Федеральное государственное образовательное бюджетное учреждение высшего образования
«ФИНАНСОВЫЙ УНИВЕРСИТЕТ
ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»
(Финансовый университет)
Департамент анализа данных, принятия решений и финансовых технологий

В.В.Щиголев

Алгебра и геометрия 1

Учебное пособие
по дисциплине «Алгебра и геометрия»

для студентов, обучающихся по направлению подготовки
01.03.02 «Прикладная математика и информатика»,
профиль «Анализ данных и принятие решений в экономике и финансах»,
(программа подготовки бакалавра).

Рассмотрено и одобрено на заседании Департамента анализа данных, принятия решений и финансовых технологий
(протокол №12 от «15» мая 2018 г.)

Москва, 2018

УДК 512
ББК 22.1я73
Щ76

Автор: Щиголев В.В., доктор физико-математических наук, профессор департамента анализа данных, принятия решений и финансовых технологий Финансового университета при Правительстве Российской Федерации

Рецензент: Тищенко А.В.., доктор физико-математических наук профессор, профессор департамента анализа данных, принятия решений и финансовых технологий Финансового университета при Правительстве Российской Федерации

Алгебра и геометрия 1: Учебное пособие по дисциплине «Алгебра и геометрия» для студентов, обучающихся по направлению подготовки 01.03.02 «Прикладная математика и информатика», профиль «Анализ данных и принятие решений в экономике и финансах», (программа подготовки бакалавра). - М.: Финансовый университет, департамент анализа данных, принятия решений и финансовых технологий, 2018. 60 с.

Учебное пособие предназначено для использования студентами в часы самостоятельной работы с целью освоения теоретических понятий высшей и линейной алгебры. Материал пособия может быть также полезен студентам-бакалаврам других направлений и профилей, магистрантам и аспирантам, осваивающим линейную алгебру.

УДК 512
ББК 22.1я73

Учебное издание

Щиголев Владимир Викторович

Алгебра и геометрия 1

Учебное пособие
по дисциплине «Алгебра и геометрия»

Компьютерный набор, верстка В.В.Щиголев
Формат 60x90/16. Гарнитура Times New Roman
Усл. п.л.3,75. Изд. № - 2018.
Заказ № ______
Электронное издание

Содержание

1. Элементарная теория множеств

Напомним, как обозначаются стандартные операции над множествами:

Обозначение	Определение	Название
$A \cup B$	Множество всех элементов, принадлежащих множеству A или множеству B; $\{x \mid x \in A \text{ или } x \in B\}$.	Объединение множеств A и B.
$A \cap B$	Множество всех элементов, принадлежащих множеству A и множеству B; $\{x \mid x \in A \text{ и } x \in B\}$.	Пересечение множеств A и B.
$A \setminus B$	Множество всех элементов, принадлежащих множеству A , но не принадлежащих множеству B; $\{x \mid x \in A \text{ и } x \notin B\}$.	Множество A минус множество B
$\mathrm{A}\Delta B$	Множество всех элементов, принадлежащих ровно одному множеству A или B; $(A\setminus B) \cup (B\setminus A)$.	Симметрическая разность множеств A и B.

Для каждого конечного множества X через $|X|$ будем обозначать количество элементов множества X. Пусть $X_1, \dots X_n$ — некоторые множества. *Декартовым произведением* $X_1 \times \dots \times X_n$ 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 $|X_1 \times \dots \times X_n| = |X_1| \cdots |X_n|$.

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 $X \xrightarrow{f} Y$.

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:

$$x_1, x_2 \in X,\ x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$$

или в эквивалентной форме

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

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

Мы зафиксируем следующие стандартные обозначения для множеств

Обозначение	**Множество**
$\mathbb{N}$	натуральные числа $\{1,2,3,\dots\}$
$\mathbb{Z}$	целые числа $\{\dots,-2,-1,0,1,2,,\dots\}$
$\mathbb{Q}$	рациональные числа $\left\{\frac{p}{q}\right\vert\ p,q \in \mathbb{Z},\ q \neq 0\}$
$\mathbb{R}$	действительные числа

Пример 1.4 Рассмотрим функцию $f\colon \mathbb{R} \to \mathbb{R}$, 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, $f^{-1}(\{-1\}) = \emptyset$. 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:*

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[rr, bend right, "g\circ f"'] & Y \arrow[r, "g"] & Z \end{tikzcd}$$

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

Пусть S — некоторое множество. *Бинарной операцией* на S называется любое отображение $f\colon S \times S \to S$. 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* $g\colon K \times S \to S$ 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 *— множе-*

ство 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

$$y = y \cdot e = y \cdot (x \cdot y') = (y \cdot x) \cdot y' = e \cdot y' = y'.$$

□

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: $(\mathbb{R}\backslash\{0\},\cdot)$ — 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 $\mathrm{i}d_X$, а обратным к f элементом является обратное отображение f^{-1}. Мы будем применять обозначение $S_n=S(\{1,2,\dots,n\})$. Эта группа называется *симметрической группой*. Мы будем использовать её в §10 для определения понятия определителя.

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

Лемма 2.3 *Пусть* $f\colon G\to H$ *— 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* $H\le G$.

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 $e_H^{-1} e_H = e_H e_H^{-1} = e$. Умножая на него предыдущее равенство слева, получаем

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

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

Доказательство следующего простого утверждения мы оставляем читателю.

Лемма 2.5 *Пусть H — подгруппа группы G. Тогда подмножества $gH = \{gh \mid h \in H\}$ (подмножества $Hg = \{hg \mid h \in H\}$) либо совпадают, либо не пересекаются.*

Подмножества gH называются *левыми смежными* классами, а подмножества Hg называются *правыми смежными* классами.

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

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

Это — коммутативное кольцо, что легко проверить умножением:

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

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

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

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

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 $\mathbb{Z}_2 = \{\overline{0}, \overline{1}\}$ (поле вычетов по модулю 2) следующими таблицами:

сложение	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{1}$
$\overline{1}$	$\overline{1}$	$\overline{0}$

умножение	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$

4. Векторные пространства

Определение 4.1 *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 (В1)–(В4) эквивалентны утверждению о том, что $(V, +)$ — абелева группа.

Пример 4.2 Основным примером векторного пространства является декартова степень поля

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

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

Векторы (наборы) складываются и умножаются на скаляр следующим образом:

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

Читатель может легко проверить справедливость аксиом (В1)–(В8). Мы называем векторное пространство из этого примера *арифметическим* векторным пространством.

Мы перейдём теперь к одному из основных понятий линейной алгебры.

Определение 4.3 *Пусть $v_1, \dots, v_n$ — векторы векторного пространства V над F. Линейной комбинацией этих векторов называется следующий вектор:*

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

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

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

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

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

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

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

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

С другой стороны, тот факт, что набор векторов $(v_1, \dots, v_n)$ линейно зависим, означает, что существуют неравные одновременно нулю скаляры $\alpha_1, \dots, \alpha_n \in F$ 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 4.5 *Если* $n > 0$, *то набор векторов* $(v_1, \dots, v_n)$ *линейно зависим тогда и только тогда, когда существует индекс* $i = 1, \dots, n$ *для которого* $v_i \in \langle v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n \rangle$.

Доказательство. Пусть набор векторов $(v_1, \dots, v_n)$ линейно зависим. Запишем линейную комбинацию $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$, в которой $\alpha_i \neq 0$ для некоторого $i = 1, \dots, n$. 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 следующей линейной оболочке векторов $\langle v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n \rangle$.

Пусть наоборот $v_i \in \langle v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n \rangle$. 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 $\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_n \in F$. 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 $(v_1, \dots, v_n)$ с ненулевыми коэффициентами

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

равна нулю. □

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

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

где $\alpha_1, \dots, \alpha_n \in F$. В этом случае говорят, что вектор v *раскладывается* по базису $(v_1, \dots, v_n)$ с *коэффициентами* $(\alpha_1, \dots, \alpha_n)$. 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 $(v_1, \dots, v_n)$ линейно независим, то $\alpha_i - \beta_i = 0$ для любого i. Следовательно, $\alpha_i = \beta_i$.

Определение 4.6 *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* $dimV = n$. *Если существуют наборы линейно независимых векторов пространства* V *сколь угодно большой длины, то пространство* V *имеет бесконечную размерность. В этом случае, мы используем обозначение* $dimV = \infty$.

Лемма 4.7 *Пусть* $dimV = n < \infty$. *Тогда существует базис пространства* V, *состоящий ровно из* n *векторов.*

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

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

Следовательно, $\alpha \neq 0$ и мы получаем
$$v = -\frac{\alpha_1}{\alpha} v_1 - \dots - \frac{\alpha_n}{\alpha} v_n \in \langle v_1, \dots, v_n \rangle.$$
□

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 4.8 *Пусть* $(g_1, \dots, g_l)$ *— линейно независимый набор векторов, элементы которого принадлежащие линейной оболочке* $\langle f_1, \dots, f_k \rangle$*. Тогда* $k \geq l$.

Доказательства. Проведём индукцию по k. В качестве базы индукции рассмотрим случай $k = 0$, так как в этом случае $\langle f_1, \dots, f_k \rangle = 0$ 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:
$$\begin{array}{rcl} g_1 & = & \alpha_{1,1} f_1 + \alpha_{1,2} f_2 + \dots + \alpha_{1,k} f_k, \\ g_2 & = & \alpha_{2,1} f_1 + \alpha_{2,2} f_2 + \dots + \alpha_{2,k} f_k, \\ & \vdots & \\ g_l & = & \alpha_{l,1} f_1 + \alpha_{l,2} f_2 + \dots + \alpha_{l,k} f_k. \end{array}$$
Если $\alpha_{1,k} = \alpha_{2,k} = \dots = \alpha_{l,k} = 0$, то $g_1, \dots, g_l$ принадлежат $\langle f_1, \dots, f_{k-1} \rangle$ и по предположению индукции $k > k - 1 \geq l$. Теперь предположим, что $\alpha_{i,k} \neq 0$ для некоторого $i = 1, \dots, l$. Тогда мы можем выразить f_k следующим образом:
$$f_k = \frac{1}{\alpha_{i,k}} g_i - \frac{\alpha_{i,1}}{\alpha_{i,k}} f_1 - \frac{\alpha_{i,2}}{\alpha_{i,k}} f_2 - \dots - \frac{\alpha_{i,k-1}}{\alpha_{i,k}} f_{k-1}.$$
Подставляя это значение во все уравнения для $g_1, \dots, g_{i-1}, g_{i+1}, \dots, g_l$ и перенося все g (с индексами) влево, а все f (с индексами) вправо получаем
$$\begin{array}{rcl} g'_1 & = & \beta_{1,1} f_1 + \beta_{1,2} f_2 + \dots + \beta_{1,k-1} f_{k-1}, \\ & \vdots & \\ g'_{i-1} & = & \beta_{i-1,1} f_1 + \beta_{i-1,2} f_2 + \dots + \beta_{i-1,k-1} f_{k-1}, \\ g'_{i+1} & = & \beta_{i+1,1} f_1 + \beta_{i+1,2} f_2 + \dots + \beta_{i+1,k-1} f_{k-1}, \\ & \vdots & \\ g'_l & = & \beta_{l,1} f_1 + \beta_{l,2} f_2 + \dots + \beta_{l,k-1} f_{k-1}, \end{array}$$

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

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

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

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

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

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

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

Следствие 4.9 *Пусть $dimV = n < \infty$. Тогда любой базис пространства V содержит ровно n элементов.*

Доказательство. Согласно лемме 4.7 в пространстве V существует базис $(g_1, \dots, g_n)$. Пусть $(f_1, \dots, f_k)$ — ещё один базис пространства V. Мы получаем $g_1, \dots, g_n \in V = \langle f_1, \dots, f_k \rangle$ и набор $(g_1, \dots, g_n)$ линейно независим. Применяя лемму 4.8 для $l = n$, получаем $k \geq n$. Однако, случай $k > n = \dim V$ невозможен по определению размерности.

Следствие 4.10 *Пусть V — векторное пространство и $V = \langle f_1, \dots, f_k \rangle$. Тогда $dimV \leq k$. Если при этом набор $(f_1, \dots, f_k)$ линейно независим, то $dimV = k$.*

Доказательство. Пусть $(g_1, \dots, g_l)$ — линейно независимый набор векторов пространства V. По лемме 4.8, получаем $l \leq k$. Это означает, что $\dim V \leq k$. С другой стороны, если набор $(f_1, \dots, f_k)$ линейно независим, то $\dim V = k$. □

Следствие 4.11 *Пусть $dimV = n < \infty$ и $(f_1, \dots, f_k)$ — линейно независимый набор векторов пространства V, где $k < n$. Тогда существуют векторы $f_{k+1}, \dots, f_n \in V$ такие, что $(f_1, \dots, f_n)$ — базис пространства V.*

Доказательство. Достаточно доказать, что можно построить f_{k+1} так, чтобы $(f_1, \dots, f_{k+1})$ был линейно независимым. Пусть $(e_1, \dots, e_n)$ — базис пространства V. Предположим, что $e_1, \dots, e_n \in \langle f_1, \dots, f_k \rangle$. По лемме 4.8 мы получаем тогда противоречие $k \geq n$. Поэтому существует $i = 1, \dots, n$, для которого $e_i \notin \langle f_1, \dots, f_k \rangle$. Мы утверждаем, что набор $(f_1, \dots, f_k, e_i)$ линейно независим. Действительно, пусть

$$\alpha_1 f_1 + \cdots + \alpha_k f_k + \lambda e_i = 0$$
для некоторых $\alpha_1, \dots, \alpha_i, \lambda \in F$. Если $\lambda \neq 0$, то получаем противоречие
$$e_i = -\frac{\alpha_1}{\lambda} f_1 - \cdots - \frac{\alpha_k}{\lambda} f_k \in \langle f_1, \dots, f_k \rangle.$$
Если же $\lambda = 0$, то получаем $\alpha_1 f_1 + \cdots + \alpha_k f_k = 0$. Отсюда из линейной независимости набора $(f_1, \dots, f_k)$, получаем $\alpha_1 = \cdots = \alpha_k = 0$. □

Определение 4.12 *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) $U \neq \emptyset$;

(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 4.13 *Пусть U — подпространство векторного пространства V и $v_1, \dots, v_n$ — векторы, принадлежащие U. Тогда $\langle v_1, \dots, v_n \rangle \subset U$.*

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 4.14 *Элементарные преобразования набора векторов не меняют их линейной оболочки.*

Доказательство. Пусть $(u_1, \dots, u_m)$ — набор векторов, полученный из набора векторов $(v_1, \dots, v_n)$ одним из элементарных преобразований. Из определения видно, что каждый вектор u_i является линейной комбинацией векторов $v_1, \dots, v_n$. Поэтому $u_i \in \langle v_1, \dots, v_n \rangle$. По лемме 4.13 получаем $\langle u_1, \dots, u_m \rangle \subset \langle v_1, \dots, v_n \rangle$.

Остаётся только доказать, что элементарные преобразования обратимы. Докажем это для преобразований (ЭВ2) и (ЭВ3). Пусть $(v_1, \dots, v_n)$ — набор векторов. Умножая i-й вектор на ненулевой скаляр α, получаем набор

$(v_1, \dots, v_{i-1}, \alpha v_i, v_{i+1}, \dots, v_n)$. Умножая i-й вектор на α^{-1}, получаем исходный набор $(v_1, \dots, v_n)$. Аналогично, умножая i-й вектор на α и добавляя получившееся произведение к j-у вектору, получаем набор $(v_1, \dots, v_{j-1}, v_j + \alpha v_i, v_{j+1}, \dots, v_n)$. 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

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

Лемма 4.15 *$V \cap U$ и $V + U$ — подпространства. При этом $V + U$ — наименьшее подпространство, содержащее объединение $V \cup U$. Если $V = \langle v_1, \dots, v_n \rangle$ и $U = \langle u_1, \dots, u_m \rangle$, то $V + U = \langle v_1, \dots, v_n, u_1, \dots, u_m \rangle$.*

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 $v + u \in W$ по свойству (П2). Этим мы доказали равенство $W = V + U$. □

Лемма 4.16 *Пусть V и U — конечномерные подпространства некоторого векторного пространства. Тогда*

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

Доказательство. Пусть $(a_1, \dots, a_n)$ — базис V и $(b_1, \dots, b_m)$ — базис U. Так как $V + U = \langle a_1, \dots, a_n, b_1, \dots, b_m \rangle$, то по следствию 4.10, получаем $\dim V + U \leq n + m$. Пространство $V \cap U$ тоже конечномерно как подпространство конечномерного пространства V или U. Поэтому все размерности в утверждении леммы конечные.

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

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

По следствию 4.11 мы можем достроить набор $(w_1, \dots, w_k)$ до базиса $(w_1, \dots, w_k, v_1, \dots, v_{n-k})$ пространства V и до базиса $(w_1, \dots, w_k, u_1, \dots, u_{m-k})$ пространства U.

Наша цель доказать, что $(w_1, \dots, w_k, v_1, \dots, v_{n-k}, u_1, \dots, u_{m-k})$ — базис суммы $V + U$. По лемме 4.15 и замечанию после леммы 4.14 мы получаем

$$V + U = \langle w_1, \dots, w_k, v_1, \dots, v_{n-k} \rangle + \langle w_1, \dots, w_k, u_1, \dots, u_{m-k} \rangle =$$
$$= \langle w_1, \dots, w_k, v_1, \dots, v_{n-k}, w_1, \dots, w_k, u_1, \dots, u_{m-k} \rangle =$$

$$= \langle w_1, \dots, w_k, v_1, \dots, v_{n-k}, u_1, \dots, u_{m-k} \rangle.$$

Теперь рассмотрим нулевую линейную комбинацию

$$\alpha_1 w_1 + \dots + \alpha_k w_k + \beta_1 v_1 + \dots + \beta_{n-k} v_{n-k} + \gamma_1 u_1 + \dots + \gamma_{m-k} u_{m-k} = 0, (1)$$

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

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

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

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

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

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

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 (1), 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

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

$\square$

5. Евклидовы пространства

Определение 5.1 *Евклидовым пространством называется линейное пространство E над полем действительных чисел $\mathbb{R}$, наделённое отображением $(\_,\_) \colon E \times E \to \mathbb{R}$, 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* $u \in E \backslash \{0\}$.

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.

Теорема 5.2 (Неравенство Коши-Буняковского) $|uv| \leq |u||v|$ *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

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

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

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

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. □

Определение 5.3 *Пусть* E *— евклидово пространство и* $(e_1, \dots, e_n)$ *— 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 $(e_1, \dots, e_n)$ легко записать координаты вектора. Действительно, пусть

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

Умножая на e_i обе части, получаем

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

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 $e_1, \dots, e_n$ следующим образом

$$e_i e_j = \delta_{i,j}.$$

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

Теорема 5.4 *В произвольном конечномерном евклидовом пространстве существует хотя бы один ортогональный базис. Более того, любой ненулевой вектор из E является элементом одного из таких базисов.*

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

$$\langle v_1, \dots, v_i \rangle = \langle e_1, \dots, e_i \rangle. \tag{2}$$

Действительно, $e_1 = v_1$, что доказывает утверждение для $i = 1$. Пусть теперь $1 < i \le n$ 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 (2). 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 4.8 получаем противоречие $i - 2 \ge i - 1$. Из доказанного факта $e_1, \dots, e_{i-1} \ne 0$ следует, что e_i — корректно определённый вектор. Отсюда

$$v_i = e_i + \alpha_{i,1} e_1 + \dots + \alpha_{i,i-1} e_{i-1} \in \langle \mathrm{e}_1, \dots, e_i \rangle.$$

Так как $v_1, \dots, v_{i-1} \in \langle \mathrm{e}_1, \dots, e_{i-1} \rangle \subset \langle \mathrm{e}_1, \dots, e_i \rangle$, то по лемме 4.13 получаем $\langle v_1, \dots, v_i \rangle \subset \langle \mathrm{e}_1, \dots, e_i \rangle$.

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,

$$e_i = v_i - \alpha_{i,1} e_1 - \dots - \alpha_{i,i-1} e_{i-1} \in \langle \mathrm{v}_1, \dots, v_i \rangle.$$

Поэтому по лемме 4.13 получаем $\langle e_1, \dots, e_i \rangle \subset \langle v_1, \dots, v_i \rangle$. Таким образом, равенство (2) 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 $V = \langle \mathrm{e}_1, \dots, e_n \rangle = \langle \mathrm{e}_1, \dots, e_{i-1}, e_{i+1}, \dots, e_n \rangle$ для некоторого i. Так как $(v_1, \dots, v_n)$ линейно независим, то по лемме 4.8, получаем противоречие $n - 1 \ge n$.

Остаётся доказать, что $e_i e_j = 0$ для различных $i \ne j$. 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 4.11 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 $e_1 = v_1 = v$. □

6. Системы линейных уравнений

Определение 6.1 *Линейным уравнением относительно переменных* $x_1, \dots, x_n$ *над полем* F *называется уравнение вида* $a_1x_1 + \dots + a_nx_n = b$, *где* $a_1, \dots, a_n, b \in F$. *Конечный набор таких уравнений называется системой линейных уравнений над полем* F. *Системы записываются так*

$$\begin{cases} a_{1,1}x_1 + \dots + a_{1,n}x_n & = & b_1; \\ \vdots & & \vdots \\ a_{m,1}x_1 + \dots + a_{m,n}x_n & = & b_m; \end{cases} \tag{3}$$

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:

$$c(a_1x_1 + \dots + a_nx_n = b) = (ca_1x_1 + \dots + ca_nx_n = cb)$$

$$\begin{aligned}(a_1x_1 + \dots + a_nx_n = b) + (a'_1x_1 + \dots + a'_nx_n = b') = \\ = ((a_1 + a'_1)x_1 + \dots + (a_n + a'_n)x_n = b + b').\end{aligned}$$

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 $0x_1 + \dots + 0x_n = 0$, и уравнением обратным к уравнению $a_1x_1 + \dots + a_nx_n = b$ является уравнение $-a_1x_1 - \dots - a_nx_n = -b$.

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) (и являются их частным случаем).

Лемма 6.2 *Элементарные преобразования сохраняют множества решений системы.*

Доказательство. Доказательство леммы 4.14 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 $x_1 = c_1, \dots, x_n = c_n$ — решение системы (3), то этот набор является решение любой системы, полученной из системы (3) одним элементарным преобразованием.

Достаточно рассмотреть преобразования (ЭС2) и (ЭС3). Сначала умножим i-е уравнение системы (3) на $\alpha \in F\backslash\{0\}$. Получаем уравнение $\alpha a_{i,1}x_1 + \dots + \alpha a_{i,n}x_n = \alpha b_i$. Проверим, что подстановка $x_1 = c_1, \dots, x_n = c_n$ даёт верное равенство:

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

В этом вычислении мы использовали верное по условию i-е уравнение системы (3)

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

Теперь умножим i-е уравнение системы (3) на $\alpha \in F$ и добавим к j-у, где $i \neq j$. Получаем уравнение $(\alpha a_{i,1} + a_{j,1})x_1 + \dots + (\alpha a_{i,n} + a_{j,n})x_n = \alpha b_i + b_j$. Проверим, что подстановка $x_1 = c_1, \dots, x_n = c_n$ даёт верное равенство:

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

В этом вычислении мы использовали верные по условию i-е и j-е уравнения системы (3)

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

□

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

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

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

Применим индукцию по количеству уравнений. База индукции — это пустая система. Теперь предположим, что наша система содержит хотя бы одно уравнение. Запишем её в виде (3). 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 (3). Для каждого $i = 2, \dots, m$ 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.

7. Матрицы

Матрицей называется прямоугольный массив элементов некоторого коммутативного 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 $A = (a_{i,j})_{i=1,\dots,m;j=1,\dots,n}$. 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.

7.1 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.

С единичной матрицей тесно связано понятие обратной 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

$$\begin{pmatrix} 0 & \cdots & 0 & b_{1,j_1} & & \cdots & & b_{1,n} \\ 0 & & \cdots & 0 & b_{2,j_2} & & \cdots & b_{2,n} \\ & & & & \cdots & & & \\ 0 & & \cdots & & 0 & b_{k,j_k} & \cdots & b_{k,n} \\ 0 & & & & \cdots & & & 0 \\ & & & & \cdots & & & \\ 0 & & & & \cdots & & & 0 \end{pmatrix}, \qquad (4)$$

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 7.1 *Ненулевые строки ступенчатой матрицы над полем линейно независимы.*

Доказательство. Пусть r_i обозначает i-у строку ступенчатой матрицы (4). Предположим, что выполнено равенство

$$\alpha_1 r_1 + \alpha_2 r_2 + \cdots + \alpha_k r_k = 0. \qquad (5)$$

Мы докажем индукцией по $l = 0, \dots, k$, что $\alpha_1 = \cdots = \alpha_l = 0$. В случае $l = 0$ доказывать нечего, так как в этом случае мы ничего не утверждаем. Пусть теперь $1 \leq l \leq k$ и утверждение верно для $l - 1$. Таким образом мы уже знаем, что $\alpha_1 = \cdots = \alpha_{l-1} = 0$. Отсюда и из равенства (5) 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 $\alpha_l = 0$. □

7.2 Операции над матрицами

Наша следующая цель — определить линейные действия над матрицами и умножения матриц. Сложение матриц 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 $A = (a_{i,k})_{i=1,\dots,m;k=1,\dots,n}$ и $B = (b_{k,j})_{k=1,\dots,n;j=1,\dots,l}$ — матрицы. Мы обозначим их произведение через $AB = C = (c_{i,j})_{i=1,\dots,m;j=1,\dots,l}$. В этом случае

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

для любых $i = 1, \dots, m$ и $j = 1, \dots, l$. 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 7.2 *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

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

Положим

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

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

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

Требуемое равенство $h_{i,j} = g_{i,j}$ получается как следствие 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 7.3 *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 7.4 $(AB)^T = B^T A^T$ *и* $(A + B)^T = A^T + B^T$.

Доказательство. Пусть

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

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 7.5 *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 7.4:

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

Второе равенство доказывается аналогично. □

Кроме транспонированной матрицы иногда бывает удобно (смотрите, например доказательство леммы 7.13) рассматривать *псевдотранспонированную* матрицу: $A^t = (a_{m+1-i,n+1-j})_{j=1,\ldots,n;i=1,\ldots,m}$, *если* $A = (a_{i,j})_{i=1,\ldots,m;j=1,\ldots,n}$.

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 7.6 *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 7.2:

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

Остальные равенства доказываются аналогично. □

7.3 Матрицы элементарных преобразований

Мы обозначим через $E_{i,j}$ или более точно $E_{i,j}^{(m,n)}$ 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 7.7 *Пусть* A *— матрица размера* $m \times n$ *и* i, j, k *— натуральные числа такие, что* $i \leq k$ *и* $j \leq k$*. 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-о столбца матрицы A, записанного на месте j-о 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 определение произведения матриц. Мы рассмотрим задачи, в которых оно естественно возникает.

Заметим, что систему (3) можно записать в следующем виде:

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

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 изменяется эта матрица при элементарных преобразованиях (ЭС1)–(ЭС5) на языке умножения матриц.

Рассмотрим следующие матрицы:

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

Лемма 7.8 *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* $X_{i,j}^{(m)}(\alpha)$ *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* $X_{i,j}^{(m)}(\alpha)$ *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 7.7.

Следствие 7.9 *Любая матрица* A *над полем может быть представлена в виде* $A = X_1 \cdots X_k B$*, где* $X_1, \dots, X_k$ *— 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 уравнений к ступенчатому виду, описанный в конце §6 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

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

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

7.4 Ранг матрицы

В этом параграфе, мы используем следующее временное определение.

Определение 7.10 *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 7.11 *Пусть A — матрица над полем F размера $m \times n$ и $c_1, \dots, c_n$ — её столбцы. Пусть B — матрица над F размера $k \times m$ и $c'_1, \dots, c'_n$ — столбцы матрицы BA. Если для некоторых $\alpha_1, \dots \alpha_n \in F$ выполнено равенство $\alpha_1 c_1 + \dots + \alpha_n c_n = 0$, то выполнено также равенство $\alpha_1 c'_1 + \dots + \alpha_n c'_n = 0$.*

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

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

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

Доказательство. По лемме 7.8 мы получаем, что $A' = BA$ для некоторой матрицы B. Произвольный столбец c_j матрицы A допускает представление

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

для некоторых элементов поля $\alpha_1, \dots, \alpha_k$. По лемме 7.11 получаем

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

Теперь предположим, что

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

для некоторых элементов поля $\alpha_1, \dots, \alpha_k$. 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 7.11 получаем

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

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

Лемма 7.13 *Пусть B — ступенчатая матрица над полем. Тогда её строчный и столбцовый ранги совпадают.*

Доказательство. Запишем матрицу B в виде (4). По лемме 7.1 мы получаем, что строчный ранг матрицы B равен k. Обозначим через $c_1, \dots, c_n$ столбцы матрицы B. Рассуждая аналогично лемме 7.1, мы можем доказать, что столбцы $c_{j_1}, c_{j_2}, \dots, c_{j_k}$ линейно независимы. Этот факт можно непосредственно извлечь из леммы 7.1, 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 $c_{j_1}, c_{j_2}, \dots, c_{j_k}$. Для этого докажем индукцией по $l = 0, \dots, k$, 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 $c_{j_1}, c_{j_2}, \dots, c_{j_l}$. Для $l = 0$ это верно так как в этом случае столбец u нулевой.

Пусть теперь $l = 1, \dots, k$ 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

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

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

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

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

Теорема 7.14 *Строчный и столбцовый ранги матрицы совпадают.*

Доказательство. Мы можем элементарными преобразованиями строк привести произвольную матрицу A над полем к ступенчатой матрице B. По следствию 7.12 мы получаем, что столбцовые ранги матриц A и B совпадают. Строчные ранги матриц A и B совпадают по лемме 4.14. Остаётся заметить, что столбцовый и строчный ранги матриц B совпадают по лемме 7.13. $\square$

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

7.5 Матрица перехода между базисами

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

$$v_j = \sum_{i=1}^{n} a_{i,j} u_i, \tag{6}$$

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

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

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 (6) в виде

$$v = uT_{u,v}. \tag{7}$$

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

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

Отсюда получаем

$$T_{u,w} = T_{u,v}T_{v,w}. \tag{8}$$

Заметим, что $T_{u,u} = E$, где E — единичная матрица размера $n \times n$. Из формулы (8) 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 связаны координаты векторов в двух различных базисах, если известна матрица перехода. Пусть $\lambda = b_1v_1 + \dots + b_nv_n$ и $\lambda = a_1u_1 + \dots + a_nu_n$, 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

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = T_{u,v} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}. \qquad (9)$$

7.6 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 (9) при помощи формулы (6). В вышеприведённых обозначениях получаем

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

Сравнивая эту формулу с разложением $\lambda = \sum_{i=1}^{n} a_i u_i$, получаем формулу

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

для любого $i = 1, \dots, n$. Эти формулы эквивалентно могут быть записаны в виде (9).

Теперь мы можем доказать формулу (8). Разложим вектор λ по третьему базису:

$$\lambda = c_1 w_1 + \dots + c_n w_n.$$

Аналогично формуле (9) 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 (9), 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 (9) 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 (8), остаётся применить следующую лемму.

Лемма 7.15 *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-му столбцу матрицы A и BX равно i-му столбцу матрицы B. По условию эти столбцы совпадают для любого $i = 1, \dots, n$. Отсюда $A = B$. □

8 Линейные отображения

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 $\alpha \in F$.

8.1 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:

$$\varphi(u) = vM_{v,u}(\varphi), \tag{10}$$

где

$$M_{v,u}(\varphi) = (\varphi_{i,j})_{i=1,\dots,m; j=1,\dots,n} \quad \text{и} \quad \varphi(u) = (\varphi(u_1), \dots, \varphi(u_n))$$

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

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

пространства U. Применяя φ к этому равенству, при помощи формулы (10) получаем

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

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

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

Мы также можем вычислить как изменится матрица линейного отображения при замене базисов. Пусть $u' = (u'_1, \dots, u'_n)$ и $v' = (v'_1, \dots, v'_m)$ — ещё два базиса пространств U и V соответственно. Согласно уравнению (7) 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 (10) 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 $T_{u,u'}^{-1}$ справа получаем

$$\varphi(u) = vT_{v,v'}M_{v',u'}(\varphi)T_{u,u'}^{-1}.$$

Сравнивая с (10), получаем $M_{v,u}(\varphi) = T_{v,v'}M_{v',u'}(\varphi)T_{u,u'}^{-1}$. Отсюда

$$M_{v',u'}(\varphi) = T_{v,v'}^{-1}M_{v,u}(\varphi)T_{u,u'}.$$

Наконец, посмотрим, как устроена матрица композиции линейных отображений. Пусть $\varphi: U \to V$ и $\psi: V \to W$ — два линейных отображения, и $u = (u_1, \dots, u_n)$, $v = (v_1, \dots, v_m)$ и $w = (w_1, \dots, w_k)$ — базисы пространств U, V и W соответственно. Вместе с формулой (10) верны аналогичные формулы

$$\psi(v) = wM_{w,v}(\psi), \qquad \psi\varphi(u) = wM_{w,u}(\psi\varphi). \tag{11}$$

Применяя к обеим частям формулы (10) отображение ψ, получаем

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

Сравнивая эту формулу с первой формулой (11), получаем

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

9 Многочлены

9.1 Определения и основные свойства

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

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

$$(a_0, a_1, \dots)(b_0, b_1, \dots) = (a_0 b_0, a_0 b_1 + a_1 b_0, , \dots, \sum_{j=0}^{i} a_j b_{i-j}, \dots).$$

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

Лемма 9.1 *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 $(a_0, a_1, \dots)$ и $(b_0, b_1, \dots)$ — 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 $\sum_{j=0}^{i} a_j b_{i-j} \neq 0$. В этом случае $a_j \neq 0$ и $b_{i-j} \neq 0$ для некоторого j. Отсюда получаем $j \leq N$ и $i - j \leq N$. Складывая, получаем $i \leq 2N$.

Докажем теперь, что множество многочленов с одной переменной является кольцом. В этом множестве последовательность из одних нулей $(0,0, \dots)$ 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 $(0,0, \dots)$ является нулём, противоположным элементом к последовательности $(a_1, a_2, \dots)$ является последовательность $(-a_1, -a_2, \dots)$. Легко заметить, что последовательность $(1,0,0, \dots)$, в которой все элементы кроме первого нулевые, является единицей. Поэтому выполнено свойство 1.

Остаётся проверить только свойства (К5) и (К6). Рассмотрим три последовательности $(a_0, a_1, \dots)$, $(b_0, b_1, \dots)$ и $(c_0, c_1, \dots)$ и определим следующие произведения:

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

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

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

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

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

откуда и следует требуемое равенство $f_l = g_l$. Этим доказано свойство (К5).

Наконец, мы положим

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

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

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 (К6) 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 $x = (0,1,0,0,\dots)$ и рассмотрим этот многочлен как переменную. Кроме того, мы отождествим любой элемент $a \in R$ с последовательностью $(a, 0,0, \dots)$. Более формально, это означает, что отображение $a \mapsto (a, 0,0, \dots)$ является инъективным гомоморфизмом. Так как $ax^n = x^n a = (0, \dots, 0, a, 0,0, \dots)$, то любой многочлен однозначно представляется в виде $a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$, где $a_0, \dots, a_n \in R$ 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 $f(\alpha)$. Таким образом

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

$\cdots + a_1\alpha + a_0$.

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:

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

В этом случае мы определим степень многочлена следующим образом:

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

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

(1) $deg(f + g) \leq max\{deg\, f, deg\, g\}$.

(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 $n \geq m$. Запишем

$$f = a_n x^n + \cdots + a_1 x + a_0, g = b_m x^m + \cdots + b_1 x + b_0, \text{где } a_n \neq 0, b_m \neq 0. \quad (12)$$

Отсюда получаем

$$f + g = a_n x^n + \cdots + a_{m+1} x^{m+1} + (a_m + b_m) x^n + \cdots + (a_1 + b_1) x + (a_0 + b_0). \quad (13)$$

Отсюда следует, что $\deg(f + g) \leq n = \max\{n, m\} = \max\{\deg f, \deg g\}$. 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 (12). Отсюда получаем равенство (13). 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 (12). 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 $\deg c = 0$. □

9.2 Деление с остатком и теорема Безу

Теорема 9.3 (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 $\deg f \geq \deg g$. Запишем многочлены f и g в виде (12). 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

$$(q' - q)g = r - r'. \tag{14}$$

Посчитаем степени многочленов в разных сторонах этого равенства, используя лемму 9.2:

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

Следовательно,

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

Учитывая наши правила сравнения и сложения для множества $\mathbb{N} \cup \{0, -\infty\}$, получаем, что вышеприведённое неравенство может выполняться только, если $\deg(q' - q) = -\infty$. Это означает, что $q' = q$. Отсюда и из (14) следует, что $r = r'$. □

Следствие 9.4 (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 9.3 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 $R = \mathbb{Z}_2$ (смотрите пример **3.7**). 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 поля.

Лемма 9.5 *Пусть* F *— поле и* $f \in F[x]$. *Предположим, что существуют попарно различные элементы* $\alpha_1, \ldots, \alpha_{n+1}$ *поля F такие, что* $n \geq$

$\deg f$ *и* $f(\alpha_i) = 0$ *для любого* $i = 1, \dots, n+1$. *Тогда* $f = 0$.

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

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

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

по лемме 9.2. Отсюда $\deg q \le n - 1$. Поэтому мы можем применить индуктивное предположение к многочлену q и получить, что $q = 0$. Отсюда $f = q(x - \alpha_{n+1}) = 0$. □

Следствие 9.6 *Пусть* F *— поле и* $f, g \in F[x]$. *Предположим, что существуют попарно различные элементы* $\alpha_1, \dots, \alpha_{n+1}$ *поля* F *такие, что* $n \ge max\{degf, degg\}$ *и* $f(\alpha_i) = g(\alpha_i)$ *для любого* $i = 1, \dots, n+1$. *Тогда* $f = g$.

Доказательство. Рассмотрим разность $h = f - g$. По лемме 9.2 получаем $\deg h \le \max\{\deg f, \deg g\} \le n$. Кроме того, $h(\alpha_i) = f(\alpha_i) - g(\alpha_i) = 0$ для любого $i = 1, \dots, n+1$. Применяя лемму 9.5 к многочлену h, получаем $h = 0$. Отсюда $f = g$. □

9.3 Рациональные корни

Теорема 9.7 *Пусть* $f = a_n x^n + \dots + a_1 x + a_0$ *— многочлен из* $\mathbb{Z}[x]$ *степени* $n \ge 1$ *и* p/q *— его корень, где* p *и* $q \ne 0$ *— 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

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

Умножая это равенство на q^n, получаем

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

Перенося все слагаемые кроме последнего в правую часть, получаем

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

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

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

(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 $k \leq n$. 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 $q^n f(m)$. 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. □

10 Определители

10.1 Симметрическая группа

Пусть n — натуральное число. Напомним, что *симметрической группой* S_n называется множество всех биекций множества $\{1, \ldots, n\}$ относительно композиции в качестве групповой операции. Элементы группы S_n называются *подстановками*. Мы будем использовать запись

$$\sigma = \begin{pmatrix} i_1 & \cdots & i_n \\ j_1 & \cdots & j_n \end{pmatrix}, \tag{15}$$

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

Заметим, что если мы переставим столбцы этой матрицы, то 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

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

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 $\sigma = (i_1, i_2, \dots, i_k)$, где $i_1, i_2, \dots, i_n$ — попарно различные числа из множества $\{1, \dots, n\}$, чтобы сказать, что σ действует следующим образом:

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

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

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

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

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 6 & 8 & 5 & 7 & 1 & 4 & 3 \end{pmatrix} = (1, 2, 6)(3, 8)(4, 5, 7) \tag{16}$$

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 $i_1, \dots, i_n$ — перестановка. Пару чисел (a, b) из множества $\{1, \dots, n\}$ назовём *инверсией* перестановки $i_1, \dots, i_n$, если $a > b$, $a = i_k$, $b = i_l$ и $k < l$. Множество всех инверсий перестановки $i_1, \dots, i_n$ обозначим через $D(i_1, \dots, i_n)$. Например,

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

Лемма 10.1 *Пусть $j_1, \dots, j_n$ — перестановка, полученная из перестановки $i_1, \dots, i_n$ перестановкой k-о и $k + 1$-о индексов, где $1 \leq k < n$. Тогда*

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

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

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

$$D(j_1, \dots, j_n) \subset D(i_1, \dots, i_n)\Delta\{(a, b)\}. \tag{17}$$

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

$$D(i_1, \dots, i_n) \subset D(j_1, \dots, j_n)\Delta\{(a, b)\}. \tag{18}$$

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 $x \in B \backslash A$. Мы собираемся применить этот факт к включению (18). Предположим, что

$$(a, b) \in (D(j_1, \dots, j_n)\Delta\{(a, b)\}) \backslash D(i_1, \dots, i_n) \tag{19}$$

Так как $(a, b) \in D(i_1, \dots, i_n) \cup D(j_1, \dots, j_n)$, то $(a, b) \in D(j_1, \dots, j_n)$ и $(a, b) \notin D(j_1, \dots, j_n) \backslash \{(a, b)\} = D(j_1, \dots, j_n)\Delta\{(a, b)\}$. Получаем противоречие с формулой (19).

Таким образом, мы доказали, что формула (19) не выполняется. Следовательно, используя ассоциативность симметрической разности, из формулы (18) получаем

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

Этим доказана формула обратная к формуле (17).

Теперь мы займёмся доказательством формулы (17). Пусть $(c, d) \in D(j_1, \dots, j_n)$. 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: $\{p, q\} \cap \{k, k + 1\} = \emptyset$. 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 $(c, d) \in D(i_1, \dots, i_n)$. Следовательно, $(c, d) \in D(i_1, \dots, i_n)\Delta\{(a, b)\}$.

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

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

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

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, $(c, d) \notin D(i_1, \dots, i_n)$ и мы опять получаем $(c, d) \in D(i_1, \dots, i_n)\Delta\{(a, b)\}$. □

Следствие 10.2 *Пусть* $j_1, \dots, j_n$ *— перестановка, полученная из перестановки* $i_1, \dots, i_n$ *перестановкой* k*-о и* $k + 1$*-о индексов, где* $1 \leq k < n$. *Тогда*

$$|D(j_1, \dots, j_n)| = |D(i_1, \dots, i_n)| \pm 1.$$

Определение 10.3 *Знак подстановки определятся следующим образом*

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

Важно заметить, что следствие 10.2 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 10.2 получаем

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

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 (15) ни верхняя ни нижняя строки матрицы на содержат инверсий. Вычислим знак транспозиции. Пусть $1 \leq a < b \leq n$. 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 10.4 $\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}\sigma \operatorname{sgn}\tau$.

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

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

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

$$\operatorname{sgn}\sigma \operatorname{sgn}\tau = (-1)^{|D(j_1,\dots,j_n)|+|D(k_1,\dots,k_n)|}(-1)^{|D(i_1,\dots,i_n)|+|D(j_1,\dots,j_n)|} =$$
$$= (-1)^{|D(i_1,\dots,i_n)|+2|D(j_1,\dots,j_n)|+|D(k_1,\dots,k_n)|} = (-1)^{|D(i_1,\dots,i_n)|+|D(k_1,\dots,k_n)|} = \operatorname{sgn}(\sigma\tau).$$

□

Следствие 10.5 $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}\sigma$.

Доказательство. По лемме 10.4 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 $\varepsilon = \pm 1$. □

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 10.4 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 10.4 представляет более практичный способ вычисления знака подстановки. Например, вычислим знак подстановки (16):

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

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 10.6 *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 $S_n = S_n^{even} \sqcup S_n^{odd}$, где S_n^{even} и S_n^{odd} — множества чётных и нечётных подстановок соответственно. Мы определим отображение $\varphi\colon S_n^{\mathrm{even}} \to S_n^{\mathrm{odd}}$ по формуле $\varphi(\sigma) = \tau\sigma$. Лемма 10.4 гарантирует, что φ корректно заданное отображение.

Рассмотрим отображение $\psi\colon S_n^{\mathrm{odd}} \to S_n^{even}$, заданное той же формулой $\psi(\sigma) = \tau\sigma$. Для любого $\sigma \in S_n^{even}$ 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 10.7 *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 также использовать 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 $n \leq 3$.

Пример 10.8 (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 10.9 (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 10.10 (Определитель матрицы 3×3) Мы имеем $S_3 = \{e, (1,2,3), (1,3,2), (1,2), (1,3), (2,3)\}$ и $\operatorname{sgn} e = \operatorname{sgn}(1,2,3) = \operatorname{sgn}(1,3,1) = 1$ и $\operatorname{sgn}(1,2) = \operatorname{sgn}(1,3) = \operatorname{sgn}(2,3) = -1$. Поэтому

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} -$$
$$= a_{1,2}a_{2,1}a_{3,3} - a_{1,3}a_{2,2}a_{3,1} - a_{1,1}a_{2,3}a_{3,2}.$$

10.2 Свойства определителя

Теорема 10.11 *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-о столбца) на фиксированный элемент кольца c. Тогда $\det A' = c\det A$.

(4) Пусть $i = 1, \dots, n$ ($j = 1, \dots, n$) — некоторый индекс такой, что $a_{i,j} = a'_{i,j} + a''_{i,j}$ для любого $j = 1, \dots, n$ ($i = 1, \dots, n$). Обозначим через A' и A'' матрицы, элементы которых вне i-й строки (j-о столбца) совпадают с соответствующими элементами матрицы A, а элемент стоящий на пересечении i-й строки и j-о столбца равен $a'_{i,j}$ и $a''_{i,j}$ соответственно для любого $j = 1, \dots, n$ ($i = 1, \dots, n$). 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 10.5 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 $i = 1, \dots, n$ такое, что $a_{i,j} = 0$ для всех $j = 1, \dots, n$. По определению 10.7 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 10.7 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 10.7 мы получаем

$$\det A = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma \;\; a_{1,\sigma(1)}\cdots a_{i-1,\sigma(i-1)}\left(a'_{i,\sigma i} + a''_{i,\sigma i}\right)a_{i+1,\sigma(i+1)}\cdots a_{\sigma(n),n} =$$

$$= \sum_{\sigma\in S_n} \operatorname{sgn}\sigma \;\; a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{i-1,\sigma(i-1)}\;a'_{i,\sigma i}\;a_{i+1,\sigma(i+1)}\cdots a_{\sigma(n),n} +$$

$$+ \sum_{\sigma\in S_n} \operatorname{sgn}\sigma \;\; a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{i-1,\sigma(i-1)}\;a''_{i,\sigma i}\;a_{i+1,\sigma(i+1)}\cdots a_{\sigma(n),n} =$$

$$= \det A' + \det A''.$$

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 $S_n = \coprod_{m=1}^{n!/2} H\sigma_m$ 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

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

$$= \sum_{i=m}^{\frac{n!}{2}} (\operatorname{sgn}\sigma_m \;\; a_{\sigma_m(1),1}a_{\sigma_m(2),2}\cdots a_{\sigma_m(n),n} +$$

$$+\operatorname{sgn}((i,j)\sigma_m)\ a_{(i,j)\sigma_m(1),1}a_{(i,j)\sigma_m(2),2}\cdots a_{(i,j)\sigma_m(n),n}).$$

Сумма в круглых скобках в правой части равна 0 при любом m, так как $\operatorname{sgn}\big((i,j)\sigma_m\big) = \operatorname{sgn}(i,j)\operatorname{sgn}\sigma_m = -\operatorname{sgn}\sigma_m$ 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 10.7.

(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 10.7.

(7) Пусть B получается из A умножением i-й строки на c и добавлением произведения к строке $k \neq i$. Тогда согласно частям (3)–(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 10.8 (Определитель блочной матрицы) *Пусть* $A = (a_{i,j})_{i,j=1}^{n}$ *— матрица и* $m = 1, \dots, n-1$ *такие, что* $a_{i,j} = 0$ *при* $i \le m < j$ *или при* $j \le m < i$. *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 $i \le m < j$. Воспользуемся определением 10.7 для вычисления $\det A$. Пусть $\sigma \in S_n$ — такой элемент, что

$$\sigma(\{1, \dots, m\}) \ne \{1, \dots, m\}$$

В этом случае $\sigma(i) > m$ для некоторого $i = 1, \dots, m$. 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

$$\det A = \sum_{\sigma \in S_{n,m}} \operatorname{sgn} \sigma \;\; a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}, \tag{20}$$

где $S_{n,m}$ — множество подстановок из S_n таких, что

$$\sigma(\{1, \dots, m\}) = \{1, \dots, m\}.$$

Заметим, что для каждой такой подстановки $\sigma(\{m+1, \dots, n\}) = \{m+1, \dots, n\}$. 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 $i \le m$.

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

$$\varphi_1(\tau)(i) = \begin{cases} \tau(i), & \text{если } i \le m; \\ i, & \text{если } i > m, \end{cases} \quad \varphi_2(\delta)(i) = \begin{cases} \delta(i-m) + m, & \text{если } i > m; \\ i, & \text{если } i \le m. \end{cases}$$

Легко понять, что любая подстановка $\sigma \in S_{n,m}$ однозначно представляется в виде $\sigma = \varphi_1(\tau)\varphi_2(\delta)$, где $\tau_1 \in S_m$ и $\delta \in S_{n-m}$.

Мы утверждаем, что

$$\operatorname{sgn} \varphi_1(\tau) = \operatorname{sgn} \tau \;\; \forall\, \tau \in S_m \text{ и } \operatorname{sgn} \varphi_2(\delta) = \operatorname{sgn} \delta \;\; \forall\, \delta \in S_{n-m}. \tag{21}$$

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 (21).

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 (21).

Теперь используя формулы (20) и (21), определения гомоморфизмов φ_1 и φ_2 и лемму 10.4, мы можем записать

$$\det A = \sum_{\tau\in S_n, \delta\in S_{n-m}} \operatorname{sgn}(\varphi_1(\tau)\varphi_2(\delta)) \times$$
$$\times\, a_{1,\varphi_1(\tau)(1)} \cdots a_{m,\varphi_1(\tau)(m)} a_{m+1,\varphi_2(\delta)(m+1)} \cdots a_{n,\varphi_2(\delta)(n)} =$$

$$= \sum_{\tau\in S_n, \delta\in S_{n-m}} (\operatorname{sgn}\tau)(\operatorname{sgn}\delta) a_{1,\tau(1)} \cdots a_{m,\tau(m)} a_{m+1,\delta(1)+m} \cdots a_{n,\delta(n-m)+m} =$$

$$= \left(\sum_{\tau\in S_n} \operatorname{sgn}\tau\, a_{1,\tau(1)} \cdots a_{m,\tau(m)} \right) \left(\sum_{\delta\in S_{n-m}} \operatorname{sgn}\delta\, a_{m+1,\delta(1)+m} \cdots a_{n,\delta(n-m)+m} \right) =$$
$$= \det A' \det A''.$$

□

Следствие 10.13 *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 10.12 и формулу для определителя матрицы размера 1×1 (смотрите пример 10.8). □

Следствие 10.14 *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 7.8, частей (6), (3) и (7) теоремы 10.11 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 10.13 и части (6) теоремы 10.11 (для первого равенства).

Определение 10.15 *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 7.14 матрица вырождена тогда и только тогда, когда вырождена её транспонированная матрица.

Лемма 10.16 *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 $\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_n$ раны 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 10.11 мы имеем $\det A = \det B$. Однако $\det B = 0$ согласно части (2) теоремы 10.11.

Предположим наоборот, что $\det A = 0$. По следствию 7.9 мы получаем представление $A = X_1 \cdots X_k B$, где $X_1, \dots, X_k$ — 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 10.14 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 10.13 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 10.17 *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 7.11 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 10.18 *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 7.9 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 10.19 *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

$$\det C = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma \prod_{i=1}^n c_{i,\sigma(i)} = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma \prod_{i=1}^n \sum_{j=1}^n a_{i,j} b_{j,\sigma(i)}. \qquad (22)$$

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 $\{1, \dots, n\}$ 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 (22), получаем

$$\det C = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma \sum_{\tau\in X_n} \left(\prod_{i=1}^n a_{i,\tau(i)}\right)\left(\prod_{k=1}^n b_{\tau(k),\sigma(k)}\right) = \\ = \sum_{\tau\in X_n} \sum_{\sigma\in S_n} \operatorname{sgn}\sigma \left(\prod_{i=1}^n a_{i,\tau(i)}\right)\left(\prod_{k=1}^n b_{\tau(k),\sigma(k)}\right) = \\ = \sum_{\tau\in X_n} \left(\prod_{i=1}^n a_{i,\tau(i)}\right)\left(\sum_{\sigma\in S_n} \operatorname{sgn}\sigma \prod_{k=1}^n b_{\tau(k),\sigma(k)}\right).$$

Сначала рассмотрим случай, когда τ не является подстановкой, то есть $\tau \in X_n \backslash S_n$. В этом случае $\tau(s) = \tau(r)$ для различных индексов $r, s = 1, \dots, n$. Рассмотрим подгруппу $H = \{e, (s, r)\}$ и разбиение $S_n = \coprod_{m=1}^{n!/2} \sigma_m H$ на левые смежные классы. Отсюда мы получаем

$$\sum_{\sigma\in S_n} \operatorname{sgn}\sigma \prod_{k=1}^n b_{\tau(k),\sigma(k)} = \\ = \sum_{m=1}^{\frac{n!}{2}} \left(\operatorname{sgn}\sigma_m \prod_{k=1}^n b_{\tau(k),\sigma_m(k)} + \operatorname{sgn}\left(\sigma_m(r,s)\right) \prod_{k=1}^n b_{\tau(k),\sigma_m(r,s)(k)}\right) =$$

$$= \sum_{m=1}^{\frac{n!}{2}} \operatorname{sgn} \sigma_m \left(\prod_{k=1}^{n} b_{\tau(k),\sigma_m(k)} - \prod_{k=1}^{n} b_{\tau(k),\sigma_m(r,s)(k)} \right) =$$
$$= \sum_{m=1}^{\frac{n!}{2}} \operatorname{sgn} \sigma_m \left(\prod_{k=1, k\neq r,s}^{n} b_{\tau(k),\sigma_m(k)} \right) \times$$
$$\times \left(b_{\tau(k),\sigma_m(k)} b_{\tau(r),\sigma_m(r)} b_{\tau(s),\sigma_m(s)} - b_{\tau(r),\sigma_m(s)} b_{\tau(s),\sigma_m(r)} \right) = 0.$$

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:
$$\mathbf{A} = \begin{pmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,n} \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} y_{1,1} & \cdots & y_{1,n} \\ \vdots & \ddots & \vdots \\ y_{n,1} & \cdots & y_{n,n} \end{pmatrix}.$$
Наша цель доказать равенство
$$det(\mathbf{AB}) = \det \mathbf{A}\, det\ \mathbf{B}, \tag{23}$$
которое является равенством двух многочленов из кольца K, которое содержится в следующем кольце:
$$\mathbb{Q}[x_{i,j}, y_{i,j}]_{i,j=1}^{n}.$$
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 10.17. По лемме 10.16 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 $X_1, \dots, X_k$ — матрицы как в лемме 10.18. Теперь по следствию 10.14 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

$$\det(AB) = \det(X_1 \cdots X_k B) = \det X_1 \det(X_2 \cdots X_k B) = \cdots$$
$$= \det X_1 \det X_2 \cdots \det X_k \det B.$$

Сравнивая, получаем $\det(AB) = \det A \det B$. Этим равенство (23) 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 $i, j = 1, \dots, n$. Применяя этот гомоморфизм к доказанному равенству (23), мы получаем утверждение теоремы. □

10.3 Разложение определителя по строке и столбцу

Пусть A — квадратная матрица размера $n \times n$ и $i, j = 1, \dots, n$. 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 10.20 (Разложение по строке и столбцу) *Пусть $A = (a_{i,j})_{i,j}^n$ — квадратная матрица размера $n \times n$. Тогда для любого $i = 1, \dots, n$ выполнено (разложение по строке)*

$$\det A = \sum_{j=1}^n a_{i,j} A_{i,j}$$

и для любого $j = 1, \dots, n$ 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 10.11), мы получаем

$$\det A=\sum_{j=1}^{n}a_{i,j}A^{i,j}, \tag{24}$$

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 10.11 и теоремой 10.12, 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 (24), 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 10.21 (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 10.11 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 10.20 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 10.22 (Формула обратной матрицы) *Пусть* A *— квадратная матрица над полем размера* $n \times n$ *такая, что* $detA \neq 0$*. 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

$$A \begin{pmatrix} A_{1,1} & A_{2,1} & \cdots & A_{n,1} \\ A_{1,2} & A_{2,2} & \cdots & A_{n,2} \\ \vdots & \vdots & & \vdots \\ A_{1,n} & A_{2,2} & \cdots & A_{n,n} \end{pmatrix} = (\det A) E, \begin{pmatrix} A_{1,1} & A_{2,1} & \cdots & A_{n,1} \\ A_{1,2} & A_{2,2} & \cdots & A_{n,2} \\ \vdots & \vdots & & \vdots \\ A_{1,n} & A_{2,2} & \cdots & A_{n,n} \end{pmatrix} A = (\det A) E.$$

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 10.20, когда $k = i$, и лемме 10.21, 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. □

Следствие 10.23 *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 10.22. Пусть наоборот A обратима. Тогда по теореме 10.19 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 10.19 получаем $\det A \neq 0$ (смотрите вычисления выше). По теореме 10.22 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 Список литературы

1. Математика в экономике. Ч.1: Линейная алгебра, аналитическая геометрия и линейное программирование: Учебник для студ. экономич. спец. вузов / А.С. Солодовников, В.А. Бабайцев, А.В. Браилов, И.Г. Шандра. — 3-е изд., перераб. и доп. — М.: Финансы и статистика; ИНФРА-М, 2011.
2. Сборник задач по курсу «Математика в экономике». В 3-х ч. Ч.1. Линейная алгебра, аналитическая геометрия и линейное программирование: учеб. пособие / под ред. В.А. Бабайцева и В.Б. Гисина. – М.: Финансы и статистика; ИНФРА-М, 2013.
3. С.В. Пчелинцев, Вопросы и задачи по линейной алгебре : для студентов все специальностей / Финансовая академия при Правительстве РФ — М. : Финакадемия, 2006 — 37с.

4. Э.Б. Винберг, Линейные представления групп — М. : Наука. Гл. ред. физ.-мат. лит., 1985 — 142 с.