

ЭКЗАМЕН ПО ЛИНЕЙНОЙ АЛГЕБРЕ

Лектор: Чубаров И. А. • Автор: Пшеничный Никита,* группа 109

1 курс • Весенний семестр 2024 г.

Аннотация

При подготовке данного файла я использовал курсы лекций И. А. Чубарова, Т. Е. ПАНОВА и О. В. КУЛИКОВОЙ, курс семинаров А. А. Клячко и книги «Курс алгебры» Э. Б. Винберга, «Введение в алгебру. Часть II: Линейная алгебра» А. И. Кострикина, «Сборник задач по аналитической геометрии и линейной алгебре» под редакцией Ю. М. Смирнова и «Задачи по линейной алгебре» А. А. Гайфуллина, А. В. Пенского и С. В. Смирнова.

В конце есть раздел с условиями (всех) и решениями (некоторых) теоретических задач Т. Е. Панова, а по тексту и в приложениях разбросаны задачи из Винберга и семинаров А. А. Клячко. Так же разбросаны по тексту ссылки на файлы Антона Александровича с крутыми методами решения почти стандартных задач по линейной алгебре. Правда, там он пишет всё очень непонятно, думаю, здесь когда-нибудь появятся комментарии к этим файлам.

Всё, что выделено ВОТ ТАКИМ ШРИФТОМ (кроме этой записи) является ссылкой или заголовком (причём, это объединение дизъюнктивно). Нумерация задач сквозная по всем вопросам и отдельная для приложений и теоретических задач.

Обо всех ошибках и опечатках пишите мне, исправлю.

Программа экзамена

1	ВЕКТОРНЫЕ ПРОСТРАНСТВА. ЛИНЕЙНАЯ ЗАВИСИМОСТЬ И НЕЗАВИСИМОСТЬ ВЕКТОРОВ. БАЗИС, РАЗМЕРНОСТЬ. ПРИМЕРЫ	6
2	МАТРИЦА ПЕРЕХОДА ОТ ОДНОГО БАЗИСА К ДРУГОМУ. ИЗМЕНЕНИЕ КООРДИНАТ ВЕКТОРА ПРИ ЗАМЕНЕ БАЗИСА	10
3	ИЗОМОРФИЗМ ВЕКТОРНЫХ ПРОСТРАНСТВ ОДИНАКОВОЙ РАЗМЕРНОСТИ	12
4	ВЕКТОРНЫЕ ПОДПРОСТРАНСТВА, РАВНОСИЛЬНОСТЬ ДВУХ СПОСОБОВ ИХ ЗАДАНИЯ. СУММА И ПЕРЕСЕЧЕНИЕ ПОДПРОСТРАНСТВ. ФОРМУЛА ГРАССМАНА	12
5	ПРЯМАЯ СУММА ПОДПРОСТРАНСТВ. ВНЕШНЯЯ ПРЯМАЯ СУММА ВЕКТОРНЫХ ПРОСТРАНСТВ	15
6	ФАКТОРПРОСТРАНСТВО, ЕГО РАЗМЕРНОСТЬ. КОРАЗМЕРНОСТЬ. СВЯЗЬ С РЕШЕНИЯМИ НЕОДНОРОДНОЙ СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ	17
7	ЛИНЕЙНЫЕ ФУНКЦИИ НА ВЕКТОРНОМ ПРОСТРАНСТВЕ, ИХ ЯДРА. ИЗМЕНЕНИЕ КОЭФФИЦИЕНТОВ ЛИНЕЙНОЙ ФОРМЫ ПРИ ЗАМЕНЕ БАЗИСА. СОПРЯЖЁННОЕ ПРОСТРАНСТВО V^* , ДУАЛЬНЫЙ БАЗИС. КАНОНИЧЕСКИЙ ИЗОМОРФИЗМ $V \simeq V^{**}$	18
8	ЛИНЕЙНЫЕ ОТОБРАЖЕНИЯ И ОПЕРАТОРЫ. ЯДРО И ОБРАЗ, СВЯЗЬ ИХ РАЗМЕРНОСТЕЙ. КРИТЕРИЙ ИНЪЕКТИВНОСТИ	21

*Telegram: @pshenikita. Последняя компиляция: 13 июня 2024 г.

9	Задание линейных отображений (операторов) матрицами. Изменение матрицы линейного отображения при переходе к другим базисам. Нахождение ядра и образа при помощи матрицы	22
10	Линейные операторы. Изменение матрицы линейного оператора при переходе к новому базису. Подобные матрицы	23
11	Векторное пространство линейных отображений. Алгебра линейных операторов. Изоморфизм алгебры матриц и алгебры линейных операторов	24
12	Инвариантные подпространства линейного оператора. Ограничение линейного оператора на инвариантное подпространство. Вид матрицы линейного оператора при наличии инвариантных подпространств	25
13	Собственные векторы и собственные значения линейного оператора. Линейная независимость собственных векторов линейного оператора, отвечающих попарно различным собственным значениям	26
14	Вычисление собственных значений и собственных векторов с помощью матрицы. Характеристический многочлен	26
15	Собственные подпространства. Неравенство между размерностью собственного подпространства и кратностью корня характеристического многочлена	27
16	Диагонализуемость (матрицы) линейного оператора. Критерии диагонализуемости и достаточное условие	27
17	Аннулирующие многочлены линейного оператора (матрицы). Минимальный многочлен. Критерий диагонализуемости в терминах минимального многочлена	29
18	Теорема Гамильтона — Кэли	31
19	Существование одномерного или двумерного инвариантного подпространства для любого линейного оператора над полем действительных чисел	32
20	Корневые подпространства. Разложение пространства в прямую сумму корневых подпространств	34
21	Жордановы клетки и матрицы, их характеристические и минимальные многочлены. Жорданов базис	36
22	Существование жорданова базиса для нильпотентного оператора (для матрицы с единственным собственным значением — характеристическим корнем)	41
23	Существование жордановой нормальной формы матрицы над алгебраически замкнутым полем	43
24	Единственность жордановой нормальной формы	44
25	Билинейные функции и их матрицы. Изменение матрицы билинейной формы при замене базиса. Ранг билинейной функции. Симметрические билинейные функции	45
26	Квадратичные формы и их матрицы. Восстановление симметрической билинейной функции по данной квадратичной функции. Диагональный вид квадратичной формы. Алгоритм Лагранжа	46

27	НОРМАЛЬНЫЙ (КАНОНИЧЕСКИЙ) ВИД КВАДРАТИЧНОЙ ФОРМЫ НАД ПОЛЯМИ ДЕЙСТВИТЕЛЬНЫХ И КОМПЛЕКСНЫХ ЧИСЕЛ. ЗАКОН ИНЕРЦИИ	49
28	ТЕОРЕМА ЯКОБИ	51
29	ПОЛОЖИТЕЛЬНО И ОТРИЦАТЕЛЬНО ОПРЕДЕЛЁННЫЕ ВЕЩЕСТВЕННЫЕ КВАДРАТИЧНЫЕ ФОРМЫ. КРИТЕРИЙ СИЛЬВЕСТРА	52
30	ЕВКЛИДОВО ПРОСТРАНСТВО. НЕРАВЕНСТВО КОШИ — БУНЯКОВСКОГО. УГОЛ МЕЖДУ ВЕКТОРАМИ И ДЛИНА ВЕКТОРА В ЕВКЛИДОВОМ ПРОСТРАНСТВЕ. НЕРАВЕНСТВО ТРЕУГОЛЬНИКА	53
31	ВЫЧИСЛЕНИЕ СКАЛЯРНОГО ПРОИЗВЕДЕНИЯ В КООРДИНАТАХ. МАТРИЦА ГРАМА И ЕЁ СВОЙСТВА	55
32	ОРТОГОНАЛЬНОСТЬ ВЕКТОРОВ. ЛИНЕЙНАЯ НЕЗАВИСИМОСТЬ СИСТЕМЫ НЕНУЛЕВЫХ ОРТОГОНАЛЬНЫХ ВЕКТОРОВ. ОРТОГОНАЛЬНЫЙ И ОРТОНОРМИРОВАННЫЙ БАЗИС. АЛГОРИТМ ОРТОГОНАЛИЗАЦИИ ГРАМА — ШМИДТА	56
33	ОРТОГОНАЛЬНЫЕ МАТРИЦЫ КАК МАТРИЦЫ ПЕРЕХОДА ОТ ОДНОГО ОРТОНОРМИРОВАННОГО БАЗИСА К ДРУГОМУ. ГРУППА $O(n)$ ОРТОГОНАЛЬНЫХ МАТРИЦ ПОРЯДКА n .	57
34	ОРТОГОНАЛЬНОЕ ДОПОЛНЕНИЕ. РАЗЛОЖЕНИЕ ЕВКЛИДОВА ПРОСТРАНСТВА В ПРЯМУЮ СУММУ ПОДПРОСТРАНСТВА И ОРТОГОНАЛЬНОГО ДОПОЛНЕНИЯ К НЕМУ. ОРТОГОНАЛЬНАЯ ПРОЕКЦИЯ	58
35	ТЕОРЕМА ПИФАГОРА. УГОЛ И РАССТОЯНИЕ МЕЖДУ ВЕКТОРОМ И ПОДПРОСТРАНСТВОМ. ОБЪЁМ n -МЕРНОГО ПАРАЛЛЕЛЕПИПЕДА	59
36	ИЗОМОРФИЗМ ЕВКЛИДОВЫХ ПРОСТРАНСТВ ОДИНАКОВОЙ РАЗМЕРНОСТИ. ИЗОМОРФИЗМ ЕВКЛИДОВА ПРОСТРАНСТВА И ЕГО СОПРЯЖЁННОГО	61
37	ЛИНЕЙНЫЕ ОПЕРАТОРЫ В ЕВКЛИДОВОМ ПРОСТРАНСТВЕ. ОПЕРАТОР, СОПРЯЖЁННЫЙ ЛИНЕЙНОМУ ОПЕРАТОРУ, ЕГО МАТРИЦА И СВОЙСТВА	62
38	САМОСОПРЯЖЁННЫЙ ЛИНЕЙНЫЙ ОПЕРАТОР, ЕГО СВОЙСТВА И МАТРИЦА. СУЩЕСТВОВАНИЕ ОРТОНОРМИРОВАННОГО БАЗИСА ИЗ СОБСТВЕННЫХ ВЕКТОРОВ	63
39	ОРТОГОНАЛЬНЫЙ ОПЕРАТОР, ЕГО СВОЙСТВА И МАТРИЦА. КАНОНИЧЕСКИЙ ВИД МАТРИЦЫ ОРТОГОНАЛЬНОГО ОПЕРАТОРА	64
40	ПОЛЯРНОЕ РАЗЛОЖЕНИЕ НЕВЫРОЖДЕННОГО ЛИНЕЙНОГО ОПЕРАТОРА В ЕВКЛИДОВОМ ПРОСТРАНСТВЕ. СИНГУЛЯРНОЕ РАЗЛОЖЕНИЕ	67
41	БИЛИНЕЙНЫЕ И КВАДРАТИЧНЫЕ ФУНКЦИИ НА ЕВКЛИДОВОМ ПРОСТРАНСТВЕ. ВЗАИМНО ОДНОЗНАЧНОЕ СООТВЕТСТВИЕ МЕЖДУ СИММЕТРИЧЕСКИМИ БИЛИНЕЙНЫМИ ФОРМАМИ И САМОСОПРЯЖЁННЫМИ ЛИНЕЙНЫМИ ОПЕРАТОРАМИ. ПРИВЕДЕНИЕ КВАДРАТИЧНОЙ ФОРМЫ К ГЛАВНЫМ ОСЯМ. ПРИВЕДЕНИЕ К ДИАГОНАЛЬНОМУ ВИДУ ПАРЫ ФОРМ, ОДНА ИЗ КОТОРЫХ ПОЛОЖИТЕЛЬНО ОПРЕДЕЛЕНА	69
42	ПОЛУТОРАЛИНЕЙНЫЕ ФУНКЦИИ И ФОРМЫ. ЭРМИТОВЫ ФОРМЫ И ЭРМИТОВЫ МАТРИЦЫ. КАНОНИЧЕСКИЙ ВИД ЭРМИТОВОЙ КВАДРАТИЧНОЙ ФОРМЫ. КРИТЕРИЙ СИЛЬВЕСТРА	71
43	УНИТАРНОЕ (ЭРМИТОВО) ПРОСТРАНСТВО. ОРТОГОНАЛИЗАЦИЯ. ОРТОГОНАЛЬНЫЕ И ОРТОНОРМИРОВАННЫЕ БАЗИСЫ. МАТРИЦА ПЕРЕХОДА МЕЖДУ ОРТОНОРМИРОВАННЫМИ БАЗИСАМИ. ГРУППА $U(n)$ УНИТАРНЫХ МАТРИЦ ПОРЯДКА n	71

44	САМОСОПРЯЖЕННЫЕ (ЭРМИТОВЫ) ОПЕРАТОРЫ В УНИТАРНОМ ПРОСТРАНСТВЕ, ИХ СВОЙСТВА. ТЕОРЕМА О СУЩЕСТВОВАНИИ ОРТОНОРМИРОВАННОГО БАЗИСА ИЗ СОБСТВЕННЫХ ВЕКТОРОВ ЭРМИТОВА ОПЕРАТОРА. ПРИВЕДЕНИЕ ЭРМИТОВОЙ КВАДРАТИЧНОЙ ФОРМЫ К ГЛАВНЫМ ОСЯМ	71
45	УНИТАРНЫЕ ОПЕРАТОРЫ В УНИТАРНОМ ПРОСТРАНСТВЕ, ИХ СВОЙСТВА. ТЕОРЕМА О СУЩЕСТВОВАНИИ ОРТОНОРМИРОВАННОГО БАЗИСА ИЗ СОБСТВЕННЫХ ВЕКТОРОВ УНИТАРНОГО ОПЕРАТОРА	71
46	АФФИННОЕ ПРОСТРАНСТВО, ОСНОВНЫЕ ПОНЯТИЯ. АФФИНИЗАЦИЯ ВЕКТОРНОГО ПРОСТРАНСТВА. АФФИННАЯ СИСТЕМА КООРДИНАТ, ИЗМЕНЕНИЕ КООРДИНАТ ПРИ ЗАМЕНЕ СИСТЕМЫ КООРДИНАТ	72
47	АФФИННО НЕЗАВИСИМЫЕ СИСТЕМЫ ТОЧЕК. БАРИЦЕНТРИЧЕСКАЯ КОМБИНАЦИЯ ТОЧЕК. ПРИМЕРЫ	73
48	АФФИННОЕ ПОДПРОСТРАНСТВО (ПЛОСКОСТЬ), ЕГО РАЗМЕРНОСТЬ. ГЕОМЕТРИЧЕСКИЙ СМЫСЛ МНОЖЕСТВА РЕШЕНИЙ НЕОДНОРОДНОЙ СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ. ЗАДАНИЕ ПЛОСКОСТИ СИСТЕМОЙ ЛИНЕЙНЫХ УРАВНЕНИЙ	74
49	ВЗАИМНОЕ РАСПОЛОЖЕНИЕ ДВУХ ПЛОСКОСТЕЙ. ПЕРЕСЕЧЕНИЕ И АФФИННАЯ ОБОЛОЧКА ДВУХ АФФИННЫХ ПЛОСКОСТЕЙ. РАЗМЕРНОСТЬ АФФИННОЙ ОБОЛОЧКИ	76
50	АФФИННОЕ ОТОБРАЖЕНИЕ И ЕГО ЛИНЕЙНАЯ ЧАСТЬ. КОМПОЗИЦИЯ АФФИННЫХ ОТОБРАЖЕНИЙ. КРИТЕРИЙ ОБРАТИМОСТИ В ТЕРМИНАХ ЛИНЕЙНОЙ ЧАСТИ. МАТРИЦА АФФИННОГО ОТОБРАЖЕНИЯ. СУЩЕСТВОВАНИЕ И ЕДИНСТВЕННОСТЬ АФФИННОГО ОТОБРАЖЕНИЯ, ДЕЙСТВУЮЩЕГО НА СИСТЕМУ ИЗ $n + 1$ АФФИННО НЕЗАВИСИМЫХ ТОЧЕК n -МЕРНОГО АФФИННОГО ПРОСТРАНСТВА	77
51	АФФИННЫЕ ПРЕОБРАЗОВАНИЯ. РАЗЛОЖЕНИЕ БИЕКТИВНОГО АФФИННОГО ПРЕОБРАЗОВАНИЯ В КОМПОЗИЦИЮ ПАРАЛЛЕЛЬНОГО ПЕРЕНОСА И ПРЕОБРАЗОВАНИЯ С НЕПОДВИЖНОЙ ТОЧКОЙ	78
52	АФФИННОЕ ЕВКЛИДОВО ПРОСТРАНСТВО. ТЕОРЕМА ОБ ИЗОМОРФИЗМЕ. РАССТОЯНИЕ МЕЖДУ ДВУМЯ ПЛОСКОСТЯМИ	79
53	ОРТОГОНАЛЬНЫЕ АФФИННЫЕ ПРЕОБРАЗОВАНИЯ (ДВИЖЕНИЯ). СОБСТВЕННЫЕ И НЕСОБСТВЕННЫЕ ДВИЖЕНИЯ. РАЗЛОЖЕНИЕ ДВИЖЕНИЯ В КОМПОЗИЦИЮ ПАРАЛЛЕЛЬНОГО ПЕРЕНОСА И ОРТОГОНАЛЬНОГО ПРЕОБРАЗОВАНИЯ С НЕПОДВИЖНОЙ ТОЧКОЙ	80
54	НЕКОТОРЫЕ АФФИННЫЕ И ЛИНЕЙНЫЕ ГРУППЫ	81
55	АФФИННО-КВАДРАТИЧНЫЕ ФУНКЦИИ. ЦЕНТР. ПРИВЕДЕНИЕ АФФИННО-КВАДРАТИЧНОЙ ФУНКЦИИ К КАНОНИЧЕСКОМУ ВИДУ ЗАМЕНОЙ АФФИННОЙ СИСТЕМЫ КООРДИНАТ	81
56	ПОВЕРХНОСТИ ВТОРОГО ПОРЯДКА (КВАДРИКИ). ПЕРЕСЕЧЕНИЕ КВАДРИКИ С ПРЯМОЙ. ЦЕНТР И ВЕРШИНА КВАДРИКИ	83
57	КВАДРИКИ В ЕВКЛИДОВОМ (ТОЧЕЧНОМ) ПРОСТРАНСТВЕ. ОРТОГОНАЛЬНАЯ КЛАССИФИКАЦИЯ КВАДРИК. АФФИННАЯ КЛАССИФИКАЦИЯ.	86
58	ОПРЕДЕЛЕНИЕ ТЕНЗОРА ТИПА (p, q) НА ВЕКТОРНОМ ПРОСТРАНСТВЕ. ЛИНЕЙНЫЕ ОПЕРАЦИИ. ОТОЖДЕСТВЛЕНИЕ ТЕНЗОРОВ МАЛЫХ ВАЛЕНТНОСТЕЙ С ГЕОМЕТРИЧЕСКИМИ ОБЪЕКТАМИ	88

59	ПРОИЗВЕДЕНИЕ ТЕНЗОРОВ. БАЗИС В ПРОСТРАНСТВЕ ТЕНЗОРОВ ТИПА (p, q) . КООРДИНАТЫ ТЕНЗОРА, ИХ ИЗМЕНЕНИЕ ПРИ ЗАМЕНЕ БАЗИСА В ОСНОВНОМ ПРОСТРАНСТВЕ. ИНВАРИАНТ. СВЕРТКИ ТЕНЗОРА	89
60	СИММЕТРИЧЕСКИЕ И КОСОСИММЕТРИЧЕСКИЕ ТЕНЗОРЫ ТИПОВ $(p, 0)$ или $(0, q)$. СИММЕТРИЗАЦИЯ И АЛЬТЕРНИРОВАНИЕ	92
61	ТЕНЗОРНАЯ АЛГЕБРА ВЕКТОРНОГО ПРОСТРАНСТВА. ВНЕШНЯЯ АЛГЕБРА И ЕЕ РАЗМЕРНОСТЬ	93
62	ТЕНЗОРЫ НА ЕВКЛИДОВОМ ПРОСТРАНСТВЕ	96
63	КОММУТИРУЮЩИЕ ОПЕРАТОРЫ НА КОНЕЧНОМЕРНОМ КОМПЛЕКСНОМ ВЕКТОРНОМ ПРОСТРАНСТВЕ	97
64	КОСОСИММЕТРИЧЕСКИЕ ФОРМЫ. СИМПЛЕКТИЧЕСКАЯ ГРУППА	97
65	ПРИЛОЖЕНИЯ	103
65.1	ПРОЕКТОРЫ	103
65.2	МЕТОД НАИМЕНЬШИХ КВАДРАТОВ	104
65.3	СПЕКТРАЛЬНОЕ РАЗЛОЖЕНИЕ САМОСОПРЯЖЁННОГО ОПЕРАТОРА	106
65.4	КОСОСИММЕТРИЧЕСКИЕ И КОСОЭРМИТОВЫ ОПЕРАТОРЫ. КАНОНИЧЕСКИЙ ВИД. ЭРМИТОВО РАЗЛОЖЕНИЕ	107
65.5	ВЫЧИСЛЕНИЕ МНОГОЧЛЕНОВ И ФУНКЦИЙ ОТ МАТРИЦ. ЭКСПОНЕНТА ЛИНЕЙНОГО ОПЕРАТОРА	108
65.6	БАРИЦЕНТРИЧЕСКИЕ КООРДИНАТЫ	108
65.7	ГЕОМЕТРИЯ АФФИННЫХ ЕВКЛИДОВЫХ ПРОСТРАНСТВ. МНОГОГРАННИКИ И ПРИЛОЖЕНИЯ	109
66	ТЕОРЕТИЧЕСКИЕ ЗАДАЧИ	110

1. ВЕКТОРНЫЕ ПРОСТРАНСТВА. ЛИНЕЙНАЯ ЗАВИСИМОСТЬ И НЕЗАВИСИМОСТЬ ВЕКТОРОВ. БАЗИС, РАЗМЕРНОСТЬ. ПРИМЕРЫ

Определение 1. *Линейным (или векторным) пространством над полем \mathbb{F} называется множество V с заданными на нём операциями сложения $+$: $(u, v) \in V \times V \mapsto (u + v) \in V$ и умножения элементов V на элементы \mathbb{F} : $(\lambda, v) \in \mathbb{F} \times V \mapsto (\lambda \cdot v) \in V$, удовлетворяющие следующим аксиомам:*

1. $v + u = u + v \quad \forall u, v \in V$;
 2. $(u + v) + w = u + (v + w) \quad \forall u, v, w \in V$;
 3. $\exists \mathbf{0} \in V : v + \mathbf{0} = v \quad \forall v \in V$;
 4. $\forall v \in V \exists (-v) \in V : (-v) + v = \mathbf{0}$;
 5. $\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v \quad \forall u, v \in V, \forall \lambda \in \mathbb{F}$;
 6. $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v \quad \forall v \in V, \forall \lambda, \mu \in \mathbb{F}$;
 7. $\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v \quad \forall v \in V, \forall \lambda, \mu \in \mathbb{F}$;
 8. $1 \cdot v = v \quad \forall v \in V$.
- $\left. \begin{array}{l} 1-4 \\ 5-8 \end{array} \right\} (V, +) \text{ — абелева группа}$
 $\left. \begin{array}{l} 5-8 \end{array} \right\} \mathbb{F} \text{ линейно действует на } V$

Определение 2. Элементы множества V называются *векторами*, элемент $\mathbf{0}$ называется *нулевым вектором*, а элемент $(-v)$ называется *противоположным к v* . Элементы \mathbb{F} называют *скалярами*.

Предложение 1.

- 1°. $\mathbf{0} \cdot v = \lambda \cdot \mathbf{0} = \mathbf{0} \quad \forall v \in V, \forall \lambda \in \mathbb{F}$;
- 2°. $(-1) \cdot v = -v \quad \forall v \in V$;
- 3°. $\lambda v = \mathbf{0} \Rightarrow \lambda = 0$ или $v = \mathbf{0}$.

Доказательство.

- 1°. $0 \cdot v + 0 \cdot v = (0 + 0) \cdot v = 0 \cdot v \Rightarrow 0 \cdot v = \mathbf{0}$. Аналогично, $\lambda \cdot \mathbf{0} + \lambda \cdot \mathbf{0} = \lambda(\mathbf{0} + \mathbf{0}) = \lambda \cdot \mathbf{0} \Rightarrow \lambda \cdot \mathbf{0} = \mathbf{0}$;
- 2°. $v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1 - 1) \cdot v = 0 \cdot v = \mathbf{0} \Rightarrow -v = (-1) \cdot v$.
- 3°. $\lambda \neq 0 \Rightarrow \mathbf{0} = \lambda^{-1} \lambda v = v$.

■

Задача 1 (А. А. Клячко¹). Найдите лишнюю аксиому и докажите, что она лишняя, а все остальные не лишние.

▷ Лишней является аксиома коммутативности сложения (первая), докажем это. Заметим, что предложение 1 было доказано без использования коммутативности, значит, им здесь можно пользоваться. Нам будет нужен ещё два пункта (везде над переходами стоят номера пунктов или аксиом, которые использовались при этом переходе):

- 4°. $v - v \stackrel{8,2^\circ}{=} 1 \cdot v + (-1) \cdot v \stackrel{6}{=} (1 - 1) \cdot v = 0 \cdot v \stackrel{1^\circ}{=} \mathbf{0} \Rightarrow \boxed{v - v = \mathbf{0}}$;
- 5°. $\mathbf{0} + v \stackrel{4^\circ}{=} (v - v) + v \stackrel{2}{=} v + \underbrace{(-v + v)}_{=0} \stackrel{3,4}{=} v \Rightarrow \boxed{\mathbf{0} + v = v}$.

Теперь мы можем вывести первую аксиому из остальных:

$$\begin{aligned}
 u + v &\stackrel{3}{=} (u + v) + \mathbf{0} \stackrel{4}{=} (u + v) + (-v + u) + (v + u) \stackrel{2}{=} ((u + v) - (v + u)) + (v + u) \stackrel{2}{=} \\
 &= (u + \underbrace{v - v}_{=0} - u) + (v + u) \stackrel{4^\circ}{=} (\underbrace{u - u}_{=0}) + (v + u) \stackrel{4^\circ}{=} \mathbf{0} + (v + u) \stackrel{5^\circ}{=} v + u.
 \end{aligned}$$

¹Ещё в первом семестре дело было...

Антон Александрович сказал, что остальные не выводятся. Но мы с Костей Зюбиным не смогли доказать это для 5-ой аксиомы. Доказательство независимости от остальных для каждой аксиомы проводится так: нужно привести пример такой структуры, в которой выполняются все аксиомы, кроме выбранной.

- **2 аксиома.** Рассмотрим множество $M = \{e, a, b\}$ с операцией $*$, заданной таблицей:

$*$	e	a	b
e	e	a	b
a	a	a	e
b	b	e	b

Заметим, что операция $*$ коммутативна, в M существует нейтральный элемент e по $*$ и каждый элемент имеет обратный по $*$. Однако эта операция не ассоциативна:

$$(b * a) * a = e * a = a, \quad b * (a * a) = b * a = e.$$

Теперь возьмём алгебраическую систему $V = (\mathbb{R} \times M, +, \star)$ с операциями, определёнными по следующим правилам:

$$u + v = (a, x) + (b, y) := (a + b, x * y), \quad \lambda \star v = \lambda \star (a, x) := (\lambda a, x).$$

Аксиома 2 не выполнена, т. к. $*$ неассоциативна. Выполнение аксиом 1, 3, 4 следует из того, что они выполняются для $+$ над \mathbb{R} и $*$ над M . Проверим выполнение остальных аксиом:

$$\begin{aligned} 5: & \quad \lambda \star (a + b, x * y) = (\lambda(a + b), x * y) = (\lambda a + \lambda b, x * y) = (\lambda a, x) + (\lambda b, y) = \lambda \star ((a, x) + (b, y)) \\ 6: & \quad (\lambda + \mu) \cdot (a, x) = ((\lambda + \mu)a, x) = (\lambda a + \mu a, x * x) = (\lambda a, x) + (\mu a, x) = \lambda \star (a, x) + \mu \star (a, x) \\ 7: & \quad (\lambda \mu) \star (a, x) = (\lambda \mu a, x) = \lambda \star (\mu a, x) \\ 8: & \quad 1 \star (a, x) = (1 \cdot a, x) = (a, x) \end{aligned}$$

- **3 аксиома.** Без третьей аксиомы нельзя ввести четвёртую, поэтому её удаление не имеет смысла.
- **4 аксиома.** Рассмотрим алгебраическую систему $V = (\mathbb{R} \cup \{\infty\}, +, \cdot)$. Доопределим сложение и умножение для ∞ следующим образом:

$$\infty + a = a + \infty := \infty, \quad \lambda \cdot \infty := \infty.$$

Выполнение аксиом 1-3 сразу вытекает из определения. Аксиома 4 не выполнена, т. к. у ∞ нет обратного по $+$. Выполнение аксиом 5-8 проверяется перебором нескольких случаев.

- **6 аксиома.** Рассмотрим алгебраическую систему $V = (\mathbb{R}, +, \star)$, в которой сложение определено так же, как в действительных числах, а умножение так:

$$\lambda \star v := v.$$

Аксиомы 1-4 выполнены, т. к. они выполнены для \mathbb{R} и $+$. Выполнение аксиом 5, 7 и 8 сразу вытекает из определения. Аксиома 6 не выполнена:

$$u + u = 1 \star u + 1 \star u, \quad (1 + 1) \star u = u.$$

- **7 аксиома.** Рассмотрим \mathbb{R} как векторное пространство над полем \mathbb{Q} с базисом $M \supset \{1, \sqrt{2}\}$. Пусть отображение $f : \mathbb{R} \rightarrow \mathbb{R}$ задаётся своими значениями на числах из M , а для других чисел определяется соотношением

$$f(q_1 v_1 + q_2 v_2 + \dots + q_n v_n) = q_1 f(v_1) + q_2 f(v_2) + \dots + q_n f(v_n),$$

где v_i — некоторые векторы из базиса M . Такое отображение является линейным, сохраняющим все рациональные числа.

Теперь возьмём алгебраическую систему $V = (\mathbb{R}, +, \star)$, в которой сложение $+$ определено естественным образом, а умножение \star определяется через f и естественное умножение \cdot :

$$\lambda \star u := f(\lambda) \cdot u.$$

Аксиомы 1-4 выполнены, т. к. они выполнены для \mathbb{R} и $+$. Выполнение аксиомы 5 проверяется непосредственно. Аксиома 6 выполнена, т. к. отображение f линейно. Проверим, что аксиома 7 не выполнена:

$$\sqrt{2} \star (\sqrt{2} \star u) = \sqrt{2} \star u = u, \quad (\sqrt{2} \star \sqrt{2}) \star u = 2u.$$

Выполнение аксиомы 8 вытекает из определения.

- **8 аксиома.** Рассмотрим алгебраическую систему $V = (\mathbb{R}, +, \star)$, в которой сложение определено естественным образом, а умножение так:

$$\lambda \star u := 0.$$

Аксиомы 1-4 выполнены, т. к. они выполнены для \mathbb{R} и $+$. Выполнение аксиом 5-7 проверяется непосредственно. Аксиома 8 не выполнена.

◀

Пример 1.

1. Множество $\{0\}$ из одного элемента является линейным пространством над любым полем.
2. Множества геометрических векторов на прямой, плоскости или пространстве являются линейными пространствами над полем \mathbb{R} .
3. Поле \mathbb{F} является векторным пространством над самим собой.
4. Поле \mathbb{C} является линейным пространством над полем \mathbb{R} , а поле \mathbb{R} является линейным пространством над полем \mathbb{Q} .

5. Пусть $\mathbb{F}^n := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} : x_i \in \mathbb{F} \right\}$ — множество столбцов фиксированной длины n из элементов поля \mathbb{F} . Операции покомпонентного сложения и умножения на скаляры

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

задают на \mathbb{F}^n структуру линейного пространства над \mathbb{F} . Его часто называют *координатным*.

Пусть V — линейное пространство над полем \mathbb{F} .

Определение 3. *Линейной комбинацией* системы векторов $\{v_i : i \in I\}$ называется формальная сумма вида $\sum_{i \in I} \lambda_i v_i$, в которой лишь конечное число скаляров λ_i отличны от нуля.

Примечание. Линейную комбинацию системы $\{v_i : i \in I\}$ можно также определить как функцию $i \in I \mapsto \lambda_i \in \mathbb{F}$, которая принимает ненулевое значение только на конечном числе индексов.

Доказательство. Пусть V — конечномерное пространство и в V существует базис e_1, \dots, e_m . Пусть $\{f_i : i \in I\}$ — другой базис. Если этот базис бесконечен, то в нём содержится конечная линейно независимая система векторов f_1, \dots, f_n , где $n > m$. При этом, т. к. e_1, \dots, e_m — базис, мы имеем $\{f_1, \dots, f_n\} \subseteq \langle e_1, \dots, e_m \rangle$, что противоречит лемме о линейной зависимости. Следовательно, базис $\{f_i : i \in I\}$ конечен, т. е. имеет вид f_1, \dots, f_n . Тогда $\{f_1, \dots, f_n\} \subseteq \langle e_1, \dots, e_m \rangle$ и $\{e_1, \dots, e_m\} \subseteq \langle f_1, \dots, f_m \rangle$. Отсюда $n = m$. ■

Лемма 3. В конечномерном пространстве любую линейно независимую систему можно дополнить до базиса.

Доказательство. Пусть $\{e_1, \dots, e_k\}$ — конечная подсистема в V . Тогда, если эта система максимальна по включению, то она базис. Иначе существует $e_{k+1} \in V$ такой, что система $\{e_1, \dots, e_k, e_{k+1}\}$ линейно независима. Продолжая процесс далее, за конечное число шагов получим базис (в силу конечномерности пространства V). ■

Примечание. Ниже изложен удобный алгоритм дополнения линейно независимой системы до базиса. Пусть

$$u_1 = \begin{pmatrix} u_{11} \\ \vdots \\ u_{1n} \end{pmatrix}, \quad u_2 = \begin{pmatrix} u_{21} \\ \vdots \\ u_{2n} \end{pmatrix}, \quad \dots, \quad u_m = \begin{pmatrix} u_{m1} \\ \vdots \\ u_{mn} \end{pmatrix}$$

линейно независимы. Обозначим

$$U = \begin{pmatrix} u_{11} & u_{21} & \dots & u_{m1} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n} & u_{2n} & \dots & u_{mn} \end{pmatrix}.$$

Отметим, что $\text{rk } U = m$, а $\text{rk}(U \mid E) = n$. Так что нужно привести матрицу $(U \mid E)$ элементарными преобразованиями строк к ступенчатому виду и дополнить столбцы матрицы U единичными столбцами, вошедшими в базис матрицы $(U \mid E)$ (в первом семестре была теорема, что при элементарных преобразованиях строк линейные зависимости между столбцами не меняются).

Лемма 4. Всякое конечномерное линейное пространство V обладает базисом. Более точно, из всякого конечного порождающего множества $S \subseteq V$ можно выбрать базис пространства V .

Доказательство. Если множество S линейно зависимо, то по лемме 1 в нём найдётся вектор, линейно выражающийся через остальные. Выкидывая этот вектор, мы получаем порождающее множество из меньшего числа векторов. Продолжая так дальше, мы в конце концов получим линейно независимое порождающее множество, т. е. базис. ■

Примечание. Чтобы сделать это на практике, выписываем векторы в матрицу по столбцам, приводим её к ступенчатому виду и те столбцы, в которых стоят лидеры, будут базисными.

Определение 9. *Размерностью* конечномерного линейного пространства V (обозначается $\dim V$) называется число элементов в базисе V . Если V бесконечномерно, то пишут $\dim V = \infty$.

Примечание. В нулевом пространстве $\{0\}$ базисом естественно считать пустое множество \emptyset . Поэтому $\dim\{0\} = 0$.

2. МАТРИЦА ПЕРЕХОДА ОТ ОДНОГО БАЗИСА К ДРУГОМУ. ИЗМЕНЕНИЕ КООРДИНАТ ВЕКТОРА ПРИ ЗАМЕНЕ БАЗИСА

Определение 1. Пусть V — линейное пространство, и e_1, \dots, e_n — базис в V . Любой вектор $x \in V$ единственным образом представляется в виде линейной комбинации базисных векторов: $x = \sum_{i=1}^n x_i e_i$. Числа $x_1, \dots, x_n \in \mathbb{F}$ называются *координатами* вектора x в базисе e_1, \dots, e_n .

Обозначения Эйнштейна. Вместо $\sum_{i=1}^n x_i e_i$ пишем $x^i e_i$ (суммирование производится по повторяющемуся индексу). В связи с этим обозначением, нам будет также удобно обозначать координаты вектора верхними индексами вместо нижних. Для произведения матриц: $c_k^i = a_j^i b_k^j$ (суммирование опять производится по повторяющемуся индексу; причём важно, что сначала повторяющийся индекс j стоит внизу, а потом сверху). Матрица (d_k^j) является обратной к (c_j^i) , если $c_j^i d_k^j = \delta_k^i$ — символ Кронекера.

Пусть в пространстве V заданы два базиса: «старый» e_1, \dots, e_n и «новый» $e'_1, \dots, e'_{n'}$. Нам будет удобно обозначать векторы нового базиса через $e'_1, \dots, e'_{n'}$. Элементы нового базиса выражаются через элементы старого: $e'_i = c_{i'}^i e_i$, $i' = 1, \dots, n'$. Эти формулы равносильны одному матричному равенству

$$(e'_1, \dots, e'_{n'}) = (e_1, \dots, e_n) \cdot \begin{pmatrix} c_{1'}^1 & \cdots & c_{n'}^1 \\ \vdots & \ddots & \vdots \\ c_{1'}^n & \cdots & c_{n'}^n \end{pmatrix}.$$

Определение 2. Матрица $C := (c_{i'}^i)$ называется *матрицей перехода* от базиса e_1, \dots, e_n к базису $e'_1, \dots, e'_{n'}$. Её столбцами являются координаты новых базисных векторов в старом базисе.

Предложение 1.

1. Матрица $C_{e' \rightarrow e} = (c_{i'}^i)$ перехода от базиса $e'_1, \dots, e'_{n'}$ к базису e_1, \dots, e_n является обратной к матрице $C_{e \rightarrow e'}$ перехода от e_1, \dots, e_n к $e'_1, \dots, e'_{n'}$, т. е. $C_{e \rightarrow e'} \cdot C_{e' \rightarrow e} = E$. В частности, матрица перехода всегда невырождена.
2. Если e_1, \dots, e_n , $e'_1, \dots, e'_{n'}$, $e''_1, \dots, e''_{n''}$ — три базиса, то для соответствующих матриц перехода выполнено $C_{e \rightarrow e'} \cdot C_{e' \rightarrow e''} = C_{e \rightarrow e''}$.

Доказательство. Первое утверждение следует из второго, если положить $e'' = e$, поэтому будем доказывать второе утверждение. Пусть $C_{e \rightarrow e'} = (c_{i'}^i)$, $C_{e' \rightarrow e''} = (c_{i''}^{i'})$, $C_{e \rightarrow e''} = (c_{i''}^i)$. Тогда

$$c_{i''}^i e_i = e_{i''} = c_{i''}^{i'} e'_{i'} = c_{i''}^{i'} c_{i'}^i e_i = c_{i'}^i c_{i''}^{i'} e_i \Rightarrow c_{i''}^i = c_{i'}^i \cdot c_{i''}^{i'}.$$

■

Примечание. Важное практическое следствие. Заметим, что в \mathbb{R}^n писать матрицу перехода от стандартного базиса к любому другому очень легко — достаточно написать базисные векторы, в которые мы хотим перейти, по столбцам матрицы. Пусть мы хотим написать матрицу перехода от базиса $a = \{a_1, a_2, \dots, a_n\}$ к базису $b = \{b_1, b_2, \dots, b_n\}$. Для этого можно написать матрицу A перехода от стандартного к a , потом матрицу B перехода от стандартного к b , а потом выдать ответ — $A^{-1}B$.

Примечание. Трюк от Александра Александровича Гайфуллина (и в Винберге находил). Чтобы найти матрицу $X = A^{-1}B$ при известных A и B , не надо искать A^{-1} , а потом умножать её на B . Домножим на A слева, получим $AX = B$. Это n систем линейных уравнений с одной и той же матрицей A (решив i -ую систему, найдём i -ый столбец X). Эти системы можно решать одновременно, записав матрицу A в правой части и приписав к ней каждый столбец B по очереди. Выглядеть будет как $(A \mid B)$. Решить системы — значит привести эту матрицу $n \times 2n$ к улучшенному ступенчатому виду: $(E \mid X)$. В правой части теперь будут стоять столбцы матрицы X .

Теорема 1 (Закон изменения координат). Пусть x^1, \dots, x^n — координаты вектора x в базисе e_1, \dots, e_n , а $x^{1'}, \dots, x^{n'}$ — в базисе $e'_1, \dots, e'_{n'}$. Тогда два набора координат связаны формулой

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

Очевидно, что нулевой столбец является её решением и что произведение любого решения на число также является решением. Докажем, что сумма решений $(u_1, \dots, u_n)^t$ и $(v_1, \dots, v_n)^t$ является решением. Подставляя её компоненты в i -ое уравнение системы, получаем

Определение 2. *Фундаментальная система решений* — это базис подпространства решений однородной СЛУ.

Тогда задать его линейной оболочкой можно, найдя ФСР, здесь приведён алгоритм, как это делать. С помощью элементарных преобразований приведём систему к улучшенному ступенчатому виду. Число ненулевых уравнений в этом ступенчатом виде равно $r = \text{rk } A$. Поэтому общее решение будет содержать r главных неизвестных и с точностью до перенумерации неизвестных будет иметь вид

[illegible]

Придавая поочерёдно одному из свободных неизвестных $x_{r+1}, x_{r+2}, \dots, x_n$ значение 1, а остальным — 0, получим следующие решения системы:

$$u_1 = \begin{pmatrix} c_{11} \\ \vdots \\ c_{r1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} c_{12} \\ \vdots \\ c_{r2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad u_{n-r} = \begin{pmatrix} c_{1,n-r} \\ \vdots \\ c_{r,n-r} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Ранг системы векторов $\{u_1, u_2, \dots, u_{n-r}\}$ равен рангу матрицы

$$\left(\begin{array}{cccc|cccc} c_{11} & c_{21} & \dots & c_{r1} & 1 & 0 & \dots & 0 \\ c_{12} & c_{22} & \dots & c_{r2} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{1,n-r} & c_{2,n-r} & \dots & c_{r,n-r} & 0 & 0 & \dots & 1 \end{array} \right)$$

Если поменять местами блоки, отделённые чертой, то получится улучшенный ступенчатый вид с количеством ступенек, равным $n - r$. Так что ранг системы векторов $\{u_1, u_2, \dots, u_{n-r}\}$ равен количеству векторов в этой системе, поэтому она линейно независима. Эта система также порождает всё подпространство решений, т. к. любая линейная комбинация вида

$$\lambda_1 u_1 + \dots + \lambda_{n-r} u_{n-r}$$

является решением, в котором свободные неизвестные имеют значения $\lambda_1, \dots, \lambda_{n-r}$.

Теперь пусть подпространство задано линейной оболочкой

$$\left\langle u_1 = \begin{pmatrix} u_{11} \\ \vdots \\ u_{1n} \end{pmatrix}, \quad u_2 = \begin{pmatrix} u_{21} \\ \vdots \\ u_{2n} \end{pmatrix}, \quad \dots, \quad u_m = \begin{pmatrix} u_{m1} \\ \vdots \\ u_{mn} \end{pmatrix} \right\rangle$$

Составим матрицу

$$U = \begin{pmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{m1} & u_{m2} & \dots & u_{mn} \end{pmatrix}$$

из строк $u_1^t, u_2^t, \dots, u_n^t$. Найдём ФСР системы $Ux = 0$ (указанным выше способом), и запишем её векторы по строкам в матрицу U' . По лемме 1 пространство решений однородной системы линейных уравнений $U'y = 0$ есть линейная оболочка строк матрицы U , а это и есть данные нам векторы. ■

Предложение 3. Пересечение $V_1 \cap V_2$ подпространств в V является подпространством в V .

Доказательство. Во-первых, $\mathbf{0} \in V_1$ и $\mathbf{0} \in V_2$, поэтому $\mathbf{0} \in V_1 \cap V_2 \neq \emptyset$. Во-вторых, $\forall u, v \in V_1 \cap V_2$, $\forall \lambda \in \mathbb{F}$ сумма $u + v$ и произведение λv также лежат в V_1 и в V_2 , а значит, и в $V_1 \cap V_2$. ■

Примечание. Аналогично доказывается, что для любого семейства подпространств $\{U_i : i \in I\}$ их пересечение $\bigcap_{i \in I} U_i$ тоже подпространство.

Определение 3 (Сумма подпространств). $V_1 + V_2 := \{v_1 + v_2 : v_1 \in V_1, v_2 \in V_2\}$.

Предложение 4. $V_1 + V_2 = \langle V_1 \cup V_2 \rangle$.

Доказательство. Включение $V_1 + V_2 \subseteq \langle V_1 \cup V_2 \rangle$ следует из того, что вектор $v_1 + v_2$ является линейной комбинацией векторов $v_1, v_2 \in V_1 \cup V_2$. Докажем обратное включение. Для этого рассмотрим линейную комбинацию $v = \lambda_1 u_1 + \dots + \lambda_n u_n$ векторов $u_1, \dots, u_n \in V_1 \cup V_2$. Можно считать, что $u_1, \dots, u_k \in V_1$ и $u_{k+1}, \dots, u_n \in V_2$. Тогда мы имеем $v = v_1 + v_2$, где $v_1 = \lambda_1 u_1 + \dots + \lambda_k u_k \in V_1$ и $v_2 = \lambda_{k+1} u_{k+1} + \dots + \lambda_n u_n \in V_2$. Следовательно, $v \in V_1 + V_2$. ■

Примечание. Само объединение $V_1 \cup V_2$ подпространств в общем случае не является подпространством. Примером служит объединение двух прямых на плоскости.

Теорема 2 (Формула Грассмана). $\dim V_1 + \dim V_2 = \dim(V_1 + V_2) + \dim(V_1 \cap V_2)$.

Доказательство. Выберем базис e_1, \dots, e_k пространства $V_1 \cap V_2$. Воспользовавшись леммой 3 из первого вопроса, можем дополнить его до базиса $e_1, \dots, e_k, f_1, \dots, f_l$ пространства V_1 и до базиса $e_1, \dots, e_k, g_1, \dots, g_m$ пространства V_2 . Тогда мы имеем $\dim(V_1 \cap V_2) = k$, $\dim V_1 = k + l$, $\dim V_2 = k + m$. Докажем, что $e_1, \dots, e_k, f_1, \dots, f_l, g_1, \dots, g_m$ — базис пространства $V_1 + V_2$. Заметим, что т. к. $V_1 + V_2 = \langle V_1 \cup V_2 \rangle$, то любой вектор из $V_1 + V_2$ выражается через эту систему векторов. Остаётся проверить, что эта система линейно независима. Пусть имеет место равенство

$$\lambda_1 e_1 + \dots + \lambda_k e_k + \mu_1 f_1 + \dots + \mu_l f_l + \nu_1 g_1 + \dots + \nu_m g_m = \mathbf{0}.$$

Перепишем его в виде

$$\lambda_1 e_1 + \dots + \lambda_k e_k + \mu_1 f_1 + \dots + \mu_l f_l = -\nu_1 g_1 - \dots - \nu_m g_m.$$

Вектор, стоящий в обеих частях этого равенства, лежит и в V_1 , и в V_2 , а значит, и в $V_1 \cap V_2$, а потому линейно выражается через e_1, \dots, e_k . Т. к. векторы $e_1, \dots, e_k, f_1, \dots, f_l$ линейно независимы по построению, получаем $\mu_1 = \dots = \mu_l = 0$. Аналогично, $\nu_1 = \dots = \nu_m = 0$. Тогда из линейной независимости e_1, \dots, e_k следует $\lambda_1 = \dots = \lambda_k = 0$. Итак, $\dim(V_1 + V_2) = k + l + m$, откуда вытекает требуемое. ■

Алгоритм вычисления базисов в $V_1 + V_2$ и $V_1 \cap V_2$ (V_1 и V_2 — подпространства конечномерного пространства V). При доказательстве формулы Грассмана мы попутно показали, что объединение базисов V_1 и V_2 — полная подсистема в $V_1 + V_2$, значит, из неё можно выделить базис (выше в вопросе 1, как). Отметим, что в случае, когда сумма $V_1 + V_2$ прямая, пересечение этих базисов пустое, и ничего дополнительно делать не нужно.

Чтобы найти базис пересечения, зададим оба подпространства в виде системы линейных уравнений. Система, состоящая из всех уравнений обеих систем, будет задавать пересечение этих подпространств. Для нахождения базиса ищем ФСР.

5. ПРЯМАЯ СУММА ПОДПРОСТРАНСТВ. ВНЕШНЯЯ ПРЯМАЯ СУММА ВЕКТОРНЫХ ПРОСТРАНСТВ

Определение 1. Сумма $V_1 + V_2$ подпространств пространства V называется *прямой* (обозначается $V_1 \oplus V_2$), если $\forall v \in V_1 + V_2$ представление $v = v_1 + v_2$, где $v_1 \in V_1, v_2 \in V_2$, единственно.

Теорема 1. Следующие условия эквивалентны для подпространств V_1 и V_2 :

1. сумма $V_1 + V_2$ прямая;
2. $V_1 \cap V_2 = \{\mathbf{0}\}$;
3. если $\mathbf{0} = v_1 + v_2$, где $v_1 \in V_1$ и $v_2 \in V_2$, то $v_1 = v_2 = \mathbf{0}$;
4. $\dim(V_1 + V_2) = \dim V_1 + \dim V_2$.

Доказательство. $1 \Rightarrow 2$. Пусть найдётся $v \in V_1 \cap V_2$, $v \neq \mathbf{0}$. Тогда $\mathbf{0} = \mathbf{0} + \mathbf{0} = v + (-v)$. Получаем, что представление вектора $\mathbf{0}$ не единственно, и сумма $V_1 + V_2$ не прямая.

$2 \Rightarrow 3$. Если существует представление $\mathbf{0} = v_1 + v_2$, где $v_1 \in V_1$ и $v_2 = (-v_1) \in V_2$ и $v_1 \neq \mathbf{0}$, т. е. $v_1 \in V_1 \cap V_2 \neq \{\mathbf{0}\}$ — противоречие.

$3 \Rightarrow 1$. Пусть у вектора $v \in V$ есть два разложения $v = u_1 + u_2 = v_1 + v_2$, где $u_1, v_1 \in V_1$ и $u_2, v_2 \in V_2$. Тогда $\mathbf{0} = (u_1 - v_1) + (u_2 - v_2)$, где $u_1 - v_1 \in V_1$ и $u_2 - v_2 \in V_2$. Следовательно, $u_1 - v_1 = u_2 - v_2 = \mathbf{0}$, т. е. два разложения совпадают.

$2 \Leftrightarrow 4$. Следствие формулы Грассмана. ■

Определение 2. Сумма $V_1 + \dots + V_n$ подпространств пространства V называется *прямой* (обозначается $V_1 \oplus \dots \oplus V_n$), если $\forall v \in V_1 + \dots + V_n$ представление $v = v_1 + v_2 + \dots + v_n$, где $v_i \in V_i$, единственно.

Теорема 2. Для подпространств V_1, \dots, V_n пространства V следующие условия эквивалентны:

1. сумма $V_1 + \dots + V_n$ прямая;
2. $V_i \cap \sum_{j \neq i} V_j = \{\mathbf{0}\}$;
3. если $\mathbf{0} = v_1 + \dots + v_n$, где $v_i \in V_i$, то $v_1 = \dots = v_n = \mathbf{0}$;
4. $\dim \left(\sum_i V_k \right) = \sum_i \dim V_i$.

Доказательство. Индукция по n с помощью предыдущей теоремы. ■

Примечание. При $n \geq 3$ условие 2 в предыдущей теореме сильнее, чем условие $V_i \cap V_j = \{\mathbf{0}\} \forall i \neq j$. Это последнее условие не гарантирует, что сумма подпространств прямая. Действительно, рассмотрим следующие три подпространства в \mathbb{R}^2 со стандартным базисом e_1, e_2 : $V_1 := \langle e_1 \rangle$, $V_2 := \langle e_2 \rangle$ и $V_3 := \langle e_1 + e_2 \rangle$. Тогда $V_i \cap V_j = \{\mathbf{0}\} \forall i \neq j$, но сумма данных подпространств не прямая, т. к.

$$e_1 + e_2 = e_1 + e_2 + \mathbf{0} = \mathbf{0} + \mathbf{0} + (e_1 + e_2).$$

Определение 3. Пусть V_1, \dots, V_n — линейные пространства над одним полем \mathbb{F} . Их *внешней прямой суммой* (обозначается как $V_1 \oplus \dots \oplus V_n$) называется линейное пространство $V_1 \times \dots \times V_n$ с операциями, определёнными покомпонентно:

$$(u_1, \dots, u_n) + (v_1, \dots, v_n) = (u_1 + v_1, \dots, u_n + v_n), \quad \lambda \cdot (v_1, \dots, v_n) = (\lambda v_1, \dots, \lambda v_n).$$

Предложение 1. Для любого пространства $U \subseteq V$ найдётся подпространство $W \subseteq V$ такое, что $V = U \oplus W$.

Определение 4. Такое подпространство W называется *прямым дополнением* к U .

Доказательство. Пусть e_1, \dots, e_k — базис в U . Его можно дополнить до базиса V векторами e_{k+1}, \dots, e_n (где $n = \dim V$), тогда искомое $W := \langle e_{k+1}, \dots, e_n \rangle$. ■

Примечание. Пространство $V = V_1 + \dots + V_m$ (V, V_1, \dots, V_m — линейные пространства над одним полем \mathbb{F}) можно превратить в прямую сумму пространств. Рассмотрим $U_i = \{(\mathbf{0}, \dots, \mathbf{0}, v_i, \mathbf{0}, \dots, \mathbf{0}) : v_i \in V_i\} \subset V$. Тогда

$$(v_1, \dots, v_m) = (v_1, \mathbf{0}, \dots, \mathbf{0}) + (\mathbf{0}, v_2, \dots, \mathbf{0}) + \dots + (\mathbf{0}, \dots, \mathbf{0}, v_m),$$

а отсюда $V = U_1 \oplus \dots \oplus U_m$.

6. ФАКТОРПРОСТРАНСТВО, ЕГО РАЗМЕРНОСТЬ. КОРАЗМЕРНОСТЬ. СВЯЗЬ С РЕШЕНИЯМИ НЕОДНОРОДНОЙ СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

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

Определение 1. *Классом смежности* вектора $v \in V$ по подпространству W называется множество $v + W := \{v + w : w \in W\}$.

Лемма 1. Равенство $v_1 + W = v_2 + W$ имеет место тогда и только тогда, когда $v_1 - v_2 \in W$.

Доказательство. \Rightarrow . Пусть $v_1 + W = v_2 + W$. Тогда $v_1 \in v_1 + W = v_2 + W$, значит $\exists w \in W : v_1 = v_2 + w$. Следовательно, $v_1 - v_2 = w \in W$.

\Leftarrow . Обратно, пусть $v := v_1 - v_2 \in W$. Докажем, что $v_1 + W \subseteq v_2 + W$. Возьмём произвольный вектор $u \in v_1 + W$. Тогда $u = v_1 + w$ для некоторого $w \in W$. Мы имеем $u = v_1 + w = v_2 + (v + w)$, где $v + w \in W$. Следовательно, $u \in v_2 + W$ и $v_1 + W \subseteq v_2 + W$. Обратное включение доказывается аналогично. ■

Предложение 1. Отношение $v_1 \sim v_2 :\Leftrightarrow v_1 - v_2 \in W$ задаёт отношение эквивалентности на V .

Доказательство. Совсем несложно проверяются все аксиомы. ■

Определение 2. *Факторпространством* линейного пространства V по подпространству W называется множество $V/W := \{v + W : v \in V\}$ с операциями сложения и умножения на скаляры:

$$(u + W) + (v + W) := (u + v) + W, \quad \lambda \cdot (v + W) := \lambda v + W.$$

Предложение 2. Приведённые выше операции определены на классах смежности корректно и задают на V/W структуру линейного пространства.

Доказательство. Проверим корректность определения операций, т. е. независимость результата операции от выбора вектора v в смежном классе $v + W$. Докажем для сложения. Если $u_1 + W = u_2 + W$ и $v_1 + W = v_2 + W$, то $u := u_1 - u_2 \in W$ и $v := v_1 - v_2 \in W$ в силу предыдущей леммы. Следовательно,

$$\begin{aligned} (u_1 + W) + (v_1 + W) &= (u_1 + v_1) + W = (u_2 + v_2) + \overbrace{(u + v)}^{\in W} + W = \\ &= (u_2 + v_2) + W = (u_2 + W) + (v_2 + W). \end{aligned}$$

Корректность определения умножения на скаляры проверяется аналогично. Теперь докажем, что V/W — линейное пространство. Аксиомы 1 и 2 сразу следуют из определения. Нулём является $0 + W = W$, а противоположным к $v + W$ является $(-v) + W$. Проверим аксиому 5:

$$\begin{aligned} \lambda \cdot ((u + W) + (v + W)) &= \lambda \cdot ((u + v) + W) = (\lambda u + \lambda v) + W = \\ &= (\lambda u + W) + (\lambda v + W) = \lambda(u + W) + \lambda(v + W). \end{aligned}$$

Оставшиеся аксиомы 6-8 проверяются аналогично. ■

Определение 3. *Коразмерностью* подпространства W линейного пространства V (обозначается через $\text{codim } W$) называется $\dim V/W$.

Теорема 1. $\text{codim } W = \dim V - \dim W$.

Доказательство. Пусть $\dim V = n$, $\dim W = k$ и e_1, \dots, e_k — базис в W . Дополним его до базиса $e_1, \dots, e_k, e_{k+1}, \dots, e_n$ в V . Докажем, что классы $e_{k+1} + W, \dots, e_n + W$ образуют базис в V/W . Вначале покажем, что они линейно независимы. Пусть $\lambda_{k+1}(e_{k+1} + W) + \dots + \lambda_n(e_n + W) = 0 + W$.

Тогда $(\lambda_{k+1}e_{k+1} + \dots + \lambda_n e_n) + W = \mathbf{0} + W$, т.е. $v := \lambda_{k+1}e_{k+1} + \dots + \lambda_n e_n \in W$. Т.к. e_1, \dots, e_k — базис в W , то можем записать $v = \lambda_1 e_1 + \dots + \lambda_k e_k$. Тогда получаем

$$\lambda_1 e_1 + \dots + \lambda_k e_k - \lambda_{k+1} e_{k+1} - \dots - \lambda_n e_n = \mathbf{0}.$$

Т. к. e_1, \dots, e_n — базис в V , то $\lambda_1 = \dots = \lambda_{k+1} = \dots = \lambda_n = 0$. Значит, классы $e_{k+1} + W, \dots, e_n + W$ линейно независимы. Осталось доказать, что эти классы порождают всё пространство. Возьмём произвольный вектор $v + W \in V/W$. Разложим вектор v по базису в V : $v = \lambda_1 e_1 + \dots + \lambda_k e_k + \lambda_{k+1} e_{k+1} + \dots + \lambda_n e_n$. Тогда

$$\begin{aligned} v + W &= (\lambda_{k+1}e_{k+1} + \dots + \lambda_n e_n) + \overbrace{(\lambda_1 e_1 + \dots + \lambda_k e_k)}^{\in W} + W = \\ &= (\lambda_{k+1}e_{k+1} + \dots + \lambda_n e_n) + W = \lambda_{k+1}(e_{k+1} + W) + \dots + \lambda_n(e_n + W). \end{aligned}$$

Итак, в базисе V/W ровно $n - k$ векторов.

Предложение 3. Совокупность всех решений произвольной совместной системы линейных уравнений есть сумма какого-либо одного её решения и подпространства решений однородной системы линейных уравнений с той же матрицей коэффициентов.

Доказательство. Пусть $u = (u_1, u_2, \dots, u_n)^t \in \mathbb{F}^n$ — частное решение неоднородной СЛУ

[illegible]

Пусть также $v = (v_1, v_2, \dots, v_n)^t \in \mathbb{F}^n$ — произвольное решение ассоциированной однородной системы

[illegible]

Тогда сумма $u + v$ является решением первой системы. Действительно,

$$\begin{aligned} a_{i1}(u_1 + v_1) + a_{i2}(u_2 + v_2) + \dots + a_{in}(u_n + v_n) &= \\ &= \underbrace{a_{i1}u_1 + a_{i2}u_2 + \dots + a_{in}u_n}_{=b_i} + \underbrace{a_{i1}v_1 + a_{i2}v_2 + \dots + a_{in}v_n}_{=0} = b_i. \end{aligned}$$

Обратно, если u' — произвольное решение неоднородной СЛУ, то $v = u' - u$ является решением ассоциированной однородной системы (проверяется так же). ■

Таким образом, решения всех СЛУ $Ax = b$ (где x и b — векторы) с фиксированной матрицей A есть факторпространство линейного пространства \mathbb{F}^n по подпространству решений ассоциированной однородной СЛУ $Ax = 0$.

7. Линейные функции на векторном пространстве, их ядра. Изменение коэффициентов линейной формы при замене базиса. Сопряжённое пространство V^* , дуальный базис. Канонический изоморфизм $V \simeq V^{**}$

Сначала см. вопрос 8 и первое определение в вопросе 11.

Определение 1. Линейное отображение $f : V \rightarrow \mathbb{F}$ из пространства V над полем \mathbb{F} в поле \mathbb{F} (одномерное векторное пространство) называется *линейной функцией* (функционалом).

Определение 2. $V^* := \text{Hom}_{\mathbb{F}}(V, \mathbb{F})$ — двойственное (сопряжённое, дуальное) пространство к V .

Пусть e_1, \dots, e_n — базис в V . Значение линейной функции $\xi \in V^*$ на любом векторе $x = x^i e_i \in V$ определяется её значениями на базисных векторах, т. к. $\xi(x) = \xi(x^i e_i) = x^i \xi(e_i)$. Определим линейные функции $\varepsilon^1, \dots, \varepsilon^n \in V^*$ по правилу $\varepsilon^i(e_j) = \delta_j^i$. Тогда для любого вектора $x = x^j e_j$ мы имеем

$$\varepsilon^i(x) = \varepsilon^i(x^j e_j) = x^j \varepsilon^i(e_j) = x^j \delta_j^i = x^i.$$

Определение 3. В связи с этим функции $\varepsilon^1, \dots, \varepsilon^n$ часто называют *координатными функциями*.

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

Доказательство. Линейная независимость: пусть $x_1 \varepsilon^1 + \dots + x_n \varepsilon^n = \mathbf{0}$. Это равенство означает, что линейная функция $\xi : x_i \varepsilon^i$ равна нулю на любом векторе из V . Вычислим её на векторе e_j :

$$\mathbf{0} = \xi(e_j) = x_i \varepsilon^i(e_j) = x_i \delta_j^i = x_j.$$

Итак, $x_1 = \dots = x_n = 0$, а значит, $\varepsilon^1, \dots, \varepsilon^n$ линейно независимы.

Теперь проверим, что $\varepsilon^1, \dots, \varepsilon^n$ порождают всё пространство V^* . Мы утверждаем, что любая линейная функция ξ представляется в виде линейной комбинации $\xi = \xi_i \varepsilon^i$, где $\xi_i = \xi(e_i)$. Действительно, для любого вектора $x = x^j e_j \in V$ мы имеем

$$\xi_i \varepsilon^i(x) = \xi_i x^i = \xi(e_i) x^i = \xi(x^i e_i) = \xi(x).$$

Таким образом, $\varepsilon^1, \dots, \varepsilon^n$ — базис в V^* . ■

Определение 4. Базис $\varepsilon^1, \dots, \varepsilon^n$ пространства V^* называется *двойственным* (сопряжённым, дуальным) базисом к e_1, \dots, e_n .

Следствие 1. $\dim V = \dim V^* \Rightarrow V \simeq V^*$.

Пусть теперь $e_{1'}, \dots, e_{n'}$ — другой базис пространства V и $C = (c_{i'}^i)$ — матрица перехода от старого базиса к новому, т. е. $e_{i'} = c_{i'}^i e_i$. Рассмотрим двойственные базисы $\varepsilon^1, \dots, \varepsilon^n$ и $\varepsilon^{1'}, \dots, \varepsilon^{n'}$.

Предложение 2. Матрица перехода от $\varepsilon^1, \dots, \varepsilon^n$ к $\varepsilon^{1'}, \dots, \varepsilon^{n'}$ есть $(C^{-1})^t$.

Доказательство. Для любого вектора $x = x^i e_i = x^{i'} e_{i'}$ мы имеем $\varepsilon^i(x) = x^i = c_{i'}^i x^{i'} = c_{i'}^i \varepsilon^{i'}(x)$. Следовательно, $\varepsilon^i = c_{i'}^i \varepsilon^{i'}$, что эквивалентно матричному соотношению

$$\begin{pmatrix} \varepsilon^1 \\ \vdots \\ \varepsilon^n \end{pmatrix} = C \cdot \begin{pmatrix} \varepsilon^{1'} \\ \vdots \\ \varepsilon^{n'} \end{pmatrix}$$

или $(\varepsilon^{1'}, \dots, \varepsilon^{n'}) = (\varepsilon^1, \dots, \varepsilon^n) \cdot (C^{-1})^t$. ■

Для построения изоморфизма между пространствами V и V^* нам необходимо выбрать базис в V (и двойственный ему базис в V^*). Разные базисы дают разные изоморфизмы. Для бесконечномерных пространств ситуация иная: пространства V и V^* **никогда не изоморфны**, пространство V^* всегда «больше». Проиллюстрируем это на примере. Обозначим через \mathbb{F}^∞ пространство финитных последовательностей из элементов поля \mathbb{F} , а через $\widehat{\mathbb{F}}^\infty$ — пространство всех бесконечных последовательностей элементов поля \mathbb{F} .

Предложение 3. $(\mathbb{F}^\infty)^* \simeq \widehat{\mathbb{F}}^\infty$.

Доказательство. Возьмём в пространстве \mathbb{F}^∞ базис $\{e_i\}_{i=1}^\infty$, $e_i = (\overbrace{0, \dots, 0}^{i-1}, 1, 0, \dots)$. Рассмотрим отображение $\mathcal{A} : (\mathbb{F}^\infty)^* \rightarrow \widehat{\mathbb{F}}^\infty$, $f \mapsto (f(e_1), f(e_2), \dots)$, которое линейной функции $f \in (\mathbb{F}^\infty)^*$ ставит в соответствие последовательность её значений на базисных векторах. Это отображение очевидно линейно. Кроме того, оно биективно: обратное отображение задаётся формулой $\mathcal{A}^{-1}(x_1, x_2, \dots) = f \in (\mathbb{F}^\infty)^*$, где $f(e_i) = x_i$. Т.к. любой элемент $y \in \mathbb{F}^\infty$ есть (конечная) линейная комбинация элементов e_i , то линейная функция f однозначно восстанавливается по её значениям на базисных векторах ■

Предложение 4. $\mathbb{Z}_2^\infty \not\cong \widehat{\mathbb{Z}}_2^\infty$.

Доказательство. \mathbb{Z}_2^∞ можно отождествить с множеством рациональных чисел на отрезке $[0; 1]$ в двоичной записи, а $\widehat{\mathbb{Z}}_2^\infty$ — с множеством вещественных чисел на $[0; 1]$ в двоичной записи. Поэтому биекции между этими множествами быть не может. ■

Следствие 2. $\mathbb{Z}_2^\infty \not\cong (\mathbb{Z}_2^\infty)^*$.

Определение 5. $V^{**} := (V^*)^*$ — второе сопряжённое пространство.

Теорема 1. Пусть V — конечномерное линейное пространство. Отображение $\varphi : x \in V \mapsto \varphi_x \in V^{**}$, где $\varphi_x(\xi) := \xi(x)$ для $\xi \in V^*$, является изоморфизмом.

Доказательство. Из определения линейных функций следует, что $\varphi_{x+y} = \varphi_x + \varphi_y$ и $\varphi_{\lambda x} = \lambda \varphi_x$. Остаётся проверить, что отображение $x \mapsto \varphi_x$ биективно. Пусть e_1, \dots, e_n — базис пространства V и $\varepsilon^1, \dots, \varepsilon^n$ — сопряжённый базис пространства V^* . Тогда

$$\omega_i(\varepsilon^j) = \varepsilon^j(e_i) = \delta_j^i,$$

так что $\omega_1, \dots, \omega_n$ — базис пространства V^{**} , сопряжённый $\varepsilon^1, \dots, \varepsilon^n$. Отображение $x \mapsto \varphi_x$ переводит вектор с координатами x_1, \dots, x_n в базисе e_1, \dots, e_n пространства V в вектор с такими же координатами в базисе $\omega_1, \dots, \omega_n$ пространства V^{**} . Следовательно, оно биективно. ■

Отметим, что этот изоморфизм не зависит от выбора базисов в V и V^{**} , поэтому его часто называют *каноническим*. Т.к. векторы можно отождествить с линейными функциями на V^* , то можно определить и значение вектора $v \in V$ на линейной функции $\xi \in V^*$ следующим образом:

$$v(\xi) := \varphi_v(\xi) = \xi(v).$$

Ввиду такой двойственности, иногда для обозначения значения $v(\xi) = \xi(v)$ используется симметричная запись $\langle v \mid \xi \rangle$.

Следствие 1. Всякий базис пространства V^* сопряжён некоторому базису пространства V .

Задача 2 (Из Винберга). Доказать, что линейные функции f_1, \dots, f_n (где $n = \dim V$) составляют базис пространства V^* тогда и только тогда, когда не существует ненулевого вектора $x \in V$, для которого $f_1(x) = \dots = f_n(x) = 0$.

▷ ⇒. Пусть f_1, \dots, f_n — базис и нашёлся такой вектор $v \in V$, $v \neq 0$, для которого $f_1(v) = \dots = f_n(v) = 0$. Этот базис сопряжён некоторому базису e_1, \dots, e_n пространства V . А это значит, что вектор v ненулевой, но все координаты в этом базисе у него нулевые. Так не бывает.

⇐. Выберем базис $\varepsilon^1, \dots, \varepsilon^n$ в V^* . Тогда $f_i = a_j^i \varepsilon^j$. Рассмотрим систему уравнений (верхними индексами обозначены координаты, а не степени)

$$\begin{cases} f_1(x) = 0, \\ f_2(x) = 0, \\ \dots\dots\dots \\ f_n(x) = 0 \end{cases} \Leftrightarrow \begin{cases} a_1^1 x^1 + a_2^1 x^2 + \dots + a_n^1 x^n = 0, \\ a_1^2 x^1 + a_2^2 x^2 + \dots + a_n^2 x^n = 0, \\ \dots\dots\dots \\ a_1^n x^1 + a_2^n x^2 + \dots + a_n^n x^n = 0, \end{cases}$$

По условию не существует ненулевого вектора x такого, что $f_1(x) = f_2(x) = \dots = f_n(x) = 0$, поэтому выписанная система не имеет решений кроме нулевого. Поэтому она определена, а значит, по правилу Крамера матрица коэффициентов невырождена, отсюда следует линейная независимость строк. А полноту можно не доказывать, потому что количество векторов правильное. ◀

Это определение и предложение после него рассказывать не нужно, оно нам пригодится в вопросе 37.

Определение 6. Пусть $\mathcal{A} : V \rightarrow W$ — линейное отображение. Отображение $\mathcal{A}^* : W^* \rightarrow V^*$, заданное формулой

$$(\mathcal{A}^*\xi)(v) := \xi(\mathcal{A}v) \quad \forall \xi \in W^*, v \in V,$$

называется *сопряжённым к \mathcal{A}* .

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

Предложение 5. Пусть e_1, \dots, e_m и f_1, \dots, f_n — базисы пространств V и W соответственно. Тогда матрица отображения $\mathcal{A} : V \rightarrow W$ в этих базисах и матрица сопряжённого отображения $\mathcal{A}^* : W^* \rightarrow V^*$ в двойственных базисах $\varphi^1, \dots, \varphi^n$ и $\varepsilon^1, \dots, \varepsilon^m$ получаются друг из друга транспонированием.

Доказательство. Пусть $A = (a_j^i)$ — матрица отображения \mathcal{A} в базисах e_1, \dots, e_m и f_1, \dots, f_n . Тогда для любого вектора $x = x^i e_i$ мы имеем

$$(\mathcal{A}^*\varphi^i)(x) = \varphi^i(\mathcal{A}x) = \varphi^i(a_k^j x^k f_j) = a_k^j x^k \varphi^i(f_j) = a_k^j x^k \delta_j^i = a_k^i x^k = a_k^i \varepsilon^k(x).$$

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

8. ЛИНЕЙНЫЕ ОТОБРАЖЕНИЯ И ОПЕРАТОРЫ. ЯДРО И ОБРАЗ, СВЯЗЬ ИХ РАЗМЕРНОСТЕЙ. КРИТЕРИЙ ИНЪЕКТИВНОСТИ

Определение 1. Пусть V и W — линейные пространства над полем \mathbb{F} . Отображение $\mathcal{A} : V \rightarrow W$ называется *линейным*, если $\forall u, v \in V, \forall \lambda \in \mathbb{F}$ выполнено $\mathcal{A}(u + v) = \mathcal{A}u + \mathcal{A}v$ и $\mathcal{A}(\lambda v) = \lambda \mathcal{A}v$.

Определение 2. Линейное отображение $\mathcal{A} : V \rightarrow V$ из пространства V в себя называется *линейным оператором*.

Биективные линейные операторы образуют группу, называемую *полной линейной группой* линейного пространства V и обозначаемую через $\text{GL}(V)$.

Определение 3. Пусть $\mathcal{A} : V \rightarrow W$ — линейное отображение. *Ядром \mathcal{A}* называется множество $\text{Ker } \mathcal{A} := \{v \in V : \mathcal{A}v = \mathbf{0}\}$. *Образом \mathcal{A}* называется множество $\text{Im } \mathcal{A} := \{\mathcal{A}v : v \in V\}$.

Предложение 1. Пусть $\mathcal{A} : V \rightarrow W$ — линейное отображение. Тогда $\text{Ker } \mathcal{A}$ — подпространство в V , а $\text{Im } \mathcal{A}$ — подпространство в W .

Доказательство. Пусть $u, v \in \text{Ker } \mathcal{A}$. Т.е. $\mathcal{A}u = \mathcal{A}v = \mathbf{0}$. Тогда $\mathcal{A}(u + v) = \mathcal{A}u + \mathcal{A}v = \mathbf{0}$ и $\mathcal{A}(\lambda u) = \lambda \mathcal{A}u = \mathbf{0}$. Следовательно, $u + v \in \text{Ker } \mathcal{A}$ и $\lambda u \in \text{Ker } \mathcal{A} \quad \forall \lambda \in \mathbb{F}$, а значит, $\text{Ker } \mathcal{A}$ — подпространство в V .

Пусть теперь $x, y \in \text{Im } \mathcal{A}$, т.е. $\exists u, v \in V : \mathcal{A}u = x, \mathcal{A}v = y$. Тогда $\mathcal{A}(u + v) = x + y$ и $\mathcal{A}(\lambda u) = \lambda x$. Следовательно, $x + y \in \text{Im } \mathcal{A}$ и $\lambda x \in \text{Im } \mathcal{A} \quad \forall \lambda \in \mathbb{F}$, а значит, $\text{Im } \mathcal{A}$ — подпространство в W . ■

Лемма 1 (Критерий инъективности). Линейное отображение $\mathcal{A} : V \rightarrow W$ инъективно тогда и только тогда, когда $\text{Ker } \mathcal{A} = \{\mathbf{0}\}$.

Доказательство. \Rightarrow . Мы знаем, что $\mathcal{A}\mathbf{0} = \mathbf{0}$, а т.к. \mathcal{A} инъективно, то $\mathbf{0}$ — единственный вектор из V , переходящий в $\mathbf{0}$, отсюда $\text{Ker } \mathcal{A} = \{\mathbf{0}\}$.

\Leftarrow . Пусть $\mathcal{A}u = \mathcal{A}v \Rightarrow \mathcal{A}(u - v) = \mathbf{0}$, значит, $u - v \in \text{Ker } \mathcal{A} = \{\mathbf{0}\}$, отсюда $u = v$. ■

Теорема 1. Пусть $\mathcal{A} : V \rightarrow W$ — линейное отображение. Тогда соответствие $v + \text{Ker } \mathcal{A} \mapsto \mathcal{A}v$ задаёт изоморфизм между факторпространством $V / \text{Ker } \mathcal{A}$ и подпространством $\text{Im } \mathcal{A}$.

Доказательство. Сначала проверим, что $v + \text{Ker } \mathcal{A} \mapsto \mathcal{A}v$ действительно корректно определяет отображение $\tilde{\mathcal{A}} : V / \text{Ker } \mathcal{A} \rightarrow \text{Im } \mathcal{A}$. Для этого нужно проверить, что если $u + \text{Ker } \mathcal{A} = v + \text{Ker } \mathcal{A}$, то $\mathcal{A}u = \mathcal{A}v$. Из равенства классов смежности следует $u - v \in \text{Ker } \mathcal{A}$, а отсюда $\mathcal{A}(u - v) = \mathbf{0}$, т. е. $\mathcal{A}u = \mathcal{A}v$. Итак, отображение $\tilde{\mathcal{A}}$ определено корректно.

Линейность и сюръективность $\tilde{\mathcal{A}}$ очевидны. Инъективность проверяется по критерию:

$$\text{Ker } \tilde{\mathcal{A}} = \{(v + \text{Ker } \mathcal{A}) \in V / \text{Ker } \mathcal{A} : \tilde{\mathcal{A}}(v + \text{Ker } \mathcal{A}) = \mathcal{A}v = \mathbf{0}\} = \text{Ker } \mathcal{A} = \mathbf{0} + \text{Ker } \mathcal{A}.$$

Итак, $\tilde{\mathcal{A}}$ задаёт изоморфизм $V / \text{Ker } \mathcal{A} \simeq \text{Im } \mathcal{A}$. ■

Следствие 1. Для всякого линейного отображения $\mathcal{A} : V \rightarrow W$ мы имеем

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

Предложение 2. Если в каких-то базисах пространств V и W линейное отображение $\mathcal{A} : V \rightarrow W$ имеет матрицу A , то

$$\dim \text{Im } \mathcal{A} = \text{rk } A.$$

Доказательство. Очевидно, что $\text{Im } \mathcal{A}$ есть линейная оболочка образов базисных векторов e_1, \dots, e_n пространства V и, значит, $\dim \text{Im } \mathcal{A}$ есть ранг системы векторов $\mathcal{A}e_1, \dots, \mathcal{A}e_n$. Но в столбцах матрицы A как раз и записаны координаты этих векторов в каком-то базисе пространства W . Следовательно, ранг этой системы векторов равен рангу матрицы A . ■

9. ЗАДАНИЕ ЛИНЕЙНЫХ ОТОБРАЖЕНИЙ (ОПЕРАТОРОВ) МАТРИЦАМИ. ИЗМЕНЕНИЕ МАТРИЦЫ ЛИНЕЙНОГО ОТОБРАЖЕНИЯ ПРИ ПЕРЕХОДЕ К ДРУГИМ БАЗИСАМ. НАХОЖДЕНИЕ ЯДРА И ОБРАЗА ПРИ ПОМОЩИ МАТРИЦЫ

Пусть $\mathcal{A} : V \rightarrow W$ — линейное отображение, e_1, \dots, e_m — базис в V , а f_1, \dots, f_n — базис в W .

Определение 1. Матрицей линейного отображения $\mathcal{A} : V \rightarrow W$ по отношению к базисам

e_1, \dots, e_m и f_1, \dots, f_n называется матрица $A = \begin{pmatrix} a_1^1 & \cdots & a_m^1 \\ \vdots & \ddots & \vdots \\ a_1^n & \cdots & a_m^n \end{pmatrix}$ размера $n \times m$, в которой i -ый

столбец состоит из координат вектора $\mathcal{A}e_i$ в базисе f_1, \dots, f_n : $\mathcal{A}e_i = a_i^j f_j$.

Предложение 1. Пусть $x = x^j e_j$ — произвольный вектор из V , а $y = y^i f_i$ — его образ в W , т. е. $y = \mathcal{A}x$. Тогда $y^i = a_j^i x^j$. Или, в матричном виде, $y = Ax$.

Доказательство. Действительно, $y^i f_i = y = \mathcal{A}x = \mathcal{A}(x^j e_j) = x^j \mathcal{A}e_j = x^j a_j^i f_i$. Т. к. $\{f_i\}_{i=1}^n$ — базис, отсюда следует, что $y^i = a_j^i x^j$. ■

Предложение 2. Пусть $\dim V = m$, $\dim W = n$. Тогда $\text{Hom}_{\mathbb{F}}(V, W) \simeq \text{Mat}_{n \times m}(\mathbb{F})$.

Доказательство. Выберем базисы e_1, \dots, e_m и f_1, \dots, f_n в V и W соответственно. Определим отображение $\text{Hom}_{\mathbb{F}}(V, W) \rightarrow \text{Mat}_{n \times m}(\mathbb{F})$, которое сопоставляет линейному отображению его матрицу в выбранных базисах. Непосредственно проверяется, что это отображение линейно. Кроме того, оно биективно: обратное отображение сопоставляет матрице $A = (a_j^i)$ линейного отображения,

определяется в координатах формулой из предыдущего предложения. Следовательно, такое отображение $\text{Hom}_{\mathbb{F}}(V, W) \rightarrow \text{Mat}_{n \times m}(\mathbb{F})$ является изоморфизмом. ■

Теорема 1 (Закон изменения матрицы линейного отображения). Имеет место соотношение $A' = D^{-1}AC$, где A — матрица линейного отображения $\mathcal{A} : V \rightarrow W$ по отношению к базисам e_1, \dots, e_m и f_1, \dots, f_n ; A' — матрица линейного отображения \mathcal{A} по отношению к базисам $e_{1'}, \dots, e_{m'}$ и $f_{1'}, \dots, f_{n'}$; $C = C_{e \rightarrow e'}$ — матрица перехода от e_1, \dots, e_m к $e_{1'}, \dots, e_{m'}$; $D = D_{f \rightarrow f'}$ — матрица перехода от f_1, \dots, f_n к $f_{1'}, \dots, f_{n'}$.

Доказательство. Пусть $C = (c_{i'}^i)$, $A = (a_i^j)$, тогда

$$\mathcal{A}e_{i'} = \mathcal{A}(c_{i'}^i e_i) = c_{i'}^i \mathcal{A}e_i = c_{i'}^i a_i^j f_j.$$

С другой стороны, если $A' = (a_{i'}^{j'})$ и $D = (d_{j'}^j)$, то

$$\mathcal{A}e_{i'} = a_{i'}^{j'} f_{j'} = a_{i'}^{j'} d_{j'}^j f_j.$$

Сравнивая два последних соотношения с учётом того, что $\{f_j\}_{j=1}^n$ — базис, получаем $a_i^j c_{i'}^i = d_{j'}^j a_{i'}^{j'}$. В матричном виде это эквивалентно $AC = DA' \Rightarrow A' = D^{-1}AC$. ■

Поиск ядра и образа линейного оператора по его матрице. Пусть имеем матрицу A оператора \mathcal{A} в каком-то базисе. Приведя её к ступенчатому виду, сможем найти базис системы столбцов матрицы A (его будут составлять столбцы, в которых есть лидеры). Вспомним, что по столбцам A написаны образы базисных векторов при отображении \mathcal{A} , а мы нашли базис этой системы. Это значит, что найденные нами столбцы есть базис $\text{Im } \mathcal{A}$.

$\text{Ker } \mathcal{A}$ — это просто пространство решений СЛУ $Ax = 0$. Чтобы найти базис ядра, нам нужно просто найти её ФСР.

10. ЛИНЕЙНЫЕ ОПЕРАТОРЫ. ИЗМЕНЕНИЕ МАТРИЦЫ ЛИНЕЙНОГО ОПЕРАТОРА ПРИ ПЕРЕХОДЕ К НОВОМУ БАЗИСУ. ПОДОБНЫЕ МАТРИЦЫ

Определение 1. Линейное отображение $\mathcal{A} : V \rightarrow V$ пространства V в себя называется *линейным оператором*.

Базис в V выбирается только один, поэтому изменение матрицы линейного оператора при переходе к новому базису выглядит так же, как в теореме 1 из 9 вопроса при $C = D$.

Определение 2. Матрицы A и B называются *подобными* тогда и только тогда, когда существует невырожденная матрица C такая, что

$$B = C^{-1}AC.$$

Предложение 1. Если A и B подобны, то $\det A = \det B$, $\text{rk } A = \text{rk } B$, $\text{tr } A = \text{tr } B$.

Доказательство. В обозначениях из определения подобия $\det B = \det(C^{-1}AC) = \det(C^{-1}C) \cdot \det A = \det A$. Второе утверждение следует из того, что ранг не меняется при домножении на невырожденную матрицу.

Непосредственным вычислением доказывается, что $\text{tr}(AC) = \text{tr}(CA)$ (теорема из первого семестра). Как следствие, $\text{tr}(C^{-1}AC) = \text{tr } A$. ■

Это значит, что определитель, ранг и след матрицы линейного оператора не меняются при любой замене базиса, таким образом, можно говорить об *определителе, ранге и следе линейного оператора*.

11. ВЕКТОРНОЕ ПРОСТРАНСТВО ЛИНЕЙНЫХ ОТОБРАЖЕНИЙ. АЛГЕБРА ЛИНЕЙНЫХ ОПЕРАТОРОВ. ИЗОМОРФИЗМ АЛГЕБРЫ МАТРИЦ И АЛГЕБРЫ ЛИНЕЙНЫХ ОПЕРАТОРОВ

Определение 1. Множество всех линейных отображений $\mathcal{A} : V \rightarrow W$ с операциями сложения и умножения на скаляры

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

является линейным пространством над полем \mathbb{F} . Оно называется *пространством линейных отображений* из V в W и обозначается $\text{Hom}_{\mathbb{F}}(V, W)$. Мы также будем пользоваться обозначением $\text{End}(V) := \text{Hom}_{\mathbb{F}}(V, V)$.

Напомним определение алгебры.

Определение 2. Алгеброй над полем \mathbb{F} называется множество \mathfrak{A} с операциями сложения, умножения и умножения на элементы поля \mathbb{F} , обладающими следующими свойствами:

1. относительно сложения и умножения на элементы поля \mathfrak{A} есть векторное пространство;
2. относительно сложения и умножения \mathfrak{A} есть кольцо;
3. $(\lambda a)b = a(\lambda b) = \lambda(ab) \quad \forall \lambda \in \mathbb{F}, \forall a, b \in \mathfrak{A}$.

На множестве $\text{End}(V)$ всех линейных операторов векторного пространства V над полем \mathbb{F} можно также определить произведение (композицию) линейных операторов $(\mathcal{A}\mathcal{B})(v) := \mathcal{A}(\mathcal{B}v)$. Очевидно, оно тоже линейно. При этом выполнены все аксиомы алгебры.

Алгебра $\text{End}(V)$ является ассоциативной и обладает единицей. Единицей этой алгебры является тождественный линейный оператор, который мы будем обозначать через id .

Определение 3. Две алгебры \mathfrak{A} и \mathfrak{A}' называются *изоморфными*, если существует изоморфизм колец из первой алгебры во вторую, являющийся одновременно и изоморфизмом векторных пространств. Пишем $\mathfrak{A} \simeq \mathfrak{A}'$.

Теорема 1. Для любого n -мерного линейного пространства V над полем \mathbb{F} выполняется $\text{End}(V) \simeq \text{Mat}_{n \times n}(\mathbb{F})$.

Доказательство. Зафиксируем базис e_1, \dots, e_n в пространстве V и рассмотрим отображение $\Phi : \text{End}(V) \rightarrow \text{Mat}_{n \times n}(\mathbb{F})$, сопоставляющее каждому линейному оператору $\mathcal{A} : V \rightarrow V$ его матрицу A в выбранном базисе.

Т.к. любой линейный оператор однозначно определяется своей матрицей (в фиксированном базисе) формулой из предложения 1 в вопросе 9, а любая квадратная матрица порядка n является матрицей порядка n является матрицей некоторого линейного оператора (столбцы этой матрицы можно считать координатами образов базисных векторов из выбранного базиса), отображение Φ биективно.

Далее пусть $\Phi(\mathcal{A}) = A = (a_i^j)$, $\Phi(\mathcal{B}) = (b_j^k)$, $\Phi(\mathcal{C}) = C = (c_i^k)$, где $\mathcal{C} = \mathcal{B}\mathcal{A}$. Тогда с одной стороны,

$$\mathcal{C}e_i = \mathcal{B}(\mathcal{A}e_i) = \mathcal{B}(a_i^j e_j) = a_i^j \mathcal{B}e_j = a_i^j b_j^k e_k,$$

а с другой —

$$\mathcal{C}e_i = c_i^k e_k.$$

Из единственности разложения вектора по базису получаем $c_i^k = b_j^k a_i^j$; в матричной записи $C = BA$. Таким образом, $\Phi(\mathcal{B}\mathcal{A}) = \Phi(\mathcal{B}) \cdot \Phi(\mathcal{A})$. Аналогично проверяются остальные аксиомы: $\Phi(\lambda \mathcal{A}) = \lambda \Phi(\mathcal{A})$ и $\Phi(\mathcal{A} + \mathcal{B}) = \Phi(\mathcal{A}) + \Phi(\mathcal{B})$. ■

Следствие 1. $\dim \text{End}(V) = (\dim V)^2$.

12. ИНВАРИАНТНЫЕ ПОДПРОСТРАНСТВА ЛИНЕЙНОГО ОПЕРАТОРА. ОГРАНИЧЕНИЕ ЛИНЕЙНОГО ОПЕРАТОРА НА ИНВАРИАНТНОЕ ПОДПРОСТРАНСТВО. ВИД МАТРИЦЫ ЛИНЕЙНОГО ОПЕРАТОРА ПРИ НАЛИЧИИ ИНВАРИАНТНЫХ ПОДПРОСТРАНСТВ

Определение 1. Подпространство $W \subseteq V$ называется *инвариантным* относительно оператора $A : V \rightarrow V$, если $A(W) \subseteq W$.

Пример 1. $\text{Ker } A$ и $\text{Im } A$ являются инвариантными подпространствами.

Пусть $W \subseteq V$ — инвариантное подпространство для оператора $A : V \rightarrow V$. Выберем базис e_1, \dots, e_k в W и дополним его до базиса $e_1, \dots, e_k, e_{k+1}, \dots, e_n$ в V . Пусть $A = (a_j^i)$ — матрица оператора A в этом базисе. Тогда $Ae_j = a_j^1 e_1 + \dots + a_j^k e_k$ при $j = 1, \dots, k$. Это означает, что матрица A имеет вид $A = \left(\begin{array}{c|c} * & * \\ 0 & * \end{array} \right)$, где в левом нижнем углу стоит матрица размера $(n - k) \times k$ из нулей. Аналогично, если имеет место разложение $V = W_1 \oplus W_2$ в прямую сумму инвариантных подпространств, $A(W_1) \subseteq W_1$, $A(W_2) \subseteq W_2$, то в подходящем базисе матрица оператора A будет иметь блочно-диагональный вид: $A = \left(\begin{array}{c|c} * & 0 \\ 0 & * \end{array} \right)$.

Определение 2. Пусть $W \subseteq V$ — инвариантное подпространство для оператора $A : V \rightarrow V$. Тогда оператор $\hat{A} : W \rightarrow W$, определённый равенством $\hat{A}w := Aw$ для $w \in W$, называется *ограничением* оператора A на подпространство W и обозначается как $A|_W$.

Определение 3. Линейный оператор $\tilde{A} : V/W \rightarrow V/W$, определённый на классах смежности по правилу $\tilde{A}(v + W) = Av + W$, называется *фактор-оператором*.

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

Предложение 1. Пусть $W \subseteq V$ — инвариантное подпространство для оператора $A : V \rightarrow V$. Пусть e_1, \dots, e_k — базис в W и $e_1, \dots, e_k, e_{k+1}, \dots, e_n$ — базис в V . Тогда матрица оператора A в этом базисе имеет вид $A = \left(\begin{array}{c|c} \hat{A} & * \\ 0 & \tilde{A} \end{array} \right)$, где \hat{A} — матрица ограничения $A|_W$ в базисе e_1, \dots, e_k , а \tilde{A} — матрица фактор-оператора в базисе $e_{k+1} + W, \dots, e_n + W$ факторпространства V/W .

Доказательство. Пусть $A = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \dots & a_n^n \end{pmatrix}$. Тогда $\hat{A} = \begin{pmatrix} a_1^1 & \dots & a_k^1 \\ \vdots & \ddots & \vdots \\ a_1^k & \dots & a_k^k \end{pmatrix}$ и $\tilde{A} = \begin{pmatrix} a_{k+1}^{k+1} & \dots & a_n^{k+1} \\ \vdots & \ddots & \vdots \\ a_{k+1}^n & \dots & a_n^n \end{pmatrix}$.

Разложим вектор по базису: $v = v^1 e_1 + \dots + v^k e_k + \dots + v^{k+1} e_{k+1} + \dots + v^n e_n$. Отметим, что в силу инвариантности подпространства W имеет место $A \cdot (v^1, \dots, v^k, 0, \dots, 0)^t = (\underbrace{*, \dots, *}_k, 0, \dots, 0)^t$. Итак,

$$\begin{aligned} Av = A \cdot v &= \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \dots & a_n^n \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_k \\ v_{k+1} \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \dots & a_n^n \end{pmatrix} \cdot \left(\begin{pmatrix} v_1 \\ \vdots \\ v_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ v_{k+1} \\ \vdots \\ v_n \end{pmatrix} \right) = \\ &= \left(\begin{array}{c|c} \hat{A} & 0 \\ 0 & \tilde{A} \end{array} \right) \cdot v + \left(\begin{array}{c|c} 0 & * \\ 0 & \tilde{A} \end{array} \right) \cdot v = \left(\begin{array}{c|c} \hat{A} & * \\ 0 & \tilde{A} \end{array} \right) \cdot v. \end{aligned}$$

Отсюда и следует, что в выбранном базисе матрица A имеет указанный вид. ■

Примечание. В конспектах Т.Е. Панова написано, что это утверждение примерно очевидно и доказывать мы его не будем. Поэтому изложенное выше доказательство я проделал сам, оно может содержать ошибки.

Следствие 1. Если пространство V представлено в виде прямой суммы ненулевых инвариантных относительно преобразования \mathcal{A} подпространств $V = W_1 \oplus \dots \oplus W_k$, то существует базис, в котором матрица преобразования имеет блочно-диагональный вид $A = \text{diag}(A_1, \dots, A_k) = \begin{pmatrix} A_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A_k \end{pmatrix}$, где A_i — матрица ограничения $\mathcal{A}|_{W_i}$ преобразования \mathcal{A} на подпространство W_i , $i = 1, \dots, k$.

13. СОБСТВЕННЫЕ ВЕКТОРЫ И СОБСТВЕННЫЕ ЗНАЧЕНИЯ ЛИНЕЙНОГО ОПЕРАТОРА.

ЛИНЕЙНАЯ НЕЗАВИСИМОСТЬ СОБСТВЕННЫХ ВЕКТОРОВ ЛИНЕЙНОГО ОПЕРАТОРА, ОТВЕЧАЮЩИХ ПОПАРНО РАЗЛИЧНЫМ СОБСТВЕННЫМ ЗНАЧЕНИЯМ

Определение 1. Ненулевой вектор $v \in V$ называется *собственным* для оператора \mathcal{A} , если $\mathcal{A}v = \lambda v$ для некоторого $\lambda \in \mathbb{F}$. Число $\lambda \in \mathbb{F}$ называется *собственным значением*, если существует собственный вектор v , для которого $\mathcal{A}v = \lambda v$.

Предложение 1. Собственные векторы, отвечающие попарно различным собственным значениям $\lambda_1, \dots, \lambda_k$ оператора \mathcal{A} , линейно независимы.

Доказательство. Докажем утверждение теоремы индукцией по k . При $k = 1$ доказывать нечего. Пусть $k > 1$ и

$$\mu_1 v_1 + \dots + \mu_{k-1} v_{k-1} + \mu_k v_k = \mathbf{0} \quad (v_i \in V_{\lambda_i}, (\mu_1, \dots, \mu_k) \neq (0, \dots, 0))$$

Применяя оператор \mathcal{A} , получаем

$$\lambda_1 \mu_1 v_1 + \dots + \lambda_{k-1} \mu_{k-1} v_{k-1} + \lambda_k \mu_k v_k = \mathbf{0}.$$

Вычитая отсюда исходное неравенство, умноженное на λ_k , получаем

$$\mu_1 (\lambda_1 - \lambda_k) v_1 + \dots + \mu_{k-1} (\lambda_{k-1} - \lambda_k) v_{k-1} = \mathbf{0},$$

откуда в силу предположения индукции следует равенство нулю всех коэффициентов. Если найдётся $\mu_j \neq 0$, то $\lambda_j - \lambda_{k-1} = 0$, однако из условия собственные значения попарно различны. Противоречие. Значит, все μ_j равны нулю, и система v_1, \dots, v_k линейно независима. ■

Следствие 1. Собственные подпространства $V_{\lambda_1}, \dots, V_{\lambda_k}$, соответствующие попарно различным собственным значениям $\lambda_1, \dots, \lambda_k$ оператора \mathcal{A} , образуют прямую сумму $V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k}$.

14. ВЫЧИСЛЕНИЕ СОБСТВЕННЫХ ЗНАЧЕНИЙ И СОБСТВЕННЫХ ВЕКТОРОВ С ПОМОЩЬЮ МАТРИЦЫ. ХАРАКТЕРИСТИЧЕСКИЙ МНОГОЧЛЕН

Определение 1. Многочлен $\chi_{\mathcal{A}}(t) := \det(\mathcal{A} - t \cdot \text{id})$ называется *характеристическим многочленом* оператора \mathcal{A} .

Так как характеристический многочлен определяется как детерминант оператора, его можно вычислять как определитель матрицы этого оператора в любом базисе:

$$\chi_{\mathcal{A}}(t) = \det(A - t \cdot E) = \det \begin{pmatrix} a_1^1 - t & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 - t & \dots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \dots & a_n^n - t \end{pmatrix}.$$

Из этой формулы ясно, что $\deg \chi_{\mathcal{A}}(t) = \dim V$, кроме того, при $\chi_{\mathcal{A}}(t) = p_n t^n + p_{n-1} t^{n-1} + \dots + p_0$, имеем $p_n = (-1)^n$, $p_{n-1} = (-1)^{n-1} \operatorname{tr} \mathcal{A}$, $p_0 = \det \mathcal{A}$.

Предложение 1. Собственные значения оператора \mathcal{A} — это в точности корни его характеристического многочлена.

Доказательство. Если λ — собственное значение, то оператор $\mathcal{A} - \lambda \cdot \operatorname{id}$ вырожден, т. е. $\det(\mathcal{A} - \lambda \cdot \operatorname{id}) = 0$, а значит, λ — корень многочлена $\chi_{\mathcal{A}}(t)$. Обратно, если $\chi_{\mathcal{A}}(\lambda) = 0$, то $\det(\mathcal{A} - \lambda \cdot \operatorname{id}) = 0$. Поэтому $\operatorname{Ker}(\mathcal{A} - \lambda \cdot \operatorname{id}) \neq \{0\}$, а значит, λ — собственное значение. ■

Поиск собственных векторов. Сначала находим корни характеристического многочлена, т. е. собственные значения. Теперь найдём подпространство V_λ для собственного значения λ . Иными словами, хотим найти все такие векторы, что $\mathcal{A}v = \lambda v$. Это равносильно решению системы уравнений с матрицей $(A - \lambda E)$. Она и задаёт искомое подпространство.

15. СОБСТВЕННЫЕ ПОДПРОСТРАНСТВА. НЕРАВЕНСТВО МЕЖДУ РАЗМЕРНОСТЬЮ СОБСТВЕННОГО ПОДПРОСТРАНСТВА И КРАТНОСТЬЮ КОРНЯ ХАРАКТЕРИСТИЧЕСКОГО МНОГОЧЛЕНА

Предложение 1. Все собственные векторы, отвечающие собственным значениям λ , и вектор 0 образуют подпространство, которое совпадает с $\operatorname{Ker}(\mathcal{A} - \lambda \cdot \operatorname{id})$.

Доказательство. По определению ядра, равенство $\mathcal{A}v = \lambda v$ имеет место тогда и только тогда, когда $v \in \operatorname{Ker}(\mathcal{A} - \lambda \cdot \operatorname{id})$. ■

Определение 1. Пусть λ — собственное значение для оператора \mathcal{A} . Подпространство $V_\lambda := \operatorname{Ker}(\mathcal{A} - \lambda \cdot \operatorname{id})$ называется *собственным подпространством*, соответствующим λ .

Предложение 2. Каждое собственное подпространство V_λ оператора \mathcal{A} является инвариантным относительно него.

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

Предложение 3. Размерность собственного подпространства V_λ не превосходит кратности λ как корня характеристического многочлена.

Доказательство. Пусть $\dim V_\lambda = k$. Выберем базис e_1, \dots, e_k в подпространстве V_λ и дополним его до базиса в V . Т. к. $\mathcal{A}e_i = \lambda e_i$ при $i = 1, \dots, k$, по предложению 1 из вопроса 12 матрица

оператора \mathcal{A} в выбранном базисе имеет вид $A = \left(\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & * \\ 0 & & \lambda & \\ \hline & 0 & & \tilde{A} \end{array} \right)$. Тогда

$$\chi_{\mathcal{A}}(t) = \det(A - t \cdot E) = (\lambda - t)^k \det(\tilde{A} - t \cdot E) = (\lambda - t)^k \chi_{\tilde{\mathcal{A}}}(t),$$

где $\tilde{\mathcal{A}}$ — фактор-оператор. Отсюда вытекает, что кратность корня λ не меньше $k = \dim V_\lambda$. ■

16. ДИАГОНАЛИЗИРУЕМОСТЬ (МАТРИЦЫ) ЛИНЕЙНОГО ОПЕРАТОРА. КРИТЕРИИ ДИАГОНАЛИЗИРУЕМОСТИ И ДОСТАТОЧНОЕ УСЛОВИЕ

Определение 1. Оператор \mathcal{A} называется *диагонализируемым*, если существует базис, в котором матрица этого оператора диагональна.

Базис, в котором матрица оператора диагональна, очевидно, состоит из собственных векторов. Поэтому оператор диагонализировать тогда и только тогда, когда для него существует базис из собственных векторов.

Теорема 1 (Критерий диагонализированности). Оператор \mathcal{A} в n -мерном пространстве V диагонализировать тогда и только тогда, когда его характеристический многочлен имеет в точности n корней (с учётом кратностей), и размерность каждого собственного подпространства V_λ равна кратности корня λ .

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

Предположим теперь, что многочлен $\chi_{\mathcal{A}}(t)$ имеет различные корни $\lambda_1, \dots, \lambda_k$, причём кратность корня λ_i равна r_i , $\sum_{i=1}^k r_i = n$ и $\dim V_{\lambda_i} = r_i$. Согласно следствию 1 из вопроса 13, пространства $V_{\lambda_1}, \dots, V_{\lambda_k}$ образуют прямую сумму, а по условию сумма их размерностей равна $n = \dim V$. Следовательно, $V = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k}$. Выбрав базис в каждом из подпространств $\{V_{\lambda_i}\}_{i=1}^k$ и взяв объединение этих базисов, мы получим базис пространства V , состоящий из собственных векторов. Итак, оператор \mathcal{A} диагонализировать. ■

Следствие 1 (Достаточное условие диагонализированности). Если характеристический многочлен $\chi_{\mathcal{A}}(t)$ имеет $n = \dim V$ различных корней, то оператор \mathcal{A} диагонализировать.

Примечание. Панов пишет, что набор собственных значений оператора \mathcal{A} часто называют его *спектром* (будет в курсе функционального анализа). Если все собственные значения имеют кратность 1 как корни характеристического многочлена, то говорят о *простом спектре*. Таким образом, операторы с простым спектром диагонализировать. Появление кратных корней является «особенностью», которая устраняется произвольно малым возмущением коэффициентов матрицы оператора. Поэтому над полем \mathbb{C} «почти все операторы диагонализировать».

К последней фразе лично я отношусь скептически, ведь (как заметил Рамиль Хакамов) любое действительное число можно произвольно малым возмущением сделать рациональным, хотя язык не повернётся сказать, что «почти все действительные числа рациональные».

UPD. Почитал, оказалось, что подмножество матриц, не являющихся диагонализировать над $\mathbb{C}^{n \times n}$, имеет нулевую меру Лебега. То есть, по факту это «почти никакие». Можно ещё сказать, что диагонализировать матрицы образуют всюду плотное подмножество в топологии Зарисского (замкнутыми называются множества нулей многочленов из некоторого выделенного множества), ведь дополнение к этому подмножеству лежит в множестве, в котором дискриминант характеристического уравнения обнуляется, т. е. на гиперповерхности. Над \mathbb{R} это не выполняется.

Если интересно почитать про геометрию дискриминанта, можно это сделать [ЗДЕСЬ](#).

Пример 1.

1. Оператор, заданный матрицей $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ в стандартном базисе в \mathbb{R}^2 не диагонализировать, т. к. его характеристический многочлен $t^2 + 1$ не имеет вещественных корней. Однако тот же оператор над \mathbb{C} диагонализировать: в базисе $f_1 = (1, i)$, $f_2 = (1, -i)$ его матрица $\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$ диагональна.
2. Оператор, заданный матрицей $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ не диагонализировать ни над каким полем по другой причине: его характеристический многочлен $(t - 1)^2$ имеет корень 1 кратности 2, но при этом размерность соответствующего собственного подпространства равна 1 (вектор $e_2 = (0, 1)$ не является собственным).

17. АННУЛИРУЮЩИЕ МНОГОЧЛЕНЫ ЛИНЕЙНОГО ОПЕРАТОРА (МАТРИЦЫ).
МИНИМАЛЬНЫЙ МНОГОЧЛЕН. КРИТЕРИЙ ДИАГОНАЛИЗИРУЕМОСТИ В ТЕРМИНАХ
МИНИМАЛЬНОГО МНОГОЧЛЕНА

Пусть $\mathcal{A} : V \rightarrow V$ — оператор. Каждому многочлену $P(t) = p_0 + p_1 t + \dots + p_n t^n \in \mathbb{F}[t]$ можно сопоставить оператор $P(\mathcal{A}) := p_0 \cdot \text{id} + p_1 \cdot \mathcal{A} + \dots + p_n \mathcal{A}^n$.

Определение 1. Такой оператор называется *многочленом от оператора \mathcal{A}* .

Определение 2. Многочлен $P(t)$ называют *аннулирующим* оператор \mathcal{A} , если $P(\mathcal{A}) = \mathcal{O}$.

Пример 1.

1. Многочлен $t - 1$ аннулирует оператор id .
2. Многочлен $t^2 - t$ аннулирует любой проектор (см. приложение про проекторы).

Предложение 1. У любого оператора существует ненулевой аннулирующий многочлен.

Доказательство. Пусть $\dim V = n$. Рассмотрим $n^2 + 1$ операторов: $\mathcal{A}^0 = \text{id}, \mathcal{A}^1 = \mathcal{A}, \mathcal{A}^2, \dots, \mathcal{A}^{n^2}$. Т. к. размерность пространства операторов равна n^2 , эти операторы линейно зависимы, т. е. существуют числа $p_0, p_1, p_2, \dots, p_{n^2}$, не все равные нулю, такие, что $p_0 \cdot \text{id} + p_1 \cdot \mathcal{A} + p_2 \cdot \mathcal{A}^2 + \dots + p_{n^2} \cdot \mathcal{A}^{n^2} = 0$. Тогда ненулевой многочлен $P(t) = p_0 + p_1 t + \dots + p_{n^2} t^{n^2}$ аннулирует \mathcal{A} . ■

Определение 3. Ненулевой многочлен со старшим коэффициентом 1 минимальной возможной степени среди аннулирующих многочленов оператора \mathcal{A} называется *минимальным многочленом* для оператора \mathcal{A} .

Лемма 1. Пусть $\mathcal{A} : V \rightarrow V$ — линейный оператор конечномерного векторного пространства V . Тогда

1. минимальный многочлен линейного оператора равен минимальному многочлену его матрицы в любом базисе;
2. минимальный многочлен является делителем любого аннулирующего многочлена;
3. минимальный многочлен единственный;
4. наборы корней у минимального многочлена и характеристического многочлена совпадают.

Доказательство.

1. Из доказательства теоремы 1 из вопроса 11 следует, что если $f(t) \in \mathbb{F}[t]$ — произвольный многочлен с коэффициентами из \mathbb{F} , то для матрицы A линейного оператора \mathcal{A} в произвольном базисе выполняется $f(A) = 0 \Leftrightarrow f(\mathcal{A}) = \mathcal{O}$. Откуда и получаем первое утверждение.
2. Пусть $f(t)$ — произвольный аннулирующий многочлен для \mathcal{A} . Поделим многочлен $f(t)$ на минимальный многочлен $\mu_{\mathcal{A}}(t)$ линейного оператора \mathcal{A} с остатком: $f(t) = \mu_{\mathcal{A}}(t)q(t) + r(t)$, при этом для остатка $r(t)$ должно выполняться $\deg r < \deg \mu_{\mathcal{A}}$. Подставляя \mathcal{A} вместо t , получаем

$$r(\mathcal{A}) = f(\mathcal{A}) - \mu_{\mathcal{A}}(\mathcal{A})q(\mathcal{A}) = 0 - 0 \cdot q(\mathcal{A}) = 0,$$

т. е. $r(t)$ аннулирует \mathcal{A} и $\deg r < \deg \mu_{\mathcal{A}}$, поэтому $r(t) = 0$, т. е. деление происходит нацело.

3. Любой из двух минимальных многочленов делит другой, а их старшие коэффициенты совпадают.
4. Т. к. характеристический многочлен $\chi_{\mathcal{A}}(t)$ линейного оператора \mathcal{A} является его аннулирующим многочленом (по теореме Гамильтона — Кэли, будет позднее), из п. 1 получаем, что $\mu_{\mathcal{A}}$ делит $\chi_{\mathcal{A}}$. Значит, корни минимального многочлена являются корнями характеристического многочлена. Обратно, пусть λ — корень характеристического многочлена. Тогда у \mathcal{A} есть

собственный вектор v с собственным значением λ . Применим к вектору v оператор $\mu_{\mathcal{A}}$. С одной стороны, расписав минимальный многочлен явно:

$$\mu_{\mathcal{A}} = t^k + b_{k-1}t^{k-1} + \dots + b_1t + b_0$$

и подставив линейный оператор \mathcal{A} вместо переменной t , получаем

$$\begin{aligned} \mu_{\mathcal{A}}(\mathcal{A})v &= (\mathcal{A}^k + b_{k-1}\mathcal{A}^{k-1} + \dots + b_1\mathcal{A} + b_0 \cdot \text{id})v = \\ &= \mathcal{A}^k v + b_{k-1}\mathcal{A}^{k-1}v + \dots + b_1\mathcal{A}v + b_0 \cdot \text{id}v = \lambda^k v + b_{k-1}\lambda^{k-1}v + \dots + b_1\lambda v + b_0v = \\ &= (\lambda^k + b_{k-1}\lambda^{k-1} + \dots + b_1\lambda + b_0)v = \mu_{\mathcal{A}}(\lambda) \cdot v. \end{aligned}$$

С другой стороны, т. к. $\mu_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$, имеем $\mu_{\mathcal{A}}(\mathcal{A})v = 0 \cdot v = \mathbf{0}$. Откуда $\mu_{\mathcal{A}}(\lambda)v = \mathbf{0}$. Т. к. $v \neq \mathbf{0}$, то число $\mu_{\mathcal{A}}(\lambda)$ равно нулю, т. е. λ — корень минимального многочлена. ■

Теорема 1 (Критерий диагонализуемости в терминах минимального многочлена). Линейный оператор $\mathcal{A} : V \rightarrow V$ конечномерного векторного пространства V над алгебраически замкнутым полем \mathbb{F} диагонализуем тогда и только тогда, когда его минимальный многочлен не имеет кратных корней.

Доказательство. Пусть сначала e_1, \dots, e_n — базис из собственных векторов для \mathcal{A} , причём $\mathcal{A}e_i = \lambda_i e_i$, т. е. матрица оператора \mathcal{A} в выбранном базисе диагональна:

$$A = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

Т. к. минимальный многочлен линейного оператора равен минимальному многочлену его матрицы в произвольном базисе, то $\mu_{\mathcal{A}}(t) = (t - \alpha_1) \dots (t - \alpha_p)$, где $\alpha_1, \dots, \alpha_p$ — все попарно различные числа $\lambda_1, \dots, \lambda_n$. Это следует из того, что выписанный многочлен является аннулирующим для матрицы A , а минимальный многочлен делит любой аннулирующий многочлен и наборы корней у минимального и характеристического многочлена совпадают.

Пусть теперь минимальный многочлен не имеет кратных корней. И пусть J — жорданова нормальная форма оператора \mathcal{A} , которая существует в силу алгебраической замкнутости поля \mathbb{F} . Позже будет написано, как искать минимальный многочлен для жордановой матрицы, а сейчас воспользуемся этим результатом:

$$\mu_{\mathcal{A}}(t) = \mu_J(t) = (t - \alpha_1)^{m_1} \dots (t - \alpha_p)^{m_p},$$

где $\alpha_1, \dots, \alpha_p$ — все попарно различные собственные значения линейного оператора \mathcal{A} , m_i — максимальный порядок жордановой клетки с α_i на диагонали в жордановой нормальной форме J . Т. к. минимальный многочлен не имеет кратных корней, получаем, что $m_i = 1$ для всех i , т. е. в J все клетки имеют порядок 1. Это и означает, что жорданова нормальная форма диагональна. ■

Задача 3 (MATH.STACKEXCHANGE). Пусть \mathbb{K} — поле и $\mathbb{F} \supseteq \mathbb{K}$ — его расширение. Пусть также $A \in \text{Mat}_{n \times n}(\mathbb{K})$ и $\mu_{A, \mathbb{K}}$ — минимальный многочлен A . Заметим, что эту матрицу A можно также рассмотреть как матрицу из $\text{Mat}_{n \times n}(\mathbb{F})$, и пусть в этом случае её минимальный многочлен есть $\mu_{A, \mathbb{F}}$. Доказать, что $\mu_{A, \mathbb{K}} = \mu_{A, \mathbb{F}}$.

Нам понадобятся две леммы.

Лемма 2. Если $A \in \text{Mat}_{n \times n}(\mathbb{K})$, то $\deg \mu_{A, \mathbb{K}} = d$ тогда и только тогда, когда выполнены следующие два условия:

1. матрицы $E, A, A^2, \dots, A^{d-1}$ линейно независимы над полем \mathbb{K} ;
2. матрицы $E, A, A^2, \dots, A^{d-1}, A^d$ линейно зависимы над полем \mathbb{K} ;

Доказательство. \Rightarrow . Пусть $\mu_{A,\mathbb{K}}(t) = a_0 + a_1t + a_2t^2 + \dots + a_dt^d$. Тогда $\mu_{A,\mathbb{K}}(A) = 0$, т. е.

$$a_0E + a_1A + a_2A^2 + \dots + a_dA^d = 0.$$

Причём, не все a_i равны нулю. Значит, выполнено условие 2. А условие 1 выполнено в силу минимальности $\mu_{A,\mathbb{K}}$. Действительно, если $E, A, A^2, \dots, A^{d-1}$ линейно зависимы, то коэффициенты этой линейной зависимости можно выбрать в качестве коэффициентов аннулирующего многочлена степени $d-1$, а такого нет, потому что $\deg \mu_{A,\mathbb{K}}$ минимальна для аннулирующих многочленов.

\Leftarrow . Эти условия равно то и означают: существует аннулирующий многочлен степени d , но не существует аннулирующего многочлена степени $d-1$. ■

Лемма 3. Если $\mathbb{F} \supseteq \mathbb{K}$ — расширение поля и $v_1, \dots, v_r \in \mathbb{K}^N \subseteq \mathbb{F}^N$. Тогда v_1, \dots, v_r линейно зависимы над полем \mathbb{K} тогда и только тогда, когда они линейно зависимы над полем \mathbb{F} .

Доказательство. \Rightarrow . Очевидно, т. к. коэффициенты этой линейной зависимости лежат как в \mathbb{K} , так и в \mathbb{F} в силу $\mathbb{K} \subseteq \mathbb{F}$.

\Leftarrow . Пусть $v_i = (v_1^i, v_2^i, \dots, v_N^i)^t$ и пусть $x_1, \dots, x_r \in \mathbb{F}^N$ — коэффициенты линейной зависимости v_1, \dots, v_r . Тогда

$$x_1 \begin{pmatrix} v_1^1 \\ \vdots \\ v_N^1 \end{pmatrix} + x_2 \begin{pmatrix} v_1^2 \\ \vdots \\ v_N^2 \end{pmatrix} + \dots + x_r \begin{pmatrix} v_1^r \\ \vdots \\ v_N^r \end{pmatrix} + x_{r+1}v'_{r+1} + \dots + x_Nv'_N = \mathbf{0},$$

где векторы v'_i ($i = r+1, \dots, N$) выбираются так, чтобы быть линейно независимыми с векторами $v_1, \dots, v_r, v'_{r+1}, \dots, v'_{i-1}$. Это можно сделать, т. к. при $i \leq N$ эти векторы образуют подпространство размерности не более N , в качестве следующего вектора можно взять любой вектор не из этого подпространства. Заметим, что в любом решении этой системы $x_{r+1} = \dots = x_N = 0$, ведь мы никак не сможем нейтрализовать слагаемое $x_iv'_i$ ($i = r+1, \dots, N$), потому что этот вектор линейно независим со всеми остальными векторами. Значит, наша дополненная векторами со штрихами система имеет ненулевое решение тогда и только тогда, когда система без штрихов имеет ненулевое решение, т. е. тогда и только тогда, когда v_1, \dots, v_r линейно независимы.

Получили СЛУ с матрицей коэффициентов $V = (v_j^i)_{j=1, \dots, N}^{i=1, \dots, r}$. Векторы v_1, \dots, v_r линейно независимы над $\mathbb{K} \Leftrightarrow$ система имеет ненулевое решение $\Leftrightarrow \det V \neq 0$ над $\mathbb{K} \Leftrightarrow \det V \neq 0$ над \mathbb{F} (тут нам неважно, над каким полем считать, все вычисления всё равно происходят внутри \mathbb{K}) $\Leftrightarrow \dots \Leftrightarrow$ векторы v_1, \dots, v_r линейно независимы над \mathbb{F} . ■

Примечание. Векторы со штрихами нужно было добавлять, чтобы сделать матрицу СЛУ квадратной, чтобы можно было считать её определитель. Сначала решение было написано без этого дополнения, но в таком виде оно, конечно, неверное.

Теперь можно и задачу решить.

▷ Понятно, что $\mu_{A,\mathbb{F}}$ делит $\mu_{A,\mathbb{K}}$. Так как их старшие коэффициенты равны, достаточно показать, что их степени равны. А это очевидное следствие из двух предыдущих лемм. ◀

18. ТЕОРЕМА ГАМИЛЬТОНА — КЭЛИ

Теорема 1 (Гамильтон, Кэли). Характеристический многочлен χ_A оператора $A: V \rightarrow V$ аннулирует этот оператор, т. е. $\chi_A(A) = 0$.

Доказательство. Для квадратной матрицы M обозначим через \widehat{M} транспонированную матрицу алгебраических дополнений. Из курса алгебры первого семестра, $M \cdot \widehat{M} = \det M \cdot E$. Теперь возьмём

в качестве M матрицу $A - t \cdot E$, где A — матрица оператора \mathcal{A} в произвольном базисе. Тогда $(A - t \cdot E)(A - t \cdot E) = \det(A - t \cdot E) \cdot E = \chi_{\mathcal{A}}(t) \cdot E$. По определению, элементы матрицы $A - t \cdot E$ являются многочленами степени не выше $n - 1$ ($n = \dim V$). Следовательно, эту матрицу можно записать в виде $A - t \cdot E = B_0 + t \cdot B_1 + t^2 \cdot B_2 + \dots + t^{n-1} B_{n-1}$, где B_i — числовые матрицы. Подставив это разложение вместе с разложением $P_{\mathcal{A}}(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n$ в формулу выше, получим

$$(A - t \cdot E)(B_0 + t \cdot B_1 + t^2 \cdot B_2 + \dots + t^{n-1} \cdot B_{n-1}) = (a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n) \cdot E.$$

Приравнявая коэффициенты при степенях t и складывая полученные уравнения, получим:

$$\begin{array}{r|l} AB_0 = a_0 E & \cdot E \\ -B_0 + AB_1 = a_1 E & \cdot A \\ -B_1 + AB_2 = a_2 E & \cdot A^2 \\ \vdots & \vdots \\ -B_{n-2} + AB_{n-1} = a_{n-1} E & \cdot A^{n-1} \\ -B_{n-1} = a_n E & \cdot A^n \end{array} \quad \xrightarrow{+} 0 = \underbrace{a_0 E + a_1 A + a_2 A^2 + \dots + a_n A^n}_{\chi_{\mathcal{A}}(A) = \chi_{\mathcal{A}}(\mathcal{A})}$$

■

19. СУЩЕСТВОВАНИЕ ОДНОМЕРНОГО ИЛИ ДВУМЕРНОГО ИНВАРИАНТНОГО ПОДПРОСТРАНСТВА ДЛЯ ЛЮБОГО ЛИНЕЙНОГО ОПЕРАТОРА НАД ПОЛЕМ ДЕЙСТВИТЕЛЬНЫХ ЧИСЕЛ

Большая часть этого вопроса — рассказ про о веществление и комплексификацию линейных пространств (взято у Панова и Винберга).

При работе с линейными операторами часто бывает удобно изменить поле скаляров. Здесь рассмотрим две такие операции: переход от \mathbb{R} к \mathbb{C} (комплексификация) и переход от \mathbb{C} к \mathbb{R} (овеществление).

Определение 1. Пусть V — линейное пространство над полем \mathbb{C} . Рассмотрим множество $V_{\mathbb{R}}$, состоящее из тех же векторов, что и V . На $V_{\mathbb{R}}$ имеется операция сложения (та же, что и на V), а вместо операции умножения на все комплексные числа оставим лишь умножение на все вещественные числа. Тогда $V_{\mathbb{R}}$ — линейное пространство над полем \mathbb{R} , которое называется *овеществлением пространства V* .

Предложение 1. Пусть e_1, \dots, e_n — базис пространства V . Тогда $e_1, \dots, e_n, ie_1, \dots, ie_n$ — базис пространства $V_{\mathbb{R}}$.

Доказательство. Проверим, что векторы $e_1, \dots, e_n, ie_1, \dots, ie_n$ линейно независимы в $V_{\mathbb{R}}$. Пусть $\lambda_1 e_1 + \dots + \lambda_n e_n + \mu_1 ie_1 + \dots + \mu_n ie_n = \mathbf{0}$ в пространстве $V_{\mathbb{R}}$, где $\lambda_k, \mu_k \in \mathbb{R}$, $k = 1, \dots, n$. Тогда в пространстве V мы имеем $(\lambda_1 + i\mu_1)e_1 + \dots + (\lambda_n + i\mu_n)e_n = \mathbf{0}$. Т.к. векторы e_1, \dots, e_n линейно независимы в V , то $\lambda_k + i\mu_k = 0 \Rightarrow \lambda_k = \mu_k = 0$, $k = 1, \dots, n$. Следовательно, векторы $e_1, \dots, e_n, ie_1, \dots, ie_n$ линейно независимы в $V_{\mathbb{R}}$. Теперь проверим, что эти векторы порождают всё пространство $V_{\mathbb{R}}$. Возьмём $v \in V_{\mathbb{R}}$ и рассмотрим его как вектор из V . Т.к. e_1, \dots, e_n — базис в V , то $v = \alpha_1 e_1 + \dots + \alpha_n e_n$, где $(\alpha_k = \lambda_k + i\mu_k) \in \mathbb{C}$ ($k = 1, \dots, n$), где $\lambda_k, \mu_k \in \mathbb{R}$. Тогда $v = \lambda_1 e_1 + \dots + \lambda_n e_n + \mu_1 ie_1 + \dots + \mu_n ie_n$. ■

Следствие 1. $\dim V_{\mathbb{R}} = 2 \dim V$.

Определение 2. Пусть V — комплексное пространство и $\mathcal{A} : V \rightarrow V$ — оператор. Тогда тот же оператор, рассматриваемый в пространстве $V_{\mathbb{R}}$ называется *овеществлением оператора \mathcal{A}* и обозначается $\mathcal{A}_{\mathbb{R}}$.

Предложение 2. Запишем матрицу оператора \mathcal{A} в базисе e_1, \dots, e_n пространства V в виде $A + iB$, где A и B — вещественные матрицы. Тогда

1. матрица оператора $\mathcal{A}_{\mathbb{R}}$ в базисе $e_1, \dots, e_n, ie_1, \dots, ie_n$ есть $\left(\begin{array}{c|c} A & -B \\ \hline B & A \end{array} \right)$;
2. $\det \mathcal{A}_{\mathbb{R}} = |\det \mathcal{A}|^2$.

Доказательство.

1. Пусть $A = (a_k^l)$ и $B = (b_k^l)$. Тогда $\mathcal{A}_{\mathbb{R}}(e_k) = \mathcal{A}(e_k) = (a_k^l + ib_k^l)e_l = a_k^l e_l + ib_k^l e_l$, $\mathcal{A}_{\mathbb{R}}(ie_k) = \mathcal{A}(ie_k) = i\mathcal{A}(e_k) = i(a_k^l + ib_k^l)e_l = -b_k^l e_l + a_k^l ie_l$.
2. Заметим, что

$$\left(\begin{array}{c|c} A & -B \\ \hline B & A \end{array} \right) \rightsquigarrow \left(\begin{array}{c|c} A - iB & -B - iA \\ \hline B & A \end{array} \right) \rightsquigarrow \left(\begin{array}{c|c} A - iB & -B - iA + i(A - iB) \\ \hline B & A + iB \end{array} \right) \rightsquigarrow \left(\begin{array}{c|c} A - iB & 0 \\ \hline B & A + iB \end{array} \right).$$

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

Определение 3. Пусть V — вещественное пространство. *Комплексной структурой* на V называется такой оператор $\mathcal{J} : V \rightarrow V$, что $\mathcal{J}^2 = -\text{id}$.

Пусть V — вещественное пространство с комплексной структурой \mathcal{J} . Введём на V операцию умножения на комплексные числа по правилу $(\lambda + i\mu)v = \lambda v + \mu \mathcal{J}v$. Тогда V превращается в комплексное пространство \tilde{V} , для которого $\tilde{V}_{\mathbb{R}} = V$, а оветствление оператора умножения на i есть \mathcal{J} . Проверка выполняется непосредственно.

Предложение 3. Пусть \mathcal{J} — комплексная структура на V . Тогда:

1. размерность вещественного пространства V чётна;
2. в подходящем базисе матрица оператора \mathcal{J} имеет вид $\left(\begin{array}{c|c} 0 & -E \\ \hline E & 0 \end{array} \right)$.

Доказательство. Т.к. любой базис V порождает \tilde{V} , то это пространство конечномерно. А т.к. $V = \tilde{V}_{\mathbb{R}}$, то $\dim V = 2 \dim \tilde{V}$ — чётно. Далее, если e_1, \dots, e_n — базис комплексного пространства \tilde{V} , то $e_1, \dots, e_n, ie_1, \dots, ie_n$ — базис пространства $\tilde{V}_{\mathbb{R}} = V$. В этом базисе оператор \mathcal{J} (оветствление оператора умножения на i) имеет указанный вид. Это прямое следствие записей сверху. ■

Определение 4. Пусть V — линейное пространство над полем \mathbb{R} . Рассмотрим пространство $V \oplus V$, состоящее из пар (u, v) , где $u, v \in V$, и введём на нём комплексную структуру следующим образом: $\mathcal{J}(u, v) := (-v, u)$. Получаемое пространство $\widetilde{V \oplus V}$ над полем \mathbb{C} называется *комплексификацией* пространства V и обозначается $V_{\mathbb{C}}$.

Предложение 4. Пусть e_1, \dots, e_n — базис пространства V . Тогда векторы $(e_1, \mathbf{0}), \dots, (e_n, \mathbf{0})$ образуют базис пространства $V_{\mathbb{C}}$.

Доказательство. Проверим линейную независимость: пусть $\alpha_1(e_1, \mathbf{0}) + \dots + \alpha_n(e_n, \mathbf{0}) = (\mathbf{0}, \mathbf{0})$ для некоторых $\alpha_k = \lambda_k + i\mu_k \in \mathbb{C}$, $k = 1, \dots, n$. Выкладка:

$$(\lambda_k + i\mu_k)(e_k, \mathbf{0}) = \lambda_k(e_k, \mathbf{0}) + \mu_k \mathcal{J}(e_k, \mathbf{0}) = (\lambda_k e_k, \mu_k e_k).$$

Подставляя это в линейную комбинацию выше, получаем

$$(\lambda_1 e_1 + \dots + \lambda_n e_n, \mu_1 e_1 + \dots + \mu_n e_n) = (\mathbf{0}, \mathbf{0}).$$

Из линейной независимости векторов e_1, \dots, e_n в V , получаем $\lambda_k = \mu_k = 0$ для всех k . А значит, $\alpha_1 = \dots = \alpha_n = 0$. Итак, $(e_1, \mathbf{0}), \dots, (e_n, \mathbf{0})$ линейно независимы. То, что они порождают всё пространство, проверяется аналогично. ■

Следствие 2. $\dim V_{\mathbb{C}} = \dim V$.

Определение 5. Пусть V — пространство над \mathbb{R} и $\mathcal{A} : V \rightarrow V$ — оператор. Оператор $\mathcal{A}_{\mathbb{C}} : V_{\mathbb{C}} \rightarrow V_{\mathbb{C}}$, заданный формулой $\mathcal{A}_{\mathbb{C}}(u, v) := (\mathcal{A}u, \mathcal{A}v)$, называется *комплексификацией оператора \mathcal{A}* .

Предложение 5. Пусть A — матрица оператора \mathcal{A} в базисе e_1, \dots, e_n . Тогда оператор $\mathcal{A}_{\mathbb{C}}$ в базисе $(e_1, 0), \dots, (e_n, 0)$ задаётся той же матрицей A .

Доказательство. $\mathcal{A}_{\mathbb{C}}(e_k, 0) = (\mathcal{A}e_k, 0) = (a_k^l e_l, 0) = a_k^l (e_l, 0)$. ■

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

Предложение 6. Пространство $(V_{\mathbb{C}})_{\mathbb{R}}$ канонически изоморфно $V \oplus V$.

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

Рассказ про оветствление и комплексификацию завершён.

Оператор $\mathcal{A} : V \rightarrow V$ в нетривиальном пространстве над полем \mathbb{C} имеет инвариантное подпространство размерности 1. Действительно, т. к. поле \mathbb{C} алгебраически замкнуто, характеристический многочлен $\chi_{\mathcal{A}}(t)$ имеет корень λ и собственный вектор $v \in V_{\lambda}$. Подпространство $\langle v \rangle$ собственное размерности 1.

Теорема 1. Оператор $\mathcal{A} : V \rightarrow V$ в нетривиальном пространстве над полем \mathbb{R} имеет инвариантное подпространство размерности 1 или 2.

Доказательство. Если характеристический многочлен $\chi_{\mathcal{A}}(t)$ имеет вещественный корень, то (аналогично) мы получаем одномерное инвариантное подпространство. Предположим, что $\chi_{\mathcal{A}}(t)$ не имеет вещественных корней. Пусть $\lambda + i\mu$ — комплексный корень, $\mu \neq 0$. Тогда $\lambda + i\mu$ — собственное значение комплексифицированного оператора $\mathcal{A}_{\mathbb{C}}$ (напомним, что в подходящих базисах матрицы операторов \mathcal{A} и $\mathcal{A}_{\mathbb{C}}$ совпадают). Возьмём соответствующий собственный вектор $u + iv \in V_{\mathbb{C}}$. Тогда $\mathcal{A}u + i\mathcal{A}v = \mathcal{A}_{\mathbb{C}}(u + iv) = (\lambda + i\mu)(u + iv) = (\lambda u - \mu v) + i(\mu u + \lambda v)$. Следовательно, $\mathcal{A}u = \lambda u - \mu v$, а $\mathcal{A}v = \mu u + \lambda v$, и линейная оболочка $\langle u, v \rangle$ является инвариантным подпространством для \mathcal{A} . ■

Следствие 1. Над полем \mathbb{R} любой линейный оператор приводим к блочно-диагональному виду, причём блоки имеют порядок не выше двух.

20. КОРНЕВЫЕ ПОДПРОСТРАНСТВА. РАЗЛОЖЕНИЕ ПРОСТРАНСТВА В ПРЯМУЮ СУММУ КОРНЕВЫХ ПОДПРОСТРАНСТВ

Определение 1. Вектор $v \in V$ называется *корневым вектором* оператора \mathcal{A} , отвечающим числу $\lambda \in \mathbb{F}$, если существует такое m , что $(\mathcal{A} - \lambda \cdot \text{id})^m v = 0$.

Обозначим через R_{λ} множество всех корневых векторов, отвечающих λ .

Предложение 1. R_{λ} является подпространством в V .

Доказательство. Пусть $u, v \in R_{\lambda}$, т. е. $(\mathcal{A} - \lambda \cdot \text{id})^l u = (\mathcal{A} - \lambda \cdot \text{id})^m v = 0$ для некоторых l и m . Тогда $(\mathcal{A} - \lambda \cdot \text{id})^l (\mu u) = 0 \ \forall \mu \in \mathbb{F}$ и $(\mathcal{A} - \lambda \cdot \text{id})^{\max\{l, m\}}(u + v) = 0$. ■

Определение 2. Подпространство $R_{\lambda} \subseteq V$ называется *корневым подпространством* для оператора \mathcal{A} , отвечающим λ .

Предложение 2. Подпространство R_{λ} нетривиально тогда и только тогда, когда λ — собственное значение оператора \mathcal{A} . При этом $V_{\lambda} \subseteq R_{\lambda}$.

Доказательство. \Leftarrow . Действительно, если λ — собственное значение, то существует $v \neq \mathbf{0}$ такой, что $(\mathcal{A} - \lambda \cdot \text{id})^1 v = \mathbf{0}$, т. е. $v \in R_\lambda$ и R_λ нетривиально. Отсюда же следует, что $V_\lambda \subseteq R_\lambda$.

\Rightarrow . Обратно, пусть R_λ содержит $u \neq \mathbf{0}$ такой, что $(\mathcal{A} - \lambda \cdot \text{id})^m u = \mathbf{0}$, причём m минимально, т. е. $v := (\mathcal{A} - \lambda \cdot \text{id})^{m-1} u \neq \mathbf{0}$. Тогда $(\mathcal{A} - \lambda \cdot \text{id})v = (\mathcal{A} - \lambda \cdot \text{id})^m u = \mathbf{0}$, т. е. v — собственный вектор, отвечающий λ . ■

Далее будем рассматривать только нетривиальные корневые подпространства.

Теорема 1. Пусть \mathcal{A} — оператор в пространстве V над алгебраически замкнутым полем, и пусть $\lambda_1, \dots, \lambda_k$ — все собственные значения оператора \mathcal{A} . Тогда $V = R_{\lambda_1} \oplus \dots \oplus R_{\lambda_k}$.

Доказательство будет опираться на три леммы.

Лемма 1. Подпространство R_λ инвариантно относительно любого оператора $\mathcal{A} - \mu \cdot \text{id}$ (в частности, относительно \mathcal{A}). Ограничение $(\mathcal{A} - \mu \cdot \text{id})|_{R_\lambda} : R_\lambda \rightarrow R_\lambda$ при $\lambda \neq \mu$ является обратимым, а при $\lambda = \mu$ нильпотентным оператором.

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

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

т. к. многочлены от оператора коммутируют. Итак, R_λ является $(\mathcal{A} - \mu \cdot \text{id})$ -инвариантным подпространством, и мы можем рассмотреть ограничение $(\mathcal{A} - \mu \cdot \text{id})|_{R_\lambda}$. Пусть $v \in \text{Ker}(\mathcal{A} - \mu \cdot \text{id})|_{R_\lambda}$, т. е. $v \in R_\lambda$ и $\mathcal{A}v = \mu v$. Тогда $(\mathcal{A} - \lambda \cdot \text{id})^m v = \mathbf{0}$ для некоторого m и $(\mathcal{A} - \lambda \cdot \text{id})v = (\mu - \lambda)v$, а значит, $(\mu - \lambda)^m v = (\mathcal{A} - \lambda \cdot \text{id})^m v = \mathbf{0}$. Следовательно, при $\mu \neq \lambda$ имеем $v = \mathbf{0}$. Тогда ядро $(\mathcal{A} - \mu \cdot \text{id})|_{R_\lambda}$ тривиально, и этот оператор инъективен (по критерию), а значит, обратим.

Наконец, если e_1, \dots, e_r — базис в R_λ и $(\mathcal{A} - \lambda \cdot \text{id})^{m_i} e_i = \mathbf{0}$, то $(\mathcal{A} - \lambda \cdot \text{id})^m v = \mathbf{0}$, $\forall v \in V$, где $m = \max\{m_1, \dots, m_r\}$. Это означает, что оператор $(\mathcal{A} - \lambda \cdot \text{id})|_{R_\lambda}$ нильпотентен. ■

Лемма 2. Корневые подпространства $R_{\lambda_1}, \dots, R_{\lambda_k}$, соответствующие различным собственным значениям $\lambda_1, \dots, \lambda_k$ образуют прямую сумму.

Доказательство. Проведём индукцию по k . При $k = 1$ доказывать нечего. Предположим, что утверждение доказано для $k - 1$ подпространств. Докажем, что соотношение $(*) : v_1 + \dots + v_k = \mathbf{0}$, где $v_i \in R_{\lambda_i}$, влечёт $v_1 = \dots = v_k = \mathbf{0}$. Имеем $(\mathcal{A} - \lambda_k \cdot \text{id})^p v_k = \mathbf{0}$ для некоторого p . Применив к $(*)$ оператор $(\mathcal{A} - \lambda_k \cdot \text{id})^p$, получим $(*) : (\mathcal{A} - \lambda_k \cdot \text{id})^p v_1 + \dots + (\mathcal{A} - \lambda_k \cdot \text{id})^p v_{k-1} = \mathbf{0}$. Т. к. подпространства $R_{\lambda_1}, \dots, R_{\lambda_{k-1}}$ инвариантны относительно $\mathcal{A} - \lambda_k \cdot \text{id}$, мы имеем $(\mathcal{A} - \lambda_k \cdot \text{id})^p v_i \in R_{\lambda_i}$, $i = 1, \dots, k-1$. По предположению индукции, из $(*)$ следует $(\mathcal{A} - \lambda_k \cdot \text{id})^p v_i = \mathbf{0}$, $i = 1, \dots, k-1$. А т. к. по предыдущей лемме оператор $\mathcal{A} - \lambda_k \cdot \text{id}$ в пространствах $R_{\lambda_1}, \dots, R_{\lambda_{k-1}}$ обратим, то $v_1 = \dots = v_{k-1} = \mathbf{0}$. Тогда из $(*)$ получаем $v_k = \mathbf{0}$. ■

Лемма 3. Размерность корневого подпространства R_λ равна кратности λ как корня характеристического многочлена оператора \mathcal{A} .

Доказательство. Обозначим через r_λ кратность корня λ . Пусть $\hat{\mathcal{A}} = \mathcal{A}|_{R_\lambda}$ — ограничение оператора \mathcal{A} на R_λ и $\tilde{\mathcal{A}} : V/R_\lambda \rightarrow V/R_\lambda$ — фактор-оператор. Тогда для характеристических многочленов мы имеем $(*) : \chi_{\mathcal{A}}(t) = \chi_{\hat{\mathcal{A}}}(t) \cdot \chi_{\tilde{\mathcal{A}}}(t) = (\lambda - t)^{\dim R_\lambda} \chi_{\tilde{\mathcal{A}}}(t)$, потому что $A = \left(\begin{array}{c|c} \hat{A} & * \\ \hline 0 & \tilde{A} \end{array} \right)$ в некотором базисе (т. к. R_λ — инвариантное подпространство). Отсюда $\dim R_\lambda \leq r_\lambda$.

Предположим, что $\dim R_\lambda < r_\lambda$. Тогда из $(*)$ следует, что λ является корнем многочлена $\chi_{\tilde{\mathcal{A}}}(t)$, т. е. собственным значением оператора $\tilde{\mathcal{A}}$. Пусть $v + R_\lambda$ — соответствующий (ненулевой) собственный вектор, т. е. $\mathcal{A}(v + R_\lambda) = \lambda(v + R_\lambda)$ или $\mathcal{A}v + R_\lambda = \lambda v + R_\lambda$. Отсюда вытекает, что $\mathcal{A}v - \lambda v = (\mathcal{A} - \lambda \cdot \text{id})v \in R_\lambda$. По определению R_λ это значит, что $\mathbf{0} = (\mathcal{A} - \lambda \cdot \text{id})^m (\mathcal{A} - \lambda \cdot \text{id})v = (\mathcal{A} - \lambda \cdot \text{id})^{m+1} v$, т. е. $v \in R_\lambda$. Но тогда $v + R_\lambda$ — нулевой вектор пространства V/R_λ . Противоречие. ■

Теперь докажем основную теорему:

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

Определение 3. Разложение $V = R_{\lambda_1} \oplus \dots \oplus R_{\lambda_k}$ называется *корневым*.

21. ЖОРДАНОВЫ КЛЕТКИ И МАТРИЦЫ, ИХ ХАРАКТЕРИСТИЧЕСКИЕ И МИНИМАЛЬНЫЕ МНОГОЧЛЕНЫ. ЖОРДАНОВ БАЗИС

Определение 1. Матрица вида $J_\lambda = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$ называется *жордановой клеткой*. Если

матрица оператора \mathcal{A} в некотором базисе является блочно-диагональной с блоками вида J_λ (возможно, соответствующих различным λ), то такая матрица называется *жордановой нормальной формой* оператора \mathcal{A} . Базис, в котором оператор имеет жорданову нормальную форму, называется *жордановым*.

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

Доказательство. Мы имеем $R_{\lambda_i} = \text{Ker}(\mathcal{A} - \lambda_i \cdot \text{id})^{m_i}$, поэтому многочлен $(t - \lambda_i)^{m_i}$ является минимальным аннулирующим для оператора $\mathcal{A}|_{R_{\lambda_i}}$. В силу теоремы о корневом разложении, $\forall v \in V$ представляется в виде $v = \sum_i v_i$, где $v_i \in R_{\lambda_i}$. Т.к. $P(\mathcal{A})$ содержит множитель $(\mathcal{A} - \lambda_i \cdot \text{id})^{m_i}$, мы имеем $P(\mathcal{A})v_i = \mathbf{0}$, т.е. $P(\mathcal{A})v = \mathbf{0}$, и многочлен $P(t)$ аннулирует оператор \mathcal{A} . С другой стороны, любой многочлен $Q(t)$, аннулирующий оператор \mathcal{A} , делится на минимальный аннулирующий многочлен для оператора $\mathcal{A}|_{R_{\lambda_i}}$, т.е. на $(t - \lambda_i)^{m_i}$, для каждого λ_i . Следовательно, $Q(t)$ делится на $P(t)$, и $P(t)$ — минимальный многочлен. ■

Про то, как искать жорданов базис, лучше всего НАПИСАЛ Антон Александрович Клячко. Однако там этот алгоритм описан очень не подробно, когда-нибудь он появится здесь со всеми обоснованиями и доказательствами. Также напишу, как всё-таки искать количество жордановых базисов над конечным полем. То, что Антон Александрович записывает клетки с единицами под собственными значениями, а не над, как это принято, **существенно**. Объяснение почему есть в теоретической задаче 3.

Как нетрудно заметить, у жордановой клетки минимальный многочлен совпадает с характеристическим. На самом деле, такие матрицы обладают особым свойством.

Задача 4 (А. А. Клячко). Докажите, что матрица коммутирует только с многочленами от себя тогда и только тогда, когда её минимальный многочлен совпадает с характеристическим.

Для начала заметим, что, не ограничивая общности, можно рассматривать матрицы сразу над алгебраически замкнутым полем (см. задачу 3 в вопросе 17), тогда у них есть жордановы формы. И понятно, что вопрос о коммутировании с какой-то матрицей равносильен вопросу о коммутировании с жордановой формой этой матрицы (просто потому что коммутирующие матрицы коммутируют в любом базисе).

Отметим также, что совпадение минимального многочлена с характеристическим равносильно тому, что у каждой жордановой клетки своё уникальное собственное значение (следствие предложения 1).

Нам понадобятся несколько технических лемм.

Лемма 1. Пусть $A = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{pmatrix}$ — жорданова матрица, причём $\chi_A(t) = \mu_A(t)$, а матрица B имеет вид

$$B = \begin{pmatrix} g_1(J_1) & & \\ & \ddots & \\ & & g_m(J_m) \end{pmatrix},$$

где g_i — многочлены. Тогда существует многочлен f такой, что $B = f(A)$.

Примечание. Это очень сильное утверждение. С помощью него мы можем «сшивать» разные многочлены, которые переводят жордановы клетки матрица A в блоки матрицы B , в единый, переводящий всю матрицу A в матрицу B . То есть, с помощью него мы сводим вопрос о коммутировании даже не к жордановым матрицам, а просто к жордановым клеткам.

Доказательство. Пусть J_i — клетка размера m_i с собственным значением λ_i на диагонали. Для каждого i положим $p_i(t) = \prod_{j \neq i} (t - \lambda_j)^{m_j}$. Т. к. все λ_i различны, $\text{НОД}(p_i, (t - \lambda_i)^{m_i}) = 1$. Поэтому существуют такие многочлены a_i и b_i , что $a_i(t)p_i(t) + b_i(t)(t - \lambda_i)^{m_i} = 1$. Отсюда $(a_i p_i)(J_j)$ равняется E при $j = i$ и 0 при $j \neq i$. Отсюда

$$(a_i p_i g_i)(A) = \begin{pmatrix} (a_i p_i g_i)(J_1) & & \\ & \ddots & \\ & & (a_i p_i g_i)(J_m) \end{pmatrix} = \begin{pmatrix} 0 & & \\ & E \cdot g_i(J_i) & \\ & & 0 \end{pmatrix} = \begin{pmatrix} 0 & & \\ & g_i(J_i) & \\ & & 0 \end{pmatrix}$$

где $g_i(J_i)$ стоит на месте i -го блока. Наконец, $B = \sum_{i=1}^m (a_i p_i g_i)(A)$. ■

Лемма 2 (Костя Зюбин). Пусть у жордановой матрицы A минимальный многочлен совпадает с характеристическим. Тогда диагональ A является многочленом от неё.

К этой лемме приведём два доказательства. Первое из них предложил Костя Зюбин (он же сформулировал лемму), второе появилось в результате обсуждения этой задачи с Николаем Юрьевичем Ероховцом.

Первое доказательство. Т. к. минимальный многочлен совпал с характеристическим, то каждому уникальному собственному значению λ_i соответствует ровно одна клетка (из предложения 1), т. е. матрица выглядит вот так:

$$A = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{pmatrix},$$

где J_i — жорданова клетка размером m_i с собственным значением λ_i , причём все такие λ_i попарно различны. Для каждого i рассмотрим многочлен $g_i(t) := \prod_{j \neq i} (t - \lambda_j)^{m_j}$. Этот многочлен, как нетрудно заметить, будет аннулировать все клетки, кроме J_i . Это значит, что

$$g_i(A) = \begin{pmatrix} g_i(J_i) & \\ & 0 \end{pmatrix}$$

(условно будем рисовать клетку в верхнем левом углу). Пусть $\mu_{G_i}(t)$ — минимальный многочлен матрицы $G_i = g_i(J_i)$ и $\lambda_i \neq 0$ (мы потом отдельно рассмотрим случай вырожденных клеток). Тогда

свободный член $\mu_{G_i}(t)$ не равен нулю. Действительно, в противном случае $\mu_{G_i}(t) = q_{G_i}(t) \cdot t \Rightarrow q_{G_i}(G_i)G_i = 0 \Rightarrow q_{G_i}(G_i) = 0 \cdot G_i^{-1} = 0$, причём $\deg q_{G_i} < \deg \mu_{G_i}$ — противоречие. Теперь будем строить искомым многочлен в несколько шагов:

$$\begin{aligned}\mu_{G_i}(g_i(A)) &= \left(\begin{array}{c|c} 0 & \\ \hline & \mu_{G_i}(0)E \end{array} \right) \\ \mu_{G_i}(g_i(A)) - \mu_{G_i}(0)E &= \left(\begin{array}{c|c} -\mu_{G_i}(0)E & \\ \hline & 0 \end{array} \right) \\ \lambda_i E - \frac{\lambda_i}{\mu_{G_i}(0)} \mu_{G_i}(g_i(A)) &= \left(\begin{array}{c|c} \lambda_i E & \\ \hline & 0 \end{array} \right).\end{aligned}$$

И вот отсюда

$$\left(\sum_i \lambda_i \right) E - \sum_{\lambda_i \neq 0} \frac{\lambda_i}{\mu_{G_i}(0)} \cdot \mu_{G_i}(g_i(A)) = \begin{pmatrix} \boxed{\lambda_1 E} & & & \\ & \boxed{\lambda_2 E} & & \\ & & \ddots & \\ & & & \boxed{\lambda_m E} \end{pmatrix}.$$

А теперь заметим, что при $\lambda_j = 0$ в получившейся матрице j -ый блок будет нулевой матрицей (там будет стоять пустая сумма), а ровно этого и хотелось. Так что найденный нами многочлен переводит A в её диагональ. ■

Второе доказательство. Рассмотрим минимальный многочлен невырожденной жордановой клетки J_i (потом поговорим про вырожденные), это $(t - \lambda_i)^{m_i}$. Подставим J_i и раскроем скобки. Получим:

$$J_i^{m_i} + \dots + (-\lambda_i)^{m_i} E = 0.$$

Заметим, что последнее слагаемое — это уже почти диагональ. Разделим обе части на $(-1)^{m_i} \lambda_i^{m_i-1}$ и перенесём в правую часть всё, кроме правого слагаемого. Получим $\lambda_i E = g_i(J_i)$. Для вырожденных клеток $g_i(t) = t^{m_i}$ (получим нулевую матрицу, а это и есть диагональ вырожденной клетки). А теперь все многочлены g_i можно «сшить» по первой лемме. ■

Лемма 3. С матрицей $D = \text{diag}(\overbrace{\lambda_1, \dots, \lambda_1}^{m_1}, \overbrace{\lambda_2, \dots, \lambda_2}^{m_2}, \dots, \overbrace{\lambda_k, \dots, \lambda_k}^{m_k})$ коммутируют такие и только такие матрицы:

$$\begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_k \end{pmatrix},$$

где B_i — блоки размера m_i .

Доказательство. То, что такие коммутируют, очевидно (проверяется непосредственно). Докажем, что коммутируют только такие. Пусть B коммутирует с D . Тогда умножив D на B слева, мы каждую строку B домножим на соответствующий коэффициент. А умножив D на B справа, мы

каждый столбец умножим на соответствующий коэффициент. Итак, получаем:

$$\begin{pmatrix} \lambda_1 b_{11} & \dots & \lambda_1 b_{1m_1} & \lambda_1 b_{1,m_1+1} & \dots & \lambda_1 b_{1,m_1+m_2} & \dots & \lambda_1 b_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \lambda_1 b_{m_1 1} & \dots & \lambda_1 b_{m_1 m_1} & \lambda_1 b_{m_1, m_1+1} & \dots & \lambda_1 b_{m_1, m_1+m_2} & \dots & \lambda_1 b_{m_1 n} \\ \lambda_2 b_{m_1+1, 1} & \dots & \lambda_2 b_{m_1+1, m_1} & \lambda_2 b_{m_1+1, m_1+1} & \dots & \lambda_2 b_{m_1+1, m_1+m_2} & \dots & \lambda_2 b_{m_1+1, n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \lambda_2 b_{m_1+m_2, 1} & \dots & \lambda_2 b_{m_1+m_2, m_1} & \lambda_2 b_{m_1+m_2, m_1+1} & \dots & \lambda_2 b_{m_1+m_2, m_1+m_2} & \dots & \lambda_2 b_{m_1+m_2, n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \lambda_k b_{n1} & \dots & \lambda_k b_{nm_1} & \lambda_k b_{n, m_1+1} & \dots & \lambda_k b_{n, m_1+m_2} & \dots & \lambda_k b_{nn} \end{pmatrix} =$$

$$= \begin{pmatrix} \lambda_1 b_{11} & \dots & \lambda_1 b_{1m_1} & \lambda_2 b_{1, m_1+1} & \dots & \lambda_2 b_{1, m_1+m_2} & \dots & \lambda_k b_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \lambda_1 b_{m_1 1} & \dots & \lambda_1 b_{m_1 m_1} & \lambda_2 b_{m_1, m_1+1} & \dots & \lambda_2 b_{m_1, m_1+m_2} & \dots & \lambda_k b_{m_1 n} \\ \lambda_1 b_{m_1+1, 1} & \dots & \lambda_1 b_{m_1+1, m_1} & \lambda_2 b_{m_1+1, m_1+1} & \dots & \lambda_2 b_{m_1+1, m_1+m_2} & \dots & \lambda_k b_{m_1+1, n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \lambda_1 b_{m_1+m_2, 1} & \dots & \lambda_1 b_{m_1+m_2, m_1} & \lambda_2 b_{m_1+m_2, m_1+1} & \dots & \lambda_2 b_{m_1+m_2, m_1+m_2} & \dots & \lambda_k b_{m_1+m_2, n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \lambda_1 b_{n1} & \dots & \lambda_1 b_{nm_1} & \lambda_2 b_{n, m_1+1} & \dots & \lambda_2 b_{n, m_1+m_2} & \dots & \lambda_k b_{nn} \end{pmatrix}.$$

Осталось приравнять элементы на соответствующих местах. Взглянем на левые верхние подматрицы $m_1 \times m_1$. Они совпадают у обеих матриц. Затем взглянем на подматрицы $M_{j=m_1+1, \dots, m_1+m_2}^{i=m_1+1, \dots, m_1+m_2}$. Они тоже совпадают! И так далее. Причём, как нетрудно увидеть, совпадают в этих матрицах только эти блоки. Получили то, что хотели. ■

Лемма 4. С жордановыми клетками коммутируют только многочлены от них².

Доказательство. Пусть $J_\lambda = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$ коммутирует с матрицей $B = (b_j^i)$. Перемножим

их с обеих сторон:

$$\begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix} = \begin{pmatrix} \lambda b_{11} + b_{21} & \lambda b_{12} + b_{22} & \dots & \lambda b_{1n} + b_{2n} \\ \lambda b_{21} + b_{31} & \lambda b_{22} + b_{32} & \dots & \lambda b_{2n} + b_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda b_{n-1, 1} + b_{n1} & \lambda b_{n-1, 2} + b_{n2} & \dots & \lambda b_{n-1, n} + b_{nn} \\ \lambda b_{n1} & \lambda b_{n2} & \dots & \lambda b_{nn} \end{pmatrix}$$

$$\begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix} \cdot \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} = \begin{pmatrix} \lambda b_{11} & b_{11} + \lambda b_{12} & \dots & b_{1, n-1} + \lambda b_{1n} \\ \lambda b_{21} & b_{21} + \lambda b_{22} & \dots & b_{2, n-1} + \lambda b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda b_{n1} & b_{n1} + \lambda b_{n2} & \dots & b_{n, n-1} + \lambda b_{nn} \end{pmatrix}.$$

Теперь приравняем соответствующие элементы получившихся матриц. Из первого столбца, $b_{i1} = 0$, $i = 2, \dots, n$. Теперь, из второго столбца, $b_{i2} = 0$, $i = 3, \dots, n$. И так далее. Получаем, что матрица B верхнетреугольная. Но это не всё, что можно про неё сказать. Приравняем клетки с координатами $(1, 2)$, получим $b_{11} + \lambda b_{12} = \lambda b_{12} + b_{22} \Rightarrow b_{11} = b_{22}$. Теперь, приравняв клетки с координатами $(2, 3)$, аналогично получим $b_{22} = b_{33}$. И так далее. Получаем, что $b_{11} = \dots = b_{nn}$. Это мы приравняли диагональ на одну выше главной. Потом приравниваем диагонали ещё на одну

²Имеется в виду, конечно, «они и только они». Однако то, что матрица коммутирует с многочленами от себя, очевидно. Интерес представляет именно то, что других коммутирующих нет.

выше, и ещё, и т. д. Получаем, что матрица B имеет вид:

$$B = \begin{pmatrix} b_0 & b_1 & \ddots & b_{n-2} & b_{n-1} \\ & b_0 & b_1 & \ddots & b_{n-2} \\ & & b_0 & \ddots & \ddots \\ & & & \ddots & b_1 \\ & & & & b_0 \end{pmatrix}.$$

Или же, $B = b_0 + b_1 N + b_2 N^2 + \dots + b_{n-1} N^{n-1}$, где $N = J_\lambda - \lambda E$ — нильпотентный оператор. А т. к. B — многочлен от $J_\lambda - \lambda E$, то это и многочлен от J_λ (подставить, раскрыть скобки). ■

Перейдём, наконец, к решению задачи.

▷ ⇐. Пусть $\chi_A(t) = \mu_A(t)$ и $BA = AB$. По лемме 2, диагональ A — это некоторый многочлен от A , поэтому B коммутирует и с диагональю A . А значит, по лемме 3, матрица B имеет блочно-диагональный вид с блоками, равными по размерам жордановым клеткам матрицы A . Коммутирование такой матрицы с жордановой матрицей A равносильно коммутированию соответствующих блоков матрицы B с жордановыми клетками A . Значит, матрица B имеет вид

$$B = \begin{pmatrix} g_1(J_1) & & \\ & \ddots & \\ & & g_m(J_m) \end{pmatrix},$$

где g_i — многочлены (по лемме 4). По лемме 1 из этих многочленов можно «сшить» многочлен f , для которого $f(A) = B$.

⇒. Пусть у матрицы A характеристический многочлен не равен минимальному. Тогда у неё какому-то собственному значению λ соответствуют (хотя бы) две клетки. Понятно, что без ограничения общности можно рассматривать случай $A = \begin{pmatrix} J_1 & \\ & J_2 \end{pmatrix}$, где J_1 и J_2 — жордановы клетки размера m_1 и m_2 соответственно с собственным значением λ на диагонали. По лемме 2 матрицы J и D коммутируют, т. е. $DJ = JD$. Предположим, что $D = f(A)$, где f — многочлен.

Теперь воспользуемся результатом задачи 1202 из «Сборника задач по аналитической геометрии и линейной алгебре» под ред. Смирнова. Утверждение³ этой задачи заключается в том, что

$$f(J_i) = \begin{pmatrix} f(\lambda) & \frac{f'(\lambda)}{1!} & \frac{f''(\lambda)}{2!} & \frac{f'''(\lambda)}{3!} & \dots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ 0 & f(\lambda) & \frac{f'(\lambda)}{1!} & \frac{f''(\lambda)}{2!} & \dots & \frac{f^{(n-2)}(\lambda)}{(n-2)!} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & f(\lambda) \end{pmatrix}.$$

А т. к. на диагоналях J_1 и J_2 собственные значения одинаковые, то на диагонали матрицы $f(A) = \begin{pmatrix} f(J_1) & \\ & f(J_2) \end{pmatrix}$ должны стоять одинаковые числа, а для D это условие не выполняется. ◀

Вспомним, что любая (невырожденная) матрица коммутирует со своей обратной: $AA^{-1} = A^{-1}A = E$. Однако есть такие невырожденные матрицы, которые коммутируют только с многочленами от себя (например, те же жордановы клетки). А это значит, что для матрицы A из условия задачи 3 верно, что A^{-1} — какой-то многочлен от A . На самом деле, это верно для всех невырожденных матриц, и доказывается это несложно. Эту задачу Антон Александрович предлагал нам ещё в первом семестре.

³По сути, полностью оно нам и не нужно. Нам достаточно понимать, что на диагонали $f(J_i)$ стоят $f(\lambda)$.

Задача 5 (А. А. Клячко). Докажите, что для любой невырожденной матрицы $A \in \text{Mat}_{n \times n}(\mathbb{F})$ существует многочлен $f \in \mathbb{F}[t]$ такой, что $A^{-1} = f(A)$.

▷ Пусть $\mu_A(t)$ — минимальный многочлен матрицы A . Заметим, что свободный член у минимального многочлена должен быть всегда ненулевой, это соображение уже несколько раз было записано. Действительно, ведь иначе можно вынести t за скобки, а т. к. оператор A невырожденный, то многочлен, который останется в скобках, аннулирует матрицу A , а степень у него меньше, чем у μ_A — противоречие. Итак, $\mu_A(A) = p_0 E + p_1 A + p_2 A^2 + \dots + p_k A^k = 0$, причём $p_0 \neq 0$. Преобразуя, получаем:

$$A \cdot \underbrace{\left(-\frac{p_1}{p_0} E - \frac{p_2}{p_0} A - \dots - \frac{p_k}{p_0} A^{k-1} \right)}_{=A^{-1}} = E.$$

◀

22. СУЩЕСТВОВАНИЕ ЖОРДАНОВА БАЗИСА ДЛЯ НИЛЬПОТЕНТНОГО ОПЕРАТОРА (ДЛЯ МАТРИЦЫ С ЕДИНСТВЕННЫМ СОБСТВЕННЫМ ЗНАЧЕНИЕМ — ХАРАКТЕРИСТИЧЕСКИМ КОРНЕМ)

Определение 1. Оператор \mathcal{A} называется *нильпотентным*, если $\mathcal{A}^k = 0$ для некоторого $k \in \mathbb{N}$. Минимальное число k , для которого $\mathcal{A}^k = 0$, называется *степенью нильпотентности* \mathcal{A} .

Пример 1. Рассмотрим оператор \mathcal{A} , заданный в базисе e_1, \dots, e_n матрицей $J_0 := \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$.

Действие этого оператора на базисные векторы описывается схемой $e_n \mapsto e_{n-1} \mapsto \dots \mapsto e_1 \mapsto 0$. Отсюда видно, что $\mathcal{A}^n = 0$, т. е. оператор \mathcal{A} нильпотентен и имеет степень n .

Сформулируем некоторые свойства нильпотентных операторов.

Предложение 1. Пусть $\mathcal{A} : V \rightarrow V$ — нильпотентный оператор, причём $\dim V = n$. Тогда

1. единственным собственным значением \mathcal{A} является 0;
2. \mathcal{A} диагонализуем тогда и только тогда, когда $\mathcal{A} = 0$;
3. степень нильпотентности \mathcal{A} не превосходит n .

Доказательство.

1. Пусть $\mathcal{A}^k = 0$ и $\mathcal{A}^{k-1} \neq 0$. Значит, существует $v \in V : u := \mathcal{A}^{k-1}v \neq 0$. Тогда $\mathcal{A}u = \mathcal{A}^k v = 0$, т. е. u — собственный вектор с собственным значением 0. Пусть $\lambda \neq 0$ — другое собственное значение, тогда по определению найдётся $w \neq 0 : \mathcal{A}w = \lambda w$. Тогда $0 = \mathcal{A}^k w = \lambda^k w$. Отсюда $0 = \lambda^k$ и $\lambda = 0$.
2. Если \mathcal{A} диагонализуем, то на диагонали его диагональной матрицы стоят собственные значения, которые все нулевые в силу п. 1. Следовательно, матрица нулевая и $\mathcal{A} = 0$.
3. Из п. 1 вытекает, что характеристический многочлен оператора \mathcal{A} есть $(-t)^n$. Тогда $\mathcal{A}^k = 0$ по теореме Гамильтона — Кэли.

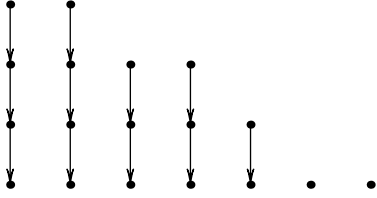
■

Теорема 1. Пусть $\mathcal{A} : V \rightarrow V$ — нильпотентный оператор. Тогда в пространстве V существует базис, в котором матрица оператора \mathcal{A} имеет блочно-диагональный вид с блоками из матриц вида J_0 (жордановы клетки с собственным значением 0) произвольных размеров. Такой вид матрицы оператора единственный с точностью до перестановки блоков.

Определение 2. Базис, существование которого утверждается в этой теореме, называется *нормальным*, а матрица оператора в таком базисе называется *нормальным видом* (*нормальной формой*) нильпотентного оператора.

Впереди нас ждёт длинное и не очень увлекательное

Доказательство. Базис, в котором матрица оператора состоит из блоков вида J_0 , удобно изображать в виде диаграммы.



В этой диаграмме точки изображают элементы нормального базиса, а стрелки описывают действие оператора \mathcal{A} . Элементы нижней строки оператор переводит в нуль, т.е. в ней стоят собственные векторы оператора (с собственным значением 0), входящие в базис. Каждый столбец соответствует одному блоку вида J_0 , причём размер блока равен высоте столбца (количеству точек в нём). Итак, нам нужно доказать существование оператора, действие оператора \mathcal{A} на элементы которого описывается

диаграммой указанного вида. Проведём индукцию по размерности пространства V . Если $\dim V = 1$, то нильпотентный оператор \mathcal{A} является нулевым, и любой ненулевой вектор в V образует нормальный базис. Пусть теперь $\dim V =: n > 1$, и пусть для размерностей, меньших n , существование нормального базиса уже доказано. Пусть $V_0 := \text{Ker } \mathcal{A}$ — подпространство собственных векторов для \mathcal{A} . Т.к. $\dim V > 0$, имеем $\dim V/V_0 < n$. Рассмотрим фактор-оператор $\tilde{\mathcal{A}} : V/V_0 \rightarrow V/V_0$, $\tilde{\mathcal{A}}(v + V_0) = \mathcal{A}v + V_0$. По индуктивному предположению $\tilde{\mathcal{A}}$ имеет нормальный базис. Можно считать его непустым: иначе $V = V_0$ и любой базис в V_0 будет нормальным для \mathcal{A} . Построим диаграмму \tilde{D} для элементов нормального базиса оператора $\tilde{\mathcal{A}}$, в каждом её столбце запишем самый верхний вектор \tilde{e}_i , $i = 1, \dots, m$ (m — количество столбцов в \tilde{D}), и положим $\tilde{e}_i =: e_i + V_0$, $e_i \in V$. Теперь построим диаграмму D из векторов пространства V следующим образом. Для $i = 1, \dots, m$ столбец с номером i будет состоять (сверху вниз) из векторов $e_i, \mathcal{A}e_i, \dots, \mathcal{A}^{h_i-1}e_i, \mathcal{A}^{h_i}e_i$, где h_i — высота i -го столбца в диаграмме \tilde{D} . Т.к. $\tilde{\mathcal{A}}^i \tilde{e}_i = \mathbf{0}$, мы имеем $\mathcal{A}^{h_i}e_i \in V_0$ и $\mathcal{A}^{h_i+1}e_i = \mathbf{0}$. Выберем базис в линейной оболочке $\langle \mathcal{A}^{h_1}e_1, \dots, \mathcal{A}^{h_m}e_m \rangle \subseteq V_0$, дополним его до базиса V_0 и поставим дополняющие векторы в качестве новых столбцов (высоты 1) в нижней строке диаграммы D ; оператор \mathcal{A} переводит их в нуль.

Таким образом, построенная диаграмма D из векторов пространства V имеет в точности такой вид, как требуется для нормального базиса. Нужно лишь проверить, что векторы, составляющие диаграмму, действительно образуют базис в V . Сначала покажем, что векторы из D порождают всё V . Пусть $v \in V$. Положим $\tilde{v} := v + V_0$. По предположению $\tilde{v} = \sum_{i=1}^m \sum_{j=0}^{h_i-1} \lambda_{ij} \tilde{\mathcal{A}}^j \tilde{e}_i$. Тогда $v -$

$\sum_{i=1}^m \sum_{j=0}^{h_i-1} \lambda_{ij} \mathcal{A}^j e_i \in V_0$. Но все векторы $\mathcal{A}^j e_i$, $j \leq h_i - 1$, лежат в строках диаграммы D , начиная со второй снизу, а подпространство V_0 порождено векторами из нижней строки D по построению. Поэтому v можно представить в виде линейной комбинации векторов из D .

Теперь доказываем линейную независимость. Сначала докажем, что векторы нижней строки линейно независимы. Действительно, если некоторая их нетривиальная линейная комбинация равна нулю, то она должна иметь вид $\sum_{i=1}^m \lambda_i \mathcal{A}^{h_i} e_i = \mathbf{0}$, ибо остальные элементы нижней строки дополняют базис линейной оболочки $\langle \mathcal{A}^{h_1} e_1, \dots, \mathcal{A}^{h_m} e_m \rangle$ до базиса V_0 . Но все $h_i \geq 1$, поэтому $\mathcal{A} \left(\sum_{i=1}^m \lambda_i \mathcal{A}^{h_i-1} e_i \right) = \mathbf{0}$, так что $\sum_{i=1}^m \lambda_i \mathcal{A}^{h_i-1} e_i \in V_0$ и $\sum_{i=1}^m \lambda_i \tilde{\mathcal{A}}^{h_i-1} \tilde{e}_i = \mathbf{0}$. Из последнего соотношения следует, что все $\lambda_i = 0$, т.к. векторы $\tilde{\mathcal{A}}^{h_i-1} \tilde{e}_i$ составляют нижнюю строку диаграммы D и являются частью базиса V/V_0 .

Наконец, покажем, что если имеется любая нетривиальная линейная комбинация векторов D , равная нулю, то из неё можно получить нетривиальную линейную зависимость между векторами нижней строки D . Отметим самую верхнюю строку D , в которой имеются ненулевые коэффициенты этой линейной комбинации. Пусть номер этой строки (считая снизу) равен h . Применим к

этой линейной комбинации оператор \mathcal{A}^{h-1} . При этом её часть, лежащая в h -ой строке, перейдёт в нетривиальную линейную комбинацию элементов нижней строки, а остальные слагаемые обратятся в нулю. Это завершает доказательство существования нормального базиса.

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

23. СУЩЕСТВОВАНИЕ ЖОРДАНОВОЙ НОРМАЛЬНОЙ ФОРМЫ МАТРИЦЫ НАД АЛГЕБРАИЧЕСКИ ЗАМКНУТЫМ ПОЛЕМ

Посмотрим, как выглядит матрица оператора $\mathcal{A}|_{R_\lambda}$ (ограничения оператора \mathcal{A} на корневое подпространство R_λ). Т. к. $(\mathcal{A} - \lambda \cdot \operatorname{id})|_{R_\lambda}$ является нильпотентным оператором, в пространстве R_λ можно выбрать нормальный базис для этого оператора. Тогда матрица оператора $(\mathcal{A} - \lambda \cdot \operatorname{id})|_{R_\lambda}$

в этом базисе будет состоять из блоков вида
$$\begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}.$$
 А значит, матрица оператора $\mathcal{A}|_{R_\lambda}$

состоит из блоков вида
$$\begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix},$$
 т. е. жордановых клеток.

Теорема 1. Для любого оператора \mathcal{A} в пространстве V над алгебраически замкнутым полем существует жорданов базис (в котором оператор имеет жорданову нормальную форму).

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

Этот билет я сдавал на коллоквиуме Антону Александровичу, и он предложил мне следующую интересную задачу.

Задача 6 (А. А. Клячко). Верно ли, что если у каждого оператора в линейном пространстве V над полем \mathbb{F} есть жорданова форма, то поле \mathbb{F} алгебраически замкнутое?

Нам пригодится вспомогательное утверждение (которое интересно и само по себе).

Лемма 1. Пусть имеем n -мерное векторное пространство V над полем \mathbb{F} . Тогда для любого многочлена $p \in \mathbb{F}[t]$ степени не больше n существует оператор \mathcal{A} такой, что $\chi_{\mathcal{A}} = p$.

Доказательство. Фиксируем некоторый базис e_1, \dots, e_n пространства V и рассмотрим оператор

с матрицей

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}$$

в этом базисе. Здесь a_1, \dots, a_n — произвольные элементы \mathbb{F} . Тогда можем вычислить его характеристический многочлен (вычисления приведены ниже). Небольшой комметарий к ним: определитель считается с помощью разложения по последней строке. Как известно, в таком разложении знак у слагаемого выбирается «в шахматном порядке», причём в левой верхней клетке матрицы стоит «+». На самом деле, «+» будет во всех клетках на главной диагонали.

$$\begin{aligned} \chi_A(t) = \det(A - t \cdot E) &= \det \begin{pmatrix} -t & 0 & \cdots & 0 & -a_0 \\ 1 & -t & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} - t \end{pmatrix} = (-1)^{n+1} \cdot (-a_0) \cdot \det \begin{pmatrix} 1 & -t & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} + \\ &+ (-1)^n \cdot (-a_1) \cdot \det \begin{pmatrix} -t & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} + (-1)^{n+1} \cdot (-a_2) \cdot \det \begin{pmatrix} -t & 0 & \cdots & 0 \\ 1 & -t & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} + \dots + (-a_{n-1} - t) \cdot \\ &\cdot \det \begin{pmatrix} -t & 0 & \cdots & 0 \\ 1 & -t & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -t \end{pmatrix} = (-1)^n a_0 + (-1)^n a_1 \cdot t + (-1)^n a_2 \cdot t^2 + \dots + (-1)^n a_{n-1} \cdot t^{n-1} + (-1)^n t^n. \end{aligned}$$

■

Теперь приступим к решению задачи:

▷ Возьмём линейный оператор $\mathcal{A} : V \rightarrow V$ и рассмотрим его матрицу A в фиксированном базисе e_1, \dots, e_n n -мерного пространства V (n можно брать любым) над полем \mathbb{F} . С одной стороны, её характеристический многочлен (из предыдущей леммы) может быть любым многочленом из $\mathbb{F}[t]$ степени не выше n . С другой стороны, характеристический многочлен жордановой матрицы J_A раскладывается на линейные множители. А т. к. $A \sim J_A$, это тот же самый многочлен. Получаем, что любой многочлен из $\mathbb{F}[t]$ раскладывается на линейные множители, что значит, что поле \mathbb{F} алгебраически замкнуто. ◀

24. Единственность жордановой нормальной формы

Теорема 1. Жорданова нормальная форма оператора единственна с точностью до перестановки клеток.

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

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

25. Билинейные функции и их матрицы. Изменение матрицы билинейной формы при замене базиса. Ранг билинейной функции. Симметрические билинейные функции

Определение 1. Пусть V — линейное пространство над полем \mathbb{F} . Функция $\mathcal{B} : V \times V \rightarrow \mathbb{F}$ называется *билинейной функцией*, если она линейна по каждому аргументу:

$$\begin{aligned}\mathcal{B}(\lambda_1 x_1 + \lambda_2 x_2, y) &= \lambda_1 \mathcal{B}(x_1, y) + \lambda_2 \mathcal{B}(x_2, y), \\ \mathcal{B}(x, \mu_1 y_1 + \mu_2 y_2) &= \mu_1 \mathcal{B}(x, y_1) + \mu_2 \mathcal{B}(x, y_2)\end{aligned}$$

для любых $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{F}$ и $x, x_1, x_2, y, y_1, y_2 \in V$.

Определение 2. Матрицей билинейной функции \mathcal{B} в базисе e_1, \dots, e_n пространства V называется квадратная матрица $B = (b_{ij})$ размера n , где $b_{ij} = \mathcal{B}(e_i, e_j)$.

Зная матрицу $B = (b_{ij})$ билинейной функции, можно восстановить значение $\mathcal{B}(x, y)$ на любой паре векторов $x = x^i e_i$ и $y = y^j e_j$:

$$\mathcal{B}(x, y) = \mathcal{B}(x^i e_i, y^j e_j) = x^i y^j \mathcal{B}(e_i, e_j) = b_{ij} x^i y^j = x^t B y.$$

Определение 3. Выражение $B(x, y) = b_{ij} x^i y^j$ называется *билинейной формой*.

Билинейная форма представляет собой однородный многочлен степени 2 от двух наборов переменных x^1, \dots, x^n и y^1, \dots, y^n , который линеен по x при фиксированных y и линеен по y при фиксированных x .

Теорема 1 (Закон изменения матрицы билинейной функции). Имеет место соотношение

$$B^t = C^t B C,$$

где B — матрица билинейной функции $\mathcal{B} : V \times V \rightarrow \mathbb{F}$ в базисе e_1, \dots, e_n , B^t — матрица в базисе $e_{1'}, \dots, e_{n'}$ и C — матрица перехода от базиса e_1, \dots, e_n к базису $e_{1'}, \dots, e_{n'}$.

Доказательство. Пусть $B = (b_{ij})$, $B^t = (b'_{ij})$, $C = (c^i_{i'})$. Мы имеем

$$b'_{ij} = \mathcal{B}(e_{i'}, e_{j'}) = \mathcal{B}(c^i_{i'} e_i, c^j_{j'} e_j) = c^i_{i'} c^j_{j'} \mathcal{B}(e_i, e_j) = c^i_{i'} b_{ij} c^j_{j'},$$

что эквивалентно матричному соотношению $B^t = C^t B C$. ■

Следствие 1. Ранг матрицы билинейной функции не зависит от базиса.

Доказательство. Т. к. матрица C обратима, $\text{rk } B^t = \text{rk}(C^t B C) = \text{rk } B$. ■

Определение 4. Рангом билинейной функции \mathcal{B} (обозначается $\text{rk } \mathcal{B}$) называется ранг её матрицы в произвольном базисе. Билинейная функция \mathcal{B} в пространстве V называется *невыврожденной*, если $\text{rk } \mathcal{B} = \dim V$.

Множество $\mathcal{B}(V)$ всех билинейных функций в пространстве V образует линейное пространство относительно операций сложения функций и умножения функций на скаляры. Сопоставление билинейной функции \mathcal{B} её матрицы B в фиксированном базисе e_1, \dots, e_n устанавливает изоморфизм между пространством $\mathcal{B}(V)$ и пространством квадратных матрицы $\text{Mat}_{n \times n}(\mathbb{F})$. Как и в случае

пространства линейных операторов $\text{Hom}(V, V)$, этот изоморфизм неканоничен, т. к. он зависит от выбора базиса.

Теорема 2. Отображение $\varphi : \text{B}(V) \rightarrow \text{Hom}(V, V^*)$, сопоставляющее билинейной функции \mathcal{B} линейной отображение $\tilde{\mathcal{B}} : V \rightarrow V^*$, задаваемое формулой

$$\tilde{\mathcal{B}}(x) := \mathcal{B}(x, \cdot) \text{ для } x \in V,$$

является каноническим изоморфизмом. Здесь $\mathcal{B}(x, \cdot) \in V^*$ — линейная функция, значение которой на векторе $y \in V$ есть $\mathcal{B}(x, y)$.

Доказательство. Отображение φ линейно, т. к. билинейная функция линейна по первому аргументу x . Кроме того, отображение φ биективно: обратное отображение φ^{-1} ставит в соответствие линейному отображению $\tilde{\mathcal{B}} : V \rightarrow V^*$ билинейную функцию \mathcal{B} , заданную по формуле $\mathcal{B}(x, y) = \tilde{\mathcal{B}}(x)(y)$. Следовательно, φ — изоморфизм. Этот изоморфизм каноничен, т. к. в его конструкции не использовался базис. ■

Определение 5. Билинейная функция $\mathcal{B} : V \times V \rightarrow \mathbb{F}$ называется *симметрической*, если $\mathcal{B}(y, x) = \mathcal{B}(x, y)$, и *кососимметрической*, если $\mathcal{B}(x, x) = 0$ для любых $x, y \in V$.

Примечание. Именно такое определение кососимметрической билинейной функции правильное, потому что над полем характеристики 2 стандартное определение $\mathcal{B}(x, y) = -\mathcal{B}(y, x)$ равносильно определению симметричности. А над полем характеристики не 2 наше определение и стандартное равносильны. Действительно, если $\mathcal{B}(x, x) = 0 \ \forall x \in V$, то $\forall u, v \in V$ выполняется

$$0 = \mathcal{B}(u + v, u + v) = \underbrace{\mathcal{B}(u, u)}_0 + \mathcal{B}(u, v) + \mathcal{B}(v, u) + \underbrace{\mathcal{B}(v, v)}_0 \Rightarrow \mathcal{B}(u, v) = -\mathcal{B}(v, u),$$

а в обратную сторону очевидно.

26. КВАДРАТИЧНЫЕ ФОРМЫ И ИХ МАТРИЦЫ. ВОССТАНОВЛЕНИЕ СИММЕТРИЧЕСКОЙ БИЛИНЕЙНОЙ ФУНКЦИИ ПО ДАННОЙ КВАДРАТИЧНОЙ ФУНКЦИИ. ДИАГОНАЛЬНЫЙ ВИД КВАДРАТИЧНОЙ ФОРМЫ. АЛГОРИТМ ЛАГРАНЖА

Определение 1. *Квадратичной формой* над \mathbb{F} называется однородный многочлен второй степени от n переменных $x = (x^1, \dots, x^n)$, т. е. многочлен вида

$$Q(x) = Q(x^1, \dots, x^n) = q_{ij}x^i x^j = \sum_{i=1}^n q_{ii}(x^i)^2 + \sum_{i < j} 2q_{ij}x^i x^j,$$

где $q_{ji} = q_{ij} \in \mathbb{F}$. Симметричная матрица $Q = (q_{ij})$ размера $n \times n$ называется *матрицей квадратичной формы*.

Если $B(x, y) = b_{ij}x^i y^j$ — симметрическая билинейная форма, то $B(x, x) = b_{ij}x^i x^j$ является квадратичной формой с матрицей B . Таким образом, квадратичная форма $B(x, x)$ полностью определяет симметрическую билинейную форму $B(x, y)$, а значит, и симметрическую билинейную функцию $\mathcal{B}(x, y)$. Это можно увидеть и не прибегая к выбору базиса: для симметрической билинейной функции имеет место соотношение

$$\mathcal{B}(x, y) = \frac{1}{2}(\mathcal{B}(x + y, x + y) - \mathcal{B}(x, x) - \mathcal{B}(y, y)).$$

Примечание. Заметим, что здесь мы должны уметь делить на 2, поэтому формула верна только для полей с характеристикой не 2.

Определение 2. Функцию $V \rightarrow \mathbb{F}$, $x \mapsto \mathcal{B}(x, x)$ называют *квадратичной функцией*.

Теорема 1. Для симметрической билинейной функции \mathcal{B} над полем характеристики, отличной от 2, существует базис, в котором её матрица диагональна. Другими словами, любую квадратичную форму $Q(x)$ линейной заменой координат $x = Cy$ можно привести к виду

$$Q(y) = r_{11}(y^1)^2 + \dots + r_{nn}(y^n)^2.$$

Мы приведём два доказательства этого факта. В первом случае мы будем работать с квадратичными формами и координатами, а во втором — с симметрическими билинейными функциями и базисами. Каждое из доказательств будет проведено таким образом, что его можно будет использовать как алгоритм.

Первое доказательство (метод Лагранжа). Пусть $Q(x) = q_{ij}x^i x^j$ — квадратичная форма. Доказательство заключается в последовательном упрощении $Q(x)$, использующем основное и два вспомогательных преобразования.

Основное преобразование производится, если в квадратичной форме $Q(x) = q_{ij}x^i x^j$ первый коэффициент q_{11} не равен нулю. Тогда имеем

$$\begin{aligned} Q(x^1, \dots, x^n) &= q_{11}(x^1)^2 + 2q_{12}x^1 x^2 + \dots + 2q_{1n}x^1 x^n + \sum_{i,j>1} q_{ij}x^i x^j = \\ &= q_{11} \left(x^1 + \frac{q_{12}}{q_{11}}x^2 + \dots + \frac{q_{1n}}{q_{11}}x^n \right)^2 - q_{11} \left(\frac{q_{12}}{q_{11}}x^2 + \dots + \frac{q_{1n}}{q_{11}}x^n \right)^2 + \sum_{i,j>1} q_{ij}x^i x^j = \\ &= q_{11} \left(x^1 + \frac{q_{12}}{q_{11}}x^2 + \dots + \frac{q_{1n}}{q_{11}}x^n \right)^2 + Q'(x^2, \dots, x^n), \end{aligned}$$

где $Q'(x^2, \dots, x^n)$ — некоторая квадратичная форма от $n - 1$ переменных. Теперь сделаем замену координат

$$\begin{aligned} u^1 &= x^1 + \frac{q_{12}}{q_{11}}x^2 + \dots + \frac{q_{1n}}{q_{11}}x^n, \\ u^2 &= x^2, \quad \dots, \quad u^n = x^n. \end{aligned}$$

В результате $Q(x)$ преобразуется к виду

$$Q(u^1, \dots, u^n) = q_{11}(u^1)^2 + Q'(u^2, \dots, u^n).$$

Если в форме $Q'(u^2, \dots, u^n)$ первый коэффициент (т. е. q'_{22}) не равен нулю, то мы снова можем применить основное преобразование, и т. д.

Первое вспомогательное преобразование производится, если $q_{11} = 0$, но существует $q_{ii} \neq 0$. В этом случае мы делаем замену $u^1 = x^i, u^i = x^1$, а остальные координаты без изменений. В результате получаем $q'_{11} \neq 0$.

Второе вспомогательное преобразование производится, если все коэффициенты q_{ii} при квадратах равны нулю, но при этом есть хотя бы один ненулевой коэффициент (в противном случае $Q(x) \equiv 0$ уже имеет нужный вид). Пусть $q_{ij} \neq 0$, где $i < j$. Произведём замену координат

$$x^i = u^i, \quad x^j = u^i + u^j, \quad x^k = u^k \text{ при } k \neq i, j.$$

В результате форма $Q(x)$ преобразуется к виду

$$Q(x) = 2q_{ij}x^i x^j + \dots = 2q_{ij}u^i(u^i + u^j) + \dots = 2q_{ij}(u^i)^2 + \dots,$$

где \dots означает члены, не содержащие квадратов. Далее мы можем применить предыдущие преобразования.

Последовательно применяя основное преобразование и (если нужно) вспомогательные преобразования, мы приводим форму $Q(x)$ к диагональному виду. ■

Второе доказательство (метод поиска базиса). Пусть $B = (b_{ij})$ — матрица билинейной функции \mathcal{B} в исходном базисе e_1, \dots, e_n .

Основное преобразование производится, если $b_{11} = \mathcal{B}(e_1, e_1) \neq 0$. Выберем новый базис следующим образом:

$$\begin{aligned} e_{1'} &= e_1, \\ e_{2'} &= e_2 - \frac{\mathcal{B}(e_1, e_2)}{\mathcal{B}(e_1, e_1)} e_1 = e_2 - \frac{b_{12}}{b_{11}} e_1, \\ &\vdots \\ e_{n'} &= e_n - \frac{\mathcal{B}(e_1, e_n)}{\mathcal{B}(e_1, e_1)} e_1 = e_n - \frac{b_{1n}}{b_{11}} e_1. \end{aligned}$$

В результате получим $\mathcal{B}(e_{1'}, e_{1'}) = 0$ при $i > 1$. Таким образом, матрица билинейной функции \mathcal{B} в новом базисе принимает вид

$$B' = \left(\begin{array}{c|ccc} b_{11} & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{B}' & \\ 0 & & & \end{array} \right),$$

где \tilde{B}' — матрица размера $(n-1) \times (n-1)$ билинейной функции \mathcal{B} на подпространстве $\langle e_{2'}, \dots, e_{n'} \rangle$. Далее мы работаем уже с этой матрицей \tilde{B}' .

Первое вспомогательное преобразование производится, если $b_{11} = 0$, но имеется $b_{ii} \neq 0$. Тогда делаем замену, меняющую местами первый и i -ый базисный векторы.

Второе вспомогательное преобразование производится, если все b_{ii} равны нулю, но при этом билинейная функция \mathcal{B} не является тождественно нулевой, т. е. $b_{ij} = \mathcal{B}(e_i, e_j) \neq 0$ для некоторых $i < j$. Произведём замену базиса

$$e_{i'} = e_i + e_j, \quad e_{j'} = e_j, \quad e_{k'} = e_k \text{ при } k \neq i, j.$$

Далее можем применить предыдущие преобразования.

Последовательно применяя основное преобразование и дополняя его в необходимых случаях вспомогательными преобразованиями, мы получаем базис f_1, \dots, f_n , в котором матрица билинейной функции \mathcal{B} имеет диагональный вид. ■

Обратим внимание, что основное и вспомогательное преобразование в обоих доказательствах — это одно и то же преобразование, просто в первом случае оно записано через координаты, а во втором — через базисы. Так что диагональные матрицы, получаемые первым и вторым методом, совпадают, как и все промежуточные матрицы.

Примечание. Если при приведении матрицы билинейной функции (квадратичной формы) к диагональному виду использовалось лишь основное преобразование, то матрица перехода от исходного базиса к базису, в котором матрица имеет диагональный вид, является верхнетреугольной.

Задача 7 (А. А. Клячко). Над полем \mathbb{Z}_2 симметрическая билинейная функция с матрицей $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ не приводится к диагональному виду заменой базиса.

Приведу здесь два решения. Первое моё:

▷ Пусть можно, и матрица замены базисов имеет вид $C = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Произведём замену:

$$C^t \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} C = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & a \\ d & b \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2ac & bc + ad \\ bc + ad & 2bd \end{pmatrix}.$$

Тут можно рассуждать по-разному:

1. Над \mathbb{Z}_2 имеем $2 = 0$, поэтому на диагонали у такой матрицы стоят нули. Поэтому если она диагональная, то нулевая, а нулевая матрица — она нулевая в любом базисе.
2. Вне диагонали должны стоять нули, поэтому $bc + ad = 0$. Над \mathbb{Z}_2 имеем $-1 = 1$, поэтому $0 = ad - bc = \det C$, но матрица перехода между базисами должна быть невырожденной.

◀

Второе — Антона Александровича:

▷ Пусть \mathcal{B} — билинейная функция с такой матрицей в базисе e_1, e_2 . Тогда

$$\mathcal{B}(x, x) = \mathcal{B}(x^1 e_1 + x^2 e_2, x^1 e_1 + x^2 e_2) = x^1 \underbrace{\mathcal{B}(e_1, e_1)}_0 + x^2 \underbrace{\mathcal{B}(e_2, e_2)}_0 + x^1 x^2 \underbrace{(\mathcal{B}(e_1, e_2) + \mathcal{B}(e_2, e_1))}_{\text{Над } \mathbb{Z}_2 \text{ это } 0} = 0.$$

Значит, при любой замене базиса на диагонали будут стоять нули. А ещё при любой замене базиса матрица должна быть симметрической. Поэтому она либо нулевая, либо такая, как есть. Ранг должен сохраняться, поэтому она такая, как есть, т. е. не диагональная.

◀

27. НОРМАЛЬНЫЙ (КАНОНИЧЕСКИЙ) ВИД КВАДРАТИЧНОЙ ФОРМЫ НАД ПОЛЯМИ ДЕЙСТВИТЕЛЬНЫХ И КОМПЛЕКСНЫХ ЧИСЕЛ. ЗАКОН ИНЕРЦИИ

Над полем \mathbb{R} квадратичную форму можно далее упростить:

Предложение 1. Для любой симметрической билинейной функции \mathcal{B} в пространстве над полем \mathbb{R} существует базис, в котором её матрица имеет диагональный вид с 1, -1 и 0 на диагонали.

Доказательство. Сначала с помощью теоремы 1 приведём квадратичную форму к виду

$$Q(u) = r_{11}(u^1)^2 + \dots + r_{nn}(u^n)^2.$$

Если $r_{ii} > 0$, то замена $y^i = \sqrt{r_{ii}}u^i$ приводит слагаемое $r_{ii}(u^i)^2$ к виду $(y^i)^2$. Если же $r_{ii} < 0$, то замена $y^i = \sqrt{-r_{ii}}u^i$ приводит слагаемое $r_{ii}(u^i)^2$ к виду $-(y^i)^2$. В результате получаем требуемый вид квадратичной формы с коэффициентами 1, -1 и 0. ■

Определение 1. Вид, описанный в последнем предложении, называется *нормальным видом* вещественной симметрической билинейной формы (вещественной квадратичной формы).

Над полем \mathbb{C} квадратичную форму можно ещё больше упростить:

Предложение 2. Для любой симметрической билинейной функции \mathcal{B} над полем \mathbb{C} существует базис, в котором её матрица имеет диагональный вид с 1 и 0 на диагонали.

Доказательство. Сначала мы с помощью последнего предложения приведём квадратичную форму к виду $(y^1)^2 + \dots + (y^p)^2 - (y^{p+1})^2 - \dots - (y^{p+q})^2$. Затем сделаем замену координат $y^k = z^k$ при $k \leq p$ и $y^k = iz^k$ при $k > p$. В результате получим требуемый вид. ■

Определение 2. Вид, описанный в последнем предложении, называется *нормальным видом* комплексной симметрической билинейной формы (комплексной квадратичной формы).

В случае симметрической билинейной формы над \mathbb{C} нормальный вид зависит только от её ранга, и поэтому мы получаем:

Предложение 3. Две комплексные симметрические билинейные формы (комплексные квадратичные формы) получаются друг из друга линейной заменой координат только и только тогда, когда их ранги совпадают.

В случае симметрических билинейных форм над \mathbb{R} ситуация сложнее: их нормальный вид не определяется одним лишь рангом, а зависит ещё от количества 1 и -1 на диагонали матрицы.

Оказывается, что нормальный вид такой формы не зависит от способа приведения к нормальному виду.

Теорема 1 (Закон инерции). Количество 1, -1 и 0 на диагонали нормального вида матрицы вещественной симметрической билинейной функции \mathcal{B} не зависит от способа приведения к нормальному виду. Другими словами, если квадратичная форма $Q(x)$ вещественной линейной заменой $x = Cy$ приводится к виду

$$(y^1)^2 + \dots + (y^p)^2 - (y^{p+1})^2 - \dots - (y^{p+q})^2,$$

а вещественной линейной заменой $x = C'z$ — к виду

$$(z^1)^2 + \dots + (z^{p'})^2 - (z^{p'+1})^2 - \dots - (z^{p'+q'})^2,$$

то мы имеем $p = p'$ и $q = q'$.

Доказательство. Пусть (x^1, \dots, x^n) — координаты в исходном базисе e_1, \dots, e_n пространства V , (y^1, \dots, y^n) — координаты в базисе f_1, \dots, f_n , а (z^1, \dots, z^n) — координаты в базисе g_1, \dots, g_n . Рассмотрим подпространства

$$\begin{aligned} U_+ &:= \langle f_1, \dots, f_p \rangle, & U_- &:= \langle f_{p+1}, \dots, f_{p+q} \rangle, & U_0 &:= \langle f_{p+q+1}, \dots, f_n \rangle, \\ W_+ &:= \langle g_1, \dots, g_{p'} \rangle, & W_- &:= \langle g_{p'+1}, \dots, g_{p'+q'} \rangle, & W_0 &:= \langle g_{p'+q'+1}, \dots, g_n \rangle \end{aligned}$$

Для ненулевого вектора $x \in U_+$ мы имеем $x = y^1 f_1 + \dots + y^p f_p$, поэтому $\mathcal{B}(x, x) = (y^1)^2 + \dots + (y^p)^2 > 0$. Аналогично, если $x \in U_- \oplus U_0$, то $\mathcal{B}(x, x) \leq 0$. Для ненулевого вектора $x \in W_+$ мы имеем $\mathcal{B}(x, x) > 0$, а для $x \in W_- \oplus W_0$ имеем $\mathcal{B}(x, x) \leq 0$. Предположим, что $p > p'$. Тогда

$$\dim U_+ + \dim(W_- \oplus W_0) = p + (n - p') > n = \dim V,$$

значит, $U_+ \cap (W_- \oplus W_0) \neq \{0\}$. Возьмём ненулевой вектор x в этом пересечении. Т. к. $x \in U_+$, имеем $\mathcal{B}(x, x) > 0$. С другой стороны, $x \in W_- \oplus W_0$ следует, что $\mathcal{B}(x, x) \leq 0$. Противоречие. Аналогично приводится к противоречию случай $p < p'$.

Следовательно, $p = p'$. Кроме того, $p + q = p' + q' = \text{rk } \mathcal{B}$, а значит, и $q = q'$. ■

Примечание. Важно понимать, что человечество на самом деле не умеет приводить квадратичные формы к какому-то адекватному виду. Мы хоть что-то знаем только про очень узкие ситуации — симметричная (кососимметричная) матрица, только над полями \mathbb{R} или \mathbb{C} и т. п. Даже над полем \mathbb{Q} понять, являются две квадратичные формы эквивалентными, сложно. Есть инвариант в виде ранга, есть замечание, что $\det A' = \det(C^t A C) = (\det C)^2 \det A$ (т. е. отношение определителей должно быть квадратом элемента поля). Но вот примерно на этом какие-то нормальные соображения заканчиваются. Ходят слухи, что в НМУ на «Алгебре 2» в 2024 году учили что-то понимать про поля типа \mathbb{Z}_p .

Определение 3. Число p называется *положительным индексом инерции*, а q — *отрицательным*. А в общем, *индексами инерции* называется пара (p, q) .

Определение 4. Разность $p - q$ между числом положительных и отрицательных диагональных элементов в нормальном виде называется *сигнатурой* вещественной симметрической билинейной функции.

Из последней теоремы следует, что сигнатура, как и ранг, являются инвариантом вещественной симметрической билинейной функции, т. е. не зависит от базиса.

Следствие 1. Две вещественные симметрические билинейные формы получаются друг из друга линейной заменой координат тогда и только тогда, когда их ранги и сигнатуры совпадают.

28. ТЕОРЕМА ЯКОБИ

Теорема Якоби позволяет (при выполнении некоторого дополнительного условия) найти нормальный вид квадратичной формы без нахождения преобразования.

Напомним, что *угловым минором порядка k* матрицы Q называется минор (определитель подматрицы), составленный из первых k строк и первых k столбцов. Угловым минором порядка k будем обозначать через $\det Q_k$.

Теорема 1 (Якоби). Предположим, что все угловые миноры матрицы Q квадратичной формы отличны от нуля до порядка $r = \operatorname{rk} Q$. Тогда существует замена координат, приводящая данную квадратичную форму к виду

$$\det Q_1 (x^1)^2 + \frac{\det Q_2}{\det Q_1} (x^2)^2 + \dots + \frac{\det Q_r}{\det Q_{r-1}} (x^r)^2.$$

Определение 1. Скажем, что для квадратичной формы имеет место *регулярный случай*, если она приводится к диагональному виду последовательным применением исключительно основного преобразования Лагранжа.

Лемма 1. Для квадратичной формы $Q(x)$ имеет место регулярный случай тогда и только тогда, когда все угловые миноры её матрицы Q отличны от нуля до порядка $r = \operatorname{rk} Q$.

Доказательство. Пусть угловые миноры до порядка r отличны от нуля. Тогда $q_{11} = \det Q_1$ — угловой минор порядка 1, который не равен нулю по предположению. Значит, на первом шаге применимо основное преобразование метода Лагранжа.

Предположим теперь, что после k -кратного применения основного преобразования метода Лагранжа матрица квадратичной формы принимает вид

$$Q' = \left(\begin{array}{ccc|ccc} q'_{11} & & 0 & & & \\ & \ddots & & & & \\ 0 & & q'_{kk} & & & \\ \hline & & & q'_{k+1,k+1} & \cdots & \\ & & 0 & \vdots & \ddots & \end{array} \right).$$

Заметим, что матрица замены координат для основного преобразования есть

$$C = \begin{pmatrix} 1 & -\frac{q_{12}}{q_{11}} & \dots & -\frac{q_{1n}}{q_{11}} \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

(см. формулы из метода замены базисов). Для угловых подматриц Q_k мы имеем $Q'_k = C_k^t Q_k C_k$, где C_k — угловая подматрица матрицы C . Т.к. $\det C_k = 1$, мы получаем $\det Q'_k = \det Q_k$, т.е. угловые миноры матрицы квадратичной формы не меняются при основном преобразовании метода Лагранжа.

Возвращаясь к матрице Q' , мы получаем $\det Q_{k+1} = \det Q'_{k+1} = q'_{11} \dots q'_{kk} q'_{k+1,k+1} \neq 0$ при $k < r$ по предположению. Следовательно, $q'_{k+1,k+1} \neq 0$, и мы снова можем применить основное преобразование.

После r -кратного применения основного преобразования мы получаем матрицу Q' , где $k = r$ и матрица в правом нижнем углу нулевая. Следовательно, квадратичная форма приведена к диагональному виду последовательным применением основного преобразования метода Лагранжа, и мы имеем регулярный случай.

Пусть теперь имеет место регулярный случай, т.е. форма приведена к диагональному виду с ненулевыми числами q'_{11}, \dots, q'_{rr} на диагонали последовательным применением основного преобразования. Тогда, т.к. угловые миноры не меняются при основном преобразовании, мы имеем $\det Q_k = \det Q'_k = q'_{11} \dots q'_{kk} \neq 0$ при $k \leq r$. ■

Теперь докажем теорему Якоби.

Доказательство. В силу предыдущей леммы, мы можем привести квадратичную форму к диагональному виду

$$q'_{11}(u^1)^2 + \dots + q'_{rr}(u^r)^2,$$

используя лишь основное преобразование метода Лагранжа. Тогда $\det Q_k = \det Q'_k = q'_{11} \dots q'_{kk}$ при $k \leq r$, т.е. $q'_{kk} = \frac{\det Q_k}{\det Q_{k-1}}$, что и требовалось. ■

Про приведение квадратичных форм к каноническому виду ЕСТЬ ФАЙЛ Антона Александровича Клячко.

Ещё Антон Александрович рассказывал на семинаре, что можно ещё и так:

Задача 8 (А. А. Клячко). Как модифицировать метод Якоби, чтобы обойти проблему с нулём на диагонали?

Теперь у нас метод не работает только если на диагонали будут стоять два нуля подряд. Оказывается, можно и эту проблему решить:

Задача 9 (А. А. Клячко). Как модифицировать метод в предыдущей задаче, чтобы он работал с двумя нулями подряд на диагонали?

И даже это не предел:

Задача 10 (А. А. Клячко). Как модифицировать метод в предыдущей задаче, чтобы он работал с тремя нулями подряд на диагонали?

А вот это уже предел, дальше нельзя.

Задача 11 (А. А. Клячко). Доказать, что проблема с четырьмя нулями на диагонали не решается.

Не уверен, что решения к этим задачам когда-нибудь здесь появятся, но я очень постараюсь.

29. ПОЛОЖИТЕЛЬНО И ОТРИЦАТЕЛЬНО ОПРЕДЕЛЁННЫЕ ВЕЩЕСТВЕННЫЕ КВАДРАТИЧНЫЕ ФОРМЫ. КРИТЕРИЙ СИЛЬВЕСТРА

Определение 1. Симметрическая билинейная функция \mathcal{B} называется *положительно определённой*, если $\mathcal{B}(x, x) > 0$ при $x \neq 0$. Соответствующая квадратичная форма $Q(x)$ удовлетворяет условию $Q(x) > 0$ при $x \neq 0$ и также называется *положительно определённой*.

Положительно определённая симметрическая билинейная функция задаёт в пространстве V скалярное произведение, т.е. превращает V в евклидово пространство.

Теорема 1 (Критерий Сильвестра). Симметрическая билинейная функция (квадратичная форма) положительно определена тогда и только тогда, когда все угловые миноры её матрицы в некотором базисе положительны.

Доказательство. Пусть все угловые миноры $\det Q_k$ матрицы квадратичной формы $Q(x)$ положительны. Тогда в силу теоремы Якоби квадратичная форма приводится к виду $Q(u) = q'_{11}(u^1)^2 + \dots + q'_{nn}(u^n)^2$, где $n = \text{rk } Q = \dim V$, а $q'_{kk} = \frac{\det Q_k}{\det Q_{k-1}} > 0$. Такая квадратичная форма положительно определена, т.к. $Q(u) > 0$ при $u \neq 0$.

Обратно, пусть $Q(x)$ положительно определена. Т.к. в любом базисе мы имеем $q_{ii} = \mathcal{B}(e_i, e_i) > 0$, всегда применимо основное преобразование метода Лагранжа. Тогда последовательно применяя основное преобразование, мы приведём квадратичную форму к виду $Q(u) = q'_{11}(u^1)^2 + \dots + q'_{nn}(u^n)^2$, где $q'_{ii} > 0$. Следовательно, $\det Q_k = \det Q'_k = q'_{11} \dots q'_{kk} > 0$ для любого k . ■

Примечание. С аналогичным доказательством, симметрическая билинейная функция отрицательно определена тогда и только тогда, когда последовательность угловых миноров её матрицы в некотором базисе знакопеременная и $\det Q_1 < 0$.

30. ЕВКЛИДОВО ПРОСТРАНСТВО. НЕРАВЕНСТВО КОШИ — БУНЯКОВСКОГО. УГОЛ МЕЖДУ ВЕКТОРАМИ И ДЛИНА ВЕКТОРА В ЕВКЛИДОВОМ ПРОСТРАНСТВЕ. НЕРАВЕНСТВО ТРЕУГОЛЬНИКА

Здесь я сразу напишу и про эрмитовы пространства, чтобы лишний раз не дублировать записи.

Определение 1. Линейное пространство над полем \mathbb{R} называется *евклидовым*, если на парах его векторов определена функция $f : V \times V \rightarrow \mathbb{R}$ (обозначаемая $(a, b) := f(a, b)$ и называемая *скалярным произведением*), удовлетворяющая следующим аксиомам:

1. *билинейность*:

$$(\lambda_1 u_1 + \lambda_2 u_2, v) = \lambda_1(u_1, v) + \lambda_2(u_2, v), \quad (u, \mu_1 v_1 + \mu_2 v_2) = \mu_1(u, v_1) + \mu_2(u, v_2)$$

$$\forall \lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{R}, \forall u, u_1, u_2, v, v_1, v_2 \in V;$$

2. *симметричность*: $(u, v) = (v, u) \forall u, v \in V$;

3. *положительная определённость*: $(v, v) \geq 0 \forall v \in V$, причём $(v, v) = 0$ только при $v = \mathbf{0}$.

Заметим, что из симметричности и линейности по одному аргументу сразу следует линейность и по другому аргументу.

Примечание. В комплексном пространстве функции с такими свойствами не бывает. Действительно, если (v, v) — положительное вещественное число, то $(iv, iv) = i^2(v, v) = -(v, v)$ — отрицательное. Таким образом, функция $V \times V \rightarrow \mathbb{R}$ (V — линейное пространство над \mathbb{C}) не может быть билинейной и положительно определённой.

Определение 2. Линейное пространство над полем \mathbb{C} называется *эрмитовым* (или *унитарным*), если на парах его векторов определена функция $f : V \times V \rightarrow \mathbb{C}$ (обозначаемая $(a, b) := f(a, b)$ и называемая *скалярным произведением*), удовлетворяющая следующим свойствам:

1. *полуторалинейность*:

$$(\lambda_1 u_1 + \lambda_2 u_2, v) = \overline{\lambda_1}(u_1, v) + \overline{\lambda_2}(u_2, v), \quad (u, \mu_1 v_1 + \mu_2 v_2) = \mu_1(u, v_1) + \mu_2(u, v_2)$$

$$\forall \lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{C}, \forall u, u_1, u_2, v, v_1, v_2 \in V;$$

2. *эрмитовость*: $(u, v) = \overline{(v, u)} \forall u, v \in V$; в частности $(v, v) \in \mathbb{R} \forall v \in V$;

3. *положительная определённость*: $(v, v) \geq 0 \forall v \in V$, причём $(v, v) = 0$ только при $v = \mathbf{0}$.

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

Пример 1.

1. Скалярное произведение векторов $u = (u^1, \dots, u^n)$ и $v = (v^1, \dots, v^n)$ в пространстве \mathbb{R}^n можно задать формулой

$$(u, v) := u^1 v^1 + u^2 v^2 + \dots + u^n v^n,$$

а скалярное произведение в \mathbb{C}^n — формулой

$$(u, v) := \overline{u^1} v^1 + \overline{u^2} v^2 + \dots + \overline{u^n} v^n.$$

Это называется *стандартным* скалярным произведением.

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

$$(A, B) := \text{tr}(\overline{A}^t B).$$

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

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

Определение 3. Пусть V — евклидово или эрмитово пространство. Для $v \in V$ величина $\sqrt{(v, v)}$ называется *длиной* вектора v и обозначается $|v|$. Векторы $u, v \in V$ такие, что $(u, v) = 0$, называются *ортогональными*, обозначается $u \perp v$.

Предложение 1. Пусть u — ненулевой вектор евклидова или эрмитова пространства V . Тогда для любого вектора $v \in V$ существует единственное разложение $v = v_1 + v_2$, где вектор v_1 коллинеарен вектору u , а вектор v_2 ортогонален u .

Доказательство. Сначала докажем единственность. Пусть $v = v_1 + v_2$ — такое разложение. Тогда имеем $v_1 = \lambda u$, $v_2 = v - \lambda u$. Условие $u \perp v_2$ влечёт

$$0 = (u, v_2) = (u, v - \lambda u) = (u, v) - \lambda(u, u).$$

Отсюда $\lambda = (u, v)/(u, u)$ и

$$v_1 = \frac{(u, v)}{(u, u)}u. \quad (*)$$

Тем самым векторы v_1 и $v_2 = v - v_1$ определены однозначно. С другой стороны, определив v_1 по этой формуле, мы получим $v_2 = (v - v_1) \perp u$. ■

Определение 4. Вектор $(*)$ называется *ортогональной проекцией* вектора v на направление вектора u и обозначается $\text{pr}_u v$, а вектор $v - \text{pr}_u v$ называется *ортогональной составляющей* вектора v относительно u и обозначается $\text{ort}_u v$.

Теорема 1 (Неравенство Коши — Буняковского). Для любых двух векторов u, v евклидова или эрмитова пространства имеет место неравенство

$$|(u, v)| \leq |u| \cdot |v|,$$

причём равенство имеет место в том и только в том случае, когда u и v коллинеарны.

Доказательство. Если $u = 0$, утверждение очевидно. Пусть $u \neq 0$. Запишем $v = v_1 + v_2$, где $v_1 = \text{pr}_u v$ и $v_2 = \text{ort}_u v$. Тогда $(v_1, v_2) = 0$, и мы имеем

$$|v|^2 = (v, v) = (v_1 + v_2, v_1 + v_2) = (v_1, v_1) + (v_1, v_2) + (v_2, v_1) + (v_2, v_2) = |v_1|^2 + |v_2|^2.$$

Отсюда $|v_1| \leq |v|$, причём равенство достигается только при $v_2 = 0$, т. е. когда вектор v коллинеарен вектору u . Осталось заметить, что $|v_1| = |\text{pr}_u v| = \frac{|(u, v)|}{|u|}$, так что неравенство $|v_1| \leq |v|$ эквивалентно требуемому. ■

Определение 5. Углом между двумя ненулевыми векторами u, v евклидова пространства называется величина

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

Неравенство Коши — Буняковского гарантирует, что угол между ненулевыми векторами всегда определён.

Следствие 1 (Неравенство треугольника). Для любых двух векторов u, v евклидова или эрмитова пространства выполнено неравенство

$$|u + v| \leq |u| + |v|.$$

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

$$(u + v, u + v) \leq (u, u) + (v, v) + 2|u||v|.$$

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

$$(u, v) + (v, u) \leq 2|u||v|,$$

которое следует из неравенства Коши — Буняковского. ■

31. ВЫЧИСЛЕНИЕ СКАЛЯРНОГО ПРОИЗВЕДЕНИЯ В КООРДИНАТАХ. МАТРИЦА ГРАМА И ЕЁ СВОЙСТВА

Определение 1. Матрицей Грама системы векторов a_1, \dots, a_k называется матрица

$$G = G(a_1, \dots, a_k) := \begin{pmatrix} (a_1, a_1) & (a_1, a_2) & \cdots & (a_1, a_k) \\ (a_2, a_1) & (a_2, a_2) & \cdots & (a_2, a_k) \\ \vdots & \vdots & \ddots & \vdots \\ (a_k, a_1) & (a_k, a_2) & \cdots & (a_k, a_k) \end{pmatrix}.$$

Пусть в евклидовом пространстве в некотором базисе заданы два вектора $u = (u^1, \dots, u^n)$ и $v = (v^1, \dots, v^n)$. Тогда

$$(u, v) = (u^i e_i, v^j e_j) = u^i v^j (e_i, e_j) = u^i v^j g_{ij} = (u^1, \dots, u^n) \cdot G \cdot (v^1, \dots, v^n)^t,$$

где $G = G(e_1, \dots, e_n)$ — матрица Грама базисных векторов.

Определение 2. Часто матрицу $G = G(e_1, \dots, e_n)$ называют *матрицей Грама скалярного произведения* (\cdot, \cdot) .

Теорема 1. Матрица G является матрицей Грама линейно независимой системы векторов в евклидовом пространстве тогда и только тогда, когда она

1. симметричная;
2. положительно определённая.

Доказательство. \Rightarrow . Если $G = (g_{ij})$ — матрица Грама скалярного произведения (\cdot, \cdot) , то $g_{ij} = (e_i, e_j) = (e_j, e_i) = g_{ji}$, т. е. $G = G^t$. Так же, $(x, x) = x^t G x > 0 \ \forall x \in V$, поэтому G задаёт положительно определённую квадратичную форму, следовательно по критерию Сильвестра все её угловые миноры положительные.

\Leftarrow . Если G симметричная и положительно определённая, то она задаёт функцию $f : V \times V \mapsto x^t G y$, удовлетворяющую аксиомам скалярного произведения, значит, G — матрица Грама этого скалярного произведения. ■

Предложение 1. Пусть G — матрица Грама системы векторов a_1, \dots, a_k , а $A = (a_j^i)$ — матрица, в столбцы которой записаны координаты векторов a_1, \dots, a_k в некотором ортонормированном базисе. Тогда имеет место соотношение $G = \overline{A}^t A$ ($G = A^t A$ в евклидовом пространстве).

Доказательство. Это вытекает из закона умножения матриц и формулы для скалярного произведения в ортонормированном базисе. ■

**32. ОРТОГОНАЛЬНОСТЬ ВЕКТОРОВ. ЛИНЕЙНАЯ НЕЗАВИСИМОСТЬ СИСТЕМЫ НЕНУЛЕВЫХ
ОРТОГОНАЛЬНЫХ ВЕКТОРОВ. ОРТОГОНАЛЬНЫЙ И ОРТОНОРМИРОВАННЫЙ БАЗИС.
АЛГОРИТМ ОРТОГОНАЛИЗАЦИИ ГРАМА — ШМИДТА**

Предложение 1. Пусть v_1, \dots, v_k — набор попарно ортогональных ненулевых векторов евклидова или эрмитова пространства. Тогда эти векторы линейно независимы.

Доказательство. Пусть некоторая линейная комбинация данных векторов равна нулю:

$$\sum_{i=1}^k \lambda_i v_i = \mathbf{0}.$$

Умножим обе части этого равенства скалярно на v_j и воспользуемся линейностью скалярного произведения по второму аргументу:

$$0 = \left(v_j, \sum_{i=1}^k \lambda_i v_i \right) = \sum_{i=1}^k \lambda_i (v_j, v_i) = \lambda_j (v_j, v_j),$$

т. к. по условию остальные слагаемые в этой сумме равны нулю. Поскольку $v_j \neq \mathbf{0}$, из положительной определённости скалярного произведения следует, что $(v_j, v_j) \neq 0$, а значит, $\lambda_j = 0$. Это выполнено $\forall j = 1, \dots, k$, следовательно, линейная комбинация $\sum_{i=1}^k \lambda_i v_i$ тривиальна. ■

Определение 1. Базис e_1, \dots, e_n евклидова или эрмитова пространства называется *ортогональным*, если его векторы попарно ортогональны. Если при этом длина каждого вектора равна 1, то базис называется *ортонормированным*.

Теорема 1. Пусть a_1, \dots, a_k — набор линейно независимых векторов пространства V . Тогда существует такой набор попарно ортогональных векторов b_1, \dots, b_k , что для каждого $i = 1, \dots, k$ линейная оболочка $\langle b_1, \dots, b_i \rangle$ совпадает с $\langle a_1, \dots, a_i \rangle$.

Доказательство. При $k = 1$ утверждение очевидно: можно взять $b_1 := a_1$. Предположим, что утверждение верно для наборов из i векторов, и докажем его для наборов из $i + 1$ вектора. Пусть b_1, \dots, b_i — ортогональный набор, построенный по набору a_1, \dots, a_i . Мы хотим, чтобы для нового вектора b_{i+1} линейная оболочка $\langle b_1, \dots, b_i, b_{i+1} \rangle$ совпадала с $\langle a_1, \dots, a_i, a_{i+1} \rangle = \langle b_1, \dots, b_i, a_{i+1} \rangle$, и поэтому будем искать b_{i+1} в виде

$$b_{i+1} = a_{i+1} + \lambda_1 b_1 + \dots + \lambda_i b_i.$$

Коэффициенты $\lambda_1, \dots, \lambda_i$ будем подбирать так, чтобы вектор b_{i+1} был ортогонален всем предыдущим векторам b_1, \dots, b_i . Умножив скалярно предыдущее равенство слева на b_j и используя то, что $(b_i, b_\ell) = 0$ при $j \neq \ell$, получаем

$$0 = (b_j, b_{i+1}) = (b_j, a_{i+1}) + \lambda_j (b_j, b_j),$$

откуда $\lambda_j = -\frac{(b_j, a_{i+1})}{(b_j, b_j)} \forall j = 1, \dots, i$. Окончательно для вектора b_{i+1} получаем

$$\begin{aligned} b_{i+1} &= a_{i+1} - \frac{(b_1, a_{i+1})}{(b_1, b_1)} b_1 - \frac{(b_2, a_{i+1})}{(b_2, b_2)} b_2 - \dots - \frac{(b_i, a_{i+1})}{(b_i, b_i)} b_i = \\ &= a_{i+1} - \text{pr}_{b_1} a_{i+1} - \text{pr}_{b_2} a_{i+1} - \dots - \text{pr}_{b_i} a_{i+1}. \end{aligned}$$

■

Определение 2. Индуктивная процедура перехода от набора a_1, \dots, a_k к ортогональному набору b_1, \dots, b_k называется *процессом ортогонализации Грама — Шмидта*.

Условие $\langle b_1, \dots, b_i \rangle = \langle a_1, \dots, a_i \rangle$ при $i = 1, \dots, k$ означает, что матрица перехода от a_1, \dots, a_k к b_1, \dots, b_k является верхнетреугольной.

Следствие 1. В евклидовом или эрмитовом пространстве V существуют ортонормированные базисы.

Доказательство. Возьмём произвольные базис и ортогонализируем его методом Грама — Шмидта, получив при этом новый базис b_1, \dots, b_n . Тогда базис, состоящий из векторов $\frac{b_1}{|b_1|}, \dots, \frac{b_n}{|b_n|}$ будет ортонормированным. ■

Предложение 2. Пусть векторы u и v имеют координаты u^1, \dots, u^n и v^1, \dots, v^n в некотором ортонормированном базисе евклидова или эрмитова пространства V . Тогда их скалярное произведение вычисляется по формуле

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

Доказательство. Пусть e_1, \dots, e_n — ортонормированный базис. Тогда

$$(u, v) = (u^i e_i, v^j e_j) = \overline{u^i}v^j (e_i, e_j) = \overline{u^i}v^j \delta_{ij} = \overline{u^1}v^1 + \overline{u^2}v^2 + \dots + \overline{u^n}v^n.$$

■

33. ОРТОГОНАЛЬНЫЕ МАТРИЦЫ КАК МАТРИЦЫ ПЕРЕХОДА ОТ ОДНОГО ОРТОНОРМИРОВАННОГО БАЗИСА К ДРУГОМУ. ГРУППА $O(n)$ ОРТОГОНАЛЬНЫХ МАТРИЦ ПОРЯДКА n .

Определение 1. Матрица перехода от одного ортонормированного базиса евклидова (соответственно, эрмитова) пространства к ортонормированному базису называется *ортогональной* (соответственно, *унитарной*).

Предложение 1. Следующие условия эквивалентны:

1. Матрица C ортогональна (соответственно, унитарна);
2. $C^t C = E$ (соответственно, $\overline{C}^t C = E$);
3. Столбцы матрицы C образуют ортонормированный базис \mathbb{R}^n (соответственно, \mathbb{C}^n);
4. $CC^t = E$ (соответственно, $C\overline{C}^t = E$);
5. Строки матрицы C образуют ортонормированный базис \mathbb{R}^n (соответственно, \mathbb{C}^n).

Доказательство. Условия 2 и 5 эквивалентны, т. к. каждое из них эквивалентно равенству $C^t = C^{-1}$ (соответственно, $\overline{C}^t = C^{-1}$). Эквивалентности $2 \Leftrightarrow 3$ и $4 \Leftrightarrow 5$ вытекают из правила умножения матриц и формул скалярного произведения в ортонормированных базисах в \mathbb{R}^n и \mathbb{C}^n .

Докажем импликацию $1 \Rightarrow 2$. Пусть e_1, \dots, e_n и $e_{1'}, \dots, e_{n'}$ — два ортонормированных базиса в эрмитовом пространстве и $C = (c_{ij}^i)$ — матрица перехода, т. е. $e_{i'} = c_{ij}^i e_i$. Тогда $(e_i, e_j) = \delta_{ij}$ и $(e_{i'}, e_{j'}) = \delta_{i'j'}$, откуда

$$\delta_{i'j'} = (e_{i'}, e_{j'}) = (c_{ij}^i e_i, c_{j'k}^j e_k) = \overline{c_{ij}^i} c_{j'k}^j (e_i, e_k) = \overline{c_{ij}^i} c_{j'k}^j \delta_{ik} = \overline{c_{ij}^i} c_{j'i}^j.$$

Согласно правилу умножения матриц, это эквивалентно матричному соотношению $E = \overline{C}^t C$ или $\overline{C}^t C = E$. Случай евклидова пространства рассматривается аналогично (убрать чёрточки над буквами).

Осталось доказать импликацию $2 \Rightarrow 1$. Пусть имеет место тождество $\overline{C}^t C = E$ или, в обозначениях Эйнштейна, $\overline{c}_{ij}^i \delta_{ij} c_{j'k}^j$. Возьмём произвольный ортонормированный базис e_1, \dots, e_n . Из соотношения $\overline{C}^t C = E$ вытекает, что матрица C невырождена, и поэтому можно рассмотреть новый базис

$e_{1'}, \dots, e_{n'}$, где $e_{i'} := c_{i'}^i e_i$. Тогда аналогично выкладке выше мы получаем $(e_{i'}, e_{j'}) = \overline{c_{i'}^i} \delta_{ij} c_{j'}^j = \delta_{i'j'}$, т. е. базис $e_{1'}, \dots, e_{n'}$ тоже ортонормирован, и C — матрица перехода от одного ортонормированного базиса к другому. ■

Определение 2. Подгруппа ортогональных матриц порядка n в группе $\text{Mat}_{n \times n}(\mathbb{R})$ обозначается через $O(n)$.

34. ОРТОГОНАЛЬНОЕ ДОПОЛНЕНИЕ. РАЗЛОЖЕНИЕ ЕВКЛИДОВА ПРОСТРАНСТВА В ПРЯМУЮ СУММУ ПОДПРОСТРАНСТВА И ОРТОГОНАЛЬНОГО ДОПОЛНЕНИЯ К НЕМУ. ОРТОГОНАЛЬНАЯ ПРОЕКЦИЯ

Определение 1. Пусть $W \subset V$ — подпространство евклидова или эрмитова пространства V . *Ортогональным дополнением* к W называется множество W^\perp , состоящее из векторов, ортогональных всем векторам из W , т. е.

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

Легко видеть, что ортогональное дополнение является подпространством.

Предложение 1. Для любого подпространства $W \subset V$ имеет место разложение $V = W \oplus W^\perp$.

Доказательство. Пусть a_1, \dots, a_k — базис в W , дополним его до базиса всего пространства V векторами a_{k+1}, \dots, a_n . Применив ортогонализацию Грама — Шмидта, получим ортогональный базис $b_1, \dots, b_k, b_{k+1}, b_n$ в V , причём его первые k векторов будут базисом в W , т. к. $\langle b_1, \dots, b_k \rangle = \langle a_1, \dots, a_k \rangle = W$. В то же время, b_{k+1}, \dots, b_n лежат в W^\perp по определению ортогонального дополнения. Итак, $\forall v \in V$ мы имеем разложение по базису

$$v = \underbrace{\lambda_1 b_1 + \dots + \lambda_k b_k}_{\in W} + \underbrace{\lambda_{k+1} b_{k+1} + \dots + \lambda_n b_n}_{\in W^\perp}$$

t. e. $V = W + W^\perp$.

Осталось доказать, что эта сумма прямая. Пусть $v \in W \cap W^\perp$. Т. к. $v \in W^\perp$, мы имеем $(v, w) = 0 \forall w \in W$. Т. к. $v \in W$, можно взять $w := v$. Тогда $(v, v) = 0$, т. е. $v = \mathbf{0}$, и наша сумма прямая. ■

Определение 2. Пусть $W \subset V$ — подпространство евклидова или эрмитова пространства. Для произвольного вектора $v \in V$ запишем разложение $v = v_1 + v_2$, где $v_1 \in W$, $v_2 \in W^\perp$. Тогда вектор v_1 называется *ортogonalной проекцией* вектора v на подпространство W и обозначается $\text{pr}_W v$, а вектор $v_2 = v - \text{pr}_W v$ называется *ортogonalной составляющей* вектора v относительно подпространства W и обозначается $\text{ort}_W v$.

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

Предложение 2. Пусть подпространство $W \subset V$ задано как линейная оболочка системы векторов: $W = \langle a_1, \dots, a_k \rangle$. Тогда проекция вектора $v \in V$ на W есть линейная комбинация $\text{pr}_W v = \lambda_1 a_1 + \dots + \lambda_k a_k$, коэффициенты которой находятся из системы линейных уравнений

[illegible]

Доказательство. Запишем $v = \text{pr}_W v + \text{ort}_W v$. Тогда $\text{ort}_W v = v - \lambda_1 a_1 - \dots - \lambda_k a_k$ ортогонален каждому из векторов a_1, \dots, a_k . Взяв скалярное произведение a_i с $\text{ort}_W v$, мы получаем

$$(a_i, v - \lambda_1 a_1 - \dots - \lambda_k a_k) = 0,$$

что эквивалентно i -му уравнению системы.

35. ТЕОРЕМА ПИФАГОРА. УГОЛ И РАССТОЯНИЕ МЕЖДУ ВЕКТОРОМ И ПОДПРОСТРАНСТВОМ. ОБЪЁМ n -МЕРНОГО ПАРАЛЛЕЛЕПИПЕДА

Теорема 1 (Пифагор). Если два вектора u и v евклидова пространства V ортогональны, то $|u + v|^2 = |u|^2 + |v|^2$.

Доказательство. Действительно,

$$|u + v|^2 = (u + v, u + v) = (u, u) + 2 \cdot \underbrace{(u, v)}_{=0} + (v, v) = |u|^2 + |v|^2.$$

■

Примечание. По индукции легко доказать $\forall k \in \mathbb{N}$, что если вектора v_1, \dots, v_k попарно ортогональны, то

$$|v_1 + \dots + v_k|^2 = |v_1|^2 + \dots + |v_k|^2.$$

Определение 1. Пусть V — евклидово пространство. Углом между ненулевым вектором $v \in V$ и подпространством $W \subset V$ называется точная нижняя грань углов между v и произвольным вектором $w \in W$:

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

Точная нижняя грань $\inf_{w \in W} \angle(v, w)$ существует, т.к. множество углов $\angle(v, w)$ ограничено снизу нулём. На самом деле, она достигается на векторе $w = \text{pr}_W v$, как показано в следующем утверждении.

Предложение 1. $\angle(v, W) = \angle(v, \text{pr}_W v)$.

Доказательство. Пусть $w \in W$ — произвольный ненулевой вектор. Обозначим $\alpha := \angle(v, \text{pr}_W v)$, $\beta := \angle(v, w)$ и $v_1 = \text{pr}_W v$. Необходимо показать, что $\alpha \leq \beta$. Т.к. $0 \leq \alpha, \beta \leq \pi$, то это требуемое неравенство равносильно $\cos \alpha \geq \cos \beta$, т.е.

$$\frac{(v, v_1)}{|v| |v_1|} \geq \frac{(v, w)}{|v| |w|}.$$

Запишем $v = v_1 + v_2$, где $v_2 = \text{ort}_W v \in W^\perp$. Тогда $(v, v_1) = (v_1 + v_2, v_1) = |v_1|^2$ и $(v, w) = (v_1 + v_2, w) = (v_1, w)$. Подставив это в последнее неравенство, получим $|v_1| \geq \frac{(v_1, w)}{|w|}$, что вытекает из неравенства Коши — Буняковского. ■

Определение 2. Определим *расстояние между векторами* u и v метрического пространства V по формуле

$$\rho(u, v) := |u - v|.$$

Легко проверить, что определённая нами функция удовлетворяет всем аксиомам метрики (неравенство треугольника легко получить из доказанного выше следствия неравенства Коши — Буняковского).

Определение 3. *Расстояние между вектором v и подпространством W* евклидова пространства определяется по формуле

$$\rho(v, W) := \inf_{w \in W} \rho(v, w).$$

Опять же, точная нижняя грань существует, т.к. множество расстояний $\rho(v, w)$ ограничено снизу нулём. И достигается она опять на векторе $\text{pr}_W v$.

Предложение 2. $\rho(v, W) = \rho(v, \text{pr}_W v) = |\text{ort}_W v|$.

Доказательство. Представим v в виде $v = v_1 + v_2$, где $v_1 = \text{pr}_W v$, $v_2 = \text{ort}_W v$. Для произвольного вектора $w \in W$ имеем

$$\rho(v, w)^2 = |v - w|^2 = |(v_1 + v_2) - w|^2 = |(v_1 - w) + v_2|^2.$$

Т. к. векторы $v_1 - w \in W$ и $v_2 \in W^\perp$ ортогональны, то по теореме Пифагора получаем

$$|(v_1 - w) + v_2|^2 = |v_1 - w|^2 + |v_2|^2 \geq |v_2|^2 = \rho^2(v, \text{pr}_W v).$$

■

Определение 4. Будем называть k -мерным параллелепипедом, натянутым на вектора a_1, \dots, a_k евклидова пространства, будем называть множество

$$\Pi(a_1, \dots, a_k) := \left\{ \sum_{i=1}^n x^i a_i : 0 \leq x^i \leq 1 \right\}.$$

Основанием этого k -мерного параллелепипеда будем называть $(k-1)$ -мерный параллелепипед $\Pi(a_1, \dots, a_{k-1})$, а его высотой будем называть длину вектора $\text{ort}_{\langle a_1, \dots, a_{k-1} \rangle} a_k$.

Определение 5. Определим k -мерный объём Vol_k параллелепипеда $\Pi(a_1, \dots, a_k)$ в евклидовом пространстве индуктивно:

1. Одномерный объём $\text{Vol}_1 \Pi(a_1) := |a_1|$ — это длина вектора;
2. $\text{Vol}_k \Pi(a_1, \dots, a_k) := \text{Vol}_{k-1} \Pi(a_1, \dots, a_{k-1}) \cdot |\text{ort}_{\langle a_1, \dots, a_{k-1} \rangle} a_k|$.

Теорема 2. $(\text{Vol}_k \Pi(a_1, \dots, a_k))^2 = \det G(a_1, \dots, a_k)$.

Доказательство. Индукция по k . При $k = 1$, очевидно, $|a_1|^2 = (a_1, a_1)$.

Пусть утверждение доказано для Vol_{k-1} , докажем его и для Vol_k . Рассмотрим разложение $a_k = \text{pr}_{\langle a_1, \dots, a_{k-1} \rangle} a_k + \text{ort}_{\langle a_1, \dots, a_{k-1} \rangle} a_k$, где $\text{pr}_{\langle a_1, \dots, a_{k-1} \rangle} a_k = \lambda_1 a_1 + \dots + \lambda_{k-1} a_{k-1}$, и обозначим $h := \text{ort}_{\langle a_1, \dots, a_{k-1} \rangle} a_k$. Тогда $(a_i, h) = 0$ при $i = 1, \dots, k-1$ и $(a_k, h) = (h, h)$. Мы имеем

$$\begin{aligned} \det G &= \det \begin{pmatrix} (a_1, a_1) & \dots & (a_1, a_k) \\ \vdots & \ddots & \vdots \\ (a_k, a_1) & \dots & (a_k, a_k) \end{pmatrix} = \det \begin{pmatrix} (a_1, a_1) & \dots & (a_1, a_{k-1}) & (a_1, \lambda_1 a_1 + \dots + \lambda_{k-1} a_{k-1} + h) \\ \vdots & \ddots & \vdots & \vdots \\ (a_k, a_1) & \dots & (a_k, a_{k-1}) & (a_k, \lambda_1 a_1 + \dots + \lambda_{k-1} a_{k-1} + h) \end{pmatrix} = \\ &= \lambda_1 \det \underbrace{\begin{pmatrix} (a_1, a_1) & \dots & (a_1, a_{k-1}) & (a_1, a_1) \\ \vdots & \ddots & \vdots & \vdots \\ (a_k, a_1) & \dots & (a_k, a_{k-1}) & (a_k, a_1) \end{pmatrix}}_{=0} + \dots + \lambda_{k-1} \det \underbrace{\begin{pmatrix} (a_1, a_1) & \dots & (a_1, a_{k-1}) & (a_1, a_{k-1}) \\ \vdots & \ddots & \vdots & \vdots \\ (a_k, a_1) & \dots & (a_k, a_{k-1}) & (a_k, a_{k-1}) \end{pmatrix}}_{=0} + \\ &+ \det \begin{pmatrix} (a_1, a_1) & \dots & (a_1, a_{k-1}) & (a_1, h) \\ \vdots & \ddots & \vdots & \vdots \\ (a_{k-1}, a_1) & \dots & (a_{k-1}, a_{k-1}) & (a_{k-1}, h) \\ (a_k, a_1) & \dots & (a_k, a_{k-1}) & (a_k, h) \end{pmatrix} = \det \begin{pmatrix} (a_1, a_1) & \dots & (a_1, a_{k-1}) & 0 \\ \vdots & \ddots & \vdots & \vdots \\ (a_{k-1}, a_1) & \dots & (a_{k-1}, a_{k-1}) & 0 \\ (a_k, a_1) & \dots & (a_k, a_{k-1}) & (h, h) \end{pmatrix} = \\ &= \begin{pmatrix} (a_1, a_1) & \dots & (a_1, a_{k-1}) \\ \vdots & \ddots & \vdots \\ (a_{k-1}, a_1) & \dots & (a_{k-1}, a_{k-1}) \end{pmatrix} \cdot (h, h) = (\text{Vol}_{k-1} \Pi(a_1, \dots, a_{k-1}))^2 |h|^2 = (\text{Vol}_k \Pi(a_1, \dots, a_k))^2. \end{aligned}$$

■

Следствие 1. Векторы a_1, \dots, a_k линейно зависимы, если и только если $\det G(a_1, \dots, a_k) = 0$.

Доказательство. \Rightarrow . Действительно, предположим, что векторы a_1, \dots, a_k линейно зависимы. Можно считать, что a_k линейно выражается через a_1, \dots, a_{k-1} . Тогда $\text{ort}_{\langle a_1, \dots, a_{k-1} \rangle} a_k = \mathbf{0}$ и, следовательно,

$$\det G = (\text{Vol}_k \Pi(a_1, \dots, a_k))^2 = (\text{Vol}_{k-1} \Pi(a_1, \dots, a_{k-1}))^2 |\text{ort}_{\langle a_1, \dots, a_{k-1} \rangle} a_k|^2 = 0.$$

\Leftarrow . Обратно, пусть $\det G = (\text{Vol}_k \Pi(a_1, \dots, a_k))^2 = 0$. Тогда, в силу индуктивного определения объёма, мы имеем $\text{Vol}_i \Pi(a_1, \dots, a_i) = 0$, а $\text{Vol}_{i-1} \Pi(a_1, \dots, a_{i-1}) \neq 0$ для некоторого i . Т.к. $\text{Vol}_i \Pi(a_1, \dots, a_i) = \text{Vol}_{i-1} \Pi(a_1, \dots, a_{i-1}) \cdot |\text{ort}_{\langle a_1, \dots, a_{i-1} \rangle} a_i|$, это означает, что $\text{ort}_{\langle a_1, \dots, a_{i-1} \rangle} a_i = \mathbf{0}$, т.е. a_i линейно выражается через a_1, \dots, a_{i-1} . ■

Следствие 2 (О геометрическом смысле определителя). Пусть $\dim V = n$ и $A = (a_j^i)$ — квадратная матрица из координат векторов a_1, \dots, a_n в некотором ортонормированном базисе. Тогда

$$\text{Vol}_n \Pi(a_1, \dots, a_n) = |\det A|.$$

Доказательство. Из предложения 1 в вопросе 31 получаем

$$(\text{Vol}_n \Pi(a_1, \dots, a_n))^2 = \det G = \det(A^t A) = (\det A)^2.$$

■

36. ИЗОМОРФИЗМ ЕВКЛИДОВЫХ ПРОСТРАНСТВ ОДИНАКОВОЙ РАЗМЕРНОСТИ. ИЗОМОРФИЗМ ЕВКЛИДОВА ПРОСТРАНСТВА И ЕГО СОПРЯЖЁННОГО

Определение 1. Два евклидовых или эрмитовых пространства V и W называются *изоморфными*, если существует изоморфизм линейных пространств $\mathcal{A} : V \rightarrow W$, сохраняющий скалярное произведение, т.е. $(\mathcal{A}u, \mathcal{A}v) = (u, v) \forall u, v \in V$.

Предложение 1. Два евклидовых или эрмитовых пространства V и W изоморфны тогда и только тогда, когда их размерности совпадают.

Доказательство. Если V и W изоморфны как евклидовы (эрмитовы) пространства, то они изоморфны как линейные пространства, а потому $\dim V = \dim W$.

Доказательство обратного утверждения аналогично доказательству соответствующего утверждения для линейных пространств: изоморфизм между евклидовыми (эрмитовыми) пространствами размерности n устанавливается при помощи биекции между базисами e_1, \dots, e_n в V и f_1, \dots, f_n в W . Для того, чтобы получаемый изоморфизм линейных пространств $\mathcal{A} : V \rightarrow W$ сохранял скалярное произведение, базисы необходимо выбрать ортонормированными. В этом случае мы имеем $f_i = \mathcal{A}e_i$ и $(e_i, e_j) = (f_i, f_j) = \delta_{ij}$. Поэтому для любых векторов $u = u^i e_i$ и $v = v^j e_j$ мы имеем

$$(\mathcal{A}u, \mathcal{A}v) = \overline{u^i} v^j (\mathcal{A}e_i, \mathcal{A}e_j) = \overline{u^i} v^j (f_i, f_j) = \overline{u^i} v^j \delta_{ij} = \sum_{i=1}^n \overline{u^i} v^i = (u, v),$$

т.е. изоморфизм \mathcal{A} сохраняет скалярное произведение. ■

Т.к. пространства V и V^* имеют одну размерность (в конечномерном случае), то они изоморфны. Однако построение изоморфизма между ними требует выбора базисов и в этом смысле неканонично. Оказывается, что в присутствии скалярного произведения можно установить изоморфизм $V \rightarrow V^*$ каноничным образом, т.е. не прибегая к выбору базисов.

Пусть V — евклидово пространство. Каждому вектору $u \in V$ сопоставим линейную функцию $\xi_u := (u, \cdot)$.

Теорема 1. Пусть V — евклидово пространство. Отображение $u \mapsto \xi_u$ устанавливает канонически изоморфизм $\mathcal{A} : V \rightarrow V^*$.

Доказательство. Линейность отображения $u \mapsto \xi_u$ вытекает из линейности скалярного произведения по первому аргументу. Т.к. $\dim V = \dim V^*$, чтобы проверить, что $\mathcal{A} : V \rightarrow V^*$ — изоморфизм, достаточно проверить, что $\text{Ker } \mathcal{A} = \{\mathbf{0}\}$. Пусть $v \in \text{Ker } \mathcal{A}$, т.е. $\mathcal{A}v = \xi_v = 0$. Тогда $\xi_v(w) = (v, w) = 0 \forall w \in W$. Но тогда и $(v, v) = 0$, значит, $v = 0$ и ядро отображения \mathcal{A} нулевое. ■

Аналогичным образом для эрмитова пространства V устанавливается канонический изоморфизм $\bar{V} \rightarrow V^*$, $u \mapsto (u, \cdot)$, где \bar{V} — комплексно сопряжённое пространство (с умножением на скаляры, определённым по формуле $\lambda \cdot v := \bar{\lambda}v$). Это позволяет отождествить два понятия «сопряжённого» пространства для эрмитова пространства V .

37. ЛИНЕЙНЫЕ ОПЕРАТОРЫ В ЕВКЛИДОВОМ ПРОСТРАНСТВЕ. ОПЕРАТОР, СОПРЯЖЁННЫЙ ЛИНЕЙНОМУ ОПЕРАТОРУ, ЕГО МАТРИЦА И СВОЙСТВА

Пусть $\mathcal{A} : V \rightarrow V$ — линейный оператор в евклидовом пространстве V . В конце вопроса 7 мы определили сопряжённое линейное отображение $\mathcal{A}^* : V^* \rightarrow V^*$ по формуле

$$(\mathcal{A}^*\xi)(v) = \xi(\mathcal{A}v) \quad \forall \xi \in V^*, v \in V,$$

При каноническом отождествлении $V \leftrightarrow V^*$, $u \leftrightarrow \xi_u = (u, \cdot)$ оператор $\mathcal{A}^* : V^* \rightarrow V^*$ переходит в оператор $\mathcal{A}^* : V \rightarrow V$, (который мы для простоты будем обозначать тем же символом \mathcal{A}^*), удовлетворяющий соотношению $\xi_{\mathcal{A}^*u} = \mathcal{A}^*\xi_u \quad \forall u \in V$.

Чтобы установить связь между $\mathcal{A} : V \rightarrow V$ и $\mathcal{A}^* : V \rightarrow V$ непосредственно, вычислим значение линейных функций $\xi_{\mathcal{A}^*u}$ и $\mathcal{A}^*\xi_u$ на векторе $v \in V$:

$$\xi_{\mathcal{A}^*u}(v) = (\mathcal{A}^*u, v), \quad (\mathcal{A}^*\xi_u)(v) = \xi_u(\mathcal{A}v) = (u, \mathcal{A}v).$$

Т. к. $\xi_{\mathcal{A}^*u} = \mathcal{A}^*\xi_u$, мы получаем $(\mathcal{A}^*u, v) = (u, \mathcal{A}v) \quad \forall u, v \in V$.

Определение 1. Пусть $\mathcal{A} : V \rightarrow V$ — оператор в евклидовом или эрмитовом пространстве V . Линейный оператор $\mathcal{A}^* : V \rightarrow V$, удовлетворяющий соотношению

$$(\mathcal{A}^*u, v) = (u, \mathcal{A}v)$$

для любых $u, v \in V$, называется *сопряжённым* к \mathcal{A} .

Соотношение $(\mathcal{A}^*u, v) = (u, \mathcal{A}v)$ определяет оператор \mathcal{A}^* однозначно. Действительно, если $(\mathcal{A}'u, v) = (u, \mathcal{A}v)$ для другого оператора $\mathcal{A}' : V \rightarrow V$, то мы получаем $((\mathcal{A}^* - \mathcal{A}')u, v) = 0 \quad \forall u, v \in V$. В частности, $((\mathcal{A}^* - \mathcal{A}')u, (\mathcal{A}^* - \mathcal{A}')u) = 0$, т. е. $(\mathcal{A}^* - \mathcal{A}')u = \mathbf{0} \quad \forall u \in V$. Следовательно, $\mathcal{A}^* = \mathcal{A}'$.

Предложение 1. Пусть A — матрица оператора $\mathcal{A} : V \rightarrow V$ в ортонормированном базисе евклидова (эрмитова) пространства V . Тогда матрица сопряжённого оператора $\mathcal{A}^* : V \rightarrow V$ в том же базисе есть A^t (соответственно, \bar{A}^t).

Доказательство. Это сразу следует из предложения 5 в вопросе 7, но мы также дадим прямое доказательство. Пусть e_1, \dots, e_n — ортонормированный базис, $A = (a_j^i)$ — матрица оператора $\mathcal{A} : V \rightarrow V$ в этом базисе, а $\tilde{A} = (\tilde{a}_j^i)$ — матрица оператора $\mathcal{A}^* : V \rightarrow V$. Тогда мы имеем

$$\begin{aligned} (\mathcal{A}e_j, e_k) &= (a_j^i e_i, e_k) = \bar{a}_j^i (e_i, e_k) = \bar{a}_j^i \delta_{ik} = \bar{a}_k^j, \\ (e_j, \mathcal{A}^*e_k) &= (e_j, \tilde{a}_k^i e_i) = \tilde{a}_k^i (e_j, e_i) = \tilde{a}_k^i \delta_{ji} = \tilde{a}_j^k. \end{aligned}$$

Т. к. $(\mathcal{A}e_j, e_k) = (e_j, \mathcal{A}^*e_k)$, мы получаем $\bar{a}_k^j = \tilde{a}_j^k$ или $\bar{A}^t = \tilde{A}$. ■

Предложение 2. Имеют место следующие равенства:

1. $(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*$, $(\lambda \mathcal{A})^* = \bar{\lambda} \mathcal{A}^*$;
2. $(\mathcal{A}\mathcal{B})^* = \mathcal{B}^* \mathcal{A}^*$.

Доказательство. Это следует из свойств операции транспонирования матриц, благодаря предыдущему предложению. ■

Докажем ключевую лемму, которая нам не раз понадобится в дальнейшем.

Лемма 1 (Важная). Если $W \subset V$ — инвариантное подпространство относительно \mathcal{A} , то W^\perp — инвариантное подпространство относительно \mathcal{A}^* .

Доказательство. Пусть $u \in W^\perp$. Тогда $\forall w \in W$ имеем $(\mathcal{A}^*u, w) = (u, \mathcal{A}w) = 0$ т.к. $\mathcal{A}w \in W$, а $u \in W^\perp$. Следовательно, $\mathcal{A}^*u \in W^\perp$. ■

38. САМОСОПРЯЖЁННЫЙ ЛИНЕЙНЫЙ ОПЕРАТОР, ЕГО СВОЙСТВА И МАТРИЦА. СУЩЕСТВОВАНИЕ ОРТОНОРМИРОВАННОГО БАЗИСА ИЗ СОБСТВЕННЫХ ВЕКТОРОВ

Определение 1. Оператор $\mathcal{A} : V \rightarrow V$ в евклидовом или эрмитовом пространстве называется *самосопряжённым*, если $\mathcal{A}^* = \mathcal{A}$, т.е. $\forall u, v \in V$ выполнено соотношение

$$(\mathcal{A}u, v) = (u, \mathcal{A}v).$$

Предложение 1. Матрица A самосопряжённого оператора \mathcal{A} в ортонормированном базисе евклидова (эрмитова) пространства симметрична (эрмитова), т.е. $A^t = A$ (соответственно, $\overline{A}^t = A$).

Если матрица оператора \mathcal{A} в некотором ортонормированном базисе симметрична (эрмитова), то оператор \mathcal{A} самосопряжён.

Доказательство. Первое утверждение вытекает из предложения 1 в вопросе 37: т.к. матрица оператора \mathcal{A}^* есть \overline{A}^t и $\mathcal{A}^* = \mathcal{A}$, мы получаем $\overline{A}^t = A$.

Докажем второе утверждение. Пусть e_1, \dots, e_n — ортонормированный базис, в котором матрица A оператора \mathcal{A} эрмитова, т.е. $\overline{A}^t = A$. Тогда из предложения 1 в вопросе 37 следует, что матрица \overline{A}^t оператора \mathcal{A}^* в том же базисе совпадает с A . Значит, $\mathcal{A}^* = \mathcal{A}$ и оператор \mathcal{A} самосопряжён. ■

В связи с этим самосопряжённые операторы в евклидовом пространстве также называют *симметрическими*, а в эрмитовом пространстве — *эрмитовыми*.

Теорема 1. Самосопряжённый оператор диагонализуем в ортонормированном базисе.

Доказательство будет опираться на лемму, которая важна сама по себе.

Лемма 1. Все корни характеристического многочлена самосопряжённого оператора \mathcal{A} вещественны.

Доказательство. Вначале докажем лемму для эрмитова пространства. В этом случае корни характеристического многочлена суть собственные значения оператора \mathcal{A} . Пусть $\lambda \in \mathbb{C}$ — такой корень и $v \neq 0$ — соответствующий собственный вектор, т.е. $\mathcal{A}v = \lambda v$. Тогда

$$\overline{\lambda}(v, v) = (\lambda v, v) = (\mathcal{A}v, v) = (v, \mathcal{A}v) = (v, \lambda v) = \lambda(v, v).$$

Т.к. $(v, v) \neq 0$, то $\overline{\lambda} = \lambda$, т.е. $\lambda \in \mathbb{R}$.

Случай евклидова пространства сводится к эрмитовому случаю при помощи комплексификации. Пусть A — матрица самосопряжённого оператора \mathcal{A} в ортонормированном базисе евклидова пространства V . Тогда матрица A вещественна и симметрична. Та же матрица A будет матрицей комплексифицированного оператора $\mathcal{A}_{\mathbb{C}}$ в соответствующем базисе эрмитова пространства $V_{\mathbb{C}}$. Этот базис также ортонормирован, а матрица A , будучи вещественной и симметричной, является эрмитовой. Следовательно, оператор $\mathcal{A}_{\mathbb{C}}$ самосопряжён, а корни его характеристического многочлена вещественны и совпадают с корнями характеристического многочлена \mathcal{A} . ■

Теперь докажем теорему 1.

Доказательство. Используем индукцию по размерности пространства V . При $\dim V = 1$ доказывать нечего. Предположим, что утверждение доказано для операторов в пространствах размерности $n - 1$, и докажем его для пространства V размерности n .

В силу предыдущей леммы у самосопряжённого оператора \mathcal{A} имеется собственный вектор v , т. е. одномерное инвариантное подпространство $W = \langle v \rangle$. В силу важной леммы, ортогональное дополнение W^\perp инвариантно относительно оператора $\mathcal{A}^* = \mathcal{A}$. Т. к. $\dim W^\perp = n - 1$, в пространстве W^\perp имеется ортонормированный базис e_1, \dots, e_{n-1} из собственных векторов оператора $\mathcal{A}|_{W^\perp}$. Тогда $e_1, \dots, e_{n-1}, \frac{v}{|v|}$ — ортонормированный базис из собственных векторов оператора \mathcal{A} . ■

Диагональный вид матрицы самосопряжённого оператора \mathcal{A} в ортонормированном базисе из собственных векторов называется *каноническим видом* самосопряжённого оператора.

Практический метод нахождения ортонормированного базиса из собственных векторов основан на следующей лемме.

Лемма 2. Собственные векторы, отвечающие различным собственным значениям самосопряжённого оператора \mathcal{A} , взаимно ортогональны.

Доказательство. Пусть $\mathcal{A}u = \lambda u$ и $\mathcal{A}v = \mu v$, где $\lambda \neq \mu$ — вещественные собственные значения. Тогда

$$\lambda(u, v) = (\lambda u, v) = (\mathcal{A}u, v) = (u, \mathcal{A}v) = (u, \mu v) = \mu(u, v),$$

откуда $(u, v) = 0$, т. к. $\lambda \neq \mu$. ■

Для нахождения ортонормированного базиса из собственных векторов самосопряжённого оператора \mathcal{A} находятся все его собственные подпространства, а затем в каждом из них выбирается ортонормированный базис. Объединение этих базисов и будет нужным базисом для \mathcal{A} . Про это хорошо НАПИСАЛ Антон Александрович Клячко.

В евклидовом пространстве верно и утверждение, обратное к предыдущей теореме.

Предложение 2. Если оператор \mathcal{A} в евклидовом пространстве диагонализировать в ортонормированном базисе, то \mathcal{A} самосопряжён.

Доказательство. Действительно, диагональная матрица симметрична, а оператор, имеющий симметричную матрицу в ортонормированном базисе евклидова пространства самосопряжён согласно предложению 1. ■

В эрмитовом пространстве класс операторов, имеющих симметричную матрицу в ортонормированном базисе, шире, чем самосопряжённые (т. к. на диагонали могут стоять не вещественные числа).

39. ОРТОГОНАЛЬНЫЙ ОПЕРАТОР, ЕГО СВОЙСТВА И МАТРИЦА. КАНОНИЧЕСКИЙ ВИД МАТРИЦЫ ОРТОГОНАЛЬНОГО ОПЕРАТОРА

Предложение 1. Следующие условия для оператора $\mathcal{A} : V \rightarrow V$ в евклидовом или эрмитовом пространстве эквивалентны:

1. Оператор \mathcal{A} сохраняет длины векторов, т. е. $|\mathcal{A}v| = |v| \ \forall v \in V$;
2. Оператор \mathcal{A} сохраняет скалярное произведение, т. е. $(\mathcal{A}u, \mathcal{A}v) = (u, v) \ \forall u, v \in V$;
3. Оператор \mathcal{A} переводит ортонормированные базисы в ортонормированные, т. е. если e_1, \dots, e_n — ортонормированный базис, то $\mathcal{A}e_1, \dots, \mathcal{A}e_n$ — также ортонормированный базис;
4. Матрица A оператора \mathcal{A} в ортонормированном базисе ортогональна (унитарна), т. е. $A^t A = E$ (соответственно, $\overline{A}^t A = E$);
5. $\mathcal{A}^* \mathcal{A} = \text{id}$, т. е. сопряжённый оператор к \mathcal{A} является его обратным.

Доказательство. Мы докажем импликации $1 \Leftrightarrow 2, 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 5 \Rightarrow 2$.

$1 \Rightarrow 2$. В евклидовом пространстве имеем $(u + v, u + v) = (u, u) + 2(u, v) + (v, v)$, откуда

$$(u, v) = \frac{1}{2}((u + v, u + v) - (u, u) - (v, v)).$$

Поэтому если оператор сохраняет квадраты длин, т. е. скалярные произведения вида (v, v) , то он сохраняет и все скалярные произведения.

В эрмитовом пространстве имеем $(u+v, u+v) = (u, u) + (u, v) + \overline{(u, v)} + (v, v) = (u, u) + 2\operatorname{Re}(u, v) + (v, v)$, откуда

$$\operatorname{Re}(u, v) = \frac{1}{2}((u+v, u+v) - (u, u) - (v, v))$$

и, аналогично,

$$\operatorname{Im}(u, v) = -\frac{1}{2}((u+iv, u+iv) - (u, u) - (v, v)).$$

Поэтому если оператор сохраняет длины, то он сохраняет и все скалярное произведение (u, v) , т. к. он сохраняет его вещественную и мнимую части.

2 \Rightarrow 1. Очевидно.

2 \Rightarrow 3. Пусть $(\mathcal{A}u, \mathcal{A}v) = (u, v)$. Тогда если e_1, \dots, e_n — ортонормированный базис, то $(\mathcal{A}e_i, \mathcal{A}e_j) = (e_i, e_j) = \delta_{ij}$, т. е. базис $\mathcal{A}e_1, \dots, \mathcal{A}e_n$ также ортонормирован.

3 \Rightarrow 4. Пусть \mathcal{A} переводит ортонормированный базис e_1, \dots, e_n в ортонормированный $\mathcal{A}e_1, \dots, \mathcal{A}e_n$, и $A = (a_j^i)$ — матрица оператора в базисе e_1, \dots, e_n . Тогда

$$\delta_{ij} = (\mathcal{A}e_i, \mathcal{A}e_j) = (a_i^k e_k, a_j^\ell e_\ell) = \bar{a}_i^k a_j^\ell (e_k, e_\ell) = \bar{a}_i^k a_j^\ell \delta_{k\ell} = \bar{a}_i^k \delta_{k\ell} a_j^\ell.$$

Это эквивалентно матричному соотношению $E = \bar{A}^t E A$ или $\bar{A}^t A = E$.

4 \Rightarrow 5. Это следует из предложения 1 в вопросе 37.

5 \Rightarrow 2. Пусть $\mathcal{A}^* \mathcal{A} = \operatorname{id}$. Тогда $(\mathcal{A}u, \mathcal{A}v) = (\mathcal{A}^* \mathcal{A}u, v) = (u, v)$, т. е. \mathcal{A} сохраняет скалярное произведение. ■

Определение 1. Оператор $\mathcal{A} : V \rightarrow V$ в евклидовом (эрмитовом) пространстве, удовлетворяющий одному из эквивалентных свойств последнего предложения, называется *ортгональным* (соответственно, *унитарным*).

Иногда ортгональные и унитарные операторы называют *изометрическими*.

Как и в случае самосопряжённых операторов, для приведения ортгонального или унитарного преобразования к каноническому виду нам понадобится утверждение об инвариантности ортгонального дополнения. Причём, эту лемму тоже будем называть «важной».

Лемма 1 (Важная). Пусть $\mathcal{A} : V \rightarrow V$ — ортгональный или унитарный оператор, а $W \subset V$ — инвариантное относительно \mathcal{A} подпространство. Тогда ортгональное дополнение W^\perp также инвариантно относительно \mathcal{A} .

Доказательство. Пусть $u \in W^\perp$. Нам надо показать, что $\mathcal{A}u \in W^\perp$, т. е. что $(\mathcal{A}u, w) = 0 \ \forall w \in W$. Мы знаем, что $\mathcal{A}(W) \subset W$. Поскольку оператор \mathcal{A} обратим, $\mathcal{A}(W) = W$. Тогда найдётся такой вектор $v \in W$, что $w = \mathcal{A}v$, а значит, $(\mathcal{A}u, w) = (\mathcal{A}u, \mathcal{A}v) = (u, v) = 0$. ■

Лемма 2. Собственные значения ортгонального (унитарного) оператора \mathcal{A} по модулю равны 1.

Доказательство. Действительно, пусть $\mathcal{A}v = \lambda v$ для $v \neq 0$. Тогда

$$(v, v) = (\mathcal{A}v, \mathcal{A}v) = (\lambda v, \lambda v) = \bar{\lambda} \lambda (v, v) = |\lambda|^2 (v, v),$$

откуда $|\lambda|^2 = 1$. ■

Теорема 1. Полностью аналогично доказательству теоремы 1 в вопросе 38 и теоремы 1 в приложении про косоэрмитовы операторы. Шаг индукции проводим, выбирая собственный вектор v и устанавливая инвариантность подпространства $\langle v \rangle^\perp$ при помощи последней леммы.

Теорема 2. Для ортгонального оператора \mathcal{A} существует ортонормированный базис, в котором его матрица блочно-диагональная с блоками размера 1 или 2, причём блоки размера 1 имеют вид

(1) или (-1) , а блоки размера 2 имеют вид $\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$, где $\varphi \neq \pi k$, $k \in \mathbb{Z}$.

Доказательство. В пространстве размерности 1 матрица и так имеет вид (1) или (-1) . В пространстве размерности 2 любая ортогональная матрица имеет вид $\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$ (если определитель равен 1) или $\begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix}$ (если определитель равен -1). В первом случае мы уже имеем требуемый вид (а оператор представляет собой поворот на угол φ в положительном направлении). Во втором случае оператор представляет собой симметрию относительно прямой под углом $\varphi/2$ к оси абсцисс. Такой оператор имеет два ортогональных собственных вектора: $(\cos \frac{\varphi}{2}, \sin \frac{\varphi}{2})$ (вектор вдоль оси симметрии) и $(-\sin \frac{\varphi}{2}, \cos \frac{\varphi}{2})$ (вектор, перпендикулярный оси симметрии). В ортонормированном базисе из этих собственных векторов операторов $\begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix}$ принимает требуемый вид $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

Далее действуем по индукции, как и при доказательстве теоремы 2 из приложения про кососимметрические операторы. Предположим, что утверждение доказано для операторов в пространствах размерности не больше $n - 1$, и докажем его для пространства V размерности n (где $n \geq 3$).

В силу теоремы 1 в вопросе 19 для оператора \mathcal{A} существует одномерное или двумерное инвариантное подпространство $W \subset V$. Из важной леммы следует, что ортогональное дополнение W^\perp также инвариантно.

По предположению индукции, в пространстве W^\perp имеется требуемый базис для ортогонального оператора $\mathcal{A}|_{W^\perp}$. Выбрав ортонормированный базис в пространстве W , как описано в начале доказательства, и взяв объединение базисов в W^\perp и W , мы получим требуемый ортонормированный базис пространства V . ■

Матрицы, описанные в двух последних теоремах, называются *каноническим видом* унитарного и ортогонального оператора.

Пример 1.

1. Как указано в доказательстве предыдущей теоремы, ортогональный оператор с матрицей $\begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix}$ имеет канонический вид $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Тот же канонический вид будет, если оператор рассматривать как унитарный.
2. Канонический вид ортогонального оператора с матрицей $\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$ — это та же самая матрица. В то же время канонический вид этого оператора, рассматриваемого как унитарный оператор, есть $\begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{i\varphi} \end{pmatrix}$, где $e^{i\varphi} = \cos \varphi + i \sin \varphi$.
3. В трёхмерном пространстве канонический вид ортогонального оператора есть

$$\begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & \pm 1 \end{pmatrix},$$

где в левом нижнем углу стоит 1 или -1 в зависимости от знака определителя оператора. Операторы, канонический вид которых имеет три блока (1) или (-1) , получаются при $\varphi = \pi k$. Если определитель положителен, то такой оператор представляет собой поворот (вокруг оси третьего вектора канонического базиса). Если же определитель отрицателен, то оператор — это «поворот с переворотом», т. е. композиция поворота и симметрии относительно плоскости, перпендикулярной оси поворота.

Отсюда, в частности, следует, что композиция двух поворотов — это снова поворот вокруг некоторой оси (т. к. в каноническом виде всегда происходит всего один поворот).

4. В четырёхмерном пространстве уже бывают независимые повороты. А именно, канонический вид ортогонального оператора общего вида с положительным определителем представляет

собой матрицу из двух блоков размера 2×2 :

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

Это композиция двух независимых поворотов: на угол φ в плоскости первого и второго базисных векторов и на угол ψ в плоскости третьего и четвёртого базисных векторов. Такой оператор не сводится к одному повороту.

Произведение ортогональных операторов, очевидно, является ортогональным оператором, и поэтому ортогональные операторы в евклидовом пространстве V образуют подгруппу в общей линейной группе $\text{GL}(V)$. Эта подгруппа называется *ортогональной подгруппой* и обозначается $O(V) \simeq O(\dim V)$.

40. ПОЛЯРНОЕ РАЗЛОЖЕНИЕ НЕВЫРОЖДЕННОГО ЛИНЕЙНОГО ОПЕРАТОРА В ЕВКЛИДОВОМ ПРОСТРАНСТВЕ. СИНГУЛЯРНОЕ РАЗЛОЖЕНИЕ

Определение 1. Самосопряжённый оператор \mathcal{A} называется *положительным*, если $(\mathcal{A}v, v) > 0$ $\forall v \in V \setminus \{0\}$.

Лемма 1. Самосопряжённый оператор \mathcal{A} положителен тогда и только тогда, когда все его собственные значения положительны.

Доказательство. \Rightarrow . Пусть $(\mathcal{A}v, v) > 0$ при $v \neq 0$. Рассмотрим собственный вектор v с собственным значением λ . Тогда $\lambda(v, v) = (\mathcal{A}v, v) > 0$, откуда $\lambda > 0$.

\Leftarrow . Обратно, пусть все собственные значения λ_i положительны. Выберем ортонормированный базис из собственных векторов e_1, \dots, e_n с собственными значениями $\lambda_1, \dots, \lambda_n$. Тогда для ненулевого вектора $v = v^i e_i$ мы имеем

$$(\mathcal{A}v, v) = (\mathcal{A}(v^i e_i), v^j e_j) = \bar{v}^i v^j (\mathcal{A}e_i, e_j) = \sum_{i,j} \lambda_i \bar{v}^i v^j (e_i, e_j) = \sum_{i=1}^n \lambda_i |v^i|^2 > 0.$$

■

Теорема 1. Для положительного оператора \mathcal{A} существует единственный положительный оператор \mathcal{P} , удовлетворяющий соотношению $\mathcal{P}^2 = \mathcal{A}$.

Доказательство. Пусть $\lambda_1, \dots, \lambda_k$ — различные собственные значения оператора \mathcal{A} и V_1, \dots, V_k — соответствующие собственные подпространства. Согласно последней лемме, все λ_i положительны. Положим $\mu_i := \sqrt{\lambda_i}$. Рассмотрим оператор \mathcal{P} , действующий в пространства V_i умножением на μ_i . Тогда $\mathcal{P}^2 = \mathcal{A}$. Оператор \mathcal{P} самосопряжён (т. к. он задаётся диагональной матрицей в ортонормированном базисе из собственных векторов оператора \mathcal{A}) и положителен в силу последней леммы.

Осталось доказать единственность оператора \mathcal{P} . Пусть оператор \mathcal{P} удовлетворяет условиям теоремы. Пусть μ_1, \dots, μ_k — его различные собственные значения и W_1, \dots, W_k — соответствующие собственные подпространства. Тогда оператор $\mathcal{P}^2 = \mathcal{A}$ действует в пространстве W_i умножением на μ_i^2 , т. е. μ_i^2 является собственным значением для \mathcal{A} . Следовательно, при подходящей нумерации имеем $\mu_i^2 = \lambda_i$ и $W_i = V_i$. Это показывает, что оператор \mathcal{P} определён однозначно. ■

Определение 2. Оператора \mathcal{P} , построенный в предыдущей теореме, называется *положительным корнем* из положительного оператора \mathcal{A} и обозначается $\sqrt{\mathcal{A}}$.

Определение 3. Самосопряжённый оператор \mathcal{A} называется *неотрицательным*, если $(\mathcal{A}v, v) \geq 0$ $\forall v \in V$. Все утверждения выше переносятся без изменений на неотрицательные операторы.

Теорема 2 (Полярное разложение). Для любого невырожденного оператора \mathcal{A} в евклидовом или эрмитовом пространстве существует единственное представление в виде

$$\mathcal{A} = \mathcal{P}\mathcal{U},$$

где \mathcal{P} — положительный, а \mathcal{U} — ортогональный (унитарный) оператор.

Доказательство. Если $\mathcal{A} = \mathcal{P}\mathcal{U}$, то $\mathcal{A}^* = \mathcal{U}^*\mathcal{P}$ и $\mathcal{A}\mathcal{A}^* = \mathcal{P}\mathcal{U}\mathcal{U}^*\mathcal{P} = \mathcal{P}^2$. Оператор $\mathcal{A}\mathcal{A}^*$ очевидно самосопряжён; кроме того, он является положительным:

$$(\mathcal{A}\mathcal{A}^*v, v) = (\mathcal{A}^*v, \mathcal{A}^*v) > 0$$

при $v \neq 0$, т. к. $\mathcal{A}^*v \neq 0$ в силу невырожденности \mathcal{A} . Поэтому положительный оператор \mathcal{P} , удовлетворяющий соотношению $\mathcal{A}\mathcal{A}^* = \mathcal{P}^2$, единственен в силу последней теоремы, а именно, $\mathcal{P} = \sqrt{\mathcal{A}\mathcal{A}^*}$. Тогда оператор $\mathcal{P}^{-1}\mathcal{A}$ также определён однозначно. Мы имеем $\mathcal{U}\mathcal{U}^* = \mathcal{P}^{-1}\mathcal{A}\mathcal{A}^*\mathcal{P}^{-1} = \mathcal{P}^{-1}\mathcal{P}^2\mathcal{P}^{-1} = \text{id}$. Следовательно, \mathcal{U} — ортогональный (унитарный) оператор и $\mathcal{A} = \mathcal{P}\mathcal{U}$. ■

Аналогично, рассмотрев положительный оператор $\mathcal{A}^*\mathcal{A}$ можно доказать существование второго полярного разложения $\mathcal{A} = \mathcal{U}'\mathcal{P}'$ (где $\mathcal{P}' = \sqrt{\mathcal{A}^*\mathcal{A}}$).

По этой теме см. теоретические задачи 21 и 22.

В одномерном эрмитовом пространстве \mathbb{C} положительные операторы — это положительные вещественные числа, а унитарные операторы — это комплексные числа, по модулю равные 1, т. е. вида $e^{i\varphi}$. Поэтому полярное разложение — это представление комплексного числа z в полярных координатах: $z = \rho e^{i\varphi}$, что объясняет название.

Теорема 3 (Сингулярное разложение). Для любого линейного отображения $\mathcal{A} : V \rightarrow V$ евклидова или эрмитова пространства V существуют такие ортонормированные базисы e_1, \dots, e_n и f_1, \dots, f_n и положительные числа $\sigma^1 \geq \sigma^2 \geq \dots \geq \sigma^r$, где $r = \dim \text{Im } \mathcal{A} = \text{rk } \mathcal{A}$, что

$$\mathcal{A}e_i = \begin{cases} \sigma^i f_i, & i \leq r, \\ 0, & i > r. \end{cases}$$

Числа $\sigma_1, \sigma_2, \dots, \sigma_r$ называют *сингулярными числами* оператора \mathcal{A} . А базисы e_1, \dots, e_n и f_1, \dots, f_n называют *сингулярными базисами* оператора \mathcal{A} .

Доказательство. Оператор $\mathcal{A}^*\mathcal{A}$ самосопряжён и неотрицателен:

$$(\mathcal{A}^*\mathcal{A})^* = \mathcal{A}^*\mathcal{A}, \quad (\mathcal{A}^*\mathcal{A}v, v) = (\mathcal{A}v, \mathcal{A}v) \geq 0$$

$\forall v \in V$. Поэтому существует ортонормированный базис собственных векторов e_1, \dots, e_m матрицы $\mathcal{A}^*\mathcal{A}$. Все её собственные значения неотрицательны. Таким образом,

$$\mathcal{A}^*\mathcal{A}e_i = (\sigma^i)^2 e_i, \quad i = 1, \dots, n.$$

Будем считать при этом, что $\sigma^1 \geq \sigma^2 \geq \dots \geq \sigma^r > 0$, $\sigma^{r+1} = \dots = \sigma^n = 0$. Значит,

$$(\mathcal{A}e_i, \mathcal{A}e_j) = (\mathcal{A}^*\mathcal{A}e_i, e_j) = (\sigma^i)^2 (e_i, e_j) \delta_{ij} = \begin{cases} 0, & i \neq j, \\ \sigma_i^2, & i = j \end{cases}$$

Значит, векторы $f_i := (\sigma^i)^{-1} \mathcal{A}e_i$ образуют ортономированный базис в V , причём $\mathcal{A}e_i = \sigma^i f_i$. ■

41. Билинейные и квадратичные функции на евклидовом пространстве.

ВЗАИМНО ОДНОЗНАЧНОЕ СООТВЕТСТВИЕ МЕЖДУ СИММЕТРИЧЕСКИМИ
БИЛИНЕЙНЫМИ ФОРМАМИ И САМОСОПРЯЖЕННЫМИ ЛИНЕЙНЫМИ ОПЕРАТОРАМИ.

ПРИВЕДЕНИЕ КВАДРАТИЧНОЙ ФОРМЫ К ГЛАВНЫМ ОСЯМ. ПРИВЕДЕНИЕ К
ДИАГОНАЛЬНОМУ ВИДУ ПАРЫ ФОРМ, ОДНА ИЗ КОТОРЫХ ПОЛОЖИТЕЛЬНО
ОПРЕДЕЛЕНА

Пусть V — евклидово пространство. Мы знаем из теоремы 1 в вопросе 36, что отображение $x \mapsto \xi_x = (x, \cdot)$ устанавливает канонический изоморфизм $V \rightarrow V^*$ между V и его двойственным пространством V^* . Это позволяет нам отождествить пространства линейных отображений $\text{Hom}(V, V)$ и $\text{Hom}(V, V^*)$. С другой стороны, $\text{Hom}(V, V)$ — это пространство $\text{End}(V)$ линейных операторов, а в силу теоремы 2 в вопросе 25, $\text{Hom}(V, V^*)$ — это пространство билинейных функций $B(V)$. Если вникнуть в построение изоморфизмов, то мы увидим, что в явном виде канонический изоморфизм между пространством операторов и пространством билинейных функций описывается следующим утверждением, которое легко доказать и непосредственно:

Предложение 1. Пусть V — евклидово пространство. Отображение $\mathcal{A} \mapsto \mathcal{B}_{\mathcal{A}} := (\mathcal{A}\cdot, \cdot)$ устанавливает изоморфизм $\psi : \text{End}(V) \rightarrow B(V)$.

Доказательство. Т.к. $\dim \text{End}(V) = \dim B(V) = n^2$, достаточно проверить, что $\text{Ker } \psi = \{\mathcal{O}\}$. Пусть $\psi(\mathcal{A}) = 0$, т.е. $\mathcal{B}_{\mathcal{A}}$ — тождественно нулевая функция. Тогда $(\mathcal{A}x, y) = 0 \ \forall x, y \in V$. В частности, $(\mathcal{A}x, \mathcal{A}x) = 0 \ \forall x \in V$, т.е. $\mathcal{A}x = \mathbf{0} \ \forall x \in V$ и $\mathcal{A} = \mathcal{O}$ — нулевой оператор. ■

Это утверждение имеет важные следствия: оно позволяет переводить утверждения об операторах в утверждения о билинейных функциях и наоборот. Одно из основных приложений заключается в следующем:

Теорема 1. Для билинейной симметрической функции в евклидовом пространстве существует ортонормированный базис, в котором её матрица диагональна. Другими словами, квадратичная форма приводится к диагональному виду ортогональным преобразованием.

Первое доказательство. Пусть \mathcal{B} — симметрическая билинейная функция и \mathcal{A} — соответствующий ей оператор, т.е. $\mathcal{B} = \mathcal{B}_{\mathcal{A}}$. Тогда из $\mathcal{B}(x, y) = \mathcal{B}(y, x)$ получаем $(\mathcal{A}x, y) = (\mathcal{A}y, x) = (x, \mathcal{A}y)$, т.е. оператор \mathcal{A} самосопряжён. Выберем ортонормированный базис e_1, \dots, e_n из собственных векторов оператора \mathcal{A} , т.е. $\mathcal{A}e_i = \lambda_i e_i$. Тогда для матрицы $B = (b_{ij})$ функции \mathcal{B} в этом базисе имеем

$$b_{ij} = \mathcal{B}(e_i, e_j) = (\mathcal{A}e_i, e_j) = (\lambda_i e_i, e_j) = \lambda_i \delta_{ij},$$

т.е. матрица B диагональна (и совпадает с матрицей оператора \mathcal{A}). ■

Второе доказательство. Посмотрим, как преобразуется матрица билинейной функции и матрица оператора при ортогональном преобразовании. Пусть B — матрица билинейной функции в некотором ортонормированном базисе. При ортогональном преобразовании с матрицей C матрица B переходит в матрицу $B' = C^t B C$. Т.к. матрица C ортогональна, то же преобразование мы можем записать в виде $B' = C^{-1} B C$. Но это — закон преобразования для матрицы оператора. Т.к. оператор с симметричной матрицей B в ортонормированном базисе самосопряжён, его можно привести к диагональному виду ортогональным преобразованием. ■

Диагональный вид, к которому приводится симметрическая билинейная функция (квадратичная форма) ортогональным преобразованием, называется *каноническим*.

Предложение 2. Канонический вид симметрической билинейной функции (квадратичной формы) единственный с точностью до перестановки диагональных элементов. Эти элементы представляют собой собственные значения матрицы квадратичной формы в любом ортонормированном базисе.

Доказательство. Пусть Q — матрица квадратичной формы в ортонормированном базисе. Тогда диагональные элементы канонического вида — это собственные значения самосопряжённого опе-

ратора с матрицей Q , т. е. корни уравнения $\det(Q - tE) = 0$. В другом ортонормированном базисе матрица квадратичной формы есть $Q' = C^t Q C$ и её собственные значения находятся из уравнения $\det(Q' - tE) = 0$. Т. к.

$$\det(Q' - tE) = \det(C^t Q C - t C^t C) = \det(C^t (Q - tE) C) = \det(C^t C) \det(Q - tE) = \det(Q - tE),$$

собственные значения матриц Q и Q' совпадают. ■

Модификация последней теоремы позволяет одновременно приводить к диагональному виду сразу две квадратичные формы, одна из которых положительно определена:

Теорема 2. Пусть даны две квадратичные формы $Q(x)$ и $B(x)$, причём форма $Q(x)$ положительно определена. Тогда существует линейная замена координат $x = Cy$, приводящая форму $Q(x)$ к нормальному виду $(y^1)^2 + \dots + (y^n)^2$, а форму $B(x)$ — к диагональному виду $\lambda_1(y^1)^2 + \dots + \lambda_n(y^n)^2$.

Доказательство. Положительно определённая симметрическая билинейная функция, соответствующая квадратичной форме $Q(x)$, превращает V в евклидово пространство. В исходных координатах матрица Грама скалярного произведения есть Q . В любом ортонормированном базисе матрица Грама (она же матрица квадратичной формы $Q(x)$) будет единичной. Согласно последней теореме, существует ортонормированный базис, в котором матрица формы $B(x)$ имеет диагональный вид. ■

Обратим внимание, что матрица C замены координат из предыдущей теоремы не является ортогональной: вместо соотношения $C^t C = E$ она удовлетворяет соотношению $C^t Q C = E$. Другими словами, столбцы матрицы C образуют ортонормированный базис относительно скалярного произведения в \mathbb{R}^n с матрицей Грама Q .

Достаточно эффективный алгоритм приведения пары квадратичных форм, одна из которых положительно определена, к диагональному виду, основан на следующем определении.

Определение 1. Пусть даны две квадратичные формы $Q(x)$ и $B(x)$, причём $Q(x)$ положительно определена. Корни уравнения $\det(B - tQ) = 0$ называются *собственными значениями пары форм $Q(x)$ и $B(x)$* .

Пусть λ — собственное значение пары форм $Q(x)$ и $B(x)$. Ненулевой вектор y , удовлетворяющий системе уравнений $(B - \lambda Q)y = 0$, называется *собственным вектором пары форм*, соответствующим собственному значению λ .

Теорема 3. Предположим, что линейная замена $x = Cy$ приводит положительно определённую форму $Q(x)$ к нормальному виду $(y^1)^2 + \dots + (y^n)^2$, а форму $B(x)$ — к диагональному виду $\lambda_1(y^1)^2 + \dots + \lambda_n(y^n)^2$. Тогда числа $\lambda_1, \dots, \lambda_n$ суть собственные значения пары форм $Q(x)$ и $B(x)$, а столбцы матрицы C образуют базис из собственных векторов пары форм, который является ортонормированным относительно скалярного произведения, задаваемого формой $Q(x)$.

Доказательство. Мы имеем $C^t Q C = E$, а $C^t B C = D$, где D — диагональная матрица с числами $\lambda_1, \dots, \lambda_n$ на диагонали. Тогда $\forall i$ матрица $D - \lambda_i E$ вырождена, следовательно,

$$0 = \det(D - \lambda_i E) = \det(C^t B C - \lambda_i C^t Q C) = \det(C^t (B - \lambda_i Q) C) = \det(C)^2 \det(B - \lambda_i Q).$$

Т. к. матрица C невырождена, отсюда следует, что $\det(B - \lambda_i Q) = 0$, т. е. λ_i — собственное значение пары форм.

Пусть c_i — i -ый столбец матрицы C . Из соотношения $C^t (B - \lambda_i Q) C = D - \lambda_i E$ мы получаем, что i -ый столбец матрицы $C^t (B - \lambda_i Q) C$ нулевой, т. е. $C^t (B - \lambda_i Q) c_i = 0$. Т. к. матрица C обратима, отсюда следует, что $(B - \lambda_i Q) c_i = 0$, т. е. c_i — собственный вектор пары форм, отвечающий собственному значению λ_i .

Наконец, соотношение $C^t Q C = E$ выражает тот факт, что столбцы матрицы C образуют ортонормированный базис относительно скалярного произведения с матрицей Грама Q . ■

42. ПОЛУТОРАЛИНЕЙНЫЕ ФУНКЦИИ И ФОРМЫ. ЭРМИТОВЫ ФОРМЫ И ЭРМИТОВЫ МАТРИЦЫ. КАНОНИЧЕСКИЙ ВИД ЭРМИТОВОЙ КВАДРАТИЧНОЙ ФОРМЫ. КРИТЕРИЙ СИЛЬВЕСТРА

Я сразу писал и про евклидовы, и про эрмитовы пространства, поэтому в этом и следующих трёх билетах мне писать нечего, могу лишь сослаться на билеты, де это уже написано.

Определение 1. Пусть V — линейное пространство над полем \mathbb{C} . Функция $\mathcal{S} : V \times V \rightarrow \mathbb{C}$ называется *полуторалинейной*, если она линейна по второму аргументу и *антILINEЙНА* (или *полулинейна*) по первому аргументу:

$$\begin{aligned}\mathcal{S}(\lambda_1 x_1 + \lambda_2 x_2, y) &= \overline{\lambda_1} \mathcal{S}(x_1, y) + \overline{\lambda_2} \mathcal{S}(x_2, y), \\ \mathcal{S}(x, \mu_1 y_1 + \mu_2 y_2) &= \mu_1 \mathcal{S}(x, y_1) + \mu_2 \mathcal{S}(x, y_2)\end{aligned}$$

для любых $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{C}$ и $x, x_1, x_2, y, y_1, y_2 \in V$. Матрица полуторалинейной функции \mathcal{S} в базисе e_1, \dots, e_n определяется как $S = (s_{ij})$, где $s_{ij} = \mathcal{S}(e_i, e_j)$.

Значение $\mathcal{S}(x, y)$ выражается через матрицу $S = (s_{ij})$ и координаты векторов $x = x^i e_i$ и $y = y^j e_j$ следующим образом:

$$\mathcal{S}(x, y) = \mathcal{S}(x^i e_i, y^j e_j) = \overline{x^i} y^j \mathcal{S}(e_i, e_j) = s_{ij} \overline{x^i} y^j = \overline{x}^t S y.$$

Определение 2. Выражение $\mathcal{S}(x, y) = s_{ij} \overline{x^i} y^j = \overline{x}^t S y$ называется *полуторалинейной формой*.

Примером полуторалинейной функции может служить эрмитово скалярное произведение.

Матрицы S и S' полуторалинейной функции \mathcal{S} в разных базисах связаны соотношением

$$S' = \overline{C}^t S C,$$

которое доказывается аналогично соотношению для билинейных функций. Отсюда следует, что ранг матрицы полуторалинейной функции не зависит от выбора базиса.

Критерий Сильвестра уже был в вопросе 29.

43. УНИТАРНОЕ (ЭРМИТОВО) ПРОСТРАНСТВО. ОРТОГОНАЛИЗАЦИЯ. ОРТОГОНАЛЬНЫЕ И ОРТОНОРМИРОВАННЫЕ БАЗИСЫ. МАТРИЦА ПЕРЕХОДА МЕЖДУ ОРТОНОРМИРОВАННЫМИ БАЗИСАМИ. ГРУППА $U(n)$ УНИТАРНЫХ МАТРИЦ ПОРЯДКА n

Определение эрмитова пространства см. в вопросе 30, про всё остальное — в вопросах 32 и 33.

44. САМОСОПРЯЖЕННЫЕ (ЭРМИТОВЫ) ОПЕРАТОРЫ В УНИТАРНОМ ПРОСТРАНСТВЕ, ИХ СВОЙСТВА. ТЕОРЕМА О СУЩЕСТВОВАНИИ ОРТОНОРМИРОВАННОГО БАЗИСА ИЗ СОБСТВЕННЫХ ВЕКТОРОВ ЭРМИТОВО ОПЕРАТОРА. ПРИВЕДЕНИЕ ЭРМИТОВОЙ КВАДРАТИЧНОЙ ФОРМЫ К ГЛАВНЫМ ОСЯМ

См. вопросы 38 и 41 (про приведение к главным осям).

45. УНИТАРНЫЕ ОПЕРАТОРЫ В УНИТАРНОМ ПРОСТРАНСТВЕ, ИХ СВОЙСТВА. ТЕОРЕМА О СУЩЕСТВОВАНИИ ОРТОНОРМИРОВАННОГО БАЗИСА ИЗ СОБСТВЕННЫХ ВЕКТОРОВ УНИТАРНОГО ОПЕРАТОРА

См. вопрос 39.

46. АФФИННОЕ ПРОСТРАНСТВО, ОСНОВНЫЕ ПОНЯТИЯ. АФФИНИЗАЦИЯ ВЕКТОРНОГО ПРОСТРАНСТВА. АФФИННАЯ СИСТЕМА КООРДИНАТ, ИЗМЕНЕНИЕ КООРДИНАТ ПРИ ЗАМЕНЕ СИСТЕМЫ КООРДИНАТ

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

Определение 1. *Аффинным пространством* называется пара (\mathfrak{A}, V) , состоящая из множества \mathfrak{A} , элементы которого называются *точками*, и векторного пространства V над полем \mathbb{F} с дополнительной операцией сложения

$$+ : \mathfrak{A} \times V \rightarrow \mathfrak{A}, \quad (p, v) \rightarrow p + v,$$

которая удовлетворяет следующим аксиомам:

1. $p + \mathbf{0} = p \quad \forall p \in \mathfrak{A}$;
2. $(p + u) + v = p + (u + v) \quad \forall p \in \mathfrak{A}$;
3. $\forall p, q \in \mathfrak{A} \exists! v \in V : p + v = q$.

Размерностью $\dim(\mathfrak{A}, V)$ аффинного пространства (\mathfrak{A}, V) называется размерность векторного пространства V .

Часто аффинным пространством называют просто множество точек \mathfrak{A} из определения выше (особенно когда из контекста понятно, какое векторное пространство V имеется в виду). Вектор $v \in V$, однозначно сопоставляемый паре точек $p, q \in \mathfrak{A}$ в силу свойства 3 обозначается \overline{pq} . Тогда из свойства 2 вытекает, что $\overline{pq} + \overline{qr} = \overline{pr}$.

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

Примечание.

1. Точки плоскости и трёхмерного пространства образуют аффинные пространства.
2. Рассмотрим совместную неоднородную систему линейных уравнений $Ax = b$, где A — матрица, а x и b — столбцы. Пусть \mathfrak{A} — множество решений x этой системы, а V — векторное пространство решений y однородной системы $Ay = 0$. Тогда (\mathfrak{A}, V) аффинное пространство. Действительно, если $x \in \mathfrak{A}$ и $y \in V$, то $A(x + y) = Ax + Ay = b$, поэтому $x + y \in \mathfrak{A}$.
3. Из всякого векторного пространства V можно получить аффинное пространство, взяв в качестве \mathfrak{A} множество векторов V ; при этом сложение точек и векторов это просто сложение векторов в исходном пространстве V . Аффинное пространство, получаемое при помощи этой процедуры (*аффинизации*) из \mathbb{F}^n , мы будем обозначать через \mathbb{A}^n .

Определение 2. *Аффинной системой координат* (или *репером*) в аффинном пространстве (\mathfrak{A}, V) называется набор $(o; e_1, \dots, e_n)$, состоящий из точки $o \in \mathfrak{A}$, называемой *началом отсчёта*, и базиса e_1, \dots, e_n в векторном пространстве V .

Координатами точки $p \in \mathfrak{A}$ в системе координат $(o; e_1, \dots, e_n)$ называются координаты x^1, \dots, x^n вектора \overline{op} в базисе e_1, \dots, e_n , т. е.

$$\overline{op} = x^1 e_1 + \dots + x^n e_n.$$

Из равенства $\overline{ab} = \overline{ob} - \overline{oa}$ следует, что если даны координаты (a_1, \dots, a_n) точки a и координаты (b_1, \dots, b_n) точки b в системе координат $(o; e_1, \dots, e_n)$, то координаты вектора \overline{ab} в базисе e_1, \dots, e_n равны $(b_1 - a_1, \dots, b_n - a_n)$. Обратно, если $b = a + v$ и даны координаты (a_1, \dots, a_n) точки a в системе координат $(o; e_1, \dots, e_n)$ и координаты (v_1, \dots, v_n) вектора v в базисе e_1, \dots, e_n , то координаты точки b в системе координат $(o; e_1, \dots, e_n)$ равны $(a_1 + v_1, \dots, a_n + v_n)$. В силу этих равенств вектор \overline{pq} можно также обозначать через $q - p$.

Примечание. Систему координат можно задать также $n + 1$ точками $\{o, a_1, \dots, a_n\}$ такими, что векторы $\overline{oa_1}, \dots, \overline{oa_n}$ образуют базис пространства V .

Пусть в аффинном пространстве (\mathfrak{A}, V) даны две аффинные системы координат $(o; e_1, \dots, e_n)$ и $(o'; e_{1'}, \dots, e_{n'})$. Рассмотрим произвольную точку $a \in \mathfrak{A}$. Из равенства $\overline{oa} = \overline{o'a} + \overline{o'o}$, учитывая выведенную нами формулу преобразования координат вектора при переходе к другому базису, получаем формулу преобразования координат точки при переходе к другой системе координат:

$$X = C_{e \rightarrow e'} X' + X_o,$$

где $C_{e \rightarrow e'}$ — матрица перехода от базиса e к базису e' , X — столбец координат точки a в $(o; e_1, \dots, e_n)$, X' — столбец координат точки A в $(o'; e_{1'}, \dots, e_{n'})$, X_o — столбец координат точки o' в $(o; e_1, \dots, e_n)$.

Определение 3. Матрицей перехода от $(o; e_1, \dots, e_n)$ к $(o'; e_{1'}, \dots, e_{n'})$ называется блочная матрица

$$\hat{C} := \begin{pmatrix} C_{e \rightarrow e'} & X_o \\ 0 & 1 \end{pmatrix}.$$

Имеем $\begin{pmatrix} X \\ 1 \end{pmatrix} = \hat{C} \begin{pmatrix} X' \\ 1 \end{pmatrix}.$

47. АФФИННО НЕЗАВИСИМЫЕ СИСТЕМЫ ТОЧЕК. БАРИЦЕНТРИЧЕСКАЯ КОМБИНАЦИЯ ТОЧЕК. ПРИМЕРЫ

Здесь я в основном основывался на Винберге, а у него хронологически плоскости появляются раньше аффинной независимости, так что сначала лучше прочитать вопрос 48.

Определение 1. Точки $p_0, p_1, \dots, p_k \in \mathfrak{A}$ называются *аффинно независимыми*, если

$$\dim \text{aff}\{p_0, p_1, \dots, p_k\} = k,$$

и *аффинно зависимыми* в противном случае.

Из доказательства теоремы 1 в следующем вопросе видно точки p_0, p_1, \dots, p_k аффинно зависимы тогда и только тогда, когда векторы $\overline{p_0 p_1}, \dots, \overline{p_0 p_k}$ линейно зависимы. В то же время, из определения ясно, что свойство точек быть аффинно зависимыми или независимыми не зависит от их нумерации.

Линейные комбинации точек аффинного пространства, вообще говоря, неопределены. Однако некоторым из них можно придать смысл. А именно,

Определение 2. Назовём *барицентрической комбинацией* точек $p_1, \dots, p_k \in \mathfrak{A}$ аффинного пространства (\mathfrak{A}, V) линейную комбинацию вида $\sum_i \lambda_i p_i$, где $\sum_i \lambda_i = 1$, и будем считать её равной точке p , определяемой равенством

$$\overline{op} = \sum_{i=1}^k \lambda_i \overline{op_i},$$

где $o \in \mathfrak{A}$.

Благодаря условию на сумму коэффициентов это определение не зависит от выбора точки o . Действительно, пусть o' — любая другая точка. Тогда

$$\overline{o'p} = \overline{o'o} + \overline{op} = \sum_{i=1}^k \lambda_i (\overline{o'o} + \overline{op_i}) = \sum_{i=1}^k \lambda_i \overline{o'p_i}.$$

В частности,

Определение 3. Центр тяжести системы точек $\{p_1, \dots, p_k\}$ определим как

$$\text{center}(p_1, \dots, p_k) = \frac{p_1 + \dots + p_k}{k}.$$

Предложение 1. Барицентрическая комбинация $\lambda p + \mu q$ для точек $p, q \in \mathfrak{A}$ аффинного пространства (\mathfrak{A}, V) есть точка r , лежащая на прямой pq и обладающая тем свойством, что

$$\overline{pr} = \frac{\mu}{\lambda} \overline{rq}$$

(если $\lambda = 0$, $\mu = 1$, то $r = q$).

Доказательство. Приняв точку r за опорную точку в определении барицентрической комбинации, мы получаем:

$$0 = \lambda \overline{rp} + \mu \overline{rq},$$

откуда и следует требуемое. ■

Вообще, барицентрические координаты — очень далеко идущая тема. Вокруг них развита довольно мощная техника для решения геометрических задач. Почитать про это можно ЗДЕСЬ и здесь в соответствующем приложении.

48. АФФИННОЕ ПОДПРОСТРАНСТВО (ПЛОСКОСТЬ), ЕГО РАЗМЕРНОСТЬ. ГЕОМЕТРИЧЕСКИЙ СМЫСЛ МНОЖЕСТВА РЕШЕНИЙ НЕОДНОРОДНОЙ СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ. ЗАДАНИЕ ПЛОСКОСТИ СИСТЕМОЙ ЛИНЕЙНЫХ УРАВНЕНИЙ

Определение 1. *Плоскостью* в аффинном пространстве (\mathfrak{A}, V) называется подмножество вида $p_0 + W$, где $p_0 \in \mathfrak{A}$, а $W \subseteq V$ — подпространство. Подпространство W часть называют *направляющим* или *касательным* к плоскости $p_0 + W$. *Размерностью* плоскости $p_0 + W$ называют размерность её касательного пространства $\dim W$.

Нульмерная плоскость есть точка. Одномерная плоскость называется *прямой*. Плоскость размерности $n - 1$ называется *гиперплоскостью*.

Очевидно, плоскость $p_0 + W$ является аффинным пространством относительно тех же операций, что и (\mathfrak{A}, V) , причём верно и обратное.

Предложение 1. Если $\mathfrak{B} \subseteq \mathfrak{A}$ и $W \subseteq V$ — подпространство, то аффинное пространство (\mathfrak{B}, W) является плоскостью.

Доказательство. Возьмём произвольную точку $p \in \mathfrak{B}$ и докажем, что множество $p + W$ совпадает с \mathfrak{B} . Из того, что (\mathfrak{B}, W) — аффинное пространство, $\forall q \in \mathfrak{B}$ существует $w \in W$ такой, что $q = p + w$. Значит, $\mathfrak{B} \subseteq (p + W)$. Теперь пусть $w \in W$. Из того, что (\mathfrak{B}, W) — аффинное пространство, $p + w \in \mathfrak{B}$. Значит, $\mathfrak{B} \supseteq (p + W)$. Таким образом, $(p + W) = \mathfrak{B}$, и (\mathfrak{B}, W) есть плоскость $p + W$. ■

Предложение 2. Пусть $(\mathfrak{B}, W) = p + W$ — плоскость и $q \in \mathfrak{B}$. Тогда $(\mathfrak{B}, W) = q + W$.

Доказательство. Докажем, что множества $p + W$ и $q + W$ совпадают. Пусть $p_1 \in (p + W)$. Тогда найдётся вектор $w \in W$ такой, что $p_1 = p + w$. Так вот, $p_1 = p + w = q + (\overline{qp} + w) \in (q + W)$. Обратное включение доказывается аналогично. ■

Определение 2. Для любого подмножества $\mathfrak{B} \subseteq \mathfrak{A}$ и любой точки $p_0 \in \mathfrak{B}$ плоскость

$$p_0 + \langle \overline{p_0 p} : p \in \mathfrak{B} \rangle$$

называется *аффинной оболочкой* множества \mathfrak{B} и обозначается через $\text{aff } \mathfrak{B}$.

Предложение 3. Если $p + W$ — плоскость, содержащая все точки $\mathfrak{B} \subseteq \mathfrak{A}$, то $(p + W) \supseteq \text{aff } \mathfrak{B}$.

Доказательство. Фиксируем точку $q \in \mathfrak{B}$. В касательном пространстве $\text{aff } \mathfrak{B}$ выберем базис $\overline{qq_1}, \dots, \overline{qq_k}$. Тогда если $q' \in \text{aff } \mathfrak{B}$, то $q' = q + \lambda_1 \overline{qq_1} + \dots + \lambda_k \overline{qq_k}$ для некоторых $\lambda_i \in \mathbb{F}$, причём все векторы $\overline{qq_i}$ содержатся в W , т. к. $q \in W$ и $q_i \in W$ по условию. Значит, $q' \in W$ ($i = 1, \dots, k$). ■

Хронологически вот здесь появляется определение аффинно независимых точек из предыдущего вопроса.

Теорема 1. Через любые $k + 1$ точек аффинного пространства проходит плоскость размерности не больше k ; при этом, если эти точки аффинно независимы, то через них проходит единственная плоскость размерности k .

Доказательство. Пусть (\mathfrak{B}, W) — плоскость и $p_0, p_1, \dots, p_k \in \mathfrak{B}$. Тогда $\text{aff}\{p_0, p_1, \dots, p_k\}$ есть плоскость размерности не больше k , проходящая через p_0, p_1, \dots, p_k . Если $\dim \text{aff}\{p_0, p_1, \dots, p_k\}$ (точки аффинно независимые), то векторы $\overline{p_0 p_1}, \dots, \overline{p_0 p_k}$ линейно независимы в касательном пространстве аффинной оболочки и эта аффинная оболочка является единственной k -мерной плоскостью, которая содержит наши точки. ■

Выше уже было про то, что множество решений неоднородной СЛУ есть аффинное пространство. Из этого следует, что это плоскость в аффинном пространстве \mathbb{A}^n . Оказывается, верно и обратное.

Теорема 2. Всякая плоскость есть множество решений некоторой системы линейных уравнений.

Доказательство. Пусть $p_0 + W$ — некоторая плоскость. Подпространство W может быть задано системой однородных линейных уравнений. Заменив свободные члены этих уравнений значениями, принимаемыми левыми частями в точке p_0 , мы получим систему линейных уравнений, задающую нашу плоскость. ■

Здесь я приведу решения двух очень похожих друг на друга задач. Решение первой есть в Винберге, оно почти дословно переписано оттуда, вторую решал я сам, так что решения отличаются.

Задача 12 (Из Винберга). Если в поле \mathbb{F} более двух элементов, то непустое подмножество $P \subseteq \mathfrak{A}$ точек аффинного пространства (\mathfrak{A}, V) (над полем \mathbb{F}) является плоскостью тогда и только тогда, когда вместе с любыми двумя различными точками оно содержит проходящую через них прямую.

▷ Необходимость очевидна, докажем достаточность. Пусть $P \subseteq \mathfrak{A}$ — непустое подмножество, обладающее указанным свойством. Фиксируем произвольную точку $p_0 \in P$ и рассмотрим подмножество $W := \{v \in V : p_0 + v \in P\} \subseteq V$. Нам нужно доказать, что W — подпространство. Ясно, что оно содержит $\mathbf{0}$. Далее, если $w \in W$ — любой ненулевой вектор и $\lambda \in \mathbb{F}$, то точка $p_0 + \lambda w$ лежит на прямой, проходящей через p_0 и $p_0 + w$, следовательно, $\lambda w \in W$. Докажем наконец, что если $w_1, w_2 \in W$ — непропорциональные векторы, то $w_1 + w_2 \in W$. Пусть $\lambda \in \mathbb{F} \setminus \{0, 1\}$. Легко видеть, что точка $p = p_0 + w_1 + w_2$ лежит на прямой проходящей через точки $p_1 = p_0 + \lambda w_1 \in P$ и $p_2 = p_0 + \frac{\lambda}{\lambda-1} w_2 \in P$, а именно

$$p = \frac{1}{\lambda} p_1 + \frac{\lambda-1}{\lambda} p_2.$$

Следовательно, $p \in P$ и $w_1 + w_2 \in W$. ◀

Задача 13 (А. А. Клячко). Если $\text{char } \mathbb{F} \neq 2$, то непустое подмножество $P \subseteq \mathfrak{A}$ точек аффинного пространства (\mathfrak{A}, V) (над полем \mathbb{F}) является плоскостью тогда и только тогда, когда вместе с любыми двумя различными точками оно содержит проходящую через них прямую.

▷ Необходимость очевидна, докажем достаточность. Пусть $P \subseteq \mathfrak{A}$ — непустое подмножество, обладающее указанным свойством. Фиксируем произвольную точку $p_0 \in P$ и рассмотрим подмножество $W := \{v \in V : p_0 + v \in P\} \subseteq V$. Нам нужно доказать, что W — подпространство. Ясно, что оно содержит $\mathbf{0}$. Далее, если $x, y \in W$, то $\text{aff}(p+x, p+y) \subseteq P$ по условию, так что $\forall \lambda \in \mathbb{F}$

$$(1-\lambda)(p+x) + \lambda(p+y) = p + (1-\lambda)x + \lambda y = p + x + \lambda(y-x) \in P.$$

Подставляя $x = \mathbf{0}$, получаем $p + \lambda y \in P$, значит, $\lambda y \in W$. А т. к. $\text{char } \mathbb{F} \neq 2$, то можем подставить $\lambda = \frac{1}{2}$, при этом получим $x + y \in P$. ◀

Здесь я воспользовался тем, что прямая через две точки есть множество барицентрических комбинаций этих двух точек. Это достаточно очевидно, однако верен и более общий факт.

Задача 14 (Из Винберга). Аффинная оболочка множества $\mathfrak{B} \subseteq \mathfrak{A}$ аффинного пространства (\mathfrak{A}, V) есть совокупность всех барицентрических комбинаций точек из \mathfrak{B} .

▷ Если $\mathfrak{B} = \emptyset$, доказывать нечего. Иначе можно фиксировать точку $o \in \mathfrak{B}$. Обозначим через $\tilde{\mathfrak{B}}$ совокупность барицентрических комбинаций точек из \mathfrak{B} , а через W — касательное подпространство к плоскости $\text{aff } \mathfrak{B}$.

Пусть $p \in \tilde{\mathfrak{B}}$, т. е. $\overline{op} = \sum_{q \in \mathfrak{B}} \lambda_i \overline{oq}$. Все векторы \overline{oq} лежат в W , а значит, и $\overline{op} \in W$. Из $o \in \mathfrak{B} \subseteq \text{aff } \mathfrak{B}$ следует $p \in \text{aff } \mathfrak{B}$, таким образом $\tilde{\mathfrak{B}} \subseteq \text{aff } \mathfrak{B}$.

Обратно, пусть $p \in \text{aff } \mathfrak{B}$. Фиксируем базис e_1, \dots, e_n в W , тогда $\overline{op} = \sum_{i=1}^n \lambda_i e_i$, причём можно считать, что $\sum_{i=1}^n \lambda_i \neq 0$. т. к. тогда можно заменить какой-нибудь базисный вектор на пропорциональный ему и сумма изменится. В этом случае заменим λ_i на $\frac{\lambda_i}{\sum_i \lambda_i}$, а базисные векторы «удлиним» на $\sum_i \lambda_i$, получим барицентрическую комбинацию. ◀

49. ВЗАИМНОЕ РАСПОЛОЖЕНИЕ ДВУХ ПЛОСКОСТЕЙ. ПЕРЕСЕЧЕНИЕ И АФФИННАЯ ОБОЛОЧКА ДВУХ АФФИННЫХ ПЛОСКОСТЕЙ. РАЗМЕРНОСТЬ АФФИННОЙ ОБОЛОЧКИ

Пусть $P_1 = p_1 + W_1$ и $P_2 = p_2 + W_2$ — плоскости.

Очевидно, что если они пересекаются и p_0 — одна из точек этого пересечения, то

$$P_1 \cap P_2 = p_0 + (W_1 \cap W_2).$$

Теорема 1. Плоскости P_1 и P_2 пересекаются тогда и только тогда, когда

$$\overline{p_1 p_2} \in W_1 + W_2.$$

Доказательство. Плоскости P_1 и P_2 пересекаются тогда и только тогда, когда существуют векторы $w_1 \in W_1$ и $w_2 \in W_2$, что

$$p_1 + w_1 = p_2 + w_2.$$

Это равенство может быть переписано в виде

$$\overline{p_1 p_2} = w_1 - w_2.$$

Поэтому существование таких векторов w_1 и w_2 как раз и означает, что $\overline{p_1 p_2} \in W_1 + W_2$. ■

Определение 1. Плоскости P_1 и P_2 называются *параллельными*, если $W_1 \subseteq W_2$ или $W_2 \subseteq W_1$ и *скрещивающимися*, если $P_1 \cap P_2 = \emptyset$ и $W_1 \cap W_2 = \emptyset$.

Теорема 2. Пусть $\pi_1 = p_1 + W_1$ и $\pi_2 = p_2 + W_2$ — плоскости. Тогда

$$\text{aff}(P_1 \cup P_2) = p_1 + \langle \overline{p_1 p_2}, W_1 + W_2 \rangle,$$

причём если $P_1 \cap P_2 \neq \emptyset$, то

$$\dim \text{aff}(P_1 \cup P_2) + \dim(W_1 + W_2),$$

иначе

$$\dim \operatorname{aff}(P_1 \cup P_2) = \dim(W_1 + W_2) + 1.$$

Доказательство. Обозначим $P := p_1 + \langle \overline{p_1 p_2}, W_1 + W_2 \rangle$. Понятно, что $P_1, P_2 \subseteq P$. Обратно, т. к. $p_1 \in \operatorname{aff}(P_1 \cup P_2)$, то $\operatorname{aff}(P_1 \cup P_2) = p_1 + X$ для некоторого подпространства $X \subseteq V$. Поскольку $p_2 \in \operatorname{aff}(P_1 \cup P_2)$, имеем $\overline{p_1 p_2} \in X$. Т. к. $\forall w_1 \in W_1$ точка $p_1 + w_1 \in P_1 \subseteq \operatorname{aff}(P_1 \cup P_2)$, то $w_1 \in X$. Далее, для любого $w_2 \in W_2$ имеем $p_2 + w_2 \in P_2 \subseteq \operatorname{aff}(P_1 \cup P_2)$ и $p_2 + w_2 = a + (\overline{p_1 p_2} + w_2) \in \operatorname{aff}(P_1 \cup P_2)$, отсюда $\overline{p_1 p_2} + w \in X$. Поскольку $\overline{p_1 p_2} \in X$, имеем $w \in X$. Получилось, что $\langle \overline{p_1 p_2}, W_1 + W_2 \rangle \subseteq X$, т. е. $P \subseteq \operatorname{aff}(P_1 \cup P_2)$, а значит, $P = \operatorname{aff}(P_1 \cup P_2)$, что и требовалось.

Из теоремы 1, если $P_1 \cap P_2 \neq \emptyset$, то $\overline{p_1 p_2} \in W_1 + W_2$ и $\dim \operatorname{aff}(P_1 \cup P_2) = \dim \langle \overline{p_1 p_2}, W_1 + W_2 \rangle = \dim(W_1 + W_2)$. А иначе $\overline{p_1 p_2} \notin W_1 + W_2$ и к $\dim(W_1 + W_2)$ нужно будет прибавить 1. ■

50. АФФИННОЕ ОТОБРАЖЕНИЕ И ЕГО ЛИНЕЙНАЯ ЧАСТЬ. КОМПОЗИЦИЯ АФФИННЫХ ОТОБРАЖЕНИЙ. КРИТЕРИЙ ОБРАТИМОСТИ В ТЕРМИНАХ ЛИНЕЙНОЙ ЧАСТИ. МАТРИЦА АФФИННОГО ОТОБРАЖЕНИЯ. СУЩЕСТВОВАНИЕ И ЕДИНСТВЕННОСТЬ АФФИННОГО ОТОБРАЖЕНИЯ, ДЕЙСТВУЮЩЕГО НА СИСТЕМУ ИЗ $n + 1$ АФФИННО НЕЗАВИСИМЫХ ТОЧЕК n -МЕРНОГО АФФИННОГО ПРОСТРАНСТВА

Пусть (\mathfrak{A}, V) и (\mathfrak{A}', V') — аффинные пространства над одним и тем же полем \mathbb{F} .

Определение 1. Аффинным отображением пространства (\mathfrak{A}, V) в пространство (\mathfrak{A}', V') называется всякое отображение $f : \mathfrak{A} \rightarrow \mathfrak{A}'$, обладающее свойством

$$f(p + x) = f(p) + \varphi(x), \quad \forall p \in \mathfrak{A}, x \in V,$$

где φ — некоторое линейное отображение пространства V в пространство V' .

Из определения вытекает, что $\varphi(\overline{pq}) = \overline{f(p)f(q)} \forall p, q \in \mathfrak{A}$. Тем самым, линейное отображение φ однозначно задаётся по f . Оно называется *дифференциалом* f и обозначается через df .

Введём на (\mathfrak{A}, V) , приняв за начало отсчёта какие-то точки $o \in \mathfrak{A}$ и $o' \in \mathfrak{A}'$. Полагая в выражении из определения аффинного отображения $p = o$, мы получаем следующее представление аффинного отображения f в векторизованной форме:

$$f(x) = \varphi(x) + b, \quad b \in V',$$

где $b := \overline{o'f(o)}$.

Обратно, как легко проверить, для любого линейного отображения $\varphi : V \rightarrow V'$ и любого вектора $b \in V'$ отображение $f(x) := \varphi(x) + b$ аффинно и его дифференциал равен φ .

Отсюда, в свою очередь, получается запись отображения f в координатах:

$$y^i = a_j^i x^j + b^i \quad (i = 1, \dots, m),$$

где x^1, \dots, x^n — координаты точки x , а y^1, \dots, y^m — координаты точки $y := f(x)$.

Определение 2. Матрицей аффинного отображения называется матрица $A = (a_j^i)$ из последней записи.

Пусть (\mathfrak{A}'', V'') — ещё одно аффинное пространство, а $g : \mathfrak{A}' \rightarrow \mathfrak{A}''$ — аффинное отображение.

Предложение 1. Отображение $gf : \mathfrak{A} \rightarrow \mathfrak{A}''$ является аффинным, причём

$$d(gf) = dg \cdot df.$$

Доказательство. При $p \in \mathfrak{A}$ и $x \in V$ имеем:

$$(gf)(p + x) = g(f(p + x)) = g(f(p) + df(x)) = g(f(p)) + dg(df(x)) = (gf)(p) + (dg \cdot df)(x).$$

■

Примечание. При $\mathbb{F} = \mathbb{R}$ дифференциал аффинного отображения есть частный случай дифференциала произвольного гладкого отображения, рассматриваемого в анализе, а последнее предложение даёт частный случай формулы для дифференциала сложной функции.

Предложение 2. Аффинное отображение биективно тогда и только тогда, когда его дифференциал биективен.

Доказательство. Выберем сначала точки отсчёта o и o' в \mathfrak{A} и \mathfrak{A}' соответственно таким образом, чтобы $f(o) = o'$. Тогда отображение f в векторизованной форме будет совпадать со своим же дифференциалом, откуда и следует доказываемое утверждение. ■

Определение 3. Биективное аффинное отображение называется *изоморфизмом* аффинных пространств. Аффинные пространства, между которыми существует изоморфизм, называются *изоморфными*.

Следствие 1. Конечномерные аффинные пространства (над одним и тем же полем) изоморфны тогда и только тогда, когда они имеют одинаковую размерность.

Очевидно, что при аффинном отображении $f : \mathfrak{A} \rightarrow \mathfrak{A}'$ всякая плоскость $P = p + W$ пространства (\mathfrak{A}, V) переходит в плоскость $f(P) = f(p) + df(W)$ пространства (\mathfrak{A}', V') . Если f биективно, то $\dim f(P) = \dim P$.

Теорема 1. Пусть $\{p_0, p_1, \dots, p_n\}$ и $\{q_0, q_1, \dots, q_n\}$ — две системы аффинно-независимых точек в n -мерном аффинном пространстве (\mathfrak{A}, V) . Тогда существует единственное аффинное отображение $f : \mathfrak{A} \rightarrow \mathfrak{A}$, переводящее p_i в $q_i \forall i = 0, 1, \dots, n$.

Доказательство. Существует единственное линейное отображение $\varphi : V \rightarrow V$, переводящее базис $\overline{p_0 p_1}, \dots, \overline{p_0 p_n}$ в базис $\overline{q_0 q_1}, \dots, \overline{q_0 q_n}$. Векторизуем пространство \mathfrak{A} , приняв за начало отсчёта точку p_0 . Тогда искомое аффинное преобразование f записывается в виде

$$f(x) := \varphi(x) + \overline{p_0 q_0}.$$

■

51. АФФИННЫЕ ПРЕОБРАЗОВАНИЯ. РАЗЛОЖЕНИЕ БИЕКТИВНОГО АФФИННОГО ПРЕОБРАЗОВАНИЯ В КОМПОЗИЦИЮ ПАРАЛЛЕЛЬНОГО ПЕРЕНОСА И ПРЕОБРАЗОВАНИЯ С НЕПОДВИЖНОЙ ТОЧКОЙ

Определение 1. Аффинное отображение аффинного пространства (\mathfrak{A}, V) в себя называется *аффинным преобразованием*.

Биективные аффинные преобразования образуют группу, называемую *полной аффинной группой* пространства (\mathfrak{A}, V) и обозначаемую через $\text{GA}(\mathfrak{A})$. В силу предложения 1 из вопроса 50 отображение

$$d : \text{GA}(\mathfrak{A}) \rightarrow \text{GL}(V)$$

является гомоморфизмом групп. Его ядро есть *группа параллельных переносов*

$$t_a : p \mapsto p + a \quad (a \in V).$$

Обозначим её через $\text{Trans}(V)$.

Предложение 1. Для любых $f \in \text{GA}(\mathfrak{A})$ и $a \in V$ имеем

$$f t_a f^{-1} = t_{df(a)}.$$

Доказательство. Применяя преобразование ft_af^{-1} к точке $q = f(p)$, получаем

$$ft_af^{-1}(q) = ft_a(p) = f(p + a) = f(p) + df(a) = q + df(a).$$

■

Если фиксировано начало отсчёта $o \in \mathfrak{A}$ и тем самым аффинное пространство (\mathfrak{A}, V) отождествлено с векторным пространством V , то группа $GL(V)$ становится подгруппой группы $GA(\mathfrak{A})$. Это не что иное, как стабилизатор точки o в группе $GA(\mathfrak{A})$. Из записи аффинных преобразований в векторизованной форме

$$f(x) = \varphi(x) + b$$

следует, что всякое аффинное преобразование $f \in GA(\mathfrak{A})$ единственным образом представляется в виде

$$f = t_b\varphi, \quad \varphi \in GL(V), b \in V.$$

Ясно, что $\varphi = df$ не зависит от выбора начала отсчёта, а вот вектор $b = \overline{of(o)}$ от этого, вообще говоря, зависит. Вспомнив, что $\varphi(\mathbf{0}) = \mathbf{0}$, мы фактически доказали следующее утверждение.

Теорема 1. Любое биективное аффинное преобразование представимо в виде композиции параллельного переноса и преобразования с неподвижной точкой.

52. АФФИННОЕ ЕВКЛИДОВО ПРОСТРАНСТВО. ТЕОРЕМА ОБ ИЗОМОРФИЗМЕ. РАССТОЯНИЕ МЕЖДУ ДВУМЯ ПЛОСКОСТЯМИ

Здесь мы будем рассматривать аффинные пространства (\mathfrak{A}, V) , ассоциированные с евклидовым пространством V , которые будем называть *аффинными евклидовыми пространствами*.

Среди всех аффинных систем координат в аффинном евклидовом пространстве выделяются системы координат, определяемые ортонормированными реперами. Они называются *прямоугольными системами координат*.

Определение 1. Расстояние ρ между точками $p, q \in \mathfrak{A}$ аффинного евклидова пространства (\mathfrak{A}, V) определяется по формуле

$$\rho(p, q) := |\overline{pq}|.$$

Определение 2. Аффинные евклидовы пространства (\mathfrak{A}, V) и (\mathfrak{A}', V') называются *изоморфными*, если между соответствующими аффинными пространствами существует *изометрический изоморфизм* Φ , т. е. такой, который сохраняет длины векторов:

$$|\overline{\Phi(p)\Phi(q)}| = |\overline{pq}| \quad \forall p, q \in \mathfrak{A}.$$

Здесь верно утверждение, аналогичное уже доказанному про аффинные пространства.

Теорема 1 (Об изоморфизме). Аффинные евклидовы пространства изоморфны тогда и только тогда, когда они имеют одинаковую размерность.

Доказательство. Чтобы доказать эту теорему, нужно перестать думать. Если размерности одинаковые, то берём любое ортогональное отображение $V \rightarrow V'$ и принимаем его за дифференциал нашего отображения, можно к нему любой вектор ещё добавить, если вдруг хочется. Вот и искомый изометрический изоморфизм. А если они изоморфны, то конечно же размерности должны быть равны, иначе между ними нельзя установить биекцию. ■

Нахождение расстояния от точки $p \in \mathfrak{A}$ до плоскости $P = p_0 + W$ с помощью векторизации сводится к нахождению расстояния от вектора $x := \overline{p_0p} \in V$ до подпространства W . А именно, пусть $x = \text{pr}_W x + \text{ort}_W x$. Приняв точку p_0 за начало отсчёта, мы получаем

$$\rho(p, P) = \rho(x, W) = |\text{ort}_W x|.$$

Точка $q := p_0 + \text{pr}_W x$ является «основанием перпендикуляра, опущенного из точки p на плоскость P ».

Теорема 2. Расстояние между плоскостями $P_1 = p_1 + W_1$ и $P_2 = p_2 + W_2$ аффинного евклидова пространства находится по формуле

$$\rho(P_1, P_2) = |\text{ort}_{W_1+W_2} \overline{p_1 p_2}|.$$

Доказательство. Разложим V в прямую сумму $V = (W_1 + W_2) \oplus (U + W)^\perp$ и представим вектор $\overline{p_1 p_2}$ в виде

$$\overline{p_1 p_2} = \text{pr}_{W_1+W_2} \overline{p_1 p_2} + \text{ort}_{W_1+W_2} \overline{p_1 p_2}.$$

Выберем любые две точки $q_1 = (p_1 + w_1) \in P_1$, $q_2 = (p_2 + w_2) \in P_2$. Тогда по теореме Пифагора получаем, что

$$\begin{aligned} \rho(q_1, q_2) &= |\overline{p_1 p_2} + w_2 - w_1|^2 = |(\text{pr}_{W_1+W_2} \overline{p_1 p_2} + w_2 - w_1) + \text{ort}_{W_1+W_2} \overline{p_1 p_2}|^2 = \\ &= |\text{pr}_{W_1+W_2} \overline{p_1 p_2} + w_2 - w_1|^2 + |\text{ort}_{W_1+W_2} \overline{p_1 p_2}|^2 \geq |\text{ort}_{W_1+W_2} \overline{p_1 p_2}|^2. \end{aligned}$$

Таким образом, $\rho(q_1, q_2) \geq |\text{ort}_{W_1+W_2} \overline{p_1 p_2}|$, причём равенство достигается для таких точек q_1 и q_2 , что

$$\text{pr}_{W_1+W_2} \overline{p_1 q_1} = w_1 - w_2.$$

Т. к. $\text{pr}_{W_1+W_2} \overline{p_1 p_2} \in W_1 + W_2$, такие векторы $w_1 \in W_1$, $w_2 \in W_2$ существуют. ■

53. ОРТОГОНАЛЬНЫЕ АФФИННЫЕ ПРЕОБРАЗОВАНИЯ (ДВИЖЕНИЯ). СОБСТВЕННЫЕ И НЕСОБСТВЕННЫЕ ДВИЖЕНИЯ. РАЗЛОЖЕНИЕ ДВИЖЕНИЯ В КОМПОЗИЦИЮ ПАРАЛЛЕЛЬНОГО ПЕРЕНОСА И ОРТОГОНАЛЬНОГО ПРЕОБРАЗОВАНИЯ С НЕПОДВИЖНОЙ ТОЧКОЙ

Определение 1. *Движением (изометрией)* аффинного евклидова пространства (\mathfrak{A}, V) называется всякое его аффинное преобразование, дифференциал которого является ортогональным оператором.

Во-первых, ясно, что всякое движение биективно. А во-вторых,

Теорема 1. Всякое движение сохраняет расстояния между точками и обратно, всякое аффинное преобразование, сохраняющее расстояния между точками, является движением.

Доказательство. \Rightarrow . Пусть f — движение, тогда $\forall p, q \in \mathfrak{A}$ имеем

$$\rho(f(p), f(q)) = |\overline{f(p)f(q)}| = |\varphi(\overline{pq})| = |\overline{pq}| = \rho(p, q),$$

т. е. f сохраняет расстояния.

\Leftarrow . Обратно, пусть f сохраняет расстояния. Тогда выполнено равенство

$$|\overline{pq}| = |\overline{f(p)f(q)}| = |\varphi(\overline{pq})|.$$

Отсюда следует, что дифференциал сохраняет длины всех векторов. Значит, он ортогонален, и f — движение. ■

Движения аффинного евклидова пространства образуют группу, обозначаемую $\text{Isom}(\mathfrak{A}, V)$.

Определение 2. Движение f называется *собственным*, если $\det df = 1$, иначе — *несобственным*.

Собственные движения образуют подгруппу в $\text{Isom}(\mathfrak{A}, V)$, обозначаемую $\text{Isom}_+(\mathfrak{A}, V)$.

Пример 1. Важным примером несобственного движения является (ортогональное) *отражение* r_H относительно гиперплоскости H . Пусть e — единичный вектор, ортогональный H . Всякую точку $p \in \mathfrak{A}$ можно единственным образом представить в виде $p = q + \lambda e$ ($q \in H$). По определению, $r_H p := q - \lambda e$. Дифференциал отражения r_H есть (ортогональное) отражение относительно направляющего подпространства гиперплоскости H в пространстве V . Пусть H_1 и H_2 — две гиперплоскости. Если они параллельны, то $dr_{H_1} = dr_{H_2}$ и, следовательно,

$$d(r_{H_1} r_{H_2}) = dr_{H_1} \cdot dr_{H_2} = \text{id}.$$

В этом случае $r_{H_1} r_{H_2}$ — параллельный перенос на удвоенный общий перпендикуляр плоскостей H_1 и H_2 . Если же H_1 и H_2 пересекаются по $(n-2)$ -мерной плоскости P , то $r_{H_1} r_{H_2}$ — поворот на вокруг P на удвоенный угол между H_1 и H_2 .

За точными определениями написанных здесь слов см. приложение про геометрию аффинных евклидовых пространств.

Теорема 2. Пусть $\Phi : \mathfrak{A} \rightarrow \mathfrak{A}$ — движение аффинного евклидова конечномерного пространства (\mathfrak{A}, V) с дифференциалом φ . Тогда найдётся вектор $u \in V$ такой, что $\varphi(u) = u$ и $\Phi = t_u \Psi$, где Ψ — движение с неподвижной точкой.

Доказательство. Пусть $a \in \mathfrak{A}$ — произвольная точка. Рассмотрим вектор

$$v := \overline{a\Phi(a)}.$$

Если существует собственный вектор линейного оператора φ с собственным значением 1, то обозначим через W собственное подпространство линейного оператора φ , отвечающее собственному значению 1, иначе $U := \{0\}$. Т. к. φ — ортогональный оператор, а W — инвариантное подпространство относительно φ , то по важной лемме подпространство W^\perp тоже является инвариантным относительно φ . Кроме того, линейный оператор $\varphi - \text{id}$ действует на W^\perp невырожденным образом. Т. к. $V = W \oplus W^\perp$, вектор v представляется в виде $v = \text{pr}_W v + \text{ort}_W v$, причём $\varphi(\text{pr}_W v) = \text{pr}_W v$. Положим $\Psi := t_u^{-1} \Phi$ (где $u := \text{pr}_W v$) и найдём для Ψ неподвижную точку в виде $b = a + w$, $w \in W^\perp$. Имеем

$$\begin{aligned} \Psi(a + w) &= (t_{-u} \Phi)(a + w) = t_{-u}(\Phi(a) + \varphi(w)) = t_{-u}((a + v) + \varphi(w)) = \\ &= a + (\text{pr}_W v + \text{ort}_W v) + \varphi(w) - \text{pr}_W v = a + \text{ort}_W v + \varphi(w) = a + w + \text{ort}_W v + (\varphi - \text{id})w. \end{aligned}$$

Хотим найти такой вектор $w \in W^\perp$, что $\text{ort}_W v + (\varphi - \text{id})w = 0$. А такой вектор существует в силу невырожденности $(\varphi - \text{id})|_{W^\perp}$ и равен $-(\varphi - \text{id})^{-1} \text{ort}_W v$. ■

54. НЕКОТОРЫЕ АФФИННЫЕ И ЛИНЕЙНЫЕ ГРУППЫ

Всё, что тут можно написать, уже было в предыдущих билетах.

55. АФФИННО-КВАДРАТИЧНЫЕ ФУНКЦИИ. ЦЕНТР. ПРИВЕДЕНИЕ АФФИННО-КВАДРАТИЧНОЙ ФУНКЦИИ К КАНОНИЧЕСКОМУ ВИДУ ЗАМЕНОЙ АФФИННОЙ СИСТЕМЫ КООРДИНАТ

Будем считать, что $\text{char } \mathbb{F} \neq 2$.

Определение 1. *Аффинно-квадратичной функцией* на аффинном пространстве (\mathfrak{A}, V) называется всякая функция $Q : \mathfrak{A} \rightarrow \mathbb{F}$, имеющая в векторизованной форме вид

$$Q(x) = q(x) + l(x) + c,$$

где q — квадратичная функция, l — линейная функция, $c \in \mathbb{F}$ — константа.

Пусть \widehat{q} — соответствующая квадратичной функции q билинейная форма.

Лемма 1. При переносе начала отсчёта o в точку $o' = o + a$ ($a \in V$) слагаемые выражения из определения аффинно-квадратичной функции преобразуются следующим образом:

$$q'(x) = q(x), \quad l'(x) = 2\widehat{q}(a, x) + l(x), \quad c' = q(a) + l(a) + c.$$

Доказательство. Непосредственно проверяем:

$$\begin{aligned} Q(o' + x) &= Q(o + a + x) = q(a + x) + l(a + x) + c = \\ &= q(a) + 2\widehat{q}(a, x) + q(x) + l(a) + l(x) + c = \\ &= q(x) + (2\widehat{q}(x) + l(x)) + (q(a) + l(a) + c). \end{aligned}$$

■

В частности, квадратичная функция q не зависит от выбора начала отсчёта. В координатах выражение аффинно-квадратичной функции принимает вид

$$Q(x) = \sum_{i,j} x^i x^j a_{ij} + \sum_i x^i b_i + c \quad (a_{ij} = a_{ji}).$$

Коэффициентам b_i и c можно придать следующий смысл:

$$c = Q(o), \quad b_i = \frac{\partial Q}{\partial x_i}(o).$$

Определение 2. Линейная функция $l(x) = \sum_i x^i b_i$ называется *дифференциалом* функции Q в точке o и обозначается $d_o Q$.

В случае $\mathbb{F} = \mathbb{R}$ это согласуется с обычным определением дифференциала.

Определение 3. Точка o называется *центром* аффинно-квадратичной функции Q , если

$$Q(o + x) = Q(o - x) \quad \forall x \in V.$$

Ясно, что это имеет место тогда и только тогда, когда $d_o Q = \mathcal{O}$. Поэтому множество всех центров функции Q задаётся системой линейных уравнений

$$\frac{\partial Q}{\partial x^1} = \dots = \frac{\partial Q}{\partial x^n} = 0.$$

Оно либо является плоскостью какой-то размерности, либо пусто. Легко видеть, что матрица коэффициентов этой системы есть удвоенная матрица квадратичной функции q . Следовательно, если q невырождена, то Q имеет единственный центр.

Теорема 1. Для любой аффинно-квадратичной функции $Q : \mathfrak{A} \rightarrow \mathbb{F}$, не являющейся аффинно-линейной, существует система координат $(o; e_1, \dots, e_n)$, в которой Q имеет один из следующих видов (*):

$$Q(a) = \lambda_1 x_1^2 + \dots + \lambda_r x_r^2 + \lambda_{r+1}, \quad Q(a) = \lambda_1 x_1^2 + \dots + \lambda_r x_r^2 + 2x_{r+1},$$

где числа $\lambda_1, \dots, \lambda_r$ отличны от нуля.

Доказательство. Существует базис e'_1, \dots, e'_n векторного пространства V , в котором матрица квадратичной функции q диагональна:

$$B = \begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix},$$

где числа $\lambda_1, \dots, \lambda_r$ отличны от нуля, т. к. функция Q не является аффинно-линейной.

Фиксируем произвольную точку $o' \in \mathfrak{A}$. В системе координат $(o'; e'_1, \dots, e'_n)$ функция Q примет вид

$$Q(o' + v) = \lambda_1 x_1'^2 + \dots = \lambda_r x_r'^2 + 2u_1' x_1' + \dots + 2u_n' x_n' + \alpha'.$$

Перенос начала координат в точку o'' , сводящийся к замене координат

$$\begin{aligned} x_i'' &= x_i' + \frac{u_i'}{\lambda_i}, & i &= 1, \dots, r, \\ x_i'' &= x_i', & i &= r+1, \dots, n, \end{aligned}$$

приводит функцию Q к виду

$$Q(o'' + v) = \lambda_1 x_1''^2 + \dots + \lambda_r x_r''^2 + 2u_{r+1}' x_{r+1}'' + \dots + 2u_n' x_n'' + \alpha''.$$

Если $u_{r+1}' = \dots = u_n' = 0$, то Q с точностью до обозначения имеет вид, как в первой строчке (*), в этом случае Q имеет центр.

Если хотя бы один из коэффициентов u_{r+1}' отличен от нуля (для определённости считаем, что это u_{r+1}'), то ещё одна замена координат

$$\begin{aligned} x_{r+1} &= u_{r+1}' x_{r+1}'' + \dots + u_n' x_n'' + \frac{\alpha''}{2}, \\ x_i &= x_i'', & i &\neq r+1 \end{aligned}$$

приведёт Q к виду, как во второй строчке (*), в этом случае Q не имеет центра. ■

56. ПОВЕРХНОСТИ ВТОРОГО ПОРЯДКА (КВАДРИКИ). ПЕРЕСЕЧЕНИЕ КВАДРИКИ С ПРЯМОЙ. ЦЕНТР И ВЕРШИНА КВАДРИКИ

Определение 1. Множество вида $\Gamma(Q) := \{p \in \mathfrak{A} : Q(p) = 0\}$, где Q — аффинно-квадратичная функция, если только оно не пусто и не является плоскостью, называется *квадрикой* или *гиперповерхностью второго порядка*.

Определение 2. Точка $o \in \mathfrak{A}$ называется *центром* квадрики, если эта квадрика симметрична относительно o , т. е. вместе со всякой точкой $o + x$ ($x \in V$) содержит точку $o - x$. Центр квадрики, лежащий на ней самой, называется её *вершиной*.

Квадрика, имеющая (хотя бы один) центр, она называется *центральной*.

Очевидно, что всякий центр аффинно-квадратичной функции Q является центром квадрики $\Gamma(Q)$. Как будет показано ниже, верно и обратное.

Предложение 1. Любая прямая либо целиком лежит на квадрике, либо пересекается с ней не более чем в двух точках.

Доказательство. Т. к. начало отсчёта o может быть выбрано в любой точке, то без ограничения общности можно считать, что прямая проходит через o . Пусть функция Q в векторизованной форме имеет вид $Q(x) = q(x) + l(x) + c$. Тогда пересечение прямой $L = o + \langle x \rangle = \{o + tx : t \in \mathbb{F}\}$ ($x \in V$) с квадрикой $\Gamma(Q)$ определяется условием

$$Q(tx) = t^2 q(x) + tl(x) + c,$$

представляющим собой квадратное уравнение относительно t . Если все коэффициенты этого уравнения равны 0, то $L \subset \Gamma(Q)$; в противном случае оно имеет не более двух корней, а это означает, что пересечение $L \cap \Gamma(Q)$ содержит не более двух точек. ■

Предложение 2. Если o — вершина квадрики Γ , то вместе с любой точкой $p \neq o$ квадрика Γ содержит всю прямую op .

Доказательство. Пусть $p = o + x$ ($x \in V$); тогда Γ содержит три различные точки $o, o + x, o - x$ прямой op и, следовательно, — всю прямую. ■

Определение 3. Всякое подмножество аффинного пространства, содержащее точку o и вместе с любой точкой $p \neq o$ всю прямую op называется *конусом* с вершиной в точке o . Квадрика называется *конической*, если она имеет (хотя бы одну) вершину.

Предложение 3. Всякая квадрика содержит точку, не являющуюся её вершиной.

Доказательство. Если бы все точки квадрики были её вершинами, то в силу предложения 2 вместе с любыми двумя точками она содержала бы проходящую через них прямую и, согласно задаче 7, была бы плоскостью, а это противоречит определению квадрики. ■

Очевидно, что пропорциональные аффинно-квадратичные функции определяют одну и ту же квадриду. Обратное утверждение не столь очевидно.

Теорема 1. Пусть Γ — квадрика в аффинном пространстве над бесконечным полем \mathbb{F} . Если $\Gamma = \Gamma(Q_1) = \Gamma(Q_2)$ для каких-то аффинно-квадратичных функций Q_1, Q_2 , то эти функции пропорциональны.

Доказательство. Возьмём в качестве начала отсчёта какую-нибудь точку o квадрики Γ , не являющуюся её вершиной. Тогда в векторизованной форме

$$Q_1(x) = q_1(x) + \ell_1(x), \quad Q_2(x) = q_2(x) + \ell_2(x),$$

где $\ell_1, \ell_2 \neq 0$. Точки пересечения прямой $\{o + tx : t \in \mathbb{F}\}$ с квадратикой Γ определяются любым из уравнений

$$t^2 q_1(x) + t \ell_1(x) = 0, \quad t^2 q_2(x) + t \ell_2(x) = 0.$$

Т. к. эти уравнения должны иметь одинаковые решения (относительно t), то при $\ell_1(x), \ell_2(x) \neq 0$ мы получаем

$$\frac{q_1(x)}{\ell_1(x)} = \frac{q_2(x)}{\ell_2(x)},$$

откуда

$$q_1(x) \ell_2(x) = q_2(x) \ell_1(x). \quad (*)$$

Как следствие⁴, это верно при всех x .

Предположим, что линейные функции ℓ_1 и ℓ_2 не пропорциональны. Тогда в подходящем базисе $\ell_1(x) = x_1, \ell_2(x) = x_2$ и равенство записывается в виде

$$q_1(x) x_2 = q_2(x) x_1.$$

Рассматривая члены в левой и правой части, мы видим, что должно быть

$$q_1(x) = l(x) x_1, \quad q_2(x) = l(x) x_2,$$

⁴Пусть поле \mathbb{F} бесконечно и $h \in \mathbb{F}[x_1, x_2, \dots, x_n]$ — какой-либо ненулевой многочлен. Если многочлены $f, g \in \mathbb{F}[x_1, \dots, x_n]$ принимают одинаковые значения при всех значениях переменных x_1, \dots, x_n , при которых многочлен h не обращается в нуль, то они равны.

Доказательство. При указанных условиях многочлены fh и gh принимают одинаковые значения вообще при всех значениях переменных, и, значит, $fh = gh$ (теорема из первого семестра). Т. к. в алгебре многочленов нет делителей нуля, то отсюда следует, что $f = g$. ■

Если поле \mathbb{F} конечно, то теорема и её доказательство тем не менее остаются в силе для многочленов, степень которых по каждому из переменных меньше числа элементов поля \mathbb{F} .

где $l(x)$ — какая-то линейная функция, и, значит,

$$Q_1(x) = (l(x) + 1)x_1, \quad Q_2(x) = (l(x) + 1)x_2.$$

Т.к. $\Gamma = \Gamma(Q_1)$, то Γ содержит гиперплоскость $x^1 = 0$. Т.к. в то же время $\Gamma = \Gamma(Q_2)$, то Q_2 должна тождественно обращаться в нуль на этой гиперплоскости. Однако ни один из её множителей $l(x) + 1$ и x_2 не обращается на ней в нуль тождественно (первых из них не обращается в нуль уже в точке o). Поскольку в алгебре многочленов нет делителей нуля, мы приходим к противоречию.

Итак, $l_2 = \lambda l_1$ ($\lambda \in \mathbb{F}^*$). Из (*) получаем тогда, что и $q_2 = \lambda q_1$, и, значит, $Q_2 = \lambda Q_1$. ■

Следствие 1. Всякий центр квадрики $\Gamma(Q)$ является также центром функции Q .

Доказательство. Если o — центр квадрики $\Gamma(Q)$, то $\Gamma(Q) = \Gamma(\bar{Q})$, где

$$\bar{Q}(o + x) = Q(o - x).$$

Следовательно, $\bar{Q} = \lambda Q$ ($\lambda \in \mathbb{F}^*$). Сравнивая члены второй степени в выражениях \bar{Q} и Q , мы видим, что должно быть $\lambda = 1$, т.е. $\bar{Q} = Q$, а это и означает, что o — центр функции Q . ■

Следствие 2. Если квадрика $\Gamma(Q)$ инвариантна относительно некоторого параллельного переноса, то и функция Q тоже инвариантна относительно этого переноса.

Доказательство. Если квадрика переходит в себя при параллельном переносе на вектор $a \in V$, то $\Gamma(Q) = \Gamma(\bar{Q})$, где

$$\bar{Q}(p) = Q(p + a).$$

Далее рассуждаем так же, как в доказательстве следствия 1. ■

Примечание. Анализируя приведённого выше доказательства с учётом замечания в последней сноске показывает, что она верна и для конечных полей, исключая только поле \mathbb{Z}_3 (напомним, что мы считаем, что $\text{char } \mathbb{F} \neq 2$). Над полем \mathbb{Z}_3 можно привести следующий контрпример:

$$Q_1 = x_1^2 + x_1x_2 + 1 = 0, \quad Q_2 = x_2^2 + x_1x_2 + 1 = 0.$$

задают одну и ту же конику в \mathbb{Z}_3^2 , состоящую из точек $(1, 1)$ и $(-1, -1)$. Однако оба следствия верны и для поля \mathbb{Z}_3 .

Определение 4. $\text{Ker } Q := \text{Ker } \hat{q} \cap \text{Ker } l$.

Предложение 4. Функция Q инвариантна относительно параллельного переноса на вектор a тогда и только тогда, когда $a \in \text{Ker } Q$.

Доказательство. Инвариантность функции Q относительно параллельного переноса на вектор a равносильна тому, что она сохраняет свой вид при переносе начала отсчёта в точку $o' = o + a$. Ввиду леммы 1 из предыдущего вопроса это происходит тогда и только тогда, когда $a \in \text{Ker } Q$. ■

Следствие 3. $\text{Ker } \hat{q} \cap \text{Ker } l$ не зависит от выбора системы координат.

Таким образом, если $\text{Ker } Q \neq \{0\}$, то квадрика $\Gamma = \Gamma(Q)$ вместе с каждой точкой $p \in \mathfrak{A}$ содержит целую плоскость $p + \text{Ker } Q$.

Определение 5. Такая квадрика называется *цилиндрической* с касательным пространством $W := \text{Ker } Q$.

Выберем базис пространства V так, чтобы последние $d := \dim W$ векторов были базисом подпространства W . Тогда выражение Q не будет содержать последних d координат. Пусть $\mathfrak{A}_0 \subset \mathfrak{A}$ — любая плоскость, касательное пространство которой натянуто на первые $n - d$ базисных векторов, и Γ_0 — квадрика, задаваемая в \mathfrak{A}_0 уравнением $Q = 0$ (т.е. $\Gamma_0 = \Gamma \cap \mathfrak{A}_0$). Тогда $\Gamma = \Gamma_0 + W$.

Определение 6. Квадрика, не являющаяся цилиндрической, называется *невырожденной*.

Предложение 5. Невырожденная квадрика имеет не более одного центра.

Доказательство. Пусть o и o' — центры квадрики Γ . Обозначим через s и s' центральные симметрии относительно o и o' соответственно. Тогда $s\Gamma = s'\Gamma = \Gamma$ и, следовательно, $ss'\Gamma = \Gamma$. Т. к.

$$d(ss') = ds \cdot ds' = (-\text{id})^2 = \text{id},$$

то ss' — (нетривиальный) параллельный перенос и, значит, квадрика Γ цилиндрическая. ■

57. КВАДРИКИ В ЕВКЛИДОВОМ (ТОЧЕЧНОМ) ПРОСТРАНСТВЕ. ОРТОГОНАЛЬНАЯ КЛАССИФИКАЦИЯ КВАДРИК. АФФИННАЯ КЛАССИФИКАЦИЯ.

Нецилиндрические квадрики можно разбить на три типа.

1. **Неконические центральные квадрики.** Выбрав начало отсчёта в центре квадрики и умножив её уравнение на подходящее число, мы приведём его к виду

$$q(x_1, \dots, x_n) = 1,$$

где q — невырожденная квадратичная функция.

2. **Конические квадрики.** Выбрав начало отсчёта в вершине квадрики, мы приведём её уравнение к виду

$$q(x_1, \dots, x_n) = 0,$$

где q — невырожденная квадратичная функция. При этом у нас ещё остаётся возможность умножить уравнение на любое число $\lambda \neq 0$.

3. **Нецентральные квадрики.** Т. к. $\text{Ker } \hat{q} \cap \text{Ker } l = \{0\}$, но $\text{Ker } \hat{q} \neq \{0\}$ (иначе квадрика была бы центральной), то $\dim \text{Ker } \hat{q} = 1$ и

$$V = \text{Ker } l \oplus \text{Ker } \hat{q}.$$

Выбрав начало отсчёта на квадрике и базис пространства V , согласованный с последним разложением, мы приведём уравнение квадрики к виду

$$u(x_1, \dots, x_{n-1}) = x_n,$$

где $u = q|_{\text{Ker } l}$ — невырожденная квадратичная функция от $n - 1$ переменных. при этом остаётся возможность умножить уравнение на любое число $\lambda \neq 0$, одновременно разделив на λ последний базисный вектор.

Возможности дальнейшего упрощения уравнения квадрики за счёт выбора подходящего базиса в пространстве V зависит от поля \mathbb{F} . При $\mathbb{F} = \mathbb{C}$ или \mathbb{R} мы можем привести квадратичную функцию q к нормальному виду.

Рассмотрим более подробно случай $\mathbb{F} = \mathbb{R}$. В этом случае уравнение невырожденной квадрики может быть приведено к одному и только одному из следующих видов:

1. **Неконические центральные квадрики:**

$$x_1^2 + \dots + x_k^2 - x_{k+1}^2 - \dots - x_n^2 = 1 \quad (0 < k \leq n).$$

2. **Конические квадрики:**

$$x_1^2 + \dots + x_k^2 - x_{k+1}^2 - \dots - x_n^2 = 0 \quad \left(\frac{n}{2} \leq k < n\right).$$

3. Нецентральные квадрики:

$$x_1^2 + \dots + x_k^2 - x_{k+1}^2 - \dots - x_{n-1}^2 = x_n \quad \left(\frac{n-1}{2} \leq k < n \right).$$

Полученный результат можно интерпретировать как классификацию вещественных квадрик с точностью до аффинных преобразований. В самом деле, если квадрики Γ_1 и Γ_2 задаются одним и тем же уравнением в аффинных системах координат, связанных с реперами $(o; e_1, \dots, e_n)$ и $(o'; e'_1, \dots, e'_n)$ соответственно, то Γ_1 переводится в Γ_2 аффинным преобразованием, переводящим репер $(o; e_1, \dots, e_n)$ в репер $(o'; e'_1, \dots, e'_n)$. Обратно, если квадрика Γ_1 переводится в квадрику Γ_2 аффинным преобразованием f , то Γ_1 и Γ_2 задаются одним и тем же уравнением в аффинных системах координат, связанных с реперами $(o; e_1, \dots, e_n)$ и $(f(o); df(e_1), \dots, df(e_n))$ соответственно.

Посмотрим теперь, к какому виду можно привести уравнение квадрики в евклидовом пространстве, если ограничиться прямоугольными системами координат. Как и в аффинной геометрии, задача сводится к случаю невырожденных квадрик. Рассмотрим, как и выше, три типа таких квадрик.

Определение 1. В произвольной размерности квадрики типа 1 при $k = n$ называются *эллипсоидами*, а при $k < n$ — *гиперболоидами*; квадрики типа 2 называются *квадратичными конусами*; квадрики типа 3 при $k = n - 1$ называются *эллиптическими параболоидами*, а при $k = n - 1$ — *гиперболическими параболоидами*.

Определение 2. С каждым параболоидом $\Gamma = \Gamma(Q)$ каноническим образом связано одномерное подпространство $\text{Ker } \hat{q} \subset V$, называемое *особым направлением* параболоида Γ .

Т. к. $\text{Ker } \hat{q} \not\subset \text{Ker } l$ при любом выборе начала отсчёта, то при $x \in \text{Ker } \hat{q}$ уравнение

$$Q(tx) = t^2 q(x) + tl(x) + c = 0$$

имеет ровно одно решение. Следовательно, любая прямая особого направления пересекает параболоид ровно в одной точке. Более того, это пересечение (по той же причине) трансверсально, ведь касание — это на самом деле кратный корень.

1. **Неконические центральные квадрики.** Из теоремы о приведении квадратичной функции к главным осям следует, что уравнение такой квадрики в прямоугольной системе координат может быть приведено к виду

$$\lambda_1 x_1^2 + \dots + \lambda_n x_n^2 = 1 \quad (\lambda_1, \dots, \lambda_n \neq 0).$$

Числа $\lambda_1, \dots, \lambda_n$ определены однозначно с точностью до перестановки.

2. **Конические квадрики.** Уравнение такой квадрики в прямоугольной системе координат может быть приведено к виду

$$\lambda_1 x_1^2 + \dots + \lambda_n x_n^2 = 0 \quad (\lambda_1, \dots, \lambda_n \neq 0).$$

Числа $\lambda_1, \dots, \lambda_n$ определены однозначно с точностью до перестановки и одновременного умножения на $\lambda \neq 0$.

3. **Нецентральные квадрики (параболоиды).** Выбрав начало отсчёта произвольно и приведя квадратичную функцию q к главным осям, мы получим прямоугольную систему координат, в котором уравнение параболоида будет иметь вид

$$\lambda_1 x_1^2 + \dots + \lambda_{n-1} x_{n-1}^2 + b_1 x_1 + \dots + b_n x_n + c = 0 \quad (\lambda_1, \dots, \lambda_n, b_n \neq 0).$$

За счёт переноса начала отсчёта по координатам x_1, \dots, x_{n-1} можно убрать линейные члены, содержащие эти координаты. При этом, вообще говоря, изменится свободный член. После этого за счёт переноса начала отсчёта по координате x_n можно убрать свободный член. Наконец, умножив уравнение на подходящее число, можно привести его к виду

$$\lambda_1 x_1^2 + \dots + \lambda_{n-1} x_{n-1}^2 = x_n \quad (\lambda_1, \dots, \lambda_{n-1} \neq 0).$$

Покажем, что начало отсчёта, при котором уравнение параболоида приводится к такому виду, определено однозначно. Для этого охарактеризуем его в инвариантных терминах.

Пусть $(o; e_1, \dots, e_n)$ — репер, в котором уравнение параболоида имеет такой вид. Тогда особое направление этого параболоида есть $\langle e_n \rangle$, а его касательная гиперплоскость в точке o задаётся уравнением $x_n = 0$. Следовательно, если базис (e_1, \dots, e_n) ортонормированный, то касательная гиперплоскость параболоида в точке o ортогональна особому направлению. Такая точка называется *вершиной* параболоида (раньше мы называли вершиной другое, это определение вводится только для параболоида), а проходящая через неё прямая особого направления — *осью* параболоида.

Предложение 1. Всякий параболоид в евклидовом пространстве имеет единственную вершину.

Доказательство. Пусть p — точка параболоида с координатами x_1, \dots, x_n . Дифференцируя каноническое уравнение параболоида, находим, что координаты нормального вектора параболоида в точке p суть

$$(2\lambda_1 x_1, \dots, 2\lambda_{n-1} x_{n-1}, -1).$$

Для того, чтобы точка p была вершиной параболоида, необходимо и достаточно, чтобы этот вектор был пропорционален e_n , а это имеет место тогда и только тогда, когда $x_1 = \dots = x_{n-1} = 0$, т. е. $p = o$. ■

Следствие 1. Коэффициенты $\lambda_1, \dots, \lambda_{n-1}$ в каноническом уравнении параболоида определены однозначно с точностью до перестановки и одновременного умножения на -1 .

Доказательство. Как мы показали, начало отсчёта, при котором уравнение параболоида приводится к требуемому виду, определено однозначно. Вектор e_n как единичный вектор особого направления определён однозначно с точностью до умножения на -1 , приводящего к умножению на -1 левой части нашего уравнения. Если вектор e_n фиксирован, то мы уже не можем умножить уравнение на число $\lambda \neq 1$, не изменив его правой части; но тогда числа $\lambda_1, \dots, \lambda_{n-1}$ определены однозначно с точностью до перестановки как собственные значения симметрического оператора, соответствующего квадратичной функции q . ■

Аналогично тому, как это было сделано применительно к аффинной классификации квадрик, полученные результаты можно интерпретировать как классификацию квадрик в аффинном евклидовом пространстве с точностью до движений.

58. ОПРЕДЕЛЕНИЕ ТЕНЗОРА ТИПА (p, q) НА ВЕКТОРНОМ ПРОСТРАНСТВЕ. ЛИНЕЙНЫЕ ОПЕРАЦИИ. ОТОЖДЕСТВЛЕНИЕ ТЕНЗОРОВ МАЛЫХ ВАЛЕНТНОСТЕЙ С ГЕОМЕТРИЧЕСКИМИ ОБЪЕКТАМИ

Здесь я основывался в основном на лекциях Панова, так что определения слегка странноваты. Но мне понравилось, так что здесь написано то, что написано.

Пусть V — линейное пространство над полем \mathbb{F} нулевой характеристики (обычно $\mathbb{F} = \mathbb{R}$ или \mathbb{C}) и V^* — двойственное пространство. Элементы $v \in V$ — это, как обычно, векторы, а элементы $\xi \in V^*$ здесь мы будем называть *ковекторами*.

Определение 1. Полилинейной функцией типа (p, q) (и валентности $p + q$) называется функция

$$\mathcal{T} : \underbrace{V \times \dots \times V}_p \times \underbrace{V^* \times \dots \times V^*}_q \rightarrow \mathbb{F}$$

от p векторных и q ковекторных аргументов, которая линейна по каждому аргументу.

Тензоры типа (p, q) образуют линейное пространство над полем \mathbb{F} , в котором сумма и умножение на скаляры определены по формуле

$$(\lambda\mathcal{T} + \mu\mathcal{S})(v_1, \dots, v_p, \xi^1, \dots, \xi^q) := \lambda\mathcal{T}(v_1, \dots, v_p, \xi^1, \dots, \xi^q) + \mu\mathcal{S}(v_1, \dots, v_p, \xi^1, \dots, \xi^q).$$

Мы будем обозначать это пространство через $\mathbb{P}_p^q(V)$.

Пример 1. Геометрическая интерпретация тензоров малых валентностей:

1. Полилинейная функция типа $(0, 0)$ — это скаляры из поля \mathbb{F} .
2. Полилинейная функция типа $(1, 0)$ — линейная функция на V , т. е. элемент V^* .
3. Полилинейная функция типа $(0, 1)$ — линейная функция на V^* , т. е. элемент из V^{**} . Однако пространства V и V^{**} канонически изоморфны. Имея в виду этот изоморфизм, полилинейную функцию типа $(0, 1)$ можно считать вектором, т. е. элементом из V .
4. Полилинейная функция типа $(2, 0)$ — это билинейная функция на V .
5. Полилинейная типа $(1, 1)$ — это линейный оператор $\mathcal{A} : V \rightarrow V$. Действительно, каждой полилинейной функции $\mathcal{T}(v, \xi)$ можно сопоставить оператор $\mathcal{A}(x) := \xi(x)$. Обратно, каждому оператору можем сопоставить полилинейную функцию $\mathcal{T}_{\mathcal{A}}(v, \xi) := \xi(\mathcal{A}v)$ типа $(1, 1)$.

Предложение 1. Сопоставление линейному оператору \mathcal{A} полилинейной функции $\mathcal{T}_{\mathcal{A}}(v, \xi) := \xi(\mathcal{A}v)$ типа $(1, 1)$ устанавливает канонический изоморфизм $\text{End}(V) \simeq \mathbb{P}_1^1(V)$.

Доказательство. Прежде всего заметим, что данное отображение $\text{End}(V) \rightarrow \mathbb{P}_1^1(V)$ линейно. Пусть $\dim V =: n$. Тогда размерность пространства $\mathbb{P}_1^1(V)$ равна n^2 — это доказывается при помощи выбора базиса так же, как и для билинейных функций. Поэтому размерности пространств $\text{End}(V)$ и $\mathbb{P}_1^1(V)$ равны, а значит, достаточно доказать, что $\text{End}(V) \rightarrow \mathbb{P}_1^1(V)$ — инъекция. Будем делать это по критерию инъективности. Пусть $\mathcal{T}_{\mathcal{A}}$ — тождественно нулевая полилинейная функция, т. е. $\xi(\mathcal{A}v) = 0 \forall v \in V, \xi \in V^*$. Получаем, что любая линейная функция обращается в нуль на векторе $\mathcal{A}v$, т. е. $\mathcal{A}v = \mathbf{0}$. Т. к. это верно $\forall v \in V$, получаем $\mathcal{A} = \mathcal{O}$. Итак, отображение $\mathcal{A} \mapsto \mathcal{T}_{\mathcal{A}}$ инъективно, а значит, задаёт изоморфизм. ■

59. ПРОИЗВЕДЕНИЕ ТЕНЗОРОВ. БАЗИС В ПРОСТРАНСТВЕ ТЕНЗОРОВ ТИПА (p, q) . КООРДИНАТЫ ТЕНЗОРА, ИХ ИЗМЕНЕНИЕ ПРИ ЗАМЕНЕ БАЗИСА В ОСНОВНОМ ПРОСТРАНСТВЕ. ИНВАРИАНТ. СВЕРТКИ ТЕНЗОРА

Зафиксируем базис e_1, \dots, e_n в пространстве V . В пространстве V^* имеется двойственный базис $\varepsilon^1, \dots, \varepsilon^n$, где $\varepsilon^i(e_j) = \delta_j^i$. Тогда любая полилинейная функция $\mathcal{T} \in \mathbb{P}_p^q(V)$ задаётся своими значениями на базисных векторах и ковекторах:

$$\begin{aligned} \mathcal{T}(v_1, \dots, v_p, \xi^1, \dots, \xi^q) &= \mathcal{T}(v_1^{i_1} e_{i_1}, \dots, v_p^{i_p} e_{i_p}, \xi_{j_1}^1 \varepsilon^{j_1}, \dots, \xi_{j_q}^q \varepsilon^{j_q}) = \\ &= v_1^{i_1} \dots v_p^{i_p} \xi_{j_1}^1 \dots \xi_{j_q}^q \mathcal{T}(e_{i_1}, \dots, e_{i_p}, \varepsilon^{j_1}, \dots, \varepsilon^{j_q}). \end{aligned}$$

Определение 1. Сопоставим полилинейной функции $\mathcal{T} \in \mathbb{T}_p^q(V)$ и базису e_1, \dots, e_n пространства V набор из n^{p+q} чисел.

$$T_{i_1, \dots, i_p}^{j_1, \dots, j_q} := \mathcal{T}(e_{i_1}, \dots, e_{i_p}, \varepsilon^{j_1}, \dots, \varepsilon^{j_q}).$$

Обозначение: $\mathcal{T} = \{T_{i_1, \dots, i_p}^{j_1, \dots, j_q}\}$.

Посмотрим, как преобразуется этот набор при заменах базисов. Пусть $C = (c_{i'}^i)$ — матрица перехода от базиса e_1, \dots, e_n к $e_{1'}, \dots, e_{n'}$. Тогда мы имеем $e_{i'} = c_{i'}^i e_i$, $\varepsilon^{j'} = c_j^{j'} \varepsilon^j$ и

$$\begin{aligned} T_{i'_1, \dots, i'_p}^{j'_1, \dots, j'_q} &= \mathcal{T}(e_{i'_1}, \dots, e_{i'_p}, \varepsilon^{j'_1}, \dots, \varepsilon^{j'_q}) = \mathcal{T}(c_{i'_1}^{i_1} e_{i_1}, \dots, c_{i'_p}^{i_p} e_{i_p}, c_{j_1}^{j'_1} \varepsilon^{j_1}, \dots, c_{j_q}^{j'_q} \varepsilon^{j_q}) = \\ &= c_{i'_1}^{i_1} \dots c_{i'_p}^{i_p} c_{j_1}^{j'_1} \dots c_{j_q}^{j'_q} \mathcal{T}(e_{i_1}, \dots, e_{i_p}, \varepsilon^{j_1}, \dots, \varepsilon^{j_q}) = c_{i'_1}^{i_1} \dots c_{i'_p}^{i_p} c_{j_1}^{j'_1} \dots c_{j_q}^{j'_q} T_{i_1, \dots, i_p}^{j_1, \dots, j_q}. \end{aligned}$$

Определение 2. Тензором типа (p, q) называется соответствие

$$\text{базисы в } V \mapsto \text{наборы из } n^{p+q} \text{ чисел } T = \{T_{i_1, \dots, i_p}^{j_1, \dots, j_q}\},$$

при котором наборы $T = \{T_{i_1, \dots, i_p}^{j_1, \dots, j_q}\}$ и $T' = \{T_{i'_1, \dots, i'_p}^{j'_1, \dots, j'_q}\}$, соответствующие различным базисам e_1, \dots, e_n и $e_{1'}, \dots, e_{n'}$ связаны соотношением

$$T_{i'_1, \dots, i'_p}^{j'_1, \dots, j'_q} = c_{i'_1}^{i_1} \dots c_{i'_p}^{i_p} c_{j_1}^{j'_1} \dots c_{j_q}^{j'_q} T_{i_1, \dots, i_p}^{j_1, \dots, j_q}.$$

Это соотношение называется *тензорным законом преобразования*. Числа $T_{i_1, \dots, i_p}^{j_1, \dots, j_q}$ называются *компонентами (координатами) тензора* T .

Тензоры типа (p, q) образуют линейное пространство $\mathbb{T}_p^q(V)$ относительно операций покомпонентного сложения и умножения на скаляры.

Полилинейная функция $\mathcal{T} \in \mathbb{P}_p^q(V)$ определяет тензор $T \in \mathbb{T}_p^q(V)$ по формуле из определения компонент тензора. Обратно, тензор определяет полилинейную функцию по формуле

$$\mathcal{T}(v_1, \dots, v_p, \xi^1, \dots, \xi^q) = v_1^{i_1} \dots v_p^{i_p} \xi_{j_1}^1 \dots \xi_{j_q}^q T_{i_1, \dots, i_p}^{j_1, \dots, j_q}.$$

Таким образом, можно отождествить пространства $\mathbb{P}_p^q(V)$ и $\mathbb{T}_p^q(V)$. Это соответствие обобщает соответствие между билинейными функциями (или линейными операторами) и их матрицами.

Далее мы не будем различать тензоры и полилинейные функции. Правда, и до этого, честно сказать, не различали.

Пример 1.

1. Скаляры $\lambda \in \mathbb{F}$ естественно считать тензорами типа $(0, 0)$. Они не меняются при замене базиса.
2. Векторы — это тензоры типа $(0, 1)$. Тензорный закон преобразования $v^{i'} = c_{i'}^i v^i$ описывает изменение координат вектора при замене базиса.
3. Ковекторы (линейные функции) — это тензоры типа $(1, 0)$. Тензорный закон преобразования $\xi_{i'} = c_{i'}^i \xi_i$ — это закон преобразования координат линейной функции при замене базиса.
4. Билинейные функции — это тензоры типа $(2, 0)$. Тензорный закон преобразования $T_{i'j'} = c_{i'}^i c_{j'}^j T_{ij}$ — это закон изменения матрицы билинейной функции.
5. Линейные операторы — это тензоры типа $(1, 1)$. Тензорный закон преобразования $T_{j'}^{i'} = c_{i'}^i c_j^{j'} T_j^i$ — это закон изменения матрицы оператора.

Определение 3. Тензорным произведением двух тензоров $\mathcal{T}_1 \in \mathbb{T}_{p_1}^{q_1}(V)$ и $\mathcal{T}_2 \in \mathbb{T}_{p_2}^{q_2}(V)$ называется тензор $\mathcal{T}_1 \otimes \mathcal{T}_2 \in \mathbb{T}_{p_1+p_2}^{q_1+q_2}(V)$, заданный по формуле

$$\begin{aligned} \mathcal{T}_1 \otimes \mathcal{T}_2(v_1, \dots, v_{p_1}, v_{p_1+1}, \dots, v_{p_1+p_2}, \xi^1, \dots, \xi^{q_1}, \xi^{q_1+1}, \dots, \xi^{q_1+q_2}) = \\ = \mathcal{T}_1(v_1, \dots, v_{p_1}, \xi^1, \dots, \xi^{q_1}) \cdot \mathcal{T}_2(v_{p_1+1}, \dots, v_{p_1+p_2}, \xi^{q_1+1}, \dots, \xi^{q_1+q_2}). \end{aligned}$$

Операция тензорного произведения, очевидно, ассоциативна и дистрибутивна относительно сложения, но не коммутативна.

Базис в пространстве тензоров $\mathbb{T}_p^q(V)$ можно задать при помощи операции тензорного произведения. Рассмотрим тензор $\varepsilon^{i_1} \otimes \dots \otimes \varepsilon^{i_p} \otimes e_{j_1} \otimes \dots \otimes e_{j_q}$. По определению, её значение на p векторах и q ковекторах задаётся формулой

$$(\varepsilon^{i_1} \otimes \dots \otimes \varepsilon^{i_p} \otimes e_{j_1} \otimes \dots \otimes e_{j_q})(v_1, \dots, v_p, \xi^1, \dots, \xi^q) = \varepsilon^{i_1}(v_1) \dots \varepsilon^{i_p}(v_p) \cdot e_{j_1}(\xi^1) \dots e_{j_q}(\xi^q)$$

и имеет координаты

$$T_{k_1, \dots, k_p}^{l_1, \dots, l_q} = (\varepsilon^{i_1} \otimes \dots \otimes \varepsilon^{i_p} \otimes e_{j_1} \otimes \dots \otimes e_{j_q})(e_{k_1}, \dots, e_{k_p}, \xi^{l_1}, \dots, \xi^{l_q}) = \delta_{k_1}^{i_1} \dots \delta_{k_p}^{i_p} \delta_{j_1}^{l_1} \dots \delta_{j_q}^{l_q}.$$

Теорема 1. Тензоры $\varepsilon^{i_1} \otimes \dots \otimes \varepsilon^{i_p} \otimes e_{j_1} \otimes \dots \otimes e_{j_q}$, отвечающие всевозможным значениям индексов $i_1, \dots, i_p, j_1, \dots, j_q$, образуют базис в пространстве $\mathbb{T}_p^q(V)$.

Доказательство. Сначала докажем линейную независимость данных тензоров. Предположим, существуют такие числа $\lambda_{i_1, \dots, i_p}^{j_1, \dots, j_q}$, что линейная комбинация $\lambda_{i_1, \dots, i_p}^{j_1, \dots, j_q} \varepsilon^{i_1} \otimes \dots \otimes \varepsilon^{i_p} \otimes e_{j_1} \otimes \dots \otimes e_{j_q}$ равна нулю. Прменив эту полилинейную функцию к аргументам $e_{k_1}, \dots, e_{k_p}, \varepsilon^{l_1}, \dots, \varepsilon^{l_q}$, получим (*)

$$\begin{aligned} 0 &= (\lambda_{i_1, \dots, i_p}^{j_1, \dots, j_q} \varepsilon^{i_1} \otimes \dots \otimes \varepsilon^{i_p} \otimes e_{j_1} \otimes \dots \otimes e_{j_q})(e_{k_1}, \dots, e_{k_p}, \varepsilon^{l_1}, \dots, \varepsilon^{l_q}) = \\ &= \lambda_{i_1, \dots, i_p}^{j_1, \dots, j_q} \delta_{k_1}^{i_1} \dots \delta_{k_p}^{i_p} \delta_{j_1}^{l_1} \dots \delta_{j_q}^{l_q} = \lambda_{k_1, \dots, k_p}^{l_1, \dots, l_q}, \end{aligned}$$

т. е. все коэффициенты линейной комбинации равны нулю.

Теперь докажем, что любой тензор $T = \{T_{i_1, \dots, i_p}^{j_1, \dots, j_q}\}$ представляется в виде линейной комбинации данных тензоров. А именно, докажем, что

$$T = T_{i_1, \dots, i_p}^{j_1, \dots, j_q} \varepsilon^{i_1} \otimes \dots \otimes \varepsilon^{i_p} \otimes e_{j_1} \otimes \dots \otimes e_{j_q}.$$

В силу полилинейности это равенство достаточно проверить на наборах базисных векторов и ковекторов, т. е. на аргументах вида $e_{k_1}, \dots, e_{k_p}, \xi^{l_1}, \dots, \xi^{l_q}$. При подстановке этих аргументов мы по определению получаем $T_{k_1, \dots, k_p}^{l_1, \dots, l_q}$, а при подстановке в правую часть мы получаем то же самое в силу выкладки, аналогичной (*). ■

Следствие 1. $\dim \mathbb{T}_p^q(V) = n^{p+q}$.

Пусть теперь $T = \{T_{i_1, \dots, i_p}^{j_1, \dots, j_q}\}$ — тензор хотя бы с одним нижним и верхним индексом, т. е. $p > 0$ и $q > 0$. Зафиксируем один верхний и один нижний индекс (пусть для простоты это будут первые индексы) и сформируем следующий новый набор из n^{p+q-2} чисел:

$$cT := \{T_{k, i_2, \dots, i_p}^{k, j_2, \dots, j_q}\},$$

где как обычно по повторяющемуся верхнему и нижнему индексу k производится суммирование.

Предложение 1. $cT = \{T_{k, i_2, \dots, i_p}^{k, j_2, \dots, j_q}\}$ является тензором типа $(p-1, q-1)$.

Доказательство. Необходимо проверить тензорный закон. Мы имеем

$$\begin{aligned} (cT)_{i'_1, \dots, i'_p}^{j'_1, \dots, j'_q} &= T_{k', i'_2, \dots, i'_p}^{k', j'_2, \dots, j'_q} = c_{k'}^{i_1} c_{i'_2}^{i_2} \dots c_{i'_p}^{i_p} c_{j'_2}^{j_2} \dots c_{j'_q}^{j_q} T_{i_1, i_2, \dots, i_p}^{j_1, j_2, \dots, j_q} = \\ &= c_{i'_2}^{i_2} \dots c_{i'_p}^{i_p} c_{j'_2}^{j_2} \dots c_{j'_q}^{j_q} c_{k'}^{i_1} c_{j'_1}^{j_1} T_{i_1, i_2, \dots, i_p}^{j_1, j_2, \dots, j_q} = c_{i'_2}^{i_2} \dots c_{i'_p}^{i_p} c_{j'_2}^{j_2} \dots c_{j'_q}^{j_q} \delta_{j_1}^{j'_1} T_{i_1, i_2, \dots, i_p}^{j_1, j_2, \dots, j_q} = \\ &= c_{i'_2}^{i_2} \dots c_{i'_p}^{i_p} c_{j'_2}^{j_2} \dots c_{j'_q}^{j_q} T_{k, i_2, \dots, i_p}^{k, j_2, \dots, j_q} = c_{i'_2}^{i_2} \dots c_{i'_p}^{i_p} c_{j'_2}^{j_2} \dots c_{j'_q}^{j_q} (cT)_{i_2, \dots, i_p}^{j_2, \dots, j_q}. \end{aligned}$$

■

Определение 4. Тензор $cT := \{T_{k,i_2,\dots,i_p}^{k,j_1,\dots,j_q}\}$ называется *свёрткой* тензора $T = \{T_{i_1,i_2,\dots,i_p}^{j_1,j_2,\dots,j_q}\}$ по (первым) верхнему и нижнему индексам.

Свёртка, очевидно, задаёт линейное отображение $\mathbb{T}_p^q(V) \rightarrow \mathbb{T}_{p-1}^{q-1}(V)$. Операцию свёртки можно проводить несколько раз до исчерпания верхних или нижних индексов. Последняя возможная свёртка называется *полной свёрткой*.

Пример 2.

1. Пусть \mathcal{A} — оператор, т. е. тензор типа $(1, 1)$. Результатом его свёртки будет тензор типа $(0, 0)$, т. е. скаляр. Этот скаляр — это сумма a_i^i диагональных элементов матрицы оператора \mathcal{A} в любом базисе, т. е. след оператора $c\mathcal{A} = \text{tr } \mathcal{A}$. Проверка тензорного закона для свёртки в данном случае сводится к проверке независимости следа от базиса.
2. Пусть $B = \{b_{i_1 i_2}\}$ — билинейная функция (тензор типа $(0, 2)$), а $u = \{u^{j_1}\}$ и $v = \{v^{j_2}\}$ — векторы (тензоры типа $(1, 0)$). Рассмотрим тензор $B \otimes u \otimes v = \{b_{i_1 i_2} u^{j_1} v^{j_2}\}$ типа $(2, 2)$. Его полная свёртка есть скаляр $b_{kl} u^k v^l = B(u, v)$ — значение билинейной функции на данной паре векторов.

60. СИММЕТРИЧЕСКИЕ И КОСОСИММЕТРИЧЕСКИЕ ТЕНЗОРЫ ТИПОВ $(p, 0)$ ИЛИ $(0, q)$. СИММЕТРИЗАЦИЯ И АЛЬТЕРНИРОВАНИЕ

Здесь мы будем рассматривать только тензоры с нижними индексами.

Определение 1. Полилинейная функция $\mathcal{T} \in \mathbb{P}_p^0(V)$ называется *симметрической*, если

$$\mathcal{T}(v_{\sigma(1)}, \dots, v_{\sigma(p)}) = \mathcal{T}(v_1, \dots, v_p),$$

и *кососимметрической*, если

$$\mathcal{T}(v_{\sigma(1)}, \dots, v_{\sigma(p)}) = \text{sgn } \sigma \mathcal{T}(v_1, \dots, v_p)$$

$\forall v_1, \dots, v_p \in V, \forall \sigma \in S_p$.

Компоненты тензора, соответствующего симметрической полилинейной функции, удовлетворяют соотношениям

$$T_{i_1, \dots, i_p} = T_{i_{\sigma(1)}, \dots, i_{\sigma(p)}},$$

а кососимметрической —

$$T_{i_1, \dots, i_p} = \text{sgn } \sigma T_{i_{\sigma(1)}, \dots, i_{\sigma(p)}}.$$

Такие тензоры называются, соответственно, *симметрическими* и *кососимметрическими*. В частности, у кососимметрического тензора отличны от нуля могут быть лишь те компоненты T_{i_1, \dots, i_p} , у которых все индексы различны.

Симметрические и кососимметрические тензоры образуют подпространства в пространстве тензоров $\mathbb{T}_p^0(V)$, которые обозначаются $S^p(V)$ и $\Lambda^p(V)$ соответственно.

Определение 2. *Симметризацией* называется линейный оператор $\text{Sym} : \mathbb{T}_p^0(V) \rightarrow \mathbb{T}_p^0(V)$, который тензору $T \in \mathbb{T}_p^0(V)$ ставит в соответствие тензор $\text{Sym } T$ с компонентами

$$(\text{Sym } T)_{i_1, \dots, i_p} := \frac{1}{p!} \sum_{\sigma \in S_p} T_{i_{\sigma(1)}, \dots, i_{\sigma(p)}}.$$

Определение 3. *Альтернированием* называется линейный оператор $\text{Alt} : \mathbb{T}_p^0(V) \rightarrow \mathbb{T}_p^0(V)$, который тензору $T \in \mathbb{T}_p^0(V)$ ставит в соответствие тензор $\text{Alt } T$ с компонентами

$$(\text{Alt } T)_{i_1, \dots, i_p} := \frac{1}{p!} \sum_{\sigma \in S_p} \text{sgn } \sigma T_{i_{\sigma(1)}, \dots, i_{\sigma(p)}}.$$

Легко видеть, что $\text{Sym } T$ — симметрический тензор, а $\text{Alt } T$ — кососимметрический $\forall T \in \mathbb{T}_p^0(V)$.

Предложение 1. Операторы Sym и Alt являются проекторами на подпространства $S^p(V)$ и $\Lambda^p(V)$.

Доказательство. Оба утверждения доказываются аналогично. Докажем второе. Достаточно показать, что $\text{Alt } T = T$ для любого кососимметрического тензора T . Мы имеем:

$$(\text{Alt } T)_{i_1, \dots, i_p} = \frac{1}{p!} \sum_{\sigma \in S_p} \text{sgn } \sigma T_{i_{\sigma(1)}, \dots, i_{\sigma(p)}} = \frac{1}{p!} \sum_{\sigma \in S_p} T_{i_1, \dots, i_p} = T_{i_1, \dots, i_p},$$

где второе равенство выполнено в силу кососимметричности T . ■

61. ТЕНЗОРНАЯ АЛГЕБРА ВЕКТОРНОГО ПРОСТРАНСТВА. ВНЕШНЯЯ АЛГЕБРА И ЕЕ РАЗМЕРНОСТЬ

Определение алгебры можно вспомнить, посмотрев в 11-ый вопрос.

Определение 1. Алгебра A называется *градуированной*, если она допускает разложение во внешнюю прямую сумму $A = \bigoplus_i A_i$, причём $A_{i_1} A_{i_2} \subseteq A_{i_1+i_2}$.

Напомним, что мы рассматриваем n -мерное векторное пространство V над полем \mathbb{F} нулевой характеристики.

Рассмотрим внешнюю прямую сумму

$$\mathbb{T}(V^*) := \mathbb{T}_0^0(V) \oplus \mathbb{T}_1^0(V) \oplus \mathbb{T}_2^0(V) \oplus \dots = \bigoplus_{p=0}^{\infty} \mathbb{T}_p^0(V).$$

Элементами этой суммы являются последовательности $(f_0, f_1, f_2, \dots) = \sum_{i \geq 0} f_i$, $f_i \in \mathbb{T}_i^0(V)$, члены которой почти все равны нулю. Тензорное умножение определяет на $\mathbb{T}(V^*)$ умножение

$$\left(\sum_{i \geq 0} f_i \right) \left(\sum_{j \geq 0} g_j \right) = \sum_{i+j \geq 0} f_i \otimes g_j,$$

где $f_i \in \mathbb{T}_i^0(V)$, $g_j \in \mathbb{T}_j^0(V)$, которое задаёт на $\mathbb{T}(V^*)$ структуру градуированной ассоциативной (но не коммутативной) алгебры с единицей.

Определение 2. Аналогично, для контравариантных тензоров определяется градуированная ассоциативная (но не коммутативная) алгебра с единицей:

$$\mathbb{T}(V) := \mathbb{T}_0^0(V) \oplus \mathbb{T}_0^1(V) \oplus \mathbb{T}_0^2(V) \oplus \dots = \bigoplus_{q=0}^{\infty} \mathbb{T}_0^q(V),$$

которая называется *тензорной алгеброй* пространства V .

Алгебру $\mathbb{T}(V^*)$ ковариантных тензоров можно понимать как тензорную алгебру векторного пространства V^* .

Определение 3. Внешним произведением кососимметрических тензоров $P \in \Lambda^p(V)$ и $Q \in \Lambda^q(V)$ называется кососимметрический тензор

$$P \wedge Q = \frac{(p+q)!}{p! \cdot q!} \text{Alt}(P \otimes Q).$$

Теорема 1. Внешнее произведение кососимметрических тензоров обладает следующими свойствами: для любых $P \in \Lambda^p(V)$, $Q \in \Lambda^q(V)$, $R \in \Lambda^r(V)$ и $\lambda, \mu \in \mathbb{F}$

1. $(\lambda P + \mu Q) \wedge R = \lambda P \wedge R + \mu Q \wedge R$ (при $p = q$);
2. $(Q \wedge P) = (-1)^{pq} P \wedge Q$;
3. $(P \wedge Q) \wedge R = P \wedge (Q \wedge R)$.

Далее будем использовать следующие обозначения: для $\sigma \in S_p$ и $P \in \mathbb{T}_p^0(V)$ обозначим через σP тензор с компонентами $(\sigma P)_{i_1, \dots, i_p} = P_{i_{\sigma(1)}, \dots, i_{\sigma(p)}}$. По определению альтернирования, $\text{Alt } P = \frac{1}{p!} \sum_{\sigma \in S_p} \text{sgn } \sigma \cdot \sigma P$, и имеет место соотношение

$$\sigma(\text{Alt } P) = \text{Alt}(\sigma P) = \text{sgn } \sigma \cdot \text{Alt } P.$$

Лемма 1. Для любых тензоров $P \in \mathbb{T}_p^0(V)$, и $Q \in \mathbb{T}_q^0(V)$ имеем

$$\text{Alt}((\text{Alt } P) \otimes Q) = \text{Alt}(P \otimes Q) = \text{Alt}(P \otimes (\text{Alt } Q)).$$

Доказательство. Докажем лишь первое равенство (второе доказывается аналогично). Поскольку операция \otimes дистрибутивна, а оператор Alt линеен, имеем

$$\text{Alt}((\text{Alt } P) \otimes Q) = \text{Alt} \left(\left(\frac{1}{p!} \sum_{\sigma \in S_p} \text{sgn } \sigma \cdot \sigma P \right) \otimes Q \right) = \frac{1}{p!} \sum_{\sigma \in S_p} \text{Alt}(\sigma P \otimes Q).$$

Каждой перестановке $\sigma \in S_p$ сопоставим перестановку $\tilde{\sigma} \in S_{p+q}$, которая на первых p индексах действует как σ , а остальные оставляет на месте, т. е.

$$\tilde{\sigma} = \begin{pmatrix} 1 & \dots & p & p+1 & \dots & p+q \\ \sigma(1) & \dots & \sigma(p) & p+1 & \dots & p+q \end{pmatrix}.$$

При этом, очевидно, $\text{sgn } \tilde{\sigma} = \text{sgn } \sigma$ и $\sigma P \otimes Q = \tilde{\sigma}(P \otimes Q)$. Тогда имеем

$$\begin{aligned} \text{Alt}((\text{Alt } P) \otimes Q) &= \frac{1}{p!} \sum_{\sigma \in S_p} \text{sgn } \sigma \cdot \text{Alt}(\tilde{\sigma}(P \otimes Q)) = \\ &= \frac{1}{p!} \sum_{\sigma \in S_p} \text{sgn } \sigma \text{sgn } \tilde{\sigma} \cdot \text{Alt}(P \otimes Q) = \frac{1}{p!} \sum_{\sigma \in S_p} \text{Alt}(P \otimes Q) = \text{Alt}(P \otimes Q). \end{aligned}$$

■

Доказательство.

1. Дистрибутивность вытекает из дистрибутивности операции \otimes и линейности оператора Alt .
2. Для доказательства антикоммутативности достаточно проверить, что имеет место соотношение $\text{Alt}(Q \otimes P) = (-1)^{pq} \text{Alt}(P \otimes Q)$. Введём перестановку

$$\delta = \begin{pmatrix} 1 & \dots & p & p+1 & \dots & p+q \\ q+1 & \dots & q+p & 1 & \dots & q \end{pmatrix}.$$

Тогда $\text{sgn } \delta = (-1)^{pq}$, т. к. δ — результат pq транспозиций. Мы имеем

$$\begin{aligned} \text{Alt}(Q \otimes P)_{i_1, \dots, i_{p+q}} &= \frac{1}{(p+q)!} \sum_{\sigma \in S_{p+q}} \text{sgn } \sigma \cdot Q_{i_{\sigma(1)}, \dots, i_{\sigma(q)}} P_{i_{\sigma(q+1)}, \dots, i_{\sigma(p+q)}} = \\ &= \frac{1}{(p+q)!} \sum_{\sigma \in S_{p+q}} \text{sgn } \delta \text{sgn } \delta \cdot \sigma P_{i_{\sigma\delta(1)}, \dots, i_{\sigma\delta(p)}} Q_{i_{\sigma\delta(p+1)}, \dots, i_{\sigma\delta(p+q)}} = \\ &= \text{sgn } \delta \frac{1}{(p+q)!} \sum_{\varphi \in S_{p+q}} \text{sgn } \varphi P_{i_{\varphi(1)}, \dots, i_{\varphi(p)}} Q_{i_{\varphi(p+1)}, \dots, i_{\varphi(p+q)}} = (-1)^{pq} \text{Alt}(P \otimes Q)_{i_1, \dots, i_{p+q}}. \end{aligned}$$

3. Мы имеем

$$\begin{aligned}(P \wedge Q) \wedge R &= \frac{(p+q+r)!}{(p+q)! \cdot r!} \text{Alt}((P \wedge Q) \otimes R) = \\ &= \frac{(p+q+r)!}{(p+q)! \cdot r!} \text{Alt}\left(\frac{(p+q)!}{p! \cdot q!} \text{Alt}(P \otimes Q) \otimes R\right) = \frac{(p+q+r)!}{p! \cdot q! \cdot r!} \text{Alt}(P \otimes Q \otimes R),\end{aligned}$$

где в последнем равенстве мы воспользовались предыдущей леммой. Равенство $P \wedge (Q \wedge R) = \frac{(p+q+r)!}{p! \cdot q! \cdot r!} \text{Alt}(P \otimes Q \otimes R)$ устанавливается аналогично. ■

Выберем теперь базис e_1, \dots, e_n в V , и пусть $\varepsilon^1, \dots, \varepsilon^n$ — двойственный базис в $V^* = \Lambda^1(V)$. Рассмотрим кососимметрические тензоры

$$\varepsilon^{i_1} \wedge \dots \wedge \varepsilon^{i_p} = p! \text{Alt}(\varepsilon^{i_1} \otimes \dots \otimes \varepsilon^{i_p}) = \sum_{\sigma \in S_p} \text{sgn } \sigma \cdot \varepsilon^{i_{\sigma(1)}} \otimes \dots \otimes \varepsilon^{i_{\sigma(p)}}. \quad (*)$$

В силу последней теоремы имеем $\varepsilon^{i_1} \wedge \dots \wedge \varepsilon^{i_p} = \text{sgn } \sigma \cdot \varepsilon^{i_{\sigma(1)}} \wedge \dots \wedge \varepsilon^{i_{\sigma(p)}}$.

Примечание. Обратим внимание, что ввиду выбора коэффициента $\frac{(p+q)!}{p! \cdot q!}$ в определении внешнего произведения, выражение $\varepsilon^{i_1} \wedge \dots \wedge \varepsilon^{i_p}$ оказалось линейной комбинацией выражений $\varepsilon^{i_{\sigma(1)}} \otimes \dots \otimes \varepsilon^{i_{\sigma(p)}}$ с целыми коэффициентами. Таким образом, кососимметрические тензоры $\varepsilon^{i_1} \wedge \dots \wedge \varepsilon^{i_p}$ определены над конечными полями (характеристики, отличной от двух) и даже над целыми числами, что удобно с алгебраической точки зрения.

Теорема 2. Кососимметрические тензоры $\{\varepsilon^{i_1} \wedge \dots \wedge \varepsilon^{i_p} : i_1 < \dots < i_p\}$ образуют базис в пространстве $\Lambda^p(V)$. В частности, $\dim \Lambda^p(V) = C_n^p$.

Доказательство. Сначала докажем, что любой кососимметрический тензор $T \in \Lambda^p(V)$ представляется в виде линейной комбинации данных тензоров. Разложим T по базису $\varepsilon^{i_1} \otimes \dots \otimes \varepsilon^{i_p}$:

$$T = T_{i_1, \dots, i_p} \varepsilon^{i_1} \otimes \dots \otimes \varepsilon^{i_p}.$$

Теперь применим к обеим частям оператор Alt . Т.к. T — кососимметрический тензор, $\text{Alt } T = T$. С другой стороны, $\text{Alt}(\varepsilon^{i_1} \otimes \dots \otimes \varepsilon^{i_p}) = \frac{1}{p!} \varepsilon^{i_1} \wedge \dots \wedge \varepsilon^{i_p}$. Итак, получаем

$$T = \frac{1}{p!} \sum_{i_1, \dots, i_p} T_{i_1, \dots, i_p} \varepsilon^{i_1} \wedge \dots \wedge \varepsilon^{i_p} = \sum_{i_1 < \dots < i_p} T_{i_1, \dots, i_p} \varepsilon^{i_1} \wedge \dots \wedge \varepsilon^{i_p},$$

что и даёт требуемое представление в виде линейной комбинации.

Предположим, что тензоры $\{\varepsilon^{i_1} \wedge \dots \wedge \varepsilon^{i_p} : i_1 < \dots < i_p\}$ линейно зависимы, т.е.

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

Тогда из (*) получаем

$$\sum_{i_1 < \dots < i_p} \lambda_{i_1, \dots, i_p} \sum_{\sigma \in S_p} \text{sgn } \sigma \cdot \varepsilon^{i_{\sigma(1)}} \otimes \dots \otimes \varepsilon^{i_{\sigma(p)}} = 0.$$

В этой сумме все тензоры $\varepsilon^{i_{\sigma(1)}} \otimes \dots \otimes \varepsilon^{i_{\sigma(p)}}$ различны, то они линейно независимы. Отсюда получаем, что $\lambda_{i_1, \dots, i_p} = 0$. ■

Определение 4. Выражение $T = \sum_{i_1 < \dots < i_p} T_{i_1, \dots, i_p} \varepsilon^{i_1} \wedge \dots \wedge \varepsilon^{i_p}$ (представляющее собой разложение кососимметрического тензора $T \in \Lambda^p(V)$ по базису $\{\varepsilon^{i_1} \wedge \dots \wedge \varepsilon^{i_p} : i_1 < \dots < i_p\}$) называется *внешней p -формой*.

Теперь рассмотрим внешнюю прямую сумму

$$\Lambda(V) = \bigoplus_{p=0}^n \Lambda^p(V).$$

Внешнее умножение определяет на $\Lambda(V)$ умножение

$$\left(\sum_{i \geq 0} f_i \right) \wedge \left(\sum_{j \geq 0} g_j \right) = \sum_{i+j \geq 0} f_i \wedge g_j,$$

где $f_i \in \Lambda^i(V)$, $g_j \in \Lambda^j(V)$, которое задаёт на $\Lambda(V)$ структуру градуированной ассоциативной алгебры с единицей. Алгебра $\Lambda(V)$ над полем \mathbb{F} называется *внешней алгеброй* пространства V . Размерность внешней алгебры

$$\dim \Lambda(V) = \dim \left(\bigoplus_{p=0}^n \Lambda^p(V) \right) = \sum_{p=0}^n \dim \Lambda^p(V) = \sum_{p=0}^n C_n^p = 2^n.$$

62. ТЕНЗОРЫ НА ЕВКЛИДОВОМ ПРОСТРАНСТВЕ

Пусть V — евклидово пространство. Тогда соответствие $v \mapsto (v, \cdot)$ задаёт канонический изоморфизм $V \rightarrow V^*$, т. е. позволяет отождествить тензоры типа $(0, 1)$ с тензорами типа $(1, 0)$. В координатах это выглядит следующим образом. Пусть $G = (g_{ij})$ — матрица Грама скалярного произведения. Т. к. скалярное произведение — это билинейная функция, G является тензором типа $(2, 0)$. Тогда при изоморфизме $V \rightarrow V^*$ вектор T с координатами T^i переходит в ковектор с координатами $T_j = g_{ij} T^i$. Таким образом, мы «опустили индекс» у тензора T при помощи фиксированного тензора типа $(2, 0)$. Эта операция обобщается следующим образом.

Определение 1. *Опускание индекса* — это линейное отображение $\mathbb{T}_p^q(V) \rightarrow \mathbb{T}_{p+1}^{q-1}(V)$ тензоров в евклидовом пространстве V , которое тензору $T = \{T_{i_1, \dots, i_p}^{j_1, \dots, j_q}\}$ типа (p, q) ставит в соответствие тензор $\{g_{ij} T_{i_1, \dots, i_p}^{j_1, \dots, j_q}\}$ типа $(p+1, q-1)$.

Пример 1. Рассмотрим изоморфизм $\psi : \text{End}(V) \rightarrow \text{B}(V)$ между пространством операторов и пространством билинейных функций, сопоставляющий оператору \mathcal{A} билинейную функцию $\mathcal{B}_{\mathcal{A}} = (\mathcal{A}, \cdot)$. В координатах это выглядит так: оператору с матрицей $A = (a_k^i)$ сопоставляется билинейная функция с матрицей $B_A = (g_{ij} a_k^i)$. Таким образом, B_A — это тензор типа $(2, 0)$, получаемый в результате опускания индекса у тензора A типа $(1, 1)$.

Для того, чтобы определить операцию, обратную к опусканию индекса, рассмотрим набора $\{g^{kl}\}$, состоящий из элементов обратной матрицы к матрице Грама $G = (g_{ij})$, т. е. $g^{kl} g_{lj} = \delta_j^k$.

Предложение 1. $\{g^{kl}\}$ является тензором типа $(0, 2)$.

Доказательство. Необходимо поверить тензорный закон. В новом базисе мы имеем $g^{k'i'} g_{i'j'} = \delta_{j'}^{k'}$. Т. к. $\{g_{ij}\}$ — тензор типа $(2, 0)$, мы имеем $g_{i'j'} = c_{i'}^i c_{j'}^j g_{ij}$. Подставив это в предыдущую формулу, получим $g^{k'i'} c_{i'}^i c_{j'}^j g_{ij} = \delta_{j'}^{k'}$. Теперь умножим обе части этого равенства на компоненты обратной матрицы $c_{k'}^{j'}$ (и просуммируем по j'): $g^{k'i'} c_{i'}^i c_{j'}^j c_{k'}^{j'} g_{ij} = \delta_{j'}^{k'} c_{k'}^{j'}$. Т. к. $c_{j'}^j c_{k'}^{j'} = \delta_k^{j'}$, отсюда получаем $g^{k'i'} c_{i'}^i g_{ik} = c_k^{k'}$. Далее умножим обе части на g^{kl} : $g^{k'i'} c_{i'}^i c_l^{l'} = c_k^{k'} c_l^{l'} g^{kl}$. Т. к. $c_{i'}^i c_l^{l'} = \delta_{i'}^{l'}$, окончательно получаем $g^{k'l'} = c_k^{k'} c_l^{l'} g^{kl}$. Это и есть тензорный закон преобразования для тензора типа $(0, 2)$. ■

Определение 2. *Поднятие индекса* — это линейное отображение $\mathbb{T}_p^q(V) \rightarrow \mathbb{T}_{p-1}^{q+1}(V)$ тензоров в евклидовом пространстве V , которое тензору $T = \{T_{i_1, \dots, i_p}^{j_1, \dots, j_q}\}$ типа (p, q) ставит в соответствие тензор $\{g^{ij} T_{i_1, \dots, i_p}^{j_1, \dots, j_q}\}$ типа $(p-1, q+1)$.

Операция опускания (или поднятия) индекса является композицией тензорного произведения с тензором $\{g_{ij}\}$ (или $\{g^{ij}\}$) и свёртки.

63. КОММУТИРУЮЩИЕ ОПЕРАТОРЫ НА КОНЕЧНОМЕРНОМ КОМПЛЕКСНОМ ВЕКТОРНОМ ПРОСТРАНСТВЕ

Теорема 1. Произведение эрмитовых операторов $\mathcal{A}\mathcal{B}$ является эрмитовым тогда и только тогда, когда \mathcal{A} и \mathcal{B} коммутируют.

Доказательство. $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A} \Leftrightarrow (\mathcal{A}\mathcal{B})^* = (\mathcal{B}\mathcal{A})^* = \mathcal{A}^*\mathcal{B}^* = \mathcal{A}\mathcal{B}$. ■

Лемма 1. Пусть \mathcal{A}, \mathcal{B} — коммутирующие операторы на комплексном пространстве V . Тогда \mathcal{A} и \mathcal{B} имеют общий собственный вектор.

Доказательство. Пусть λ — собственное значение оператора \mathcal{A} , а V_λ — соответствующее ему собственное подпространство. Как известно, оно инвариантно относительно \mathcal{A} . Тогда оно инвариантно и относительно оператора \mathcal{B} . Действительно, пусть $v \in V_\lambda$, тогда

$$\mathcal{A}(\mathcal{B}v) = \mathcal{B}(\mathcal{A}v) = \mathcal{B}(\lambda v) = \lambda(\mathcal{B}v),$$

следовательно, $\mathcal{B}v \in V_\lambda$. Оператор $\mathcal{B}|_{V_\lambda}$ имеет собственный вектор $u \in V_\lambda$: $\mathcal{B}u = \mu u$, μ — собственное значение \mathcal{B} . Таким образом, $\mathcal{A}u = \lambda u$, $\mathcal{B}u = \mu u$, т. е. u — общий собственный вектор. ■

Теорема 2. Два эрмитовых (унитарных) оператора \mathcal{A}, \mathcal{B} в эрмитовом пространстве V одновременно приводятся к диагональному виду в некотором ортонормированном базисе тогда и только тогда, когда они коммутируют.

Доказательство. \Rightarrow . Предположим, что \mathcal{A} и \mathcal{B} диагонализированы в общем ортонормированном базисе, мы приходим к выводу о перестановочности их матриц A, B в этом базисе. Но очевидно, что коммутирующие матрицы коммутируют в любом базисе. Это устанавливается цепочкой равенств

$$C^{-1}AC \cdot C^{-1}BC = C^{-1}ABC = C^{-1}BAC = C^{-1}BC \cdot C^{-1}AC.$$

Таким образом, коммутируют и сами операторы.

\Leftarrow . Обратно, пусть $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$. Тогда по предыдущей лемме операторы \mathcal{A} и \mathcal{B} имеют общий собственный вектор v . Без ограничения общности, можно считать $|v| = 1$. Подпространство $W = \langle v \rangle^\perp$ размерности $n-1$ инвариантно относительно \mathcal{A} и \mathcal{B} , т. к. они эрмитовы или унитарны (важная лемма). Ограничения \mathcal{A} и \mathcal{B} на W будут коммутирующими эрмитовыми (унитарными) операторами. Индукция по размерности приводит к явной конструкции ортонормированного базиса, в котором \mathcal{A} и \mathcal{B} запишутся в диагональной форме. ■

64. КОСОСИММЕТРИЧЕСКИЕ ФОРМЫ. СИМПЛЕКТИЧЕСКАЯ ГРУППА

Пусть \mathcal{B} — кососимметрическая билинейная функция в пространстве над полем характеристики не 2. Соответствующую ей кососимметрическую билинейную форму $B(x, y) = b_{ij}x^i y^j$, где $b_{ji} = -b_{ij}$, можно представить в виде

$$B(x, y) = \sum_{i < j} b_{ij}(x^i y^j - x^j y^i).$$

Теорема 1. Для любой кососимметрической билинейной функции \mathcal{B} над полем характеристики не 2, существует базис, в котором её матрица блочно-диагональная с блоками размера 1 или 2, причём блоки размера 1 нулевые, а блоки размера 2 имеют вид $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Другими словами, любую кососимметрическую билинейную форму $B(x, y)$, линейной заменой координат можно привести к виду

$$(x^1 y^2 - x^2 y^1) + (x^3 y^4 - x^4 y^3) + \dots + (x^{2k-1} y^{2k} - x^{2k} y^{2k-1}).$$

Доказательство. Проведём индукцию по размерности пространства V .

При $\dim V = 1$ доказывать нечего, т. к. кососимметрическая функция нулевая (здесь пользуемся тем, что характеристика поля не равна 2).

Пусть $\dim V = 2$. Тогда матрица кососимметрической функции в произвольном базисе e_1, e_2 имеет вид $\begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}$, где $b = b_{12} = \mathcal{B}(e_1, e_2)$. Пусть $b \neq 0$ (иначе мы уже имеем два блока из нулей). Тогда в новом базисе $e_{1'} = e_1$ и $e_{2'} = \frac{1}{b}e_2$ мы имеем

$$b'_{12} = \mathcal{B}(e_{1'}, e_{2'}) = \mathcal{B}\left(e_1, \frac{1}{b}e_2\right) = \frac{1}{b}\mathcal{B}(e_1, e_2) = 1,$$

т. е. матрица кососимметрической формы имеет требуемый вид $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Теперь предположим, что утверждение уже доказано для пространств размерности меньше n , и докажем его для размерности n . Можно считать, что функция \mathcal{B} не является тождественно нулевой (иначе доказывать нечего). Пусть $b_{12} = \mathcal{B}(e_1, e_2) \neq 0$ для некоторого базиса e_1, \dots, e_n . Попробуем заменить базисные векторы так, чтобы новые векторы $e_{1'}, \dots, e_{n'}$ удовлетворяли соотношениям

$$\mathcal{B}(e_{1'}, e_{2'}) = 1, \quad \mathcal{B}(e_{1'}, e_{i'}) = \mathcal{B}(e_{2'}, e_{i'}) = 0 \quad \text{при } i \geq 3. \quad (*)$$

Новый базис будем искать в виде

$$e_{1'} = e_1, \quad e_{2'} = ce_2, \quad e_{i'} = e_i + c_i^1 e_1 + c_i^2 e_2 \quad \text{при } i \geq 3.$$

Подставив эти соотношения в $(*)$, получим $c = \frac{1}{b_{12}}$ и

$$0 = \mathcal{B}(e_{1'}, e_{i'}) = \mathcal{B}(e_1, e_i + c_i^1 e_1 + c_i^2 e_2) = b_{1i} + c_i^2 b_{12},$$

откуда $c_{i'}^2 = -\frac{b_{1i}}{b_{12}}$. Аналогично,

$$0 = \mathcal{B}(e_{2'}, e_{i'}) = \mathcal{B}(ce_2, e_i + c_i^1 e_1 + c_i^2 e_2) = c(b_{2i} - c_i^1 b_{12}),$$

откуда $c_{i'}^1 = \frac{b_{2i}}{b_{12}}$. Окончательно, наша замена базиса имеет вид

$$e_{1'} = e_1, \quad e_{2'} = \frac{1}{b_{12}}e_2, \quad e_{i'} = e_i + \frac{b_{2i}}{b_{12}}e_1 - \frac{b_{1i}}{b_{12}}e_2 \quad \text{при } i \geq 3.$$

Ввиду соотношений $(*)$, в новом базисе матрица билинейной функции \mathcal{B} будем иметь вид

$$B' = \left(\begin{array}{cc|ccc} 0 & 1 & 0 & \dots & 0 \\ -1 & 0 & 0 & \dots & 0 \\ \hline 0 & 0 & & & \\ \vdots & \vdots & & \tilde{B}' & \\ 0 & 0 & & & \end{array} \right),$$

где матрица \tilde{B}' – матрица билинейной функции \mathcal{B} на подпространстве $\langle e_{3'}, \dots, e_{n'} \rangle$. Т. к. это пространство имеет размерность $n-2$, по предположению индукции в нём существует требуемый базис $e_{3'"}, \dots, e_{n''}$. Тогда в базисе $e_{1'}, e_{2'}, e_{3'"}, \dots, e_{n''}$ исходного пространства V матрица кососимметрической функции \mathcal{B} имеет требуемый вид. ■

Определение 1. Вид, описанный в последней теореме, называется *нормальным видом* кососимметрической билинейной формы.

Следствие 1. Ранг кососимметрической билинейной функции — чётное число; кососимметрическая билинейная функция в пространстве нечётной размерности всегда вырождена.

Ясно, что нормальный вид кососимметрической билинейной формы зависит только от её ранга, и поэтому мы получаем:

Предложение 1. Две кососимметрические билинейные формы получаются друг из друга линейной заменой координат тогда и только тогда, когда их ранги совпадают.

Про приведение кососимметрической билинейной формы к каноническому виду ПИСАЛ Антон Александрович Клячко.

Евклидово пространство можно определить как вещественное линейное пространство с фиксированной невырожденной симметрической билинейной функцией. Аналогично вводится следующее определение.

Определение 2. *Симплектическим пространством* называется вещественное пространство V с фиксированной невырожденной кососимметрической билинейной функцией, которая обозначается $\omega : V \times V \rightarrow \mathbb{R}, (u, v) \mapsto \omega(u, v)$.

Из последнего следствия вытекает, что размерность симплектического пространства V всегда чётна; мы фиксируем обозначение $\dim V = n = 2m$.

Определение 3. *Ортогональным дополнением* подпространства U в симплектическом пространстве V называется подпространство

$$U^\perp := \{v \in V : \omega(u, v) = 0 \ \forall u \in U\}.$$

Предложение 2. Для любых подпространств U и W симплектического пространства V имеем:

1. $\dim U^\perp = n - \dim U$;
2. $(U^\perp)^\perp = U$;
3. $(U + W)^\perp = U^\perp \cap W^\perp$;
4. $(U \cap W)^\perp = U^\perp + W^\perp$.

Доказательство.

1. Выберем базис e_1, \dots, e_k в U и дополним его до базиса e_1, \dots, e_n в V . Пусть в этом базисе кососимметрическая функция ω задаётся матрицей $B = (b_{ij})$, т. е. $\omega(e_i, e_j) = b_{ij}$. Тогда для вектора $x = x^j e_j$ условие принадлежности подпространству U^\perp записывается k линейными уравнениями

$$\omega(e_i, x) = \omega(e_i, x^j e_j) = b_{ij} x^j = 0, \quad i = 1, \dots, k.$$

Матрица этой системы уравнений представляет собой $(k \times n)$ -матрицу, составленную из первых k строк матрицы B . Т. к. матрица B невырождена (имеет ранг n), ранг матрицы системы равен k , а значит, размерность пространства решений равна $n - k = n - \dim U$, что и требовалось.

2. Т. к. ω — кососимметрическая функция, равенство $\omega(u, v) = 0$ для любых $u \in U$ и $v \in U^\perp$ означает, что $U \subseteq (U^\perp)^\perp$. Из утверждения 1 вытекает, что $\dim(U^\perp)^\perp = n - (n - \dim U) = \dim U$, поэтому $U = (U^\perp)^\perp$.
3. Докажем, что $(U + W)^\perp \subseteq U^\perp \cap W^\perp$. Пусть $v \in (U + W)^\perp$, т. е. $\omega(u + w, v) = 0 \ \forall u \in U, w \in W$. Полагая $w = \mathbf{0}$ и $u = \mathbf{0}$ мы получим, соответственно, $v \in U^\perp$ и $v \in W^\perp$, откуда $v \in U^\perp \cap W^\perp$. Теперь докажем, что $U^\perp \cap W^\perp \subseteq (U + W)^\perp$. Пусть $v \in U^\perp \cap W^\perp$. Тогда $\omega(u, v) = 0 \ \forall u \in U$ и $\omega(w, v) = 0 \ \forall w \in W$. Следовательно, $\omega(u + w, v) = 0$, т. е. $v \in (U + W)^\perp$.

4. Ясно, что $U^\perp \subseteq (U \cap W)^\perp$ и $W^\perp \subseteq (U \cap W)^\perp$, откуда $U^\perp + W^\perp \subseteq (U \cap W)^\perp$. Далее, имеем

$$\begin{aligned} \dim(U^\perp + W^\perp) &= \dim U^\perp + \dim W^\perp - \dim(U^\perp \cap W^\perp) = \\ &= \dim U^\perp + \dim W^\perp - \dim(U + W)^\perp = \\ &= n - \dim U + n - \dim W - n + \dim(U + W) = n - \dim(U \cap W) = \dim(U \cap W)^\perp. \end{aligned}$$

Поэтому $(U \cap W)^\perp = U^\perp + W^\perp$.

■

Примечание. Последнее предложение выполнено для произвольной невырожденной билинейной функции, но при этом необходимо различать левое и правое ортогональные дополнения.

В отличие от евклидовых пространств, равенство $V = U \oplus U^\perp$ может не иметь места в симплектическом пространстве.

Определение 4. Подпространство $U \subseteq V$ называется *изотропным*, если $\omega|_U = \mathcal{O}$.

Другими словами, U изотропно, если $U \subseteq U^\perp$. Для изотропного пространства мы имеем $\dim U \leq \dim U^\perp = n - \dim U$, откуда $\dim U \leq m = \frac{n}{2}$.

Определение 5. Изотропное подпространство L максимальной размерности $m = \frac{\dim V}{2}$ называется *лагранжевым*. Для такого подпространства имеем $L = L^\perp$.

Из теоремы о нормальном виде кососимметрической функции вытекает, что существует базис e_1, \dots, e_{2m} , для которого

$$\omega(e_i, e_j) = \begin{cases} 1, & \text{если } i \text{ нечётно и } j = i + 1, \\ 0, & \text{для всех остальных } i \text{ и } j > i. \end{cases}$$

Рассмотрим новый базис, получаемый перенумерацией векторов e_1, \dots, e_{2m} :

$$a_i = e_{2i-1}, \quad b_i = e_{2i}, \quad i = 1, \dots, m.$$

Определение 6. Такой базис $a_1, \dots, a_m, b_1, \dots, b_m$ удовлетворяет соотношениям

$$\omega(a_i, b_j) = \delta_{ij}, \quad \omega(a_i, a_j) = \omega(b_i, b_j) = 0$$

и называется *симплектическим* или *гамильтоновым* базисом.

Матрица функции ω в симплектическом базисе имеет вид $\begin{pmatrix} 0 & E \\ -E & 0 \end{pmatrix}$.

Из существования симплектического базиса сразу вытекает существование лагранжевых подпространств. Для каждого подмножества индексов $I = \{i_1, \dots, i_k\} \subseteq \{1, \dots, m\}$ (возможно, пустого) положим

$$L_I := \langle a_i : i \in I, b_j : j \notin I \rangle.$$

Ясно, что $\dim L_I = m$ и L_I — лагранжево подпространство. В частности, $L_{\{1, \dots, m\}} = \langle a_1, \dots, a_m \rangle$ и $L_\emptyset = \langle b_1, \dots, b_m \rangle$.

Определение 7. Лагранжевы подпространства вида L_I называются *стандартными* для данного симплектического базиса.

Лемма 1. Любое изотропное подпространство U содержится в некотором лагранжевом.

Доказательство. Пусть $\dim U = k$. Рассмотрим $\omega|_{U^\perp}$ — ограничение кососимметрической функции ω на подпространство U^\perp . Т.к. $\omega|_U = \mathcal{O}$ и $U \subseteq U^\perp$, кососимметрическая функция $\omega|_{U^\perp}$ вырождена (если $U \neq \{0\}$), а её ранг равен

$$\dim U^\perp - \dim(U^\perp)^\perp = \dim U^\perp - \dim U = n - k - k = 2(m - k).$$

Приведя билинейную функцию $\omega|_{U^\perp}$ к нормальному виду, получим, что в некотором базисе $a_1, \dots, a_{m-k}, b_1, \dots, b_{m-k}, c_1, \dots, c_k$ её матрица будет иметь вид

$$\begin{pmatrix} 0 & E & 0 \\ -E & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

При этом $U = \langle c_1, \dots, c_k \rangle$. Рассмотрим $L = \langle b_1, \dots, b_{m-k}, c_1, \dots, c_k \rangle$. Тогда $U \subseteq L$ и $\dim L = m$. Наконец, L лагранжево, т. к. матрица ограничения $\omega|_L$ есть правый нижний квадрат из нулей в матрице ограничения $\omega|_{U^\perp}$. ■

Теорема 2. Для любого лагранжева подпространства $L \subseteq V$ существует такое лагранжево подпространство \tilde{L} , что

$$V = L \oplus \tilde{L},$$

причём, подпространство \tilde{L} можно выбрать среди стандартных подпространств L_I , отвечающих произвольному наперёд заданному симплектическому базису.

Доказательство. Рассмотрим подпространство $L_{\{1, \dots, m\}} = \langle a_1, \dots, a_m \rangle$ и пусть I — такое подмножество индексов, что векторы a_i , $i \in I$, порождают в $L_{\{1, \dots, m\}}$ подпространство, дополнительное к $L_{\{1, \dots, m\}} \cap L$:

$$(L_{\{1, \dots, m\}} \cap L) \oplus \langle a_i : i \in I \rangle = L_{\{1, \dots, m\}}$$

(для этого нужно выбрать базис e_1, \dots, e_k в $L_{\{1, \dots, m\}} \cap L$ и затем выбросить из семейства векторов $e_1, \dots, e_k, a_1, \dots, a_m$ все векторы, линейно выражающиеся через предыдущие). Т. к. $L_{\{1, \dots, m\}} \cap L \subseteq L = L^\perp$ и $\langle a_i : i \in I \rangle \subseteq L_I = L_I^\perp$, мы имеем

$$L_{\{1, \dots, m\}} = (L_{\{1, \dots, m\}} \cap L) \oplus \langle a_i : i \in I \rangle \subseteq L^\perp + L_I^\perp = (L \cap L_I)^\perp.$$

Следовательно, $L \cap L_I \subseteq L_{\{1, \dots, m\}}^\perp = L_{\{1, \dots, m\}}$. Тогда

$$L \cap L_I = (L \cap L_{\{1, \dots, m\}}) \cap (L_I \cap L_{\{1, \dots, m\}}) = (L \cap L_{\{1, \dots, m\}}) \cap \langle a_i : i \in I \rangle = \{0\}.$$

Т. к. $\dim L + \dim L_I = \dim V$, мы получаем, что $V = L \oplus L_I$. ■

Определение 8. Пусть V — вещественное пространство. Скажем, что на V задано *обобщённое скалярное произведение*, если на V задана фиксированная невырожденная билинейная функция \mathcal{B} . По аналогии с обычным скалярным произведением будем обозначать обобщённое скалярное произведение через $(x, y) := \mathcal{B}(x, y)$.

Определение 9. Вектор $v \in V$ называется *изотропным*, если $(v, v) = 0$.

Примерами являются евклидово и симплектическое пространства. В евклидовом пространстве изотропным является только вектор 0 , а в симплектическом — вообще все вектора.

Определение 10. Оператор $\mathcal{A} : V \rightarrow V$ *сохраняет обобщённое скалярное произведение*, если $(\mathcal{A}x, \mathcal{A}y) = (x, y) \forall x, y \in V$.

Предложение 3. Пусть обобщённое скалярное произведение задаётся матрицей G в некотором базисе. Тогда оператор, заданный матрицей A в том же базисе, сохраняет обобщённое скалярное произведение тогда и только тогда, когда выполнено соотношение $A^t G A = G$.

Доказательство. Пусть в базисе e_1, \dots, e_n имеем $\mathcal{A}e_i = a_i^j e_j$ и $g_{ij} = (e_i, e_j)$. Тогда для матриц $A = (a_i^j)$ и $G = (g_{ij})$ имеем

$$(\mathcal{A}e_i, \mathcal{A}e_j) = (a_i^k e_k, a_j^l e_l) = a_i^k a_j^l (e_k, e_l) = a_i^k g_{kl} a_j^l = (A^t G A)_{ij}.$$

Поэтому соотношения $(\mathcal{A}e_i, \mathcal{A}e_j) = (e_i, e_j)$ для всех i, j эквивалентны матричному соотношению $A^t G A = G$. ■

Предложение 4. Операторы, сохраняющие обобщённое скалярное произведение, образуют группу относительно композиции.

Доказательство. Очевидно, что тождественный оператор сохраняет обобщённое скалярное произведение, и композиция операторов с этим свойство также обладает этим свойством. Кроме того, если оператор \mathcal{A} , сохраняющий скалярное произведение, обратим, то \mathcal{A}^{-1} также сохраняет скалярное произведение:

$$(\mathcal{A}^{-1}x, \mathcal{A}^{-1}y) = (\mathcal{A}\mathcal{A}^{-1}x, \mathcal{A}\mathcal{A}^{-1}y) = (x, y).$$

Существование обратного оператора вытекает из соотношения $A^tGA = G$. Действительно, вычисление определителя даёт $(\det A)^2 \det G = \det G$; т. к. матрица G невырождена (потому что билинейная функция \mathcal{B} невырождена), получаем $\det A \neq 0$. ■

Определение 11. Группа операторов, сохраняющих симплектическое скалярное произведение в \mathbb{R}^{2m} , называется *симплектической группой* и обозначается $\text{Sp}(2m)$. В гамильтоновом базисе матрица A симплектического оператора удовлетворяет соотношению $A^tGA = G$, где $G = \begin{pmatrix} 0 & E \\ -E & 0 \end{pmatrix}$ с блоками размера m . Такие матрицы называются *симплектическими*.

65. ПРИЛОЖЕНИЯ

Здесь описаны некоторые сюжеты, которые я у кого-то видел, но которые, к сожалению, не вошли в программу экзамена.

65.1. ПРОЕКТОРЫ

Определение 1. Пусть пространство V представимо в виде прямой суммы двух подпространств: $V = V_1 \oplus V_2$. Тогда $\forall v \in V$ имеется единственное разложение $v = v_1 + v_2$, $v_1 \in V_1$, $v_2 \in V_2$. Оператор $\mathcal{P} : V \rightarrow V$, переводящий вектор $v = v_1 + v_2$ в вектор v_1 , называется *проектором* на V_1 вдоль V_2 .

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

Теорема 1. Оператор $\mathcal{A} : V \rightarrow V$ является проектором тогда и только тогда, когда $\mathcal{A}^2 = \mathcal{A}$.

Доказательство. \Rightarrow . Если \mathcal{A} — проектор на V_1 вдоль V_2 , то для $v = v_1 + v_2$ имеем

$$\mathcal{A}^2 v = \mathcal{A}(\mathcal{A}(v_1 + v_2)) = \mathcal{A}v_1 = v_1 = \mathcal{A}v,$$

т. е. $\mathcal{A}^2 = \mathcal{A}$.

\Leftarrow . Положим $V_1 := \text{Im } \mathcal{A}$, $V_2 := \text{Ker } \mathcal{A}$. Мы покажем, что \mathcal{A} — проектор на V_1 вдоль V_2 .

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

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

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

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

т. е. $V = V_1 \oplus V_2$.

Рассмотрим теперь произвольный вектор $v \in V$ и представим его в виде $v = v_1 + v_2$, $v_1 \in V_1 = \text{Im } \mathcal{A}$, $v_2 \in V_2 = \text{Ker } \mathcal{A}$. Тогда существует $u \in V$ такой, что $v_1 = \mathcal{A}u$, а $\mathcal{A}v_2 = \mathbf{0}$. Мы имеем

$$\mathcal{A}v = \mathcal{A}v_1 + \mathcal{A}v_2 = \mathcal{A}(\mathcal{A}u) = \mathcal{A}^2 u = \mathcal{A}u = v_1.$$

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

Приведём здесь один простой факт из теории множеств.

Предложение 1. Отображение $f : X \rightarrow X$ тождественно на своём образе тогда и только тогда, когда $f^2 = f$.

Доказательство. \Rightarrow . $\forall x \in X \underbrace{f(f(x))}_{\in \text{Im } f} = f(x) \Rightarrow f^2 = f$.

\Leftarrow . $\forall x \in X \ f(f(x)) = f(x)$. Таким образом, $\forall (y := f(x)) \in \text{Im } f \ f(y) = y \Rightarrow f|_{\text{Im } f} = \text{id}$. ■

Таким образом, линейный оператор $\mathcal{A} : V \rightarrow V$ является проектором тогда и только тогда, когда он тождественен на своём образе.

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

Задача 1 (А. А. Клячко). Докажите, что оператор \mathcal{B} коммутирует с проектором \mathcal{A} тогда и только тогда, когда $\text{Im } \mathcal{A}$ и $\text{Ker } \mathcal{A}$ являются инвариантными подпространствами для \mathcal{B} .

$\triangleright \Rightarrow$. Пусть $v \in \text{Im } \mathcal{A}$. Тогда существует $u \in V$ такой, что $\mathcal{A}u = v$. При этом

$$\mathcal{B}v = \mathcal{B}(\mathcal{A}u) = \mathcal{A}(\mathcal{B}u) \in \text{Im } \mathcal{A},$$

т. е. подпространство $\text{Im } \mathcal{A}$ инвариантно для \mathcal{B} . Теперь пусть $w \in \text{Ker } \mathcal{A}$. Тогда

$$\mathcal{A}(\mathcal{B}w) = \mathcal{B}(\mathcal{A}w) = \mathcal{B}\mathbf{0} = \mathbf{0} \in \text{Ker } \mathcal{A},$$

т. е. подпространство $\text{Ker } \mathcal{A}$ инвариантно для \mathcal{B} .

\Leftarrow . Т. к. \mathcal{A} — проектор, то $V = \text{Im } \mathcal{A} + \text{Ker } \mathcal{B}$, т. е. $\forall v \in V$ можем записать $v = v_1 + v_2$, где $v_1 \in \text{Im } \mathcal{A}$, $v_2 \in \text{Ker } \mathcal{A}$. Так что силу линейности операторов нам достаточно проверить коммутирование отдельно на векторах из $\text{Im } \mathcal{A}$ и из $\text{Ker } \mathcal{A}$.

Пусть $u \in \text{Im } \mathcal{A}$, тогда $\mathcal{B}u \in \text{Im } \mathcal{A}$. Как обсуждалось выше, проектор тождественен на своём образе, так что

$$\mathcal{A}(\mathcal{B}u) = \mathcal{B}u, \quad \mathcal{B}(\mathcal{A}u) = \mathcal{B}u.$$

Теперь пусть $w \in \text{Ker } \mathcal{A}$, тогда $\mathcal{B}w \in \text{Ker } \mathcal{A}$. Поэтому

$$\mathcal{A}(\mathcal{B}w) = \mathbf{0}, \quad \mathcal{B}(\mathcal{A}w) = \mathcal{B}\mathbf{0} = \mathbf{0}.$$

◀

Задача 2 (А. А. Клячко). Докажите, что два проектора \mathcal{A} и \mathcal{B} коммутируют тогда и только тогда, когда пространство V раскладывается в прямую сумму своих подпространств следующим образом:

$$V = \text{Ker } \mathcal{A} \cap \text{Ker } \mathcal{B} \oplus \text{Im } \mathcal{A} \cap \text{Ker } \mathcal{B} \oplus \text{Ker } \mathcal{A} \cap \text{Im } \mathcal{B} \oplus \text{Im } \mathcal{A} \cap \text{Im } \mathcal{B}.$$

$\triangleright \Rightarrow$. Проверим, что для каждого из подпространств в прямой сумме выполнено, что его пересечение с суммой остальных тривиально (см. теорему 2 в вопросе 5, пункт 2). Для простоты изложения введём обозначения для векторов соответствующих подпространств:

$$v_1 \in \text{Ker } \mathcal{A} \cap \text{Ker } \mathcal{B}, \quad v_2 \in \text{Im } \mathcal{A} \cap \text{Ker } \mathcal{B}, \quad v_3 \in \text{Ker } \mathcal{A} \cap \text{Im } \mathcal{B}, \quad v_4 \in \text{Im } \mathcal{A} \cap \text{Im } \mathcal{B}.$$

Предположим, что $v_1 = v_2 + v_3 + v_4$ (то есть, что v_1 лежит в первом подпространстве и в сумме трёх остальных). В первом равенстве подействуем оператором \mathcal{AB} на обе части. При этом обнулится всё, кроме $\mathcal{A}(\mathcal{B}v_4)$ (остальные векторы лежат в ядре одного из операторов \mathcal{A} и \mathcal{B}). Поэтому равенство принимает вид $\mathcal{A}(\mathcal{B}v_4) = \mathbf{0}$, т. е. $v_4 \in \text{Ker } \mathcal{A}$, но $v_4 \in \text{Im } \mathcal{A}$, причём $\text{Ker } \mathcal{A} \cap \text{Im } \mathcal{A} = \{\mathbf{0}\}$. Значит, $v_4 = \mathbf{0}$. Теперь равенство принимает вид $v_1 = v_2 + v_3$. Подействуем оператором \mathcal{A} на обе его части, получим $\mathcal{A}v_2 = \mathbf{0}$. Таким образом, $v_2 \in \text{Ker } \mathcal{A} \cap \text{Im } \mathcal{A} \Rightarrow v_2 = \mathbf{0}$. Итак, выражение приняло вид $v_1 = v_3$. Но тогда получаем, что вектор v_1 лежит одновременно и в $\text{Ker } \mathcal{B}$, и в $\text{Im } \mathcal{B}$, то есть, он тоже нулевой. И наконец, $v_1 = \mathbf{0}$, что и требовалось.

Остальные случаи разбираются аналогично.

\Leftarrow . Заметим, что $v_2 \in \text{Im } \mathcal{A}$, а проектор \mathcal{A} тождественен на своём образе, так что $\mathcal{A}v_2 = v_2$. Отсюда, $\mathcal{B}(\mathcal{A}v_2) = \mathcal{B}v_2 = \mathbf{0}$. Аналогично, $\mathcal{A}(\mathcal{B}v_3) = \mathbf{0}$. А вектор v_4 лежит в образах обоих проекторов, поэтому $\mathcal{A}(\mathcal{B}v_4) = \mathcal{B}(\mathcal{A}v_4) = v_4$.

$$\begin{aligned} \mathcal{A}(\mathcal{B}(v_1 + v_2 + v_3 + v_4)) &= \mathcal{A}(\mathcal{B}v_3 + \mathcal{B}v_4) = \overbrace{\mathcal{A}(\mathcal{B}v_3)}^{\mathbf{0}} + \mathcal{A}(\mathcal{B}v_4) = \mathcal{A}(\mathcal{B}v_4) = v_4, \\ \mathcal{B}(\mathcal{A}(v_1 + v_2 + v_3 + v_4)) &= \mathcal{B}(\mathcal{A}v_2 + \mathcal{A}v_4) = \underbrace{\mathcal{B}(\mathcal{A}v_2)}_{\mathbf{0}} + \mathcal{B}(\mathcal{A}v_4) = \mathcal{B}(\mathcal{A}v_4) = v_4. \end{aligned}$$

Таким образом, \mathcal{A} и \mathcal{B} коммутируют. ◀

65.2. МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

Часто на практике при исследовании какого-нибудь природного или социального явления делается допущение, что это явление описывается линейной формулой. Точнее, мы предполагаем,

что некоторая величина b линейно зависит от других величин a_1, \dots, a_n , и мы хотим найти эту зависимость

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

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

[illegible]

в которой число неизвестных меньше числа уравнений. Такая система, как правило, несовместна. Поэтому находится «наилучшее приближённое» решение x^1, \dots, x^n , для которого отклонение значений b^i от $a_i^j x^j = a_1^i x^1 + \dots + a_n^i x^n$ будет наименьшим.

Метод наименьших квадратов решает эту задачу нахождения наилучшего приближённого решения в предположении, что в качестве меры отклонения берётся сумма квадратов разностей величин $a_1^i x^1 + \dots + a_n^i x^n$ и b^i .

Определение 1. *Псевдорешением системы*

[illegible]

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

$$(a_j^1 x^j - b^1)^2 + (a_j^2 x^j - b^2)^2 + \dots + (a_j^m x^j - b^m)^2$$

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

Пусть $A = (a_j^i)$ — матрица СЛУ, $a_1, \dots, a_n \in \mathbb{R}^m$ — вектор-столбцы этой матрицы, а $b \in \mathbb{R}^m$ — вектор правых частей.

Теорема 1. Псевдорешение системы $Ax = b$ находится как решение системы

[illegible]

Другими словами, псевдорешение — это набор коэффициентов в разложении проекции $\text{pr}_{\langle a_1, \dots, a_n \rangle} b$ по векторам a_1, \dots, a_n , а квадратичное отклонение псевдорешения — это квадрат длины вектора $\text{ort}_{\langle a_1, \dots, a_n \rangle} b$.

Доказательство. Квадратичное отклонение набора x^1, \dots, x^n — это по определению квадрат длины вектора $a_j x^j - b$, т. е. квадрат расстояния между векторами b и $a_j x^j \in \langle a_1, \dots, a_n \rangle$. Мы знаем из предложения 2 в вопросе 35, что расстояние между b и точкой $a_j x^j$ подпространства $\langle a_1, \dots, a_n \rangle$ минимально, когда $a_j x^j$ — это проекция вектора b на $\langle a_1, \dots, a_n \rangle$. Коэффициенты в разложении проекции по векторам подпространства находятся из указанной системы (предложение 2 в вопросе 34), а минимальное расстояние равно $|\text{ort}_{\langle a_1, \dots, a_n \rangle} b|$. ■

65.3. СПЕКТРАЛЬНОЕ РАЗЛОЖЕНИЕ САМОСОПРЯЖЁННОГО ОПЕРАТОРА

Определение 1. Проектор $\mathcal{P} : V \rightarrow V$ на U вдоль W называется *ортгогональным*, если $W = U^\perp$. Такой проектор будем обозначать через pr_U .

Это обозначение вполне согласуется с предыдущими: если $V = U \oplus U^\perp$, то $\forall v \in V$ мы имеем $v = \text{pr}_U v + \text{ort}_U v$.

Предложение 1. Проектор $\mathcal{P} : V \rightarrow V$ является самосопряжённым оператором тогда и только тогда, когда он ортгогонален.

Доказательство. Пусть $\mathcal{P} = \text{pr}_U$ — ортгогональный проектор. Выберем ортонормированный базис в U и дополним его до ортонормированного базиса в V . Тогда в этом ортонормированном базисе матрица оператора pr_U диагональна (с единицами и нулями на диагонали), а значит, оператор pr_U самосопряжён.

Обратно, пусть \mathcal{P} — самосопряжённый проектор на U вдоль W . Возьмём произвольные векторы $u \in U = \text{Im } \mathcal{P}$ и $w \in W = \text{Ker } \mathcal{P}$. Тогда $u = \mathcal{P}v$ для некоторого $v \in V$ и $\mathcal{P}w = \mathbf{0}$. Мы имеем

$$(u, w) = (\mathcal{P}v, w) = (v, \mathcal{P}w) = 0,$$

откуда получаем $W = U^\perp$ и $\mathcal{P} = \text{pr}_U$. ■

Пример 1. Пусть $u = (u^1, \dots, u^n)^t \in \mathbb{R}^n$ — ненулевой вектор-столбец и $\langle u \rangle$ — одномерное подпространство. Тогда матрица оператора $\text{pr}_{\langle u \rangle}$ в стандартном базисе \mathbb{R}^n есть

$$\frac{1}{|u|^2} uu^t = \frac{1}{|u|^2} \begin{pmatrix} u^1 \\ u^2 \\ \vdots \\ u^n \end{pmatrix} \cdot (u^1 \quad u^2 \quad \dots \quad u^n) = \frac{1}{|u|^2} \begin{pmatrix} u^1 u^1 & u^1 u^2 & \dots & u^1 u^n \\ u^2 u^1 & u^2 u^2 & \dots & u^2 u^n \\ \vdots & \vdots & \dots & \vdots \\ u^n u^1 & u^n u^2 & \dots & u^n u^n \end{pmatrix}.$$

Напомним, что спектром оператора \mathcal{A} называется множество его собственных значений. Для каждого собственного значения λ рассмотрим ортгогональный проектор pr_{V_λ} на соответствующее собственное подпространство V_λ .

Теорема 1 (Спектральное разложение). Пусть \mathcal{A} — самосопряжённый оператор. Тогда имеет место разложение

$$\mathcal{A} = \sum_{\lambda} \lambda \text{pr}_{V_\lambda},$$

где сумма берётся по всем собственным значениям. При этом проекторы pr_{V_λ} удовлетворяют соотношениям $\text{pr}_{V_\lambda} \mathcal{A} = \mathcal{A} \text{pr}_{V_\lambda} = \lambda \text{pr}_{V_\lambda}$ и $\text{pr}_{V_\lambda} \text{pr}_{V_\mu} = \mathcal{O}$ при $\lambda \neq \mu$.

Доказательство. Оператор \mathcal{A} коммутирует с проектором pr_{V_λ} , потому что $\text{Im } \text{pr}_{V_\lambda} = V_\lambda$ и $\text{Ker } \text{pr}_{V_\lambda} = V_\lambda^\perp$ являются инвариантными подпространствами для \mathcal{A} (задача 1 в приложении про проекторы). Действительно, первое очевидно, а второе следует из самосопряжённости оператора: если $v \in V_\lambda^\perp$, а $u \in V_\lambda$, то

$$(\mathcal{A}v, u) = (v, \mathcal{A}u) = (v, \lambda u) = \lambda(v, u) = 0.$$

А композиция проекторов на разные собственные подпространства тождественно нулевая, потому что (из решения задачи 2 в приложении про проекторы) их действие на вектор $v \in V$ есть его компонента из подпространства $\text{Im } \text{pr}_{V_\lambda} \cap \text{Im } \text{pr}_{V_\mu} = \{\mathbf{0}\}$. Коммутируют они по той же задаче (там очень вырожденный случай получается).

Разложению $V = \bigoplus_{\lambda} V_\lambda$ в прямую сумму собственных подпространств соответствует разложение тождественного оператора в сумму ортгогональных проекторов:

$$\text{id} = \sum_{\lambda} \text{pr}_{V_\lambda}.$$

Умножив это соотношение слева на \mathcal{A} и используя соотношение $\mathcal{A} \operatorname{gr}_{V_\lambda} = \lambda \operatorname{gr}_{V_\lambda}$, получим требуемое. ■

65.4. КОСОСИММЕТРИЧЕСКИЕ И КОСОЭРМИТОВЫ ОПЕРАТОРЫ. КАНОНИЧЕСКИЙ ВИД. ЭРМИТОВО РАЗЛОЖЕНИЕ

Определение 1. Оператор $\mathcal{A} : V \rightarrow V$ в евклидовом (эрмитовом) пространстве называется *кососимметрическим* (соответственно, *косоэрмитовым*), если $\mathcal{A}^* = -\mathcal{A}$, т. е. $\forall u, v \in V$ выполнено соотношение

$$(\mathcal{A}u, v) = -(u, \mathcal{A}v).$$

Предложение 1. Матрица A кососимметрического (косоэрмитова) оператора \mathcal{A} в ортонормированном базисе евклидова (эрмитова) пространства кососимметрична (косоэрмитова), т. е. $A^t = A$ (соответственно, $\bar{A}^t = A$).

Если матрица оператора \mathcal{A} в некотором ортонормированном базисе кососимметрична (косоэрмитова), то оператор \mathcal{A} кососимметричен (косоэрмитов).

Доказательство. То же, что и для самосопряжённых операторов. ■

Теорема 1. Для косоэрмитова оператора \mathcal{A} существует ортонормированный базис, в котором его матрица диагональна с чисто мнимыми числами на диагонали. Другими словами, для косоэрмитова оператора существует ортонормированный базис из собственных векторов, а все собственные значения — чисто мнимые.

Доказательство. Доказательство, как и в случае самосопряжённого оператора, основано на важной лемме. Будем вести индукцию по размерности пространства V . При $\dim V = 1$ доказывать нечего. Предположим, что утверждение доказано для операторов в пространствах размерности $n - 1$, и докажем его для размерности n .

Выберем собственный вектор v для \mathcal{A} , т. е. одномерное инвариантное подпространство $W = \langle v \rangle$. В силу важной леммы ортогональное дополнение W^\perp инвариантно относительно оператора \mathcal{A}^* , а значит, оно инвариантно и относительно $\mathcal{A} = -\mathcal{A}^*$. Т. к. $\dim W^\perp = n - 1$, в пространстве W^\perp имеется ортонормированный базис e_1, \dots, e_{n-1} из собственных векторов оператора $\mathcal{A}|_{W^\perp}$. Тогда $e_1, \dots, e_{n-1}, \frac{v}{|v|}$ — ортонормированный базис из собственных векторов оператора \mathcal{A} .

Пусть D — диагональная матрица косоэрмитова оператора \mathcal{A} в ортонормированном базисе из собственных векторов. Т. к. $\mathcal{A}^* = -\mathcal{A}$, получаем $\bar{D}^t = -D$. Следовательно, диагональные элементы матрицы D (собственные числа) удовлетворяют соотношению $\bar{\lambda} = -\lambda$, т. е. являются чисто мнимыми. ■

Теорема 2. Для кососимметрического оператора \mathcal{A} существует ортонормированный базис, в котором его матрица блочно-диагональна с блоками размера 1 или 2, причём блоки размера 1 нулевые, а блоки размера 2 имеют вид $\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}$ с ненулевыми $a \in \mathbb{R}$.

Доказательство. В пространстве размерности 1 или 2 доказывать нечего, т. к. кососимметрическая матрица и так имеет там требуемый вид. Предположим, что утверждение доказано для операторов в пространствах размерности не больше $n - 1$, и докажем его для пространства V размерности n (где $n \geq 3$).

В силу теоремы 1 в вопросе 19 для оператора \mathcal{A} существует одномерное или двумерное инвариантное подпространство $W \subset V$. Как и в случае косоэрмитовых операторов, из важной леммы следует, что ортогональное дополнение W^\perp также инвариантно.

По предположению индукции, в пространстве W^\perp имеется требуемый базис для оператора $\mathcal{A}|_{W^\perp}$. Выбрав произвольный ортонормированный базис в W и взяв объединение базисов W^\perp и W , мы получим ортонормированный базис пространства V , в котором матрица оператора \mathcal{A} состоит

из блоков требуемого вида и ещё одного блока размера 1 или 2 — матрицы оператора $\mathcal{A}|_W$. Этот последний блок — кососимметрическая матрица размера 1 или 2, т. е. она тоже имеет требуемый вид. ■

Теорема 3 (Эрмитово разложение). Для любого оператора \mathcal{A} в эрмитовом пространстве существует единственное представление в виде

$$\mathcal{A} = \mathcal{R} + i\mathcal{I},$$

где \mathcal{R} и \mathcal{I} — эрмитовы операторы.

Доказательство. Сначала докажем единственность. Если $\mathcal{A} = \mathcal{R} + i\mathcal{I}$ — эрмитово разложение, то $\mathcal{A}^* = \mathcal{R}^* - i\mathcal{I}^* = \mathcal{R} - i\mathcal{I}$. Из этих двух соотношений получаем

$$\mathcal{R} = \frac{1}{2}(\mathcal{A}^* + \mathcal{A}), \quad \mathcal{I} = \frac{i}{2}(\mathcal{A}^* - \mathcal{A}),$$

т. е. операторы \mathcal{R} и \mathcal{I} определены однозначно и эрмитово разложение единственно.

С другой стороны, операторы \mathcal{R} и \mathcal{I} , задаваемыми предыдущими формулами, очевидно, эрмитовы (самосопряжены), так что эрмитово разложение существует. ■

В одномерном эрмитовом пространстве \mathbb{C} эрмитовы операторы — это вещественные числа, а операторы \mathcal{R} и \mathcal{I} в эрмитовом разложении — это вещественная и мнимая части комплексного числа.

65.5. ВЫЧИСЛЕНИЕ МНОГОЧЛЕНОВ И ФУНКЦИЙ ОТ МАТРИЦ. ЭКСПОНЕНТА ЛИНЕЙНОГО ОПЕРАТОРА

65.6. БАРИЦЕНТРИЧЕСКИЕ КООРДИНАТЫ

Пусть p_0, p_1, \dots, p_n — аффинно независимые точки n -мерного аффинного пространства (\mathfrak{A}, V) . Тогда каждая точка $p \in \mathfrak{A}$ единственным образом представляется в виде

$$p = \sum_{i=0}^n x^i p_i, \quad \text{где} \quad \sum_{i=0}^n x^i = 1.$$

В самом деле, это равенство можно переписать в виде

$$\overline{p_0 p} = \sum_{i=1}^n x^i \overline{p_0 p_i},$$

откуда следует, что в качестве x^1, \dots, x^n можно (и должно) взять координаты вектора $\overline{p_0 p}$ в базисе $\overline{p_0 p_1}, \dots, \overline{p_0 p_n}$; после этого определить x^0 равенством $x^0 = 1 - \sum_{i=1}^n x^i$

Определение 2. Числа x_0, x_1, \dots, x_n называются *барицентрическими координатами* точки p относительно p_0, p_1, \dots, p_n .

Предложение 2. Пусть $f : \mathfrak{A} \rightarrow \mathfrak{A}'$ — аффинное отображение. Тогда

$$f\left(\sum_i \lambda_i p_i\right) = \sum_i \lambda_i f(p_i)$$

для любой барицентрической комбинации $\sum_i \lambda_i p_i$ точек p_1, \dots, p_k .

Доказательство. Векторизуем пространство \mathfrak{A} . Тогда получим

$$f\left(\sum_i \lambda_i p_i\right) = \varphi\left(\sum_i \lambda_i p_i\right) + b = \sum_i \lambda_i (\varphi(p_i) + b) = \sum_i \lambda_i f(p_i).$$

■

В частности, центр тяжести системы точек при аффинном отображении переходит в центр тяжести их образов.

Предложение 3. Барицентрические координаты суть аффинно-линейные функции.

Доказательство. Пусть x^0, x^1, \dots, x^n — барицентрические координаты относительно аффинно-независимых точек p_0, p_1, \dots, p_n . Если векторизовать пространство \mathfrak{A} , приняв точку p_0 за начало отсчёта, то x^1, \dots, x^n будут обычными координатами относительно базиса $\overline{p_0 p_1}, \dots, \overline{p_0 p_n}$. Следовательно, x^1, x^2, \dots, x^n — аффинно-линейные функции. А т.к. $x^0 = 1 - \sum_{i=1}^n x^i$, то x^0 — также аффинно-линейная функция. ■

65.7. ГЕОМЕТРИЯ АФФИННЫХ ЕВКЛИДОВЫХ ПРОСТРАНСТВ. МНОГОГРАННИКИ И ПРИЛОЖЕНИЯ

66. ТЕОРЕТИЧЕСКИЕ ЗАДАЧИ

Здесь представлены теоретические задачи Тараса Евгеньевича Панова, к некоторым также написаны решения. Рекомендую проверять их на адекватность.

Задача 1. Докажите, что пространство, двойственное к пространству $\mathbb{R}[t]$ всех многочленов от одной переменной над \mathbb{R} , не изоморфно пространству $\mathbb{R}[t]$.

▷ В пространстве $\mathbb{R}[t]$ есть счётный базис: $1, x^1, x^2, x^3, \dots$. Докажем, что в пространстве $(\mathbb{R}[t])^*$ счётного базиса нет. Для этого найдём в нём несчётную линейно независимую систему. Положим $\xi_x : \mathbb{R}[t] \rightarrow \mathbb{F}$, $\xi_x(p) = p(x)$ и рассмотрим систему векторов $\{\xi_x(p) : x \in \mathbb{R}\}$; докажем, что она линейно независима:

$$\sum_{i \in I} \lambda_i \xi_{x_i} = 0.$$

Подставим поочерёдно многочлены p_j , корнями которого являются все x_i , кроме j -го (можно, т.к. почти все λ_i нулевые). Итак для каждого j получим $\lambda_j = 0$. Значит, наша система линейно независима, и в $(\mathbb{R}[t])^*$ нет счётного базиса. ◀

Задача 2. $(V_{\mathbb{R}})_{\mathbb{C}}$ канонически изоморфно $V \oplus \bar{V}$, где \bar{V} — комплексно сопряжённое пространство, в котором сложение то же, что и в V , а умножение $*$ на скаляры определяется как $\lambda * v := \bar{\lambda}v$.

▷ Так как \mathbb{C} — алгебраически замкнутое поле, комплексная структура \mathcal{J} должна иметь $n = \dim(V_{\mathbb{R}})_{\mathbb{C}}$ собственных значений, каждое из которых удовлетворяет $\lambda^2 = -1$, т.е. $\lambda = \pm 1$. Значит, можем записать $(V_{\mathbb{R}})_{\mathbb{C}} = V^+ \oplus V^-$, где V^+ и V^- — собственные подпространства для $+i$ и $-i$ соответственно. Первое из них изоморфно V , а второе — \bar{V} . ◀

Задача 3. Найдите все инвариантные подпространства оператора, матрица которого есть жорданова клетка J_{λ} .

▷ Для удобства обозначим $e_0 := \mathbf{0}$, а через \mathcal{A} — данный оператор, который в базисе e_1, \dots, e_n имеет матрицу J_{λ} . Тогда $\forall i = 1, \dots, n$ имеем

$$J_{\lambda} \cdot e_i = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \lambda & \ddots \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix} \cdot e_i = \lambda e_i + e_{i-1}.$$

Отсюда следует, что подпространства $\langle e_0, e_1, \dots, e_i \rangle \ \forall i = 0, 1, \dots, n$ являются инвариантными для оператора с матрицей J_{λ} .

Докажем, что других инвариантных подпространств нет. Пусть V_k — инвариантное подпространство размерности $k \in \{1, \dots, n-1\}$. Рассмотрим ограничение оператора \mathcal{A} на это инвариантное подпространство и выберем в нём жорданов базис по алгоритму Антона Александровича, чтобы в нём матрица имела вид

$$\underbrace{\begin{pmatrix} \lambda & & & \\ 1 & \lambda & & \\ & 1 & \lambda & \\ & & \ddots & \ddots \\ & & & 1 & \lambda \end{pmatrix}}_{k \times k}.$$

Как я уже писал, то, что единицы стоят снизу, а не сверху, имеет значение, и вот какое. Если единицы стоят сверху, то собственным вектором является e_1 . А если снизу, то e_n . В алгоритме Антона Александровича мы из более высоких векторов получаем более низкие, поэтому хотим,

чтобы первым стоял вектор максимальной высоты, а собственный — последним. Поэтому мы и единицы ставим внизу. А до этого я их ставил вверх, чтобы было поэстетичнее. Правда, теперь у нас будет «перевернутая» нумерация, но в целом такие вещи полезны для тренировки понимания того, что происходит.

Итак, в качестве e'_1 берём случайный вектор, например, e_k . Конечно, мы взяли не совсем случайный вектор, а заведомо хороший, но будем считать, что нам просто повезло. Следующий вектор получаем следующим образом:

$$(J_\lambda - \lambda E)e'_1 = J_\lambda e_k - \lambda e_k = \cancel{\lambda e_k} + e_{k-1} - \cancel{\lambda e_k} = e_{k-1}.$$

И дальше так можем продолжать, пока не дойдём до вектора $e'_k = e_1$. Таким образом, $V_k = \langle e'_1, \dots, e'_k \rangle = \langle e_k, \dots, e_1 \rangle$. ◀

Задача 4. Докажите, что для любого оператора \mathcal{A} существует оператор \mathcal{B} такой, что $\mathcal{A}\mathcal{B}\mathcal{A} = \mathcal{A}$.

▷ Выберем базис e_1, \dots, e_k в $\text{Im } \mathcal{A}$ и дополним его до базиса $e_1, \dots, e_k, e_{k+1}, \dots, e_n$ пространства V . Существуют векторы f_1, \dots, f_k такие, что $\mathcal{A}f_i = e_i \ \forall i = 1, \dots, k$. Определим оператор \mathcal{B} на базисных векторах e_1, \dots, e_k так, чтобы $\mathcal{B}e_i = f_i \ \forall i = 1, \dots, k$, а на остальных базисных векторах можно его доопределить произвольным образом. Получаем

$$\mathcal{A}(\mathcal{B}(\mathcal{A}v)) = \mathcal{A}(\mathcal{B}\lambda^i e_i) = \mathcal{A}(\lambda^i f_i) = \lambda^i e_i = \mathcal{A}v$$

$\forall v \in V$, так что $\mathcal{A}\mathcal{B}\mathcal{A} = \mathcal{A}$. Заметим, что если оператор \mathcal{A} невырожденный, то сконструированный нами оператор \mathcal{B} есть в точности \mathcal{A}^{-1} . ◀

Задача 5. Докажите, что в конечномерном пространстве над полем нулевой характеристики не существует операторов \mathcal{A} и \mathcal{B} , удовлетворяющий соотношению $\mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A} = \text{id}$. Существуют ли такие операторы над полем положительной характеристики?

▷ В конечномерных пространствах над полем нулевой характеристики такого не бывает. Действительно, взяв след от левой части равенства, получим $\text{tr}(\mathcal{B}\mathcal{A} - \mathcal{A}\mathcal{B}) = \text{tr } \mathcal{B} \text{tr } \mathcal{A} - \text{tr } \mathcal{A} \text{tr } \mathcal{B} = 0$. С другой, стороны, берём след от правой части равенства, получаем $\text{tr id} = n$.

Ясно, что равенство $n = 0$ может выполняться при $\text{char } \mathbb{F} = \dim V = n$. Чтобы сильно не задумываться, возьмём поле \mathbb{Z}_2 и небольшим перебором найдём матрицы операторов, удовлетворяющие требуемому свойству:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Так что над полями положительной характеристики такое бывает, а нулевой — не бывает. ◀

Задача 6. Постройте пример операторов в бесконечномерном пространстве над \mathbb{R} , удовлетворяющий соотношению $\mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A} = \text{id}$.

▷ Я не особо знаю бесконечномерных пространств над \mathbb{R} , кроме $\mathbb{R}[t]$, так что искомый пример лежит где-то там. ◀

Задача 7. Пусть \mathcal{A}, \mathcal{B} — операторы. Докажите, что характеристические многочлены операторов $\mathcal{A}\mathcal{B}$ и $\mathcal{B}\mathcal{A}$ совпадают.

▷ Докажем более сильное утверждение: если $A \in \text{Mat}_{m \times n}(\mathbb{F})$ и $B \in \text{Mat}_{n \times m}(\mathbb{F})$, то характеристические многочлены матриц AB и BA удовлетворяют следующему условию:

$$t^n \det(t \cdot \begin{matrix} E \\ m \times m \end{matrix} - AB) = t^m \det(t \cdot \begin{matrix} E \\ n \times n \end{matrix} - BA).$$

Определим

$$C := \begin{pmatrix} t \cdot \begin{smallmatrix} E \\ m \times m \end{smallmatrix} & A \\ B & \begin{smallmatrix} E \\ n \times n \end{smallmatrix} \end{pmatrix}, \quad D := \begin{pmatrix} \begin{smallmatrix} E \\ m \times m \end{smallmatrix} & 0 \\ -B & t \cdot \begin{smallmatrix} E \\ n \times n \end{smallmatrix} \end{pmatrix}.$$

Тогда

$$\det CD = t^n \det(t \cdot \begin{smallmatrix} E \\ m \times m \end{smallmatrix} - AB),$$

$$\det DC = t^m \det(t \cdot \begin{smallmatrix} E \\ n \times n \end{smallmatrix} - BA).$$

С учётом $\det CD = \det DC$ получаем требуемое. ◀

Задача 8. Докажите, что оператор $\text{id} + \mathcal{AB}$ обратим тогда и только тогда, когда $\text{id} + \mathcal{BA}$ обратим.

▷ ◀