АЛГЕБРА И ГЕОМЕТРИЯ 1

Щиголев В.В.

1. Элементарная теория множеств

Для каждого конечного множества X через $|X|$ будем обозначать количество элементов множества X. Пусть $X_1, \dots X_n$ — конечный набор множеств. **Декартовым произведением** $X_1 \times \cdots \times X_n$ называется множество всех упорядоченных наборов $(x_1, \dots, x_n)$ длинны n, где $x_i \in X_i$. В случае, когда все множества X_i конечны, верно равенство $|X_1 \times \cdots \times X_n| = |X_1| \cdots |X_n|$.

Пример 1.1. Пусть $X = \{1, 2, 5\}$ и $Y = \{2, 4\}$. Тогда $X \times Y = \{(1,2), (2,2), (5,2), (1,4), (2,4), (5,4)\}$. Количество элементов равно $|X \times Y| = 6 = 3 \cdot 2 = |X| \cdot |Y|$.

Используя понятие декартового произведения, легко определить понятие отображения (функции). Пусть X и Y — два множества. **Отображением** из X в Y называется подмножество f декартового произведения $X \times Y$ такое, что для каждого элемента $x \in X$ существует единственный элемент $y \in Y$, для которого $(x, y) \in f$. Этот элемент y обозначается через $f(x)$. Неформально говоря, отображение — это закон сопоставляющий каждому элементу из X единственный элемент из Y. Утверждение о том, что f — отображение из X в Y кратко записывается в виде $f : X \to Y$ или $X \overset{f}{\to} Y$.

Пример 1.2. Пусть $X = \{1, -2, 2\}$ и $Y = \{1, 3, 4\}$. Рассмотрим функцию $f : X \to Y$, заданную правилом $f(x) = x^2$. Тогда $f = \{(1,1), (-2,4), (2,4)\}$.

Отображения множеств могут быть следующих трёх видов:

(1) **Инъективное** отображение $f : X \to Y$ — это отображение, которое различным элементам из X сопоставляет различные элементы из Y:

$$x_1, x_2 \in X \text{ и } x_1 \neq x_2 \Longrightarrow f(x_1) \neq f(x_2)$$

или в эквивалентной форме

$$f(x_1) = f(x_2) \Longrightarrow x_1 = x_2.$$

(2) **Сюръективное** отображение $f : X \to Y$ — это отображение, которое в каждый элемент из Y переводит хотя бы один элемент из X:

$$\forall y \in Y \ \exists x \in X \ f(x) = y.$$

(3) **Биективное** отображение — это отображение, являющееся одновременно инъективным и сюръективным.

Биективное отображение мы ещё будем называть **биекцией**. Кроме того, под биекцией множества X мы будем понимать биективное отображение из X в X.

Введённые понятия тесно связаны с понятиями образа и прообраза. Пусть $f : X \to Y$ — отображение. Определим **образ** $f(A)$ множества $A \subset X$ как множество всех элементов вида $f(a)$, где $a \in A$:

$$f(A) = \{f(a) \mid a \in A\}.$$

По определению $f(A) \subset Y$ и $f(\{x\}) = \{f(x)\}$. Образ $f(X)$ всего множества X называется **образом отображения** f и обозначается через $\operatorname{im} f$.

Определим **прообраз** $f^{-1}(B)$ подмножества $B \subset Y$ как множество всех $x \in X$, для которых $f(x) \in B$:

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\}.$$

По определению $f^{-1}(B) \subset X$.

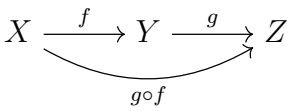
Доказательство следующего простого утверждение мы оставляем в качестве упражнения.

Лемма 1.3. *Пусть $f : X \to Y$ — отображение множеств.*

(1) *f инъективно тогда и только тогда, когда для любого элемента $y \in Y$ прообраз $f^{-1}(y)$ состоит не более чем из одного элемента.*
(2) *f сюръективно тогда и только тогда, когда для любого элемента $y \in Y$ прообраз $f^{-1}(y)$ не пуст.*
(3) *f сюръективно тогда и только тогда, когда $f(X) = Y$.*
(4) *f биективно тогда и только тогда, когда для любого элемента $y \in Y$ прообраз $f^{-1}(y)$ состоит ровно из одного элемента.*

Пример 1.4. Рассмотрим функцию $f : \mathbb{R} \to \mathbb{R}$, заданную формулой $f(x) = x^2$. Это отображение не инъективно и не сюръективно, например, $f(-1) = f(1) = 1$ и $f(x) = -1$ не имеет решения, иначе говоря, $f^{-1}(\{-1\}) = \varnothing$. Приведём ещё примеры: $f([-2, 3]) = [0, 9]$ и $f^{-1}([-3, 2]) = [-\sqrt{2}, \sqrt{2}]$.

Для двух отображений $X \xrightarrow{f} Y$ и $Y \xrightarrow{g} Z$ можно определить их **композицию** $X \xrightarrow{g \circ f} Z$ по правилу $(g \circ f)(x) = g(f(x))$. Последняя запись показывает, почему мы пишем $g \circ f$, а не $f \circ g$: аргумент пишется справа и порядок символов g, f и x не меняется. Это определение ещё можно продемонстрировать **коммутативной диаграммой**:



Коммутативность этой диаграммы заключается в том, что результат прохождения из одного узла в другой по всевозможным путям одинаков. В данной диаграмме только их X в Z ведёт больше одного, а точнее два, пути. Равенство результатов, как раз и есть формула $(g \circ f)(x) = g(f(x))$.

Рассмотрим пример биекции $f : X \to Y$. Мы можем определить **обратное отображение** $f^{-1} : Y \to X$, если всякий раз, когда $f(x) = y$ мы определяем $f^{-1}(y) = x$. С точки зрения композиции отображений обратное отображение характеризуется свойствами $f^{-1} \circ f = \mathrm{id}_X$ и $f \circ f^{-1} = \mathrm{id}_Y$. Здесь и далее id_X обозначает **тождественное** отображение на X, то есть отображение из X в X, оставляющее все элементы на месте: $\mathrm{id}_X(x) = x$.

Пусть S — некоторое множество. **Бинарной операцией** на S называется любое отображение $f : S \times S \to S$. Для бинарных операций мы будем всегда использовать **инфиксное обозначение**: результат действия операции f на пару (a, b) обозначается через afb. Аналогичным понятием является понятия **внешней бинарной операции** $g : K \times S \to S$ для произвольных множеств K и S. В этом случае мы ещё говорим, что множество K **действует** на S и используем инфиксное обозначение agb для результата действия g на (a, b).

Пример 1.5. Рассмотрим следующие бинарные операции на $\mathbb{R}^2$: $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ и $(x_1, y_1) - (x_2, y_2) = (x_1 - x_2, y_1 - y_2)$. Кроме того, $\mathbb{R}$ действует на $\mathbb{R}^2$ по правилу $\alpha \cdot (x, y) = (\alpha x, \alpha y)$.

2. Группы

Определение 2.1. *Группой называется пара $(G, \cdot)$, где G — множество и $\cdot$ — бинарная операция на G, удовлетворяющие следующим свойствам:*

(Г1) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ *для любых* $x, y, z \in G$.
(Г2) *Существует элемент* $e \in G$ *такой, что* $x \cdot e = e \cdot x = x$ *для любого* $x \in G$.
(Г3) *Для любого элемента* $x \in G$ *существует элемент* $y \in G$, *для которого* $x \cdot y = y \cdot x = e$.

Бинарная операция, удовлетворяющая свойству (Г1) называется **ассоциативной**. Элемент e из свойства (Г2) называется **единицей** группы G или её **нейтральным** элементом, а элемент y из свойства (Г3) называется **обратным к элементу** x. Бинарная операция $\cdot$ называется **групповой операцией**.

Лемма 2.2. *Элементы e и y определяемые условиями* (Г2) *и* (Г3) *соответственно определены однозначно.*

Доказательство. Пусть e' — ещё одна единица группы G. Мы знаем, что $e \cdot x = x$ и $y \cdot e' = y$ для любых $x, y \in G$. Подставляя $x = e'$ и $y = e$, получаем $e' = e \cdot e' = e$.

Для доказательства второго утверждения рассмотрим ещё один обратный элемент y' к элементу x. По свойствам (Г1)–(Г3) получаем

$$y \overset{(Г2)}{=\!=} y \cdot e \overset{(Г3)}{=\!=} y \cdot (x \cdot y') \overset{(Г1)}{=\!=} (y \cdot x) \cdot y' \overset{(Г3)}{=\!=} e \cdot y' \overset{(Г2)}{=\!=} y'.$$

Здесь и далее сверху равенства мы пишем обозначение свойства, на применении которого это равенство основано. □

Иногда, мы обозначаем единицу группы G через e_G, если мы хотим подчеркнуть тот факт, что e_G — единица именно группы G.

Скажем ещё несколько слов по поводу обозначений. В силу свойства (Г1) порядок постановки скобок в выражениях, содержащих несколько (более одной) операций $\cdot$ не имеет значения и мы будем эти скобки опускать. Наиболее общепринятыми являются **мультипликативная** и **аддитивная** системы обозначений. Их основные свойства приведены в следующей таблице:

	мультипликативная	**аддитивная**
групповая операция	не пишется, $\cdot, *, \bullet$ и тому подобное	$+$
единица	e или 1	0
обратный элемент	x^{-1}	$-x$

Мы обычно будем применять мультипликативную систему записи за исключением того, случая когда выполнено следующее свойство:

(Г4) $x \cdot y = y \cdot x$ для любых элементов $x, y \in G$.

В этом случае группа называется **абелевой** и мы часто, хотя и не всегда, применяем аддитивную систему записи. Бинарная операция, удовлетворяющая свойству (Г4) называется **коммутативной**.

Наконец, допуская вольность речи, мы называем само множество G группой, если из контекста очевидно какая групповая операция подразумевается на G.

Рассмотрим примеры групп. Пары $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$ и $(\mathbb{R}, +)$ являются группами причём абелевыми. Их единицей является ноль числовой оси, а элемент обратный к x это $-x$. Действительно, $x + 0 = 0 + x = x$ и $x + (-x) = (-x) + x = 0$ для любого элемента x из

$\mathbb{Z}$, $\mathbb{Q}$ или $\mathbb{R}$ соответственно. С другой стороны, $(\mathbb{Z}, -)$ не является группой, потому что операция $-$ не ассоциативна:

$$(a-b)-c = a-b-c, \qquad a-(b-c) = a-b+c$$

и эти выражения в общем случае не равны.

Пара $(\mathbb{R}, \cdot)$ тоже не является группой, так как 0 не имеет обратного: уравнение $x \cdot 0 = 1$ не имеет решения. Однако эту проблему легко устранить, удалив этот 0 из области рассмотрения: $(\mathbb{R} \setminus \{0\}, \cdot)$ — уже группа.

Простейшим примером неабелевой группы является группа $S(X)$ всех биекций множества X. Групповой операцией является операция композиции $gf = g \circ f$. Единицей этой группы является тождественное отображение id_X, а обратным к f элементом является обратное отображение f^{-1} (см. параграф 1). Мы будем применять обозначение $S_n = S(\{1, 2, \ldots, n\})$. Эта группа называется *симметрической группой*. Мы будем использовать её в параграфе 11 для определения понятия определеителя.

Гомоморфизмом групп называется отображение $f : G \to H$, где H и G — группы, такое, что $f(xy) = f(x)f(y)$ для любых элементов $x, y \in G$.

Лемма 2.3. *Пусть $f : G \to H$ — гомоморфизм групп. Тогда $f(e) = e$ и $f(x^{-1}) = f(x)^{-1}$ для любого $x \in G$.*

Доказательство. Докажем первое утверждение. Из равенства $e = ee$ получаем $f(e) = f(ee) = f(e)f(e)$. Умножаем, на $f(e)^{-1}$ слева, получаем $f(e)^{-1}f(e) = f(e)^{-1}(f(e)f(e)) = (f(e)^{-1}f(e))f(e)$. Так как $f(e)^{-1}f(e) = e$, то $e = ef(e) = f(e)$. □

Наконец, заметим, что отмена свойства (Г3) приводит к понятию **моноида**, а отмена свойств (Г2) и (Г3) приводит к понятию **полугруппы**. Таким образом, моноид — это полугруппа с нейтральным элементом. Он единственен в силу леммы 2.2. Например, множества неотрицательных и положительных целых чисел моноид и полугруппа относительно сложения соответственно.

Определение 2.4. *Подгруппой группы G называется подмножество $H \subset G$ замкнутое относительно групповой операции ($h_1, h_2 \in H \Rightarrow h_1h_2 \in H$) и само являющееся группой относительно ограничения на неё этой операции. Этот факт записывается в виде $H \leqslant G$.*

Таким образом, в подгруппе H имеется единица e_H и обратный элемент h_H^{-1} для каждого элемента $h \in H$ определённые с внутренней точки зрения. Но можно рассмотреть единицу e группы G и обратный элемент h^{-1} к элементу $h \in H$ с точки зрения большой группы G. Естественно спросить: верны ли равенства

$$e_H = e, \qquad h_H^{-1} = h^{-1}.$$

Докажем первое равенство. Запишем $e_H^2 = e_H$. Теперь рассмотрим обратный элемент e_H^{-1} к элементу e_H в большой группе G, то есть, такой элемент, что $e_H^{-1}e_H = e_He_H^{-1} = e$. Умножая на него предыдущее равенство слева, получаем

$$e_H = ee_H = e_H^{-1}e_He_H = e_H^{-1}e_H = e.$$

Теперь из равенства $h_H^{-1}h = hh_H^{-1} = e_H = e$ и единственности обратного элемента (лемма 2.2) получаем $h_H^{-1} = h^{-1}$.

3. Кольца

Определение 3.1. *Кольцом называется тройка $(R, +, \cdot)$, где R — множество и $+,\cdot$ — бинарные операции на R, удовлетворяющая следующим свойствам:*

(К1) $(x + y) + z = x + (y + z)$ *для любых $x, y, z \in R$.*

(К2) *Существует элемент* $0 \in R$ *такой, что* $x + 0 = 0 + x = x$ *для любого* $x \in R$.
(К3) *Для любого элемента* $x \in R$ *существует элемент* $y \in R$, *для которого* $x + y = y + x = 0$.
(К4) $x + y = y + x$ *для любых* $x, y \in R$.
(К5) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ *для любых* $x, y, z \in R$.
(К6) $x \cdot (y + z) = x \cdot y + x \cdot z$ *и* $(y + z) \cdot x = y \cdot x + z \cdot x$ *для любых* $x, y, z \in R$.

Можно сказать, что $(R, +)$ — абелева группа, $(R, \cdot)$ — полугруппа и обе эти операции связаны соотношениями **дистрибутивности** (К6).

Если дополнительно выполнено свойство

(К7) Существует элемент $1 \in R$ такой, что $x \cdot 1 = 1 \cdot x = x$ для любого $x \in R$,

то кольцо называется **кольцом с единицей**, а если выполнено свойство

(К8) $xy = yx$ для любых $x, y \in R$,

то кольцо называется **коммутативным**.

Мы всегда в дальнейшем будем предполагать, что любое рассматриваемое кольцо является кольцом с единицей и операцию умножения $\cdot$ мы будем опускать. С другой стороны, коммутативность умножения (К8) мы будем предполагать далеко не всегда.

Лемма 3.2. *Пусть* R *— кольцо. Тогда* $0x = x0 = 0$ *для любого* $x \in R$.

Доказательство. Из свойства (К2) получаем равенство $0 + 0 = 0$. Умножая его на x слева, получаем $x(0+0) = x$. Применяя свойство дистрибутивности (К6) к левой части, получаем $x0+x0 = x0$. Добавляя теперь $-x0$ к обеим частям, получаем $-x0+x0+x0 = -x0 + x0$. Отсюда $0 + x0 = 0$ и $x0 = 0$. Аналогично доказывается, что $0x = 0$. □

Пример 3.3. Множества $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ являются коммутативными кольцами относительно (обычных) сложения и умножения. Можно сконструировать также промежуточные кольца. Например, пусть n — целое число, не являющиеся полным квадратом. Положим

$$\mathbb{Z}\big[\sqrt{n}\big] = \big\{a + b\sqrt{n} \mid a, b \in \mathbb{Z}\big\}.$$

Это — коммутативное кольцо, что легко проверить умножением:

$$(a + b\sqrt{n})(c + d\sqrt{n}) = ac + nbd + (ad + bc)\sqrt{n}.$$

Мы получаем $\mathbb{Z} \subsetneq \mathbb{Z}\big[\sqrt{n}\big] \subsetneq \mathbb{R}$.

Пример 3.4. Приведём пример некоммутативного кольца. Для этого возьмём любое коммутативное кольцо R и рассмотрим множество матриц размера 2×2:

$$M_2(R) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \;\middle|\; a, b, c, d \in R \right\}.$$

Введём операции сложения и умножения следующим образом:

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} + \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & d_1 + d_2 \end{pmatrix}.$$

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1a_2 + b_1c_2 & a_1b_2 + b_1d_2 \\ c_1a_2 + d_1c_2 & c_1b_2 + d_1d_2 \end{pmatrix}.$$

Оставляем читателю проверить аксиомы (К1)–(К8).

В дальнейшем мы определим матрицы любого размера $m \times n$ для натуральных m и n. Их можно будет складывать и умножать при согласованности размеров.

Гомоморфизмом колец называется отображение $f : R \to S$, где R и S — кольца, такое, что $f(xy) = f(x)f(y)$ и $f(x + y) = f(x) + f(y)$ для любых элементов $x, y \in R$ и $f(1)$=1.

Лемма 3.5. *Пусть $f : R \to S$ — гомоморфизм колец. Тогда $f(0) = 0$ и $f(-x) = -f(x)$ для любого $x \in R$.*

Доказательство. Эти факты следуют из леммы 2.3 и того факта, что $(R, +)$ — группа. □

Заметим, что условие $f(1) = 1$ нельзя вывести из двух предыдущих. Действительно, отображение $f : \mathbb{Z} \to \mathbb{Z}$, заданное формулой $f(x) = 0$, удовлетворяет свойствам $f(xy) = f(x)f(y)$ и $f(x + y) = f(x) + f(y)$ однако гомоморфизмом не является. В связи с этим примером возникает вопрос: может ли в кольце выполняться равенство $0 = 1$? Да оно может выполняться и в этом случае $x = 1x = 0x = 0$. То есть, всё кольцо состоит из одного 0.

Определение 3.6. *Полем называется коммутативное ненулевое кольцо, в котором для каждого ненулевого элемента существует обратный.*

Пример 3.7. Полями являются множества $\mathbb{Q}$ и $\mathbb{R}$ относительно операций сложения и умножения. Поля могут быть конечными. Например зададим поле следующими таблицами

4. Многочлены и алгебраические уравнения

4.1. **Определения и основные свойства.** *Пусть R — кольцо. Рассмотрим множество всех бесконечных счётных последовательностей $(a_0, a_1, \ldots)$ элементов из R, содержащих только конечное количество ненулевых элементов. Мы можем складывать и умножать эти последовательности по следующим правилам:*

$$(a_0, a_1, \ldots) + (b_0, b_1, \ldots) = (a_0 + b_0, a_1 + b_1, \ldots).$$

$$(a_0, a_1, \ldots)(b_0, b_1, \ldots) = (a_0b_0, a_0b_1 + a_1b_0, a_0b_2 + a_1b_1 + a_2b_0, \ldots, \sum_{j=0}^{i} a_jb_{i-j}, \ldots).$$

Полученное множество последовательностей с операциями сложения и умножения и есть множество многочленов с одной переменной над R.

Лемма 4.1. *Множество многочленов с одной переменной над R является кольцом относительно введённых выше операций умножения и сложения.*

Доказательство. Сначала надо проверить корректность задания операций сложения и умножения. Очевидно, проверки требует только конечность ненулевых элементов в результирующих последовательностях. Пусть $(a_0, a_1, \ldots)$ и $(b_0, b_1, \ldots)$ — два многочлена над R. По определению существует индекс N такой, что $a_i = 0$ и $b_i = 0$ при $i > N$. В этом случае $a_i + b_i = 0$ при $i > N$. Предположим, что $\sum_{j=0}^{i} a_jb_{i-j} \neq 0$. В этом случае $a_j \neq 0$ и $b_{i-j} \neq 0$ для некоторого j. Отсюда получаем $j \leqslant N$ и $i - j \leqslant N$. Складывая, получаем $i \leqslant 2N$.

Докажем теперь, что множество многочленов с одной переменной является кольцом. В этом множестве последовательность из одних нулей $(0, 0, \ldots)$ является нулём. Так как сложение определяется почленно, то выполнение свойств (К1)–(К4) следует из выполнения аналогичных свойств для кольца R. При этом последовательность из одних нулей $(0, 0, \ldots)$ является нулём, противоположным элементом к последовательности $(a_1, a_2, \ldots)$ является последовательность $(-a_1, -a_2, \ldots)$. Легко заметить, что последовательность $(1, 0, 0, \ldots)$, в которой все элементы кроме первого нулевые, является единицей. Поэтому выполнено свойство (К7).

Остаётся проверить только свойства (K5) и (K6). Рассмотрим три последовательности $(a_0, a_1, \ldots)$, $(b_0, b_1, \ldots)$ и $(c_0, c_1, \ldots)$ и определим следующие произведения:

$$(a_0, a_1, \ldots)(b_0, b_1, \ldots) = (d_0, d_1, \ldots), \quad (b_0, b_1, \ldots)(c_0, c_1, \ldots) = (e_0, e_1, \ldots)$$

$$(d_0, d_1, \ldots)(c_0, c_1, \ldots) = (f_1, f_2, \ldots), \quad (a_0, a_1, \ldots)(e_0, e_1, \ldots) = (g_0, g_1, \ldots)$$

для соответсвующих $d_i, e_j, f_k, g_l \in R$. Мы получаем

$$f_l = \sum_{\substack{p,k\geqslant 0\\ p+k=l}} d_p c_k = \sum_{\substack{p,k\geqslant 0\\ p+k=l}} \left(\sum_{\substack{i,j\geqslant 0\\ i+j=p}} a_i b_j \right) c_j = \sum_{\substack{i,j,k\geqslant 0\\ i+j+k=l}} a_i b_j c_k,$$

$$g_l = \sum_{\substack{i,q\geqslant 0\\ i+q=l}} a_i e_q = \sum_{\substack{i,q\geqslant 0\\ i+q=l}} a_i \left(\sum_{\substack{j,k\geqslant 0\\ j+k=q}} b_j c_k \right) = \sum_{\substack{i,j,k\geqslant 0\\ i+j+k=l}} a_i b_j c_k.$$

откуда и следует требуемое равенство $f_l = g_l$. Этим доказно свойство (K5).

Наконец, мы положим

$$(b_0, b_1, \ldots) + (c_0, c_1, \ldots) = (h_0, h_1, \ldots), \quad (a_0, a_1, \ldots)(b_0, b_1, \ldots) = (s_0, s_1, \ldots)$$

$$(a_0, a_1, \ldots)(c_0, c_1, \ldots) = (t_0, t_1, \ldots), \quad (a_0, a_1, \ldots)(h_0, h_1, \ldots) = (r_0, r_1, \ldots)$$

Мы получаем

$$a_i = \sum_{j=0}^{i} a_i h_{i-j} = \sum_{j=0}^{i} a_i (b_{j-i} + c_{j-i}) = \sum_{j=0}^{i} (a_i b_{j-i} + a_i c_{j-i}) = \sum_{j=0}^{i} a_i b_{j-i} + \sum_{j=0}^{i} a_i c_{j-i} = s_i + t_i.$$

Этим первая часть свойства (K6) доказана. Вторая часть доказывается аналогично. □

Возможно, что данное нами определение кольца многочленов отличается от привычного (неформального) определения, знакомого читателю, в котором фигурировала переменная. Однако, мы можем записать многочлены из нашего определения более привычно, если положим $x = (0, 1, 0, 0, \ldots)$ и рассмотрим этот многочлен как переменную. Кроме того, мы отождествим любой элемент $a \in R$ с последовательностью $(a, 0, 0, \ldots)$. Более формально, это означает, что отображение $a \mapsto (a, 0, 0, \ldots)$ является инъективным гомоморфизмом. Так как $ax^n = x^n a = (0, \ldots, 0, a, 0, 0, \ldots)$, то любой многочлен однозначно представляется в виде $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, где $a_0, \ldots, a_n \in R$ и $a_n \neq 0$. Нулевой многочлен при этом представляется в виде пустой суммы, которая по определению равна нулю. Поэтому мы будем записывать нулевой многочлен просто символом 0. В дальнейшем мы будем работать с многочленами именно в этом виде. При этом само кольцо многочленов от одной переменной над R будем обозначать через $R[x]$ (или с любой другой переменной в квадратных скобках). Это кольцо содержит R как подкольцо.

Мы можем подставить вместо переменной x произвольный элемент из R (или из его расширения). Результат подстановки обозначается через $f(a)$. Таким образом

$$f = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \Longrightarrow f(\alpha) = a_n \alpha^n + a_{n-1} \alpha^{n-1} + \cdots + a_1 \alpha + a_0$$

Заметим, что для коммутативного кольца R выполняется соотношение

$$(fg)(\alpha) = f(\alpha) g(\alpha)$$

(в общем случае это не так: приведите примеры). Если $f(\alpha) = 0$, то мы говорим, что α — *корень* многочлена f.

Далее мы сконцентрируемся на случае, когда кольцо R — коммутативно и не имеет делителей нуля:

$$ab = ba \text{ для любых } a, b \in R, \qquad ab = 0 \Longrightarrow a = 0 \text{ или } b = 0.$$

В этом случае мы определим степень многочлена следующим образом:

$$\deg f = \begin{cases} n, & \text{если } f = a_n x^n + \cdots + a_1 x + a_0,\ n \geqslant 0,\ a_n \neq 0 \\ -\infty, & \text{если } f = 0. \end{cases}$$

В первом случае многочлен $a_n x^m$ называется старшим членом многочлена f. Степень многочлена $f \neq 0$ равна степени его старшего члена.

Мы видим, что степени многочленов могут быть любыми элементами множества $\mathbb{N} \cup \{0, -\infty\}$. Это множество имеет естественный линейный порядок: элементы отличные от $-\infty$ сравниваются как целые числа, и $-\infty < n$ для любого $n \in \mathbb{N} \cup \{0\}$. Кроме того, на множестве $\mathbb{N} \cup \{0, -\infty\}$ определена бинарная операция сложения, совпадающая с обычной операцией сложения на целых числах, и такая, что $-\infty + x = -\infty$ для любого $x \in \mathbb{N} \cup \{0, -\infty\}$. Например

$$1 < 2, \quad -\infty < 5, \quad 2 + 1 = 3, \quad -\infty + 5 = -\infty, \quad -\infty + (-\infty) = -\infty.$$

Исходя из порядка на множестве $\mathbb{N} \cup \{0, -\infty\}$ мы можем определить максимум и минимум конечного множества его элементов.

Лемма 4.2. *Пусть R — коммутативное кольцо без делителей нуля и $f, g \in R[x]$. Тогда выполнены следующие свойства:*

(1) $\deg(f + g) \leqslant \max\{\deg f, \deg g\}$.
(2) *Если* $\deg f > \deg g$, *то* $\deg(f + g) = \deg f$.
(3) $\deg(fg) = \deg f + \deg g$.
(4) $\deg(cf) = \deg f$ *для любого* $c \in R \setminus \{0\}$.

Доказательство. (1) Это утверждение очевидно, если $f = 0$ или $g = 0$. Действительно, в случае $f = 0$ получаем

$$\deg(f + g) = \deg g = \max\{\deg g, -\infty\} = \max\{\deg g, \deg f\}.$$

Здесь мы использовали тот факт, что $-\infty$ — наименьший элемент множества $\mathbb{N} \cup \{0, -\infty\}$.

Пусть теперь $f \neq 0$ и $g \neq 0$. Положим $n = \deg f$ и $m = \deg g$. Без ограничения общности можно считать, что $n \geqslant m$. Запишем

$$f = a_n x^n + \cdots + a_1 x + a_0, \quad g = b_m x^m + \cdots + b_1 x + b_0, \qquad \text{где } a_n \neq 0 \text{ и } b_m \neq 0. \tag{1}$$

Отсюда получаем

$$f + g = a_n x^n + \cdots + a_{m+1} x^{m+1} + (a_m + b_m) x^n + \cdots + (a_1 + b_1) x + (a_0 + b_0). \tag{2}$$

Из этой записи следует, что $\deg(f + g) \leqslant n = \max\{n, m\} = \max\{\deg f, \deg g\}$. Заметим, что в случае, когда $m = n$ и $a_n = -b_n$ мы имеем строгое неравенство $\deg(f + g) < n = \max\{\deg f, \deg g\}$.

(2). Предположим, что $\deg f > \deg g$. Опять случай $g = 0$ очевиден. Поэтому предположим, что $f \neq 0$ и $g \neq 0$ и запишем f и g в виде (1). Отсюда получаем равенство (2). Так как $n > m$, то старший член суммы $f + g$ равен $a_n x^n$ и $\deg(f + g) = n = \deg f$.

(3) В случае $f = 0$ получаем

$$\deg(fg) = \deg 0 = -\infty = -\infty + \deg g = \deg f + \deg g.$$

Аналогично рассматривается случай $g = 0$. Поэтому мы рассмотрим случай, когда $f \neq 0$ и $g \neq 0$. В этом случае имеет место представление (1). Умножая многочлены f и g друг на друга, получаем

$$fg = a_n b_m x^{n+m} + \cdots + \sum_{i=\max\{0,k-m\}}^{\min\{k,n\}} a_i b_{k-i} x^k + \cdots + a_0 b_0.$$

Так как $a_n \neq 0$ и $b_n \neq 0$ и кольцо R не содержит делителей нуля, то $a_n b_m \neq 0$. Отсюда и из вышеприведённой формулы следует, что $\deg(fg) = n + m = \deg f + \deg g$.

(4) Следует из (3), так как $\deg c = 0$. □

4.2. **Деление с остатком и теорема Безу.**

Теорема 4.3 (о делении с остатком)**.** *Пусть R — поле и $f, g \in R[x]$ и $f \neq 0$. Тогда существуют единственные многочлены $q, r \in R[x]$ такие, что*

$$f = qg + r, \quad \deg r < \deg g.$$

Доказательство. Существование. Проведём индукцию по степени многочлена f. Базой индукции является случай $\deg f < \deg g$. В этом случае искомыми многочленами являются $q = 0$ и $r = f$.

Пусть теперь $\deg f \geqslant \deg g$. Запишем многочлены f и g в виде (1). Рассмотрим многочлен

$$h = f - \frac{a_n}{b_m} x^{n-m} g.$$

Старшие члены обоих многочленов f и $a_n/b_m x^{n-m} g$ равны $a_n x^n$. Поэтому $\deg h < \deg f$.

Применим предположении индукции к паре h и g. Получаем

$$h = q'g + r, \quad \text{где } \deg r < \deg g.$$

Теперь мы можем выразить из этого равенства многочлен f:

$$f = h + \frac{a_n}{b_m} x^{n-m} g = q'g + r + \frac{a_n}{b_m} x^{n-m} g = \left(q' + \frac{a_n}{b_m} x^{n-m}\right) g + r.$$

Следовательно, достаточно положить

$$q = q' + \frac{a_n}{b_m} x^{n-m}.$$

Единственность. Пусть $f = qg + r$ и $f = q'g + r'$ — два таких представления. Вычитая второе равенство из первого, получаем $0 = (q - q')g + r - r'$. Перенося первое слагаемое в левую сторону, получаем

$$(q' - q)g = r - r'. \tag{3}$$

Посчитаем степени многочленов в разных сторонах этого равенства, используя лемму 4.2:

$$\deg(q' - q)g = \deg(q' - q) + \deg g, \quad \deg(r - r') \leqslant \max\{\deg r, \deg r'\} < \deg g.$$

Следовательно,

$$\deg(q' - q) + \deg g < \deg g.$$

Учитывая наши правила сравнения и сложения для множества $N \cup \{0, -\infty\}$, получаем, что вышеприведённое неравенство может выполняться только если $\deg(q' - q) = -\infty$. Это означает, что $q' = q$. Отсюда и из (3) следует, что $r = r'$. □

Следствие 4.4 (Теорема Безу)**.** *Пусть R — поле, $f \in R[x]$ и $\alpha \in R$. Тогда*

(1) *Остаток от деления f на $x - \alpha$ равен $f(\alpha)$.*
(2) *Многочлен f делится на $x - \alpha$ тогда и только тогда, когда $f(\alpha) = 0$.*

Доказательство. (1) Согласно теореме 4.3 запишем деление с остатком

$$f = q(x-\alpha) + r, \quad \deg r < 1 = \deg(x-\alpha).$$

Степень многочлена r может быть равна либо 0 либо $-\infty$. В любом случае $r \in R$. Подставляя α вместо x, получаем

$$f(\alpha) = q(\alpha)(\alpha-\alpha) + r(\alpha) = r(\alpha) = r.$$

(2) Предположим, что f делится на $x-\alpha$. Тогда $f = q(x-\alpha)$ для некоторого $q \in R[x]$. Подставляя α вместо x, получаем $f(\alpha) = q(\alpha)(\alpha-\alpha) = 0$. Наоборот, предположим, что $f(\alpha) = 0$. Запишем деление с остатком согласно части (1)

$$f = q(x-\alpha) + f(\alpha) = q(x-\alpha).$$

□

Обсудим вопрос о функциональном равенстве многочленов: *можно ли утверждать, что $f = g$, если известно, что $f(\alpha) = g(\alpha)$ для любого $\alpha \in R$?*

Этот вопрос, заданный в таком виде, имеет очевидный ответ нет. Действительно пусть $R = \mathbb{Z}_2$ (кольцо вычетов по модулю 2). Рассмотрим многочлены $f = x$ и $g = x^2$ очевидно $f(\alpha) = g(\alpha)$ для любого $\alpha \in \mathbb{Z}_2$, так как $0^2 = 0$ и $1^2 = 1$. Однако при этом $f \neq g$.

Ситуация изменится, если мы рассмотрим достаточно большие кольца R и потребуем отсутствие делителей нуля.

Лемма 4.5. *Пусть R — коммутативное кольцо без делителей нуля и $f \in R[x]$. Предположим, что существуют попарно различные элементы $\alpha_1, \ldots, \alpha_{n+1}$ кольца R такие, что $n \geqslant \deg f$ и $f(\alpha_i) = 0$ для любого $i = 1, \ldots, n+1$. Тогда $f = 0$.*

Доказательство. Так как кольцо R можно считать подкольцом своего поля частных, то мы без ограничения общности будем считать, что R — поле.

Проведём индукцию по n. Если $n = -1$, то $\deg f \leqslant -1$. Следовательно, $\deg f = -\infty$ и $f = 0$.

Предположим теперь, что $n \geqslant 0$ и утверждение верно для меньших n. Рассмотрим деление с остатком По теореме Безу получаем $f = q(x-\alpha_{n+1})$ для некоторого $q \in R[x]$. Мы получаем $0 = f(\alpha_i) = q(\alpha_i)(\alpha_i - \alpha_{n+1})$ для любого $i = 1, \ldots, n$. Сокращая на ненулевой элемент $\alpha_i - \alpha_{n+1}$, получаем $q(\alpha_i) = 0$ для любого $i = 1, \ldots, n$. Кроме того,

$$n \geqslant \deg f = \deg q + \deg(x-\alpha_{n+1}) = \deg q + 1$$

по лемме 4.2. Отсюда $\deg q \leqslant n-1$. Поэтому мы можем применить индуктивное предположение к многочлену q и получить, что $q = 0$. Отсюда $f = q(x-\alpha_{n+1}) = 0$. □

Следствие 4.6. *Пусть R — коммутативное кольцо без делителей нуля и $f, g \in R[x]$. Предположим, что существуют попарно различные элементы $\alpha_1, \ldots, \alpha_{n+1}$ кольца R такие, что $n \geqslant \max\{\deg f, \deg g\}$ и $f(\alpha_i) = g(\alpha_i)$ для любого $i = 1, \ldots, n+1$. Тогда $f = g$.*

Доказательство. Рассмотрим разность $h = f - g$. По лемме 4.2 получаем $\deg h \leqslant \max\{\deg f, \deg g\} \leqslant n$. Кроме того, $h(\alpha_i) = f(\alpha_i) - g(\alpha_i) = 0$ для любого $i = 1, \ldots, n+1$. Применяя лемму 4.5 к многочлену h, получаем $h = 0$. Отсюда $f = g$. □

4.3. **Рациональные корни.**

Теорема 4.7. *Пусть $f = a_n x^n + \cdots + a_1 x + a_0$ — многочлен из $\mathbb{Z}[x]$ степени $n \geqslant 1$ и p/q — его корень, где p и $q \neq 0$ — взаимно простые целые числа. Тогда*

(1) *p делит a_0.*

(2) *q делит a_n.*
(3) *$p - mq$ делит $f(m)$ для любого $m \in \mathbb{Z}$.*

Доказательство. (1) Запишем условие $f(p/q) = 0$ более подробно

$$a_n \left(\frac{p}{q}\right)^n + a_{n-1}\left(\frac{p}{q}\right)^{n-1} + \cdots + a_1\frac{p}{q} + a_0 = 0.$$

Умножая это равенство на q^n, получаем

$$a_n p^n + a_{n-1}p^{n-1}q + \cdots + a_1 pq^{n-1} + a_0 q^n = 0.$$

Перенося все слагаемые кроме последнего в правую часть, получаем

$$a_0 q^n = -a_n p^n - a_{n-1}p^{n-1}q - \cdots - a_1 pq^{n-1} = -p(a_n p^{n-1} + a_{n-1}p^{n-2}q + \cdots + a_1 q^{n-1}).$$

Следовательно, p делит $a_0 q^n$. Так как p и q взаимно простые, то p делит a_0.

(2) Эта часть доказывается аналогично части (1) с применением равенства

$$a_n p^n = -a_{n-1}p^{n-1}q - \cdots - a_1 pq^{n-1} - a_0 q^n = -q(a_{n-1}p^{n-1} + \cdots + a_1 pq^{n-2} + a_0 q^{n-1}).$$

(3) Вспомним про тождество

$$x^k - y^k = (x-y)(x^{k-1} + x^{k-2}y + \cdots + xy^{k-2} + y^{k-1}),$$

которое выполняется для любых рациональных x, y и натурального k. Нам понадобится только случай, когда $k \leqslant n$. Выполняя подстановку

$$x \mapsto m, \quad y \mapsto \frac{p}{q}$$

и умножая обе части на q^n, получаем

$$q^n\left(m^k - \left(\frac{p}{q}\right)^k\right) = q\left(m - \frac{p}{q}\right)q^{n-1}\left(m^{k-1} + m^{k-2}\left(\frac{p}{q}\right) + \cdots + m\left(\frac{p}{q}\right)^{k-2} + \left(\frac{p}{q}\right)^{k-1}\right) =$$
$$= (mq-p)(m^{k-1}q^{n-1} + m^{k-2}pq^{n-2} + \cdots + mp^{k-2}q^{n-k+1} + p^{k-1}q^{n-k}).$$

Обозначая

$$c_k = (m^{k-1}q^{n-1} + m^{k-2}pq^{n-2} + \cdots + mp^{k-2}q^{n-k+1} + p^{k-1}q^{n-k}),$$

получаем

$$q^n\left(m^k - \left(\frac{p}{q}\right)^k\right) = (mq-p)c_k.$$

Эта формула остаётся верной для $k = 0$, если мы положим $c_0 = 0$. Заметим, что все $c_k \in \mathbb{Z}$.

Теперь мы можем вернуться к нашему многочлену f:

$$q^n\left(f(m) - f\left(\frac{p}{q}\right)\right) = q^n\left(\sum_{k=0}^n a_k m^k - \sum_{k=0}^n a_k\left(\frac{p}{q}\right)^k\right) = \sum_{k=0}^n a_k q^n\left(m^k - \left(\frac{p}{q}\right)^k\right) =$$
$$= \sum_{k=0}^n a_k(mq-p)c_k = (mq-p)\sum_{k=0}^n a_k c_k.$$

Таким образом, если p/q — корень многочлена f, то $mp - q$ делит $q^n f(m)$. Согласно ... для того чтобы доказать, что $mp - q$ делит $f(m)$ остаётся заметить, что числа q^n и $mq - p$ взаимно простые. Действительно, предположим, что простое число r делит q^n и $mq - p$ одновременно. Тогда r делит q и, следовательно, делит p. Это невозможно, так как числа p и q взаимно простые. □

5. Комплексные числа

5.1. **Построение поля комплексных чисел.** Мы рассмотрим расширение поля действительных чисел, построив поле так называемых *комплексных чисел*. По определению комплексное число — это пара действительных чисел (a, b). Комплексные числа можно складывать и умножать по следующим правилам:

$$(a, b) + (c, d) = (a + c, b + d), \quad (a, b)(c, d) = (ac - bd, ad + bc). \tag{4}$$

Множество всех таких пар мы обозначим через $\mathbb{C}$, имея ввиду, что на нём действуют вышеопределённые бинарные операции.

Теорема 5.1. *Множество $\mathbb{C}$ является полем относительно операций сложения и умножения (4).*

Доказательство. Докажем сначала, что $\mathbb{C}$ — коммутативное кольцо (с единицей). Для этого существует два способа.

Способ 1. Проверим вручную свойства (К1)–(К6). Свойства (К1)–(К4) следуют из соответствующих свойств сложения в кольце действительных чисел, ввиду того, что операция сложения в $\mathbb{C}$ определена независимо по каждой координате. Нулём является пара $(0, 0)$, противоположным к паре (x, y) является пара $(-x, -y)$.

Свойство (К5) следует из следующих вычислений

$$\begin{aligned}\big((a, b)(c, d)\big)(f, g) &= (ac - bd, ad + bc)(f, g) = \\ &= \big((ac - bd)f - (ad + bc)g, (ac - bd)g + (ad + bc)f\big) = \\ &= (acf - bdf - adg - bcg, acg - bdg + adf + bcf).\end{aligned}$$

$$\begin{aligned}(a, b)\big((c, d)(f, g)\big) &= (a, b)\big(cf - dg, cg + df\big) = \\ &= \big(a(cf - dg) - b(cg + df), a(cg + df) + b(cf - dg)\big) = \\ &= (acf - adg - bcg - bdf, acg + adf + bcf - bdg).\end{aligned}$$

Как легко видеть эти выражения совпадают.

Проверим свойство (К6). Получаем

$$\begin{aligned}(a, b)\big((c, d) + (f, g)\big) &= \\ &= (a, b)(c + f, d + g) = \big(a(c + f) - b(d + g), a(d + g) + b(c + f)\big) = \\ &= (ac + af - bd - bg, ad + ag + bc + bf).\end{aligned}$$

$$\begin{aligned}(a, b)(c, d) + (a, b)(f, g) &= (ac - bd, ad + bc) + (af - bg, ag + bf) = \\ &= (ac - bd + af - bg, ad + bc + ag + bf).\end{aligned}$$

Эти выражения тоже совпадают.

Кроме того, в кольце $\mathbb{C}$ существует единица (свойство (К7)) — это элемент $(1, 0)$:

$$(1, 0)(a, b) = (1 \cdot a - 0 \cdot b, 1 \cdot b + 0 \cdot a) = (a, b),$$

$$(a, b)(1, 0) = (a \cdot 1 - b \cdot 0, a \cdot 0 + b \cdot 1) = (a, b),$$

и это кольцо коммутативно (свойство (К8)):

$$(a, b)(c, d) = (ac - bd, ad + bc), \qquad (c, d)(a, b) = (ca - db, cb + da).$$

Способ 2. Отождествим пару (a, b) со следующей матрицей:

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Используя правила умножения матриц (см. §9), легко проверить, что такое отождествление сохраняет сумму и произведение:

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} + \begin{pmatrix} c & -d \\ d & c \end{pmatrix} = \begin{pmatrix} a+c & -b-d \\ b+d & a+c \end{pmatrix},$$

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} c & -d \\ d & c \end{pmatrix} = \begin{pmatrix} ac-bd & -ad-bc \\ bc+ad & -bd+ac \end{pmatrix}.$$

При этом единице $(1, 0)$ в кольце $\mathbb{C}$ соответствует единичная матрица E, а противоположному элементу соответствует противоположная матрица. Теперь тот факт, что $\mathbb{C}$ — кольцо следует из того, что $M_2(\mathbb{R})$ — кольцо. Фактически мы отождествили $\mathbb{C}$ с подкольцом кольца $M_2(\mathbb{R})$.

Теперь остаётся доказать, что $\mathbb{C}$ — поле. Это следует следующего вычисления

$$(a, b)\left(\frac{a}{\sqrt{a^2+b^2}}, -\frac{b}{\sqrt{a^2+b^2}}\right) = (1, 0).$$

верного для любых действительных a и b одновременно не равных нулю. □

Теперь мы можем ввести обозначение $i = (0, 1)$. Легко проверить, что $i^2 = (-1, 0)$.

Лемма 5.2. *Отображение $\iota : \mathbb{R} \to \mathbb{C}$, заданное правилом $\iota(a) = (a, 0)$ является инъективным гомоморфизмом колец.*

Доказательство. Отображение ι инъективно по определению. Остаётся доказать, что ι сохраняет операции сложения и умножения. Это следует из следующих вычислений:

$$\iota(a) + \iota(b) = (a, 0) + (b, 0) = (a + b, 0) = \iota(a + b),$$

$$\iota(a)\iota(b) = (a, 0)(b, 0) = (ab - 00, a0 + 0b) = (ab, 0) = \iota(ab).$$

□

Теперь мы можем отождествить каждое действительно число a с его образом $\iota(a)$. При этом мы получаем $i^2 = -1$. Так как для любого комплексного числа (a, b) выполнено

$$(a, b) = (a, 0) + (b, 0)(0, 1) = (a, 0) + (b, 0)i,$$

то мы можем представить любое комплексное число z (с использованием вышеупомянутого отождествления) в виде

$$z = a + bi, \text{ для соответствующих } a, b \in \mathbb{R}.$$

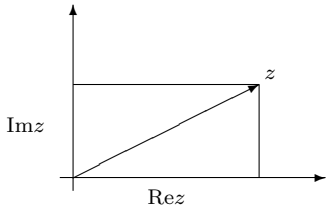
Такое представление называется *алгебраической формой записи* комплексного числа. Мы будем использовать обозначения $a = \mathrm{Re}z$ и $b = \mathrm{Im}z$. Эти числа называются соответственно *действительной* и *мнимой* частью комплексного числа z.

Замечание. Мы ввели комплексные числа как пары действительных чисел. Существуют и альтернативные способы. Например, мы могли бы ввести поле комплексных чисел как подкольцо кольца матриц $M_2(\mathbb{R})$ как в доказательстве теоремы 5.1. Наверное, самым естественным способом введения поля комплексных чисел было бы присоединение к полю действительных чисел $\mathbb{R}$ элемента i, удовлетворяющего свойству $i^2 = -1$. Эту интуитивную идею можно превратить в строго определение, если рассмотреть фактор кольцо

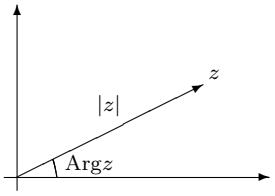
$$\mathbb{C} = \mathbb{R}[x]/(x^2 + 1)$$

и обозначить через i образ элемента x при естественном гомоморфизме $\mathbb{R}[x] \to \mathbb{C}$. Остаётся ещё обсудить вопрос почему $\mathbb{R}$ вкладывается в $\mathbb{C}$ как подкольцо и почему любой элемент из $\mathbb{C}$ можно единственным образом представать в виде $a+bi$ для $a, b \in \mathbb{R}$. Обсуждение этих вопросов увело бы нас слишком далеко от основного содержания этих лекций. Поэтому мы остановились на более наглядном определении поля комплексных чисел в виде множества пар действительных чисел с операциями (4).

5.2. **Модуль и аргумент.** Мы будем изображать комплексные числа на плоскости, считая первую компоненту абсциссой, а вторую ординатой.



Рассмотрим, что произойдёт, если задать туже точку z на плоскости в полярной системе координат.



Формулой мы можем записать

$$|z| = \sqrt{(\mathrm{Re}z)^2 + (\mathrm{Im}z)^2}.$$

Это число называется *модулем* комплексного числа z. Рассмотрим теперь случай $z \neq 0$. Как известно угол в полярных координатах определён не однозначно, а с точностью до $2\pi k$. Используя наши знания о факторкольцах, мы можем определить *аргумент* $\mathrm{Arg}z$ как элемент из факторгруппы $\mathbb{R}/2\pi\mathbb{Z}$, каждый представитель $\varphi \in \mathrm{Arg}z$ которого удовлетворяет соотношениям

$$\cos\varphi = \frac{\mathrm{Re}z}{|z|}, \quad \sin\varphi = \frac{\mathrm{Im}z}{|z|}.$$

В прочем, мы будем понимать сами функции cos и sin как функции из $\mathbb{R}/2\pi\mathbb{Z}$ в $\mathbb{R}$, получаемые из обычных тригонометрических функций следующим образом:

$$\cos(\varphi + 2\pi\mathbb{Z}) = \cos\varphi, \quad \sin(\varphi + 2\pi\mathbb{Z}) = \sin\varphi.$$

Тогда мы можем записать

$$\cos\mathrm{Arg}z = \frac{\mathrm{Re}z}{|z|}, \quad \sin\mathrm{Arg}z = \frac{\mathrm{Im}z}{|z|}.$$

Мы получили, что любое комплексное число z можно записать в виде

$$z = r(\cos\varphi + i\sin\varphi),$$

где $r = |z|$ и $\varphi \in \mathrm{Arg}z$, если $z \neq 0$. Эта запись не однозначна по двум причинам: во первых даже если $z \neq 0$, то φ определён с точностью до $2\pi k$; если $z = 0$, то в качествеφ можно взять любое число. Эта запись называется *тригонометрической формой записи* комплексного числа

5.3. **Сложение и умножение.** Обе формы записи комплексного числа: алгебраическая и тригонометрическая имеют свои преимущества. Вот первой форме легко складывать числа, а во второй их умножать:

$$(a+bi)+(c+di)=(a+c)+(b+d)i.$$

$$r(\cos\varphi+i\sin\varphi)\cdot r'(\cos\varphi'+i\sin\varphi')=rr'(\cos(\varphi+\varphi')+\sin(\varphi+\varphi')).$$

Последняя формула следует из формул косинуса и синуса суммы. Таким образом мы заключаем, что модуль произведения двух комплексных чисел равен произведению их модулей. При этом, если оба числа ненулевые, то аргумент их произведения равен сумме их аргументов. В качестве следствия из последней формулы получаем формулу *Муавра*

$$(r(\cos\varphi+i\sin\varphi))^n=r^n(\cos(n\varphi)+i\sin(n\varphi)).$$

Кроме того, полезно (и легко запомнить) формулу *Эйлера*

$$e^{i\varphi}=\cos\varphi+i\sin\varphi.$$

Из этой формулы легко "получить" формулы косинуса и синуса суммы, если считать, что экспонента суммы равна произведению экспонент (что выглядит вполне естественно и известно из школьного курса математики). Здесь кавычки стоят потому, что как раз формулы косинуса и синуса суммы позволяют установить правило для экспоненты суммы. Однако такое соображение бывает полезно тем, кто забыл тригонометрические формулы, но помнит формулу Эйлера (что, согласитесь, гораздо проще).

Теперь мы можем определить

$$e^{a+ib}=e^ae^{ib}=e^a(\cos b+i\sin b).$$

Таким образом может быть записано любое комплексное число, кроме нуля. Такая форма записи комплексного числа называется *экспоненциальной*.

5.4. **Сопряжение.** Сопряжённым к комплексному числу $z=a+ib$ называется число $\bar z=a-ib$. В тригонометрической форме получаем

$$\overline{r(\cos\varphi+i\sin\varphi)}=r(\cos\varphi-i\sin\varphi)=r(\cos(-\varphi)+i\sin(-\varphi)).$$

Отсюда следует, что

$$\overline{e^z}=e^{\bar z}.$$

Через сопряжённое число удобно выразить действительную и мнимые части, модуль и обратное число:

$$\mathrm{Re}z=\frac{z+\bar z}{2},\qquad \mathrm{Im}z=\frac{z-\bar z}{2i},\qquad |z|=\sqrt{z\bar z},\qquad z^{-1}=\frac{\bar z}{|z|^2}.$$

Кроме того $\bar{\bar z}=z$ и $|\bar z|=|z|$.

Теорема 5.3. *Любой автоморфизм поля* $\mathbb{R}$ *тождественный.*

Доказательство. Пусть $\varphi:\mathbb{R}\to\mathbb{R}$ — автоморфизм поля $\mathbb{R}$. Докажем последовательно следующие утверждения:

(1) $\varphi(x)>0$, если $x>0$.
(2) φ монотонное.
(3) $\varphi(n)=n$ для любого $n\in\mathbb{Z}$.
(4) $\varphi(q)=q$ для любого $q\in\mathbb{Q}$.
(5) $\varphi(x)=x$ для любого $x\in\mathbb{R}$.

(1) $\varphi(x) = \varphi(\sqrt{x}^2) = \varphi(\sqrt{x})^2 > 0$.
(2) Пусть $x < y$. Тогда $y - x > 0$. Согласно пункту (1) получаем $\varphi(y - x) > 0$. Отсюда $\varphi(y) - \varphi(x) > 0$ и $\varphi(y) > \varphi(x)$.
(3) Пусть n — натуральное число. Тогда

$$\varphi(n) = \varphi(1 + \cdots + 1) = \varphi(1) + \cdots + \varphi(1) = n\varphi(1) = n.$$

Отсюда $\varphi(-n) = -\varphi(n) = -n$. Остаётся заметить, что $\varphi(0) = 0$.
(4) Пусть $q = a/b$, где $a, b \in \mathbb{Z}$ и $b \neq 0$. Согласно пункту (3), получаем

$$a = \varphi(a) = \varphi(qb) = \varphi(q)\varphi(b) = \varphi(q)b.$$

Отсюда $\varphi(q) = a/b = q$.
(5) Пусть $s < x < r$ для $r, s \in \mathbb{Q}$. Согласно пунктам (2) и (3) получаем

$$s = \varphi(s) < \varphi(x) < \varphi(r) = r.$$

Так как эти неравенства верны для любых рациональных r и s удовлетворяющих неравенствам $s < x < r$, то $\varphi(x) = x$. □

Теорема 5.4. *Отображение из $\mathbb{C}$ в себя, заданное формулой $z \mapsto \bar{z}$, является единственным нетождественным автоморфизмом поля $\mathbb{C}$ переводящим $\mathbb{R}$ в себя.*

Доказательство. Докажем сначала, что сопряжение — автоморфизм. Для этого заметим, что

$$\overline{(a + ib) + (c + id)} = \overline{a + c + i(b + d)} = a + c - i(b + d) = a - ib + c - id = \overline{a + ib} + \overline{c + id}.$$

$$\overline{(a + ib)(c + id)} = \overline{ac - bd + i(ad + bc)} = ac - bd - i(ad + bc) =$$
$$= ac - (-b)(-d) + i(a(-d) + (-b)c) = (a - ib)(c - id) = \overline{(a + ib)(c + id)}.$$

Пусть теперь $\varphi : \mathbb{C} \to \mathbb{C}$ — автоморфизм, переводящий $\mathbb{R}$ в себя. Применяя теорему 5.3 к ограничению φ на $\mathbb{R}$, получаем, что $\varphi(x) = x$ для любого $x \in \mathbb{R}$. Теперь мы можем заметить, что

$$\varphi(i)^2 = \varphi(i^2) = \varphi(-1) = -1.$$

Так как уравнение $z^2 + 1 = 0$ имеет в $\mathbb{C}$ только два решения $z = i$ или $z = -i$, то $\varphi(i) = i$ или $\varphi(i) = -i$. В первом случае получаем

$$\varphi(a + ib) = \varphi(a) + \varphi(i)\varphi(b) = a + ib$$

для любых $a, b \in \mathbb{R}$, и φ — тождественный автоморфизм. Во втором случае получаем

$$\varphi(a + ib) = \varphi(a) + \varphi(i)\varphi(b) = a - ib$$

для любых $a, b \in \mathbb{R}$, и φ — автоморфизм сопряжения. □

5.5. **Основная теорема алгебры.** Поле $\mathbb{F}$ называется *алгебраически замкнутым*, если любой многочлен $f \in \mathbb{F}[x]$ степени больше нуля (непостоянный многочлен) имеет хотя бы один корень в $\mathbb{F}$.

Лемма 5.5. *Пусть $\mathbb{F}$ — алгебраически замкнутое поле. Тогда любой многочлен $f \in \mathbb{F}[x]$ представляется в виде $c(x - \alpha_1) \cdots (x - \alpha_n)$ для некоторых $c, \alpha_1, \ldots, \alpha_n \in \mathbb{F}$.*

Доказательство. Если $f = 0$, то такое представление существует при $c = 0$ и любых n и $\alpha_1, \ldots, \alpha_n$. Теперь предположим, что $\deg f > 0$ и докажем утверждение индукцией по этой степени. База индукции, это случай $\deg f = 1$. В это случае $f = a_1x - a_0 = a_1(x - a_0/a_1)$ и утверждение выполнено.

Пусть теперь $\deg f > 1$. По определению f имеет корень, скажем α_1. По теореме Безу $f = (x - \alpha_1)g$ для некоторого $g \in \mathbb{F}[x]$. Мы знаем, что

$$\deg f = \deg(x - \alpha_1) + \deg g = 1 + \deg g,$$

откуда $\deg g = \deg f - 1$ и $1 \leqslant \deg g < \deg f$. Поэтому к g применимо утверждение леммы:

$$g = c(x - \alpha_2) \cdots (x - \alpha_n)$$

для некоторых $c, \alpha_2, \ldots, \alpha_n \in \mathbb{F}$. Выражая отсюда f, получаем требуемое разложение. □

Теорема 5.6 (Основная теорема алгебры)**.** *Поле $\mathbb{C}$ алгебраически замкнуто.*

Существует много доказательств этой теоремы от чисто топологических до почти полностью алгебраических. Мы не будем приводить их здесь.

6. Векторные пространства

Определение 6.1. *Векторным пространство называется четвёрка $(V, F, +, \cdot)$, где V — множество, F — поле, $+$ — бинарная операция на V и $\cdot : F \times V \to V$ — внешняя бинарная операция, удовлетворяющая следующим свойствам:*

(B1) $(u + v) + w = u + (v + w)$ *для любых* $u, v, w \in V$.
(B2) *Существует элемент* $0 \in V$ *такой, что* $v + 0 = 0 + v = v$ *для любого* $v \in R$.
(B3) *Для любого элемента* $v \in V$ *существует элемент* $u \in V$*, для которого* $v + u = u + v = 0$.
(B4) $u + v = v + u$ *для любых* $v, u \in V$.
(B5) $(\alpha\beta) \cdot v = \alpha \cdot (\beta \cdot v)$ *для любых* $\alpha, \beta \in F$ *и* $v \in V$.
(B6) $1 \cdot v = v$ *для любого* $v \in V$.
(B7) $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$ *для любых* $\alpha, \beta \in F$ *и* $v \in V$.
(B8) $\alpha \cdot (v + u) = \alpha \cdot v + \alpha \cdot u$ *для любых* $\alpha \in F$ *и* $v, u \in V$.

Элементы множества V называются **векторами**, а элементы поля F **скалярами**. Допуская вольность речи, мы будем называть само V векторным пространством, если понятно о каком поле и операциях идёт речь. Как обычно, следуя нашей традиции, мы опускаем операцию умножения на скаляр $\cdot$ в наших обозначениях. Заметим, что иногда в литературе векторное пространство называется линейным. Легко заметить, что аксиомы (B1)–(B4) эквивалентны утверждению о том, что $(V, +)$ — абелева группа.

Пример 6.2. Основным примером векторного пространства является декартова степень поля

$$F^n = \underbrace{F \times \cdots \times F}_{n \text{ раз}}.$$

Элементами этого множества являются наборы $(x_1, \ldots, x_n)$, где $x_i \in F$.

Векторы (наборы) складываются и умножаются на скаляр следу.щим образом:

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n),$$
$$\alpha(x_1, \ldots, x_n) = (\alpha x_1, \ldots, \alpha x_n).$$

Читатель может легко проверить справедливость аксиом (B1)–(B8). Мы называем векторное пространство из этого примера **арифметическим** векторным пространством.

Определение 6.3. *Пусть $v_1, \ldots, v_n$ — векторы векторного пространства V над F. Линейной комбинацией этих векторов называется следующий вектор:*

$$\alpha_1 v_1 + \cdots + \alpha_n v_n,$$

где $\alpha_1, \ldots, \alpha_n \in F$. Множество всех линейных комбинаций векторов $v_1, \ldots, v_n$ обозначается через $\langle v_1, \ldots, v_n \rangle$.

Мы будем также говорить, что $\langle v_1, \ldots, v_n \rangle$ является линейной комбинацией набора векторов $(v_1, \ldots, v_n)$.

Нам часто будет важно рассматривать не только вектор $\alpha_1 v_1 + \cdots + \alpha_n v_n$, то и саму эту форму записи. Это удобно сделать следующим образом.

Определение 6.4. *Линейной комбинацией набора векторов $(v_1, \ldots, v_n)$ векторного пространства V над F с коэффициентами $(\alpha_1, \ldots \alpha_n) \in F^n$ называется вектор*

$$\alpha_1 v_1 + \cdots + \alpha_n v_n.$$

Если набор векторов $(v_1, \ldots, v_n)$ таков, что только его линейная комбинация с нулевыми коэффициентами равна нулю, то набор $(v_1, \ldots, v_n)$ называется линейно независимым. В противном случае набор $(v_1, \ldots, v_n)$ называется линейно зависимым.

Формулой линейную независимость набора $(v_1, \ldots, v_n)$ можно записать так

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0 \Rightarrow \alpha_1 = \cdots = \alpha_n = 0 \quad \forall \alpha_1, \ldots \alpha_n \in F.$$

С другой стороны, тот факт, что набор векторов $(v_1, \ldots, v_n)$ линейно зависим, означает, что существуют неравные одновременно нулю скаляры $\alpha_1, \ldots, \alpha_n \in F$ такие, что $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$. Исходя из последнего замечания, пустой набор векторов всегда линейно независим. Кроме того, линейная оболочка нуля векторов равна нулю: $\langle \rangle = 0$.

Лемма 6.5. *Если $n > 0$, то набор векторов $(v_1, \ldots, v_n)$ линейно зависим тогда и только тогда, когда существует индекс $i = 1, \ldots, n$ для которого $v_i \in \langle v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \rangle$.*

Доказательство. Пусть набор векторов $(v_1, \ldots, v_n)$ линейно зависим. Запишем линейную комбинацию $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$, в которой $\alpha_i \neq 0$ для некоторого $i = 1, \ldots, n$. Так как F — поле, то мы можем записать

$$v_i = -\frac{\alpha_1}{\alpha_i} v_1 - \cdots - \frac{\alpha_{i-1}}{\alpha_i} v_{i-1} - \frac{\alpha_{i+1}}{\alpha_i} v_{i+1} - \cdots - \frac{\alpha_n}{\alpha_i} v_n.$$

Последний элемент принадлежит линейной оболочке $\langle v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \rangle$.

Пусть наоборот $v_i \in \langle v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \rangle$. Запишем

$$v_i = \alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \cdots + \alpha_n v_n$$

для некоторых $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n \in F$. Отсюда получаем

$$-\alpha_1 v_1 - \cdots - \alpha_{i-1} v_{i-1} - v_1 - \alpha_{i+1} v_{i+1} - \cdots - \alpha_n v_n = 0.$$

Следовательно, линейная комбинация набора $(v_1, \ldots, v_n)$ с ненулевыми коэффициентами

$$(-\alpha_1, \ldots, -\alpha_{i-1}, -1, -\alpha_{i+1}, \ldots, -\alpha_n)$$

равна нулю. □

Набор векторов $(v_1, \ldots, v_n)$ называются **базисом** векторного пространства V, если он линейно независим и линейная оболочка $\langle v_1, \ldots, v_n \rangle$ равна V. Таким образом, каждый вектор $v \in V$ можно представить в виде

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n,$$

где $\alpha_1, \ldots, \alpha_n \in F$. В этом случае говорят, что вектор v **раскладывается** по базису $(v_1, \ldots, v_n)$ с **коэффициентами** $(\alpha_1, \ldots, \alpha_n)$. Коэффициенты разложения определены однозначно. Действительно, пусть $v = \beta_1 v_1 + \cdots + \beta_n v_n$ — ещё одно разложение. Вычитая второе из первого, получаем $(\alpha_1 - \beta_1) v_1 + \cdots + (\alpha_n - \beta_n) v_n = 0$. Так как $(v_1, \ldots, v_n)$ линейно независим, то $\alpha_i - \beta_i = 0$ для любого i. Следовательно, $\alpha_i = \beta_i$.

Определение 6.6. *Векторное пространство V имеет размерность $n \in \mathbb{Z}$, если максимум длин линейно независимых наборов векторов этого пространства равен n. В этом случае, мы используем обозначение* $\dim V = n$. *Если существуют наборы линейно независимых векторов пространства V сколь угодно большой длины, то пространство V имеет бесконечную размерность. В этом случае, мы используем обозначение* $\dim V = \infty$.

Лемма 6.7. *Пусть* $\dim V = n < \infty$. *Тогда существует базис пространства V, состоящий ровно из n векторов.*

Доказательство. По определению в V существует линейно независимый набор векторов $(v_1, \ldots, v_n)$. Мы утверждаем, что он является базисом пространства V. Действительно, возьмём произвольный вектор $v \in V$ и рассмотрим набор $(v_1, \ldots, v_n, v)$. Этот набор имеет длину $n+1$ и поэтому он линейно зависим. Следовательно

$$\alpha_1 v_1 + \cdots + \alpha_n v_n + \alpha v = 0$$

для некоторых не равных одновременно нулю скаляров $\alpha_1, \ldots, \alpha_n, \alpha$. Предположим, что $\alpha = 0$. Тогда $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$. Так как набор скаляров $(\alpha_1, \ldots, \alpha_n)$ в рассматриваемом случае ненулевой, то мы получаем противоречие с линейной независимостью набора $(v_1, \ldots, v_n)$.

Следовательно, $\alpha \neq 0$ и мы получаем

$$v = -\frac{\alpha_1}{\alpha} v_1 - \cdots - \frac{\alpha_n}{\alpha} v_n \in \langle v_1, \ldots, v_n \rangle.$$

□

Возникает вопрос: *все ли базисы пространства размерности n состоят из n векторов?* Так же хорошо бы было иметь критерий, по которому *можно было бы установить, что пространство n имеет размерность n, на основании мощности какого-нибудь базиса.*

На оба этих вопроса можно ответить при помощи следующего результата.

Лемма 6.8. *Пусть $(g_1, \ldots, g_l)$ — линейно независимый набор векторов, элементы которого принадлежащие линейной оболочке $\langle f_1, \ldots, f_k \rangle$. Тогда $k \geqslant l$.*

Доказательство. Проведём индукцию по k. В качестве базы индукции рассмотрим случай $k = 0$, так как в этом случае $\langle f_1, \ldots, f_k \rangle = \langle\rangle = 0$ и следовательно $l = 0$, потому что набор, содержащий нулевой вектор, линейно зависим.

Пусть теперь $k > 0$ и лемма верна для меньших значений k. Мы можем считать, что $l > 0$, так как иначе наше утверждение превратиться в верное утверждение $k > l = 0$ Мы запишем следующие разложения:

$$\begin{aligned} g_1 &= \alpha_{1,1} f_1 + \alpha_{1,2} f_2 + \cdots + \alpha_{1,k} f_k, \\ g_2 &= \alpha_{2,1} f_1 + \alpha_{2,2} f_2 + \cdots + \alpha_{2,k} f_k, \\ &\vdots \\ g_l &= \alpha_{l,1} f_1 + \alpha_{l,2} f_2 + \cdots + \alpha_{l,k} f_k. \end{aligned}$$

Если $\alpha_{1,k} = \alpha_{2,k} = \cdots = \alpha_{l,k} = 0$, то $g_1, \ldots, g_l$ принадлежат $\langle f_1, \ldots, f_{k-1} \rangle$ и по предположению индукции $k > k - 1 \geqslant l$. Теперь предположим, что $\alpha_{i,k} \neq 0$ для некоторого $i = 1, \ldots, l$. Тогда мы можем выразить f_k следующим образом:

$$f_k = \frac{1}{\alpha_{i,k}} g_i - \frac{\alpha_{i,1}}{\alpha_{i,k}} f_1 - \frac{\alpha_{i,2}}{\alpha_{i,k}} f_2 - \cdots - \frac{\alpha_{i,k-1}}{\alpha_{i,k}} f_{k-1}.$$

Подставляя это значение во все уравнения для $g_1, \ldots, g_{i-1}, g_{i+1}, \ldots, g_l$ и перенося все g (с индексами) влево, а все f (с индексами) вправо получаем

$$\begin{aligned} g'_1 &= \beta_{1,1}f_1 + \beta_{1,2}f_2 + \cdots + \beta_{1,k-1}f_{k-1}, \\ &\vdots \\ g'_{i-1} &= \beta_{i-1,1}f_1 + \beta_{i-1,2}f_2 + \cdots + \beta_{i-1,k-1}f_{k-1}, \\ g'_{i+1} &= \beta_{i+1,1}f_1 + \beta_{i+1,2}f_2 + \cdots + \beta_{i+1,k-1}f_{k-1}, \\ &\vdots \\ g'_l &= \beta_{l,1}f_1 + \beta_{l,2}f_2 + \cdots + \beta_{l,k-1}f_{k-1}, \end{aligned}$$

где

$$g'_m = g_m - \frac{\alpha_{m,k}}{\alpha_{i,k}}g_i \quad \text{и} \quad \beta_{m,j} = \alpha_{m,j} - \frac{\alpha_{m,k}\alpha_{i,j}}{\alpha_{i,k}}.$$

Таким образом, $g'_1, \ldots, g'_{i-1}, g'_{i+1}, \ldots, g'_l \in \langle f_1, \ldots, f_{k-1}\rangle$.

Мы утверждаем, что набор $(g'_1, \ldots, g'_{i-1}, g'_{i+1}, \ldots, g'_l)$ линейно независим. Действительно, пусть

$$\lambda_1 g'_1 + \cdots + \lambda_{i-1}g'_{i-1} + \lambda_{i+1}g'_{i+1} + \cdots + \lambda_l g'_l = 0$$

для некоторых $\lambda_1, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_l \in F$. Перепишем уравнение выше в виде

$$\lambda_1 g_1 + \cdots + \lambda_{i-1}g_{i-1} - \frac{\lambda_1\alpha_{1,k}+\cdots+\lambda_{i-1}\alpha_{i-1,k}+\lambda_{i+1}\alpha_{i+1,k}+\cdots+\lambda_l\alpha_{l,k}}{\alpha_{i,k}}g_i + \lambda_{i+1}g_{i+1} + \cdots + \lambda_l g_l = 0.$$

В силу линейной независимости набора векторов $(g_1, \ldots, g_l)$ все коэффициенты этой линейной комбинации равны нулю. В частности, $\lambda_1 = \cdots = \lambda_{i-1} = \lambda_{i+1} = \cdots = \lambda_l = 0$.

Теперь рассматривая набор $(g'_1, \ldots, g'_{i-1}, g'_{i+1}, \ldots, g'_l)$ и векторы $f_1, \ldots, f_{k-1}$, мы по индуктивному предположению заключаем, что $k - 1 \geqslant l - 1$. Отсюда $k \geqslant l$. □

Следствие 6.9. *Пусть* $\dim V = n < \infty$. *Тогда любой базис пространства* V *содержит ровно* n *элементов.*

Доказательство. Согласно лемме 6.7 в пространстве V существует базис $(g_1, \ldots, g_n)$. Пусть $(f_1, \ldots, f_k)$ — ещё один базис пространства V. Мы получаем $g_1, \ldots, g_n \in V = \langle f_1, \ldots, f_k\rangle$ и набор $(g_1, \ldots, g_n)$ линейно независим. Применяя лемму 6.8 для $l = n$, получаем $k \geqslant n$. Однако, случай $k > n = \dim V$ невозможен по определению размерности. □

Следствие 6.10. *Пусть* V *— векторное пространство и* $V = \langle f_1, \ldots, f_k\rangle$. *Тогда* $\dim V \leqslant k$. *Если при этом набор* $(f_1, \ldots, f_k)$ *линейно независим, то* $\dim V = k$.

Доказательство. Пусть $(g_1, \ldots, g_l)$ — линейно независимый набор векторов пространства V. По лемме 6.8, получаем $l \leqslant k$. Это означает, что $\dim V \leqslant k$. С другой стороны, если $(f_1, \ldots, f_k)$ линейно независим, то $\dim V = k$. □

Следствие 6.11. *Пусть* $\dim V = n < \infty$ *и* $(f_1, \ldots, f_k)$ *— линейно независимый набор векторов пространства* V, *где* $k < n$. *Тогда существуют векторы* $f_{k+1}, \ldots, f_n \in V$ *такие, что* $(f_1, \ldots, f_n)$ *— базис пространства* V.

Доказательство. Достаточно доказать, что можно построить f_{k+1} так, чтобы $(f_1, \ldots, f_{k+1})$ был линейно независимым. Пусть $(e_1, \ldots, e_n)$ — базис пространства V. Предположим, что $e_1, \ldots, e_n \in \langle f_1, \ldots, f_k\rangle$. По лемме 6.8 мы получаем тогда противоречие $k \geqslant n$. Поэтому существует $i = 1, \ldots, n$, для которого $e_i \notin \langle f_1, \ldots, f_k\rangle$. Мы утверждаем, что набор $(f_1, \ldots, f_k, e_i)$ линейно независим. Действительно, пусть

$$\alpha_1 f_1 + \cdots + \alpha_k f_k + \lambda e_i = 0$$

для некоторых $\alpha_1, \ldots, \alpha_i, \lambda \in F$. Если $\lambda \neq 0$, то получаем противоречие

$$e_i = -\frac{\alpha_1}{\lambda} f_1 - \cdots - \frac{\alpha_k}{\lambda} f_k \in \langle f_1, \ldots, f_k \rangle$$

Если же $\lambda = 0$, то получаем $\alpha_1 f_1 + \cdots + \alpha_k f_k = 0$. Отсюда из линейной независимости набора $(f_1, \ldots, f_k)$, получаем $\alpha_1 = \cdots = \alpha_k = 0$. □

Определение 6.12. *Подпространством векторного пространства V называется его непусто подмножество U замкнутое относительно операций сложения и умножения на скаляры. Другими словами U должно удовлетворять следующим свойствам:*

(П1) $U \neq \varnothing$;
(П2) *если* $u, v \in U$, *то* $v + u \in U$;
(П3) *если* $\alpha \in F$ *и* $v \in U$, *то* $\alpha v \in U$.

На самом деле мы уже встречались с подпространствами: линейная оболочка векторов является подпространством. Более того, подпространство само является векторным пространством относительно операций ограничения. Поэтому оно само является линейной оболочкой любого своего базиса.

Лемма 6.13. *Пусть U — подпространство векторного пространства V и $v_1, \ldots, v_n$ — векторы, принадлежащие U. Тогда $\langle v_1, \ldots, v_n \rangle \subset U$.*

Доказательство. Заметим, что $0 = 0 \cdot u \in U$ для любого вектора $u \in U$, который существует в силу (П1). Это доказывает лемму в случае $n = 0$. С другой стороны, для $n > 0$ утверждение следует из свойств (П2) и (П3). □

Элементарными преобразованиями конечного набора векторов называются следующие действия:

(ЭВ1) Поменять местами два вектора.
(ЭВ2) Умножить один из векторов на ненулевой скаляр.
(ЭВ3) Умножить один из векторов на скаляр и прибавить получившееся произведение к другому вектору.
(ЭВ4) Добавить нулевой вектор.
(ЭВ5) Удалить нулевой вектор.

Лемма 6.14. *Элементарные преобразования набора векторов не меняют их линейной оболочки.*

Доказательство. Пусть $(u_1, \ldots, u_m)$ набор векторов, полученный из набора векторов $(v_1, \ldots, v_n)$ одним из элементарных преобразований. Из определения видно, что каждый вектор u_i является линейной комбинацией векторов $v_1, \ldots, v_n$. Поэтому $u_i \in \langle v_1, \ldots, v_n \rangle$. По лемме 6.13 получаем $\langle u_1, \ldots, u_m \rangle \subset \langle v_1, \ldots, v_n \rangle$.

Остаётся только доказать, что элементарные преобразования обратимы. Докажем это для преобразований (ЭВ2) и (ЭВ3). Пусть $(v_1, \ldots, v_n)$ — набор векторов. Умножая i-й вектор на ненулевой скаляр α, получаем набор $(v_1, \ldots, v_{i-1}, \alpha v_i, v_{i+1}, \ldots, v_n)$. Умножая i-й вектор на α^{-1}, получаем исходный набор $(v_1, \ldots, v_n)$. Аналогично, умножая i-й вектор на α и добавляя получившееся произведение к j-у вектору, получаем набор $(v_1, \ldots, v_{j-1}, v_j + \alpha v_i, v_{j+1}, \ldots, v_n)$. Умножая i-й вектор на $-\alpha$ и добавляя получившееся произведение к j-у вектору, получаем исходный набор. □

Из этой леммы, например, следует, что линейная оболочка не меняется при удалении одного из нескольких одинаковых векторов.

Легко проверить, что (теоретико множественное) **пересечение** подпространств тоже является подпространством. С другой стороны, объединение подпространств не

обязано быть подпространством (приведите примеры). В место этого, можно рассмотреть **сумму** подпространств. Пусть V и U — подпространства некоторого векторного пространства. Положим

$$V + U = \{v + u \mid v \in V \text{ и } u \in U\}.$$

Лемма 6.15. *$V \cap U$ и $V + U$ — подпространства. При этом $V + U$ — наименьшее подпространство, содержащее объединение $V \cup U$. Если $V = \langle v_1, \ldots, v_n\rangle$ и $U = \langle u_1, \ldots, u_m\rangle$, то $V + U = \langle v_1, \ldots, v_n, u_1, \ldots, u_m\rangle$.*

Доказательство. Любой вектор $v \in V$ имеет представление $v = v + 0$ и при этом $0 \in U$. Это доказывает, что $V \subset V + U$. Аналогично $U \subset V + U$. Отсюда $V \cup U \subset V + U$. Пусть наоборот W такое линейное подпространство, что $V \cup U \subset W \subset V + U$. Возьмем произвольный вектор из $V + U$. По определению он имеет вид $v + u$, где $v \in V \subset W$ и $u \in U \subset W$. Так как W — подпространство, то $v + u \in W$. Этим мы доказали равенство $W = V + U$. □

Лемма 6.16. *Пусть V и U — конечномерные подпространства некоторого векторного пространства. Тогда*

$$\dim V + U = \dim V + \dim U - \dim V \cap U.$$

Доказательство. Пусть $(a_1, \ldots, a_n)$ — базис V и $(b_1, \ldots, b_m)$ — базис U. Так как $V + U = \langle a_1, \ldots, a_n, b_1, \ldots, b_m\rangle$, то по следствию 6.10, получаем $\dim V + U \leqslant n + m$. Пространство $V \cap U$ тоже конечномерно как подпространство конечномерного пространства V или U. Поэтому все размерности в утверждении леммы конечные.

Пусть $(w_1, \ldots, w_k)$ — базис пересечение $V \cap U$. Так как $V \cap U$ — подпространство пространств V и U, то

$$k = \dim V \cap U \leqslant \dim V = n, \quad k = \dim V \cap U \leqslant \dim U = m.$$

По следствию 6.11 мы получаем, что набор $(w_1, \ldots, w_k)$ можно достроить до базиса $(w_1, \ldots, w_k, v_1, \ldots, v_{n-k})$ пространства V и до базиса $(w_1, \ldots, w_k, u_1, \ldots, u_{m-k})$ пространства U.

Наша цель доказать, что $(w_1, \ldots, w_k, v_1, \ldots, v_{n-k}, u_1, \ldots, u_{m-k})$ — базис суммы $V + U$. По лемме 6.15 и замечанию после леммы 6.14 мы получаем

$$\begin{aligned} V + U = \langle w_1, \ldots, w_k, v_1, \ldots, v_{n-k}\rangle + \langle w_1, \ldots, w_k, u_1, \ldots, u_{m-k}\rangle = \\ = \langle w_1, \ldots, w_k, v_1, \ldots, v_{n-k}, w_1, \ldots, w_k, u_1, \ldots, u_{m-k}\rangle = \\ = \langle w_1, \ldots, w_k, v_1, \ldots, v_{n-k}, u_1, \ldots, u_{m-k}\rangle. \end{aligned}$$

Теперь рассмотрим нулевую линейную комбинацию

$$\alpha_1 w_1 + \cdots + \alpha_k w_k + \beta_1 v_1 + \cdots + \beta_{n-k} v_{n-k} + \gamma_1 u_1 + \cdots + \gamma_{m-k} u_{m-k} = 0, \tag{5}$$

где $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_{n-k}, \gamma_1, \ldots, \gamma_{m-k} \in F$. Перенося, получаем

$$\gamma_1 u_1 + \cdots + \gamma_{m-k} u_{m-k} = -\beta_1 v_1 - \cdots - \beta_{n-k} v_{n-k} - \alpha_1 w_1 - \cdots - \alpha_k w_k.$$

Левая часть принадлежит пространству U, а правая пространству V. Поэтому обе части принадлежат пересечению $V \cap U$. Поэтому возможно разложение

$$\gamma_1 u_1 + \cdots + \gamma_{m-k} u_{m-k} = \lambda_1 w_1 + \cdots + \lambda_k w_k$$

для некоторых $\lambda_1, \ldots, \lambda_k \in F$. Перенося в одну сторону, получаем линейную зависимость

$$\gamma_1 u_1 + \cdots + \gamma_{m-k} u_{m-k} - \lambda_1 w_1 - \cdots - \lambda_k w_k = 0$$

элементов базиса $w_1, \dots, w_k, u_1, \dots, u_{m-k}$ пространства U. Поэтому все её коэффициенты нулевые. В частности $\gamma_1 = \dots = \gamma_{m-k} = 0$. Возвращаясь к уравнению (5), получаем

$$\alpha_1 w_1 + \dots + \alpha_k w_k + \beta_1 v_1 + \dots + \beta_{n-k} v_{n-k} = 0.$$

Отсюда $\alpha_1 = \dots = \alpha_k = \beta_1 = \dots = \beta_{n-k} = 0$, так как $(w_1, \dots, w_k, v_1, \dots, v_{n-k})$ — базис пространства V.

Таким образом, мы доказали, что

$$\dim V + U = n + m - k = \dim V + \dim U - \dim V \cap U.$$

□

7. Евклидовы пространства

Определение 7.1. *Евклидовым пространством называется линейное пространство E над полем действительных чисел $\mathbb{R}$, наделённое отображением $(\_\,,\_) : E \times E \to \mathbb{R}$, которое удовлетворяет следующим свойствам:*

(Евк1) $(u, v) = (v, u)$ *для любых* $u, v \in E$;
(Евк2) $(\alpha u + \beta v, w) = \alpha(u, w) + \beta(v, w)$ *для любых* $u, v, w \in E$ *и* $\alpha, \beta \in \mathbb{R}$;
(Евк3) $(u, u) > 0$ *для любого* $u \in E \setminus \{0\}$.

Заметим, что симметричность (Евк1) доказывает аналог свойства (Евк2) для второго аргумента $(w, \alpha u + \beta v) = \alpha(w, u) + \beta(w, v)$. Кроме того, из свойства (Евк2) получаем $(0, v) = (0+0, v) = (0, v)+(0, v)$. Сокращая на $(0, v)$, получаем $(0, v) = 0$. В дальнейшем, мы часто будем писать $u \cdot v$ или uv вместо (u, v). Следуя этому обозначению мы будем писать v^2 вместо (v, v).

Для каждого вектора $v \in E$ определим его *длину*

$$|v| = \sqrt{v^2}.$$

Извлечение корня корректно, так как $v^2 > 0$. Возводя это равенство в квадрат, получаем

$$|v|^2 = v^2.$$

Пусть $\lambda \in \mathbb{R}$ и $v \in E$. Пользуясь свойством линейности (Евк2), мы получаем

$$|\lambda v| = \sqrt{(\lambda v)^2} = \sqrt{\lambda^2 v^2} = \sqrt{\lambda^2}\sqrt{v^2} = |\lambda||v|.$$

В последней формуле $|\lambda|$ обозначает (обычный) модуль числа, а $|v|$ обозначает длину вектора.

Основным примером (конечномерных) евклидовых пространств является пространство $\mathbb{R}^n$ со следующим скалярным произведением:

$$(x_1, \dots, x_n)(y_1, \dots, y_n) = x_1 y_1 + \dots + x_n y_n.$$

Легко проверить, что свойства (Евк1)–(Евк3) выполняются.

Теорема 7.2 (Неравенство Коши-Буняковского)**.** $|uv| \leqslant |u||v|$ *для любых* $u, v \in E$. *При этом равенство достигается только при пропорциональных* u *и* v.

Доказательство. В случае, когда $u = 0$, векторы u и v пропорциональны и обе части неравенства обращаются в ноль. Поэтому достаточно доказывать теорему в случае $u \neq 0$.

Рассмотрим функцию $f : \mathbb{R} \to \mathbb{R}$, заданную формулой

$$f(t) = (tu + v)^2.$$

Раскрывая скобки по свойству (Евк2) и пользуясь свойством (Евк3), получаем

$$f(t) = u^2 t^2 + 2(uv)t + v^2 \geqslant 0.$$

Получившаяся парабола имеет ветви, направленные вверх. Поэтому квадратный трёхчлен имеет неположительный дискриминант: $D = (2(uv))^2 - 4u^2v^2 \leqslant 0$. Отсюда, получаем $(uv)^2 \leqslant v^2u^2$. Извлекая квадратный корень, получаем

$$|uv| \leqslant \sqrt{u^2}\sqrt{v^2} = |u||v|.$$

Если выполнено равенство $|uv| = |u||v|$, то мы получаем $D = 0$ и следовательно квадратное уравнение $f(t) = 0$ имеет решение $t = t_0$. Это означает, что $(t_0u + v)^2 = 0$. По свойству (Евк3) получаем $t_0v + u = 0$ и следовательно $u = -t_0v$.

Пусть наоборот $v = \lambda u$. В этом случае, получаем

$$|uv| = |u(\lambda u)| = |\lambda u^2| = |\lambda||u^2| = |\lambda|u^2 = |\lambda||u|^2 = |u|(|\lambda||u|) = |u||\lambda u| = |u||v|.$$

В этой записи важно внимательно следить, где $|\ |$ обозначают модуль действительного числа, а где длину вектора. □

Для любого подпространства U евклидового пространства E положим

$$U^\perp = \{e \in E \mid eu = 0 \text{ для любого } u \in U\}.$$

Заметим, что $U \cap U^\perp = \{0\}$. Это следует из части (Евк3) определения евклидового пространства.

Лемма 7.3. (1) *Пусть U — подпространство евклидового пространства E. Тогда $U^\perp$ — тоже подпространство пространства E*

(2) *Пусть U и V — подпространства евклидового пространства E такие, что $U \subset V$. Тогда $V^\perp \subset U^\perp$.*

Доказательство. (1) Множество $U^\perp$ непустой, так как $0 \in U^\perp$. Пусть теперь $e, f \in U^\perp$. Возьмём произвольный вектор $u \in U$. Тогда получаем $(e + f)u = eu + fu = 0$. Следовательно, $e + f \in U^\perp$. Пусть теперь α — произвольное действительное число. Тогда $(\alpha e)u = \alpha(eu) = 0$. Этим доказано, что $\alpha e \in U^\perp$.

(2) Пусть $e \in V^\perp$. Возьмём произвольный вектор $u \in U$. Тогда $u \in V$. Следовательно по определению $eu = 0$. Отсюда $e \in U^\perp$. □

Определение 7.4. *Пусть E — евклидово пространство и $(e_1, \ldots, e_n)$ — его базис. Этот базис называется ортогональным, если $e_ie_j = 0$ для различных i и j. Если дополнительно $|e_i| = 1$, то этот базис называется ортонормированным.*

В ортонормированном базисе $(e_1, \ldots, e_n)$ легко записать координаты вектора. Действительно, пусть

$$v = \alpha_1e_1 + \cdots + \alpha_ne_n.$$

Умножая на e_i обе части, получаем

$$e_iv = \alpha_1e_ie_1 + \cdots + \alpha_{i-1}e_ie_{i-1} + \alpha_ie_i^2 + \alpha_{i+1}e_ie_{i+1} + \cdots + \alpha_ne_ie_n = \alpha_ie_i^2.$$

Отсюда $\alpha_i = e_iv/e_i^2$.

Используя символ Кронекера

$$\delta_{i,j} = \begin{cases} 1, & \text{если } i = j; \\ 0, & \text{если } i \neq j, \end{cases}$$

мы можем записать условие ортонормированности базиса $e_1, \ldots, e_n$ следующим образом

$$e_ie_j = \delta_{i,j}.$$

Если у нас имеется ортогональный базис $(f_1, \ldots, f_n)$ пространства, то из него легко получить ортонормированный базис $(e_1, \ldots, e_n)$, положив $e_i = f_i/f_i^2$.

Теорема 7.5. (1) *В произвольном конечномерном евклидовом пространстве существует хотя бы один ортогональный базис. Более того, любой ненулевой вектор из E является элементом одного из таких базисов.*

(2) *Для любого подпространства U конечномерного евклидового пространства E выполнено $E = U \oplus U^{\perp}$.*

Доказательство. Мы докажем эту теорему при помощи процесса ортогонализации. Пусть $(v_1, \dots, v_n)$ — произвольный базис пространства E. Мы построим базис $(e_1, \dots, e_n)$ по индукции следующим образом:

$$e_1 = v_1, \quad e_i = v_i - \alpha_{i,1}e_1 - \dots - \alpha_{i,i-1}e_{i-1}, \text{ где } \alpha_{i,j} = \frac{v_i e_j}{e_j^2}.$$

В этом определение есть одна проблема, мы делим на возможно нулевое число e_j^2. Мы разрешим эту проблему следующим образом: индукцией по $i = 1, \dots, n$ мы докажем, что

$$\langle v_1, \dots, v_i \rangle = \langle e_1, \dots, e_i \rangle. \tag{6}$$

Действительно, $e_1 = v_1$, что доказывает утверждение для $i = 1$. Пусть теперь $1 < i \leqslant n$ и утверждение выполнено для $i - 1$ вектора, то есть

$$\langle v_1, \dots, v_{i-1} \rangle = \langle e_1, \dots, e_{i-1} \rangle.$$

Докажем равенство (6). Во первых заметим, что если бы $e_j = 0$ для некоторого $j = 1, \dots, i-1$, то мы получили бы, что векторы линейно независимого набора $(v_1, \dots, v_{i-1})$ принадлежат линейной оболочке $i-2$ векторов $\langle e_1, \dots, e_{j-1}, e_{j+1}, \dots, e_{i-1} \rangle$. По лемме 6.8 получаем противоречие $i - 2 \geqslant i - 1$. Из доказанного факта $e_1, \dots, e_{i-1} \neq 0$ следует, что e_i — корректно определённый вектор. Отсюда

$$v_i = e_i + \alpha_{i,1}e_1 + \dots + \alpha_{i,i-1}e_{i-1} \in \langle e_1, \dots, e_i \rangle.$$

Так как $v_1, \dots, v_{i-1} \in \langle e_1, \dots, e_{i-1} \rangle \subset \langle e_1, \dots, e_i \rangle$, то по лемме 6.13 получаем $\langle v_1, \dots, v_i \rangle \subset \langle e_1, \dots, e_i \rangle$.

Наоборот, по предположению индукции $e_1, \dots, e_{i-1} \in \langle v_1, \dots, v_{i-1} \rangle \subset \langle v_1, \dots, v_i \rangle$ и $v_i \in \langle v_1, \dots, v_i \rangle$. Следовательно,

$$e_i = v_i - \alpha_{i,1}e_1 - \dots - \alpha_{i,i-1}e_{i-1} \in \langle v_1, \dots, v_i \rangle.$$

Поэтому по лемме 6.13 получаем $\langle e_1, \dots, e_i \rangle \subset \langle v_1, \dots, v_i \rangle$. Таким образом, равенство (6) доказано и наш алгоритм может продолжаться, пока не будут построены все n векторов $e_1, \dots, e_n$. Из этого равенства для $i = n$ следует, что $\langle e_1, \dots, e_n \rangle = V$. Если бы набор $(e_1, \dots, e_n)$ был линейно зависимым, то $V = \langle e_1, \dots, e_n \rangle = \langle e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_n \rangle$ для некоторого i. Так как $(v_1, \dots, v_n)$ линейно независим, то по лемме 6.8, получаем противоречие $n - 1 \geqslant n$.

Остаётся доказать, что $e_i e_j = 0$ для различных $i \neq j$. Без ограничения общности можно считать, что $i > j$. Мы докажем это утверждение индукцией по i. Для $i = 1$ это утверждение очевидно, так как не существует j меньшего i. В противном случае по индуктивному предположению получаем

$$e_i e_j = v_i e_j - \alpha_{i,1}e_1 e_j - \dots - \alpha_{i,j-1}e_{j-1}e_j - \alpha_{i,j}e_j^2 - \\ - \alpha_{i,j+1}e_{j+1}e_j - \dots - \alpha_{i,i-1}e_{i-1}e_j = v_i e_j - \alpha_{i,j}e_j^2 = v_i e_j - \frac{v_i e_j}{e_j^2}e_j^2 = 0.$$

Наконец, заметим, что если v — ненулевой вектор из E, то набор, состоящий из одного вектора v линейно независим. Отсюда по следствию 6.11 получаем, что существует базис $(v_1, \dots, v_n)$ пространства E такой, что $v_1 = v$. В процессе ортогонализации мы получим ортогональный базис $(e_1, \dots, e_n)$ для которого $e_1 = v_1 = v$.

Часть (1) автоматически следует из этой конструкции. Докажем теперь часть (2). По следствию 6.11 существует такой базис $(v_1, \dots, v_n)$ пространства E, что набор $(v_1, \dots, v_m)$ образует базис пространства U, где $m = \dim U$. Применим процесс ортогонализации к набору $(v_1, \dots, v_n)$ и получим ортогональный базис $(e_1, \dots, e_n)$. По уравнению (6) получаем $\langle e_1, \dots, e_m \rangle = \langle v_1, \dots, v_m \rangle = U$. Мы утверждаем, что $U^\perp = \langle e_{m+1}, \dots, e_n \rangle$. Действительно, в силу ортогональности базиса $(e_1, \dots, e_n)$ получаем $e_{m+1}, \dots, e_n \in U^\perp$. Следовательно, $\langle e_{m+1}, \dots, e_n \rangle \subset U^\perp$. Наоборот, пусть $v \in U^\perp$. Разложим этот вектор по базису $v = \alpha_1 e_1 + \dots + \alpha_n e_n$. Предположим, что $\alpha_i \neq 0$ для какого-нибудь $i = 1, \dots, m$. Умножая на e_i, получаем

$$e_i v = \alpha_1 e_i e_1 + \dots + \alpha_{i-1} e_i e_{i-1} + \alpha_i e_i^2 + \alpha_{i+1} e_i e_{i+1} + \dots + \alpha_n e_i e_n = \\ = 0 + \dots + 0 + \alpha_i e_i^2 + 0 + \dots + 0 = \alpha_i e_i^2 \neq 0,$$

что противоречит тому, что $v \in U^\perp$. Таким образом $\alpha_1 = \dots = \alpha_m = 0$ и $v = \alpha_{m+1} e_{m+1} + \dots + \alpha_n e_n \in \langle e_{m+1}, \dots, e_n \rangle$. Отсюда получаем

$$E = \langle e_1, \dots, e_m \rangle \oplus \langle e_{m+1}, \dots, e_n \rangle = U \oplus U^\perp.$$

□

Следствие 7.6. *Пусть E — конечномерное евклидово пространство и U — его подпространство. Тогда $\dim U^\perp = \dim E - \dim U$ и $(U^\perp)^\perp = U$.*

Доказательство. Первое утверждение следует непосредственно из разложения $E = U \oplus U^\perp$ предыдущей теоремы. Для того чтобы доказать второе утверждение заметим, что

$$U \subset (U^\perp)^\perp. \tag{7}$$

Действительно, возьмём произвольный вектор $u \in U$. Нам требуется доказать, что $u \in (U^\perp)^\perp$. Это утверждение эквивалентно тому, что

$$uv = 0 \text{ для любого } v \in U^\perp.$$

Но это утверждение выполнено в силу определения пространства $U^\perp$ и коммутативности скалярного произведения. Кроме того в силу уже доказанного равенства для размерностей, мы получаем

$$\dim (U^\perp)^\perp = \dim E - \dim U^\perp = \dim E - (\dim E - \dim U) = \dim U.$$

Отсюда в силу (7) и ... мы получаем требуемое равенство $(U^\perp)^\perp = U$. □

Замечание. Включение (7) выполнено в любом (не обязательно конечномерном) евклидовом пространстве.

Следствие 7.7. *Пусть U и V — подпространства евклидова пространства E. Тогда $(U + V)^\perp = U^\perp \cap V^\perp$.*

Доказательство. Заметим, что $U \subset U + V$. Отсюда по лемме 7.3 получаем $(U + V)^\perp \subset U^\perp$. Аналогично, $(U + V)^\perp \subset V^\perp$. Таким образом получаем

$$(U + V)^\perp \subset U^\perp \cap V^\perp. \tag{8}$$

Докажем обратное включение. Пусть $e \in U^\perp \cap V^\perp$. Выберем произвольный вектор $w \in U + V$. Тогда $w = u + v$ для некоторых $u \in U$ и $v \in V$. Отсюда получаем

$$ew = e(u + v) = eu + ev = 0 + 0 = 0.$$

Этим доказано, что $e \in (U + V)^\perp$. □

Следствие 7.8. *Пусть U и V — подпространства конечномерного евклидова пространства E. Тогда $U + V = (U^\perp \cap V^\perp)^\perp$.*

Доказательство. Возьмём ортогональные дополнения от обеих частей основного равенства следствия 7.7. Используя следствие 7.6 получаем

$$U + V = ((U + V)^\perp)^\perp = (U^\perp \cap V^\perp)^\perp.$$

□

Замечание. Без условия конечномерности мы получаем всего лишь включение

$$U + V \subset (U^\perp \cap V^\perp)^\perp.$$

8. Системы линейных уравнений

Определение 8.1. *Линейным уравнением относительно переменных $x_1, \dots, x_n$ над полем F называется уравнение вида $a_1x_1 + \cdots + a_nx_n = b$, где $a_1, \dots, a_n, b \in F$. Конечный набор таких уравнений называется системой линейных уравнений над полем F. Системы записываются так*

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n &= b_1; \\ \vdots & \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n &= b_m; \end{cases} \tag{9}$$

Решением этой системы называется такой набор элементов поля $(c_1, \dots, c_m) \in F$, что подстановка $x_1 = c_1, \dots, x_n = c_n$ приводит все уравнения системы к верным равенствам.

Заметим, что линейные уравнения можно умножать на элементы поля и складывать:

$$c(a_1x_1 + \cdots + a_nx_n = b) \quad \Longrightarrow \quad ca_1x_1 + \cdots + ca_nx_n = cb$$

$$(a_1x_1 + \cdots + a_nx_n = b) + (a'_1x_1 + \cdots + a'_nx_n = b') \quad \Longrightarrow$$
$$\Longrightarrow \quad (a_1 + a'_1)x_1 + \cdots + (a_n + a'_n)x_n = b + b'.$$

При этом, как мы видим, получаются опять линейные уравнения. Легко заметить, что множество всех линейных уравнений от фиксированных n переменных представляет собой $n + 1$-мерное векторное пространство. Нулём этого пространства, очевидно, является нулевое уравнение $0x_1 + \cdots + 0x_n = 0$, и уравнением обратным к уравнению $a_1x_1 + \cdots + a_nx_n = b$ является уравнение $-a_1x_1 - \cdots - a_nx_n = -b$.

Элементарными преобразованиями конечной системы линейных уравнений называются следующие действия:

(ЭС1) Поменять местами два уравнения.
(ЭС2) Умножить одно из уравнений на ненулевой элемент поля.
(ЭС3) Умножить одно из уравнений на элемент поля и прибавить получившееся произведение к другому уравнению.
(ЭС4) Добавить нулевое уравнение.
(ЭС5) Удалить нулевое уравнение.

Эти преобразование системы линейных уравнений находятся в полной аналогии с уже встречавшимися у нас преобразованиями набора векторов (ЭВ1)–(ЭВ5) на странице 21 (и являются их частным случаем).

Лемма 8.2. *Элементарные преобразования сохраняют множества решений системы.*

Доказательство. Доказательство леммы 6.14 показывает, что линейные преобразования обратимы: если вторая системы получается из первой элементарными преобразованиями, то и первая получается из второй элементарными преобразованиями.

Поэтому достаточно доказать следующее утверждение: если $x_1 = c_1, \ldots, x_n = c_n$ — решение системы (9), то этот набор является решение любой системы, полученной из системы (9) одним элементарным преобразованием.

Достаточно рассмотреть преобразования (ЭС2) и (ЭС3). Сначала умножим i-е уравнение системы (9) на $\alpha \in F \setminus \{0\}$. Получаем уравнение $\alpha a_{i,1}x_1 + \cdots + \alpha a_{i,n}x_n = \alpha b_i$. Проверим, что подстановка $x_1 = c_1, \ldots, x_n = c_n$ даёт верное равенство:

$$\alpha a_{i,1}c_1 + \cdots + \alpha a_{i,n}c_n = \alpha(a_{i,1}c_1 + \cdots + a_{i,n}c_n) = \alpha b_i.$$

В этом вычислении мы использовали верное по условию i-е уравнение системы (9)

$$a_{i,1}c_1 + \cdots + a_{i,n}c_n = b_i.$$

Теперь умножим i-е уравнение системы (9) на $\alpha \in F$ и добавим к j-у, где $i \neq j$. Получаем уравнение $(\alpha a_{i,1} + a_{j,1})x_1 + \cdots + (\alpha a_{i,n} + a_{j,n})x_n = \alpha b_i + b_j$. Проверим, что подстановка $x_1 = c_1, \ldots, x_n = c_n$ даёт верное равенство:

$$\begin{aligned}(\alpha a_{i,1} + a_{j,1})c_1 + \cdots + (\alpha a_{i,n} + a_{j,n})c_n = \alpha(a_{i,1}c_1 + \cdots + a_{i,n}c_n) + \\ + a_{j,1}c_1 + \cdots + a_{j,n}c_n = \alpha b_i + b_j.\end{aligned}$$

В этом вычислении мы использовали верные по условию i-е и j-е уравнения системы (9)

$$a_{i,1}c_1 + \cdots + a_{i,n}c_n = b_i, \quad a_{j,1}c_1 + \cdots + a_{j,n}c_n = b_j.$$

□

Опишем теперь процесс решения системы линейных уравнений при помощи элементарных преобразований, так же известный как **метод Гаусса**. Для этого опишем процесс приведения системы к **ступенчатой системе**, то есть, к системе вида

$$\begin{cases} a_{1,j_1}x_{j_1} + \cdots + a_{1,j_2}x_{j_2} + \cdots + a_{1,j_k}x_{j_k} + \cdots + a_{1,n}x_n = b_1; \\ \qquad\qquad a_{2,j_2}x_{j_2} + \cdots + a_{2,j_k}x_{j_k} + \cdots + a_{1,n}x_n = b_2; \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \vdots \\ \qquad\qquad\qquad\qquad a_{k,j_k}x_{j_k} + \cdots + a_{k,n}x_n = b_k; \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad 0 = b_{k+1}; \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \vdots \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad 0 = b_l, \end{cases}$$

где $j_1 < \cdots < j_k$ и $a_{1,j_1} \neq 0$, $a_{2,j_2} \neq 0 \ldots$, $a_{k,j_k} \neq 0$. Возможны все случаи параметров k и l в пределах неравенств $0 \leqslant k \leqslant l$.

Применим индукцию по количеству уравнений. База индукции — это пустая система. Теперь предположим, что наша система содержит хотя бы одно уравнение. Запишем её в виде (9). Если все коэффициенты $a_{i,j} = 0$, то система содержит только уравнения вида $0 = b_i$. Эта система уже ступенчатая.

Предположим теперь, что существует $a_{i,j} \neq 0$. Выберем в качестве j_1 наименьший индекс, для которого существует индекс i_1 такой, что $a_{i_1,j_1} \neq 0$. Меняя местами первое и i_1-е уравнения (преобразование (ЭС1)), без ограничения общности можно считать, что $a_{1,j_1} \neq 0$ в системе (9). Для каждого $i = 2, \ldots, m$ умножим первое уравнение на $-a_{i,j_1}/a_{1,j_1}$ и добавим его к i-у уравнению (преобразование (ЭС3)). В результате

получаем систему вида

$$\begin{cases} a_{1,j_1}+ & a_{1,j_1+1}x_{j_1+1}+\cdots+a_{1,n}x_n &= b_1; \\ & a'_{2,j_1+1}x_{j_1+1}+\cdots+a'_{2,n}x_n &= b'_2; \\ & \vdots & \vdots \\ & a'_{m,j_1+1}x_{j_1+1}+\cdots+a'_{m,n}x_n &= b'_m; \end{cases}$$

Собственно вычисления требуют элементы поля $a'_{i,j}$ и b'_i. Теперь мы забываем про первое уравнение и повторяем наш (рекурсивный) алгоритм к оставшимся уравнениям.

9. Матрицы

Матрицей называется прямоугольный массив элементов некоторого кольца. Матрица может быть записана в следующем виде

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}.$$

Мы обычно обозначаем эту матрицу с помощью записи $A = (a_{i,j})_{i=1,\ldots,m;j=1,\ldots,n}$. Кроме того, мы будем обозначать элемент, стоящий на пересечении i-й и j-й строки матрицы, путём взятия обозначения матрицы в скобки и приписывания нижнего индекса i, j.

9.1. **Виды матриц.** Особую роль будут играть *квадратные матрицы*, то есть матрицы размера $n \times n$. Про такую матрицу мы будем говорить, что она является *квадратной матрицей размера* n. Для такой матрицы мы будем применять обозначение $A = (a_{i,j})_{i=1}^n$. В квадратных матрицах принято выделять *главную диагональ*, состоящую из позиций на пересечении i-й строки и i-о столбца и *побочную диагональ*, состоящую из позиций на пересечении i-й строки и $n+1-i$-о столбца, где n — размер матрицы.

Особым видом квадратной матрицы является *единичная матрица*

$$E = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

В этой матрице на главной диагонали стоят единицы, а в остальных местах стоят нули. Мы будем обозначать также через $E^{(n)}$ единичную матрицу размера n, если нам надо уточнить размер матрицы.

Матрица A^{-1} называется обратной матрицей квадратной матрицы A, если выполнены равенства $AA^{-1} = A^{-1}A = E$. Не для любой матрицы существует обратная. Если же обратная матрица A^{-1} существует, то матрица A называется *обратимой*.

Ступенчатой матрицей мы будем называть матрицу вида

$$B = \begin{pmatrix} 0 & \cdots & 0 & b_{1,j_1} & & & \cdots & & & b_{1,n} \\ 0 & & \cdots & & 0 & b_{2,j_2} & & \cdots & & b_{2,n} \\ & & & & & \cdots & & & & \\ 0 & & & \cdots & & & 0 & b_{k,j_k} & \cdots & b_{k,n} \\ 0 & & & & & \cdots & & & & 0 \\ & & & & & \cdots & & & & \\ 0 & & & & & \cdots & & & & 0 \end{pmatrix}, \tag{10}$$

где $j_1 < j_2 < \cdots < j_k$ и $b_{1,j_1} \neq 0$, $b_{2,j_2} \neq 0$,...,$b_{k,j_k} \neq 0$.

Лемма 9.1. *Ненулевые строки ступенчатой матрицы над полем линейно независимы.*

Доказательство. Пусть r_i обозначает i-у строку ступенчатой матрицы (10). Предположим, что выполнено равенство

$$\alpha_1 r_1 + \alpha_2 r_2 + \cdots + \alpha_k r_k = 0. \tag{11}$$

Мы докажем индукцией по $l = 0, \ldots, k$, что $\alpha_1 = \cdots = \alpha_l = 0$. В случае $l = 0$ доказывать нечего, так как в этом случае мы ничего не утверждаем. Пусть теперь $1 \leqslant l \leqslant k$ и утверждение верно для $l - 1$. Таким образом мы уже знаем, что $\alpha_1 = \cdots = \alpha_{l-1} = 0$. Отсюда и из равенства (11) получаем

$$\alpha_l r_l + \alpha_{l+1} r_{l+1} + \cdots + \alpha_k r_k = 0.$$

Посмотрев на j_l-ю компоненту мы получаем равенство

$$\alpha_l b_{l,j_l} + \alpha_{l+1} 0 + \cdots + \alpha_k 0 = 0.$$

Так как $b_{l,j_l} \neq 0$ и мы работаем над полем, то $\alpha_l = 0$. □

9.2. **Операции над матрицами.** Наша следующая цель — определить линейные действия над матрицами и умножения матриц. Сложение матриц определяется определяется поэлементно:

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} + \begin{pmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & \ddots & \vdots \\ b_{m,1} & \cdots & b_{m,n} \end{pmatrix} = \begin{pmatrix} a_{1,1} + b_{1,1} & \cdots & a_{1,n} + b_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} + b_{m,1} & \cdots & a_{m,n} + b_{m,n} \end{pmatrix}$$

и умножение на элемент кольца $x \in R$ также

$$x \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} = \begin{pmatrix} xa_{1,1} & \cdots & xa_{1,n} \\ \vdots & \ddots & \vdots \\ xa_{m,1} & \cdots & xa_{m,n} \end{pmatrix},$$

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} x = \begin{pmatrix} a_{1,1}x & \cdots & a_{1,n}x \\ \vdots & \ddots & \vdots \\ a_{m,1}x & \cdots & xa_{m,n}x \end{pmatrix}$$

Умножение матриц — несколько более сложная операция. Мы дадим сейчас точное определение, а затем обсудим различные задачи, которые мотивируют такое определение. Пусть $A = (a_{i,k})_{i=1,\ldots,m;k=1,\ldots,n}$ и $B = (b_{k,j})_{k=1,\ldots,n;j=1,\ldots,l}$ — матрицы. Мы обозначим их произведение через $AB = C = (c_{i,j})_{i=1,\ldots,m;j=1,\ldots,l}$. В этом случае

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

для любых $i = 1, \ldots, m$ и $j = 1, \ldots, l$. Матрицы C имеет размер $m \times l$. Таким образом *произведение матриц AB определено тогда и только тогда, когда число столбцов матрицы A равно числу строк матрицы B. При этом число строк произведения AB равно числу строк матрицы A и число столбцов произведения AB равно числу столбцов матрицы B.*

Лемма 9.2. *Умножение матриц ассоциативно и дистрибутивно: для любых матриц подходящего размера $(AB)C = A(BC)$, $(A + B)C = AC + BC$ и $A(B + C) = AB + AC$.*

Доказательство. Докажем свойство ассоциативности. Рассмотрим матрицы

$$A = (a_{i,p})_{i=1,\dots,n;p=1,\dots,m}, \quad B = (b_{p,q})_{p=1,\dots,m;q=1,\dots,k}, \quad C = (c_{q,j})_{q=1,\dots,k;j=1,\dots,l}.$$

Положим

$$D = AB, \qquad F = BC, \qquad G = DC, \qquad H = AF.$$

Мы будем использовать следующие обозначения для элементов этих матриц:

$$D = (d_{i,q})_{i=1,\dots,n;q=1,\dots,k}, \quad F = (f_{p,j})_{p=1,\dots,m;j=1,\dots,l},$$

$$G = (g_{i,j})_{i=1,\dots,n;j=1,\dots,l}, \quad H = (h_{i,j})_{i=1,\dots,n;j=1,\dots,l}.$$

Наша цель — доказать, что $G = H$. Для любых $i = 1, \dots, n$ и $j = 1, \dots, l$ получаем

$$h_{i,j} = \sum_{p=1}^{m} a_{i,p} f_{p,j} = \sum_{p=1}^{m} a_{i,p} \left(\sum_{q=1}^{k} b_{p,q} c_{q,l} \right) = \sum_{p=1}^{m} \sum_{q=1}^{k} a_{i,p} b_{p,q} c_{q,l}.$$

$$g_{i,j} = \sum_{q=1}^{k} d_{i,q} c_{q,j} = \sum_{q=1}^{k} \left(\sum_{p=1}^{m} a_{i,p} b_{p,q} \right) c_{q,j} = \sum_{q=1}^{k} \sum_{p=1}^{m} a_{i,p} b_{p,q} c_{q,j}.$$

Требуемое равенство $h_{i,j} = g_{i,j}$ получается как следсвтие перестановочности независимых суммирований $\sum_{p=1}^{m}$ и $\sum_{q=1}^{k}$.

Теперь докажем дистрибутивность. Рассмотрим матрицы

$$A = (a_{i,p})_{i=1,\dots,n;p=1,\dots,m}, \quad B = (b_{i,p})_{i=1,\dots,n;p=1,\dots,m}, \quad C = (c_{p,j})_{p=1,\dots,m;j=1,\dots,k}.$$

Положим

$$D = A + B, \qquad F = AC, \qquad G = BC, \qquad H = DC.$$

Мы будем использовать следующие обозначения для элементов этих матриц:

$$D = (d_{i,p})_{i=1,\dots,n;p=1,\dots,m}, \quad F = (f_{i,j})_{i=1,\dots,n;j=1,\dots,k},$$

$$G = (g_{i,j})_{i=1,\dots,n;j=1,\dots,k}, \quad H = (h_{i,j})_{i=1,\dots,n;j=1,\dots,k}.$$

Наша цель — доказать, что $H = F + G$. Для любых $i = 1, \dots, n$ и $j = 1, \dots, k$ получаем

$$h_{i,j} = \sum_{p=1}^{m} d_{i,p} c_{p,j} = \sum_{p=1}^{m} (a_{i,p} + b_{i,p}) c_{p,j} = \sum_{p=1}^{m} (a_{i,p} c_{p,j} + b_{i,p} c_{p,j}) =$$

$$= \sum_{p=1}^{m} a_{i,p} c_{p,j} + \sum_{p=1}^{m} b_{i,p} c_{p,j} = f_{i,j} + g_{i,j}.$$

И так мы доказали, что $(A + B)C = AC + BC$. Равенство $A(B + C) = AB + AC$ доказывается аналогично. □

Определение 9.3. *Транспонированной матрицей матрицы* $A = (a_{i,j})_{i=1,\dots,m;j=1,\dots,n}$ *называется матрица* $A^T = (a_{i,j})_{j=1,\dots,n;i=1,\dots,m}$.

Таким образом, если A имеет размер $m \times n$, то A^T имеет размер $n \times m$.

Лемма 9.4. $(AB)^T = B^T A^T$ *и* $(A + B)^T = A^T + B^T$.

Доказательство. Пусть $A = (a_{i,p})_{i=1,\dots,m;p=1,\dots,n}$ и $B = (b_{p,j})_{p=1,\dots,n;j=1,\dots,k}$. Положим $C = AB$ и $D = B^T A^T$. Мы будем использовать следующие обозначения для элементов этих матриц:

$$A^T = (a^T_{p,i})_{p=1,\dots,n;i=1,\dots,m}, \quad B^T = (b^T_{j,p})_{j=1,\dots,k;p=1,\dots,n},$$

$$C = (c_{i,j})_{i=1,\dots,m;j=1,\dots,k}, \quad C^T = (c^T_{j,i})_{j=1,\dots,k;i=1,\dots,m}, \quad D = (d_{j,i})_{j=1,\dots,k;i=1,\dots,m}.$$

Таким образом $a^T_{p,i} = a_{i,p}$, $b^T_{j,p} = b_{p,j}$ и $c^T_{j,i} = c_{i,j}$. Наша цель доказать, что $C^T = D$.

Для произвольных $j = 1, \dots, k$ и $i = 1, \dots, m$ мы получаем

$$d_{j,i} = \sum_{p=1}^{n} b^T_{j,p} a^T_{p,i} = \sum_{p=1}^{n} b_{p,j} a_{i,p} = \sum_{p=1}^{n} a_{i,p} b_{p,j} = c_{i,j} = c^T_{j,i}.$$

Доказательство второго равенства оставляем читателю. □

Следствие 9.5. *Для любой обратимой матрицы A, выполнено равенство $(A^{-1})^T = (A^T)^{-1}$.*

Доказательство. Нам требуется доказать следующие равенства $A^T(A^{-1})^T = E$ и $(A^{-1})^T A^T = E$. Докажем первое с помощью леммы 9.4:

$$A^T(A^{-1})^T = (A^{-1}A)^T = E^T = E.$$

Второе равенство доказывается аналогично. □

Кроме транспонированной матрицы иногда бывает удобно (смотрите, например доказательство леммы 9.13) рассматривать псевдотранспонированную матрицу: $A^t = (a_{m+1-i,n+1-j})_{j=1,\dots,n;i=1,\dots,m}$, если $A = (a_{i,j})_{i=1,\dots,m;j=1,\dots,n}$.

Заметим, что для любой матрицы A размера $m \times n$ мы получаем $E^{(m)}A = AE^{(n)} = A$.

Умножение на скаляр $x \in R$ легко переписать в терминах умножения матриц:

$$xA = (xE)A = A(xE).$$

Эти факты легко получаются, если заметить, что в матрице xE все элементы равны нулю кроме диагональных, которые равны x. Так как 1 коммутирует со всеми элементами кольца R, то $Ex = xE$.

Следствие 9.6. *Для любых матриц A и B подходящего размера и элемента кольца $x \in R$ выполнены равенства $(xA)B = x(AB)$, $(Ax)B = A(xB)$ и $(AB)x = A(Bx)$.*

Доказательство. Докажем первое равенство при помощи леммы 9.2:

$$(xA)B = ((xE)A)B = (xE)(AB) = x(AB).$$

Остальные равенства доказываются аналогично. □

9.3. **Матрицы элементарных преобразований.** Мы обозначим через $E_{i,j}$ или более точно $E^{(m,n)}_{i,j}$ матрицу размера $m \times n$ в которой на всех местах стоит 0 кроме пересечения i-й строки и j-о столбца, где стоит 1. Произвольная матрица A размера $m \times n$ имеет следующее разложение:

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j} E_{i,j}.$$

Это разложение мы будем часть применять в различных доказательствах. Матрицы $E_{i,j}$ мы будем называть *матричными единицами*. Легко понять как устроено умножение матриц на матричные единицы.

Лемма 9.7. *Пусть A — матрица размера $m \times n$ и i, j, k — натуральные числа такие, что $i \leqslant k$ и $j \leqslant k$. Матрица $E_{i,j}^{(k,m)}A$ состоит из j-й строки матрицы A, записанной на месте i-й строки, и нулями в остальных местах. Аналогично, матрица $AE_{i,j}^{(n,k)}$ состоит из i-го столбца матрицы A, записанного на месте j-го столбца, и нулями в остальных местах.*

Доказательство. Обозначим $B = E_{i,j}^{(k,m)}A$. Мы используем следующую запись для элементов этих матриц:

$$A = (a_{j,q})_{j=1,\dots,m;q=1,\dots,n}, \quad B = (b_{p,q})_{p=1,\dots,k;q=1,\dots,n}.$$

Для $p = 1, \dots, k$ и $q = 1, \dots, n$ таких, что $p \neq j$ и мы получаем

$$b_{p,q} = \sum_{s=1}^{m} (E_{i,j})_{p,s} a_{s,q} = \sum_{s=1}^{m} 0 a_{s,q} = 0.$$

С другой стороны, для любого $q = 1, \dots, n$ получаем

$$b_{i,q} = \sum_{s=1}^{m} (E_{i,j})_{i,s} a_{s,q} = (E_{i,j})_{i,j} a_{j,q} = a_{j,q}.$$

Утверждение о матрице $AE_{i,j}^{(n,k)}$ доказывается аналогично. □

Теперь мы посвятим некоторое время тому, чтобы "оправдать" наше опредение произведения матриц. Мы рассмотрим задачи, в которых оно естественно возникает.

Заметим, что систему (9) можно записать в следующем виде:

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Таким образом, вся информация о системе кроме названия переменных содержится в следующей матрице:

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & \cdots & a_{m,n} & b_m \end{pmatrix}$$

которую мы назовём *расширенной матрицей системы*. Мы опишем как из изменяется эта матрица при элементарных преобразованиях (ЭС1)–(ЭС5) на языке умножения матриц.

Рассмотрим следующие матрицы:

- $X_{i,j}^{(n)} = E_{i,j} + E_{j,i} + \sum\limits_{\substack{k=1 \\ k \neq i, k \neq j}}^{n} E_{k,k}$ размера $n \times n$, где $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$ и $i \neq j$.
- $X_{i}^{(n)}(\alpha) = \alpha E_{i,i} + \sum\limits_{\substack{k=1 \\ k \neq i}}^{n} E_{k,k}$ размера $n \times n$, где $\alpha \in R$ и $1 \leqslant i \leqslant n$.
- $X_{i,j}^{(n)}(\alpha) = E + \alpha E_{i,j}$ размера $n \times n$, где $\alpha \in R$, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$ и $i \neq j$.
- $Y_{m,n} = \sum\limits_{i=1}^{\min\{m,n\}} E_{i,i}$ размера $m \times n$.

Лемма 9.8. *Пусть A — матрица размера $m \times n$.*

(1) *Умножение матрицы A на матрицу $X_{i,j}^{(m)}$ слева меняет местами i-ю и j-ю строки, а умножение на $X_{i,j}^{(n)}$ справа меняет местами i-й и j-й столбцы. Таким образом, преобразование* (ЭС1) *соответствует умножению расширенной матрицы системы на $X_{i,j}^{(m)}$ слева.*

(2) *Умножение матрицы A на матрицу $X_i^{(m)}(\alpha)$ слева умножает i-ю строку на α, а умножение на $X_i^{(n)}(\alpha)$ справа умножает i-й столбец на α. Таким образом, преобразование* (ЭС2) *соответствует умножению расширенной матрицы системы на $X_i^{(m)}(\alpha)$ слева.*

(3) *Умножение матрицы A на матрицу $X_{i,j}^{(n)}(\alpha)$ слева умножает j-ю строку на α и прибавляет результат умножения к i-й строке, а умножение на $X_{i,j}^{(n)}(\alpha)$ справа умножает i-й столбец на α и прибавляет результат умножения к j-у столбцу. Таким образом, преобразование* (ЭС3) *соответствует умножению расширенной матрицы системы на $X_{i,j}^{(n)}(\alpha)$ слева.*

(4) *Умножение матрицы A на матрицу $Y_{m-1,m}$ слева удаляет нижнюю строку, а умножение на $Y_{n,n-1}$ справа удаляет самый правый столбец. Наоборот, умножение матрицы A на матрицу $Y_{m+1,m}$ слева добавляет нулевую строку снизу, а умножение на $Y_{n,n+1}$ справа добавляет нулевой столбец справа. Таким образом, преобразование* (ЭС4) *соответствует умножению расширенной матрицы системы на $Y_{m+1,m}$ слева, а преобразование* (ЭС5) *соответствует умножению расширенной матрицы системы на $Y_{m-1,m}$ слева.*

Доказательство. Достаточно применим лемму 9.7. □

Следствие 9.9. *Любая матрица A над полем может быть представлена в виде $A = X_1 \cdots X_k B$, где $X_1, \ldots, X_k$ — матрицы вида $X_{i,j}^{(n)}$, $X_i^{(n)}(\alpha)$ или $X_{i,j}^{(n)}(\beta)$, где $\alpha \neq 0$ и B — ступенчатая матрица.*

Доказательство. Метод Гаусса приведения системы линейных линейных уравнений к ступенчатому виду, описанный в конце параграфа 8 без изменения применим к матрице A. Записывая каждое линейное преобразование в виде умножения на соответствующую матрицу Y_i, получаем $Y_1 \cdots Y_k A = B$, где B — ступенчатая матрица. Легко проверить, что

$$\left(X_{i,j}^{(n)}\right)^{-1} = X_{i,j}^{(n)}, \quad X_i^{(n)}(\alpha)^{-1} = X_i^{(n)}(\alpha^{-1}), \quad X_{i,j}^{(n)}(\beta)^{-1} = X_{i,j}^{(n)}(-\beta).$$

Поэтому мы можем записать $A = X_1 \cdots X_k B$, где $X_i = (Y_{k-i+1})^{-1}$. □

9.4. **Ранг матрицы.** В этом параграфе, мы используем следующее временное определение.

Определение 9.10. *Пусть A — матрица над полем. Её строчным рангом (столбцовым) рангом называется размерность линейной оболочки строк (столбцов).*

Наша цель — доказать, что строчный и столбцовый ранги совпадают.

Лемма 9.11. *Пусть A — матрица над полем F размера $m \times n$ и $c_1, \ldots, c_n$ — её столбцы. Пусть B — матрица над F размера $k \times m$ и $c'_1, \ldots, c'_n$ — столбцы матрицы BA. Если для некоторых $\alpha_1, \ldots \alpha_n \in F$ выполнено равенство $\alpha_1 c_1 + \cdots + \alpha_n c_n = 0$, то выполнено также равенство $\alpha_1 c'_1 + \cdots + \alpha_n c'_n = 0$.*

Доказательство. Пусть $C = BA$, $A = (a_{i,j})_{i=1,\ldots,m;j=1,\ldots,n}$, $B = (b_{r,i})_{r=1,\ldots,k;i=1,\ldots,m}$ и $C = (c_{r,j})_{r=1,\ldots,k;j=1,\ldots,n}$. Запишем линейную зависимость столбцов следующим образом:

$$\sum_{j=1}^{n} \alpha_j a_{i,j} = 0 \quad \text{для любого } i = 1, \ldots, m.$$

Отсюда получаем

$$\sum_{j=1}^{n} \alpha_j c_{r,j} = \sum_{j=1}^{n} \alpha_j \sum_{i=1}^{n} b_{r,i} a_{i,j} = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_j b_{r,i} a_{i,j} = \sum_{i=1}^{n} b_{r,i} \sum_{j=1}^{n} \alpha_j a_{i,j} = 0.$$

□

Следствие 9.12. *Пусть A — матрица над полем, A' — матрица полученная из A элементарными преобразованиями строк $c_1, \ldots, c_n$ и $c'_1, \ldots, c'_n$ — столбцы матриц A и A' соответственно. Если $(c_{j_1}, \ldots, c_{j_k})$ — базис линейной оболочки, натянутой на столбцы матрицы A, то $(c'_{j_1}, \ldots, c'_{j_k})$ — базис линейной оболочки, натянутой на столбцы матрицы A'.*

Доказательство. По лемме 9.8 мы получаем, что $A' = BA$ для некоторой матрицы B. Произвольный столбец c_j матрицы A допускает представление

$$c_j = \alpha_1 c_{j_1} + \cdots + \alpha_k c_{j_k}$$

для некоторых элементов поля $\alpha_1, \ldots, \alpha_k$. По лемме 9.11 получаем

$$c'_j = \alpha_1 c'_{j_1} + \cdots + \alpha_k c'_{j_k}.$$

Теперь предположим, что

$$\alpha_1 c'_{j_1} + \cdots + \alpha_k c'_{j_k} = 0$$

для некоторых элементов поля $\alpha_1, \ldots, \alpha_k$. Так как линейные преобразования обратимы, то A также получается из A' элементарными преобразованиями строк. Поэтому $A = B'A'$ для некоторой матрицы B'. Следовательно, по лемме 9.11 получаем

$$\alpha_1 c_{j_1} + \cdots + \alpha_k c_{j_k} = 0.$$

Так как набор векторов $(c_{j_1}, \ldots, c_{j_k})$ линейно независим (как базис), то $\alpha_1 = \cdots = \alpha_k = 0$. □

Лемма 9.13. *Пусть B — ступенчатая матрица над полем. Тогда её строчный и столбцовый ранги совпадают.*

Доказательство. Запишем матрицу B в виде (10). По лемме 9.1 мы получаем, что строчный ранг матрицы B равен k. Обозначим через $c_1, \ldots, c_n$ столбцы матрицы B. Рассуждая аналогично лемме 9.1, мы можем доказать, что столбцы $c_{j_1}, c_{j_2}, \ldots, c_{j_k}$ линейно независимы. Этот факт можно непосредственно извлечь из леммы 9.1, если рассмотреть матрицу, получающуюся из матрицы $(c_{j_1} c_{j_2}, \cdots c_{j_k})$ операцией псевдотранспонирования.

Теперь покажем, что все остальные столбцы c_j выражаются через линейные комбинации столбцов $c_{j_1}, c_{j_2}, \ldots, c_{j_k}$. Для этого докажем индукцией по $l = 0, \ldots, k$, что любой столбец

$$u = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix},$$

где m — количество строк матрицы B, имеющий нули в строках с номерами больше чем l выражается через линейную комбинации столбцов $c_{j_1}, c_{j_2}, \ldots, c_{j_l}$. Для $l = 0$ это верно так как в этом случае столбец u нулевой.

Пусть теперь $l = 1, \ldots, k$ и утверждение верно для меньших значений l. Рассмотрим столбец

$$u' = u - \frac{u_l}{b_{l,j_l}} c_{j_l}.$$

Элементы этого столбца, находящиеся в строках с номером больше чем $l-1$, нулевые. Применяя предположение индукции к u', получаем

$$u' = \alpha_1 c_{j_1} + \cdots + \alpha_{l-1} c_{j_{l-1}}$$

для некоторых элементов поля $\alpha_1, \ldots, \alpha_{l-1}$. Выражая из этого уравнения u, получаем

$$u = \alpha_1 c_{j_1} + \cdots + \alpha_{l-1} c_{j_{l-1}} + \frac{u_l}{b_{l,j_l}} c_{j_l}.$$

Таким образом $c_{j_1}, c_{j_2}, \ldots, c_{j_k}$ — базис линейной оболочки столбцов матрицы A. Отсюда получаем, что k — столбцовый ранг матрицы B. □

В дальнейшем мы будем называть совпадающие строчный и столбцовый ранги матрицы A просто *рангом* и обозначать это число через $\operatorname{rank} A$.

Теорема 9.14. *Строчный и столбцовый ранги матрицы совпадают.*

Доказательство. Мы можем элементарными преобразованиями строк привести произвольную матрицу A над полем к ступенчатой матрице B. По следствию 9.12 мы получаем, что столбцовые ранги матриц A и B совпадают. Строчные ранги матриц A и B совпадают по лемме 6.14. Остаётся заметить, что столбцовый и строчный ранги матриц B совпадают по лемме 9.13. □

9.5. **Матрица перехода между базисами.** Пусть V — конечномерное векторное пространство над полем F и пусть $u = (u_1, \ldots, u_n)$ и $v = (v_1, \ldots, v_n)$ — два его базиса. Мы можем выразить элементы второго (нового) базиса, через элементы первого (старого) базиса.

$$v_j = \sum_{i=1}^{n} a_{i,j} u_i, \tag{12}$$

где $a_{i,j} \in F$. Мы можем записать элементы поля $a_{i,j}$ в виде матрицы

$$T_{u,v} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}, \tag{13}$$

которую мы называем матрицей перехода от базиса u к базису v.

Мы немного расширим понятие матрицы и согласимся писать в неё не только скаляры но и векторы. Кроме этого, мы разрешим умножать векторы на скаляры справа по очевидному правилу $w\alpha = \alpha w$. Теперь мы можем записать уравнения (12) в виде

$$v = uT_{u,v}. \tag{14}$$

В этом уравнении базисы u и v рассматриваются как матрицы размера $1 \times n$. Заметим, что матрица $T_{u,v}$ определяется из этого уравнения однозначно.

Предположим, что $w = (w_1, \ldots, w_n)$ — третий базис пространства V. Мы получаем

$$w = vT_{v,w} = uT_{u,v}T_{v,w}.$$

Отсюда получаем

$$T_{u,w} = T_{u,v}T_{v,w}. \tag{15}$$

Заметим, что $T_{u,u} = E$, где E — единичная матрица размера $n \times n$. Из формулы (15) получаем $E = T_{u,u} = T_{u,v}T_{v,u}$. Отсюда следует, что матрица $T_{u,v}$ всегда обратима и

$$T_{v,u} = T_{u,v}^{-1}.$$

Действуя аналогичным способом, мы можем ответь на вопрос, как связнанны координаты векторов в двух различных базисах, если известна матрица перехода. Пусть $\lambda = b_1v_1 + \cdots + b_nv_n$ и $\lambda = a_1u_1 + \cdots + a_nu_n$, где $b_i, a_i \in F$. Получаем

$$(\lambda) = v \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = u\, T_{u,v} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Здесь слева стоит матрица размере 1×1. Сравнивая это уравнение с уравнением

$$(\lambda) = u \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

и используя однозначность разложение по базису, получаем

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = T_{u,v} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}. \tag{16}$$

9.6. **Координатный подход.** Возможно, простые доказательства предыдущего раздела могут вызвать недоверие у некоторых читателей: мы умножали матрицы, составленные из векторов, и умножали сами векторы на скаляры с обратной стороны.

Сначала приведём прямое доказательство формулы (16) при помощи формулы (12). В вышеприведённых обозначениях получаем

$$\lambda = \sum_{j=1}^{n} b_jv_j = \sum_{j=1}^{n} b_j \sum_{i=1}^{n} a_{i,j}u_i = \sum_{i,j=1}^{n} b_ja_{i,j}u_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{i,j}b_j \right) u_i.$$

Сравнивая эту формулу с разложеним $\lambda = \sum\limits_{i=1}^{n} a_iu_i$, получаем формулу

$$a_i = \sum_{j=1}^{n} a_{i,j}b_j$$

для любого $i = 1, \ldots, n$. Эти формулы эквивалентно могут быть записанны в виде (16).

Теперь мы можем доказать формулу (15). Разложим вектор λ по третьему базису:

$$\lambda = c_1w_1 + \cdots + c_nw_n.$$

Аналогично формуле (16) мы получаем

$$\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = T_{v,w} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

Подставляя эту формулу в (16), получаем

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = T_{u,v}T_{v,w} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

Ещё аналогично формуле (16) мы получаем

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = T_{u,w} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

Отсюда

$$T_{u,v}T_{v,w} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = T_{u,w} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

Для получения формулы (15), остаётся применить следующую лемму.

Лемма 9.15. *Пусть A и B— матрицы размера $m \times n$ такие, что $AX = BX$ для любой матрицы X размера $n \times 1$. Тогда $A = B$.*

Доказательство. Достаточно рассмотреть случай, когда

$$X = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

где 1 стоит в строке i. Тогда получаем, что AX равно i-у столбцу матрицы A и BX равно i-у столбцу матрицы B. По условию эти столбцы совпадают для любого $i = 1, \ldots, n$. Отсюда $A = B$. □

10. Линейные отображения

10.1. **Определение и операции.** Пусть U и V — векторные пространства над полем F. Отображение $\varphi : U \to V$ называется линейным, если выполняются следующие два свойства:

(Л1) $\varphi(u + u') = \varphi(u) + \varphi(u')$ для любых $u, u' \in U$;

(Л2) $\varphi(\alpha u) = \alpha\varphi(u)$ для любого $u \in U$ для $\alpha \in F$.

Из этих определений следует, что отображение, которое отправляет любой элемент из U в нулевой вектор из V является линейным отображением. Оно называется *нулевым* и обозначается также через 0.

Пусть $\varphi : U \to V$ — линейное отображение и $c \in \mathbb{F}$ — скаляр. Мы рассмотрим *скалярное кратное* $c\varphi : U \to V$, определённое формулой

$$c\varphi(u) = (c\varphi)(u)$$

Легко проверить, что отображение $c\varphi$ линейное:

$$(c\varphi)(u + u') = c\varphi(u + u') = c(\varphi(u) + \varphi(u')) = c\varphi(u) + c\varphi(u') = (c\varphi)(u) + (c\varphi)(u');$$

$$(c\varphi)(\alpha u) = c\varphi(\alpha u) = c\alpha\varphi(u) = \alpha c\varphi(u) = \alpha(c\varphi)(u).$$

Заметим, что в последнем вычислении мы применили коммутативность поля $\mathbb{F}$.

Для двух линейных отображений $\varphi, \psi : U \to V$ определим их *сумму* $\varphi + \psi : U \to V$ по формуле

$$(\varphi + \psi)(u) = \varphi(u) + \psi(u).$$

Легко проверить, что отображение $\varphi+\psi$ линейное:

$$(\varphi+\psi)(u+u')=\varphi(u+u')+\psi(u+u')=\varphi(u)+\varphi(u')+\psi(u)+\psi(u')=$$
$$=(\varphi(u)+\psi(u))+(\varphi(u')+\psi(u'))=(\varphi+\psi)(u)+(\varphi+\psi)(u');$$
$$(\varphi+\psi)(\alpha u)=\varphi(\alpha u)+\psi(\alpha u)=\alpha\varphi(u)+\alpha\psi(u)=\alpha(\varphi(u)+\psi(u))=\alpha(\varphi+\psi)(u).$$

Наконец, для двух линейных отображений $\varphi: U\to V$ и $\psi: V\to W$ определим их *композицию* $\psi\varphi: U\to W$ как композицию отображений множеств (см. §1). Легко проверить, что отображение $\psi\varphi$ линейное:

$$(\psi\varphi)(u+u')=\psi(\varphi(u+u'))=\psi(\varphi(u)+\varphi(u'))=\psi(\varphi(u))+\psi(\varphi(u'))=\psi\varphi(u)+\psi\varphi(u');$$
$$(\psi\varphi)(\alpha u)=\psi(\alpha\varphi(u))=\alpha\psi\varphi(u).$$

Заметим следующие свойства введённых операций:

- $0\varphi=0$, где в левой части стоит умножение оператора на скаляр;
- $c0=0$, где в левой части стоит умножение оператора на скаляр;
- $\varphi+0=0+\varphi=\varphi$;
- $\varphi0=0$, где в левой части стоит композиция операторов;
- $0\psi=0$, где в левой части стоит композиция операторов;
- $\varphi+\psi=\psi+\varphi$;
- $(\varphi+\psi)+\rho=\psi+(\varphi+\rho)$;
- $(c+c')\varphi=c\varphi+c'\varphi$;
- $c(\varphi+\psi)=c\varphi+c\psi$;
- $(\varphi\psi)\rho=\psi(\varphi\rho)$;
- $(c\varphi)\psi=\varphi(c\psi)=c(\varphi\psi)$;
- $\varphi(\psi+\rho)=\varphi\psi+\varphi\rho$;
- $(\varphi+\psi)\rho=\varphi\rho+\psi\rho$.

В формулах выше мы считаем, что φ,ψ и ρ — линейные операторы и c и c' — скаляры. Мы так же используем стандартное соглашение о том, что операция умножения на скаляр и композиции выполняются до операции сложения.

Мы, в частности, получаем, что $\varphi+(-1)\varphi=0$. Это даёт основание определить $-\varphi=(-1)\varphi$. Мы, как обычно, используем сокращённое обозначение $\varphi-\psi=\varphi+(-\psi)$.

Линейное отображение из векторного пространства в само себя называется *линейным оператором*.

10.2. **Матрица линейного отображения.** Пусть $\varphi: U\to V$ — линейное отображение, $u=(u_1,\dots,u_n)$ — базис пространства U и $v=(v_1,\dots,v_m)$ — базис пространства V. Мы можем записать

$$\varphi(u_j)=\sum_{i=1}^{m}\varphi_{i,j}v_i.$$

Тоже самое в матричной форме можно записать следующим образом:

$$\varphi(u)=vM_{v,u}(\varphi), \tag{17}$$

где

$$\varphi(u)=(\varphi(u_1),\dots,\varphi(u_n))\quad\text{и}\quad M_{v,u}(\varphi)=\begin{pmatrix}\varphi_{1,1}&\cdots&\varphi_{1,n}\\ \vdots&\ddots&\vdots\\ \varphi_{m,1}&\cdots&\varphi_{m,n}\end{pmatrix}.$$

При помощи матрицы $M_{v,u}(\varphi)$ можно ответить на вопрос, как преобразуются координаты вектора при действии линейного оператора φ. Действительно, рассмотрим произвольный вектор

$$\lambda = a_1 u_1 + \cdots + a_n u_n$$

пространства U. Применяя φ к этому равенству, при помощи (17) получаем

$$\big(\varphi(\lambda)\big) = \big(a_1\varphi(u_1) + \cdots + a_n\varphi(u_n)\big) = \varphi(u)\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} = vM_{v,u}(\varphi)\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}.$$

Это означает, что координаты $(b_1, \ldots, b_n)$ вектора $\varphi(\lambda)$ в базисе v вычисляются по правилу

$$\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = M_{v,u}(\varphi)\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}. \tag{18}$$

10.3. **Замена базисов.** Мы можем теперь вычислить как изменится матрица линейного отображения при замане базисов. Пусть $u' = (u'_1, \ldots, u'_n)$ и $v' = (v'_1, \ldots, v'_m)$ — ещё два базиса пространств U и V соответственно. Согласно уравнению (14) получаем

$$u' = uT_{u,u'}, \quad v' = vT_{v,v'}.$$

Применяя φ к первому уравнению и используя линейность φ, получаем

$$\varphi(u') = \varphi(u)T_{u,u'}.$$

Аналогично (17) запишем

$$\varphi(u') = v'M_{v',u'}(\varphi),$$

Подставляя в это уравнение выражения для $\varphi(u')$ и v', полученные выше, получаем

$$\varphi(u)T_{u,u'} = vT_{v,v'}M_{v',u'}(\varphi).$$

Умножая это уравнение на $T^{-1}_{u,u'}$ справа получаем

$$\varphi(u) = vT_{v,v'}M_{v',u'}(\varphi)T^{-1}_{u,u'}.$$

Сравнивая с (17), получаем $M_{v,u}(\varphi) = T_{v,v'}M_{v',u'}(\varphi)T^{-1}_{u,u'}$. Отсюда

$$M_{v',u'}(\varphi) = T^{-1}_{v,v'}M_{v,u}(\varphi)T_{u,u'}. \tag{19}$$

10.4. **Матрицы скалярного кратного, суммы и композиции.** Пусть $\varphi : U \to V$ — линейное отображение и u и v — базисы конечномерных пространств U и V. Умножая обе части формулы (17) на скаляр c, получаем: в левой части

$$c\varphi(u) = c\varphi(u)$$

и в правой части

$$cvM_{v,u}(\varphi) = v(cM_{v,u}(\varphi)).$$

Равенство этих двух выражений показывает, что

$$M_{v,u}(c\varphi) = cM_{v,u}(\varphi). \tag{20}$$

Пусть $\psi : U \to V$ — ещё один линейный оператор. Наряду с формулой (17) верна также формула $\psi(u) = vM_{v,u}(\psi)$. Складывая эти формулы получаем: в левой части

$$\varphi(u) + \psi(u) = (\varphi + \psi)(u)$$

и в правой части

$$vM_{v,u}(\varphi) + vM_{v,u}(\psi) = v(M_{v,u}(\varphi) + M_{v,u}(\psi)).$$

Равенство этих двух выражений показывает, что

$$M_{v,u}(\varphi+\psi)=M_{v,u}(\varphi)+M_{v,u}(\psi). \tag{21}$$

Наконец, посмотрим как устроена матрица композиции линейных отображений. Пусть $\rho : V \to W$ — линейное отображения, и w — базис конечномерного линейного пространства W. Вместе с формулой (17) верны аналогичные формулы

$$\rho(v)=wM_{w,v}(\rho),\qquad \rho\varphi(u)=wM_{w,u}(\rho\varphi). \tag{22}$$

Применяя к обеим частям формулы (17) отображение ρ, получаем

$$\rho\varphi(u)=\rho(v)M_{v,u}(\varphi)=wM_{w,v}(\rho)M_{v,u}(\varphi).$$

Сравнивая эту формулу с первой формулой (22), получаем

$$M_{w,u}(\rho\varphi)=M_{w,v}(\rho)M_{v,u}(\varphi). \tag{23}$$

Для линейного оператора φ на V и базиса $v=(v_1,\dots,v_n)$ этого пространства положим $M_v(\varphi)=M_{v,v}(\varphi)$.

10.5. **Ядро и образ.** Пусть $\varphi : U \to V$ — линейный оператор. Рассмотрим следующие множество

$$\ker\varphi=\{u\in U \mid \varphi(u)=0\}.$$

Кроме этого множества, можно рассмотреть обычный теоретико множественный образ $\operatorname{im}\varphi$ отображения φ.

Лемма 10.1. *Пусть $\varphi : U \to V$ — линейный оператор. Множества $\ker\varphi$ и $\operatorname{im}\varphi$ являются линейными подпространствами в U и V соответственно.*

Доказательство. Так как $\varphi(0)=0$, то $0\in\ker\varphi$ и $0\in\operatorname{im}\varphi$. Пусть теперь $u,u'\in\ker\varphi$. Получаем

$$\varphi(u+u')=\varphi(u)+\varphi(u')=0+0=0.$$

Следовательно, $u+u'\in\ker\varphi$. Для любого $\alpha\in\mathbb{F}$, получаем

$$\varphi(\alpha u)=\alpha\varphi(u)=\alpha 0=0.$$

Следовательно, $\alpha u\in\ker\varphi$.

Пусть теперь $v,v'\in\operatorname{im}\varphi$. Тогда $v=\varphi(u)$ и $v'=\varphi(u')$. Тогда получаем

$$\varphi(u+u')=\varphi(u)+\varphi(u')=v+v'.$$

Этим доказано, что $v+v'\in\operatorname{im}\varphi$. Для любого $\alpha\in\mathbb{F}$, получаем

$$\varphi(\alpha u)=\alpha\varphi(u)=\alpha v.$$

Отсюда $\alpha v\in\operatorname{im}\varphi$. □

Лемма 10.2. *Пусть $\varphi : U \to V$ — линейный оператор, где U — конечномерное пространство. Тогда $\ker\varphi$ и $\operatorname{im}\varphi$ конечномерны и при этом*

$$\dim U=\dim\ker\varphi+\dim\operatorname{im}\varphi.$$

Доказательство. Пространство $\ker\varphi$ конечномерно, так как U конечномерно. Пусть $(w_1,\dots,w_k)$ — базис пространства $\ker\varphi$ и $(u_1,\dots,u_m)$ — базис пространства U. Мы утверждаем, что

$$\operatorname{im}\varphi=\langle\varphi(u_1),\dots,\varphi(u_m)\rangle.$$

Действительно линейная оболочка справа принадлежит $\operatorname{im}\varphi$, так как ей принадлежат все векторы $\varphi(u_i)$. Пусть $v\in\operatorname{im}\varphi$. Тогда $v=\varphi(u)$ для некоторого $u\in U$. Разложим этот вектор по базису

$$u=\alpha_1u_1+\dots+\alpha_mu_m.$$

Применяя φ, получаем

$$v = \varphi(u) = \alpha_1\varphi(u_1) + \cdots + \alpha_m\varphi(u_m) \in \langle\varphi(u_1), \ldots, \varphi(u_m)\rangle.$$

Выберем такие индексы $i_1, \ldots, i_n \in \{1, \ldots, m\}$, что $(\varphi(u_{i_1}), \ldots, \varphi(u_{i_n}))$ — базис $\operatorname{im}\varphi$. Заметим, что из условия линейной независимости элементов базиса следует, что числа $i_1, \ldots, i_n$ попарно различны.

Мы утверждаем, что $(w_1, \ldots, w_k, u_{i_1}, \ldots, u_{i_n})$ — базис пространства U. Действительно пусть $u \in U$. Рассмотрим разложение по базису пространства $\operatorname{im}\varphi$:

$$\varphi(u) = \alpha_1\varphi(u_{i_1}) + \cdots + \alpha_n\varphi(u_{i_n}) = \varphi(\alpha_1 u_{i_1} + \cdots + \alpha_n u_{i_n}).$$

Перенося в одну сторону получаем,

$$\varphi(u - \alpha_1 u_{i_1} - \cdots - \alpha_n u_{i_n}) = 0.$$

Следовательно, $u - \alpha_1 u_{i_1} - \cdots - \alpha_n u_{i_n} \in \ker\varphi$. Этот вектор можно разложить по базису пространства $\ker\varphi$:

$$u - \alpha_1 u_{i_1} - \cdots - \alpha_n u_{i_n} = \beta_1 w_1 + \cdots + \beta_k w_k.$$

Перенося все векторы кроме u вправо получаем

$$u = \beta_1 w_1 + \cdots + \beta_k w_k + \alpha_1 u_{i_1} + \cdots + \alpha_n u_{i_n}.$$

Теперь предположим, что

$$u = \beta'_1 w_1 + \cdots + \beta'_k w_k + \alpha'_1 u_{i_1} + \cdots + \alpha'_n u_{i_n}.$$

Вычитая из верхнего равенства нижнее получаем

$$0 = (\beta_1 - \beta'_1)w_1 + \cdots + (\beta_k - \beta'_k)w_k + (\alpha_1 - \alpha'_1)u_{i_1} + \cdots + (\alpha_n - \alpha'_n)u_{i_n}.$$

Применяя к этому равенству φ, получаем

$$0 = \varphi(0) = (\beta_1 - \beta'_1)\varphi(w_1) + \cdots + (\beta_k - \beta'_k)\varphi(w_k) + (\alpha_1 - \alpha'_1)\varphi(u_{i_1}) + \cdots + (\alpha_n - \alpha'_n)\varphi(u_{i_n}) = \\ = (\alpha_1 - \alpha'_1)\varphi(u_{i_1}) + \cdots + (\alpha_n - \alpha'_n)\varphi(u_{i_n}).$$

Так как набор $(\varphi(u_{i_1}), \ldots, \varphi(u_{i_n}))$ линейно независимый, то получаем $\alpha_i = \alpha'_i$ для всех $i = 1, \ldots, n$. Следовательно, мы можем записать

$$0 = (\beta_1 - \beta'_1)w_1 + \cdots + (\beta_k - \beta'_k)w_k.$$

Так как набор $(w_1, \ldots, w_k)$ линейно независимый, то $\beta_i = \beta'_i$ для всех $i = 1, \ldots, k$.

Теперь требуемое утверждение, следует из того, что $\dim U = k + n$, $\dim\ker\varphi = k$ и $\dim\operatorname{im}\varphi = n$. □

11. Определители

11.1. **Симметрическая группа.** Пусть n — натуральное число. Назовём *симметрической группой* S_n множество всех биекций множества $\{1, \ldots, n\}$ относительно композиции в качестве групповой операции. Элементы группы S_n называются *подстановками*.

Мы будем использовать запись

$$\sigma = \begin{pmatrix} i_1 & \cdots & i_n \\ j_1 & \cdots & j_n \end{pmatrix},$$

чтобы сказать, что подстановка σ — это отображение из действующее следующим образом: $\sigma(i_k) = j_k$ для любого $k = 1, \ldots, n$. В этой записи $i_1, \ldots, i_n$ и $j_1, \ldots, j_n$ — последовательности, полученные из последовательности $1, \ldots, n$ некоторыми перестановками её элементов. В связи с этим мы будем называть в дальнейшем такие последовательности *перестановками* (множества $\{1, \ldots, n\}$).

Заметим, что если переставим столбцы этой матрицы, что мы получим ту же самую подстановку. Например,

$$\begin{pmatrix} 3 & 5 & 1 & 2 & 4 \\ 1 & 3 & 2 & 5 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 4 & 3 & 2 & 1 \\ 3 & 4 & 1 & 5 & 2 \end{pmatrix}.$$

Приведём пример умножения перестановок в такой записи:

$$\begin{pmatrix} 2 & 3 & 1 & 5 & 4 \\ 3 & 1 & 5 & 4 & 2 \end{pmatrix} \begin{pmatrix} 1 & 5 & 3 & 4 & 2 \\ 2 & 1 & 4 & 5 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 2 & 4 & 5 \end{pmatrix}.$$

Поясним, как получается такая формула. Возьмём 1. Первая (правая) подстановка левой части переводит её в 2. Далее вторая (левая) подстановка левой части переводит эту 2 в 3. Таким образом получаем первый столбец матрицы правой части. Повторяя эту процедуру для элементов $2, 3, 4, 5$, получаем остальные столбцы.

Подстановка, которая каждый элемент оставляет на месте, называется *тождественной подстановкой*. Вот её запись в виде матрицы

$$e = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix} \tag{24}$$

Каждая подстановка σ имеет обратную подстановку σ^{-1} в том смысле, что $\sigma\sigma^{-1} = \sigma^{-1}\sigma = e$. Обратная подстановка задаётся следующей формулой:

$$\begin{pmatrix} i_1 & \cdots & i_n \\ j_1 & \cdots & j_n \end{pmatrix}^{-1} = \begin{pmatrix} j_1 & \cdots & j_n \\ i_1 & \cdots & i_n \end{pmatrix}.$$

Важным примером подстановки являются циклы. Мы будем записывать цикл в виде $\sigma = (i_1, i_2, \ldots, i_k)$, где $i_1, i_2, \ldots, i_n$ — попарно различные числа из множества $\{1, \ldots, n\}$, чтобы сказать, что σ действует следующим образом:

$$\sigma(i_1) = i_2, \sigma(i_2) = i_3, \ldots, \sigma(i_{k-1}) = i_k, \sigma(i_k) = i_1$$

и $\sigma(i) = i$ для любого $i \in \{1, \ldots, n\} \setminus \{i_1, \ldots, i_k\}$. Два цикла $(i_1, \ldots, i_k)$ и $(j_1, \ldots, j_l)$ называются *независимыми*, если $\{i_1, \ldots, i_k\} \cap \{j_1, \ldots, j_l\} = \varnothing$. Очевидно, что произведение независимых циклов перестановочно:

$$(i_1, \ldots, i_k)(j_1, \ldots, j_l) = (j_1, \ldots, j_l)(i_1, \ldots, i_k).$$

Очевидно, что каждая подстановка может быть единственных образом записана в виде произведения независимых циклов. Например,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 6 & 8 & 5 & 7 & 1 & 4 & 3 \end{pmatrix} = (1, 2, 6)(3, 8)(4, 5, 7). \tag{25}$$

В дальнейшем мы будем называть цикл вида (i_1, i_2) *транспозицией*.

Наша следующая цель — ввести понятие знака подстановки, которое нам понадобиться для понятия определителя матрицы. Пусть $i_1, \ldots, i_n$ — перестановка. Пару чисел (a, b) из множества $\{1, \ldots, n\}$ назовём *инверсией* перестановки $i_1, \ldots, i_n$, если $a > b$, $a = i_k$, $b = i_l$ и $k < l$. Множество всех инверсий перестановки $i_1, \ldots, i_n$ обозначим через $\mathrm{D}(i_1, \ldots, i_n)$. Например,

$$\mathrm{D}(2, 4, 1, 3) = \{(2, 1), (4, 1), (4, 3)\}.$$

Лемма 11.1. *Пусть $j_1, \ldots, j_n$ — перестановка, полученная из перестановки $i_1, \ldots, i_n$ перестановкой k-о и $k+1$-о индексов, где $1 \leqslant k < n$. Тогда*

$$\mathrm{D}(j_1, \ldots, j_n) = \mathrm{D}(i_1, \ldots, i_n) \Delta \{(a, b)\},$$

где $a > b$ и $\{i_k, i_{k+1}\} = \{a, b\}$.

Доказательство. Сначала заметим, что эту формулу достаточно доказать в одну сторону:

$$\mathrm{D}(j_1,\dots,j_n) \subset \mathrm{D}(i_1,\dots,i_n)\Delta\{(a,b)\}. \tag{26}$$

Действительно, так как перестановка $i_1,\dots,i_n$ тоже получена из перестановки $j_1,\dots,j_n$ перестановкой k-о и $k+1$-о индексов и $\{j_k,j_{k+1}\}=\{i_k,i_{k+1}\}=\{a,b\}$, то мы получаем

$$\mathrm{D}(i_1,\dots,i_n) \subset \mathrm{D}(j_1,\dots,j_n)\Delta\{(a,b)\}. \tag{27}$$

Теперь заметим следующий факт о симметрической разности: если $A\subset B\subset X$ и $x\in X$, то $A\Delta\{x\}\subset B\Delta\{x\}$ за исключением случая, когда $x\in B\setminus A$. Мы собираемся применить этот факт к включению (27). Предположим, что

$$(a,b)\in(\mathrm{D}(j_1,\dots,j_n)\Delta\{(a,b)\})\setminus\mathrm{D}(i_1,\dots,i_n). \tag{28}$$

Так как $(a,b)\in\mathrm{D}(i_1,\dots,i_n)\cup\mathrm{D}(j_1,\dots,j_n)$, то $(a,b)\in\mathrm{D}(j_1,\dots,j_n)$ и $(a,b)\notin\mathrm{D}(j_1,\dots,j_n)\setminus\{(a,b)\}=\mathrm{D}(j_1,\dots,j_n)\Delta\{(a,b)\}$. Получаем противоречие с формулой (28).

Таким образом, мы доказали, что формула (28) не выполняется. Следовательно, используя ассоциативность симметрической разности, из формулы (27) получаем

$$\mathrm{D}(i_1,\dots,i_n)\Delta\{(a,b)\}\subset(\mathrm{D}(j_1,\dots,j_n)\Delta\{(a,b)\})\,\Delta\{(a,b)\}=$$

$$=\mathrm{D}(j_1,\dots,j_n)\Delta\,(\{(a,b)\}\Delta\{(a,b)\})=\mathrm{D}(j_1,\dots,j_n)\Delta\varnothing=\mathrm{D}(j_1,\dots,j_n).$$

Этим доказана формула обратная к формуле (26).

Теперь мы займёмся доказательством формулы (26). Пусть $(c,d)\in\mathrm{D}(j_1,\dots,j_n)$. Мы запишем $c=j_p>d=j_q$ для соответствующих $p<q$. Напомним, как вычислять элемент элементы перестановки $j_1,\dots,j_n$ через элементы перестановки $i_1,\dots,i_n$:

$$i_m=\begin{cases} j_{k+1}, & \text{если } m=k;\\ j_k, & \text{если } m=k+1;\\ j_m, & \text{если } m\neq k \text{ и } m\neq k+1.\end{cases}$$

Случай 1: $\{p,q\}\cap\{k,k+1\}=\varnothing$. В этом случае $(c,d)\neq(a,b)$ и

$$(c,d)=(j_p,j_q)=(i_p,i_q)\in D(i_1,\dots,i_n).$$

Следовательно, $(c,d)\in D(i_1,\dots,i_n)\Delta\{(a,b)\}$.

Случай 2: $p<q=k$. В этом случае $(c,d)\neq(a,b)$ и $(c,d)=(j_p,j_k)=(i_p,i_{k+1})$. Так как $c>d$ и $p<k+1$, то $(c,d)\in\mathrm{D}(i_1,\dots,i_n)$. Следовательно, $(c,d)\in\mathrm{D}(i_1,\dots,i_n)\Delta\{(a,b)\}$.

Случай 3: $p<k$ *и* $q=k+1$. В этом случае $(c,d)\neq(a,b)$ и $(c,d)=(j_p,j_{k+1})=(i_p,i_k)$. Так $p<k$, то $(c,d)\in\mathrm{D}(i_1,\dots,i_n)$. Следовательно, $(c,d)\in\mathrm{D}(i_1,\dots,i_n)\Delta\{(a,b)\}$.

Случай 4: $k+1=p<q$. В этом случае $(c,d)\neq(a,b)$ и $(c,d)=(j_{k+1},j_q)=(i_k,i_q)$. Так $k<q$, то $(c,d)\in\mathrm{D}(i_1,\dots,i_n)$. Следовательно, $(c,d)\in\mathrm{D}(i_1,\dots,i_n)\Delta\{(a,b)\}$.

Случай 5: $k=p$ *и* $k+1<q$. В этом случае $(c,d)\neq(a,b)$ и $(c,d)=(j_k,j_q)=(i_{k+1},i_q)$. Так как $k+1<q$, то $(c,d)\in\mathrm{D}(i_1,\dots,i_n)$. Следовательно, $(c,d)\in\mathrm{D}(i_1,\dots,i_n)\Delta\{(a,b)\}$.

Случай 6: $k=p$ *и* $k{+}1=q$. В этом случае $(c,d)=(a,b)$ и $(c,d)=(j_k,j_k{+}1)=(i_k,i_{k+1})$. Следовательно, $(c,d)\notin\mathrm{D}(i_1,\dots,i_n)$ и мы опять получаем $(c,d)\in\mathrm{D}(i_1,\dots,i_n)\Delta\{(a,b)\}$. □

Следствие 11.2. *Пусть $j_1,\dots,j_n$ — перестановка, полученная из перестановки $i_1,\dots,i_n$ перестановкой k-о и $k+1$-о индексов, где $1\leqslant k<n$. Тогда*

$$|\,\mathrm{D}(j_1,\dots,j_n)|=|\,\mathrm{D}(i_1,\dots,i_n)|\pm1.$$

Определение 11.3. *Знак подстановки определятся следующим образом*

$$\operatorname{sgn}\begin{pmatrix} i_1 & \cdots & i_n\\ j_1 & \cdots & j_n\end{pmatrix}=(-1)^{|\,\mathrm{D}(i_1,\dots,i_n)|+|\,\mathrm{D}(j_1,\dots,j_n)|}.$$

Важно заметить, что следствие 11.2 гарантирует, что знак подстановки определён формулой выше корректно, то есть, не зависит от формы записи подстановки в виде матрицы. Действительно, пусть

$$\begin{pmatrix} i_1 & \cdots & i_n \\ j_1 & \cdots & j_n \end{pmatrix} = \begin{pmatrix} i'_1 & \cdots & i'_n \\ j'_1 & \cdots & j'_n \end{pmatrix},$$

где матрица в правой части получается из матрицы в левой части перестановкой двух соседних столбцов. Тогда в силу следствия 11.2 получаем

$$(-1)^{|\,\mathrm{D}(i'_1,\ldots,i'_n)|+|\,\mathrm{D}(j'_1,\ldots,j'_n)|} = (-1)^{|\,\mathrm{D}(i_1,\ldots,i_n)|\pm1+|\,\mathrm{D}(j_1,\ldots,j_n)|\pm1} = (-1)^{|\,\mathrm{D}(i_1,\ldots,i_n)|+|\,\mathrm{D}(j_1,\ldots,j_n)|}.$$

Из этих равенств и того факта, что любая перестановка столбцов матрицы получается последовательностью перестановок соседних столбцов, получается утверждение о корректности определения знака подстановки.

Приведём примеры явного применения этой формулы. Во первых очевидно, что знак тождественной подстановки равен 1, так как в записи (24) ни верхняя ни нижняя строки матрицы на содержат инверсий. Вычислим знак транспозиции. Пусть $1 \leqslant a < b \leqslant n$. Запишем цикл (a, b) в виде матрицы

$$(a,b) = \begin{pmatrix} 1 & a-1 & a & a+1 & \cdots & b-1 & b & b+1 & \cdots & n \\ 1 & a-1 & b & a+1 & \cdots & b-1 & a & b+1 & \cdots & n \end{pmatrix}.$$

Первая строчка не содержит инверсий. Все инверсии нижней строчки следующие:

$$(b,a),(b,a+1),\cdots,(b,b-1),(a+1,a),(a+1,a),\cdots,(b-1,a).$$

Всего этих инверсий $(b-a)+(b-1-a) = 2(b-a)+1$. Так как это число всегда нечётное, то

$$\operatorname{sgn}(a,b) = -1.$$

Лемма 11.4. $\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}\sigma \operatorname{sgn}\tau$.

Доказательство. Переставляя столбцы матриц, если это необходимо мы можем записать подстановки σ и τ в следующем виде

$$\sigma = \begin{pmatrix} j_1 & \cdots & j_n \\ k_1 & \cdots & k_n \end{pmatrix}, \quad \tau = \begin{pmatrix} i_1 & \cdots & i_n \\ j_1 & \cdots & j_n \end{pmatrix}.$$

Тогда мы получаем

$$\sigma\tau = \begin{pmatrix} j_1 & \cdots & j_n \\ k_1 & \cdots & k_n \end{pmatrix}\begin{pmatrix} i_1 & \cdots & i_n \\ j_1 & \cdots & j_n \end{pmatrix} = \begin{pmatrix} i_1 & \cdots & i_n \\ k_1 & \cdots & k_n \end{pmatrix}.$$

Отсюда получаем

$$\begin{aligned}\operatorname{sgn}\sigma \operatorname{sgn}\tau &= (-1)^{|\,\mathrm{D}(j_1,\ldots,j_n)|+|\,\mathrm{D}(k_1,\ldots,k_n)|}(-1)^{|\,\mathrm{D}(i_1,\ldots,i_n)|+|\,\mathrm{D}(j_1,\ldots,j_n)|} = \\ &= (-1)^{|\,\mathrm{D}(i_1,\ldots,i_n)|+2|\,\mathrm{D}(j_1,\ldots,j_n)|+|\,\mathrm{D}(k_1,\ldots,k_n)|} = (-1)^{|\,\mathrm{D}(i_1,\ldots,i_n)|+|\,\mathrm{D}(k_1,\ldots,k_n)|} = \operatorname{sgn}(\sigma\tau).\end{aligned}$$

□

Следствие 11.5. $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}\sigma$.

Доказательство. По лемме 11.4 получаем

$$1 = \operatorname{sgn} e = \operatorname{sgn}(\sigma\sigma^{-1}) = (\operatorname{sgn}\sigma)(\operatorname{sgn}\sigma^{-1}).$$

Отсюда

$$\operatorname{sgn}(\sigma^{-1}) = \frac{1}{\operatorname{sgn}\sigma} = \operatorname{sgn}\sigma.$$

Здесь мы использовали тот факт, что $1/\varepsilon = \varepsilon$ для $\varepsilon = \pm1$. □

Теперь мы можем вычислить знак любого цикла $(i_1, \dots, i_k)$. Для этого заметим следующее представление

$$(i_1, \dots, i_k) = (i_1, i_2)(i_2, i_3) \cdots (i_{k-1}, i_k).$$

Отсюда по лемме 11.4 получаем

$$\operatorname{sgn}(i_1, \dots, i_k) = \operatorname{sgn}(i_1, i_2) \operatorname{sgn}(i_2, i_3) \cdots \operatorname{sgn}(i_{k-1}, i_k) = (-1)^{k-1}.$$

Это правило вместе с леммой 11.4 представляет более практичный способ вычисления знака подстановки. Например, вычислим знак подстановки (25):

$$\operatorname{sgn} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 6 & 8 & 5 & 7 & 1 & 4 & 3 \end{pmatrix} = \operatorname{sgn}(1,2,6) \operatorname{sgn}(3,8) \operatorname{sgn}(4,5,7) = \\ = (-1)^{3-1}(-1)^{2-1}(-1)^{3-1} = -1.$$

Введём ещё немного терминологии: подстановки знака 1 называются *чётными*, а подстановки знака -1 *нечётными*. В группе S_1 есть только одна тождественная подстановка и она, как мы уже знаем, чётная. Это значит, что в группе S_1 все подстановки чётные. Однако в группах S_n при $n > 2$ ситуация другая.

Лемма 11.6. *Пусть n — натуральное число больше 1. Тогда в группе S_n одинаковое число чётных и нечётных подстановок.*

Доказательство. Рассмотрим подстановку $\tau = (1, 2)$. Условие $n > 1$ гарантирует, что τ — корректно заданный элемент из S_n. Рассмотрим разбиение $S_n = S_n^{\text{even}} \sqcup S_n^{\text{odd}}$, где S_n^{even} и S_n^{odd} — множества чётных и нечётных подстановок соответственно. Мы определим отображение $\varphi : S_n^{\text{even}} \to S_n^{\text{odd}}$ по формуле $\varphi(\sigma) = \tau\sigma$. Лемма 11.4 гарантирует, что φ корректно заданное отображение.

Рассмотрим отображение $\psi : S_n^{\text{odd}} \to S_n^{\text{even}}$, заданное той же формулой $\psi(\sigma) = \tau\sigma$. Для любого $\sigma \in S_n^{\text{even}}$ получаем

$$\psi\varphi(\sigma) = \psi(\tau\sigma) = \tau^2\sigma = \sigma.$$

Следовательно, $\psi\varphi = \mathrm{id}$. Аналогично доказывается, что $\varphi\psi = \mathrm{id}$. Этим доказано, что φ — биекция, откуда следует утверждение леммы. □

Определение 11.7. *Пусть $A = (a_{i,j})_{i,j=1}^n$ — матрица с элементами из коммутативного кольца R. Тогда её определителем называется следующий элемент из R:*

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \; a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}.$$

Мы будем использовать также следующее обозначение:

$$\begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{vmatrix} = \det \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}.$$

Прежде чем формулировать свойства определителя, вычислим определитель матриц размера $n \times n$, где $n \leqslant 3$.

Пример 11.8 (Определитель матрицы 1×1)**.** Мы имеем $S_1 = \{e\}$ и $\operatorname{sgn} e = 1$. Поэтому

$$|a_{1,1}| = \operatorname{sgn} e \; a_{1,e(1)} = a_{1,1}.$$

Пример 11.9 (Определитель матрицы 2×2)**.** Мы имеем $S_2 = \{e, (1,2)\}$ и $\operatorname{sgn} e = 1$, $\operatorname{sgn}\sigma = -1$, где мы обозначили $\sigma = (1, 2)$. Поэтому

$$\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = \operatorname{sgn} e \; a_{1,e(1)} a_{2,e(2)} + \operatorname{sgn}\sigma \; a_{1,\sigma(1)} a_{2,\sigma(2)} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}.$$

Пример 11.10 (Определитель матрицы 3×3)**.** Мы имеем

$$S_3 = \{e, (1,2,3), (1,3,2), (1,2), (1,3), (2,3)\},$$

$$\operatorname{sgn} e = \operatorname{sgn}(1,2,3) = \operatorname{sgn}(1,3,1) = 1, \quad \operatorname{sgn}(1,2) = \operatorname{sgn}(1,3) = \operatorname{sgn}(2,3) = -1.$$

Поэтому

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = a_{1,1}a_{2,2}a_{3,3}+a_{1,2}a_{2,3}a_{3,1}+a_{1,3}a_{2,1}a_{3,2}-a_{1,2}a_{2,1}a_{3,3}-a_{1,3}a_{2,2}a_{3,1}-a_{1,1}a_{2,3}a_{3,2}.$$

11.2. **Свойства определителя.**

Теорема 11.11. *Пусть $A = (a_{i,j})_{i,j=1}^n$ — матрица.*

(1)
$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \; a_{\sigma(1),1}a_{\sigma(2),2}\cdots a_{\sigma(n),n} = \det A^T.$$

(2) *Если матрица A содержит нулевую строку или нулевой столбец, то* $\det A = 0$.
(3) *Пусть матрица A' получается из A умножением всех элементов i-й строки (j-го столбца) на фиксированный элемент кольца c. Тогда* $\det A' = c \det A$.
(4) *Пусть $i = 1, \ldots, n$ ($j = 1, \ldots, n$) — некоторый индекс такой, что $a_{i,j} = a'_{i,j} + a''_{i,j}$ для любого $j = 1, \ldots, n$ ($i = 1, \ldots, n$). Обозначим через A' и A'' матрицы, элементы которых вне i-й строки (j-го столбца) совпадают с соответствующими элементами матрицы A, а элемент стоящий на пересечении i-й строки и j-го столбца равен $a'_{i,j}$ и $a''_{i,j}$ соответственно для любого $j = 1, \ldots, n$ ($i = 1, \ldots, n$). Тогда* $\det A = \det A' + \det A''$.
(5) *Если матрица A содержит две одинаковые строки или два одинаковых столбца, то* $\det A = 0$.
(6) *Если матрица B получается из A перестановкой двух строк или двух столбцов, то* $\det B = -\det A$.
(7) *Если строки (столбцы) матрицы B получается из строк (столбцов) матрицы A элементарным преобразованием вида* (ЭВ3)*, то* $\det B = \det A$.

Доказательство. (1) Переставляя множители в каждом произведении $a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{n,\sigma(n)}$, так чтобы вторые индексы возрастали от 1 до n, мы согласно следствию 11.5 можем записать

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \; a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{n,\sigma(n)} = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \; a_{\sigma^{-1}(1),1}a_{\sigma^{-1}(2),2}\cdots a_{\sigma^{-1}(n),n} =$$
$$= \sum_{\sigma \in S_n} \operatorname{sgn}\sigma^{-1} \; a_{\sigma^{-1}(1),1}a_{\sigma^{-1}(2),2}\cdots a_{\sigma^{-1}(n),n}.$$

Когда σ пробегает множество S_n, то σ^{-1} пробегает тоже самое множество. Поэтому, заменяя σ^{-1} на σ в правой части формулы выше, получаем требуемую формулу для определителя.

(2) Пусть существует некоторое $i = 1, \ldots, n$ такое, что $a_{i,j} = 0$ для всех $j = 1, \ldots, n$. По определению 11.7 получаем

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \; a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{i-1,\sigma(i-1)}\, 0 \, a_{i+1,\sigma(i+1)}\cdots a_{\sigma(n),n} = 0.$$

В случае, когда матрица A содержит нулевой столбец утверждение получается аналогично с применением части (1).

(3) По определению 11.7 мы получаем

$$\det A' = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma\ a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{i-1,\sigma(i-1)}\,ca_{i,\sigma i}\,a_{i+1,\sigma(i+1)}\cdots a_{\sigma(n),n} =$$
$$= c\sum_{\sigma\in S_n} \operatorname{sgn}\sigma\ a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{i-1,\sigma(i-1)}\,a_{i,\sigma i}\,a_{i+1,\sigma(i+1)}\cdots a_{\sigma(n),n} = c\det A.$$

В случае, когда матрица A' получается из матрицы A умножением на c каждого элемента столбца, утверждение доказывается аналогично с использованием части (1).

(4) Предположим, что $a_{i,j} = a'_{i,j} + a''_{i,j}$ для любого $j = 1,\dots,n$. По определению 11.7 мы получаем

$$\det A = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma\ a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{i-1,\sigma(i-1)}\,(a'_{i,\sigma i} + a''_{i,\sigma i})\,a_{i+1,\sigma(i+1)}\cdots a_{\sigma(n),n} =$$
$$= \sum_{\sigma\in S_n} \operatorname{sgn}\sigma\ a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{i-1,\sigma(i-1)}\,a'_{i,\sigma i}\,a_{i+1,\sigma(i+1)}\cdots a_{\sigma(n),n} +$$
$$+ \sum_{\sigma\in S_n} \operatorname{sgn}\sigma\ \ a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{i-1,\sigma(i-1)}\,a''_{i,\sigma i}\,a_{i+1,\sigma(i+1)}\cdots a_{\sigma(n),n} \ = \ \det A' \ + \ \det A''.$$

В случае, когда $a_{i,j} = a'_{i,j} + a''_{i,j}$ для любого $i = 1,\dots,n$, утверждение доказывается аналогично с использованием части (1).

(5) Пусть i-я и j-я строки матрицы A совпадают. Это означает, что $a_{i,k} = a_{j,k}$ для любого $k = 1,\dots,n$. Рассмотрим подгруппу $H = \{e,(i,j)\}$ группы S_n и разбиение $S_n = \bigsqcup_{m=1}^{n!/2} H\sigma_m$ на правые смежные классы. Здесь $\{\sigma_m\}_{m=1}^{n!/2}$ — множество представителей смежных классов. Замечая, что $a_{(i,j)l,k} = a_{l,k}$ для любых $k,l = 1,\dots,n$, согласно части (1) мы можем записать

$$\det A = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma\ a_{\sigma(1),1}a_{\sigma(2),2}\cdots a_{\sigma(n),n} = \sum_{i=m}^{n!/2}\Big(\operatorname{sgn}\sigma_m\ a_{\sigma_m(1),1}a_{\sigma_m(2),2}\cdots a_{\sigma_m(n),n} +$$
$$+ \operatorname{sgn}((i,j)\sigma_m)\ a_{(i,j)\sigma_m(1),1}a_{(i,j)\sigma_m(2),2}\cdots a_{(i,j)\sigma_m(n),n}\Big).$$

Сумма в круглых скобках в правой части равна 0 при любом m, так как $\operatorname{sgn}((i,j)\sigma_m) = \operatorname{sgn}(i,j)\operatorname{sgn}\sigma_m = -\operatorname{sgn}\sigma_m$ и $a_{(i,j)\sigma_m(k),k} = a_{\sigma_m(k),k}$ для любого $k = 1,\dots,n$.

В случае, когда совпадают столбцы матрицы A, утверждение доказывается аналогично с использованием определения 11.7.

(6) Пусть B получается из A перестановкой i-й и j-й строки. В этом случае мы получаем $b_{k,l} = a_{(i,j)k,l}$ для любых $k,l = 1,\dots,n$. Согласно части (1) мы получаем

$$\det B = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma\ b_{\sigma(1),1}b_{\sigma(2),2}\cdots b_{\sigma(n),n} = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma\ a_{(i,j)\sigma(1),1}a_{(i,j)\sigma(2),2}\cdots a_{(i,j)\sigma(n),n} =$$
$$= -\sum_{\sigma\in S_n} \operatorname{sgn}((i,j)\sigma)\ a_{(i,j)\sigma(1),1}a_{(i,j)\sigma(2),2}\cdots a_{(i,j)\sigma(n),n}.$$

Когда σ пробегает множество S_n, то $(i,j)\sigma$ пробегает тоже самое множество. Поэтому, заменяя $(i,j)\sigma$ на σ в правой части формулы выше, получаем, что правая часть равна $-\det A$, как и требовалось.

В случае, когда B получается из A перестановкой столбцов утверждение доказывается аналогично с использованием определения 11.7.

(7) Пусть B получается из A умножением i-й строки на c и добавлением произведения к строке $k \neq i$. Тогда согласно частям (4), (3) и (5) мы получаем

$$\det B = \begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & & \vdots \\ a_{k-1,1} & a_{k-1,2} & \cdots & a_{k-1,n} \\ a_{k,1}+ca_{i,1} & a_{k,2}+ca_{i,2} & \vdots & a_{k,n}+ca_{i,n} \\ a_{k+1,1} & a_{k+1,2} & \cdots & a_{k+1,n} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix} =$$

$$= \det A + \begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & & \vdots \\ a_{k-1,1} & a_{k-1,2} & \cdots & a_{k-1,n} \\ ca_{i,1} & ca_{i,2} & \vdots & ca_{i,n} \\ a_{k+1,1} & a_{k+1,2} & \cdots & a_{k+1,n} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} = \end{vmatrix} = \det A + c \begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & & \vdots \\ a_{k-1,1} & a_{k-1,2} & \cdots & a_{k-1,n} \\ a_{i,1} & a_{i,2} & \vdots & a_{i,n} \\ a_{k+1,1} & a_{k+1,2} & \cdots & a_{k+1,n} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} = \end{vmatrix}.$$

Определитель последней матрицы равен 0, так как её i-я строка совпадает с k-й. □

Теорема 11.12 (Определитель блочной матрицы)**.** *Пусть $A = (a_{i,j})_{i,j=1}^{n}$ — матрица и $m = 1, \ldots, n-1$ такие, что $a_{i,j} = 0$ при $i \leqslant m < j$ или при $j \leqslant m < i$. Рассмотрим блоки $A' = (a_{i,j})_{i,j=1}^{m}$ и $A'' = (a_{i+m,j+m})_{i,j=1}^{n-m}$. Тогда*

$$\det A = \det A' \det A''.$$

Доказательство. Рассмотрим случай, когда $a_{i,j} = 0$ при $i \leqslant m < j$. Воспользуемся определением 11.7 для вычисления $\det A$. Пусть $\sigma \in S_n$ — такой элемент, что

$$\sigma(\{1, \ldots, m\}) \neq \{1, \ldots, m\}$$

В этом случае $\sigma(i) > m$ для некоторого $i = 1, \ldots, m$. Следовательно, $a_{i,\sigma(i)} = 0$ и всё произведение $a_{1,\sigma(1)}a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$ равно нулю. Эти рассуждения показывают, что

$$\det A = \sum_{\sigma \in S_{n,m}} \operatorname{sgn}\sigma \; a_{1,\sigma(1)}a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}, \tag{29}$$

где $S_{n,m}$ — множество подстановок из S_n таких, что

$$\sigma(\{1, \ldots, m\}) = \{1, \ldots, m\}.$$

Заметим, что для каждой такой подстановки $\sigma(\{m+1, \ldots, n\}) = \{m+1, \ldots, n\}$. Следовательно, имеет место однозначное разложение $\sigma = \sigma_1\sigma_2$, где $\sigma_1, \sigma_2 \in S_n$ и $\sigma_1(i) = i$ при $i > m$ и $\sigma_2(i) = i$ при $i \leqslant m$.

Мы сформулируем этот факт несколько по-другому. Рассмотрим гомоморфизмы

$$\varphi_1 : S_m \to S_n, \quad \varphi_2 : S_{n-m} \to S_n,$$

заданные формулами

$$\varphi_1(\tau)(i) = \begin{cases} \tau(i), & \text{если } i \leqslant m, \\ i, & \text{если } i > m, \end{cases} \qquad \varphi_2(\delta)(i) = \begin{cases} \delta(i-m)+m, & \text{если } i > m, \\ i, & \text{если } i \leqslant m, \end{cases}$$

Легко понять, что любая подстановка $\sigma \in S_{n,m}$ однозначно представляется в виде $\sigma = \varphi_1(\tau)\varphi_2(\delta)$, где $\tau \in S_m$ и $\delta \in S_{n-m}$.

Мы утверждаем, что

$$\operatorname{sgn}\varphi_1(\tau) = \operatorname{sgn}\tau \text{ для любой } \tau \in S_m \text{ и } \operatorname{sgn}\varphi_2(\delta) = \operatorname{sgn}\delta \text{ для любой } \delta \in S_{n-m}. \qquad (30)$$

Действительно, запишем $\tau \in S_n$ в виде

$$\tau = \begin{pmatrix} 1 & 2 & \cdots & m \\ i_1 & i_2 & \cdots & i_m \end{pmatrix}$$

Тогда получаем

$$\varphi_1(\tau) = \begin{pmatrix} 1 & 2 & \cdots & m & m+1 & \cdots & n \\ i_1 & i_2 & \cdots & i_m & m+1 & \cdots & n \end{pmatrix}$$

Очевидно, что множества инверсий в нижних строчках обеих матриц совпадают. В частности совпадают и их количества, откуда следует первая из формул (30).

Аналогично, запишем $\delta \in S_{n-m}$ в виде

$$\delta = \begin{pmatrix} 1 & 2 & \cdots & n-m \\ j_1 & j_2 & \cdots & j_{n-m} \end{pmatrix}$$

Тогда получаем

$$\varphi_2(\delta) = \begin{pmatrix} 1 & \cdots & n & n-m+1 & n-m+2 & \cdots & n \\ 1 & \cdots & n & j_1+m & j_2+m & \cdots & j_{n-m}+m \end{pmatrix}$$

Очевидно, что (c, d) является инверсией нижней строки первой матрицы тогда и только тогда, когда $(c+m, d+m)$ является инверсией нижней строки второй матрицы, в которой все инверсии имеют такой вид. Отсюда следует вторая из формул (30).

Теперь используя формулы (29) и (30), определения гомоморфизмов φ_1 и φ_2 и лемму 11.4, мы можем записать

$$\det A = \sum_{\tau_\in S_n, \delta \in S_{n-m}} \operatorname{sgn}(\varphi_1(\tau)\varphi_2(\delta)) a_{1,\varphi_1(\tau)(1)} \cdots a_{m,\varphi_1(\tau)(m)} a_{m+1,\varphi_2(\delta)(m+1)} \cdots a_{n,\varphi_2(\delta)(n)} =$$

$$= \sum_{\tau \in S_n, \delta \in S_{n-m}} \operatorname{sgn}(\tau)\operatorname{sgn}(\delta) a_{1,\tau(1)} \cdots a_{m,\tau(m)} a_{m+1,\delta(1)+m} \cdots a_{n,\delta(n-m)+m} =$$

$$= \left(\sum_{\tau \in S_n} \operatorname{sgn}(\tau) a_{1,\tau(1)} \cdots a_{m,\tau(m)} \right) \left(\sum_{\delta \in S_{n-m}} \operatorname{sgn}(\delta) a_{m+1,\delta(1)+m} \cdots a_{n,\delta(n-m)+m} \right) =$$

$$= \det A' \det A''.$$

□

Следствие 11.13. *Определитель верхнетреугольной (нижнетреугольной) матрицы равен произведению её диагональных элементов.*

Доказательство. Достаточно выделить в верхнетреугольной (нижнетреугольной) матрице блоки размера 1×1, лежащие на диагонали, и применить теорему 11.12 и формулу для определителя матрицы размера 1×1 (см. пример 11.8). □

Следствие 11.14. *Пусть X — матрица вида $X_{i,j}^{(n)}$, $X_i^{(n)}(\alpha)$ или $X_{i,j}^{(n)}(\alpha)$ и A — матрица размера $n \times n$. Тогда* $\det(XA) = \det X \det A$.

Доказательство. Утверждение следует из леммы 9.8, частей (6), (3) и (7) теоремы 11.11 и следующих равенств

$$\det X_{i,j}^{(n)} = -1, \quad \det X_i^{(n)}(\alpha) = \alpha, \quad \det X_{i,j}^{(n)}(\alpha) = 1,$$

которые следуют из следствия 11.13 и части (6) теоремы 11.11 (для первого равенства). □

Определение 11.15. *Квадратная матрица над полем называется вырожденной, если её ранг меньше её размера.*

Другими словами, квадратная матрица над полем вырождена тогда и только тогда, когда её строки (столбцы) являются линейно зависимыми. По теореме 9.14 матрица вырождена тогда и только тогда, когда вырождена её транспонированная матрица.

Лемма 11.16. *Квадратная матрица над полем вырождена тогда и только тогда, когда её определитель равен нулю.*

Доказательство. Предположим, что A вырождена. Тогда одна из её строк r_i является линейной комбинацией оставшихся строк:

$$r_i = \alpha_1 r_1 + \cdots + \alpha_{i-1} r_{i-1} + \alpha_{i+1} r_{i+1} + \cdots + \alpha_n r_n,$$

где не все коэффициенты $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n$ равны нулю. Применяя соответствующие элементарные преобразования вида (ЭВ3) к строкам матрицы A, мы можем добиться того, что i-я строка получившейся матрицы B будет нулевой. В нашей записи выше мы должны умножить первую строку на $-\alpha_1$ и прибавить к i-й, умножить вторую строку на $-\alpha_2$ и прибавить к i-й и т.д. Согласно части (7) теоремы 11.11 мы имеем $\det A = \det B$. Однако $\det B = 0$ согласно части (2) теоремы 11.11.

Предположим наоборот, что $\det A = 0$. По следствию 9.9 мы получаем представление $A = X_1 \cdots X_k B$, где $X_1, \ldots, X_k$ — матрицы вида $X_{i,j}^{(n)}$, $X_i^{(n)}(\alpha)$ или $X_{i,j}^{(n)}(\beta)$ для $\alpha \neq 0$ и B — ступенчатая матрица. По следствию 11.14 получаем

$$0 = \det A = \det X_1 \cdots \det X_k \det B.$$

Так как $\det X_i \neq 0$ и мы работаем над полем, то $\det B = 0$. Так как матрица B ступенчатая, то по следствию 11.13 у неё должна быть нулевая строка (внизу). Следовательно, $\operatorname{rank} B < n$, где $n \times n$ размер матриц A и B. Однако элементарные преобразования строк не меняют ранга матрицы. Следовательно,

$$\operatorname{rank} A = \operatorname{rank} B < n$$

и матрица A вырожденная. □

Лемма 11.17. *Пусть A и B — квадратные матрицы над полем одинакового размера. Если матрица B вырожденная, то матрицы AB и BA тоже вырожденные.*

Доказательство. Так как матрица B вырожденная, то существует нетривиальная (в которой не все коэффициенты равны нулю) линейная комбинация её столбцов равная нулю. По лемме 9.11 столбцы матрицы AB удовлетворяют тем же линейным уравнениям, что и столбцы матрицы B, то есть являются линейно зависимыми. Следовательно, матрицы AB вырожденная.

Для того, чтобы доказать вырожденность матрицы BA, надо заметить, что матрица B^T вырожденная. По уже доказанному мы получаем, что матрица $A^T B^T = (BA)^T$ вырожденная. Поэтому вырождена и матрица BA. □

Лемма 11.18. *Любая невырожденная матрица A над полем может быть представлена в виде $A = X_1 \cdots X_k$, где $X_1, \ldots, X_k$ — матрицы вида $X_{i,j}^{(n)}$, $X_i^{(n)}(\alpha)$ или $X_{i,j}^{(n)}(\beta)$, где $\alpha \neq 0$.*

Доказательство. Согласно следствию 9.9 мы имеем представление $A = X_1 \cdots X_m B$, где $X_1, \ldots, X_m$ матрицы вида, указанного в формулировке леммы и B — ступенчатая

матрица. Так как $\operatorname{rank} A = \operatorname{rank} B$, то $\det B \neq 0$ и B — верхнетреугольная матрица с ненулевыми элементами на диагонали. Запишем её в следующем виде

$$B = \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n-1} & b_{1,n} \\ 0 & b_{2,2} & \cdots & b_{2,n-1} & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & b_{n-1,n-1} & b_{n-1,n} \\ 0 & 0 & \cdots & 0 & b_{n,n} \end{pmatrix}$$

Мы опишем как привести эту матрицу к единичной элементарными преобразованиями. Сначала поделим последнюю строку на $b_{n,n}$ (преобразование (ЭВ2)). Затем для каждого $i = 1, \ldots, n-1$ умножим получившуюся n-у строку на $-b_{i,n}$ и добавим к i-й строке (преобразование (ЭВ3). В результате мы получим матрицу

$$\begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n-1} & 0 \\ 0 & b_{2,2} & \cdots & b_{2,n-1} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & b_{n-1,n-1} & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Теперь забудем о последней строке и последнем столбце и повторим описанный выше процесс. В результате мы придём к единичной матрице. Записывая все произведённые элементарные преобразования в виде умножения слева на соответствующие матрицы, получаем требуемый результат. □

Теорема 11.19. *Пусть A и B — квадратные матрицы одинакового размера над произвольным коммутативным кольцом. Тогда* $\det(AB) = \det A \det B$.

Доказательство 1. Пусть $A = (a_{i,j})_{i,j=1}^n$ и $B = (b_{i,j})_{i,j=1}^n$. Положим $C = AB$ и пусть $C = (c_{i,j})_{i,i=1}^n$. По определению мы получаем

$$\det C = \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \prod_{i=1}^n c_{i,\sigma(i)} = \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \prod_{i=1}^n \sum_{j=1}^n a_{i,j} b_{j,\sigma(i)}. \tag{31}$$

В этом произведении, мы хотим воспользоваться законом дистрибутивности, чтобы раскрыть произведение сумм в этой формуле. Мы получаем

$$\prod_{i=1}^n \sum_{j=1}^n a_{i,j} b_{j,\sigma(i)} = \sum_{\tau \in X_n} \prod_{i=1}^n a_{i,\tau(i)} b_{\tau(i),\sigma(i)} = \sum_{\tau \in X_n} \left(\prod_{i=1}^n a_{i,\tau(i)} \right) \left(\prod_{k=1}^n b_{\tau(k),\sigma(k)} \right),$$

где X_n — множество всех отображений из множества $\{1, \ldots, n\}$ в себя. Заменим, что $S_n \subset X_n$ и что при этом множество X_n больше чем S_n при $n > 1$, так как оно содержит отображения, не являющиеся биекциями. Подставляя это выражение в формулу (31), получаем

$$\begin{aligned} \det C = \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \sum_{\tau \in X_n} \left(\prod_{i=1}^n a_{i,\tau(i)} \right) \left(\prod_{k=1}^n b_{\tau(k),\sigma(k)} \right) = \\ = \sum_{\tau \in X_n} \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \left(\prod_{i=1}^n a_{i,\tau(i)} \right) \left(\prod_{k=1}^n b_{\tau(k),\sigma(k)} \right) = \\ = \sum_{\tau \in X_n} \left(\prod_{i=1}^n a_{i,\tau(i)} \right) \left(\sum_{\sigma \in S_n} \operatorname{sgn} \sigma \prod_{k=1}^n b_{\tau(k),\sigma(k)} \right). \end{aligned}$$

Сначала рассмотрим случай, когда τ не является подстановкой, то есть $\tau \in X_n \setminus S_n$. В этом случае $\tau(s) = \tau(r)$ для различных индексов $r, s = 1, \ldots, n$. Рассмотрим подгруппу $H = \{e, (s, r)\}$ и разбиение $S_n = \bigsqcup_{m=1}^{n!/2} \sigma_m H$ на левые смежные классы. Отсюда мы получаем

$$\sum_{\sigma \in S_n} \operatorname{sgn}\sigma \prod_{k=1}^{n} b_{\tau(k),\sigma(k)} = \sum_{m=1}^{n!/2} \left(\operatorname{sgn}\sigma_m \prod_{k=1}^{n} b_{\tau(k),\sigma_m(k)} + \operatorname{sgn}(\sigma_m(r,s)) \prod_{k=1}^{n} b_{\tau(k),\sigma_m(r,s)(k)} \right) =$$

$$\sum_{m=1}^{n!/2} \operatorname{sgn}\sigma_m \left(\prod_{k=1}^{n} b_{\tau(k),\sigma_m(k)} - \prod_{k=1}^{n} b_{\tau(k),\sigma_m(r,s)(k)} \right) =$$

$$\sum_{m=1}^{n!/2} \operatorname{sgn}\sigma_m \left(\prod_{\substack{k=1 \\ k \neq r,s}}^{n} b_{\tau(k),\sigma_m(k)} \right) \left(b_{\tau(r),\sigma_m(r)} b_{\tau(s),\sigma_m(s)} - b_{\tau(r),\sigma_m(s)} b_{\tau(s),\sigma_m(r)} \right) = 0.$$

Отсюда, переупорядочивая множители в произведении $\prod_{k=1}^{n} b_{\tau(k),\sigma(k)}$, получаем

$$\det C = \sum_{\tau \in S_n} \left(\prod_{i=1}^{n} a_{i,\tau(i)} \right) \left(\sum_{\sigma \in S_n} \operatorname{sgn}\sigma \prod_{k=1}^{n} b_{\tau(k),\sigma(k)} \right) =$$

$$= \sum_{\tau \in S_n} \left(\prod_{i=1}^{n} a_{i,\tau(i)} \right) \left(\sum_{\sigma \in S_n} \operatorname{sgn}\sigma \prod_{k=1}^{n} b_{k,\sigma\tau^{-1}(k)} \right) =$$

$$= \sum_{\tau \in S_n} \operatorname{sgn}\tau \left(\prod_{i=1}^{n} a_{i,\tau(i)} \right) \left(\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma\tau^{-1}) \prod_{k=1}^{n} b_{k,\sigma\tau^{-1}(k)} \right) =$$

$$= \sum_{\tau \in S_n} \operatorname{sgn}\tau \left(\prod_{i=1}^{n} a_{i,\tau(i)} \right) \left(\sum_{\mu \in S_n} \operatorname{sgn}\mu \prod_{k=1}^{n} b_{k,\mu(k)} \right) =$$

$$= \left(\sum_{\tau \in S_n} \operatorname{sgn}\tau \prod_{i=1}^{n} a_{i,\tau(i)} \right) \left(\sum_{\mu \in S_n} \operatorname{sgn}\mu \prod_{k=1}^{n} b_{k,\mu(k)} \right) = \det A \det B.$$

□

Доказательство 2. Рассмотрим следующее коммутативное кольцо от нескольких переменных:

$$K = \mathbb{Z}[x_{i,j}, y_{i,j}]_{i,j=1}^{n}$$

и следующие две матрицы с элементами из этого кольца:

$$\mathbf{A} = \begin{pmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,n} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} y_{1,1} & \cdots & y_{1,n} \\ \vdots & \ddots & \vdots \\ y_{n,1} & \cdots & y_{n,n} \end{pmatrix}.$$

Наша цель — доказать равенство

$$\det(\mathbf{AB}) = \det \mathbf{A} \det \mathbf{B}, \tag{32}$$

которое является равенством двух многочленов из кольца K, которое содержится в следующем кольце:

$$\mathbb{Q}[x_{i,j}, y_{i,j}]_{i,j=1}^{n}.$$

Согласно ..., нам достаточно доказать, что эти многочлены равны при поставке вместо переменных $x_{i,j}$ и $y_{i,j}$ любых рациональных чисел. Другими словами нам достаточно

доказать, что $\det(AB) = \det A \det B$ для любых матриц A и B над рациональными числами.

Рассмотрим сначала случай, когда матрица A вырожденная. В этом случае матрица AB вырождена по лемме 11.17. По лемме 11.16 получаем

$$\det(AB) = 0 = 0 \det B = \det A \det B.$$

Теперь рассмотрим случай, когда A невырожденная матрица. Мы получаем

$$A = X_1 \cdots X_k,$$

где $X_1, \ldots, X_k$ — матрицы как в лемме 11.18. Теперь по следствию 11.14 получаем

$$\det A = \det(X_1 \cdots X_k) = \det X_1 \det(X_2 \cdots X_k) = \cdots = \det X_1 \det X_2 \cdots \det X_k.$$

и

$$\det(AB) = \det(X_1 \cdots X_k B) = \det X_1 \det(X_2 \cdots X_k B) = \cdots$$
$$\cdots = \det X_1 \det X_2 \cdots \det X_k \det B.$$

Сравнивая, получаем $\det(AB) = \det A \det B$. Этим равенство (32) доказано.

Чтобы завершить доказательство, достаточно заметить, что для любого набора элементов $\{a_{i,j}\}_{i,j=1}^n$ и $\{b_{i,j}\}_{i,j=1}^n$ произвольного коммутативного кольца R существует (единственный) гомоморфизм колец $\varphi : K \to R$ такой, что $\varphi(x_{i,j}) = a_{i,j}$ и $\varphi(y_{i,j}) = b_{i,j}$ для любых $i, j = 1, \ldots, n$. Применяя этот гомоморфизм к доказанному равенству (32), мы получаем утверждение теоремы. □

Следствие 11.20. *Пусть A — обратимая матрица. Тогда* $\det A^{-1} = 1/\det A$.

Доказательство. Из равенства $A^{-1}A = E$ и теоремы 11.19, получаем

$$1 = \det E = \det(A^{-1}A) = \det(A^{-1}) \det(A).$$

Отсюда следует утверждение следствия. □

11.3. **Разложение определителя по строке и столбцу.** Пусть A — квадратная матрица размера $n \times n$ и $i, j = 1, \ldots, n$. Мы рассмотрим матрицу $A(i, j)$, получающуюся из A вычёркиванием i-й строки и j-о столбца. Мы положим

$$A_{i,j} = (-1)^{i+j} \det A(i, j).$$

Теорема 11.21 (Разложение по строке и столбцу)**.** *Пусть $A = (a_{i,j})_{i,j}^n$ — квадратная матрица размера $n \times n$. Тогда для любого $i = 1, \ldots, n$ выполнено (разложение по строке)*

$$\det A = \sum_{j=1}^{n} a_{i,j} A_{i,j}$$

и для любого $j = 1, \ldots, n$ выполнено (разложение по столбцу)

$$\det A = \sum_{i=1}^{n} a_{i,j} A_{i,j}.$$

Доказательство. Докажем первую формулу. Рассмотрим следующее разложение i-й строки

$$(a_{i,1}, \cdots, a_{i,n}) = \sum_{j=1}^{n} a_{i,j} e_j,$$

где $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ с единицей на j-м месте. Пользуясь линейностью определителя по строке (части (2) и (3) теоремы 11.11), мы получаем

$$\det A = \sum_{j=1}^{n} a_{i,j} A^{i,j}, \tag{33}$$

где

$$A^{i,j} = \begin{vmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix}.$$

Наша цель перенести единицу стоящую на пересечении i-й строки и j-о столбца в левый верхний угол, переставляя строки и столбцы матрицы. Переставляя её с каждым нулём i-й строки слева, мы перенесём единицу в начала i-й строки, потратив $j-1$ перестановку. Аналогично, переставляя эту единицу с каждым элементом первого столбца выше, мы перенесём её в левый верхний угол, потратив при этом ещё $i-1$ перестановку. Отсюда пользуясь частью (6) теоремы 11.11 и теоремой 11.12, получаем

$$A^{i,j} = (-1)^{j-1+i-1} \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ a_{1,j} & a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,j} & a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,j} & a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n,j} & a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix} =$$

$$= (-1)^{i+j} \det A(i,j) = A_{i,j}.$$

Подставляя полученное равенство в формулу (33), получаем требуемое разложение определителя.

Формула разложения определителя по столбцу получается аналогичными рассуждениями с j-м столбцом или применением уже полученной формулы разложения по строке к транспонированной матрице. □

Лемма 11.22 (Фальшивое разложение)**.** *Пусть $A = (a_{i,j})_{i,j}^{n}$ — квадратная матрица размера $n \times n$. Тогда для любых различных индексов $i, k = 1, \dots, n$ выполнено*

$$\sum_{j=1}^{n} a_{k,j} A_{i,j} = 0$$

и для любых различных индексов $k, j = 1, \dots, n$ выполнено

$$\sum_{i=1}^{n} a_{i,k} A_{i,j} = 0.$$

Доказательство. Для доказательства первой формулы вместо матрицы A рассмотрим матрицу B, полученную из матрицы A заменой её i-й строки на её k-ю строку. Таким образом i-я и k-строки матрицы B совпадают. Согласно части (5) теоремы 11.11 получаем $\det B = 0$. При этом $B(i,j) = A(i,j)$ так как после вычёркивания i-й строки из

обеих матриц A и B мы получим одну и туже матрицу. Отсюда мы получаем $A_{i,j} = B_{i,j}$. По теореме 11.21 получаем

$$0 = \det B = \sum_{j=1}^{n} a_{k,j} B_{i,j} = \sum_{j=1}^{n} a_{k,j} A_{i,j}.$$

Вторая формула доказывается аналогично. □

Естественно возникает вопрос: зачем нужно фальшивое разложение? Следующая теорема даёт на него ответ.

Теорема 11.23 (Формула обратной матрицы)**.** *Пусть A — квадратная матрица над полем размера $n \times n$ такая, что $\det A \neq 0$. Тогда*

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} A_{1,1} & A_{2,1} & \cdots & A_{n,1} \\ A_{1,2} & A_{2,2} & \cdots & A_{n,2} \\ \vdots & \vdots & & \vdots \\ A_{1,n} & A_{2,2} & \cdots & A_{n,n} \end{pmatrix}.$$

Доказательство. Нам требуется доказать формулы

$$A \begin{pmatrix} A_{1,1} & A_{2,1} & \cdots & A_{n,1} \\ A_{1,2} & A_{2,2} & \cdots & A_{n,2} \\ \vdots & \vdots & & \vdots \\ A_{1,n} & A_{2,2} & \cdots & A_{n,n} \end{pmatrix} = (\det A)E, \quad \begin{pmatrix} A_{1,1} & A_{2,1} & \cdots & A_{n,1} \\ A_{1,2} & A_{2,2} & \cdots & A_{n,2} \\ \vdots & \vdots & & \vdots \\ A_{1,n} & A_{2,2} & \cdots & A_{n,n} \end{pmatrix} A = (\det A)E.$$

Докажем первую формулу. Для этого умножим k-у строку первом матрицы левой части на i-й столбец второй матрицы левой части (при помощи стандартного скалярного произведения). По теореме 11.21, когда $k = i$, и лемме 11.22, когда $k \neq i$, получаем (разложение по строке)

$$\sum_{j=1}^{n} a_{k,j} A_{i,k} = (\det A)\delta_{k,j}.$$

Этим доказывается первое равенство. Второе равенство доказывается аналогично с использованием разложения (обычного и фальшивого) по столбцу. □

В дальнейшем (см. теорему 12.2) нам будет полезно использовать матрицу

$$B = \begin{pmatrix} A_{1,1} & A_{2,1} & \cdots & A_{n,1} \\ A_{1,2} & A_{2,2} & \cdots & A_{n,2} \\ \vdots & \vdots & & \vdots \\ A_{1,n} & A_{2,2} & \cdots & A_{n,n} \end{pmatrix},$$

упомянутую в предыдущей теореме. Она называется *присоединённой* матрицей матрицы A. Её вхождения многочлены степени $n-1$ от коэффициентов матрицы A и она удовлетворяет следующим равенствам:

$$AB = BA = (\det A)E. \tag{34}$$

Следствие 11.24. *Квадратная матрица A над полем обратима тогда и только тогда, когда её определитель не равен нулю. Если $AB = E$ ($BA = E$) для некоторой матрицы того же размера, то $BA = E$ ($AB = E$) и, следовательно, $B = A^{-1}$.*

Доказательство. Если $\det A \neq 0$, то A^{-1} вычисляется по формуле, приведённой в теореме 11.23. Пусть наоборот A обратима. Тогда по теореме 11.19 получаем

$$1 = \det E = \det(AA^{-1}) = \det A \det A^{-1}.$$

Отсюда $\det A \neq 0$.

Предположим теперь, что $AB = E$. Применяя к этому равенству det, по теореме 11.19 получаем $\det A \neq 0$ (смотрите вычисления выше). По теореме 11.23 существует обратная матрица A^{-1}. Умножая равенство $AB = E$ на A^{-1} слева, получаем

$$B = EB = A^{-1}AB = A^{-1}E = A^{-1}.$$

По определению обратной матрицы, мы получаем $BA = A^{-1}A = E$. Второе утверждение доказывается аналогично.

□

11.4. **Миноры.** Пусть $A = (a_{i,j})_{i=1,\dots,m;j=1,\dots,n}$ — матрица над коммутативным кольцом. Для любой пары подмножеств

$$I = \{i_1, \dots, i_k\} \subset \{1, \dots, m\}, \quad J = \{j_1, \dots, j_l\} \subset \{1, \dots, n\}$$

мы рассмотрим *подматрицу*

$$A(I, J) = \begin{pmatrix} a_{i_1,j_1} & a_{i_1,j_2} & \cdots & a_{i_1,j_l} \\ a_{i_2,j_1} & a_{i_2,j_2} & \cdots & a_{i_2,j_l} \\ \vdots & \vdots\ddots & \vdots & \\ a_{i_k,j_1} & a_{i_k,j_2} & \cdots & a_{i_k,j_l} \end{pmatrix}$$

Если $|I| = |J|$, то мы можем рассмотреть определитель этой матрицы

$$A[I, J] = \det A(I, J).$$

Этот элемент называется *минором*. Минор называется *главным*, если $I = J$ и *угловым*, если $I = J = \{1, \dots, r\}$ для некоторого r. Для главных миноров мы используем обозначение $A[I] = A[I, I]$. Аналогичное обозначение используем для подматриц $A(I) = A(I, I)$. Мы уже имели дело с подматрицами и минорами в §11.3:

$$A(i, j) = A(\{1, \dots, i-1, i+1, \dots, n\}, \{1, \dots, j-1, j+1, \dots, n\}),$$

$$A_{i,j} = (-1)^{i+j} A[\{1, \dots, i-1, i+1, \dots, n\}, \{1, \dots, j-1, j+1, \dots, n\}].$$

Теорема 11.25. *Ранг матрицы над полем равен максимальному размеру её ненулевого минора.*

Доказательство. Пусть $A = (a_{i,j})_{i=1,\dots,m,j=1,\dots,n}$ — матрица над полем, k — её ранг и l — размер её максимального ненулевого минора. Обозначим через $r_1, \dots, r_m$ строки матрицы A. Пусть $r_{i_1}, \dots, r_{i_k}$ — базис линейной оболочки строк, где $1 \leqslant i_1 < \dots < i_k \leqslant n$. Рассмотрим матрицу B со строками $r_{i_1}, \dots, r_{i_k}$. Обозначим через $c_1, \cdots, c_n$ её столбцы. По теореме 9.14, размерность линейной оболочки столбцов матрицы B тоже равна k. Следовательно, существуют столбцы $c_{j_1}, \dots, c_{j_k}$, являющиеся базисом этой линейной оболочки, где $1 \leqslant j_1 < \cdots < j_k \leqslant n$. Пусть — матрица со столбцами $c_{j_1}, \dots, c_{j_k}$. Это — невырожденная матрица размера $k \times k$. По лемме 11.16 её определитель не равен нулю. Замечая, что $C = A(\{i_1, \dots, i_k\}, \{j_1, \dots, j_k\})$, получаем $k \leqslant l$.

Наоборот пусть $A[\{i_1, \dots, i_l\}, \{j_1, \dots, j_l\}] \neq 0$ для некоторых $1 \leqslant i_1 < \dots < i_l \leqslant n$ и $1 \leqslant j_1 < \cdots < j_l \leqslant n$. По лемме 11.16 строки подматрицы $A(\{i_1, \dots, i_l\}, \{j_1, \dots, j_l\})$ линейно независимы. Следовательно линейно независимы и строки исходной матрицы с номерами $i_1, \dots, i_l$. Отсюда $k \geqslant l$. □

12. Линейные операторы

Линейным оператором называется линейное отображение φ из векторного пространства в само себя. Таким образом, мы имеем возможность рассматривать один и тот же базис v на области определения и на области значений оператора. Матрицу оператора относительно этого базиса мы обозначим через $M_v(\varphi)$. Иначе говоря, $M_v(\varphi) = M_{v,v}(\varphi)$ в обозначениях пункта 10.2.

12.1. **Собственные значения и векторы операторов.** Пусть $\varphi : V \to V$ — линейный оператор на векторном пространстве V над полем $\mathbb{F}$. Вектор $\lambda \in V$ называется *собственным*, если он ненулевой и

$$\varphi(\lambda) = t\lambda \tag{35}$$

для некоторого $t \in \mathbb{F}$. В этом случае элемент поля t называется *собственным значением* оператора φ, соответствующим собственному вектору λ. Далее мы будем рассматривать только тот случай, когда пространство V конечномерно.

Пусть $v = (v_1, \ldots, v_n)$ — какой-нибудь базис пространства V. Запишем наш собственный вектор в виде

$$\lambda = a_1u_1 + \cdots + a_nu_n,$$

а результат действия оператора φ на него в виде

$$\varphi(\lambda) = b_1u_1 + \cdots + b_nu_n$$

для соответствующих $a_i, b_i \in \mathbb{F}$. Согласно формуле (18) верно равенство

$$\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = M_v(\varphi) \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}.$$

Таким образом, раскладывая обе части равенства (35), по базису v мы получаем

$$M_v(\varphi) \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} = t \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} = tE \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}.$$

Перенося всё в левую сторону, мы получаем

$$(M_v(\varphi) - tE) \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} = 0. \tag{36}$$

По лемме 11.16 получаем $\det(M_v(\varphi) - tE) = 0$. С другой стороны, если это уравнение выполняется некоторого $t \in \mathbb{F}$, то по той же лемме 11.16 существует некоторый ненулевой набор $(a_1, \ldots, a_n)$ элементов поля $\mathbb{F}$ для которого выполнено уравнение (36). Полагая $\lambda = a_1u_1 + \cdots + a_nu_n$, мы получаем $\varphi(\lambda) = t\lambda$.

Заметим, что многочлен $\det(M_v(\varphi) - tE)$ не зависит от выбора базиса v. Действительно, пусть v' — ещё один базис пространства V. По формуле (19) с помощью теоремы 11.19 и следствия 11.20 получаем

$$\begin{aligned} \det(M'_v(\varphi) - tE) &= \det(T^{-1}_{v,v'} M_v(\varphi) T_{v,v'} - tE) = \det(T^{-1}_{v,v'}(M_v(\varphi) - tE)T_{v,v'}) = \\ &= \det(T_{v,v'})^{-1} \det(M_v(\varphi) - tE) \det(T_{v,v'}) = \det(M_v(\varphi) - tE). \end{aligned} \tag{37}$$

Получившийся многочлен называется *характеристическим многочленом* оператора φ и обозначается $\chi(\varphi)$ или $\chi(\varphi, t)$, если мы хотим указать переменную.

Наряду с собственными векторами значениями операторов мы можем рассматривать собственные значения квадратных матриц. Пусть A — квадратная матрица размера $n \times n$ над полем $\mathbb{F}$. Рассмотрим оператор на арифметическом пространстве $\mathbb{F}^n$, заданный формулой $\varphi(\lambda) = A\lambda$. Его собственные векторы, собственные значения и характеристический многочлен назовём собственными векторами, собственными значениями и характеристическим значением матрицы A. Выбирая стандартный базис в $\mathbb{F}^n$, получаем, что характеристический многочлен матрицы A равен $\det(A - tE)$.

Пример. ...

Теорема 12.1. *Для любой квадратной матрицы размера $n \times n$ над полем выполнено равенство*

$$\chi(A,t) = (-1)^n t^n + (-1)^{n-1} a_{n-1} t^{n-1} + (-1)^{n-2} a_{n-2} t_{n-2} + \cdots - a_1 t + a_0,$$

где

$$a_k = \sum_{I \subset \{1,\ldots,n\}, |I| = n-k} A[I].$$

Доказательство. Пусть a_i обозначает i-й столбец матрицы A и через e_i — столбец высоты n, в котором 1 стоит на i-м месте. Тогда i-й столбец матрицы $A - tE$ равен $a_i - te_i$. Используя линейность определителя по столбцам, получаем

$$\begin{aligned}\chi(A,t) &= |a_1 - te_1, a_2 - te_2, \ldots, a_n - te_n| = \\ &= \sum_{k=0}^{n} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} |a_1, \ldots, a_{i_1 - 1}, -te_1, a_{i_1+1}, \ldots, a_{i_k - 1}, -te_{i_k}, a_{i_k+1}, \ldots, a_n| = \\ &= \sum_{k=0}^{n} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} (-t)^k |a_1, \ldots, a_{i_1 - 1}, e_1, a_{i_1+1}, \ldots, a_{i_k - 1}, e_{i_k}, a_{i_k+1}, \ldots, a_n|.\end{aligned}$$

Используя k раз разложение определителя по столбцам (теорема 11.21) получаем

$$|a_1, \ldots, a_{i_1 - 1}, e_1, a_{i_1+1}, \ldots, a_{i_k - 1}, e_{i_k}, a_{i_k+1}, \ldots, a_n| = A[I],$$

где $I = \{1, \ldots, n\} \setminus \{i_1, \ldots, i_k\}$. Заметим, что знак определителя не изменяется, так как все ненулевые элементы столбцов, по которым производится разложение находятся на диагонали. Наконец остаётся заметить, что коэффициент при t^n равен

$$(-1)^n [e_1, e_2, \ldots, e_n] = (-1)^n.$$

□

Из вычисления (37) мы видим, что коэффициенты характеристического многочлена (как и сам многочлен) не измениться, если матрицу заменить на сопряжённую. Из вышеприведённой теоремы следует, что свободный член характеристического многочлена равен определителю. Кроме того, имеет своё название первый коэффициент характеристического многочлена умноженного на $(-1)^n$:

$$\operatorname{tr} A = a_{1,1} + a_{2,2} + \cdots + a_{n,n}.$$

Этот элемент называется *следом* матрицы A.

Теорема 12.2 (Гамильтона-Келли)**.** *Любая квадратная матрица над полем является корнем своего характеристического многочлена.*

Доказательство. Рассмотрим матрицу A размера $n \times n$ над полем $\mathbb{F}$. Запишем её характеристический многочлен в виде

$$\chi(A,t) = \sum_{i=0}^{n} c_i t^i, \tag{38}$$

где $c_i \in \mathbb{F}$. Пусть B — присоединённая матрица матрицы $A - tE$. Вхождения этой матрицы — многочлены из $\mathbb{F}[t]$ степени на выше чем $n-1$. Поэтому справедливо разложение

$$B = \sum_{i=0}^{n-1} t^i B_i,$$

где $B_0, \ldots, B_{n-1}$ — матрицы над исходным полем $\mathbb{F}$. Нам будет удобно определить $B_n = 0$ и $B_{-1} = 0$. Используя характеристическое уравнение (34) для присоединённой матрицы, получаем

$$\chi(A,t)E = \det(A-tE)E = (A-tE)B = (A-tE)\sum_{i=0}^{n-1} t^i B_i =$$
$$= \sum_{i=0}^{n-1} t^i AB_i - \sum_{i=0}^{n-1} t^{i+1} B_i = \sum_{i=0}^{n} t^i AB_i - \sum_{i=0}^{n} t^i B_{i-1} = \sum_{i=0}^{n} t^i (AB_i - B_{i-1}).$$

Сравнивая с (38), получаем $c_i E = AB_i - B_{i-1}$. Следовательно, результат подстановки матрицы A вместо переменной t в характеристическом многочлене равен

$$\chi(A,A) = \sum_{i=0}^{n} c_i A^i = \sum_{i=0}^{n} (c_i E)A^i = \sum_{i=0}^{n} A^i (c_i E) =$$
$$= \sum_{i=0}^{n} A^i (AB_i - B_{i-1}) = \sum_{i=0}^{n} (A^{i+1}B_i - A^i B_{i-1}) = \sum_{i=0}^{n} A^{i+1} B_i - \sum_{i=0}^{n} A^i B_{i-1} =$$
$$= \sum_{i=1}^{n+1} A^i B_{i-1} - \sum_{i=0}^{n} A^i B_{i-1} = A^{n+1} B_n - B_{-1} = 0.$$

□

Следствие 12.3. *Любой оператор на конечномерном векторном пространстве является корнем своего характеристического многочлена.*

Доказательство. Пусть φ — оператор на конечномерном векторном пространстве V и v — какой-нибудь базис пространства V. По определению $\chi(\varphi,t) = \det(M_v(\varphi) - tE)$. Запишем в этот многочлен в виде

$$\chi(\varphi,t) = \sum_{i=0}^{n} c_i t^i.$$

По теореме 12.2 получаем $\sum_{i=0}^{n} c_i M_v(\varphi)^i = 0$. Пользуясь формулами 20, 21 и 23, получаем

$$M_v\left(\sum_{i=0}^{n} c_i \varphi^i\right) = 0.$$

Так как матрица оператора равна нулю тогда и только тогда, когда сам оператор равен нулю, то мы получаем требуемое равенство

$$\sum_{i=0}^{n} c_i \varphi^i = 0.$$

□

12.2. **Диагонализируемые операторы.** Пусть φ — оператор на конечномерном векторном пространстве V над полем $\mathbb{F}$. Рассмотрим многочлен $f \in \mathbb{F}[t]$

$$f(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0,$$

где $a_0, \ldots, a_n \in \mathbb{F}$. Как мы видели выше (на примере матриц), мы можем вместо переменной t подставить оператор φ в многочлене f, то есть рассмотреть оператор

$$f(\varphi) = a_n \varphi^n + a_{n-1} \varphi^{n-1} + \cdots + a_1 \varphi + a_0 \,\mathrm{id},$$

где $\mathrm{id} : V \to V$ — тождественный оператор. Если $f(\varphi) = 0$, то мы говорим, что f *аннулирует* оператор φ. Следствие 12.3 показывает, что характеристический многочлен аннулирует свой оператор.

Определение 12.4. *Ненулевой многочлен, аннулирующий оператор φ, со старшим коэффициентом* 1 *минимальной возможной степени называется минимальным многочленом оператора φ.*

Заметим, что минимальным многочленом нулевого оператора является многочлен 1.

Лемма 12.5. (1) *Минимальным многочлен для любого оператора существует и единственный.*

(2) *Минимальный многочлен делит любой аннулирующий многочлен.*

Доказательство. (1) Для доказательства существования достаточно найти хотя бы один ненулевой многочлен, аннулирующий оператор φ, со старшим коэффициентом 1. Согласно следствию 12.3 и теореме 12.1 получаем, что таким многочленом является характеристический многочлен операторы φ, умноженный на $(-1)^n$, где n – размерность пространства, на котором действует φ.

Предположим теперь, что f и g — два различных минимальных многочлена для φ. По определению эти многочлены имеют одинаковые степени. Вычитая, получаем, что $f-g$ тоже аннулирует φ. Пусть a — старший коэффициент многочлена $f-g$. Мы получили, что $(f-g)/a$ — ненулевой многочлен, аннулирующий φ, степени строго меньше чем f. Мы приходим к противоречию с определением минимального многочлена.

(2) Пусть f и g — соответственно минимальный и произвольный аннулирующий многочлены оператора φ. Выполним деление с остатком $g = qf + r$ (см. теорему 4.3). В этом случае многочлен r тоже аннулирует φ, но имеет меньшую степень чем f. Если $r \neq 0$, то, деля его на его старший коэффициент, приходим к противоречию как в части (1). □

Определение 12.6. *Оператор φ на конечномерном векторном пространстве V над полем $\mathbb{F}$ называется диагонализируемым в базисе $(v_1, \ldots, v_n)$ пространства V, если $\varphi(v_i) = \lambda_i v_i$ для соответствующих элементов $\lambda_i \in \mathbb{F}$. Оператор называется диагонализируемым, если он диагонализируем в каком-нибудь базисе.*

Лемма 12.7. *Пусть φ — оператор на конечномерном векторном пространстве V, аннулируемый многочленом вида $(x-\alpha_1)\cdots(x-\alpha_m)$. Тогда существует некоторый базис $(v_1, \ldots, v_n)$ пространства V такой, что $\varphi(v_i) = \lambda_i v_i$ для соответствующих скаляров λ_i и $\{\lambda_1, \ldots, \lambda_n\} \subset \{\alpha_1, \ldots, \alpha_m\}$. В частности φ диагонализируем.*

Доказательство. Мы докажем лемму индукцией по m. База индукции $m = 1$ вполне очевидна. Действительно, в этом случае $\varphi = \alpha_1 \operatorname{id}$ и нам подходит любой базис пространства V. В обозначениях формулировки леммы мы должны положить $n = \dim V$ и $\lambda_1 = \cdots = \lambda_n = \alpha_1$. В этом случае, очевидно, $\{\lambda_1, \ldots, \lambda_n\} = \{\alpha_1\}$.

Теперь предположим, что $m > 1$ и лемма верна для аннулирующих многочленов меньшей степени. Положим $U = \operatorname{im}(\varphi - \alpha_m \operatorname{id})$. Тогда $\varphi(U) \subset U$. Действительно, пусть $u \in U$. По определению $u = \varphi(v) - \alpha v$ для некоторого $v \in V$. Получаем

$$\varphi(u) = \varphi^2(v) - \alpha\varphi(v) = (\varphi - \alpha_m \operatorname{id})(\varphi(v)) \in U.$$

Кроме того многочлен $(t - \alpha_1) \cdots (t - \alpha_{m-1})$ аннулирует $\varphi|_U$. Следовательно, по предположению индукции, применённому к ограничению $\varphi|_U : U \to U$, получаем, что существует некоторый базис $(u_1, \ldots, u_k)$ пространства U такой, что $\varphi(u_i) = \lambda_i u_i$ для некоторых скаляров λ_i причём $\{\lambda_1, \ldots, \lambda_k\} \subset \{\alpha_1, \ldots, \alpha_m\}$.

Мы утверждаем, что $\ker(\varphi - \alpha_m \operatorname{id}) \cap U = 0$. Действительно, пусть v — произвольный элемент этого пересечения. Так как $v \in \ker(\varphi - \alpha_m \operatorname{id})$, то $\varphi(v) = \alpha_m v$. Отсюда и из того, что $v \in U$, получаем (см. вычисление $f(\varphi)(v_i)$ в первой части доказательства)

$$0 = (\varphi - \alpha_1 \operatorname{id}) \cdots (\varphi - \alpha_{m-1} \operatorname{id})(v) = (\alpha_m - \alpha_1) \cdots (\alpha_m - \alpha_{m-1})v.$$

Сокращая на ненулевой скаляр $(\alpha_m - \alpha_1) \cdots (\alpha_m - \alpha_{m-1})$, получаем $v = 0$. Положим $U = \operatorname{im}(\varphi - \alpha_m \operatorname{id})$. Тогда $\varphi(U) \subset U$. Действительно, пусть $u \in U$. По определению $u = \varphi(v) - \alpha v$ для некоторого $v \in V$. Получаем

$$\varphi(u) = \varphi^2(v) - \alpha\varphi(v) = (\varphi - \alpha_m \operatorname{id})(\varphi(v)) \in U. \tag{39}$$

По лемме 10.2 получаем

$$\dim(\ker(\varphi - \alpha_m \operatorname{id}) \oplus U) = \dim \ker(\varphi - \alpha_m \operatorname{id}) + \dim \operatorname{im}(\varphi - \alpha_m \operatorname{id}) = \dim V.$$

Согласно ... получаем $\ker(\varphi - \alpha_m \operatorname{id}) \oplus U = V$. Остаётся только выбрать некоторый базис $v_1, \ldots, v_{n-k}$ пространства $\ker(\varphi - \alpha_m \operatorname{id})$ и положить $v_{n-k+1} = u_1, \ldots, v_n = u_k$. □

Теорема 12.8. *Оператор диагонализируем тогда и только тогда, когда его минимальный многочлен имеет вид $(x - \alpha_1) \cdots (x - \alpha_m)$ для попарно различных скаляров $\alpha_1, \ldots, \alpha_m$.*

Доказательство. Предположим сначала, что φ диагонализируемый оператор на пространстве V. Пусть $(v_1, \ldots, v_n)$ такой базис пространства V, что $\varphi(v_i) = \lambda_i v_i$. Выберем такие попарно различные скаляры $\alpha_1, \ldots, \alpha_m$, что

$$\{\alpha_1, \ldots, \alpha_m\} = \{\lambda_1, \ldots, \lambda_n\}.$$

Положим $f = (t - \alpha_1) \cdots (t - \alpha_m)$. Мы утверждаем, что f аннулирует оператор φ. Действительно,

$$\begin{aligned} f(\varphi)(v_i) = (\varphi - \alpha_1 \operatorname{id}) \cdots (\varphi - \alpha_m \operatorname{id})(v_i) = (\varphi - \alpha_1 \operatorname{id}) \cdots (\varphi - \alpha_{m-1} \operatorname{id})(\varphi(v_i) - \alpha_m v_i) = \\ = (\varphi - \alpha_1 \operatorname{id}) \cdots (\varphi - \alpha_{m-1} \operatorname{id})((\lambda_i - \alpha_m)v_i) = (\lambda_i - \alpha_m)(\varphi - \alpha_1 \operatorname{id}) \cdots (\varphi - \alpha_{m-1} \operatorname{id})(v_i) = \\ \cdots = (\lambda_i - \alpha_1) \cdots (\lambda_i - \alpha_m)v_i = 0. \end{aligned}$$

Последнее равенство выполнено так как $\lambda_i = \alpha_j$ для некоторого j. Так как $f(\varphi)$ переводит в ноль элементы некоторого базиса, то $f(\varphi) = 0$.

Предположим теперь, что g — многочлен со старшим коэффициентом 1, аннулирующий φ, степени $k < m$. Запишем

$$g = t^k + a_{k-1}t^{k-1} + \cdots + a_1 t + a_0.$$

Отсюда получаем

$$0 = g(\varphi)(v_i) = (\varphi^k + a_{k-1}\varphi^{k-1} + \cdots + a_1\varphi + a_0\,\mathrm{id})(v_i) = (\lambda_i^k + a_{k-1}\lambda_i^{k-1} + \cdots + a_1\lambda_i + a_0)v_i.$$

Таким образом, мы получаем, что $\alpha_1, \ldots, \alpha_m$ — корни g, что противоречит лемме 4.5.

Вторая часть следует из леммы 12.7. □

12.3. **Сопряжённые операторы в евклидовых пространствах.**

Определение 12.9. *Пусть E — евклидово пространство со скалярным произведением $(\cdot,\cdot)$ и φ и ψ — операторы на E. Мы будем говорить, что эти операторы сопряжённые, если $(\varphi(v), u) = (v, \psi(u))$ для любых $u, v \in E$.*

Лемма 12.10. *Пусть $e = (e_1, \ldots, e_n)$ — ортонормированный базис евклидова пространства E. Операторы φ и ψ на пространстве E сопряжены тогда и только тогда, когда $M_e(\psi) = M_e(\varphi)^T$.*

Доказательство. Запишем

$$M_e(\varphi) = \begin{pmatrix} \varphi_{1,1} & \cdots & \varphi_{1,n} \\ \vdots & \ddots & \vdots \\ \varphi_{m,1} & \cdots & \varphi_{m,n} \end{pmatrix}, \quad M_e(\psi) = \begin{pmatrix} \psi_{1,1} & \cdots & \psi_{1,n} \\ \vdots & \ddots & \vdots \\ \psi_{m,1} & \cdots & \psi_{m,n} \end{pmatrix}$$

Тогда

$$\varphi(e_j) = \sum_{i=1}^{m} \varphi_{i,j} e_i, \quad \psi(e_k) = \sum_{l=1}^{m} \psi_{l,k} e_l.$$

Предположим сначала, что операторы φ и ψ сопряжены. Получаем

$$(\varphi(e_j), e_k) = \left(\sum_{i=1}^{m} \varphi_{i,j} e_i, e_k \right) = \sum_{i=1}^{m} \varphi_{i,j}(e_i, e_k) = \varphi_{k,j}.$$

С другой стороны,

$$(e_j, \psi(e_k)) = \sum_{l=1}^{m} \psi_{l,k}(e_j, e_l) = \psi_{j,k}.$$

Так как $(\varphi(e_j), e_k) = (e_j, \psi(e_k))$, то $\varphi_{k,j} = \psi_{j,k}$. Это и означает, что $M_e(\psi) = M_e(\varphi)^T$.

Предположим теперь, что $M_e(\psi) = M_e(\varphi)^T$. Рассуждения выше, прочитанные в обратном порядке, означают, что

$$(\varphi(e_j), e_k) = (e_j, \psi(e_k)). \tag{40}$$

Возьмём теперь два произвольных вектора $v, u \in E$ и разложим их по базису

$$v = \sum_{j=1}^{n} a_j e_j, \quad u = \sum_{k=1}^{n} b_k e_k.$$

Отсюда получаем

$$(\varphi(v), u) = \left(\sum_{j=1}^{n} a_j \varphi(e_j), \sum_{k=1}^{n} b_k e_k \right) = \sum_{j=1}^{n} \sum_{k=1}^{n} a_j b_k (\varphi(e_j), e_k) =$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{n} a_j b_k (e_j, \psi(e_k)) = \left(\sum_{j=1}^{n} a_j e_j, \sum_{k=1}^{n} b_k \psi(e_k) \right) = (v, \psi(u)).$$

□

Оператор сопряжённый сам себе (это означает, что φ и φ — сопряжённые операторы) называется *самосопряжённым*. Согласно предыдущей лемме самосопряжённые операторы и только они имеют симметрическую матрицу в любом ортонормированном базисе. Более того для того, чтобы оператор был самосопряжённым достаточно, чтобы он имел симметрическую матрицу в одном из ортонормированных базисов.

12.4. **Сопряжённые операторы в эрмитовых пространствах.**

12.5. **Диагонализируемость самосопряжённого оператора.** Сначала мы докажем следующее вспомогательное утверждение.

Лемма 12.11. *Пусть E — конечномерное ненулевое евклидово пространство и φ — самосопряжённый оператор на E. Тогда φ имеет хотя бы один собственный вектор.*

Доказательство. Пусть $v = (v_1, \dots, v_n)$ — ортонормированный базис евклидова пространства E. По лемме 12.10 матрица $M_e(\varphi)$ симметрическая. Рассмотрим эрмитово пространство $\mathbb{C}^n$ со стандартным скалярным произведением $(\cdot|\cdot)$ и оператор ψ на нём, заданный формулой $\psi(v) = M_e(\varphi)v$, где $v \in \mathbb{C}^n$ воспринимается как столбец высоты n. Согласно ... матрица $M_e(\varphi)$ является матрицей оператора ψ в стандартном базисе пространства $\mathbb{C}^n$:

$$e_1 = (1, 0, \dots, 0), \quad e_2 = (0, 1, 0, \dots, 0), \quad \dots, e_n = (0, \dots, 0, 1).$$

Матрица $M_e(\varphi)$ вещественная, поэтому

$$\overline{M_e(\varphi)}^T = M_e(\varphi)^T = M_e(\varphi).$$

По ... оператор ψ самосопряжённый. Как любой оператор на конечномерном ненулевом комплексном пространстве ψ имеет собственный вектор, скажем, v. Следовательно, $\psi(v) = \lambda v$ для соответствующего $\lambda \in \mathbb{C}$. Отсюда по свойствам **??** и **??** получаем

$$\lambda(v|v) = (\lambda v|v) = (\psi(v)|v) = (v|\psi(v)) = (v|\lambda v) = \bar{\lambda}(v|v).$$

Сокращая на ненулевое число $(v|v)$, получаем $\lambda = \bar{\lambda}$. Следовательно, λ — вещественное число.

По ... получаем $\det(M_e(\varphi) - \lambda E) = 0$. Опять применяя ..., получаем что λ — собственное значение оператора φ, а следовательно существует и собственный вектор оператора φ (соответствующий этому собственному значению). □

Теорема 12.12. *Пусть E — конечномерное евклидово пространство и φ — самосопряжённый оператор на E. Тогда оператор φ диагонализируем в некотором ортонормированном базисе.*

Доказательство. Применим индукцию по размерности E. В случае $\dim E = 0$ утверждение очевидно. Пусть теперь $\dim E > 0$. Тогда по лемме 12.11 существует собственный вектор v оператора φ. Запишем $\varphi(v) = \lambda v$ для соответствующего $\lambda \in \mathbb{C}$. Положим $U = \langle v \rangle^\perp$. Мы утверждаем, что U инвариантно относительно φ. Действительно, пусть $u \in U$. Тогда

$$(\varphi(u), v) = (u, \varphi(v)) = (u, \lambda v) = \lambda(u, v) = 0.$$

Следовательно, мы можем рассмотреть ограничение $\varphi|_U$. Согласно части (2) теоремы 7.5 мы получаем $E = \langle v \rangle \oplus U$. По лемме 6.16 мы получаем

$$\dim E = \dim\langle v \rangle + \dim U = 1 + \dim U.$$

Отсюда $\dim U = \dim E - 1 < \dim E$. Следовательно, к ограничению $\varphi|_U$ применимо предположение индукции и этот оператор диагонализируем, скажем, в ортонормированном базисе $e_1, \dots, e_n$. Тогда, как легко видеть, оператор φ диагонализируем в ортонормированном базисе $e_1, \dots, e_n, v$. □

12.6. **Существование Жордановой нормальной формы.**

Лемма 12.13. *Пусть $\varphi : V \to V$ — оператор на конечномерном пространстве V над алгебраически замкнутым полем $\mathbb{F}$ и $\lambda_1, \dots, \lambda_k$ — попарно различные собственные значения оператора φ. Тогда существую натуральные числа $n_1, \dots, n_k$ такие, что*

$$V = \ker(\varphi - \lambda_1 \operatorname{id})^{n_1} \oplus \ker(\varphi - \lambda_2 \operatorname{id})^{n_2} \oplus \dots \oplus \ker(\varphi - \lambda_k \operatorname{id})^{n_k}.$$

Доказательство. Проведём индукцию по количеству собственных значений оператора φ. Случай $k = 0$ соответствует нулевому пространству V. В этом случае утверждение верно, так как пустая прямая сумма равна нулевому подпространству.

Предположим теперь, что $V \neq 0$. В этом случае $k > 0$ в силу алгебраической замкнутости поля $\mathbb{C}$. Рассмотрим подпространства

$$W_i = \ker(\varphi - \lambda_1 \operatorname{id})^i.$$

Мы получаем $W_i \subset W_{i+1}$. Действительно, пусть $w \in W_i$. Тогда $(\varphi - \lambda_1 \operatorname{id})^i(w) = 0$. Применяя, к этому равенству $\varphi - \lambda_1 \operatorname{id}$, получаем $(\varphi - \lambda_1 \operatorname{id})^{i+1}(w) = 0$. Следовательно, $w \in W_{i+1}$. Так как V конечномерно, то существует такое n_1, что

$$W_{n_1} = W_{n_1+1} = W_{n_1+2} = \dots .$$

Мы утверждаем, что

$$W_{n_1} \cap \operatorname{im}(\varphi - \lambda_1 \operatorname{id})^{n_1} = 0.$$

Действительно, пусть v — элемент из этого пересечения. Тогда $v = (\varphi - \lambda_1 \operatorname{id})^{n_1}(u)$ для некоторого $u \in V$, так как $v \in \operatorname{im}(\varphi - \lambda_1 \operatorname{id})^{n_1}$. С другой стороны, так как $v \in \ker(\varphi - \lambda_1 \operatorname{id})^{n_1}$, то

$$0 = (\varphi - \lambda_1 \operatorname{id})^{n_1}(v) = (\varphi - \lambda_1 \operatorname{id})^{2n_1}(u).$$

Отсюда $u \in W_{2n_1} = W_{n_1}$ и $0 = (\varphi - \lambda_1 \operatorname{id})^{n_1}(u) = v$.

По лемме 10.2, получаем (см. доказательство леммы 12.7)

$$V = W_{n_1} \oplus U_{n_1}, \tag{41}$$

где $U_{n_1} = \operatorname{im}(\varphi - \lambda_1 \operatorname{id})^{n_1}$. Заметим, что подпространства W_{n_1} и U_{n_1} инвариантны относительно φ. Эти утверждения доказываются вычислением, аналогичным вычислению (39).

Рассмотрим оператор ψ, являющийся оператором ограничения φ на U_{n_1}. Пусть λ — его собственное значение с собственным вектором $v \neq 0$. Тогда λ — также собственное значение оператора φ. Поэтому $\lambda = \lambda_i$ для некоторого $i = 1, \dots, k$. Если $i = 1$, то $v \in W_1 \subset W_{n_1}$. Так как $v \in U_{n_1}$, то мы получаем противоречие $v = 0$. Таким образом, мы доказали, что $i > 1$. С другой стороны, пусть $v_i \neq 0$ собственный вектор оператора φ с собственным значением λ_i, где $i > 1$. В силу формулы (41) существует разложение

$$v_i = w_i + u_i, \tag{42}$$

где $w_i \in W_{n_1}$ и $u_i \in U_{n_1}$. Мы получаем

$$(\varphi - \lambda_1 \operatorname{id})^{n_1}(v_i) = (\lambda_i - \lambda_1)^{n_1} v_i \neq 0,$$

Следовательно, $v_i \notin W_{n_1}$. Поэтому $u_i \neq 0$. Умножая, равенство (42) на λ_i, получаем

$$\lambda_i v_i = \lambda_i w_i + \lambda_i u_i,$$

а применяя к равенству (42) оператор φ, получаем

$$\varphi(v_i) = \varphi(w_i) + \varphi(u_i).$$

Так как левые части этого равенства равны и сумма (41) прямая, то $\varphi(u_i) = \lambda_i u_i$. Это равенство можно переписать в виде $\psi(u_i) = \lambda_i u_i$. Таким образом λ_i — собственное значение оператора ψ.

И так мы можем применить индуктивное предположение к оператору ψ. Получаем

$$U_{n_1} = (\psi - \lambda_2 \operatorname{id})^{n_2} \oplus \cdots \oplus (\psi - \lambda_k \operatorname{id})^{n_k}.$$

для некоторых натуральных чисел $n_2, \ldots, n_k$. Подставляя это равенство в формулу (41), получаем

$$V = \ker(\varphi - \lambda_1 \operatorname{id})^{n_1} \oplus \ker(\psi - \lambda_2 \operatorname{id})^{n_2} \oplus \cdots \oplus \ker(\psi - \lambda_k \operatorname{id})^{n_k}.$$

Нам остаётся доказать, что

$$\ker(\psi - \lambda_i \operatorname{id})^{n_i} = \ker(\varphi - \lambda_i \operatorname{id})^{n_i}$$

для любого $i = 2, \ldots, k$. По определению левая часть содержится в правой. Пусть теперь v — вектор из правой части. Рассмотрим разложение $v = w + u$, где $w \in W_{n_1}$ и $u \in U_{n_1}$. Предположим, что $w \neq 0$. Применяя оператор $(\psi - \lambda_i \operatorname{id})^{n_i}$, получаем

$$0 = (\varphi - \lambda_i \operatorname{id})^{n_i}(v) = (\varphi - \lambda_i \operatorname{id})^{n_i}(w) + (\varphi - \lambda_i \operatorname{id})^{n_i}(u).$$

Так как W_{n_1} и U_{n_1} инвариантны относительно φ, то

$$(\varphi - \lambda_i \operatorname{id})^{n_i}(w) \in W_{n_1}, \quad (\varphi - \lambda_i \operatorname{id})^{n_i}(u) \in U_{n_1}.$$

В силу того, что сумма (41) прямая, получаем $(\varphi - \lambda_i \operatorname{id})^{n_i}(w) = 0$. Таким образом, мы получаем, что

$$\ker(\varphi - \lambda_1 \operatorname{id})^{n_1} \cap \ker(\varphi - \lambda_i \operatorname{id})^{n_i} \neq 0.$$

Это пространство инвариантно относительно φ. Поэтому существует его ненулевой собственный вектор h из этого пространства с собственным значением λ. Мы получаем

$$(\varphi - \lambda_1 \operatorname{id})^{n_1}(h) = (\lambda - \lambda_1)^{n_1} h, \quad (\varphi - \lambda_i \operatorname{id})^{n_i}(h) = (\lambda - \lambda_i)^{n_i} h.$$

Так как $\lambda_1 \neq \lambda_i$, то оба вектора в правых частях равенств выше не могут быть одновременно равны нулю. Приходим к противоречию. Таким образом $w = 0$ и $v = u \in U_{n_1}$. Отсюда $v \in \ker(\psi - \lambda_i \operatorname{id})^{n_i}$. □

Лемма 12.14 (О нильпотентном операторе)**.** *Пусть $\varphi : V \to V$ — оператор на конечномерном пространстве V над алгебраически замкнутым полем $\mathbb{F}$ такой, что $T^n = 0$ для некоторого натурального n. Тогда существуют такие векторы $u_1, \ldots, u_k$ и натуральные числа $a_1, \ldots, a_k$, что $T^{a_1}(u_1) = \cdots = T^{a_k}(u_k) = 0$ и набор векторов*

$$(u_1, T(u_1), \ldots, T^{a_1-1}(u_1), u_2, T(u_2), \ldots, T^{a_2-1}(u_2), \ldots, u_k, T(u_k), \ldots, T^{a_k-1}(u_k))$$

образует базис пространства V.

Доказательство. Проведём индукцию по размерности пространства V. Если $\dim V = 0$, то $V = 0$ и мы можем взять пустой набор векторов, что соответствует $k = 0$.

Предположим, что $\dim V > 0$. Если $\ker T$ был бы равен нулю, то мы получили бы, что T — взаимнооднозначное отображение и как следствие $T^n \neq 0$. Теперь предположим, что $\ker T \neq 0$. Отсюда и из ... получаем $\dim \operatorname{im} T = \dim V - \dim \ker T < \dim V$. Таким образом, мы можем применить индуктивное предположение к пространству $\operatorname{im} T$. Пусть $v_1 \ldots, v_l \in \operatorname{im} T$ и $b_1, \ldots, b_l \in \mathbb{N}$ такие, что $T^{b_1}(v_1) = \cdots = T^{b_k}(v_l) = 0$ и набор

$$(v_1, T(v_1), \ldots, T^{b_1-1}(v_1), v_2, T(v_2), \ldots, T^{b_2-1}(v_2), \ldots, v_l, T(v_l), \ldots, T^{b_l-1}(v_l))$$

образует базис пространства $\operatorname{im} T$. Выберем векторы $w_1, \ldots, w_l \in V$ так, что $T(w_i) = v_i$ для любого $i = 1, \ldots, l$. В этом случае $T^{b_i+1}(w_i) = T^{b_i}(v_i) = 0$. Так как

$$(T^{b_1}(w_1), \ldots, T^{b_l}(w_l)) = (T^{b_1-1}(v_1), \ldots, T^{b_l-1}(v_l)),$$

то этот набор линейно независим как поднабор базиса. Кроме того, все векторы принадлежат ядру $\ker T$. Согласно ... мы можем дополнить этот набор до базиса $(T^{b_1}(w_1), \dots, T^{b_l}(w_l), z_1, \dots$ пространства $\ker T$. Мы утверждаем, что набор

$$(w_1, T(w_1), \dots, T^{b_1}(w_1), w_2, T(w_2), \dots, T^{b_2}(w_2), \dots, w_l, T(w_l), \dots, T^{b_l}(w_l), z_1, \dots, z_m)$$

образует базис пространства V.

Возьмём сначала произвольный вектор $v \in V$. Рассмотрим разложение

$$T(v) = \alpha_{1,0}v_1+\alpha_{1,1}T(v_1)+\dots+\alpha_{1,b_1-1}T^{b_1-1}(v_1)+\dots+\alpha_{l,0}v_l+\alpha_{l,1}T(v_l)+\dots+\alpha_{l,b_l-1}T^{b_l-1}(v_l).$$

Отсюда следует, что вектор

$$h = v-(\alpha_{1,0}w_1+\alpha_{1,1}T(w_1)+\dots+\alpha_{1,b_1-1}T^{b_1-1}(w_1)+\dots+\alpha_{l,0}w_l+\alpha_{l,1}T(w_l)+\dots+\alpha_{l,b_l-1}T^{b_l-1}(w_l)).$$

принадлежит пространству $\ker T$. Записывая разложение

$$h = \alpha_{1,b_1}T^{b_1}(w_1) + \dots + \alpha_{l,b_l}T^{b_l}(w_l) + \beta_1 z_1 + \dots \beta_m z_m,$$

и перенося все векторы кроме v в правую сторону, получим представление вектора v в виде линейной комбинации.

□

13. Квадратичные формы

13.1. **Определение.** Сначала мы дадим определение квадратичной формы над произвольным полем, а затем сосредоточимся над случаем, когда это поле имеет характеристику 2.

Определение 13.1. *Пусть $\mathbb{F}$ — поле и V — векторное пространство над $\mathbb{F}$. Квадратичной формой на V называется отображение $q : V \to \mathbb{F}$ такое, что $q(cv) = c^2q(v)$ для любых $c \in \mathbb{F}$ и $v \in V$ и отображение $(\cdot,\cdot) : V \times V \to \mathbb{F}$, заданное формулой $(u, v) = q(u + v) - q(u) - q(v)$, билинейное.*

Заметим, что $(u, v) = (v, u)$ вне зависимости от того, билинейная ли она. Форму $(\cdot,\cdot)$ назовём ассоциированной с q.

Мы можем описать квадратичные формы на конечномерных векторных пространствах более явно.

Лемма 13.2. *Пусть $(v_1, \dots, v_n)$ — базис векторного пространства V. Квадратичные формы и только они определяются формулой*

$$q(x_1v_1 + \dots + x_nv_n) = \sum_{1\leqslant i\leqslant j\leqslant n} a_{i,j}x_ix_j \quad \forall x_1, \dots, x_n \in \mathbb{F} \tag{43}$$

для некоторых $a_{i,j} \in \mathbb{F}$.

Доказательство. Проверим сначала, что отображение q, заданное таким образом, является квадратичной формой. Для произвольного $c \in \mathbb{F}$, получаем

$$q(c(x_1v_1 + \dots + x_nv_n)) = q((cx_1)v_1 + \dots + (cx_n)v_n) = \sum_{1\leqslant i\leqslant j\leqslant n} a_{i,j}(cx_i)(cx_j) =$$
$$= c^2 \sum_{1\leqslant i\leqslant j\leqslant n} a_{i,j}x_ix_j = c^2q(x_1v_1 + \dots + x_nv_n).$$

Теперь вычислим отображение $(v,u) = q(v+u) - q(v) - q(u)$. Для $v = x_1v_1 + \cdots + x_nv_n$ и $u = y_1v_1 + \cdots + y_nv_n$ Получаем

$$\begin{aligned}(v,u) = q(u+v) - q(u) - q(v) = q((x_1+y_1)v_1 + \cdots + (x_n+y_n)v_n) - \\ -q(x_1v_1 + \cdots + x_nv_n) - q(y_1v_1 + \cdots + y_nv_n) = \\ = \sum_{1\leqslant i\leqslant j\leqslant n} a_{i,j}(x_i+y_i)(x_j+y_j) - \sum_{1\leqslant i\leqslant j\leqslant n} a_{i,j}x_ix_j - \sum_{1\leqslant i\leqslant j\leqslant n} a_{i,j}y_iy_j \\ = \sum_{1\leqslant i\leqslant j\leqslant n} a_{i,j}(x_ix_j + y_ix_j + x_iy_j + y_iy_j) - \sum_{1\leqslant i\leqslant j\leqslant n} a_{i,j}x_ix_j - \sum_{1\leqslant i\leqslant j\leqslant n} a_{i,j}y_iy_j \\ = \sum_{1\leqslant i\leqslant j\leqslant n} a_{i,j}(y_ix_j + x_iy_j).\end{aligned} \tag{44}$$

Возьмём произвольный элемент $c \in \mathbb{F}$. Тогда получаем $cv = cx_1v_1 + \cdots + cx_nv_n$ и

$$(cv,u) = \sum_{1\leqslant i\leqslant j\leqslant n} a_{i,j}(y_icx_j + cx_iy_j) = c\sum_{1\leqslant i\leqslant j\leqslant n} a_{i,j}(y_ix_j + x_iy_j) = c(v,u).$$

Теперь возьмём ещё один вектор $v' = cx'_1v_1 + \cdots + cx'_nv_n$. Тогда $v + v' = (x_1+x'_1)v_1 + \cdots + (x_n+x'_n)v_n$. Отсюда получаем

$$\begin{aligned}(v+v',u) = \sum_{1\leqslant i\leqslant j\leqslant n} a_{i,j}(y_i(x_j+x'_j) + (x_i+x'_i)y_j) = \\ = \sum_{1\leqslant i\leqslant j\leqslant n} a_{i,j}(y_ix_j + x_iy_j) + \sum_{1\leqslant i\leqslant j\leqslant n} a_{i,j}(y_ix'_j + x'_iy_j) = (v,u) + (v',u).\end{aligned}$$

Линейность по второму аргументу отображения $(\cdot,\cdot)$ следует из симметричности этого отображения.

Теперь пусть $q : V \to \mathbb{F}$ — билинейная форма и $(\cdot,\cdot)$ — билинейная форма ассоциированная с q. Положим $a_{i,i} = q(v_i)$ и $a_{i,j} = (v_i,v_j)$ для $i < j$. Мы докажем индукцией по $k = 1,\ldots,n$, что

$$q(x_1v_1 + \cdots + x_kv_k) = \sum_{1\leqslant i\leqslant j\leqslant k} a_{i,j}x_ix_j.$$

Базой индукции является случай $k = 1$. Получаем

$$q(x_1v_1) = x_1^2q(v_1) = x_1^2a_{1,1} = a_{1,1}x_1^2 = \sum_{1\leqslant i\leqslant j\leqslant 1} a_{i,j}x_ix_j.$$

Теперь предположим, что $1 \leqslant k < n$ и что доказываемая формула верна для k. Используя билинейность формы $(\cdot,\cdot)$ и предположение индукции, мы получаем

$$\begin{aligned}q(x_1v_1 + \cdots + x_kv_k + x_{k+1}v_{k+1}) = \\ = (x_1v_1 + \cdots + x_kv_k, x_{k+1}v_{k+1}) + q(x_1v_1 + \cdots + x_kv_k) + q(x_{k+1}v_{k+1}) = \\ = \sum_{1\leqslant i\leqslant k} x_ix_{k+1}(v_i,v_{k+1}) + \sum_{1\leqslant i\leqslant j\leqslant k} a_{i,j}x_ix_j + x_{k+1}^2q(v_{k+1}) = \\ = \sum_{1\leqslant i\leqslant k} a_{i,k+1}x_ix_{k+1} + \sum_{1\leqslant i\leqslant j\leqslant k} a_{i,j}x_ix_j + a_{k+1,k+1}x_{k+1}^2 = \sum_{1\leqslant i\leqslant j\leqslant k+1} a_{i,j}x_ix_j.\end{aligned}$$

Требуемый результат получается при $k = n$. □

Заметим, что представление квадратичной формы q в виде (43) единственно. Действительно, сначала мы получаем

$$q(v_i) = q(0v_1 + \cdots + 0v_{i-1} + 1v_i + 0v_{i+1} + \cdots + 0v_n) = a_{i,i}.$$

Теперь применим формулу (44) в случае, когда $i < j$, $x_1 = \cdots = x_{i-1} = x_{i+1} = \cdots = x_n = 0$, $x_i = 1$ и $y_1 = \cdots = y_{j-1} = y_{j+1} = \cdots = y_n = 0$, $y_j = 1$. Получаем

$$(v_i, v_j) = a_{i,j}(y_i x_j + x_i y_j) = a_{i,j}.$$

Отождествляя матрицы размера 1×1 с элементами поля мы получаем

$$q(x_1 v_1 + \cdots + x_n v_n) = X^T A X, \tag{45}$$

где

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & a_{n,n} \end{pmatrix}.$$

Замечание. Для произвольной квадратной матрицы B с элементами из поля $\mathbb{F}$ отображение $q : V \to \mathbb{F}$, заданное формулой

$$q(x_1 v_1 + \cdots + x_n v_n) = X^T B X,$$

является квадратичной формой. В этом можно убедиться непосредственными вычислениями аналогичными вычислениям доказательства леммы 13.2. С другой стороны, можно заметить, что

$$X^T B X = X^T A X,$$

где

$$A = \begin{pmatrix} b_{1,1} & b_{1,2} + b_{2,1} & b_{1,3} + b_{3,1} & \cdots & b_{1,n} + b_{n,1} \\ 0 & b_{2,2} & b_{2,3} + b_{3,2} & \cdots & b_{2,n} + b_{n,2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & b_{n,n} \end{pmatrix}$$

Отсюда и из леммы 13.2 следует, что q является квадратичной формой.

13.2. **Представление с помощью симметрической матрицы.** Предположим теперь, что $\mathbb{F}$ — поле характеристики не равной двум. В представлении (45) мы использовали верхнетреугольную матрицу A. Однако нам удобнее использовать симметрическую матрицу. В данном случае мы можем добиться такого представления следующим приёмом. Транспонируя равенство (45) используя тот факт, что матрица размера $1 \times \times 1$ всегда симметрическая, получаем

$$q(x_1 v_1 + \cdots + x_n v_n) = q(x_1 v_1 + \cdots + x_n v_n)^T = (X^T A X)^T = X^T A^T (X^T)^T = X^T A^T X.$$

Теперь складывая это равенство с равенством (45), получакм

$$2q(x_1 v_1 + \cdots + x_n v_n) = X^T A X + X^T A^T X = X^T (A + A^T) X.$$

Деля на 2 (что мы можем сделать в поле характеристики не равной 2), получаем

$$q(x_1 v_1 + \cdots + x_n v_n) = X^T \frac{A + A^T}{2} X.$$

Теперь замечаем, что

$$\left(\frac{A + A^T}{2}\right)^T = \frac{A^T + A}{2} = \frac{A + A^T}{2}.$$

Следовательно, матрица $(A + A^T)/2$ симметрическая.

Лемма 13.3. *Пусть $(v_1, \dots, v_n)$ — базис векторного пространства V над полем $\mathbb{F}$ характеристики не равной двум. Представление квадратичной формы $q : V \to \mathbb{F}$ в виде*

$$q(x_1v_1 + \dots + x_nv_n) = X^TAX, \text{ где } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

с симметрической матрицей A существует и единственно.

Доказательство. Существование уже доказано непосредственно перед этой леммой. Поэтому докажем единственность. Запишем равенство из формулировки более подробно

$$q(x_1v_1 + \dots + x_nv_n) = \sum_{i,i=1}^{n} a_{i,j}x_ix_j = \sum_{i=1}^{n} a_{i,i}x_i^2 + \sum_{1 \leqslant i < j \leqslant n} 2a_{i,j}x_ix_j.$$

Отсюда получаем $q(v_i) = a_{i,i}$. Для $i < j$ получаем

$$(v_i, v_j) = q(v_i + v_j) - q(v_i) - q(v_j) = (a_{i,i} + a_{j,j} + 2a_{i,j}) - a_{i,i} - a_{j,j} = 2a_{i,j}.$$

Так как 2 обратима в $\mathbb{F}$, то отсюда мы можем вычислить $a_{i,j}$. □

Мы обозначим матрицу A из этой леммы относящуюся к базису $v = (v_1, \dots, v_n)$ через q_v. Наша следующая цель — понять, как меняется эта матрица при замене базиса.

Лемма 13.4. *Пусть $v = (v_1, \dots, v_n)$ и $u = (u_1, \dots, u_n)$ — базисы векторного пространства V над полем $\mathbb{F}$ характеристики не равной двум. Тогда*

$$q_v = (T_{u,v})^T q_u T_{u,v}$$

(напоминаем, что $T_{u,v}$ — матрица перехода от базиса u к базису v).

Доказательство. Пусть λ — произвольный вектор из V. Запишем его в обоих базисах

$$\lambda = x_1v_1 + \dots + x_nv_n, \qquad \lambda = y_1u_1 + \dots + y_nu_n.$$

Введём следующие обозначения для столбцов координат

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

По определению матриц q_v и q_u выполнены равенства

$$q(\lambda) = X^Tq_vX, \tag{46}$$

$$q(\lambda) = Y^Tq_uY, \tag{47}$$

С другой стороны по формуле (16) получаем $Y = T_{u,v}X$. Подставляя Y в формулу (47), получаем

$$q(\lambda) = (T_{u,v}X)^Tq_uT_{u,v}X = X^T(T_{u,v})^Tq_uT_{u,v}X.$$

Сравнивая получившееся выражение с формулой (46), замечая, что матрица $(T_{u,v})^Tq_uT_{u,v}$ симметрическая, и пользуясь единственностью леммы 13.3, получаем требуемое равенство. □

13.3. **Ортогональные матрицы.** Квадратная матрица P над полем $\mathbb{F}$ называется *ортогональной* тогда и только тогда, когда $P^TP = E$. Напомним, что E обозначает единичную матрицу такого же размера как P.

Лемма 13.5. (1) *Если матрица P ортогональная, то* $\det P = 1$ *или* $\det P = -1$.
(2) *Матрица P ортогональная тогда и только тогда, когда P обратима и $P^{-1} = P^T$.*
(3) *Матрица P ортогональная тогда и только тогда, когда $PP^T = E$.*

Доказательство. (1) Применяя теорему 11.19 и часть (1) теоремы 11.11, получаем

$$1 = \det(P^TP) = \det(P^T)\det(P) = \det(P)^2.$$

Заметим, что уравнение $x^2 = 1$ имеет только два решения $x = 1$ и $x = -1$ над любым полем. Это следует из разложения $x^2 - 1 = (x-1)(x+1)$. Следовательно, $\det P = 1$ или $\det P = -1$.

(2) Если P ортогональная матрица, то согласно части (1) получаем $\det P \neq 0$. По следствию 11.24, матрица P обратима. Умножая равенство $P^TP = E$ справа на P^{-1}, получаем

$$P^{-1} = EP^{-1} = (P^TP)P^{-1} = P^T(PP^{-1}) = P^TE = P^T.$$

Наоборот, пусть P обратима и $P^{-1} = P^T$. По определению обратной матрицы $P^TP = P^{-1}P = E$.

(3) Пусть матрица P ортогональна. Согласно части (2) получаем $P^{-1} = P^T$. Отсюда $PP^T = PP^{-1} = E$. Пусть, наоборот, $PP^T = E$. Переписывая наши рассуждения частей (1) и (2) задом наперёд, получаем, что P обратима и $P^{-1} = P^T$. Поэтому $P^TP = P^{-1}P = E$. □

Заметим, что уравнение $PP^T = E$ эквивалентно тому, что строки $r_1, \ldots, r_n$ матрицы P удовлетворяют соотношению $(r_i, r_j) = \delta_{i,j}$, где $\delta_{i,j}$ — символ Кронекера и $(\cdot,\cdot) : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ — стандартное произведение

$$(x_1, \ldots, x_n)(y_1, \ldots, y_n) = x_1y_1 + \cdots + x_ny_n.$$

Аналогично равенство $P^TP = E$ эквивалентно тому, что столбцы $c_1, \ldots, c_n$ матрицы P удовлетворяют соотношению $(c_i, c_j) = \delta_{i,j}$.

Нам потребуются ортогональные матрицы над полем вещественных чисел. В этом случае ортогональные матрицы можно описать, как матрицы перехода.

Лемма 13.6. (1) *Пусть $u = (u_1, \ldots, u_n)$ и $v = (v_1, \ldots, v_n)$ — два ортонормированных базиса евклидова пространства E. Тогда матрица перехода $T_{u,v}$ от базиса u к базису v ортогональная.*
(2) *Пусть $u = (u_1, \ldots, u_n)$ и $v = (v_1, \ldots, v_n)$ — два базиса евклидова пространства E такие, что и ортонормированный и матрица перехода $T_{u,v}$ от базиса u к базису v ортогональная. Тогда v тоже ортонормированный.*

Доказательство. (1) Запишем матрицу $T_{u,v}$ в виде (13). При этом выполнено уравнение (12). В силу ортонормированности базисов u и v получаем

$$\delta_{j,k} = (v_j, v_k) = \left(\sum_{i=1}^n a_{i,j}u_i, \sum_{l=1}^n a_{l,k}u_l\right) = \sum_{i=1}^n\sum_{l=1}^n a_{i,j}a_{l,k}(u_i, u_l) =$$

$$= \sum_{i=1}^n\sum_{l=1}^n a_{i,j}a_{l,k}\delta_{i,l} = \sum_{i=1}^n a_{i,j}a_{i,k}.$$

Это уравнение эквивалентно тому, что $(T_{u,v})^T T_{u,v} = E$. Следовательно, матрица $T_{u,v}$ ортогональная.

(2) Если прочитать цепочку равенств выше справа налево, то мы получим $\delta_{j,k} = (v_j, v_k)$, что означает ортонормированность базиса v. □

13.4. **Канонический вид вещественной квадратичной формы в евклидовом пространстве.**

Теорема 13.7. *Пусть E — конечномерное евклидово пространство и $q : E \to \mathbb{R}$ — квадратичная форма. Тогда существует такой ортонормированный базис $e_1, \ldots, e_n$ пространства E и вещественные числа $\lambda_1, \ldots, \lambda_n$, что*

$$q(x_1 e_1 + \ldots + x_n e_n) = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2$$

для любых $x_1, \ldots, x_n \in \mathbb{R}$. Числа $\lambda_1, \ldots, \lambda_n$ определены однозначно с точностью до перестановки.

Доказательство. Обозначим скалярное произведение на E через $(\cdot, \cdot)$. Выберем произвольный ортонормированный базис $v = (v_1, \ldots, v_n)$ пространства E. Запишем матрицу квадратичной формы q в этом базисе в виде

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}.$$

Напомним, что $a_{i,j} = a_{j,i}$.

Пусть φ — оператор на E с матрицей q_v. По лемме 12.10 оператор φ самосопряжённый. Мы утверждаем, что

$$(\varphi(u), u) = q(u) \tag{48}$$

для любого вектора $u \in E$. Действительно, по построению

$$\varphi(v_j) = \sum_{i=1}^{m} a_{i,j} v_i.$$

Используя ортонормированность базиса v, получаем $(\varphi(v_j), v_k) = a_{k,j}$. Пусть теперь $u = x_1 v_1 + \cdots + x_n v_n$. Тогда

$$(\varphi(u), u) = \left(\sum_{j=1}^{k} x_j \varphi(v_j), \sum_{k=1}^{n} v_k \right) = \sum_{j=1}^{k} \sum_{k=1}^{n} x_j x_k (\varphi(v_j), v_k) = \sum_{j=1}^{k} \sum_{k=1}^{n} x_j x_k a_{k,j} = q(u).$$

Эти равенство (48) доказано. По теореме 12.12 оператор φ диагонализируем в некотором ортонормированном базисе $e_1, \ldots, e_n$. Пусть $\varphi(e_i) = \lambda_i e_i$. Согласно формуле (48) мы получаем

$$\begin{aligned} q(x_1 e_1 + \ldots + x_n e_n) &= (\varphi(x_1 e_1 + \ldots + x_n e_n), x_1 e_1 + \ldots + x_n e_n) = \\ &= (x_1 \varphi(e_1) + \ldots + x_n \varphi(e_n), x_1 e_1 + \ldots + x_n e_n) = \\ &= (x_1 \lambda_1 e_1 + \ldots + x_n \lambda_n e_n, x_1 e_1 + \ldots + x_n e_n) = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2. \end{aligned}$$

Заметим, что мы получили

$$q_e = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Для доказательства единственности предположим, что существует другой ортонормированный базис $e_1, \ldots, e_n$ пространства E и вещественные числа $\lambda'_1, \ldots, \lambda'_n$, что

$$q(x_1 e'_1 + \ldots + x_n e'_n) = \lambda'_1 x_1^2 + \cdots + \lambda'_n x_n^2$$

для любых $x_1, \ldots, x_n \in \mathbb{R}$. В этом случае

$$q_{e'} = \begin{pmatrix} \lambda'_1 & 0 & \cdots & 0 \\ 0 & \lambda'_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda'_n \end{pmatrix}.$$

По лемме 13.4 получаем $q_{e'} = (T_{e,e'})^T q_e T_{e,e'}$. Согласно части (1) леммы 13.6 матрица перехода $T_{e,e'}$ ортогональная. Следовательно, $q_{e'} = (T_{e,e'})^{-1} q_e T_{e,e'}$ и матрицы $q_{e'}$ и q_e сопряжены. Согласно ... эти матрицы имеют одинаковые с точностью до перестановки собственные значения, что и доказывает утверждение единственности этой теоремы. □

Конструкция первой части доказательства позволяет выписать формулы для координат. Пусть $v \in E$ и $v = x_1 v_1 + \cdots + x_n v_n$ и $v = y_1 e_1 + \cdots + y_n e_n$. Запишем координаты в столбцы

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

по формуле 16 получаем $X = T_{v,e} Y$. Отсюда

$$u = vX,\ u = eY \Longrightarrow Y = (T_{v,e})^T X,\ q(u) = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2. \tag{49}$$

Следствие 13.8. *Пусть E — конечномерное евклидово пространство и $q : E \to \mathbb{R}$ — квадратичная форма. Тогда существует такой ортогональный базис $(f_1, \ldots, f_n)$ пространства E и целые числа i и j, что $1 \leqslant i \leqslant j \leqslant n$ и*

$$q(x_1 f_1 + \ldots + x_n f_n) = x_1^2 + \cdots + x_i^2 - x_{i+1}^2 - \cdots - x_j^2$$

для любых $x_1, \ldots, x_n \in \mathbb{R}$. Числа i и j определены однозначно.

Доказательство. Пусть $(e_1, \ldots, e_n)$ — такой ортонормированный базис пространства E, который требуется в теореме 13.7. Переставляя, если необходимо, элементы базиса, мы можем добиться того, что

$$\lambda_1, \ldots, \lambda_i > 0, \quad \lambda_{i+1}, \ldots, \lambda_j < 0, \quad \lambda_{j+1} = \cdots = \lambda_n = 0.$$

Положим

$$f_1 = \frac{e_1}{\sqrt{\lambda_1}}, \ldots, f_i = \frac{e_i}{\sqrt{\lambda_i}},\ f_{i+1} = \frac{e_{i+1}}{\sqrt{-\lambda_{i+1}}}, \ldots, f_j = \frac{e_j}{\sqrt{-\lambda_j}},\ f_{j+1} = e_{j+1}, \ldots, f_n = e_n.$$

Мы получаем

$$\begin{aligned} &q(x_1 f_1 + \ldots + x_n f_n) = \\ &= q\left(\frac{x_1}{\sqrt{\lambda_1}} e_1 + \ldots + \frac{x_i}{\sqrt{\lambda_i}} e_i + \frac{x_{i+1}}{\sqrt{-\lambda_{i+1}}} e_{i+1} + \cdots + \frac{x_j}{\sqrt{-\lambda_j}} e_j + x_1 e_{j+1} + \ldots + x_n e_n\right) = \\ &= \lambda_1 \left(\frac{x_1}{\sqrt{\lambda_1}}\right)^2 + \cdots + \lambda_i \left(\frac{x_i}{\sqrt{\lambda_i}}\right)^2 + \lambda_{i+1} \left(\frac{x_{i+1}}{\sqrt{-\lambda_{i+1}}}\right)^2 + \cdots + \lambda_j \left(\frac{x_j}{\sqrt{-\lambda_j}}\right)^2 = \\ &= x_1^2 + \cdots + x_i^2 - x_{i+1}^2 - \cdots - x_j^2. \end{aligned}$$

Предположим теперь, что $(f'_1, \ldots, f'_n)$ — ещё один ортогональный базис такой, что

$$q(x_1 f'_1 + \ldots + x_n f'_n) = x_1^2 + \cdots + x_{i'}^2 - x_{i'+1}^2 - \cdots - x_{j'}^2$$

для любых $x_1, \ldots, x_n \in \mathbb{R}$. Рассмотрим соответствующий ортонормированный базис $(e'_1, \ldots, e'_n)$, где $e'_k = f'_k/|f'_k|$. Мы получаем

$$q(x_1 e'_1 + \ldots + x_n e'_n) = q\left(\frac{x_1}{|f'_1|} f'_1 + \cdots + \frac{x_n}{|f'_n|} f'_n\right) =$$
$$= \frac{x_1^2}{|f'_1|^2} + \cdots + \frac{x_{i'}^2}{|f'_{i'}|^2} - \frac{x_{i'+1}^2}{|f'_{i'+1}|^2} - \cdots - \frac{x_{j'}^2}{|f'_{j'}|^2} + 0x_{j'+1}^2 + \cdots + 0x_n^2.$$

Из теоремы 13.7 следует, что последовательности чисел (длины n) $\lambda_1, \ldots, \lambda_n$ и

$$\frac{1}{|f'_1|^2}, \ldots, \frac{1}{|f'_{i'}|^2}, -\frac{1}{|f'_{i'+1}|^2}, \ldots, -\frac{1}{|f'_{j'}|^2}, 0, \ldots, 0$$

получаются одна из другой перестановкой. Учитывая знаки, получаем, что это возможно только, если $i = i'$ и $j = j'$. □

13.5. **Пример для $\mathbb{R}^n$.** Пусть квадратичная форма q на $\mathbb{R}^3$ задана формулой

$$q(x_1, x_2, x_3) = 23\,{x_1}^2 + 20\,x_1x_2 - 18\sqrt{3}\,x_1x_3 - 4\,{x_2}^2 + 20\sqrt{3}\,x_2x_3 + 5\,{x_3}^2. \qquad (50)$$

Запишем её матрицу в стандартном базисе $v = (v_1, v_2, v_3)$, где

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Получаем

$$q_v = \begin{pmatrix} 23 & 10 & -9\sqrt{3} \\ 10 & -4 & 10\sqrt{3} \\ -9\sqrt{3} & 10\sqrt{3} & 5 \end{pmatrix}.$$

Рассматривая оператор с такой матрицей в базисе v мы находим следующие собственные значения: $\lambda_1 = 16$, $\lambda_2 = -24$, $\lambda_3 = 32$ и следующие собственные векторы:

$$e_1 = \begin{pmatrix} -\frac{\sqrt{2}}{4} \\ -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{6}}{4} \end{pmatrix}, \quad e_2 = \begin{pmatrix} \frac{\sqrt{2}}{4} \\ -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{6}}{4} \end{pmatrix}, \quad e_3 = \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ 0 \\ -\frac{1}{2} \end{pmatrix}$$

Записывая эти столбцы в одну матрицу, получаем

$$T_{v,e} = \begin{pmatrix} -\frac{\sqrt{2}}{4} & \frac{\sqrt{2}}{4} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{6}}{4} & \frac{\sqrt{6}}{4} & \frac{1}{2} \end{pmatrix}.$$

Пусть

$$(x_1, x_2, x_3) = y_1 e_1 + y_2 e_2 + y_3 e_3.$$

Тогда согласно формулам (49) получаем

$$\begin{cases} y_1 = \dfrac{-\sqrt{2}x_1 - 2\sqrt{2}x_2 - \sqrt{6}x_3}{4} \\ y_2 = \dfrac{\sqrt{2}x_1 - 2\sqrt{2}x_2 + \sqrt{6}x_3}{4} \\ y_3 = \dfrac{-\sqrt{3}x_1 + x_3}{2}, \end{cases} \qquad q(x_1, x_2, x_3) = 16y_1^2 - 24y_2^2 + 32y_3^2.$$

Читатель может подставить y_1, y_2, y_3 из левой части в правую и проверить, что в действительности получается исходная формула (50).