \rangle \bigoplus \ldots \bigoplus \langle \mathbf{e}_n \rangle$, где $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ — некоторый базис в V.

Наряду с внутренней прямой суммой, которой в основном будем заниматься, можно рассмотреть *внешнюю прямую сумму* векторных пространств $U_1, U_2, \ldots, U_k$ как декартово произведение $U_1 \times U_2 \times \ldots \times U_k$, на котором операции сложения и умножения на константу определены «покомпонентно». Конструкции внутренней и внешней прямой суммы по сути эквивалетны (связаны естественным изоморфизмом).

В случае рассмотрения суммы $U = \sum\limits_{i=1}^{k} U_i$ через $\overline{U_i}$ обозначаем сумму всех рассматриваемых пространств, за исключением U_i, т.е. $\overline{U_i} = U_1 + \ldots + U_{i-1} + U_{i+1} + \ldots + U_k$.

Теорема 3.1 (критерий-1 прямой суммы). *Пусть* $U_i \leqslant V$, $i = 1, \ldots, k$. *Сумма подпространств* $U_1, \ldots, U_k$ *— прямая сумма* $\Leftrightarrow U_i \cap \overline{U_i} = O$, $i = 1, 2, \ldots, k$.

$\triangleright$ $\boxed{\Rightarrow}$ Предположим противное, скажем условие $U_i \cap \overline{U_i} = O$ не выполнено для $i = 1$. Это значит, что нашелся ненулевой вектор $\mathbf{a}_1 \in U_1 \cap \overline{U_1}$. Имеем $\mathbf{a}_1 = \mathbf{a}_2 + \ldots + \mathbf{a}_k$ для некоторых $\mathbf{a}_i \in U_i$, $i = 2, \ldots, k$. Но тогда $\mathbf{0} = \mathbf{0} + \ldots + \mathbf{0} = \mathbf{a}_1 + (-\mathbf{a}_2) + \ldots + (-\mathbf{a}_k)$ — два различных разложения нулевого вектора в сумму векторов из U_i вопреки определению прямой суммы.

$\boxed{\Leftarrow}$ Предположим противное, сумма $U = \sum\limits_{i=1}^{k} U_i$ не является прямой суммой, тогда для некоторого вектора $\mathbf{a} \in U$ найдутся два различных разложения $\mathbf{a} = \sum\limits_{i=1}^{k} \mathbf{a}_i = \sum\limits_{i=1}^{k} \mathbf{b}_i$, где $\mathbf{a}_i \in U_i$, $\mathbf{b}_i \in U_i$, $i = 1, \ldots, k$. Разложения различны, значит хотя бы для одного i имеем $\mathbf{a}_i \neq \mathbf{b}_i$, скажем $\mathbf{a}_1 \neq \mathbf{b}_1$. Но тогда $\mathbf{a}_1 - \mathbf{b}_1 = \sum\limits_{i=2}^{k} (\mathbf{b}_i - \mathbf{a}_i)$. В левой части равенства — ненулевой вектор из U_1, а в правой части — вектор из $\overline{U_1}$, поэтому $U_1 \cap \overline{U_1} \neq O$. Противоречие. $\square$

Указанным критерием часто удобно пользоваться в случае двух подпространств. Этот случай выделим отдельным следствием.

Следствие. *Пусть* $U_1 \leqslant V$, $U_2 \leqslant V$. *Тогда* $U_1 + U_2$ *— прямая сумма* $\Leftrightarrow U_1 \cap U_2 = O$.

В отличие случая двух подпространств, условие тривиальности попарных пересечений $U_1 \cap U_2 = U_2 \cap U_3 = U_3 \cap U_1 = O$ не является достаточным для того, чтобы сумма подпространств $U_1 + U_2 + U_3$ была прямой суммой. Контрпримером могут служить три различных одномерных пространства, лежащие в двумерном пространстве.

Упражнение. Сумма подпространств $U_1, \ldots, U_k$ — прямая сумма $\Leftrightarrow \forall\ \mathbf{a}_i \in U_i$, $\mathbf{a}_i \neq \mathbf{0}$ система $\mathbf{a}_1, \ldots, \mathbf{a}_k$ линейно независима.

Упражнение. Пусть $U_i \leqslant V$ таковы, что $U_1 \oplus U_2 \oplus U_3$ — прямая сумма. Тогда $U = U_2 \oplus U_3$ — прямая сумма и $U_1 \oplus U$ — также прямая сумма.

Упражнение. Пусть $U_i \leqslant V$, $i = 1, \ldots, k$, — подпространства, для которых выполнено: $U_j \cap (\sum\limits_{i=1}^{j-1} U_i) = O$ для всех $j = 2, 3, \ldots, k$. Тогда $\sum\limits_{i=1}^{k} U_i$ — прямая сумма.

Теорема 3.2 (критерий-2 прямой суммы). *Пусть* $U_i \leqslant V$, $\dim U_i = n_i < \infty$, $\mathrm{e}^{(i)}$ *— базис в* U_i, $i = 1, \ldots, k$; $U = \sum\limits_{i=1}^{k} U_i$. *Тогда следующие условия эквивалентны:*

1) $U = \bigoplus\limits_{i=1}^{k} U_i$;

2) система из $\sum\limits_{i=1}^{k} n_i$ *векторов* $\bigcup\limits_{i=1}^{k} \mathrm{e}^{(i)}$ *— базис в* U;

3) $\dim U = \sum\limits_{i=1}^{k} n_i$.

$\triangleright$ Имеем $U_i = \langle \mathrm{e}^{(i)} \rangle$, тогда согласно предложению 3.1, $U = \langle \mathrm{e} \rangle$, где $\mathrm{e} = \bigcup\limits_{i=1}^{k} \mathrm{e}^{(i)}$.

2) $\Leftrightarrow$ 3) По следствию из основной теоремы 2.1 имеем $\dim U = \operatorname{rg} \mathrm{e}$. Значит (см. предложение 2.4) $\dim U = \sum\limits_{i=1}^{k} n_i \Leftrightarrow \operatorname{rg} \mathrm{e} = \sum\limits_{i=1}^{k} n_i \Leftrightarrow$ система e линейно независима.

1) $\Rightarrow$ 2) Предположим противное, и система e линейно зависима. Запишем нетривиальную линейную комбинацию векторов из e, равную $\mathbf{0}$: $\sum\limits_{i=1}^{k} \ell_i = \mathbf{0}$, где ℓ_i — линейная комбинация векторов из $\mathrm{e}^{(i)}$. Хотя бы одна из линейных комбинаций ℓ_i нетривиальная, пусть это ℓ_1, тем самым $\ell_1 \neq \mathbf{0}$. Тогда $\ell_1 = -\ell_2 - \ldots - \ell_k$. Здесь ℓ_i равно некоторому вектору из U_i, значит $\ell_1 \in U_1 \cap \overline{U_1}$ в противоречие с теоремой 3.1.

2) $\Rightarrow$ 1) Предположим противное, U не являетяся прямой суммой подпространств U_i, и скажем (см. теорему 3.1) $\exists\, \mathbf{a} \neq \mathbf{0}$: $\mathbf{a} \in U_1 \cap \overline{U_1}$. Тогда $\mathbf{a} = \ell_1 = \ell_2 + \ldots + \ell_k$, где $\ell_i \in U_i$, т.е. ℓ_i равно некоторой линейной комбинации векторов из $\mathrm{e}^{(i)}$. Перенося в левую часть, получаем $\ell_1 - \ell_2 - \ldots - \ell_k = \mathbf{0}$ (в левой части нетривиальная линейная комбинация, поскольку уже ℓ_1 — нетривиальная линейная комбинация), откуда e — линейно зависимая система. Противоречие. □

Прямое дополнение. Проекции

> **Определение.** Если $U_1 \leqslant V$ и $U_2 \leqslant V$ таковы, что $U_1 \bigoplus U_2 = V$, то подпространство U_2 называют *прямым дополнением* подпространства U_1 (в векторном пространстве V).

Подпространства U_1 и U_2 входят в определение симметрично, поэтому: U_1 — прямое дополнение для $U_2 \Leftrightarrow U_2$ — прямое дополнение для U_1. Отметим, что прямое дополнение ничего общего не имеет с понятием теоретико-множественного дополнения.

Предложение 3.3. *Пусть* $\dim V = n < \infty$. *Тогда сумма размерностей подпространства и любого его прямого дополнения равна* n.

$\triangleright$ Следует из теоремы 3.2. □

Предложение 3.4. *Пусть* $\dim V = n < \infty$. *Для любого подпространства* $U \leqslant V$ *существует прямое дополнение.*

$\triangleright$ Выберем в U базис $\mathbf{e}_1, \ldots, \mathbf{e}_k$. Согласно предложению 2.8, систему $\mathbf{e}_1, \ldots, \mathbf{e}_k$ можно дополнить до базиса $\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{e}_{k+1}, \ldots, \mathbf{e}_n$ пространства V. Тогда из теоремы 3.2 вытекает, что подпространство $W = \langle \mathbf{e}_{k+1}, \ldots, \mathbf{e}_n \rangle$ таково, что $U \oplus W = V$. □

Заметим, что для одного пространства может существовать много прямых дополнений (достаточно посмотреть на геометрический пример $U \leqslant V$, где $\dim U = 1$, $\dim V = 2$).

Если $V = U_1 \bigoplus U_2$, то, согласно определению прямой суммы, любой вектор $\mathbf{a} \in V$ однозначно представляется в виде суммы $\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_2$, где $\mathbf{a}_i \in U_i$, $i = 1, 2$. Вектор $\mathbf{a}_1$ называется *проекцией* вектора $\mathbf{a}$ на подпространство U_1 *вдоль* U_2 (или *параллельно* U_2).

Для $U \leqslant V$ определяется *фактор-пространство* V/U как фактор-множество относительно эквивалентности $\mathbf{a} \sim \mathbf{b} \Leftrightarrow \mathbf{b} - \mathbf{a} \in U$; элементы V/U — смежные классы вида $\mathbf{a} + U$ с естественно определенными операциями $(\mathbf{a}+U)+(\mathbf{b}+U) := (\mathbf{a}+\mathbf{b})+U$, $\lambda(\mathbf{a}+U) := \lambda\mathbf{a}+U$. Имеется естественный изоморфизм $W \to V/U$, где W — прямое дополнение подпространства $U \leqslant V$; он задается как $\mathbf{a} \mapsto \mathbf{a} + U$.

Формула размерностей суммы и пересечения

Теорема 3.3 (формула Грассмана)**.** *Пусть* $U_1 \leqslant V$, $U_2 \leqslant V$. *Тогда*

$$\boxed{\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = \dim U_1 + \dim U_2}.$$

$\triangleright$ Достаточно рассмотреть случай $\dim U_i < \infty$ (иначе в обеих частях формулы — бесконечность).

Согласно предложению 3.4, можем выбрать $W \leqslant U_2$ так, что

$$(U_1 \cap U_2) \oplus W = U_2. \tag{1.1}$$

Тогда $U_1 + U_2 = U_1 + ((U_1 \cap U_2) + W) = (U_1 + (U_1 \cap U_2)) + W = U_1 + W$. Кроме того, поскольку $W \leqslant U_2$, имеем $U_1 \cap W = U_1 \cap U_2 \cap W = (U_1 \cap U_2) \cap W = O$ (из (1.1) по следствию из теоремы 3.1). Получаем, что $U_1 + W$ — прямая сумма:

$$U_1 + U_2 = U_1 \oplus W. \tag{1.2}$$

Согласно теореме 3.2, из равенств (1.1) и (1.2) следует, что
$\dim W = \dim U_2 - \dim(U_1 \cap U_2) = \dim(U_1 + U_2) - \dim U_1$, откуда следует требуемая формула размерностей. ☐

Упражнение. Пусть $\dim V = 4$. Существуют ли подпространства U_1 и U_2 такие, что $\dim U_1 = \dim U_2 = 3$ и $\dim(U_1 \cap U_2) = 1$?

Примеры

I.1. Сумма двух непараллельных отрезков — параллелограмм. (Здесь, как обычно, точки отождествляем с концами радиус-векторов.)

I.2. Пусть V — геометрическое трехмерное векторное пространство, $U_1 \leqslant V$, $\dim U_1 = 1$ (т.е. U_1 — прямая, проходящая через O), $U_2 \leqslant V$, $\dim U_2 = 2$ (т.е. U_2 — плоскость, проходящая через O). Если $U_1 \not\subset U_2$, то $V = U_1 \oplus U_2$.
Понятно, как геометрически разложить радиус-вектор $\overrightarrow{OA}$ в сумму проекций: через A проведем прямую, параллельную U_1; пусть она пересекает U_2 в точке A_2. Тогда $\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_2$, где $\mathbf{a}_1 = \overrightarrow{A_2A}$, $\mathbf{a}_2 = \overrightarrow{OA_2}$.

II.1. Множество решений совместной СЛУ (над $\mathbb{R}$) $AX = b$ с n неизвестными имеет вид $\mathrm{Sol}(AX = b) = \{X_0\} + \mathrm{Sol}(AX = O)$. Если $r = \operatorname{rg} A$, то $\mathrm{Sol}(AX = b)$ — это $(n-r)$-мерная плоскость (или $(n-r)$-мерное *линейное многообразие*) в пространстве $\mathbb{R}^n$.

II.2. Пусть $\mathbf{M}_{n\times n} = \mathbf{M}_{n\times n}(\mathbb{F})$, где $\mathbb{F}$ — поле характеристики, не равной 2. Тогда
$\mathbf{M}_{n\times n} = \mathbf{M}^+_{n\times n} \bigoplus \mathbf{M}^-_{n\times n}$.
Действительно, $A = \dfrac{A + A^T}{2} + \dfrac{A - A^T}{2}$ и $\mathbf{M}^+_{n\times n} \cap \mathbf{M}^-_{n\times n} = O$.

II.3. Пусть $A \in \mathbf{M}_{n\times m}(\mathbb{R})$, $A = (a_{\bullet 1} \dots a_{\bullet m})$. Положим $U = \langle a_{\bullet 1} \dots a_{\bullet m} \rangle$, $\dim U = \operatorname{rg} A = r$. Пусть $W = \mathrm{Sol}(A^T X = O)$. Тогда $\dim W = n - r$. Кроме того $U \cap W = O$. Действительно (естественное объяснение получается также в **??** главы 5), если $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ таков, что $Y \in U \cap W$, то есть $Y^T Y = O$, откуда $\sum\limits_{i=1}^{n} y_i^2 = 0$, значит $Y = O$.

III.1. Несложно показать, что $\mathbf{F} = \mathbf{F}^+ \bigoplus \mathbf{F}^-$, где $\mathbf{F}^+$, $\mathbf{F}^-$ — подпространства четных и нечетных функций $\mathbb{R} \to \mathbb{R}$.
Заметим, что $e^x = \operatorname{ch} x + \operatorname{sh} x$, при этом $f(x) = \operatorname{ch} x$ — четная функция, а $g(x) = \operatorname{sh} x$ — нечетная функция. Поэтому $f(x) = \operatorname{ch} x$ — это проекция функции $y(x) = e^x$ на $\mathbf{F}^+$ вдоль $\mathbf{F}^-$.

§ 4. Понятие аффинного пространства

Идея конструкции абстрактного точечного (аффинного) пространства — в сопоставлении
«точка $\leftrightarrow$ радиус-вектор».

Определение и свойства

Аффинным пространством, ассоциированным с векторным пространством V, называется множество S (элементы которого будем называть {точками}) с отображением $S \times V \to S$ (откладываение от точки вектора, обозначать будем «$+$»), такие что
A1) $p + (\mathbf{a} + \mathbf{b}) = (p + \mathbf{a}) + \mathbf{b} \qquad (\forall p \in S, \forall \mathbf{a}, \mathbf{b} \in V\ );$
A2) $p + \mathbf{0} = p; \qquad (\forall p \in S);$
A3) $\forall p, q \in S$ существует единстивенный $\mathbf{a} \in V$ такой, что $p + \mathbf{a} = q$.

Вектор $\mathbf{a}$ из A3) называем *вектором, соединяющим точки* p и q, и обозначаем $\overline{pq}$.

Видим «правило треугольника»: $\overline{pq} + \overline{qr} = \overline{pr}$.

V естественно является аффинным пространством, ассоциированным с V.

Наоборот, если в афиинном пространстве зафиксировать точку o («начало отсчета»), то возникает биекция (*векторизация*) $S \to V$:

$$p \mapsto \overline{op}.$$

(зависит от выбора начала отсчета)

(o, e) — *аффинная система координат* в S (или *репер*)
(в аналитической геометрии было: ДСК)

Координаты точки $p \in S$ в ДСК (o, e) —
это координаты вектора $\overline{op}$ в базисе e.

координаты точки $p + \mathbf{a}$???

координаты вектора $\overline{pq}$

Замена координат:

$$X = SX' + \gamma$$

Примеры подмножеств, конструкций

Плоскость, или *линейное подмногообразие* —

$$p + U,$$

где $U \leqslant V$.
Размерность k-мерная плоскость ($k = 0$ — точка, $k = 1$ — прямая, $k = n-1$ — гиперплоскость).

Любые $k+1$ точек из S принадлежат некоторой плоскости размерности $\leqslant k$.
$p_0 + \langle \overline{p_0p_1}, \overline{p_0p_2}, \ldots, \overline{p_0p_k} \rangle$.

Аффинная зависимость, независимость системы из $k+1$ точек.
Связь с линейной зависимостью векторов.

Линейные комбинации точек (чтобы была корректность, сумма коэффициентов должна равняться 1, пример — центр масс.)

паралелелепипед, симмлекс.

выпуклые линейные комбинации, выпуклая оболочка.

Глава 2

Линейные отображения

В этой главе рассматриваются векторные пространства V, $\widetilde{V}$ и т.д. над произвольным полем $\mathbb{F}$. (При этом для обозначения операции сложения в разных пространствах мы позволяем себе использовать один и тот же символ «+»; то же касается нулевых векторов, и т.д.)

В ситауциях, когда используется специфика поля $\mathbb{F}$, оговариваем это отдельно. (Например, для $\mathbb{F} = \mathbb{C}$ умножение линейного отображения на константу может быть определено особым способом.)

§ 1. Определение. Операции над линейными отображениями. Изоморфизм

Определение и его следствия

> **Определение.** Отображение $\varphi : V \to \widetilde{V}$ называется *линейным*, если $\forall\ \mathbf{a}, \mathbf{b} \in V$ и $\forall\ \lambda \in \mathbb{F}$ выполняются равенства
> L1. $\varphi(\mathbf{a} + \mathbf{b}) = \varphi(\mathbf{a}) + \varphi(\mathbf{b})$,
> L2. $\varphi(\lambda \mathbf{a}) = \lambda \varphi(\mathbf{a})$.

Линейное отображение также называют *гомоморфизмом* векторных пространств. Множество всех линейных отображений $V \to \widetilde{V}$ обычно обозначают $\mathrm{Hom}(V; \widetilde{V})$. Мы чаще будем использовать чуть более короткое обозначение $L(V, \widetilde{V})$.

Линейное отображение $V \to V$ (т.е. в частном случае $\widetilde{V} = V$) называют также *линейным преобразованием* или *линейным оператором.*

Линейное отображение $V \to \mathbb{F}$ называют также *линейным функционалом* или *линейной функцией.* Это соответствует случаю $\dim \widetilde{V} = 1$, в этом случае $\widetilde{V}$ можно отождествить с полем констант $\mathbb{F}$.

> Примером линейного отображения является *нулевое* отображение $0 : V \to \widetilde{V}$ такое, что $\forall$ $\mathbf{a} \in V$ выполнено $0(\mathbf{a}) = \mathbf{0}$.

Предложение 1.1. *Пусть* $\varphi \in L(V, \widetilde{V})$. *Тогда* $\forall\ \mathbf{a}_i \in V$ *и* $\forall\ \lambda_i \in \mathbb{F}$ *выполнено*

$$\boxed{\varphi\left(\sum_{i=1}^{k} \lambda_i \mathbf{a}_i\right) = \sum_{i=1}^{k} \lambda_i \varphi(\mathbf{a}_i)}. \tag{2.1}$$

▷ Следует из многократного применения L1, L2. □

Отметим, что L1 и L2 представляют собой частные случаи фомулы (2.1). Зафиксируем несложные, но важные следствия определения и предыдущего предложения.

Предложение 1.2. *Пусть* $\varphi \in L(V, \widetilde{V})$. *Тогда*
1). $\varphi(\mathbf{0}) = \mathbf{0}$;
2). $\forall\ \mathbf{a} \in V$ *выполнено* $\varphi(-\mathbf{a}) = -\varphi(\mathbf{a})$.

$\triangleright$ 1). Следует из L2 для $\lambda = 0$.
2). Следует из L2 для $\lambda = -1$. $\square$

Предложение 1.3. *1). Если* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ *— линейно зависимая система векторов, то* $\varphi(\mathbf{a}_1), \varphi(\mathbf{a}_2), \ldots, \varphi(\mathbf{a}_k)$ *— тоже линейно зависимая система векторов.*
2). $\forall\ \mathcal{A} \subset V$ *выполнено* $\operatorname{rg}\varphi(\mathcal{A}) \leqslant \operatorname{rg}\mathcal{A}$.

$\triangleright$ 1). Следует из предложения 1.1 и пункта 1) предложения 1.2.
2). Следует из 1). $\square$

Предложение 1.4 (образ подпространства)**.** *Пусть* $\varphi \in L(V, \widetilde{V})$, $U \leqslant V$ *и* $U = \langle \mathcal{A} \rangle$. *Тогда*
1). $\varphi(U) \leqslant \widetilde{V}$, *более того,* $\varphi(U) = \langle \varphi(\mathcal{A}) \rangle$. *В частности, если* $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ *— базис в* V, *то (напомним, что* $\varphi(V)$ *также обозначается* $\operatorname{Im}\varphi$*)*

$$\boxed{\operatorname{Im}\varphi = \langle \varphi(\mathbf{e}_1), \varphi(\mathbf{e}_2), \ldots, \varphi(\mathbf{e}_n) \rangle}.$$

2). $\dim \varphi(U) \leqslant \dim U$.

$\triangleright$ 1). Следует из предложения 1.1.
2). Следует из 1) (и является частным случаем пункта 2 предложения 1.3). $\square$

В следующем предложении отметим отдельно свойства линейного вложения (инъективного отображения).

Предложение 1.5. *Пусть* $\varphi \in L(V, \widetilde{V})$ *и* φ *инъективно. Тогда*
1). Если $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ *— линейно независимая система векторов, то* $\varphi(\mathbf{a}_1), \varphi(\mathbf{a}_2), \ldots, \varphi(\mathbf{a}_k)$ *— тоже линейно независимая система векторов.*
2). $\forall\ \mathcal{A} \subset V$ *выполнено* $\operatorname{rg}\varphi(\mathcal{A}) = \operatorname{rg}\mathcal{A}$. *В частности, для* $U \leqslant V$ *выполнено* $\dim \varphi(U) = \dim U$.

$\triangleright$ 1). Пусть $\sum\limits_{i=1}^{k} \lambda_i \varphi(\mathbf{a}_i) = \mathbf{0}$, тогда $\varphi\left(\sum\limits_{i=1}^{k} \lambda_i \mathbf{a}_i\right) = \mathbf{0}$. Но $\varphi(\mathbf{0}) = \mathbf{0}$, поэтому в силу инъективности, $\sum\limits_{i=1}^{k} \lambda_i \mathbf{a}_i = \mathbf{0}$. Отсюда, поскольку $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ — линейно независимая система, имеем $\lambda_1 = \ldots = \lambda_k = 0$, то есть $\sum\limits_{i=1}^{k} \lambda_i \varphi(\mathbf{a}_i)$ — тривиальная линейная комбинация.
2). Следует из 1). $\square$

Следующая теорема показывает, что линейное отображение определено, причем единственным образом, образами базисных векторов.

Теорема 1.1. *Пусть* $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ *— базис в* V, *и* $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ *— фиксированные векторы из* $\widetilde{V}$. *Тогда*
1). существует единственное линейное отображение $\varphi : V \to \widetilde{V}$ *такое, что* $\varphi(\mathbf{e}_i) = \mathbf{c}_i$, $i = 1, 2, \ldots, n$.
2). При этом φ *инъективно* $\Leftrightarrow$ *система* $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ *линейно независима.*

$\triangleright$ 1). *Единственность.* Пусть $\mathbf{a} \in V$ разложен по базису e: $\mathbf{a} = \sum\limits_{i=1}^{n} x_i \mathbf{e}_i$. Тогда, в силу (2.1), однозначно получаем

$$\varphi(\mathbf{a}) = \sum_{i=1}^{n} x_i \varphi(\mathbf{e}_i) = \sum_{i=1}^{n} x_i \mathbf{c}_i. \tag{2.2}$$

Существование. Очевидно, что формула (2.2) удовлетворяет условию $\varphi(\mathbf{e}_i) = \mathbf{c}_i$, $i = 1, 2, \ldots, n$. Достаточно доказать, что она действительно определяет линейное отображение. Для произвольных векторов $\mathbf{a}, \mathbf{b} \in V$, таких, что $\mathbf{a} = \sum\limits_{i=1}^{n} x_i \mathbf{e}_i$, $\mathbf{b} = \sum\limits_{i=1}^{n} y_i \mathbf{e}_i$, по определению отображения φ (с

учетом предложения 2.9 главы 1) имеем: $\varphi(\mathbf{a}) = \sum\limits_{i=1}^{n} x_i\mathbf{c}_i$, $\varphi(\mathbf{b}) = \sum\limits_{i=1}^{n} y_i\mathbf{c}_i$, $\varphi(\mathbf{a}+\mathbf{b}) = \sum\limits_{i=1}^{n}(x_i + y_i)\mathbf{c}_i$, $\varphi(\lambda\mathbf{a}) = \sum\limits_{i=1}^{n}(\lambda x_i)\mathbf{c}_i$. Теперь легко видеть, что L1 и L2 выполнены.
2). $\boxed{\Rightarrow}$ Следует из предложения 1.5.
$\boxed{\Leftarrow}$ Пусть $\varphi(\mathbf{a}) = \varphi(\mathbf{b})$ для некоторых векторов $\mathbf{a} \neq \mathbf{b}$, тогда $\varphi(\mathbf{a} - \mathbf{b}) = \mathbf{0}$. Разложим ненулевой вектор $\mathbf{a}-\mathbf{b}$ по базису e: $\mathbf{a}-\mathbf{b} = \sum\limits_{i=1}^{n} \lambda_i\mathbf{e}_i$. Тогда $\varphi(\mathbf{a}-\mathbf{b}) = \sum\limits_{i=1}^{n} \lambda_i\mathbf{c}_i = \mathbf{0}$. Получаем, что нетривиальная линейная комбинация векторов $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ равна $\mathbf{0}$. Противоречие. □

Утверждение теоремы 1.1 дает понимание, насколько мы свободны в определении линенйного отображения, оно бывает полезно при конструировании линейных отображений с заданными свойствами.

Упражнение. Может ли для некоторых $\varphi \in L(V, V)$ и $\mathbf{a} \in V$ выполняться одновременно условия $\varphi(\mathbf{a}) \neq \mathbf{0}$, $\varphi(\varphi(\mathbf{a})) = \mathbf{0}$?

Изоморфизм. Композиция (произведение). Обратное отображение.

Определение. Отображение $\varphi : V \to \widetilde{V}$ называется *изоморфизмом*, если оно линейно и биективно.

Определение. Векторные пространства V и $\widetilde{V}$ называются *изоморфными*, если существует изоморфизм $\varphi : V \to \widetilde{V}$.

Тот факт, что V и $\widetilde{V}$ изоморфны, обозначаем $V \cong \widetilde{V}$.

Предложение 1.6. *1). Пусть $\varphi \in L(V, \widetilde{V})$, $\psi \in L(\widetilde{V}, \widehat{V})$. Тогда $\psi\varphi \in L(V, \widehat{V})$.*
2). Если кроме того φ и ψ — изоморфизмы, то $\psi\varphi$ — также изоморфизм.

▷ 1). Проверим L1 для отображения $\psi\varphi$. Из определения композиции и условия L1 для отображений ψ и φ имеем: $\forall\ \mathbf{a}, \mathbf{b} \in V$ выполнено
$(\psi\varphi)(\mathbf{a} + \mathbf{b}) = \psi(\varphi(\mathbf{a} + \mathbf{b})) = \psi(\varphi(\mathbf{a}) + \varphi(\mathbf{b})) = \psi(\varphi(\mathbf{a})) + \psi(\varphi(\mathbf{b})) = (\psi\varphi)(\mathbf{a}) + (\psi\varphi)(\mathbf{b})$.
Аналогично проверяется L2.
2). Следует из 1) и того, что композиция биективных отображений биективна. □

Говорят, что два преобразования $\varphi, \psi \in L(V, V)$ *перестановочные* (или *коммутируют*), если $\varphi\psi = \psi\varphi$. Отметим, что композиция линейных преобразований вообще говоря не подчиняется тождеству $\varphi\psi = \psi\varphi$. Нетрудно привести соответствующие примеры (например, используя замечание после теоремы 1.1).

Предложение 1.7. *Если отображение $\varphi : V \to \widetilde{V}$ — изоморфизм, то φ^{-1} — также изоморфизм.*

▷ Для данных векторов $\mathbf{a}, \mathbf{b} \in \widetilde{V}$ однозначно определены векторы $\mathbf{c} = \varphi^{-1}(\mathbf{a})$ и $\mathbf{d} = \varphi^{-1}(\mathbf{b})$.
Так как $\varphi(\mathbf{c}+\mathbf{d}) = \varphi(\mathbf{c})+\varphi(\mathbf{d}) = \mathbf{a}+\mathbf{b}$, то $\mathbf{c}+\mathbf{d} = \varphi^{-1}(\mathbf{a}+\mathbf{b})$, то есть $\varphi^{-1}(\mathbf{a}+\mathbf{b}) = \varphi^{-1}(\mathbf{a})+\varphi^{-1}(\mathbf{b})$.
Далее, $\varphi(\lambda\mathbf{c}) = \lambda\varphi(\mathbf{c}) = \lambda\mathbf{a}$, откуда $\varphi^{-1}(\lambda\mathbf{a}) = \lambda\mathbf{c} = \lambda\varphi^{-1}(\mathbf{a})$. □

Для целого неотрицательного k определим k-ую степень преобразования $\varphi : V \to V$ как $\varphi^k = \underbrace{\varphi\varphi\ldots\varphi}_{k \text{ букв } \varphi}$ при $k > 0$ и как тождественное преобразование I_V при $k = 0$. При этом справедливы равенства $\varphi^{m+k} = \varphi^m\varphi^k$ и $\varphi^{mk} = (\varphi^m)^k$. Если кроме того φ — изоморфизм, то можно определить k-ую степень и для отрицательных k как $\varphi^k = (\varphi^{-1})^{-k}$. Нетрудно проверить, что в этом случае равенства $\varphi^{m+k} = \varphi^m\varphi^k$ и $\varphi^{mk} = (\varphi^m)^k$ остаются в силе для всех $m, k \in \mathbb{Z}$.

Очевидно $V \cong V$ (пример изоморфизма — тождественное преобразование $I_V : V \to V$). Предложения 1.6 и 1.7 показывают, что отношение «быть изоморфными» симметрично и транзитивно, т.е. $V \cong \widetilde{V} \Rightarrow \widetilde{V} \cong V$ и $V \cong \widetilde{V}$, $\widetilde{V} \cong \widehat{V} \Rightarrow V \cong \widehat{V}$. Тем самым, все векторные пространства можно мыслить разбитыми на классы эквивалентности (так что два пространства из одного класса эквивалентности изоморфны, а из разных — не изоморфны). Следующая теорема дает классификацию конечномерных векторных пространств.

Теорема 1.2. *Пусть* $\dim V < \infty$, $\dim \widetilde{V} < \infty$. *Тогда*

$$V \cong \widetilde{V} \ \Leftrightarrow\ \dim V = \dim \widetilde{V}.$$

$\triangleright$ $\boxed{\Rightarrow}$ Если $\varphi : V \to \widetilde{V}$ — изоморфизм, то φ инъективно и сюръективно. Значит, согласно предложению 1.5, имеем $\dim V = \dim(\varphi(V)) = \dim \widetilde{V}$.

$\boxed{\Leftarrow}$ Пусть $\dim V = \dim \widetilde{V} = n$. В предложении 2.9 главы 1 мы фактически сталкивались с изоморфизмом — сопоставлением вектору его координатного столбца (в фиксированном базисе). Значит, у нас есть «эталонное» n-мерное пространство $\mathbb{F}^n = \mathbf{M}_{n\times 1}$, так что $V \cong \mathbf{M}_{n\times 1}$, $\widetilde{V} \cong \mathbf{M}_{n\times 1}$, и следовательно, $V \cong \widetilde{V}$. $\square$

В следующей теореме соберем условия, эквивалентные изоморфности данного линейного отображения в конечномерном случае.

Теорема 1.3. *Пусть* $\dim V = \dim \widetilde{V} = n < \infty$, $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ *— базис в* V. *Пусть* $\varphi \in L(V, \widetilde{V})$. *Тогда следующие условия эквивалентны:*
1) φ *— изоморфизм;*
2) φ *— инъекция;*
3) φ *— сюръекция;*
4) $\varphi(\mathbf{e}_1), \varphi(\mathbf{e}_2), \ldots, \varphi(\mathbf{e}_n)$ *— базис в* $\widetilde{V}$.

$\triangleright$ 1) $\Rightarrow$ 2) Очевидно.

2) $\Rightarrow$ 3) Согласно предложению 1.5, $\dim(\varphi(V)) = \dim V = n$. Значит, по следствию 2 из теоремы 2.1, имеем $\varphi(V) = V$, то есть φ сюръективно.

3) $\Rightarrow$ 4) Как мы знаем (см. предложение 1.4), $\varphi(V) = \langle \varphi(\mathbf{e}_1), \varphi(\mathbf{e}_2), \ldots, \varphi(\mathbf{e}_n)\rangle$. Из того, что $\dim \varphi(V) = n$, следует, что $\mathrm{rg}(\varphi(\mathbf{e}_1), \varphi(\mathbf{e}_2), \ldots, \varphi(\mathbf{e}_n)) = n$, значит, система векторов $\varphi(\mathbf{e}_1), \varphi(\mathbf{e}_2), \ldots, \varphi(\mathbf{e}_n)$ линейно независима, т.е. является базисом в V.

4) $\Rightarrow$ 1) Согласно теореме 1.1, φ инъективно. А поскольку $\varphi(V) = \langle \varphi(\mathbf{e}_1), \varphi(\mathbf{e}_2), \ldots, \varphi(\mathbf{e}_n)\rangle$, φ сюръективно. $\square$

Линейные операции на $\mathrm{Hom}(V; \widetilde{V})$.

> **Определение.** *Суммой отображений* $\varphi, \psi \in L(V, \widetilde{V})$ называется такое отображение $\eta : V \to \widetilde{V}$, что $\forall \mathbf{a} \in V$ выполнено
> $\eta(\mathbf{a}) = \varphi(\mathbf{a}) + \psi(\mathbf{a})$.

> **Определение.** Произведением отображения $\varphi \in L(V, \widetilde{V})$ на константу $\lambda \in \mathbb{F}$ называется такое отображение $\psi : V \to \widetilde{V}$, что $\forall \mathbf{a} \in V$ выполнено
> $\psi(\mathbf{a}) = \lambda\varphi(\mathbf{a})$.

Результат операции сложения и умножения на число называется, как обычно, суммой и произведением на число, и обозначаются обычным образом: $\varphi + \psi$ и $\lambda\varphi$. Таким образом, определения означают выполнение следующих равенств, которые выглядят вполне естественно:

$$(\varphi + \psi)(\mathbf{a}) = \varphi(\mathbf{a}) + \psi(\mathbf{a});$$

$$(\lambda\varphi)(\mathbf{a}) = \lambda\varphi(\mathbf{a}).$$

При $\mathbb{F} = \mathbb{C}$ умножение линейного отображения на константу может быть определено особым способом (и использованием комплексного сопряжения). *Второй вариант* определения:

$$(\lambda\varphi)(\mathbf{a}) = \overline{\lambda}\varphi(\mathbf{a}).$$

Предложение 1.8. *Если $\varphi, \psi \in L(V, \widetilde{V})$, то $\varphi + \psi \in L(V, \widetilde{V})$ и $\lambda\varphi \in L(V, \widetilde{V})$.*

$\rhd$ Проверяется непосредственно. Например, проверим L2 для $\lambda\varphi$ со вторым вариантом определения. Имеем $(\lambda\varphi)(\mu\mathbf{a}) = \overline{\lambda}\varphi(\mu\mathbf{a}) = \overline{\lambda}\mu\varphi(\mathbf{a}) = \mu(\overline{\lambda}\varphi(\mathbf{a})) = \mu(\lambda\varphi)(\mathbf{a}))$. Тем самым, $\lambda\varphi$ удовлетворяет L2. $\square$

Предложение 1.9. *Множество $L(V, \widetilde{V})$ является векторным пространством относительно введенных выше операций сложения и умножения на число.*

$\rhd$ Проверяется непосредственно. $\square$

При $\mathbb{F} = \mathbb{C}$ множество $L(V, \widetilde{V})$ наделяется структурой векторного пространства двумя способами, в соответствии с двумя способами определения умножения на константу. В случае рассмотрения второго варианта вместо $L(V, \widetilde{V})$ пишем $\overline{L}(V, \widetilde{V})$.

Предложение 1.10. *Пусть $\varphi, \psi \in L(V, \widetilde{V})$, $\widetilde{\varphi}, \widetilde{\psi} \in L(\widetilde{V}, \widehat{V})$. Тогда*
1. $\widetilde{\varphi}(\varphi + \psi) = \widetilde{\varphi}\varphi + \widetilde{\varphi}\psi$;
2. $(\widetilde{\varphi} + \widetilde{\psi})\varphi = \widetilde{\varphi}\varphi + \widetilde{\psi}\varphi$;
3. $\lambda(\widetilde{\varphi}\varphi) = (\lambda\widetilde{\varphi})\varphi = \widetilde{\varphi}(\lambda\varphi)$.

$\rhd$ Проверяется непосредственно. $\square$

Функции от операторов.

На множестве линейных операторов $L(V, V)$ определены операции сложения, умножения на константы и умножения. Предложения 1.9 и 1.10 вметсте с условием ассоциативности умножения (которая выполнена для композиции произвольных отображений) означают, что $L(V, V)$ имеет структуру *ассоциативной алгебры.* Поэтому $L(V, V)$ можем называть алгеброй линейных операторов на пространстве V.

Если $\varphi \in L(V, V)$, а $f \in \mathbb{F}[X]$ — некоторый многочлен, так что $f(x) = \sum\limits_{i=0}^{m} a_i x^i$, то положим $f(\varphi) = \sum\limits_{i=0}^{m} a_i \varphi^i$ (напомним, что $\varphi^0 = I_V$ — тождественное преобразование). Преобразование $f(\varphi)$ перестановочно с φ. Более общо, если f и g — два многочлена, то $f(\varphi)$ и $g(\varphi)$ — перестановочные преобразования.

Формально, при фиксированном $\varphi \in L(V, V)$ подстановка $f \mapsto f(\varphi)$ задает гомоморфизм алгебр $\mathbb{F}[X] \to L(V, V)$. Образ — множество многочленов от φ — является коммутативной подалгеброй в $L(V, V)$. Ниже мы говорим о ядре этого гомоморфизма.

Говорят, что многочлен $f \in \mathbb{F}[X]$ является *аннулирующим* для оператора $\varphi \in L(V, V)$ (или f аннулирует φ), если $f(\varphi) = 0$ (нулевой оператор). При $\dim V < n$ для данного $\varphi \in L(V, V)$ имеются ненулевые аннулирующие многочлены. Отложив доказательство этого факта, установим структуру множества многочленов, аннулирующих φ.

> **Определение.** Ненулевой многочлен $\mu \in \mathbb{F}[X]$ называется *минимальным* многочленом оператора $\varphi \in L(V, V)$, если μ аннулирует φ и имеет минимальную степень среди всех аннулирующих φ многочленов.

Предложение 1.11. *Пусть $\mu \in \mathbb{F}[X]$ — минимальный многочлен оператора $\varphi \in L(V, V)$ и $f \in \mathbb{F}[X]$. Тогда f аннулирует $\varphi \Leftrightarrow f \,\vdots\, \mu$ (делится в кольце $\mathbb{F}[X]$).*

$\rhd$ Разделим f на μ с остатком: $f = q\mu + r$, где $q, r \in \mathbb{F}[X]$, $\deg r < \deg \mu$. Тогда равенство остается верным после подстановки φ: $f(\varphi) = q(\varphi)\mu(\varphi) + r(\varphi)$. Поскольку $\mu(\varphi) = 0$, имеем $f(\varphi) = r(\varphi)$. Значит, $f(\varphi) = 0 \Leftrightarrow r(\varphi) = 0$. Условие $r(\varphi) = 0$ означает, что r аннулирует φ и его степень меньше степени минимального многочлена, т.е. $r = 0$ (нулевой многочлен). Таким образом, $f(\varphi) = 0 \Leftrightarrow r = 0 \Leftrightarrow f \,\vdots\, \mu$. $\square$

Следствие. *Для оператора $\varphi \in L(V, V)$ минимальный многочлен единственный, с точностью до умножение на ненулевую константу*

$\rhd$ Если μ и μ' — оба минимальные многочлены для φ, то, согласно предложению 1.11, $\mu' \vdots \mu$ и $\mu \vdots \mu'$. $\Box$

Многочлен являются частным случаем степенного ряда, или более общо, ряда Лорана. Можно определять функции от оператора φ через значение соответствующего операторного ряда $\sum a_i \varphi^i$, правда во мноих случаях для обеспечения сходимости такого ряда требуются дополнительные условия на φ (кроме того, сходимость можно понимать в разных смыслах). Так, для линейных операторов $\varphi : V \to V$, где V — пространство над $\mathbb{R}$ или $\mathbb{C}$, можно говорить об $\exp(\varphi)$ и т.д.

Примеры

Пусть $V = U_1 \bigoplus U_2$. Тогда для каждого вектора $\mathbf{a}$ имеется единственное разложение $\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_2$, где $\mathbf{a}_1$ — проекция $\mathbf{a}$ на U_1 вдоль U_2, $\mathbf{a}_2$ — проекция $\mathbf{a}$ на U_2 вдоль U_1.

Отображение $\varphi : V \to V$ такое, что $\varphi(\mathbf{a}) = \mathbf{a}_1$, называется *проектированием* на U_1 вдоль (или параллельно) U_2.

Отображение $\psi : V \to V$ такое, что $\varphi(\mathbf{a}) = \mathbf{a}_1 - \mathbf{a}_2$, называется *отражением* (или *симметрией*) относительно U_1 вдоль (или параллельно) U_2.

Легко проверить, что отражение и проектирование — линейные преобразования, причем отражение является изоморфизмом.

I. Пусть $\varphi : \mathbf{R}^3 \to \mathbf{R}^3$ — движение, для которого начало координат O — неподвижная точка (отождествляем точки с радиус-векторами), например, φ — поворот вокруг оси (прямой, проходящей через O).
Линейные биективные преобразования плоскости $\varphi : \mathbf{R}^2 \to \mathbf{R}^2$ — аффинные преобразовани, для которых O — неподвижная точка (см. ниже параграф §. 4.

II.1. Пусть $A \in \mathbf{M}_{m\times n}$ — фиксированная матрица. Определим $\varphi \in L(\mathbf{M}_{n\times p}, \mathbf{M}_{m\times p})$ правилом $\varphi(X) = AX$. Нетрудно проверить, что φ линейно (ниже увидим, что в некотором смысле к этому примеру можно свести любое линейное отображение конечномерных пространств).
Аналогично можно определить отображение домножение справа на фиксированную матрицу.

II.2. Транспонирование матриц $m \times n$ — пример изоморфизма $\mathbf{M}_{m\times n} \to \mathbf{M}_{n\times m}$.

III.1. Пусть $\mathbf{C}[a, b]$ — пространство всех непрерывных функций, определенных на отрезке $[a, b]$. Для различных отрезков $[a, b]$ пространства $\mathbf{C}[a, b]$ изоморфны. Например, изоморфизм между пространствами $\mathbf{C}[-2, 0]$, $\mathbf{C}[0, 1]$ определяется правилом $\varphi(f(x)) = f(2x - 2)$.

III.2. Пусть $V = \mathbf{C}^1(\mathbb{R})$. *Дифференцирование* $d : V \to V$ задается правилом $(d(f))(x) = f'(x)$. Нетрудно видеть, что d — линейное преобразование. Можно определить *линейный дифференциальный оператор* — многочлен от d.
В случае функций многих переменных (для $V = \mathbf{C}^1(\mathbb{R}^n)$) можно рассмотреть линейные операторы взятия частной производной $\dfrac{\partial}{\partial x_i}$ (эти операторы — не перестановочные, но становятся таковыми при оганичении на подпространство $\mathbf{C}^2(\mathbb{R}^n)$).
Линейное отображение $\mathbf{C}(\mathbb{R}) \to \mathbf{C}(\mathbb{R})$ *интегрирование* с переменным верхним пределом задается правилом $\varphi(f(x)) = \int\limits_0^x f(t)\,dt$.

III.3. Пусть $V = \mathbf{F}(\mathbb{N}) = \{(a_1, a_2, \ldots) \mid a_i \in \mathbb{F}\}$ — пространство числовых последовательностей. *Оператор сдвига* $\varphi : V \to V$ определяется равенством $\varphi(a_1, a_2, \ldots) = (a_2, a_3, \ldots)$. Определим *оператор первой разности* $\Delta = \varphi - I_V$, так что $\Delta(a_1, a_2, \ldots) = (a_2 - a_1, a_3 - a_2, \ldots)$.
Для $k \in \mathbb{N}$ степень Δ^k называют *оператором k-ой разности.*

IV. Пусть $\mathbb{F}$ — подполе некоторого поля $\mathbb{K}$, а $\varphi : \mathbb{K} \to \mathbb{K}$ — изоморфизм поля $\mathbb{K}$ на себя (автоморфизм), при котором все элементы поля $\mathbb{F}$ неподвижны. Тогда φ является линейным оператором $\mathbb{K} \to \mathbb{K}$, где $\mathbb{K}$ рассматривается как линейное пространство над $\mathbb{F}$.

§ 2. Матрица линейного отображения

В этом параграфе занимаемся только конечномерными векторными пространствами. Полагаем $\dim V = n$, $\dim \widetilde{V} = m$, $\dim \widehat{V} = p$.

Определение. Координатная запись линейного отображения

Определение. Пусть в пространствах V и $\widetilde{V}$ зафиксированы базисы $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ и $\mathrm{f} = (\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m)$ соответственно. *Матрицей линейного отображения* $\varphi \in L(V, \widetilde{V})$ в паре базисов e и f называется матрица $A \in \mathbf{M}_{m\times n}$, столбцы $a_{\bullet 1}, a_{\bullet 2}, \ldots, a_{\bullet n}$ которой — координатные столбцы соответственно векторов $\varphi(\mathbf{e}_1), \varphi(\mathbf{e}_2), \ldots, \varphi(\mathbf{e}_n)$ в базисе f.

Матрицу линейного преобразования $\varphi \in L(V, V)$ в паре совпадающих базисов e и e будем также называть короче: матрица φ в базисе e.

Компактно определение можно записать как $(\varphi(\mathbf{e}_1)\, \varphi(\mathbf{e}_2)\, \ldots\, \varphi(\mathbf{e}_n)) = \mathrm{f}A$. Тот факт, что A — матрица линейного отображения φ в паре базисов e и f, будем обозначать $\varphi \xrightarrow[\mathrm{e,f}]{} A$.

Предложение 2.1. *Пусть в пространствах V и $\widetilde{V}$ зафиксированы базисы e и f соответственно. Тогда отображение $\varphi \xrightarrow[\mathrm{e,f}]{} A$ — биекция между $L(V, \widetilde{V})$ и $\mathbf{M}_{m\times n}$.*

$\triangleright$ По определению матрица A несет в себе полную информацию об образах базисных векторов. Остается воспользоваться теоремой 1.1. $\square$

Выведем формулу координатной записи линейного отображения.

Теорема 2.1. *Пусть $\varphi \xrightarrow[\mathrm{e,f}]{} A$. Если $\mathbf{a} = \mathrm{e}X$, $\varphi(\mathbf{a}) = \mathrm{f}Y$, то*

$$\boxed{Y = AX}.$$

$\triangleright$ Так как $\mathbf{a} = \sum\limits_{i=1}^{n} x_i \mathbf{e}_i$, то $\varphi(\mathbf{a}) = \sum\limits_{i=1}^{n} x_i \varphi(\mathbf{e}_i)$. Заменим в этом равенстве векторы на их координатные столбцы в базисе f, получим: $Y = \sum\limits_{i=1}^{n} x_i a_{\bullet i}$, где, как обычно, $a_{\bullet i}$ обозначает i-й столбец матрицы A. Правая часть последнего равенства равна AX, что и требовалось уствновить. $\square$

Формула из предыдущей теоремы фактически эквивалентна определению матрицы линейного отображения. Более, точно, справедливо следующее

Следствие. *Пусть дано отображение $\varphi : V \to \widetilde{V}$ (априори не известно, что оно линейное) и матрица $A \in M_{m\times n}$. Пусть $\forall\ \mathbf{a} \in V$ координатные столбцы X и Y векторов $\mathbf{a} = \mathrm{e}X$ и $\varphi(\mathbf{a}) = \mathrm{f}Y$ связаны равенством $Y = AX$. Тогда $\varphi \in L(V, \widetilde{V})$, причем $\varphi \xrightarrow[\mathrm{e,f}]{} A$.*

$\triangleright$ Возьмем $\psi \in L(V, \widetilde{V})$ такое, что $\psi \xrightarrow[\mathrm{e,f}]{} A$ (такое ψ существует, согласно предложению 2.1). Тогда φ и ψ имеют одну и ту же координатную запись $Y = AX$, т.е. $\varphi = \psi$. $\square$

Матрица перехода и матрица преобразования

Переход от базиса $\mathrm{e} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ к $\mathrm{e}' = (\mathbf{e}'_1, \dots, \mathbf{e}'_n)$ формально не связан с линейными преобразованиями. Однако, согласно теореме 1.1, по паре базисов e и e′ можно определить единственное линейное отображение $\varphi : V \to V$ такое, что $\varphi(\mathbf{e}_i) = \mathbf{e}'_i$ для $i = 1, \dots, n$ (при этом φ является изоморфизмом — см. теорему 1.3). Отметим следующую связь между матрицей линейного преобразования и матрицей перехода.

Предложение 2.2. *Пусть* $\dim V = n < \infty$, e *и* e′ *— базисы в* V. *Пусть изоморфизм* $\varphi : V \to V$ *таков, что* $\varphi(\mathbf{e}_i) = \mathbf{e}'_i$ *для* $i = 1, \dots, n$; $\varphi \xrightarrow[\mathrm{e,e}]{} A$. *Тогда* A *— матрица перехода от* e *к* e′.

▷ Достаточно сопоставить определения матрицы линейного отображения и матрицы перехода. □

Связь между операциями над отображениями и матрицами

Предложение 2.1 может быть усилено следующим образом.

Теорема 2.2. *Пусть* e *и* f *— базисы пространств* V *и* $\widetilde{V}$ *соответственно. Тогда отображение* $\varphi \xrightarrow[\mathrm{e,f}]{} A$ *является изоморфизмом линейных пространств* $L(V, \widetilde{V})$ *и* $\mathbf{M}_{m\times n}$.

▷ Непосредственная проверка. □

Следствие. $\dim L(V, \widetilde{V}) = mn$.

Аналогично, для векторного пространства V над $\mathbb{C}$ пространство $\overline{L}(V, \widetilde{V})$ (со вторым способом определения отображения на константу) изоморфно $\mathbf{M}_{m\times n}(\mathbb{C})$; изоморфизм устанавливается как $\varphi \mapsto \overline{A}$.

Предложение 2.3. *Пусть* $\varphi \in L(V, \widetilde{V})$, $\widetilde{\varphi} \in L(\widetilde{V}, \widehat{V})$, e, f, g *— базисы пространств* V, $\widetilde{V}$, $\widehat{V}$ *соответственно. Пусть* $\varphi \xrightarrow[\mathrm{e,f}]{} A$, $\widetilde{\varphi} \xrightarrow[\mathrm{f,g}]{} \widetilde{A}$. *Тогда* $\widetilde{\varphi}\varphi \xrightarrow[\mathrm{e,g}]{} \widetilde{A}A$.

▷ Пусть $\mathbf{a} \in V$ — произвольный вектор, $\mathbf{a} = \mathrm{e}X$, $\varphi(\mathbf{a}) = \mathrm{f}Y$, $\widetilde{\varphi}(\varphi(\mathbf{a})) = \mathrm{g}Z$. Тогда дважды пользуясь теоремой 2.1, имеем $Y = AX$, $Z = \widetilde{A}Y$, откуда $Z = \widetilde{A}(AX) = (\widetilde{A}A)X$. Отсюда следует требуемое (см. следствие из теоремы 2.1). □

Следствие 1. *Пусть* $\varphi \xrightarrow[\mathrm{e,f}]{} A$. *Тогда* φ *— изоморфизм* $\Leftrightarrow$ A *обратима. Если* φ *— изоморфизм, то* $\varphi^{-1} \xrightarrow[\mathrm{f,e}]{} A^{-1}$.

▷ Следует из того, что $I_V \xrightarrow[\mathrm{e,e}]{} E$. □

Следствие 2. *Пусть* $\varphi \in L(V, V)$ *и* $\varphi \xrightarrow[\mathrm{e,e}]{} A$. *Тогда для любого многочлена* p *имеем* $p(\varphi) \xrightarrow[\mathrm{e,e}]{} p(A)$.

Следствие 3. *Пусть* e *— базис пространства* V. *Тогда отображение* $\varphi \xrightarrow[\mathrm{e,e}]{} A$ *является изоморфизмом алгебр* $L(V, V)$ *и* $\mathbf{M}_{n\times n}(\mathbb{F})$.
В частности группа (по умножению) изоморфизмов $V \to V$ *изоморфна* $GL_n(\mathbb{F})$ *(группе невырожденных матриц).*

Установленное соответствие между операциями над линейными отображенями и операциями над матрицами может быть применено в обе стороны: некоторые задачи об отображениях могут быть сведены к вопросам о матрицах, и наоборот (см., например, упражнения в конце следующего параграфа).

Многие понятие могут быть определены одновременно для линейных операторов и для матриц. Напрмер: говорят о *ранге отображения*, имея в виду ранг матрицы этого отображения (ниже мы увидим, что этот ранг не зависит от выбора базисов); *невырожденным* называют оператор $\varphi \in L(V, V)$, матрица которого невырожденная (это условие эквивалентно тому, что φ — изоморфизм); с другой стороны, можем говорить об аннулирующих и минмальных многочленах для матриц из $\mathbf{M}_{n\times n}(\mathbb{F})$.

Изменение матрицы при замене базиса

Теорема 2.3. *Пусть в V выбраны базисы* e *и* e', *связанные матрицей перехода S:* $\mathrm{e}' = \mathrm{e}S$; *в* $\widetilde{V}$ *выбраны базисы* f *и* f', *связанные матрицей перехода R:* $\mathrm{f}' = \mathrm{f}R$. *Пусть* $\varphi \in L(V, \widetilde{V})$ *таково, что* $\varphi \xrightarrow[\mathrm{e,f}]{} A$ *и* $\varphi \xrightarrow[\mathrm{e',f'}]{} A'$. *Тогда*

$$\boxed{A' = R^{-1}AS}.$$

В частности, если $V = \widetilde{V}$, $\mathrm{e} = \mathrm{f}$ *и* $\mathrm{e}' = \mathrm{f}'$, *то*

$$\boxed{A' = S^{-1}AS}.$$

$\triangleright$ Пусть $\mathbf{a} \in V$ — произвольный вектор. Пусть $\mathbf{a} = \mathrm{e}X = \mathrm{e}'X'$, $\varphi(\mathbf{a}) = \mathrm{f}Y = \mathrm{f}'Y'$. Тогда по теореме 2.1 имеем $Y = AX$, $Y' = A'X'$ и (по теореме 2.4, глава 1) $X = SX'$, $Y = RY'$. Отсюда $RY' = ASX' \Rightarrow Y' = R^{-1}ASX'$ или $Y' = (R^{-1}AS)X'$. Получаем требуемое: $R^{-1}AS = A'$ (см. следствие из теоремы 2.1). $\square$

Следствие. *Ранг матрицы линейного отображения* $\varphi \in L(V, \widetilde{V})$ *не зависит от выбора базисов в пространствах* V *и* $\widetilde{V}$.

$\triangleright$ Следует из того, что A' получается из A домножением слева и справа на невырожденные матрицы. $\square$

Инвариантная характеризация ранга матрицы линейного отображения дается ниже в следствии из теоремы 3.1.

Заметим также, что матрицу $S^{-1}AS$ иногда называют *подобной* или *сопряженной* матрице A. Нетрудно проверить, что отношение подобия — это отношение эквивалентности на множестве $\mathbf{M}_{n\times n}(\mathbb{F})$. Это согласуется с тем фактом, что подобные матрицы соответствуют одному и тому же оператору, в разных базисах.

Примеры

Пусть $V = U_1 \bigoplus U_2$. Введем базис $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ в V, *согласованный* $U_1 \bigoplus U_2$, так, что $\mathbf{e}_1, \ldots, \mathbf{e}_k$ — базис в U_1, а $\mathbf{e}_{k+1}, \ldots, \mathbf{e}_n$ — базис в U_2.
Пусть $\varphi : V \to V$ — проектирование на U_1 вдоль U_2. Тогда (по определению матрицы линейного отображения): $\varphi \xrightarrow[\mathrm{e,e}]{} \begin{pmatrix} E_k & O \\ O & O \end{pmatrix}$.
Пусть $\psi : V \to V$ — отражение относительно U_1 вдоль U_2. Тогда $\psi \xrightarrow[\mathrm{e,e}]{} \begin{pmatrix} E_k & O \\ O & -E_{n-k} \end{pmatrix}$.

> III.1. Пусть $V = \mathbf{P}_n = \langle 1, x, x^2, \ldots, x^n \rangle$ и $\mathrm{e} = (1, x, x^2, \ldots, x^n)$ — стандартный базис в V. Оператор дифференцирования $d : V \to V$ имеет в базисе e матрицу $\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 2 & \ldots & 0 \\ & & \ldots & & \\ 0 & 0 & 0 & \ldots & n \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$
> (поскольку $(x^{k+1})' = (k+1)x^k$, имеем $d(e_{i+1}) = (i+1)e_i$).

§ 3. Образ и ядро

Образ

В предложении 1.4 мы видели, что при линейном отображении $\varphi \in L(V, \widetilde{V})$ образ $\varphi(U)$ подпространства $U \leqslant V$ является подпространством в $\widetilde{V}$. Для образа отображения φ наряду с обозначением $\varphi(V)$ используют обозначение $\operatorname{Im}\varphi$. Очевидно, $\varphi \in L(V, \widetilde{V})$ является сюръективным $\Leftrightarrow$ $\widetilde{V} = \operatorname{Im}\varphi$.

Теорема 3.1 (координатное описание образа)**.** *Пусть* $\varphi \in L(V, \widetilde{V})$, e, f *— базисы в* V *и* $\widetilde{V}$. *Пусть* $\varphi \xrightarrow[\mathrm{e,f}]{} A$. *Для вектора* $\mathbf{b} \in \widetilde{V}$, $\mathbf{b} = \mathrm{f}Y$ *выполнено:* $\mathbf{b} \in \operatorname{Im}\varphi \Leftrightarrow Y \in \langle a_{\bullet 1}, a_{\bullet 2}, \ldots, a_{\bullet n}\rangle$.

▷ Следует из сопоставления определения матрицы линейного отображения и предложения 1.4. □

Иначе говоря, теорема утверждает, что в терминах координатных столбцов $\operatorname{Im}\varphi$ задается как линейная оболочка столбцов матрицы линейного отображения. Это, в частности, дает еще одно объяснение того факта, что $\operatorname{rg} A$ не зависит от выбора базисов в пространствах V и $\widetilde{V}$.

Следствие. *В условиях теоремы* $\boxed{\dim \operatorname{Im}\varphi = \operatorname{rg} A}$.

Ядро

Покажем вначале, что при линейном отображении полный прообраз подпространства является подпространством.

Предложение 3.1. *Пусть* $\varphi \in L(V, \widetilde{V})$ *и* $\widetilde{U} \leqslant \widetilde{V}$. *Тогда* $\{\mathbf{a} \in V \mid \varphi(\mathbf{a}) \in \widetilde{U}\} \leqslant V$.

▷ Положим $U = \{\mathbf{a} \in V \mid \varphi(\mathbf{a}) \in \widetilde{U}\}$ и проверим для U свойства П1 и П2.
Пусть $\mathbf{a}, \mathbf{b}$ — произвольные векторы из U. Тогда $\varphi(\mathbf{a}) \in \widetilde{U}$, $\varphi(\mathbf{b}) \in \widetilde{U}$. Из свойства П1 для $\widetilde{U}$ получаем, что $\varphi(\mathbf{a}) + \varphi(\mathbf{b}) \in \widetilde{U}$, то есть $\varphi(\mathbf{a} + \mathbf{b}) \in \widetilde{U}$. Но последнее включение означает, что $\mathbf{a} + \mathbf{b} \in U$.
П2 проверяется аналогично. □

> **Определение.** *Ядром* линейного отображения $\varphi \in L(V, \widetilde{V})$ называется подмножество $\{\mathbf{a} \in V \mid \varphi(\mathbf{a}) = \mathbf{0}\}$.

Обозначение для ядра — $\operatorname{Ker}\varphi$. Опеределение можно переформулировать так: $\operatorname{Ker}\varphi$ — это полный прообраз нулевого подпространства, поэтому из предложения 3.1 вытекает

Следствие. *Если* $\varphi \in L(V, \widetilde{V})$, *то* $\operatorname{Ker}\varphi \leqslant V$.

Для выяснения, является ли φ инъективным, полезен следующий критерий.

Предложение 3.2. $\varphi \in L(V, \widetilde{V})$ *инъективно* $\Leftrightarrow \operatorname{Ker}\varphi = O$.

▷ Следующее доказательство похоже на доказательство пункта 2 теоремы 1.1.
$\boxed{\Rightarrow}$ Пусть $\mathbf{a} \in \operatorname{Ker}\varphi$, то есть $\varphi(\mathbf{a}) = \mathbf{0}$. Но также $\varphi(\mathbf{0}) = \mathbf{0}$, и поскольку φ инъективно, имеем $\mathbf{a} = \mathbf{0}$.
$\boxed{\Rightarrow}$ Пусть $\varphi(\mathbf{a}) = \varphi(\mathbf{b})$. Тогда $\varphi(\mathbf{a} - \mathbf{b}) = \mathbf{0}$. С учетом $\operatorname{Ker}\varphi = O$ имеем $\mathbf{a} - \mathbf{b} = \mathbf{0}$, то есть $\mathbf{a} = \mathbf{b}$. Тем самым, $\varphi(\mathbf{a}) = \varphi(\mathbf{b}) \Rightarrow \mathbf{a} = \mathbf{b}$, и инъективность доказана. □

Теорема 3.2 (координатное описание ядра)**.** *Пусть* $\varphi \in L(V, \widetilde{V})$, e, f *— базисы в* V *и* $\widetilde{V}$. *Пусть* $\varphi \xrightarrow[\mathrm{e,f}]{} A$. *Для вектора* $\mathbf{a} \in \widetilde{V}$, $\mathbf{a} = \mathrm{f}X$ *выполнено:* $\mathbf{a} \in \operatorname{Ker}\varphi \Leftrightarrow AX = O$.

▷ Достаточно сопоставить определение ядра с формулой $Y = AX$ (см. теорему 2.1). □

Иначе говоря, теорема утверждает, что в терминах координатных столбцов $\operatorname{Ker}\varphi$ задается как общее решение $\operatorname{Sol}(AX = O)$ однородной системы линейных уравнений с матрицей коэффициентов A.

Следствие. *В условиях теоремы* $\boxed{\dim \operatorname{Ker}\varphi = n - \operatorname{rg} A}$.

Связь между размерностями ядра и образа

Установим формулу, связывающую размерности ядра и образа.

Теорема 3.3. *Пусть* $\varphi \in L(V, \widetilde{V})$ *и* $\dim V = n < \infty$. *Тогда*

$$\boxed{\dim \operatorname{Ker} \varphi + \dim \operatorname{Im} \varphi = n}. \tag{2.3}$$

$\triangleright$ Можно считать, что $\dim \widetilde{V} < \infty$ (иначе заменим $\widetilde{V}$ на любое конечномерное подпространство, содержащее $\operatorname{Im} \varphi$, замена не влияет на Im и Ker). Рассмотрим матрицу A отображения φ в некоторых базисах. По следствию из теоремы 3.1 имеем $\dim \operatorname{Im} \varphi = \operatorname{rg} A$, а по следствию из теоремы 3.2 имеем $\dim \operatorname{Ker} \varphi = n - \operatorname{rg} A$. Отсюда вытекет нужная формула размерностей. $\square$

Следствие. *Пусть* $\varphi \in L(V, \widetilde{V})$, $\dim V < \infty$ *и* $U \leqslant V$. *Тогда*

$$\dim U \leqslant \dim \varphi(U) + \dim \operatorname{Ker} \varphi, \tag{2.4}$$

причем равенство достигается $\Leftrightarrow U \geqslant \operatorname{Ker} \varphi$.

$\triangleright$ Ограничение φ на U (т.е. $\varphi|_U \in L(U, \widetilde{V})$), обозначим ψ. Применив (2.3) к ψ, имеем $\dim \varphi(U) + \dim \operatorname{Ker} \psi = \dim U$, при этом $\operatorname{Ker} \psi = (\operatorname{Ker} \varphi) \cap U \leqslant \operatorname{Ker} \varphi$, откуда следуют нужные утверждения. $\square$

Отметим, что (2.3) находится в согласии с эквивалентностью условий 1), 2), 3) в теореме 1.3.

Мы доказали теорему 3.3, перейдя с «языка отображений» на «матричный язык». Более «концептуально» — получить формулу размерностей (2.3) как следствие *теоремы о гомоморфизме* $V/\operatorname{Ker} \varphi \cong \operatorname{Im} \varphi$ (это стандартная теорема курса алгебры для многих алгебраических систем), которая вытекает из естественной биекции $\mathbf{a} + \operatorname{Ker} \varphi \mapsto \varphi(\mathbf{a})$. Как следствие теоремы о гомоморфизме можно получить биекцию между подпространствами в V, содержашими $\operatorname{Ker} \varphi$, и подпространствами в $\operatorname{Im} \varphi$ (что соответсвует случаю равенства в (2.4)).

В следующем упражнении намечен еще один путь доказательства формулы из теоремы 3.3 без привлечения матрицы линейного отображения.

Упражнение. Пусть $\varphi \in L(V, \widetilde{V})$ и $\dim V = n < \infty$. Пусть $\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n$ — базис в $\operatorname{Ker} \varphi$. Дополним его до базиса $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ пространства V. Докажите, что тогда $\operatorname{Im} \varphi = \langle \varphi(\mathbf{e}_1), \ldots, \varphi(\mathbf{e}_r) \rangle$, причем $\varphi(\mathbf{e}_1), \ldots, \varphi(\mathbf{e}_r)$ — линейно независимая система.

Предлагаем еще одно упражнение о простейшем виде матрицы линенйного отображения.

Упражнение. Пусть $\varphi \in L(V, \widetilde{V})$. Докажите, что в пространствах V и $\widetilde{V}$ можно выбрать базисы e и f так, что

$$\varphi \underset{\mathrm{e,f}}{\longrightarrow} \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}.$$

(Утверждение последнего упражнения не следует путать со случаем выбора одного и того же базиса e = f для матрицы линейного преобразования, как обычно будет происходить в главе 3.)

Примеры

Пусть $V = U_1 \bigoplus U_2$. Пусть $\varphi : V \to V$ — проектирование на U_1 вдоль U_2. Тогда $\operatorname{Im} \varphi = U_1$, $\operatorname{Ker} \varphi = U_2$.

III.1. Пусть $V = \mathbf{C}^\infty(\mathbb{R})$. Общее решение U линейного дифференциального уравнения $x^{(n)} + a_{n-1}(t) x^{(n-1)} + \ldots + a_1(t) x' + a_0(t) x = 0$ (где $a_i(t)$ — непрерывные функции $\mathbb{R} \to \mathbb{R}$) является ядром линейного дифференциального оператора $d^n + a_{n-1} d^{n-1} + \ldots + a_1 d + a_0 d^0$. Теорема существования и единственности из курса дифференциальных уравнений устанавливает изоморфизм между решением $x \in U$ и столбцом *начальных условий* $\begin{pmatrix} x(t_0) \\ x'(t_0) \\ \ldots \\ x^{(n-1)}(t_0) \end{pmatrix}$, тем самым, $\dim U = n$.

> III.2. Пусть $V = \mathbf{F}(\mathbb{N}) = \{(a_1, a_2, \dots) \,|\, a_i \in \mathbb{F}\}$ — пространство числовых последовательностей. Фибоначчиевы последовательности могут быть заданы как ядро линейного оператора $\varphi^2 - \varphi - \varphi^0$, где $\varphi : V \to V$ — оператор сдвига.

Ядро и образ произведения отображений

Начнем с двух почти очевидных предложений.

Предложение 3.3. *Пусть* $\varphi \in L(V, \widetilde{V})$, $\psi \in L(\widetilde{V}, \widehat{V})$. *Тогда*

$$\operatorname{Im}(\psi\varphi) \leqslant \operatorname{Im}\psi.$$

$\triangleright$ Это частный случай включения $\psi(U) \leqslant \psi(V)$ для $U = \varphi(V)$. $\square$

Предложение 3.4. *Пусть* $\varphi \in L(V, \widetilde{V})$, $\psi \in L(\widetilde{V}, \widehat{V})$. *Тогда*

$$\operatorname{Ker}(\psi\varphi) \geqslant \operatorname{Ker}\varphi.$$

$\triangleright$ Если $\mathbf{a} \in \operatorname{Ker}\varphi$, то $\varphi(\mathbf{a}) = \mathbf{0} \Leftrightarrow (\psi\varphi)(\mathbf{a}) = \psi(\varphi(\mathbf{a})) = \mathbf{0} \Leftrightarrow \mathbf{a} \in \operatorname{Ker}(\psi\varphi)$. $\square$

Упражнение. Переформулируйте и докажите теоремы об оценке ранга произведения матриц $\operatorname{rg}(AB) \leqslant \operatorname{rg} A$, $\operatorname{rg}(AB) \leqslant \operatorname{rg} B$ в терминах линейных отображений.

Зафиксируем еще несколько фактов, которые будем использовать далее.

Предложение 3.5. *Пусть* $\varphi \in L(V, \widetilde{V})$, $\psi \in L(\widetilde{V}, \widehat{V})$, $\dim V < \infty$. *Тогда*

$$\dim\operatorname{Ker}(\psi\varphi) \leqslant \dim\operatorname{Ker}\varphi + \dim\operatorname{Ker}\psi. \tag{2.5}$$

$\triangleright$ Пусть $U = \operatorname{Ker}(\psi\varphi)$. Тогда $\varphi(U) \leqslant \operatorname{Ker}\psi$. Согласно следствию из теоремы 3.3, имеем $\dim U = \dim\varphi(U) + \dim\operatorname{Ker}\varphi \leqslant \dim\operatorname{Ker}\psi + \dim\operatorname{Ker}\varphi$, что и требовалось. $\square$

Отметим, что равенство в (2.5) достигается $\Leftrightarrow \varphi(U) = \operatorname{Ker}\psi$, или, поскольку U является полным прообразом $\operatorname{Ker}\psi$ относительно φ, эквивалентно, $\operatorname{Ker}\psi \leqslant \operatorname{Im}\varphi$.

Упражнение. а) Для матриц $A \in \mathbf{M}_{m\times n}$, $B \in \mathbf{M}_{n\times k}$ докажите *неравенство Сильвестра*

$$\operatorname{rg}(AB) \geqslant \operatorname{rg} A + \operatorname{rg} B - n.$$

б) Каков максимальный ранг матрицы $A \in \mathbf{M}_{n\times n}$, если $A^2 = O$?

в)* Для матриц $A \in \mathbf{M}_{m\times n}$, $B \in \mathbf{M}_{n\times k}$, $C \in \mathbf{M}_{k\times l}$ докажите *неравенство Фробениуса* (обобщающее неравенство Сильвестра)

$$\operatorname{rg}(ABC) + \operatorname{rg} B \geqslant \operatorname{rg}(AB) + \operatorname{rg}(BC).$$

В качестве следствия предложения 3.4 имеем

Предложение 3.6. *Пусть* $\varphi \in L(V, V)$, $\psi \in L(V, V)$, *причем* $\varphi\psi = \psi\varphi$. *Тогда*

$$\operatorname{Ker}(\psi\varphi) \geqslant \operatorname{Ker}\varphi + \operatorname{Ker}\psi. \tag{2.6}$$

$\triangleright$ Так как $\operatorname{Ker}(\psi\varphi) \geqslant \operatorname{Ker}\varphi$ и $\operatorname{Ker}(\psi\varphi) = \operatorname{Ker}(\varphi\psi) \geqslant \operatorname{Ker}\psi$, то $\operatorname{Ker}(\psi\varphi) \geqslant \langle\operatorname{Ker}\varphi, \operatorname{Ker}\psi\rangle = \operatorname{Ker}\varphi + \operatorname{Ker}\psi$. $\square$

Следствие. *Пусть* $\varphi \in L(V, V)$, $\psi \in L(V, V)$, $\dim V < \infty$, *причем* $\varphi\psi = \psi\varphi$. *Если* $\operatorname{Ker}\varphi \cap \operatorname{Ker}\psi = O$, *то* $\operatorname{Ker}(\psi\varphi) = \operatorname{Ker}\varphi \oplus \operatorname{Ker}\psi$.

$\triangleright$ Из (2.6), при условии $\operatorname{Ker}\varphi \cap \operatorname{Ker}\psi = O$, следует, что $\operatorname{Ker}(\psi\varphi) \geqslant \operatorname{Ker}\varphi \oplus \operatorname{Ker}\psi$ и $\dim\operatorname{Ker}(\psi\varphi) \geqslant \dim\operatorname{Ker}\varphi + \dim\operatorname{Ker}\psi$. Тогда из (2.5) вытекет требуемое. $\square$

Предложение 3.7. *Пусть* $\varphi \in L(V, V)$, *а* $f_1, \ldots, f_k \in \mathbb{F}[X]$ *— попарно взаимно простые многочлены. Тогда* $\sum\limits_{i=1}^{k} \operatorname{Ker}(f_i(\varphi))$ *— прямая сумма.*
Если, кроме того, $\dim V < \infty$, *то* $\operatorname{Ker}(f_1(\varphi)) \oplus \ldots \oplus \operatorname{Ker}(f_k(\varphi)) = \operatorname{Ker}(f_1(\varphi) \ldots f_k(\varphi))$.

$\triangleright$ Докажем в случае $k = 2$. Этого будет достаточно, так как случай произвольного k можно разобрать, применив утверждение для $k = 2$ к многочленам f_1 и $f_2 \ldots f_k$, далее к f_2 и $f_3 \ldots f_k$ и т.д.
Так как f_1 и f_2 взаимно просты, существуют многочлены $g_1, g_2 \in \mathbb{F}[X]$ такие, что $g_1 f_1 + g_2 f_2 = 1$. Тогда из этого равенсва многочленов имеем $g_1(\varphi) f_1(\varphi) + g_2(\varphi) f_2(\varphi) = I_V$.
Применяя это операторное равенство для $\mathbf{a} \in \operatorname{Ker}(f_1(\varphi)) \cap \operatorname{Ker}(f_2(\varphi))$, получаем $\mathbf{0} = \mathbf{a}$, значит $\operatorname{Ker}(f_1(\varphi)) \cap \operatorname{Ker}(f_2(\varphi)) = O$, тем самым утверждение про прямую сумму доказано.
Второе утверждение сразу получаем в силу следствия из предложения 3.6. $\square$

§ 4. Аффинные преобразования (на плоскости).

Определение и основные свойства

S — аффинное пространство.
далее рассматриваем двумерный случай: $S = \mathcal{P}$ (плоскость)
V — двумерное векторное пространство (радиус-векторы на плоскости).

> **Определение.** Преобразование плоскости $f : \mathcal{P} \to \mathcal{P}$ называется линейным (линейно-аффинным) преобразованием, если
> $\exists\ \varphi \in L(V, V)$ такое, что $\forall\ M, N \in \mathcal{P}$ выполнено
> $$\overrightarrow{f(M)f(N)} = \varphi(\overrightarrow{MN}).$$

Соответствующее линейное отображение $\varphi \in L(V, V)$ обозначим $\widetilde{f}$; иногда $\widetilde{f}$ называют *дифференциалом* отображения f.

> **Определение.** Линейное преобразование $f : \mathcal{P} \to \mathcal{P}$ называется *аффинным*, если оно биективно.

Предложение 4.1. .
Пусть $f : \mathcal{P} \to \mathcal{P}$ *— линейное преобразование. Тогда*
f *аффинно* $\Leftrightarrow$ $\widetilde{f}$ *биективно.*

Предложение 4.2. .
Если $f : \mathcal{P} \to \mathcal{P}$ *— аффинное преобразование, то* f^{-1} *также аффинное преобразование, причем* $\widetilde{f^{-1}} = \widetilde{f}^{-1}$.

Предложение 4.3. .
1) Если $f : \mathcal{P} \to \mathcal{P}$ *и* $g : \mathcal{P} \to \mathcal{P}$ *— линейные преобразования, то* gf *также линейное, причем* $\widetilde{gf} = \widetilde{g}\widetilde{f}$.
2) Если $f : \mathcal{P} \to \mathcal{P}$ *и* $g : \mathcal{P} \to \mathcal{P}$ *— аффинные преобразования, то* gf *также аффинное.*

Как видим, аффинные преобразования образуют группу относительно композиции.

Координатная запись линейного преобразования.

Зафиксируем ДСК (O, e).

Пусть в определении линейного преобразования $f : \mathcal{P} \to \mathcal{P}$ дифференциал $\widetilde{f} : V \to V$ имеет матрицу A, а $f(O)$ имеет координатный столбец C.

Предложение 4.4. *Пусть X и Y — координатные столбцы точки M ее образа $f(M)$. Тогда*

$$\boxed{Y = AX + C}. \tag{2.7}$$

Видим, что C — это координатный столбец для $f(O)$. Координатные записи для f $(Y = AX+C)$ и $\widetilde{f}$ $(Y = AX)$ отличаются «сдвигом на C». Иначе, f может быть представлено в виде композиции преобразования с неподвижной точкой O и сдвига (параллельного переноса).

Предложение 4.5. *(критерий аффинности) Пусть дифференциал $\widetilde{f}$ линейного преобразования $f : \mathcal{P} \to \mathcal{P}$ имеет в базисе* e *матрицу A. Тогда*
f аффинно (то есть биективно) $\Leftrightarrow |A| \neq 0$.

Теорема 4.1. *Пусть даны точки $A, B, C, K, L, M \in \mathcal{P}$, причем точки A, B и C не лежат на одной прямой.*
Тогда существует единственное линейное преобразование $f : \mathcal{P} \to \mathcal{P}$ такое, что $f(A) = K$, $f(B) = L$, $f(C) = M$; при этом f аффинно $\Leftrightarrow$ K, L, M не лежат на одной прямой.

Теорема 4.2 (связь аффинного преобразования с заменой координат). *Пусть $f : \mathcal{P} \to \mathcal{P}$ — аффинное преобразование. Тогда если M имеет координатный столбец X в ДСК $(O, \mathbf{e_1}, \mathbf{e_2})$, то $f(M)$ имеет тот же координатный столбец X в ДСК $(f(O), \widetilde{f}(\mathbf{e_1}), \widetilde{f}(\mathbf{e_2}))$.*

Примеры

Пусть l — прямая. Выберем ПДСК $(O, \mathbf{e_1}, \mathbf{e_2})$ так, что $O \in l$, $\mathbf{e_1} \parallel l$.

$$\begin{cases} y_1 = x_1 \\ y_2 = \lambda x_2. \end{cases}$$

— при $\lambda = 0$ (ортогональным) проектированием на прямую l;
— при $\lambda = 1$ — тождественное преобразование;
— при $\lambda = -1$ — *осевая симметрия* (или *отражение*) относительно прямой l;
— при $\lambda > 0$ — *сжатие* к прямой l с коэффициентом λ (слово "сжатие"употребляется и в случае $\lambda > 1$, а иногда оно заменяется на слово "растяжение");
— при $\lambda < 0$ — композиция симметрии относительно l и сжатия к l.

Пусть M — некоторая точка. Выберем ДСК $(O, \mathbf{e_1}, \mathbf{e_2})$ так, что $O = M$.

$$\begin{cases} y_1 = \lambda x_1 \\ y_2 = \lambda x_2. \end{cases}$$

— при $\lambda \neq 0$ — гомотетия с коэффициентом λ.

Пусть $\mathbf{c}$ — некоторый вектор, имеющий координаты $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$. в некоторой системе координат $(O, \mathbf{e_1}, \mathbf{e_2})$.

$$\begin{cases} y_1 = x_1 + c_1 \\ y_2 = x_2 + c_2 \end{cases}$$

— параллельный перенос.

Рассмотрим преобразование $f : \mathcal{P} \to \mathcal{P}$ поворота на угол φ (против часовой стрелки вокруг точки M). Выберем прямоугольную систему координат $(O, \mathbf{e_1}, \mathbf{e_2})$ так, что $O = M$.

$$\begin{cases} y_1 = \cos\varphi\, x_1 - \sin\varphi\, x_2 \\ y_2 = \sin\varphi\, x_1 + \cos\varphi\, x_2. \end{cases}$$

Из указанных примеров композицией можно получать новые (на самом деле все) аффинные преобразования. Известно, что всякое афинное преобразование может быть разложено в виде произведения движения (=композиция поворота или отражения и сдвига) и сжатий к перпендикулярным осям.

Геометрические свойства

Предложение 4.6. *Пусть $f : \mathcal{P} \to \mathcal{P}$ — аффинное преобразование. Тогда*
1) образ прямой — прямая, образ отрезка — отрезок;
2) параллельные прямые переходят в параллельные прямые;
3) отношение длин параллельных отрезков сохраняется;
4) образ центрально-симметричной фигуры — центрально-симметричная фигура.

Теорема 4.3. *(изменение площадей) Пусть дифференциал $\widetilde{f}$ линейного преобразования $f : \mathcal{P} \to \mathcal{P}$ имеет в базисе* e *матрицу A. Тогда*

$$\boxed{S(f(\Pi)) = |\det A| \cdot S(\Pi)}.$$

Таким образом, модуль $\det A$ отвечает за изменение площадей, а знак $\det A$ — за сохранение или изменение ориентации. Это согласуется с предложением о критерии аффинности (случай $\det A = 0$ соответсвует вырождению образа парлеллограмма).

Теорема 4.4. *Образом алгебраической кривой порядка n при аффинном преобразовании является алгебраическая кривая порядка n.*

Глава 3

Структура линейного преобразования.

В этой главе V — векторное пространство над полем $\mathbb{F}$, а φ — фиксированное линейное преобразование $V \to V$.

§ 1. Инвариантные подпространства

Важную информацию о линейном преобразовании $\varphi \in L(V,V)$ можно узнать, изучая инвариантные подпространства (то есть инвариантные относительно φ подмножества V, являющиеся подпространствами). Напомним, что $U \subset V$ называется инвариантным относительно φ, если $\varphi(U) \subset U$, т.е. если $\forall\ \mathbf{a} \in U$ выполнено $\varphi(\mathbf{a}) \in U$. Всякий раз, когда имеется инвариантное относительно $\varphi \in L(V,V)$ подпространство $U \leqslant V$, можем говорить об ограничении (или сужении) $\varphi|_U \in L(U,U)$. Кроме того, условие инвариантности U относительно φ дает возможность корректно определить ***фактор-оператор*** $\breve{\varphi} \in L(V/U, V/U)$ по естественному правилу $\breve{\varphi}(\mathbf{a}+U) = \varphi(\mathbf{a}) + U$.

Предложение 1.1. *Пусть $U \leqslant V$, $U = \langle \mathbf{a}_1, \ldots, \mathbf{a}_k \rangle$. Тогда U инвариантно относительно $\varphi \in L(V,V) \Leftrightarrow \varphi(\mathbf{a}_1), \ldots, \varphi(\mathbf{a}_k) \in U$.*

$\triangleright$ Следует из предложения 1.4 главы 2. $\square$

Предложение 1.2. *Пусть подпространства $U_i \leqslant V$, $i = 1, \ldots, k$, инвариантны относительно $\varphi \in L(V,V)$. Тогда $\sum\limits_{i=1}^{k} U_i$ и $\bigcap\limits_{i=1}^{k} U_i$ тоже инвариантны относительно φ.*

$\triangleright$ $\square$

Отметим, что предыдущее предложение верно и для произвольных подмножеств $U_1, U_2, \ldots, U_k \subset V$.

Предложение 1.3. *Пусть $\varphi \in L(V,V)$, $\psi \in L(V,V)$ таковы, что $\varphi\psi = \psi\varphi$. Тогда $\operatorname{Ker}\psi$, $\operatorname{Im}\psi$ инвариантны относительно φ.*

$\triangleright$ 1). Пусть $\mathbf{a} \in \operatorname{Ker}\psi$, так что $\psi(\mathbf{a}) = \mathbf{0}$. Тогда $\psi(\varphi(\mathbf{a})) = \varphi(\psi(\mathbf{a})) = \varphi(\mathbf{0}) = \mathbf{0}$, откуда $\varphi(\mathbf{a}) \in \operatorname{Ker}\psi$.
2). Пусть $\mathbf{b} \in \operatorname{Im}\psi$, так что $\mathbf{b} = \psi(\mathbf{a})$ для некоторого вектора $\mathbf{a}$. Тогда $\varphi(\mathbf{b}) = \varphi(\psi(\mathbf{a})) = \psi(\varphi(\mathbf{a}))$, откуда $\varphi(\mathbf{b}) \in \operatorname{Im}\psi$. $\square$

Следствие. *Пусть $\varphi \in L(V,V)$ и $f \in \mathbb{F}[X]$. Тогда $\operatorname{Ker} f(\varphi)$, $\operatorname{Im} f(\varphi)$ инвариантны относительно φ.*

Предложение 1.4. *Пусть $\varphi \in L(V,V)$ и $\lambda_0 \in \mathbb{F}$. Пусть подпространство $U \leqslant V$ таково, что $U \geqslant \operatorname{Im}(\varphi - \lambda_0)$. Тогда U инвариантно относительно φ.*

$\triangleright$ Для любого вектора $\mathbf{a}$ выполнено $(\varphi - \lambda_0)(\mathbf{a}) \in \operatorname{Im}(\varphi - \lambda_0) \leqslant U$. Тогда если $\mathbf{a} \in U$, то $\varphi(\mathbf{a}) = (\varphi - \lambda_0)(\mathbf{a}) + \lambda_0\mathbf{a} \in U$ (оба слагаемых принадлежат в U). $\square$

Предыдущее предложение согласуется с соответствием между инвариантными подпространствами сопряженных операторов (см. главу 6).

Предложение 1.5. *Пусть $\varphi \in L(V,V)$ — изоморфизм. Пусть подпространство $U \leqslant V$ инвариантно относительно $\varphi \in L(V,V)$ и $\dim U < \infty$. Тогда U инвариантно относительно φ^{-1}.*

$\rhd$ Сужение $\varphi|_U : U \to U$ является инъекцией, поэтому (см. теорему 1.3 главы 2) является изоморфизмом. Значит $(\varphi|_U)^{-1} : U \to U$ — сужение φ^{-1} на U. $\square$

Иногда вид матрицы линейного преобразования говорит о наличии некоторых инвариантных подпространств.

Предложение 1.6. *Пусть $\dim V = n < \infty$ и $\varphi \in L(V,V)$ имеет в базисе $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ матрицу $A = (a_{ij})$. Подпространство $U_k = \langle \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k \rangle$ инвариантно относительно $\varphi \Leftrightarrow a_{ij} = 0$ для всех $i = k+1, \ldots, n$, $j = 1, 2, \ldots, k$ (то есть матрица A имеет блочно-треугольный вид $\begin{pmatrix} B & C \\ O & D \end{pmatrix}$, где $O \in \mathbf{M}_{(n-k)\times k}$ — нулевая матрица).*
Кроме того, в таком случае B — матрица сужения $\varphi|_{U_k}$ в базисе $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k$ пространства U_k.

$\rhd$ По определению матрицы линейного отображения, $a_{ij} = 0$ для всех $i = k+1, \ldots, n \Leftrightarrow \varphi(\mathbf{e}_j) \in \langle \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k \rangle$. Остается воспользоваться предложением 1.1. $\square$

В условиях предложения 1.6 матрица D соответствует матрице фактор-оператора (в базисе $\mathbf{e}_i + U$, $i = k+1, \ldots, n$).

Следствие. *Пусть $\varphi \in L(V,V)$, $\varphi \xrightarrow[\mathrm{e,e}]{} A$. Тогда A верхнетреугольная $\Leftrightarrow$ все подпространства $\langle \mathbf{e}_1, \ldots, \mathbf{e}_k \rangle$, $k = 1, \ldots, n$, инвариантны относительно φ.*

Аналогично предложению 1.6, матрица A (где $\varphi \xrightarrow[\mathrm{e,e}]{} A$) имеет блочно-треугольный вид $\begin{pmatrix} B & O \\ C & D \end{pmatrix}$, где $O \in \mathbf{M}_{k\times(n-k)} \Leftrightarrow \langle \mathbf{e}_{k+1}, \mathbf{e}_{k+2}, \ldots, \mathbf{e}_n \rangle$ инвариантно относительно φ. Соответственно, блочно-диагональная структура A (с квадратными блоками по диагонали) означает разложение V в прямую сумму инвариантных подпространств, каждое из которых — линейная оболочка нескольких подряд идущих базисных векторов. В частности, A диагональна тогда и только тогда, когда все одномерные подпространства $\langle e_i \rangle$, $i = 1, \ldots, n$, инвариантны относительно φ.

В последнем случае структура оператора φ вполне понятна: рассматриваемый базис таков, что на каждом из них φ действует умножением на константу.

Предложение 1.7. *Пусть $\dim V = n < \infty$, $\varphi \in L(V,V)$ и $f \in \mathbb{F}[X]$ — многочлен степени k такой, что оператор $f(\varphi)$ вырожден. Тогда существует ненулевое подпростанство размерности не больше k, инвариантное относительно φ.*

$\rhd$ Пусть $\mathbf{a}$ — ненулевой вектор из $\operatorname{Ker} f(\varphi)$. Покажем, что подпространство $U = \langle \mathbf{a}, \varphi(\mathbf{a}), \varphi^2(\mathbf{a}), \ldots, \varphi^{k-1}(\mathbf{a}) \rangle$ инвариантно относительно φ. Условия из предложения 1.1, очевидно выполнены для всех порождающих U векторов, кроме возможно, $\varphi^{k-1}(\mathbf{a})$. Итак, нужно проверить, что $\varphi^k(\mathbf{a}) \in U$.
Пусть $f(x) = \alpha_k x^k + \alpha_{k-1} x^{k-1} + \ldots + \alpha_1 x + \alpha_0$, $\alpha_k \neq 0$. Тогда
$(f(\varphi))(\mathbf{a}) = \alpha_k \varphi^k(\mathbf{a}) + \alpha_{k-1}\varphi^{k-1}(\mathbf{a}) + \ldots + \alpha_1 \varphi(\mathbf{a}) + \alpha_0 \mathbf{a} = \mathbf{0}$, откуда видим, что $\varphi^k(\mathbf{a})$ линейно выражается через $\mathbf{a}, \varphi(\mathbf{a}), \varphi^2(\mathbf{a}), \ldots, \varphi^{k-1}(\mathbf{a})$, т.е. $\varphi^k(\mathbf{a}) \in U$. $\square$

Ниже как следствие предложения 1.7 мы установим существование инвариантных подпространств малой размерности в случаях $\mathbb{F} = \mathbb{C}$ ($\dim = 1$) и $\mathbb{F} = \mathbb{R}$ ($\dim \leqslant 2$).

§ 2. Собственные векторы. Диагонализируемость.

В этом параграфе полагаем, что поле $\mathbb{F}$ — либо $\mathbb{R}$, либо $\mathbb{C}$. Хотя теорию несложно обобщить на случай произвольного поля $\mathbb{F}$, используя наряду с $\mathbb{F}$ *алгебраическое замыкание* — расширение $\mathbb{K}$ поля $\mathbb{F}$, для которого каждый многочлен из $\mathbb{K}[X]$ имеет корень.

Собственные значения. Собственные векторы и собственные подпространства.

Определение. Константа $\lambda_0 \in \mathbb{F}$ называется *собственным значением* преобразования φ, если $\exists\, \mathbf{a} \in V$ такой, что $\mathbf{a} \neq \mathbf{0}$ и
$$\varphi(\mathbf{a}) = \lambda_0 \mathbf{a}. \tag{3.1}$$

Определение. Если для $\mathbf{a} \neq \mathbf{0}$ выполнено (3.1), то вектор $\mathbf{a}$ называется *собственным вектором* преобразования φ, отвечающим собственному значению λ_0.

Предложение 2.1. *Ненулевой вектор* $\mathbf{a}$ *является собственным вектором преобразования* $\varphi \Leftrightarrow$ *одномерное подпространство* $\langle \mathbf{a} \rangle$ *инвариантно относительно* φ.

$\rhd$ В силу предложения 1.1, условие инвариантности $\langle \mathbf{a} \rangle$ эквивалентно выполнению (3.1) для некоторого $\lambda_0 \in \mathbb{F}$. $\square$

Предложение 2.2. *Ненулевой вектор* $\mathbf{a}$ *является собственным вектором преобразования* φ, *отвечающим собственному значению* $\lambda_0 \Leftrightarrow \mathbf{a} \in \operatorname{Ker}(\varphi - \lambda_0)$.

$\rhd$ Достаточно заметить, что (3.1) эквивалентно условию $(\varphi - \lambda_0)(\mathbf{a}) = \mathbf{0}$. $\square$

Отметим, что пара предложений выше согласуется с предложением 1.7: наличие одномерных инвариантных подпространств для φ связано с вырожденностью оператора $\varphi - \lambda_0$ (многочлена от φ первой степени).

Определение. Пусть λ_0 — собственное значение преобразования $\varphi \in L(V, V)$. Подпространство $\operatorname{Ker}(\varphi - \lambda_0)$ называется *собственным подпространством* для преобразования φ (отвечающим собственному значению λ_0).

Собственное подпространство условимся обозначать V_{λ_0}. Таким образом, по определению $V_{\lambda_0} = \operatorname{Ker}(\varphi - \lambda_0)$ — содержит все собственные векторы, отвечающие собственному значению λ_0, и $\mathbf{0}$.

Иногда, в других курсах и книгах, собственными подпространствами называют также и подпространства $\operatorname{Ker}(\varphi - \lambda_0)$. Отметим, что введенные формально новые понятия, такие как собственное подпространство, по сути лишь новая терминология для важных частных случаев определенных ранее понятий.

Предложение 2.3. *Пусть* $\lambda_1, \lambda_2, \ldots, \lambda_k$ — *различные собственные значения для* $\varphi \in L(V, V)$. *Тогда* $\sum\limits_{i=1}^{k} V_{\lambda_i}$ — *прямая сумма.*

$\rhd$ Это предложение сразу следует из предложения 3.7 главы 2 — достаточно рассмотреть линейные многочлены $f_i(x) = x - \lambda_i$.

Приведем еще одно, независимое, доказательство. Предположим, противное, и пусть, скажем, $\mathbf{a}_1 \in V_{\lambda_1} \cap \sum\limits_{i=2}^{k} V_{\lambda_i}$, $\mathbf{a}_1 \neq \mathbf{0}$ (см. теорему 3.1, глава 1). Тогда $\mathbf{a}_1 = \sum\limits_{i=2}^{k} \mathbf{a}_i$ для некоторых $\mathbf{a}_i \in V_{\lambda_i}$. Применим к этому равенству преобразование $\psi = \prod\limits_{i=2}^{k} (\varphi - \lambda_i)$. Так как $(\varphi - \lambda_i)(\mathbf{a}_i) = \mathbf{0}$ и в произведении $\prod\limits_{i=2}^{k} (\varphi - \lambda_i)$ преобразования перестановочны, то $\psi(\mathbf{a}_i) = \mathbf{0}$ для $i = 2, \ldots, k$. С другой стороны, $\psi(\mathbf{a}_1) = \prod\limits_{i=2}^{k} (\lambda_1 - \lambda_i)\mathbf{a}_1 \neq \mathbf{0}$. Противоречие. $\square$

Характеристический многочлен

Далее в этом параграфе полагаем $\dim V = n < \infty$.

> **Определение.** Пусть φ имеет в некотором базисе матрицу A. Функция $|A-\lambda E|$ переменной λ называется *характеристическим многочленом* оператора φ.

Для характеристического многочлена примем обозначение $\chi_\varphi(\lambda)$.

> **Определение.** Уравнение $\chi_\varphi(\lambda) = 0$ называется *характеристическим уравнением*, а его (комплексные) корни — *характеристическими числами* преобразования φ.

Характеристическое уравнение дает условие вырожденности оператора $\varphi-\lambda$; это важное соображение еще будет зафиксировано ниже (см. теорему 2.1). А сперва сделаем некоторые наблюдения про $\chi_\varphi(\lambda)$.

Предложение 2.4. *$\chi_\varphi(\lambda)$ является многочленом степени n:*

$$\chi_\varphi(\lambda) = (-1)^n\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_{n-1}\lambda + a_n, \tag{3.2}$$

где $a_i \in \mathbb{F}$. При этом

$$a_1 = (-1)^{n-1}\operatorname{tr} A, \quad a_n = |A|.$$

$\triangleright$ Формулу вида (3.2) можно получить, пользуясь явным разложением определителя $|A-\lambda E|$. Отсюда сразу можно получить $a_n = \chi_\varphi(0) = |A|$.

Далее, диагональные элементы матрицы $|A-\lambda E|$ — линейные функции от λ (равные $a_{ii}-\lambda$), а внедиагональные элементы — константы, поэтому в явном разложении определителя все слагаемые, кроме произведения диагональных элементов будут содержать λ в степени не превышающей $n-2$:

$$\chi_\varphi(\lambda) = \prod_{i=1}^{n}(a_{ii}-\lambda) + g(\lambda),$$

где $\deg g \leqslant n-2$. Отсюда в $\chi_\varphi(\lambda)$ коэффициент при λ^n равен $(-1)^n$, а коэффициент при λ^{n-1} равен $a_1 = (-1)^{n-1}\sum\limits_{i=1}^{n} a_{ii} = (-1)^{n-1}\operatorname{tr} A$. $\square$

Теперь разложим (над полем комплексных чисел) многочлен $\chi_\varphi(\lambda)$ на линейные множители:

$$\chi_\varphi(\lambda) = (-1)^n\prod_{i=1}^{n}(\lambda-\mu_i), \tag{3.3}$$

где $\mu_i \in \mathbb{C}$ — характеристические числа.

Среди μ_i могут быть равные числа. Группируя равные характеристические числа, примем далее такие обозначения: пусть $\lambda_1, \lambda_2, \ldots, \lambda_k$ — попарно различные характеристические числа, а $s_1, s_2, \ldots, s_k$ — их (алгебраические) *кратности*, так что $\sum\limits_{i=1}^{k} s_i = n$. Тем самым, формула (3.3) принимает вид

$$\chi_\varphi(\lambda) = (-1)^n\prod_{i=1}^{k}(\lambda-\lambda_i)^{s_i}. \tag{3.4}$$

Предложение 2.5. *1). Сумма (с учетом кратности) всех характеристических чисел равна* $\operatorname{tr} A$.
2). Произведение (с учетом кратности) всех характеристических чисел равна $|A|$.

$\triangleright$ Достаточно приравнять соответствующие коэффициенты в (3.2) и (3.3) (или же воспользоваться теоремой Виета). $\square$

Отметим следующий важный частный случай.

Предложение 2.6. *Пусть матрица A оператора φ (в некотором базисе) — верхнетреугольная. Тогда характеристические числа совпадают с числами, расположенными на диагонали матрицы A.*

$\triangleright$ Это ясно, поскольку определитель верхнетреугольной матрицы $A - \lambda E$ равен произведению диагональных элементов. $\square$

Следствие. *Пусть в случае $\mathbb{F} = \mathbb{R}$ оператор φ в некотором базисе имеет верхнетреугольную матрицу. Тогда все характеристические числа φ вещественные.*

Следующее предложение обосновывает корректность обозначения $\chi_\varphi(\lambda)$.

Предложение 2.7. *Характеристический многочлен преобразования φ не зависит от выбора базиса в V.*

$\triangleright$ Нужно доказать, что если $\varphi \underset{\mathrm{e,e}}{\longrightarrow} A$ и $\varphi \underset{\mathrm{e',e'}}{\longrightarrow} A'$, то $|A - \lambda E| = |A' - \lambda E|$. По теореме 2.3 $A' = S^{-1}AS$, поэтому
$|A' - \lambda E| = |S^{-1}AS - \lambda E| = |S^{-1}AS - \lambda S^{-1}ES| = |S^{-1}(A - \lambda E)S| = |S^{-1}| \cdot |A - \lambda E| \cdot |S| =$
$= |S^{-1}| \cdot |S| \cdot |A - \lambda E| = |S^{-1}S| \cdot |A - \lambda E| = |E| \cdot |A - \lambda E| = |A - \lambda E|$. $\square$

Следствие. *Определитель, след, набор характеристических чисел (с учетом кратностей) матрицы оператора φ не зависят от выбора базиса.*

Теорема 2.1. *Для константы $\lambda_0 \in \mathbb{F}$ следующие условия эквивалентны: λ_0 — собственное значение для $\varphi \Leftrightarrow \lambda_0$ — характеристическое число (т.е. корень χ_φ).*

$\triangleright$ λ_0 — собственное значение $\Leftrightarrow \operatorname{Ker}(\varphi - \lambda_0) \neq O \Leftrightarrow$ (переходя к координатам) СЛУ $(A - \lambda_0 E)X = O$ имеем ненулевое решение $\Leftrightarrow$ квадратная матрица $(A - \lambda_0 E)$ вырожденная $\Leftrightarrow$ $|A - \lambda_0 E| = 0$. $\square$

Таким образом, в случае $\mathbb{F} = \mathbb{C}$ «собственное значение» и «характеристическое число» — это одно и то же. В случае $\mathbb{F} = \mathbb{R}$ разница только в том, что характеристическое число может не принадлежать $\mathbb{R}$; в этом случае собственные значения — это в точности вещестенные характеристические числа.

Теперь мы готовы зафиксировать результат об инвариантных подпространствах малой размерности.

Теорема 2.2. *1). Пусть $\mathbb{F} = \mathbb{C}$. Тогда для φ существует одномерное инвариантное подпространство (или эквивалентно, существует собственный вектор).*
2). Пусть $\mathbb{F} = \mathbb{R}$ и n нечетно. Тогда для φ существует одномерное инвариантное подпространство.
3). Пусть $\mathbb{F} = \mathbb{R}$. Тогда для φ существует ненулевое инвариантное подпространство размерности не выше 2.

$\triangleright$ 1) Следует из теоремы 2.1, поскольку χ_φ имеет (комплексный) корень.
2) Следует из теоремы 2.1, поскольку при нечетном n многочлен $\chi_\varphi \in \mathbb{R}[X]$ имеет вещественный корень.
3) Если $\chi_\varphi \in \mathbb{R}[X]$ имеет вещественный корень, найдем, как и в пункте 2), одномерное инвариантное подпространство.
Иначе, пусть $\lambda_0 \in \mathbb{C} \setminus \mathbb{R}$ — корень χ_φ. Многочлен $f(x) = x^2 - (\lambda_0 + \overline{\lambda_0})x + \lambda_0\overline{\lambda_0}$ принадлежит $\mathbb{R}[X]$ (отметим, что $\overline{\lambda_0}$ — тоже корень χ_φ). При этом оператор $f(\varphi)$ вырожден, так как его матрица $f(A) = (A - \lambda_0 E)(A - \overline{\lambda_0}E)$ вырождена (здесь, как обычно, A — матрица оператора φ). Действительно, $\det f(A) = |A - \lambda_0 E| \cdot |A - \overline{\lambda_0}E| = 0$. Теперь нужное утверждение следует из предложения 1.7. $\square$

Теорема 2.3. *Пусть* $\lambda_i \in \mathbb{F}$. *является корнем кратности* s_i *характеристического многочлена* $\chi_\varphi(\lambda)$ *преобразования* φ. *Тогда*

$$1 \leqslant \dim V_{\lambda_i} \leqslant s_i.$$

$\rhd$ Пусть $\dim V_{\lambda_i} = s$. Выберем в $\dim V_{\lambda_i}$ базис $(\mathbf{e}_1, \ldots, \mathbf{e}_s)$ и дополним его до базиса $\mathrm{e} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ в пространстве V. Тогда $\varphi \xrightarrow[\mathrm{e,e}]{} A$, где A имеет блочный вид $\begin{pmatrix} \lambda_i E_s & C \\ O & D \end{pmatrix}$. Пользуясь формулой определителя с углом нулей, имеем
$\chi_\varphi(\lambda) = |A - \lambda E| = |\lambda_i E_s - \lambda E_s| \cdot |D - \lambda E_{n-s}| = (\lambda_i - \lambda)^s p(\lambda)$, где p — многочлен.
Тем самым, $s_i \geqslant s$. $\square$

Величину $\dim V_{\lambda_i}$ иногда называют *геометрической кратностью* собственного значения λ_i. Таким образом, в теореме 2.3 утверждается, что геометрическая кратность корня не превышает алгебраической кратности.

Диагонализируемость

Определение. Преобразование φ называется *диагонализируемым*, если в V существует базис, в котором матрица φ имеет диагональный вид.

Ранее мы видели (см. следствие из предложения 2.6), что в случае наличие характеристических чисел, лежащих вне поля $\mathbb{F}$, препятствует не только диагонализируемости, но даже треугольному виду матрицы преобразования. Поэтому в критерии ниже полагаем, что все характеристические числа оператора φ принадлежат $\mathbb{F}$ (в случае $\mathbb{F} = \mathbb{C}$ это условие не несет в себе никаких ограничений, т.е. выполнено всегда).

Теорема 2.4 (критерий-1 диагонализируемости). *Пусть оператор* φ *имеет попарно различные характеристические числа* $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{F}$ *кратностей* $s_1, s_2, \ldots, s_k$ *соответственно. Тогда следующие условия эквивалентны:*
1) φ *диагонализируем;*
2) В V *существует базис из собственных векторов;*
3) $\dim V_{\lambda_i} = s_i$ *для* $i = 1, 2, \ldots, k$*;*
4) $V = \bigoplus_{i=1}^{k} V_{\lambda_i}$.

$\rhd$ 1) $\Leftrightarrow$ 2). Очевидно следует из определения матрицы линейного преобразования.
2) $\Rightarrow$ 3). В базисе из собственных пусть t_i векторов соответствуют λ_i $(i = 1, \ldots, k)$, так что $t_1 + \ldots + t_k = n$. Тогда в V_{λ_i} содержится линейно независимую систему из t_i векторов, откуда $\dim V_{\lambda_i} \geqslant t_i$. С учетом теоремы 2.3, имеем

$$t_i \leqslant \dim V_{\lambda_i} \leqslant s_i, \quad i = 1, \ldots, k.$$

Суммируем эти неравенства. Поскольку $s_1 + \ldots + s_k = t_1 + \ldots + t_k = n$, все неравенства должны обращаться в равенства, т.е. $\dim V_{\lambda_i} = s_i$ для всех $i = 1, \ldots, k$.
3) $\Rightarrow$ 4). Следует из теоремы 2.3 с учетом $s_1 + \ldots + s_k = n$.
4) $\Rightarrow$ 2). В каждом V_{λ_i} выберем базис. Объединение этих базисов будет нужным базисом в V (см. критерий-2 прямой суммы — теорема 3.2 главы 1), и этот базис составлен из собственных векторов. $\square$

Следствие. *Если* $\chi_\varphi(\lambda)$ *имеет* n *различных вещественных корней, принадлежащих* $\mathbb{F}$, *то* φ *диагонализируемо.*

Это следствие очерчивает достаточно большой класс диагонализируемых операторов.

Вообще, определять диагонализируем оператор или нет, часто бывает удобно, используя условие 3) из теоремы 2.4.

Посмотрим, что означает условие диагонализируемости в переводе «на матричный язык». В силу формулы изменения матрицы при замене базиса (см. теорему 2.3 главы 2), для диагонализируемости матрицы A (матрицы оператора φ в некотором базисе) необходимо и достаточно существование такой невырожденной матрицы S, что матрица $S^{-1}AS$ (подобная матрице A) диагональна.

Примеры

I. Преобразование $\mathbb{R}^2 \to \mathbb{R}^2$ поворота на угол, не кратный π, не диагонализиремо (нет собственных векторов, или поскольку характеристические числа не вещественны).
Сжатие к прямой (проходящей через O), проектирование на прямую, отражение относительно прямой, напротив, диагонализируемые.

II. Матрица $\begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix}$ не диагонализируема (как для $\mathbb{F} = \mathbb{R}$, так и для $\mathbb{F} = \mathbb{C}$), так как λ_0 — корень кратности 2, но $\dim V_{\lambda_0} = 1$.
(Объяснение без использования теоремы 2.4 такое: если бы это преобразование было диагонализируемым, то диагональный вид был бы нулевой матрицей, и значит, преобразование было бы нулевым, что неверно.) Этот пример — частный случай жордановой клетки, общая теория на этот счет — в следующем параграфе.

III. Пусть $V = \mathbf{P}_n = \langle 1, x, x^2, \ldots, x^n \rangle$ и Оператор дифференцирования $d : V \to V$ имеет единственное характеристическое число 0 кратности $n + 1$ (см. матрицу d в стандартном базисе — пример в конце параграфа § 2.). При $n \geqslant 1$ преобразование d не является диагонализируемым, как и любой линейный дифференциальный оператор вида $d - \lambda I_V$.

§ 3. Структура недиагонализируемых операторов.

В этом параграфе V — векторное пространство над полем $\mathbb{F}$, $\dim V = n < \infty$, а φ — фиксированное линейное преобразование $V \to V$. Предполагается (если не оговаривается противное), что все характеристические числа оператора φ принадлежат $\mathbb{F}$ (для $\mathbb{F} = \mathbb{C}$ это выполнено всегда).

Характеристические числа обозначаем $\lambda_1, \lambda_2, \ldots, \lambda_k$, а $s_1, s_2, \ldots, s_k$ — их кратности соответственно.

Треугольный вид

Лемма 3.1. *Существует базис, в котором матрица φ верхнетреугольная, с заданным порядком расстановки характеристических чисел по диагонали.*

$\triangleright$ СХЕМА. Требуется найти флаг инвариантных подпространств $\langle \mathbf{e}_1, \ldots, \mathbf{e}_k \rangle$, $k = 1, \ldots, n$.

Индукция. Найдем $(n-1)$-мерное инвариантное подпространство $U \geqslant \operatorname{Im}(\varphi - \lambda_i)$ (см. предложение 1.4). Возьмем любой $\mathbf{e}_n \notin U$. Тогда $\varphi(\mathbf{e}_n) = \lambda_i \mathbf{e}_n + \mathbf{b}$, где $\mathbf{b} \in U$. Тогда для ограничения $\varphi|_U : U \to U$ характеристический многочлен равен $\dfrac{\chi_\varphi(\lambda)}{\lambda_i - \lambda}$. Применяя предположение индукции к U, находим в U нужные базисные векторы $\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}$. $\square$

Упражнение. Пусть $f \in \mathbb{F}[X]$. Найдите характеристические числа (и их кратности) оператора $f(\varphi)$ (зная характеристические числа (и их кратности) оператора φ.

Корневые подпространства. Теорема Гамильтона-Кэли.

Определяемые ниже *корневые подпространства* можно мыслить как обобщение собственных подпространств $V_{\lambda_i} = \operatorname{Ker}(\varphi - \lambda_i)$.

Для каждого характеристического числа λ_i рассмотрим систему вложенных подпространств:

$$O \leqslant \operatorname{Ker}(\varphi - \lambda_i) \leqslant \operatorname{Ker}(\varphi - \lambda_i)^2 \leqslant \operatorname{Ker}(\varphi - \lambda_i)^3 \leqslant \dots \tag{3.5}$$

Так как $\dim V < \infty$, в этой цепочке начиная с некоторого шага наступит стабилизация. Обозначим m_i номер этого шага, так что $\operatorname{Ker}(\varphi - \lambda_i)^{m_{i-1}} \neq \operatorname{Ker}(\varphi - \lambda_i)^{m_i} = \operatorname{Ker}(\varphi - \lambda_i)^{m_{i+1}} = \dots$.

Подпространство $\operatorname{Ker}(\varphi - \lambda_i)^{m_i}$ назовем *корневым*, отвечающим собственному значению λ_i. Обозначаем это корневое подпространство V^{λ_i}. Равество $V^{\lambda_i} = \operatorname{Ker}(\varphi - \lambda_i)^m$ будет верно для всех $m \geqslant m_i$, и наоборот, m_i — минимальное натуральное m, для которого верно это равенство.

Следующие утверждения о $\operatorname{Ker}(\varphi - \lambda_i)^{t_i}$ (в частности о V^{λ_i}) можно сравнить с теоремами 2.3, 2.3 и 2.4 о V_{λ_i}.

Лемма 3.2. *Для любых* $t_i \in \mathbb{Z}_+$ *сумма* $\sum\limits_{i=1}^{k} \operatorname{Ker}(\varphi - \lambda_i)^{t_i}$ *— прямая сумма.*
В частности, $\sum\limits_{i=1}^{k} V^{\lambda_i}$ *— прямая сумма.*

▷ Это предложение сразу следует из предложения 3.7 главы 2 — достаточно рассмотреть многочлены $f_i(x) = (x - \lambda_i)^{t_i}$. □

Лемма 3.3. $\dim \operatorname{Ker}(\varphi - \lambda_i)^{s_i} \geqslant s_i$.

▷ СХЕМА. Согласно лемме 3.1, существует базис $\mathrm{e} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$, в котором матрица оператора φ верхнетреугольная, причем константны, равные λ_i, расположены в первых s_i диагональных клетках. Это будет означать что для $j = 1, \dots, s_i$ выполнено $\varphi(\mathbf{e}_j) = \lambda_i \mathbf{e}_j + \sum\limits_{x=1}^{j-1} c_x \mathbf{e}_x$, или $(\varphi - \lambda_i)(\mathbf{e}_j) \in \langle \mathbf{e}_1, \dots, \mathbf{e}_{j-1} \rangle$; отсюда $(\varphi - \lambda_i)^j(\mathbf{e}_j) = 0$, в частности, $\mathbf{e}_j \in \operatorname{Ker}(\varphi - \lambda_i)^{s_i}$ для $j = 1, \dots, s_i$. Требуемое доказано, поскольку в $\operatorname{Ker}(\varphi - \lambda_i)^{s_i}$ нашлась линейно независимая система из s_i векторов. □

Следствие. $\dim V^{\lambda_i} \geqslant s_i$.

Теорема 3.1. *Справедливы следующие утверждения:*

$$1). \quad \boxed{V \;=\; \bigoplus_{i=1}^{k} \dim V^{\lambda_i}}; \tag{3.6}$$

$$2). \quad \boxed{\dim V^{\lambda_i} = s_i};$$

3) $m_i \leqslant s_i$ *(т.е. стабилизация в (3.5) наступает на* s_i*-м шаге или ранее),* $i = 1, \dots, k$

▷ 1), 2). Из леммы 3.2, с учетом последнего следствия $\dim V^{\lambda_i} \geqslant s_i$, размерность прямой суммы $\bigoplus\limits_{i=1}^{k} \dim V^{\lambda_i}$ не меньше $\sum\limits_{i=1}^{k} s_i = n = \dim V$, значит эта сумма обязана совпадать с V, а во все неравенства обязаны обращаться в равенство.
3). Следует из 2) с учетом леммы 3.3. □

Предложение 3.1. *Минимальным многочленом для оператора* φ *является многочлен*

$$\mu(\lambda) = \prod_{i=1}^{k} (\lambda - \lambda_i)^{m_i}.$$

▷ Из предложения 3.7 главы 2 следует, что

$$\operatorname{Ker}\Big(\prod_{i=1}^{k} (\varphi - \lambda_i)^{t_i}\Big) \;=\; \bigoplus_{i=1}^{k} \operatorname{Ker}(\varphi - \lambda_i)^{t_i}. \tag{3.7}$$

Если $t_i = m_i$, $i = 1, \ldots, k$, то в правой части (3.7) прямая сумма корневых подпространств, т.е. V, значит $\prod\limits_{i=1}^{k}(\varphi - \lambda_i)^{m_i}$ — нулевой оператор, т.е. $\mu(\lambda)$ действительно аннулирует φ.

Согласно предложению 1.11 главы 2, минимальный многочлен для φ является делителем μ, а значит имеет вид $\prod_{i=1}^{k}(\lambda - \lambda_i)^{t_i}$, где $t_i \leqslant m_i$, $i = 1, \ldots, k$.

Пусть хотя бы одно из этих неравенств строгое, например, $t_1 < m_1$. Тогда $\dim \mathrm{Ker}(\varphi - \lambda_i)^{t_1} < s_1$ и $\dim \mathrm{Ker}(\varphi - \lambda_i)^{t_i} \leqslant s_i$, $i = 2, \ldots, k$. Отсюда следует, что размерность прямой суммы в правой части (3.7) строго меньше $\sum\limits_{i=1}^{k} s_i = n$, а значит $\mathrm{Ker}(\prod\limits_{i=1}^{k}(\varphi - \lambda_i)^{t_i}) \neq V$. Следовательно, при $t_1 < m_1$ многочлен $\prod_{i=1}^{k}(\lambda - \lambda_i)^{t_i}$ не аннулирует φ. $\square$

Предложение 3.2 (критерий-2 диагонализируемости). *Следующие условия на φ эквивалентны:*
1) φ диагонализируем;
2) $V_{\lambda_i} = V^{\lambda_i}$ для всех $i = 1, 2, \ldots, k$;
3) μ_φ раскладывается (в $\mathbb{F}[X]$) на различные линейные множители.

$\triangleright$ 1) $\Leftrightarrow$ 2). Достаточно составить условие 4) в критерии-1 диагонализируемости (теорема 2.4), (3.6) и очевидные включения $V_{\lambda_i} \leqslant V^{\lambda_i}$.
2) $\Leftrightarrow$ 3). Условие 2) означает, что $m_i = 1$ для всех $i = 1, 2, \ldots, k$. Остается воспользоваться предыдущим предложением 3.1. $\square$

Упражнение. Оператор $\varphi \in L(\mathbb{R}^n, \mathbb{R}^n)$ удовлетворяет равенству $\varphi^3 = \varphi$. Докажите, что φ диагонализируем.

Упражнение. Пусть φ диагонализируем, а $U \leqslant V$ — инвариантное подпространство. Докажите, что оператор $\varphi|_U$ ограничения на U тоже диагонализируем, как и фактор-оператор.

Теорема 3.2 (Теорема Гамильтона-Кэли). *Для всякого $\varphi \in L(V, V)$ выполнено*

$$\chi_\varphi(\varphi) = 0.$$

$\triangleright$ Из описания минимального многочлена μ_φ (предложение 3.1) следует, что χ_φ делится на μ_φ, и значит, многочлен χ_φ аннулирует φ. $\square$

Итак, в случае $\mathbb{F} = \mathbb{C}$ теорема Гамильтона-Кэли доказана для всех операторов, а в случае $\mathbb{F} = \mathbb{R}$ — верна также для всех операторов, но доказана для операторов, удовлетворяющих условию, сформулировнному в начале параграфа: все характеристические числа оператора φ принадлежат $\mathbb{R}$. Устранить этот недостаток можно, посмотрев на теорему Гамильтона-Кэли как формальное на тождество для матриц $A \in \mathbf{M}_{n\times n}(\mathbb{F})$. Так как это тождество верно для всех матриц из $\mathbf{M}_{n\times n}(\mathbb{C})$, то оно верно и для всех матриц из $\mathbf{M}_{n\times n}(\mathbb{R}) \subset \mathbf{M}_{n\times n}(\mathbb{C})$.

Упражнение. Матрица $A \in \mathbf{M}_{n\times n}$ таковы, что $A^k = O$ при некотором $k \in \mathbb{N}$. Докажите, что $A^n = O$.

Упражнение.* Пусть $\mu \in \mathbb{R}[X]$ — минимальный многочлен матрицы $A \in \mathbf{M}_{n\times n}(\mathbb{R})$. Докажите, что μ является минимальным многочленом для A, при рассмотрении A как $A \in \mathbf{M}_{n\times n}(\mathbb{C})$ (и соответственно минимальный среди многочленов над $\mathbb{C}$).
Попробуйте обобщить это утверждение для произвольного поля и его расширения $\mathbb{F} \subset \mathbb{K}$.

ЖНФ: определения и формулировки.

Жордановой клеткой, отвечающей константе λ_0, называют квадратную матрицу вида

$$\begin{pmatrix} \lambda_0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & \lambda_0 & 1 & \ldots & 0 \\ & & & \ldots & & \\ & & & \ldots & & \\ 0 & 0 & \ldots & 0 & 0 & \lambda_0 \end{pmatrix}.$$

Блочно-диагональная матрица, у которой каждый диагональный блок является жордановой клеткой, называется *жордановой матрицей*, или матрицей, имеющей *жорданову нормальную форму* (жнф). Очевидно, жордановы матрицы являются верхнетреугольными.

Базис, в котором φ имеет жнф, называют *жорданов базис.*

«Прочитаем» жнф (по определению матрицы линейного преобразования) и поймем, что значит жорданов базис. Каждая жорданова клетка $t \times t$ соответствует так называемой *жордановой цепочке* длины t:

$$\mathbf{0} \leftarrow \mathbf{e}_1 \leftarrow \mathbf{e}_2 \leftarrow \ldots \leftarrow \mathbf{e}_t.$$

Здесь стрелочкой обозначено применение оператора $\varphi - \lambda_i$. Каждая жорданова цепочка начинается с собственного вектора, следующие векторы называется *присоединенными* ($\mathbf{e}_{i+1}$ — присоединенный для $\mathbf{e}_i$).

Основными утверждениями здесь являются следующие.

Теорема 3.3 (существование ЖНФ). *В V существует базис, в котором матрица φ — жорданова.*

Теорему существования можно переформулировать так: существует базис, являющийся объединением жордановых цепочек.

Теорема 3.4 (единственность ЖНФ). *Для данного φ единственна ЖНФ (с точностью до перестановки жордановых клеток по диагонали).*

Теорему единственности можно переформулировать так: для каждого λ_i количество жордановых цепочек данной длины для жорданова базиса определено однозначно.

Единственность жнф выполнена несмотря на то, что, как увидим в доказательстве сущесвтвования, при конструировании жорданова базиса может быть достаточно степеней свободы.

Доказательство существования ЖНФ

Достаточно в каждом корневом подпространстве V^{λ_i} построить базис, являющийся объединением жордановых цепочек. Далее в доказательстве существования рассматриваем нильпотентный оператор $\psi : V^{\lambda_i} \to V^{\lambda_i}$, являющийся ограничением оператора $\varphi - \lambda_i$ на V^{λ_i}.

СХЕМА: построение жордановых цепочкек «сверху вниз» в каждом корневом подпространстве, исходя из «башни» 3.5, которую обозначим используя ψ:

$$O \leqslant \operatorname{Ker}\psi \leqslant \operatorname{Ker}\psi^2 \leqslant \operatorname{Ker}\psi^3 \leqslant \ldots \leqslant \operatorname{Ker}\psi^{m_i} = V^{\lambda_i}. \tag{3.8}$$

Индуктивно выберем подпространства $W_{m_i+1} = O$, далее W_{m_i}, W_{m_i-1} и т.д. такие, что

$$W_{t+1} \oplus \operatorname{Ker}\psi^t = \operatorname{Ker}\psi^{t+1}. \tag{3.9}$$

По W_{t+1} определим W_t ($t \geqslant 1$) так.

Если для W_{t+1} верно (3.9), то для $\psi(W_{t+1})$ выполнено:

i) $\psi(W_{t+1}) \leqslant \operatorname{Ker}\psi^t$;

(это следует только из того, что $W_{t+1} \leqslant \operatorname{Ker}\psi^{t+1}$.)

ii) $\dim(\psi(W_{t+1})) = \dim W_{t+1}$;

(это верно, так как $W_{t+1} \cap \operatorname{Ker}\psi = O$.)

iii) $\psi(W_{t+1}) \cap \operatorname{Ker}\psi^{t-1} = O$. (proof)

Из iii) следует, что можно определить $W_t \geqslant \psi(W_{t+1})$ так, чтобы выполнялось (3.9), т.е. $W_t \oplus \operatorname{Ker}\psi^{t-1} = \operatorname{Ker}\psi^t$.

Пользуясь определенными подпространствами W_t, несложно построить жордановы цепочки «сверху вниз». На t-м «слое», имея базис в $\psi(W_{t+1})$, дополним его до базиса в W_t, применяя ψ, «спускаем» на этаж ниже, получая базис в $\psi(W_t)$.

Упражнение. а) Докажите, что если для некотрого t выполнено $\operatorname{Ker}(\varphi - \lambda_i)^t = \operatorname{Ker}(\varphi - \lambda_i)^{t+1}$, то $\operatorname{Ker}(\varphi - \lambda_i)^t = V^{\lambda_i}$.

б) Докажите, что последовательность $\dim(\operatorname{Ker}\varphi^t) - \dim(\operatorname{Ker}\varphi^{t-1})$, $t = 1, 2, \ldots$, невозрастающая.

Доказательство единственности жнф

Пусть дан жорданов базис B. Сперва поймем, что множество B_{λ_i} векторов из B, принадлежащих цепочкам, отвечающим данному λ_i, имеет мощность s_i. (Иначе говоря, сумма длин жордановых цепочек, отвечающих λ_i, равна s_i.)

Пусть $|B_{\lambda_i}| = s'_i$. Так как $B_{\lambda_i} \subset V^{\lambda_i}$, то $s'_i \leqslant \dim V^{\lambda_i} = s_i$. С другой стороны, всего векторов в жордановом базисе n, т.е. $\sum\limits_{i=1}^{k} s'_i = n = \sum\limits_{i=1}^{k} s_i$. Значит, каждое нерваенство $s'_i \leqslant s_i$ обращается в равенство.

Кроме того, мы видим, что $\langle B_{\lambda_i} \rangle = V^{\lambda_i}$.

Далее, как в доказательстве существования рассматриваем $\psi : V^{\lambda_i} \to V^{\lambda_i}$, являющийся ограничением оператора $\varphi - \lambda_i$ на V^{λ_i}. ψ действует на жордановых цепочках как «спуск на один этаж».

Поэтому $\psi(V^{\lambda_i})$ равно линейной оболочке всех векторов из B_{λ_i}, не являющихся последними векторами своих жордановых цепочек. Аналогично, $\psi^2(V^{\lambda_i})$ равно линейной оболочке всех векторов из B_{λ_i}, не являющихся последними и предпоследними векторами своих жордановых цепочек, и т.д.

Обозначим c_d количество жордановых цепочек длины d, отвечающих λ_i. Их наблюдений выше имеем равенства:

$$\dim V^{\lambda_i} - \dim(\operatorname{Im}\psi) = c_1 + c_2 + c_3 \ldots + c_{m_i},$$

$$\dim(\operatorname{Im}\psi) - \dim(\operatorname{Im}\psi^2) = c_2 + c_3 \ldots + c_{m_i},$$

$$\dim(\operatorname{Im}\psi^2) - \dim(\operatorname{Im}\psi^3) = c_3 \ldots + c_{m_i},$$

и т.д. Отсюда можно выразить c_i через (инвариантные) характеристики φ.

Отметим, что для величины m_i (шаг, на котором происходит стабилизация ядер в (3.5)) теперь есть еще два эквивалетных описания: максимальная длина жордановой цепочки, отвечающей λ_i, в жордановом базисе, или максимальный размер жордановой клетки, отвечающей λ_i, в жнф. В этих терминах можно теперь формулировать, например, описание минимального многочлена (см. предложение 3.2).

О связи с теорией линейных дифференциальных уравнений

Пусть $V_0 = \mathbf{C}^\infty$ и $d : V_0 \to V_0$ — оператор дифференцирования.

Множество $V \leqslant V_0$ решений однородного дифф.ур.

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1 y' + a_0 y = 0$$

— это $V = \operatorname{Ker} p(d)$, где $p(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$.

V инвариантно относительно d, поэтому рассмотрим d как сужение $d : V \to V$.
Тогда p — аннулирующий многочлен для d.

Для каждого корня λ_i (кратности s_i) многочлена p попробуем найти собственный вектор f оператора d:
$d(f) = \lambda_i f$ — ответ единственный (с точностью до пропорциональности): $f(x) = e^{\lambda_i x}$.

Ясно, что $f \in V$, и так как $\operatorname{Ker}(d - \lambda_i)^{s_i} \leqslant \operatorname{Ker} p(d) = V$,
для корня λ_i вся жорданова цепочка длины s_i лежит в V:

$$e^{\lambda_i x} \leftarrow (1 + x)e^{\lambda_i x} \leftarrow (1 + x + \frac{x^2}{2})e^{\lambda_i x} \leftarrow (1 + x + \frac{x^2}{2} + \frac{x^3}{3!})e^{\lambda_i x} \leftarrow \ldots .$$

Из теории ДУ известно, что $\dim V = n$, поэтому найденный нами жорданов базис для оператора d —это базис в пространстве решений V.

Связь с теорией линейных рекуррент

Пусть V_0 — пространство всех последовательностей комплексных чисел $f = (f_0, f_1, f_2, \ldots)$
$\delta : V_0 \to V_0$ — оператор сдвига: $\delta(f) = g$ так что $g_i = f_{i+1}$.

Рассматриваем подпространство $V \leqslant V_0$
решений однородной рекурренты степени n:

$$f_{k+n} + a_{k-1} f_{k+n-1} + \ldots + a_1 f_{k+1} + a_0 f_k = 0.$$

$\dim V = n$ и это $V = \operatorname{Ker} p(\delta)$,
где $p(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$.

V инвариантно относительно δ, поэтому рассмотрим δ как сужение $\delta : V \to V$.
Тогда p — аннулирующий многочлен для δ.

Для каждого корня λ_i (кратности s_i) многочлена p попробуем найти собственный вектор f оператора δ:
$\delta(f) = \lambda_i f$ — ответ единственный (с точностью до пропорциональности):
$f = (1, \lambda_i, \lambda_i^2, \ldots)$ или $f_k = \lambda_i^k$ — геометрическая прогрессия.

Ясно, что $f \in V$, и так как $\operatorname{Ker}(\delta - \lambda_i)^{s_i} \leqslant \operatorname{Ker} p(\delta) = V$,
для корня λ_i имеется одна жорданова цепочка и одна «строго растущая» башня $\operatorname{Ker}(\delta - \lambda_i) \leqslant \ldots \leqslant \operatorname{Ker}(\delta - \lambda_i)^{s_i}$.

Для предъявления базиса (на этот раз не будем брать жорданов базис) в корневом подпространстве достаточно взять по вектору «с каждого этажа».

Возьмем $g \in V_0$ вида $g_k = q_{t-1}(k)\lambda_i^k$, где q_r — многочлен степени r.
такое $g \in \operatorname{Ker}(\delta - \lambda_i)^t \setminus \operatorname{Ker}(\delta - \lambda_i)^{t-1}$.

Глава 4

Билинейные и квадратичные формы

На протяжении всей этой главы V обозначает данное векторное пространство над полем $\mathbb{R}$ или $\mathbb{C}$. Будем одновременно развивать теорию *билинейных* форм в случае пространства над $\mathbb{R}$ и *полуторалинейных* форм в случае пространства над $\mathbb{C}$. В большинстве случаев определения, формулировки и доказательства аналогичны. Знак комплексного сопряжения при работе в пространстве над $\mathbb{R}$ можно игнорировать. Терминология для пространства над $\mathbb{C}$ приводится в скобках.

§ 1. Билинейные формы. Матрица билинейной формы

Определение

> **Определение.** Отображение $\beta : V \times V \to \mathbb{R}$ ($\beta : V \times V \to \mathbb{C}$) называется *билинейным* (*полуторалинейным*), или *билинейной (полуторалинейной) функцией*, или *билинейной (полуторалинейной) формой* на пространстве V, если $\forall\ \mathbf{a}, \mathbf{a}_1, \mathbf{a}_2, \mathbf{b}, \mathbf{b}_1, \mathbf{b}_2 \in V$ и $\forall\ \lambda \in \mathbb{R}$ ($\mathbb{C}$) выполняются равенства
> B1.1. $\beta(\mathbf{a}_1 + \mathbf{a}_2, \mathbf{b}) = \beta(\mathbf{a}_1, \mathbf{b}) + \beta(\mathbf{a}_2, \mathbf{b})$,
> B1.2. $\beta(\lambda\mathbf{a}, \mathbf{b}) = \lambda\beta(\mathbf{a}, \mathbf{b})$,
> B2.1. $\beta(\mathbf{a}, \mathbf{b}_1 + \mathbf{b}_2) = \beta(\mathbf{a}, \mathbf{b}_1) + \beta(\mathbf{a}, \mathbf{b}_2)$,
> B2.2. $\beta(\mathbf{a}, \lambda\mathbf{b}) = \overline{\lambda}\beta(\mathbf{a}, \mathbf{b})$.

Множество всех билинейных форм на пространстве V над полем $\mathbb{F}$ обозначают $\mathrm{Hom}(V, V; \mathbb{F})$. Мы будем пользоваться более коротким обозначением $\mathcal{B}(V)$ для множества всех билинейных (полуторалинейных) форм на пространстве V.

Матрица и билинейной формы. Координатная запись

Если $\dim V = n < \infty$ и в V зафиксирован некоторый базис $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$, то билинейной (полуторалинейной) форме можно сопоставить матрицу $n \times n$ следующим образом.

> **Определение.** *Матрицей билинейной (полуторалинейной) формы* $\beta \in \mathcal{B}(V)$ в базисе e называется матрица $B = (b_{ij}) \in \mathbf{M}_{n\times n}$ такая, что $(b_{ij}) = \beta(\mathbf{e}_i, \mathbf{e}_j)$ для всех $i = 1, \ldots, n$, $j = 1, \ldots, n$.

Тот факт, что B — матрица билинейной (полуторалинейной) формы β в базисе e будем обозначать $\beta \xrightarrow[\mathrm{e}]{} B$. Посмотрев на определение, матрицу B можно неформально назвать таблицей билинейного умножения.

Теорема 1.1 (координатная запись). *Пусть* $\beta \in \mathcal{B}(V)$ *и* $\beta \xrightarrow[\mathrm{e}]{} B$. *Пусть* $\mathbf{a} = \mathrm{e}X$, $\mathbf{b} = \mathrm{e}Y$. *Тогда*

$$\boxed{\beta(\mathbf{a}, \mathbf{b}) = X^T B \overline{Y}}.$$

▷ Раскроем $\beta(\mathbf{a},\mathbf{b}) = \beta(\sum\limits_{i=1}^{n} x_i\mathbf{e}_i, \sum\limits_{i=j}^{n} y_j\mathbf{e}_j)$, пользуясь линейностью по первому и (полулинейностью) по второму аргументам:
$\beta(\mathbf{a},\mathbf{b}) = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} x_i\overline{y_j}\beta(\mathbf{e}_i,\mathbf{e}_j) = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} x_i b_{ij}\overline{y_j}$. Полученная двойная сумма и есть единственный элемент матрицы $X^T B\overline{Y}$ (эту матрицу размера 1×1 мы отождествляем с числом, записанным в единственной ее ячейке). □

Формула из предыдущей теоремы фактически эквивалентна определению матрицы билинейной формы. Более, точно, справедливо следующее

Предложение 1.1. *Пусть дано отображение* $\beta : V \times V \to \mathbb{R}$ *(*$\beta : V \times V \to \mathbb{C}$*) (априори не известно, что билинейное) и матрица* $B \in M_{n\times n}$*. Пусть* $\forall\ \mathbf{a},\mathbf{b} \in V$*, имеющих координатные столбцы* X *и* Y *в базисе* e*, выполнено* $\beta(\mathbf{a},\mathbf{b}) = X^T B\overline{Y}$*. Тогда* $\beta \in \mathcal{B}(V)$*, причем* $\beta \xrightarrow[\mathrm{e}]{} B$*.*

▷ Непосредственно проверяется, что отображение β, заданное как $\beta(\mathbf{a},\mathbf{b}) = X^T B\overline{Y}$, удовлетворяет равенствам B1.1—B2.2 из опеределения, поэтому $\beta \in \mathcal{B}(V)$.

Кроме того, $\mathbf{e}_i = \mathrm{e}E_{\bullet i}$ (где $E_{\bullet i}$ — i-й столбец единичной матрицы, и из правил перемножения матриц $\beta(\mathbf{e}_i,\mathbf{e}_j) = E^T_{\bullet i} B\overline{E_{\bullet j}} = b_{ij}$. Поэтому в самом деле $\beta \xrightarrow[\mathrm{e}]{} B$. □

В зависимости от ситуации удобно пользоваться как определением матрицы билинейной формы, так и координатной записью $X^T B\overline{Y}$.

Как следствие предложения 1.1 получаем, что соответствие $\beta \xrightarrow[\mathrm{e}]{} B$ (зависящее от выбора базиса e) является взаимно-однозначным соответствием между $\mathcal{B}(V)$ и $\mathbf{M}_{n\times n}$.

Отметим, что после введения на множестве $\mathcal{B}(V)$ естественных операций сложения и умножения на константу, $\mathcal{B}(V)$ превращается в векторное пространство, при этом соответствие $\beta \xrightarrow[\mathrm{e}]{} B$ является изоморфизмом.

Изменение матрицы при замене базиса

Теорема 1.2. *Пусть в* V *выбраны базисы* e *и* e$'$*, связанные матрицей перехода* S*:* $\mathrm{e}' = \mathrm{e}S$*. Пусть* $\beta \in \mathcal{B}(V)$ *таково, что* $\beta \xrightarrow[\mathrm{e}]{} B$ *и* $\beta \xrightarrow[\mathrm{e}']{} B'$*. Тогда*

$$\boxed{B' = S^T B\overline{S}}.$$

▷ Пусть $\mathbf{a},\mathbf{b} \in V$ — произвольные векторы. Пусть $\mathbf{a} = \mathrm{e}X = \mathrm{e}'X'$, $\mathbf{b} = \mathrm{e}Y = \mathrm{e}'Y'$. Тогда по теореме имеем $\beta(\mathbf{a},\mathbf{b}) = X^T B\overline{Y}$ и $\beta(\mathbf{a},\mathbf{b}) = X'^T B\overline{Y'}$ Подставляя $X = SX'$, $Y = SY'$ (см. теорему 2.4, глава 1), имеем $\beta(\mathbf{a},\mathbf{b}) = (SX')^T B\overline{SY'} = X'^T S^T B\overline{SY'} = X'^T(S^T B\overline{S})\overline{Y'}$. В силу предложения 1.1 получаем требуемое: $B' = S^T B\overline{S}$. □

Следствие 1. *В обозначениях теоремы* $\operatorname{rg} B = \operatorname{rg} B'$*, т.е. ранг матрицы билинейной формы не зависит от выбора базиса*

▷ Достаточно заметить, что матрица перехода невырожденная, а умножение на невырожденную матрицу не меняет ранг. □

Следствие 2. *В обозначениях теоремы определители* $|B|$ *и* $|B'|$ *отличаются на положительный вещественный множитель.*

▷ По правилу произведения определителей: $|B'| = |S^T|\cdot|B|\cdot|\overline{S}| = |S|\cdot|B|\cdot\overline{|S|} = |z|^2\cdot|B|$, где $z = |S|$. □

Следствие 1 позволяет корректно ввести ранг $\operatorname{rg}\beta$ билинейной формы.

В случае, если S — *элементарная матрица*, преобразование $B \to S^T B\overline{S}$ соответствует следующим *двойным элементарным преобразовниям*: выполнеяется элементарное преобразование строк (ему соответствует домножение слева на матрицу S^T), а затем *соответствующее* элементарное преобразование столбцов строк (ему соответствует домножение справа на матрицу $\overline{S}$). Виды двойных элементарных преобразований: прибавим к i-й строке j-ю, умноженную на λ, а затем прибавим к i-му столбцу j-й, умноженный на $\overline{\lambda}$; поменяем местами i-ю и j-ю строки, а затем поменяем местами i-й и j-й столбцы; i-ю строку умножим на λ, а затем i-й столбец умножим на $\overline{\lambda}$.

Сужение

Если $U \leqslant V$ и $\beta \in \mathcal{B}(V)$, то можно рассмотреть *сужение* $\beta\,|_{U\times U}\colon U \times U \to \mathbb{R}$. Очевидно, сужение билинейной формы на подпространство U является билинейной формой на подпространстве U.

Для матрицы B обозначим через B_k левую верхнюю угловую подматрицу $k \times k$, так что $B_k = (b_{ij})$ для всех $i = 1, \ldots, k$, $j = 1, \ldots, k$. В частности, $B_1 = (b_{11})$, $B_n = B$.

Если в пространстве V выбран базис $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$, то обозначим через $\mathrm{e}^{(k)}$ упорядоченную систему $(\mathbf{e}_1, \ldots, \mathbf{e}_k)$ — это базис пространства $U_k = \langle \mathbf{e}_1, \ldots, \mathbf{e}_k \rangle$. (При $k = n$ имеем $U_k = V$.)

Предложение 1.2. *Пусть $\beta \in \mathcal{B}(V)$ и $\beta \xrightarrow[\mathrm{e}]{} B$. Тогда $\beta\,|_{U\times U} \xrightarrow[\mathrm{e}^{(k)}]{} B_k$.*

$\triangleright$ Сразу следует из определения. $\square$

Примеры (скалярное произведение, на функциях с плотностью), пр-во минковского.

§ 2. Симметричные билинейные и квадратичные формы

Симметричные билинейные формы

Определение. Билинейная (полуторалинейная) форма β на пространстве V называется *симметричной* (*эрмитовой* или *эрмитово симметричной*), если $\forall\ \mathbf{a}, \mathbf{b} \in V$

$$\beta(\mathbf{a}, \mathbf{b}) = \overline{\beta(\mathbf{b}, \mathbf{a})}.$$

Множество всех симметричных (эрмитовых) билинейных (полуторалинейных) форм на пространстве V обозначаем $\mathcal{B}_{sym}(V)$.

Отметим, сразу, что $\forall\ \mathbf{a} \in V$ значение $\beta(\mathbf{a}, \mathbf{a})$ вещественно (это утверждение содержательно для комплексного пространства).

Теорема 2.1. *Пусть $\dim V < \infty$, e — базис в V. Пусть $\beta \in \mathcal{B}(V)$, $\beta \xrightarrow[\mathrm{e}]{} B$. Тогда $\beta \in \mathcal{B}_{sym}(V) \Leftrightarrow B^* = B$ (где, как обычно, $B^* = \overline{B^T}$).*

$\triangleright$ $\boxed{\Rightarrow}$ Надо доказать, что $\overline{b_{ji}} = b_{ij}$ или что $\overline{\beta(\mathbf{e}_j, \mathbf{e}_i)} = \beta(\mathbf{e}_i, \mathbf{e}_j)$ для всевозможных пар индексов. Но это сразу следует из определения.

$\boxed{\Leftarrow}$ Пусть $\mathbf{a}, \mathbf{b} \in V$ — произвольные векторы, $\mathbf{a} = \mathrm{e}X$, $\mathbf{b} = \mathrm{e}Y$. Тогда $\beta(\mathbf{a}, \mathbf{b}) = X^T B \overline{Y}$. Также $\beta(\mathbf{b}, \mathbf{a}) = Y^T B \overline{X}$ или (транспонируем матрицу 1×1) $\beta(\mathbf{b}, \mathbf{a}) = \overline{X}^T B^T Y = \overline{X^T B^* \overline{Y}} = \overline{X^T B \overline{Y}}$. Отсюда $\overline{\beta(\mathbf{b}, \mathbf{a})} = \beta(\mathbf{a}, \mathbf{b})$, что и требовалось. $\square$

Квадратичные формы

Определение. Пусть $\beta \in \mathcal{B}_{sym}(V)$. Отображение $k : V \to \mathbb{R}$ ($k : V \to \mathbb{C}$), заданное правилом $k(\mathbf{a}) = \beta(\mathbf{a}, \mathbf{a})$ называется *квадратичной (эрмитовой) формой* или *квадратичной функцией*, порожденной билинейной формой β.

Множество всех (эрмитовых) квадратичных форм обозначим $\mathcal{K}(V)$.

Определение дает возможность говорить о (эрмитово) симметричной матрице (эрмитовой) квадратичной форме. Координатная запись квадратичной формы имеет вид $k(\mathbf{a}) = X^T B \overline{X}$ или $k(\mathbf{a}) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ij} x_i \overline{x_j}$, где $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ — координатный столбец вектора $\mathbf{a}$. Заметим, что диагональные элементы матрицы B — это значения квадратичной формы на базисных векторах: $b_{ii} = k(\mathbf{e}_i)$.

Замечание. *В вещественном случае породить квадратичную форму можно было бы и произвольной (не обязательно симметричной) билинейной формой, однако это не изменило бы запас квадратичных форм: билинейная форма с матрицей B порождает ту же форму, что и симметричная билинейная форма с матрицей* $\dfrac{B+B^T}{2}$.

Теорема 2.2. *Данная (эрмитово) квадратичная форма k порождается ровно одной билинейной формой $\beta \in \mathcal{B}_{sym}(V)$.*

$\triangleright$ Достаточно показать, как значение $\beta(\mathbf{a},\mathbf{b})$ на заданных векторах $\mathbf{a}$, $\mathbf{b}$ определяется только по значениям квадратичной формы.

1) Доказательство для $\mathbb{R}$. Заметим, что
$k(\mathbf{a}+\mathbf{b}) = \beta(\mathbf{a}+\mathbf{b},\mathbf{a}+\mathbf{b}) = \beta(\mathbf{a},\mathbf{a}) + \beta(\mathbf{a},\mathbf{b}) + \beta(\mathbf{b},\mathbf{a}) + \beta(\mathbf{b},\mathbf{b}) = k(\mathbf{a}) + 2\beta(\mathbf{a},\mathbf{b}) + k(\mathbf{b})$, поэтому $\beta(\mathbf{a},\mathbf{b}) = \dfrac{k(\mathbf{a}+\mathbf{b}) - k(\mathbf{a}) - k(\mathbf{b})}{2}$.

2) Доказательство для . Имеем
$k(\mathbf{a}+\mathbf{b}) = \beta(\mathbf{a}+\mathbf{b},\mathbf{a}+\mathbf{b}) = \beta(\mathbf{a},\mathbf{a}) + \beta(\mathbf{a},\mathbf{b}) + \beta(\mathbf{b},\mathbf{a}) + \beta(\mathbf{b},\mathbf{b}) = k(\mathbf{a}) + \beta(\mathbf{a},\mathbf{b}) + \beta(\mathbf{b},\mathbf{a}) + k(\mathbf{b})$, отсюда

$$\beta(\mathbf{a},\mathbf{b}) + \beta(\mathbf{b},\mathbf{a}) = k(\mathbf{a}+\mathbf{b}) - k(\mathbf{a}) - k(\mathbf{b}).$$

Далее
$k(\mathbf{a}+i\mathbf{b}) = \beta(\mathbf{a}+i\mathbf{b},\mathbf{a}+i\mathbf{b}) = \beta(\mathbf{a},\mathbf{a}) - i\beta(\mathbf{a},\mathbf{b}) + i\beta(\mathbf{b},\mathbf{a}) + \beta(\mathbf{b},\mathbf{b}) = k(\mathbf{a}) - i(\beta(\mathbf{a},\mathbf{b}) - \beta(\mathbf{b},\mathbf{a})) + k(\mathbf{b})$, отсюда

$$\beta(\mathbf{a},\mathbf{b}) - \beta(\mathbf{b},\mathbf{a}) = ik(\mathbf{a}+i\mathbf{b}) + ik(\mathbf{a}) + ik(\mathbf{b}).$$

Складывая получаенные равенства, получаем выражение $\beta(\mathbf{a},\mathbf{b})$ только через значения $k(\mathbf{a}+\mathbf{b})$, $k(\mathbf{a}+i\mathbf{b})$, $k(\mathbf{a})$, $k(\mathbf{b})$. $\square$

Следствие. *Тождественно нулевая (эрмитова) квадратичная форма порождается лишь тождественно нулевой (эрмитово) симметричной формой.*

Теорема 2.2 показывает, что в определении фактически устанавливается взаимно-однозначное соответствие между множествами $\mathcal{B}_{sym}(V)$ и $\mathcal{K}(V)$. Ниже отождествляем эти множества.

§ 3. Диагональный и канонический вид квадратичной формы

Определение. Говорят, что (эрмитова) квадратичная форма k имеет *диагональный вид* в базисе е конечномерного пространства V, если матрица формы k в базисе е диагональна.

Определение. Диагональный вид (эрмитовой) квадратичной формы называется *каноническим*, если каждый диагональный элемент матрицы равен одному из чисел $1, 0, -1$.

Диагональный вид формы k в базисе е означает, что $k(\mathbf{a}) = \sum\limits_{i=1}^{n} d_i x_i \overline{x_i}$ или $k(\mathbf{a}) = \sum\limits_{i=1}^{n} d_i |x_i|^2$, где $d_i \in \mathbb{R}$. Заметим, что от диагонального вида легко перейти к каноническому, выполнив простую замену координат: $x_i = x_i'/\sqrt{|d_i|}$ для всех i с условием $d_i \neq 0$.

Докажем следующую основную теорему.

Теорема 3.1. *Пусть $\dim V = n < \infty$, $k \in \mathcal{K}(V)$. Тогда существует базис, в котором k имеет диагональный вид.*

$\triangleright$ Достаточно научиться выполнять двойные элементарные преобразования, приводящие к матрице B', у которой все элементы первой строки и первого столбца, за исключением возможно b'_{11} равны нулю. Далее с матрицей B' можно продолжить аналогичные двойные элементарные преобразования, не затрагивающие первые строку и столбец, и т.д.

1) Пусть $b_{11} \neq 0$. Последовательно проведем для всех $m = 2, 3, \ldots, n$, для которых $b_{m1} \neq 0$, следующие двойные элементарные преобразовния: вычтем из m-й строки 1-ю строку, умноженную

на b_{m1}/b_{11}, а затем вычтем из m-го столбца 1-й столбец, умноженный на $b_{1m}/b_{11} = \overline{b_{m1}}/b_{11}$. Двойное элементарное преобразование соответствует следующей операции с матрицами: $B \to C^T B \overline{C}$, где C — элементарная матрица, т.е. соответствует замене базиса. После указанной серии двойных элементарных преобразований все элементы перой строки и первого столбца, за исключением b_{11} станут равными нулю.

2) Пусть $b_{11} = 0$, но для некоторого m верно $b_{m1} \neq 0$. Тогда сведем ситуацию к случаю 1), предварительно выполнив следующее двойное элементарное преобразование: прибавим к 1-й строке m-ю строку, умноженную на $\overline{b_{m1}}$, а затем прибавим к 1-му столбцу m-й столбец, умноженный на $b_{m1} = \overline{b_{1m}}$. После этого в левом верхнем углу окажется число $2|b_{m1}|^2 \neq 0$. $\square$

Следствие. *Пусть* $\dim V = n < \infty$, $k \in \mathcal{K}(V)$. *Тогда существует базис, в котором* k *имеет канонический вид.*

Алгоритм, описанный в доказательстве теоремы 1.2, позволяет также вести «протокол», т.е. отслеживать преобразования базиса. Координатные столбцы исходного базиса образуют единичную матрицу (это исходная матрица перехода). Далее при каждом двойном преобразовании с текущей матрицей перехода S проделываем только столбцовое преобразование.

§ 4. Знакоопределенные формы. Индексы инерции

Определение. (Эрмитова) квадратичная форма k на пространстве V называется *положительно определенной*, если $\forall \mathbf{a} \in V$, $\mathbf{a} \neq \mathbf{0}$, выполнено $k(\mathbf{a}) > 0$.

Определение. (Эрмитова) квадратичная форма k на пространстве V называется *положительно полуопределенной*, если $\forall \mathbf{a} \in V$ выполнено $k(\mathbf{a}) \geqslant 0$.

Аналогично определяются *отрицательно определенные* и *отрицательно полуопределенные* формы. Конечно, существуют квадратичные формы, не принадлежащие ни к одному из определенных типов, например $x_1^2 - x_2^2$.

Упражнение. Форма k положительно определена и имеет в некотором базисе матрицу B. Может ли диагональный элемент матрицы B быть неположительным?

Следующая несложная теорема показывает, как выяснить по диагональному виду формы, является ли она положительно определенной, положительно полуопределенной, и т.д.

Теорема 4.1. *Пусть* $k \in \mathcal{K}(V)$, e — *базис в* V *и* $k \xrightarrow[\mathrm{e}]{} \operatorname{diag}(d_1, d_2, \ldots, d_n)$. *Тогда*
k *положительно определена* $\Leftrightarrow d_i > 0$ *для всех* $i = 1, \ldots, n$;
k *положительно полуопределена* $\Leftrightarrow d_i \geqslant 0$ *для всех* $i = 1, \ldots, n$;
k *отрицательно определена* $\Leftrightarrow d_i < 0$ *для всех* $i = 1, \ldots, n$;
k *отрицательно полуопределена* $\Leftrightarrow d_i \leqslant 0$ *для всех* $i = 1, \ldots, n$.

$\rhd$ Докажем для положительно определенных форм (в остальных случаях доказательство аналогично).

$\boxed{\Rightarrow}$ Имеем $d_i = k(\mathbf{e}_i) > 0$ (из опрделения положительной определенности).

$\boxed{\Leftarrow}$ Пусть все $d_i > 0$ и $\mathbf{a} \neq \mathbf{0}$ — произвольный вектор, $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ — его координатный столбец.

Тогда $k(\mathbf{a}) = \sum\limits_{i=1}^{n} d_i |x_i|^2 > 0$, поскольку хотя бы одна координата x_i ненулевая. $\square$

Следствие. *Определитель матрицы положительно определенной формы положителен.*

$\rhd$ Определитель матрицы положительно определенной формы в базисе, где она имеет диагональный вид, положителен. А значит, по следствию 2 из теоремы 1.2, определитель положиетелен и для произвольного базиса. $\square$

> **Определение.** *Положительным индексом инерции* квадратичной формы k называется наибольшее целое число p, для которого существует такое подпространство $U \leqslant V$, $\dim U = p$, что сужение $k\,|_U$ является положительно определенной формой.

Аналогично определяется отрицательный индекс инерции. Положительный и отрицательный индекс формы k обозначаем p (или $p(k)$) и q (или $q(k)$) соответственно. Очевидно, форма k положительно определена $\Leftrightarrow$ $p(k) = n$ (где $n = \dim V$), форма k положительно полуопределена $\Leftrightarrow$ $q(k) = 0$.

Теорема 4.2 (об индексах инерции). *Пусть $k \in \mathcal{K}(V)$, e — базис в V и $k \xrightarrow[\mathrm{e}]{} \mathrm{diag}(d_1, d_2, \ldots, d_n)$. Тогда $p(k)$ равно количеству положительных чисел среди $d_1, d_2, \ldots, d_n$, а $q(k)$ — количеству отрицательных чисел среди $d_1, d_2, \ldots, d_n$.*

$\triangleright$ Пусть p' — количество положительных среди чисел $d_1, \ldots, d_n$. Докажем, что $p = p'$, где $p = p(k)$. Для отрицательного индекса инерции рассуждения будут аналогичны.

Можно считать, что p' первых d_i положительны (этого можно добиться перестановкой векторов в базисе e), т.е. $d_1 > 0, \ldots, d_{p'} > 0$, $d_{p'+1} \leqslant 0, \ldots, d_n \leqslant 0$. Положим $U_{p'} = \langle \mathbf{e}_1, \ldots, \mathbf{e}_{p'} \rangle$, $W_{p'} = \langle \mathbf{e}_{p'+1}, \ldots, \mathbf{e}_n \rangle$. Заметим, что $k\,|_{U_{p'}}$ положительно определена, а $k\,|_{W_{p'}}$ отрицательно полуопределена. Так как $\dim U_{p'} = p'$, имеем $p \geqslant p'$.

Предположим, что $p > p'$, тогда рассмотрим подпространство $U \leqslant V$ такое, что $\dim U = p$ и $k\,|_U$ положительно определена. Имеем $\dim U + \dim W_{p'} = p + (n - p') > n$. Но $\dim(U + W_{p'}) \leqslant \dim V = n$, значит по теореме 3.3 главы 1 получаем $\dim(U \cap W_{p'}) > 0$, то есть найдется ненулевой вектор $\mathbf{a} \in U \cap W_{p'}$. Поскольку $\mathbf{a} \in U$, имеем $k(\mathbf{a}) > 0$, а так как $\mathbf{a} \in W_{p'}$, имеем $k(\mathbf{a}) \leqslant 0$. Противоречие. $\square$

Следствие. *Сумма индексов инерции $p(k) + q(k)$ равна рангу $\mathrm{rg}\, k$.*

Для выяснения, является ли данная форма положительно определенной, без приведения ее к диагональному виду, иногда применяется следующий критерий.

Теорема 4.3 (критерий Сильвестра). *Пусть $k \in \mathcal{K}(V)$, e — базис в V и $k \xrightarrow[\mathrm{e}]{} B$. Тогда k положительно определена $\Leftrightarrow$ $|B_i| > 0$ для всех $i = 1, 2, \ldots, n$.*

$\triangleright$ $\boxed{\Leftarrow}$ Если k положительно определена, то, очевидно $k\,|_U$ тоже положительно определена для любого подпространства $U \leqslant V$. Поскольку B_i — матрица сужения k на подпространство $U_i = \langle \mathbf{e}_1, \ldots, \mathbf{e}_i \rangle$ (см. предложение 1.2), достаточно применить следствие из теоремы 4.1.

$\boxed{\Rightarrow}$ Предположим противное и найдем минимальное m, для которого форма $\tilde{k} = k\,|_{U_m}$ не является положительно определенной (где $U_m = \langle \mathbf{e}_1, \ldots, \mathbf{e}_m \rangle$). По выбору m получаем, что форма $k\,|_{U_{m-1}} = \tilde{k}\,|_{U_{m-1}}$ положительно определена, значит $p(\tilde{k}) \geqslant m - 1$. Так как $p(\tilde{k}) < m$ (иначе $\tilde{k}$ была бы положительно определенной), имеем $p(\tilde{k}) = m - 1$. Пусть $D = \mathrm{diag}(d_1, \ldots, d_m)$ — диагональный вид (в некотором базисе) формы $\tilde{k}$. Тогда среди чисел $d_1, \ldots, d_m$ ровно $m - 1$ положительных — все кроме одного, значит $|D| = d_1 d_2 \ldots d_m \leqslant 0$. Но это противоречит условию $|B_m| > 0$, поскольку D и B_m — матрицы одной и той же формы $\tilde{k}$ в разных базисах (см. следствие 2 из теоремы 1.2). $\square$

Чтобы выяснить, является ли данная форма k отрицательно определенной, можно проверить на положительную определенность форму $(-k)$.

Упражнение. Пусть $k \in \mathcal{K}(V)$, e — базис в V и $k \xrightarrow[\mathrm{e}]{} B$. Тогда k положительно полуопределена $\Rightarrow$ $|B_i| \geqslant 0$ для всех $i = 1, 2, \ldots, n$. Обратное утверждение неверно.

Упражнение. Пусть $k \in \mathcal{K}(V)$, положительно определена. Тогда сумма всех элементов ее матрицы больше 0.

Упражнение. Пусть положительный индекс инерции формы k не равен 0. Докажите, что существует базис, в котором в матрице формы k все диагональные элементы положительные (а внедиагональные — какие угодно).

§ 5. Кососимметричные формы

В этом параграфе V — векторное пространство над $\mathbb{R}$.

Определение. Билинейная (полуторалинейная) форма β на пространстве V называется *кососимметричной*, если $\forall\ \mathbf{a}, \mathbf{b} \in V$

$$\beta(\mathbf{a}, \mathbf{b}) = -\beta(\mathbf{b}, \mathbf{a}).$$

Множество всех симметричных (эрмитовых) билинейных (полуторалинейных) форм на пространстве V обозначаем $\mathcal{B}_{Alt}(V)$.

Для $\beta \in \mathcal{B}_{Alt}(V)$, очевидно, $\beta(\mathbf{a}, \mathbf{a}) = 0$.

Теорема 5.1. *Пусть* $\dim V < \infty$, e *— базис в* V. *Пусть* $\beta \in \mathcal{B}(V)$, $\beta \xrightarrow[\mathrm{e}]{} B$. *Тогда* $\beta \in \mathcal{B}_{Alt}(V) \Leftrightarrow B^T = -B$.

$\triangleright$ Аналогично доказательству теоремы 2.1. $\square$

Теорема 5.2 (канонический вид)**.** *Пусть* $\dim V = n < \infty$, $\beta \in \mathcal{B}_{Alt}(V)$. *Тогда существует базис, в котором* β *имеет блочно-диагональный вид, в котором по диагонали встречаются блоки вида* (0) *или* $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

$\triangleright$ СХЕМА: ДВОЙНЫМИ ЭЛЕМЕНТАРНЫМИ ПРЕОБРАЗОВАНИЯМИ. $\square$

Глава 5

Евклидовы и унитарные пространства

§ 1. Скалярное произведение. Матрица Грама

Определения

Билинейную (эрмитово) симметричную положительно определенную функцию называют также *скалярным произведением.*

Определение. *Евклидовым пространством* называется векторное пространство над $\mathbb{R}$, в котором зафиксировано скалярное произведение.

Определение. *Унитарным пространством* называется векторное пространство над $\mathbb{C}$, в котором зафиксировано скалярное произведение.

Будем развивать теорию евклидовых и унитарных пространств параллельно (все случаи, когда имеются отличия, будем отмечать). В этой главе, если не оговорено противное, предполагается, что мы работаем в евклидовы (унитарные) пространстве $\mathcal{E}$.

Для обозначения скалярного произведения векторов $\mathbf{a}$ и $\mathbf{b}$ будем использовать $(\mathbf{a}, \mathbf{b})$ (опуская β в записи $\beta(\mathbf{a}, \mathbf{b})$).

Определение. *Длиной*, или *нормой* вектора $\mathbf{a}$ называют величину $\sqrt{(\mathbf{a}, \mathbf{a})}$.

Длину обозначают $|\mathbf{a}|$ или $||\mathbf{a}||$. Из положительной определенности скалярного произведения вытекает, что $\forall\, \mathbf{a} \in \mathcal{E}$ выполнено $|\mathbf{a}| \geqslant 0$, причем $|\mathbf{a}| = 0 \Leftrightarrow \mathbf{a} = \mathbf{0}$.

Операцию деления вектора на его длину (т.е. переходу к единичному вектору того же направления) иногда называют нормированием («отнормируем вектор»).

Определение. Говорят, что векторы $\mathbf{a}$ и $\mathbf{b}$ ортогональны, если $(\mathbf{a}, \mathbf{b}) = 0$.

Обозначение для ортогональности векторов: $\mathbf{a} \perp \mathbf{b}$. Нетрудно заметить, что существует единственный вектор, ортогональный любому вектору — это $\mathbf{0}$.

Заметим, что любое подпространство U евклидова (унитарного) пространства также естественным образом становится евклидовым (унитарным) (в качестве скалярного произведения на U выступает сужение скалярного произведения на объемлющем пространстве).

Матрица Грама

Определение. *Матрицей Грама* системы векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ называется матрица $(\gamma_{ij}) \in \mathbf{M}_{k\times k}$ такая, что $\gamma_{ij} = (\mathbf{a}_i, \mathbf{a}_j)$.

Таким образом, матрица Грама представляет собой «таблицу умножения» базисных векторов. Обозначение для матрицы Грама: $\Gamma(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)$.

Отметим, что понятие матрицы Грама не является абсолютно новым. Скажем, если $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ — базис, то матрица $\Gamma(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ совпадает с матрицей билинейной формы скалярного произведения в базисе e (см. определение из § **??** главы 4). Более общо, если $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ — линейно независимая система векторов, то $\Gamma(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)$ совпадает с матрицей скалярного произведения на подпространстве $U = \langle \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k \rangle$ (в базисе $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$).

Следующая теорема говорит о том, что зная матрицу Грама, можно вычислять скалярное произведение, а значит, длины и (как увидим далее) другие метрические характеристики.

Теорема 1.1 (Скалярное произведение). *Пусть* $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ *— базис в* $\mathcal{E}$*, и* $\Gamma = \Gamma(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ *— матрица Грама. Если векторы* $\mathbf{a}, \mathbf{b} \in \mathcal{E}$ *таковы, что* $\mathbf{a} = \mathrm{e}X$ *и* $\mathbf{b} = \mathrm{e}Y$*, то*

$$\boxed{(\mathbf{a}, \mathbf{b}) = X^T \Gamma \overline{Y}.}$$

$\triangleright$ Это частный случай теоремы из главы 4. $\square$

Предложение 1.1. *Если система векторов* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ *линейно независима, то* $|\Gamma(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)| > 0$*; иначе* $|\Gamma(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)| = 0$.

$\triangleright$ 1) Пусть система векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ линейно независима. Тогда, как было отмечено выше, $\Gamma(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)$ — это матрица скалярного произведения на подпространстве $U = \langle \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k \rangle$ (в базисе $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$). То есть $\Gamma(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)$ — это матрица положительно определенной билинейной симметричной формы, значит, $|\Gamma(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)| > 0$ по следствию из теоремы 4.1 главы 4.

2) Пусть система векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ линейно зависима, т.е. $\sum\limits_{i=1}^{n} \lambda_i \mathbf{a}_i = \mathbf{0}$ для некоторого ненулевого столбца $\lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$. Домножив это равенство скалярно на каждый из векторов $\mathbf{a}_j$, $j = 1, \ldots, k$, и воспользовавшись линейностью, получим $\sum\limits_{i=1}^{n} \lambda_i (\mathbf{a}_i, \mathbf{a}_j) = 0$. Получается, что столбец λ является нетривиальным решением системы линейных уравнений с (квадратной) матрицей коэффициентов $\Gamma(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)^T$. Значит, эта матрица вырожденная, откуда $|\Gamma(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)| = 0$. $\square$

На самом деле геометрический смысл (для случая $\mathbb{R}$) определителя $|\Gamma(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)|$ — это квадрат объема k-мерного параллелепипеда, построенного на векторах $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$).

Следствие 1 (Неравенство Коши-Буняковского-Шварца (КБШ)). $\forall\ \mathbf{a}, \mathbf{b} \in \mathcal{E}$ *выполнено*

$$|\mathbf{a}| \cdot |\mathbf{b}| \geqslant |(\mathbf{a}, \mathbf{b})|.$$

$\triangleright$ Неравенство $|\Gamma(\mathbf{a}, \mathbf{b})| \geqslant 0$ имеет вид $(\mathbf{a}, \mathbf{a})(\mathbf{b}, \mathbf{b}) - (\mathbf{a}, \mathbf{b})(\mathbf{b}, \mathbf{a}) \geqslant 0 \Leftrightarrow (\mathbf{a}, \mathbf{a})(\mathbf{b}, \mathbf{b}) - (\mathbf{a}, \mathbf{b})\overline{(\mathbf{a}, \mathbf{b})} \geqslant 0 \Leftrightarrow |\mathbf{a}|^2 \cdot |\mathbf{b}|^2 - |(\mathbf{a}, \mathbf{b})|^2 \geqslant 0$. $\square$

Отметим, что неравенство КБШ позволяет корректно определить угол между векторами Евклидова пространства.

Следствие 2 (Неравенство треугольника). $\forall\ \mathbf{a}, \mathbf{b} \in \mathcal{E}$ *выполнено*

$$|\mathbf{a}| + |\mathbf{b}| \geqslant |\mathbf{a} + \mathbf{b}|.$$

$\triangleright$ Имеем $|\mathbf{a} + \mathbf{b}|^2 = (\mathbf{a} + \mathbf{b}, \mathbf{a} + \mathbf{b}) = |\mathbf{a}|^2 + |\mathbf{b}|^2 + (\mathbf{a}, \mathbf{b}) + (\mathbf{b}, \mathbf{a}) =$
$= |\mathbf{a}|^2 + |\mathbf{b}|^2 + (\mathbf{a}, \mathbf{b}) + \overline{(\mathbf{a}, \mathbf{b})} \leqslant |\mathbf{a}|^2 + |\mathbf{b}|^2 + 2|(\mathbf{a}, \mathbf{b})|$, что по неравенству КБШ не больше чем $|\mathbf{a}|^2 + |\mathbf{b}|^2 + 2|(\mathbf{a}| \cdot |\mathbf{b})| = (|\mathbf{a}| + |\mathbf{b}|)^2$. $\square$

II. *Стандартное* скалярное произведение на пространстве столбцов высоты n определяется как $(\mathbf{a}, \mathbf{b}) = \sum\limits_{i=1}^{n} x_i \overline{y_i} = X^T \overline{Y}$, где $\mathbf{a} = \mathrm{e}X$, $\mathbf{b} = \mathrm{e}Y$.

III. В пространстве $C[a, b]$ непрерывных (комплекснозначных) функций, опеределенных на отрезке $[a, b]$, можно задать скалярное произведение как
$(f, g) = \int\limits_{a}^{b} f(x)\overline{g(x)}\, dx$. варианты с весами. ПРИМЕРЫ - к БИЛИНЕЙНЫМИ ФУНКЦИЯМ

§ 2. Ортогональные системы векторов. Ортогональное дополнение. Ортогонализация

Ортогональные системы векторов. ОНБ

Определение. Система векторов $\mathbf{a}_1, \ldots, \mathbf{a}_k$ в евклидовом (унитарном) пространстве называется *ортогональной*, если $\mathbf{a}_i \perp \mathbf{a}_j$ для всех $1 \leqslant i < j \leqslant k$.

Определение. Ортогональная система векторов $\mathbf{a}_1, \ldots, \mathbf{a}_k$ называется *ортонормированной*, если $|\mathbf{a}_i| = 1$ для всех $1 \leqslant i \leqslant k$.

В частности, можно говорить об ортогональных базисах и ортонормированных базисах (далее используем сокращение ОНБ).

Предложение 2.1 (теорема Пифагора)**.** *Пусть* $\mathbf{a}_1, \ldots, \mathbf{a}_k$ *— ортогональная система векторов. Тогда* $|\mathbf{a}_1 + \ldots + \mathbf{a}_k|^2 = |\mathbf{a}_1|^2 + \ldots + |\mathbf{a}_k|^2$.

$\rhd$ $|\mathbf{a}_1 + \ldots + \mathbf{a}_k|^2 = (\mathbf{a}_1 + \ldots + \mathbf{a}_k, \mathbf{a}_1 + \ldots + \mathbf{a}_k)|$. Раскрывая по линейности с учетом того, что $(\mathbf{a}_i, \mathbf{a}_j) = 0$ при $i \neq j$, получаем $(\mathbf{a}_1, \mathbf{a}_1) + \ldots + (\mathbf{a}_k, \mathbf{a}_k) = |\mathbf{a}_1|^2 + \ldots + |\mathbf{a}_k|^2$. $\Box$

Отметим следующий почти очевидный критерий ортогональности в терминах матрицы Грама.

Предложение 2.2. *Система векторов* $\mathbf{a}_1, \ldots, \mathbf{a}_k$ *является ортогональной* $\Leftrightarrow \Gamma(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ *— диагональная матрица.*
Система векторов $\mathbf{a}_1, \ldots, \mathbf{a}_k$ *является ортонормированной* $\Leftrightarrow \Gamma(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ *— единичная матрица.*

$\rhd$ Сразу следует из определения матрицы Грама. $\Box$

Следствие 1. *Ортогональная система ненулевых векторов линейно независима.*

$\rhd$ Следует из предыдущего предложения с учетом предложения 1.1. $\Box$

Следствие 2. *Для конечномерного евклидова пространства: ортогональный базис — это базис, в котором форма скалярного произведения имеет диагональный вид, ОНБ — это базис, в котором форма скалярного произведения имеет канонический вид.*

Следствие 3. *В конечномерном евклидовом пространстве существет ОНБ.*

$\rhd$ Ввиду следствия 2, это частный случай следствия из теоремы 3.1 главы 4. $\Box$

Следствие 4. *Пусть* $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ *— ОНБ в* $\mathcal{E}$*. Если векторы* $\mathbf{a}, \mathbf{b} \in \mathcal{E}$ *таковы, что* $\mathbf{a} = \mathrm{e}X$ *и* $\mathbf{b} = \mathrm{e}Y$*, то*

$$\boxed{(\mathbf{a}, \mathbf{b}) = X^T \overline{Y}.}$$

$\rhd$ Это утверждение — частный случай теоремы 4.1. $\Box$

Переход от ОНБ к ОНБ. Ортогональные и унитарные матрицы

> **Определение.** Комплексная матрица Q размера $n \times n$ называется *унитарной*, если $Q^*Q = E$.

> **Определение.** Вещественная матрица Q размера $n \times n$ называется *ортогональной*, если $Q^TQ = E$.

Множество всех ортогональных и унитарных матриц $n \times n$ обозначаем соответственно O_n и U_n. Сравнивая определения, получаем $O_n \subset U_n$ и более того, $O_n = U_n \cap \mathbf{M}_{n\times n}(\mathbb{R})$ (т.е. ортогональные матрицы это в точной вещественные унитарные матрицы).

Предложение 2.3. *Для комплексной матрицы Q размера $n \times n$ следующие условия эквивалентны:*
1) $Q \in U_n$;
2) $\exists\, Q^{-1}$ и $Q^{-1} = Q^*$;
3) столбцы матрицы Q образуют ОНБ в унитарном пространстве столбцов $\mathbb{C}^n = \mathbf{M}_{n\times 1}$, наделенном стандартным скалярным произведением $(X, Y) = X^T\overline{Y}$.

▷ Очевидно, 1) ⇔ 2).
Пусть $Q_1, \dots, Q_n$ — столбцы матрицы Q. Равеноство $Q^*Q = E$ означает, что $\overline{Q_i^T}Q_j = \delta_{ij}$ или $Q_i^T\overline{Q_j} = \delta_{ij}$. Значит, 1) ⇔ 3). □

Видим, в частности, что обращать ортогональную матрицу легко: достаточно ее транспонировать. На практике свойство 3), возможно, наиболее просто для проверки, является ли матрица Q ортогональной (унитарной).

Предложение 2.4. *Пусть $Q \in U_n$. Тогда $Q^T \in U_n$, $\overline{Q} \in U_n$, $Q^* \in U_n$.*

▷ □

Видим, в частности, что строки ортогональной (унитарной) матрицы — тоже ОНБ в пространстве строк со стандартным скалярным произведением.

Предложение 2.5 (Групповое свойство)**.** *Пусть $Q, R \in U_n$. Тогда $QR \in U_n$, $Q^{-1} \in U_n$.*

▷ □

Предложение 2.6. *Пусть $Q \in U_n$. Тогда $|\det Q| = 1$.*

▷ Имеем $\det(Q^*Q) = \det E = 1$. Отсюда $\det Q^* \cdot \det Q = 1 \Leftrightarrow \overline{\det Q^T} \cdot \det Q = 1 \Leftrightarrow \overline{\det Q} \cdot \det Q = 1 \Leftrightarrow |\det Q|^2 = 1$. □

Имеется следующая связь между ортогональными (унитарными) матрицами и переходом от ОНБ к ОНБ.

Теорема 2.1. *Пусть $\dim \mathcal{E} = n < \infty$, $\mathrm{e} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ — ОНБ в $\mathcal{E}$, $\mathrm{e}' = (\mathbf{e}'_1, \dots, \mathbf{e}'_n)$ — некоторый базис в $\mathcal{E}$. Пусть S — матрица переходв от базиса e к базису e'. Тогда*
Тогда φ является ортогональным ⇔ S — ортогональная (унитарная) матрица.

▷ Это сразу следует из определения матрицы перехода и условия 3) в предложении 2.3. □

Таким образом, ортогональные (унитарные) матрицы — в точности матрицы перехода от ОНБ к ОНБ. Иногда именно в теорминах матрицы перехода естественно интерпретировать ортогональные (унитарные) матрицы. Например, можно дать другое доказательство предожения 2.5. УПР.?

Ортогональные подпространства

> **Определение.** Множества U_1 и U_2 евклидова (унитарного) пространства называются *ортогональными*, если $\forall\, \mathbf{a}_1 \in U_1$ и $\forall\, \mathbf{a}_2 \in U_2$ выполнено $\mathbf{a}_1 \perp \mathbf{a}_2$.

Ортогональность двух векторов получается как частный случай (вектор считаем одноэлементным подмножеством). Обозначение для ортогональности прежнее, например $\mathbf{a} \perp U$ — значит вектор $\mathbf{a}$ ортогонален любому вектору из U. Чаще нас будут интересовать ортогональные подпространства. Почти очевидно, что ортогональные подпространства пересекаются тривиально. Более общий факт дает следующая теорема.

Теорема 2.2. *Пусть $U_1, \ldots, U_k$ — попарно ортогональные подпространства евклидова (унитарного) пространства $\mathcal{E}$. Тогда сумма $U_1 + \ldots + U_k$ является прямой суммой.*

$\triangleright$ Предположим противоное, пусть скажем найдется ненулевой $\mathbf{a} \in U_1 \cap (U_2 + \ldots + U_k)$ (пользуемся критерием из теоремы ... главы....). Тогда $\mathbf{a} = \mathbf{a}_2 + \ldots + \mathbf{a}_k$, где $\mathbf{a}_i \in U_i$. Перенесем все слагаемые в левую часть и вычеркнем нулевые векторы, тогда получем, что сумма ненулевых попарно ортогональных векторов равна $\mathbf{0}$. Это противоречит следствию 1 из предложения **??**. $\Box$

Предложение 2.7 (признак ортогональности). *Пусть $U_1 = \langle \mathbf{a}_1, \ldots, \mathbf{a}_k \rangle$, $U_2 = \langle \mathbf{b}_1, \ldots, \mathbf{b}_l \rangle$. Тогда $U_1 \perp U_2 \Leftrightarrow \mathbf{a}_i \perp \mathbf{b}_j$, $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, l$.*

$\triangleright$ $\boxed{\Rightarrow}$ Очевидно из определения ортогональности подпространств.

$\boxed{\Leftarrow}$ Пусть $\mathbf{a} \in U_1$, $\mathbf{b} \in U_2$. Тогда существуют разложения $\mathbf{a} = \sum\limits_{i=1}^{k} \alpha_i \mathbf{a}_i$, $\mathbf{b} = \sum\limits_{j=1}^{l} \beta_j \mathbf{b}_j$. Тогда из линейности скалярного произведения получаем $(\mathbf{a}, \mathbf{b}) = \sum\limits_{i=1}^{k} \sum\limits_{j=1}^{l} \alpha_i \overline{\beta_j} (\mathbf{a}_i, \mathbf{b}_j) = 0$, т.е. $\mathbf{a} \perp \mathbf{b}$. $\Box$

Ортогональное дополнение. Ортогональная проекция

> **Определение.** *Ортогональным дополнением* подпространства U (в евклидовом (унитарном) пространстве $\mathcal{E}$) называется множество всех векторов, ортогональных U.

Обозначение для ортогонального дополнения: $U^{\perp}$. Итак, $U^{\perp} = \{\mathbf{a} \,|\, \mathbf{a} \perp U\}$. Очевидно, $U \perp U^{\perp}$, значит согласно теореме 2.2, имеем $U \bigoplus U^{\perp}$. В случае $\dim U < \infty$ можно усилить так.

Теорема 2.3. *Пусть $U \leqslant \mathcal{E}$, $\dim U = k < \infty$. Тогда*

$\boxed{U \bigoplus U^{\perp} = \mathcal{E}}$.

$\triangleright$ Представим произвольный вектор $\mathbf{a} \in \mathcal{E}$ в виде суммы $\mathbf{a}_1 + \mathbf{a}_2$, где $\mathbf{a}_1 \in U$, $\mathbf{a}_1 \perp U$.

Возьмем в U ортогональный базис $\mathbf{b}_1, \ldots, \mathbf{b}_k$ и представим произвольный вектор $\mathbf{a} \in V$ в виде $\mathbf{a} = \alpha_1 \mathbf{b}_1 + \ldots + \alpha_k \mathbf{b}_k + \mathbf{c}$. Достаточно подобрать коэффициенты $\alpha_1, \ldots, \alpha_k$ так, чтобы $\mathbf{c} = \mathbf{a} - \alpha_1 \mathbf{b}_1 - \ldots - \alpha_k \mathbf{b}_k$ был ортогонален U. Последнее равносильно (см. предложение 3.1) тому, что $(\mathbf{c}, \mathbf{b}_i) = 0$, $i = 1, \ldots, k$. Ввиду ортогональности $\mathbf{b}_i \perp \mathbf{b}_j$, $i \neq j$, имеем $(\mathbf{c}, \mathbf{b}_i) = 0 \Leftrightarrow (\mathbf{a}, \mathbf{b}_i) - \alpha_i (\mathbf{b}_i, \mathbf{b}_i) = 0 \Leftrightarrow \alpha_i = \dfrac{(\mathbf{a}, \mathbf{b}_i)}{(\mathbf{b}_i, \mathbf{b}_i)}$. $\Box$

Следствие 1. *Если $\dim \mathcal{E} = n < \infty$, $U \leqslant \mathcal{E}$ и $\dim U = k$, то $\dim U^{\perp} = n - k$.*

Следствие 2. *Если $\dim \mathcal{E} = n < \infty$ и $U \leqslant \mathcal{E}$, то $(U^{\perp})^{\perp} = U$.*

$\triangleright$ Так как $U \perp U^{\perp}$, то $U \subset (U^{\perp})^{\perp}$. Но по следствию 1, $\dim U = \dim (U^{\perp})^{\perp} = n - \dim U^{\perp}$. Значит (см....), $(U^{\perp})^{\perp} = U$. $\Box$

Упражнение. Для $U_i \leqslant \mathcal{E}$, $\dim U_i < \infty$, $i = 1, 2$ докажите, что $(U_1 + U_2)^{\perp} = U_1^{\perp} \cap U_2^{\perp}$.

Следствие 3. *Пусть $\dim \mathcal{E} = n < \infty$. Тогда данную ортогональную систему ненулевых векторов можно дополнить до ортогонального базиса.*

$\rhd$ Пусть $\mathbf{b}_1, \ldots, \mathbf{b}_k$ — данная ортогональная система, положим $U = \langle \mathbf{b}_1, \ldots, \mathbf{b}_k \rangle$. В силу 2.2, $\mathbf{b}_1, \ldots, \mathbf{b}_k$ — ортогональный базис в U. Пусть $\mathbf{b}_{k+1}, \ldots, \mathbf{b}_n$ — ортогональный базис в $U^\perp$. Тогда $\mathbf{b}_1, \ldots, \mathbf{b}_k, \mathbf{b}_{k+1}, \ldots, \mathbf{b}_n$ — ортогональный базис в $\mathcal{E}$. $\square$

Формула $U \bigoplus U^\perp = \mathcal{E}$ оправдывает термин «ортогональное дополнение». Это частный случай прямого дополнения (см. ....). В таком случае упростим терминологию и обозначения.

> **Определение.** *Ортогональной проекцией* вектора $\mathbf{a}$ на подпространство $U \leqslant \mathcal{E}$ называется проекция $\mathbf{a}$ на U вдоль $U^\perp$.

Обозначение для ортогональной проекции: $\operatorname{pr}_U \mathbf{a}$

Предложение 2.8 (формула проекции). *Пусть $\mathbf{a} \in \mathcal{E}$, $U \leqslant \mathcal{E}$ и $\mathbf{b}_1, \ldots, \mathbf{b}_k$ — ортогональный базис в U. Тогда*

$$\boxed{\operatorname{pr}_U \mathbf{a} = \sum_{i=1}^{k} \frac{(\mathbf{a}, \mathbf{b}_i)}{(\mathbf{b}_i, \mathbf{b}_i)} \mathbf{b}_i.}$$

$\rhd$ В доказательстве теоремы 2.3 мы уже нашли нужное выражение для $\operatorname{pr}_U \mathbf{a}$ (в виде $\alpha_1 \mathbf{b}_1 + \ldots + \alpha_k \mathbf{b}_k$, где $\alpha_i = \dfrac{(\mathbf{a}, \mathbf{b}_i)}{(\mathbf{b}_i, \mathbf{b}_i)}$. $\square$

Проекции могут естественно возникать в задачах на экстремум:

Предложение 2.9 (расстояние до подпространства). *Пусть $U \leqslant \mathcal{E}$, $\dim U < \infty$. Тогда $\min\limits_{\mathbf{x} \in U} |\mathbf{a} - \mathbf{x}| = |\operatorname{pr}_{U^\perp} \mathbf{a}|$.*

$\rhd$ Представим $\mathbf{a}$ как $\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_2$, где $\mathbf{a}_1 = \operatorname{pr}_U \mathbf{a}$, $\mathbf{a}_2 = \operatorname{pr}_{U^\perp} \mathbf{a}$. Тогда для $\mathbf{x} \in U$ имеем $|\mathbf{a} - \mathbf{x}|^2 = |\mathbf{a}_1 + \mathbf{a}_2 - \mathbf{x}|^2 = |(\mathbf{a}_1 - \mathbf{x}) + \mathbf{a}_2|^2$. Так как $(\mathbf{a}_1 - \mathbf{x}) \perp \mathbf{a}_2$, по теореме Пифагора $|(\mathbf{a}_1 - \mathbf{x}) + \mathbf{a}_2|^2 = |\mathbf{a}_1 - \mathbf{x}|^2 + |\mathbf{a}_2|^2 \geqslant |\mathbf{a}_2|^2$. При этом неравенство обращается в равенство при $\mathbf{x} = \mathbf{a}_1$. $\square$

Следующее предложение показывает, что при работе в с координатами в ОНБ очень легко получать описание ортогонального дополнения.

Предложение 2.10 (Ортогональное дополнение в координатах). *Пусть* e *— ОНБ в $\mathcal{E}$, $U = \langle \mathbf{a}_1, \ldots, \mathbf{a}_k \rangle$, где $\mathbf{a}_1, \ldots, \mathbf{a}_k$ — векторы с с координатными столбцами $X_1, \ldots, X_k$ в ОНБ* e*, Φ — матрица со столбцами X_i: $\Phi = (X_1 \ \ldots \ X_k)$. Тогда $U^\perp$ задается (в коодинатах в ОНБ* e*) как* $\operatorname{Sol}(\Phi^* X = O)$.

$\rhd$ Система $\Phi^* X = O$ состоит из уравнений вида $\overline{X_i^T} X = 0$. Эти уравнения означают, что вектор $\mathbf{x}$ с координатным столбцом X ортогоналей $\mathbf{a}_i$, $i = 1, \ldots, k$. В силу предложения 3.1, это эквивалентно условию $\mathbf{b} \perp U$. $\square$

Следствие (Теорема Фредгольма для СЛУ). *Пусть дана СЛУ $AX = b$ (с матрицей коэффициентов A размера $m \times n$. Тогда $AX = b$ совместна $\Leftrightarrow$ $\forall Y_0 \in \operatorname{Sol}(A^* Y = O)$ выполнено «условие ортогональности»: $b^* Y = 0$.*

$\rhd$ Рассмотрим пространство $\mathcal{E} = \mathbb{M}_{m \times 1}$ столбцов высоты m со стандартным скалярным произведением. Столбцы $a_{\bullet 1}, \ldots, a_{\bullet n}$ матрицы A и столбец b лежат в $\mathcal{E}$. Условие совместности системы $AX = b$ эквивалентно тому, что $b \in U$, где $U = \langle a_{\bullet 1}, \ldots, a_{\bullet n} \rangle$. Далее, $U^\perp = \operatorname{Sol}(A^* Y = O)$, поэтому «условие ортогональности» $\forall Y_0 \in \operatorname{Sol}(A^* Y = O)$ эквивалентно тому, что $b \perp U^\perp$. Но очевидно, $b \in U \Leftrightarrow b \perp U^\perp$. $\square$

Ортогонализация

Пусть изначально подпространство задано как линейная оболочка произвольной конечной системы векторов: $U = \langle \mathbf{a}_1, \ldots, \mathbf{a}_k \rangle$, а требуется найти ортогональный (или ОНБ) бизис в U (например, чтобы после этого была возможность использовать формулу для $\operatorname{pr}_U \mathbf{a}$). Фактически это та же процедура нахождения базиса, в котором форма скалярного произведения имеет диагональный (канонический вид). Но ввиду важности этого частного случая, дадим несколько другое описание прцедуры (с геометрической точки зрания). Алгоритм называется *ортогонализацей Грама-Шмитда.*

Положим $U_i = \langle \mathbf{a}_1, \ldots, \mathbf{a}_i \rangle$, так что $U_{i+1} = U_i \cup \{\mathbf{a}_{i+1}\}$. Идея следующая: будет по очереди «поправлять» векторы $\mathbf{a}_i$, заменяя на $\mathbf{b}_i = \operatorname{pr}_{U_{i-1}^{\perp}} \mathbf{a}_i$ так, что $\mathbf{b}_i \perp U_{i-1}$ и $U_i = \langle \mathbf{b}_1, \ldots, \mathbf{b}_i \rangle$.

Получим явные формулы. Первый ненулевой вектор $\mathbf{a}_t$ из списка $\mathbf{a}_1, \ldots, \mathbf{a}_k$ является ортогональным базисом в U_t. Пусть на некотором шаге мы уже имеем ортогональный базис $\mathbf{b}_1, \ldots, \mathbf{b}_\ell$ в U_m. Заменяем $\mathbf{a}_{m+1}$ на вектор $\mathbf{a}_{m+1} - \operatorname{pr}_{U_m} \mathbf{a}_{m+1}$, т.е. на $\mathbf{a}_{m+1} - \sum\limits_{i=1}^{\ell} \dfrac{(\mathbf{a}_{m+1}, \mathbf{b}_i)}{(\mathbf{b}_i, \mathbf{b}_i)} \mathbf{b}_i$. Если последний вектор нулевой (это соответсвует случаю $\mathbf{a}_{m+1} \in U_m$), то пропустим его, если он ненулевой, то объявим $\mathbf{b}_{\ell+1} = \mathbf{a}_{m+1} - \sum\limits_{i=1}^{\ell} \dfrac{(\mathbf{a}_{m+1}, \mathbf{b}_i)}{(\mathbf{b}_i, \mathbf{b}_i)} \mathbf{b}_i$, тем самым достраивая ортогональный базис в U_{m+1}.

§ 3. Сопряженное преобразование. Самосопряженные преобразования.

Сопряженное преобразование

Рассмотрим линейное преобразовние $\varphi : \mathcal{E} \to \mathcal{E}$.

> **Определение.** Линейное преобразование $\psi : \mathcal{E} \to \mathcal{E}$ называют *сопряженным* преобразованию φ, если $\forall\ \mathbf{a}, \mathbf{b} \in \mathcal{E}$ выполнено
>
> $$\boxed{(\varphi(\mathbf{a}), \mathbf{b}) = (\mathbf{a}, \psi(\mathbf{b}))}.$$

Обозначение для сопряженного преобразования: φ^*. Из определения неясно, существует ли φ^* и единственно ли оно. Этот недостаток определения будет устранен после следующего предложения.

Предложение 3.1. *Пусть* $\dim \mathcal{E} < \infty$, e *— ОНБ в* $\mathcal{E}$. *Пусть* $\varphi, \psi \in L(\mathcal{E}, \mathcal{E})$, $\varphi \underset{\text{e,e}}{\longrightarrow} A$, $\psi \underset{\text{e,e}}{\longrightarrow} B$. *Тогда*

$$\psi = \varphi^* \Leftrightarrow \boxed{B = A^*}.$$

$\triangleright$ Пусть $\mathbf{a} = \mathrm{e}X$, $\mathbf{b} = \mathrm{e}Y$. Имеем: $(\varphi(\mathbf{a}), \mathbf{b}) = (AX)^T \overline{Y} = X^T A^T \overline{Y} = \beta_1(\mathbf{a}, \mathbf{b})$, где β_1 — билинейная (полуторалинейная) форма с матрицей A^T.
$(\mathbf{a}, \psi(\mathbf{b})) = X^T \overline{BY} = X^T \overline{BY} = \beta_2(\mathbf{a}, \mathbf{b})$, где β_2 — билинейная (полуторалинейная) форма с матрицей $\overline{B}$.
Теперь утверждение теоремы означает, что равноство форм $\beta_1 = \beta_2$ эквивалентно равенству их матриц (в одном базисе). $\square$

Упражнение. Сформулируйте аналог предыдущего предложения в произвольном базисе с матрицей Грама Г.

Следствие 1. *Пусть* $\dim \mathcal{E} < \infty$. *Для данного* $\varphi \in L(\mathcal{E}, \mathcal{E})$ *существует и единственное* φ^*.

Следствие 2. *Пусть* $\dim \mathcal{E} < \infty$, $\varphi, \psi \in L(\mathcal{E}, \mathcal{E})$. *Тогда*
1) $(\varphi^*)^* = \varphi$;
2) $(\varphi\psi)^* = \psi^*\varphi^*$; *3)* $\operatorname{rg}(\varphi) = \operatorname{rg}(\varphi^*)$; *4)* $\overline{\chi_\varphi(\lambda)} = \chi_{\varphi^*}(\overline{\lambda})$.

$\triangleright$ Введем ОНБ e и перейдем к матрицам преобразования в этом ОНБ. Свойства 1) — 3) сразу следуют из соответсвующих свойств для матриц.
4) $\overline{\chi_\varphi(\lambda)} = \overline{|A-\lambda E|}$ $\chi_{\varphi^*}(\overline{\lambda}) = |A^* - \overline{\lambda}E| = |\overline{A^T - \lambda E}| = \overline{|(A-\lambda E)^T|} = \overline{|A-\lambda E|}$. $\square$

Теорема 3.1. *Пусть* $U \leqslant \mathcal{E}$, $\dim\mathcal{E} < \infty$, $\varphi \in L(\mathcal{E},\mathcal{E})$. *Тогда*
U *инвариантно относительно* $\varphi \Leftrightarrow U^\perp$ *инвариантно относительно* φ^*.

$\triangleright$ $\boxed{\Rightarrow}$ Пусть $\mathbf{b} \in U^\perp$. Требуется понять, что $\varphi^*(\mathbf{b}) \in U^\perp$, то есть что $\forall\ \mathbf{a} \in U$ выполнено $(\mathbf{a}, \varphi^*(\mathbf{b})) = 0$. Но $(\mathbf{a}, \varphi^*(\mathbf{b})) = (\varphi(\mathbf{a}), \mathbf{b}) = 0$, так как $\varphi(\mathbf{a}) \in U$ ввиду инвариантности U относительно φ.
$\boxed{\Leftarrow}$ Аналогично ввиду $(\varphi^*)^* = \varphi$ и $(U^\perp)^\perp = U$. $\square$

Теорема 3.2 (Теорема Фредгольма). *Пусть* $\dim\mathcal{E} = n < \infty$, $\varphi \in L(\mathcal{E},\mathcal{E})$. *Тогда*

$$\boxed{\operatorname{Ker}\varphi^* = (\operatorname{Im}\varphi)^\perp}.$$

$\triangleright$ Во-первых можно заметить, что подпространства $\operatorname{Ker}\varphi^*$ и $(\operatorname{Im}\varphi)^\perp$ имеют равные размерности. Действительно, по теор..... $\dim\operatorname{Ker}\varphi^* = n - \operatorname{rg}\varphi^* = n - \operatorname{rg}\varphi$, а по .... $\dim(\operatorname{Im}\varphi)^\perp = n - \dim(\operatorname{Im}\varphi) = n - \operatorname{rg}\varphi$.

Значит, согласно ....., достаточно доказать включение $\operatorname{Ker}\varphi^* \in (\operatorname{Im}\varphi)^\perp$. Пусть $\mathbf{b} \in \operatorname{Ker}\varphi^*$. Покажем, что $\mathbf{b} \in (\operatorname{Im}\varphi)^\perp$, т.е. $\forall\ \mathbf{c} \in \operatorname{Im}\varphi$ выполнено $\mathbf{c} \perp \mathbf{b}$. Поскольку $\mathbf{c} \in \operatorname{Im}\varphi$, найдем $\mathbf{a} \in \mathcal{E}$ такой, что $\mathbf{c} = \varphi(\mathbf{a})$. Тогда $(\mathbf{c}, \mathbf{b}) = (\varphi(\mathbf{a}), \mathbf{b}) = (\mathbf{a}, \varphi^*(\mathbf{b})) = (\mathbf{a}, \mathbf{0}) = 0$, откуда $\mathbf{c} \perp \mathbf{b}$, что и требовалось. $\square$

Самосопряженные преобразования

Определение. Линейное преобразование $\varphi : \mathcal{E} \to \mathcal{E}$ называется *самосопряженным*, если $\varphi^* = \varphi$.

Иначе говоря, $\varphi \in L(\mathcal{E},\mathcal{E})$ самосопряженное $\Leftrightarrow \forall\ \mathbf{a}, \mathbf{b} \in \mathcal{E}$ выполнено $(\varphi(\mathbf{a}), \mathbf{b}) = (\mathbf{a}, \varphi(\mathbf{b}))$. Тогда нетрудно заметить, что если U — инвариантное подпространство для самосопряженного φ, то сужение $\varphi\,|_U\colon U \to U$ тоже является самосопряженным преобразованием.

Теорема 3.3. *Пусть* $\varphi \in L(\mathcal{E},\mathcal{E})$, e — *ОНБ*, $\varphi \xrightarrow[\text{e,e}]{} A$. *Тогда* φ — *самосопряженное* $\Leftrightarrow A^* = A$.

$\triangleright$ Это частный случай предложения 3.1. $\square$

Теорема 3.4. *Если* $\varphi \in L(\mathcal{E},\mathcal{E})$ — *самосопряженное преобразование, то все его характеристические числа вещественные.*

$\triangleright$ 1) Докажем вначале утверждение для унитарного пространства (над $\mathbb{C}$). Пусть λ_0 — характеристическое число, а $\mathbf{a}$ — собственный вектор, соответствующий собственному значению λ_0. Запишем равенство $(\varphi(\mathbf{a}), \mathbf{a}) = (\mathbf{a}, \varphi(\mathbf{a}))$ (оно следует из определения самосопряженного преобразования). Так как $\varphi(\mathbf{a}) = \lambda_0\mathbf{a}$, то $(\lambda_0\mathbf{a}, \mathbf{a}) = (\mathbf{a}, \lambda_0\mathbf{a})$, откуда $\lambda_0(\mathbf{a}, \mathbf{a}) = \overline{\lambda_0}(\mathbf{a}, \mathbf{a}) \Rightarrow \lambda_0 = \overline{\lambda_0} \Rightarrow \lambda_0 \in \mathbb{R}$.
2) Покажем, как случай евклидова пространства (над $\mathbb{R}$) свести к разобранному. С учетом теоремы 3.3 получаем, что в п.1) доказано, что для каждой эрмитово-симметричной матрицы A уравнение $|A - \lambda E| = 0$ имеет лишь вещественные корни. Этого достаточно, так как φ в любом ОНБ имеет симметричную матрицу. (частный случай эрмитово-симметричной). $\square$

Предложение 3.2. *Пусть* $\varphi \in L(\mathcal{E},\mathcal{E})$ — *самосопряженное преобразование,* $\lambda_i \neq \lambda_j$ — *его различные собственные значения. Тогда* $V_{\lambda_i} \perp V_{\lambda_j}$

$\triangleright$ Пусть $\mathbf{a}_i \in V_{\lambda_i}$ и $\mathbf{a}_j \in V_{\lambda_j}$. Тогда $(\varphi(\mathbf{a}_i), \mathbf{a}_j) = (\mathbf{a}_i, \varphi(\mathbf{a}_j)) \Rightarrow (\lambda_i\mathbf{a}_i, \mathbf{a}_j) = (\mathbf{a}_i, \lambda_j\mathbf{a}_j) \Rightarrow \lambda_i(\mathbf{a}_i, \mathbf{a}_j) = \lambda_j(\mathbf{a}_i, \mathbf{a}_j)$. Отсюда с учетом $\lambda_i \neq \lambda_j$ получаем $(\mathbf{a}_i, \mathbf{a}_j) = 0$, что и требовалось. $\square$

Теорема 3.5 (Основная теорема о самосопряженных преобразованиях). *Пусть* $\dim\mathcal{E} < \infty$. *Для самоспоряженного преобразования* $\varphi \in L(\mathcal{E},\mathcal{E})$ *существует ОНБ из собственных векторов.*

$\triangleright$ Проведем индукцию по $\dim \mathcal{E}$. База $\dim \mathcal{E} = 1$ очевидна.

Пусть $\dim \mathcal{E} = n$ и предположим, для размерности $n-1$ утверждение теоремы верно. Рассмотрим собственный вектор и нормируем его, получим $\mathbf{e}_n$ такой, что $|\mathbf{e}_n| = 1$ и $\varphi(\mathbf{e}_n) = \lambda_n \mathbf{e}_n$. Подпространство $\langle \mathbf{e}_n \rangle$ размерности 1 инвариантно, значит $U = \langle \mathbf{e}_n \rangle^\perp$ — $(n-1)$-мерное инвариантное подпространство. Как мы замечали, сужение φ на U — самосопряженное, и по предположению индукции в U существует ОНБ $\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}$ из собственных векторов. Но $\mathbf{e}_n \perp \mathbf{e}_i$ для всех $i = 1, \ldots, n-1$, поэтому $\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}, \mathbf{e}_n$ — искомый ОНБ в $\mathcal{E}$. $\square$

Последняя теорема фактически усиливает предложение 3.2: получается, что для самосопряженного преобразования $\mathcal{E} = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \ldots \oplus V_{\lambda_k}$, где V_{λ_i}, $i = 1, \ldots, k$, — попарно ортогональные собственные подпространства.

Геометрическая интерпретация последней теоремы: самосопряженное преобразование — композиция «обобщенных растяжений» (т.е. с любым вещественным коэффициентом) вдоль ортогональных осей. В частности, ортогональные проектирования и отражения являются самосопряженными преобразованиями.

§ 4. Изометрии. Ортогональные и унитарные преобразования

Изометрия (изоморфизм евклидовых (унитарных) пространств)

Пусть даны два евклидовых (унитарных) пространства $\mathcal{E}_1$ и $\mathcal{E}_2$.

> **Определение.** Отображение $\varphi : \mathcal{E}_1 \to \mathcal{E}_2$ называется *изометрией*, или *изоморфизмом евклидовых (унитарных) пространств*, если φ — изоморфизм векторных пространств с дополнительным условием: $\forall\, \mathbf{a}, \mathbf{b} \in \mathcal{E}_1$ выполнено $(\varphi(\mathbf{a}), \varphi(\mathbf{b})) = (\mathbf{a}, \mathbf{b})$.

Итак, изоморфизм екввлидовых пространств — это обычный изоморфизм векторных пространств, который вдобавок сохраняет скалярное произведение.

> **Определение.** Два евклидовых (унитарных) пространства $\mathcal{E}_1$ и $\mathcal{E}_2$ называются *изоморфными*, если существует изоморфизм $\varphi : \mathcal{E}_1 \to \mathcal{E}_2$.

Обозначение $\cong$ (как и в случае обычного изоморфизма). Как и для обычного изоморфизма показывается, что отношение $\cong$ — отношение эквивалентности.

Предложение 4.1. *Пусть* $\dim \mathcal{E}_1 = \dim \mathcal{E}_2 = n < \infty$ *и* $\varphi \in L(\mathcal{E}_1 \to \mathcal{E}_2)$, $\mathrm{e} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ — *ОНБ в* $\mathcal{E}_1$. *Тогда* φ *является изоморфизмом евклидовых (унитарных) пространств* $\Leftrightarrow$ $(\varphi(\mathbf{e}_1), \ldots, \varphi(\mathbf{e}_n))$ — *ОНБ в* $\mathcal{E}_2$.

$\triangleright$ $\boxed{\Rightarrow}$ Сразу следует из определения изоморфизма евклидовых (унитарных) пространств: $\varphi(\mathbf{e}_i), \varphi(\mathbf{e}_j)) = (\mathbf{e}_i, \mathbf{e}_j)) = \delta_{ij}$.

$\boxed{\Leftarrow}$ Так как φ переводит базис в базис, то φ — (обычный) изоморфизм векторных пространств (см.....). Возьмем произвольные векторы $\mathbf{a}$ и $\mathbf{b}$ из V и разложим их по базису e: $\mathbf{a} = \sum\limits_{i=1}^{n} x_i \mathbf{e}_i$, $\mathbf{b} = \sum\limits_{i=1}^{n} y_i \mathbf{e}_i$. Тогда $\varphi(\mathbf{a}) = \sum\limits_{i=1}^{n} x_i \varphi(\mathbf{e}_i)$, $\varphi(\mathbf{b}) = \sum\limits_{i=1}^{n} y_i \varphi(\mathbf{e}_i)$. Тогда по следствию 4 из предложения 2.2: $(\mathbf{a}, \mathbf{b}) = \sum\limits_{i=1}^{n} x_i \overline{y_i} = (\varphi(\mathbf{a}), \varphi(\mathbf{b}))$. $\square$

Теорема 4.1. *Два конечномерных евклидовых (унитарных) пространства* $\mathcal{E}_1$ *и* $\mathcal{E}_2$ *изоморфны* $\Leftrightarrow$ $\dim \mathcal{E}_1 = \dim \mathcal{E}_2$.

$\triangleright$ $\boxed{\Rightarrow}$ Если $\dim \mathcal{E}_1 \neq \dim \mathcal{E}_2$, то $\mathcal{E}_1$ и $\mathcal{E}_2$ не изоморфны даже как обычные векторные пространства. $\boxed{\Leftarrow}$ Если $\dim \mathcal{E}_1 = \dim \mathcal{E}_2$, то достаточно взять ОНБ в каждом из пространств и взять линейное отображение $\varphi : \mathcal{E}_1 \to \mathcal{E}_2$, переводящее ОНБ в ОНБ. Согласно предложению 3.1, φ будет изоморфизмом. $\square$

Ортогональные и унитарные преобразования

Определение. Линейное преобразование $\varphi : \mathcal{E} \to \mathcal{E}$ называется ортогональным (унитарным), если $\forall\, \mathbf{a}, \mathbf{b} \in V$ выполнено

$$(\varphi(\mathbf{a}), \varphi(\mathbf{b})) = (\mathbf{a}, \mathbf{b}).$$

Ортогональное преобразование сохраняется скалярное произведение. Геометрически это означает сохранение длин и (в случае евклидова пространства) углов, т.е. если мыслить себе векторы как радиус-векторы с началом в O, то происходит «движение» с неподвижной точкой O. Для конечномерных евклидовых (унитарных) пространств ортогональные (унитарные) преобразование — это в точности изоморфизмы на себя:

Предложение 4.2. *Пусть* $\dim \mathcal{E} = n < \infty$, $\varphi \in L(\mathcal{E} \to \mathcal{E})$. *Тогда* φ *— ортогональное (унитарное)* $\Leftrightarrow$ φ *— изоморфизм евклидовых (унитарных) пространств.*

$\triangleright$ $\boxed{\Rightarrow}$ Достаточно доказать биективность φ. Но из определелния ортогональности преобразования следует, что φ переводит ОНБ в некоторую ортонормированную систему из n векторов, т.е. в ОНБ.
$\boxed{\Leftarrow}$ Очевидно по определению изоморфизма. $\square$

Следствие 1. *Пусть* $\dim \mathcal{E} = n < \infty$, $\varphi \in L(\mathcal{E} \to \mathcal{E})$, $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ *— ОНБ в* $\mathcal{E}$. *Тогда* φ *является ортогональным (унитарным) преобразованием* $\Leftrightarrow$ $(\varphi(\mathbf{e}_1), \ldots, \varphi(\mathbf{e}_n))$ *— ОНБ в* $\mathcal{E}$.

Следствие 2. *Пусть* $\dim \mathcal{E} = n < \infty$, $\mathrm{e} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ *— ОНБ в* $\mathcal{E}$. *Пусть* $\varphi \in L(\mathcal{E} \to \mathcal{E})$ *так, что* $\varphi \underset{\mathrm{e,e}}{\longrightarrow} A$. *Тогда* φ *является ортогональным (унитарным) преобразованием* $\Leftrightarrow$ A *— ортогональная (унитарная) матрица.*

Следствие 3. *Пусть* $\dim \mathcal{E} = n < \infty$, $\varphi \in L(\mathcal{E} \to \mathcal{E})$. *Тогда* φ *является ортогональным (унитарным) преобразованием* $\Leftrightarrow$ φ *обратимо, причем* $\varphi^* = \varphi^{-1}$.

$\triangleright$ Докажем в терминах матриц (хотя можно вывести непосредственно из определения). Пусть e — ОНБ в $\mathcal{E}$ так, что $\varphi \underset{\mathrm{e,e}}{\longrightarrow} A$. Тогда $\varphi^* \underset{\mathrm{e,e}}{\longrightarrow} A^*$, $\varphi^{-1} \underset{\mathrm{e,e}}{\longrightarrow} A^{-1}$. Поскольку $A^* = A^{-1}$, получаем $\varphi^* = \varphi^{-1}$. $\square$

Предложение 4.3 (Групповые свойства). *Пусть* $\dim \mathcal{E} = n < \infty$, $\varphi, \psi \in L(\mathcal{E}, \mathcal{E})$ *— ортогональные (унитарные) преобразования. Тогда преобразования* $\varphi\psi$ *и* φ^{-1} *— также ортогональные (унитарные).*

$\triangleright$ Можно вывести из определения или из предложения... (ортог. матрицы). $\square$

Предложение 4.4. *Пусть* $\varphi \in L(\mathcal{E}, \mathcal{E})$ *— ортогональное (унитарное) и* λ_0 *— его характеристическое число. Тогда* $|\lambda_0| = 1$.

$\triangleright$ 1) Докажем вначале утверждение для унитарного пространства (над $\mathbb{C}$). Пусть $\mathbf{a}$ — собственный вектор, соответствующий собственному значению λ_0. Равенство $(\varphi(\mathbf{a}), \varphi(\mathbf{a}) = (\mathbf{a}, \mathbf{a})$ принимает вид $(\lambda_0 \mathbf{a}, \lambda_0 \mathbf{a}) = (\mathbf{a}, \mathbf{a})$, откуда $\lambda_0 \overline{\lambda_0} (\mathbf{a}, \mathbf{a}) = (\mathbf{a}, \mathbf{a}) \Rightarrow \lambda_0 \overline{\lambda_0} = 1 \Rightarrow |\lambda_0| = 1$.
2) Покажем, как случай евклидова пространства (над $\mathbb{R}$) свести к разобранному. В п.1) доказано, что для каждой унитарной матрицы A уравнение $|A - \lambda E| = 0$ имеет лишь корни по модулю равные 1. Этого достаточно, так как φ в любом ОНБ имеет ортогональную матрицу (частный случай унитарной матрицы). $\square$

В унитарном (над $\mathbb{C}$) пространстве имеется результат, а аналогичный теореме 3.5.

Теорема 4.2 (канонический вид унитарного преобразовния). *Пусть* $\mathcal{E}$ *— унитарное пространство,* $\dim \mathcal{E} < \infty$. *Для унитарного преобразования* $\varphi \in L(\mathcal{E}, \mathcal{E})$ *существует ОНБ из собственных векторов.*

$\triangleright$ Доказательство повторяет доказательство теоремы 3.5 с следующим небольшим отличием в обосновании инвариантности $U = \langle \mathbf{e}_n \rangle^{\perp}$ относительно φ: по теореме 3.1 U инвариантно относительно $\varphi^* = \varphi^{-1}$; но тогда по предложению .... главы ... , U инвариантно и относительно φ. $\square$

Аналогом последней теоремы для евклидова пространства (над $\mathbb{R}$) является следующее утверждение о каноническом виде ортогонального преобразования. Пусть $\varphi \in L(\mathcal{E}, \mathcal{E})$ — ортогональное. Тогда $\mathcal{E}$ — прямая сумма попарно ортогональных инвариантных подпространств размерности 1 или 2.
Доказать это по той же схеме, что и теорему 3.5 используя, что у любого преобразования вещественного пространства найдется одномерное либо двумерное инвариантное подпространство.

Полярное разложение

Теорема 4.3. *Пусть* $\dim \mathcal{E} = n < \infty$, $\varphi, \in L(\mathcal{E}, \mathcal{E})$. *Тогда существуют самосопряженное преобразование* ψ *и ортогональное (унитарное) преобразование* θ *такие, что*

$$\varphi = \psi\theta.$$

$\triangleright$ $\square$

Мы получили следующую информацию о геометрии произвольного преобразования евклидова пространства: это композиция некоторого «движения» и «обощенных растяжений» к ортогональным осям.

Упражнение. Выведите из теоремы существование разложения $\varphi = \theta_1\psi_1$, где ψ_1 — самосопряженное, θ_1 — ортогональное.

§ 5. Билинейные формы в евклидовом пространстве

Связь между билинейными формами и преобразованиями

Пусть $\mathcal{E}$ — евклидово (унитарное) пространство. Каждому линейному оператору $\varphi \in L(\mathcal{E}, \mathcal{E})$ сопоставим отображение $\beta_\varphi : \mathcal{E} \times \mathcal{E} \to \mathbb{R}$ по формуле

$$\boxed{\beta_\varphi(\mathbf{a}, \mathbf{b}) = (\mathbf{a}, \varphi(\mathbf{b}))}. \qquad (*)$$

Предложение 5.1. β_φ, *определенное* $(*)$ — *билинейная (полуторалинейная) форма.*

$\triangleright$ $\square$

Предложение 5.2. *Пусть* $\dim \mathcal{E} < \infty$, $\varphi \in L(\mathcal{E}, \mathcal{E})$, e — *некоторый ОНБ и* $\varphi \underset{\mathrm{e,e}}{\longrightarrow} A$. *Тогда* $\varphi_\beta \underset{\mathrm{e}}{\longrightarrow} \overline{A}$.

$\triangleright$ Запишем в координатах (как обычно, полагая $\mathbf{a} = \mathrm{e}X$, $\mathbf{b} = \mathrm{e}Y$):
$\beta(\mathbf{a}, \mathbf{b}) = (\mathbf{a}, \varphi(\mathbf{b})) = X^T\overline{AY} = X^T\overline{AY}$. Мы видим координатную запись билинейной формы с матрицей $B = \overline{A}$. Соглаасно предложению..., все доказано. $\square$

Мы видим, что сопоставление $\varphi \to \beta_\varphi$ согласуется матрицами в ОНБ. Это объясняется также и тем, что законы преобразования при переходе от ОНБ к ОНБ (с некоторой ортогональной (унитарной) матрицей перехода S) для матрицы оператора A и (комплексно-сопряженной) матрицы билинейной формы $\overline{B}$ выглядят одинаково: $A \to S^{-1}AS$, $\overline{B} \to \overline{S^T}\overline{B}S$.

Следствие 1. *Пусть* $\dim \mathcal{E} < \infty$. *Сопоставление* $\varphi \to \beta_\varphi$ *задает биекцию* $L(\mathcal{E}, \mathcal{E}) \to \mathcal{B}(\mathcal{E})$.

Следствие 2. *Пусть* $\dim \mathcal{E} < \infty$. *Сопоставление* $\varphi \to \beta_\varphi$ *задает биекцию между множеством самосопряженных операторов и* $\mathcal{B}_{sym}(\mathcal{E})$ *(или* $\mathcal{K}(\mathcal{E})$*).*

Таким образом, сопоставление $\varphi \to \beta_\varphi$ дает возможножность сводить изучение билинейных форм к изучению операторов и наоборот. В частности, изучение квадратичных (эрмитровых) форм может быть сведено к изучению самосопряженных операторов.

Иногда термины *положительная определенность* или *положительность* мы будет применять и к самосопряженным преобразованиям (имея в виду указанную биекцию).

Приведение к главным осям

Теорема 5.1. *Пусть k — квадратичная форма в евклидовом (унитарном) пространстве $\mathcal{E}$. Тогда существует ОНБ, в котором k имеет диагональный вид.*

$\triangleright$ СХЕМА: ссылаемся на соответсвующую теорему о с/с операторах. $\square$

Упражнение. Пусть дана квадратичная форма k на конечномерном пространстве V (без евклидовой структуры). Пусть k в некотором базисе имеет матрицу B. Тогда k положительно определена $\Leftrightarrow$ все корни уравнения $|B - \lambda E| = 0$ положительны.

Задачу нахождение указанного диагонального вида и соответствующего ОНБ иногда называют *приведением формы к главным осям.*

Приложение к классификации кривых второго порядка

Предложение 5.3. *Пусть ℓ — кривая второго порядка на плоскости. Тогда существует ПДКС Oxy, в котором ℓ задается уравнением $Ax^2 + By^2 + 2Dx + 2Ey + F = 0$.*

$\triangleright$ $\square$

Предложение 5.4. *Пусть ℓ — поверхность второго порядка в пространстве. Тогда существует ПДКС $Oxyz$, в котором ℓ задается уравнением $Ax^2 + By^2 + Cz^2 + 2Dx + 2Ey + 2Fz + G = 0$.*

$\triangleright$ $\square$

Пара форм в векторном пространстве V (без евклидовой структуры)

Следствие (Теорема о паре форм)**.** *Пусть в векторном пространстве V (без евклидовой или унитарной структуры) заданы две формы $\beta \in \mathcal{B}_{sym}$ и $g \in \mathcal{B}_{sym}$, причем g положительно определена. Тогда существует базис, в котором g имеет канонический вид, а β имеет диагональный вид (в частности, обе формы β и g имеют диагональный вид).*

$\triangleright$ Зафиксируем билинейную (полуторалинейную) симметричную форму, которая порождает g, как скалярное произведение. При этом V превратилось в евклидово (унитарное) пространство. Теперь достаточно воспользоваться теоремой 5.1, поскольку ОНБ — это базис, в котором g имеет канонический вид (см. следствие 2 предложения 2.2). $\square$

АЛГОРИТМ приведения пары форм к диагональному виду.

Пусть $\beta \xrightarrow[\mathrm{e}]{} B$ и $g \xrightarrow[\mathrm{e}]{} G$. Как и в доказательстве теоремы, считаем V евклидовым, где скалярное произведение задается как $(\mathbf{a}, \mathbf{b}) = g(\mathbf{a}, \mathbf{b})$; G — матрица Грама данного базиса e. Нам требуется найти базис e$'$, для которого $\beta \xrightarrow[\mathrm{e}']{} diag$ и $g \xrightarrow[\mathrm{e}']{} E$.

Переформулировка (с учетом предложения

Равенство $(*)$ расписывается в координатах в базисе e как $X^T B\overline{Y} = X^T G(\overline{AY})$, где $\varphi \xrightarrow[\mathrm{e,e}]{} A$. Отсюда (в силу...) $B = G\overline{A}$ и $A = \overline{G^{-1}B}$. Далее проделываются шаги алгоритма нахождения ОНБ, в котором самосопряженное преобразование имеет диагональный вид.

I. (с.з. для φ).

$|\overline{G^{-1}B} - \lambda E| = 0$, или эквивалентный вид (домножив на $|G|$ и используя, что $\lambda \in \mathbb{R}$),

$|B - \lambda G| = 0$.

После нахождений λ_i (с кратностями s_i) уже известен диагональный вид $diag$.

II. (с.подпр. для φ).

СЛУ $(\overline{G^{-1}B} - \lambda_i E)X = 0 \Leftrightarrow (\overline{B - \lambda_i G})X = 0$. Поэтому:

$V_{\lambda_i} = \mathrm{Sol}((\overline{B - \lambda_i G})X = 0)X = 0)$.

ФСР $=$ базис V_{λ_i}.

III. Внутри каждого V_{λ_i}

ортогонализуем базис,

а затем нормируем базис.

(применяя формулы, помним, что $(\mathbf{a}, \mathbf{b}) = X^T G\overline{Y}$, т.е. матрица Грама $= G$).

Задача. Приведите пример пары квадратичных форм, для которых не существует базиса, в котором они обе имеют диагональный вид.

Указание. Можно подобрать формы, для которых инвариант $|F - \lambda G|$ имеет невещественный корень.

Глава 6

Сопряженное пространство

Пусть V — векторное пространство над $\mathbb{R}$ (или $\mathbb{C}$).

§ 1. Связь с общей теорией линейных отображений

Определение

> **Определение.** Пространство $\overline{L}(V,\widetilde{V})$, где $\widetilde{V}=\mathbb{R}$ (или $\mathbb{C}$)) называется *сопряженным (или двойственным)* пространством для пространства V.

Таким образом, сопряженное пространство — это частный случай пространства $\overline{L}(V,\widetilde{V})$, где $\dim\widetilde{V}=1$. Элементы сопряженного пространства — линейные функции (функционалы), поэтому сопряженное пространство также называют *пространством линейных функций.* Сопряженное пространство для пространства V обозначается V^*.

Предложение 1.1. *Если* $\dim V=n<\infty$*, то* $\dim V^*=n$.

$\triangleright$ Это частный случай следствия из теоремы 2.2 главы 2. $\square$

Вся теория о линейных отображениях переносится на частный случай пространства V^*. В частности, если зафиксировать базис $\mathrm{e}=(\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n)$ в пространстве V то каждая линейная функция ℓ получает в соответствие матрицу $1\times n$, то есть строку $(\ell(\mathbf{e}_1),\ell(\mathbf{e}_2),\dots,\ell(\mathbf{e}_n))$. (Здесь мы считаем, что в пространстве $\widetilde{V}=\mathbb{R}$ (или $\mathbb{C}$)) зафиксирован базис — число 1.)

Если $\mathbf{a}\in V$ и $\ell\in V^*$, то наряду с записью $\ell(\mathbf{a})$ будем использовать запись $\langle\mathbf{a},\ell\rangle$. Эта запись удобна, так как имеется линейность по каждому аргументу, т.е. $\forall\ \mathbf{a},\mathbf{a}_1,\mathbf{a}_2\in V$; $\forall\ \ell,\ell_1,\ell_2\in V^*$; $\forall\lambda\in\mathbb{R}(\mathbb{C})$ выполнено:
$\langle\mathbf{a}_1+\mathbf{a}_2,\ell\rangle=\langle\mathbf{a}_1,\ell\rangle+\langle\mathbf{a}_2,\ell\rangle$, $\langle\lambda\mathbf{a},\ell\rangle=\lambda\langle\mathbf{a},\ell\rangle$;
$\langle\lambda\mathbf{a},\ell_1+\ell_2\rangle=\langle\mathbf{a},\ell_1\rangle+\langle\mathbf{a},\ell_2\rangle$, $\langle\mathbf{a},\lambda\ell\rangle=\overline{\lambda}\langle\mathbf{a},\ell\rangle$.

Иными, словами, скобка $\langle\cdot,\cdot\rangle$ задает билинейную (полуторалинейную) функцию $V\times V^*\to\mathbb{R}(\mathbb{C})$.

Взаимный базис

Везде при работе с базисами и координатами полагаем, что $\dim V=n<\infty$. Чтобы в дальнейшем была возможность использовать сокращенную тензорную запись суммирования, координаты векторов пространства V будем нумеровать верхним индексом, а для пространства V^* — наоборот, векторы нумеруем верхними индексами, а координаты — нижними.

> **Определение.** Базис $\mathrm{e}^*=(\mathbf{e}^1,\mathbf{e}^2,\dots,\mathbf{e}^n)$ пространства V^* называется *взаимным (или биортогональным, или двойственным)* для базиса $\mathrm{e}=(\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n)$ пространства V, если $\langle\mathbf{e}_i,\mathbf{e}^j\rangle=\delta_i^j$ $(i=1,2,\dots,n,\ j=1,2,\dots,n)$.

Предложение 1.2. *Пусть* $\dim V=n$. *Тогда для любого базиса в* V *существует единственный взаимный базис в* V^*.

$\rhd$ Для каждого конкретного j вектор $\mathbf{e}^j \in V^*$ определяется однозначно как линейное отображение $V \to \mathbb{R}(\mathbb{C})$, которому в базисе e соответствует строка $E_j = (0\,0\ldots 1\,0\ldots 0)$ (единица на j-м месте). Строки $E_1,\ldots,E_n$ образуют базис в $\mathbf{M}_{1\times n}$, поэтому $\mathbf{e}^1,\mathbf{e}^2,\ldots,\mathbf{e}^n$ образуют базис пространства V^*. $\square$

Следствие. *Пусть* $\dim V = n$ *и* $\mathbf{a} \in V$. *Если* $\forall\ \ell \in V^*$ *выполнено* $\langle \mathbf{a}, \ell\rangle = 0$, *то* $\mathbf{a} = \mathbf{0}$.

$\rhd$ От противного, если $\mathbf{a} \neq \mathbf{0}$, то $\mathbf{a}$ можно включить в некоторый базис e, так что $\mathbf{a} = \mathbf{e}_1$. Тогда, если $\ell = \mathbf{e}^1$, то условие $\langle \mathbf{a}, \ell\rangle = 0$ нарушится. $\square$

Предложение 1.3. *Пусть* $\dim V = n < \infty$, e^* *— взаимный базис для базиса* e. *Пусть векторы* $\mathbf{a} \in V$, $\ell \in V^*$ *разложены по этим базисам:* $\mathbf{a} = x^i\mathbf{e}_i$, $\ell = y_j\mathbf{e}^j$. *Тогда* $\langle \mathbf{a}, \ell\rangle = x^i\overline{y_i}$.

$\rhd$ Достаточно раскрыть скобку по линейности и воспользоваться биортогональностью базисов: $\langle \mathbf{a}, \ell\rangle = \langle x^i\mathbf{e}_i, y_j\mathbf{e}^j\rangle = x^i\overline{y_j}\langle \mathbf{e}_i, \mathbf{e}^j\rangle = x^i\overline{y_i}$. $\square$

Таким образом, значение $\langle \mathbf{a}, \ell\rangle$ равно *свертке* $x^i\overline{y_i}$. В частности, видим, как вектор из взаимного базиса действует на произвольном векторе из V:

Следствие. *Пусть* $\dim V = n < \infty$, e^* *— взаимный базис для базиса* e *и* $\mathbf{a} = x^i\mathbf{e}_i$. *Тогда* $\langle \mathbf{a}, \mathbf{e}^i\rangle = x^i$.

Предложение 1.4. *Пусть* $\dim V = n$. *Пусть* e *и* e' *— базисы в* V, *а* e^* *и* e'^* *— их взаимные базисы в* V^* *соответственно. Пусть* S *— матрица перехода от* e *к* e', *а* C *— матрица перехода от* e^* *к* e'^*. *Тогда*

$$C = (S^{-1})^*.$$

$\rhd$ Во встречающихся матрицах перехода обозначаем верхним индексом номер строки, а нижним — номер столбца. Так, если $S = (s_i^k)$ — матрица перехода от $\mathrm{e} = (\mathbf{e}_1,\ldots,\mathbf{e}_n)$ к $\mathrm{e}' = (\mathbf{e}'_1,\ldots,\mathbf{e}'_n)$, то согласно определению матрицы перехода, $\mathbf{e}'_i = \sum\limits_{k=1}^{n} s_i^k\mathbf{e}_k$, или используя тензорную запись суммирования,

$$\mathbf{e}'_i = s_i^k\mathbf{e}_k.$$

Аналогично, если $C = (c_j^m)$ — матрица перехода от $\mathrm{e}^* = (\mathbf{e}^1,\ldots,\mathbf{e}^n)$ к $\mathrm{e}'^* = (\mathbf{e}'^1,\ldots,\mathbf{e}'^n)$, то $\mathbf{e}'^j = \sum\limits_{m=1}^{n} c_j^m\mathbf{e}^m$. Чтобы здесь также использовать тензорное суммирование, введем матрицу $R = C^T$, так что $R = (r_m^j)$ и $c_j^m = r_m^j$. Тогда

$$\mathbf{e}'^j = r_m^j\mathbf{e}^m$$

Условие биортогональности запишется как $\delta_i^j = \langle \mathbf{e}'_i, \mathbf{e}'^j\rangle = \langle s_i^k\mathbf{e}_k, r_m^j\mathbf{e}^m\rangle$. Раскрывая по линейности и используя биортогональность e и e^*, получаем условия $\delta_i^j = s_i^k\overline{r_k^j}$, равносильные матричному равенству $\overline{R}S = E$, откуда $R = \overline{S}^{-1} = \overline{S^{-1}}$, отсюда $C = R^T = (S^{-1})^*$. $\square$

Работа одновременно в V и V^* иногда дает больше, чем работа в одном из них. Например, следующее предложение является признаком линейной независимости (при $n = k$ признаком базиса в V).

Предложение 1.5. *Пусть система векторов* $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2,\ldots,\mathbf{e}_k)$ *пространства* V *и система векторов* $\ell = (\ell^1,\ell^2,\ldots,\ell^k)$ *пространства* V^* *биортогональны, т.е. таковы, что* $\langle \mathbf{e}_i, \ell^j\rangle = \delta_i^j$ *для всех* $i, j \in \{1,\ldots,k\}$. *Тогда* e *и* ℓ *— линейно независимые системы в* V *и* V^* *соответственно.*
В частности, если $\dim V = n = k$, *то* e *— базис в* V, *а* ℓ *— взаимный базис для* e.

$\rhd$ Достаточно доказать линейную независимость системы векторов $\mathbf{e}_1, \mathbf{e}_2,\ldots,\mathbf{e}_k$ (линейную независимость системы векторов $\ell^1,\ell^2,\ldots,\ell^k$ можно доказать аналогично).
Пусть $\lambda^i\mathbf{e}_i = \mathbf{0}$. Применив к этому равенству $\ell^j \in V^*$, имеем $\lambda^j = 0$. Таким образом, все коэффициенты в этой линейной комбинации должны быть равны 0. $\square$

Примеры

III.1. (интегральный функционал) Пусть $V = C[a,b]$, и $f_0 \in V$. Определим $\widetilde{f_0} \in V^*$: $\forall\ g \in V$ положим $\langle g, \widetilde{f_0}\rangle = \int\limits_a^b g(x) f_0(x)\, dx$.
Другой пример: δ-функция — линейная функция $\delta_\alpha \in V^*$ такая, что $\forall\ g \in V$ выполнено: $\langle g, \delta_\alpha\rangle = g(\alpha)$.

III.2. Пусть $V = \mathbf{P}_n$ (многочлены степени $\leqslant n$); $\alpha_0, \alpha_1, \ldots, \alpha_n$ — различные числа. Для системы линейных функнций $\ell^i = \delta_{\alpha_i}$, $i = 0, 1, \ldots, n$, биортогональной будет система многочленов

$$p_i(x) = \frac{\prod\limits_{k\neq i}(x - \alpha_k)}{\prod\limits_{k\neq i}(\alpha_i - \alpha_k)},$$

$i = 0, 1, \ldots, n$. Из предложения 1.5 следует, что $(\ell^0, \ell^1, \ldots, \ell^n)$ — взаимный базис для базиса $(p_0, p_1, \ldots, p_n)$.
Разложение многочлена $f \in \mathbf{P}_n$ по p_i имеет вид

$$\sum_{i=0}^{n} f(\alpha_i)\, p_i.$$

Мы получили так называемый *интерполяционный многчлен* Лагранжа, дающий явную формулу для многочлена степени не выше n, принимающего в заданных $n+1$ точках предписанные значения.

§ 2. Естественные изоморфизмы

Во многих рассуждениях у нас появлялись изоморфизмы, которые бы изменились при другом выборе базиса (как, скажем, сопоставление вектору его координатного столбца). В ситуациях, когда этого не происходит, т.е. если изоморфизм не меняется при замене базиса, будем говорить, что изоморфизм естественный.

Теорема 2.1. *Отображение, сопоставляющее вектору* $\mathbf{a} \in V$ *отображение* $\widetilde{\mathbf{a}} : V^* \to \mathbb{R}(\mathbb{C})$ *по правилу*

$$\langle \ell, \widetilde{\mathbf{a}}\rangle = \overline{\langle \mathbf{a}, \ell\rangle} \tag{6.1}$$

(для любого $\ell \in V^*$*), является инъективным гомоморфизмом (вложением)* $V \to V^{**}$.

$\triangleright$ Для доказательства достаточно проверить следующие утверждения.

1) $\widetilde{\mathbf{a}}$, заданное правилом (6.1), линейно, т.е. действительно $\widetilde{\mathbf{a}} \in V^{**}$.

2) Соответствие $\mathbf{a} \to \widetilde{\mathbf{a}}$ линейно, т.е. проверить, что $\widetilde{\mathbf{a}_1 + \mathbf{a}_2} = \widetilde{\mathbf{a}_1} + \widetilde{\mathbf{a}_2}$ и $\widetilde{\lambda\mathbf{a}} = \lambda\widetilde{\mathbf{a}}$.

3) Тривиальность ядра отображения $\mathbf{a} \to \widetilde{\mathbf{a}}$. Если $\widetilde{\mathbf{a}} = \mathbf{0}$ (то есть для любого $\ell \in V^*$ выполнено $\langle \ell, \widetilde{\mathbf{a}}\rangle = 0$, то $\mathbf{a} = \mathbf{0}$. $\square$

Следствие. *В случае* $\dim V = n < \infty$ *указанное соответствие* $\mathbf{a} \to \widetilde{\mathbf{a}}$ *является естественным изоморфизмом* $V \to V^{**}$.

Соответствие $\mathbf{a} \to \widetilde{\mathbf{a}}$ позволяет отождествить далее конечномерное пространство со своим дважды сопряженным.

Теорема 2.2. *Пусть* $\mathcal{E}$ *— евклидово (унитарное) пространство. Отображение* $\mathcal{E} \to \mathcal{E}^*$, *сопоставляющее вектору* $\mathbf{a} \in \mathcal{E}$ *отображение* $\widehat{\mathbf{a}} : V^* \to \mathbb{R}(\mathbb{C})$ *по правилу*

$$\langle \mathbf{x}, \widehat{\mathbf{a}}\rangle = (\mathbf{x}, \mathbf{a}) \tag{6.2}$$

(для любого $\mathbf{x} \in \mathcal{E}$*), является инъективным гомоморфизмом (вложением)* $\mathcal{E} \to \mathcal{E}^*$.

$\triangleright$ Для доказательства достаточно проверить следующие утверждения.

1) $\widehat{\mathbf{a}}$, заданное правилом 6.2, линейно, т.е. действительно $\widehat{\mathbf{a}} \in \mathcal{E}^*$.

2) Соответствие $\mathbf{a} \to \widehat{\mathbf{a}}$ линейно, т.е. проверить, что $\widehat{\mathbf{a}_1 + \mathbf{a}_2} = \widehat{\mathbf{a}_1} + \widehat{\mathbf{a}_2}$ и $\widehat{\lambda\mathbf{a}} = \lambda\widehat{\mathbf{a}}$.

3) Тривиальность ядра отображения $\mathbf{a} \to \widehat{\mathbf{a}}$. Если $\widehat{\mathbf{a}} = \mathbf{0}$ (то есть для любого $\mathbf{x} \in V$ выполнено $\langle \mathbf{x}, \widehat{\mathbf{a}} \rangle = 0$), то $\mathbf{a} = \mathbf{0}$. $\square$

Следствие. *В случае* $\dim V = n < \infty$ *указанное соответствие* $\mathbf{a} \to \widehat{\mathbf{a}}$ *является естественным изоморфизмом* $\mathcal{E} \to \mathcal{E}^*$.

Соответствие $\mathbf{a} \to \widehat{\mathbf{a}}$ позволяет отождествить далее конечномерное евклидово (унитарное) пространство со своим сопряженным. В этом смысле теорию евклидовых (унитарных) пространств можно считать частным случаем теории двойственности.

§ 3. Биортогональность. Соответствие между подпространствами в V и V^*

Биортогональность.

> **Определение.** Множества $U \subset V$ и $W \subset V^*$ называются *биортогональными*, если $\forall\, \mathbf{a} \in U$ и $\forall\, \ell \in W$ выполнено $\langle \mathbf{a}, \ell \rangle = 0$.

Для биортогональности векторов и множеств из V и V^* сохраним обозначение $\perp$.

Предложение 3.1 (признак биортогональности). *Пусть* $U = \langle \mathbf{a}_1, \ldots, \mathbf{a}_k \rangle \leqslant V$, $W = \langle \ell_1, \ldots, \ell_l \rangle \leqslant V^*$. *Тогда* $U \perp W \Leftrightarrow \mathbf{a}_i \perp \ell_j$, $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, l$.

$\triangleright$ Аналогично доказательству предложения 3.1. $\square$

Биортогональное дополнение

> **Определение.** Подмножество $W = \{\ell \in V^* \,|\, U \perp \ell\}$ пространства V^* называется *аннулятором*, или биортогональным дополнением подпространства U.

Для евклидова пространства отождествление (6.2) превращает определение аннулятора в определение ортогонального дополнения. Для аннулятора примем обозначение $U^\perp$.

Предложение 3.2. *Для любого* $U \leqslant V$ *подмножество* $U^\perp$ *является подпространтвом в* V^*.

$\triangleright$ $\square$

Для любого $W \leqslant V^*$ подмножество $W^\perp$ будет является подпространтвом в V^{**}. Но в конечномерном случае в силу отождествления $V \cong V^{**}$ заданного (6.1), мы можем отождествить $W^\perp$ с подпространтвом в V, которое также называется *нуль-пространством* подпространства W. Зафиксируем также определение нуль-пространства, не использующее отождествление $V \cong V^{**}$.

> **Определение.** Подмножество $U = \{\mathbf{a} \in V \,|\, \mathbf{a} \perp W\}$ пространства V называется *нуль-пространством* подпространства $W \leqslant V^*$.

Для нуль-пространства сохраняем обозначение $W^\perp$.

Предложение 3.3. *Для любого* $W \leqslant V^*$ *подмножество* $W^\perp$ *является подпространтвом в* V.

$\triangleright$ $\square$

Теорема 3.1. *Пусть* $\dim V = n < \infty$, $U \leqslant V$. *Тогда*
1) $(U^\perp)^\perp = U$; *2)* $\dim U + \dim U^\perp = n$.

$\triangleright$ Пусть $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k$ — базис в U. Дополним его до базиса $\mathrm{e} = (\mathbf{e}_1, \ldots, \mathbf{e}_k, \ldots, \mathbf{e}_n)$ в V. Мы покажем, что для $U^\perp$ имеется следующее описание, из которого вытеают оба утверждения: $U^\perp = \langle \mathbf{e}^{k+1}, \mathbf{e}^{k+2}, \ldots, \mathbf{e}^n \rangle$, где $\mathbf{e}^1, \ldots, \mathbf{e}^n$ — взаимный базис для баззиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ пространства V. $\ldots \square$

Упражнение. Пусть $\dim V < \infty$. Для $U_i \leqslant V$, $i = 1, 2$, докажите, что $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$ и аналогично, $(U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp$.

§ 4. Сопряженное преобразование

Определение. Преобразование $\psi : V^* \to V^*$ называют *сопряженным* преобразованию $\varphi \in L(V, V)$, если $\forall\ \mathbf{a} \in V$, и $\forall\ \ell \in V^*$ выполнено

$$\boxed{\langle \varphi(\mathbf{a}), \ell \rangle = \langle \mathbf{a}, \psi(\ell) \rangle}. \tag{6.3}$$

Обозначение для сопряженного преобразования: φ^*. Для евклидова пространстве, в силу отождествления (6.2), это определения согласуется с определением из главы 5. Из определения φ^* однозначно определено. Проверим, что оно линейно.

Предложение 4.1. *Пусть* $\varphi \in L(V, V)$. *Тогда* $\varphi^* \in L(V^*, V^*)$.

$\triangleright$ $\square$

Предложение 4.2. *Пусть* $\dim V < \infty$, e — *базис в* V, *и* e^* — *его взаимный базис в* V^*. *Пусть* $\varphi \in L(V, V)$, $\varphi \underset{\mathrm{e,e}}{\longrightarrow} A$. $\varphi^* \underset{\mathrm{e^*,e^*}}{\longrightarrow} A^*$. *Тогда*

$\triangleright$ Аналогично предложению 3.1. $\square$

Следствие 1. *Пусть* $\dim V < \infty$, $\varphi, \psi \in L(V, V)$. *Тогда*
1) $(\varphi^*)^* = \varphi$;
2) $(\varphi\psi)^* = \psi^*\varphi^*$;
3) $\underline{\mathrm{rg}(\varphi)} = \mathrm{rg}(\varphi^*)$;
4) $\chi_\varphi(\lambda) = \chi_{\varphi^*}(\overline{\lambda})$.

$\triangleright$ Введем в V и V^* взаимные базисы и перейдем к матрицам в этих базисах. Далее доказываем так же, как следствие 2 из предложения 3.1 главы 5. $\square$

Теорема 4.1. *Пусть* $U \leqslant V$, $\dim V < \infty$, $\varphi \in L(V, V)$. *Тогда*
U *инвариантно относительно* $\varphi \Leftrightarrow U^\perp$ *инвариантно относительно* φ^*.

$\triangleright$ Аналогично теореме 3.1 главы 5. $\square$

Теорема 4.2 (Теорема Фредгольма). *Пусть* $\dim V = n < \infty$, $\varphi \in L(V, V)$. *Тогда*

$$\boxed{\operatorname{Ker} \varphi^* = (\operatorname{Im} \varphi)^\perp}.$$

$\triangleright$ Аналогично теореме 3.2 главы 5. $\square$

Отметим, что все содержание этого параграфа обобщается на случай линейного отображения $\varphi \in L(V, \widetilde{V})$. Для $\varphi \in L(V, \widetilde{V})$ так же формула (6.3) позволяет определить *сопряженное отображение* $\varphi^* \in L(\widetilde{V}^*, V^*)$, для которого выполняются аналоги всех предложений и теорем из этого параграфа.

Глава 7

Тензоры

В этой главе рассматриваем векторные пространства над полем $\mathbb{F}$. (где возникает деление (симметризация и пр.) предполагаем, что характеристика поля равна 0).

§ 1. Полилинейные отображения

Определение и координантая запись

Отображение $\beta : V_1 \times V_2 \times \ldots \times V_k \to \mathbb{F}$ *полилинейно*, если $\beta(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)$ линейно по каждому из аргументов.

Множество всех полилинейных отображений обозначим $\mathrm{Hom}(V_1, V_2, \ldots, V_k; \mathbb{F})$.

Примеры. В старых обозначениях для V над $\mathbb{R}$:
$\mathrm{Hom}(V, V; \mathbb{R}) = \mathcal{B}(V)$;
$\mathrm{Hom}(V; \mathbb{R}) = V^*$;
$\mathrm{Hom}(V^*; \mathbb{R}) = V$ (с учетом отождествления V и V^{**}).

Пусть $\dim V_t = n_t$ и в каждом V_t зафиксирован базис: в V_1 — базис $\mathrm{e} = (\mathbf{e}_1, \ldots, \mathbf{e}_{n_1})$, в V_2 — базис $\mathrm{f} = (\mathbf{f}_1, \ldots, \mathbf{f}_{n_2})$, и т.д.

Сопоставим каждому $\beta \in \mathrm{Hom}(V_1, V_2, \ldots, V_k; \mathbb{F})$ массив из $n_1 \ldots n_k$ констант $b_{i_1 i_2 \ldots i_k}$, $i_t \in \{1, \ldots, n_t\}$ (далее для упрощения обозначений пишем $b_{ij\ldots}$):

$$\beta \underset{\mathrm{e},\mathrm{f},\ldots}{\longrightarrow} b_{ij\ldots}, \tag{7.1}$$

где $\boxed{b_{ij\ldots} = \beta(\mathbf{e}_i, \mathbf{f}_j, \ldots)}$.

$b_{ij\ldots}$ называются *компонентами* полилинейного отображения в базисах $\mathrm{e}, \mathrm{f}, \ldots$.

Теорема 1.1 (координатная запись)**.** *Пусть* $\beta \in \mathrm{Hom}(V_1, V_2, \ldots, V_k; \mathbb{F})$ *и* $\beta \underset{\mathrm{e},\mathrm{f},\ldots}{\longrightarrow} b_{ij\ldots}$. *Пусть векторы* $\mathbf{a} \in V_1$, $\mathbf{b} \in V_2, \ldots,$ *разложены по базисам:* $\mathbf{a} = x^i \mathbf{e}_i$, $\mathbf{b} = y^j \mathbf{f}_j$, $\ldots$ *Тогда*

$$\boxed{\beta(\mathbf{a}, \mathbf{b}, \ldots) = b_{ij\ldots} x^i y^j \ldots}. \tag{7.2}$$

$\triangleright$ Раскроем $\beta(\mathbf{a}, \mathbf{b}, \ldots)$, ... $\square$

(7.1) задает взаимно-однозначное соответствие $\mathrm{Hom}(V_1, V_2, \ldots, V_k; \mathbb{F}) \to \mathbb{F}^{n_1 n_2 \ldots n_k}$

Замена базиса

Во встречающихся матрицах перехода обозначаем верхним индексом номер строки, а нижним — номер столбца, так если $S = (s^i_j)$ — матрица перехода от базиса $(\mathbf{e}_1, \ldots)$ к базису $(\mathbf{e}'_1, \ldots)$, то замена базиса происходит по правилу

$$\mathbf{e}'_j = s^i_j \mathbf{e}_i,$$

а, скажем, замена координат вектора — по правилу

$$x^i = s^i_j x'^j.$$

Теорема 1.2. *Пусть в* V_1 *выбраны базисы* e *и* e′, *связанные матрицей перехода* s^i_j, *в* V_2 *выбраны базисы* f *и* f′, *связанные матрицей перехода* t^i_j, *и т.д. Пусть* $\beta \in T(V_1, V_2, \dots, V_k)$ *так, что*

$$\beta \xrightarrow[\mathrm{e,f},\dots]{} b_{ij\dots},$$

$$\beta \xrightarrow[\mathrm{e',f'},\dots]{} b'_{ij\dots}$$

Тогда

$$\boxed{b'_{ij\dots} = b_{kl\dots} s^k_i t^l_j \dots}. \tag{7.3}$$

(Координатное определение тензора)

Линейные операции

$\mathrm{Hom}(V_1, V_2, \dots, V_k; \mathbb{F})$ — векторное пространство над $\mathbb{F}$.

§ 2. Тензоры и основные операции над ними

Тензор как полилинейная функция

> **Определение.** Тензором типа (p, q) над векторным пространством V называют полилинейное отображение из множества $\mathrm{Hom}(\underbrace{V^*, \dots, V^*}_{p}, \underbrace{V, \dots, V}_{q}; \mathbb{F})$.

Более заумное название тензора типа (p, q): p раз контравариантный и q раз ковариантный.

Обозначение (помимо $\mathrm{Hom}(\underbrace{V^*, \dots, V^*}_{p}, \underbrace{V, \dots, V}_{q}; \mathbb{F})$):

$$T^p_q(V).$$

(или просто T^p_q, если ясно, о каком V идет речь).

Для $\tau \in T^p_q(V)$ определено значение $\tau(\underbrace{\ell, \dots}_{p}, \underbrace{\mathbf{a}, \mathbf{b}, \dots}_{q})$ на упорядоченном наборе из p линейных функционалах (ковекторах) и q векторах.

$T^p_q(V)$ — векторное пространство.

Что такое $T^0_0(V)$?

Выписывая компоненты в базисах, условимся во всех q копиях V выбирать один и тот же базис $\mathrm{e} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$, а во всех p копиях V^* выбирать базис $\mathrm{e}^* = (\mathbf{e}^1, \dots, \mathbf{e}^n)$, биортогональный базису e. Компоненты пишем с p верхними и q нижними индексами, для возможности использовать тензорное суммирование.

Для $\tau \in T^p_q(V)$ соответствие (7.1) приобретает вид:

$\tau \xrightarrow[\mathrm{e}^*,\dots,\mathrm{e},\mathrm{e},\dots]{} t^{j_1 j_2 \dots j_p}_{i_1 i_2 \dots i_q}$, или (сокращая обозначения)

$$\tau \xrightarrow[\mathrm{e}]{} t^{j\dots}_{ik\dots}.$$

Массив из n^{p+q} констант $t^{j\dots}_{ik\dots} = \tau(\underbrace{\mathbf{e}^j, \dots,}_{q} \underbrace{\mathbf{e}_i, \mathbf{e}_k, \dots,}_{p})$ — компоненты тензора τ в базисе e.

Теперь формула (7.2) приобретает вид

$$\tau(\underbrace{\ell, \dots}_{p}, \underbrace{\mathbf{a}, \mathbf{b}, \dots}_{q}) = t^{j\dots}_{ik\dots} x^i y^k \dots z_j \dots,$$

где $\mathbf{a} = x^i \mathbf{e}_i$, $\mathbf{b} = y^k \mathbf{e}_k$, …, $\ell = z_j \mathbf{e}^j$, …— разложения векторов ($p + q$ аргументов, к которым применяется τ) по базисам e и e*

Формула (7.3) изменения компонент при переходе от базиса e к e′ и соответственно от базиса e* к e′*:

$$\boxed{t'^{j\dots}_{ik\dots} = t^{j'\dots}_{i'k'\dots}\, s^{i'}_i s^{k'}_k \dots r^j_{j'} \dots},$$

где (как, скажем, и в доказательства предложения 1.4 главы 6), $S = (s^{i'}_i)$ — матрица перехода от e к e′, а $R = (r^j_{j'}) = S^{-1}$ — транспонированная к матрице перехода от e* к e′* (транспонирование связано с нашей договоренностью о верхних и нижних индексах при работе в V^*).

Тензорное умножение

> **Определение.** Тензорным произведением тензора $\alpha \in T^p_q$ на тензор $\beta \in T^{p'}_{q'}$ называют тензор $\gamma \in T^{p+p'}_{q+q'}$ такой, что
>
> $$\gamma(\underbrace{\ell, \dots,}_{p} \underbrace{m, \dots,}_{p'} \underbrace{\mathbf{a}, \mathbf{b}, \dots,}_{q} \underbrace{\mathbf{c}, \mathbf{d}, \dots}_{q'}) \;=\; \alpha(\underbrace{\ell, \dots,}_{p} \underbrace{\mathbf{a}, \mathbf{b}, \dots}_{q}) \;\cdot\; \beta(\underbrace{m, \dots,}_{p'} \underbrace{\mathbf{c}, \mathbf{d}, \dots}_{q'}).$$

Обозначение: $\alpha \otimes \beta$.

Ассоциативность, дистрибутивность относительно сложения.

Нет коммутативности.

«организуем» алгебру

$$\bigoplus_{q=0}^{\infty} T^0_q$$

$$\bigoplus_{p,q} T^p_q.$$

В координатах — «независимое умножение». Пусть

$$\alpha \underset{\mathrm{e}}{\longrightarrow} a^{j_1 j_2 \dots j_p}_{i_1 i_2 \dots i_q},$$

$$\beta \underset{\mathrm{e}}{\longrightarrow} b^{j'_1 j'_2 \dots j'_{p'}}_{i'_1 i'_2 \dots i'_{q'}}.$$

Тогда

$$\gamma \underset{\mathrm{e}}{\longrightarrow} c^{j_1 j_2 \dots j_p\, j'_1 j'_2 \dots j'_{p'}}_{i_1 i_2 \dots i_q\, i'_1 i'_2 \dots i'_{q'}} \;=\; a^{j_1 j_2 \dots j_p}_{i_1 i_2 \dots i_q} \;\cdot\; b^{j'_1 j'_2 \dots j'_{p'}}_{i'_1 i'_2 \dots i'_{q'}}.$$

Разложимый тензор — тензор, равный произведению тензоров валентности 1 (векторов и ковекторов).

> Как действует тензор $\mathbf{e}^3 \otimes \mathbf{e}^1 \otimes \mathbf{e}^3 \in T^0_3$?
> $(\mathbf{e}^3 \otimes \mathbf{e}^1 \otimes \mathbf{e}^3)(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) = \delta^3_i \delta^1_j \delta^3_k$.
> $(\mathbf{e}^3 \otimes \mathbf{e}^1 \otimes \mathbf{e}^3)(\mathbf{a}, \mathbf{b}, \mathbf{c}) = x^3 y^1 z^3$, где $\mathbf{a} = x^i \mathbf{e}_i$, $\mathbf{b} = y^j \mathbf{e}_j$, $\mathbf{c} = z^k \mathbf{e}_k$.

Тензоры вида $\underbrace{\mathbf{e}_j \otimes \dots}_{p} \otimes \underbrace{\mathbf{e}^i \otimes \mathbf{e}^k \otimes \dots}_{q}$ — «стандартный» (т.е. согласованный с e) базис в T^p_q, так что

$$\tau = t^{j\dots}_{ik\dots}\; \underbrace{\mathbf{e}_j \otimes \dots}_{p} \otimes \underbrace{\mathbf{e}^i \otimes \mathbf{e}^k \otimes \dots}_{q},$$

где $\tau \underset{\mathrm{e}}{\longrightarrow} t^{j\dots}_{ik\dots}$.

Свертка

Зафиксируем некоторый базис и зададим *свертку* (по паре первых аргументов) как линейное отображение

$$T_q^p \to T_{q-1}^{p-1}$$

по следующему правилу. Достаточно определить значение свертки вначале на «стандартном базисе» $T_q^p(V)$, связанным с базисом e пространства V (далее продолжается по линейности):

$$\underbrace{\mathbf{e}_j \otimes \mathbf{e}_l \otimes \ldots}_{p} \otimes \underbrace{\mathbf{e}^i \otimes \mathbf{e}^k \otimes \ldots}_{q} \mapsto \delta^i_j \ \underbrace{\mathbf{e}_l \otimes \ldots}_{p-1} \otimes \underbrace{\mathbf{e}^k \otimes \ldots}_{q-1}. \tag{7.4}$$

Из линейности легко следует, что

$$\underbrace{\mathbf{a} \otimes \mathbf{b} \otimes \ldots}_{p} \otimes \underbrace{\ell \otimes m \otimes \ldots}_{q} \mapsto \langle \mathbf{a}, \ell \rangle \ \underbrace{\mathbf{b} \otimes \ldots}_{p-1} \otimes \underbrace{m \otimes \ldots}_{q-1}.$$

В частности, это означает, что (7.4) работает и для другого выбора базиса.

В координатах переход к свертке выглядит просто:

$$t^{jl\ldots}_{ik\ldots} \mapsto t^{il\ldots}_{ik\ldots}.$$

Примеры.

– Значение линейной формы на векторе: $\mathbf{a} \otimes \ell$ — далее свертка (в результате $l_i x^i$).

– Значение билинейной формы $\mathbf{a} \otimes \mathbf{b} \otimes \beta$ — далее свертка (в результате $b_{ij} x^i y^j$)

—Матрица $A = (a^i_j)$ линейного отображения $\varphi : V \to V$ — тензор из T^1_1.
$\operatorname{tr} A = a^i_i$ — свертка.
Матрица произведения двух отображений: $a^i_j b^j_k$.

— В евклидовом пространстве g_{ij} — метрический Тензор (матрица Грама).
Тензорное домножение на g_{ij} с последующей сверткой по одному из аргументов — «опускание индекса».
Например: $b_{jk} = a^i_j g_{ik}$.
(связь между операторами и билинейными формами в еквлидовом пространстве).

§ 3. Симметричные и кососимметричные тензоры.

Перестановка аргументов в функциях нескольких переменных.

Если X и Y — произвольные множества, то на множестве $\mathcal{F}$ отображений $\underbrace{X \times X \times \ldots \times X}_{q} \to Y$ естественным образом действует группа перестановок $S_q = S_{\{1,2,\ldots,q\}}$, так что $\sigma \in S_q$ переставляет аргументы отображения $f \in \mathcal{F}$:

$$(\sigma f)(x_1, \ldots, x_q) = f(x_{\sigma(1)}, \ldots, x_{\sigma(q)}).$$

Очевидно, $f \mapsto \sigma f$ задает биекцию $\mathcal{F} \to \mathcal{F}$. Далее считаем, что Y — векторное пространство (над некоторым полем $\mathbb{F}$), тогда $f \mapsto \sigma f$ задает изоморфизм $\mathcal{F} \to \mathcal{F}$.

Симметрические функции. Симметризация.

Определение. Функция $f \in \mathcal{F}$ называется *симметрической*, если $\forall\ \sigma \in S_q$ выполнено

$$\sigma f = f.$$

Подмножество всех симметрических $f \in \mathcal{F}$ можно обозначить $\mathcal{F}^+$.

Можно определять более общее понятие — симметричность относительно действия заданной подгруппы в $H \leqslant S_q$, заменяя в определении симметричности условие $\forall\ \sigma \in S_q$ на $\forall\ \sigma \in H$. Например, если $H = S_{\{1,2,\dots,p\}}$, получаем условие симметричности относительно первых p аргументов (и аналогичнно для $H = S_{\{p+1,p+2,\dots,q\}}$ — условие симметричности относительно последних $q-p$ аргументов).

С помощью действия S_q на $\mathcal{F}$ (в случае поля нулевой характеристики) можем определить *усреднение*, или *симметризацию* как

$$\mathrm{Sym}(f) = \frac{1}{q!}\sum_{\sigma\in S_q}\sigma f,$$

или более общо, симметризацию относительно действия заданной подгруппы в $H \leqslant S_q$ как

$$\mathrm{Sym}_H(f) = \frac{1}{|H|}\sum_{\sigma\in H}\sigma f.$$

Следующее предложение показывает, что $\mathrm{Sym} : \mathcal{F} \to \mathcal{F}$ является проектором на $\mathcal{F}^+$.

Предложение 3.1. *1) $\forall\ f \in \mathcal{F}$ выполнено $\mathrm{Sym}(f) \in \mathcal{F}^+$;*
2) если $f \in \mathcal{F}^+$, то $\mathrm{Sym}(f) = f$.

Лемма. *Пусть $\sigma_0 \in H \leqslant S_q$ (H — подгруппа в S_q). Тогда $\forall\ f \in \mathcal{F}$ выполнено*

$$\mathrm{Sym}_H(\sigma_0 f) = \mathrm{Sym}_H(f).$$

Лемма. *Пусть $H \leqslant G \leqslant S_q$ (две подгруппы S_q, одна вложена в другую). Тогда $\forall\ f \in \mathcal{F}$ выполнено*

$$\mathrm{Sym}_G(\mathrm{Sym}_H(f)) = \mathrm{Sym}_G(f).$$

Кососимметрические функции. Альтернирование.

Вспомним, что каждой перестановке приписан *знак* $\varepsilon(\sigma)$, равный ± 1 (в зависимости от четности σ).

Определение. Функция $f \in \mathcal{F}$ называется *кососимметрической*, если $\forall\ \sigma \in S_q$ выполнено

$$\sigma f = \varepsilon(\sigma)\cdot f.$$

Подмножество всех кососимметрических $f \in \mathcal{F}$ можно обозначить $\mathcal{F}^-$.

Упражнение. Докажите, что $\mathcal{F}^+ \oplus \mathcal{F}^-$ — прямая сумма. Однако, равенство $\mathcal{F} = \mathcal{F}^+ \oplus \mathcal{F}^-$, верное в случае функций от $q = 2$ переменных вообще говоря, перестает быть верным при $q > 2$.

Аналогично симметризации можно определить «усреднение со знаком», или *альтернирование* как

$$\mathrm{Alt}(f) = \frac{1}{q!}\sum_{\sigma\in S_q}\varepsilon(\sigma)\cdot\sigma f.$$

Более общо, для подгруппы в $H \leqslant S_q$

$$\mathrm{Alt}_H(f) = \frac{1}{|H|} \sum_{\sigma \in H} \varepsilon(\sigma) \cdot \sigma f.$$

Отметим, что если H состоит только из четных перестановок, то $\mathrm{Alt}_H(f) = \mathrm{Sym}_H(f)$.

Следующее предложение покавается, $\mathrm{Sym} : \mathcal{F} \to \mathcal{F}$ является проектором на $\mathcal{F}^-$.

Предложение 3.2. *1)* $\forall\ f \in \mathcal{F}$ *выполнено* $\mathrm{Alt}(f) \in \mathcal{F}^-$*;*
2) если $f \in \mathcal{F}^-$*, то* $\mathrm{Alt}(f) = f$.

Лемма. *Пусть* $\sigma_0 \in H \leqslant S_q$ *(H — подгруппа в S_q). Тогда* $\forall\ f \in \mathcal{F}$ *выполнено*

$$\mathrm{Alt}_H(\sigma_0 f) = \varepsilon(\sigma_0) \cdot \mathrm{Alt}_H(f).$$

Лемма. *Пусть* $H \leqslant G \leqslant S_q$ *(две подгруппы S_q, одна вложена в другую). Тогда* $\forall\ f \in \mathcal{F}$ *выполнено*

$$\mathrm{Alt}_G(\mathrm{Alt}_H(f)) = \mathrm{Alt}_G(f).$$

Применение к тензорам.

Так как тензоры из являются функциями $\underbrace{V \times V \times \ldots \times V}_{q} \to \mathbb{F}$, согласно общей излагаемой теории, на пространстве $T_q^0 = T_q^0(V)$ определяется изоморфизм $T_q^0 \to T_q^0$ перестановки аргументов $\tau \mapsto \sigma\tau$. Далее определяются линейные проекторы $T_q^0 \to T_q^0$ симметризация $\tau \mapsto \mathrm{Sym}(\tau)$ и альтернирование $\tau \mapsto \mathrm{Alt}(\tau)$.

Множества симметрических и кососимметрических тензоров обозначаем (наряду с $T_q^0(V)^+$ и $T_q^0(V)^-$) через $S_q^0(V)$ и $\Lambda_q^0(V)$ соответственно. Ясно, что $S_q^0(V)$ и $\Lambda_q^0(V)$ — подпространства в $T_q^0(V)$.

В координтанатах: пусть $\tau \underset{\mathrm{e}}{\longrightarrow} t_{ik\ldots}$. Тогда

$$\sigma\tau \underset{\mathrm{e}}{\longrightarrow} t_{\sigma(i)\sigma(k)\ldots}.$$

(Простой пример: транспонирование матрицы билинейной формы: $b_{ij} \mapsto b_{ji}$.)

Для обозначения компонент тензора после симметризации и альтернирования применяются иногда круглые и квадратные скобки. Так,

$$\mathrm{Sym}(\tau) \underset{\mathrm{e}}{\longrightarrow} t_{(ik\ldots)} = \frac{1}{q!} \sum_{\sigma \in S_q} t_{\sigma(i)\sigma(k)\ldots};$$

$$\mathrm{Alt}(\tau) \underset{\mathrm{e}}{\longrightarrow} t_{[ik\ldots]} = \frac{1}{q!} \sum_{\sigma \in S_q} \varepsilon(\sigma) \cdot t_{\sigma(i)\sigma(k)\ldots}.$$

Рассмотрим тензоры $\alpha \in T_p^0$ и $\beta \in T_q^0$. Тогда, согласно определению,

$$(\alpha \otimes \beta)(\underbrace{\mathbf{a}, \mathbf{b}, \ldots}_{p}, \underbrace{\mathbf{c}, \mathbf{d}, \ldots}_{q}) = \alpha(\underbrace{\mathbf{a}, \mathbf{b}, \ldots}_{p}) \cdot \beta(\underbrace{\mathbf{c}, \mathbf{d}, \ldots}_{q}) = (\beta \otimes \alpha)(\underbrace{\mathbf{c}, \mathbf{d}, \ldots}_{q}, \underbrace{\mathbf{a}, \mathbf{b}, \ldots}_{p}).$$

Отсюда

$$\beta \otimes \alpha = \sigma_{q,p}(\alpha \otimes \beta),$$

где перестановка $\sigma_{q,p}$ задается как $\begin{pmatrix} 1 & 2 & \ldots & q & q+1 & q+2 & \ldots & p+q \\ p+1 & p+2 & \ldots & p+q & 1 & 2 & \ldots & p \end{pmatrix}$.

Так как, $\varepsilon(\sigma_{q,p}) = (-1)^{pq}$, с учетом лемм и , получаем следующее.

Предложение 3.3. $\forall\ \alpha \in T_p^0$ *и* $\beta \in T_q^0$ *выполнено*

$$\mathrm{Sym}(\alpha \otimes \beta) = \mathrm{Sym}(\beta \otimes \alpha).$$

Предложение 3.4. *$\forall\ \alpha \in T_p^0$ и $\beta \in T_q^0$ выполнено*

$$\operatorname{Alt}(\alpha \otimes \beta) = (-1)^{pq} \operatorname{Alt}(\beta \otimes \alpha).$$

Предложение 3.5. *$\forall\ \alpha \in T_p^0$ и $\beta \in T_q^0$ выполнено*

$$\operatorname{Sym}((\operatorname{Sym}(\alpha) \otimes \beta) = \operatorname{Sym}(\alpha \otimes \operatorname{Sym}(\beta)) = \operatorname{Sym}(\alpha \otimes \beta).$$

$\triangleright$ СХЕМА. Заметим, что $\operatorname{Sym}(\alpha) \otimes \beta = \operatorname{Sym}_H(\alpha \otimes \beta)$, где $S_p \cong H \leqslant S_{p+q}$ — подгруппа перестановок, для которых последние q элементов неподвижны. $\square$

Предложение 3.6. *$\forall\ \alpha \in T_p^0$ и $\beta \in T_q^0$ выполнено*

$$\operatorname{Alt}((\operatorname{Alt}(\alpha) \otimes \beta) = \operatorname{Alt}(\alpha \otimes \operatorname{Alt}(\beta)) = \operatorname{Alt}(\alpha \otimes \beta).$$

Симметрические тензоры. Базис в S_q^0.

Обозначим $q! \operatorname{Sym}(\mathbf{e}^{i_1} \otimes \mathbf{e}^{i_2} \otimes \ldots \otimes \mathbf{e}^{i_q})$ коротко $\mathbf{e}^{i_1}\mathbf{e}^{i_2}\ldots\mathbf{e}^{i_q}$.

> Например,
> $\mathbf{e}^3\mathbf{e}^1\mathbf{e}^3 = 2(\mathbf{e}^3 \otimes \mathbf{e}^1 \otimes \mathbf{e}^3 + \mathbf{e}^1 \otimes \mathbf{e}^3 \otimes \mathbf{e}^3 + \mathbf{e}^3 \otimes \mathbf{e}^3 \otimes \mathbf{e}^1)$.
> $\mathbf{e}^3\mathbf{e}^1\mathbf{e}^2 = \mathbf{e}^1 \otimes \mathbf{e}^2 \otimes \mathbf{e}^3 + \mathbf{e}^1 \otimes \mathbf{e}^3 \otimes \mathbf{e}^2 + \mathbf{e}^2 \otimes \mathbf{e}^1 \otimes \mathbf{e}^3 + \mathbf{e}^2 \otimes \mathbf{e}^3 \otimes \mathbf{e}^1 + \mathbf{e}^3 \otimes \mathbf{e}^1 \otimes \mathbf{e}^2 + \mathbf{e}^3 \otimes \mathbf{e}^2 \otimes \mathbf{e}^1$.

Из-за коммутативности $\mathbf{e}^{i_1}\mathbf{e}^{i_2}\ldots\mathbf{e}^{i_q}$ приводится к виду $(\mathbf{e}^1)^{k_1}\ldots(\mathbf{e}^n)^{k_n}$, где k_i — целые неотрицательные числа, сумма которых равна q.

Предложение 3.7. *1) $\{(\mathbf{e}^1)^{k_1}\ldots(\mathbf{e}^n)^{k_n} \mid k_i \in \mathbb{Z}_+, \sum k_i = q\}$ — базис в пространстве S_q^0;*
2) $\dim S_q^0 = C_{n+q-1}^q$.

Симметрическая алгебра.

Для $\tau \in S_p^0$, $\tau' \in S_q^0$ определяется *симметрическое умножение* по правилу:

$$\tau \vee \tau' = \frac{(p+q)!}{p!q!} \operatorname{Sym}(\tau \otimes \tau').$$

(иногда определяется без нормирующего множителя). Из предложений 3.3 и 3.5 следует коммутативность и ассоциативность симметрического умножения.

Иногда знак $\vee$ опускается, что согласуется с обозначением $\mathbf{e}^{i_1}\mathbf{e}^{i_2}\ldots\mathbf{e}^{i_q}$:

$$\mathbf{e}^{i_1} \vee \mathbf{e}^{i_2} \vee \ldots \vee \mathbf{e}^{i_q} = \mathbf{e}^{i_1}\mathbf{e}^{i_2}\ldots\mathbf{e}^{i_q}.$$

$$S(V) = \bigoplus_{q=0}^{\infty} S_q^0.$$

$S(V)$ — *симметрическая алгебра* относительно $+, \vee$. (ассоциативная, коммутативная)

Кососимметрические тензоры. Базис в Λ_q^0.

Обозначим (временно) $q! \operatorname{Alt}(\mathbf{e}^{i_1} \otimes \mathbf{e}^{i_2} \otimes \ldots \otimes \mathbf{e}^{i_q})$ коротко $\mathbf{e}^{i_1} \circ \mathbf{e}^{i_2} \circ \ldots \circ \mathbf{e}^{i_q}$.

> Например,
> $\mathbf{e}^3 \circ \mathbf{e}^1 \circ \mathbf{e}^3 = 0$;
> $\mathbf{e}^1 \circ \mathbf{e}^2 \circ \mathbf{e}^3 = \mathbf{e}^1 \otimes \mathbf{e}^2 \otimes \mathbf{e}^3 - \mathbf{e}^1 \otimes \mathbf{e}^3 \otimes \mathbf{e}^2 + \mathbf{e}^2 \otimes \mathbf{e}^3 \otimes \mathbf{e}^1 - \mathbf{e}^2 \otimes \mathbf{e}^1 \otimes \mathbf{e}^3 + \mathbf{e}^3 \otimes \mathbf{e}^1 \otimes \mathbf{e}^2 - \mathbf{e}^3 \otimes \mathbf{e}^2 \otimes \mathbf{e}^1$.

В частности, в трехмерном пространстве $(\mathbf{e}^1 \circ \mathbf{e}^2 \circ \mathbf{e}^3)(\mathbf{a}, \mathbf{b}, \mathbf{c})$ = детерминант из координатных столбцов векторов $\mathbf{a}, \mathbf{b}, \mathbf{c}$.

Из-за кососкоммутативности $\mathbf{e}^{i_1} \circ \mathbf{e}^{i_2} \circ \ldots \circ \mathbf{e}^{i_q}$ приводится к виду 0 либо $\pm\mathbf{e}^{j_1} \circ \mathbf{e}^{j_2} \circ \ldots \circ \mathbf{e}^{j_q}$, где $j_1 < j_2 < \ldots < j_q$.

Предложение 3.8. *1) $\{\mathbf{e}^{j_1} \circ \mathbf{e}^{j_2} \circ \ldots \circ \mathbf{e}^{j_q} \mid j_1 < j_2 < \ldots < j_q\}$ — базис в пространстве Λ_q^0;*
2) $\dim \Lambda_q^0 = C_n^q$ при $q \leqslant n$ и $\Lambda_q^0 = O$ при $q > n$.

Внешняя алгебра.

Для $\tau \in \Lambda^0_p$, $\tau' \in \Lambda^0_q$ определяется *внешнее умножение* по правилу:

$$\tau \wedge \tau' = \frac{(p+q)!}{p!q!}\,\mathrm{Alt}(\tau \otimes \tau').$$

(иногда определяется без нормирующего множителя). Из предложений 3.4 и 3.6 следует кососоммутативность:

$$\tau \wedge \tau' = (-1)^{pq}\tau' \wedge \tau$$

(в частности, при нечетном p $\tau \wedge \tau = 0$) и ассоциативность внешнего умножения.

$$\mathbf{e}^{i_1} \circ \mathbf{e}^{i_2} \circ \ldots \circ \mathbf{e}^{i_q} = \mathbf{e}^{i_1} \wedge \mathbf{e}^{i_2} \wedge \ldots \wedge \mathbf{e}^{i_q}.$$

$$\Lambda(V) = \bigoplus_{q=0}^{n} \Lambda^0_q.$$

$\Lambda(V)$ — *внешняя алгебра* относительно $+, \wedge$. (ассоциативная, кососоммутативная)