

Лекции по алгебре

Н.Ю.Золотых

8 января 2001 г.

Фраза на странице 63 имеет copyright © S.I.Sorochan, 1998. Размещено с разрешения автора.

Оглавление

1. Счетные и континуальные множества	7
1.1. Множества	7
1.2. Отображения	7
1.3. Свойства счетных множеств	9
1.4. Теорема Кантора	9
1.5. Несчетность множества действительных чисел	10
1.6. Декартово произведение	10
2. Комплексные числа	11
2.1. Понятие комплексного числа	11
2.2. Тригонометрическая форма комплексного числа	13
2.3. Комплексно сопряженные числа	15
2.4. Неравенство треугольника	16
2.5. Корни из комплексных чисел	17
2.6. Корни из единицы	18
2.7. Квадратные корни из комплексных чисел	21
2.8. Уравнения второй, третьей и четвертой степени	22
2.9. Вычисление сумм и произведений	26
2.10. Числовые кольца и поля	30
3. Линейные пространства	31
3.1. Аксиоматическое определение линейного пространства	31
3.2. Примеры линейных пространств	31
3.3. Простейшие следствия из аксиом	32
3.4. Линейные подпространства	32
3.5. Линейные комбинации и линейные оболочки	33
3.6. Линейная зависимость векторов	34
3.7. Базис линейного пространства	36
3.7.1. Размерность и базис арифметического пространства	37
3.7.2. Размерность и базис пространства геометрических векторов	37
3.7.3. Размерность пространства многочленов	37
3.8. Координаты векторов	37
3.9. Изоморфизм линейных пространств	38
3.10. Размерность подпространства	39
3.11. Сумма и пересечение подпространств	40
3.12. Прямая сумма подпространств	41
3.12.1. Проекция вектора на подпространство	42
3.13. Изменение координат при замене базиса	42

4. Матрицы и системы линейных уравнений	43
4.1. Метод Гаусса решения систем линейных уравнений	43
4.2. Ранг матрицы	46
4.3. Пространство решений системы линейных однородных уравнений	47
4.4. Линейное многообразие	48
4.5. Множество решений системы линейных уравнений общего вида	48
4.6. Описание подпространства и линейного многообразия в виде системы линейных уравнений	49
5. Определители	50
5.1. Перестановки и подстановки	50
5.2. Определитель. Комбинаторное определение	51
5.3. Свойства определителя	51
5.4. Миноры. Теорема Лапласа	51
5.5. Полилинейные знакопеременные функции	51
5.6. Минорный ранг	52
5.7. Сумма определителей	53
5.8. Теорема Бинэ–Коши	53
6. Линейные отображения и преобразования	55
6.1. Определения и примеры	55
6.2. Матрица линейного оператора	56
6.3. Операции с линейными отображениями	57
6.4. Изменение матрицы оператора при замене базисов	58
6.5. Ядро и образ оператора	59
6.6. Линейные преобразования	59
6.7. Собственные числа и собственные векторы преобразования	60
6.7.1. Выражение коэффициентов характеристического многочлена через главные миноры матрицы	62
6.7.2. Матрица Фробениуса	62
6.8. Диагонализуемость линейного преобразования	63
6.9. Аннулирующий многочлен	65
6.9.1. Метод Крылова построения минимального многочлена	66
6.10. Жорданова форма линейного преобразования	68
6.10.1. Определения	68
6.10.2. Цель	68
6.10.3. Разложение пространства в прямую сумму корневых подпространств	68
6.10.4. Построение Жорданова базиса	70
6.10.5. Единственность Жордановой формы	72
6.11. Матричные функции	73
6.11.1. Значение многочлена от матрицы	73
6.11.2. Интерполяционный многочлен	74
6.11.3. Функции от матриц	74
7. Билинейные функции	77
7.1. Определения	77
7.2. Матрица билинейной функции	78
7.3. Эрмитовы функции	80
7.4. Знакоопределенные эрмитовы функции	83
8. Евклидовы пространства	85
8.1. Определения	85
8.2. Матрица Грама	86
8.3. Ортогональность	87
8.4. Ортогональный и ортонормированный базисы	88
8.5. Унитарный изоморфизм	88
8.6. Ортогональные суммы и ортогональные дополнения	89

8.7. Метод нахождения проекции и перпендикуляра	90
8.8. Процесс ортогонализации Грама-Шмидта	92
8.9. Объем параллелепипеда	92
8.10. Неравенство Адамара	93
8.11. Метрические задачи в унитарных пространствах	94
8.12. Нормальное решение системы линейных уравнений	95
8.13. Псевдорешение несовместной системы линейных уравнений	95
9. Линейные преобразования унитарных и евклидовых пространств	96
9.1. Сопряженные преобразования	96
9.2. Теорема Ш'ура	97
9.3. Нормальные преобразования	98
9.4. Унитарные и ортогональные преобразования	100
9.5. Эрмитовы (самосопряженные) и симметрические преобразования	102
9.6. Косоэрмитово преобразование	103
9.7. Разложения преобразований	103

Список иллюстраций

2.1.	Сложение и вычитание комплексных чисел	13
2.2.	Модуль и аргумент комплексного числа	14
2.3.	Комплексно сопряженные числа	16
2.4.	Неравенство треугольника	17
2.5.	Значения $\sqrt[n]{\zeta}$ ($n = 5$)	18
2.6.	Значения $\sqrt[6]{1}$	19

Глава 1

Счетные и континуальные множества

1.1. Множества

Множество — неопределяемое понятие, поэтому оперирование с ним, как и всегда в математике с неопределяемыми понятиями, регулируется рядом аксиом. Мы не будем на них останавливаться: это увело бы нас слишком далеко в теорию множеств, а примем так называемую “наивную” точку зрения на понятие множества. Просто уясним для себя, что под множеством мы будем понимать произвольную совокупность объектов.

Конечные множества можно задать перечислив все его элементы, например: $A = \{a_1, a_2, \dots, a_n\}$. Под мощностью конечного множества мы будем понимать число элементов в этом множестве. Мощность множества A обозначается $|A|$. Для нашего примера $|A| = n$.

Со школьного курса математики вам известны некоторые *бесконечные* числовые множества: натуральных чисел $\mathbf{N}$, целых чисел $\mathbf{Z}$, рациональных чисел $\mathbf{Q}$, действительных чисел $\mathbf{R}$. Если перенести определение мощности как числа элементов также и на бесконечные множества, то все бесконечные множества окажутся равномощными: все будут иметь “мощность бесконечность”. Ниже мы познакомимся с более продуктивным способом определить мощность для бесконечных множеств.

Говорят, что множество B *включено* в множество A и записывают $B \subseteq A$, если любой элемент из B принадлежит также A . В этом случае B называют *подмножеством* множества A . В курсе дискретной математики, вы узнаете, что всего в любом конечном множестве A содержится $2^{|A|}$ его подмножеств (разумеется, считается также пустое множество $\emptyset$ и все множество A). Совокупность всех подмножеств произвольного множества A обычно обозначается 2^A , поэтому справедлива легко запоминаемая формула:

$$|2^A| = 2^{|A|}. \quad (1.1)$$

Пусть $A = \{1, 2, 3\}$, тогда $2^A = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$. Легко видеть, что $|A| = 3$, $|2^A| = 2^3 = 8$.

1.2. Отображения

Правило, или закон, по которому каждому элементу x множества A ставится в соответствие единственный элемент y множества B называется *отображением*. Элемент y называется *образом* элемента x , а x — *прообразом* элемента y . Тот факт, что отображение φ действует из множества A в множество B , обозначают $\varphi : A \rightarrow B$. Для отображения φ образ элемента x обозначают φx или $\varphi(x)$.

Если A и B — некоторые подмножества множества $\mathbf{R}$, то отображение называется (числовой) *функцией*. Отображение $\mathbf{N} \rightarrow \mathbf{R}$ называется *последовательностью*.

Отображение φ называется *инъективным*, или просто *инъекцией*, если образы любых двух различных друг от друга элементов x и y из A различны, т.е. $\forall x, y \ x \neq y, x \in A, y \in A \ \varphi x \neq \varphi y$.

Отображение φ называется *сюръективным*, или просто *сюръекцией*, если для произвольного элемента $h \in B$ найдется по крайней мере один прообраз.

Отображение φ называется *биективным* (*биекцией*), или *взаимно однозначным*, если оно одновременно и инъективно, и сюръективно. Как легко видеть, это условие эквивалентно тому что, любой элемент из B имеет ровно один прообраз.

Упражнение 1. Докажите, что для существования инъекции $A \rightarrow B$ необходимо и достаточно существование сюръекции $B \rightarrow A$.

Заметим, что если существуют инъекция $A \rightarrow B$ и сюръекция $A \rightarrow B$, то существует и биекция $A \rightarrow B$. Это интуитивно ясное утверждение, между тем, имеет достаточно сложное доказательство и мы на нем не останавливаемся.

Пусть множества A и B конечны. Легко видеть, что тогда

- для существования инъекции $A \rightarrow B$ необходимо и достаточно, чтобы $|A| \leq |B|$;
- для существования сюръекции $A \rightarrow B$ необходимо и достаточно, чтобы $|A| \geq |B|$;
- для существования биекции $A \rightarrow B$ необходимо и достаточно, чтобы $|A| = |B|$.

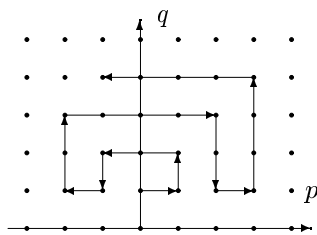
Эти утверждения позволяют теперь дать следующие определения. Для произвольных множеств A и B (в том числе бесконечных) будем говорить, что мощность A меньше или равна мощности B , и записывать $|A| \leq |B|$, если существует инъекция $A \rightarrow B$. Будем говорить, что мощность A больше или равна мощности B , и записывать $|A| \geq |B|$, если существует сюръекция $A \rightarrow B$. И наконец, будем говорить, что мощности множеств A и B равны (множества A и B равномощны), и записывать $|A| = |B|$, если существует биекция $A \rightarrow B$.

Докажем теперь, что $|\mathbf{Z}| = |2\mathbf{Z}|$, где $2\mathbf{Z}$ — множество четных целых чисел. Действительно, биекция $\varphi: \mathbf{Z} \rightarrow 2\mathbf{Z}$ задается, например, формулой $\varphi(x) = 2x$. Легко видеть, что любое четное число имеет при этом отображении единственный прообраз. Также нетрудно установить равномощность множеств $\mathbf{Z}$ и $\mathbf{N}$:

$$\begin{array}{ccccccc} \dots & -2 & -1 & 0 & 1 & 2 & \dots \\ & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\ \dots & 5 & 3 & 1 & 2 & 4 & \dots \end{array}$$

В первой строке приведен фрагмент ряда целых чисел. Каждому целому числу x поставлено в соответствие натуральное $\psi(x)$, написанное снизу. Общее правило легко угадывается.

Упражнение 2. Привести явную формулу для отображения $\psi(x)$.



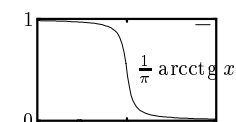
Для построения биективного отображения $\mathbf{Q} \rightarrow \mathbf{N}$ рассмотрим кривую, начальный фрагмент которой изображен на рисунке ??, а дальнейшее поведение легко угадывается. Кривая будет замечать все точки с координатами (p, q) , где $p \in \mathbf{Z}$, $q \in \mathbf{N}$. Каждой такой точке поставим в соответствие рациональное число $\frac{p}{q}$. Это (вспомогательное) отображение будет сюръективным, но не будет инъективным: например, точкам $(2, 1)$ и $(4, 2)$ будет соответствовать одно и то же число $\frac{2}{1} = \frac{4}{2}$. Будем двигаться по кривой по направлению, указанно-

му стрелкой, и каждой встретившейся точке (p, q) присваивать порядковый номер $\varphi(p, q)$: $\varphi(0, 1) = 1$, $\varphi(1, 1) = 2$, $\varphi(1, 2) = 3$, $\varphi(0, 2) = 4$, $\varphi(-1, 1) = 5$ и т.д.

Упражнение 3. Привести явную формулу для $\varphi(p, q)$.

Для построения биекции $\mathbf{Q} \rightarrow \mathbf{N}$ осталось лишь чуть-чуть скорректировать отображение φ : при обходе кривой будем нумеровать лишь точки, для которых $(p, q) = 1$, оставляя без номера все остальные.

Покажем, что множество $\mathbf{R}$ равномощно множеству точек произвольно выбранного на числовой прямой интервала (a, b) : равномощность множеств $\mathbf{R}$ и $(0, 1)$ устанавливается, например, взаимно однозначным отображением $\varphi(x) = \frac{1}{\pi} \operatorname{arctg} x$, а равномощность множеств $(0, 1)$ и (a, b) (а так же $[0, 1]$ и $[a, b]$) — биекцией $\psi(x) = a + (b - a)x$ (см. рис. ??). Функции $\varphi(x)$ и $\psi(x)$ — строго монотонные и, следовательно, как известно из математического анализа, — взаимно однозначные.



Упражнение 4. Построить биекцию $[0, 1] \rightarrow (0, 1)$.

Подведем итог. Мы рассмотрели несколько бесконечных множеств. Оказалось, что $|\mathbf{N}| = |\mathbf{Z}| = |\mathbf{Q}|$ и $|\mathbf{R}| = |(a, b)|$, однако пока не решен вопрос “ $|\mathbf{N}| \stackrel{?}{=} |\mathbf{R}|$ ”. Так как $\mathbf{N} \subseteq \mathbf{R}$, то, конечно, $|\mathbf{N}| \leq |\mathbf{R}|$. Скоро мы увидим, что биекции $\mathbf{N} \rightarrow \mathbf{R}$ не существует и, таким образом, $|\mathbf{N}| < |\mathbf{R}|$. Множества, равномощные множеству натуральных чисел называются *счетными*, или множествами мощности $\aleph_0$ (читается ‘алеф-0’). Множества, равномощные множеству действительных чисел называются *континуальными* (мощности континуума), или множествами мощности $\aleph_1$ (читается ‘алеф-1’).

1.3. Свойства счетных множеств

Теорема 1.3.1. Любое бесконечное множество A содержит счетное подмножество B .

ДОКАЗАТЕЛЬСТВО. Действительно, выберем из множества A произвольный элемент a_1 , так как A бесконечно, то мы можем в A выбрать другой элемент $a_2 \neq a_1$, далее выбираем в A элемент a_3 , отличный от a_1 и a_2 и т.д. Пусть $B = \{a_1, a_2, a_3, \dots\}$. Очевидно, B — счетно. ■

Теорема 1.3.2. Всякое бесконечное подмножество B счетного множества A само счетно.

ДОКАЗАТЕЛЬСТВО. Так как A счетно, то мы можем считать, что $A = \{a_1, a_2, a_3, \dots\}$. Найдем наименьший индекс i_1 (номер в ряду элементов), такой, что $a_{i_1} \in B$, далее выберем наименьший индекс i_2 , отличный от i_1 , для которого $a_{i_2} \in B$ и т.д. Таким образом, $B = \{a_{i_1}, a_{i_2}, a_{i_3}, \dots\}$. Очевидно, B — счетно. ■

Теорема 1.3.3. Объединение конечного или счетного множества счетных множеств — счетно.

ДОКАЗАТЕЛЬСТВО. Докажем утверждение для случая счетного множества счетных множеств. Пусть

$$\begin{aligned} A_1 &= \{a_{11}, a_{12}, a_{13}, \dots\} \\ A_2 &= \{a_{21}, a_{22}, a_{23}, \dots\} \\ A_3 &= \{a_{31}, a_{32}, a_{33}, \dots\} \\ &\vdots \end{aligned}$$

— семейство счетных множеств. Для установления биекции $A_1 \cup A_2 \cup \dots \rightarrow \mathbb{N}$ необходимо предложить некоторое правило обхода элементов этих множеств, например:

$$\begin{array}{ccccccc} a_{11} & \rightarrow & a_{12} & & a_{13} & \rightarrow & a_{14} \dots \\ & & \downarrow & & \uparrow & & \downarrow \\ a_{21} & \leftarrow & a_{22} & & a_{23} & & a_{24} \dots \\ \downarrow & & & & \uparrow & & \downarrow \\ a_{31} & \rightarrow & a_{32} & \rightarrow & a_{33} & & a_{34} \dots \\ & & & & & & \downarrow \\ a_{31} & \leftarrow & a_{32} & \leftarrow & a_{33} & \leftarrow & a_{44} \dots \\ \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

Следуя по стрелке, мы присваиваем элементам множеств A_i порядковые номера и тем самым строим отображение в $\mathbb{N}$. Элементы, встретившиеся неоднократно, конечно, пропускаются (это возможно, так как множества A_i не обязаны не пересекаться). ■

1.4. Теорема Кантора

Для произвольного конечного множества A из (1.1) следует неравенство $|A| < |2^A|$. Оказывается, это неравенство справедливо и для произвольного бесконечного множества.

Теорема 1.4.1 (Кантор). Для произвольного множества A

$$|A| < |2^A|.$$

ДОКАЗАТЕЛЬСТВО. Неравенство $|A| \leq |2^A|$ устанавливается сюръекцией $\varphi x = \{x\}$.

Чтобы доказать, что $|A| \neq |2^A|$, покажем, что допущение о существовании биекции $\psi : A \rightarrow 2^A$ приводит к противоречию. Действительно, пусть ψ — биекция из A в 2^A . Рассмотрим множество

$$I = \{x \in A \mid x \notin \psi x\} \quad (1.2)$$

(множество I содержит те и только те элементы, которые не содержатся в своем образе). Так как ψ — биективное отображение, то для любого элемента из 2^A найдется прообраз в A . Пусть для $I \in 2^A$ таким прообразом будет некоторый элемент i : $\psi i = I$. Теперь попробуем решить вопрос ‘ $i \in I$ ’.

1) Если $i \in I$, тогда по (1.2) $i \notin \psi i$. Однако $\psi i = I$, т.е. $i \in I$.

2) Пусть $i \notin I$, тогда по (1.2) получаем $i \in \psi i$, т.е. $i \in I$.

Оба раза мы приходим к противоречию, следовательно, не верно наше предположение о существовании биекции ψ . ■

1.5. Несчетность множества действительных чисел

Основная задача данного раздела — доказать, что $|\mathbf{R}| > |\mathbf{N}|$.

Теорема 1.5.1 (Несчетность множества $\mathbf{R}$). $|2^{\mathbf{N}}| = |\mathbf{R}|$.

Доказательство. Сперва установим равномощность множеств $2^{\mathbf{N}}$ и $[0, 1]$. Пусть¹ для любого $A \in 2^{\mathbf{N}}$

$$\varphi(A) = (0, \alpha_1 \alpha_2 \alpha_3 \dots)_2, \text{ где } \alpha_i = \begin{cases} 0, & \text{если } i \in A, \\ 1, & \text{если } i \notin A \end{cases}.$$

В частности, $\varphi(\emptyset) = 0$, $\varphi(\mathbf{N}) = 1$. Очевидно, что φ — сюръекция, однако, следующие примеры:

$$\begin{aligned} \varphi(\{1, 3\}) &= (0, 101(0))_2 = \frac{1}{2} + \frac{1}{8} = \frac{5}{8}, \\ \varphi(\{1, 4, 5, 6, \dots\}) &= (0, 100(1))_2 = \frac{1}{2} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \dots = \frac{1}{2} + \frac{1}{8} = \frac{5}{8} \end{aligned}$$

— показывают, что φ не является инъекцией: двум разным подмножествам множества натуральных чисел соответствует одно и то же число, или, что эквивалентно, существует $\alpha \in [0, 1]$, обладающее двумя прообразами. Это возможно, если в двоичной системе счисления α представимо как бесконечная дробь с периодом (1) или (0). Легко видеть, что множество N таких чисел счетно: оно есть объединение счетного числа конечных множеств, состоящих из чисел, период (1) которых начинается с первого, второго и т.д. места после запятой. Счетным является также множество M прообразов всех чисел из N ($M = \{x \in 2^{\mathbf{N}} : \varphi(x) \in N\}$). Пусть ψ — некоторая биекция из M в N . Скорректируем теперь отображение φ . Для любого $A \in 2^{\mathbf{N}}$ положим

$$\theta(A) = \begin{cases} \psi(A), & \text{если } A \in M, \\ \varphi(A), & \text{если } A \notin M. \end{cases}$$

Легко видеть, что θ — биекция из $2^{\mathbf{N}}$ в $[0, 1]$. Так как $|[0, 1]| = |\mathbf{R}|$, то $|2^{\mathbf{N}}| = |\mathbf{R}|$. ■

Теперь из теоремы Кантора получаем

Утверждение 1.5.1. $|\mathbf{R}| > |\mathbf{N}|$.

1.6. Декартово произведение

Декартовым произведением $A_1 \times A_2 \times \dots \times A_n$ множеств $A_1, A_2, \dots, A_n$ называется множество всех упорядоченных наборов $(a_1, a_2, \dots, a_n)$, таких, что $a_i \in A_i$ ($i = 1, 2, \dots, n$). Декартовой n -ой степенью A^n множества A называется множество всех упорядоченных наборов $(a_1, a_2, \dots, a_n)$, таких, что $a_i \in A$ ($i = 1, 2, \dots, n$).

Упражнение 5. Доказать, что $|\mathbf{Z}^n| = |\mathbf{N}|$, $|\mathbf{R}^n| = |\mathbf{R}|$ для любого натурального n .

¹Через

$$(0, \alpha_1 \alpha_2 \alpha_3 \dots)_2 = \frac{\alpha_1}{2} + \frac{\alpha_2}{4} + \frac{\alpha_3}{8} + \dots + \frac{\alpha_i}{2^i} + \dots$$

мы обозначили представление действительного числа в двоичной системе счисления.

Глава 2

Комплексные числа

В 1545 г. итальянский математик Дж. Кардано (G. Cardano) в книге “Великое искусство, или об алгебраических правилах” для системы уравнений

$$\begin{cases} x + y = 10, \\ xy = 40. \end{cases}$$

привел два корня $5 + \sqrt{-15}$, $5 - \sqrt{-15}$. Никакой практической ценности в таком решении Кардано не увидел. Скорее, этот пример был приведен как некий курьез. Интересно отметить, что в том же сочинении была впервые опубликована найденная С. Ферро и Н. Тарталья (S. Ferro, N. Tartaglia) формула для решения кубического уравнения

$$x^3 + px + q = 0. \quad (2.1)$$

Его корни предлагалось находить по формуле

$$x = \sqrt[3]{u} + \sqrt[3]{v}, \quad (2.2)$$

где u, v есть решения системы

$$\begin{cases} u + v = -q, \\ uv = -\frac{p^3}{27}. \end{cases} \quad (2.3)$$

В некоторых случаях, несмотря на то, что уравнение (2.1) заведомо имело корни, система (2.3) действительных решений не имела и формула (2.2) считалась бесполезной. Другой итальянец Р. Бомбелли (R. Bombelli) в своем труде “Алгебра” (1572) показал, что действительные корни уравнения (2.1) в таких случаях выражаются через радикалы от “мнимых” величин, он разработал простейшие правила действия с ними и подошел, таким образом, к созданию теории комплексных чисел.

2.1. Понятие комплексного числа

Комплексным числом называется упорядоченная¹ пара действительных чисел (a, b) . Два комплексных числа (a, b) и (c, d) *равны* тогда и только тогда, когда $a = c$, $b = d$. На множестве всех комплексных чисел $\mathbf{C}$ определены операции сложения и умножения по правилам:

$$(a, b) + (c, d) = (a + c, b + d), \quad (a, b) \cdot (c, d) = (ac - bd, ad + bc). \quad (2.4)$$

Легко проверить, что эти операции обладают следующими свойствами:

- ассоциативность: $(a, b) + [(c, d) + (e, f)] = [(a, b) + (c, d)] + (e, f)$, $(a, b) [(c, d)(e, f)] = [(a, b)(c, d)] (e, f)$;
- коммутативность: $(a, b) + (c, d) = (c, d) + (a, b)$, $(a, b)(c, d) = (c, d)(a, b)$;
- дистрибутивность: $(a, b) [(c, d) + (e, f)] = (a, b)(c, d) + (a, b)(e, f)$.

¹ Упорядоченная в том смысле, что задано, какое число в паре — первое, какое — второе.

Докажем, например, дистрибутивность. Пользуясь правилами (2.4) сложения и умножения комплексных чисел, а также законами (ассоциативность, коммутативность, дистрибутивность), справедливыми для вещественных чисел, получаем: $(a, b) [(c, d) + (e, f)] = (a, b)(c+e, d+f) = (a(c+e) - b(d+f), a(d+f) + (ac + ae - bd - bf, ad + af + bc + be))$; с другой стороны, $(a, b)(c, d) + (a, b)(e, f) = (ac - bd, ad + bc) + (ae - bf, af + be) = (ac - bd + ae - bf, ad + bc + af + be)$; левая и правая части совпали.

Упражнение 6. Доказать остальные свойства операций сложения и умножения комплексных чисел.

Нулем называется такое комплексное число (x, y) , что для произвольного числа (a, b) выполняется равенство $(a, b) + (x, y) = (a, b)$. Из определения получаем $a + x = a$, $b + y = b$, откуда $x = 0$, $y = 0$. Следовательно, нулем является пара $(0, 0)$ (и только она). Числом *противоположным* к (a, b) называется такая пара (x, y) , что $(a, b) + (x, y) = (0, 0)$. Противоположное число обозначается $-(a, b)$. Нетрудно видеть, что $-(a, b) = (-a, -b)$. *Разностью* (или числом, полученным в результате *вычитания*) комплексных чисел (c, d) и (a, b) называется решение (x, y) уравнения $(a, b) + (x, y) = (c, d)$. Разность обозначается $(c, d) - (a, b)$ и, очевидно, равна $(c - a, d - b)$. Легко видеть, что разность $(c, d) - (a, b)$ есть сумма (c, d) и числа, противоположного к (a, b) , т.е. $(c, d) - (a, b) = (c, d) + [-(a, b)]$.

Единицей называется такое комплексное число (x, y) , что для произвольного числа (a, b) выполняется равенство $(a, b)(x, y) = (a, b)$. Из определения произведения получаем

$$\begin{cases} ax - by = a, \\ bx + ay = b. \end{cases}$$

В случае, если $a^2 + b^2 \neq 0$, т.е. $(a, b) \neq (0, 0)$, имеем единственное решение предыдущей системы: $x = 1$, $y = 0$. Таким образом, $(a, b)(1, 0) = (a, b)$ для любого (a, b) в том числе, как нетрудно проверить, и для $(a, b) = (0, 0)$. Следовательно, $(1, 0)$ — единица. Числом *обратным* к (a, b) называется такая пара (x, y) , что $(a, b)(x, y) = (1, 0)$. Обратное число обозначается $(a, b)^{-1}$. Система

$$\begin{cases} ax - by = 1, \\ bx + ay = 0 \end{cases}$$

в случае, когда $(a, b) \neq (0, 0)$, имеет единственное решение

$$(a, b)^{-1} = \left(\frac{a}{a^2 + b^2}, -\frac{b}{a^2 + b^2} \right), \quad (2.5)$$

а в случае $(a, b) = (0, 0)$ — неразрешима. *Частным* (или числом, полученным в результате *деления*) комплексных чисел (c, d) и (a, b) называется решение (x, y) уравнения $(a, b)(x, y) = (c, d)$. Частное обозначается $(c, d)/(a, b)$. Его мы можем получить из системы

$$\begin{cases} ax - by = c, \\ bx + ay = d. \end{cases} \quad (2.6)$$

Домножая первое уравнение на a , второе — на b и складывая, получаем: $(a^2 + b^2)x = ac + bd$, затем, домножая первое уравнение системы (2.6) на b , второе — на a и вычитая первое из второго, получаем: $(a^2 + b^2)y = ad - cb$. В случае $a^2 + b^2 \neq 0$ имеем:

$$\frac{(c, d)}{(a, b)} = \left(\frac{ac + bd}{a^2 + b^2}, \frac{ad - cb}{a^2 + b^2} \right), \quad (2.7)$$

если же $a^2 + b^2 = 0$, то результат деления, как легко видеть из (2.6), не определен. Сравнивая (2.5) и (2.7), получаем, что частное $(c, d)/(a, b)$ есть произведение (c, d) на величину, обратную к (a, b) : $(c, d)/(a, b) = (c, d) \cdot (a, b)^{-1}$.

Отождествим комплексные числа вида $(a, 0)$ с действительными. А именно, положим $(a, 0) = a$. Из правил сложения и умножения комплексных чисел следует, что $(a, 0) + (b, 0) = (a + b, 0)$, $(a, 0)(b, 0) = (ab, 0)$ для любых a, b . Таким образом, результаты выполнения арифметических операций над числами такого вида не зависят от того, как они были получены: по правилам сложения и умножения комплексных чисел, рассмотренным в данном разделе, или по законам действительных чисел. Обозначим $i = (0, 1)$. Число i называется *мнимой единицей*. Легко проверить, что $(0, b) = bi$ и, следовательно, $(a, b) = a + bi$.

Запись $a + bi$ называется *алгебраической формой комплексного числа* (a, b) . Ее использование освобождает нас от заучивания правил арифметических операций (2.4): легко проверить, что $i^2 =$

-1^2 , поэтому *работать*³ с комплексными числами можно как с алгебраическими двучленами, зависящими от символа i , с заменой, где необходимо, i^2 на -1 . Например, для нахождения произведения $(a + bi)(c + di)$ раскроем скобки и получим $ac + adi + bci + bdi^2$, так как $i^2 = -1$, то $(a + bi)(c + di) = ac - bd + (ad + bc)i$. Для нахождения частного $(c + di)/(a + bi)$ домножим числитель и знаменатель дроби на число $c - di$, называемое *сопряженным* к $c + di$, получим:

$$\frac{c + di}{a + bi} = \frac{(c + di)(a - bi)}{(c + di)(a - bi)} = \frac{ac + bd + i(ad - bc)}{a^2 + b^2} = \frac{ac + bd}{a^2 + b^2} + \frac{ad - bc}{a^2 + b^2}i.$$

Результат совпадает с (2.7).

Пример 1.

- а) $(4 + i)(5 + 3i) + (3 + i)(3 - 2i) = 20 + 12i + 5i - 3 + 9 - 6i + 3i + 2 = 28 + 14i$;
 б)

$$\begin{aligned} \frac{(5 + i)(7 - 6i)}{3 + i} &= \frac{35 - 30i + 7i + 6}{3 + i} = \\ &= \frac{41 - 23i}{3 + i} = \frac{(41 - 23i)(3 - i)}{(3 + i)(3 - i)} = \\ &= \frac{123 - 41i - 69i - 23}{10} = \frac{100 - 110i}{10} = 10 - 11i. \end{aligned}$$

Упражнение 7. Вычислить

- а) $\frac{(5+i)(3+5i)}{2i}$;
 б) $\frac{(2+i)(4+i)}{1+i}$.

Обычно комплексное число обозначают одной буквой: $z = a + bi$, здесь a называется *действительной частью* числа z , b — *мнимой частью* комплексного числа мнимая часть. Используется обозначение: $\operatorname{Re} z = a$, $\operatorname{Im} z = b$.

2.2. Геометрическая интерпретация и тригонометрическая форма записи комплексных чисел

Если на плоскости задана прямоугольная система координат, то произвольному комплексному числу $(a, b) = a + bi$ можно сопоставить точку с координатами (a, b) . Такое отображение, очевидно, является биекцией множества $\mathbf{C}$ на множество всех точек плоскости. Плоскость, точки которой проинтерпретированы таким образом, называется *комплексной плоскостью*. Ось абсцис (называемая в данном случае действительной осью и обозначаемая Re) при такой интерпретации будет соответствовать множеству действительных чисел. Ось ординат (называемая в данном случае мнимой осью и обозначаемая Im) — множеству *чисто мнимых* чисел, т.е. чисел вида bi , где b — любое вещественное. Начало координат соответствует нулю. *Наглядную геометрическую интерпретацию имеют сложение и вычитание комплексных чисел — это просто сложение и вычитание из радиус-векторов по правилу параллелограмма* (рис. 2.1).

Назовем *модулем*, или *абсолютной величиной* Число комплексное, абсолютная величина, комплексного числа $z = a + bi$ расстояние r от точки z до начала координат (рис. 2.2). Модуль числа обозначается $|z|$. Очевидно, определение модуля, известное для вещественных чисел, остается верным и для комплексных. Далее, $|z| = 0$ тогда и только тогда, когда $z = 0$. *Аргументом* не равного нулю числа z назовем угол φ , отсчитываемый в положительном направлении (против часовой стрелки), между направлением оси Re и радиус-вектором точки z . Аргумент числа z обозначается $\arg z$. Аргумент определен неоднозначно с точностью до полного оборота, т.е. с точностью до слагаемого $2\pi k$, где $k \in \mathbf{Z}$. Очевидно, по паре (r, φ) комплексное число z определяется однозначно.

Применяя теорему Пифагора и известные формулы тригонометрии, получаем $|a + bi| = \sqrt{a^2 + b^2}$, $a = r \cos \varphi$, $b = r \sin \varphi$, или же

$$a + bi = r(\cos \varphi + i \sin \varphi).$$

Запись $r(\cos \varphi + i \sin \varphi)$ называется *тригонометрической формой* комплексного числа $a + bi$.

Пример 2. Найти тригонометрическую форму числа:

² таким образом, i — корень многочлена $x^2 + 1 = 0$

³ слово не совсем подходит по контексту

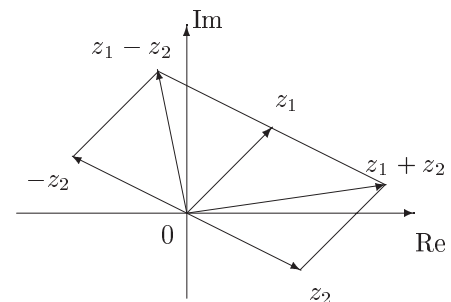


Рис. 2.1. Сложение и вычитание комплексных чисел

- а) $8 = 8(\cos 0 + i \sin 0)$;
 б) $i = \cos \frac{\pi}{2} + i \sin \frac{\pi}{2}$;
 в) $-8 = 8(\cos \pi + i \sin \pi)$;
 д) $1 - i = \sqrt{2} \left(\cos \left(-\frac{\pi}{2} \right) + i \sin \left(-\frac{\pi}{2} \right) \right)$, действительно, $|1 - i| = \sqrt{2}$, одним из решений системы

$$\begin{cases} 1 = \sqrt{2} \cos \varphi, \\ -1 = \sqrt{2} \sin \varphi \end{cases}$$

является $\varphi = -\pi/2$;

- е) $-\sqrt{3} - i = 2 \left(\cos \left(-\frac{5\pi}{6} \right) + i \sin \left(-\frac{5\pi}{6} \right) \right)$;
 ф) $\cos \alpha - i \sin \alpha = \cos(-\alpha) + i \sin(-\alpha)$;
 г) $\sin \alpha + i \cos \alpha = \cos \left(\frac{\pi}{2} - \alpha \right) + i \sin \left(\frac{\pi}{2} - \alpha \right)$;
 х) $1 + \cos \varphi + i \sin \varphi$, если $\varphi \in [-\pi, \pi]$: воспользовавшись формулами двойного угла, получаем: $2 \cos^2 \frac{\varphi}{2} + i 2 \cos \frac{\varphi}{2} \sin \frac{\varphi}{2} = 2 \cos \frac{\varphi}{2} \left(\cos \frac{\varphi}{2} + i \sin \frac{\varphi}{2} \right)$.

Упражнение 8. Найти тригонометрическую форму числа:

- а) $1 - i$;
 б) $1 - i\sqrt{3}$;
 в) $\frac{1 + i \operatorname{tg} \alpha}{1 - i \operatorname{tg} \alpha}$.

Пример 3. Описать множество точек, изображающих комплексные числа z , удовлетворяющие уравнению $|z - z_0| = r$ (числа $r > 0$, z_0 фиксированы).

1 способ. Пусть $z = x + yi$, $z_0 = x_0 + y_0i$, тогда исходное уравнение будет эквивалентно уравнению $(x - x_0)^2 + (y - y_0)^2 = r^2$. Множество точек, ему удовлетворяющих, есть окружность радиуса r с центром в точке $z_0 = x_0 + y_0i$.
 2 способ. Нетрудно видеть, что $|z - z_0|$ есть расстояние между точками z и z_0 . Таким образом, речь идет о всех точках z , для которых расстояние до фиксированной точки z_0 есть постоянная величина r . Описанное геометрическое место точек — окружность.

Пример 4. Описать множество точек, изображающих комплексные числа z , удовлетворяющие уравнению $|z - z_1| = |z - z_2|$ (числа z_1, z_2 — фиксированы).

Используя геометрическую интерпретацию, которую мы дали в предыдущем примере модулю разности двух комплексных чисел, получаем, что описанная уравнением $|z - z_1| = |z - z_2|$ совокупность есть множество точек, равноудаленных от z_1 и z_2 , т.е. серединный перпендикуляр отрезка $[z_1, z_2]$.

Упражнение 9. Изобразить на комплексной плоскости множество точек, соответствующих числам z , таким, что:

- а) $1 \leq |z - 2 + i| < 2$;
 б) $|\arg z| < \pi/4$;
 в) $\operatorname{Im} z = 3$.

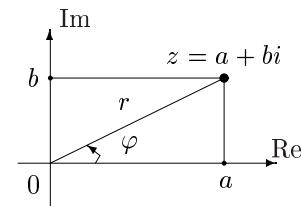


Рис. 2.2. Модуль и аргумент комплексного числа

Умножение и деление комплексных чисел в тригонометрической форме. Рассмотрим произведение двух комплексных чисел, записанных в тригонометрической форме: $z_1 z_2 = r(\cos \varphi + i \sin \varphi) \cdot \rho(\cos \psi + i \sin \psi)$. Применяя формулы для суммы и разности тригонометрических функций, получаем: $z_1 z_2 = r\rho(\cos \varphi \cos \psi - \sin \varphi \sin \psi + i(\sin \varphi \sin \psi + \cos \varphi \cos \psi)) = r\rho[\cos(\varphi + \psi) + i \sin(\varphi + \psi)]$. В конце полученной цепочки равенств имеем комплексное число, записанное в тригонометрической форме. Его модуль равен $r\rho$, аргумент равен $\varphi + \psi$. Иными словами,

$$|z_1 z_2| = |z_1| |z_2|, \quad \arg(z_1 z_2) = \arg(z_1) + \arg(z_2).$$

В геометрической интерпретации произведению $z_1 z_2$ соответствует точка, радиус-вектор которой получен поворотом радиус-вектора точки z_1 на угол $\arg(z_2)$ и растяжением в $|z_2|$ раз.

Нетрудно проверить, что, если $\rho \neq 0$, то

$$[\rho(\cos \psi + i \sin \psi)]^{-1} = \rho^{-1} [\cos(-\psi) + i \sin(-\psi)],$$

поэтому

$$\frac{r(\cos \varphi + i \sin \varphi)}{\rho(\cos \psi + i \sin \psi)} = \frac{r}{\rho} [\cos(\varphi - \psi) + i \sin(\varphi - \psi)],$$

или

$$\left| \frac{z_1}{z_2} \right| = \frac{|z_1|}{|z_2|}, \quad \arg \frac{z_1}{z_2} = \arg z_1 - \arg z_2.$$

В геометрической интерпретации частному z_1/z_2 соответствует точка, радиус-вектор которой получен поворотом радиус-вектора точки z_1 на угол $-\arg(z_2)$ и сжатием в $|z_2|$ раз.

Пример 5. Выполнить действия:

a)

$$\begin{aligned}
 (1 + i\sqrt{3})(1 + i)(\cos \varphi + i \sin \varphi) &= \\
 &= 2 \left(\cos \frac{\pi}{3} + i \sin \frac{\pi}{3} \right) \sqrt{2} \left(\cos \frac{\pi}{4} + i \sin \frac{\pi}{4} \right) (\cos \varphi + i \sin \varphi) = \\
 &= 2\sqrt{2} \left(\cos \left(\frac{7\pi}{12} + \varphi \right) + i \sin \left(\frac{7\pi}{12} + \varphi \right) \right);
 \end{aligned}$$

b)

$$\begin{aligned}
 \frac{\cos \varphi + i \sin \varphi}{\cos \psi - i \sin \psi} &= \\
 &= \frac{\cos \varphi + i \sin \varphi}{\cos(-\psi) + i \sin(-\psi)} = \\
 &= \cos(\varphi + \psi) + i \sin(\varphi + \psi).
 \end{aligned}$$

Пример 6. Найти тригонометрическую форму числа $1 + \cos \varphi + i \sin \varphi$, если $\varphi \in [-\pi, \pi]$.
 Данный пример уже был рассмотрен (пример 2 з). Приведем другой способ:

$$\begin{aligned}
 1 + \cos \varphi + i \sin \varphi &= \\
 &= (\cos 0 + i \sin 0) + (\cos \varphi + i \sin \varphi) = \\
 &= \left(\cos \left(-\frac{\varphi}{2} \right) + i \sin \left(-\frac{\varphi}{2} \right) \right) \left(\cos \frac{\varphi}{2} + i \sin \frac{\varphi}{2} \right) + \left(\cos \frac{\varphi}{2} + i \sin \frac{\varphi}{2} \right)^2 = \\
 &= 2 \cos \frac{\varphi}{2} \left(\cos \frac{\varphi}{2} + i \sin \frac{\varphi}{2} \right).
 \end{aligned}$$

Пусть n — произвольное натуральное число. Как и для вещественных чисел, будем говорить, что ζ есть n -ая степерь числа z и записывать $\zeta = z^n$, если

$$\zeta = \underbrace{z \cdot z \cdot \dots \cdot z}_{n \text{ сомножителей}}.$$

Мы уже определили z^{-1} как число, обратное к $z \neq 0$. Для произвольного $z \neq 0$ определим $z^0 = 1$ и $z^{-n} = (z^{-1})^n$. Докажем, что

$$(z^{-1})^n = (z^n)^{-1}.$$

Действительно, $(z^{-1})^n z^n = z^{-1} \cdot \dots \cdot z^{-1} z \cdot \dots \cdot z = (z^{-1} \cdot \dots \cdot (z^{-1} z) \cdot \dots \cdot z) = 1$.

Упражнение 10. Докажите, что $(z^m)^n = z^{mn}$, $z^m z^n = z^{m+n}$, $z_1^n z_2^n = (z_1 z_2)^n$ для произвольных целых m, n .

Используя метод математической индукции, теперь легко доказать *формулу Муавра* (A. de Moivre, 1736):

$$[r(\cos \varphi + i \sin \varphi)]^n = r^n [\cos(n\varphi) + i \sin(n\varphi)],$$

справедливую для произвольного целого n .

Пример 7. Вычислить:

$$\begin{aligned}
 (1 + i\sqrt{3})^{150} &= 2 \left(\cos \frac{\pi}{3} + i \sin \frac{\pi}{3} \right)^{150} = \\
 &= 2^{150} \left(\cos \frac{150\pi}{3} + i \sin \frac{150\pi}{3} \right) = \\
 &= 2^{150} [\cos(50\pi) + i \sin(50\pi)] = \\
 &= 2^{150} (\cos 0 + i \sin 0) = 2^{150}.
 \end{aligned}$$

Пример 8. Доказать, что если $z + \frac{1}{z} = 2 \cos \theta$, то

$$z^m + \frac{1}{z^m} = 2 \cos m\theta. \quad (2.8)$$

Исходное уравнение эквивалентно квадратному уравнению $z^2 - 2z \cos \theta + 1 = 0$. Его корни:

$$z_{1,2} = \cos \theta \pm \sqrt{\cos^2 \theta - 1} = \cos \theta \pm i \sin \theta.$$

Поэтому $z_{1,2}^n = \cos n\theta \pm i \sin n\theta$, $z_{1,2}^{-n} = \cos n\theta \mp i \sin n\theta$, откуда сразу следует доказываемое равенство.

Упражнение 11. Вычислить:

$$\begin{aligned}
 \text{a) } & \frac{(-1+i\sqrt{3})^{15}}{(1-i)^{20}} + \frac{(-1-i\sqrt{3})^{15}}{(1+i)^{20}}; \\
 \text{b) } & (1 + \cos \varphi + i \sin \varphi)^n;
 \end{aligned}$$

2.3. Комплексно сопряженные числа

Напомним, что число $a - bi$ называется (*комплексно*) *сопряженным* к числу $a + bi$. Для числа сопряженного к z используется обозначение $\bar{z}$. Нетрудно видеть, что $\bar{\bar{z}} = z$. Если число $z \neq 0$ задано в тригонометрической форме $z = r(\cos \varphi + i \sin \varphi)$, то $\bar{z} = r(\cos \varphi - i \sin \varphi) = r(\cos(-\varphi) + i \sin(-\varphi))$ (см. рис. 2.3). Таким образом, аргументы взаимно сопряженных чисел отличаются знаком, а модули совпадают: $|\bar{z}| = |z|$, $\arg \bar{z} = -\arg z$.

Выполняются следующие свойства операции сопряжения:

$$\begin{aligned}\overline{z_1 + z_2} &= \bar{z}_1 + \bar{z}_2, & \overline{z_1 - z_2} &= \bar{z}_1 - \bar{z}_2, \\ \overline{z_1 \cdot z_2} &= \bar{z}_1 \cdot \bar{z}_2, & \overline{z_1 / z_2} &= \bar{z}_1 / \bar{z}_2; \\ z\bar{z} &= |z|^2, & z + \bar{z} &= 2 \operatorname{Re} z.\end{aligned}$$

Докажем, например, первое свойство. Пусть $z_1 = a+bi$, $z_2 = c+di$, тогда $\overline{z_1 + z_2} = \overline{(a+c) + (b+d)i} = (a+c) - (b+d)i$, с другой стороны, $\bar{z}_1 + \bar{z}_2 = a-bi + c-di = (a+c) - (b+d)i$. Левая и правая части доказываемого равенства совпадают.

Упражнение 12. Доказать остальные свойства операции комплексного сопряжения.

Пример 9. Решить уравнение $\bar{z} = z^3$.

1 способ. Представим z в алгебраической форме: $z = x + iy$. Уравнение переписывается следующим образом:

$$x - iy = x^3 + 3x^2yi - 3xy^2 - y^3i.$$

Приравняв действительные и мнимые части, получаем совокупность систем:

$$\begin{cases} x = x^3 - 3xy^2, \\ -y = 3x^2y - y^3; \end{cases} \quad \text{или} \quad \begin{cases} x(x^2 - 3y^2 - 1) = 0, \\ y(y^2 - 3x^2 - 1) = 0; \end{cases}$$

откуда

$$\begin{aligned} & \begin{cases} x = 0, \\ y = 0; \end{cases} \quad \text{или} \quad \begin{cases} x = 0, \\ y^2 - 3x^2 = 1; \end{cases} \\ & \text{или} \quad \begin{cases} x^2 - 3y^2 = 1, \\ y = 0; \end{cases} \quad \text{или} \quad \begin{cases} x^2 - 3y^2 = 1, \\ y^2 - 3x^2 = 1; \end{cases} \end{aligned}$$

т.е.

$$\begin{aligned} & \begin{cases} x = 0, \\ y = 0; \end{cases} \quad \text{или} \quad \begin{cases} x = 0, \\ y^2 = 1; \end{cases} \\ & \text{или} \quad \begin{cases} x^2 = 1, \\ y = 0; \end{cases} \quad \text{или} \quad \begin{cases} x^2 - 3y^2 = 1, \\ y^2 - 3x^2 = 1. \end{cases} \end{aligned}$$

Третья система несовместна в $\mathbf{R}$. Действительно, умножая первое уравнение на 3 и складывая его со вторым, мы получаем $8y^2 = -4$, что невозможно для действительных y . Из первых трех систем получаем решения: $0, \pm i, \pm 1$ соответственно.

2 способ. Запишем z в тригонометрической форме: $z = r(\cos \varphi + i \sin \varphi)$. Уравнение примет вид:

$$r(\cos(-\varphi) + i \sin(-\varphi)) = r^3(\cos 3\varphi + i \sin 3\varphi),$$

откуда

$$\begin{cases} r = r^3, \\ -\varphi = 3\varphi + 2\pi k, \quad k \in \mathbf{Z}; \end{cases} \quad \text{т.е.} \quad \begin{cases} r = 0 \text{ или } r = 1, \\ \varphi = \frac{\pi k}{2}, \quad k \in \mathbf{Z}. \end{cases}$$

Последняя система дает решения: $0, \pm i, \pm 1$.

Упражнение 13. Решить уравнения:

- $|z| + z = 8 + 4i$;
- $\bar{z} = z^2$;
- $|z| - iz = 1 - 2i$;
- $z^2 = \bar{z}^3$;
- $z^2 + z|z| + |z|^2 = 0$.

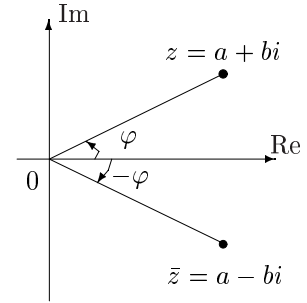


Рис. 2.3. Комплексно сопряженные числа

2.4. Неравенство треугольника

Теорема 2.4.1 (Неравенство треугольника). Для произвольных комплексных z_1, z_2 выполняются неравенства

$$|z_1| - |z_2| \leq ||z_1| - |z_2|| \leq |z_1 + z_2| \leq |z_1| + |z_2|.$$

Доказательство. Если одно из чисел, участвующих в неравенствах, равно нулю, то неравенства становятся очевидными. Поэтому будем считать, что $z_1 \neq 0, z_2 \neq 0$.

Вначале докажем, что $|z_1 + z_2| \leq |z_1| + |z_2|$. Представим числа $z_1, z_2, z_1 + z_2$ в тригонометрической форме:

$$\begin{aligned} z_1 &= r(\cos \varphi + i \sin \varphi), \\ z_2 &= \rho(\cos \psi + i \sin \psi), \\ z_1 + z_2 &= R(\cos \Phi + i \sin \Phi). \end{aligned}$$

Таким образом,

$$r(\cos \varphi + i \sin \varphi) + \rho(\cos \psi + i \sin \psi) = R(\cos \Phi + i \sin \Phi).$$

Приравнявая действительные и мнимые части в последнем равенстве, получаем:

$$\begin{aligned} r \cos \varphi + \rho \cos \psi &= R \cos \Phi, \\ r \sin \varphi + \rho \sin \psi &= R \sin \Phi. \end{aligned}$$

Теперь, умножая первое равенство на $\cos \varphi$, а второе — на $\sin \varphi$ и складывая, приходим к

$$r(\cos \varphi \cos \varphi + \sin \varphi \sin \varphi) + \rho(\cos \psi \cos \varphi + \sin \psi \sin \varphi) = R(\cos^2 \Phi + \sin^2 \Phi).$$

Вспоминая формулы тригонометрии, получаем:

$$r \cos(\varphi - \varphi) + \rho \cos(\psi - \varphi) = R.$$

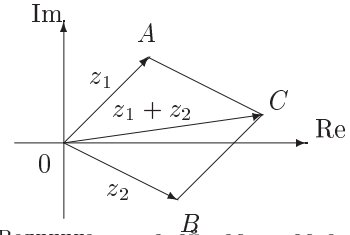
Так как $|\cos(\varphi - \varphi)| \leq 1$, $|\cos(\psi - \varphi)| \leq 1$, то $r + \rho \geq R$, т.е. $|z_1| + |z_2| \geq |z_1 + z_2|$.

Воспользуемся только что доказанным неравенством для чисел $z_1 + z_2$ и $-z_2$. Получаем: $|(z_1 + z_2) - z_2| \leq |z_1 + z_2| + |-z_2|$, т.е. $|z_1| \leq |z_1 + z_2| + |z_2|$, откуда $|z_1| - |z_2| \leq |z_1 + z_2|$. Аналогично можно получить $|z_2| - |z_1| \leq |z_1 + z_2|$, а, следовательно, $||z_1| - |z_2|| \leq |z_1 + z_2|$.

Неравенство $|z_1| - |z_2| \leq ||z_1| - |z_2||$ очевидно. ■

Замечание 2.4.1. Замечание Дадим геометрическую интерпретацию неравенству $|z_1 + z_2| \leq |z_1| + |z_2|$, только что доказанному чисто алгебраически.

На рис. 2.4 стороны треугольника OAC равны $OA = |z_1|$, $OB = AC = |z_2|$, $OC = |z_1 + z_2|$. Неравенство $|z_1 + z_2| \leq |z_1| + |z_2|$ выражает тот факт, что сторона треугольника не превосходит суммы двух других сторон.



Пример 10. Доказать, что если $|z| < \frac{1}{2}$, то

$$|(1+i)z^3 + iz| < \frac{3}{4}.$$

Из неравенства треугольника получаем: $|(1+i)z^3 + iz| \leq |(1+i)z^3| + |iz|$. Величина в правой части последнего неравенства по правилу вычисления модуля произведения равна $|1+i||z|^3 + |i||z| = \sqrt{2}|z|^3 + |z|$. Так как $|z| < \frac{1}{2}$, то

$$\sqrt{2}|z|^3 + |z| < \sqrt{2} \frac{1}{2^3} + \frac{1}{2} < \frac{3}{4}. \quad \text{Рис. 2.4. Неравенство треугольника}$$

Проследив всю цепочку рассмотренных неравенств, получаем доказываемое.

Упражнение 14. Доказать, что если $|z| \leq 2$, то $1 \leq |z^2 - 5| \leq 9$.

2.5. Корни из комплексных чисел

Пусть $n \in \mathbf{N}$. Число z называется *корнем n -ой степени* из числа ζ , если $z^n = \zeta$. Легко видеть, что $\zeta = 0$ обладает единственным (нулевым) значением корня произвольной натуральной степени. Пусть $\zeta \neq 0$.

Представим ζ в тригонометрической форме: $\zeta = \rho(\cos \psi + i \sin \psi)$. Мы ищем такое $z = r(\cos \varphi + i \sin \varphi)$, что $z^n = \zeta$. Воспользовавшись формулами Муавра, последнее равенство перепишем в виде

$$r^n [\cos(n\varphi) + i \sin(n\varphi)] = \rho(\cos \psi + i \sin \psi).$$

Так как для ненулевого комплексного числа модуль определен однозначно, а аргумент с точностью до $2\pi k$, где $k \in \mathbf{Z}$, то $r^n = \rho$, а $n\varphi = \psi + 2\pi k$. Получаем:

$$r = \sqrt[n]{\rho}, \quad \varphi = \frac{\psi + 2\pi k}{n}$$

(для вычисления r используется арифметическое значение корня $\sqrt[n]{\rho}$). Итак, для произвольного $k \in \mathbf{Z}$ каждое из чисел

$$z = \sqrt[n]{\rho} \left(\cos \frac{\psi + 2\pi k}{n} + i \sin \frac{\psi + 2\pi k}{n} \right) \quad (2.9)$$

является значением корня n -ой степени из числа $\zeta = \rho(\cos \psi + i \sin \psi)$.

Выясним, есть ли среди чисел (2.9) совпадающие. Разделим произвольное $k \in \mathbf{Z}$ на n с остатком, т.е. представим k в виде $k = np + q$ (напомним, что в данном случае p называется *частным*, q — *остатком*, $0 \leq q < n$). Подставляя это выражение для k в (2.9), получаем:

$$\begin{aligned} z &= \sqrt[n]{\rho} \left(\cos \frac{\psi + 2\pi(np + q)}{n} + i \sin \frac{\psi + 2\pi(np + q)}{n} \right) = \\ &= \sqrt[n]{\rho} \left(\cos \frac{\psi + 2\pi q}{n} + i \sin \frac{\psi + 2\pi q}{n} \right). \end{aligned}$$

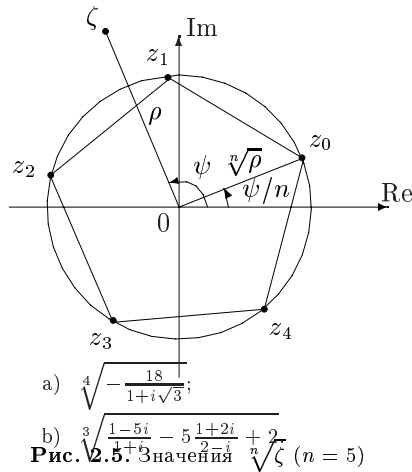
С другой стороны, при значениях $k = 0, 1, \dots, n-1$ все числа в (2.9), различны: их аргументы различны и отличаются не более, чем на 2π .

Подводя итог, получаем, что *произвольное ненулевое комплексное число $\zeta = \rho(\cos \psi + i \sin \psi)$ обладает n различными значениями корня n -ой степени из единицы; все эти значения можно получить по формуле:*

$$\sqrt[n]{\zeta} = \sqrt[n]{\rho} \left(\cos \frac{\psi + 2\pi k}{n} + i \sin \frac{\psi + 2\pi k}{n} \right), \quad (k = 0, 1, \dots, n-1). \quad (2.10)$$

Заметим, что знаки радикалов, встречающиеся в (2.10), имеют разный смысл: в правой части $\sqrt[n]{\cdot}$ означает арифметическое значение корня из положительного (действительного) числа, в левой — множество *всевозможных* значений корня из комплексного числа.

Из (2.10) легко видеть, что на комплексной плоскости все значения корня находятся на одинаковом расстоянии ($\sqrt[n]{\rho}$) от точки 0, кроме того, угол с вершиной в 0 между направлениями на соседние значения корня (центральный угол) постоянен и равен $2\pi/n$. Таким образом, *точки комплексной плоскости, соответствующие всем значениям корня степени $n \geq 3$ из одного и того же числа, находятся в вершинах правильного n -угольника* (см. рис. 2.5).



Пример 11. Найти все значения $\sqrt[3]{-8}$.

Представим -8 в тригонометрической форме: $-8 = 8(\cos \pi + i \sin \pi)$. По формуле (2.10) имеем:

$$\sqrt[3]{-8} = \sqrt[3]{8} \left(\cos \frac{\pi + 2\pi k}{3} + i \sin \frac{\pi + 2\pi k}{3} \right) \quad (k = 0, 1, 2).$$

Итак, $\sqrt[3]{-8}$ имеет следующие значения:

$$\begin{aligned} z_0 &= 2 \left(\cos \frac{\pi}{3} + i \sin \frac{\pi}{3} \right) = 1 + i\sqrt{3}, \\ z_1 &= 2 (\cos \pi + i \sin \pi) = -2, \\ z_2 &= 2 \left(\cos \frac{5\pi}{3} + i \sin \frac{5\pi}{3} \right) = 1 - i\sqrt{3}. \end{aligned}$$

Заметим, что значение корня $z_1 = -2$ известно нам со школы.

Упражнение 15. Вычислить

Утверждение 2.5.1. *Множество всех значений корня $\sqrt[n]{\zeta\xi}$ можно получить домножая произвольное частное значение корня $\sqrt[n]{\zeta}$ на все значения $\sqrt[n]{\xi}$ ($\zeta \neq 0$, $\xi \neq 0$).*

ДОКАЗАТЕЛЬСТВО. Пусть z — частное значение корня $\sqrt[n]{\zeta}$, т.е. $z^n = \zeta$, поэтому $(zx)^n = z^n x^n = \zeta \xi$ для любого значения x корня $\sqrt[n]{\xi}$, следовательно, zx — частное значение $\sqrt[n]{\zeta\xi}$. Когда x пробегает все n значений $\sqrt[n]{\xi}$, произведение zx принимает n *разных* значений $\sqrt[n]{\zeta\xi}$, т.е. все значения $\sqrt[n]{\zeta\xi}$. ■

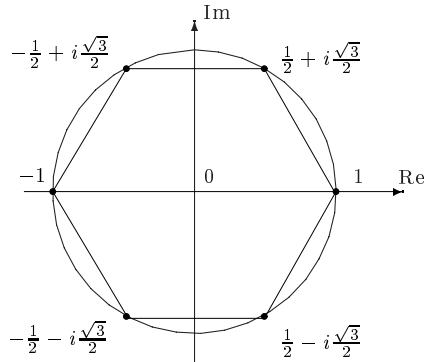
2.6. Корни из единицы

По формуле (2.10) корнями n -ой степени из $1 = \cos 0 + i \sin 0$ являются числа

$$\varepsilon_k = \cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n}. \quad (2.11)$$

Пример 12.

- а) $\sqrt[3]{1} = 1$ (одно значение).
 б) $\sqrt[2]{1} = \{1, -1\}$ (два значения).
 в) $\sqrt[3]{1}$. $\varepsilon_0 = 1$, $\varepsilon_1 = \cos \frac{2\pi}{3} + i \sin \frac{2\pi}{3} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$, $\varepsilon_2 = \cos \frac{4\pi}{3} + i \sin \frac{4\pi}{3} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$.
 г) $\sqrt[4]{1} = \{\pm 1, \pm i\}$.
 д) $\sqrt[5]{1} = \left\{ \pm 1, \pm \frac{1}{2} \pm i\frac{\sqrt{3}}{2} \right\}$ (рис. 2.6).

Рис. 2.6. Значения $\sqrt[5]{1}$

Упражнение 16. Найти все значения $\sqrt[12]{1}$.

Пример 13. Вычислить $\sqrt[3]{-8}$.

Данный пример уже рассматривался (пример 11). Приведем другой способ его решения, основанный на свойстве, установленном в конце предыдущего параграфа. Одним из частных значений $\sqrt[3]{-8}$, очевидно, является -2 . С другой стороны, $\sqrt[3]{1}$ имеет значения: $1, -\frac{1}{2} \pm i\frac{\sqrt{3}}{2}$. Поэтому все значений $\sqrt[3]{-8}$ исчерпываются тремя числами: $-2, 1 \pm i\sqrt{3}$.

Пример 14. Найти сумму всех корней n -ой степени из 1.

При $n = 1$ данная сумма, очевидно, равна 1. При $n > 1$ из (2.11) непосредственно следует, что $\varepsilon_k = (\varepsilon_1)^k$, поэтому упомянутая сумма есть сумма начального отрезка геометрической прогрессии. Имеем:

$$\begin{aligned} 1 + \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_{n-1} &= \\ &= 1 + \varepsilon_1 + (\varepsilon_1)^2 + \dots + (\varepsilon_1)^{n-1} = \\ &= \frac{1 - (\varepsilon_1)^n}{1 - \varepsilon_1} = 0, \end{aligned}$$

так как $(\varepsilon_1)^n = 1$.

Упражнение 17.

- а) Найти произведение всех корней n -ой степени из 1.
 б) Найти сумму и произведение s -ых степеней всех корней n -ой степени из 1.

Упражнение 18. Вычислить $1 + 2\varepsilon + 3\varepsilon^2 + \dots + n\varepsilon^{n-1}$, если ε — корень n -ой степени из 1.

Упражнение 19. Решить уравнения:

- а) $(x+1)^n - (x-1)^n = 0$;
 б) $(x+i)^n + (x-i)^n = 0$.

Утверждение 2.6.1. Произведение и частное любых двух значений корня n -ой степени из 1 является корнем n -ой степени из 1.

Доказательство. Пусть $\varepsilon, \varepsilon'$ — некоторые значения корня n -ой степени из 1, тогда $\varepsilon^n = 1$ и $\varepsilon'^n = 1$, следовательно, $(\varepsilon\varepsilon')^n = \varepsilon^n\varepsilon'^n = 1$ и $(\varepsilon/\varepsilon')^n = \varepsilon^n/\varepsilon'^n = 1$, таким образом, $\varepsilon\varepsilon', \varepsilon/\varepsilon'$ являются значениями корня n -ой степени из 1. ■

Первообразные корни. Корень ε n -ой степени из 1 называется *первообразным*, если $\varepsilon^m \neq 1$ для любого натурального $m < n$ (т.е. не является корнем из единицы никакой меньшей степени). В данном случае говорят также, что ε *принадлежит показателю n* . Из определения сразу вытекает, что произвольное число ε может принадлежать лишь одному показателю. Легко видеть, что

$$\varepsilon_1 = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$$

является первообразным корнем.

Утверждение 2.6.2. Для того, чтобы корень ε n -ой степени из 1 являлся первообразным, необходимо и достаточно, чтобы величины

$$\varepsilon, \varepsilon^2, \dots, \varepsilon^{n-1}, \varepsilon^n = 1 \quad (2.12)$$

исчерпывали все значения $\sqrt[n]{1}$.

Замечание 2.6.1. Утверждение о том, что величины (2.12) исчерпывают все значения корня n -ой степени из 1 эквивалентно тому, что они попарно различны.

ДОКАЗАТЕЛЬСТВО.

НЕОБХОДИМОСТЬ. Предположим противное: среди величин (2.12) нашлось две равных, например, $\varepsilon^k = \varepsilon^l$ для некоторых натуральных k, l , причем $1 \leq k < l \leq n$. Тогда $\varepsilon^{l-k} = 1$, $l - k > 0$, следовательно, ε не является первообразным.

ДОСТАТОЧНОСТЬ. Так как величины (2.12) исчерпывают все значения корня n -ой степени из 1, то, в частности, $\varepsilon^m \neq \varepsilon^n = 1$ для всякого натурального $m < n$, следовательно, ε — первообразный. ■

Утверждение 2.6.3. Пусть ε — первообразный корень n -ой степени из 1, тогда для того, чтобы $\varepsilon^m = 1$, необходимо и достаточно, чтобы m было кратно n .

ДОКАЗАТЕЛЬСТВО.

НЕОБХОДИМОСТЬ. Разделим m с остатком на n : имеем $m = np + r$ для некоторых натуральных p и r ($0 \leq r < n$), поэтому

$$1 = \varepsilon^m = \varepsilon^{np} \varepsilon^r = \varepsilon^r.$$

Итак, $\varepsilon^r = 1$. Так как ε — первообразный и $0 \leq r < n$, то $r = 0$.

ДОСТАТОЧНОСТЬ. Имеем $m = np$, следовательно, $\varepsilon^m = (\varepsilon^n)^p = 1$. ■

Следствие 2.6.1. Если корень ε m -ой степени из 1 принадлежит показателю n , то m кратно n .

Утверждение 2.6.4. Пусть ε — первообразный корень n -ой степени из 1, тогда для того, чтобы ε^k был также первообразным n -ой степени, необходимо и достаточно, чтобы $(n, k) = 1$.

ДОКАЗАТЕЛЬСТВО.

НЕОБХОДИМОСТЬ. Предположим противное: ε и ε^k — первообразные, однако $(n, k) = d > 1$. Имеем $n = n'd$, $k = k'd$ для некоторых натуральных n', k' , причем, так как $d > 1$, то $n' < n$. Однако

$$(\varepsilon^k)^{n'} = \varepsilon^{kn'} = \varepsilon^{k'dn'} = \varepsilon^{k'n} = (\varepsilon^n)^{k'} = 1.$$

Так как $n' < n$, то ε^k не является первообразным корнем.

ДОСТАТОЧНОСТЬ. Предположим теперь, что ε — первообразный корень, $(n, k) = 1$, однако $(\varepsilon^k)^m = 1$ для некоторого натурального $m < n$ (т.е. ε^k не является первообразным). Из утверждения 2 следует, что km кратно n , однако $(n, k) = 1$, поэтому m кратно n , что невозможно, так как $m < n$. ■

Следствие 2.6.2 (Критерий). Величина

$$\varepsilon_k = \cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n} \quad (k = 0, 1, \dots, n-1)$$

является первообразным корнем n -ой степени из 1 тогда и только тогда, когда $(n, k) = 1$.

ДОКАЗАТЕЛЬСТВО. Применим к первообразному корню $\varepsilon = \varepsilon_1$ утверждение 3. ■

Пример 15. Найти все первообразные корни из 1 степени:

- 1, ответ: 1;
- 2, один первообразный корень: $\varepsilon_1 = -1$;
- 3, два первообразных корня: $\varepsilon_{1,2} = -\frac{1}{2} \pm i\frac{\sqrt{3}}{2}$;
- 4, два первообразных корня: $\varepsilon_{1,3} = \pm i$;
- 6, два первообразных корня: $\varepsilon_{1,5} = \frac{1}{2} \pm i\frac{\sqrt{3}}{2}$;
- 8, выпишем все натуральные числа, не превосходящие $n = 8$ и взаимно простые с ним: 1, 3, 5, 7; первообразными корнями являются величины:

$$\begin{aligned} \varepsilon_1 &= \cos \frac{2\pi}{8} + i \sin \frac{2\pi}{8} = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}, \\ \varepsilon_3 &= \cos \frac{3 \cdot 2\pi}{8} + i \sin \frac{3 \cdot 2\pi}{8} = -\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}, \\ \varepsilon_5 &= \cos \frac{5 \cdot 2\pi}{8} + i \sin \frac{5 \cdot 2\pi}{8} = -\frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2}, \\ \varepsilon_7 &= \cos \frac{7 \cdot 2\pi}{8} + i \sin \frac{7 \cdot 2\pi}{8} = \frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2}. \end{aligned}$$

Упражнение 20. Найти все первообразные корни из 1 степени 12;

Упражнение 21. Вычислить сумму $1 + \varepsilon + \varepsilon^2 + \dots + \varepsilon^{n-1}$, если ε — первообразный корень степени $2n$ из 1.

Из утверждения 3 следует, что число первообразных корней n -ой степени из 1 совпадает с количеством $\varphi(n)$ натуральных чисел, не превосходящих n и взаимно простых с ним. Функция $\varphi(n)$ называется *функцией Эйлера*, она составляет предмет пристального изучения в теории чисел.

Круговые многочлены. Круговым многочленом показателя n , или многочленом деления круга, или циклотомическим многочленом, называется выражение

$$\Phi_n(x) = (x - \eta_1)(x - \eta_2) \dots (x - \eta_s),$$

где $\eta_1, \eta_2, \dots, \eta_s$ — первообразные корни n -ой степени из 1.

Пример 16.

- a) $\Phi_1(x) = x + 1$;
- b) $\Phi_2(x) = x - 1$;
- c) $\Phi_3(x) = (x + \frac{1}{2} - i\frac{\sqrt{3}}{2})(x + \frac{1}{2} + i\frac{\sqrt{3}}{2}) = x^2 + x + 1$;
- d) $\Phi_4(x) = (x - i)(x + i) = x^2 + 1$;
- e) $\Phi_6(x) = (x - \frac{1}{2} - i\frac{\sqrt{3}}{2})(x - \frac{1}{2} + i\frac{\sqrt{3}}{2}) = x^2 - x + 1$;

Утверждение 2.6.5.

$$x^n - 1 = \prod_{m|n} \Phi_m(x)$$

(произведение берется по всем натуральным делителям m числа n).

Доказательство. Чтобы перечислить все корни n -ой степени из 1, воспользуемся следующей процедурой. Для каждого делителя m числа n выпишем все корни, принадлежащие показателю m . Из утверждения 2 следует, что таким образом мы перечислим все корни n -ой степени из 1 и каждый по одному разу. Доказываемое теперь следует из разложения

$$x^n - 1 = (x - 1)(x - \varepsilon_1) \dots (x - \varepsilon_{n-1}).$$

Пример 17. Выписать $\Phi_5(x)$, $\Phi_{10}(x)$.

- a) Делителями $n = 5$ являются числа 1, 5, поэтому $x^5 - 1 = \Phi_1 \Phi_5$, откуда

$$\Phi_5 = \frac{x^5 - 1}{\Phi_1} = x^4 + x^3 + x^2 + x + 1.$$

- b) Делителями $n = 10$ являются числа 1, 2, 5, 10, поэтому $x^{10} - 1 = \Phi_1 \Phi_2 \Phi_5 \Phi_{10}$, откуда

$$\begin{aligned} \Phi_{10} &= \frac{x^{10} - 1}{\Phi_1 \Phi_2 \Phi_5} = \frac{(x^{10} - 1)\Phi_1}{\Phi_1 \Phi_2 (x^5 - 1)} = \frac{x^5 + 1}{\Phi_2} = \frac{x^5 + 1}{x - 1} = \\ &= x^4 - x^3 + x^2 - x + 1. \end{aligned}$$

Упражнение 22. Найти

- a) $\Phi_{12}(x)$;
- b) $\Phi_{15}(x)$.

Упражнение 23. Найти $\Phi_p(x)$, если p — простое число.

2.7. Квадратные корни из комплексных чисел

Квадратным корнем из комплексного числа $z = a + bi$ является такое число $x + iy$, что $(x + iy)^2 = a + bi$. Пусть $z \neq 0$, тогда: $a + bi = x^2 + 2xyi - y^2$, или

$$\begin{cases} a = x^2 - y^2, \\ b = 2xy. \end{cases} \quad (2.13)$$

Возведем оба уравнения системы в квадрат и прибавим к первому второе: $a^2 + b^2 = x^4 + 2x^2y^2 + y^4$, откуда $a^2 + b^2 = (x^2 + y^2)^2$. Так как $x^2 + y^2 > 0$, то $x^2 + y^2 = \sqrt{a^2 + b^2}$. Рассмотрим это уравнение вместе с первым уравнением системы (2.13):

$$\begin{cases} x^2 + y^2 = \sqrt{a^2 + b^2}, \\ x^2 - y^2 = a. \end{cases}$$

Отсюда,

$$\begin{aligned} x^2 &= \frac{1}{2}(a + \sqrt{a^2 + b^2}), \\ y^2 &= \frac{1}{2}(-a + \sqrt{a^2 + b^2}). \end{aligned} \quad (2.14)$$

Каждое из этих двух соотношений дает два разных значения для x и y . Комбинируя их, мы можем получить четыре различных комплексных числа, однако не все они удовлетворяют системе (2.13): как видно из второго уравнения, *знаки x и y должны совпадать, если $b > 0$, и различаться,*

если $b < 0$. Если $b = 0$, т.е. число z — вещественное, то либо x , либо y равно нулю. В предыдущем разделе мы видели, что корень n -ой степени из произвольного ненулевого комплексного числа имеет ровно n значений. Таким образом, для $n = 2$ эти значения получаются по формулам (2.14), скомбинированным с приведенным правилом выбора знака.

Пример 18. Найти все значения $\sqrt[4]{2 - i\sqrt{12}}$.

Сначала найдем все значения корня квадратного из $2 - i\sqrt{12}$. Из (2.14) имеем $x^2 = 3$, $y^2 = 1$. Так как мнимая часть подкоренного числа отрицательна, то $\sqrt{2 - i\sqrt{12}} = \pm(\sqrt{3} - i)$. Вычислим теперь $\sqrt{\sqrt{3} - i}$. Получаем $x^2 = \frac{1}{2}(\sqrt{3} + 2)$, $y^2 = \frac{1}{2}(-\sqrt{3} + 2)$. Отсюда $\sqrt{\sqrt{3} - i} = \pm \left(\sqrt{\frac{\sqrt{3} + 2}{2}} - i\sqrt{\frac{-\sqrt{3} + 2}{2}} \right) = \pm \left(\frac{1 + \sqrt{3}}{2} + \frac{1 - \sqrt{3}}{2}i \right)$. Теперь необходимо найти $\sqrt{-\sqrt{3} + i}$. Легко видеть, что оба значения этого корня отличаются от соответствующих значений $\sqrt{\sqrt{3} - i}$ множителем i . Итак, значениями корня являются $\pm \left(\frac{1 + \sqrt{3}}{2} + \frac{1 - \sqrt{3}}{2}i \right)$, $\pm \left(\frac{\sqrt{3} - 1}{2} + \frac{\sqrt{3} + 1}{2}i \right)$.

Упражнение 24. Найти

- a) $\sqrt{-8i}$,
- b) $\sqrt{3 - 4i}$.

2.8. Уравнения второй, третьей и четвертой степени

Квадратные уравнения. Уравнением второй степени называется уравнение $ax^2 + bx + c = 0$, где $a \neq 0$, $a, b, c \in \mathbb{C}$. Делением его на a получаем уравнение, равносильное исходному:

$$x^2 + px + q = 0. \quad (2.15)$$

Для его решения воспользуемся способом выделения полного квадрата. В правой части имеем:

$$x^2 + 2x\frac{p}{2} + \left(\frac{p}{2}\right)^2 - \left(\frac{p}{2}\right)^2 + q = \left(x + \frac{p}{2}\right)^2 - \left(\frac{p}{2}\right)^2 + q.$$

Уравнение примет вид

$$\left(x + \frac{p}{2}\right)^2 = \left(\frac{p}{2}\right)^2 + q, \quad \text{или } x + \frac{p}{2} = \sqrt{\left(\frac{p}{2}\right)^2 + q}.$$

Квадратный корень в последней формуле имеет два значения, отличающиеся знаком. Примем для них обозначение $\pm \sqrt{}$. Итак, уравнение (2.15) имеет в общем случае два решения

$$x_{1,2} = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 + q}.$$

Пример 19. $x^2 - (2 + i)x + (-1 + 7i) = 0$.

$$x_{1,2} = \frac{2 + i}{2} \pm \sqrt{\frac{3 + 4i}{4} + 1 - 7i} = \frac{2 + i}{2} \pm \frac{\sqrt{7 - 24i}}{2}.$$

Для нахождения всех значений квадратного корня воспользуемся методами из предыдущего раздела:

$$\begin{aligned} x^2 &= \frac{1}{2}(7 + \sqrt{7^2 + 24^2}) = 16, \\ y^2 &= \frac{1}{2}(-7 + \sqrt{7^2 + 24^2}) = 9. \end{aligned}$$

Отсюда $\sqrt{7 - 24i} = \pm(4 - 3i)$, следовательно,

$$x_{1,2} = \frac{2 + i}{2} \pm \frac{4 - 3i}{2}.$$

Итак, $x_1 = 3 - i$, $x_2 = -1 + 2i$.

Упражнение 25. Решить уравнения:

- a) $x^2 - (3 - 2i)x + (5 - 5i) = 0$;
- b) $(2 + i)x^2 - (5 - i)x + (2 - 2i) = 0$.

Упражнение 26. Решить уравнения:

- a) $x^4 - 3x^2 + 4 = 0$;
- b) $x^4 - 30x^2 + 289 = 0$.

Упражнение 27. Составить формулу для решения биквадратного уравнения $x^4 + px + q = 0$ с вещественными коэффициентами, удобную в случае $\frac{p^2}{4} - q < 0$.

Кубические уравнения. После замены $y = x - \frac{a}{3}$ в уравнении третьей степени $y^3 + ay^2 + by + c = 0$ исчезает член с квадратом неизвестной y . Уравнение переписывается:

$$x^3 + px + q = 0. \quad (2.16)$$

Его решения будем искать в виде

$$x = \alpha + \beta, \quad (2.17)$$

где α, β — некоторые комплексные числа, связанные кроме (2.17) другим соотношением, которое мы определим ниже. После подстановки (2.17) в (2.16) получаем:

$$\alpha^3 + 3\alpha^2\beta + 3\alpha\beta^2 + p\alpha + p\beta + q = 0,$$

или

$$\alpha^3 + \beta^3 + \alpha(3\alpha\beta + p) + \beta(3\alpha\beta + p) + q = 0,$$

или

$$\alpha^3 + \beta^3 + (\alpha + \beta)(3\alpha\beta + p) + q = 0. \quad (2.18)$$

Пусть

$$3\alpha\beta + p = 0, \quad (2.19)$$

тогда (2.18) примет вид

$$\alpha^3 + \beta^3 = -q. \quad (2.20)$$

Переписывая (2.20) и возводя (2.19) в куб, получим систему

$$\begin{cases} \alpha^3 + \beta^3 = -q, \\ \alpha^3\beta^3 = -\frac{p^3}{27}. \end{cases}$$

Используя теорему Виета, получаем, что α^3, β^3 являются решениями следующего квадратного уравнения

$$z^2 + qz - \frac{p^3}{27} = 0,$$

поэтому

$$\alpha = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}, \quad \beta = \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}. \quad (2.21)$$

Среди всевозможных значений комбинаций α, β необходимо выбрать лишь те, которые удовлетворяют условию (2.19). Легко видеть, что таким образом будет получено 3 решения (для каждого α из (2.19) можно определить единственное β). На практике из (2.21) выбирают какую-нибудь пару α_1, β_1 , удовлетворяющую (2.19); решениями уравнения (2.16) являются числа

$$\begin{aligned} x_1 &= \alpha_1 + \beta_1, \\ x_2 &= \alpha_1\varepsilon + \beta_1\varepsilon^2 = -\frac{\alpha_1 + \beta_1}{2} - i\frac{\alpha_1 - \beta_1}{2}\sqrt{3}, \\ x_3 &= \alpha_1\varepsilon^2 + \beta_1\varepsilon = -\frac{\alpha_1 + \beta_1}{2} + i\frac{\alpha_1 - \beta_1}{2}\sqrt{3}, \end{aligned} \quad (2.22)$$

где $\varepsilon = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$.

Выражения (2.21, 2.22) называются *формулами Кардано*.

Упражнение 28. Проверьте, что x_2, x_3 удовлетворяют условию (2.19).

Пример 20. $y^3 + 3y^2 - 6y + 4 = 0$.

После подстановки $y = x - 1$ исходное уравнение примет вид: $x^3 - 9x + 12 = 0$. Имеем:

$$\begin{aligned} \alpha &= \sqrt[3]{-\frac{12}{2} + \sqrt{\frac{144}{4} - \frac{9^3}{27}}} = \sqrt[3]{-6 + \sqrt{9}} = \\ &= \sqrt[3]{-6-3} = -\sqrt[3]{9}, \\ \beta &= \sqrt[3]{-6+3} = -\sqrt[3]{9}. \end{aligned}$$

Поэтому

$$\begin{aligned} x_1 &= -\left(\sqrt[3]{9} + \sqrt[3]{9}\right), \\ x_{2,3} &= \frac{\sqrt[3]{9} + \sqrt[3]{9}}{2} \pm i\frac{\sqrt[3]{9} - \sqrt[3]{9}}{2}\sqrt{3}. \end{aligned}$$

Отсюда

$$\begin{aligned} y_1 &= -\left(1 + \sqrt[3]{9} + \sqrt[3]{9}\right), \\ y_{2,3} &= \frac{\sqrt[3]{9} + \sqrt[3]{9} - 2}{2} \pm i\frac{\sqrt[3]{9} - \sqrt[3]{9}}{2}\sqrt{3}. \end{aligned}$$

Кубические уравнения с действительными коэффициентами. Пусть в уравнении (2.16) p, q — действительные числа. В зависимости от знака выражения $q^2/4 + p^3/27$ в формулах Кардано (2.21) приходится вычислять корень квадратный либо из положительного, либо из нулевого, либо из отрицательного (действительного) числа. Рассмотрим каждый из этих случаев.

1) $q^2/4 + p^3/27 > 0$. Под знаком кубического радикала в (2.21) стоят действительные числа. Вещественные значения этих корней дадут действительный корень $\alpha_1 + \beta_1$. Два других корня, вычисленных по формулам (2.22), очевидно, будут взаимно сопряженными комплексными (см. пример 20).

2) $q^2/4 + p^3/27 = 0$. В этом случае $\alpha_1 = \beta_1$, поэтому $x_1 = 2\alpha_1, x_2 = x_3 = -\alpha_1$. Все корни действительные.

3) $q^2/4 + p^3/27 < 0$ (так называемый “неприводимый случай”). Пусть

$$z = -\frac{q}{2} + i\sqrt{-\frac{q^2}{4} - \frac{p^3}{27}} = r(\cos \varphi + i \sin \varphi)$$

(в данном случае радикал означает арифметическое значение корня). Определим модуль r :

$$r = \sqrt{\frac{q^2}{4} - \frac{q^2}{4} - \frac{p^3}{27}} = \sqrt{-\frac{p^3}{27}},$$

и аргумент φ :

$$r \cos \varphi = -\frac{q}{2}, \text{ следовательно, } \cos \varphi = -\frac{q}{2} \sqrt{-\frac{27}{p^3}}.$$

Из формул Кардано (2.21) следует, что

$$\alpha_1 = \sqrt[3]{z} = \sqrt[3]{r} \left(\cos \frac{\varphi}{3} + i \sin \frac{\varphi}{3} \right), \quad \beta_1 = \sqrt[3]{\bar{z}} = \sqrt[3]{r} \left(\cos \frac{\varphi}{3} - i \sin \frac{\varphi}{3} \right).$$

Легко проверить, что $\alpha_1 \beta_1 = -p/3$, поэтому по формулам (2.22) получаем:

$$\begin{aligned} x_1 &= \alpha_1 + \beta_1 = 2\sqrt[3]{r} \cos \frac{\varphi}{3} = 2\sqrt{-\frac{p}{3}} \cos \frac{\varphi}{3}, \\ x_{2,3} &= \alpha \varepsilon^{1,2} + \beta \varepsilon^{2,1} = \\ &= \sqrt[3]{r} \left[\cos \left(\frac{\varphi}{3} \pm \frac{2\pi}{3} \right) + i \sin \left(\frac{\varphi}{3} \pm \frac{2\pi}{3} \right) \right] + \\ &\quad + \sqrt[3]{r} \left[\cos \left(-\frac{\varphi}{3} \mp \frac{2\pi}{3} \right) + i \sin \left(-\frac{\varphi}{3} \mp \frac{2\pi}{3} \right) \right] = \\ &= 2\sqrt[3]{r} \cos \left(\frac{\varphi}{3} \pm \frac{2\pi}{3} \right) = -2\sqrt[3]{r} \cos \left(\frac{\varphi}{3} \mp \frac{\pi}{3} \right) = \\ &= -2\sqrt[3]{-\frac{p}{3}} \cos \left(\frac{\varphi}{3} \mp \frac{\pi}{3} \right). \end{aligned}$$

Итак, в “неприводимом” случае все три корня уравнения вещественные и различные и могут быть найдены по формулам:

$$\begin{aligned} x_1 &= 2\sqrt{-\frac{p}{3}} \cos \frac{\varphi}{3}, \\ x_{2,3} &= -2\sqrt[3]{-\frac{p}{3}} \cos \left(\frac{\varphi}{3} \mp \frac{\pi}{3} \right), \\ \text{где } \cos \varphi &= -\frac{q}{2} \sqrt{-\frac{27}{p^3}}. \end{aligned} \tag{2.23}$$

(так называемое “тригонометрическое” решение).

Заметим, что в данном случае несмотря на то, что все три корня действительные, нам не удалось их выразить через радикалы от *действительных* же чисел: в формулах (2.21) приходится извлекать кубический корень из комплексного (не действительного) числа, а в (2.23) встречаются трансцендентные функции. Можно показать, что в *общем* случае (т.е. для уравнения с *буквенными* коэффициентами) это невозможно, хотя в *частных* примерах иногда удается представить вещественные корни кубического уравнения через радикалы от действительных чисел (см. примеры ниже).

Пример 21. $x^3 - 6x + 4 = 0$.

По формулам (2.21) получаем

$$\begin{aligned} \alpha &= \sqrt[3]{-2 + \sqrt{4 - 8}} = \sqrt[3]{-2 + 2i}, \\ \beta &= \sqrt[3]{-2 - \sqrt{4 - 8}} = \sqrt[3]{-2 - 2i}. \end{aligned}$$

Здесь кубический корень удается извлечь “точно”. Например, $\alpha_1 = 1 + i$, $\beta_1 = 1 - i$. Прямой подстановкой в (2.19) убеждаемся, что условие (2.19) выполнено. Решения:

$$\begin{aligned} x_1 &= (1 + i) + (1 - i) = 2, \\ x_2 &= (1 + i)\left(-\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) + (1 - i)\left(-\frac{1}{2} - \frac{\sqrt{3}}{2}i\right) = -1 - \sqrt{3}, \\ x_3 &= (1 + i)\left(-\frac{1}{2} - \frac{\sqrt{3}}{2}i\right) + (1 - i)\left(-\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) = -1 + \sqrt{3}. \end{aligned}$$

Пример 22. $x^3 - 19x + 30 = 0$.

По формулам (2.23) получаем (приближенные вычисления с точностью до 4 значащих цифр):

$$\begin{aligned} p &= -19, \quad q = 30, \\ \cos \varphi &= -\frac{30}{2} \sqrt{\frac{27}{19^3}} \approx -0,9412, \\ \text{отсюда } \varphi &\approx 160^\circ 16', \quad \frac{\varphi}{3} \approx 53^\circ 25', \\ x_1 &= 2\sqrt{\frac{19}{3}} \cos 53^\circ 25' \approx 3,002, \\ x_2 &= -2\sqrt{\frac{19}{3}} \cos 113^\circ 25' \approx 1,999, \\ x_3 &= -2\sqrt{\frac{19}{3}} \cos 6^\circ 35' \approx -5,002. \end{aligned}$$

Легко проверить, что точными значениями решений уравнения являются целые числа 3, 2, -5.

Уравнения четвертой степени. Опишем *способ Феррари* (L. Ferrari, 1545) для решения уравнения четвертой степени

$$y^4 + ay^3 + by^2 + cy + d = 0. \quad (2.24)$$

Легко проверить, что для произвольного значения параметра λ уравнение (2.24) можно представить в виде:

$$\left(x^2 + \frac{a}{2}x + \frac{\lambda}{2}\right)^2 - \left[\left(\frac{a^2}{4} + \lambda - b\right)x^2 + \left(\frac{a\lambda}{2} - c\right)x + \left(\frac{\lambda^2}{4} - d\right)\right] = 0. \quad (2.25)$$

Подберем λ так, чтобы выражение в квадратных скобках было квадратом двучлена, зависящего от x . Для этого необходимо приравнять к нулю дискриминант этого выражения:

$$\left(\frac{a\lambda}{2} - c\right)^2 - 4\left(\frac{a^2}{4} + \lambda - b\right)\left(\frac{\lambda^2}{4} - d\right) = 0. \quad (2.26)$$

Таким образом, взяв любой корень λ уравнения (2.26), мы сможем разложить левую часть (2.25) на линейные множители и тем самым отыскать корни исходного уравнения. Вспомогательное кубическое уравнение (2.26) называется *резольвентой*.

Пример 23. $x^4 - 2x^3 + 2x^2 + 4x - 8 = 0$.

Резольвента имеет вид:

$$\left(-\frac{2\lambda}{2} - 4\right)^2 - 4\left(\frac{(-2)^2}{4} + \lambda - 2\right)\left(\frac{\lambda^2}{4} + 8\right) = 0.$$

После раскрытия скобок и приведения подобных получаем:

$$\lambda^3 - 2\lambda^2 + 24\lambda - 48 = 0,$$

далее,

$$\lambda^2(\lambda - 2) + 24(\lambda - 2) = 0,$$

и, наконец,

$$(\lambda - 2)(\lambda^2 + 24) = 0.$$

Одним из корней является $\lambda = 2$. Воспользовавшись (2.25), левую часть исходного уравнения запишем в виде

$$\begin{aligned} (x^2 - x + 1)^2 - (x^2 - 6x + 9) &= \\ = (x^2 - x + 1)^2 - (x - 3)^2 &= \\ = (x^2 - x + 1 - x + 3)(x^2 - x + 1 + x - 3) &= \\ = (x^2 - 2x + 4)(x^2 - 2). \end{aligned}$$

Раскладывая на линейные множители, получаем следующие корни: $1 \pm i\sqrt{3}, \pm\sqrt{2}$.

Уравнения высших степеней. Мы с вами увидели, что корни алгебраического уравнения степени не выше 4 с буквенными коэффициентами можно выразить через коэффициенты этого уравнения при помощи конечного числа действий сложения, вычитания, умножения, деления и знаков корня (говорят, что общее уравнение степени не выше 4 *разрешимо в радикалах*). Н. Абель (N. Abel) в 1826 г. показал, что для алгебраических уравнений степени выше 4 этого сделать нельзя, иными словами, общее уравнение степени 5 и выше неразрешимо в радикалах. Результат Абеля не исключал возможности, что корни каждого *конкретного* уравнения (с числовыми коэффициентами) степени выше 4 выражаются с помощью некоторой комбинации арифметических операций и операций извлечения корня над коэффициентами исходного уравнения. Уравнения любой степени частных видов решаются в радикалах (например, *двучленное уравнение* $x^n - a = 0$ ⁴). Полное решение вопроса о том, при каких условиях алгебраическое уравнение разрешимо в радикалах было получено Э. Галуа (E. Galois) в 1830 г. Из результатов Галуа, например, следует, что для любой степени выше 4 найдутся уравнения с рациональными (и даже целыми) коэффициентами, неразрешимые в радикалах. Простым примером такого уравнения может служить $x^5 - 4x - 2 = 0$.

Пример 24. Выразить в радикалах $\cos \frac{2\pi}{5}, \sin \frac{2\pi}{5}$.

Число $\cos \frac{2\pi}{5} + i \sin \frac{2\pi}{5}$ является первообразным корнем 5-ой степени из 1 и поэтому одним из решений уравнения $x^5 - 1 = 0$ или $\Phi_5(x) = 0$. Таким образом, задача сводится к разысканию решений уравнения $x^4 + x^3 + x^2 + x + 1 = 0$. Это уравнение *возвратное*. Так как $x = 0$ не является его корнем, то мы можем записать:

$$x^2 + x + 1 + \frac{1}{x} + \frac{1}{x^2} = 0, \quad (2.27)$$

или

$$\left(x^2 + \frac{1}{x^2}\right) + \left(x + \frac{1}{x}\right) + 1 = 0.$$

Пусть

$$u = x + \frac{1}{x}, \quad (2.28)$$

⁴Его корни $x = \sqrt[n]{a}$.

тогда $u^2 = x^2 + 2 + \frac{1}{x^2}$ и (2.27) примет вид $u^2 + u - 1 = 0$. Корни последнего уравнения: $u_{1,2} = -\frac{1}{2} \pm \frac{\sqrt{5}}{2}$. Теперь из (2.28) имеем:

$$\begin{aligned} x_{1,2} &= \frac{\sqrt{5}}{4} - \frac{1}{4} \pm i \frac{1}{4} \sqrt{2} \sqrt{5 + \sqrt{5}}, \\ x_{3,4} &= -\frac{\sqrt{5}}{4} - \frac{1}{4} \pm i \frac{1}{4} \sqrt{2} \sqrt{5 - \sqrt{5}}. \end{aligned}$$

Очевидно, $\cos \frac{2\pi}{5} + i \sin \frac{2\pi}{5} = x_1$. Таким образом,

$$\cos \frac{2\pi}{5} = \frac{\sqrt{5}}{4} - \frac{1}{4}, \quad \sin \frac{2\pi}{5} = \frac{1}{4} \sqrt{2} \sqrt{5 + \sqrt{5}}.$$

Упражнение 29. Разработать метод построения при помощи циркуля и линейки правильного пятиугольника, вписанного в окружность заданного радиуса.

Указание. Использовать выражение для $\cos \frac{2\pi}{5}$. Разработать способ построения отрезка длины $\sqrt{5}$.

Пример 25. Составить алгебраическое уравнение с целыми коэффициентами, корнем которого является длина стороны правильного 14-угольника, вписанного в круг радиуса 1.

Обозначим сторону упомянутого 14-угольника через d . Легко видеть, что $d = 2 \sin \frac{\pi}{14} = 2 \cos \frac{3\pi}{7}$. Пусть

$$\varepsilon = \cos \frac{3\pi}{7} + i \sin \frac{3\pi}{7} = \cos \frac{3 \cdot 2\pi}{14} + i \sin \frac{3 \cdot 2\pi}{14},$$

тогда $d = \varepsilon + \varepsilon^{-1}$. Заметим, что ε является первообразным корнем 14-ой степени из 1 и поэтому корнем кругового многочлена

$$\begin{aligned} \Phi_{14}(x) &= \frac{x^{14} - 1}{\Phi_1 \Phi_2 \Phi_7} = \frac{(x^{14} - 1)\Phi_1}{\Phi_1 \Phi_2 (x^7 - 1)} = \frac{x^7 + 1}{x + 1} = \\ &= x^6 - x^5 + x^4 - x^3 + x^2 - x + 1. \end{aligned}$$

Поделив уравнение $x^6 - x^5 + x^4 - x^3 + x^2 - x + 1 = 0$ на x^3 , получим

$$\left(x^3 + \frac{1}{x^3}\right) - \left(x^2 + \frac{1}{x^2}\right) + \left(x + \frac{1}{x}\right) - 1 = 0,$$

или

$$\left(x + \frac{1}{x}\right)^3 - \left(x + \frac{1}{x}\right)^2 - 2\left(x + \frac{1}{x}\right) + 1 = 0,$$

т.е. $d^3 - d^2 - 2d + 1 = 0$. Это и есть исходное уравнение.

Упражнение 30. Решить уравнения:

- $z^5 + z^4 + z^3 + z^2 + z + 1 = 0$;
- $z^4 - 4z^3 + 7z^2 - 16z + 12 = 0$;

Упражнение 31. Решить уравнения:

- $z^4 + 6z^3 + 9z^2 + 10z = 0$;
- $z^4 + 2z^2 - 24z + 72 = 0$.

2.9. Вычисление сумм и произведений с помощью комплексных чисел

Пример 26. Выразить $\cos 5x$ через $\cos x$ и $\sin x$ ($x \in \mathbf{R}$).

Пусть $z = \cos x + i \sin x$. Найдем z^5 двумя способами. По формуле бинома имеем:

$$\begin{aligned} (\cos x + i \sin x)^5 &= \\ &= \cos^5 x + 5i \cos^4 x \sin x - 10 \cos^3 x \sin^2 x - \\ &\quad - 10i \cos^2 x \sin^3 x + 5 \cos x \sin^4 x + i \sin^5 x. \end{aligned}$$

С другой стороны, по формуле Муавра: $(\cos x + i \sin x)^5 = \cos 5x + i \sin 5x$. Приравняв правые части полученных равенств и выделяя действительную часть, получаем:

$$\cos 5x = \cos^5 x - 10 \cos^3 x \sin^2 x + 5 \cos x \sin^4 x.$$

Упражнение 32. Выразить через $\cos x$ и $\sin x$:

- $\cos 8x$;
- $\sin 7x$.

Упражнение 33. Выразить $\operatorname{tg} 6x$ через $\operatorname{tg} x$.

Пример 27. Вычислить $\sin x + \sin 2x + \dots + \sin nx$ ($x \in \mathbf{R}$).

Пусть $z = \sin x + i \cos x$, тогда рассматриваемая сумма есть мнимая часть выражения $z + z^2 + z^3 + \dots + z^n$, которое, используя формулы для суммы геометрической прогрессии, формулу Муавра и простейшие тригонометрические

формулы, преобразуем следующим образом:

$$\begin{aligned}
 z + z^2 + z^3 + \dots + z^n &= z \cdot \frac{z^n - 1}{z - 1} = \\
 &= (\cos x + i \sin x) \frac{\cos nx + i \sin nx - 1}{\cos x + i \sin x - 1} = \\
 &= (\cos x + i \sin x) \frac{2 \sin^2 \frac{nx}{2} - 2i \sin nx \cos \frac{x}{2}}{2 \sin^2 \frac{x}{2} - 2i \sin x \cos \frac{x}{2}} = \\
 &= (\cos x + i \sin x) \frac{-i \sin \frac{nx}{2} (i \sin \frac{nx}{2} + \cos \frac{nx}{2})}{-i \sin \frac{x}{2} (i \sin \frac{x}{2} + \cos \frac{x}{2})} = \\
 &= \frac{\left(\cos \frac{(n+1)x}{2} + i \sin \frac{(n+1)x}{2} \right)}{\sin \frac{x}{2}} \sin \frac{x}{2} \cdot \sin \frac{nx}{2}.
 \end{aligned}$$

Выделяя мнимую часть, получаем

$$\sin x + \sin 2x + \dots + \sin nx = \frac{\sin \frac{(n+1)x}{2} \sin \frac{nx}{2}}{\sin \frac{x}{2}}.$$

Выделяя действительную часть, можно получить

$$\cos x + \cos 2x + \dots + \cos nx = \frac{\cos \frac{(n+1)x}{2} \sin \frac{nx}{2}}{\sin \frac{x}{2}}.$$

Упражнение 34. Вычислить сумму

$$1 + a \cos \varphi + a^2 \cos 2\varphi + \dots + a^k \cos k\varphi.$$

Пример 28. Вычислить

$$\sin x + \binom{n}{1} \sin 2x + \dots + \binom{n}{n} \sin(n+1)x.$$

Пусть $z = \sin x + i \cos x$. Преобразуем выражение, мнимая часть которого есть рассматриваемая сумма:

$$\begin{aligned}
 z + \binom{n}{1} z^2 + \dots + \binom{n}{n} z^{n+1} &= \\
 &= z \left(1 + \binom{n}{1} z + \dots + \binom{n}{n} z^n \right) = \\
 &= z(1+z)^n = (\cos x + i \sin x)(1 + \cos x + i \sin x)^n = \\
 &= (\cos x + i \sin x) \left(2 \cos^2 \frac{x}{2} + i 2 \sin \frac{x}{2} \cos x \right)^n = \\
 &= 2^n (\cos x + i \sin x) \cos^n \frac{x}{2} \left(\cos \frac{x}{2} + i \sin \frac{x}{2} \right)^n = \\
 &= 2^n (\cos x + i \sin x) \cos^n \frac{x}{2} \left(\cos \frac{nx}{2} + i \sin \frac{nx}{2} \right).
 \end{aligned}$$

Мы последовательно применили формулу бинома Ньютона, тригонометрические формулы двойного угла и, наконец, использовали формулу Муавра. Теперь, выделяя мнимую часть в последнем выражении, получаем

$$\begin{aligned}
 \sin x + \binom{n}{1} \sin 2x + \dots + \binom{n}{n} \sin(n+1)x &= \\
 &= 2^n \cos^n \frac{x}{2} \sin \frac{(n+2)x}{2}.
 \end{aligned}$$

Пример 29. Вычислить

$$\sin^2 x + \sin^2 3x + \dots + \sin^2(2n-1)x.$$

Так как $\sin^2 y = \frac{1}{2} - \frac{1}{2} \cos 2y$, то

$$\begin{aligned}
 \sin^2 x + \sin^2 3x + \dots + \sin^2(2n-1)x &= \\
 &= \frac{1}{2}n - \frac{1}{2}(\cos 2x + \cos 6x + \dots + \cos 2(2n-1)x).
 \end{aligned}$$

Для вычисления суммы в круглых скобках положим $z = \cos x + i \sin x$ и рассмотрим сумму

$$\begin{aligned}
 z^2 + z^6 + \dots + z^{2(2n-1)} &= z^2 \cdot \frac{z^{2(2n-1)} - 1}{z^4 - 1} = \\
 &= z^{2n} \cdot \frac{z^{2n} - z^{-2n}}{z^2 - z^{-2}} = \\
 &= (\cos 2nx + i \sin 2nx) \frac{\sin 2nx}{\sin 2x}.
 \end{aligned}$$

Выделяя действительную часть, получаем

$$\cos 2x + \cos 6x + \dots + \cos 2(2n-1)x = \frac{\cos 2nx \sin 2nx}{2 \sin 2x} = \frac{\sin 4nx}{4 \sin 2x}.$$

Поэтому

$$\sin^2 x + \sin^2 3x + \dots + \sin^2(2n-1)x = \frac{n}{2} - \frac{\sin 4nx}{4 \sin 2x}.$$

Упражнение 35. Показать, что

$$\begin{aligned} \text{a) } \cos^2 x + \cos^2 2x + \dots + \cos^2 nx &= \frac{n}{2} + \frac{\cos(n+1)x \sin nx}{2 \sin x}; \\ \text{b) } \sin^2 x + \sin^2 2x + \dots + \sin^2 nx &= \frac{n}{2} - \frac{\cos(n+1)x \sin nx}{2 \sin x}. \end{aligned}$$

Пример 30. Вычислить:

$$\begin{aligned} \text{a) } \cos x + 2 \cos 2x + 3 \cos 3x + \dots + n \cos nx, \\ \text{b) } \sin x + 2 \sin 2x + 3 \sin 3x + \dots + n \sin nx. \end{aligned}$$

Пусть $z = \cos x + i \sin x$. Сумма а) совпадает с действительной частью выражения $z + 2z^2 + 3z^3 + \dots + nz^n$, сумма б) — с его мнимой частью. Для нахождения $z + 2z^2 + 3z^3 + \dots + nz^n$ рассмотрим произведение $(1-z)(z + 2z^2 + 3z^3 + \dots + nz^n)$, в котором, раскрывая скобки, получим

$$\begin{aligned} z + 2z^2 + 3z^3 + \dots + nz^n - z^2 - 2z^3 - \dots - (n-1)z^n - nz^{n+1} = \\ = z + z^2 + z^3 + \dots + z^n - nz^{n+1} = \\ = z \frac{1-z^{n+1}}{1-z} - nz^{n+1}. \end{aligned}$$

Следовательно,

$$\begin{aligned} z + 2z^2 + 3z^3 + \dots + nz^n &= \\ &= \frac{z}{1-z} \left(\frac{1-z^{n+1}}{1-z} - nz^{n+1} \right) = \\ &= \frac{z(1-z^{n+1}(1+n) + nz^{n+1})}{(1-z)^2}. \end{aligned}$$

Для отделения мнимой и действительной части в последнем выражении удобно ввести величину $\zeta = \cos \frac{x}{2} + i \sin \frac{x}{2}$, заметив, что $\zeta^2 = z$. Итак,

$$\begin{aligned} z + 2z^2 + 3z^3 + \dots + nz^n &= \\ &= \frac{1-z^{n+1}(1+n) + nz^{n+1}}{(\zeta^{-1}-\zeta)^2} = \\ &= \frac{1-z^{n+1}(1+n) + nz^{n+1}}{-2i \sin \frac{x}{2}} = \\ &= \frac{1-(1+n)(\cos nx + i \sin nx) + n(\cos(n+1)x + i \sin(n+1)x)}{-4 \sin^2 \frac{x}{2}}. \end{aligned}$$

Выделяя действительную и мнимую части, соответственно, получаем

$$\begin{aligned} \cos x + 2 \cos 2x + 3 \cos 3x + \dots + n \cos nx &= \\ &= \frac{(n+1) \cos nx - n \cos(n+1)x - 1}{4 \sin^2 \frac{x}{2}}, \\ \sin x + 2 \sin 2x + 3 \sin 3x + \dots + n \sin nx &= \\ &= \frac{(n+1) \sin nx - n \sin(n+1)x}{4 \sin^2 \frac{x}{2}}. \end{aligned}$$

Упражнение 36. Доказать, что

$$\begin{aligned} \text{a) } \cos \frac{\pi}{n} + \cos \frac{3\pi}{n} + \dots + \cos \frac{(2n-1)\pi}{n} &= 0; \\ \text{b) } \sin \frac{\pi}{n} + \sin \frac{3\pi}{n} + \dots + \sin \frac{(2n-1)\pi}{n} &= 0; \end{aligned}$$

Пример 31. Выразить $\sin^3 x$ через $\sin$ углов, кратных x .

Пусть $z = \cos x + i \sin x$, тогда нетрудно проверить, что $\sin x = \frac{z-z^{-1}}{2i}$ и, следовательно, по формуле бинома Ньютона,

$$\sin^3 x = \frac{z^3 - 3z^2z^{-1} + 3zz^{-2} - z^{-3}}{-8i} = \frac{z^3 - 3z + 3z^{-1} - z^{-3}}{-8i}.$$

Подставляя вместо z его выражение через x и приводя подобные, получаем

$$\sin^3 x = \frac{3 \sin x - \sin 3x}{4}.$$

Пример 32. Выразить $\cos^4 x \cdot \sin^3 x$ через $\sin$ кратных углов.

Пусть $z = \cos x + i \sin x$, тогда $\cos x = \frac{z+z^{-1}}{2}$, $\sin x = \frac{z-z^{-1}}{2i}$ и, следовательно,

$$\begin{aligned} \cos^4 x \cdot \sin^3 x &= \left(\frac{z+z^{-1}}{2} \right)^4 \left(\frac{z-z^{-1}}{2i} \right)^3 = \\ &= -\frac{1}{128i} (z^2 - z^{-2})^3 (z + z^{-1}) = \\ &= -\frac{1}{128i} (z^6 - 3z^2 + 3z^{-2} - z^{-6}) (z + z^{-1}) = \\ &= -\frac{1}{128i} (z^7 + z^5 - 3z^3 - 3z + 3z^{-1} + 3z^{-3} - z^{-5} - z^{-7}) = \\ &= -\frac{1}{128i} (z^7 - z^{-7} + z^5 - z^{-5} - 3(z^3 - z^{-3}) - 3(z - z^{-1})) = \\ &= -\frac{1}{64} (\sin 7x + \sin 5x - 3 \sin 3x - 3 \sin x). \end{aligned}$$

Упражнение 37. Выразить через тригонометрические функции кратных углов:

$$\text{a) } \sin^4 x;$$

b) $\sin^3 x \cos^5 x$.

Упражнение 38. Доказать, что

a) $\prod_{k=1}^{n-1} \sin \frac{\pi k}{2n} = \frac{\sqrt{n}}{2^{n-1}};$

b) $\prod_{k=1}^n \sin \frac{\pi k}{2n+1} = \frac{\sqrt{2n+1}}{2^n}.$

Пример 33. Вычислить

a) $1 - \binom{n}{2} + \binom{n}{4} - \binom{n}{6} + \dots$

b) $\binom{n}{1} - \binom{n}{3} + \binom{n}{5} - \binom{n}{7} + \dots$

Разлагая $(1+i)^n$ по формуле бинома:

$$(1+i)^n = 1 + \binom{n}{1}i + \binom{n}{2}i^2 + \binom{n}{3}i^3 + \dots + \binom{n}{n}i^n,$$

видим, что выражение а) есть действительная часть числа $(1+i)^n$, тогда как б) — его мнимая часть. Так как

$$(1+i)^n = 2^{n/2} \left(\cos \frac{n\pi}{4} + i \sin \frac{n\pi}{4} \right),$$

то

a) $1 - \binom{n}{2} + \binom{n}{4} - \binom{n}{6} + \dots = 2^{n/2} \cos \frac{n\pi}{4},$

b) $\binom{n}{1} - \binom{n}{3} + \binom{n}{5} - \binom{n}{7} + \dots = 2^{n/2} \sin \frac{n\pi}{4}.$

Пример 34. Вычислить

$$\binom{n}{1} - \frac{1}{3}\binom{n}{3} + \frac{1}{9}\binom{n}{5} - \frac{1}{27}\binom{n}{7} + \dots$$

Разложим $(1+i/\sqrt{3})^n$ по формуле бинома:

$$\left(1 + \frac{i}{\sqrt{3}}\right)^n = 1 + \binom{n}{1}\frac{i}{\sqrt{3}} + \binom{n}{2}\frac{-1}{3} + \binom{n}{3}\frac{-i}{3\sqrt{3}} + \binom{n}{4}\frac{1}{9} + \binom{n}{5}\frac{i}{9\sqrt{3}} + \dots$$

С другой стороны, по формуле Муавра:

$$\left(1 + \frac{i}{\sqrt{3}}\right)^n = \frac{2^n}{3^{n/2}} \left(\cos \frac{n\pi}{6} + i \sin \frac{n\pi}{6} \right).$$

Приравнявая мнимые части в обоих выражениях, получаем

$$\frac{1}{\sqrt{3}} \left(\binom{n}{1} - \frac{1}{3}\binom{n}{3} + \frac{1}{9}\binom{n}{5} - \frac{1}{27}\binom{n}{7} + \dots \right) = \frac{2^n}{3^{n/2}} \sin \frac{n\pi}{6}.$$

Отсюда

$$\binom{n}{1} - \frac{1}{3}\binom{n}{3} + \frac{1}{9}\binom{n}{5} - \frac{1}{27}\binom{n}{7} + \dots = \frac{2^n}{3^{(n-1)/2}} \sin \frac{n\pi}{6}.$$

Пример 35. Вычислить

$$1 + \binom{n}{3} + \binom{n}{6} + \dots$$

Рассмотрим выражение $(1+1)^n + (1+\varepsilon)^n + (1+\varepsilon^2)^n$, где $\varepsilon = \cos \frac{2\pi}{3} + i \sin \frac{2\pi}{3}$. Раскладывая все три бинома и вынося за скобки общие биномиальные коэффициенты, получаем:

$$\begin{aligned} (1+1+1) + \binom{n}{1}(1+\varepsilon+\varepsilon^2) + \binom{n}{2}(1+\varepsilon^2+\varepsilon^4) + \\ + \binom{n}{3}(1+1+1) + \binom{n}{4}(1+\varepsilon^4+\varepsilon^8) + \dots = \\ = 3 \left(1 + \binom{n}{3} + \binom{n}{6} + \dots \right). \end{aligned}$$

С другой стороны,

$$\begin{aligned} (1+1)^n + (1+\varepsilon)^n + (1+\varepsilon^2)^n = \\ = 2^n + \left(\cos \frac{\pi}{3} + i \sin \frac{\pi}{3} \right)^n + \left(\cos \frac{-\pi}{3} + i \sin \frac{-\pi}{3} \right)^n = \\ = 2^n + \cos \frac{n\pi}{3} + i \sin \frac{n\pi}{3} + \cos \frac{-n\pi}{3} + i \sin \frac{-n\pi}{3}. \end{aligned}$$

Следовательно,

$$1 + \binom{n}{3} + \binom{n}{6} + \dots = \frac{1}{3} \left(2^n + 2 \cos \frac{n\pi}{3} \right).$$

Упражнение 39. Вычислить суммы:

a) $\binom{n}{0} + \binom{n}{4} + \binom{n}{8} + \dots;$

b) $\binom{n}{1} + \binom{n}{5} + \binom{n}{9} + \dots;$

c) $\binom{n}{1} - \frac{1}{3}\binom{n}{5} + \frac{1}{9}\binom{n}{9} + \dots$

2.10. Числовые кольца и поля

Множество $A \subseteq \mathbf{C}$ *замкнуто* относительно сложения, если для любых двух произвольных a и b из A их сумма $a + b$ также принадлежит A . Аналогично определяется замкнутость по остальным арифметическим операциям.

Непустое числовое множество $R \subseteq \mathbf{C}$ называется (*числовым*) *кольцом*, если оно замкнуто относительно сложения, вычитания и умножения. Множество $\mathbf{N}$ замкнуто относительно сложения и умножения, однако не замкнуто относительно вычитания, и, следовательно, не является кольцом. Множества $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$, как не трудно видеть, — кольца.

Пусть R — некоторое кольцо, $a \in R$. Так как R замкнуто относительно вычитания, то $0 = a - a$ принадлежит R . Итак, любое кольцо содержит 0.

Числовое кольцо F называется (*числовым*) *полем*, если множество его ненулевых элементов не пусто и замкнуто относительно деления. Множество $\mathbf{Z}$ уже не является полем, однако $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$ — поля, в них деление на ненулевой элемент возможно и не выводит за пределы множества.

Теорема 2.10.1 (Минимальность поля $\mathbf{Q}$). *Для любого числового поля F верно, что $\mathbf{Q} \subseteq F$.*

Доказательство. Из определения поля следует, что $0 \in F$ и $a \in F$ для некоторого ненулевого a . Вследствие замкнутости F по операции деления имеем $1 = a/a \in F$. Далее, используя замкнутость по сложению, получаем, что числа $2 = 1 + 1$, $3 = 1 + 1 + 1$ и т.д. принадлежат F , следовательно $\mathbf{N} \subseteq F$. Справедливо также, что $-1 = 0 - 1 \in F$, $-2 = 0 - 2 \in F$ и т.д. Таким образом, $\mathbf{Z} \subseteq F$. И, наконец, $\frac{p}{q}$ тоже принадлежит F для любых $p \in \mathbf{N}$, $q \in \mathbf{Z}$, поэтому $\mathbf{Q} \subseteq F$. ■

Пример 36. Рассмотрим множество чисел вида $\alpha + \beta\sqrt{3}$, где $\alpha \in \mathbf{Z}$, $\beta \in \mathbf{Z}$. Оно является кольцом, но не является полем. Если же положить $\alpha \in \mathbf{Q}$, $\beta \in \mathbf{Q}$, то получим поле: замкнутость относительно деления следует из равенств

$$\frac{\alpha + \beta\sqrt{3}}{\gamma + \delta\sqrt{3}} = \frac{(\alpha + \beta\sqrt{3})(\gamma - \delta\sqrt{3})}{(\gamma + \delta\sqrt{3})(\gamma - \delta\sqrt{3})} = \frac{\alpha\gamma - 3\beta\delta}{\gamma^2 - 3\delta^2}. \quad (2.29)$$

Следует отметить, что равенство $\gamma + \delta\sqrt{3} = 0$ возможно лишь в случае, когда $\gamma = 0$ и $\delta = 0$. Действительно, если $\gamma + \delta\sqrt{3} = 0$ и $\gamma = 0$, то $\delta = 0$; если же $\gamma + \delta\sqrt{3} = 0$ и $\delta \neq 0$, то $\sqrt{3} = -\frac{\gamma}{\delta}$, что невозможно, так как $\sqrt{3}$ — иррациональное число. Итак, $\gamma + \delta\sqrt{3} \neq 0$, а значит, $\gamma - \delta\sqrt{3} \neq 0$ и поэтому домножение числителя и знаменателя дроби на это число в (2.29) допустимо. Выкладки в (2.29) есть ни что иное как *освобождение от иррациональности в знаменателе*, в идейном плане ничем не отличающееся от процедуры деления комплексных чисел с помощью домножения на сопряженное к знаменателю.

Линейные пространства

3.1. Аксиоматическое определение линейного пространства

Пусть V — непустое множество, F — поле. V называется *линейным* (или *векторным*) *пространством* над полем F , если

- 1) задано *правило сложения*, ставящее в соответствие любым двум элементам a, b из V элемент c из V , называемый *суммой* и обозначаемый $c = a + b$;
- 2) задано *правило умножения на число*, ставящее в соответствие каждому a из V и каждому α из F элемент d из V , обозначаемый $d = \alpha a$ (или $d = \alpha \cdot a$);
- 3) выполняются следующие свойства (*аксиомы линейного пространства*):
 - а) $\forall a, b \in V \ a + b = b + a$ (*коммутативность*),
 - б) $\forall a, b, c \in V \ a + (b + c) = (a + b) + c$ (*ассоциативность*),
 - в) $\exists q \ \forall a \in V \ a + q = a$ (q называется *нулевым* элементом),
 - г) $\forall a \in V \ \exists b \in V \ a + b = q$ (b называется элементом, *противоположным* к a),
 - д) $\forall a \in V \ 1 \cdot a = a$,
 - е) $\forall a \in V \ \forall \alpha, \beta \in F \ \alpha(\beta a) = (\alpha\beta a)$,
 - ж) $\forall a, b \in V \ \forall \alpha \in F \ \alpha(a + b) = \alpha a + \alpha b$,
 - з) $\forall a \in V \ \forall \alpha, \beta \in F \ (\alpha + \beta)a = \alpha a + \beta a$.

Элементы множества V называются *векторами*, элементы множества F называются *скалярами* (или, просто, числами). Далее, как правило, векторы мы будем обозначать латинскими буквами, скаляры — греческими.

3.2. Примеры линейных пространств

Пространство свободных векторов

Пространство радиус-векторов

Арифметическое пространство

Пространство многочленов

3.3. Простейшие следствия из аксиом

- 1) В любом линейном пространстве существует единственный нулевой элемент.

Предположим, что существует два нулевых элемента $q_1, q_2 \in V$, тогда положив в аксиоме 3в $a = q_1$, $q = q_2$, получаем $q_1 + q_2 = q_1$. С другой стороны, положив $a = q_2$, $q = q_1$ и воспользовавшись аксиомой 3а, получаем $q_1 + q_2 = q$. Приравняв правые части полученных равенств, получаем, $q_1 = q_2$.

Далее нулевой элемент (нулевой вектор, или нуль-вектор) обозначается o^1 .

- 2) Для любого вектора $a \in V$ существует единственный противоположный элемент.

Пусть b_1, b_2 — элементы пространства V , противоположные вектору a . По аксиоме 3г

$$b_1 + (a + b_2) = b_1 + o = b_1.$$

Аналогично,

$$(b_1 + a) + b_2 = o + b_2 = b_2.$$

По аксиоме 3г правые части приведенных равенств совпадают, поэтому $b_1 = b_2$.

Вектор, противоположный к вектору a будем обозначать $-a$.

- 3) Для любых a, b из V уравнение $a + x = b$ имеет и единственное решение.

Подстановкой легко убеждаемся, что $x = (-a) + b$ является решением уравнения. Для доказательства единственности предположим, что имеется два решения x_1, x_2 . Тогда $a + x_1 = a + x_2 = b$. Прибавляя ко всем частям равенства $(-a)$, получаем $(-a) + a + x_1 = (-a) + a + x_2$, откуда $x_1 = x_2$.

Назовем *разностью* векторов b и a решение уравнения $a + x = b$. Разность обозначим $b - a$. Мы установили, что $b - a = (-a) + b = b + (-a)$.

- 4) $0 = o$ для любого $a \in V$.

Имеем $a + 0a = 1a + 0a = (1 + 0)a = 1a = a$. Таким образом, $a + 0a = a$, откуда $0a = a - a = o$.

- 5) $\alpha o = o$ для любого $\alpha \in F$.

Пусть $a \in V$. Имеем $\alpha a + \alpha o = \alpha(a + o) = \alpha a$. Таким образом, $\alpha a + \alpha o = \alpha a$, откуда $\alpha o = \alpha a - \alpha a = o$.

- 6) $(-\alpha)a = -(\alpha a)$.

Необходимо показать, что вектор $(-\alpha)a$ является противоположным к αa . Действительно, $\alpha a + (-\alpha)a = (\alpha - \alpha)a = 0a = o$.

- 7) Как следствие получаем $(-1)a = -a$.

- 8) $(\alpha - \beta)a = \alpha a - \beta a$

- 9) $\alpha(a - b) = \alpha a - \alpha b$

- 10) Если $\alpha a = 0$, то $\alpha = 0$ или $a = 0$.

Если $\alpha = 0$, то утверждение справедливо. Пусть $\alpha \neq 0$, тогда $a = \frac{1}{\alpha}(\alpha a) = \frac{1}{\alpha} \cdot o = o$.

3.4. Линейные подпространства

Пусть V — линейное пространство над полем F . Множество $W \subseteq V$ называется (линейным) *подпространством*, если оно является пространством над F относительно операций сложения векторов и умножения их на числа, определенных в V .

Теорема 3.4.1 (Критерий линейного подпространства). *Для того, чтобы непустое подмножество W линейного пространства V являлось подпространством необходимо и достаточно, чтобы*

¹Нулевой вектор $o \in V$ следует отличать от числа $0 \in F$.

- 1) $a + b \in W$ для любых a, b из W (замкнутость относительно сложения векторов),
- 2) $\alpha a \in W$ для любого α из F и любого a из W (замкнутость относительно умножения векторов на числа).

ДОКАЗАТЕЛЬСТВО.

НЕОБХОДИМОСТЬ. Условия 1–2 включены в определение линейного пространства.

ДОСТАТОЧНОСТЬ. Проверим, что все 8 аксиом линейного пространства в условиях теоремы выполнены. Проверка аксиом 3а–3б, 3д–3з тривиальна.

Покажем, что $0 \in W$ (аксиома 3в). Пусть $\alpha = 0$, $a \in W$, тогда $0 = 0a = \alpha a \in W$ по условию 2.

Покажем, что $-a \in W$ для любого $a \in W$ (аксиома 3г). Пусть $\alpha = -1$, тогда $-a = (-1)a = \alpha a \in W$ по условию 2. ■

Пример 37. Подпространство радиус-векторов, концы которых лежат на прямой, проходящей через полюс.

Подпространство арифметических векторов с нулевой первой компонентой.

Подпространство четных (нечетных) многочленов.

Подпространство многочленов степени не выше n

3.5. Линейные комбинации и линейные оболочки

Линейной комбинацией векторов $a_1, \dots, a_n$ с коэффициентами $\alpha_1, \dots, \alpha_n$ называется выражение $\alpha_1 a_1 + \dots + \alpha_n a_n$.

Множество всех линейных комбинаций векторов $a_1, \dots, a_n$ называется *линейной оболочкой* этих векторов и обозначается $L(a_1, \dots, a_n)$. Формально:

$$L(a_1, \dots, a_n) = \{\alpha_1 a_1 + \dots + \alpha_n a_n : \alpha_1, \dots, \alpha_n \in F\}.$$

Теорема 3.5.1. Пусть $a_1, \dots, a_n$ — произвольные векторы из V . Тогда $L(a_1, \dots, a_n)$ — подпространство в V и для любого другого подпространства W , содержащего векторы $a_1, \dots, a_n$, справедливо $L(a_1, \dots, a_n) \subseteq W$.

ДОКАЗАТЕЛЬСТВО. Докажем сперва, что $L(a_1, \dots, a_n)$ — подпространство в V . Действительно, если $a = \alpha_1 a_1 + \dots + \alpha_n a_n$ и $b = \beta_1 a_1 + \dots + \beta_n a_n$, то $a + b = (\alpha_1 + \beta_1)a_1 + \dots + (\alpha_n + \beta_n)a_n$. Таким образом, $a + b \in L(a_1, \dots, a_n)$. Если $\alpha \in F$, то $\alpha a = (\alpha \alpha_1)a_1 + \dots + (\alpha \alpha_n)a_n$. Таким образом, $\alpha a \in L(a_1, \dots, a_n)$. Достаточные условия в критерии линейного подпространства выполнены.

Пусть теперь W — произвольное подпространство, содержащее векторы $a_1, \dots, a_n$. Тогда для любых чисел $\alpha_1, \dots, \alpha_n$ из F справедливо $\alpha_j a_j \in W$ ($j = 1, \dots, n$) (замкнутость относительно операции умножения на число) и $\alpha_1 a_1 + \dots + \alpha_n a_n \in W$ (замкнутость относительно операции сложения), т.е. $b \in W$ для произвольного $b \in L(a_1, \dots, a_n)$. ■

Говорят, что вектор b *линейно выражается* через систему $a_1, \dots, a_n$, если $b = \alpha_1 a_1 + \dots + \alpha_n a_n$ для некоторых $\alpha_1, \dots, \alpha_n$, иными словами, $b \in L(a_1, \dots, a_n)$. Говорят, что система $b_1, \dots, b_m$ *линейно выражается* через систему $a_1, \dots, a_n$, если для всех $i \in \{1, \dots, m\}$ вектор b_i линейно выражается через систему $a_1, \dots, a_n$.

Утверждение 3.5.1. Система $b_1, \dots, b_m$ линейно выражается через систему $a_1, \dots, a_n$ тогда и только тогда, когда $L(b_1, \dots, b_m) \subseteq L(a_1, \dots, a_n)$.

ДОКАЗАТЕЛЬСТВО.

НЕОБХОДИМОСТЬ. Если система $b_1, \dots, b_m$ линейно выражается через систему $a_1, \dots, a_n$, т.е. $b_i \in L(a_1, \dots, a_n)$ для всех $i \in \{1, \dots, m\}$, то по замкнутости $\beta_1 b_1 + \dots + \beta_m b_m \in L(a_1, \dots, a_n)$ для любых $\beta_1, \dots, \beta_m$ из F .

ДОСТАТОЧНОСТЬ. Пусть $L(b_1, \dots, b_m) \subseteq L(a_1, \dots, a_n)$. Так как $b_i \in L(b_1, \dots, b_m)$, то $b_i \in L(a_1, \dots, a_n)$ ($i = 1, \dots, m$). Т.е. Система $b_1, \dots, b_m$ линейно выражается через систему $a_1, \dots, a_n$. ■

Следствие 3.5.1. Если система $b_1, \dots, b_m$ линейно выражается через систему $a_1, \dots, a_n$, а система $a_1, \dots, a_n$ линейно выражается через систему $b_1, \dots, b_m$, то система $a_1, \dots, a_n$ линейно выражается через систему $b_1, \dots, b_m$.

Системы $a_1, \dots, a_n$ и $b_1, \dots, b_m$ называются *эквивалентными*, если $a_1, \dots, a_n$ линейно выражаются через $b_1, \dots, b_m$, а $b_1, \dots, b_m$ линейно выражаются через $a_1, \dots, a_n$.

Следствие 3.5.2. Системы $a_1, \dots, a_n$ и $b_1, \dots, b_m$ эквивалентны тогда и только тогда, когда $L(a_1, \dots, a_n) = L(b_1, \dots, b_m)$.

Рефлексивность, симметричность, транзитивность.

3.6. Линейная зависимость векторов

Линейная комбинация $\alpha_1 a_1 + \dots + \alpha_n a_n$ векторов $a_1, \dots, a_n \in V$ называется *тривиальной*, если $\alpha_1 = \dots = \alpha_n = 0$. Говорят, что система векторов $a_1, \dots, a_n$ *линейно зависима*, если существует нетривиальная комбинация этих векторов, равная нулю, иными словами, если найдутся числа $\alpha_1, \dots, \alpha_n \in F$, такие, что $\alpha_j \neq 0$ для некоторого $j \in \{1, \dots, n\}$ и

$$\alpha_1 a_1 + \dots + \alpha_n a_n = 0. \quad (3.1)$$

В противном случае система называется *линейно независимой*. В силу важности вводимых терминов переформулируем определение линейной независимости. Система $a_1, \dots, a_n$ называется линейно независимой, если равенство (3.1) возможно лишь в случае $\alpha_1 = \dots = \alpha_n = 0$.

Утверждение 3.6.1 (Система из одного вектора). Система, состоящая из одного вектора линейно зависима тогда и только тогда, когда вектор нулевой.

Доказательство. Для системы, состоящей лишь из одного вектора a , равенство (3.1) примет вид $\alpha a = 0$, откуда $\alpha = 0$, но тогда комбинация тривиальная, или $a = 0$, что и утверждается. ■

Утверждение 3.6.2. Система из двух векторов линейно зависима тогда и только тогда, когда векторы пропорциональны.

Доказательство. Равенство (3.1) принимает вид $\alpha_1 a_1 + \alpha_2 a_2 = 0$, причем $\alpha_1 \neq 0$ или(и) $\alpha_2 \neq 0$. В первом случае имеем $a_1 = (-\alpha_2/\alpha_1)a_2$, во втором имеем $a_2 = (-\alpha_1/\alpha_2)a_1$. ■

Следствие 3.6.1. Два вектора геометрического пространства V_2 (V_3) линейно зависимы тогда и только тогда, когда они коллинеарны.

Утверждение 3.6.3. Если подсистема некоторой системы линейно зависима, то и вся система линейно зависима.

Доказательство. Пусть $a_1, \dots, a_n$ — исходная система, $a_{j_1}, \dots, a_{j_m}$ — линейно зависимая подсистема ($1 \leq j_1 < j_2 < \dots < j_m \leq n$). По определению линейной зависимости найдутся такие $\alpha_{j_1}, \dots, \alpha_{j_m}$, не все равные нулю, что

$$\alpha_{j_1} a_{j_1} + \dots + \alpha_{j_m} a_{j_m} = 0.$$

В последнем равенстве добавим к левой части тривиальную комбинацию векторов, не вошедших в подсистему:

$$0a_{i_1} + \dots + 0a_{i_{n-m}},$$

где

$$\{i_1, \dots, i_{n-m}\} = \{1, \dots, n\} \setminus \{j_1, \dots, j_m\}.$$

Полученная комбинация будет нулевой, но не тривиальной. ■

Следствие 3.6.2. Любая подсистема линейно независимой системы линейно независима.

Теорема 3.6.1 (Критерий линейной зависимости). Система $a_1, \dots, a_n$, где $n \geq 2$, линейно зависима тогда и только тогда, когда

$$a_j = L(a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n)$$

для некоторого $j \in \{1, \dots, n\}$.

ДОКАЗАТЕЛЬСТВО.

НЕОБХОДИМОСТЬ. Пусть система $a_1, \dots, a_n$ линейно зависима, тогда $\alpha_1 a_1 + \dots + \alpha_n a_n = 0$, причем $\alpha_j \neq 0$ для некоторого j , откуда

$$a_j = \left(-\frac{\alpha_1}{\alpha_j}\right) a_1 + \dots + \left(-\frac{\alpha_{j-1}}{\alpha_j}\right) a_{j-1} + \left(-\frac{\alpha_{j+1}}{\alpha_j}\right) a_{j+1} + \dots + \left(-\frac{\alpha_n}{\alpha_j}\right) a_n.$$

ДОСТАТОЧНОСТЬ. Пусть

$$a_j = \alpha_1 a_1 + \dots + \alpha_{j-1} a_{j-1} + \alpha_{j+1} a_{j+1} + \dots + \alpha_n a_n$$

для некоторых $\alpha_1, \dots, \alpha_{j-1}, \alpha_{j+1}, \alpha_n$. Тогда

$$\alpha_1 a_1 + \dots + \alpha_{j-1} a_{j-1} + \alpha_j a_j + \alpha_{j+1} a_{j+1} + \dots + \alpha_n a_n = 0$$

где $\alpha_j = -1$. Получили нулевую нетривиальную (так как $\alpha_j = -1 \neq 0$) комбинацию. ■

Теорема 3.6.2 (Усиленный критерий линейной зависимости). Система $a_1, \dots, a_n$ линейно зависима тогда и только тогда, когда $a_1 = 0$ или $a_j = L(a_1, \dots, a_{j-1})$ для некоторого $j \in \{1, \dots, j-1, j+1, \dots, n\}$.

ДОКАЗАТЕЛЬСТВО.

НЕОБХОДИМОСТЬ. Пусть система $a_1, \dots, a_n$ линейно зависима, тогда $\alpha_1 a_1 + \dots + \alpha_n a_n = 0$. Пусть j — максимальный индекс, такой, что $\alpha_j \neq 0$, тогда $\alpha_1 a_1 + \dots + \alpha_j a_j = 0$, откуда

$$a_j = \left(-\frac{\alpha_1}{\alpha_j}\right) a_1 + \dots + \left(-\frac{\alpha_{j-1}}{\alpha_j}\right) a_{j-1}.$$

ДОСТАТОЧНОСТЬ. Если $n \geq 2$, то воспользуемся предыдущим критерием. Если $n = 1$, то $a_1 = 0$. ■

Следствие 3.6.3. Если система $a_1, \dots, a_n$ линейно независима, а система $a_1, \dots, a_n, b_1, \dots, b_m$ — линейно зависима, то найдется такое $j \in 1, \dots, m$, что

$$b_j = L(a_1, \dots, a_n, b_1, \dots, b_{j-1}).$$

Лемма 3.6.1 (Теорема о замене). Если система $a_1, \dots, a_n$ линейно выражается через систему $b_1, \dots, b_m$, причем $n \geq m$, то первая система линейно зависима.

ДОКАЗАТЕЛЬСТВО. Предположим противное. Пусть в условиях теоремы первая система линейно независима. Обозначим через B_0 систему $b_1, \dots, b_m$.

Так как вектор a_1 линейно выражается через систему B_0 , то объединенная система a_1, B_0 линейно зависима и эквивалентна B_0 . По предыдущему следствию в новой системе найдется такой вектор b_{i_1} из B_0 , что $b_{i_1} \in L(a_1, B_0 \setminus b_{i_1})$, поэтому B_0 эквивалентна системе a_1, B_1 , где $B_1 = B_0 \setminus b_{i_1}$.

Так как вектор a_2 линейно выражается через систему B_0 , эквивалентную a_1, B_1 , то объединенная система a_1, a_2, B_1 линейно зависима и эквивалентна B_0 . По предыдущему следствию в новой системе найдется такой вектор b_{i_2} из B_1 , что $b_{i_2} \in L(a_1, a_2, B_1 \setminus b_{i_2})$, поэтому B_0 эквивалентна системе a_1, a_2, B_2 , где $B_2 = B_0 \setminus b_{i_1}, b_{i_2}$.

Продолжая эти рассуждения далее (можно провести индукцию), мы получим цепочку эквивалентных систем

$$B_0 \sim a_1, B_1 \sim a_1, a_2, B_2 \sim \dots \sim a_1, \dots, a_m, B_m,$$

где $B_j = B_{j-1} \setminus b_{i_j}$ ($j = 1, \dots, m$). В частности, B_m — пустая система. Поэтому система $B_0 = b_1, \dots, b_m$ эквивалентна системе $a_1, \dots, a_m$. Следовательно, вся первая система $a_1, \dots, a_n$, выражающаяся через систему B_0 , выражается через свою подсистему $a_1, \dots, a_n$, следовательно, она линейно зависима, что противоречит условию. ■

Следствие 3.6.4. Эквивалентные линейно независимые системы содержат одинаковое число векторов.

Базой системы векторов называется ее любая линейно независимая подсистема, эквивалентная исходной системе.

Следствие 3.6.5. *Любая ненулевая система векторов $a_1, \dots, a_m$ имеет базу.*

ДОКАЗАТЕЛЬСТВО. Пусть a_{i_1} — некоторый ненулевой вектор из V . Система, состоящая из одного вектора a_{i_1} , линейно независима. Поэтому, если она эквивалентна исходной системе, то является ее базой. В противном случае найдется вектор $a_{i_2} \notin L(a_{i_1})$. По усиленному критерию линейной зависимости система a_{i_1}, a_{i_2} — независимая. Если она эквивалентна исходной системе, то является ее базой. В противном случае найдется вектор $a_{i_3} \notin L(a_{i_1}, a_{i_2})$ и т.д. Описанный процесс оборвется, так как исходная система конечна. ■

Следствие 3.6.6. *Для любой системы число векторов в произвольной базе одинаково.*

Число векторов в базе называется *рангом* системы.

3.7. Базис линейного пространства

Система векторов $a_1, \dots, a_n$ линейного пространства V называется *полной*, если $V = L(a_1, \dots, a_n)$. Линейное пространство, в котором существует полная система, называется *конечномерным*.

Базисом пространства V называется полная линейно независимая система.

Утверждение 3.7.1. *В любом ненулевом конечномерном пространстве V существует базис.*

ДОКАЗАТЕЛЬСТВО. Пространство V конечномерное, поэтому $V = L(a_1, \dots, a_n)$ для некоторых $a_1, \dots, a_n$. В качестве базиса пространства возьмем базу системы $a_1, \dots, a_n$. ■

Утверждение 3.7.2. *Для любого пространства число векторов в любом базисе постоянно.*

ДОКАЗАТЕЛЬСТВО. Число векторов в базисе конечномерного пространства V называется *размерностью пространства V* и обозначается $\dim V$. Размерность нулевого пространства считается равной нулю. ■

Утверждение 3.7.3. *Для того, чтобы линейно независимая система $a_1, \dots, a_n$ была базисом, необходимо и достаточно, чтобы любая система $b_1, \dots, b_{n+1}$ из $n+1$ вектора была линейно зависимой.*

ДОКАЗАТЕЛЬСТВО.

НЕОБХОДИМОСТЬ. Система $b_1, \dots, b_{n+1}$ выражается через линейно независимую систему $a_1, \dots, a_n$, поэтому $b_1, \dots, b_{n+1}$ зависима по лемме о замене.

ДОСТАТОЧНОСТЬ. Пусть b — произвольный вектор из V . Рассмотрим систему $a_1, \dots, a_n, b$. По условию она линейно зависима. Теперь из усиленного критерия линейной зависимости следует, что $b \in L(a_1, \dots, a_n)$, т.е. система $a_1, \dots, a_n$ — полная. ■

ЗАМЕЧАНИЕ. Утверждение 3 позволяет дать следующее определение базиса, эквивалентное исходному: базис — это наибольшая (по мощности) линейно независимая система.²

Утверждение 3.7.4. *Для того, чтобы полная система $a_1, \dots, a_n$ была базисом, необходимо и достаточно, чтобы либо $n = 1$ и $a_1 \neq 0$, либо любая система $b_1, \dots, b_{n-1}$ из $n-1$ вектора не была бы полной.*

ДОКАЗАТЕЛЬСТВО.

НЕОБХОДИМОСТЬ. Пусть система $b_1, \dots, b_{n-1}$ полная, тогда линейно независимая система $a_1, \dots, a_n$ выражается через $b_1, \dots, b_{n-1}$, что противоречит лемме о замене.

ДОСТАТОЧНОСТЬ. Ни для какого $j = \{1, \dots, n\}$ система $a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n$ не является базисом, поэтому $a_j \notin \{a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n\}$. Линейная независимость системы векторов $a_1, \dots, a_n$ следует теперь из критерия линейной независимости. ■

ЗАМЕЧАНИЕ. Утверждение 4 позволяет дать следующее определение базиса, эквивалентное исходному: базис — это наименьшая (по мощности) полная система.³

² Такое определение не означает, что базис пространства определен единственным образом

³См. примечание 2

Утверждение 3.7.5. Пусть $\dim V = n$ и система $a_1, \dots, a_n$ — линейно независима, тогда она полная (и, следовательно, базис).

ДОКАЗАТЕЛЬСТВО. Так как $\dim V = n$, то любая система из $n + 1$ вектора линейно независима. Поэтому $a_1, \dots, a_n$ полная по теореме. ■

Утверждение 3.7.6. Пусть $\dim V = n$ и система $a_1, \dots, a_n$ — полная, тогда она линейно независима (и, следовательно, базис).

ДОКАЗАТЕЛЬСТВО. Так как $\dim V = n$, то любая система из $n - 1$ вектора полная. Поэтому $a_1, \dots, a_n$ линейно независима по теореме. ■

3.7.1. Размерность и базис арифметического пространства

3.7.2. Размерность и базис пространства геометрических векторов

3.7.3. Размерность пространства многочленов

3.8. Координаты векторов

Пусть $e_1, \dots, e_n$ — базис пространства V над полем F . Тогда для произвольного вектора a из V найдутся такие числа $\alpha_1, \dots, \alpha_n$ из F , что $a = \alpha_1 e_1 + \dots + \alpha_n e_n$. Коэффициенты $\alpha_1, \dots, \alpha_n$ разложения вектора по базису называются *координатами вектора*.

Утверждение 3.8.1. В произвольном базисе координаты вектора определены однозначно.

ДОКАЗАТЕЛЬСТВО. Пусть $e_1, \dots, e_n$ — базис пространства V . Предположим, что для вектора a нашлось два разложения по базису

$$a = \alpha_1 e_1 + \dots + \alpha_n e_n = \beta_1 e_1 + \dots + \beta_n e_n.$$

Откуда получаем

$$(\alpha_1 - \beta_1)e_1 + \dots + (\alpha_n - \beta_n)e_n = 0.$$

Так как система $e_1, \dots, e_n$ — линейно независимая, то все коэффициенты в полученной линейной комбинации равны 0, откуда $\alpha_j = \beta_j$ ($j = 1, \dots, n$). ■

Координаты удобно записывать в столбец, который мы будем называть *координатным столбцом* и обозначать $[a]_e$. Таким образом, если $a = \alpha_1 e_1 + \dots + \alpha_n e_n$, то

$$[a]_e = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Координатный столбец можно рассматривать как вектор арифметического пространства F^n . Имея в виду эту интерпретацию, получаем

Утверждение 3.8.2. Для произвольных векторов a, b из V и произвольного α из F справедливо

$$\begin{aligned} [a + b]_e &= [a]_e + [b]_e, \\ [\alpha a]_e &= \alpha [a]_e. \end{aligned}$$

ДОКАЗАТЕЛЬСТВО. Пусть

$$\begin{aligned} a &= \alpha_1 e_1 + \dots + \alpha_n e_n, \\ b &= \beta_1 e_1 + \dots + \beta_n e_n. \end{aligned}$$

Тогда

$$a + b = (\alpha_1 + \beta_1)e_1 + \dots + (\alpha_n + \beta_n)e_n,$$

т.е. $[a + b]_e = [a]_e + [b]_e$. Второе равенство доказывается аналогично. ■

ЗАМЕЧАНИЕ. Словесная формулировка утверждения выглядит следующим образом: при сложении векторов их координаты складываются, при умножении на число умножаются на это число.

3.9. Изоморфизм линейных пространств

Пусть пространства V и V' заданы над одним и тем же полем F . Биекция $\varphi : V \rightarrow V'$ называется *изоморфизмом*, если для любых векторов a, b из V и любого α из F справедливо

- 1) $\varphi(a + b) = \varphi a + \varphi b$,
- 2) $\varphi(\alpha a) = \alpha(\varphi a)$.

Свойство 3.9.1. *Отображение, обратное к изоморфизму является изоморфизмом.*

ДОКАЗАТЕЛЬСТВО. Отображение $\varphi^{-1} : V' \rightarrow V$, обратное к изоморфизму $\varphi : V \rightarrow V'$, является биекцией. Осталось показать выполнение свойств 1, 2 для отображения φ^{-1} . Пусть a', b' — произвольные векторы из V' , тогда $\varphi^{-1}a' \in V$ и $\varphi^{-1}b' \in V$. Из определения изоморфизма φ имеем

$$\begin{aligned}\varphi(\varphi^{-1}a' + \varphi^{-1}b') &= \varphi\varphi^{-1}a' + \varphi\varphi^{-1}b', \\ \varphi(\varphi^{-1}a' + \varphi^{-1}b') &= a' + b'.\end{aligned}$$

Применяя к обеим частям последнего равенства отображение φ^{-1} , получаем

$$\varphi^{-1}a' + \varphi^{-1}b' = \varphi^{-1}a' + \varphi^{-1}b'.$$

Равенство $\varphi^{-1}(\alpha a) = \alpha(\varphi^{-1}a)$ доказывается аналогично. ■

Свойство 3.9.2. *Образом нулевого вектора пространства V при изоморфизме φ является нулевой вектор пространства V' , т.е.⁴ $\varphi o = o$.*

ДОКАЗАТЕЛЬСТВО. Из определения изоморфизма

$$\varphi o = \varphi(0a) = 0\varphi a = o.$$
■

Свойство 3.9.3. *При изоморфизме линейно зависящая система отображается в линейно зависящую систему.*

ДОКАЗАТЕЛЬСТВО. Пусть система $a_1, \dots, a_n$ линейно зависящая, тогда найдутся такие числа $\alpha_1, \dots, \alpha_n$, что

$$\alpha_1 a_1 + \dots + \alpha_n a_n = o, \tag{3.2}$$

причем $\alpha_j \neq 0$ для некоторого $j \in \{1, \dots, n\}$. Применим отображение φ к обеим частям равенства (3.2). Из определения изоморфизма получаем

$$\alpha_1 (\varphi a_1) + \dots + \alpha_n (\varphi a_n) = o.$$

Нами получена нетривиальная нулевая комбинация векторов $\varphi a_1, \dots, \varphi a_n$, значит векторы линейно зависимы. ■

Свойство 3.9.4. *При изоморфизме линейно независимая система отображается в линейно независимую систему.*

ДОКАЗАТЕЛЬСТВО. Утверждение является следствием свойств 1, 3. ■

Свойство 3.9.5. *При изоморфизме $\varphi : V \rightarrow V'$ базис пространства V отображается в базис пространства V' .*

⁴Заметим, что в следующем равенстве o в левой части — вектор пространства V , тогда как o в правой части — вектор пространства V'

ДОКАЗАТЕЛЬСТВО. По свойству 3 базис отображается в линейно независимую систему. При этом по свойству 4 любая линейно зависимая система отображается в линейно зависимую. Таким образом базис отображается в наибольшую линейно независимую систему, т.е. в базис. ■

Пространства V и V' называются *изоморфными*, если между ними существует изоморфизм $\varphi : V \rightarrow V'$. Если V и V' изоморфны, то пишут $V \cong V'$.

Утверждение 3.9.1. *Отношение изоморфности пространств является отношением эквивалентности.*

ДОКАЗАТЕЛЬСТВО.

РЕФЛЕКСИВНОСТЬ. Очевидно, что тождественное преобразование $\varepsilon : V \rightarrow V'$, определяемое формулой $\varepsilon a = a$ является изоморфизмом. Таким образом, $V \cong V$.

СИММЕТРИЧНОСТЬ. Если $V \cong V'$, то $V' \cong V$ по свойству 1.

ТРАНЗИТИВНОСТЬ. Докажем, что если $V \cong V'$ и $V' \cong V''$, то $V \cong V''$. Действительно, если φ и φ' — изоморфизмы из V в V' и из V' в V'' , то произведение изоморфизмов $\psi = \varphi\varphi'$, определяемой формулой $\psi a = \varphi(\varphi' a)$ есть изоморфизм V в V'' . ■

Теорема 3.9.1 (Критерий изоморфности пространств). *Для того, чтобы пространства V и V' , заданные над одним и тем же полем F были изоморфны необходимо и достаточно, чтобы $\dim V = \dim V'$.*

ДОКАЗАТЕЛЬСТВО.

НЕОБХОДИМОСТЬ. Необходимость следует из свойства 5.

ДОСТАТОЧНОСТЬ. Пусть $\dim V = n$. Покажем, что $V \cong F^n$. Действительно, из утверждения 2 предыдущего раздела, следует, что отображение $\varphi : V \rightarrow F^n$, ставящее в соответствие вектору из V столбец его координат в некотором фиксированном базисе, является изоморфизмом. Аналогично доказывается, что $V' \cong F^n$. Так как изоморфность — симметричное и транзитивное отношение, то $V \cong V'$. ■

3.10. Размерность подпространства

Утверждение 3.10.1. *Для любого ненулевого подпространства V_1 конечномерного линейного пространства V существует базис.*

ДОКАЗАТЕЛЬСТВО. Пусть a_1 — некоторый ненулевой вектор из V_1 . Система, состоящая из одного вектора a_1 , линейно независима. Поэтому, если $V_1 = L(a_1)$, то система образует базис пространства V_1 . В противном случае в V_1 найдется вектор $a_2 \notin L(a_1)$. По усиленному критерию линейной зависимости система a_1, a_2 — независимая. Поэтому, если $V_1 = L(a_1, a_2)$, то система образует базис пространства V_1 . В противном случае в V_1 найдется вектор $a_3 \notin L(a_1, a_2)$ и т.д. Описанный процесс оборвется по крайней мере через n шагов, где $n = \dim V$, так как любая система из $n + 1$ вектора линейно зависима. ■

Утверждение 3.10.2. *Пусть V_1, V_2 — конечномерные подпространства, причем $V_1 \subseteq V_2$, тогда $\dim V_1 \leq \dim V_2$.*

ДОКАЗАТЕЛЬСТВО. Пусть $a_1, \dots, a_l, b_1, \dots, b_m$ — базисы подпространств V_1 и V_2 соответственно. Так как $V_1 \subseteq V_2$, то первая система линейно выражается через вторую и поэтому по лемме о замене $l \leq m$. ■

Утверждение 3.10.3. *Пусть V_1, V_2 — конечномерные подпространства, причем $V_1 \subseteq V_2$ и $\dim V_1 = \dim V_2$, тогда $V_1 = V_2$.*

ДОКАЗАТЕЛЬСТВО. Базис $a_1, \dots, a_l$ подпространства V_1 является линейно независимой системой векторов из V_2 . Так как $l = \dim V_1 = \dim V_2$, то $a_1, \dots, a_l$ является базисом подпространства V_2 и, следовательно, $V_1 = L(a_1, \dots, a_l) = V_2$. ■

3.11. Сумма и пересечение подпространств

Суммой подпространств V_1 и V_2 пространства V называется множество

$$V_1 + V_2 = \{x = y + z : y \in V_1, z \in V_2\}.$$

Пересечение подпространств понимается в обычном теоретико-множественном смысле:

$$V_1 \cap V_2 = \{x : x \in V_1, x \in V_2\}.$$

Утверждение 3.11.1. Сумма подпространств одного и того же пространства является подпространством.

ДОКАЗАТЕЛЬСТВО. Так как $o \in V_1$, $o \in V_2$, то $o + o \in V_1 + V_2$ и $V_1 + V_2 \neq \emptyset$. Докажем замкнутость относительно сложения векторов и умножения их на числа. Пусть $x \in V_1 + V_2$ и $x' \in V_1 + V_2$, тогда найдутся такие векторы y, y' из V_1 и z, z' из V_2 , что $x = y + z$, $x' = y' + z'$. Так как $y + y' \in V_1$ и $z + z' \in V_2$, то $x + x' = (y + y') + (z + z') \in V_1 + V_2$. Так как $\alpha y \in V_1$ и $\alpha z \in V_2$, то $\alpha x = \alpha y + \alpha z \in V_1 + V_2$. ■

Утверждение 3.11.2. Пересечение подпространств одного и того же пространства является подпространством.

ДОКАЗАТЕЛЬСТВО. Так как $o \in V_1$, $o \in V_2$, то $o \in V_1 \cap V_2$ и $V_1 \cap V_2 \neq \emptyset$. Докажем замкнутость относительно сложения векторов и умножения их на числа. Пусть $x \in V_1 \cap V_2$ и $x' \in V_1 \cap V_2$. Так как $x + x' \in V_1$ и $x + x' \in V_2$, то $x + x' \in V_1 \cap V_2$. Так как $\alpha x \in V_1$ и $\alpha x \in V_2$, то $\alpha x \in V_1 \cap V_2$. ■

Лемма 3.11.1. Любую линейно независимую систему векторов можно дополнить до базиса конечномерного пространства.

ДОКАЗАТЕЛЬСТВО. Утверждение доказывается аналогично доказательству утверждения 1 из предыдущего раздела. ■

Теорема 3.11.1 (Размерность суммы подпространств). Если V_1, V_2 — конечномерные подпространства пространства V , то

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

ДОКАЗАТЕЛЬСТВО. Пусть $a_1, \dots, a_k$ — базис подпространства $V_1 \cap V_2$. Векторами $b_1, \dots, b_l$ дополним систему $a_1, \dots, a_k$ до базиса пространства V_1 . Векторами $c_1, \dots, c_m$ дополним систему $a_1, \dots, a_k$ до базиса пространства V_2 . Для доказательства теоремы покажем, что система

$$a_1, \dots, a_k, b_1, \dots, b_l, c_1, \dots, c_m$$

образует базис пространства $V_1 + V_2$.

Легко проверить полноту системы. Для доказательства ее линейной независимости рассмотрим нулевую линейную комбинацию:

$$\underbrace{\alpha_1 a_1 + \dots + \alpha_k a_k + \beta_1 b_1 + \dots + \beta_l b_l}_{-c \in V_1} + \underbrace{\gamma_1 c_1 + \dots + \gamma_m c_m}_{c \in V_2} = 0. \quad (3.3)$$

Пусть

$$c = \gamma_1 c_1 + \dots + \gamma_m c_m. \quad (3.4)$$

Легко видеть, что $c \in V_2$ и тогда по (3.3) имеем $c \in V_1$. Следовательно, $c \in V_1 \cap V_2$ и поэтому найдутся такие $\delta_1, \dots, \delta_k$, что

$$c = \delta_1 a_1 + \dots + \delta_k a_k. \quad (3.5)$$

Из (3.4, 3.5) получаем два разложения вектора c по базису пространства V_2 :

$$\begin{aligned} c &= 0a_1 + \dots + 0a_k + \gamma_1 c_1 + \dots + \gamma_m c_m, \\ c &= \delta_1 a_1 + \dots + \delta_k a_k + 0c_1 + \dots + 0c_m. \end{aligned}$$

Так как разложение по базису единственно, то все коэффициенты в этих разложениях нулевые, поэтому $c = 0$ и, следовательно, нулевая линейная комбинация (3.3) — тривиальная. ■

где $a_i \in V_i$ ($i = 1, \dots, s$) и $a_{i'} \neq 0$ для некоторого i' . Каждый вектор a_i разложим по базису $a_{i1}, \dots, a_{in_i}$:

$$o = \sum_{j=1}^{n_1} \alpha_{1j} a_{1j} + \dots + \sum_{j=1}^{n_s} \alpha_{sj} a_{sj}.$$

В приведенной линейной комбинации найдутся ненулевые коэффициенты⁵, поэтому система A — линейно зависима, что противоречит условию. ■

Следствие 3.12.1. Сумма $V_1 + V_2$ прямая тогда и только тогда, когда $V_1 \cap V_2 = \{0\}$.

Доказательство. Утверждение следует из доказанного критерия и теоремы о размерности суммы подпространств. ■

3.12.1. Проекция вектора на подпространство

Пусть все пространство V раскладывается в прямую сумму подпространств: $V = V_1 \dot{+} V_2$. Тогда для произвольного $x \in V$ определяются единственным образом векторы $y \in V_1$, $z \in V_2$, такие, что $x = y + z$. Вектор y называется *проекцией* вектора x на подпространство V_1 параллельно V_2 . Обозначение: $y = \text{pr}_{V_1 \| V_2} x$. Очевидно, что $z = \text{pr}_{V_2 \| V_1} x$.

3.13. Изменение координат при замене базиса

Пусть $e_1, \dots, e_n$ и $f_1, \dots, f_n$ — два базиса пространства V . Каждый вектор второго базиса разложим по первому базису:

$$f_j = \alpha_{1j} e_1 + \dots + \alpha_{nj} e_n \quad (j = 1, \dots, n).$$

Найдем связь координат произвольного вектора x в этих базисах. Пусть

$$x = x_1 e_1 + \dots + x_n e_n = x'_1 f_1 + \dots + x'_n f_n.$$

Имеем

$$x = \sum_{j=1}^n x'_j f_j = \sum_{j=1}^n x'_j \sum_{i=1}^n \alpha_{ij} e_i = \sum_{i=1}^n \underbrace{\left(\sum_{j=1}^n \alpha_{ij} x'_j \right)}_{x_i} e_i$$

Так как координаты в любом фиксированном базисе определяются однозначно, то

$$x_i = \sum_{j=1}^n \alpha_{ij} x'_j.$$

⁵ среди чисел $\alpha_{i'1}, \dots, \alpha_{i'n_{i'}}$

Матрицы и системы линейных уравнений

4.1. Метод Гаусса решения систем линейных уравнений

Рассмотрим произвольную систему m линейных уравнений n неизвестными над полем F :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1; \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2; \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m, \end{cases} \quad (4.1)$$

где $a_{ij} \in F$, $b_i \in F$ ($i = 1, \dots, m$; $j = 1, \dots, n$). Решением или частным решением системы (4.1) называется набор $x = (x_1, \dots, x_n)^T \in F^n$ значений неизвестных x_j , при подстановке которых в систему получаем m верных тождеств. Множество всех частных решений системы будем называть *общим решением*. Если множество решений пусто, то система называется *несовместной*. Две системы назовем *эквивалентными* если их множества решений совпадают.

Коэффициенты левой части системы (4.1) запишем в матрицу A , коэффициенты правой части системы — в столбец b :

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Матрицу A будем называть *основной матрицей* системы. Блочную матрицу $(A|b)$ будем называть *расширенной матрицей* системы.

Основная идея метода Гаусса решения системы линейных уравнений — ее упрощение с помощью серии преобразований строк расширенной матрицы, переводящих систему уравнений в эквивалентную.

Пример 38. Решим систему линейных уравнений

$$\begin{cases} x_1 + 2x_2 - x_3 + x_5 = 1; \\ 3x_1 + 4x_2 - x_3 + 2x_4 - x_5 = 1; \\ 4x_1 + 6x_2 - 2x_3 + 2x_4 = 2; \\ x_1 + 2x_2 - x_3 + x_4 + x_5 = 0. \end{cases} \quad (4.2)$$

Запишем расширенную матрицу:

$$\left(\begin{array}{ccccc|c} 1 & 2 & -1 & 0 & 1 & 1 \\ 3 & 4 & -1 & 2 & -1 & 1 \\ 4 & 6 & -2 & 2 & 0 & 2 \\ 1 & 2 & -1 & 1 & 1 & 0 \end{array} \right).$$

Из второй строки вычтем первую, умноженную на 3. Из третьей строки вычтем первую, умноженную на 4. Из четвертой строки вычтем первую. Легко видеть, что данные преобразования соответствуют процедуре исключения неизвестной x_1 : из первого уравнения мы выражаем x_1 и полученное выражение подставляем в остальные уравнения. Матрица примет вид:

$$\left(\begin{array}{ccccc|c} 1 & 2 & -1 & 0 & 1 & 1 \\ 0 & -2 & 2 & 2 & -4 & -2 \\ 0 & -2 & 2 & 2 & -4 & -2 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array} \right).$$

Вторую строку прибавим к первой и вычтем из третьей. Затем поделим вторую строку на -2 . Матрица примет вид:

$$\left(\begin{array}{ccccc|c} 1 & 0 & 1 & 2 & -3 & -1 \\ 0 & 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array} \right).$$

Из первой строки вычтем четвертую, умноженную на 2. Ко второй строке прибавим четвертую. Переставим третью и четвертую строки. Матрица примет вид:

$$\left(\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & -3 & 1 \\ 0 & 1 & -1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right). \quad (4.3)$$

Таким образом, исходная система уравнений эквивалентна системе:

$$\begin{cases} x_1 + x_3 - 3x_5 = 1 \\ x_2 - x_3 + 2x_5 = 0 \\ x_4 = -1 \\ 0 = 0 \end{cases} \quad (4.4)$$

Выразим x_1, x_2, x_4 через x_3, x_5 :

$$\begin{aligned} x_1 &= 1 - x_3 + 3x_5 \\ x_2 &= 0 + x_3 - 2x_5 \\ x_4 &= -1 \end{aligned} \quad (4.5)$$

Легко видеть, что какие бы значения не принимали *свободные неизвестные* x_3, x_5 всегда найдутся значения *зависимых неизвестных* при которых все уравнения системы (4.2), и, следовательно, системы (4.4), превращаются в верные равенства. С другой стороны, любое решение исходной системы (4.2) удовлетворяет также и системе (4.5). Таким образом, общее решение системы (4.2) можно записать в виде:

$$\begin{aligned} x_1 &= 1 - t_1 + 3t_2; \\ x_2 &= t_1 - 2t_2; \\ x_3 &= t_1; \\ x_4 &= -1; \\ x_5 &= t_2, \end{aligned} \quad (4.6)$$

где t_1, t_2 — произвольные числа¹. Решение (4.6) можно записать в “векторной форме”:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} + t_1 \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 3 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Следующие преобразования строк матрицы будем называть *элементарными преобразованиями*

- 1) перестановка (транспозиция) строк i и j ($i \neq j$),
- 2) домножение строки i на число α ($\alpha \neq 0$),
- 3) прибавление к строке i строки j , умноженной на α .

В дальнейшем нам понадобятся также элементарные преобразования столбцов матрицы.

Лемма 4.1.1. *Элементарные преобразования строк расширенной матрицы системы линейных уравнений не меняют множества решений.*

ДОКАЗАТЕЛЬСТВО. Докажем что преобразование третьего типа не меняет множества решений. Пусть $x' = (x'_1, \dots, x'_n)^T$ — частное решение системы (4.1) ■

Обозначим через e_j столбец, у которого все, кроме j -ой компоненты равны нулю, а j -ая компонента равна 1. Будем говорить, что матрица $A \in F^{m \times n}$ имеет *простейший вид* если ее столбцы с номерами $j_1, \dots, j_r$, где $1 \leq j_1 < \dots < j_r \leq n$, суть столбцы $e_1, \dots, e_r$ соответственно и, если $m > r$, то последние $m - r$ строк — нулевые. Например, матрица (4.3) имеет простейший вид.

Лемма 4.1.2. *С помощью элементарных преобразований строк произвольную матрицу можно привести к простейшему виду.*

¹Форма записи ответа, в общем случае не единственна

что $j_1 = 1, \dots, j_r = r$, тогда имеем:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_{10} \\ c_{20} \\ \vdots \\ c_{r0} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + t_1 \cdot \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{r1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + t_2 \cdot \begin{pmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{r2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + t_{n-r} \cdot \begin{pmatrix} c_{1n-r} \\ c_{2n-r} \\ \vdots \\ c_{rn-r} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \quad (4.9)$$

Мы доказали следующее утверждение:

Теорема 4.1.1. Если система линейных уравнений с расширенной матрицей $(A|b)$ совместна, то найдутся такие векторы $c_0, c_1, \dots, c_{n-r}$, что

$$M(A, b) = \{x = c_0 + t_1 c_1 + \dots + t_{n-r} c_{n-r} : t_1, \dots, t_{n-r} \in F\}.$$

В частности, если $n = r$, то решение c_0 единственное.

ДОКАЗАТЕЛЬСТВО.

4.2. Ранг матрицы

Пусть $A \in F^{m \times n}$. Рассмотрим столбцы $a_1, \dots, a_n$ матрицы $A = (a_1, \dots, a_n)$ как векторы пространства F^m , а строки $\bar{a}_1, \dots, \bar{a}_m$ как векторы пространства F^n . Базу (соответственно ранг) системы $a_1, \dots, a_n$ назовем *столбцовой базой* соответственно *столбцовым рангом* матрицы A . Базу (соответственно ранг) системы $\bar{a}_1, \dots, \bar{a}_m$ назовем *строчечной базой* соответственно *строчечным рангом* матрицы A .

Утверждение 4.2.1. Строчечный и столбцовый ранги матрицы простейшего вида совпадают.

ДОКАЗАТЕЛЬСТВО. Непосредственно проверяется, что столбцы $j_1, \dots, j_r$ матрицы простейшего вида (т.е. векторы $e_1, \dots, e_r$) образуют столбцовую базу. С другой стороны, легко видеть, что строки $1, 2, \dots, r$ образуют строчечную базу. Таким образом, строчечный и столбцовый ранг матрицы простейшего вида равен r . ■

Утверждение 4.2.2. Пусть матрица $A' = (a'_1, \dots, a'_n)$ получена из матрицы $A = (a_1, \dots, a_n)$ серией элементарных преобразований ее строк. Если $a_{j_1}, \dots, a_{j_r}$ — столбцовая база матрицы A , то $a'_{j_1}, \dots, a'_{j_r}$ — столбцовая база матрицы A' .

ДОКАЗАТЕЛЬСТВО. Докажем, что система $a'_{j_1}, \dots, a'_{j_r}$ линейно независима. Так как система $a_{j_1}, \dots, a_{j_r}$ линейно независима, то система линейных уравнений (рассмотренная относительно неизвестных $\alpha_1, \dots, \alpha_r$)

$$\alpha_1 a_{j_1} + \dots + \alpha_r a_{j_r} = 0 \quad (4.10)$$

имеет единственное (нулевое) решение. С матрицей этой системы осуществим те же преобразования, что и с матрицей A . Очевидно, что получим систему

$$\alpha'_1 a'_{j_1} + \dots + \alpha'_r a'_{j_r} = 0.$$

Эта система эквивалентна (4.10), поэтому она тоже имеет единственное (нулевое) решение, что означает линейную независимость векторов $a'_{j_1}, \dots, a'_{j_r}$.

Теперь покажем, что для произвольного $j \in \{1, \dots, n\}$ вектор a'_j линейно выражается через $a'_{j_1}, \dots, a'_{j_r}$. Система

$$\alpha_1 a_{j_1} + \dots + \alpha_r a_{j_r} = a_j$$

совместна, поэтому совместной будет система

$$\alpha_1 a'_{j_1} + \dots + \alpha_r a'_{j_r} = a'_j,$$

полученная из исходной той же серией элементарных преобразований, с помощью которых мы из матрицы A получили A' . ■

Следствие 4.2.1. При элементарных преобразованиях строк столбцовый ранг матрицы не изменяется.

Аналогично можно доказать

Следствие 4.2.2. При элементарных преобразованиях столбцов строчечный ранг матрицы не изменяется.

Утверждение 4.2.3. Пусть матрица A' получена из матрицы A серией элементарных преобразований ее строк. Если $\bar{a}_1, \dots, \bar{a}_m$ — строки матрицы A , $\bar{a}'_1, \dots, \bar{a}'_m$ — строки матрицы A' , то $L(\bar{a}_1, \dots, \bar{a}_m) = L(\bar{a}'_1, \dots, \bar{a}'_m)$.

Доказательство. Рассмотрим, например, преобразование третьего типа. Легко видеть, что системы

$$\begin{aligned} &\bar{a}_1, \dots, \bar{a}_i, \dots, \bar{a}_j, \dots, \bar{a}_m; \\ &\bar{a}_1, \dots, \bar{a}_i + \alpha \bar{a}_j, \dots, \bar{a}_j, \dots, \bar{a}_m \end{aligned}$$

эквивалентны. ■

Следствие 4.2.3. При элементарных преобразованиях строк строчечный ранг матрицы не изменяется.

Аналогично можно доказать

Следствие 4.2.4. При элементарных преобразованиях столбцов столбцовый ранг не изменяется.

Теорема 4.2.1. Строчечный и столбцовый ранги произвольной матрицы совпадают.

Доказательство. Любую матрицу серией элементарных преобразований, сохраняющих столбцовый и строчечный ранги, можно привести к простейшему виду. Но для матрицы простейшего вида эти ранги равны, следовательно, они будут равны и для исходной матрицы. ■

Столбцовый $\equiv$ строчечный ранг назовем просто *рангом* матрицы и обозначим $\text{rank } A$.

Теорема 4.2.2 (Кронекер–Капелли). Система линейных уравнений с основной матрицей A и расширенной матрицей $(A|b)$ совместна тогда и только тогда, когда $\text{rank } A = \text{rank}(A|b)$.

Доказательство. Систему уравнений можно записать в виде $a_1 x_1 + \dots + a_n x_n = b$, где $a_1, \dots, a_n$ — столбцы матрицы A . Очевидно, что совместность системы эквивалентна условию $b \in L(a_1, \dots, a_n)$, что в свою очередь эквивалентно условию $\text{rank } \{a_1, \dots, a_n\} = \text{rank } \{a_1, \dots, a_n, b\}$. ■

4.3. Пространство решений системы линейных однородных уравнений

Система линейных уравнений с нулевой правой частью ($b = 0$) называется *однородной*. Исследуем множество $M(A, 0)$.

Теорема 4.3.1. Пусть $A \in F^{m \times n}$, $\text{rank } A = r$. Множество $M(A, 0)$ решений системы однородных уравнений с матрицей A есть линейное подпространство в F^n размерности $n - r$.

Доказательство. Утверждение сразу следует из формул (4.9), описывающих множество $M(A, 0)$. ■

Базис пространства $M(A, 0)$ называется *фундаментальной системой решений* системы линейных однородных уравнений.

Следствие 4.3.1. Любая линейно независимая система из $n - r$ решений системы линейных однородных уравнений составляет ее фундаментальную систему решений.

4.4. Линейное многообразие

Рассмотрим линейное пространство V над полем F . Пусть a_0 — произвольный вектор, L — подпространство в V . Под записью $a_0 + L$ будем понимать множество всевозможных векторов вида $a_0 + x$, где x — произвольный вектор из L . Множество $a_0 + L$ называется *линейным многообразием* порожденным вектором a_0 и подпространством L . Подпространство L называется также *несущим* подпространством для многообразия $a_0 + L$.

Утверждение 4.4.1. *Следующие три утверждения эквивалентны:*

- 1) $a_0 + L = b_0 + L$,
- 2) $b_0 \in a_0 + L$,
- 3) $b_0 - a_0 \in L$.

ДОКАЗАТЕЛЬСТВО. Импликации $1 \rightarrow 2$, $2 \rightarrow 3$ очевидны. Докажем импликацию $3 \rightarrow 1$. Пусть $b_0 - a_0 \in L$, тогда $b_0 = a_0 + l$ для некоторого $l \in L$. Рассмотрим произвольный вектор $a = a_0 + x$ из $a_0 + L$, $x \in L$. Имеем $a = b_0 + (x - l)$, поэтому $a \in b_0 + L$, откуда $a_0 + L \subseteq b_0 + L$. Аналогично можно показать, что $b_0 + L \subseteq a_0 + L$. ■

Замечание 4.4.1. Импликация $2 \rightarrow 1$ означает, что многообразие порождается любым своим представителем. Эквивалентность $1 \sim 3$ дает критерий совпадения двух линейных многообразий с одинаковым несущим подпространством.

Утверждение 4.4.2. Пусть a_0, b_0 — векторы, L, L' — подпространства. Тогда, если $a_0 + L = b_0 + L'$, то $L = L'$.

ДОКАЗАТЕЛЬСТВО. Из предыдущего утверждения следует, что, если $a_0 + L = b_0 + L'$, то $a_0 + L = a_0 + L'$, поэтому для любого x из L найдется такой y из L' , что $a_0 + x = a_0 + y$, откуда $x = y$, т.е. $x \in L'$, а, следовательно, $L \subseteq L'$. Аналогично можно показать, что $L' \subseteq L$. ■

Замечание 4.4.2. Словесная формулировка утверждения: направляющее подпространство линейного многообразия определяется однозначно.

Из доказанного утверждения следует корректность следующего определения. *Размерность* линейного многообразия $a_0 + L$ называется размерность его несущего подпространства L : $\dim(a_0 + L) = \dim L$.

Линейное многообразие размерности 1 называется *прямой* размерности 2 — *плоскостью* иногда плоскостью называют произвольное линейное многообразие, размерности $n - 1$ — *гиперплоскостью*.

4.5. Множество решений системы линейных уравнений общего вида

Теорема 4.5.1. Пусть $A \in F^{m \times n}$, $b \in F^m$, $\text{rank } A = r$. Множество $M(A, b)$ решений системы уравнений с расширенной матрицей $(A|b)$ есть линейное многообразие в F^n размерности $n - r$.

ДОКАЗАТЕЛЬСТВО. Утверждение сразу следует из формул (4.9), описывающих множество $M(A, b)$. ■

Теорема 4.5.2. Пусть $A \in F^{m \times n}$, $b \in F^m$, $a_0 \in M(A, 0)$, тогда

$$M(A, b) = x_0 + M(A, 0). \quad (4.11)$$

ДОКАЗАТЕЛЬСТВО. Утверждение сразу следует из формул (4.9). ■

Замечание 4.5.1. Из формулы (4.11) следует, что несущим подпространством множества решений системы линейных неоднородных уравнений является линейное пространство решений соответствующей однородной системы.

Замечание 4.5.2. Можно дать следующую словесную формулировку теоремы: общее решение $M(A, b)$ неоднородной системы уравнений есть ее частное решение x_0 плюс общее решение $M(A, 0)$ соответствующей однородной системы.

4.6. Описание подпространства и линейного многообразия арифметического пространства в виде множества решений системы линейных уравнений

Ранее мы видели, что множество решений однородной (соответственно неоднородной) системы линейных уравнений есть линейное подпространство (соответственно многообразие) арифметического пространства. В этом разделе мы покажем обратное: любое линейное подпространство (соответственно многообразие) арифметического пространства является множеством решений некоторой однородной (соответственно неоднородной) системы линейных уравнений.

Теорема 4.6.1. *Для любого подпространства L арифметического пространства F^n существует матрица $A \in F^{m \times n}$, такая, что $L = M(A, 0)$.*

ДОКАЗАТЕЛЬСТВО. Пусть $L = L(b_1, \dots, b_s)$, где $b_i = (b_{i1}, \dots, b_{in})^\top$ ($i = 1, \dots, s$). Рассмотрим систему линейных однородных уравнений

$$\begin{cases} b_{11}x_1 + b_{12}x_2 + \dots + b_{1n}x_n = 0; \\ b_{21}x_1 + b_{22}x_2 + \dots + b_{2n}x_n = 0; \\ \dots \\ b_{s1}x_1 + b_{s2}x_2 + \dots + b_{sn}x_n = 0. \end{cases} \quad (4.12)$$

Пусть векторы

$$\begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{pmatrix}, \dots, \begin{pmatrix} a_{m1} \\ a_{m2} \\ \vdots \\ a_{mn} \end{pmatrix}$$

образуют фундаментальную систему решений для (4.12). Покажем, что матрица $A = (a_{ij})$ — искомая. Иными словами, покажем, что множество решений системы (4.1) есть заданное линейное подпространство. Действительно, этой системе удовлетворяют координаты каждого вектора $b_1, \dots, b_s$, следовательно, этой системе удовлетворяют координаты произвольного вектора из L . Таким образом, $L \subseteq M(A, 0)$. Легко видеть, что ранг системы (4.12) равен размерности пространства L , поэтому размерность m пространства решений системы (4.12) равна $m = n - \dim L$. Откуда

$$\dim M(A, 0) = n - m = n - (n - \dim L) = \dim L.$$

Итак, $L \subseteq M(A, 0)$ и $\dim M(A, 0) = \dim L$. Следовательно, $L = M(A, 0)$. ■

Теорема 4.6.2. *Для любого линейного многообразия $a_0 + L$ арифметического пространства F^n существуют матрица $A \in F^{m \times n}$ и вектор $b \in F^m$, такие, что $a_0 + L = M(A, b)$.*

ДОКАЗАТЕЛЬСТВО. Для подпространства L построим однородную систему линейных уравнений с матрицей A , такую, что $L = M(A, 0)$. Положим $b = Aa_0$. Из теоремы 4.5.1 следует, что $a_0 \in M(A, b)$. ■

Определители

5.1. Перестановки и подстановки

Напомним, что *перестановкой* попарно различных элементов $k_1, \dots, k_n$ называется упорядоченный набор $\pi = (i_1, \dots, i_n)$, в котором $i_p \neq i_s$ при $p \neq s$ и для каждого p найдется t , такое, что $i_p = i_t$. Далее, если не оговорено противное, рассматриваются перестановки элементов $1, \dots, n$.

Транспозицией элементов i_p и i_s ($p \neq s$) в перестановке $i_1, \dots, i_p, \dots, i_s, \dots, i_n$ называется ее преобразование, при котором перестановка переходит в перестановку $i_1, \dots, i_s, \dots, i_p, \dots, i_n$ (элементы i_p и i_s поменялись местами, остальные элементы остались на прежних местах).

Утверждение 5.1.1. *Все $n!$ перестановок из n элементов можно расположить в таком порядке, что каждая следующая перестановка получается из предыдущей одной транспозицией, причем начинать можно с любой перестановки.*

ДОКАЗАТЕЛЬСТВО. Докажем теорему методом математической индукции по параметру n . Основание индукции очевидно. Теперь предположим, что все $(n-1)!$ перестановок из n элементов можно получить с помощью транспозиций. Вначале в качестве исходной рассмотрим перестановку $1, 2, \dots, n$. По предположению индукции с помощью только транспозиций можно получить все перестановки элементов $2, \dots, n$. В начало каждой такой перестановки припишем 1. В результате получим все $(n-1)!$ перестановок из n элементов с первым элементом 1. В последней из выписанных перестановок совершим транспозицию элементов 1 и 2. Выпишем все перестановки элементов, стоящих на 2-ом, 3-ом, $\dots$, n -ом местах. В начало каждой такой перестановки припишем 2. В результате получим все $(n-1)!$ перестановок из n элементов с первым элементом 2. В последней из выписанных перестановок совершим транспозицию элементов 2 и 3 и т.д.

Суммарное количество всех выписанных перестановок — $n(n-1)! = n!$, причем все эти перестановки попарно различны (на первом месте стояли по очереди все n элементов).

Нам осталось показать, что для того, чтобы получить все перестановки с помощью одних транспозиций, можно начинать с любой перестановки. Для доказательства рассмотрим исходную перестановку и переобозначим ее элементы перенумеровав их. ■

Будем говорить, что элементы i_p и i_s , где $p < s$, образуют *инверсию* перестановке $\pi = (i_1, \dots, i_n)$, если $i_p > i_s$. Общее число инверсий в перестановке обозначим $\sigma(\pi) = \sigma(i_1, \dots, i_n)$. В случае $n = 1$ примем $\sigma(\pi) = 1$. Перестановка π называется *четной* если $\sigma(\pi)$ — четно, перестановка π называется *нечетной* если $\sigma(\pi)$ — нечетно.

Утверждение 5.1.2. *Транспозиция меняет четность перестановки.*

ДОКАЗАТЕЛЬСТВО. Вначале рассмотрим транспозицию соседних элементов i_p и i_{p+1} :

$$(i_1, \dots, i_{p-1}, i_p, i_{p+1}, i_{p+2}, \dots, i_n) \rightarrow (i_1, \dots, i_{p-1}, i_{p+1}, i_p, i_{p+2}, \dots, i_n).$$

Легко видеть, что если $i_p < i_{p+1}$, то число инверсий увеличивается на 1 (появляется новая инверсия элементов i_p, i_{p+1}); если $i_p > i_{p+1}$, то число инверсий уменьшается на 1 (исчезает инверсия элементов i_p, i_{p+1}). Итак, транспозиция соседних элементов меняет четность.

$$\begin{array}{ll}
\rightarrow (i_1, \dots, i_{p-1}, i_p, i_{p+1}, i_{p+2}, \dots, i_{s-2}, i_{s-1}, i_s, i_{s+1}, i_n) & \rightarrow \\
\rightarrow (i_1, \dots, i_{p-1}, i_{p+1}, i_p, i_{p+2}, \dots, i_{s-2}, i_{s-1}, i_s, i_{s+1}, i_n) & \rightarrow \\
\rightarrow (i_1, \dots, i_{p-1}, i_{p+1}, i_{p+2}, i_p, \dots, i_{s-2}, i_{s-1}, i_s, i_{s+1}, i_n) & \rightarrow \\
\vdots & \vdots \\
\rightarrow (i_1, \dots, i_{p-1}, i_{p+1}, i_{p+2}, i_{p+3}, \dots, i_p, i_{s-1}, i_s, i_{s+1}, i_n) & \rightarrow \\
\rightarrow (i_1, \dots, i_{p-1}, i_{p+1}, i_{p+2}, i_{p+3}, \dots, i_{s-1}, i_p, i_s, i_{s+1}, i_n) & \rightarrow \\
\rightarrow (i_1, \dots, i_{p-1}, i_{p+1}, i_{p+2}, i_{p+3}, \dots, i_{s-1}, i_s, i_p, i_{s+1}, i_n) & \rightarrow \\
\rightarrow (i_1, \dots, i_{p-1}, i_{p+1}, i_{p+2}, i_{p+3}, \dots, i_s, i_{s-1}, i_p, i_{s+1}, i_n) & \rightarrow \\
\vdots & \vdots \\
\rightarrow (i_1, \dots, i_{p-1}, i_{p+1}, i_{p+2}, i_s, \dots, i_{s-2}, i_{s-1}, i_p, i_{s+1}, i_n) & \rightarrow \\
\rightarrow (i_1, \dots, i_{p-1}, i_{p+1}, i_s, i_{p+2}, \dots, i_{s-2}, i_{s-1}, i_p, i_{s+1}, i_n) & \rightarrow \\
\rightarrow (i_1, \dots, i_{p-1}, i_s, i_{p+1}, i_{p+2}, \dots, i_{s-2}, i_{s-1}, i_p, i_{s+1}, i_n) & \rightarrow \\
\rightarrow (i_1, \dots, i_{p-1}, i_s, i_{p+1}, i_{p+2}, \dots, i_{s-2}, i_{s-1}, i_p, i_{s+1}, i_n). & \rightarrow
\end{array}$$

Из предыдущих двух утверждений получаем

$$\tau = \begin{pmatrix} k_1 & k_2 & \dots & k_n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}, \quad (5.1)$$
$$\tau = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix},$$

ДОКАЗАТЕЛЬСТВО. Для получения всех таблиц, определяющих одну и ту же подстановку достаточно рассмотреть произвольную перестановку столбцов таблицы τ . Эти перестановки можно получить с помощью лишь одних транспозиций столбцов. Каждая такая транспозиция меняет четность обеих перестановок и, следовательно, сохраняет четность величины $\sigma(\tau)$. ■

Рассмотрим функцию $f(a_1, \dots, a_n)$ от n векторных переменных $a_1, \dots, a_n$ с числовыми значениями, где $a_j \in F^n$ ($j = 1, \dots, n$). Такая функция называется *линейной по каждому аргументу* или

полилинейной, если

$$\begin{aligned} f(a_1, \dots, a'_j + a''_j, \dots, a_n) &= f(a_1, \dots, a'_j, \dots, a_n) + f(a_1, \dots, a''_j, \dots, a_n), \\ f(a_1, \dots, \alpha a_j, \dots, a_n) &= \alpha f(a_1, \dots, a_j, \dots, a_n) \end{aligned}$$

для любых входящих сюда векторов, любых чисел α_1, α_2 из поля F и любого $j \in \{1, \dots, n\}$. Функция f называется *знакопеременной* если

$$f(a_1, \dots, a_j, \dots, a_i, \dots, a_n) = -f(a_1, \dots, a_i, \dots, a_j, \dots, a_n)$$

для любых векторов и любых $i, j \in \{1, \dots, n\}$, $i \neq j$. Функция f называется *нормированной* если

$$f(e_1, \dots, e_n) = 1,$$

где e_j — столбец, у которого j -ая компонента равна 1, а остальные равны 0.

Из доказанных результатов следует, что определитель есть полилинейная знакопеременная нормированная функция от своих столбцов. Верно и обратное утверждение:

Теорема 5.5.1. *Любая полилинейная знакопеременная нормированная функция $f(a_1, \dots, a_n)$ есть определитель матрицы, составленной из столбцов $a_1, \dots, a_n$.*

Доказательство. По индукции для полилинейной функции мы можем доказать

$$f\left(a_1, \dots, a_{i-1}, \sum_{i=1}^m \alpha_i b_i, a_{i+1}, \dots, a_n\right) = \sum_{i=1}^m \alpha_i f(a_1, \dots, a_{i-1}, b_i, a_{i+1}, \dots, a_n). \quad (5.2)$$

Далее, пусть $i_1, \dots, i_n$ — некоторая перестановка чисел $1, \dots, n$. Тогда, легко видеть, что

$$f(e_{i_1}, \dots, e_{i_n}) = (-1)^{\sigma(i_1, \dots, i_n)}. \quad (5.3)$$

Докажем теперь, что если $a_i = a_j$, где $i \neq j$, то знакопеременная функция $f(a_1, \dots, a_n)$ равна нулю. Действительно, при транспозиции i -го и j -го аргумента функция не меняется, но по свойству знакопеременности меняет знак, следовательно,

$$f(a_1, \dots, a_i, \dots, a_i, \dots, a_n) = 0. \quad (5.4)$$

Пусть $a_j = (a_{1j}, \dots, a_{nj})^\top$ ($j = 1, \dots, n$), тогда

$$a_j = \sum_{i=1}^n a_{ij} e_i. \quad (5.5)$$

Из (5.2–5.5) получаем

$$\begin{aligned} f(a_1, \dots, a_n) &= \\ &= f\left(\sum_{i_1=1}^n a_{i_1 1} e_{i_1}, \dots, \sum_{i_n=1}^n a_{i_n n} e_{i_n}\right) = \\ &= \sum_{i_1=1}^n \dots \sum_{i_n=1}^n a_{i_1 1} \dots a_{i_n n} f(e_{i_1}, \dots, e_{i_n}) = \\ &= \sum_{i_1, \dots, i_n} (-1)^{\sigma(i_1, \dots, i_n)} a_{i_1 1} \dots a_{i_n n} = \\ &= \det A. \end{aligned}$$

■

Таким образом, мы можем дать новое эквивалентное предыдущим определение: определитель матрицы есть полилинейная нормированная знакопеременная функция от столбцов матрицы (третья точка зрения на определитель).

5.6. Минорный ранг

Назовем *базисным минором* матрицы ненулевой минор наибольшего порядка. Базисный минор матрицы в общем случае определяется неоднозначно, но однозначно определяется его порядок, называемый *минорным рангом* матрицы. В этом разделе мы увидим, что определение минорного ранга эквивалентно другим определениям ранга матрицы и, таким образом,

$$\text{столбцовый ранг} = \text{строчечный ранг} = \text{минорный ранг}.$$

Теорема 5.6.1. *Минор, стоящий на пересечении базисных строк и столбцов матрицы является базисным.*

ДОКАЗАТЕЛЬСТВО. Пусть $i_1, \dots, i_r, j_1, \dots, j_r$ — номера базисных строк и столбцов матрицы A соответственно. Обозначим B подматрицу, образованную элементами, стоящими на пересечении базисных строк и столбцов. Обозначим C подматрицу, образованную всеми строками матрицы A и столбцами $j_1, \dots, j_r$. По условию столбцы матрицы C линейно независимы, поэтому $\text{rank } C = r$. По условию каждая строка матрицы A (и, следовательно, C) линейно выражается через строки $i_1, \dots, i_r$. Так как $\text{rank } C = r$, то строки $i_1, \dots, i_r$ линейно независимы, откуда $\text{rank } B = r$. ■

Минор $M_{j_1, \dots, j_r, j}^{i_1, \dots, i_r, i}$ называется *окаймляющим* отношением к минору $M_{j_1, \dots, j_r}^{i_1, \dots, i_r}$, если $i \notin \{i_1, \dots, i_r\}$, $j \notin \{j_1, \dots, j_r\}$.

Теорема 5.6.2 (об окаймляющих минорах). *Пусть $M_{j_1, \dots, j_r}^{i_1, \dots, i_r} \neq 0$ и $M_{j_1, \dots, j_r, j}^{i_1, \dots, i_r, i} = 0$ для всех $i \notin \{i_1, \dots, i_r\}$, $j \notin \{j_1, \dots, j_r\}$, тогда ранг матрицы равен r , строки с номерами $i_1, \dots, i_r$ и столбцы с номерами $j_1, \dots, j_r$ — базисные.*

ДОКАЗАТЕЛЬСТВО. Для простоты будем считать, что $\{i_1, \dots, i_r\} = \{1, \dots, r\}$, $\{j_1, \dots, j_r\} = \{1, \dots, r\}$. Обозначим B подматрицу, образованную элементами, стоящими на пересечении этих строк и столбцов. Обозначим подматрицу, образованную элементами, стоящими на пересечении строк и столбцов с номерами $1, \dots, r, i$ и $1, \dots, r, j$ соответственно. Обозначим D подматрицу, образованную элементами, стоящими на пересечении строк с номерами $1, \dots, r, i$ и всех столбцов исходной матрицы A .

Так как $\det B \neq 0$, то строки и столбцы матрицы B — линейно независимы. Отсюда в частности следует, что первые r столбцов матрицы C линейно независимы. Но вся система столбцов матрицы C линейно зависима, $\det C = 0$, поэтому ее последний столбец линейно выражается через предыдущие столбцы. Следовательно, каждый столбец матрицы D линейно выражается через первые r столбцов. Так как эти столбцы линейно независимы, то $\text{rank } D = r$. Однако первые r строк матрицы D также линейно независимы и поэтому образуют строчечную базу. Итак, произвольная строка матрицы D (и, следовательно, A) линейно выражается через линейно независимую систему строк с номерами $1, \dots, r$, поэтому эта система является строчечной базой для всей матрицы A . Утверждение про столбцы матрицы доказывается аналогично. ■

Следствие 5.6.1. *Если минор $M_{j_1, \dots, j_r}^{i_1, \dots, i_r}$ базисный, то строки с номерами $i_1, \dots, i_r$ и столбцы с номерами $j_1, \dots, j_r$ — базисные.*

Следствие 5.6.2. *Если $M_{j_1, \dots, j_r}^{i_1, \dots, i_r} \neq 0$ и $M_{j_1, \dots, j_r, j}^{i_1, \dots, i_r, i} = 0$ для всех $i \notin \{i_1, \dots, i_r\}$, $j \notin \{j_1, \dots, j_r\}$, то минор $M_{j_1, \dots, j_r}^{i_1, \dots, i_r}$ базисный.*

5.7. Сумма определителей

Из свойства линейности определителя получаем.

Теорема 5.7.1 (о сумме определителей). *Определитель суммы двух матриц A и B порядка n равен сумме всех 2^n определителей матриц, получающихся из A заменой части столбцов A соответствующими столбцами (т.е. столбцами с теми же номерами) из B (включая сами матрицы A и B).*

Аналогичная теорема справедлива и для строк матриц.

5.8. Теорема Бинэ–Коши

Теорема 5.8.1 (Бинэ–Коши). *Пусть $A = (a_{ij}) \in F^{m \times n}$, $B = (b_{ij}) \in F^{n \times m}$. Тогда*

$$\det(AB) = \sum_{1 \leq k_1 < k_2 < \dots < k_m \leq n} M_{k_1, k_2, \dots, k_m}^{1, 2, \dots, m}(A) \cdot M_{1, 2, \dots, m}^{k_1, k_2, \dots, k_m}(B).$$

В частности, если $m > n$, то $\det(AB) = 0$. Если $m = n$, то

$$\det(AB) = \det A \det B.$$

(последнее равенство составляет содержание теоремы об умножении определителей.

ДОКАЗАТЕЛЬСТВО. Пусть $C = A \cdot B = (c_{ij})$, где $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$ ($i, j = 1, \dots, m$). Тогда

$$\det C = \begin{vmatrix} \sum_{i_1=1}^n a_{1i_1} b_{i_1 1} & \dots & \sum_{i_m=1}^n a_{1i_m} b_{i_m m} \\ \dots & \dots & \dots \\ \sum_{i_1=1}^n a_{mi_1} b_{i_1 1} & \dots & \sum_{i_m=1}^n a_{mi_m} b_{i_m m} \end{vmatrix}.$$

Пользуясь свойством линейности определителя, получаем

$$\det C = \sum_{i_1=1}^n \sum_{i_1=1}^n \dots \sum_{i_m=1}^n b_{i_1 1} \dots b_{i_m m} \begin{vmatrix} a_{1i_1} & \dots & a_{1i_m} \\ \dots & \dots & \dots \\ a_{mi_1} & \dots & a_{mi_m} \end{vmatrix}.$$

В правой части равенства имеем сумму n^m определителей по всем наборам $(i_1, \dots, i_m)$, где $1 \leq i_j \leq n$. Из них определители с одинаковыми столбцами заведомо равны нулю, поэтому

$$\det C = \sum_{i_1, \dots, i_m} b_{i_1 1} \dots b_{i_m m} \begin{vmatrix} a_{1i_1} & \dots & a_{1i_m} \\ \dots & \dots & \dots \\ a_{mi_1} & \dots & a_{mi_m} \end{vmatrix}.$$

В правой части равенства имеем сумму $n(n-1) \dots (n-m+1)$ определителей по всем размещениям $i_1, \dots, i_m$ элементов $1, \dots, n$. Каждое такое размещение есть перестановка элементов $k_1, \dots, k_m$, где $k_1 \leq \dots \leq k_m$. Поэтому

$$\det C = \sum_{1 \leq k_1 < k_2 < \dots < k_m \leq n} \sum_{i_1, \dots, i_m} b_{i_1 1} \dots b_{i_m m} \begin{vmatrix} a_{1i_1} & \dots & a_{1i_m} \\ \dots & \dots & \dots \\ a_{mi_1} & \dots & a_{mi_m} \end{vmatrix}.$$

Внешняя сумма берется по всем сочетаниям $k_1, \dots, k_m$ из элементов $1, \dots, n$, а внутренняя — по всем перестановкам $i_1, \dots, i_m$ элементов $k_1, \dots, k_m$. Переставим столбцы определителей в порядке возрастания $k_1, \dots, k_m$ их индексов $i_1, \dots, i_m$:

$$\det C = \sum_{1 \leq k_1 < k_2 < \dots < k_m \leq n} \sum_{i_1, \dots, i_m} b_{i_1 1} \dots b_{i_m m} (-1)^{\sigma(i_1, \dots, i_m)} \begin{vmatrix} a_{1k_1} & \dots & a_{1k_m} \\ \dots & \dots & \dots \\ a_{mk_1} & \dots & a_{mk_m} \end{vmatrix}.$$

Так как

$$\sum_{i_1, \dots, i_m} b_{i_1 1} \dots b_{i_m m} (-1)^{\sigma(i_1, \dots, i_m)} = \begin{vmatrix} b_{k_1 1} & \dots & a_{k_1 m} \\ \dots & \dots & \dots \\ b_{k_m 1} & \dots & a_{k_m m} \end{vmatrix},$$

то

$$\det C = \sum_{1 \leq k_1 < k_2 < \dots < k_m \leq n} \begin{vmatrix} a_{1k_1} & \dots & a_{1k_m} \\ \dots & \dots & \dots \\ a_{mk_1} & \dots & a_{mk_m} \end{vmatrix} \begin{vmatrix} b_{k_1 1} & \dots & a_{k_1 m} \\ \dots & \dots & \dots \\ b_{k_m 1} & \dots & a_{k_m m} \end{vmatrix}.$$

Пример 39.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{vmatrix} = \\ = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & b_{12} \\ b_{31} & b_{32} \end{vmatrix} + \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \cdot \begin{vmatrix} b_{21} & b_{22} \\ b_{31} & b_{32} \end{vmatrix}.$$

Пример 40. Пусть a, b, c, x, y, z — векторы геометрического трехмерного пространства. Далее $(\cdot, \cdot)$, $[\cdot, \cdot]$, $(\cdot, \cdot, \cdot)$ обозначают скалярное, векторное и смешанное произведение векторов соответственно. Справедливы равенства:

$$(a, b, c)(x, y, z) = \begin{vmatrix} (a, x) & (a, y) & (a, z) \\ (b, x) & (b, y) & (b, z) \\ (c, x) & (c, y) & (c, z) \end{vmatrix}, \\ ([a, b], [x, y]) = \begin{vmatrix} (a, x) & (a, y) \\ (b, x) & (b, y) \end{vmatrix}, \\ (a, b, c)(x, y) = \begin{vmatrix} a & b & c \\ (a, x) & (b, x) & (c, x) \\ (a, y) & (b, y) & (c, y) \end{vmatrix}.$$

Линейные отображения и преобразования

6.1. Определения и примеры

Рассмотрим два линейных пространства V, W , заданных над одним и тем же полем F . Отображение $\varphi : V \rightarrow W$ называется *линейным оператором* (*линейным отображением*), если выполнены следующие свойства (свойства линейности):

- 1) $\varphi(x + y) = \varphi x + \varphi y$ для произвольных векторов $x, y \in V$.
- 2) $\varphi(\alpha x) = \alpha \varphi x$ для произвольного вектора $x \in V$ и произвольного числа $\alpha \in F$.

Отображение $\varphi : V \rightarrow V$ называется *преобразованием*.

Примеры.

Нулевой оператор. Для произвольных пространств V, W , заданных над полем F определим $\varphi x = 0$ для произвольного $x \in V$. Отображение φ , очевидно, является линейным.

Тождественное преобразование. Для произвольного пространства V определим преобразование $\varepsilon : V \rightarrow V$ следующим образом: для произвольного $x \in V$ положим $\varepsilon x = x$. Отображение ε , очевидно, является линейным.

Преобразование проектирования. Преобразование $\varphi : V \rightarrow V$, ставящее в соответствие вектору x его проекцию $\text{pr}_{V_1 \| V_2} x$ является линейным.

Действительно, пусть $x = x_1 + x_2$, $y = y_1 + y_2$; $x_1, y_1 \in V_1$, $x_2, y_2 \in V_2$. Тогда $x + y = (x_1 + y_1) + (x_2 + y_2)$, $x_1 + y_1 \in V_1$, $x_2 + y_2 \in V_2$ и поэтому

$$\text{pr}_{V_1 \| V_2}(x + y) = x_1 + y_1 = \text{pr}_{V_1 \| V_2} x + \text{pr}_{V_1 \| V_2} y.$$

Таким образом, $\varphi(x + y) = \varphi x + \varphi y$.

С другой стороны, $\alpha x = \alpha x_1 + \alpha x_2$, поэтому

$$\text{pr}_{V_1 \| V_2} \alpha x = \alpha \text{pr}_{V_1 \| V_2} x.$$

Следовательно, $\varphi(\alpha x) = \alpha \varphi x$.

Упражнение 40. Дать (геометрическую) интерпретацию преобразования проектирования в пространствах V_2 и V_3 .

Из свойств линейности оператора по индукции легко вывести следующее обобщение:

$$\varphi \left(\sum_{j=1}^s \alpha_j a_j \right) = \sum_{j=1}^s \alpha_j \varphi a_j. \quad (6.1)$$

В свойстве 2) полагая $\alpha = 0$ получаем $\varphi 0 = 0$.

6.2. Матрица линейного оператора

Обозначим через $\Phi(V, W)$ множество всех линейных операторов, действующих из V в W . В этом разделе мы дадим обозрение множества $\Phi(V, W)$.

1. Пусть $e_1, e_2, \dots, e_n$ — некоторый базис пространства V . Для произвольного вектора $x = \sum_{j=1}^n \alpha_j e_j \in V$ из (6.1) получаем:

$$\varphi x = \sum_{j=1}^n \alpha_j \varphi e_j. \quad (6.2)$$

Таким образом, линейный оператор восстанавливается однозначно по образам базисных векторов.

2. Пусть $e_1, e_2, \dots, e_n$ и $f_1, f_2, \dots, f_m$ — базисы пространств V и W соответственно. Так как $\varphi e_j \in W$ ($j = 1, 2, \dots, n$), то φe_j раскладывается по базису $f_1, f_2, \dots, f_m$. Пусть $[\varphi e_j]_{\mathbf{f}} = (\alpha_{1j}, \alpha_{2j}, \dots, \alpha_{mj})^T$ — столбец координат вектора φe_j в базисе $f_1, f_2, \dots, f_m$. Матрица, j -ый столбец которой есть столбец координат $[\varphi e_j]_{\mathbf{f}}$ ($j = 1, 2, \dots, n$), называется *матрицей оператора* φ , построенной в базисах $e_1, e_2, \dots, e_n$ и $f_1, f_2, \dots, f_m$. Обозначение: $[\varphi]_{f,e}$. Из определения получаем

$$[\varphi]_{f,e} = (\alpha_{ij}) \in F^{m \times n}.$$

3. Переходя в выражении (6.2) к координатам, получим:

$$[\varphi x]_{\mathbf{f}} = \sum_{j=1}^n \alpha_j [\varphi e_j]_{\mathbf{f}}.$$

Последнее равенство в матричной форме приобретает вид

$$[\varphi x]_{\mathbf{f}} = [\varphi]_{f,g} [x]_{\mathbf{g}}. \quad (6.3)$$

4. По аналогии с (6.3) рассмотрим равенство

$$[\varphi x]_{\mathbf{f}} = A [x]_{\mathbf{e}}, \quad (6.4)$$

где e, f — два фиксированных базиса пространств V, W соответственно, A — произвольная матрица из $F^{m \times n}$. Приведенная формула каждому вектору $x \in V$ ставит в соответствие вектор $\varphi x \in W$, таким образом, *определяет* оператор $\varphi : V \rightarrow W$. Исследуем его.

Лемма 6.2.1. *Отображение $\varphi : V \rightarrow W$, заданное формулой (6.4), является линейным, причем A — его матрица: $[\varphi]_{f,e} = A$.*

ДОКАЗАТЕЛЬСТВО. Для произвольных $x, y \in V$ имеем

$$[\varphi(x+y)]_{\mathbf{f}} = A[x+y]_{\mathbf{e}} = A([x]_{\mathbf{e}} + [y]_{\mathbf{e}}) = A[x]_{\mathbf{e}} + A[y]_{\mathbf{e}} = [\varphi x]_{\mathbf{f}} + [\varphi y]_{\mathbf{f}},$$

следовательно, $\varphi(x+y) = \varphi x + \varphi y$. Для произвольного $x \in V$ и произвольного $\alpha \in F$ имеем

$$[\varphi(\alpha x)]_{\mathbf{f}} = A[\alpha x]_{\mathbf{e}} = A(\alpha [x]_{\mathbf{e}}) = \alpha A[x]_{\mathbf{e}} = \alpha [\varphi x]_{\mathbf{f}},$$

следовательно, $\varphi(\alpha x) = \alpha \varphi x$. Итак, оператор φ — линейный.

Проверим теперь, что $[\varphi]_{f,e} = A$. Подставим e_j ($j = 1, 2, \dots, n$) в (6.4): $[\varphi e_j]_{\mathbf{f}} = A[e_j]_{\mathbf{e}}$. Так как j -ая компонента столбца $[e_j]_{\mathbf{e}}$ равна 1, а остальные компоненты равны 0, то $A[e_j]_{\mathbf{e}}$ есть j -ый столбец матрицы A , значит A , по определению, есть матрица оператора φ . ■

5. Переформулировка результатов пп. 2,4 приводит нас к следующему.

Следствие 6.2.1. *Отображение, ставящее в соответствие всякому оператору его матрицу, является биекцией из $\Phi(V, W)$ в $F^{m \times n}$.*

6.3. Операции с линейными отображениями

Определения.

- Суммой двух операторов $\varphi, \psi \in \Phi(V, W)$ называется оператор χ , такой, что $\chi x = \varphi x + \psi x$ для произвольного $x \in V$. Обозначение для суммы: $\chi = \varphi + \psi$.
- Произведением оператора $\varphi \in \Phi(V, W)$ на число $\alpha \in F$ называется оператор χ , такой, что $\chi x = \alpha(\varphi x)$ для произвольного $x \in V$. Обозначение для произведения оператора на число: $\chi = \alpha\varphi$.
- Рассмотрим три пространства U, V, W , заданных над одним и тем же полем F . Пусть ψ, φ — линейные операторы, действующие из U в V и из V в W соответственно: $U \xrightarrow{\psi} V \xrightarrow{\varphi} W$. Произведением операторов φ и ψ называется оператор χ , такое, что $\chi x = \varphi(\psi x)$. Обозначение для произведения операторов: $\chi = \varphi\psi$.

Итак,

$$\begin{aligned}(\varphi + \psi)x &= \varphi x + \psi x, \\ (\alpha\varphi)x &= \alpha(\varphi x), \\ (\varphi\psi)x &= \varphi(\psi x).\end{aligned}$$

Линейность результата операций.

Утверждение 6.3.1. 1) Пусть $\varphi, \psi \in \Phi(V, W)$, $\alpha \in F$, тогда операторы $\varphi + \psi$, $\alpha\varphi$ — линейные, причем $[\varphi + \psi] = [\varphi] + [\psi]$, $[\alpha\varphi] = \alpha[\varphi]$.

2) Пусть $\varphi \in \Phi(V, W)$, $\psi \in \Phi(U, V)$, тогда оператор $\varphi\psi$ — линейный, причем $[\varphi\psi] = [\varphi][\psi]$.

Доказательство. Все свойства доказываются аналогично. Приведем для примера два способа доказательства п. 2.

1 СПОСОБ. Пусть

$$\begin{aligned}g_1, g_2, \dots, g_l &\text{ — базис пространства } U, \\ &\quad \downarrow \psi \\ e_1, e_2, \dots, e_n &\text{ — базис пространства } V, \\ &\quad \downarrow \varphi \\ f_1, f_2, \dots, f_m &\text{ — базис пространства } W;\end{aligned}\tag{6.5}$$

$$[\varphi]_{f,e} = (\alpha_{ij}) \in F^{m \times n}, [\psi]_{e,g} = (\beta_{jk}) \in F^{n \times l}.\tag{6.6}$$

Теперь получаем $[(\varphi\psi)x]_f = [\varphi(\psi x)]_f = [\varphi]_{f,e}[\psi x]_e = [\varphi]_{f,e}[\psi]_{e,g}[x]_g$. Обозначим $A = [\varphi]_{f,e}[\psi]_{e,g}$, тогда $[(\varphi\psi)x]_f = A[x]_g$. Применяя лемму из предыдущего раздела, получаем, что оператор $\varphi\psi$ — линейный и $[\varphi\psi]_{f,g} = [\varphi]_{f,e}[\psi]_{e,g}$.

2 СПОСОБ. Докажем вначале линейность. Для произвольных $x, y \in U$ имеем $(\varphi\psi)(x + y) = \varphi(\psi(x + y)) = \varphi(\psi x + \psi y) = \varphi(\psi x) + \varphi(\psi y) = (\varphi\psi)x + (\varphi\psi)y$. Для произвольных $x \in U$, $\alpha \in F$ имеем $(\varphi\psi)(\alpha x) = \varphi(\psi(\alpha x)) = \varphi(\alpha\psi x) = \alpha\varphi(\psi x) = \alpha(\varphi\psi)x$. Итак, оператор $\varphi\psi$ — линейный.

Теперь вычислим $[\varphi\psi]_{f,g} = (\gamma_{ik})$. В системе обозначений (6.5, 6.6) для $k = 1, 2, \dots, l$ получаем $(\varphi\psi)g_k = \varphi(\psi g_k) = \varphi \sum_{j=1}^n \beta_{jk} e_j = \sum_{j=1}^n \beta_{jk} \varphi e_j = \sum_{j=1}^n \beta_{jk} \sum_{i=1}^m \alpha_{ij} f_i = \sum_{i=1}^m f_i \sum_{j=1}^n \alpha_{ij} \beta_{jk}$. По определению матрицы линейного оператора получаем, что $\gamma_{ik} = \sum_{j=1}^n \alpha_{ij} \beta_{jk}$, или $[\varphi\psi]_{f,g} = [\varphi]_{f,e}[\psi]_{e,g}$. ■

ЗАМЕЧАНИЕ. Сумма, произведение и произведение на число линейных операторов являются операторами линейными в том числе и для бесконечномерного пространства. 2-ой способ доказательства проходит и для этого случая.

Другие свойства линейных операций. Линейные операции над операторами — это сложение операторов и умножение их на числа из поля.

Утверждение 6.3.2. Для любых $\varphi, \psi, \chi \in \Phi(V, W)$, $\alpha, \beta \in F$ справедливы равенства:

- 1) $\varphi + \psi = \psi + \varphi$,
- 2) $\varphi + (\psi + \chi) = (\varphi + \psi) + \chi$,

- 3) $\varphi + 0 = \varphi$, где 0 — нулевой оператор,
 4) $\varphi + (-1)\varphi = 0$ (таким образом, $(-1)\varphi$ — оператор противоположный оператору φ),
 5) $1\varphi = \varphi$,
 6) $\alpha(\beta\varphi) = (\alpha\beta)\varphi$,
 7) $(\alpha + \beta)\varphi = \alpha\varphi + \beta\varphi$,
 8) $\alpha(\varphi + \psi) = \alpha\varphi + \alpha\psi$.

ДОКАЗАТЕЛЬСТВО. Доказательство приведенных выше свойств не должно вызывать затруднений. Докажем для примера двумя способами последнее свойство.

1 СПОСОБ (ПО ОПРЕДЕЛЕНИЮ). Для произвольного $x \in V$ имеем $[\alpha(\varphi + \psi)x]_{\mathbf{f}} = [\alpha(\varphi + \psi)]_{\mathbf{f}, \mathbf{e}}[x]_{\mathbf{e}} = [\alpha\varphi]_{\mathbf{f}, \mathbf{e}}[x]_{\mathbf{e}} + [\alpha\psi]_{\mathbf{f}, \mathbf{e}}[x]_{\mathbf{e}} = [\alpha\varphi x]_{\mathbf{f}} + [\alpha\psi x]_{\mathbf{f}}$, откуда вытекает доказываемое.

2 СПОСОБ (ЧЕРЕЗ СВОЙСТВА МАТРИЦ). Для произвольного $x \in V$ имеем $(\alpha(\varphi + \psi))x = \alpha((\varphi + \psi)x) = \alpha(\varphi x + \psi x) = \alpha(\varphi x) + \alpha(\psi x) = (\alpha\varphi)x + (\alpha\psi)x$, т.е. $\alpha(\varphi + \psi) = \alpha\varphi + \alpha\psi$. ■

ЗАМЕЧАНИЕ. 2-ой способ доказательства проходит также и для случая бесконечномерного линейного пространства.

Следствие 6.3.1. Множество $\Phi(V, W)$ относительно ранее введенных операций сложения операторов и умножения их на числа является линейным пространством. Отображение, ставящее в соответствие всякому преобразованию его матрицу, является изоморфизмом $\Phi(V, W)$ в $F^{m \times n}$.

ДОКАЗАТЕЛЬСТВО. Замкнутость операций сложения операторов и умножения их на числа вытекает из утверждения о линейности результата операций. Аксиомы линейного пространства доказаны в предыдущем утверждении. Биективность указанного в формулировке следствия вытекает из утверждения п. 5 предыдущего раздела. Сохранение операций следует из утверждения о линейности результата операций. ■

Другие свойства операций умножения линейных операторов.

Утверждение 6.3.3. Пусть U, V, W, Q — линейные пространства, заданные над одним полем F . Тогда для любых линейных операторов $\varphi, \varphi_1 \in \Phi(V, W)$, $\psi, \psi_1 \in \Phi(U, V)$, $\chi \in \Phi(Q, U)$, и произвольных $\alpha, \beta \in F$ справедливы равенства:

- 1) $\varphi(\psi\chi) = (\varphi\psi)\chi$,
 2) $\alpha(\varphi\psi) = (\alpha\varphi)\psi = \varphi(\alpha\psi)$,
 3) $(\varphi + \varphi_1)\psi = \varphi\psi + \varphi_1\psi$,
 4) $\varphi(\psi + \psi_1) = \varphi\psi + \varphi\psi_1$.

ДОКАЗАТЕЛЬСТВО. Как и свойства линейных операций сформулированные свойства легко могут быть выведены из определения, либо из соответствующих свойств матриц. ■

6.4. Изменение матрицы оператора при замене базисов

Пусть $\varphi \in \Phi(V, W)$,

$$\left. \begin{array}{l} e_1, e_2, \dots, e_n \\ e'_1, e'_2, \dots, e'_n \end{array} \right\} \text{ два базиса пространства } V;$$

$$\left. \begin{array}{l} f_1, f_2, \dots, f_m \\ f'_1, f'_2, \dots, f'_m \end{array} \right\} \text{ два базиса пространства } W.$$

Изучим, как меняется матрица преобразования φ при переходе от базисов $e_1, e_2, \dots, e_n$ и $f_1, f_2, \dots, f_m$ к базисам $e'_1, e'_2, \dots, e'_n$ и $f'_1, f'_2, \dots, f'_m$ соответственно. Из равенств

$$\begin{aligned} [x]_e &= Q_{e \rightarrow e'} [x]_{e'}, \\ [\varphi x]_f &= Q_{f \rightarrow f'} [\varphi x]_{f'}, \\ [\varphi x]_f &= [\varphi]_{f,e} [x]_e \end{aligned}$$

легко получается формула

$$[\varphi x]_{f'} = Q_{f \rightarrow f'}^{-1} [\varphi]_{f,e} Q_{e \rightarrow e'} [x]_{e'},$$

справедливая для произвольного $x \in V$. Обозначив $A = Q_{f \rightarrow f'}^{-1} [\varphi]_{f,e} Q_{e \rightarrow e'}$, получаем

$$[\varphi]_{f',e'} = Q_{f \rightarrow f'}^{-1} [\varphi]_{f,e} Q_{e \rightarrow e'}.$$

6.5. Ядро и образ оператора

Образом, или *множеством значений* оператора $\varphi \in \Phi(V, W)$, называется множество векторов из W , для которых в V существует по крайней мере один прообраз. Обозначение: φV или $\text{Im } \varphi$. Итак, по определению,

$$\varphi V = \{y \in W : \exists x \in V, \varphi x = y\}.$$

Ядром, или *нуль-пространством*, оператора $\varphi \in \Phi(V, W)$, называется множество векторов из V , обращающихся в $0 \in W$. Обозначение: $\text{Ker } \varphi$. Итак, по определению,

$$\text{Ker } \varphi = \{x \in V : \varphi x = 0\}.$$

Утверждение 6.5.1. *Образ и ядро линейного оператора $\varphi \in \Phi(V, W)$ являются подпространствами в W, V соответственно.*

ДОКАЗАТЕЛЬСТВО. 1) Очевидно, что $\varphi V \neq \emptyset$. Для $y_1, y_2 \in \varphi V$ найдутся такие $x_1, x_2 \in V$, что $\varphi x_1 = y_1, \varphi x_2 = y_2$. Так как $y_1 + y_2 = \varphi(x_1 + x_2)$ и $x_1 + x_2 \in V$, то $y_1 + y_2 \in \varphi V$. Пусть $y = \varphi x \in V, \alpha \in F$, тогда $\alpha y = \alpha \varphi x = \varphi(\alpha x) \in \varphi V$.

2) Так как $\varphi 0 = \varphi 0 \cdot 0 = 0 \varphi = 0$, то $0 \in \varphi V$, поэтому $\varphi V \neq \emptyset$. Пусть $x_1 \in \text{Ker } \varphi, x_2 \in \text{Ker } \varphi$, тогда $\varphi(x_1 + x_2) = \varphi x_1 + \varphi x_2 = 0$, следовательно, $x_1 + x_2 \in \text{Ker } \varphi$. Пусть $x \in \text{Ker } \varphi, \alpha \in F$, тогда $\varphi(\alpha x) = \alpha \varphi x = 0$, следовательно, $\alpha x \in \text{Ker } \varphi$. ■

Размерность образа линейного оператора называется его *рангом* и обозначается $\text{rank } \varphi$, размерность ядра линейного оператора называется *дефектом* и обозначается $\text{def } \varphi$.

Теорема 6.5.1 (Ранг и дефект оператора). 1) Пусть $e_1, e_2, \dots, e_n$ и $f_1, f_2, \dots, f_m$ — произвольные базисы пространств V, W соответственно, тогда $\text{rank } \varphi = [\varphi]_{f,e}$.

2) Имеет место равенство $\text{def } \varphi + \text{rank } \varphi = \dim V$.

ДОКАЗАТЕЛЬСТВО. 1) Рассмотрим множество $[\varphi V]$ координатных столбцов всех векторов из $\varphi V = \{y = \varphi x : x \in V\}$. Имеем $[\varphi V] = \{y = [\varphi]_{f,e} x : x \in F^n\}$, таким образом, $[\varphi V]$ является линейной оболочкой столбцов матрицы $[\varphi]_{f,e}$ и имеет размерность $\text{rank} [\varphi]_{f,e}$. Очевидно, $\dim V = \dim [\varphi V]$.

2) Рассмотрим множество $[\text{Ker } \varphi]$ координатных столбцов всех векторов из $\text{Ker } \varphi$. Имеем $[\text{Ker } \varphi] = \{x \in F^n : [\varphi]_{f,e} x = 0\}$, таким образом, $[\text{Ker } \varphi]$ есть множество решений системы линейных уравнений, поэтому $\dim [\text{Ker } \varphi] = n - \text{rank} [\varphi]_{f,e}$, или $\text{def } \varphi = n - \text{rank } \varphi$. ■

6.6. Линейные преобразования

Напомним, что линейным преобразованием называется линейный оператор $\varphi : V \rightarrow V$.

- Преобразование называется *тождественным* (единичным) и обозначается ε , если $\varepsilon x = x$ для любого $x \in V$.
- Преобразование ψ_1 называется *левым обратным* для преобразования φ , если $\psi_1 \varphi = \varepsilon$.

- Преобразование ψ_2 называется *правым обратным* для преобразования φ , если $\varphi\psi_2 = \varepsilon$.
- Преобразование называется *обратным* для преобразования φ и обозначается φ^{-1} , если $\varphi\varphi^{-1} = \varphi\varphi^{-1} = \varepsilon$. Преобразование φ называется *невырожденным*, если для него существует обратное преобразование.

Матрицей преобразования φ в базисе $e_1, e_2, \dots, e_n$ называется матрица, составленная из координатных столбцов $[\varphi e_j]_e$ ($j = 1, 2, \dots, n$). Для матрицы преобразования используется обозначение $[\varphi]_e$. Для произвольных φ, ψ из $\Phi(V, V)$ и произвольного α из F , используя свойства матриц линейных операторов, получаем

$$\begin{aligned} [\alpha\varphi]_e &= \alpha[\varphi]_e, \\ [\varphi + \psi]_e &= [\varphi]_e + [\psi]_e, \\ [\varphi\psi]_e &= [\varphi]_e[\psi]_e. \end{aligned} \quad (6.7)$$

В пространстве $\Phi(V, V)$ определим операцию возведения в степень:

$$\varphi^m = \underbrace{\varphi\varphi \dots \varphi}_{m \text{ раз}}.$$

Из (6.7) по индукции получаем $[\varphi^m] = [\varphi]^m$.

Значением многочлена

$$f(\lambda) = a_0\lambda^m + a_1\lambda^{m-1} + \dots + a_{m-1}\lambda + a_m \in F[\lambda] \quad (6.8)$$

от преобразования φ назовем преобразование

$$f(\varphi) = a_0\varphi^m + a_1\varphi^{m-1} + \dots + a_{m-1}\varphi + a_m\varepsilon.$$

Значением многочлена (6.8) *от матрицы* $A \in F^{n \times n}$ назовем матрицу

$$f(A) = a_0A^m + a_1A^{m-1} + \dots + a_{m-1}A + a_mE.$$

Из (6.7) по индукции получаем $[f(\varphi)]_e = f([\varphi]_e)$.

6.7. Собственные числа и собственные векторы преобразования

Подпространство W линейного пространства V называется *инвариантным относительно преобразования* φ , если $\varphi W \subseteq W$, т.е. $\varphi x \in W$ для любого x из W . *Сужением преобразования* (индуцированным преобразованием) φ на инвариантное подпространство W называется преобразование $\psi: W \rightarrow W$ такое, что $\psi x = \varphi x$ для любого $x \in W$. Обозначение: $\phi|_W = \psi$. Очевидно, преобразование $\phi|_W$ — линейное.

Упражнение 41. Докажите, что образ и ядро линейного преобразования являются инвариантными пространствами. Опишите соответствующие индуцированные преобразования.

Вектор $x \in V$ называется *собственным вектором* преобразования φ , если для некоторого $\lambda \in F$ выполнены условия

$$\begin{aligned} \varphi x &= \lambda x, \\ x &\neq 0. \end{aligned} \quad (6.9)$$

Число λ при этом называется *собственным значением* (числом) преобразования φ . Говорят также, что собственный вектор x *принадлежит* или *относится* к собственному значению λ .

Обозначим через V_λ множество всех собственных векторов, принадлежащих собственному значению λ , дополненное нулевым вектором, иными словами:

$$V_\lambda = \{x \in V : \lambda x\}.$$

Утверждение 6.7.1. *Множество V_λ является подпространством пространства V .*

Доказательство. Очевидно $V_\lambda \neq \emptyset$. Далее, если $x, y \in V_\lambda$, $\alpha \in F$, то

$$\begin{aligned} \varphi(x + y) &= \varphi x + \varphi y = \lambda x + \lambda y = \lambda(x + y), \\ \varphi(\alpha x) &= \alpha(\varphi x) = \alpha(\lambda x) = \lambda(\alpha x). \end{aligned}$$

поэтому $x + y \in V_\lambda$, $\alpha x \in V_\lambda$. ■

Подпространство V_λ называется *собственным пространством* преобразования φ .

Упражнение 42. Доказать что V_λ инвариантно относительно φ .

Утверждение 6.7.2. Если x — собственный вектор, то для любого ненулевого $\alpha \in F$ вектор αx — тоже собственный, относящийся к тому же собственному числу.

ДОКАЗАТЕЛЬСТВО.

$$\varphi(\alpha x) = \alpha(\varphi x) = \alpha(\lambda x) = \lambda(\alpha x).$$

■

Утверждение 6.7.3. Собственные векторы — это в точности базисные векторы одномерных инвариантных подпространств.

ДОКАЗАТЕЛЬСТВО. Пусть x — собственный вектор. По предыдущему утверждению αx для любого $\alpha \neq 0$ — тоже собственный, относящийся к тому же собственному числу, поэтому подпространство $L(x)$ инвариантно.

Пусть $L(x)$, где x — некоторый ненулевой вектор, инвариантно, т.е. $\varphi x \in L(x)$, или $\varphi x = \lambda x$ для некоторого $\lambda \in F$, следовательно, вектор x — собственный. ■

Упражнение 43. Дайте геометрическую интерпретацию понятию собственного вектора.

Исследуем задачу нахождения собственных векторов преобразования φ . Пусть $e_1, \dots, e_n$ — произвольный базис пространства V . Условия (6.9), очевидно, эквивалентны системе

$$\begin{cases} (\varphi - \lambda \varepsilon)x = 0, \\ x \neq 0; \end{cases} \Leftrightarrow \begin{cases} [\varphi - \lambda \varepsilon]_{\mathbf{e}}[x]_{\mathbf{e}} = 0, \\ x \neq 0; \end{cases}$$

или

$$\begin{cases} ([\varphi]_{\mathbf{e}} - \lambda E)[x]_{\mathbf{e}} = 0, \\ x \neq 0. \end{cases} \quad (6.10)$$

Итак, вектор x является собственным тогда и только тогда, когда его координатный столбец $[x]_{\mathbf{e}}$ является нетривиальным решением квадратной системы линейных уравнений (6.10). Для существования такого x необходимо и достаточно, чтобы

$$\det([\varphi]_{\mathbf{e}} - \lambda E) = 0. \quad (6.11)$$

По аналогии с (6.11) для произвольной матрицы $A \in F^{n \times n}$ относительно неизвестного λ рассмотрим уравнение

$$\det(A - \lambda E) = 0,$$

называемое *характеристическим уравнением матрицы* A . Его корни совпадают с корнями многочлена

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

из $F[\lambda]$. Этот многочлен называется *характеристическим многочленом матрицы* A .

Лемма 6.7.1. Характеристические многочлены подобных матриц совпадают.

ДОКАЗАТЕЛЬСТВО. Пусть $B = Q^{-1}AQ$ для некоторой невырожденной Q . Тогда

$$\begin{aligned} \det(\lambda E - B) &= \det(Q^{-1}\lambda EQ - Q^{-1}AQ) \\ &= \det(Q^{-1}(\lambda E - A)Q) \\ &= \underbrace{\det(Q^{-1})}_{1/\det Q} \det(\lambda E - A) \det Q \\ &= \det(\lambda E - A). \end{aligned}$$

■

Из леммы следует, что уравнение (6.11), записанное для одно и того же преобразования φ в разных базисах, имеет один и тот же вид и поэтому корректны следующие определения. Уравнение (6.11) называется *характеристическим уравнением преобразования* φ . Левая часть этого уравнения

$$\chi(\lambda) = \det(\lambda E - [\varphi]_{\mathbf{e}})$$

называется *характеристическим многочленом преобразования* φ .

Из всего вышесказанного следует

Теорема 6.7.1. Для того, что $\lambda \in F$ являлось собственным числом преобразования φ необходимо и достаточно, чтобы оно являлось корнем характеристического многочлена этого преобразования.

6.7.1. Выражение коэффициентов характеристического многочлена через главные миноры матрицы

Миноры матрицы $A \in F^{n \times n}$ вида $M_A(j_1, \dots, j_k)$ называются *главными минорами* матрицы A . Сумма диагональных элементов (сумма главных миноров первого порядка) матрицы A называется ее *следом* и обозначается $\text{tr } A$.

Лемма 6.7.2. *Определитель суммы двух матриц A и B порядка n равен сумме всех 2^n определителей матриц, получающихся из A заменой части столбцов соответствующими столбцами из B .*

Доказательство. Пусть $a_1, \dots, a_n$ — столбцы матрицы A , и $b_1, \dots, b_n$ — столбцы матрицы B . Имеем

$$\begin{aligned} \det(A+B) &= \det(a_1+b_1, a_2+b_2, \dots, a_n+b_n) \\ &= \det(a_1, a_2+b_2, \dots, a_n+b_n) + \det(b_1, a_2+b_2, \dots, a_n+b_n) \\ &= \det(a_1, a_2, \dots, a_n+b_n) + \det(a_1, b_2, \dots, a_n+b_n) \\ &\quad + \det(b_1, a_2, \dots, a_n+b_n) + \det(b_1, b_2, \dots, a_n+b_n) \\ &= \text{и т.д.} \end{aligned}$$

Каждый раз расщепляя определители по очередному столбцу мы получим в конце цепочки равенств 2^n определителей, удовлетворяющих доказываемым свойствам. ■

Теорема 6.7.2. *Коэффициент s_k многочлена*

$$\det(A - \lambda E) = (-\lambda)^n + s_1(-\lambda)^{n-1} + \dots + s_{n-1}(-\lambda) + s_n$$

равен сумме главных миноров порядка k . В частности, $s_1 = \text{tr } A$, $s_n = \det A$.

Доказательство. Применим предыдущую лемму к матрицам A и $(-\lambda E)$. В указанной сумме каждый из определителей имеет следующий вид:

- столбцы с номерами $j_1, \dots, j_k$, где $1 \leq j_1 < \dots < j_k \leq n$, совпадают с соответствующими столбцами матрицы A (столбцы первой группы),
- в остальных столбцах на диагонали стоит $-\lambda$, на остальных местах 0 (столбцы второй группы).

Раскладывая каждый определитель по столбцам второй группы, мы получим $\lambda^{n-k} M_A(j_1, \dots, j_k)$. Для окончания доказательства осталось собрать слагаемые с одинаковым множителем λ^{n-k} . ■

6.7.2. Матрица Фробениуса

Матрица вида

$$F(a_1, \dots, a_n) = \begin{pmatrix} -a_1 & -a_2 & \dots & -a_{n-1} & -a_n \\ 1 & 0 & & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ & & \ddots & & \\ 0 & 0 & & 1 & 0 \end{pmatrix}$$

называется *матрицей Фробениуса*.

Утверждение 6.7.4. *Характеристический многочлен матрицы Фробениуса $F(a_1, \dots, a_n)$ равен*

$$\Phi(a_1, \dots, a_n; \lambda) = \lambda^n + a_1 \lambda^{n-1} + a_2 \lambda^{n-2} + \dots + a_{n-1} \lambda + a_n.$$

Доказательство. Доказательство проведем индукцией по n . Основание индукции легко проверяется. Найдем теперь $\det(\lambda E - F(a_1, \dots, a_n))$. Раскладывая этот определитель по последнему столбцу и пользуясь предположением индукции, получаем

$$\begin{aligned} \det(\lambda E - F(a_1, \dots, a_n)) &= a_n(-1)^{1+n}(-1)^{n-1} + (-\lambda)((-\lambda)^{n-1} + a_1(-\lambda)^{n-2} + \dots + a_{n-1}) \\ &= \Phi(a_1, \dots, a_n; \lambda). \end{aligned}$$

Ввиду предыдущего утверждения матрица Фробениуса $F(a_1, \dots, a_n)$ называется также *сопровождающей матрицей* многочлена $\Phi(a_1, \dots, a_n; \lambda)$.

Следствие 6.7.1. *Всякий многочлен степени n со старшим коэффициентом 1 может быть характеристическим многочленом некоторой квадратной матрицы порядка n .*

6.8. Диагонализируемость линейного преобразования

Преобразование называется *диагонализируемым*, если существует базис, в котором матрица преобразования имеет диагональный вид.

Теорема 6.8.1 (Очень простое, но очень важное утверждение). *Преобразование диагонализуемо тогда и только тогда, когда существует базис из собственных векторов преобразования.*

ДОКАЗАТЕЛЬСТВО.

НЕОБХОДИМОСТЬ. Пусть в базисе $e_1, \dots, e_n$

$$[\varphi]_{\mathbf{e}} = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Так как j -ый столбец матрицы $[\varphi]_{\mathbf{e}}$ есть

$$[\varphi e_j]_{\mathbf{e}} = (0, \dots, 0, \lambda_j, 0, \dots, 0)^{\top},$$

то $\varphi e_j = \lambda_j e_j$, т.е. e_j — собственный вектор ($j = 1, \dots, n$).

ДОСТАТОЧНОСТЬ. Пусть базис $e_1, \dots, e_n$ состоит из собственных векторов, тогда $\varphi e_j = \lambda_j e_j$ ($j = 1, \dots, n$), откуда

$$[\varphi]_{\mathbf{e}} = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Ну, ЛИБО ЭТО ПОНИМАЕШЬ, ЛИБО НЕТ¹.

Пусть характеристический многочлен преобразования φ линейного пространства V , заданного над полем F , имеет вид

$$\begin{aligned} \chi(\lambda) &= (\lambda - \lambda_1)^{k_1} \cdot \dots \cdot (\lambda - \lambda_s)^{k_s} p(\lambda), \\ &\text{где } \lambda_i \in F, \lambda_i \neq \lambda_j \text{ при } i \neq j, \\ &f(\lambda) \text{ корней в } F \text{ не имеет.} \end{aligned}$$

Назовем *алгебраической кратностью* собственного значения λ_i величину k_i (т.е. его кратность, как корня характеристического многочлена).

Назовем *геометрической кратностью* собственного значения λ_i размерность собственного подпространства, принадлежащего собственному значению λ_i :

$$d_i = \dim V_{\lambda_i},$$

иными словами, величину $\text{def}(\varphi - \lambda_i \varepsilon)$, или, что то же, максимальное число линейно независимых решений системы

$$([\varphi]_{\mathbf{e}} - \lambda_i E)x = 0.$$

Лемма 6.8.1. *Геометрическая кратность собственного числа не превосходит алгебраической.*

ДОКАЗАТЕЛЬСТВО. Пусть геометрическая кратность собственного значения λ_0 равна d_0 . Следовательно, найдется система линейно независимых собственных векторов $e_1, \dots, e_{d_0}$, относящихся к собственному значению λ_0 . Дополним ее до базиса и в нем построим матрицу преобразования:

$$[\varphi]_{\mathbf{e}} = \left(\begin{array}{ccc|c} \lambda_0 & & & B \\ & \ddots & & \\ & & \lambda_0 & \\ \hline & & 0 & C \end{array} \right).$$

¹© S.I.Sorochan, 1998. Размещено с разрешения автора

Очевидно, что

$$\chi_\varphi = \det(\lambda E - [\varphi]_e) = (\lambda - \lambda_0)^{d_0} p(\lambda), \quad (6.12)$$

где $p(\lambda) \in F[\lambda]$. Из (6.12) получаем, что алгебраическая кратность собственного значения λ_0 не меньше d_0 . ■

Лемма 6.8.2. *Собственные векторы, относящиеся к разным собственным числам, линейно независимы.*

Доказательство. Пусть $\lambda_1, \dots, \lambda_s$ — некоторые попарно различные собственные числа преобразования φ . В i -ой строке следующей таблицы выписана произвольная линейно независимая система собственных векторов, относящихся к собственному значению λ_i :

$$\begin{aligned} e_{11}, e_{12}, \dots, e_{1d_1} &\rightarrow \lambda_1; \\ e_{21}, e_{22}, \dots, e_{2d_2} &\rightarrow \lambda_2; \\ \dots & \\ e_{s1}, e_{s2}, \dots, e_{sd_s} &\rightarrow \lambda_s. \end{aligned} \quad (6.13)$$

Покажем, что объединенная система собственных векторов (6.13) линейно независима.

Доказательство проведем индукцией по s . При $s = 1$ система линейно независима по предположению. Предположим теперь, что векторы, стоящие в первых $s - 1$ строках таблицы (6.13), линейно независимы и покажем, что все векторы в (6.13) линейно независимы. Рассмотрим произвольную равную нулю линейную комбинацию этих векторов:

$$\sum_{i=1}^s \sum_{j=1}^{d_i} \alpha_{ij} e_{ij} = 0. \quad (6.14)$$

Применим к обеим частям (6.14) преобразование φ :

$$\sum_{i=1}^s \sum_{j=1}^{d_i} \alpha_{ij} \underbrace{\varphi e_{ij}}_{\lambda_i e_{ij}} = 0, \quad (6.15)$$

откуда

$$\sum_{i=1}^s \sum_{j=1}^{d_i} \alpha_{ij} \lambda_i e_{ij} = 0. \quad (6.16)$$

Теперь домножим обе части (6.14) на λ_s :

$$\sum_{i=1}^s \sum_{j=1}^{d_i} \alpha_{ij} \lambda_s e_{ij} = 0. \quad (6.17)$$

Вычитая (6.17) из (6.16), получаем

$$\sum_{i=1}^{s-1} \sum_{j=1}^{d_i} \alpha_{ij} (\lambda_i - \lambda_s) e_{ij} = 0.$$

В последнем равенстве имеем нулевую линейную комбинацию векторов, по предположению индукции линейно независимых. Поэтому все коэффициенты этой линейной комбинации равны нулю:

$$\alpha_{ij} \underbrace{(\lambda_i - \lambda_s)}_{\neq 0} = 0 \quad (i = 1, \dots, s-1),$$

Откуда

$$\alpha_{ij} = 0 \quad (i = 1, \dots, s-1; j = 1, \dots, d_i).$$

Теперь в правой части (6.14) остается лишь одна сумма:

$$\sum_{j=1}^{d_s} \alpha_{sj} e_{sj} = 0.$$

Так как векторы $e_{s1}, \dots, e_{sd_s}$ линейно независимые, то и

$$\alpha_{sj} = 0 \quad (j = 1, \dots, d_s).$$

Итак, все коэффициенты в произвольной линейной комбинации (6.14) равны нулю, поэтому система (6.13) линейно независима. ■

Из двух предыдущих лемм и утверждения получаем следующий результат.

Теорема 6.8.2 (Критерий диагонализированности). *Преобразование диагонализировано тогда и только тогда, когда сумма геометрических кратностей всех собственных чисел совпадает с размерностью пространства.*

6.9. Аннулирующий многочлен

Говорят, что многочлен $f(\lambda)$ *аннулирует* (обращает в ноль) преобразование φ на подпространстве $W \subseteq V$, если $W \subseteq \text{Ker } f(\varphi)$, иными словами, $f(\varphi)x = 0$ для произвольного вектора $x \in W$. В данном случае $f(\lambda)$ называется также *аннулирующим многочленом* преобразования φ на подпространстве (относительно подпространства) W .

Аннулирующий многочлен минимальной степени со старшим коэффициентом 1 называется *минимальным аннулирующим*, или просто *минимальным*.

Пример 41.

- а) Минимальным многочленом, аннулирующим тождественное преобразование ε на произвольном подпространстве, является, очевидно, $f(\lambda) = \lambda - 1$.
- б) Пусть

$$[\varphi] = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

В данном случае первый базисный вектор e_1 собственный, поэтому относительно подпространства $L(e_1)$ аннулирующим (и минимальным) многочленом, является $\lambda - 1$. Относительно всего пространства, легко проверить, аннулирующим является многочлен $f(\lambda) = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2$. Действительно,

$$[f(\varphi)] = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^2 - 2 \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Заметим, что многочлен $f(\lambda)$ совпадает с характеристическим.

Утверждение 6.9.1. *Минимальный многочлен существует и единственен.*

Доказательство.

Существование. Вначале мы докажем существование аннулирующего многочлена. В линейном пространстве $\Phi(V, V)$ рассмотрим систему векторов $\varepsilon, \varphi, \varphi^2, \dots, \varphi^{n^2}$. Так как $\dim \Phi(V, V) = n^2$, то рассматриваемая система линейно зависима, т.е. найдутся такие коэффициенты $\alpha_0, \alpha_1, \dots, \alpha_{n^2}$, что

$$\alpha_0 \varepsilon + \alpha_1 \varphi + \dots + \alpha_{n^2} \varphi^{n^2} = 0.$$

Очевидно, что многочлен $f(\lambda) = \alpha_0 + \alpha_1 \lambda + \dots + \alpha_{n^2} \lambda^{n^2}$ аннулирует преобразование φ на любом подпространстве $W \subseteq V$. Отсюда следует существование минимального аннулирующего многочлена.

Единственность. Пусть $f_1(\lambda), f_2(\lambda)$ – минимальные многочлены преобразования φ на подпространстве W . Докажем, что многочлен $f_1(\lambda) - f_2(\lambda)$ является аннулирующим. Действительно,

$$(f_1(\varphi) - f_2(\varphi))x = f_1(\varphi)x - f_2(\varphi)x = 0$$

для произвольного вектора $x \in W$. У разности $f_1(\lambda) - f_2(\lambda)$ старшие члены взаимно уничтожаются, следовательно, либо $f_1(\lambda) - f_2(\lambda) = 0$, либо степень $f_1(\lambda) - f_2(\lambda)$ меньше степени многочленов $f_1(\lambda), f_2(\lambda)$. Последнее однако невозможно, так как $f_1(\lambda), f_2(\lambda)$ – минимальные, следовательно, $f_1(\lambda) - f_2(\lambda) = 0$, и поэтому $f_1(\lambda) = f_2(\lambda)$. ■

Утверждение 6.9.2. Пусть $f(\lambda)$ – минимальный многочлен преобразования φ относительно подпространства W . Тогда множество всех многочленов, аннулирующих φ на W есть множество $\{f(\lambda)p(\lambda) : p(\lambda) \in F[\lambda]\}$.

Доказательство. Так как $p(\varphi)f(\varphi)x = p(\varphi)0 = 0$ для произвольного $x \in W$, то $f(\lambda)p(\lambda)$ аннулирующий.

Покажем, что других аннулирующих многочленов, кроме многочленов вида $f(\lambda)p(\lambda)$, нет. Для этого поделим произвольный аннулирующий многочлен $g(\lambda)$ с остатком на $f(\lambda)$:

$$g(\lambda) = f(\lambda)p(\lambda) + r(\lambda), \quad (6.18)$$

причем либо $r(\lambda) = 0$, либо $\deg r(\lambda) < \deg f(\lambda)$. Так как $r(\varphi)x = g(\varphi)x - p(\lambda)f(\lambda)x = 0$ для произвольного $x \in W$, то $r(\lambda)$ – также аннулирующий, поэтому $r(\lambda) = 0$. Теперь из (6.18) получаем $g(\lambda) = f(\lambda)p(\lambda)$. ■

Утверждение 6.9.3. Пусть многочлены $f_1(\lambda)$, $f_2(\lambda)$ являются минимальными аннулирующими многочленами для преобразования φ на подпространствах W_1 , W_2 соответственно, причем $W_1 \subseteq W_2$. Тогда $f_2 \mid f_1$.

Доказательство. Так как $W_1 \subseteq W_2$, то $f_2(\lambda)$ аннулирует φ на W_1 . Доказываемое теперь следует из предыдущего утверждения. ■

Утверждение 6.9.4. Пусть многочлены $f_1(\lambda)$, $f_2(\lambda)$ являются минимальными аннулирующими многочленами для преобразования φ на подпространствах W_1 , W_2 соответственно. Тогда $(f_1(\lambda), f_2(\lambda))$ является минимальным аннулирующим многочленом для преобразования φ на подпространстве $W_1 + W_2$.

Доказательство. Сначала докажем, что $f(\lambda) = (f_1(\lambda), f_2(\lambda))$ аннулирует преобразование φ на подпространстве $W_1 + W_2$. Пусть

$$f(\lambda) = p_1(\lambda)f_1(\lambda) = p_2(\lambda)f_2(\lambda).$$

Для произвольного вектора $x = x_1 + x_2 \in W$, где $x_1 \in W_1$, $x_2 \in W_2$, имеем $f(\varphi)x = f(\varphi)(x_1 + x_2) = f(\varphi)x_1 + f(\varphi)x_2 = p_1(\varphi)f_1(\varphi)x_1 + p_2(\varphi)f_2(\varphi)x_2 = 0$. Итак, многочлен $f(\lambda)$ – аннулирующий на $W_1 + W_2$.

Докажем теперь, что $f(\lambda)$ – минимальный многочлен, т.е. из всех аннулирующих многочленов многочлен $f(\lambda)$ имеет минимальную степень. Так как $W_1 \subseteq W_1 + W_2$ и $W_2 \subseteq W_1 + W_2$, то по предыдущему утверждению произвольный аннулирующий на подпространстве $W_1 + W_2$ многочлен должен делиться без остатка как на многочлен $f_1(\lambda)$, так и на многочлен $f_2(\lambda)$. Из всех многочленов, удовлетворяющих этим свойствам, минимальную степень имеет $f(\lambda)$. Следовательно, этот многочлен минимальный аннулирующий. ■

6.9.1. Метод Крылова построения минимального многочлена

В данном разделе мы опишем способ нахождения минимального многочлена, аннулирующего преобразование φ на всем пространстве V .

Пусть $e_1, e_2, \dots, e_n$ – базис пространства V . Для произвольного $i \in \{1, 2, \dots, n\}$ в силу конечномерности пространства найдется натуральное s , такое, что система

$$e_i, \varphi e_i, \varphi^2 e_i, \dots, \varphi^{s-1} e_i$$

линейно независима, а система

$$e_i, \varphi e_i, \varphi^2 e_i, \dots, \varphi^s e_i$$

линейно зависима. Пусть

$$\alpha_0 e_i + \alpha_1 \varphi e_i + \dots + \alpha_{s-1} \varphi^{s-1} e_i + \varphi^s e_i = 0$$

для некоторых α_j ($j = 0, 1, \dots, s-1$). Легко видеть, что многочлен

$$f_i(\lambda) = \alpha_0 + \alpha_1 \lambda + \dots + \alpha_{s-1} \lambda^{s-1} + \lambda^s$$

является минимальным для $L(e_i)$.

По утверждению 4 минимальным многочленом, аннулирующим φ на всем пространстве V , является $(f_1(\lambda), \dots, f_n(\lambda))$.

Пример 42. Построить минимальный аннулирующий многочлен преобразования φ , заданного в некотором базисе матрицей

$$[\varphi] = \begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}.$$

Для базиса e_1, e_2, e_3 , в котором задана матрица преобразования, имеем

$$\begin{aligned} [e_1] &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, & [\varphi e_1] &= \begin{pmatrix} 0 \\ -4 \\ -2 \end{pmatrix}, & [\varphi^2 e_1] &= \begin{pmatrix} -4 \\ -16 \\ -8 \end{pmatrix}; \\ [e_2] &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, & [\varphi e_2] &= \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}, & [\varphi^2 e_2] &= \begin{pmatrix} 4 \\ 12 \\ 4 \end{pmatrix}; \\ [e_3] &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, & [\varphi e_3] &= \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}. \end{aligned}$$

В каждой строке приведенной таблицы векторы линейно зависимы. Найдем их нулевые нетривиальные комбинации:

$$\begin{aligned} 4e_1 - 4\varphi e_1 + \varphi^2 e_1 &= 0 &\rightarrow f_1(\lambda) &= 4 - 4\lambda + \lambda^2 = (\lambda - 2)^2; \\ 4e_2 - 4\varphi e_2 + \varphi^2 e_2 &= 0 &\rightarrow f_2(\lambda) &= 4 - 4\lambda + \lambda^2 = (\lambda - 2)^2; \\ 4e_3 - 4\varphi e_3 &= 0 &\rightarrow f_3(\lambda) &= 4 - 4\lambda + \lambda^2 = \lambda - 2. \end{aligned}$$

Минимальный многочлен имеет вид $f(\lambda) = (f_1(\lambda), f_2(\lambda), f_3(\lambda)) = (\lambda - 2)$. Как и в примере 41 в данном случае характеристический многочлен $\chi(\lambda) = (\lambda - 2)^3$ является аннулирующим. В следующей теореме утверждается, что это справедливо для любого преобразования.

Теорема 6.9.1 (Гамильтон–Кэли). Любое линейное преобразование является корнем своего характеристического многочлена

Доказательство. Мы покажем, что произвольная матрица A является корнем своего характеристического многочлена $\chi_A(\lambda)$, т.е. $\chi_A(A) = 0$.

Рассмотрим квадратную матрицу $(\lambda E - A)^* = (\alpha_{ij})$, в которой $\alpha_{ij} = (-1)^{i+j} M_A(j, i)$. Имеем

$$(\lambda E - A)^*(\lambda E - A) = \det(\lambda E - A)E.$$

В полученное равенство подставим $\lambda = A$, тогда в его правой части получаем $\chi_A(A)$, а в левой 0. ■

Замечание 1. Доказанная теорема равносильна утверждению, что характеристический многочлен преобразования является аннулирующим многочленом того же преобразования во всем пространстве

Замечание 2. Приведенное доказательство не основывалось на результатах данного раздела и является другим доказательством существования аннулирующего многочлена.

Теорема 6.9.2 (Теорема о корнях аннулирующего многочлена). Пусть многочлен

$$\chi(\lambda) = (\lambda - \lambda_1)^{k_1} \cdot \dots \cdot (\lambda - \lambda_s)^{k_s} \\ (\lambda_i \neq \lambda_j \text{ при } i \neq j)$$

является характеристическим многочленом преобразования φ . Тогда минимальным многочленом, аннулирующим φ на всем пространстве, является многочлен

$$\mu(\lambda) = (\lambda - \lambda_1)^{l_1} \cdot \dots \cdot (\lambda - \lambda_s)^{l_s}, \quad (6.19)$$

где $l_1, \dots, l_s$ – некоторые натуральные числа, такие, что $1 \leq l_i \leq k_i$ ($i = 1, 2, \dots, s$).

ДОКАЗАТЕЛЬСТВО. По теореме Гамильтона-Кэли характеристический многочлен является аннулирующим. Минимальный многочлен является его делителем и, поэтому, имеет вид (6.19), причем $l_i \leq k_i$ ($i = 1, 2, \dots, s$).

Теперь докажем, что $l_i \geq 1$. Для этого рассмотрим собственный вектор x . Имеем $\varphi x = \lambda_i x$. Легко видеть, что многочлен $f_i(\lambda) = \lambda - \lambda_i$ является минимальным аннулирующим преобразование φ на подпространстве $L(x)$. Необходимое неравенство следует теперь из утверждения 4. ■

ЗАМЕЧАНИЕ. Из доказанной теоремы следует, что множество корней минимального многочлена без учета их кратностей есть в точности множество характеристических чисел. Кратности, с которыми они встречаются в минимальном многочлене не превосходят алгебраических кратностей характеристических чисел.

6.10. Канонический вид преобразования комплексного линейного пространства (Жорданова форма линейного преобразования)

6.10.1. Определения

Жордановой клеткой называется квадратная матрица порядка n , следующего вида:

$$J_n(\lambda_0) = \begin{pmatrix} \lambda_0 & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_0 & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda_0 & \ddots & 0 & 0 \\ \dots & \dots & \dots & \ddots & \ddots & \dots \\ 0 & 0 & 0 & \dots & \lambda_0 & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_0 \end{pmatrix}.$$

Жордановой матрицей J называется блочно-диагональная матрица, составленная из жордановых клеток:

$$J = \text{diag}(J_{k_1}(\lambda_1), J_{k_2}(\lambda_1), \dots, J_{k_s}(\lambda_1)).$$

Базис пространства V называется *Жордановым*, если в этом базисе матрица преобразования является Жордановой (имеет *Жорданову форму*). В данном случае базис также называют каноническим базисом преобразования, а саму матрицу преобразования в этом базисе — *каноническим видом*, или *Жордановой формой* преобразования.

6.10.2. Цель

Теорема 6.10.1 (Жордан). Для любого преобразования φ комплексного линейного n -мерного пространства существует Жорданов базис. Жорданова форма определена единственным с точностью до перестановки жордановых клеток образом.

Теорема 6.10.2 (Жордан (матричная формулировка)). Для любой матрицы $A \in \mathbb{C}^{n \times n}$ найдется такая невырожденная матрица Q , что $J = Q^{-1}AQ$ есть жорданова матрица. Матрица J определяется единственным образом с точностью до перестановки жордановых клеток.

6.10.3. Разложение пространства в прямую сумму корневых подпространств

Рассмотрим минимальный многочлен

$$f(\lambda) = (\lambda - \lambda_1)^{k_1} (\lambda - \lambda_2)^{k_2} \dots (\lambda - \lambda_s)^{k_s},$$

аннулирующий преобразование φ на всем пространстве V .

Множество $P_i = \text{Ker}(\varphi - \lambda_i)^{k_i}$ называется *корневым подпространством*, принадлежащим собственному числу (корню) λ_i ($i = 1, 2, \dots, s$).

Теорема 6.10.3. Пространство V есть прямая сумма корневых подпространств P_i :

$$V = P_1 + P_2 + \dots + P_s.$$

ДОКАЗАТЕЛЬСТВО. Легко видеть, что многочлены

$$f_i(\lambda) = \frac{f(\lambda)}{(\lambda - \lambda_i)^{k_i}}, \quad (i = 1, 2, \dots, s)$$

взаимно просты, поэтому найдутся такие $u_1(\lambda), u_2(\lambda), \dots, u_s(\lambda)$, что

$$1 = \sum_{i=1}^s f_i(\lambda) u_i(\lambda),$$

откуда

$$\varepsilon = \sum_{i=1}^s f_i(\varphi) u_i(\varphi),$$

следовательно, для любого $x \in V$

$$x = \sum_{i=1}^s \underbrace{f_i(\varphi) u_i(\varphi) x}_{x_i}.$$

Обозначим $x_i = f_i(\varphi) u_i(\varphi) x$, тогда

$$x = \sum_{i=1}^s x_i. \quad (6.20)$$

Докажем, что формула (6.20) задает разложение произвольного вектора x по прямой сумме $P_1 + \dots + P_s$.

Сперва покажем, что $x_i \in P_i$. Действительно,

$$(\varphi - \lambda_i \varepsilon)^{k_i} x_i = \underbrace{(\varphi - \lambda_i \varepsilon)^{k_i} f_i(\varphi)}_{f(\varphi)=0} u_i(\varphi) x = 0.$$

Теперь докажем, что рассматриваемая сумма прямая. Для этого достаточно установить единственность разложения по пространствам P_i ($i = 1, 2, \dots, s$) для нулевого вектора. Пусть

$$0 = \sum_{j=1}^s x_j, \quad \text{где } x_j \in P_j \quad (6.21).$$

Применим к обеим частям равенства (6.21) преобразование $f_i(\varphi)$, тогда

$$0 = \sum_{j=1}^s f_i(\varphi) x_j = f_i(\varphi) x_i,$$

последнее равенство справедливо в силу

$$f_i(\varphi) x_j = \prod_{\substack{\nu=1 \\ \nu \neq i}} (\varphi - \lambda_\nu \varepsilon)^{k_{n_\nu}} x_j = 0.$$

Так как многочлены $f_i(\lambda)$ и $(\lambda - \lambda_i)^{k_i}$ взаимно простые, то существуют такие $u(\lambda), v(\lambda)$, что

$$1 = u(\lambda) f_i(\lambda) + v(\lambda) (\lambda - \lambda_i)^{k_i},$$

поэтому

$$x_i = u(\varphi) \underbrace{f_i(\varphi) x_i}_0 + v(\varphi) \underbrace{(\varphi - \lambda_i \varepsilon)^{k_i} x_i}_{0, \text{ т.к. } x_i \in P_i} = 0.$$

Итак, компонента разложения $x_i = 0$ ($i = 1, 2, \dots, s$) определяется однозначно, поэтому рассматриваемая сумма прямая. ■

ЗАМЕЧАНИЕ 1. Вместо $f(\lambda)$ в теореме можно рассматривать любой аннулирующий многочлен преобразования φ . Формулировка теоремы и доказательство при этом никак не изменятся.

ЗАМЕЧАНИЕ 2. Корневые пространства P_i ($i = 1, 2, \dots, s$) инвариантны относительно преобразования φ .

ЗАМЕЧАНИЕ 3. Сужение преобразования φ на подпространство P_i (индуцированное преобразование) имеет одно собственное значение λ_i .

6.10.4. Построение Жорданова базиса

В силу замечаний 2, 3 предыдущего пункта мы можем на время ограничиться рассмотрением линейного преобразования φ пространства V с одним собственным числом λ_0 . В этом случае минимальный многочлен имеет вид $f(\lambda) = (\lambda - \lambda_0)^k$. Обозначим $\psi = \varphi - \lambda_0 \varepsilon$, тогда

$$\psi^k = 0. \quad (6.22)$$

Назовем *высотой* вектора $x \in V$ такое $h \geq 0$, что $\psi^{h-1}x \neq 0$, но $\psi^h x = 0$. В нашем случае высота произвольного вектора не превосходит k . Существует единственный вектор высоты 0 — нулевой вектор. Векторы высоты 1 это в точности собственные векторы преобразования.

Обозначим $L_h = \text{Ker } \psi^h$. Иными словами, L_h — это множество векторов высоты $\leq h$. Имеем

$$\{0\} = L_0 \subseteq \underbrace{L_1}_{\substack{\text{собственное} \\ \text{подпространство}}} \subseteq \dots \subseteq L_k = L_{k+1} = \dots = V.$$

Следующая лемма показывает, что $L_h \neq L_{h+1}$ при $h \leq k$, и поэтому,

$$\{0\} = L_0 \subset L_1 \subset \dots \subset L_k = L_{k+1} = \dots = V.$$

Лемма 6.10.1. Если $L_h = L_{h+1}$, то $L_h = L_l$ для любого $l \geq h$.

Доказательство. Воспользуемся методом доказательства от противного. Пусть

$$L_h = L_{h+1} \neq L_{h+2},$$

тогда найдется такой $x \in V$, что

$$x \notin L_h = L_{h+1}, \quad x \in L_{h+2}.$$

Используя определение пространств L_{h+2} и L_{h+1} , соответственно получаем

$$\begin{aligned} \psi^{h+2}x = 0 &\Rightarrow \psi^{h+1}\psi x = 0 \Rightarrow \psi x \in L_{h+1}, \\ \psi^{h+1}x \neq 0 &\Rightarrow \psi^h\psi x \neq 0 \Rightarrow \psi x \notin L_h. \end{aligned}$$

Итак, $\psi x \in L_{h+1}$, однако $\psi x \notin L_h$, что невозможно, так как $L_h = L_{h+1}$. ■

Опишем алгоритм построения Жорданова базиса.

- на предварительном шаге необходимо построить базисы пространств $L_1, L_2, \dots, L_k$;
- далее найти такую линейно независимую систему $a_{11}, \dots, a_{1t_1}$, для которой справедливо соотношение

$$L_k = L_{k-1} + L(a_{11}, \dots, a_{1t_1});$$

- для каждого $h = 2, 3, \dots, k$ необходимо найти линейно независимую систему $a_{h1}, \dots, a_{ht_h}$ такую, что

$$\begin{aligned} L_{k-h+1} &= L_{k-h} \\ &+ L(\psi^{h-1}a_{11}, \dots, \psi^{h-1}a_{1t_1}, \dots, \psi a_{h-1,1}, \dots, \psi a_{h-1,t_{h-1}}, a_{h1}, \dots, a_{ht_h}). \end{aligned}$$

Необходимые соотношения еще раз приведены в таблице 6.1

В настоящем пункте мы докажем, что построенная система векторов

$$\begin{array}{ccccccc} a_{11} & \dots & a_{1t_1} & & & & \\ \psi a_{11} & \dots & \psi a_{1t_1} & \psi a_{21} & \dots & \psi a_{2t_2} & \\ \vdots & & & & & & \\ \psi^{k-1}a_{11} & \dots & \psi^{k-1}a_{1t_1} & \psi^{k-2}a_{21} & \dots & \psi^{k-2}a_{2t_2} & \dots a_{k1} \dots a_{kt_k} \end{array} \quad (6.23)$$

образует Жорданов базис пространства V .

Сперва докажем возможность построения системы (6.23).

Таблица 6.1. Схема построения Жорданова базиса

$$\begin{aligned}
L_k &= L_{k-1} + L(a_{11}, \dots, a_{1t_1}), \\
L_{k-1} &= L_{k-2} + L(\psi a_{11}, \dots, \psi a_{1t_1}, a_{21}, \dots, a_{2t_2}), \\
L_{k-2} &= L_{k-3} + L(\psi^2 a_{11}, \dots, \psi^2 a_{1t_1}, \psi a_{21}, \dots, \psi a_{2t_2}, a_{31}, \dots, a_{3t_3}), \\
&\vdots \\
L_{k-h} &= L_{k-h+1} + L(\psi^{h-2} a_{11}, \dots, \psi^{h-2} a_{1t_1}, \dots, a_{h-1,1}, \dots, a_{h-1,t_{h-1}}), \\
L_{k-h+1} &= L_{k-h} + L(\psi^{h-1} a_{11}, \dots, \psi^{h-1} a_{1t_1}, \dots, \psi a_{h-1,1}, \dots, \psi a_{h-1,t_{h-1}}, a_{h1}, \dots, a_{ht_h}), \\
&\vdots \\
L_1 &= L_0 + L(\psi^{k-1} a_{11}, \dots, \psi^{k-1} a_{1t_1}, \dots, \psi a_{k-1,1}, \dots, \psi a_{k-1,t_k}, a_{k1}, \dots, a_{kt_k}).
\end{aligned}$$

Лемма 6.10.2. Системы векторов

$$a_{h1}, \dots, a_{ht_h} \quad (h = 1, 2, \dots, k),$$

удовлетворяющие вышеупомянутым условиям, существуют.

Доказательство. Лемма доказывается индукцией по h .

Основание индукции. Система $a_{11}, \dots, a_{1t_1}$ является дополнением базиса пространства L_{k-1} до базиса пространства L_k . Так как $L_{k-1} \subset L_k$, то такая система существует.

Индуктивный переход $h-1 \rightarrow h$. Пусть $u_1, \dots, u_p$ — произвольный базис L_{k-h+1} , $v_1, \dots, v_r$ — произвольный базис L_{k-h+2} . Обозначим векторы

$$\psi^{h-2} a_{11}, \dots, \psi^{h-2} a_{1t_1}, \dots, a_{h-1,1}, \dots, a_{h-1,t_{h-1}}$$

через $w_1, \dots, w_q$ соответственно. По предположению индукции векторы

$$u_1, \dots, u_p, w_1, \dots, w_q$$

составляют базис подпространства L_{k-h} ($(h-1)$ -ый шаг алгоритма). Докажем, что векторы

$$v_1, \dots, v_r, \psi w_1, \dots, \psi w_q$$

образуют линейно независимую систему и, следовательно, могут быть дополнены (векторами $a_{h1}, \dots, a_{ht_h}$) до базиса пространства L_{k-h+1} . Для этого рассмотрим линейную комбинацию

$$\sum_{j=1}^r \alpha_j v_j + \sum_{i=1}^q \beta_i \psi w_i = 0. \quad (6.24)$$

Применим преобразование ψ^{k-h+2} к обеим частям (6.24):

$$\sum_{j=1}^r \alpha_j \underbrace{\psi^{k-h+2} v_j}_0 + \psi^{k-h+1} \sum_{i=1}^q \beta_i w_i = 0.$$

Получаем:

$$\psi^{k-h+1} \sum_{i=1}^q \beta_i w_i = 0,$$

следовательно, по определению пространства L_{k-h+1} $\sum_{i=1}^q \beta_i w_i \in L_{k-h+1}$. Раскладывая эту сумму по базису пространства L_{k-h+1} , получаем

$$\sum_{i=1}^q \beta_i w_i = \sum_{j=1}^p \gamma_j u_j \quad (6.25)$$

для некоторых $\gamma_1, \dots, \gamma_p$. Так как по предположению индукции векторы $u_1, \dots, u_p, w_1, \dots, w_q$ линейно независимы, то в (6.25)

$$\beta_i = 0 \quad (i = 1, \dots, q); \quad \gamma_j = 0 \quad (j = 1, \dots, p).$$

Теперь из (6.24) получаем

$$\sum_{j=1}^r \alpha_j v_j = 0.$$

Так как векторы $v_1, \dots, v_r$ линейно независимы, то $\alpha_j = 0$, ($j = 1, \dots, r$). Итак, в линейной комбинации (6.24) коэффициенты равны нулю и, следовательно, векторы $v_1, \dots, v_r, \psi w_1, \dots, \psi w_q$ линейно независимы. ■

Лемма 6.10.3. Система (6.23) образует базис пространства V .

ДОКАЗАТЕЛЬСТВО. По построению L_1 есть линейная оболочка векторов, стоящих в нижней строке таблицы (6.23). Подпространство L_2 есть линейная оболочка векторов, стоящих в двух последних строках. Поднимаясь так далее снизу вверх по таблице (6.23) мы получаем, что пространство $L_k = V$ есть линейная оболочка векторов системы (6.23). ■

Вертикальный ряд векторов в таблице (6.23) назовем *цепочкой*. Векторы в одной цепочке будем рассматривать в последовательности от собственного вектора к вектору максимальной высоты (снизу вверх). Покажем, что каждой цепочке соответствует одна Жорданова клетка. Таким образом, вся совокупность векторов (6.23) является Жордановым базисом.

Лемма 6.10.4. Каждой цепочке векторов $a_1, \dots, a_r$ таблицы (6.23) в матрице преобразования соответствует Жорданова клетка.

ДОКАЗАТЕЛЬСТВО. Для доказательства достаточно заметить, что вектор a_1 — собственный, поэтому $\varphi a_1 = \lambda_0 a_1$, и $a_j = \psi a_{j+1}$ ($j = 2, 3, \dots, r$), поэтому $\varphi a_{j+1} = \lambda_0 a_{j+1} + a_j$. ■

В заключение пункта отметим, что в случае нескольких собственных значений преобразования φ необходимо проделать те же вычисления с каждым собственным числом.

6.10.5. Единственность Жордановой формы

Лемма 6.10.5. Для произвольного натурального i

$$\text{rank}(J_n(\lambda_0) - \lambda E)^i = \begin{cases} n - i, & \text{если } \lambda = \lambda_0, \quad n - i > 0, \\ 0, & \text{если } \lambda = \lambda_0, \quad n - i \leq 0, \\ n, & \text{если } \lambda \neq \lambda_0, \quad n - i \leq 0. \end{cases}$$

Обозначим через $l_j = l_j(\lambda_0)$ число Жордановых клеток вида $J_n(\lambda_0)$ в Жордановой форме J преобразования φ . Пусть $r_j = r_j(\lambda_0) = \text{rank}(\varphi - \lambda_0 \varepsilon)^j$. Следующая лемма показывает, что Жорданова форма преобразования определена однозначно с точностью до перестановки диагональных клеток.

Теорема 6.10.4. Для числа Жордановых клеток порядка j в Жордановой форме J справедливо соотношение

$$l_j(\lambda_0) = r_{j-1}(\lambda_0) - 2r_j(\lambda_0) + r_{j+1}(\lambda_0).$$

ДОКАЗАТЕЛЬСТВО. Выразим величины

$$r_j(\lambda_0) = \text{rank}(\varphi - \lambda_0 \varepsilon)^j = \text{rank}(J - \lambda_0 E)^j$$

через $l_j(\lambda_0)$. При возведении блочно-диагональной матрицы в степень каждый блок возводится в степень независимо и матрица сохраняет блочную структуру. Таким образом, величина $r_j(\lambda_0)$ равна сумме рангов каждого блока. Обозначим через $\bar{r}$ сумму рангов блоков, соответствующих собственным числам, отличным от λ_0 . Пользуясь леммой, получаем

$$r_j(\lambda_0) = l_{j+1} + 2l_{j+2} + \dots + (p-j)l_p + \bar{r} = \sum_{i=j+1}^p (i-j)l_i + \bar{r},$$

откуда

$$\begin{aligned}
 & r_{j-1} - 2r_j + r_{j+1} \\
 &= \sum_{i=j}^p (i-j)l_i - 2 \sum_{i=j+1}^p (i-j)l_i + \sum_{i=j+2}^p (i-j)l_i \\
 &= l_j + 2l_{j+1} + \sum_{i=j+2}^p (i-j)l_i \\
 &\quad - 2 \left(l_{j+1} + \sum_{i=j+2}^p (i-j)l_i \right) \\
 &\quad + \sum_{i=j+2}^p (i-j)l_i \\
 &= l_j.
 \end{aligned}$$

■

Итак, число Жордановых клеток заданного порядка с заданным числом на диагонали полностью определяется формулами из предыдущей теоремы и зависит только от преобразования, но никак не от базиса. Поэтому *Жорданова форма (но не базис) линейного преобразования определены однозначно с точностью до перестановок диагональных клеток.*

6.11. Матричные функции

6.11.1. Значение многочлена от матрицы

Утверждение 6.11.1. Если матрицы A, B подобны, причем $B = Q^{-1}AQ$, тогда для любого многочлена $g(\lambda) \in \mathbb{C}[\lambda]$ выполнено соотношение

$$g(B) = Q^{-1}g(A)Q. \quad (6.26)$$

Утверждение 6.11.2. Для произвольного многочлена $g(\lambda) \in \mathbb{C}[\lambda]$ и произвольной блочно-диагональной матрицы $\text{diag}(A_1, \dots, A_p)$

$$g(\text{diag}(A_1, \dots, A_p)) = \text{diag}(g(A_1), \dots, g(A_p)). \quad (6.27)$$

Утверждение 6.11.3.

$$(J_n(\lambda_0))^m = \begin{pmatrix} \lambda_0^m & \binom{m}{1}\lambda_0^{m-1} & \binom{m}{2}\lambda_0^{m-2} & \dots & \binom{m}{n-1}\lambda_0^{m-n+1} \\ 0 & \lambda_0^m & \binom{m}{1}\lambda_0^{m-1} & \dots & \binom{m}{n-2}\lambda_0^{m-n+2} \\ \dots & & & & \\ 0 & 0 & 0 & \dots & \lambda_0^m \end{pmatrix}. \quad (6.28)$$

ДОКАЗАТЕЛЬСТВО. Утверждение легко доказывается индукцией по m . ■

Утверждение 6.11.4. Для произвольного многочлена $g(\lambda) \in \mathbb{C}[\lambda]$ выполняется соотношение

$$g(J_n(\lambda_0)) = \begin{pmatrix} g(\lambda_0) & \frac{g'(\lambda_0)}{1!} & \frac{g''(\lambda_0)}{2!} & \dots & \frac{g^{(n-1)}(\lambda_0)}{(n-1)!} \\ 0 & g(\lambda_0) & \frac{g'(\lambda_0)}{1!} & \dots & \frac{g^{(n-2)}(\lambda_0)}{(n-2)!} \\ \dots & & & & \\ 0 & 0 & 0 & \dots & g(\lambda_0) \end{pmatrix}. \quad (6.29)$$

ДОКАЗАТЕЛЬСТВО. Если $g(\lambda) = \lambda^m$, то формула 6.28 превращается в 6.29. Далее доказательство не должно вызывать затруднений. ■

Пример 43. Для матрицы

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$$

найти A^{2000} .

Характеристический многочлен в данном случае совпадает с минимальным и имеет вид $f(\lambda) = \lambda^2 - 4\lambda + 4 = (\lambda - 2)^2$, поэтому

$$J = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}, \quad Q = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix}, \quad Q^{-1} = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix}, \quad J = Q^{-1}AQ.$$

Отсюда

$$A = QJQ^{-1}, \quad A^{2000} = QJ^{2000}Q^{-1} = Q \begin{pmatrix} 2^{2000} & 2000 \cdot 2^{1999} \\ 0 & 2^{2000} \end{pmatrix} Q^{-1}.$$

6.11.2. Интерполяционный многочлен Лагранжа-Сильвестера

Рассмотрим таблицу чисел

$$\begin{array}{c|cccc} \lambda_1 & c_{10} & c_{11} & \dots & c_{1,k_1-1} \\ \lambda_2 & c_{20} & c_{21} & \dots & c_{2,k_2-1} \\ \vdots & & & & \\ \lambda_s & c_{s0} & c_{s1} & \dots & c_{s,k_s-1} \end{array}, \quad (6.30)$$

в которой $\lambda_i \neq \lambda_j$ при $i \neq j$. Многочлен $g(x) \in \mathbf{C}[\lambda]$, удовлетворяющий условиям

$$g^{(j)}(\lambda_i) = c_{ij} \quad (i = 1, 2, \dots, s; j = 0, 1, \dots, k_i - 1) \quad (6.31)$$

назовем *интерполяционным многочленом Лагранжа-Сильвестера*, относящимся к интерполяционной таблице (6.30).

Теорема 6.11.1. *Для любой таблицы интерполяции (6.30) существует единственный интерполяционный многочлен Лагранжа-Сильвестера степени меньшей*

$$m = k_1 + \dots + k_s.$$

ДОКАЗАТЕЛЬСТВО. Пусть

$$g(\lambda) = f_1 \lambda^{m-1} + f_2 \lambda^{m-2} + \dots + f_{m-1} \lambda + f_m$$

— искомый интерполяционный многочлен. Из условий (6.31) получаем систему линейных уравнений

$$\begin{cases} f_1 \lambda_1^{m-1} + f_2 \lambda_1^{m-2} + \dots + f_{m-1} \lambda_1 + f_m & = c_{10}, \\ (m-1)f_1 \lambda_1^{m-2} + (m-2)f_2 \lambda_1^{m-3} + \dots + f_{m-1} & = c_{11}, \\ \vdots & \\ f_1 \lambda_s^{m-1} + f_2 \lambda_s^{m-2} + \dots + f_{m-1} \lambda_s + f_m & = c_{s0}, \\ (m-1)f_1 \lambda_s^{m-2} + (m-2)f_2 \lambda_s^{m-3} + \dots + f_{m-1} & = c_{s1}, \\ \vdots & \end{cases} \quad (6.32)$$

рассматриваемую относительно неизвестных $f_1, \dots, f_m$. Необходимо доказать, что система (6.32) имеет единственное решение. Система квадратная и поэтому достаточно доказать существование и единственность решения в частном случае: когда, например, ее правая часть нулевая. Итак, пусть $c_{ij} = 0$ ($i = 1, \dots, s; j = 0, \dots, k_i - 1$). В этом случае необходимо найти интерполяционный многочлен Лагранжа-Сильвестера $g(\lambda)$, для которого $g^{(j)}(\lambda_i) = 0$ ($i = 1, 2, \dots, s; j = 0, 1, \dots, k_i - 1$), иными словами, многочлен, для которого λ_i является корнем кратности не меньше k_i . Таким образом, многочлен $g(\lambda)$ степени меньше m имеет с учетом кратности m корней. Существует лишь один такой многочлен $g(\lambda) = 0$. Итак, система (6.32) с нулевой правой частью имеет единственное решение, следовательно, единственное решение будет существовать и при любой другой правой части. ■

6.11.3. Функции от матриц

Пусть

$$f(\lambda) = (\lambda - \lambda_1)^{k_1} \dots (\lambda - \lambda_s)^{k_s}$$

— минимальный многочлен матрицы A . Будем говорить, что произвольная функция $g(\lambda) : \mathbf{C} \rightarrow \mathbf{C}$ задана на спектре матрицы A , если существует набор значений

$$\begin{aligned} &g(\lambda_1), g'(\lambda_1), \dots, g^{k_1-1}(\lambda_1); \\ &g(\lambda_2), g'(\lambda_2), \dots, g^{k_2-1}(\lambda_2); \\ &\dots \\ &g(\lambda_s), g'(\lambda_s), \dots, g^{k_s-1}(\lambda_s). \end{aligned} \quad (6.33)$$

В данном случае говорят также, что *функция $g(\lambda)$ на спектре матрицы A принимает значения (6.33)*. Обозначим через $g_A(\lambda)$ многочлен минимальной степени, совпадающий на спектре матрицы

A с функцией $g(\lambda)$ (так называемый интерполяционный многочлен Лагранжа-Сильвестера, заданный на спектре матрицы A). Значение $g(A)$ функции $g(\lambda)$ определим как значение многочлена $g_A(\lambda)$ от матрицы A :

$$g(A) = g_A(A).$$

Сразу отметим, что, по предыдущей теореме, *степень интерполяционного многочлена Лагранжа-Сильвестера, заданного на спектре матрицы A , меньше степени минимального многочлена этой матрицы.*

Пример 44. Для матрицы

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$$

найти e^A .

Минимальный многочлен $f(\lambda) = (\lambda - 2)^2$ имеет вторую степень, поэтому ищем интерполяционный многочлен в виде

$$g_A(\lambda) = f_1\lambda + f_2.$$

На спектре матрицы A функция $g(\lambda) = e^\lambda$ принимает следующие значения: $g(2) = e^2$, $g'(2) = e^2$, поэтому

$$\begin{aligned} g_A(2) &= 2f_1 + f_2 = e^2, \\ g'_A(2) &= f_1 = e^2; \end{aligned}$$

откуда получаем $f_1 = e^2$, $f_2 = -e^2$, следовательно, $g_A(\lambda) = e^2\lambda - e^2$, поэтому

$$e^A = g_A(A) = e^2A - e^2E = \begin{pmatrix} 0 & e^2 \\ -e^2 & 2e^2 \end{pmatrix}.$$

Следующее утверждение показывает корректность введенного определения для того случая, когда функция $g(\lambda)$ сама является многочленом.

Теорема 6.11.2. Пусть $g(\lambda)$, $h(\lambda)$ — многочлены из $\mathbf{C}[\lambda]$, A — матрица из $\mathbf{C}^{n \times n}$. Для того, чтобы $g(A) = h(A)$ необходимо и достаточно, чтобы значения многочленов $g(\lambda)$, $h(\lambda)$ совпадали на спектре матрицы A .

Доказательство. Пусть

$$f(\lambda) = (\lambda - \lambda_1)^{k_1} \cdot \dots \cdot (\lambda - \lambda_s)^{k_s}$$

— минимальный многочлен матрицы A , тогда

$$\begin{aligned} &g(A) = h(A) \\ \Leftrightarrow &g(A) - h(A) = 0 \\ \Leftrightarrow &g(\lambda) - h(\lambda) \text{ — аннулирующий многочлен матрицы } A \\ \Leftrightarrow &g(\lambda) - h(\lambda) = p(\lambda)f(\lambda) \text{ для некоторого } p(\lambda) \in \mathbf{C}[\lambda] \\ \Leftrightarrow &\lambda_i \text{ — корень } g(\lambda) - h(\lambda) \text{ кратности } \geq k_i \text{ } (i = 1, 2, \dots, s) \\ \Leftrightarrow &g^{(j)}(\lambda_i) - h^{(j)}(\lambda_i) = 0 \text{ } (i = 1, \dots, s; j = 0, \dots, k_i - 1). \end{aligned}$$

■

Пример 45. Для матрицы

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$$

найти A^{2000} .

Этот пример уже рассматривался. Приведем другой способ его решения. Для функции $g(\lambda) = \lambda^{2000}$ найдем интерполяционный многочлен

$$g_A(\lambda) = f_1\lambda + f_2,$$

заданный на спектре матрицы A . Имеем

$$\begin{aligned} g_A(2) &= 2f_1 + f_2 = 2^{2000}, \\ g'_A(2) &= f_1 = 2000 \cdot 2^{1999}; \end{aligned}$$

откуда $f_1 = 2000 \cdot 2^{1999}$, $f_2 = -1999 \cdot 2^{2000}$, следовательно,

$$g_A(\lambda) = 2000 \cdot 2^{1999}\lambda - 1999 \cdot 2^{2000} = 2^{2000}(1000\lambda - 1999).$$

Подставляя $\lambda = A$ в $g_A(\lambda)$, получаем

$$A^{2000} = 2^{2000} \cdot \begin{pmatrix} 1 & 1000 \\ -1000 & 1001 \end{pmatrix}.$$

Следующие утверждения являются аналогами утверждений 1, 2, 4 из раздела 6.11.1.

Утверждение 6.11.5. *Если матрицы A, B подобны, причем $B = Q^{-1}AQ$, тогда для любой функции $g(\lambda)$, определенной на спектре матрицы A , $g(B)$ существует, причем*

$$g(B) = Q^{-1}g(A)Q.$$

Доказательство. Так как A и B подобны, то у них совпадают минимальные многочлены и, следовательно, наборы значений функции $g(\lambda)$ на их спектре, поэтому $g(B)$ существует и $g_A(\lambda) = g_B(\lambda)$. Теперь по утверждению 1 получаем

$$g(B) = g_B(B) = g_A(B) = Q^{-1}g_A(A)Q = Q^{-1}g(A)Q.$$

■

Утверждение 6.11.6. *Для произвольной функции $g(\lambda)$, определенной на спектре блочно-диагональной матрицы $\text{diag}(A_1, \dots, A_p)$ функция $g(A_i)$ существует для каждого $i = 1, 2, \dots, p$ и*

$$g(\text{diag}(A_1, \dots, A_p)) = \text{diag}(g(A_1), \dots, g(A_p)).$$

Доказательство.

Пусть $f_i(\lambda)$ — минимальный многочлен матрицы A_i ($i = 1, 2, \dots, p$), а $f(\lambda)$ — минимальный многочлен матрицы A . Так как $f(\lambda) = (f_1(\lambda), \dots, f_p(\lambda))$, то все корни многочлена $f_i(\lambda)$ являются корнями многочлена $f(\lambda)$, причем в последнем каждый из корней имеет не меньшую кратность (спектр матрицы A_i содержится в спектре матрицы A). Следовательно, $g(A_i)$ существует, $g_A(\lambda) = g_{A_i}(\lambda)$ и поэтому, используя утверждение 2, получаем

$$\begin{aligned} g(\text{diag}(A_1, \dots, A_p)) &= g_A(\text{diag}(A_1, \dots, A_p)) \\ &= \text{diag}(g_A(A_1), \dots, g_A(A_p)) \\ &= \text{diag}(g_{A_1}(A_1), \dots, g_{A_p}(A_p)) \\ &= \text{diag}(g(A_1), \dots, g(A_p)). \end{aligned}$$

■

Утверждение 6.11.7. *Пусть функция $g(\lambda)$ определена на спектре матрицы $J_n(\lambda_0)$, тогда справедлива формула (6.29).*

Доказательство. Для краткости обозначим $A = J_n(\lambda_0)$. Минимальный многочлен матрицы A равен $f(\lambda) = (\lambda - \lambda_0)^n$, поэтому интерполяционный многочлен $g_A(\lambda)$ удовлетворяет n условиям вида

$$g_A^{(j)}(\lambda_0) = g^{(j)}(\lambda_0) \quad (j = 0, 1, \dots, n-1),$$

откуда по утверждению 4 получаем доказываемое.

■

Билинейные функции

7.1. Определения

Рассмотрим линейное пространство V , заданное над полем $F \subseteq \mathbf{C}$. Функцию

$$f : V \times V \rightarrow F \quad (7.1)$$

назовем *билинейной*, если для любых x, y, z из V и любых α, β из F выполнены соотношения¹

- 1) $f(x + y, z) = f(x, z) + f(y, z)$,
- 2) $f(\alpha x, y) = \alpha f(x, y)$,
- 3) $f(x, y + z) = f(x, y) + f(x, z)$,
- 4) $f(x, \beta y) = \bar{\beta} f(x, y)$.

ЗАМЕЧАНИЕ. Свойства 1-2 означают линейность по первому аргументу. Если $F \subseteq \mathbf{R}$ (вещественный случай) имеем $\beta = \bar{\beta}$ и четвертое свойство примет вид:

$$4' \quad f(x, \beta y) = \beta f(x, y).$$

Тогда свойства 3-4 также означают линейность (по второму аргументу). Если $\mathbf{R} \subset F$ (комплексный случай), то мы можем сказать, что свойства 3-4 означают “полулинейность” по второму аргументу. В силу этого замечания билинейные функции иногда называют *полуторалинейными*. Впрочем тогда билинейными функциями называют функции, обладающие свойствами 1-3,4’.

Обозначим множество всех билинейных функций, действующих в пространстве V , через $\mathcal{F}(V, F)$.

Пусть $f(x, y) \in \mathcal{F}(V, F)$, тогда функция $f(x, x)$ одного аргумента $x \in V$ называется *квадратичной*.

Если $V = F^n$, то билинейные (квадратичные) функции иногда называют билинейными (квадратичными) *формами*.

Пример 46. Легко видеть, что следующие функции являются билинейными

- a) $f(x, y) = x_1 y_1 + \dots + x_n y_n$ в $\mathbf{R}^n$;
- b) $f(x, y) = x_1 \overline{y_1} + \dots + x_n \overline{y_n}$ в $\mathbf{C}^n$;
- c) $f(x, y) = (x, y)$ в евклидовых пространствах $\mathbf{V}_2$ и $\mathbf{V}_3$.

Упражнение 44.

- a) Доказать, что $f(0, x) = 0$ для любой билинейной функции f и любого вектора $x \in V$.
- b) Доказать, что для любых $x_1, \dots, x_m, y_1, \dots, y_l$ из V и любых $\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_l$ из F справедливо

$$f\left(\sum_{i=1}^m \alpha_i x_i, \sum_{j=1}^l \beta_j y_j\right) = \sum_{i=1}^m \alpha_i \sum_{j=1}^l \bar{\beta}_j f(x_i, y_j).$$

- c) Доказать, что для любой матрицы $A \in \mathbf{C}^{n \times n}$ отображение $f : \mathbf{C}^n \times \mathbf{C}^n \rightarrow \mathbf{C}$, заданное формулой $f(x, y) = x^T A \bar{y}$, является билинейной функцией.
- d) Привести пример отображения $\mathbf{R}^n \times \mathbf{R}^n \rightarrow \mathbf{R}$, не являющегося билинейной функцией.

¹ Формула (7.1) означает, что f — функция двух векторных аргументов $x, y \in V$, принимающая числовые значения $f(x, y) \in F$.

7.2. Матрица билинейной функции

В этом разделе мы дадим обозрение всего множества $\mathcal{F}(V, F)$.

1. Определение. Пусть $e_1, \dots, e_n$ — базис пространства V . Матрицу

$$A = (a_{ij}) \in F^{n \times n},$$

в которой

$$a_{ij} = f(e_i, e_j) \quad (i = 1, \dots, n; j = 1, \dots, n),$$

назовем *матрицей билинейной функции* $f \in \mathcal{F}(V, F)$ в базисе e и обозначим $[f]_e = A$.

2. Выражение значения билинейной функции через координаты векторов. Пусть

$$[x]_e = (x_1, \dots, x_n)^\top, \quad [y]_e = (y_1, \dots, y_n)^\top,$$

тогда

$$f(x, y) = f\left(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j\right) = \sum_{i=1}^n x_i \sum_{j=1}^n y_j f(e_i, e_j),$$

или на матричном языке

$$f(x, y) = [x]_e^\top [f]_e \overline{[y]_e}. \quad (7.2)$$

3. Пусть $e_1, \dots, e_n$ — произвольный базис пространства V , A — произвольная матрица из $F^{n \times n}$. Непосредственной проверкой свойств 1-4 убеждаемся, что функция

$$f(x, y) = [x]_e^\top A \overline{[y]_e} \quad (7.3)$$

билинейная. Кроме того, по формуле (7.3)

$$f(e_i, e_j) = (0, \dots, 1, \dots, 0) A \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = a_{ij},$$

следовательно, $A = [f]_e$. Итак, формула (7.2) задает общий вид билинейной функции.

4. Переформулировка результатов пп. 2,3 приводит нас к следующему.

Следствие 7.2.1. *Отображение, ставящее в соответствие всякой билинейной функции его матрицу в некотором фиксированном базисе, является биекцией из $\mathcal{F}(V, F)$ в $F^{n \times n}$.*

5. Связь матриц билинейной функции в разных базисах. Пусть $e_1, e_2, \dots, e_n$ и $e'_1, e'_2, \dots, e'_n$ — два базиса пространства V . Исследуем, как меняется матрица билинейной функции $f \in \mathcal{F}(V, F)$ при переходе от первого базиса ко второму. Для произвольных векторов x, y из V имеем

$$\begin{aligned} f(x, y) &= [x]_e^\top [f]_e \overline{[y]_e} \\ &= (Q_{e \rightarrow e'} [x]_{e'})^\top [f]_e \overline{Q_{e \rightarrow e'} [y]_{e'}} \\ &= [x]_{e'}^\top Q_{e \rightarrow e'}^\top [f]_e \overline{Q_{e \rightarrow e'} [y]_{e'}}. \end{aligned}$$

Сравнивая последнее выражение с (7.3), из утверждения пункта 3 получаем

$$[f]_{e'} = Q_{e \rightarrow e'}^\top [f]_e \overline{Q_{e \rightarrow e'}}.$$

Так как $Q_{e \rightarrow e'}$ — матрица невырожденная, то $\text{rank}[f]_{e'} = \text{rank}[f]_e$. Таким образом, ранг матрицы билинейной функции не зависит от базиса и называется *рангом билинейной функции*. Обозначение: $\text{rank } f$.

Матрицы A и B называются *конгруэнтными*, если существует такая невырожденная матрица Q , что $B = Q^\top A Q$. Матрицы A и B называются *соединенными*, если существует такая невырожденная матрица Q , что $B = Q^\top A \overline{Q}$. Таким образом, соединенные матрицы являются матрицами одной билинейной функции в разных базисах.

6. Элементарные преобразования матрицы, сохраняющие конгруэнтность и соединенность. Напомним, что три типа элементарных преобразований со строками матрицы A можно осуществить домножая A слева на матрицы элементарных преобразований:

- E_{ij} осуществляет перестановку строк,
- $E_i(\alpha)$ — домножение строки на число α ,
- $E_{ij}(\alpha)$ — прибавление к i -ой строке j -ой, умноженной на α .

Каждая из вышеперечисленных матриц получается из единичной в результате того же самого преобразования, которое она осуществляет:

$$E_{ij} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & 0 & & & 1 \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 & & \\ & & & & & & & 1 & 0 & & \\ & & & & & & & & & 1 & \\ & & & & & & & & & & \ddots & \\ & & & & & & & & & & & 1 \end{pmatrix},$$

$$E_i(\alpha) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \alpha & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix},$$

$$E_{ij}(\alpha) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \dots & \alpha & & \\ & & & & \vdots & & \\ & & & & & 1 & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}.$$

Элементарные преобразования со столбцами матрицы A можно осуществить домножая A справа на те же матрицы:

- E_{ij} осуществляет перестановку столбцов,
- $E_i(\alpha)$ — домножение столбца на число α ,
- $E_{ij}(\alpha)$ — прибавление к j -му столбцу i -го, умноженного на α .

Так как

$$\frac{E_{ij}^\top}{E_{ij}} = E_{ij}, \quad \frac{E_i(\alpha)^\top}{E_i(\alpha)} = E_i(\alpha), \quad \frac{E_{ij}(\alpha)^\top}{E_{ij}(\alpha)} = E_{ji}(\alpha),$$

$$\frac{E_{ij}^\top}{E_{ij}} = E_{ij}, \quad \frac{E_i(\alpha)^\top}{E_i(\alpha)} = E_i(\bar{\alpha}), \quad \frac{E_{ij}(\alpha)^\top}{E_{ij}(\alpha)} = E_{ji}(\bar{\alpha}),$$

то каждая из следующих пар элементарных преобразований переводят матрицу A в конгруэнтную:

- перестановка строк с номерами i, j вместе с перестановкой столбцов с номерами i, j .
- умножение строки на α , умножение столбца на α .

- прибавление к i -ой строке j -ой, умноженной на α , прибавление к i -му столбцу j -го, умноженного на α ,

Каждая из следующих пар элементарных преобразований переводит матрицу A в соединенную:

- перестановка строк с номерами i, j , перестановка столбцов с номерами i, j .
- умножение строки на α , умножение столбца на $\bar{\alpha}$.
- прибавление к i -ой строке j -ой, умноженной на α , прибавление к i -му столбцу j -го, умноженного на $\bar{\alpha}$.

7.3. Эрмитовы функции

Билинейную функцию f назовем *эрмитовой*, если

$$f(x, y) = \overline{f(y, x)}$$

для любых x, y из V . В вещественном случае последняя формула примет вид $f(x, y) = f(y, x)$ и эрмитова форма называется *симметрической*.

Матрица $A = a_{ij} \in F^{n \times n}$ называется *эрмитовой*, если $F^\top = \bar{F}$, таким образом, $a_{ij} = \overline{a_{ji}}$, в частности, $a_{ii} \in \mathbf{R}$ ($i = 1, \dots, n$; $j = 1, \dots, n$). Матрица $A \in F^{n \times n}$ называется *симметрической*, если $F^\top = F$.

Утверждение 7.3.1. *Для того, чтобы билинейная функция f была эрмитовой необходимо и достаточно, чтобы матрица $[f]_e$ была эрмитовой (каков бы ни был базис $e_1, \dots, e_n$).*

Доказательство. Пусть f — эрмитова билинейная функция. Тогда для любого базиса $e_1, \dots, e_n$ имеем $f(e_i, e_j) = \overline{f(e_j, e_i)}$, следовательно, матрица $[f]_e$ — эрмитова.

Пусть теперь для некоторого базиса $e_1, \dots, e_n$ матрица $[f]_e$ — эрмитова. Тогда для любых x, y из V получаем

$$\begin{aligned} \overline{f(x, y)} &= \overline{[x]_e^\top [f]_e [y]_e} \\ &= \overline{[x]_e}^\top \overline{[f]_e} \overline{[y]_e} \\ &= ([x]_e^\top [f]_e [y]_e)^\top \\ &= [y]_e^\top [f]_e [x]_e \\ &= f(y, x). \end{aligned}$$

■

Теорема 7.3.1 (Лагранж). *Любая матрица A соединена с некоторой диагональной.*

Доказательство. Опишем *алгоритм Лагранжа* приведения матрицы A к диагональному виду с помощью элементарных преобразований, сохраняющих соединенность.

I. ОБЩИЙ СЛУЧАЙ: $a_{11} \neq 0$. Выполним над матрицей A следующие элементарные преобразования:

$$\begin{aligned} (i) &:= (i) - \frac{a_{i1}}{a_{11}}(1) \quad (i = 2, 3, \dots, n), \\ [j] &:= [j] - \frac{\bar{a}_{j1}}{a_{11}}[1] \quad (j = 2, 3, \dots, n). \end{aligned}$$

Очевидно, что после этих преобразований матрица A перейдет в соединенную матрицу следующего вида

$$\left(\begin{array}{c|c} a_{11} & 0 \\ \hline 0 & B^{(1)} \end{array} \right). \quad (7.4)$$

II. ОСОБЫЙ СЛУЧАЙ: $a_{11} = 0$. Возможны следующие варианты:

1. $a_{i1} = a_{1i} = 0$ ($i = 2, 3, \dots, n$). Матрица A уже имеет вид (7.4). В данном случае никаких действий производить не нужно.

2. $\exists k \ a_{k1} = \overline{a_{1k}} \neq 0$.

- 1) $a_{kk} \neq 0$. Переставляем строчки с номерами 1 и k , переставляем столбцы с номерами 1 и k . После таких преобразований матрица A переходит в соединенную матрицу, над которой выполняем преобразования шага I.
- 2) $a_{kk} = 0$.
 - а) $a_{k1} = \alpha + i\beta$, $a_{1k} = \alpha - i\beta$, $\alpha \neq 0$. Выполним над матрицей A следующие элементарные преобразования

$$(1) := (1) + (k), \quad [1] := [1] + [k].$$

Матрица A перейдет в соединенную матрицу $A' = (a'_{ij})$, в которой

$$a'_{11} = a_{k1} + a_{1k} = 2\alpha \neq 0.$$

Над матрицей A' выполним преобразования шага I.

- б) $a_{k1} = i\beta$, $a_{1k} = -i\beta$, $\beta \neq 0$. Выполним над матрицей A следующие элементарные преобразования:

$$(1) := (1) + i(k), \quad [1] := [1] - i[k].$$

Матрица A перейдет в соединенную матрицу $A' = (a'_{ij})$, в которой

$$a'_{11} = ia_{k1} - ia_{1k} = -2\beta \neq 0.$$

Над матрицей A' выполним преобразования шага I.

После выполнения описанных здесь действий матрица A перейдет в соединенную матрицу (7.4). Далее достаточно те же действия провести с матрицей $B^{(1)}$ и т.д. ■

Следствие 7.3.1. Любая матрица $A \in \mathbb{C}^{n \times n}$ соединена с некоторой диагональной матрицей с диагональными элементами $0, \pm 1$.

ДОКАЗАТЕЛЬСТВО. Можем считать, что матрица A уже имеет диагональный вид:

$$A = \text{diag}(d_1, \dots, d_n).$$

Для каждого $i = 1, 2, \dots, n$, если $d_i \neq 0$, выполним следующие преобразования с матрицей A :

$$(i) := \frac{(i)}{\sqrt{|d_i|}}; \quad [i] := \frac{[i]}{\sqrt{|d_i|}}.$$

При этом исходная матрица переходит в соединенную матрицу, обладающую вышеуказанным свойством. ■

Базис $e_1, e_2, \dots, e_n$ называется *каноническим* для эрмитовой функции $f(x, y)$, если

$$f(e_i, e_j) = 0 \quad (i, j = 1, 2, \dots, n; i \neq j),$$

иными словами, если матрица $[f]_e$, называемая в данном случае *каноническим видом функции f* , диагональна.

Следствие 7.3.2. Для любой эрмитовой билинейной функции $f(x, y)$ существует канонический базис.

ДОКАЗАТЕЛЬСТВО. Для матрицы $[f]_e$ построим соединенную с ней диагональную матрицу B такую, что $B = Q^T [f]_e \bar{Q}$ для некоторой невырожденной матрицы Q . Осталось рассмотреть Q как матрицу перехода к новому базису e' , тогда $[f]_{e'} = B$, следовательно, базис e' — канонический ■

Базис $e_1, e_2, \dots, e_n$ называется *нормальным* для эрмитовой функции $f(x, y)$, если

$$f(e_i, e_j) = \begin{cases} 0, & \text{если } i \neq j, \\ \pm 1 \text{ или } 0, & \text{если } i = j. \end{cases}$$

В данном случае матрица $[f]_e$ называется *нормальным видом функции f* .

Следствие 7.3.3. Для любой эрмитовой билинейной функции $f(x, y)$ существует нормальный базис.

Доказательство. Доказательство следствия аналогично доказательству предыдущего утверждения. ■

Пусть $A = [f]$ есть матрица билинейной функции f в некотором базисе. Обозначим через Δ_i миноры матрицы $A = (a_{ij})$ следующего вида

$$\Delta_i = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1i} \\ a_{21} & a_{22} & \dots & a_{2i} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ii} \end{vmatrix}.$$

Миноры Δ_i называются *угловыми минорами* матрицы A .

Теорема 7.3.2 (Якбби). Пусть $\Delta_i \neq 0$ ($i = 1, 2, \dots, r = \text{rank } f$). Существует канонический базис $e_1, \dots, e_n$, для которого

$$f(e_i, e_i) = \begin{cases} \Delta_1, & \text{если } i = 1, \\ \frac{\Delta_i}{\Delta_{i-1}}, & \text{если } i = 2, \dots, r. \end{cases}$$

Доказательство. Докажем по индукции, что в нашем случае алгоритм Лагранжа не изменяет угловых миноров матрицы A . Так как $a_{11} = \Delta_1 \neq 0$, то на первом этапе выполняются преобразования шага I (общий случай). Эти преобразования (строка вычитается из строк, расположенных ниже; столбец вычитается из столбцов, расположенных правее) не изменяют угловых миноров матрицы A . По окончании преобразований матрица A переходит в матрицу вида (7.4), в которой $\Delta_2 = a'_{11}b'_{11} \neq 0$, откуда $b'_{11} \neq 0$.

На k -ом этапе матрица приобретает вид $\text{diag}(a'_{11}, \dots, a'_{kk}, B_k)$. По предположению индукции угловые миноры этой матрицы совпадают с Δ_i . Так как $\Delta_{k+1} = a'_{11} \cdot \dots \cdot a'_{kk} \cdot b'_{11} \neq 0$, то $b'_{11} \neq 0$ и снова в Алгоритме Лагранжа выполняются преобразования шага I (общий случай), которые не изменяют угловых миноров.

Окончательно получаем матрицу $\text{diag}(d_1, \dots, d_r, 0, \dots, 0)$, в которой $\Delta_i = d_1 \cdot \dots \cdot d_i$ ($i = 1, 2, \dots, r$), откуда получаем доказываемое. ■

Теорема 7.3.3 (Закон инерции). Нормальный вид билинейной функции определен однозначно с точностью до перестановок диагональных элементов.

Доказательство. Пусть $e_1, \dots, e_n, e'_1, \dots, e'_n$ — два базиса пространства V , такие, что

$$f(x, y) = \sum_{j=1}^t x_j \bar{y}_j - \sum_{j=t+1}^r x_j \bar{y}_j, \quad (7.5)$$

$$f(x, y) = \sum_{j=1}^{t'} x'_j \bar{y}'_j - \sum_{j=t'+1}^{r'} x'_j \bar{y}'_j, \quad (7.6)$$

где

$$\begin{aligned} [x]_e &= (x_1, \dots, x_n)^\top, & [x]_{e'} &= (x'_1, \dots, x'_n)^\top, \\ [y]_e &= (y_1, \dots, y_n)^\top, & [y]_{e'} &= (y'_1, \dots, y'_n)^\top. \end{aligned}$$

Имеем $\text{rank } f = r = r'$. Предположим, что $t > t'$. Обозначим

$$L_1 = L(e_1, \dots, e_t), \quad L_2 = L(e_{t'+1}, \dots, e_n).$$

Имеем

$$\dim(L_1 \cap L_2) = \underbrace{\dim L_1}_t + \underbrace{\dim L_2}_{n-t'} - \underbrace{\dim(L_1 + L_2)}_{\leq} \geq t - t' > 0,$$

поэтому найдется $x \neq 0$ такой, что $x \in L_1 \cap L_2$. Так как $x \in L_1$, то $x_j = 0$ ($j = t+1, \dots, n$), поэтому из (7.5) получаем

$$f(x, x) = \sum_{j=1}^t x_j \bar{x}_j = \sum_{j=1}^t |x_j|^2 > 0.$$

Однако, так как $x \in L_2$, то $x'_j = 0$ ($j = 1, \dots, t'$), поэтому из (7.6) получаем

$$f(x, x) = \sum_{j=t'+1}^r x'_j \bar{x}'_j = - \sum_{j=1}^t |x'_j|^2 \leq 0.$$

Полученное противоречие завершает доказательство теоремы. ■

Назовем *положительным индексом инерции* s_+ (*отрицательным индексом инерции* s_-) число положительных (отрицательных) диагональных элементов в каноническом виде билинейной функции. *Сигнатурой эрмитовой билинейной функции* назовем величину $\sigma = s_+ - s_-$. Из теоремы инерции следует, что положительный и отрицательный индекс инерции, а, следовательно, сигнатура, есть величина, не зависящая от базиса.

7.4. Знакоопределенные эрмитовы функции

Эрмитова функция $f(x, y)$ называется

- *положительно определенной* ($f > 0$), если $f(x, x) > 0$,
- *положительно полуопределенной* ($f \geq 0$), если $f(x, x) \geq 0$,
- *отрицательно определенной* ($f < 0$), если $f(x, x) < 0$,
- *отрицательно полуопределенной* ($f \leq 0$), если $f(x, x) \leq 0$

для любого ненулевого вектора x .

Утверждение 7.4.1. • $f > 0$ тогда и только тогда, когда $s_+ = n$,

- $f \geq 0$ тогда и только тогда, когда $s_- = 0$,
- $f < 0$ тогда и только тогда, когда $s_- = n$,
- $f \leq 0$ тогда и только тогда, когда $s_+ = 0$.

Доказательство. Докажем, например, первое утверждение.

Необходимость. Пусть $s_+ < n$, тогда в некотором (нормальном) базисе

$$f(x, y) = \sum_{j=1}^{s_+} x_j \bar{y}_j - \sum_{j=s_++1}^r x_j \bar{y}_j.$$

Для вектора $x \neq 0$ такого, что $x_j = 0$ ($j = 1, \dots, s_+$) и $x_j \neq 0$ ($j = s_++1, \dots, n$), получаем $f(x, x) \leq 0$, что противоречит условию.

Достаточность. Пусть $s_+ = n$, тогда в некотором (нормальном) базисе

$$f(x, y) = \sum_{j=1}^n x_j \bar{y}_j,$$

откуда $f(x, x) > 0$ для любого ненулевого x . ■

Теорема 7.4.1 (Критерий Сильвестера). Эрмитова функция положительно определена тогда и только тогда, когда в любом базисе $\Delta_i > 0$ ($i = 1, \dots, n$).

ДОКАЗАТЕЛЬСТВО.

НЕОБХОДИМОСТЬ. Пусть функция положительно определена, тогда в любом базисе диагональные элементы матрицы этой функции положительны. Следовательно, во время приведения алгоритмом Лагранжа матрицы к каноническому виду никогда не возникает особого случая, поэтому угловые миноры Δ_i не изменяются. Однако по предыдущему утверждению $\Delta_i = d_1 \cdot \dots \cdot d_n > 0$.

ДОСТАТОЧНОСТЬ. Достаточность немедленно следует из теоремы Якоби. ■

Евклидовы пространства

8.1. Определения

Линейное пространство V , заданное над полем $F \subseteq \mathbf{C}$, называется *унитарным*, если задано отображение $V \times V \rightarrow F$, ставящее каждой паре векторов $x, y \in V$ число $(x, y) \in F$, называемое *скалярным произведением*, обладающее следующими свойствами:

- 1) $(x, y) = \overline{(y, x)}$,
- 2) $(x + y, z) = (x, z) + (y, z)$,
- 3) $(\alpha x, y) = \alpha(x, y)$,
- 4) $x \neq 0 \Rightarrow (x, x) > 0$

для любых x, y, z из V и α из F . Свойства (1-4) называются аксиомами унитарного пространства.

В случае $F \subseteq \mathbf{R}$ 1-ая аксиома приобретает вид

$$1') (x, y) = (y, x)$$

и унитарное пространство называется *евклидовым*. Свойства (1', 2-4) называются аксиомами евклидова пространства.

Утверждение 8.1.1. *Скалярное произведение в унитарном пространстве есть положительно определенная эрмитова билинейная функция. И наоборот, любую положительно определенную эрмитову (симметрическую) функцию можно выбрать в качестве скалярного произведения.*

Доказательство. Аксиомы 2, 3 означают линейность по первому аргументу. Далее,

$$\begin{aligned} (x, y + z) &= \overline{(y + z, x)} && \text{(по аксиоме 1)} \\ &= \overline{(y, x) + (z, x)} && \text{(по аксиоме 2)} \\ &= \overline{(y, x)} + \overline{(z, x)} \\ &= (x, y) + (x, z) && \text{(по аксиоме 2)} \end{aligned}$$

для любых x, y, z из V . Аналогично можно показать, что

$$(x, \alpha y) = \overline{\alpha}(x, y)$$

для любых x, y , из V и α из F . Итак скалярное произведение (x, y) является билинейной функцией. По аксиоме 1 эта функция эрмитова и по аксиоме 4 — положительно определенная.

Вторая часть утверждения очевидна. ■

Пример 47. Следующие эрмитовы билинейные функции являются симметрическими и, следовательно, их можно выбрать в качестве скалярного произведения:

- а) стандартное скалярное произведение $(x, y) = |x||y|\cos\varphi$ в пространствах $\mathbf{V}_2, \mathbf{V}_3$;

b)

$$(x, y) = \sum_{j=1}^n x_j y_j, \quad \text{где } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

в пространстве $\mathbf{R}^n$;

c)

$$(x, y) = \sum_{j=1}^n x_j \bar{y}_j,$$

в пространстве $\mathbf{C}^n$;

d)

$$(f, g) = \int_a^b f(t)g(t)dt$$

в пространстве $\mathbf{R}(a, b)$ непрерывных функций, заданных на отрезке $[a, b]$.Приведенные функции будем называть *стандартным скалярным произведением* в соответствующих пространствах.

8.2. Матрица Грама

Матрицей Грама системы векторов $a_1, \dots, a_k$ называется матрица

$$\Gamma(a_1, \dots, a_k) = (\alpha_{ij}) \in F^{k \times k}, \quad \text{где } \alpha_{ij} = (a_i, a_j).$$

Если $e_1, \dots, e_n$ — базис, то $\Gamma(e_1, \dots, e_n)$ — матрица билинейной функции $f(x, y) = (x, y)$, записанная в этом базисе. Поэтому справедливо утверждение:**Утверждение 8.2.1.**

$$(x, y) = [x]_{\mathbf{e}}^T \Gamma(e_1, \dots, e_n) \overline{[y]_{\mathbf{e}}}.$$

Утверждение 8.2.2 (Свойство определителя Грама). 1) Если система $a_1, \dots, a_k$ линейно независима, то $\det \Gamma(a_1, \dots, a_k) > 0$.2) Если система $a_1, \dots, a_k$ линейно зависима, то $\det \Gamma(a_1, \dots, a_k) = 0$.**Доказательство.** 1) Дополним линейно независимую систему $a_1, \dots, a_k$ до базиса пространства V . По критерию Сильвестера k -ый угловой минор матрицы билинейной функции $f(x, y) = (x, y)$, построенной в этом базисе, положителен:

$$\Delta_k = \det \Gamma(a_1, \dots, a_k) > 0.$$

2) Предположим теперь, что система $a_1, \dots, a_k$ линейно зависима, тогда найдутся числа $\alpha_1, \dots, \alpha_k$, не равные нулю одновременно, такие, что

$$\sum_{j=1}^k \alpha_j a_j = 0. \quad (8.1)$$

Скалярно домножая справа обе части равенства (8.1) на a_i ($i = 1, \dots, k$), получаем

$$\begin{cases} \alpha_1(a_1, a_1) + \alpha_2(a_2, a_1) + \dots + \alpha_k(a_k, a_1) = 0, \\ \alpha_2(a_1, a_2) + \alpha_2(a_2, a_2) + \dots + \alpha_k(a_k, a_2) = 0, \\ \dots \\ \alpha_k(a_1, a_k) + \alpha_k(a_2, a_k) + \dots + \alpha_k(a_k, a_k) = 0. \end{cases} \quad (8.2)$$

Рассмотрим равенства (8.2) как систему линейных уравнений относительно $\alpha_1, \dots, \alpha_k$. Система имеет нетривиальное (ненулевое) решение, следовательно, определитель этой системы, совпадающий с $\det \Gamma(a_1, \dots, a_k)$, равен нулю. ■Величина $|x| = \sqrt{(x, x)}$ называется *длиной* или *нормой* вектора x . Часто для норм векторов используют обозначение $\|x\|$. Из 4-ой аксиомы унитарного пространства следует, что $|x| = 0$ тогда и только тогда, когда $x = 0$.

Теорема 8.2.1 (Неравенство Коши-Буняковского-Шварца). Для любых векторов a, b унитарного пространства V выполнено неравенство

$$|(a, b)| \leq |a| \cdot |b|.$$

Доказательство. По свойству определителя Грама

$$0 \leq \Gamma(a, b) = \begin{vmatrix} (a, a) & (a, b) \\ (b, a) & (b, b) \end{vmatrix} = (a, a)(b, b) - (a, b)(b, a) = |a|^2|b|^2 - |(a, b)|^2,$$

Откуда $|(a, b)|^2 \leq |a|^2|b|^2$. ■

Пример 48. Рассмотрим неравенство Коши-Буняковского-Шварца в пространствах $\mathbf{V}_2, \mathbf{V}_3, \mathbf{R}^n, \mathbf{R}(a, b)$, с введенными в них стандартными скалярными произведениями:

- а) $|(a, b)| \leq |a||b|$ в пространствах $\mathbf{V}_2, \mathbf{V}_3$ следует также из определения $(a, b) = |a||b|\cos\varphi$;
 б)

$$\left(\sum_{j=1}^n x_j y_j \right)^2 \leq \sum_{j=1}^n x_j^2 \sum_{j=1}^n y_j^2$$

- в пространстве $\mathbf{R}^n$;
 в))

$$\left(\int_a^b f(t)g(t)dt \right)^2 \leq \int_a^b f^2(t)dt \int_a^b g^2(t)dt$$

в пространстве $\mathbf{R}(a, b)$.

Упражнение 45. Запишите неравенство Коши-Буняковского-Шварца в пространстве $\mathbf{C}^n$ со стандартным скалярным произведением.

Теорема 8.2.2 (Неравенство треугольника). Для любых векторов a, b унитарного пространства V выполнено неравенство

$$|a + b| \leq |a| + |b|.$$

Доказательство.

$$\begin{aligned} |a + b|^2 &= (a + b, a + b) \\ &= (a, a) + (a, b) + (b, a) + (b, b) \\ &= |a|^2 + (a, b) + \overline{(a, b)} + |b|^2 \\ &= |a|^2 + 2\operatorname{Re}(a, b) + |b|^2 \\ &\leq |a|^2 + 2|(a, b)| + |b|^2 \\ &\leq |a|^2 + 2|a||b| + |b|^2 && \text{так как } |(a, b)| \leq |a||b| \\ &= (|a| + |b|)^2. \end{aligned}$$

■

Упражнение 46. Дайте геометрическую интерпретацию неравенству треугольника в пространствах $\mathbf{V}_2, \mathbf{V}_3$.

8.3. Ортогональность

Векторы a, b называются *ортогональными*, или *перпендикулярными*, если $(a, b) = 0$. Обозначение: $a \perp b$. Система векторов $a_1, \dots, a_k$ называется *ортогональной*, если $(a_i, a_j) = 0$ при $i \neq j$, $i = 1, \dots, k, j = 1, \dots, k$. Система векторов $a_1, \dots, a_k$ называется *ортонормированной*, если

$$(a_i, a_j) = \begin{cases} 0, & \text{если } i \neq j \quad (\text{т.е. } a_i \perp a_j), \\ 1, & \text{если } i = j \quad (\text{т.е. } |a_i| = 1). \end{cases}$$

Теорема 8.3.1 (Теорема Пифагора). Если $a \perp b$, то $|a + b|^2 = |a|^2 + |b|^2$.

Доказательство.

$$|a + b|^2 = (a + b, a + b) = (a, a) + \underbrace{(a, b)}_0 + \underbrace{(b, a)}_0 + (b, b) = |a|^2 + |b|^2.$$

■

Теорема 8.3.2 (Теорема Пифагора (обобщение)). Если система $a_1, \dots, a_k$ ортогональна, то

$$|a_1 + \dots + a_k|^2 = |a_1|^2 + \dots + |a_k|^2.$$

ДОКАЗАТЕЛЬСТВО. Доказательство легко проводится по индукции. ■

Утверждение 8.3.1. Ортогональная система ненулевых векторов линейно независима. Ортонормированная система линейно независима.

ДОКАЗАТЕЛЬСТВО. Очевидно, матрица Грама ортогональной системы ненулевых векторов — диагональная с ненулевыми диагональными элементами. Матрица Грама ортонормированной системы единичная. По свойству определителя Грама в обоих случаях системы линейно независимы. ■

8.4. Ортогональный и ортонормированный базисы

Базис $e_1, \dots, e_n$ называется *ортогональным (ортонормированным)*, если он представляет собой ортогональную (соответственно ортонормированную) систему.

В ортогональном базисе матрица Грама диагональная, поэтому

$$(x, y) = [x]_{\mathbf{e}}^{\top} \underbrace{\Gamma(e_1, \dots, e_n)}_{\text{diag}(d_1, \dots, d_n)} [\bar{y}]_{\mathbf{e}} = d_1 x_1 \bar{y}_1 + \dots + d_n x_n \bar{y}_n,$$

где $[x]_{\mathbf{e}} = (x_1, \dots, x_n)^{\top}$, $[\bar{y}]_{\mathbf{e}} = (y_1, \dots, y_n)^{\top}$.

В ортонормированном базисе матрица Грама единичная, поэтому

$$(x, y) = x_1 \bar{y}_1 + \dots + x_n \bar{y}_n.$$

Теорема 8.4.1. В любом унитарном пространстве существует ортогональный (ортонормированный) базис.

ДОКАЗАТЕЛЬСТВО. Канонический базис билинейной функции $f(x, y) = (x, y)$ является ортогональным, а нормальный базис той же билинейной функции — ортонормированным. Утверждение следует теперь из теоремы Лагранжа. ■

Ниже мы опишем алгоритм нахождения ортонормированного базиса любого подпространства унитарного пространства.

Утверждение 8.4.1. Пусть $e_1, \dots, e_n$ — ортогональный (ортонормированный) базис и $a = \alpha_1 e_1 + \dots + \alpha_n e_n$, тогда

$$\alpha_j = \frac{(a, e_j)}{(e_j, e_j)}, \quad (\text{соответственно } \alpha_j = (a, e_j)). \quad (8.3)$$

ДОКАЗАТЕЛЬСТВО. Домножая равенство $a = \alpha_1 e_1 + \dots + \alpha_n e_n$ справа на e_j ($j = 1, \dots, n$) и раскладывая по множителям, получаем $(a, e_j) = \alpha_1 (e_1, e_j) + \dots + \alpha_n (e_n, e_j)$. Так как базис ортогональный (ортонормированный), то от всей суммы в правой части этого равенства останется только одно слагаемое, откуда и следуют доказываемые формулы. ■

Пример 49. В пространствах $\mathbf{V}_2, \mathbf{V}_3$ со стандартным скалярным произведением j -ая координата вектора в ортонормированном базисе $e_1, \dots, e_n$ равна $(x, e_j) = |x| \cos \varphi$, где φ — угол между вектором x и e_j .

8.5. Унитарный изоморфизм

Унитарные пространства V и V' , заданные над полем F называются *унитарно изоморфными*, если существует такой изоморфизм $\varphi: V \rightarrow V'$, при котором

$$(\varphi x, \varphi y) = (x, y)$$

для любых векторов x, y из V .

Упражнение 47. Доказать, что унитарный изоморфизм обладает свойствами рефлексивности, симметричности и транзитивности.

Теорема 8.5.1. *Унитарные пространства, заданные над одним и тем же полем изоморфны тогда и только тогда, когда равны их размерности.*

ДОКАЗАТЕЛЬСТВО. Необходимость следует из свойств изоморфизма линейных пространств. Для доказательства достаточности покажем, что произвольное унитарное пространство V , заданное над полем F унитарно изоморфно арифметическому пространству F^n со стандартным скалярным произведением далее необходимо воспользоваться транзитивностью унитарного изоморфизма. Выберем в V произвольный ортонормированный базис $e_1, \dots, e_n$ и определим изоморфизм φ по правилу: $\varphi x = [x]_e$. Тогда

$$(\varphi x, \varphi y) = \sum_{j=1}^n x_j y_j = (x, y).$$

■

8.6. Ортогональные суммы и ортогональные дополнения

Будем говорить, что множества S и T векторов унитарного пространства V *ортогональны* и писать $S \perp T$, если $(a, b) = 0$ для любых $a \in S$, $b \in T$. Условие $\{a\} \perp T$ будем записывать $a \perp T$.

Утверждение 8.6.1.

$$a \perp L(a_1, \dots, a_k) \Leftrightarrow a \perp a_i \quad (i = 1, \dots, k).$$

ДОКАЗАТЕЛЬСТВО. Необходимость очевидна. Докажем достаточность. Пусть

$$x = x_1 a_1 + \dots + x_k a_k \in L(a_1, \dots, a_k),$$

тогда

$$(x, a) = x_1 \underbrace{(a_1, a)}_0 + \dots + x_k \underbrace{(a_k, a)}_0 = 0.$$

■

Будем говорить, что множества $S_i \subseteq V$ ($i = 1, \dots, k$) *ортогональны*, если $S_i \perp S_j$ ($i, j = 1, \dots, k$; $i \neq j$). Сумму ортогональных подпространств W_i ($i = 1, \dots, k$) назовем *ортогональной* и будем обозначать

$$W_1 + \dots + W_k = W_1 \oplus \dots \oplus W_k.$$

Утверждение 8.6.2. *Ортогональная сумма подпространств является прямой суммой:*

$$W_1 \oplus \dots \oplus W_k = W_1 \dot{+} \dots \dot{+} W_k.$$

ДОКАЗАТЕЛЬСТВО. Покажем, что пересечение ортогональных подпространств содержит лишь нулевой вектор. Пусть

$$x \in W_1 \cap \dots \cap W_k.$$

Тогда $(x, x) = 0$, откуда $x = 0$.

■

Ортогональным дополнением $W^\perp$ подпространства $W \subseteq V$ называется множество всех векторов из V , ортогональных с каждым вектором из W :

$$W^\perp = \{x \in V : x \perp W\}.$$

Теорема 8.6.1. *Для любого подпространства W унитарного пространства V ортогональное дополнение является подпространством и*

$$V = W \oplus W^\perp.$$

ДОКАЗАТЕЛЬСТВО. Выберем в V произвольный ортонормированный базис $e_1, \dots, e_n$, а в W произвольный базис $a_1, \dots, a_k$, тогда условия ортогональности произвольного вектора $x \in V$ подпространству W примут вид $(x, a_i) = 0$ ($i = 1, \dots, k$) или в координатах

$$\begin{cases} x_1 \bar{a}_{11} + x_2 \bar{a}_{12} + \dots + x_n \bar{a}_{1k} = 0, \\ x_1 \bar{a}_{21} + x_2 \bar{a}_{22} + \dots + x_n \bar{a}_{2k} = 0, \\ \dots \\ x_1 \bar{a}_{k1} + x_2 \bar{a}_{k2} + \dots + x_n \bar{a}_{kk} = 0, \end{cases} \quad (8.4)$$

где

$$\begin{aligned} [a_i]_{\mathbf{e}} &= (a_{i1}, \dots, a_{ik})^\top \quad (i = 1, \dots, k) \\ [x]_{\mathbf{e}} &= (x_{i1}, \dots, x_{ik})^\top. \end{aligned}$$

Множество $M(A, 0)$ решений системы линейных уравнений (8.4) есть множество столбцов координат всех векторов из $W^\perp$. Матрица этой системы составлена из координат базисных векторов $a_1, \dots, a_k$ и поэтому имеет ранг k , откуда $\dim W^\perp = \dim M(A, 0) = n - k$, и так как сумма $W \oplus W^\perp$ — прямая, то $\dim(W \oplus W^\perp) = k + (n - k) = n$, иными словами $W \oplus W^\perp = V$. ■

Следствие 8.6.1. Для любого вектора a из V существуют и единственны векторы $b \in W$ и $c \in W^\perp$, такие, что $a = b + c$.

Согласно формулировке следствия называется *ортогональной проекцией* вектора a на подпространство W , вектор c называется *перпендикуляром* или *ортогональной составляющей* вектора a на подпространство W . Обозначения: $b = \text{pr}_W a$, $c = \text{ort}_W a$. Итак,

$$a = \underbrace{\text{pr}_W a}_{\in W} + \underbrace{\text{ort}_W a}_{\in W^\perp}.$$

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

Упражнение 48. Пусть W, W_1, W_2 — подпространства унитарного пространства V . Доказать утверждения

- 1) $V^\perp = \{0\}$,
- 2) $(W^\perp)^\perp = W$,
- 3) $W_1 \subseteq W_2 \Leftrightarrow W_2^\perp \subseteq W_1^\perp$,
- 4) $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$,
- 5) $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$.

8.7. Метод нахождения проекции и перпендикуляра

Пусть $a_1, \dots, a_k$ — произвольный базис подпространства W , a — произвольный вектор из V . Для некоторых пока не определенных $\alpha_1, \dots, \alpha_k$ из F имеем

$$\begin{aligned} \text{pr}_W a &= \alpha_1 a_1 + \dots + \alpha_k a_k; \\ \text{ort}_W a &= a - \text{pr}_W a; \\ (\text{ort}_W a, a_j) &= 0; \quad (j = 1, \dots, k). \end{aligned} \quad (8.5)$$

Откуда

$$(a - \text{pr}_W a, a_j) = 0 \Rightarrow (\text{pr}_W a, a_j) = (a, a_j) \quad (j = 1, \dots, k).$$

Перепишем последние равенства:

$$\begin{cases} \alpha_1 (a_1, a_1) + \alpha_2 (a_2, a_1) + \dots + \alpha_k (a_k, a_1) = (a, a_1), \\ \alpha_1 (a_1, a_2) + \alpha_2 (a_2, a_2) + \dots + \alpha_k (a_k, a_2) = (a, a_2), \\ \dots \\ \alpha_1 (a_1, a_k) + \alpha_2 (a_2, a_k) + \dots + \alpha_k (a_k, a_k) = (a, a_k) \end{cases} \quad (8.6)$$

Матрица системы линейных уравнений (8.6) совпадает с транспонированной матрицей Грама $\Gamma(a_1, \dots, a_k)$, причем векторы $a_1, \dots, a_k$ линейно независимы, поэтому она невырождена и система (8.6) имеет единственное решение. Векторы $\text{pr}_W a$ и $\text{ort}_W a$ находятся теперь по двум первым формулам (8.5).

ЗАМЕЧАНИЕ 1. Рассуждения данного пункта являются новым независимым доказательством теоремы о разложении пространства в прямую сумму подпространства и его ортогонального дополнения.

ЗАМЕЧАНИЕ 2. Заметим, что описанный здесь алгоритм работает и в случае, когда $a_1, \dots, a_k$ — линейно зависимые и $W = L(a_1, \dots, a_k)$. В силу существования проекции система (8.6) имеет решение, но не единственное. Однако по любому решению вектор $\text{pr}_W a$ определяется однозначно.

ЗАМЕЧАНИЕ 3. Пусть $a_1, \dots, a_k$ — столбцы матрицы $A \in F^{n \times k}$, рассматриваемые как векторы унитарного пространства F^n со стандартным скалярным произведением, $a \in F^n$, тогда

$$\begin{aligned}\Gamma(a_1, \dots, a_k) &= A^\top \bar{A}, \\ \Gamma(a_1, \dots, a_k)^\top &= \bar{A}^\top A\end{aligned}$$

и система (8.6) приобретает вид:

$$\bar{A}^\top A y = \bar{A}^\top a \text{ где } y = (\alpha_1, \dots, \alpha_k)^\top. \quad (8.7)$$

С помощью решения y последней системы можно найти искомую проекцию:

$$\text{pr}_W a = A y.$$

ЗАМЕЧАНИЕ 4. Так как

$$\text{ort}_W a = a - \alpha_1 a_1 - \dots - \alpha_k a_k,$$

то

$$L(a_1, \dots, a_k, a) = L(a_1, \dots, a_k, \text{ort}_W a)$$

и, следовательно,

- 1) если $a \notin L(a_1, \dots, a_k)$, то $\text{ort}_W a \neq 0$,
- 2) если $a \in L(a_1, \dots, a_k)$, то $\text{ort}_W a = 0$.

ЗАМЕЧАНИЕ 5. Если система $a_1, \dots, a_k$ — ортогональная, то (8.6) приобретает особо упрощенный вид

$$\alpha_i (a_i, a_i) = (a, a_i) \quad (i = 1, \dots, k),$$

откуда

$$\alpha_i = \frac{(a, a_i)}{(a_i, a_i)} \quad (i = 1, \dots, k),$$

поэтому

$$\begin{aligned}\text{pr}_W a &= \frac{(a, a_1)}{(a_1, a_1)} a_1 + \frac{(a, a_2)}{(a_2, a_2)} a_2 + \dots + \frac{(a, a_k)}{(a_k, a_k)} a_k, \\ \text{ort}_W a &= a - \frac{(a, a_1)}{(a_1, a_1)} a_1 - \frac{(a, a_2)}{(a_2, a_2)} a_2 - \dots - \frac{(a, a_k)}{(a_k, a_k)} a_k.\end{aligned} \quad (8.8)$$

Ср. с (8.3). По теореме Пифагора

$$|a|^2 = |\text{pr}_W a|^2 + |\text{ort}_W a|^2, \quad (8.9)$$

поэтому $|a|^2 \geq |\text{pr}_W a|^2$. Теперь, применяя обобщенную теорему Пифагора к правой части последнего неравенства, получаем *неравенство Бесселя*:

$$|a|^2 = \frac{|(a, a_1)|^2}{(a_1, a_1)} + \dots + \frac{|(a, a_k)|^2}{(a_k, a_k)}$$

Из (8.9) следует, что равенство здесь возможно тогда и только тогда, когда $\text{ort}_W a = 0$, т.е. $a \in W$ (*равенство Парсеваля*).

Теорема 8.7.1 (Теорема о длине перпендикуляра). Пусть $a_1, \dots, a_k$ — базис подпространства $W \subseteq V$, тогда для произвольного вектора a из унитарного пространства V справедливо равенство

$$|\text{ort}_W a| = \sqrt{\frac{\det G(a_1, \dots, a_k, a)}{\det G(a_1, \dots, a_k)}}. \quad (8.10)$$

ДОКАЗАТЕЛЬСТВО. Пусть $\text{pr}_W a = \alpha_1 a_1 + \dots + \alpha_k a_k$. Вычтем из последнего столбца определителя Грама $\det G(a_1, \dots, a_k, a)$ линейную комбинацию остальных столбцов с коэффициентами $\bar{\alpha}_1, \dots, \bar{\alpha}_k$. Тогда

$$\begin{aligned} \det G(a_1, \dots, a_k, a) &= \begin{vmatrix} (a_1, a_1) & (a_1, a_2) & \dots & (a_1, a_k) & (a_1, a) \\ (a_2, a_1) & (a_2, a_2) & \dots & (a_2, a_k) & (a_2, a) \\ \dots & \dots & \dots & \dots & \dots \\ (a_k, a_1) & (a_k, a_2) & \dots & (a_k, a_k) & (a_k, a) \\ (a, a_1) & (a, a_2) & \dots & (a, a_k) & (a, a) \end{vmatrix} \\ &= \begin{vmatrix} (a_1, a_1) & (a_1, a_2) & \dots & (a_1, a_k) & (a_1, a - \alpha_1 a_1 - \dots - \alpha_k a_k) \\ (a_2, a_1) & (a_2, a_2) & \dots & (a_2, a_k) & (a_2, a - \alpha_1 a_1 - \dots - \alpha_k a_k) \\ \dots & \dots & \dots & \dots & \dots \\ (a_k, a_1) & (a_k, a_2) & \dots & (a_k, a_k) & (a_k, a - \alpha_1 a_1 - \dots - \alpha_k a_k) \\ (a, a_1) & (a, a_2) & \dots & (a, a_k) & (a, a - \alpha_1 a_1 - \dots - \alpha_k a_k) \end{vmatrix} \end{aligned}$$

Так как $a - \alpha_1 a_1 - \dots - \alpha_k a_k = \text{ort}_W a$ ($(a, \text{ort}_W a) = (\text{ort}_W a, \text{ort}_W a)$ и $(a_i, \text{ort}_W a) = 0$), то

$$\det G(a_1, \dots, a_k, a) = \begin{vmatrix} (a_1, a_1) & \dots & (a_1, a_k) & 0 \\ \dots & \dots & \dots & \dots \\ (a_k, a_1) & \dots & (a_k, a_k) & 0 \\ (a, a_1) & \dots & (a, a_k) & (\text{ort}_W a, \text{ort}_W a) \end{vmatrix}.$$

Теперь, раскрывая определитель по последнему столбцу, получаем

$$\det G(a_1, \dots, a_k, a) = (\text{ort}_W a, \text{ort}_W a) \det G(a_1, \dots, a_k),$$

откуда и следует доказываемая формула. ■

8.8. Процесс ортогонализации Грама-Шмидта

Опишем процедуру нахождения ортогонального базиса $b_1, \dots, b_k$ подпространства W . Пусть $a_1, \dots, a_k$ — произвольный базис подпространства W . Положим

$$\begin{aligned} W_i &= L(a_1, \dots, a_i) \quad (i = 1, \dots, k) \\ b_i &= \begin{cases} a_1, & \text{если } i = 1, \\ \text{ort}_{W_{i-1}} a_i, & \text{если } i = 2, \dots, k. \end{cases} \end{aligned} \quad (8.11)$$

Имеем $W_i = L(b_1, \dots, b_i)$ и поэтому $\text{ort}_{W_{i-1}}$ можно вычислять учитывая замечание 5 предыдущего раздела.

ЗАМЕЧАНИЕ 1. Из теоремы Пифагора следует, что $|b_i| \leq |a_i|$ ($i = 1, \dots, k$).

ЗАМЕЧАНИЕ 2. Описанная процедура годится и для случая, когда $W = L(a_1, \dots, a_k)$, но векторы не обязательно являются линейно независимыми. По замечанию 4 из предыдущего параграфа, $b_i = 0$ тогда и только тогда, когда $a_i \in W_i$.

8.9. Объем параллелепипеда

Параллелепипедом, построенным на системе векторов $a_1, \dots, a_k$ линейного пространства V , называется множество

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

Пример 50. В пространствах $\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3$ множества $\Pi(a_1), \Pi(a_1, a_2), \Pi(a_1, a_2, a_3)$ являются соответственно отрезком, параллелограммом и параллелепипедом, построенным на векторах a_1, a_2, a_3 .

Определим b_i, W_i согласно (8.11). *Объемом параллелепипеда*, построенного на системе векторов $a_1, \dots, a_k$ унитарного пространства V , назовем величину

$$V(a_1, \dots, a_k) = \begin{cases} |a_1|, & \text{если } i = 1, \\ V(a_1, \dots, a_k) \cdot |\text{ort}_{W_{k-1}} a_k|, & \text{если } i \geq 2. \end{cases} \quad (8.12)$$

Пример 51. В геометрических пространствах радиус векторов (8.12) превращается в знакомую формулу для вычисления площади параллелограмма (“основание умножить на высоту”) и параллелепипеда (“площадь основания умножить на высоту”).

Применяя несколько раз формулы (8.12, 8.10), получаем

$$\begin{aligned} V(a_1, \dots, a_k) &= |a_1| \cdot |\operatorname{pr}_{a_1} a_2| \cdot |\operatorname{pr}_{W_2} a_3| \cdot \dots \cdot |\operatorname{pr}_{W_{k-1}} a_k| \\ &= \sqrt{\det \Gamma(a_1)} \cdot \frac{\det \Gamma(a_1, a_2)}{\det \Gamma(a_1)} \cdot \dots \cdot \frac{\det \Gamma(a_1, \dots, a_k)}{\det \Gamma(a_1, \dots, a_{k-1})} \\ &= \sqrt{\det \Gamma(a_1, \dots, a_k)}. \end{aligned}$$

Теперь мы можем указать *геометрический смысл* определителя Грама: *определитель Грама системы векторов является квадратом объема параллелепипеда, построенного на этой системе.*

Так как $|b_i| = |\operatorname{ort}_{W_{i-1}} a_i| \leq |a_i|$, то

$$V(a_1, \dots, a_k) \leq |a_1| \cdot \dots \cdot |a_k|.$$

Последнее неравенство тоже имеет прозрачный геометрический смысл: *объем параллелепипеда не превосходит произведения длин его ребер.*

Подведем итог всему сказанному в данном пункте.

Теорема 8.9.1. Пусть $b_1, \dots, b_k$ — векторы, полученные из системы $a_1, \dots, a_k$ с помощью процесса ортогонализации, тогда

$$\begin{aligned} V(a_1, \dots, a_k) &= V(b_1, \dots, b_k) \\ &= \sqrt{\det \Gamma(a_1, \dots, a_k)} = \sqrt{\det \Gamma(b_1, \dots, b_k)} \\ &= |b_1| \cdot \dots \cdot |b_k| \leq |a_1| \cdot \dots \cdot |a_k|. \end{aligned}$$

Последнее неравенство превращается в равенство тогда и только тогда, когда система векторов $a_1, \dots, a_k$ — ортогональная или один из векторов a_i нулевой.

8.10. Неравенство Адамара

Во многих задачах необходимо уметь оценивать сверху величину определителя, исходя из величины его элементов. Неравенство Адамара, которое мы выведем, является одной из таких оценок.

В арифметическом пространстве F^n введем стандартное скалярное произведение. Пусть $a_1, \dots, a_k$ — столбцы матрицы $A \in F^{n \times k}$. Тогда

$$\Gamma(a_1, \dots, a_k) = A^T \overline{A}.$$

Из результатов предыдущего параграфа получаем, что

$$\sqrt{\det(A^T \overline{A})} \leq |a_1| \cdot \dots \cdot |a_k|.$$

Если матрица A квадратная $n \times n$, то $\det A = \det A^T$ и $\det \overline{A} = \overline{\det A}$, поэтому

$$|\det A| \leq |a_1| \cdot \dots \cdot |a_n|. \quad (8.13)$$

Если теперь $A = (a_{ij}) \in F^{n \times n}$ и $|a_{ij}| \leq \alpha$, то (8.13) примет вид

$$|\det A| \leq (\alpha \sqrt{n})^n. \quad (8.14)$$

Неравенства (8.13, 8.14) называются *неравенствами Адамара*.

Упражнение 49. Найти максимальную величину определителя третьего порядка с элементами ± 1 .

Неравенство (8.14) при $\alpha = 1$ достигается на матрицах специального вида (*матрицах Адамара*):

$$H_1 = (1), \quad H_i = \left(\begin{array}{c|c} H_{i-1} & -H_{i-1} \\ \hline H_{i-1} & H_{i-1} \end{array} \right) \text{ при } i \geq 2.$$

Легко доказать, что столбцы (и строки) матрицы H_i ортогональны и поэтому $\det H_i$ достигает своего максимального значения.

8.11. Метрические задачи в унитарных пространствах

В этом разделе векторы иногда мы будем называть точками. Вектор l называется наклонной из точки x к подпространству W , если $l = x - y$ для некоторого $y \in W$.

Теорема 8.11.1 (Теорема о длине наклонной). *Минимум*

$$\min_{y \in W}$$

достигается в единственной точке $y = \text{pr}_W x$. Иными словами,

$$\begin{aligned} \min_{y \in W} |x - y| &= |x - \text{pr}_W x| = |\text{ort}_W x|, \\ \forall z \in W, z \neq \text{pr}_W x \quad \min_{y \in W} |x - y| &< |x - z|. \end{aligned}$$

ДОКАЗАТЕЛЬСТВО. Для произвольного $y \in W$ по теореме Пифагора получаем

$$\begin{aligned} |x - y| &= |\text{ort}_W x + \underbrace{\text{pr}_W x - y}_{\in W}| \\ &= \sqrt{|\text{ort}_W x|^2 + |\text{pr}_W x - y|^2} \\ &\geq |\text{ort}_W x|, \end{aligned}$$

причем равенство возможно тогда и только тогда, когда $|\text{pr}_W x - y| = 0$, т.е. $y = \text{pr}_W x$. ■

Упражнение 50. Дайте геометрическую интерпретацию теореме о длине наклонной.

Расстоянием между двумя точками x, y называется величина

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

Расстоянием между двумя множествами S, T векторов унитарного пространства называется величина

$$\rho(S, T) = \inf_{\substack{x \in S \\ y \in T}} |x - y|.$$

Теорема 8.11.2. Пусть U, W — подпространства, а x, x_0, y_0 — произвольные векторы унитарного пространства V , тогда

- 1) $\rho(x, W) = |\text{ort}_W x|$,
- 2) $\rho(x, x_0 + W) = |\text{ort}_W(x - x_0)|$,
- 3) $\rho(x_0 + W, y_0 + U) = |\text{ort}_{W+U}(x_0 - y_0)|$.

ДОКАЗАТЕЛЬСТВО. Первая и вторая формулы являются частными случаями третьей, поэтому докажем третью:

$$\begin{aligned} \rho(x_0 + W, y_0 + U) &= \inf_{\substack{x \in x_0 + W \\ y \in y_0 + U}} |x - y| = \inf_{\substack{x \in W \\ y \in U}} |(x_0 + x) - (y_0 + y)| \\ &= \inf_{\substack{x \in W \\ y \in U}} |(x_0 - y_0) - \underbrace{(y - x)}_{z \in W+U}| = \inf_{z \in W+U} |(x_0 - y_0) - z| \\ &= \rho(x_0 - y_0, W + U). \end{aligned}$$

По теореме о длине наклонной $\rho(x_0 - y_0, W + U) = |\text{ort}_{W+U}(x_0 - y_0)|$. ■

8.12. Нормальное решение системы линейных уравнений

Пусть $A \in F^{m \times n}$, $b \in F^m$. Рассмотрим унитарное пространство F^n с введенным некоторым образом скалярным произведением. Решение $x' \in F^n$ системы линейных уравнений $Ax = b$ минимальной длины называется *нормальным решением* системы. Множество всех решений системы есть линейное многообразие $x_0 + W$, где x_0 — частное решение системы $Ax = b$, а W — множество всех решений системы $Ax = 0$. Пусть x_0 — частное решение системы $Ax = b$, M — множество всех решений однородной системы $Ax = 0$, тогда $x_0 + M$ — общее решение системы $Ax = b$.

Из определения нормального решения получаем

$$|x'| = \min_{x \in x_0 + W} |x| = \min_{x \in W} |x_0|.$$

По теореме о длине наклонной нормальное решение определяется единственным образом и равно

$$x' = \text{ort}_W x_0 = \text{pr}_{W^\perp} x_0.$$

Теперь предположим, что скалярное произведение в F^n стандартное. Тогда $W^\perp = L(\bar{a}'_1, \dots, \bar{a}'_m)$ (см. теорему о разложении пространства в прямую сумму подпространства и его ортогонального дополнения) и, кроме того, $\Gamma(\bar{a}'_1, \dots, \bar{a}'_m) = A\bar{A}^\top$, откуда $\Gamma^\top = A\bar{A}^\top$. Теперь система (8.6) для нахождения $\text{pr}_{W^\perp} x_0$ примет вид

$$\begin{aligned} A\bar{A}^\top y &= x_0 A, \text{ или } A\bar{A}^\top y = b \\ \text{где } y &= (\alpha_1, \dots, \alpha_m)^\top. \end{aligned} \quad (8.15)$$

По решению y системы (8.15) из (8.15) можно найти нормальное решение:

$$x' = \bar{A}^\top y.$$

8.13. Псевдорешение несовместной системы линейных уравнений

Пусть $A \in F^{m \times n}$, $b \in F^m$. Назовем *невязкой системы* $Ax = b$ на векторе $x \in F^n$ столбец $Ax - b \in F^m$. Предположим, что в F^m введено скалярное произведение. *Псевдорешением* системы $Ax = b$ называется вектор x' , на котором норма невязки $|Ax - b|$ минимальна:

$$|Ax' - b| = \inf_{x \in F^n} |Ax - b|.$$

Заметим, что если система совместна, то любое ее решение является псевдорешением и других псевдорешений нет.

Теорема 8.13.1. *Множество всех псевдорешений системы $Ax = b$ есть множество решений системы $Ax = \text{pr}_W b$, где $W = L(a_1, \dots, a_n)$ и $a_1, \dots, a_n$ — столбцы матрицы A .*

ДОКАЗАТЕЛЬСТВО. Имеем

$$\begin{aligned} & \inf_{x \in F^n} \overbrace{|Ax - b|}^{y \in F^m} \\ &= \inf_{y \in W} |y - b| \\ &= |y - \text{pr}_W b|. \end{aligned}$$

Последнее равенство следует из теоремы о длине наклонной. ■

Теперь предположим, что скалярное произведение в F^m стандартное. Тогда $\Gamma(\bar{a}_1, \dots, \bar{a}_n) = A^\top \bar{A}$, откуда $\Gamma^\top = \bar{A}^\top A$. Теперь система (8.6) для нахождения $\text{pr}_W b$ примет вид

$$\bar{A}^\top Ax = \bar{A}^\top b \quad \Rightarrow \quad \text{pr}_W b = Ax.$$

Последнее равенство показывает, что множество решений системы $\bar{A}^\top Ax = \bar{A}^\top b$ совпадает со множеством всех псевдорешений системы $Ax = b$.

Линейные преобразования унитарных и евклидовых пространств

9.1. Сопряженные преобразования

Рассмотрим линейное преобразование φ унитарного или евклидова пространства V . Отображение $V \rightarrow V$ называется *сопряженным* к преобразованию φ и обозначается φ^* , если

$$(\varphi x, y) = (x, \varphi^* y) \quad (9.1)$$

для произвольных x, y из V .

Теорема 9.1.1. *Для любого преобразования φ унитарного или евклидова пространства V преобразование φ^* существует, единственно и является линейным, причем*

$$[\varphi^*]_e = \overline{\Gamma_e^{-1} [\varphi]_e^\top \Gamma_e}. \quad (9.2)$$

ДОКАЗАТЕЛЬСТВО.

СУЩЕСТВОВАНИЕ. Покажем, что линейное преобразование ψ с матрицей

$$[\psi]_e = \overline{\Gamma_e^{-1} [\varphi]_e^\top \Gamma_e}$$

является сопряженным к преобразованию φ . Используя правила выражения скалярного произведения через координаты векторов и матрицу Грама и правила выражения координатного столбца образа вектора через координаты прообраза и матрицу преобразования, для произвольных x, y из V получаем

$$\begin{aligned} (x, \psi y) &= [x]^\top \Gamma \overline{[\psi] [y]} \\ &= [x]^\top \Gamma \Gamma_e^{-1} \overline{[\varphi]_e^\top \Gamma_e [y]} \\ &= ([\varphi] [x])^\top \Gamma [y] \\ &= (\varphi x, y). \end{aligned}$$

ЕДИНСТВЕННОСТЬ И ПРОЧ. Выражая скалярное произведение через координаты векторов, равенство (9.1) перепишем в виде

$$\begin{aligned} &([\varphi] [x])^\top \Gamma [y] = [x]^\top \Gamma \overline{[\varphi^*] [y]}, \\ \text{откуда} \quad &[x]^\top [\varphi]^\top \Gamma [y] = [x]^\top \Gamma \overline{[\varphi^*] [y]}, \\ \text{так как } x \text{ — любой, то} \quad &[\varphi]^\top \Gamma [y] = \Gamma \overline{[\varphi^*] [y]}, \\ \text{следовательно,} \quad &[\varphi^*] [y] = \underbrace{\Gamma^{-1} [\varphi]^\top \Gamma [y]}_{A = [\varphi^*]}. \end{aligned}$$

Таким образом, преобразование φ^* единственно, является линейным и имеет матрицу (9.2). ■

ЗАМЕЧАНИЕ. Если базис ортонормированный, то матрица Грама единичная и формула (9.2) приобретает особенно простой вид:

$$[\varphi^*]_e = \overline{[\varphi]_e}^\top.$$

Матрица $\overline{A}^\top$ называется *сопряженной* к матрице $A \in F^{n \times n}$ и обозначается A^* . Таким образом, в ортонормированном базисе матрица сопряженного преобразования сопряжена с матрицей исходного преобразования.

Теорема 9.1.2 (Свойства операции сопряжения). Для произвольных преобразований φ , ψ и любого скаляра $\alpha \in F$ справедливы следующие свойства:

- 1) $(\varphi^*)^* = \varphi$;
- 2) $(\varphi + \psi)^* = \varphi^* + \psi^*$;
- 3) $(\alpha\varphi)^* = \overline{\alpha}\varphi^*$;
- 4) $(\varphi\psi)^* = \psi^*\varphi^*$.

ДОКАЗАТЕЛЬСТВО. Свойства можно вывести из соответствующих свойств матриц. Проиллюстрируем на примере свойства 1) другой способ доказательства. Для произвольных x, y из V имеем:

$$\begin{aligned} (\varphi x, y) &= (x, \varphi^* y) = \overline{(\varphi^* y, x)} \\ &= \overline{(y, \varphi^{**} x)} = (\varphi^{**} x, y). \end{aligned}$$

Откуда $(\varphi x - \varphi^{**} x, y) = 0$. Последнее равенство справедливо для произвольного вектора y . В частности, для $y = \varphi x - \varphi^{**} x$ имеем

$$(\varphi x - \varphi^{**} x, \varphi x - \varphi^{**} x) = 0,$$

поэтому $\varphi x = \varphi^{**} x$, где x — любой вектор из V , поэтому $\varphi = \varphi^{**}$. ■

9.2. Теорема Шүра

Лемма 9.2.1. Ортогональное дополнение к собственному вектору линейного преобразования φ унитарного или евклидова пространства является инвариантным подпространством для сопряженного преобразования φ^* .

ДОКАЗАТЕЛЬСТВО. Пусть x — собственный вектор, λ — соответствующее ему собственное число: $\varphi x = \lambda x$. Покажем, что

$$\varphi^* x^\perp \subseteq x^\perp.$$

Действительно, для произвольного $y \in x^\perp$ имеем

$$(x, \varphi^* y) = (\varphi x, y) = \lambda(x, y) = 0.$$

■

Лемма 9.2.2. Для любого линейного преобразования φ комплексного n -мерного линейного пространства V существует инвариантное подпространство размерности $n - 1$.

ДОКАЗАТЕЛЬСТВО.

СПОСОБ 1. См. теорему Жордана.

СПОСОБ 2. В пространстве V введем скалярное произведение. Для преобразования φ^* найдем собственный вектор x . Подпространство $x^\perp$, разумеется, имеет размерность $n - 1$ и по предыдущей лемме инвариантно относительно преобразования $\varphi^{**} = \varphi$. ■

ЗАМЕЧАНИЕ. Второй способ доказательства дает, конечно, более простой алгоритм нахождения $(n - 1)$ -мерного инвариантного подпространства.

Лемма 9.2.3. Для любого линейного преобразования φ комплексного n -мерного линейного пространства V существует система инвариантных подпространств V_i ($i = 1, 2, \dots, n$), такая, что $\dim V_i = i$ и

$$V_1 \subset V_2 \subset \dots \subset V_n = V.$$

ДОКАЗАТЕЛЬСТВО.

СПОСОБ 1. См. теорему Жордана.

СПОСОБ 2. В V найдем $(n-1)$ -мерное инвариантное (относительно φ) подпространство. Рассмотрим сужение $\varphi|_{V_{n-1}}$ преобразования φ на подпространство V_{n-1} . В нем по лемме 2 снова найдем инвариантное подпространство V_{n-2} размерности на 1 меньше и т.д. ■

Теорема 9.2.1. *Шур Для любого линейного преобразования φ унитарного пространства существует ортонормированный базис, в котором матрица преобразования верхнетреугольная.*

ДОКАЗАТЕЛЬСТВО.

СПОСОБ 1. Построим жорданов базис $e_1, \dots, e_n$ преобразования φ . С помощью процесса ортогонализации ортонормируем этот базис. Полученный базис обозначим через $f_1, \dots, f_n$. Заметим, что матрица перехода $Q_{e \rightarrow f}$ верхнетреугольная. Имеем

$$[\varphi]_f = Q_{e \rightarrow f}^{-1} [\varphi]_e Q_{e \rightarrow f},$$

причем все три матрицы в правой части верхнетреугольные. Поэтому матрица $[\varphi]_f$ тоже верхнетреугольная и $f_1, \dots, f_n$ — искомый базис.

СПОСОБ 2. По предыдущей лемме существуют инвариантные подпространства V_i , такие, что $V_1 \subset V_2 \subset \dots \subset V_n$ и $\dim V_i = i$. Пусть f_1 — произвольный нормированный вектор из V_1 , f_2 — нормированный вектор из V_2 , ортогональный V_1 , $\dots$, f_n — нормированный вектор из V_n , ортогональный V_{n-1} . Легко видеть, что базис $f_1, \dots, f_n$ удовлетворяет условиям теоремы. ■

9.3. Нормальные преобразования

Преобразование φ евклидова или унитарного пространства называется *нормальным*, если

$$\varphi^* \varphi = \varphi \varphi^*.$$

Матрица $A \in F^{n \times n}$ называется *нормальной*, если $A^* A = A A^*$.

Теорема 9.3.1. *Критерий нормального преобразования унитарного пространства] Преобразование φ унитарного пространства является нормальным тогда и только тогда, когда существует ортонормированный базис пространства, состоящий из собственных векторов.*

ДОКАЗАТЕЛЬСТВО.

ДОСТАТОЧНОСТЬ. Пусть $e_1, \dots, e_n$ — система собственных векторов преобразования φ , составляющая ортонормированный базис пространства, тогда

$$[\varphi]_e = \text{diag}(\lambda_1, \dots, \lambda_n), \quad [\varphi^*]_e = \text{diag}(\bar{\lambda}_1, \dots, \bar{\lambda}_n).$$

Легко видеть, что $[\varphi^*]_e [\varphi]_e = [\varphi]_e [\varphi^*]_e$, и, следовательно, $\varphi^* \varphi = \varphi \varphi^*$.

НЕОБХОДИМОСТЬ. Пусть φ — нормальное преобразование. По теореме Шура для него существует ортонормированный базис $e_1, \dots, e_n$, в котором $[\varphi]_e = (\beta_{ij})$ — верхнетреугольная матрица, т.е. $\beta_{ij} = 0$ при $i > j$. Итак,

$$[\varphi]_e = \begin{pmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1n} \\ 0 & \beta_{22} & \dots & \beta_{2n} \\ \dots & & & \\ 0 & 0 & \dots & \beta_{nn} \end{pmatrix} \quad \text{и} \quad [\varphi^*]_e = \begin{pmatrix} \bar{\beta}_{11} & 0 & \dots & 0 \\ \bar{\beta}_{12} & \bar{\beta}_{22} & \dots & 0 \\ \dots & & & \\ \bar{\beta}_{1n} & \bar{\beta}_{2n} & \dots & \bar{\beta}_{nn} \end{pmatrix}.$$

Из определения нормального преобразования получаем

$$[\varphi][\varphi^*] - [\varphi^*][\varphi] = 0.$$

Выразим через β_{ij} диагональные элементы матрицы, стоящей в правой части приведенного равенства:

$$\begin{aligned} |\beta_{11}|^2 + |\beta_{12}|^2 + \dots + |\beta_{1n}|^2 - |\beta_{11}|^2 &= 0, \\ |\beta_{22}|^2 + |\beta_{23}|^2 + \dots + |\beta_{2n}|^2 - |\beta_{12}|^2 - |\beta_{22}|^2 &= 0, \\ &\dots \\ |\beta_{nn}|^2 - |\beta_{1n}|^2 - |\beta_{2n}|^2 - \dots - |\beta_{nn}|^2 &= 0. \end{aligned}$$

Последовательно получаем, что все внедиагональные элементы равны нулю. ■

Следствие 9.3.1. Собственные подпространства нормального преобразования унитарного пространства попарно ортогональны.

Следствие 9.3.2. Для произвольной нормальной матрицы $A \in \mathbf{C}^{n \times n}$ существует ортогональная $Q \in \mathbf{C}^{n \times n}$, такая, что $Q^{-1}AQ$ диагональна (также говорят, что матрица A ортогонально подобна диагональной).

Теорема 9.3.2. Критерий нормального преобразования евклидова пространства Преобразование φ евклидова пространства является нормальным тогда и только тогда, когда существует ортонормированный базис пространства, в котором матрица преобразования φ блочно диагональная с блоками двух типов:

- 1) блоки первого порядка;
- 2) блоки второго порядка вида

$$\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}.$$

ДОКАЗАТЕЛЬСТВО.

ДОСТАТОЧНОСТЬ. Пусть в некотором ортонормированном базисе $e_1, \dots, e_n$ матрица преобразования φ имеет блочно-диагональный вид

$$[\varphi]_e = \text{diag} \left(\lambda_1, \dots, \lambda_k, \begin{pmatrix} \alpha_1 & -\beta_1 \\ \beta_1 & \alpha_1 \end{pmatrix}, \dots, \begin{pmatrix} \alpha_m & -\beta_m \\ \beta_m & \alpha_m \end{pmatrix} \right).$$

Тогда

$$[\varphi]_e = \text{diag} \left(\bar{\lambda}_1, \dots, \bar{\lambda}_k, \begin{pmatrix} \bar{\alpha}_1 & \bar{\beta}_1 \\ -\bar{\beta}_1 & \bar{\alpha}_1 \end{pmatrix}, \dots, \begin{pmatrix} \bar{\alpha}_m & \bar{\beta}_m \\ -\bar{\beta}_m & \bar{\alpha}_m \end{pmatrix} \right).$$

Поэтому

$$[\varphi\varphi^*]_e = \left(|\lambda_1|^2, \dots, |\lambda_k|^2, \begin{pmatrix} |\alpha_1|^2 + |\beta_1|^2 & 0 \\ 0 & |\alpha_1|^2 + |\beta_1|^2 \end{pmatrix}, \dots, \begin{pmatrix} |\alpha_m|^2 + |\beta_m|^2 & 0 \\ 0 & |\alpha_m|^2 + |\beta_m|^2 \end{pmatrix} \right) = [\varphi^*\varphi]_e,$$

Следовательно, $\varphi\varphi^* = \varphi^*\varphi$.

НЕОБХОДИМОСТЬ. Пусть φ — нормальное преобразование евклидова пространства, $e_1, \dots, e_n$ — произвольный ортонормированный базис этого пространства. Рассмотрим преобразование $\Phi: \mathbf{C}^n \rightarrow \mathbf{C}^n$, заданное формулой $\Phi u = Au$, где $u \in \mathbf{C}^n$, $A = [\varphi]_e$. Пусть в пространстве $\mathbf{C}^n$ определено стандартное скалярное произведение. Так как матрица A — нормальная, то преобразование Φ — нормальное. Так как матрица A — вещественная, то, не нарушая общности можно считать, что система собственных значений имеет вид

$$\lambda_1, \dots, \lambda_t, \lambda_{t+1}, \dots, \lambda_k, \\ \bar{\lambda}_{t+1}, \dots, \bar{\lambda}_k,$$

где $\lambda_i \in \mathbf{R}$ ($i = 1, \dots, t$), $\lambda_i \neq \lambda_j$ ($i \neq j$).

Для каждого вещественного собственного значения λ_i ($i = 1, \dots, t$) построим ортонормированный базис $u_{i1}, \dots, u_{is_i}$ собственного подпространства, составленный из векторов с вещественными компонентами. Так как матрица A и собственное значение λ_i — вещественные, то это возможно.

Для пары комплексно сопряженных собственных значений $\alpha_i \pm i\beta_i \notin \mathbf{R}$ ($i = t+1, \dots, k$) построим ортогональный базис собственного подпространства

$$u_{i1} = x_{i1} + iy_{i1}, \dots, u_{is_i} = x_{is_i} + iy_{is_i} \quad (9.3)$$

и

$$v_{i1} = x_{i1} - iy_{i1}, \dots, v_{is_i} = x_{is_i} - iy_{is_i} \quad (9.4)$$

соответственно, где $x_{ij}, y_{ij} \in \mathbf{R}^n$ ($i = t+1, \dots, k$; $j = 1, \dots, s_i$).

Таким образом, имеем:

Собственные числа	Собственные векторы
$\lambda_1 :$	$u_{11}, \dots, u_{1s_1}$
$\vdots$	$\vdots$
$\lambda_t :$	$u_{t1}, \dots, u_{ts_t}$
$\alpha_{t+1} \pm i\beta_{t+1} :$	$x_{t+1\,1} \pm iy_{t+1\,1}, \dots, x_{t+1\,s_{t+1}} \pm iy_{t+1\,s_{t+1}}$
$\vdots$	$\vdots$
$\alpha_k \pm i\beta_k :$	$x_{k1} \pm iy_{k1}, \dots, x_{ks_k} \pm iy_{ks_k}$

1. Из формул (9.3,9.4) следует, что

$$x_{ij} = \frac{u_{ij} + v_{ij}}{2}, \quad y_{ij} = \frac{u_{ij} - v_{ij}}{2i},$$

поэтому

$$L(u_{ij}, v_{ij}) = L(x_{ij}, y_{ij}). \quad (9.5)$$

2а. Докажем, что $(x_{ij}, y_{ij}) = 0$. Имеем

$$0 = (u_{ij}, v_{ij}) = (x_{ij} + iy_{ij}, x_{ij} - iy_{ij}) = (x_{ij}, x_{ij}) + i(x_{ij}, y_{ij}) + i(y_{ij}, x_{ij}) - (y_{ij}, y_{ij}),$$

откуда, приравнявая нулю действительные и мнимые части, получаем $(x_{ij}, x_{ij}) = (y_{ij}, y_{ij})$ и $(x_{ij}, y_{ij}) = 0$.

2б. При $i \neq p$ или(и) $j \neq q$ имеем $L(u_{ij}, v_{ij}) \perp L(u_{pq}, v_{pq})$, откуда, в силу (9.5), получаем

$$(x_{ij}, x_{pq}) = (x_{ij}, y_{pq}) = (y_{ij}, y_{pq}) = 0.$$

3. Имеем

$$\Phi x_{ij} + i\Phi y_{ij} = \Phi u_{ij} = (\alpha_{ij} + i\beta_{ij})(x_{ij} + iy_{ij}) = (\alpha_{ij}x_{ij} - \beta_{ij}y_{ij}) + i(\alpha_{ij}y_{ij} + \beta_{ij}x_{ij}),$$

откуда, приравнявая действительные и мнимые части, получаем

$$\Phi x_{ij} = \alpha_{ij}x_{ij} - \beta_{ij}y_{ij}, \quad \Phi y_{ij} = \alpha_{ij}y_{ij} + \beta_{ij}x_{ij}.$$

4. Теперь легко видеть, что система f_{ij} ($i = 1, \dots, k; j = 1, \dots, s_i$), f'_{ij} ($i = t+1, \dots, k; j = 1, \dots, s_i$), где

$$\begin{aligned} [f_{ij}]_e &= u_{ij} & (i = 1, \dots, t; j = 1, \dots, s_i), \\ [f_{ij}]_e &= x_{ij}/|x_{ij}| & (i = t+1, \dots, k; j = 1, \dots, s_i), \\ [f_{ij}]_e &= y_{ij}/|y_{ij}| & (i = t+1, \dots, k; j = 1, \dots, s_i), \end{aligned}$$

составляет искомый базис. ■

9.4. Унитарные и ортогональные преобразования

Преобразование φ унитарного (соответственно евклидова) пространства называется *унитарным* (соответственно *ортогональным*), если

$$\varphi\varphi^* = \varphi^*\varphi = E. \quad (9.6)$$

Легко видеть, что унитарное (ортогональное) преобразование является нормальным. Условие (9.6), очевидно, эквивалентно утверждению, что

$$(\varphi x, \varphi y) = (x, y) \quad (9.7)$$

для любых векторов x, y и поэтому *унитарные (ортогональные) преобразования — это в точности те преобразования, которые сохраняют¹ скалярное произведение*. Из (9.6) также следует, что для унитарного (ортогонального) преобразования φ обратное φ^{-1} существует и $\varphi^{-1} = \varphi^*$.

Матрица $A \in \mathbf{C}^{n \times n}$ называется *унитарной*, если $A^\top A = E$.

¹ в смысле формулы (9.7)

Теорема 9.4.1. Преобразование φ унитарного (евклидова) пространства унитарно (ортогонально) тогда и только тогда, когда

$$(\varphi x, \varphi x) = (x, x) \quad (9.8)$$

для любого вектора x .

ДОКАЗАТЕЛЬСТВО. Необходимость очевидна. Докажем достаточность. Для унитарного (евклидова) пространства легко проверить равенства

$$(x, y) = \frac{1}{4} ((x + y, x + y) - (x - y, x - y) + i(x + iy, x + iy) - i(x - iy, x - iy))$$

и

$$(x, y) = \frac{1}{4} ((x + y, x + y) - (x - y, x - y))$$

соответственно, откуда $(\varphi x, \varphi y) = (x, y)$. ■

ЗАМЕЧАНИЕ. Для случая евклидова геометрического пространства $\mathbf{V}_3$ можно дать следующее доказательство достаточности условий теоремы. Формула (9.8) означает, что преобразование сохраняет модули (т.е. длины) векторов, а значит расстояние между точками. Сохранение углов следует теперь из теоремы косинусов. Скалярное произведение выражается через длины векторов (которые не изменяются) и косинус угла между ними (который также остается постоянным).

Теорема 9.4.2. Преобразование φ унитарного (евклидова) пространства унитарно (ортогонально) тогда и только тогда, когда образы векторов произвольного (какого-нибудь) ортонормированного базиса ортонормированы.

ДОКАЗАТЕЛЬСТВО.

НЕОБХОДИМОСТЬ. Пусть φ — унитарное (ортогональное) преобразование, $e_1, \dots, e_n$ — ортонормированный базис. Тогда $(\varphi e_i, \varphi e_j) = (e_i, e_j)$, т.е. система $\varphi e_1, \dots, \varphi e_n$ — ортонормирована.

ДОСТАТОЧНОСТЬ. Пусть $e_1, \dots, e_n, \varphi e_1, \dots, \varphi e_n$ — две ортонормированные системы, $x = \alpha_1 e_1 + \dots + \alpha_n e_n, y = \beta_1 e_1 + \dots + \beta_n e_n$ — произвольные векторы. Тогда

$$\begin{aligned} (\varphi x, \varphi y) &= (\alpha_1 \varphi e_1 + \dots + \alpha_n \varphi e_n, \beta_1 \varphi e_1 + \dots + \beta_n \varphi e_n) = \alpha_1 \bar{\beta}_1 + \dots + \alpha_n \bar{\beta}_n \\ &= (\alpha_1 e_1 + \dots + \alpha_n e_n, \beta_1 e_1 + \dots + \beta_n e_n) = (x, y). \end{aligned}$$
■

Теорема 9.4.3 (Свойства собственных чисел ортогонального (унитарного) преобразования).

Все собственные числа ортогонального (унитарного) преобразования по абсолютной величине равны 1.

ДОКАЗАТЕЛЬСТВО. Для собственного вектора x , такого, что $\varphi x = \lambda x$, имеем

$$(x, x) = (\varphi x, \varphi x) = (\lambda x, \lambda x) = \lambda \bar{\lambda} (x, x),$$

откуда $\lambda \bar{\lambda} = |\lambda|^2 = 1$. ■

Теорема 9.4.4 (Критерий унитарного преобразования). Если все собственные числа нормального преобразования φ унитарного пространства по абсолютной величине равны 1, то φ — унитарное.

ДОКАЗАТЕЛЬСТВО. Так как φ — нормальное, то существует ортонормированный базис $e_1, \dots, e_n$, для которого

$$[\varphi]_e = \text{diag}(\lambda_1, \dots, \lambda_n),$$

откуда

$$[\varphi^*]_e = \text{diag}(\bar{\lambda}_1, \dots, \bar{\lambda}_n).$$

По условию $|\lambda_i| = 1$, поэтому $[\varphi]_e [\varphi^*]_e = E$, т.е. $\varphi \varphi^* = \varepsilon$. ■

Теорема 9.4.5 (Критерий ортогонального преобразования). Преобразование φ евклидова пространства является ортогональным тогда и только тогда, когда существует ортонормированный базис пространства, в котором матрица преобразования φ блочно диагональная с блоками двух типов:

1) блоки первого порядка вида

$$(\pm 1);$$

2) блоки второго порядка вида

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

ДОКАЗАТЕЛЬСТВО. Теорема доказывается аналогично доказательству критерия нормального преобразования евклидова пространства. ■

9.5. Эрмитовы (самосопряженные) и симметрические преобразования

Преобразование φ унитарного (соответственно евклидова) пространства называется *эрмитовым* $\equiv$ *самосопряженным* (соответственно *симметрическим*), если $\varphi = \varphi^*$. Легко видеть, что эрмитово (симметрическое) преобразование является нормальным.

Теорема 9.5.1 (Свойства собственных чисел эрмитового преобразования). Все собственные значения эрмитового преобразования вещественны.

ДОКАЗАТЕЛЬСТВО. Для собственного вектора x , такого, что $\varphi x = \lambda x$, имеем

$$\lambda(x, x) = (\varphi x, x) = (x, \varphi x) = \overline{\lambda}(x, x),$$

откуда $\lambda = \overline{\lambda} \in \mathbf{R}$. ■

Теорема 9.5.2 (Критерий эрмитового преобразования). Если все собственные числа нормального преобразования φ унитарного пространства вещественны, то φ эрмитово.

ДОКАЗАТЕЛЬСТВО. Так как φ — нормальное, то существует ортонормированный базис $e_1, \dots, e_n$, для которого

$$[\varphi]_e = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Так как по условию $\lambda_i \in \mathbf{R}$, то $[\varphi]_e = [\varphi^*]_e$, т.е. $\varphi = \varphi^*$. ■

Теорема 9.5.3 (Критерий симметрического преобразования). Преобразование φ евклидова пространства тогда и только тогда является симметрическим, когда для него существует ортонормированный базис из собственных векторов.

ДОКАЗАТЕЛЬСТВО. Ну, вообще-то, очевидно. ■

Эрмитово (симметрическое) преобразование называется

- *положительным*, или *положительно определенным*, если $(x, \varphi x) > 0$ для любого ненулевого вектора x ,
- *неотрицательным*, или *положительно полуопределенным*, если $(x, \varphi x) \geq 0$ для любого вектора x ,
- *отрицательным*, или *отрицательно определенным*, если $(x, \varphi x) < 0$ для любого ненулевого вектора x ,
- *неположительным*, или *отрицательно полуопределенным*, если $(x, \varphi x) \leq 0$ для любого вектора x .

Теорема 9.5.4 (Критерий положительного (неотрицательного) преобразования). Эрмитово или симметрическое преобразование положительно определено (соответственно полуопределено) тогда и только тогда, когда все его собственные значения положительны (соответственно неотрицательны).

ДОКАЗАТЕЛЬСТВО. Пусть $e_1, \dots, e_n$ — ортонормированный базис из собственных векторов эрмитова или симметрического преобразования φ , причем $\varphi e_i = \lambda_i e_i$. Пусть $[x]_e = (x_1, \dots, x_n)^\top$, $[y]_e = (y_1, \dots, y_n)^\top$, тогда $(x, y) = x_1 \bar{y}_1 + \dots + x_n \bar{y}_n$ и $(x, \varphi x) = \lambda_1 |x_1|^2 + \dots + \lambda_n |x_n|^2$, откуда и следует утверждение теоремы. ■

9.6. Косоэрмитово преобразование

Преобразование φ унитарного пространства называется *косоэрмитовым*, если $\varphi^* = -\varphi$.

Утверждение 9.6.1. Преобразование φ унитарного пространства является косоэрмитовым тогда и только тогда, когда $\varphi = i\psi$ для некоторого эрмитова преобразования ψ .

ДОКАЗАТЕЛЬСТВО.

НЕОБХОДИМОСТЬ. Пусть $\psi = -i\varphi$, где ψ — косоэрмитово преобразование, тогда

$$(\psi)^* = (-i\varphi)^* = i\varphi^* = -i\varphi = \psi,$$

таким образом, ψ — эрмитово. Легко видеть, что $\varphi = i\psi$.

ДОСТАТОЧНОСТЬ. Достаточность проверяется непосредственно. ■

9.7. Разложения преобразований

Произвольное комплексное число можно представить либо в алгебраической форме $\alpha + i\beta$, либо в тригонометрической $r(\cos \varphi + i \sin \varphi)$. Следующие две теоремы о разложении преобразований являются аналогами этих представлений.

Теорема 9.7.1 (Разложение преобразования в прямую сумму эрмитова и косоэрмитова преобразования). Для любого преобразования φ унитарного пространства существуют единственные эрмитовы преобразования ψ_1, ψ_2 , такие, что $\varphi = \psi_1 + i\psi_2$.

ДОКАЗАТЕЛЬСТВО.

СУЩЕСТВОВАНИЕ. Пусть

$$\psi_1 = \frac{1}{2}(\varphi + \varphi^*), \quad \psi_2 = \frac{1}{2i}(\varphi - \varphi^*). \quad (9.9)$$

Легко проверить, что данные преобразования эрмитовы и $\varphi = \psi_1 + i\psi_2$.

ЕДИНСТВЕННОСТЬ. Пусть $\varphi = \psi_1 + i\psi_2$ и преобразования ψ_1, ψ_2 эрмитовы. Тогда $\varphi^* = \psi_1 - i\psi_2$, откуда следуют формулы (9.9). По этим формулам ψ_1, ψ_2 определяются однозначно. ■

Теорема 9.7.2 (Полярное разложение). Для любого преобразования φ унитарного (соответственно евклидова) пространства существуют неотрицательное самосопряженное (соответственно симметрическое) преобразование ψ и унитарное (соответственно ортогональное) преобразование η , такие, что $\varphi = \psi\eta$. Если φ — невырождено, то преобразования ψ, η определяются единственным образом.

ДОКАЗАТЕЛЬСТВО. Во-первых, исследуем преобразование $\varphi\varphi^*$. Докажем, что оно самосопряженное (симметрическое) и неотрицательное. Действительно,

$$(\varphi\varphi^*)^* = (\varphi^*)^* \varphi^* = \varphi\varphi^*.$$

Далее,

$$(x, \varphi\varphi^* x) = (\varphi x, \varphi x) \geq 0.$$

для любого вектора x .

Во-вторых, докажем что если φ — невырождено, то преобразования ψ , η определяются единственным образом. Пусть $\varphi = \psi\eta$, $\psi^* = \psi$, $\psi \geq 0$, $\eta^* = \eta^{-1}$. Тогда $\varphi\varphi^* = \psi\eta(\psi\eta)^* = \psi^2$, откуда положительное самосопряженное преобразование ψ определяется единственным образом. Теперь $\eta = \psi^{-1}\varphi$.

Наконец, покажем, как построить преобразования ψ и η^2 . Пусть $e_1, \dots, e_n$ — ортонормированный базис из собственных векторов преобразования $\varphi\varphi^* \geq 0$. Имеем

$$\varphi\varphi^*e_i = k_i^2e_i.$$

Не нарушая общности, будем считать, что

$$k_i > 0 \quad (i = 1, \dots, m), \quad k_i = 0 \quad (i = m+1, \dots, n).$$

Пусть³

$$\psi e_i = k_i e_i \quad (i = 1, \dots, n). \quad (\star)$$

Положим

$$w_i = \frac{\varphi^* e_i}{k_i} \quad (i = 1, \dots, m) \quad (\star\star)$$

и докажем, что векторы $w_1, \dots, w_m$ образуют ортонормированный базис пространства $W = L(w_1, \dots, w_m)$. Действительно,

$$k_i k_j (w_i, w_j) = (\varphi^* e_i, \varphi^* e_j) = (e_i, \varphi\varphi^* e_j) = k_j^2 (e_i, e_j).$$

Пусть

$$\eta w_i = e_i \quad (i = 1, \dots, n). \quad (\star\star\star)$$

Легко видеть, что ψ — неотрицательное самосопряженное (симметрическое), η — унитарное (ортогональное). Проверим, что $\varphi = \psi\eta$. Имеем

$$\begin{aligned} \text{Для } i = 1, \dots, m: \quad & \psi\eta w_i = \psi e_i = k_i e_i \quad \text{и} \quad \varphi w_i = \varphi\varphi^* e_i / k_i = k_i e_i; \\ \text{для } i = m+1, \dots, n: \quad & \psi\eta w_i = \psi e_i = 0 \quad \text{и} \quad (\varphi w_i, \varphi w_i) = (w_i, \underbrace{\varphi^* \varphi w_i}_{\in \text{Im } \varphi^* = W}) = 0. \end{aligned}$$

■

ЗАМЕЧАНИЕ. Применим доказанную теорему для преобразования φ^* , тогда $\varphi^* = \psi\eta$, откуда $\varphi = \eta^* \psi$, причем η^* — унитарное (ортогональное), ψ — самосопряженное (симметрическое) неотрицательное преобразование.

²Из предыдущего абзаца следует способ нахождения $\psi = \sqrt{\varphi\varphi^*}$ и $\eta = \psi^{-1}\varphi$ по произвольному невырожденному преобразованию φ . Приведенный далее способ уже не требует невырожденности φ .

³Таким образом, как и для невырожденного преобразования имеем $\psi = \sqrt{\varphi\varphi^*}$. Далее символами $\star$ отмечены основные формулы, используемые в алгоритме нахождения полярного разложения.

Предметный указатель

- Значение многочлена от матрицы 48
Адамара матрицы 82
Адамара неравенства 82
Аннулирование многочленом преобразования 53
Ассоциативность 31
База системы векторов 35
Базис канонический 69
Базис линейного пространства 36
Базис нормальный 69
Базис ортогональный 76
Бесселя неравенство 79
Биекция 8
Билинейная функция эрмитова 68
Билинейная функция симметрическая 68
Билинейная функция, ранг 66
Цепочка векторов 60
Частное 12
Частное 18
Чисел комплексных равенство 11
Число комплексное, аргумент 13
Число комплексное 11
Число обратное 12
Число противоположное 12
Число собственное 48
Длина вектора 75
Единица мнимая 12
Единица 12
Эйлера функция 20
Эквивалентность чистем векторов 33
Эрмитова функция, канонический вид 69
Эрмитова функция, нормальный вид 69
Феррари способ 25
Форма квадратичная 65
Формулы Кардано, неприводимый случай 24
Фробениуса матрица 50
Функция билинейная 65
Функция квадратичная 65
Функция от матрицы 63
Функция полуторалинейная 65
Функция 7
Индекс инерции, отрицательный 71
Индекс инерции, положительный 71
Инъекция 7
Изоморфизм линейных пространств 37
Кардано формулы 23
Кольцо числовое 30
Комбинация линейная 33
Коммутативность 31
Комплексных чисел деление 12
Комплексное число чисто мнимое 13
Комплексное число сопряженное 13
Комплексное число, извлечение корня натуральной степени 17
Комплексное число, сопряженное 15
Комплексное число, тригонометрическая форма записи 13
Комплексного числа алгебраическая форма 12
Комплексного числа действительная часть 13
Комплексного числа модуль 13
Корень первообразный 19
Кратность собственного значения, алгебраическая 51
Кратность собственного значения, геометрическая 51
Лагранжа алгоритм 68
Лагранжа–Сильвестра интерполяционный многочлен 62
Линейная комбинация, тривиальная 34
Линейные пространства, изоморфные 38
Линейное преобразование, характеристический многочлен 49
Линейное преобразование, характеристическое уравнение 49
Линейное пространство, аксиомы 31
Линейное пространство, конечномерное 36
Матрица билинейной функции 66
Матрица эрмитова 68
Матрица нормальная 87
Матрица оператора 44
Матрица преобразования 48
Матрица симметрическая 68
Матрица сопровождающая 51
Матрица сопряженная 86
Матрица унитарная 90
Матрица, канонический вид 56
Матрицы элементарных преобразований 67
Матрицы конгруэнтные 66
Матрицы соединенные 66
Минор главный 50
Миноры угловые 70
Многочлен аннулирующий минимальный 53

- Многочлен аннулирующий 53
 Многочлен циклотомический 21
 Многочлен деления круга 21
 Многочлен характеристический 49
 Многочлен круговой 21
 Многочлен минимальный 53
 Множеств включение 7
 Множества векторов ортогональные 77
 Множество континуальное 9
 Множество счетное 9
 Множество 7
 Муавра формула 15
 Невязка системы 83
 Независимость линейная 34
 Норма вектора 75
 Нормальное решение 83
 Нуль 12
 Оболочка линейная 33
 Образ отображения 7
 Оператор линейный 43
 Оператора дефект 47
 Оператора множество значений 47
 Оператора нуль-пространство 47
 Оператора образ 47
 Оператора произведение на число 45
 Оператора ранг 47
 Оператора ядро 47
 Операторов сумма 45
 Определенность отрицательная 71
 Определенность отрицательная 71
 Определенность положительная 71
 Ортогональная составляющая проекции 78
 Ортогональное дополнение 78
 Остаток 18
 Освобождение от иррациональности в знаменателе 30
 Отображение биективное 8
 Отображение инъективное 7
 Отображение сопряженное 85
 Отображение сюръективное 7
 Отображение взаимно однозначное 8
 Отображение 7
 Параллелепипеда объем 81
 Параллелепипед 80
 Парсеваля равенство 80
 Пересечение подпространств 39
 Перпендикулярность 75
 Перпендикуляр 78
 Плоскость комплексная 13
 Подмножество 7
 Подпространство инвариантное 48
 Подпространство корневое 56
 Подпространство 32
 Показатель корня 19
 Поле числовое 30
 Полуопределенность положительная 71
 Последовательность 7
 Преобразование эрмитово, неотрицательное 91
 Преобразование эрмитово, неположительное 91
 Преобразование эрмитово, отрицательно определенное 91
 Преобразование эрмитово, отрицательно полуопределенное 91
 Преобразование эрмитово, отрицательное 91
 Преобразование эрмитово, положительно определенное 91
 Преобразование эрмитово, положительно определенное 91
 Преобразование эрмитово, положительное 91
 Преобразование эрмитово 91
 Преобразование индуцированное 48
 Преобразование косоэрмитово 92
 Преобразование левое обратное 47
 Преобразование линейных пространств 43
 Преобразование невырожденное 47
 Преобразование нормальное 87
 Преобразование обратное 47
 Преобразование ортогональное 89
 Преобразование правое обратное 47
 Преобразование самосопряженное 91
 Преобразование симметрическое 91
 Преобразование тождественное 47
 Преобразование унитарное 89
 Преобразование, диагонализированное 51
 Проекция ортогональная 78
 Проекция вектора 42
 Произведение декартово 10
 Произведение скалярное 73
 Пространство евклидово 73
 Пространство линейное, правило сложения, правило умножения на число 31
 Пространство линейное, правило сложения 31
 Пространство линейное 31
 Пространство собственное 48
 Пространство унитарное 73
 Пространство векторное 31
 Пространство 31
 Псевдорешение 83
 Ранг системы векторов 35
 Расстояние между множествами 82
 Расстояние между точками 82
 Разложение вектора по подпространствам 40
 Размерность линейного пространства 36
 Разность комплексных чисел 12
 Разность векторов 32
 Разрешение в радикалах 25
 Резольвента 25
 Сигнатура 71
 Система ортогональная 75
 Система ортонормированная 75
 Система векторов, полная 36
 Скалярное произведение стандартное 74
 Скаляр 31
 След матрицы 50

- Собственный вектор, относится к собственному числу 48
- Собственный вектор, принадлежность собственному числу 48
- Спектр матрицы 63
- Степень декартова 10
- Столбец координатный 37
- Сумма подпространств ортогональная 77
- Сумма подпространств, прямая 40
- Сумма подпространств 39
- Сумма векторов 31
- Сужение преобразования 48
- Сюръекция 7
- Унитарные пространства ортогональные 77
- Унитарных пространств изоморфизм 77
- Уравнение двучленное 25
- Уравнение характеристическое 49
- Уравнение возвратное 25
- Вектор линейного пространства 31
- Вектор нулевой 31
- Вектор противоположный 31
- Вектор собственный 48
- Вектора координаты 37
- Векторы ортогональные 75
- Вычитание комплексных чисел 12
- Выразимость линейная систем векторов 33
- Выразимость линейная векторов 33
- Высота вектора 58
- Замкнутость 30
- Зависимость линейная 34
- Значение многочлена от преобразования 48
- Значение собственное 48
- Жорданов базис 56
- Жорданова форма преобразования 56
- Жорданова форма 56
- Жорданова клетка 56
- Жорданова матрица 56
- п 48
- Базис ортонормированный 76