Введение в линейную алгебру

П.А. Кожевников

17 июня 2017 г.

Оглавление

1 Векторные пространства **7**
§ 1. Векторные пространства и подпространства 7
Аксиомы и их следствия 7
Подпространства 8
Линейные комбинации 9
Линейная оболочка 9
Примеры 10
§ 2. Линейная зависимость. Размерность и ранг. Базис 11
Линейная зависимость 11
Ранг и размерность 12
Базис 14
Координаты 15
Замена базиса и координат 16
Примеры 16
§ 3. Сумма подпространств. Прямая сумма 17
Сумма подпространств 17
Прямая сумма 18
Прямое дополнение. Проекции 20
Формула размерностей суммы и пересечения 20
Примеры 21

2 Линейные отображения **23**
§ 1. Определение. Операции над линейными отображениями. Изоморфизм 23
Определение и его следствия 23
Композиция (произведение). Обратное отображение. Изоморфизм . . 25
Сумма и произведение на число 26
Примеры 27
§ 2. Матрица линейного отображения 28
Определение. Координатная запись линейного отображения 28
Матрица перехода и матрица преобразования 29
Связь между операциями над отображениями и матрицами 29
Изменение матрицы при замене базиса 30
Примеры 30
§ 3. Образ и ядро 30
Образ 30
Ядро 31
Связь между размерностями ядра и образа 32
Примеры 32
§ 4. Сопряженное пространство 33
Определение 33

Взаимный базис 33
Примеры 34
§ 5. Структура линейного преобразования 34
Инвариантные подпространства 34
Собственные значения, собственные векторы, собственные подпространства. 35
Характеристический многочлен 36
Диагонализируемость и треугольный вид 38
Примеры 39

3 Билинейные и квадратичные формы **41**
§ 1. Билинейные формы. Матрица билинейной формы 41
Определение 41
Матрица и билинейной формы. Координатная запись 41
Изменение матрицы при замене базиса 42
Сужение 43
§ 2. Симметричные билинейные формы. Квадратичные формы 43
Симметричные билинейные формы 43
Квадратичные формы 44
§ 3. Диагональный и канонический вид квадратичной формы 45
§ 4. Знакоопределенные формы. Индексы инерции 45

4 Евклидовы и унитарные пространства **49**
§ 1. Скалярное произведение. Матрица Грама 49
Определения 49
Матрица Грама 50
§ 2. Ортогональные системы векторов. Ортогональное дополнение. Ортогонализация 51
Ортогональные системы векторов. ОНБ 51
Переход от ОНБ к ОНБ. Ортогональные и унитарные матрицы . . . 52
Ортогональные подпространства 53
Ортогональное дополнение. Ортогональная проекция 53
Ортогонализация 55
§ 3. Сопряженное преобразование. Самосопряженные преобразования. 55
Сопряженное преобразование 55
Самосопряженные преобразования 56
§ 4. Изоморфизм. Ортогональные и унитарные преобразования 58
Изоморфизм евклидовых (унитарных) пространств 58
§ 5. Ортогональные и унитарные преобразования 58
Полярное разложение 60
§ 6. Квадратичные формы 60
Список обозначений61

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 $\psi : \mathbb{R} \to \mathbb{R}$, что $\forall x \in \mathbb{R}\ \psi(x) = 2x$.

> Пусть V — произвольное множество, на котором определены: бинарная операция (которую называем *сложением*, результат применения сложения к $\mathbf{a}, \mathbf{b} \in V$ обозначаем $\mathbf{a} + \mathbf{b}$) и унарная операция для каждого $\lambda \in \mathbb{R}$ (которую называем *умножением* на λ, результат применения умножения на λ к $\mathbf{a} \in V$ обозначаем $\lambda \cdot \mathbf{a}$ или $\lambda\mathbf{a}$).
> V называется *векторным пространством* (или линейным пространством), 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. $\exists\ \mathbf{o} \in V\ \forall\ \mathbf{a} \in V$: $\mathbf{a} + \mathbf{o} = \mathbf{o} + \mathbf{a} = \mathbf{a}$;
> V3. $\forall\ \mathbf{a} \in V\ \exists\ \mathbf{x} \in V$: $\mathbf{x} + \mathbf{a} = \mathbf{a} + \mathbf{x} = \mathbf{o}$;
> V4. $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a} \quad (\forall\ \mathbf{a}, \mathbf{b} \in V)$;
> V5. $\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b} \quad (\forall\ \mathbf{a}, \mathbf{b} \in V,\ \forall\ \lambda \in \mathbb{R}\ )$;
> V6. $(\lambda + \mu)\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a} \quad (\forall\ \mathbf{a} \in V,\ \forall\ \lambda, \mu \in \mathbb{R}\ )$;
> V7. $1 \cdot \mathbf{a} = \mathbf{a} \quad (\forall\ \mathbf{a} \in V)$;
> V8. $(\lambda\mu)\mathbf{a} = \lambda(\mu\mathbf{a}) \quad (\forall\ \mathbf{a} \in V,\ \forall\ \lambda, \mu \in \mathbb{R}\ )$.

На протяжении главы V будет обозначать векторное пространство. Элементы векторного пространства V будем называть *векторами* (независимо от их природы). Вектор $\mathbf{o}$ из аксиомы V2 называется *нулевым* вектором (также $\mathbf{o}$ можно назвать нейтральным элементом относительно сложения). 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 $\mathbf{a} - \mathbf{a} = \mathbf{o}$. [1]

[1]Аксиомы V1 и V4 называются свойстами *ассоциативности* и *коммутативности* сложения, V5 и V6 можно назвать свойством линейности, или дистрибутивности (по векторам и по константам соответственно), V7 — условие нормировки, V8 — смешанная ассоциативность умножения.
В аксиомах V1—V4 встречается лишь операция сложения. Множество V с операцией "+", удовлетворяю-

Замечание. *На самом деле здесь мы определили лишь векторные пространства над полем* $\mathbb{R}$. *Можно изменить множество констант на поле* $\mathbb{C}$ *комплексных чисел или вообще на произвольное поле* $\mathbb{F}$. *Теория в этой главе обобщается без каких-либо изменений на векторные пространства над произвольным полем.*

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 еще некоторые следствия аксиом.

Предложение 1.1. *1. Нулевой вектор единственный;*
2. $\forall\ \mathbf{a} \in V$ *противоположный ему вектор единственный;*
3. $\mathbf{a}+\mathbf{b}=\mathbf{a}+\mathbf{c} \Leftrightarrow \mathbf{b}=\mathbf{c} \quad (\forall\ \mathbf{a},\mathbf{b},\mathbf{c} \in V)$ *— закон сокращения;*
4. $0 \cdot \mathbf{a} = \lambda \cdot \mathbf{o} = \mathbf{o} \quad (\forall\ \mathbf{a} \in V,\ \forall\ \lambda \in \mathbb{R})$*;*
5. $-(\lambda\mathbf{a}) = (-\lambda)\mathbf{a} = \lambda(-\mathbf{a}) \quad (\forall\ \mathbf{a} \in V,\ \forall\ \lambda \in \mathbb{R})$*;*
6. $\lambda(\mathbf{a}-\mathbf{b}) = \lambda\mathbf{a} - \lambda\mathbf{b} \quad (\forall\ \mathbf{a},\mathbf{b} \in V,\ \forall\ \lambda \in \mathbb{R})$*;*
7. $(\lambda-\mu)\mathbf{a} = \lambda\mathbf{a} - \mu\mathbf{a} \quad (\forall\ \mathbf{a} \in V,\ \forall\ \lambda,\mu \in \mathbb{R})$ *.*

$\triangleright$ 1. Пусть $\mathbf{o}$ и $\mathbf{o}'$ — два нулевых вектора. Тогда $\mathbf{o} = \mathbf{o}+\mathbf{o}' = \mathbf{o}'$, то есть $\mathbf{o}$ и $\mathbf{o}'$ совпадают.

2. Пусть векторы $\mathbf{x}$ и $\mathbf{y}$ оба являются противоположными для вектора $\mathbf{a}$. Тогда $\mathbf{x} = \mathbf{x}+\mathbf{o} = \mathbf{x}+(\mathbf{a}+\mathbf{y}) = (\mathbf{x}+\mathbf{a})+\mathbf{y} = \mathbf{o}+\mathbf{y} = \mathbf{y}$, то есть $\mathbf{x}$ и $\mathbf{y}$ совпадают.

3. $\mathbf{a}+\mathbf{b} = \mathbf{a}+\mathbf{c} \Rightarrow -\mathbf{a}+(\mathbf{a}+\mathbf{b}) = -\mathbf{a}+(\mathbf{a}+\mathbf{c}) \Rightarrow (-\mathbf{a}+\mathbf{a})+\mathbf{b} = (-\mathbf{a}+\mathbf{a})+\mathbf{c} \Rightarrow \mathbf{o}+\mathbf{b} = \mathbf{o}+\mathbf{c} \Rightarrow \mathbf{b} = \mathbf{c}$. Обратное следствие очевидно.

4. $0 \cdot \mathbf{a} = (0+0) \cdot \mathbf{a} = 0 \cdot \mathbf{a} + 0 \cdot \mathbf{a}$. С другой стороны, $0 \cdot \mathbf{a} = \mathbf{o} + 0 \cdot \mathbf{a}$. По закону сокращения (см. 3) $0 \cdot \mathbf{a} = \mathbf{o}$. Аналогично $\lambda \cdot \mathbf{o} = \lambda(\mathbf{o}+\mathbf{o}) = \lambda \cdot \mathbf{o} + \lambda \cdot \mathbf{o}$, откуда $\lambda \cdot \mathbf{o} = \mathbf{o}$.

5. $(-\lambda)\mathbf{a} + \lambda\mathbf{a} = (-\lambda+\lambda)\mathbf{a} = 0 \cdot \mathbf{a} = \mathbf{o}$, поэтому $(-\lambda)\mathbf{a}$ является противоположным для $\lambda\mathbf{a}$, то есть равен $-(\lambda\mathbf{a})$. Аналогично $\lambda(-\mathbf{a}) + \lambda\mathbf{a} = \lambda(-\mathbf{a}+\mathbf{a}) = \lambda \cdot \mathbf{o} = \mathbf{o}$, поэтому $\lambda(-\mathbf{a}) = -(\lambda\mathbf{a})$.

6. Вытекает из утверждения 5 и определения вычитания векторов: $\lambda(\mathbf{a}-\mathbf{b}) = \lambda(\mathbf{a}+(-\mathbf{b})) = \lambda\mathbf{a} + \lambda(-\mathbf{b}) = \lambda\mathbf{a} - \lambda\mathbf{b}$.

7. Аналогично 6. $\square$

Подпространства

> Непустое подмножество U векторного пространства V называется *подпространством*, если $\forall\ \mathbf{a},\mathbf{b} \in U$, $\forall\ \lambda \in \mathbb{R}$ 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 вектороное пространство V содержит *тривиальные* подпространства V и $O = \{\mathbf{o}\}$ (*нулевое* подпространство).

Предложение 1.2. *Пусть* $U \leqslant V$*, тогда* $\mathbf{o} \in U$*.*

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

щее только аксиоме V1, называется *полугруппой*; удовлетворяющее аксиомам V1—V3 — *группой*; аксиомам V1—V4 — *коммутативной*, или *абелевой* группой.

Предложение 1.3. *Пересечение подпространств является подпространством.*

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

> 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 $\mathbf{o}$.

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

$\triangleright$ 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.

> *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 $\mathbf{a}_1, \dots, \mathbf{a}_k$ обозначается $\langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle$. Формальная запись опеделения:
$\langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle := \left\{ \sum\limits_{i=1}^{k} \lambda_i \mathbf{a}_i \,|\, \lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R} \right\}$.

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

> *Линейной оболочкой* системы векторов $\mathcal{A}$ называется множество всех векторов, каждый из которых линейно выражается через несколько векторов из $\mathcal{A}$.
> Условимся также считать, что $\langle \varnothing \rangle = O$.

Обозначение линейной оболочки: $\langle \mathcal{A} \rangle$. Очевидно, для любого подмножества $\mathcal{A} \subset V$ выполнено $\mathcal{A} \subset \langle \mathcal{A} \rangle$. Отметим, что везде рассматриваются линейные комбинации из конечного количества слагаемых, хотя «длина» линейной комбинации не ограничивается. Формально, $\langle \mathcal{A} \rangle := \left\{ \sum\limits_{i=1}^{k} \lambda_i \mathbf{a}_i \,|\, k \in \mathbb{N},\, \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k \in \mathcal{A},\, \lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R} \right\}$.

Предложение 1.5. *$\langle \mathcal{A} \rangle$ 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 $\mathbf{a}_i, \mathbf{b}_j \in \mathcal{A}$, $\alpha_i, \beta_j \in \mathbb{R}$. 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 1.4 показывает, что тогда и $\langle \mathcal{A} \rangle \subset U$. Таким образом, предложение 1.5 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. 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 $\mathbf{M}_{m \times n}$ матриц $m \times n$ (заполненных вещественными числами) — векторное пространство относительно операций сложения матриц и умножения на число. $\mathbb{R}^n$ отождествляем с множеством столбцов $\mathbf{M}_{n \times 1}(\mathbb{R})$.
> Пусть $\mathbf{M}^+_{n \times n} = \{A \in \mathbf{M}_{n \times n} \,|\, A^T = A\}$ — множество симметричных матриц $n \times n$, $\mathbf{M}^-_{n \times n} = \{A \in \mathbf{M}_{n \times n} \,|\, A^T = -A\}$ — множество кососимметричных матриц $n \times n$.
> Тогда $\mathbf{M}^+_{n \times n} \leqslant \mathbf{M}_{n \times n}$, $\mathbf{M}^-_{n \times n} \leqslant \mathbf{M}_{n \times n}$.

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 $A_i X = O$, $i = 1, 2, \ldots, k$ — несколько однородных СЛУ относительно $X \in \mathbf{M}_{n\times 1}$, то СЛУ, представляющая собой объединение этих систем, имеет в качестве множества решений пересечение $\bigcap\limits_{i=1}^{k} \mathrm{Sol}(A_i X = O)$.

III.1. Пусть S — некоторое множество. Множество $\mathbf{F}(S)$ всех функций $f : S \to \mathbb{R}$ — векторное пространство относительно операций сложения функций и умножения функции на число.
Пусть $\mathbf{F}^+(\mathbb{R}) = \{f \in \mathbf{F}(\mathbb{R}) \,|\, \forall\, x \in \mathbb{R}\ f(-x) = f(x)\}$ — 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) многоленов $\mathbf{P}$ может быть задано линейной оболочкой: $\mathbf{P} = \langle 1, x, x^2, x^3, \ldots\rangle$.
$\mathbf{P} \geqslant \mathbf{P}_n = \langle 1, x, x^2, \ldots, x^n\rangle$, где $\mathbf{P}_n$ — множество многочленов f степени $\deg f \leqslant n$.

III.3. В множестве $\mathbf{F}(\mathbb{N}) = \{(f_1, f_2, \ldots) \,|\, f_i \in \mathbb{R}\}$ 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 n раз дифференцируемых функций $\mathbb{R} \to \mathbb{R}$.

§ 2. Линейная зависимость. Размерность и ранг. Базис

Линейная зависимость

Система векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ называется *линейно зависимой*, если некоторая их нетривиальная линейная комбинация равна $\mathbf{o}$, 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 $\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\ldots+\lambda_k\mathbf{a}_k=\mathbf{o}$, 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 $\mathbf{o}$. $\square$

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 $\mathbf{o}$. 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 $\mathbf{o}$.

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

Следствие. *Любая система векторов, содержащая* $\mathbf{o}$, *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 $\mathbf{o}$. 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 $\sum\limits_{i=1}^{k}(\lambda_i-\mu_i)\mathbf{a}_i=\mathbf{o}$. 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

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.

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. $\square$

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 $\mu\mathbf{a} + \sum\limits_{i=1}^{r} \lambda_i \mathbf{a}_i = \mathbf{o}$. 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 $\operatorname{rg}(\mathbf{a}_1, \ldots$ $\ldots, \mathbf{a}_k) = \operatorname{rg}(\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{b})$.*

$\triangleright$ 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 $U = W$.*

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. $\square$

Базис

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

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.

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. (То есть если $\mathbf{a} = \mathrm{e}X$ $\mathbf{b} = \mathrm{e}Y$, то $\mathbf{a} + \mathbf{b} = \mathrm{e}(X + Y)$, $\lambda\mathbf{a} = \mathrm{e}(\lambda X)$.*[2]

$\triangleright$ 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 возникает линейное взаимно-однозначное соответствие между V и $\mathbf{M}_{n\times 1}$ — сопоставление вектору его координатного столбца.

[2]Заметим, что утверждение предложения означает возможность раскрытия скобок естественным образом в компактных обозначениях.

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 $(n = \dim V)$.

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.

Из определения ясно, что матрица перехода от базиса e к самому себе — это единичная матрица.

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'.

$\triangleright$ По условию $\mathbf{a} = \mathrm{e}X = \mathrm{e}'X'$. Так как $\mathrm{e}' = \mathrm{e}S$, имеем $\mathrm{e}X = \mathrm{e}SX'$,[3] и из закона сокращения (e — линейно независимая система) получаем $X = SX'$. $\square$

Предложение 2.10. *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.

$\triangleright$ $\mathrm{e}'' = \mathrm{e}'R = (\mathrm{e}S)R = \mathrm{e}(SR)$. $\square$

Следствие. *Матрица* S *перехода от базиса* e *к базису* e' *обратима, и* S^{-1} *является матрицей перехода от базиса* e' *к базису* e.

$\triangleright$ Положив в предложении 2.10 $\mathrm{e}'' = \mathrm{e}$, получим $SR = E$. $\square$

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 $\dim \mathbf{M}_{m\times n} = mn$. В частности, $\dim \mathbf{M}_{n\times 1} = n$.

II.2. Подпространство $U \leqslant \mathbb{R}^n = \mathbf{M}_{n\times 1}$ 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, составленная из столбцов $a_{\bullet 1}, \ldots, A_{\bullet k}$.

[3] На самом деле здесь используем ассоциативность умножения матриц.

II.3. Подпространство $U \leqslant \mathbb{R}^n$ может быть задано как множество решений некоторой однородной системы линейных уравнений $AX = O$, т.е. $U = \mathrm{Sol}(AX = O)$. При этом если $\mathrm{rg}\, A = r$, то $\dim U = n - r$. ФСР (фундаментальная система решений) системы $AX = O$ — это базис в $U = \mathrm{Sol}(AX = O)$.
В указанном виде может быть задано любое подпространство $U \leqslant \mathbb{R}^n$ (имеется алгоритм получения СЛУ, для которой данная линейная оболочка столбцов явлется множеством решений; см. также предложение 2.10 главы **??**).

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

III.2. Рассмотрим пространство фибоначчиевых последовательностей $V = \{(f_1, f_2, \ldots) \mid f_{i+1} = f_i + f_{i-1},\ i = 2, 3, \ldots\}$. Каждая последовательность из V однозначно задается первыми двумя членами f_1, f_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 $(f_1, f_2, \ldots) \in V$ совпадает с линейной комбинацией $f_1 e_1 + f_2 e_2$ (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 $\lambda_{(i)} = (1, \lambda_i, \lambda_i^2, \ldots)$, $i = 1, 2$, образуют базис в V (очевидно, $\lambda_{(1)}$, $\lambda_{(2)}$ — линейно независимая система).
Линейное выражение (обычной) последовательности Фибоначчи $f = (1, 1, 2, 3, 5, 8, \ldots)$ через $\lambda_{(i)}$ позволит найти явную формулу для чисел Фибоначчи f_n. Равенство $f = c_1 \lambda_{(1)} + c_2 \lambda_{(2)}$ 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, $f_n = c_1 \lambda_1^{n-1} + c_2 \lambda_2^{n-1} = \frac{\lambda_1^n - \lambda_2^n}{\sqrt{5}}$.

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

Упражнение. (Треугольный базис) Дана верхнетреугольная матрица $A \in \mathbf{M}_{n\times n}$, у которой на главной диагонали ненулевые числа. Докажите, что ее столбцы образуют базис в $\mathbf{M}_{n\times 1}$.

§ 3. Сумма подпространств. Прямая сумма

Сумма подпространств

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 $\{\mathbf{a}_1 + \mathbf{a}_2 + \ldots + \mathbf{a}_k \mid \mathbf{a}_i \in \mathcal{A}_i\}$, $i = 1, \ldots, k$.

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 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* $\sum\limits_{i=1}^{k} U_i = \langle \mathcal{A}_1 \bigcup \mathcal{A}_2 \bigcup \ldots \bigcup \mathcal{A}_k \rangle$.

$\triangleright$ 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 $\mathcal{A}_1 \bigcup \mathcal{A}_2 \bigcup \ldots \bigcup \mathcal{A}_k$. $\square$

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

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

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.

> Сумма U подпространств $U_1, U_2, \ldots, U_k$ называется *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.

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 $\mathbf{o} = \mathbf{o} + \ldots + \mathbf{o} = \mathbf{a}_1 + (-\mathbf{a}_2) + \ldots + (-\mathbf{a}_k)$ — 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. $\square$

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

Упражнение. Сумма подпространств $U_1, \ldots, U_k$ — прямая сумма $\Leftrightarrow \forall\ \mathbf{a}_i \in U_i$, $\mathbf{a}_i \neq \mathbf{o}$ 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.

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. 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 подпространстве 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 $\mathbf{o}$: $\sum\limits_{i=1}^{k} \ell_i = \mathbf{o}$, 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 $\ell_1 \neq \mathbf{o}$. 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) $\exists\ \mathbf{a} \neq \mathbf{o}$: $\mathbf{a} \in U_1 \cap \overline{U_1}$. 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 $\ell_1 - \ell_2 - \ldots - \ell_k = \mathbf{o}$ (в левой части нетривиальная линейная комбинация, поскольку уже ℓ_1 — нетривиальная линейная комбинация), откуда e — линейно зависимая система. Противоречие. $\square$

Прямое дополнение. Проекции

> 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. $\square$

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 $U \oplus W = V$. $\square$

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).

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. $\square$

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), $U_2 \leqslant V$, $\dim U_2 = 2$ (т.е. U_2 — плоскость). 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 $AX = b$ с n неизвестными имеет вид $\{X_0\} + \mathrm{Sol}(AX = O)$. Если $r = \mathrm{rg}\, A$, то $\mathrm{Sol}(AX = b)$ — это $(n-r)$-мерная плоскость (или $(n-r)$-мерное *линейное многообразие*) в пространстве $\mathbb{R}^n$.

II.2. $\mathbf{M}_{n\times n} = \mathbf{M}^+_{n\times n} \bigoplus \mathbf{M}^-_{n\times n}$.
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 $A \in \mathbf{M}_{n\times m}$, $A = (a_{\bullet 1} \dots a_{\bullet m})$. Положим $U = \langle a_{\bullet 1} \dots a_{\bullet m} \rangle$, $\dim U = \mathrm{rg}\, A = r$. 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 **??**), 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.
Заметим, что $e^x = \mathrm{ch}\, x + \mathrm{sh}\, x$, при этом $f(x) = \mathrm{ch}\, x$ — четная функция, а $g(x) = \mathrm{sh}\, x$ — нечетная функция. Поэтому $f(x) = \mathrm{ch}\, x$ — это проекция функции $y(x) = e^x$ на $\mathbf{F}^+$ вдоль $\mathbf{F}^-$.

Глава 2

Линейные отображения

§ 1. Определение. Операции над линейными отображениями. Изоморфизм

Определение и его следствия

В этой главе векторные пространства (над $\mathbb{R}$ или над $\mathbb{C}$) обозначаем V, $\widetilde{V}$, $\widetilde{\widetilde{V}}$.

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

Через $L(V, \widetilde{V})$ обозначаем множество всех линейных отображений $V \to \widetilde{V}$.

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

Если $\widetilde{V} = V$, то линейное отображение $V \to \widetilde{V}$ называют также *линейным преобразованием* или *линейным оператором*.

Если $\dim \widetilde{V} = 1$, то линейное отображение $V \to \widetilde{V}$ называют также *линейным функционалом* или *линейной функцией*.

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

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

$\triangleright$ Следует из многократного применения L1, L2. $\square$

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. $\varphi(\mathbf{o}) = \mathbf{o}$;
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* $\varphi(V) = \langle \varphi(\mathbf{e}_1), \varphi(\mathbf{e}_2), \ldots, \varphi(\mathbf{e}_n) \rangle$*.*
2. $\dim \varphi(U) \leqslant \dim U$*.*

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

В следущем 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 $\sum\limits_{i=1}^{k} \lambda_k \varphi(\mathbf{a}_k) = \mathbf{o}$, тогда $\varphi\left(\sum\limits_{i=1}^{k} \lambda_i \mathbf{a}_i\right) = \mathbf{o}$. Но $\varphi(\mathbf{o}) = \mathbf{o}$, поэтому в силу инъективности, $\sum\limits_{i=1}^{k} \lambda_i \mathbf{a}_i = \mathbf{o}$. Отсюда, поскольку $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ — линейно независимая система, имеем $\lambda_1 = \ldots = \lambda_k = 0$, то есть $\sum\limits_{i=1}^{k} \lambda_k \varphi(\mathbf{a}_k)$ — 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 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 **??**) 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 $\varphi(\mathbf{a} - \mathbf{b}) = \mathbf{o}$. 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 $\varphi(\mathbf{a}-\mathbf{b})=\sum\limits_{i=1}^{n}\lambda_i\mathbf{c}_i=\mathbf{o}$. Получаем, что нетривиальная линейная комбинация векторов $\mathbf{c}_1,\mathbf{c}_2,\ldots,\mathbf{c}_n$ равна $\mathbf{o}$. Противоречие. □

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

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

Композиция (произведение). Обратное отображение. Изоморфизм

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

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

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 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: $I_V:V\to V$). 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 $V\cong\widetilde{V}$, $\widetilde{V}\cong\widetilde{\widetilde{V}}\Rightarrow V\cong\widetilde{\widetilde{V}}$. Тем самым, все векторные пространства разбиваются 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 **??** мы фактически сталкивались с изоморфизмом — сопоставлением вектору его координатного столбца (в фиксированном базисе). Значит, у нас есть «эталонное» n-мерное пространство $\mathbb{R}^n = \mathbf{M}_{n\times 1}$, 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, которые означают, что рассматриваемое отображение — изоморфизм.

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 — *сюръеция;*
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. $\square$

Сумма и произведение на число

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

> Произведением отображения $\varphi \in L(V, \widetilde{V})$ на число λ называется такое отображение $\psi : V \to \widetilde{\varphi}$, что $\forall \mathbf{a} \in V$ выполнено
> (1 вариант) $\psi(\mathbf{a}) = \lambda\varphi(\mathbf{a})$;
> (2 вариант) $\psi(\mathbf{a}) = \overline{\lambda}\varphi(\mathbf{a})$.

Результат операции сложения и умножения на число называется, как обычно, суммой и произведением на число, и обозначаются обычным образом: $\varphi + \widetilde{\varphi}$ и $\lambda\varphi$.

Так, для пространства над $\mathbb{C}$ имеются два варианта определения $\lambda\varphi$.

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

$\triangleright$ 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. $L(V,\widetilde{V})$ *— векторное пространство относительно введенных выше операций сложения и умножения на число (при каждом из двух вариантов определения умножения на число).*

$\triangleright$ Проверяется непосредственно. $\square$

Далее через $L(V,\widetilde{V})$ обозначаем векторное пространство линейных отображений $V \to \widetilde{V}$ с первым вариантом определения умножения на число, а через $\overline{L}(V,\widetilde{V})$ — то же пространство со вторым вариантом определения умножения на число.

Предложение 1.10. *Пусть* $\varphi, \psi \in L(V,\widetilde{V})$, $\widetilde{\varphi}, \widetilde{\psi} \in L(\widetilde{V},\widetilde{\widetilde{V}})$. *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. $\lambda(\widetilde{\varphi}\varphi) = (\lambda\widetilde{\varphi})\varphi = \widetilde{\varphi}(\lambda\varphi)$.

$\triangleright$ Проверяется непосредственно. $\square$

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

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 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 U_2.

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

> I. Пусть дано аффинное преобразование плоскости $\varphi : \mathbf{R}^2 \to \mathbf{R}^2$, для которого начало координат — неподвижная точка (как обычно, отождествляем с радиус-векторами). Тогда φ — линейное и биективное преобразование.

> 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 $V = \mathbf{C}^\infty(\mathbb{R})$. *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 $V = \mathbf{C}^\infty(\mathbb{R})$) можно рассмотреть линейные операторы взятия частной производной $\dfrac{\partial}{\partial x_i}$.
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 $V = \mathbf{F}(\mathbb{N}) = \{(f_1, f_2, \ldots) \mid f_i \in \mathbb{R}\}$ — пространство числовых последовательностей. *Оператор сдвига (влево)* $\varphi : V \to V$ определяется равенством $\varphi(f_1, f_2, \ldots) = (f_2, f_3, \ldots)$. Определим *оператор первой разности* $\Delta = \varphi - I_V$, так что $\Delta(f_1, f_2, \ldots) = (f_2 - f_1, f_3 - f_2, \ldots)$.
Для $k \in \mathbb{N}$ степень Δ^k называют *оператором k-ой разности.*

§ 2. Матрица линейного отображения

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

Определение. Координатная запись линейного отображения

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 $\varphi \underset{\mathrm{e,f}}{\longrightarrow} A$.

Предложение 2.1. *Пусть в пространствах V и $\widetilde{V}$ зафиксированы базисы* e *и* f *соответственно. Тогда отображение $\varphi \underset{\mathrm{e,f}}{\longrightarrow} A$ — 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 $\varphi \underset{\mathrm{e,f}}{\longrightarrow} A$. 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 $\varphi \xrightarrow[\mathrm{e,f}]{} A$.*

$\rhd$ 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 $\mathrm{e} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ к $\mathrm{e}' = (\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ 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 $i = 1, \ldots, n$ (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 $i = 1, \ldots, n$; $\varphi \xrightarrow[\mathrm{e,e}]{} A$. Тогда A — матрица перехода от* e *к* e$'$.

$\rhd$ Достаточно сопоставить определения матрицы линейного отображения и матрицы перехода. $\square$

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 $\mathbf{M}_{m\times n}$. (Аналогично, $\varphi \to \overline{A}$ — изоморфизм линейных пространств $\varphi \in \overline{L}(V, \widetilde{V})$ и $\mathbf{M}_{m\times n}$).*

$\rhd$ Непосредственная проверка. $\square$

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

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

$\rhd$ 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). $\square$

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

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

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

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.

Упражнение. Получите доказательство теоремы о ранге произведения матриц, используя линейные отображения.

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

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

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 **??**) $X = SX'$, $Y = RY'$. 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 3.2.

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. *Пусть* $\varphi \in L(V, \widetilde{V})$, *и* $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ *— базис в* V. *Тогда* $\operatorname{Im}\varphi = \langle \varphi(\mathbf{e}_1), \varphi(\mathbf{e}_2), \ldots, \varphi(\mathbf{e}_n) \rangle$.

$\triangleright$ Следует из 1.2. $\square$

Теорема 3.2 (координатное описание образа)**.** *Пусть* $\varphi \in L(V, \widetilde{V})$, e, f *— базисы в* V *и* $\widetilde{V}$. *Пусть* $\varphi \underset{\mathrm{e,f}}{\longrightarrow} A$. *Для вектора* $\mathbf{b} \in \widetilde{V}$, $\mathbf{b} = \mathrm{f}Y$ *выполнено:* $\mathbf{b} \in \operatorname{Im}\varphi \Leftrightarrow Y \in \langle a_{\bullet 1}, a_{\bullet 2}, \ldots, a_{\bullet n} \rangle$.

$\triangleright$ Следует из сопоставления определелния матрицы линейного отображения и предложения 1.4. $\square$

Иначе говоря, теорема утверждает, что в терминах координатных столбцов $\operatorname{Im}\varphi$ задается как линейная оболочка столбцов матрицы линейного отображения.

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

В частности, мы получаем еще одно обоснование того, что $\operatorname{rg} A$ не зависит от выбора базисов в пространствах V и $\widetilde{V}$.

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* $\{\mathbf{a} \in V \mid \varphi(\mathbf{a}) \in \widetilde{U}\} \leqslant V$.

$\triangleright$ 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. $\square$

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

Обозначение для ядра — $\operatorname{Ker}\varphi$. Опеределение можно переформулировать так: $\operatorname{Ker}\varphi$ — это полный прообраз нулевого подпространства.

Предложение 3.2. *Если* $\varphi \in L(V, \widetilde{V})$, *то* $\operatorname{Ker}\varphi \leqslant V$.

$\triangleright$ Это частный случай предложения 3.1. $\square$

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

Предложение 3.3. $\varphi \in L(V, \widetilde{V})$ *инъективно* $\Leftrightarrow \operatorname{Ker}\varphi = O$.

$\triangleright$ $\boxed{\Rightarrow}$ Пусть $\mathbf{a} \in \operatorname{Ker}\varphi$, то есть $\varphi(\mathbf{a}) = \mathbf{o}$. Но также $\varphi(\mathbf{o}) = \mathbf{o}$, и поскольку φ инъективно, имеем $\mathbf{a} = \mathbf{o}$.
$\boxed{\Rightarrow}$ Пусть $\varphi(\mathbf{a}) = \varphi(\mathbf{b})$. Тогда $\varphi(\mathbf{a} - \mathbf{b}) = \mathbf{o}$. С учетом $\operatorname{Ker}\varphi = O$ имеем $\mathbf{a} - \mathbf{b} = \mathbf{o}$, то есть $\mathbf{a} = \mathbf{b}$. Тем самым, $\varphi(\mathbf{a}) = \varphi(\mathbf{b}) \Rightarrow \mathbf{a} = \mathbf{b}$, и инъективность доказана. [1] $\square$

Теорема 3.3 (координатное описание ядра)**.** *Пусть* $\varphi \in L(V, \widetilde{V})$, e, f *— базисы в* V *и* $\widetilde{V}$. *Пусть* $\varphi \underset{\mathrm{e,f}}{\longrightarrow} A$. *Для вектора* $\mathbf{a} \in \widetilde{V}$, $\mathbf{a} = \mathrm{f}X$ *выполнено:* $\mathbf{a} \in \operatorname{Ker}\varphi \Leftrightarrow AX = O$.

[1] Можно заметить, что доказательство предложения 3.3 родственно доказательству пункта 2 теоремы 1.1.

$\triangleright$ Достаточно сопоставить определение ядра и формулу $Y = AX$ (см. теорему 2.1). $\square$

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

Уствноаим формулу, связывающую размерности ядра и образа.

Теорема 3.4. *Пусть* $\varphi \in L(V, \widetilde{V})$ *и* $\dim V = n < \infty$. *Тогда*

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

$\triangleright$ 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 3.2 имеем $\dim \operatorname{Im}\varphi = \operatorname{rg} A$, а по следствию из теоремы 3.3 имеем $\dim \operatorname{Ker}\varphi = n - \operatorname{rg} A$. Отсюда вытекет нужная формула размерностей. $\square$

Из теоремы 3.4 можно, в частности, по-другому доказать эквивалнтность условий 1), 2), 3) в теореме 1.3.

В следующем упражнении наметим также другой путь доказательства формулы из теоремы 3.4 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 об 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

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

2

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, $\dim U = n$.

[2] Утверждение в упражнении не следует путать со случаем выбора одного и того же базиса e = f для линейного преобразования — см. параграф **??** .

III.2. Пусть $V = \mathbf{F}(\mathbb{N}) = \{(f_1, f_2, \ldots) \mid f_i \in \mathbb{R}\}$ — пространство числовых последовательностей.
Фибоначчиевы последовательности могут быть заданы как ядро линейного оператора $\varphi^2 - \varphi - \varphi^0$, где $\varphi : V \to V$ — оператор сдвига.

§ 4. Сопряженное пространство

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 4.1. *Если* $\dim V = n < \infty$, *то* $\dim V^* = n$.

$\triangleright$ Это частный случай следствия из теоремы 2.2. $\square$

Вся теория о линейных отображениях переносится на частный случай пространства V^*. В частности, если зафиксировать базис $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ в пространстве V то каждая линейная функция φ получает в соответствие матрицу $1 \times n$, то есть строку $(\varphi(\mathbf{e}_1), \varphi(\mathbf{e}_2), \ldots, \varphi(\mathbf{e}_n))$. (Здесь мы считаем, что в пространстве $\widetilde{V} = \mathbb{R}$ (или $\mathbb{C}$)) зафиксирован базис — число 1.)

Если $\mathbf{a} \in V$ и $\varphi \in V^*$, то наряду с записью $\varphi(\mathbf{a})$ будем использовать запись $\langle \mathbf{a}, \varphi \rangle$. Эта запись удобна, так как имеется линейность по каждому аргументу, т.е. $\forall\ \mathbf{a}, \mathbf{a}_1, \mathbf{a}_2 \in V$; $\forall\ \varphi, \varphi_1, \varphi_2 \in V^*$; $\forall \lambda \in \mathbb{R}(\mathbb{C})$ выполнено:
$\langle \mathbf{a}_1 + \mathbf{a}_2, \varphi \rangle = \langle \mathbf{a}_1, \varphi \rangle + \langle \mathbf{a}_2, \varphi \rangle$, $\langle \lambda \mathbf{a}, \varphi \rangle = \lambda \langle \mathbf{a}, \varphi \rangle$;
$\langle \lambda \mathbf{a}, \varphi_1 + \varphi_2 \rangle = \langle \mathbf{a}, \varphi_1 \rangle + \langle \mathbf{a}, \varphi_2 \rangle$, $\langle \mathbf{a}, \lambda \varphi \rangle = \overline{\lambda} \langle \mathbf{a}, \varphi \rangle$.

Взаимный базис

Базис $\ell = (\ell_1, \ell_2, \ldots, \ell_n)$ пространства V^* называется *взаимным (или биортогональным, или двойственным)* для базиса $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ пространства V, если $\langle \mathbf{e}_i, \ell_j \rangle = \delta_{ij}$. [3]

Таким образом, если ℓ — взаимный базис для базиса e и $\mathbf{a} = \mathrm{e} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, то $\langle \mathbf{a}, \ell_i \rangle = x_i$.

Предложение 4.2. *Пусть* $\dim V = n$. *Тогда*
1. для любого базиса в V *существует единственный взаимный базис в* V^**;*
2. Если базисы e *и* e′ *в* V *связаны матрицей перехода* S *(то есть* $\mathrm{e}' = \mathrm{e}S$*), то взаимные базисы* ℓ *и* ℓ' *в* V^* *связаны матрицей перехода* $(\overline{S}^T)^{-1}$ *(то есть* $\ell' = \ell(\overline{S}^T)^{-1}$*).*

$\triangleright$ 1. Для каждого конкретного j вектор $\ell_j \in V^*$ определяется однозначно как вектор, 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 $\ell_1, \ell_2, \ldots, \ell_n$ образуют базис пространства V^*.
2. ВЫКЛАДКА... $\square$

Предложение 4.3. *Пусть* $\dim V = n$. *Пусть система векторов* $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ *пространства* V *и система векторов* $\ell = (\ell_1, \ell_2, \ldots, \ell_n)$ *пространства* V^* *(априори не известно, что системы линейно независимы) таковы, что* $\langle \mathbf{e}_i, \ell_j \rangle = \delta_{ij}$ *для всех* $i, j \in \{1, \ldots, n\}$. *Тогда* e *— базис в* V, *а* ℓ *— взаимный базис для* e.

$\triangleright$ Достаточно доказать линейную независимость системы векторов $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ (линейную независимость системы векторов $\ell_1, \ell_2, \ldots, \ell_n$ можно доказать аналогично). Пусть $\sum\limits_{i=1}^{n} \lambda_i \mathbf{e}_i = \mathbf{o}$. Применив к этому равенству $\ell_j \in V^*$, имеем (с учетом $\langle \mathbf{e}_i, \ell_j \rangle = \delta_{ij}$) $\lambda_j \langle \mathbf{e}_j, \ell_j \rangle = 0 \Rightarrow \lambda_j = 0$. $\square$

Примеры

III.1. 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 $V = \mathbf{P}_n$; $\alpha_0, \alpha_1, \ldots, \alpha_n$ — различные числа. Для системы линейных функций $\ell_i = \delta_{\alpha_i}$, $i = 0, 1, \ldots, n$, 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 4.3 следует, что $(\ell_0, \ell_1, \ldots, \ell_n)$ — взаимный базис для базиса $(p_0, p_1, \ldots, p_n)$. Разложение многочлена $p \in \mathbf{P}_n$ по p_i имеет вид

$$\sum_{i=0}^{n} p(\alpha_i) p_i.$$

Это так называемый *интерполяционный многчлен* Лагранжа.

§ 5. Структура линейного преобразования

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 преобразования $\varphi \in L(V,V)$ подпространство $U \leqslant V$, можем говорить о сужении $\varphi|_U \in L(U,U)$.

Предложение 5.1. *Пусть* V *— векторное пространство,* $U \leqslant V$, $U = \langle \mathbf{a}_1, \ldots, \mathbf{a}_k \rangle$. *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. $\square$

Предложение 5.2. *Пусть* V *— векторное пространство, и 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* $\bigcap\limits_{i=1}^{k} U_i$ *инвариантны относительно* φ.

$\triangleright$ TO BE PROVED $\square$

Замечание. *Предыдущее предложение верно и для произвольных подмножеств* $U_1, U_2, \dots, U_k \subset V$.

Предложение 5.3. *Пусть* V *— векторное пространство, и* $\varphi, \psi \in L(V, V)$ *таковы, что* $\varphi\psi = \psi\varphi$. *Тогда* $\operatorname{Ker}\psi$, $\operatorname{Im}\psi$ *инвариантны относительно* φ.

$\triangleright$ TO BE PROVED $\square$

Предложение 5.4. *Пусть* V *— векторное пространство над* $\mathbb{R}(\mathbb{C})$, $\varphi \in L(V, V)$ *и* $\lambda \in \mathbb{R}(\mathbb{C})$. *Пусть подпространство* $U \leqslant V$ *таково, что* $U \geqslant \operatorname{Im}(\varphi - \lambda)$. *Тогда* U *инвариантно относительно* φ.

$\triangleright$ TO BE PROVED $\square$

Предложение 5.5. *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}.

$\triangleright$ Сужение $\varphi|_U : U \to U$ является инъекцией, поэтому (см. теорему 1.3) 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 5.6. *Пусть* V *— векторное пространство,* $\dim V = n < \infty$, *и* $\varphi \in L(V, V)$ *имеет в базисе* $\mathrm{e} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ *матрицу* $A = (a_{ij})$. *Подпространство* $U_k = \langle \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k \rangle$ *инвариантно относительно* $\varphi \Leftrightarrow a_{ij} = 0$ *для всех* $i = k+1, \dots, n$, $j = 1, 2, \dots, k$ *(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* $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k$ *пространства* U_k.

$\triangleright$ По определению матрицы линейного отображения, $a_{ij} = 0$ для всех $i = k+1, \dots, n$ $\Leftrightarrow \varphi(\mathbf{e}_j) \in \langle \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k \rangle$. Остается воспользоваться предложением 5.1. $\square$

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

Аналогично последнему предложению, A имеет блочно-треугольный вид $\begin{pmatrix} B & O \\ C & D \end{pmatrix}$, где $O \in \mathbf{M}_{k\times(n-k)} \Leftrightarrow \langle \mathbf{e}_{k+1}, \mathbf{e}_{k+2}, \dots, \mathbf{e}_n \rangle$ инвариантно относительно φ.

Упражнение. Пусть $\varphi \in L(V, V)$, $\varphi \xrightarrow[\mathrm{e,f}]{} A$, где A — блочно-диагональныая (с квадратными блоками размеров $n_i \times n_i$ по диагонали). Какие инвариантные относительно φ подпространства можно указать?

Собственные значения, собственные векторы, собственные подпространства.

Следующие определения связаны с изучением одномерных инвариантных подпространств.

Пусть $\varphi \in L(V,V)$ — векторное пространство над $\mathbb{R}(\mathbb{C})$. Число $\lambda_0 \in \mathbb{R}(\mathbb{C})$ называется *собственным значением* преобразования φ, если $\exists\, \mathbf{a} \in V$ такой, что $\mathbf{a} \neq \mathbf{o}$ и $\varphi(\mathbf{a}) = \lambda_0\mathbf{a}$.

Если $\mathbf{a} \neq \mathbf{o}$ и $\varphi(\mathbf{a}) = \lambda_0\mathbf{a}$, то вектор $\mathbf{a}$ называется *собственным вектором* преобразования φ, отвечающим собственному значению λ_0.

Предложение 5.7. *Пусть $\varphi \in L(V,V)$, $\mathbf{a} \in V$, $\mathbf{a} \neq \mathbf{o}$. Вектор $\mathbf{a}$ является собственным вектором преобразования $\varphi \Leftrightarrow$ одномерное подпространство $\langle \mathbf{a} \rangle$ инвариантно относительно φ.*

$\triangleright$ TO BE PROVED $\square$

Предложение 5.8. *Пусть $\varphi \in L(V,V)$, $\mathbf{a} \in V$, $\mathbf{a} \neq \mathbf{o}$. Вектор $\mathbf{a}$ является собственным вектором преобразования φ, отвечающим собственному значению $\lambda_0 \Leftrightarrow \mathbf{a} \in \mathrm{Ker}(\varphi - \lambda_0)$.*

$\triangleright$ TO BE PROVED $\square$

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

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

Предложение 5.9. *Пусть $\lambda_1, \lambda_2, \dots, \lambda_k$ — различные собственные значения для $\varphi \in L(V,V)$. Тогда $\sum\limits_{i=1}^{k} V_{\lambda_i}$ — прямая сумма.*

$\triangleright$ Предположим, противное, и пусть, скажем, $\mathbf{a}_1 \in V_{\lambda_1} \cap \sum\limits_{i=2}^{k} V_{\lambda_i}$, $\mathbf{a}_1 \neq \mathbf{o}$ (см. теорему 3.1, глава **??**). 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 $(\varphi - \lambda_i)(\mathbf{a}_i) = \mathbf{o}$ и в произведении $\prod\limits_{i=2}^{k}(\varphi - \lambda_i)$ преобразования перестановочны, то $\psi(\mathbf{a}_i) = \mathbf{o}$ для $i = 2, \dots, k$. С другой стороны, $\psi(\mathbf{a}_1) = \prod\limits_{i=2}^{k}(\lambda_1 - \lambda_i)\mathbf{a}_1 \neq \mathbf{o}$. Противоречие. $\square$

Характеристический многочлен

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

Пусть преобразование $\varphi \in L(V,V)$ имеет в некотором базисе матрицу A. Функция $|A - \lambda E|$ переменной λ называется *характеристическим многочленом* преобразования φ.

Уравнение $|A - \lambda E| = 0$ называется *характеристическим уравнением*, а его (комплексные) корни — *характеристическими числами* преобразования φ.

Обозначение для характеристического многочлена: $\chi_\varphi(\lambda)$.

Раскрытие определителя $|A-\lambda E|$ показывает, что характеристический многочлен действительно является многочленом, и его степень равна n:

$$\chi_\varphi(\lambda) = (-1)^n\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_{n-1}\lambda + a_n. \tag{2.3}$$

При этом (из раскрытия определителя $|A - \lambda E|$) нетрудно видеть, что $a_1 = (-1)^{n-1}\sum\limits_{i=1}^{n} a_{ii} = (-1)^{n-1}\operatorname{tr} A$, $a_n = \chi_\varphi(0) = |A|$.

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

$$\chi_\varphi(\lambda) = (-1)^n(\lambda-\lambda_1)^{s_1}(\lambda-\lambda_2)^{s_2}\ldots(\lambda-\lambda_k)^{s_k}, \tag{2.4}$$

где $\lambda_1, \lambda_2, \ldots, \lambda_k$ — различные характеристические числа, а $s_1, s_2, \ldots, s_k$ — их кратности (так что $\sum\limits_{i=1}^{k} s_i = n$).

Предложение 5.10. *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 (2.3) и (2.4) (или же воспользоваться теоремой Виета). $\square$

Отметим, что если матрица A — верхнетреугольная, то характеристические числа совпадают с числами, расположенными на диагонали.

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

Предложение 5.11. *Характеристический многочлен преобразования* $\varphi \in L(V,V)$ *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 преобразования* $\varphi \in L(V,V)$ *не зависят от выбора базиса.*

Следующие теоремы проясняют связь между характеристическим многочленом и предыдущим разделом.

Теорема 5.1. *Пусть* V *— векторное пространство над* $\mathbb{R}(\mathbb{C})$, *и пусть* $\lambda_0 \in \mathbb{R}(\mathbb{C})$. *Тогда* λ_0 *— собственное значение для* $\varphi \in L(V,V) \Leftrightarrow \lambda_0$ *— характеристическое число.*

$\triangleright$ λ_0 — собственное значение $\Leftrightarrow$ $\operatorname{Ker}(\varphi - \lambda_0) \neq O$ $\Leftrightarrow$ (переходя к координатам) $\operatorname{Sol}((A-\lambda_0E)X = O) \neq O \Leftrightarrow$ квадратная матрица $A - \lambda_0E$ вырожденная $\Leftrightarrow |A-\lambda_0E| = 0$. $\square$

Упражнение. Докажите, что в нечетномерном вещественном пространстве у любого линейного преобразования существует собственный вектор.

Теорема 5.2. *Пусть* V *— векторное пространство над* $\mathbb{R}(\mathbb{C})$, $\lambda_0 \in \mathbb{R}(\mathbb{C})$. *Пусть* λ_0 *является корнем кратности* s_0 *характеристического многочлена* $\chi_\varphi(\lambda)$ *преобразования* $\varphi \in L(V,V)$. *Тогда* $1 \leqslant \dim V_{\lambda_0} \leqslant s_0$.

$\triangleright$ Пусть $\dim V_{\lambda_0} = s$. Выберем в $\dim V_{\lambda_0}$ 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 $\varphi \underset{\mathrm{e,e}}{\longrightarrow} A$, где A имеет блочный вид $\begin{pmatrix} \lambda_0E_s & C \\ O & D \end{pmatrix}$. Тогда $\chi_\varphi(\lambda) = |A - \lambda E| = |\lambda_0E_s - \lambda E_s|\cdot|D - \lambda E_{n-s}| = (\lambda_0 - \lambda)^s p(\lambda)$, где p — многочлен. $\square$

Диагонализируемость и треугольный вид

> Преобразование $\varphi \in L(V,V)$ называется *диагонализируемым*, если в V существует базис, в котором матрица φ имеет диагональный вид.

Теорема 5.3. *Пусть V — векторное пространство над $\mathbb{R}(\mathbb{C})$. Пусть $\varphi \in L(V,V)$, имеет различные характеристические числа $\lambda_1, \lambda_2, \ldots, \lambda_k$ кратностей $s_1, s_2, \ldots, s_k$. Тогда следующие условия эквивалентны:*

1) φ диагонализируемо;

2) В V существует базис из собственных векторов;

3) $\dim V_{\lambda_i} = s_i$ для $i = 1, 2, \ldots, k$ (в частности, в случае V над $\mathbb{R}$ необходимо все $\lambda_1, \lambda_2, \ldots, \lambda_k$ вещественные).

4) $V = \bigoplus\limits_{i=1}^{k} V_{\lambda_i}$.

$\triangleright$ 1) $\Leftrightarrow$ 2) Очевидно следует из определения матрицы линейного преобразования.
2) $\Rightarrow$ 3) TO BE PROVED
3) $\Rightarrow$ 4) TO BE PROVED
4) $\Rightarrow$ 2) TO BE PROVED $\square$

Следствие. *Пусть V — векторное пространство над $\mathbb{R}$, $\varphi \in L(V,V)$. Если $\chi_\varphi(\lambda)$ имеет n различных вещественных корней, то φ диагонализируемо.*

Исследование на диагонализируемость часто удобно проводить, используя условие 3) из теоремы 5.3.

Пусть $\varphi \in L(V,V)$ имеет матрицу A в некотором базисе. Тогда для диагонализируемости φ необходимо и достаточно существование такой невырожденной матрицы S, что $S^{-1}AS$ является диагональной матрицей. Поэтому иногда говорят о диагонализируемости квадратной матрицы A.

Теорема 5.4. *Пусть $\varphi \in L(V,V)$ — произвольное линейное преобразование векторного пространства над $\mathbb{C}$, либо $\varphi \in L(V,V)$ — линейное преобразование векторного пространства над $\mathbb{R}$, имеющее лишь вещественные характеристические числа. Тогда в V существует базис, в котором матрица φ верхнетреугольная.*

$\triangleright$ TO BE PROVED $\square$

Пусть $J_t(\lambda_0)$ — матрица $t \times t$, у которой по главной диагонали — одинаковые числа, на следующей диагонали над главной — единицы, а остальные элементы — нули:

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

$J_t(\lambda_0)$ называется *жордановой клеткой*. Блочно-диагональная матрица, у которой каждый диагональный блок является жордановой клеткой, называется *жордановой матрицей*, или матрицей, имеющей *жорданову нормальную форму* (жнф).

Теорему 5.4 можно усилить: заменив «верхнетреугольная» на «жорданова», получим формулировку теоремы о существовании жнф.

Теорема 5.5 (Теорема Гамильтона-Кэли). *Для всякого $\varphi \in L(V,V)$ выполнено*

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

Упражнение. Выведите теорему Гамильтона-Кэли из теоремы Жордана.

Примеры

Не все преобразования диагонализируемы.

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).

II. Матрица $\begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix}$ не диагонализируема ни над $\mathbb{R}$, ни над $\mathbb{C}$, так как λ_0 — корень кратности 2, но $\dim V_{\lambda_0} = 1$.
(Объяснение без использования теоремы 5.3 такое: если бы это преобразование было диагонализируемым, то диагональный вид был бы нулевой матрицей, и значит, преобразование было бы нулевым, что неверно.)

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 из **??**). При $n > 1$ преобразование d не является диагонализируемым.

Глава 3

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

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 (полуторалинейных) форм на пространстве V обозначаем $\mathcal{B}(V)$.

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.

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 $\beta \underset{\mathrm{e}}{\rightarrow} B$. Посмотрев на определение, матрицу B можно неформально назвать таблицей билинейного умножения.

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

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

$\triangleright$ 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). $\square$

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 $\beta \xrightarrow[\mathrm{e}]{} B$.*

$\triangleright$ 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 $\beta(\mathbf{e}_i,\mathbf{e}_j) = E_{\bullet i}^T B\overline{E_{\bullet j}} = b_{ij}$. Поэтому в самом деле $\beta \xrightarrow[\mathrm{e}]{} B$. $\square$

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 $\mathbf{M}_{n\times n}$. [1]

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*

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

$\triangleright$ 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 **??**), 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: $B' = S^T B\overline{S}$. $\square$

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

$\triangleright$ Достаточно заметить, что матрица перехода невырожденная, а умножение на невырожденную матрицу не меняет ранг. $\square$

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

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

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

[1] Сверх того, 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.

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* $\beta\mid_{U\times U}\colon U \times U \to \mathbb{R}$. 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 $i = 1, \dots, k$, $j = 1, \dots, k$. В частности, $B_1 = (b_{11})$, $B_n = B$.

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

Предложение 1.2. *Пусть $\beta \in \mathcal{B}(V)$ и $\beta \underset{\mathrm{e}}{\longrightarrow} B$. Тогда $\beta\mid_{U\times U} \underset{\mathrm{e}^{(k)}}{\longrightarrow} B_k$.*

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

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

§ 2. Симметричные билинейные формы. Квадратичные формы

Симметричные билинейные формы

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.[2]

Теорема 2.1. *Пусть* $\dim V < \infty$, e *— базис в V. Пусть $\beta \in \mathcal{B}(V)$, $\beta \underset{\mathrm{e}}{\longrightarrow} B$. 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. $\square$

Упражнение. Определите *кососимметричные билинейные формы*, найдите условия на матрицу, эквивалентные кососимметричности формы.

[2]Это утверждение содержательно для комплексного пространства.

Квадратичные формы

Пусть $\beta \in \mathcal{B}_{sym}(V)$. Отображение $k : V \to V$, 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 взаимнооднозначное соответствие между множествами $\mathcal{B}_{sym}(V)$ и $\mathcal{K}(V)$. Ниже отождествляем эти множества.

§ 3. Диагональный и канонический вид квадратичной формы

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

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 e 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

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

(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.[3]

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 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 $i = 1, \ldots, n$.

$\triangleright$ 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 $\mathbf{a} \neq \mathbf{o}$ — 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.*

$\triangleright$ 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. $\square$

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 и* $k \xrightarrow[\mathrm{e}]{} \operatorname{diag}(d_1, d_2, \ldots, d_n)$. *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 $p \geqslant p'$.

[3]Конечно, существуют квадратичные формы, не принадлежащие ни к одному из определенных типов.

Предположим, что $p > p'$, тогда рассмотрим подпространство $U \leqslant V$ такое, что $\dim U = p$ и $k\mid_U$ 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 **??** 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 $\operatorname{rg} k$.*

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 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 $k\mid_U$ 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 $\tilde{k} = k\mid_{U_m}$ не является положительно определенной (где $U_m = \langle \mathbf{e}_1, \ldots, \mathbf{e}_m \rangle$). По выбору m получаем, что форма $k\mid_{U_{m-1}} = \tilde{k}\mid_{U_{m-1}}$ 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 $D = \operatorname{diag}(d_1, \ldots, d_m)$ — 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 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.

Глава 4

Евклидовы и унитарные пространства

§ 1. Скалярное произведение. Матрица Грама

Определения

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

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

> *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 $\beta(\mathbf{a}, \mathbf{b})$).

> *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 $|\mathbf{a}| = 0 \Leftrightarrow \mathbf{a} = \mathbf{0}$. НОРМИРОВАТЬ!

> 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 $\mathbf{o}$.

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

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 своеобразую «таблицу умножения» для скалярного произведения. 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 3). 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 3. $\square$

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* $|\Gamma(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)| = 0$.[1]

$\triangleright$ 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 3.

2) Пусть система векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ линейно зависима, т.е. $\sum\limits_{i=1}^{n} \lambda_i \mathbf{a}_i = \mathbf{o}$ 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 $|\Gamma(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)| = 0$. $\square$

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.

[1] На самом деле геометрический смысл определителя $|\Gamma(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)|$ — это ориентированный объем k-мерного параллелепипеда, построенного на векторах $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$).

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

Упражнение. ОБОБЩ для Билин функций

§ 2. Ортогональные системы векторов. Ортогональное дополнение. Ортогонализация

Ортогональные системы векторов. ОНБ

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 $1 \leqslant i < j \leqslant k$.

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* $|\mathbf{a}_1 + \ldots + \mathbf{a}_k|^2 = |\mathbf{a}_1|^2 + \ldots + |\mathbf{a}_k|^2$.

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

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.*

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

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

$\triangleright$ Следует из предыдущего предложения с учетом предложения 1.1. $\square$

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.*

$\triangleright$ Ввиду следствия 2, это частный случай следствия из теоремы 3.1 главы 3. $\square$

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*

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

$\triangleright$ Это утверждение — частный случай теоремы 4.1. $\square$

Переход от ОНБ к ОНБ. Ортогональные и унитарные матрицы

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

|| 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* $(X, Y) = X^T\overline{Y}$.

$\triangleright$ Очевидно, 1) $\Leftrightarrow$ 2).
Пусть $Q_1, \ldots, Q_n$ — 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) $\Leftrightarrow$ 3). $\square$

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* $Q^T \in U_n$, $\overline{Q} \in U_n$, $Q^* \in U_n$.

$\triangleright$ $\square$

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* $QR \in U_n$, $Q^{-1} \in U_n$.

$\triangleright$ $\square$

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

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

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

§ 2.ОРТОГОНАЛЬНЫЕ СИСТЕМЫ ВЕКТОРОВ. ОРТОГОНАЛЬНОЕ ДОПОЛНЕНИЕ.

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

$\triangleright$ Это сразу следует из определения матрицы перехода и условия 3) в предложении 2.3. $\square$

Таким образом, ортогональные (унитарные) матрицы — в точности матриы 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

> 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 $\mathbf{o}$. Это противоречит следствию 1 из предложения **??**. $\square$

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. $\mathbf{a} \perp \mathbf{b}$. $\square$

Ортогональное дополнение. Ортогональная проекция

> *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 4.1) 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 $(\mathbf{c}, \mathbf{b}_i) = 0$ $\Leftrightarrow (\mathbf{a}, \mathbf{b}_i) - \alpha_i (\mathbf{b}_i, \mathbf{b}_i) = 0 \Leftrightarrow \alpha_i = \dfrac{(\mathbf{a}, \mathbf{b}_i)}{(\mathbf{b}_i, \mathbf{b}_i)}$. $\square$

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....), $(U^{\perp})^{\perp} = U$. $\square$

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

$\triangleright$ 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.

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

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

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*

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

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

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

Проекции могут естественно возникать в задачах на экстремум:

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

$\triangleright$ Представим $\mathbf{a}$ как $\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_2$, где $\mathbf{a}_1 = \mathrm{pr}_U \mathbf{a}$, $\mathbf{a}_2 = \mathrm{pr}_{U^{\perp}} \mathbf{a}$. 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* $\mathcal{E}$, $U = \langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle$*, где* $\mathbf{a}_1, \dots, \mathbf{a}_k$ *— векторы с с координатными столбцами* $X_1, \dots, X_k$ *в ОНБ* e*,* Φ *— матрица со столбцами* X_i*:* $\Phi = (X_1 \ \dots \ X_k)$*. Тогда* $U^\perp$ *задается (в коодинатах в ОНБ* e*) как* $\mathrm{Sol}(\Phi^* X = O)$.

$\triangleright$ 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 $\mathbf{a}_i$, $i = 1, \dots, k$. В силу предложения 4.1, 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* $\Leftrightarrow \forall\, Y_0 \in \mathrm{Sol}(A^* Y = O)$ *выполнено «условие ортогональности»:* $b^* Y = 0$.

$\triangleright$ Рассмотрим пространство $\mathcal{E} = \mathbb{M}_{m \times 1}$ столбцов высоты m со стандартным скалярным произведением. Столбцы $a_{\bullet 1}, \dots, a_{\bullet n}$ матрицы A и столбец b лежат в $\mathcal{E}$. Условие совместности системы $AX = b$ эквивалентно тому, что $b \in U$, где $U = \langle a_{\bullet 1}, \dots, a_{\bullet n} \rangle$. Далее, $U^\perp = \mathrm{Sol}(A^* Y = O)$, поэтому «условие ортогональности» $\forall\, Y_0 \in \mathrm{Sol}(A^* Y = O)$ 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: $U = \langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle$, а требуется найти ортогональный (или ОНБ) бизис в U (например, чтобы после этого была возможность использовать формулу для $\mathrm{pr}_U \mathbf{a}$). 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 $U_i = \langle \mathbf{a}_1, \dots, \mathbf{a}_i \rangle$, так что $U_{i+1} = U_i \cup \{\mathbf{a}_{i+1}\}$. Идея следующая: будет по очереди «поправлять» векторы $\mathbf{a}_i$, заменяя на $\mathbf{b}_i = \mathrm{pr}_{U_{i-1}^\perp} \mathbf{a}_i$ так, что $\mathbf{b}_i \perp U_{i-1}$ и $U_i = \langle \mathbf{b}_1, \dots, \mathbf{b}_i \rangle$.

Получим явные формулы. Первый ненулевой вектор $\mathbf{a}_t$ из списка $\mathbf{a}_1, \dots, \mathbf{a}_k$ является ортогональным базисом в U_t. Пусть на некотором шаге мы уже имеем ортогональный базис $\mathbf{b}_1, \dots, \mathbf{b}_\ell$ в U_m. Заменяем $\mathbf{a}_{m+1}$ на вектор $\mathbf{a}_{m+1} - \mathrm{pr}_{U_m} \mathbf{a}_{m+1}$, 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 $\varphi : \mathcal{E} \to \mathcal{E}$.

> 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* $\varphi, \psi \in L(\mathcal{E}, \mathcal{E})$, $\varphi \underset{\mathrm{e,e}}{\longrightarrow} A$, $\psi \underset{\mathrm{e,e}}{\longrightarrow} B$. *Тогда*

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

$\rhd$ Пусть $\mathbf{a} = \mathrm{e}X$, $\mathbf{b} = \mathrm{e}Y$. Имеем: $(\varphi(\mathbf{a}), \mathbf{b}) = (AX)^T\overline{Y} = X^TA^T\overline{Y} = \beta_1(\mathbf{a}, \mathbf{b})$, 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)* $\mathrm{rg}(\varphi) = \mathrm{rg}(\varphi^*)$; *4)* $\overline{\chi_\varphi(\lambda)} = \chi_{\varphi^*}(\overline{\lambda})$.

$\rhd$ 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^*.

$\rhd$ $\boxed{\Rightarrow}$ 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*

$$\boxed{\operatorname{Ker}\varphi^* = (\operatorname{Im}\varphi)^\perp}.$$

$\rhd$ Во-первых можно заметить, что подпространства $\operatorname{Ker}\varphi^*$ и $(\operatorname{Im}\varphi)^\perp$ имеют равные размерности. Действительно, по теор..... $\dim\operatorname{Ker}\varphi^* = n - \mathrm{rg}\,\varphi^* = n - \mathrm{rg}\,\varphi$, а по .... $\dim(\operatorname{Im}\varphi)^\perp = n - \dim(\operatorname{Im}\varphi) = n - \mathrm{rg}\,\varphi$.

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 $(\mathbf{c}, \mathbf{b}) = (\varphi(\mathbf{a}), \mathbf{b}) = (\mathbf{a}, \varphi^*(\mathbf{b})) = (\mathbf{a}, \mathbf{o}) = 0$, откуда $\mathbf{c} \perp \mathbf{b}$, что и требовалось. $\square$

Самосопряженные преобразования

|| 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, $\varphi \xrightarrow[\mathrm{e,e}]{} A$. 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 самоспряженного 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. Изоморфизм. Ортогональные и унитарные преобразования

Изоморфизм евклидовых (унитарных) пространств

Пусть даны два евклидовых (унитарных) пространства $\mathcal{E}_1$ и $\mathcal{E}_2$.

> Отображение $\varphi : \mathcal{E}_1 \to \mathcal{E}_2$ называется *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.

> 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.... (см. ....).

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 (такое φ существует по ....). Согласно предложению **??**, φ будет изоморфизмом. $\square$

§ 5. Ортогональные и унитарные преобразования

> 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 5.1. *Пусть* $\dim \mathcal{E} = n < \infty$, $\varphi \in L(\mathcal{E} \to \mathcal{E})$. *Тогда* φ *— ортогональное (унитарное)* $\Leftrightarrow \varphi$ *— 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* $\Leftrightarrow \varphi$ *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 5.2 (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 5.3. *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 5.1 (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 5.2. *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.

§ 6. Квадратичные формы

Теорема 6.1. *Пусть* k *— квадратичная форма в евклидовом (унитарном) пространстве* $\mathcal{E}$. *Тогда существует ОНБ, в котором* k *имеет диагональный вид.*

$\triangleright$ $\square$

ЧЕРЕЗ присоедин. преобразование? (пример подъема индекса?) или через матрицы?

Следствие (Теорема о паре форм)**.** *Пусть в векторном пространстве* V *(без евклидовой или унитарной структуры) заданы квадратичные формы* f *и* g*, причем* g *положительно определена. Тогда существует базис, в котором* g *имеет канонический вид, а* f *имеет 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 6.1, поскольку ОНБ — это базис, в котором g имеет канонический вид (см. следствие 2 предложения 2.2). $\square$

Упражнение. 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 хотя бы один не вещественный корень.

Список используемых обозначений

$\Rightarrow$	знак следствия
$\Leftrightarrow$	знак эквивалентности (равносильности)
$\forall$	квантор всеобщности
$\exists$	квантор существования
$a \in A$	a принадлежит множеству A
$a \notin A$	a не принадлежит множеству A
$\varnothing$	пустое множество
$A = \{a_1, a_2, \ldots, a_n\}$	множество с элементами $a_1, a_2, \ldots, a_n$
$\mathbb{N}$	множество натуральных чисел
$\mathbb{Z}$	множество целых чисел
$\mathbb{R}$	множество действительных чисел
$\mathbb{C}$ — множество комплексных чисел	
$B \subset A$	B является подмножеством множества A
$B \subseteq A$	B является подмножеством множества A
$\{a \in A \mid \mathcal{X}\}$	подмножество A, заданное условием $\mathcal{X}$
$\bigcup$	объединение (множеств)
$\bigcap$	пересечение (множеств)
$A \setminus B$	разность множеств A и B
$\times$	декартово произведение (множеств)
$\sum$	знак суммирования
$f : X \to Y$	отображение из X в Y
$f(X')$	образ подмножества X' при оботображении f
I	тождественное преобразование
I_X	тождественное преобразование $X \to X$
δ_{ij}	символ Кронекера, $\delta_{ij} = 1$ при $i = j$ и $\delta_{ij} = 0$ при $i \neq j$
$\mathbf{M}_{m\times n}$	множество матриц размера $m \times n$
$\mathbf{M}_{m\times n}(K)$	множество матриц размера $m \times n$ с элементами из множества K
$\mathbf{M}^{+}_{n\times n}$	симметрические матрицы $n \times n$
$\mathbf{M}^{-}_{n\times n}$	кососимметрические матрицы $n \times n$
(a_{ij})	матрица с элементами a_{ij}
$a_{i\bullet}$	стока с номером i матрицы (a_{ij})
$a_{\bullet j}$	столбец с номером j матрицы (a_{ij})
$\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$	диагональная матрица
E	единичная матрица
E_n	единичная матрица порядка n
$A + B$	сумма матриц A и B
λA	произведение матрицы A на число λ

O	нулевая матрица (некоторого размера)
$-A$	матрица, противоположная матрице A
A^T	матрица, транспонированная к матрице A
AB	произведение матриц A и B
A^{-1}	матрица, обратная к матрице A
$\overline{A}$	матрица, комплексно сопряженная для матрицы A, т.е если $A =$
A^*	сопряженная матрица, т.е $A^* = \overline{A^T}$
$\operatorname{rg} A$	ранг матрицы A
$\operatorname{tr} A$	след матрицы A
$\lvert A\rvert$ или $\det A$	определитель (детерминант) матрицы A
$AX = b$	матричная запись системы линейных уравнений
$\operatorname{Sol}(AX = b)$	общее решение (множество всех решений) системы линейных у
$\mathbf{a}$	вектор $\mathbf{a}$ (элемент векторного пространства)
$\mathbf{o}$	нулевой вектор
$-\mathbf{a}$	вектор, противоположный вектору $\mathbf{a}$
$U \leqslant V$	U является подпространством в векторном пространстве V
$\langle \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k \rangle$	линейная оболочка системы векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$
$\langle \mathcal{A} \rangle$	линейная оболочка системы векторов $\mathcal{A}$
$\operatorname{rg} \mathcal{A}$	ранг системы векторов $\mathcal{A}$
$\operatorname{rg}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)$	ранг конечной системы векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$
$\dim V$	размерность векторного пространства V
$\mathbf{a} = \mathrm{e}X$	вектор $\mathbf{a}$ имеет координатный столбец X в базисе e (или $\mathbf{a}$ раск
$\mathrm{e}' = \mathrm{e}S$	S — матрица перехода от базиса e к базису e′
$\mathcal{A}_1 + \mathcal{A}_2 + \ldots + \mathcal{A}_k$ или $\sum\limits_{i=1}^{k} \mathcal{A}_i$	сумма (по Минковскому)
$U_1 \bigoplus U_2 \bigoplus \ldots \bigoplus U_k$ или $\bigoplus\limits_{i=1}^{k} U_i$	прямая сумма подпространств
$L(V, \widetilde{V})$	множество всех линейных отображений $V \to \widetilde{V}$
$V \cong \widetilde{V}$	V изоморфно $\widetilde{V}$
$\varphi \underset{\mathrm{e,f}}{\longrightarrow} A$	A — матрица линейного отображения φ в паре базисов e и f
V^*	сопряженное (двойственное) пространство
$\langle \mathbf{a}, \varphi \rangle$	значение функционала $\varphi \in V^*$ на векторе $\mathbf{a} \in V$
$\operatorname{Im} \varphi$	образ отображения φ
$\operatorname{Ker} \varphi$	ядро отображения φ