Экзамен по алгебре

Лектор: Гайфуллин С. А. • Автор: Пшеничный Никита[*], группа 109

I курс • Осенний семестр 2023 г.

Аннотация

При подготовке этого файла я использовал: курс лекций С. А. Гайфуллина, записи лекций Ю. Г. Прохорова на *teach-in*, книгу «Курс алгебры» Э. Б. Винберга и курс семинаров А. А. Клячко.

Экзаменационные вопросы

[*]Telegram: @pshenikita

1 Системы линейных уравнений. Матрица коэффициентов и расширенная матрица коэффициентов. Элементарные преобразования. Эквивалентные системы. Элементарные преобразования переводят СЛУ в эквивалентную

Определение 1.1. Матрицей размера $m \times n$ над полем $\mathcal{K}$ называется прямоугольная таблица из элементов поля $\mathcal{K}$, имеющая m **строк** и n **столбцов**:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Определение 1.2. Линейным уравнением с неизвестными $x_1, x_2, \dots, x_n$ над полем $\mathcal{K}$ называется уравнение вида

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b,$$

где **коэффициенты** $a_1, a_2, \dots, a_n$ и **свободный член** b являются элементами поля $\mathcal{K}$.

Система m линейных уравнений с n неизвестными в общем виде пишется так:

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

Определение 1.3. Матрица

$$\mathcal{A} := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

называется **матрицей коэффициентов**, а матрица

$$\widetilde{\mathcal{A}} := \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right)$$

— **расширенной матрицей коэффициентов** системы $(*)$.

Определение 1.4. Решение системы — это элемент поля $\mathcal{K}^n$, который при подстановке вместо $(x_1, \dots, x_n)$ обращает каждое уравнение в верное равенство. Система уравнений называется **совместной**, если она имеет хотя бы одно решение, и **несовместной** в противном случае. Совместная система называется **определённой**, если её решение единственно.

Определение 1.5. Элементарными преобразованиями системы линейных уравнений называются преобразования следующих трёх типов:

1. прибавление к одному уравнению другого, умноженного на число;
2. перестановка двух уравнений;
3. умножение одного уравнения на число, отличное от нуля.

Элементарными преобразованиями строк матрицы называются преобразования следующих трёх типов:

1. прибавления к одной строке другой, умноженной на число;

2. перестановка двух строк;
3. умножение одной строки на число, отличное от нуля.

Определение 1.6. Две СЛУ называются **эквивалентными**, если множества их решений совпадают. Обозначение для расширенных матриц: $\widetilde{\mathcal{A}} \sim \widetilde{\mathcal{B}}$.

Теорема 1.1. Если одну систему можно перевести в другую элементарными преобразованиями, то они эквивалентны.

Докажем сначала вспомогательное утверждение:

Лемма 1.1. Преобразование, обратное к элементарному, тоже элементарное.

Доказательство. Докажем для каждого преобразования:

1. («прибавить i-ое уравнение к j-му с коэффициентом λ»)$^{-1}$ = (прибавить i-ое к j-му с коэффициентом $-\lambda$»)
2. Преобразование второго типа обратно к самому себе
3. («умножить i-ое уравнение на $c \neq 0$»)$^{-1}$ = («умножить i-ое уравнение на $c^{-1} \neq 0$»)

■

Теперь докажем теорему:

Доказательство. Докажем, что каждое решение первой системы является решением второй. Это очевидно для преобразований второго и третьего типа, докажем и для первого:

$$(a_{j1} + \lambda a_{i1})x_1 + \ldots + (a_{jn} + \lambda a_{in})x_n = \underbrace{a_{j1}x_1 + \ldots + a_{jn}x_n}_{=0} + \lambda \underbrace{(a_{i1}x_1 + \ldots + a_{in}x_n)}_{=0} = 0.$$

Далее, из леммы 1.1 обратное к элементарному преобразвоание тоже элементарное. Поэтому всякое решение второй системы является решением первой. Значит, множества решений этих систем совпадают, т. е. они эквивалентны. ■

Примечание. Обратное утверждение неверно — если СЛУ эквивалентны, то они не всегда переводятся друг в друга элементарными преобразованиями. Контрпример — две системы с пустыми множествами решений:

$$A = \left(\begin{array}{cc|c} 0 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right) \equiv B = \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 0 & 1 \end{array}\right).$$

Никаким элементарным преобразованием строк нельзя занулить первый столбец матрицы $\mathcal{B}$.

2 Экзотические уравнения. Свободные и главные переменные. Ступенчатый и улучшенный ступенчатый вид матрицы. Метод Гаусса решения СЛУ

Определение 2.1. Лидером (**ведущим элементом**) ненулевой строки $(a_1, a_2, \ldots, a_n) \in \mathcal{K}^n$ её первый ненулевой элемент.

Определение 2.2. Матрица называется **ступенчатой**, если номера ведущих элементов её ненулевых строк образуют строго возрастающую последовательность, а нулевые строки расположены в конце:

$$\begin{pmatrix} * & * & * & * & * \\ & * & * & * & * \\ & & & * & * \\ & & 0 & & * \end{pmatrix}$$

Матрица называется улучшенной ступенчатой, если она ступенчатая, её ведущие элементы равны 1, а элементы над ними равны 0:

$$\begin{pmatrix} \boxed{1} & 0 & * & 0 & 0 \\ & \boxed{1} & * & 0 & 0 \\ & & & \boxed{1} & 0 \\ & & 0 & & \boxed{1} \end{pmatrix}$$

Теорема 2.1. Всякую матрицу путём элементарных преобразований строк можно привести к ступенчатому виду.

Доказательство. Если данная матрица нулевая, то она уже ступенчатая. Если она ненулевая, то пусть j_1 — номер её первого ненулевого столбца. Переставив, если нужно, строки, добьёмся того, чтобы $a_{1j_1} \neq 0$. После этого прибавим к каждой строке, начиная со второй, первую строку, умноженную на подходящее число, с таким расчётом, чтобы все элементы j_1-го столбца, кроме первого, стали равными нулю. Теперь рассмотрим матрицу без первой строки и первый j_1 столбцов. Приведём её таким же методом к ступенчатому виду. Продолжая процесс таким же образом, мы получим ступенчатую матрицу. ■

Процесс, проводимый нами в доказательстве последней теоремы, называется *прямым ходом метода Гаусса*.

Примечание. Для приведения матрицы к ступенчатому виду достаточно преобразований первого типа. Действительно, преобразования второго типа нужны были нам лишь для того, чтобы поднять на первую строчку ненулевой элемент. Однако, эту задачу можно выполнить преобразованием первого типа. Пусть $a_{i_1 j_1} \neq 0$. Тогда прибавим i_1 строку расширенной матрицы к первой. Теперь получаем $a_{0j_1} = a_{i_1 j_1} \neq 0$. А далее действуем так же.

Теперь приведём расширенную матрицу коэффициентов СЛУ к улучшенному ступенчатому виду. По теореме 1.1 множество её решений не изменилось, а решать её в таком виде куда проще.

Определение 2.3. Пусть $j_1, j_2, \ldots, j_r$ — номера ведущих коэффициентов ненулевых уравнений системы. Неизвестные $x_{j_1}, x_{j_2}, \ldots, x_{j_r}$ назовём **главными**, а остальные — **свободными**.

Определение 2.4. Уравнения вида

$$0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = b,$$

где $b \neq 0$ назовём **экзотическими**.

Теорема 2.2. Для каждого набора значений свободных переменных существует единственный набор значений главных переменных, который дополняет данный набор до решения совместной системы.

Доказательство. Идём по строкам ступечатой расширенной матрицы коэффициентов СЛУ снизу вверх. Рассмотрим первую встреченную нами ненулевую строку, её номер i_r. В соответствующем уравнении ровно одна главная переменная — x_{j_r}. Её можно выразить через свободные переменные и свободный член уравнения единственным образом. Теперь возьмём следующую строку. Её номер i_{r-1}, и в ней ровно две главные неизвестные — x_{j_r} и $x_{j_{r-1}}$. Выразим $x_{j_{r-1}}$ через x_{j_r}, свободные переменные и свободный член. А теперь подставим выражения x_{j_r} через свободные переменные, полученное на предыдущем шаге. Получается, $x_{j_{r-1}}$ тоже выражается через свободный член и свободные переменные единственным образом. Продолжив процесс, получим выражения для всех главных переменных через свободные. ■

Процесс, проводимый нами в доказательстве последней теоремы, называется *обратным ходом метода Гаусса*.

Примечание. Как следствие, любая СЛУ над бесконечным полем имеет либо ни одного, либо одно, либо бесконечно много решений. А над конечным — либо ни одного, либо одно, либо $|\mathcal{K}|^{n-r}$, где r — количество ступенек в ступенчатой матрице (иными словами, $n - r$ — количество свободных переменных). Доказательство того, что это число r определенно корректно, будет позднее.

3 Единственность улучшенного ступенчатого вида матрицы. Понятие ранга матрицы (через ступенчатый вид) и его корректность

Теорема 3.1. Каждая матрица имеет единственный улучшенный ступенчатый вид.

Доказательство. Рассмотрим матрицу A. Допустим, она может быть приведена элементарными преобразованиями к двум улучшенным ступенчатым видам B и C. Наша цель — доказать, что $B = C$. Рассмотрим однородную систему уравнений с матрицей коэффициентов A. Расширенная матрица коэффициентов $\widetilde{A} = (A \mid 0)$ приводится теми же элементарными преобразованиями к виду ступенчатым видам $\widetilde{B} = (B \mid 0)$ и $\widetilde{C} = (C \mid 0)$. Это означает, что однородные системы с матрицами коэффициентов B и C эквивалентны. Пусть $B \neq C$. Отбросим нулевые строки в матрицах $\widetilde{B}$ и $\widetilde{C}$, при этом системы останутся эквивалентными. Будем идти по строкам матриц $\widetilde{B}$ и $\widetilde{C}$, пока не дойдём до места, где будет различие. Попадаем в один из трёх случаев:

1. *Строки с первой по $(k-1)$-ую снизу в матрицах $\widetilde{B}$ и $\widetilde{C}$ совпадают, а лидеры k-ой строки снизу в матрицах $\widetilde{B}$ и $\widetilde{C}$ имеют различные позиции.* Пусть лидеры стоят в столбцах p и q соответственно. Не ограничивая общности, можем считать, что $p < q$. Из того, что строки ниже, чем k-ые снизу у матриц $\widetilde{B}$ и $\widetilde{C}$ совпадают, следует, что разделение переменный x_i при $i > q$ на главные и свободные в системах $\widetilde{B}$ и $\widetilde{C}$ одинаково. Положим все свободные переменные с номерами $> q$ равными нулю. В системе $(C \mid 0)$ переменная x_q главная и, следвоательно (т. к. система однородная) при указанному задании переменных она также обязательно равна нулю. В системе $(B \mid 0)$ переменная x_q свободная, а потому при указанному задании переменных она может принимать значение 1. Таким образом, нашли решение одной системы, которое не является решением другой. Противоречие с эквивалентностью систем $(B \mid 0)$ и $(C \mid 0)$.

2. *Строки с первой по $(k-1)$-ую снизу в матрицах $\widetilde{B}$ и $\widetilde{C}$ совпадают, лидеры k-ой строки снизу у этих матриц имеют одинаковые позиции p, но есть номер $s > p$ такой, что в s-ом столбце в рассматриваемой строке у матриц $\widetilde{B}$ и $\widetilde{C}$ стоят различные числа.* Пусть эти числа b и c соответственно. Не ограничивая общности, $b \neq 0$. Тогда x_s — свободная переменая для системы $(B \mid 0)$, т. к. в столбцах, соответствующих главным переменным стоят нули за счёт улучшенного ступенчатого вида. А значит, x_s — свободная переменнная и для $(C \mid 0)$. Положим все свободные переменные, кроме x_s, равными нулю, а $x_s = 1$. Тогда из системы $(B \mid 0)$ получаем, что $x_s = -b$, а из системы $(C \mid 0)$ — что $x_s = -c$, но $b \neq c$. Противоречие.

3. *Проходя снизу вверхн по ненулевым строкам матриц $\widetilde{B}$ и $\widetilde{C}$, мы всё время видели, что очередные строки совпадают, но строки в одной из матриц закончились, а в другой — нет.* Пусть строки закончились в матрице $\widetilde{B}$, тогда в матрице $\widetilde{C}$ есть ещё одна нулевая строка. Из того, что строки ниже, чем данная, у матриц $\widetilde{B}$ и $\widetilde{C}$ совпадают, следует, что разделение переменных x_i при $i > q$ на главные и свободные в соответствующих системах одинаково. Но в системе $(B \mid 0)$ переменная x_q свободная (нет строки, где лидер имеет позицию q), а в системе $(C \mid 0)$ — свободная. Далее можем поступить так же, как и в первом случае.

■

Примечание. Отсюда следует, что количество ненулевых строк в любом ступенчатом виде данной матрицы одинаково. Действительно, ведь для любого ступенчатого вида можно построить улучшенный ступенчатый с таким же количеством ненулевых строк — достаточно каждую ненулевую строку разделить на её ведущий элемент, а затем вычесть её из всех строк выше с подходящим коэффициентом. А ступенчатый вид единственный.

Определение 3.1 (Ступенчатый ранг матрицы). Количество ненулевых строк в ступенчатом виде данной матрицы A называется **рангом матрицы** A и обозначается $\operatorname{rk} A$.

4 Векторное пространство. Простейшие свойства из аксиом. Подпространство. Критерий того, что подмножество является подпространством

Определение 4.1. Векторным пространством над полем $\mathcal{K}$ называется такое множество V с операциями слоежния и умножения на элементы поля $\mathcal{K}$, обладающими следующими свойствами:

$$\left.\begin{array}{l} 1.\ \forall a,b,c \in V\ (a+b)+c = a+(b+c) \\ 2.\ \exists \mathbf{0} \in V : \forall v \in V\ v+\mathbf{0} = v \\ 3.\ \forall v \in V\, \exists (-v) \in V : (-v)+v = \mathbf{0}. \\ 4.\ \forall a,b \in V\ a+b = b+a \end{array}\right\} V \text{ — абелева группа по } +$$

5. $\forall \lambda \in \mathbb{R}, u,v \in V\ \lambda \cdot (u+v) = \lambda \cdot u + \lambda \cdot v$
6. $\forall \lambda,\mu \in \mathbb{R}, v \in V\ (\lambda+\mu)\cdot v = \lambda \cdot v + \mu \cdot v$
7. $\forall \lambda,\mu \in \mathbb{R}, v \in V\ (\lambda\mu)\cdot v = \lambda \cdot (\mu \cdot v)$
8. $\forall v \in V\ 1 \cdot v = v$

Примечание. Докажем, что аксиома 4 выводится через остальные. Сделаем это в несколько шагов (над знаками равенства указан номер аксиомы или пункта доказательства, благодаря которой сделан переход):

$1^*.\ 0 \cdot v = (0+0)\cdot v \overset{6}{=} 0\cdot v + 0\cdot v \Rightarrow \boxed{0\cdot v = \mathbf{0}}$

$2^*.\ (-1)\cdot v + v \overset{8}{=} (-1\cdot v) + 1\cdot v \overset{6}{=} (-1+1)\cdot v = 0\cdot v \overset{1^*}{=} \mathbf{0} \Rightarrow \boxed{(-1)\cdot v = -v}$

$3^*.\ v - v \overset{8,2^*}{=\!=} 1\cdot v + (-1)\cdot v \overset{6}{=} (1-1)\cdot v = 0\cdot v = \mathbf{0} \Rightarrow \boxed{v - v = \mathbf{0}}$

Теперь выведем 4 аксиому из остальных:

$$u+v \overset{2}{=} (u+v)+\mathbf{0} \overset{3}{=} (u+v)+(-(v+u)+(v+u)) \overset{1}{=} ((u+v)-(v+u))+(v+u) =$$
$$= (u+\underbrace{v-v}_{=0}-u)+(v+u) \overset{3^*}{=} \underbrace{(u-u)}_{=0}+(v+u) = v+u.$$

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

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

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

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

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

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

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

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

5 : $\lambda \star (a+b, x*y) = (\lambda(a+b), x*y) = (\lambda a + \lambda b, x*y) = (\lambda a, x) + (\lambda b, y) = \lambda \star ((a,x)+(b,y))$
6 : $(\lambda+\mu)\cdot(a,x) = ((\lambda+\mu)a, x) = (\lambda a + \mu a, x*x) = (\lambda a, x) + (\mu a, x) = \lambda \star (a,x) + \mu \star (a,x)$
7 : $(\lambda\mu) \star (a,x) = (\lambda\mu a, x) = \lambda \star (\mu a, x)$
8 : $1 \star (a,x) = (1\cdot a, x) = (a,x)$

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

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

  Выполнение аксиом 1, 2 и 4 сразу вытекает из определения. Аксиома 3 не выполнена, т. к. у ∞ нет обратного по $+$. Выполнение аксиом 5-8 проверяется перебором нескольких случаев.
- **6 аксиома**. Рассмотрим алгебраическую систему $V = (\mathbb{R}, +, \star)$, в которой сложение определено так же, как в действительных числах, а умножение так:

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

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

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

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

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

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

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

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

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

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

  $$\lambda \star u := \mathbf{0}.$$

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

Теорема 4.1. Множество строк $\mathcal{K}^n$ является векторным пространством.

Доказательство. Легко видеть, что для него выполняется каждая из аксиом. ■

Теорема 4.2 (Свойства векторного пространства).

1. В любом векторном пространстве нулевой вектор единственный
2. Для любого вектора v противоположный вектор $-v$ единственный
3. Для любого вектора выполнено $0 \cdot v = \mathbf{0}$
4. Для любого числа λ выполнено $\lambda \cdot \mathbf{0} = \mathbf{0}$
5. Для любого вектора v выполнено $(-1) \cdot v = -v$
6. Для любого вектора v выполнено $-(-v) = v$

7. Для любых векторов u и v $-(u+v) = (-u) + (-v)$
8. $\lambda v = \mathbf{0} \Rightarrow (\lambda = 0) \wedge (v = \mathbf{0})$

Доказательство.

1. Пусть есть два нулевых вектора — $\mathbf{0}_1$ и $\mathbf{0}_2$. Тогда $\mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_1$ с одной стороны (т. к. $\mathbf{0}_2$ нулевой) и $\mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_2$ (т. к. $\mathbf{0}_1$ нулевой). Отсюда $\mathbf{0}_1 = \mathbf{0}_2$.
2. Пусть таких векторов два: $(-v)_1$ и $(-v)_2$. Тогда имеем
$$(-v)_1 = \underbrace{(-v)_1 + v}_{=\mathbf{0}} + (-v)_2 = (-v)_2.$$
3. Доказывалось ранее.
4. $\lambda \cdot \mathbf{0} = \lambda \cdot (\mathbf{0} + \mathbf{0}) = \lambda \cdot \mathbf{0} + \lambda \cdot \mathbf{0}$. Вычитая $\lambda \cdot \mathbf{0}$ из обеих частей равенства, получаем требуемое.
5. Доказывалось ранее.
6. $-(-v) + (-v) = (-1) \cdot (-v) + (-1) \cdot v = (-1) \cdot \underbrace{(-v + v)}_{=\mathbf{0}} = \mathbf{0}$
7. $-(u+v) = (-1) \cdot (u+v) = (-1)u + (-1)v = (-u) + (-v)$.
8. Допустим, $\lambda \neq 0$. Тогда существует число $1/\lambda \neq 0$. Имеем
$$\mathbf{0} = \frac{1}{\lambda}\mathbf{0} = \frac{1}{\lambda}(\lambda v) = \left(\frac{1}{\lambda} \cdot \lambda\right) v = v.$$

■

Определение 4.2. Подмножество U векторного пространства V над полем $\mathcal{K}$ называется **подпространством** в V, если оно является векторным пространством над полем $\mathcal{K}$.

Теорема 4.3 (Критерий подпространства). Подмножество U векторного пространства V над полем $\mathcal{K}$ является подпространством тогда и только тогда, когда:

1. $\forall u_1, u_2 \in U \ (u_1 + u_2) \in U$
2. $\forall u \in U, \lambda \in \mathcal{K} \ \lambda u \in U$

Доказательство. $\Rightarrow$. Пусть U — подпространство в V. Тогда сумма двух его элементов лежит снова в U и умножение любого элемента на число лежит в U.

$\Leftarrow$. Наоборот, пусть выполнены условия теоремы. Тогда выполнение аксиом 1, 4-8 сразу следует из того, что они выполнены над V. Чтобы доказать справедливость аксиомы 2, нужно, чтобы $\mathbf{0} \in V$. Это так в силу того, что $U \ni 0 \cdot u = \mathbf{0}$. А для выполнения аксиомы 3 требуется, что $\forall u \in U \ (-u) \in U$. Это выполнется, т. к. $U \ni (-1) \cdot u = -u$. ■

5 Понятие линейной зависимости системы векторов. Три свойства линейной зависимости

Определение 5.1. Пусть $v_1, \ldots, v_k$ — векторы из векторного пространства V, а $\lambda_1, \ldots, \lambda_k$ — элементы поля $\mathcal{K}$. Тогда **линейной комбинацией** векторов $v_1, \ldots, v_k$ с коэффициентами $\lambda_1, \ldots, \lambda_k$ — это выражение
$$\lambda_1 v_1 + \ldots + \lambda_k v_k.$$

При $\lambda_1 = \ldots = \lambda_k = 0$ линейная комбинация называется **тривиальной**.

Определение 5.2. Конечная система векторов называется **линейно зависимой**, если существует нетривиальная нулевая линейная комбинация. Бесконечная система векторов называется **линейно зависимой**, если в ней можно выделить конечную линейно зависимую подсистему. Система векторов называется **линейно независимой**, если она не линейно зависима.

Теорема 5.1 (Три свойства линейной зависимости).

1. Пусть $\mathcal{S}$ и $\mathcal{S}'$ — две системы векторов, причём $\mathcal{S} \subseteq \mathcal{S}'$. Тогда если $\mathcal{S}$ линейно зависима, то и $\mathcal{S}'$ линейно зависима.
2. Система V линейно зависима тогда и только тогда, когда в ней есть вектор v, который линейно выражается через остальные.
3. Пусть система V линейно независима, а система $V \cup \{u\}$ линейно зависима. Тогда существует единственный набор элементов поля $\mathcal{K}$: $\mu_1, \ldots, \mu_k$ такой, что $u = \mu_1 v_1 + \ldots + \mu_k v_k$, где $\{v_1, \ldots, v_k\} \subseteq V$ (равенство достигается для конечной системы).

Примечание. Свойство 2, как видно, из формулировки, является критерием линейной зависимости.

Доказательство. Докажем сначала для конечных систем:

1. $\mathcal{S} = \{v_1, \ldots, v_k\}$ линейно зависима, поэтому в ней есть вектор u, выражающийся линейно через некоторые векторы $v_1, \ldots, v_k$. Однако заметим, что $\{u, v_1, \ldots, v_k\} \subseteq \mathcal{S} \subseteq \mathcal{S}'$. Поэтому u выражается через векторы системы $\mathcal{S}'$: коэффициенты при $v_1, \ldots, v_k$ не меняем, а коэффициенты при остальных векторах выбираем нулевыми.
2. $V = \{v_1, \ldots, v_k\}$ линейно зависима, значит, существует нетривиальная линейная комбинация, такая что
$$\lambda_1 v_1 + \ldots + \lambda_k v_k = 0, \quad (\lambda_1, \ldots, \lambda_k) \neq (0, \ldots, 0).$$
Существует j такой, что $\lambda_j \neq 0$, поэтому можно выразить v_j:
$$v_j = \sum_{i \neq j} \left(-\frac{\lambda_i}{\lambda_j}\right) v_i.$$
Обратно, пусть существует вектор $v_j = \sum\limits_{i \neq j} \lambda_i v_i$. Тогда
$$\lambda_1 v_1 + \ldots + (-1) v_j + \ldots + \lambda_k v_k = 0.$$
Эта линейная комбинация нетривиальная, т. к. $-1 \neq 0$.
3. Существует нетривиальная линейная комбинация
$$\lambda_1 v_1 + \ldots + \lambda_k v_k + \mu u = 0 \qquad (*)$$
При этом $\mu \neq 0$, т. к. иначе останется линейная комбинация $\lambda_1 v_1 + \ldots + \lambda_k v_k = 0$, а она обязательно тривиальная (т. к. система $\{v_1, \ldots, v_k\}$ линейно независима), а поэтому и комбинация $(*)$ тривиальная. Итак, $\mu \neq 0$, поэтому можно так же, как в пункте 2, выразить его через остальные:
$$u = \sum_i \left(-\frac{\lambda_i}{\mu}\right) v_i.$$

А теперь — для бесконечных:

1. $\mathcal{S}$ линейно зависима, а значит, в ней можно выделить конечную линейно зависимую подсистему. А из того, что $\mathcal{S}' \supseteq \mathcal{S}$, в $\mathcal{S}'$ можно выделить эту же линейно зависимую подсистему
2. Выделим в V конечную линейно зависимую систему и применим утверждение теоремы для неё.
3. Выделим в $V \cup \{u\}$ линейно зависимую подсистему и применим утверждение теоремы для неё.

■

6 Однородные системы с количеством неизвестных большим количества уравнений. Основная лемма о линейной зависимости

Определение 6.1. СЛУ называется однородной, если свободные члены во всех уравнениях равны нулю.

Примечание. Однородные СЛУ всегда совместны, в качестве решения подходит строка $(0, \ldots, 0) \in \mathcal{K}^n$.

Лемма 6.1. Если в однородной системе неизвестных больше, чем уравнений, то она имеет ненулевое решение.

Доказательство. Количество главных переменных равно количеству лидеров, а оно, в свою очередь, не превышает количества уравнений (строк матрицы коэффициентов), а оно меньше количества неизвестных. Значит, количество главных переменных меньше количества всех переменных, значит, среди переменных есть свободные, т. е. решений этой системы бесконечно много. Среди них можно выбрать ненулевое. ■

Лемма 6.2 (Основная лемма о линейной зависимости). Пусть $\{v_1, \ldots, v_n\}$ и $\{u_1, \ldots, u_m\}$ — две системы векторов из V. Допустим, что каждый вектор v_i линейно выражается через систему векторов $\{u_1, \ldots, u_m\}$ и при этом $n > m$. Тогда система $\{v_1, \ldots, v_n\}$ линейно зависима.

Доказательство. Из условия, существуют такие λ_{ij}, что

$$v_i = \sum_{j=1}^{m} \lambda_{ij} u_j.$$

Составим линейную комбинацию

$$\sum_{i=1}^{n} \mu_i v_i = \sum_{i=1}^{n} \left(\mu_i \sum_{j=1}^{m} \lambda_{ij} u_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{m} (\mu_i \lambda_{ij} u_j) = \sum_{j=1}^{m} \sum_{i=1}^{n} (\mu_i \lambda_{ij} u_j) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} \mu_i \lambda_{ij} \right) u_j.$$

Эта линейная комбинация (уж точно) равна нулю, если $\sum\limits_{i=1}^{n} \mu_i \lambda_{ij} = 0 \ \forall j$. Докажем, что мы можем подобрать подходящие μ_i. Имеем однородную СЛУ с переменными μ_i (n штук) и коэффициентами λ_{ij} (m штук), причём $n > m$, значит (по предыдущей лемме), у этой системы есть ненулевое решение. Итак, существуют такие коэффициенты μ_i, не все равные нулю, что $\mu_1 v_1 + \ldots + \mu_n v_n = 0$. ■

7 Решения однородной системы образуют подпространство. Линейная оболочка: опредление и доказательство того, что это подпространство

Теорема 7.1. Совокупность всех решений системы однородных линейных уравнений с n неизвестными является подпространством пространства $\mathcal{K}^n$.

Доказательство. Рассмотрим произвольную систему однородных линейных уравнений:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0, \end{cases}$$

Очевидно, что нулевая строка является её решением и что произведение любого решения на число также является решением. Докажем, что сумма решений $(u_1, \ldots, u_n)$ и $(v_1, \ldots, v_n)$ является решением. Подставляя её компоненты в i-ое уравнение системы, получаем:

$$a_{i1}(u_1 + v_1) + \ldots + a_{in}(u_n + v_n) = \underbrace{a_{i1}u_1 + \ldots + a_{in}u_n}_{=0} + \underbrace{a_{i1}v_1 + \ldots + a_{in}v_n}_{=0} = 0.$$

■

Теорема 7.2. Совокупность всех решений произвольной совместной системы линейных уравнений есть сумма какого-либо одного её решения и подпространства решений системы однородных линейных уравнений с той же матрицей коэффициентов.

Доказательство. Пусть $u \in \mathcal{K}^n$ — какое-либо фиксированное решение системы

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

Аналогично предыдущему доказывается, что сумма решения u и произвольного решения однородной системы является решением системы $(*)$. Обратно, если u' — любое решение системы $(*)$, то $v = u' - u$ — решение однородной системы. ■

Примечание (из Винберга). Неопределённые системы линейных уравнений могут иметь различную «степень неопределённости», каковой естественно считать число свободных неизвестных в общем решении системы. Однако одна и та же система линейных уравнений может допускать различные общие решения, в которых разные неизвестные играют роль свободных, и закономерен вопрос, будет ли число свободных неизвестных всегда одним и тем же. Положительный ответ на этот вопрос даётся с помощью понятия размерности векторного пространства.

Определение 7.1. Пусть $S \subset V$ (V — векторное пространство). Совокупность всевозможных линейных комбинаций векторов из S называется **линейной оболочкой** множества S и обозначается через $\langle S \rangle$. Говорят, что пространство V **порождается** множеством S, если $\langle S \rangle = V$.

Определение 7.2. Пространство называется **конечномерным**, если оно порождается конечным количеством векторов и **бесконечномерным**, если оно порождается бесконечным количеством векторов.

Примечание. Я не заметил важной ремаки в конспектах Сергей Александровича — в курсе мы считаем вообще все пространства конечномерными. С бесконечномерными возникают большие проблемы — у нас даже нет определения линейной зависимости для бесконечного количества векторов. В дальнейшем мы будем доказывать утверждения для конечных систем, подразумевая, что для бесконечных они тоже верны (чаще всего достаточно убрать верхний предел суммирования). Цитата из Винберга по этому поводу: «Понятия базиса и размерности могут быть перенесены на бесконечномерные векторные пространства. Чтобы это сделать, надо определить, что такое линейная комбинация бесконечной системы векторов. В чисто алгебраической ситуации нет иного выхода, кроме как ограничиться рассмотрением линейных комбинаций, в которых лишь конечное число коэффициентов отлично от нуля». И далее следует определение линейной комбинации для бесконечной системы векторов как выражение вида $\sum_{i \in I} \lambda_i a_i$, в котором лишь конечное число коэффициентов λ_i отлично от нуля, так что сумма фактически является конечной и, таким образом, имеет смысл. Потом можно определить и линейную зависимости, и базис, и т. д. Но понятно, что так решается проблема только для счётных множеств. Что делать с несчётными, неясно совсем.

Теорема 7.3. Пусть $S \subset V$ (V — векторное пространство). Тогда линейная оболочка $\langle S \rangle$ является подпространством в V.

Доказательство. По теореме 4.2 достаточно проверить, что линейная оболочка замкнута относительно операций сложения векторов и умножения вектора на число. Это сразу следует из определения. ■

Утверждение (упражнение из Винберга). $\langle S \rangle$ — это наименьшее подпространство в V, содержащее S.

Доказательство. Предположим, что из $\langle S \rangle$ можно убрать элемент $v' = \lambda_1 v_1 + \ldots + \lambda_k v_k$. Но мы можем сконструировать v' из $v_1, \ldots, v_k$ с помощью операций умножения на число и сложения векторов, поэтому (из теоремы 4.2) он должен лежать в подпространстве. Противоречие. Значит, никакой элемент $\langle S \rangle$ нельзя убрать, чтобы оно осталось подпространством. ■

8 Базис системы векторов: эквивалентность 4 определений. Стандартный базис в $\mathbb{R}^n$. Дополнение линейно независимой системы до базиса. Выбор базиса из полной системы

Определение 8.1. Подсистема $L \subseteq S$ называется **полной**, если любой вектор из S является линейной комбинацией векторов из L.

Утверждение. $L \subseteq S$ полна тогда и только тогда, когда $\langle L \rangle = \langle S \rangle$.

Доказательство. $\Rightarrow$. Заметим, что $\langle L \rangle \subseteq \langle S \rangle$ (просто потому что $L \subseteq S$). Любой вектор из S является линейной комбинацией векторов из L, а значит, линейные комбинации векторов из S также являются линейными комбинациями векторов из L. А это значит, что $\langle S \rangle \subseteq \langle L \rangle$. Отсюда получаем $\langle L \rangle = \langle S \rangle$.

$\Leftarrow$. Каждый вектор из S принадлежит линейной оболочке S (просто потому что $S \subseteq \langle S \rangle$). А т. к. $\langle S \rangle = \langle L \rangle$, то этот вектор принадлежит линейной оболочке L, т. е. линейно выражается через векторы из L. ■

Теорема 8.1. Пусть B — подсистема векторов в системе S. Тогда следующие условия эквивалентны:

1. B — максимальная по включению линейно независимая подсистема векторов из S;
2. B — полная линейно независимая подсистема векторов из S;
3. B — минимальная по включению полная подсистема векторов из S
4. Каждый вектор из S линейно выражается через векторы из B, причём единственным образом.

Определение 8.2. Подсистема подмножества векторного пространства, удовлетворяющая условиям предыдущей теоремы, называется **базисом** этого подмножества.

Доказательство. Докажем, что из каждого следующего пункта следует предыдущий:

$1 \Rightarrow 2$. Пусть B — базис. Тогда B линейно независима, нужно лишь доказать, что B — полная подсистема. Возьмём $v \in S \setminus B$. Тогда система $B \cup \{v\}$ линейно зависима по определению базиса. По свойству линейной зависимости, вектор v является линейной комбинацией векторов из v. Значит, система B полна по определению.

$2 \Rightarrow 3$. По условию система B полная. Допустим, что для некоторого $v \in B$ система $B \setminus \{v\}$ также является полной. То есть, найдутся такие $e_1, \ldots, e_k \in B$ и $\lambda_1, \ldots, \lambda_k \in \mathcal{K}$, что $v = \sum_i^k \lambda_i e_i$. Но тогда B линейно зависима (по критерию линейной зависимости).

$3 \Rightarrow 4$. Пусть B — минимальная по включению полная подсистема. Тогда по определению полной подсистемы каждый вектор из S линейно выражается через B. Допустим, что есть вектор $v \in S$, такой что он выражается двумя различными способами

$$v = \sum_{i=1}^{n} \lambda_i e_i = \sum_{i=1}^{n} \mu_i e_i, \quad e_i \in B,\ \lambda_i, \mu_i \in \mathcal{K}.$$

Получаем $0 = \sum_{i=1}^{n} (\lambda_i - \mu_i) e_i$ — нетривиальная линейная комбинация. По критерию линейно зависимости, найдётся e_j, который выражается через остальные. Тогда система $B \setminus \{e_j\}$ полная. Получили противоречие с её минимальностью.

$4 \Rightarrow 1$. Допустим, что B линейно зависима. Тогда по критерию линейной зависимости существует $v \in B$, такой что

$$v = \sum_{i=1}^{n} \lambda_i e_i, \quad e_i \in B \setminus \{v\}, \lambda_i \in \mathcal{K}.$$

Получаем два линейных выражения v через B (второе имеет вид $v = 1 \cdot v$). Противоречие, значит, B линейно независима. Докажем теперь, что она максимальна по включению. Рассмотрим $u \in S \setminus B$. По условию, вектор u линейно выражается через B. Тогда по критерию линейной зависимости $B \cup \{u\}$ линейно зависима. ■

Определение 8.3. Пусть $\{e_1, \ldots, e_k\}$ — базис векторного пространства V и $v = v_1 e_1 + v_2 e_2 + \ldots + v_k e_k$. Тогда

столбец $\begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix}$ называется **координатами** вектора v в базисе $\{e_1, \ldots, e_k\}$. Так как $V \sim K^k$, то можно писать

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix}.$$

Определение 8.4. Размерностью конечномерного векторного пространства V называется число векторов $\dim V$ в его базисе.

Теорема 8.2. Определение выше корректно. Иными словами, все базисы конечномерного векторного пространства V содержат одно и то же число векторов.

Доказательство. Если бы в пространстве V существовали два базиса из разного числа векторов, то тот из них, в котором больше векторов, был бы линейно зависим по основной лемме. ■

Утверждение (задача из Винберга). Число базисов n-мерного векторного пространства над полем из q элементов равно

$$\prod_{k=0}^{n-1}(q^n - q^k).$$

Доказательство. Всего у нас в таком пространстве q^n векторов. Нам нужно выбрать среди них q линейно независимых, они и будут базисом. Для выбора первого вектора есть $q^n - 1$ возможностей, нам подходят все, кроме нулевого. Для выбора второго — $q^n - q$, нам подходят все, кроме коллинеарных первому, а таких q. И так далее посчитаем искомую велечину для $k - 1$ векторов. Теперь, чтобы посчитать количество способов выбрать k-ый линейно независимых вектор, посчитаем количество способов выбрать k-ый линейно зависимый вектор и вычтем из общего количества векторов. Итак, мы знаем, что система из $k - 1$ выбранных векторов линейно независима, а из k — линейно зависима. Отсюда (по третьему свойству линейной зависимости) k-ый вектор выражается через остальные единственным образом. А коэффициентов в этом выражении можно выбрать q^{k-1} штук (по q для каждого из $k - 1$ векторов). Итак, количество возможностей выбрать k-ый вектор линейно независимым равняется $q^n - q^{k-1}$. Теперь считаем ответ по правилу произведения (эти выборы не зависят друг от друга) и получаем требуемое. ■

Определение 8.5. Стандартным базисом в $\mathcal{K}^n$ называется базис, состоящий из строк единичной матрицы размера $n \times n$.

Примечание. Можете посмотреть интересные задачи 5-9 в Винберге, мне лень оформлять их решения.

Теорема 8.3. В конечномерном пространстве любую линейно независимую систему можно дополнить до базиса.

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

Теорема 8.4. Из любой конечной полной подсистемы S (в которой есть ненулевой вектор) можно выбрать базис S.

Доказательство. Воспользуемся тем, что базис — минимальная по включению полная подсистема. Пусть нам дана полная конечная подсистема $\{v_1, \ldots, v_n\}$. Если её можно уменьшить (выкинуть один из векторов) так, чтобы система осталась полной, то сделаем это. Иначе она уже базис. А т. к. система конечна, мы не можем бесконечно её уменьшать. Поэтому через конечное число шагов дойдём до непустой (т. к. в S был ненулевой вектор) минимальной по включению полной системы, т. е. до базиса. ■

Теорема 8.5 (Свойство монотонности размерности). Всякое подпространство U конечномерного векторного пространства V также конечномерно, причём $\dim U \leqslant \dim V$. Более того, если $U \neq V$, то $\dim U < \dim V$.

Доказательство. Пусть $\{e_1, \ldots, e_k\}$ — базис подпространства U, т. е. максимальная по включению линейно независимая система векторов из этого подпространства. Тогда $\dim U = k$. Линейно независимую систему $\{e_1, \ldots, e_k\}$ можно дополнить до базиса всего пространства V. Следовательно, если $U \neq V$, то $\dim V > k$. ■

Лемма 8.1. Всякое векторное пространство V над полем $\mathcal{K}$ изоморфно пространству $\mathcal{K}^{\dim V}$.

Доказательство. Пусть $\{e_1, \ldots, e_n\}$ — базис пространства V ($n := \dim V$). Рассмотрим отображение

$$\varphi : V \to K^n,$$

ставящее в соответствие каждому вектору строку из его коэффициентов в линейном разложении в базисе $\{e_1, \ldots, e_n\}$. Иными словами,

$$\sum_{i=1}^{n} x_i e_i \overset{\varphi}{\mapsto} (x_1, \ldots, x_n).$$

Очевидно, что оно биективно (т. к. каждый вектор линейно выражается через базисные единственным образом). Легко убедиться и в том, что такое отображение линейно. А значит, это изоморфизм. ■

Теорема 8.6 (из Винберга). Конечномерные векторные пространства над одним и тем же полем изоморфны тогда и только тогда, когда они имеют одинаковую размерность.

Доказательство. Если $f : V \to U$ — изоморфизм векторных пространств и $\{e_1, \ldots, e_k\}$ — базис пространства V, то $\{f(e_1), \ldots, f(e_k)\}$ — базис пространства U (*образ базиса является базисом образа*) в силу линейности функции f:

$$f\left(\sum_i \lambda_i e_i\right) = \sum_i \lambda_i f(e_i).$$

Так что, $\dim U = \dim V$. Обратно, согласно предыдущей лемме, всякое n-мерное пространство изоморфно $\mathcal{K}^n$, а значит, они изоморфны между собой. ■

9 Ранг системы векторов и размерность векторного пространства. Связь ранга системы и размерности линейной оболочки. Строчный, столбцовый и ступенчатый ранги матрицы, их совпадение

Определение размерности векторного пространства было дано ранее (определение 8.3).

Определение 9.1. Рангом системы векторов называется размерность её линейной оболочки.

Ранее доказывалось (теорема 8.2), что это определение корректно.

Определение 9.2. Системы векторов $\{a_1, \ldots, a_n\}$ и $\{b_1, \ldots, b_m\}$ называются **эквивалентными**, если каждый из векторов b_j линейно выражается через $a_1, \ldots, a_n$ и, наоборот, каждый из векторов a_i линейно выражается через $b_1, \ldots, b_m$.

Лемма 9.1. Ранги эквивалентных систем векторов равны.

Доказательство. Каждый вектор из $\{a_1, \ldots, a_n\}$ линейно выражается через векторы из $\{b_1, \ldots, b_m\}$, значит, линейные комбинации вида $\sum_i \lambda_i a_i$ тоже линейно выражаются через векторы из $\{b_1, \ldots, b_m\}$. Отсюда $\langle a_1, \ldots, a_n\rangle \subseteq \langle b_1, \ldots, b_m\rangle$. Аналогично доказывается обратное, а отсюда

$$\langle a_1, \ldots, a_n\rangle = \langle b_1, \ldots, b_m\rangle$$

А из равенства линейных оболочек очевидно вытекает равенство их размерностей. ■

Определение 9.3. Строчным рангом матрицы A называется ранг системы её строк, **столбцовым рангом** — ранг системы её столбцов.

Теорема 9.1. Строчный, столбцовый и ступенчатый ранги совпадают.

Сначала докажем две вспомогательные леммы.

Лемма 9.2. Каждый вид ранга не меняется при элементарных преобразованиях строк.

Доказательство. Докажем утверждение отдельно для каждого вида ранга:

- **Строчный ранг**. Пусть мы пришли элементарным преобразованием от матрицы A к матрице A'. Очевидно, что системы строк этих матриц эквивалентны. А потому их ранги равны.
- **Столбцовый ранг**. Докажем, что линейные зависимости между столбцами матрицы не меняются при элементарных преобразованиях строк. Линейная зависимость между какими-то столбцами матрицы может пониматься как линейная зависимость между всеми её столбцами, в которую остальные столбца входят с нулевыми коэффициентами. Следовательно, если какие-то столбцы матрицы линейно зависимы, то они останутся линейно зависимыми после любых элементарных преобразований строк. Так как элементарные преобразований обратимы, то и наоборот: если какие-то столбцы линейно независимы, то они и останутся линейно независимыми. Значит, если какие-то столбцы матрицы составляют максимальную линейно независимую систему её столбцов, то после любых элементарных преобразований строк столбцы с теми же номерами буду составлять максимальную линейно независимую систему столбцов полученной матрицы, и поэтому ранг матрицы не изменится.
- **Ступенчатый ранг**. Утверждение для него сразу следует из единственности улучшенного ступенчатого вида.

■

Лемма 9.3. В ступенчатом виде все виды ранга совпадают.

Доказательство. Докажем, что ненулевые строки образуют базис системы строк (т. е. просто, что они линейной независимы). Пусть лидеры имеют номера $j_1, \ldots, j_r$. Уберём из системы ненулевых строк строку с номером i. Тогда из улучшенного ступенчатого вида матрицы, у оставшихся векторов j_i-ая координата равна 0, а значит, никакой их линейной комбинацией нельзя получить i-ую строку. Таким образом, система ненулевых строк — линейно независима и максимальна по включению (остальные строки нулевые), а значит, является базисом.

Теперь докажем, что столбцы с номерами $j_1, j_2, \ldots, j_r$ образуют базис системы столбцов. Действительно, эта система полна и минимальна по включению, то есть, базис.

Ступенчатый ранг матрицы равен количеству её ненулевых строк, что, в свою очередь, равно строчному рангу. Также, ступенчатый ранг равен количеству главных переменных, а это, как мы показали, и есть столбцовый ранг. ■

Теперь совсем легко доказать теорему 9.1:

Доказательство. Приведём матрицу к улучшенному ступенчатому виду, каждый вид ранга при этом не изменится. А в улучшенному ступенчатому виде все ранги совпадают. ■

10 Свойства ранга матрицы. Теорема Кронекера — Капелли. Критерий определённости системы

Определение 10.1. Пусть A — матрица $m \times n$. Матрица $B = A^T$ размера $n \times m$ называется **транспонированной** к матрице A, если

$$b_{ij} = a_{ij}.$$

Примечание. Ясно, что $(A^T)^T = A$.

Теорема 10.1 (Свойства ранга матрицы). Пусть A — матрица $m \times n$. Тогда

1. $\operatorname{rk} A \leqslant \min\{m, n\}$
2. $\operatorname{rk} A = \operatorname{rk} A^T$
3. Если к матрице добавить k столбцов (строк), то ранг не уменьшится и увеличится не более, чем на k

Доказательство.

1. $\operatorname{rk} A$ равен рангу системы S строк матрицы A, т. е. $\dim\langle S\rangle$. Система S, очевидно, полна в $\langle S\rangle$, а значит, из неё можно выделить базис, в нём будет не больше векторов, чем в исходной системе, т. е. не больше m. Аналогично доказывается $\operatorname{rk} A \leqslant n$.
2. Ранг системы строк A равен рангу системы столбцов A^T (это одна и та же система).
3. Базис системы строк (столбцов) матрицы A — это линейно независимая система. Значит, в новой матрице её можно дополнить до базиса. Другие строки матрицы A в него уже не попадут (они образуют линейно зависимую систему со старым базисом), поэтому в него могут добавиться только новые строки.

■

Теорема 10.2 (Кронекер — Капелли). СЛУ совместна тогда и только тогда, когда ранг матрицы её коэффициентов равен рангу расширенной матрицы её коэффициентов.

Доказательство. Приведём расширенную матрицу коэффициентов к улучшенному ступенчатому виду (ранг и совместность СЛУ при этом не меняются).

$\Rightarrow$. Если СЛУ совместна, то в ней нет экзотических уравнений, т. е. нет лидеров в последнем столбце матрицы. А, как мы доказывали ранее, базис системы столбцов состоит из тех столбцов, в которых есть лидеры. Поэтому последний столбец расширенной матрицы выражается линейно через базис системы столбцов матрицы коэффициентов. Значит, их ранги равны.

$\Leftarrow$. Если ранг матрицы коэффициентов равен рангу расширенной матрицы, то последний столбец расширенной матрицы линейно выражается через базис системы столбцов матрицы коэффициентов (иначе нужно было бы добавить его в базис и ранг расширенной матрицы был бы больше на 1). А значит, экзотических уравнений нет, и система совместна. ■

Теорема 10.3 (Критерий определённости системы). Совместная система линейных уравнений является определённой тогда и только тогда, когда ранг матрицы её коэффициентов равен числу неизвестных.

Доказательство. Ранг матрицы коэффициентов равен количеству ненулевых строк в ступенчатом виде, т. е. количеству главных переменных. Условие теоремы равносильно тому , что главных переменных столько же, сколько всех переменных, т. е. тому, что система определена. ■

11 Фундаментальная система решений и алгоритм её поиска. Размерность пространства решений однородной СЛУ. Структура решений неоднородной СЛУ

Определение 11.1. Фундаментальная система решений — это базиса подпространства решений однородной СЛУ.

Примечание. Доказательство следующей теоремы даёт практический способ построения ФСР.

Теорема 11.1. Размерность пространства решений системы однородных линейных уравнений с n неизвестными и матрицей коэффициентов A равна $n - \operatorname{rk} A$.

Доказательство. Рассмотрим систему уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0, \end{cases}$$

С помощью элементарных преобразований приведём её к ступечантому виду. Число ненулевых уравнений в этом ступенчатом виде равно $r = \operatorname{rk} A$. Поэтому общее решение будет содержать r главных неизвестных и

с точностью до перенумерации неизвестных будет иметь вид

$$\begin{cases} x_1 = c_{11}x_{r+1} + c_{12}x_{r+2} + \ldots + c_{1,n-r}x_n, \\ x_2 = c_{21}x_{r+1} + c_{22}x_{r+2} + \ldots + c_{2,n-r}x_n, \\ \ldots \\ x_r = c_{r1}x_{r+1} + c_{r2}x_{r+2} + \ldots + c_{r,n-r}x_n. \end{cases}$$

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

$$\begin{aligned} &u_1 = (c_{11}, c_{21}, \ldots, c_{r1}, 1, 0, \ldots, 0), \\ &u_2 = (c_{12}, c_{22}, \ldots, c_{r2}, 0, 1, \ldots, 0), \\ &\ldots \\ &u_{n-r} = (c_{1,n-r}, c_{2,n-r}, \ldots, c_{r,n-r}, 0, 0, \ldots, 0, 1). \end{aligned}$$

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

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

Элементарные преобразования столбцов (как и элементарные преобразования строк) сохраняют ранг матрицы. Заметим, что поменяв местами «блоки», отделённые друг от друга чертой, мы приведём матрицу к улучшенному ступенчатому виду. Как можно видеть, все её строки в улучшенном ступенчатом виде ненулевые, а потому, во-первых, ранг этой матриц равен количеству строк, т. е. $n - r$, а во-вторых, система $\{u_1, \ldots, u_{n-r}\}$ линейно независима. Также, эта система полна, т. к. любая линейная комбинация вида

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

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

Структура решений неоднородной СЛУ — это теорема 7.2.

12 Линейное отображение. Его матрица в фиксированных базисах. Образ заданного вектора. Изоморфизм. Любое конечномерное пространство изоморфно пространству строк

Определение 12.1. Отображение $\varphi : U \to V$ (U и V — векторные пространства) называется **линейным**, если $\varphi(u_1 + u_2) = \varphi(u_1) + \varphi(u_2)$ и $\varphi(\lambda u) = \lambda\varphi(u)$.

Определение 12.2. Фиксируем базис $e = \{e_1, \ldots, e_n\}$ в U и базис $f = \{f_1, \ldots, f_m\}$. Матрица $A(\varphi, e, f) = (a_{ij})$, по столбцам которой стоят координаты образов базисных векторов из e в базисе f, называется **матрицей линейного отображения** φ.

Примечание. Размеры матрицы линейного отображения — $m \times n$.

Утверждение. $\varphi(v) = A \cdot v$, где A — матрица линейного отображения φ.

Доказательство. С одной стороны,

$$\varphi(v) = \varphi\left(\sum_i^k v_i e_i\right) = \sum_i^k v_i \varphi(e_i).$$

А с другой стороны,

$$A \cdot v = \begin{pmatrix} f(e_1) & f(e_2) & \ldots & f(e_k) \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{pmatrix} = \sum_i^k v_i \varphi(e_i).$$

Отсюда сразу следует $\varphi(v) = A \cdot v$. ■

Определение 12.3. Векторные пространства U и V над полем $\mathcal{K}$ называются **изоморфными**, если существует линейное биективное отображение $\varphi : U \to V$. Само отображение φ называется при этом **изоморфизмом** пространств V и U.

См. также лемму 8.1.

13 Операции над линейными отображениями и над матрицами, связь между ними. Матричная запись СЛУ

Примечание. Здесь векторные пространства U, V, W взяты над одним полем $\mathcal{K}$ и, говоря «число», имеем в виду произвольный элемент $\mathcal{K}$.

Определение 13.1 (Операции над линейными отображениями). Пусть $\varphi : U \to V, \psi : U \to V, \xi : V \to W$. Определим следующие операции над ними:

1. **Сумма**. $(\varphi + \psi)(u) := \varphi(u) + \psi(u)$. Проверим, что $\varphi + \psi$ линейно:

$$(\varphi + \psi)(u_1 + u_2) = \varphi(u_1 + u_2) + \psi(u_1 + u_2) = \varphi(u_1) + \psi(u_1) + \varphi(u_2) + \psi(u_2) = (\varphi + \psi)(u_1) + (\varphi + \psi)(u_2),$$
$$(\varphi + \psi)(\lambda u) = \varphi(\lambda u) + \psi(\lambda u) = \lambda(\varphi(u) + \psi(u)) = \lambda \cdot (\varphi + \psi)(u).$$

2. **Умножение на число**. $(c\varphi)(u) = c \cdot \varphi(u)$. Нетрудно проверить, что $c\varphi$ линейно.
3. **Композиция**. $(\xi \circ \varphi)(u) = \xi(\varphi(u))$. Проверим линейность:

$$(\xi \circ \varphi)(u_1 + u_2) = \xi(\varphi(u_1 + u_2)) = \xi(\varphi(u_1) + \varphi(u_2)) = \xi(\varphi(u_1)) + \xi(\varphi(u_2)) = (\xi \circ \varphi)(u_1) + (\xi \circ \varphi)(u_2),$$
$$(\xi \circ \varphi)(\lambda u) = \xi(\varphi(\lambda u)) = \xi(\lambda\varphi(u)) = \lambda\xi(\varphi(u)) = \lambda(\xi \circ \varphi)(u).$$

Определение 13.2 (Операции над матрицами). Пусть A, B и C — матрицы линейных преобразований φ, ψ и ξ соответственно в фиксированных базиса $e = \{e_1, \ldots, e_n\}$ в U, $f = \{f_1, \ldots, f_m\}$ в V, $s = \{s_1, \ldots, s_k\}$ в W. Определим следующие операции над ними:

1. **Суммой матриц** A и B называется матрица преобразования $\varphi + \psi$. Обозначается $A + B$. Заметим, что суммировать можно лишь матрицы одинакового размера.
2. **Произведением числа λ и матрицы** A называется матрица преобразования $\lambda\varphi$. Обозначается λA.
3. **Произведением матриц** C и A называется матрица преобразования $\xi\circ\varphi$. Обозначается CA. Заметим, что умножать можно лишь матрицы размеров $k \times m$ и $m \times n$.

Утверждение. Формулы для операций над матрицами $A(\varphi, e, f) = (a_{ij})$ и $B(\psi, e, f) = (b_{ij})$ и $C(\xi, f, s) = (c_{ij})$ (обозначения — из предыдущих определений) пишутся так:

1. $(A + B)_{ij} = (a_{ij} + b_{ij})$ (при этом $\underset{m\times n}{A}$ и $\underset{m\times n}{B}$);
2. $(\lambda A)_{ij} = (\lambda a_{ij})$;
3. $(CA)_{ij} = \sum\limits_{t=1}^{m} c_{it} \cdot a_{tj}$ (при этом $\underset{m\times n}{A}$ и $\underset{k\times m}{C}$);

Доказательство. Столбца матрицы линейного преобразования — это образы базисных векторов, поэтому чтобы понять, как выглядит матрица преобразования, достаточно посмотреть на то, куда переходит базис. Разберём отдельно все операции:

1. Сложение:

$$(\varphi + \psi)(e_j) = \varphi(e_j) + \psi(e_j) = \sum_{i=1}^{m} a_{ij} f_i + \sum_{i=1}^{m} a_{ij} f_i = \sum_{i=1}^{m} (a_{ij} + b_{ij} f_i).$$

Коэффициенты при f_i и стоит в новой матрице на месте (i, j). Таким образом, $(A + B)_{ij} = (a_{ij} + b_{ij})$.

2. Умножение на число:
$$(\lambda\varphi)(e_j) = \lambda\left(\sum_{i=1}^{m} a_{ij} f_i\right) = \sum_{j=1}^{m} (\lambda a_{ij}) f_i,$$
отсюда $(\lambda A)_{ij} = (\lambda a_{ij})$.

3. Произведение:
$$(\xi \circ \varphi)(e_j) = \xi\left(\sum_{t=1}^{m} a_{tj} f_t\right) = \sum_{t=1}^{m} a_{tj}\xi(f_t) = \sum_{t=1}^{m} a_{tj}\left(\sum_{i=1}^{k} c_{it} s_i\right) = \sum_{t=1}^{m}\sum_{i=1}^{k} a_{tj} c_{it} s_i = \sum_{i=1}^{k}\left(\sum_{t=1}^{m} c_{it} a_{tj}\right) s_i,$$
отсюда $(CA)_{ij} = \sum\limits_{t=1}^{m} c_{it} \cdot a_{tj}$.

■

Теперь СЛУ с матрицей коэффициентов A, неизвестными $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ и свободными членами $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ можно записать так:
$$AX = B.$$

Теорема 13.1. Соответствие $\varphi \to A(\varphi, e, f)$ задаёт биекцию между линейными отображениями $U \to V$ и матрицами $m \times n$.

Доказательство. В курсе аналитической геометрии мы доказывали (см. билет про скалярное произведение в моём файле) теорему о том, что $f(*) = (u, *)$ — общий вид линейной функции. Причём, в доказательстве мы предъявляли вектор $u = (\varphi(e_1), \ldots, \varphi(e_n))$. Отсюда сразу следует утверждение теоремы, ведь есть расписать координаты векторов $\varphi(e_i)$ в базисе f по столбцам, u станет в точности матрицей линейного преобразования φ в фиксированных базисах e и f. ■

14 Свойства операций над матрицами. Связь с транспонированием

Легко видеть, что биекция, существование которой доказывается в теореме 13.1, является линейным отображением (из определения операций над матрицами). Также легко видеть, что множество линейных отображений — это векторное пространство. А значит, и множество матриц фиксированного размера является векторным простраством. При этом, $\dim \underset{m\times n}{\mathrm{Mat}} = mn$.

Теорема 14.1 (Свойства умножения матриц).

1. $(AB)C = A(BC)$
2. $A(B + C) = AB + AC$
3. $(A + B)C = AC + BC$
4. $\lambda(AB) = (\lambda A)B = A(\lambda B)$

Доказательство. В силу теоремы 13.1 нам достаточно показать эти свойства для соответствующих операций над линейными отображениями.

1. Пусть $\varphi : W \to L$, $\psi : V \to W$, $\xi : U \to V$. Тогда
$$((\varphi \circ \psi) \circ \xi)(u) = (\varphi \circ \psi)(\xi(u)) = \varphi(\psi(\xi(u))) = \varphi((\psi \circ \xi)(u)) = (\varphi \circ (\psi \circ \xi))(u),$$
отсюда $(\varphi \circ \psi) \circ \xi = \varphi \circ (\psi \circ \xi)$.

2. Пусть $\varphi : V \to W$, $\psi : V \to W$ и $\xi : U \to V$. Тогда
$$((\varphi + \psi) \circ \xi)(u) = (\varphi + \psi)(\xi(u)) = \varphi(\xi(u)) + \psi(\xi(u)) = (\varphi \circ \xi)(u) + (\psi \circ \xi)(u).$$

3. Аналогично предыдущему пункту.

4. Пусть $\varphi : V \to W$, $\psi : U \to V$. Тогда

$$(\lambda(\varphi \circ \psi))(u) = \lambda\varphi(\psi(u)) = (\lambda\varphi \circ \psi)(u) = (\varphi \circ \lambda\psi)(u).$$

■

Теорема 14.2 (Связь с транспонированием).

1. $(A + B)^T = A^T + B^T$
2. $(\lambda A)^T = \lambda A^T$
3. $(AB)^T = B^T A^T$

Доказательство. Эти свойства докажем, расписав их по формулам, которые мы вывели ранее:

1. $(A + B)^T_{ij} = (A + B)_{ji} = a_{ji} + b_{ji} = (A^T + B^T)_{ij}$.
2. $(\lambda A)^T = (\lambda A)_{ji} = (\lambda a_{ij})^T = \lambda a_{ji} = (\lambda A)^T$.
3. $(AB)^T_{ij} = (AB)_{ji} = \sum\limits_{t=1}^{m} b_{jt}a_{ti} = (B^T A^T)_{ij}$.

■

15 Вывод обобщённой ассоциативности для ассоциативной операции

Примечание. Здесь для удобства часто будем называть операцию умножением.

Определение 15.1. Ассоциативная операция называется **обобщённо ассоциативной**, если произведение любого количества сомножителей не зависит от расстановки скобок в этом произведении.

Определение 15.2. Назовём **стандартной** такую расстановку скобок, в которой операции выполняются слева направо.

Теорема 15.1. Ассоциативная операция всегда обобщённо ассоциативна.

Доказательство. Доказательство проведём индукцией по количеству множителей k:

База индукции ($k = 3$). Верно в силу ассоциативности.

Шаг индукции. Пусть для $k < n$ утверждение верно. Докажем его для $k = n$. Докажем, что для произвольной расстановки скобок результат будет таким же, как и для стандартной. Рассмотрим некоторую расстановку скобок в произведении $x_1 * \ldots * x_n$. Пусть последнее умножение перемножает скобки A и B. Рассмотрим два случая:

1. $B = x_n$. Тогда для количество множителей в скобке A равно $n - 1 < n$, а потому по предположению индукции можем стандартно расставить скобки в ней. Но тогда и во всё выражении они будут расставлены стандартно и теорема доказана.
2. $B \neq x_n$. Внутри B по предположению индукции расстановку скобок можно принять стандартной, а потому $B = C * x_n$. А отсюда
$$A * B = A * (C * x_n) = (A * C) * x_n$$
из ассоциативности, и мы попали в предыдущий случай.

■

16 Верхние оценки на ранг суммы и произведения матриц

Антон Александрович задавал нам на дом вывести и нижние оценки, поэтому они здесь будут.

Утверждение. $|\operatorname{rk} A - \operatorname{rk} B| \leqslant \operatorname{rk}(A+B) \leqslant \operatorname{rk} A + \operatorname{rk} B$.

Доказательство. Сначала докажем оценку сверху. Строки матрицы $A+B$ — это линейные комбинации строк матриц A и B. А значит, базис строк $A+B$ — линейно независимая система, линейно выражающаяся через базисы строк A и B. А значит, по основной лемме о линейной зависимости $\operatorname{rk}(A+B) \leqslant \operatorname{rk} A + \operatorname{rk} B$.

Оценка снизу равносильна системе

$$\begin{cases} \operatorname{rk} A - \operatorname{rk} B \leqslant \operatorname{rk}(A+B), \\ \operatorname{rk} B - \operatorname{rk} A \leqslant \operatorname{rk}(A+B) \end{cases}$$

Докажем первое из неравенств системы, второе доказывается аналогично. Итак, из оценки сверху:

$$\operatorname{rk}(A+B) + \operatorname{rk} B = \operatorname{rk}(A+B) + \operatorname{rk}(-B) \leqslant \operatorname{rk} A.$$

Перенеся слагаемые в нужные стороны, получим то, что хотели. ■

Примечание. $\operatorname{rk} A = \operatorname{rk}(\lambda A)$ $(\lambda \neq 0)$, т. к. по сути умножение матрицы на ненулевое число — это умножение каждой из её строк на это число, а это преобразование является элементарным.

Утверждение. $\operatorname{rk} A + \operatorname{rk} B - n \leqslant \operatorname{rk} AB \leqslant \min\{\operatorname{rk} A, \operatorname{rk} B\}$, где $\underset{m\times n}{A}$ и $\underset{n\times k}{B}$.

Лемма 16.1. Столбцы AB — линейные комбинации столбцов A с коэффициентами из столбцов B. Строки BA — это линейные комбинации строк B с коэффициентами из строк A.

Примечание. Далее за $X^{(i)}$ будем обозначать i-ый столбец матрицы X, а за $X_{(i)}$ — её i-ую строку.

Доказательство. Докажем утверждение непосредственным умножением:

$$(AB)^{(j)} = \begin{pmatrix} \sum\limits_{t=1}^{n} a_{1t} b_{tj} \\ \sum\limits_{t=1}^{n} a_{2t} b_{tj} \\ \vdots \\ \sum\limits_{t=1}^{n} a_{mt} b_{tj} \end{pmatrix} = \sum_{t=1}^{n} b_{tj} A^{(t)}.$$

Второе утверждение доказывается аналогично. ■

Теперь докажем оценки на ранг произведения:

Доказательство. Сначала докажем верхнюю оценку. Система столбцов матрицы AB линейно выражается через столбцы A, поэтому (из основной леммы о линейной зависимости) $\operatorname{rk} AB \leqslant \operatorname{rk} A$. Аналогично, $\operatorname{rk} AB \leqslant \operatorname{rk} B$. Отсюда сразу следует требуемое. Теперь докажем и нижнюю оценку. Для этого рассмотрим матрицу

$$\begin{pmatrix} \underset{n\times n}{E} & \underset{n\times k}{0} \\ \underset{m\times n}{0} & \underset{m\times k}{AB} \end{pmatrix}$$

размера $(m+n) \times (n+k)$. Как нетрудно заметить, её ранг равен $\operatorname{rk} AB + \operatorname{rk} E$. Из леммы 16.1, можно производить элементарные преобразования не над отдельными её строками, а над блоками из матриц (Антон Александрович называл их «гиперстроками»), ведь такие преобразования можно разбить на цепочки преобразований обычных строк. Итак, элементарными преобразованиями (которые, как известно, не меняют ранг) переведём нашу матрицу в такую:

$$\begin{pmatrix} E & 0 \\ 0 & AB \end{pmatrix} \rightsquigarrow \begin{pmatrix} E & 0 \\ A & AB \end{pmatrix} \rightsquigarrow \begin{pmatrix} E & -B \\ A & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} E & B \\ A & 0 \end{pmatrix}$$

Теперь докажем, что $\operatorname{rk}\begin{pmatrix} E & B \\ A & 0 \end{pmatrix} \geqslant \operatorname{rk} A + \operatorname{rk} B$. Приведём нашу блочную матрицу к улучшенному ступенчатому виду. Для этого можно брать строки единичной матрицы и вычитать их из строк матрицы A, обнуляя их. Это возможно, т. к. строки матрицы E образуют стандартный базис в $\mathcal{K}^n$, а потому строки A точно выражаются как линейные комбинации строк E. Что же будет при этом происходить со вторым «гиперстолбцом»? Там мы просто каждый раз будем из нулевой матрицы вычитать B с каким-то коэффициеном. А потому результат будем λB. Итак, улучшенный ступенчатый вид нашей матрицы:

$$\begin{pmatrix} E & B \\ 0 & \lambda B \end{pmatrix}.$$

Её ранг равен количеству ненулевых строк, т. е. $n + \operatorname{rk} B$. А из неравенства $\operatorname{rk} A \geqslant n$ получаем то, что хотели. Итак,

$$\operatorname{rk} AB + n = \operatorname{rk} AB + \operatorname{rk} E = \operatorname{rk}\begin{pmatrix} E & 0 \\ 0 & AB \end{pmatrix} \geqslant \operatorname{rk} A + \operatorname{rk} B.$$

Перенеся слагаемые в нужные стороны, получим то, что хотели. ■

17 Правая и левая обратные матрицы. Критерий существования. Обратная матрица, её единственность и критерий существования

Определение 17.1. Пусть A — матрица $m \times n$. Матрица B размера $n \times m$ называется **левой обратной** к матрице A, если $BA = E$ и **правой обратной** к матрице A, если $AB = E$.

Определение 17.2. Пусть A — матрица $n \times n$. Матрица B размера $n \times n$ называется **обратной** к матрице A, если $AB = BA = E$. Обозначается A^{-1}.

Теорема 17.1 (Критерий существования левой/правой обратной). Левая обратная к матрице $\underset{n\times m}{A}$ существует тогда и только тогда, когда $\operatorname{rk} A = n$. Если $m = n$ и левая обратная к матрице A существует, то она единственна. Те же утверждения для правой обратной.

Доказательство. Рассмотрим матрицу BA. Если $\operatorname{rk} A < n$, то $\operatorname{rk} AB \leqslant \operatorname{rk} A < n$, но $\operatorname{rk} E = n$, а значит, $AB \neq E$. Противоречие.

Пусть теперь $\operatorname{rk} A = n$. Тогда строки A — полная система в $\mathcal{K}^n$, т. к. базис системы строк A является базисом пространства $\mathcal{K}^n$ (в этом базисе все векторы линейно независимы и их количество правильное). Строки матрицы BA — это линейные комбинации строк A с коэффициентами из строк B. Причём, т. к. система строк A полна, то существуют коэффициенты (строки матрицы B), линейная комбинация строк A с которыми даст строки единичной матрицы. Иными словами, существует такая матрица B, что $BA = E$.

А если $m = n$ и $\operatorname{rk} A = n$, то вся система строк является базисом $\mathcal{K}^n$, а разложение по базису каждого вектора единственно. Поэтому существует единственный набор коэффициентов, линейная комбинация строк A с которыми даст строки единичной матрицы. Иными словами, существует единственная матрица B, такая что $BA = E$. Для правой обратной аналогично. ■

Примечание. Отсюда следует, что к матрице A существует и левая, и правая обратные (причём, единственные) тогда и только тогда, когда $\operatorname{rk} A = m = n$, т. е. матрица A квадратная и её ранг равен размеру. Теперь легко видеть, что если B — левая обратная к A, а C — правая, то $B = C$:

$$B = B(AC) = (BA)C = C.$$

Отсюда получаем слеудющую теорему:

Теорема 17.2. К матрице A существует (и притом, только одна) обратная матрица тогда и только тогда, когда она квадратная и её ранг равен размеру.

Теорема 17.3. Одно из равенств $AB = E$ или $BA = E$ влечёт другое.

Доказательство. Пусть выполнено $AB = E$. Тогда B — правая обратная к A. Так как правая обратная существует, то $\operatorname{rk} A = n$. А отсюда следует, что и левая обратная существует. Выше обсуждалось, что эти матрицы обязаны быть равны. ■

Утверждение (Задача Антона Александровича). A^{-1} (если существует) является многочленом от A.

Доказательство. Рассмотрим матрицы $A^{n^2}, A^{n^2-1}, \ldots, A^2, A, E$. Их всего n^2+1 штук. А размерность пространства $\underset{n\times n}{\mathrm{Mat}} \ni A$ равна n^2. Значит, эти матрицы линейно зависимы. Иными словами, существует их нетривиальная нулевая линейная комбинация. Пусть её коэффициенты — $\lambda_1, \lambda_2, \ldots, \lambda_{n^2+1}$. Тогда рассмотрим многочлен

$$\lambda_1 X^{n^2} + \lambda_2 X^{n^2-1} + \ldots + \lambda_{n^2} X + \lambda_{n^2+1} E.$$

Из определения коэффициентов λ_i, A является корнем этого многочлена. Значит, множество многочленов, аннулирующих A непусто. Выберем из него многочлен наименьшей степени, пусть это

$$f(X) = \mu_1 X^k + \mu_2 X^{k+1} + \ldots + \mu_k X + \mu_{k+1} E.$$

Предположим, что $\mu_{k+1} = 0$. Тогда имеем

$$\mu_1 A^k + \mu_2 A^{k+1} + \ldots + \mu_k A = 0.$$

Домножим на A^{-1} справа и получим многочлен степени $k-1$, аннулирующий матрицу A. Противоречие с тем, что минимальная степень многочлена с таким свойством — k.

Итак, $\mu_{k+1} \neq 0$. Тогда домножим равенство

$$\mu_1 A^k + \mu_2 A^{k+1} + \ldots + \mu_k A + \mu_{k+1} E = 0$$

справа на A^{-1} и выразим её:

$$A^{-1} = \sum_{i=1}^{k} \left(-\frac{\mu_i}{\mu_{k+1}} \right) A^{k-i}.$$

■

Определение 17.3. Матрица называется **вырожденной** (**обратимой**), если у неё нет обратной.

Примечание. В решении теоретических задач часто полезно помнить, что улучшенный ступенчатый вид матрицы — это E тогда и только тогда, когда она невырождена.

18 Элементарные матрицы. Умножение на элементарные матрицы слева и справа. Матрица, обратная к произведению. Обратная к транспонированной матрице

Определение 18.1. Непосредственным вычислением проверяется, что элементарные преобразования строк какой-либо матрицы A равносильные её умножению слева на **элементарные матрицы** следующих трёх типов:

$$E + cE_{ij} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & c & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}, \quad P_{ij} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & \ddots & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}, \quad Q_i(c) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & c & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix},$$

где $i \neq j$ и $c \neq 0$, а все элементы этих матриц, не выписанные явно, такие же, как у единичной матрицы.

Умножение на элементарные матрицы справа дают нам элементарные преобразования столбцов. Так, умножение матрицы A слева на $E + cE_{ij}$ $(i \neq j)$ приводит к тому, что к i-ой строке прибавляется j-ая

строка, умноженная на c. А если умножить ту же матрицу на A справа, то к j-му столбцу прибавляется i-ый столбец, умноженный на c.

Примечание. Из Винберга. Метод Гаусса в матричной интерпретации состоит в последовательном умножении уравнения

$$AX = B$$

слева на элементарные матрицы, имеющем целью приведения матрицы A к улучшенному ступенчатому виду. Используя вместо элементарных матриц какие-либо другие матрицы, можно получить другие методы решения систем линейных уравнений, которые, быть может, не столь просты в теоретическом отношении, но, скажем, более надёжны при приближённых вычислениях (в случае $\mathcal{K} = \mathbb{R}$). Таков, например, метод вращений, при котором в качестве элементарных берутся матрицы вида

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & \cos\alpha & \cdots & -\sin\alpha & & \\ & & \vdots & \ddots & \vdots & & \\ & & \sin\alpha & \cdots & \cos\alpha & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}.$$

Утверждение.

1. $(AB)^{-1} = B^{-1}A^{-1}$;
2. $(A^T)^{-1} = (A^{-1})^T$.

Доказательство. Нужно просто перемножить и удостовериться, что получается E:

1. $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = E$;
2. $(A^{-1})^T A^T = (AA^{-1})^T = E^T = E$.

Как следствие, если A и B обратимы, то и AB тоже. ■

19 Алгоритм поиска обратной матрицы. Разложение невырожденной матрицы в произведение элементарных

Чтобы найти обратную матрицу, нам нужно решить матричное уравнение

$$AX = E.$$

Для этого нам нужно решить n систем уравнений с одной и той же матрицей коэффициентов A, столбцы свободных членов которых составляют матрицу E. Эти системы можно решать одновременно методом Гаусса. После приведения матрицы коэффициентов к единичной матрице (что возможно в силу её невырожденности) преобразованные столбцы свободных членов составят искомую матрицу A^{-1}.

$$(A \mid E) \rightsquigarrow \ldots \rightsquigarrow (E \mid A^{-1}).$$

А, как уже обсуждалось, метод Гаусса заключается в домножении матрицу на элементарные слева. А т. к. улучшенный ступенчатый вид невырожденной матрицы единичный, то получаем

$$A = U_1A_1 = U_2U_1A_2 = \ldots = U_N \ldots U_1E = U_N \ldots U_1,$$

где U_i — элементарные матрицы.

Александр Александрович Гайфуллин рассказал полезный трюк. Часто пригождается решать матричные уравнения типа

$$AX = B.$$

Решением является A^{-1}. Однако искать обратную к A, а затем умножать её на B может быть затруднительно. Можно поступить тем же способом, которым мы пользовались ранее, заменив единичную матрицу на B (ведь по сути ничего не меняется, нам всё ещё нужно решить n СЛУ с n неизвестными):

$$(A \mid B) \rightsquigarrow \ldots \rightsquigarrow (E \mid A^{-1}B).$$

Ещё это можно объяснить следующим образом (так это делал Сергей Александрович): элементарные преобразования строк — это умножения слева на элементарные матрицы. Мы делаем над обеими частями матрицы $(A \mid B)$ одни и те же элементарные преобразования, в результате слева получаем $U_N \ldots U_1 \cdot A = E$, значит, $U_N \ldots U_1 = A^{-1}$, а справа $U_N \ldots U_1 B = A^{-1}B$.

20 Подстановки и перестановки. Их количество. Произведение подстановок, его свойства. Разложение подстановки на произведение независимых циклов

Определение 20.1. Зафиксируем множество $\Omega_n = \{1, 2, \ldots, n\}$. **Подстановкой** будем называть биекцию $\sigma : \Omega_n \to \Omega_n$, а **перестановкой** — упорядоченный набор элементов Ω_n. Обозначать подстановки будем так:

$$\begin{pmatrix} 1 & \ldots & i & \ldots & n \\ \sigma(1) & \ldots & \sigma(i) & \ldots & \sigma(n) \end{pmatrix}.$$

Примечание. Между перестановками и подстановками можно установить биекцию, поэтому эти понятия взаимозаменяемы.

Определение 20.2. Произведением подстановок называется их композиция.

Теорема 20.1. Перестановки порядка n образуют группу. Она обозначается S_n.

Доказательство. Проверим выполнение всех аксиом:

1. Ассоциативность выполняется, т. к. она выполняется для всех отображений
2. Тождественная подстановка
$$\varepsilon = \begin{pmatrix} 1 & \ldots & i & \ldots & n \\ 1 & \ldots & i & \ldots & n \end{pmatrix}$$
является нейтральным элементом.
3. Если поменять местами две строки в записи перестановки σ, то мы получим такую, которая в композиции с σ даст тождественную. Это и есть обратная подстановка.

■

Утверждение. $|S_n| = n!$

Доказательство. Нам нужно выбрать n элементов из множества $\{1, 2, \ldots, n\}$. У нас есть n способов выбрать первый элемент, затем $n - 1$ способов выбрать второй элемент и т. д. Таким образом, количество способов выбрать уникальную подстановку равно

$$n \cdot (n-1) \cdot \ldots \cdot 1 = n!$$

■

Примечание. Из лекций Ю. Г. Прохорова. В общем случае перестановки не коммутируют. Однако есть условие, при котором это свойство всё-таки выполняется.

Определение 20.3. Пусть $\sigma \in S_n$. Элемент $i \in \Omega_n$ называется **неподвижным**, если $\sigma(i) = i$.

Всё множество Ω_n разбивается на два подмножества:

$$\Omega_n = F(\sigma) \sqcup M(\sigma),$$

где $F(\sigma)$ — множество неподвижных элементов, а $M(\sigma)$ — подвижных.

Лемма 20.1. $i \in M(\sigma) \Rightarrow \sigma(i) \in M(\sigma)$.

Доказательство. Действительно, пусть $i \in M(\sigma)$ (обозначим $\sigma(i) = j$) и $\sigma(i) \in F(\sigma)$. Тогда

$$\sigma(i) = \sigma(\sigma(i)) = \sigma(j),$$

при этом $i \neq j$, однако σ — биекция. Противоречие. ■

Теорема 20.2. Если $M(\sigma_1) \cap M(\sigma_2) = \varnothing$, то $\sigma_1 \circ \sigma_2 = \sigma_2 \circ \sigma_1$.

Доказательство. Множества $F(\sigma_1)$ и $M(\sigma_1)$ не пересекаются и полностью покрывают собой множество Ω_n. То же верно и для множеств $F(\sigma_2)$ и $M(\sigma_2)$. Из этого следует, что $M(\sigma_2) \subseteq F(\sigma_1)$ и $M(\sigma_1) \subseteq F(\sigma_2)$ (чтобы убедиться в этом, можно нарисовать картинку).

Не ограничивая общности, пусть $i \in M(\sigma_1)$. Тогда $i \notin M(\sigma_2)$ и, как следствие, $i \in F(\sigma_2)$. А по предыдущей лемме, $\sigma_1(i) \in M(\sigma_1)$ и аналогично получаем $\sigma_1(i) \in F(\sigma_2)$. Итак,

$$(\sigma_1 \circ \sigma_2)(i) = \sigma_1(\sigma_2(i)) = \sigma_1(i), \quad (\sigma_2 \circ \sigma_1)(i) = \sigma_2(\sigma_1(i)) = \sigma_1(i).$$

Таким образом, значения $\sigma_1 \circ \sigma_2$ и $\sigma_2 \circ \sigma_1$ совпадают в каждой точке, а значит, они равны. ■

Определение 20.4. Подстановка $\sigma \in S_n$ называется **циклом**, если существуют числа $i_1, \ldots, i_k \in \Omega_n$, такие что $\sigma(i_1) = i_1, \sigma(i_2) = i_3, \ldots, \sigma(i_{k-1}) = i_k, \sigma(i_k) = i_1$. Остальные элементы отображаются сами в себя. При этом, k называется **длиной цикла**. Сокращённая запись цикла — $[i_1, i_2, \ldots, i_k]$.

Определение 20.5. Циклы σ_1 и σ_2 называются **независимыми**, если

$$M(\sigma_1) \cap M(\sigma_2) = \varnothing.$$

Примечание. Такие циклы, как уже доказано, коммутируют.

Теорема 20.3. $\forall \sigma \in S_n$ существуют независимые циклы $\sigma_1, \ldots, \sigma_k$ такие, что

$$\sigma = \sigma_1 \cdot \ldots \cdot \sigma_N.$$

при этом такое разложение единственно с точностью до порядка множителей.

Доказательство. Доказательство проведём индукцией по $|M(\sigma)|$. База индукции очевидна, докажем шаг. Рассмотрим какой-то $i \in \Omega_n$ и введём следующие обозначения:

$$i_0 := i, \quad i_1 := \sigma(i), \quad i_2 := \sigma^2(i), \quad \ldots, \quad i_k := \sigma^k(i).$$

Иными словами, $i_k = \sigma(i_{k-1})$. Элементов i_j бесконечно много, однако множество Ω_n, в которое они все входят, конечно. А потому последовательность должна с какого-то момента зациклиться, при этом наименьший положительный период r не превосходит n. Итак, имеем

$$\sigma^t(i) = \sigma^{t+r}(i), \quad \varepsilon(i) = \sigma^r(i), \qquad (*)$$

отсюда $\sigma^r(i) = i$, причём r — минимальное положительное число с таким свойством. Утверждается, что тогда числа $i_0, i_1, \ldots, i_{r-1}$ попарно различны. Действительно, если это не так, то, произведя заново выкладку $(*)$, придём к противоречию с минимальностью r. Теперь рассмотрим перестановку

$$\widehat{\sigma} = [i_0, \ldots, i_{r-1}]^{-1} \circ \sigma.$$

Заметим, что

$$\begin{gathered}\widehat{\sigma}(i_0) = [i_{r-1}, \ldots, i_0](\sigma(i_0)) = [i_{r-1}, \ldots, i_0](i_1) = i_0,\\ \widehat{\sigma}(i_1) = [i_{r-1}, \ldots, i_0](\sigma(i_1)) = [i_{r-1}, \ldots, i_0](i_2) = i_1,\\ \vdots\\ \widehat{\sigma}(i_{r-1}) = [i_{r-1}, \ldots, i_0](\sigma(i_{r-1})) = [i_{r-1}, \ldots, i_0](i_0) = i_{r-1}.\end{gathered}$$

Значит, $i_1, \ldots, i_{r-1} \in F(\widehat{\sigma})$. Заметим при этом, что образы остальных элементов такие же, ведь цикл $[i_{r-1}, \ldots, i_0]$ никуда их не переставляет: $\widehat{\sigma}(j) = \sigma(j)$ при $j \notin \{i_0, \ldots, i_{r-1}\}$. Проанализируем, что мы получили. У подстановки σ элементы $i_0, \ldots, i_{r-1}$ были подвижными, т. к., из уже доказанного, числа $\{i_0, \ldots, i_{r-1}\}$ попарно различны. А у новой подстановки $\widehat{\sigma}$ эти элементы стали неподвижными, а остальные — какими были, такими и остались. Значит, $|M(\widehat{\sigma})| < |M(\sigma)|$, пока σ не тождественна. Значит, по предположению индукции, для $\widehat{\sigma}$ существует разложение на независимые циклы. А в произведении с циклом $[i_{r-1}, \ldots, i_0]$ они дадут σ. ■

21 Инверсии. Чётность перестановки и подстановки. Знак подстановки, изменение чётности при умножениии на транспозицию. Разложение подстановки на транспозиции. Знак произведения подстановок

Определение 21.1. В подстановке σ положение двух элементов $\sigma(i)$ и $\sigma(j)$ $(i < j)$ называется **порядком**, если $\sigma(i) < \sigma(j)$ и **инверсией**, если $\sigma(i) > \sigma(j)$.

Определение 21.2. Чётностью подстановки называется чётность общего числа инверсий в ней. **Знак подстановки** $\operatorname{sgn}\sigma := (-1)^{\text{чётность } \sigma}$.

Примечание. Определения для перестановок даются так же, как для соответствующих им подстановок.

Лемма 21.1. При умножении подстановки σ на транспозицию $[i, j]$ слева меняет местами числа i и j, а справа — меняет местами числа $\sigma(i)$ и $\sigma(j)$.

Доказательство. Если $\sigma(x) \notin \{i, j\}$, то $([i, j] \circ \sigma)(x) = \sigma(x)$. Если же $\sigma(x) = i$, то $([i, j] \circ \sigma)(x) = j$ и наоборот. Это доказывает первое утверждение.

Если $x \notin \{i, j\}$, то $[i, j](x) = x$. Теперь $(\sigma \circ [i, j])(i) = \sigma(j)$, а $(\sigma \circ [i, j])(j) = \sigma(i)$. ■

Теорема 21.1. Чётность подстановки меняется при умножении на транспозицию.

Доказательство. При транспозиции соседних элементов меняется взаимное расположение только этих элементов, так что число инверсий изменяется (увеличивается или уменьшается) на 1; следовательно, чётность меняется. Теперь заметим, что

$$[i, j] = [i, i+1] \circ [i+1, i+2] \circ \ldots \circ [j-1, j] \circ [j-2, j-1] \circ \ldots \circ [i, i+1].$$

То есть, транспозиция несоседних элементов раскладывается в произведение $2(j-i)+1$ соседних элементов. При каждой из них меняется чётность подстановки, а всего их нечётное количество, поэтому в итоге чётность изменится. ■

Теорема 21.2. Любая подстановка раскладывается в произведение транспозиций.

Доказательство. Любая подстановка раскладывается в произведение независимых циклов. А каждый цикл раскладывается в произведение транспозиций:

$$[i_1, i_2, \ldots, i_k] = [i_2 \circ i_1] \circ [i_3 \circ i_2] \circ \ldots \circ [i_k, i_{k-1}].$$

■

Примечание. Из двух предыдущих теорем следует, что если $\sigma = \sigma_1 \cdot \ldots \cdot \sigma_N$ — разложение подстановки в произведене транспозиций, то $\operatorname{sgn}\sigma = (-1)^N$. А для цикла (разложение на транспозиции которого приводилось выше) $\operatorname{sgn}[i_1, \ldots, i_k] = (-1)^{k-1}$.

Утверждение (Задача из Кострикина). Доказать, что любая чётная подстановка представима в виде произведения циклов длины 3.

Доказательство. Чётная подстановка раскладывается в произведения чётного числа транспозиций. Осталось заметить, что

$$[i_1, i_2] \circ [i_3, i_4] = [i_1, i_2] \circ \underbrace{[i_1, i_3] \circ [i_1, i_3]}_{\varepsilon} \circ [i_3, i_4] = [i_1, i_2, i_3] \circ [i_1, i_3, i_4].$$

■

Теорема 21.3. $\operatorname{sgn}(\sigma \circ \delta) = \operatorname{sgn}\sigma \cdot \operatorname{sgn}\delta$.

Доказательство. Разложим подстановки в условии в произведения подстановок:

$$\sigma = \sigma_1 \circ \ldots \circ \sigma_N, \quad \delta = \delta_1 \circ \ldots \circ \delta_M,$$

Тогда

$$\operatorname{sgn}(\sigma \circ \delta) = \operatorname{sgn}(\sigma_1 \circ \ldots \circ \sigma_N \circ \delta_1 \circ \ldots \circ \delta_M) = (-1)^{N+M} = (-1)^N \cdot (-1)^M = \operatorname{sgn}\sigma \cdot \operatorname{sgn}\delta.$$

■

Примечание. Как следствие, $\operatorname{sgn}\sigma^{-1} = \operatorname{sgn}\sigma$, т. к.

$$1 = \operatorname{sgn}\varepsilon = \operatorname{sgn}(\sigma \circ \sigma^{-1}) = \operatorname{sgn}\sigma \cdot \operatorname{sgn}\sigma^{-1}$$

и $\operatorname{sgn}(\sigma \circ \delta) = \operatorname{sgn}(\delta \circ \sigma)$. Отсюда же можно извлечь ещё одно полезное следствие:

Теорема 21.4. При $n > 1$ число чётных подстановок равно числу нечётных.

Доказательство. Докажем, что отображение $\sigma \overset{\varphi}{\mapsto} [1, 2] \circ \sigma$ является биекцией из множества чётных подстановок в множество нечётных. Из уже доказанного, если σ чётная, то $[1, 2] \circ \sigma$ нечётная. Докажем же, что φ — биекция. Во-первых, φ — сюръекция, действительно, каждая нечётная подстановка σ является образом чётной подстановки $[1, 2] \circ \sigma$. Во-вторых, φ — инъекция, так как, очевидно, $\sigma \neq \delta \Rightarrow [1, 2] \circ \sigma \neq [1, 2] \circ \delta$. ■

Примечание. Множество чётных подстановок порядка n обозначается A_n. Из уже доказанного, легко увидеть, что они образуют группу по операции композиции.

22 Чётность цикла. Чётность произвольной подстановки через декремент

Чётность цикла была раньше.

Определение 22.1. Декремент подстановки σ — это число $d(\sigma)$, равное n за вычетом количества независимых циклов в разложении σ.

Теорема 22.1. $\operatorname{sgn}\sigma = (-1)^{d(\sigma)}$.

Доказательство. Пусть $k_1, \ldots, k_m$ — длины независимых циклов в разложении σ. Тогда

$$\operatorname{sgn}\sigma = \prod_{i=1}^{m} (-1)^{k_i - 1} = (-1)^{\sum\limits_{i=1}^{m} (k_i - 1)} = (-1)^{n-m} = (-1)^{d(\sigma)}.$$

■

23 Формула определителя квадратной матрицы. Определитель транспонированной матрицы. Линейность и кососимметричность определителя как функции от строк и столбцов матрицы

Определение 23.1. Определителем квадратной матрицы $A = (a_{ij})$ порядка n называется число

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \cdot a_{1\sigma(1)} \dots a_{n\sigma(n)}.$$

Теорема 23.1. $\det A^T = \det A$.

Примечание. Как следствие, любой свойство определителя матрицы по отношению к её строкам верно также и для её столбцов.

Доказательство. Заметим, что т. к. S_n — группа, то если σ пробегает S_n то и σ^{-1} пробегает S_n.

$$\det A^T = \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \cdot a_{\sigma(1)1} \dots a_{\sigma(n)n} = \left\{\delta := \sigma^{-1}\right\} = \sum_{\delta \in S_n} \operatorname{sgn} \delta \cdot a_{1\delta(1)} \dots a_{n\delta(n)} = \det A.$$

■

Теорема 23.2. Определитель является полилинейной кососимметричной функцией строк матрицы.

Доказательство. Линейность определителя по каждой из строк матрицы вытекает из того, что для любого i его можно представить в виде

$$\det A = \sum_j a_{ij} u_j,$$

где $u_1, \dots, u_n$ не зависят от элементов i-ой строки матрицы (видно из формулы определителя).

Для проверки кососимметричности посмотрим, что происходит при перестановке i-ой и j-ой строк матрицы. Из доказательства теоремы 21.4 видно, что отображение $\varphi : \sigma \in S_n \mapsto \sigma \circ [i, j] \in S_n$ биективно, при этом $\operatorname{sgn} \sigma = -\operatorname{sgn} \varphi(\sigma)$ и $\varphi(\sigma) = \varphi^{-1}(\sigma)$. Поэтому подстановки можно разбить на такие пары $(\sigma, \varphi(\sigma))$ (как следствие, если σ пробегает S_n, то и $\varphi(\sigma)$ её пробегает). Обозначим за $A_{(i)\leftrightarrow(j)}$ матрицу, у которой переставлены местами строки с номерами i и j. Тогда

$$\det A_{(i)\leftrightarrow(j)} = \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \cdot a_{1\sigma(1)} \dots a_{i\sigma(j)} \dots a_{j\sigma(i)} \dots a_{n\sigma(n)} =$$
$$= \left\{\delta := \varphi(\sigma)\right\} = \sum_{\delta \in S_n} (-\operatorname{sgn} \delta) \cdot a_{1\delta(1)} \dots a_{i\delta(i)} \dots a_{j\delta(j)} \dots a_{n\delta(n)} = -\det A.$$

■

24 Определитель матрицы с нулевой строкой/столбцом. Определитель матрицы с пропорциональными строками. Изменение определителя при элементарных преобразованиях строк/столбцов. Определитель треугольной матрицы. Алгоритм вычисления определителя с помощью элементарных преобрвазований. Эквивалентные условия невырожденности матрицы. Определитель матрицы с углом нулей

Лемма 24.1. Определитель матрицы со строкой нулей равен нулю.

Доказательство. Пусть i-ая строка матрицы нулевая. В каждом слагаемом из формулы определителя присутствует множитель $a_{i\sigma(i)} = 0$. Поэтому каждое слагаемое равно нулю, поэтому нулю равна вся сумма, т. е. определитель. ■

Теорема 24.1 (Изменения определителя при элементарных преобразованиях). При элементарных преобразованиях

- 1-го типа определитель не изменяется;
- 2-го типа определитель умножается на -1;
- 3-го типа умножается на ненулевую константу.

Доказательство. Докажем через определение, а потом обсудим, почему лучше делать так, а не как рассказывал Сергей Александрович на лекциях.

1. Пусть к i-ой строке прибавили j-ую с коэффициентом λ. Тогда определитель новой матрицы равен

$$\sum_{\sigma\in S_n} \operatorname{sgn}\sigma \cdot a_{1\sigma(1)} \dots (a_{i\sigma(i)} + \lambda a_{j\sigma(i)}) \dots \dots a_{n\sigma(n)} = \det A + \sum_{\sigma\in S_n} \operatorname{sgn}\sigma \cdot a_{1\sigma(1)} \dots a_{j\sigma(i)} \dots a_{j\sigma(j)} \dots a_{n\sigma(n)}.$$

   Посмотрим на образовавшуюся сумму. Как мы уже знаем, каждой подстановке σ можно биективно сопоставить подстановку $\delta = \sigma \circ [i,j]$ (при этом $\sigma(i) = \delta(j)$, $\sigma(j) = \delta(i)$, а остальные элементы переходят туда же, куда и раньше). Поэтому возникает биекция между слагаемыми нашей суммы:

$$\operatorname{sgn}\sigma \cdot a_{1\sigma(1)} \dots a_{j\sigma(i)} \dots a_{j\sigma(j)} \dots a_{n\sigma(n)} \mapsto -\operatorname{sgn}\sigma a_{1\delta(1)} \dots a_{j\delta(j)} \dots a_{i\delta(i)} \dots a_{n\delta(n)}.$$

   Они отличаются только знаком, поэтому в сумме дают 0. А значит, и вся сумма равна 0 и остаётся только $\det A$.

2. Это буквально кососимметричность.

3. Пусть i-ую строку матрицы умножили на $c \neq 0$. Тогда в каждом слагаемом формулы определителя новой матрицы будет присутствовать ровно один множитель из i-ой строки, $c \cdot a_{i\sigma(i)}$. Поэтому, вынеся из каждого слагаемого константу c, получим, что определитель новой матрицы равен $c \cdot \det A$.

■

Примечание. На лекциях Сергей Александрович давал другое доказательство того, что определитель не изменяется при элементарных преобразованиях первого типа. Сначала доказывалась лемма, что определитель с двумя одинаковыми строками равен нулю. Доказательство было такое: если поменять эти строки местами, то из кососимметричности знак определителя изменится, а сама матрица при этом не поменяется. Поэтому $\det A = -\det A$, а отсюда $\det A = 0$. Проблема вот в чём: это доказательство не работает для матриц над полем с характеристикой 2, потому что над этим полем равенство $2 \cdot \det A = 0$ не влечёт $\det A = 0$, т. к. 2 — это и есть 0 (примером такого поля служит $\mathbb{Z}_2$). Мне на это указал принимающий на коллоквиуме, и я придумал доказательство выше.

Теорема 24.2. Определитель треугольной матрицы равен произведению элементов на её главной диагонали.

Лемма 24.2. Для любой нетождественной подстановки $\sigma \in S_n$ существует $i \in \Omega_n$ такой, что $i > \sigma(i)$.

Доказательство. Докажем утверждение индукцией по n.

База индукции ($n = 2$). Единственная нетождественная подстановка порядка 2 — это $[1,2]$. При этом, $2 > [1,2](2) = 1$.

Шаг. Пусть утверждение верно для любого $n < m$. Пусть $\sigma \in S_m$ — нетождественая подстановка. Если $\sigma(m) = m$, то можно рассмотреть её без последнего элемента. Так как если $i \neq m$, то $i \in \Omega_{m-1}$ и $\sigma(i) \in \Omega_{m-1}$, то имеем подстановку $\sigma' \in S_{m-1}$ такую, что $\sigma'(j) = \sigma(j)$ для каждого $j \in \Omega_{m-1}$. Если для неё не существует такого i, что $i > \sigma'(i)$, то по предположению индукции она тождественная. Но тогда и σ тождественная.

А если $\sigma(m) \neq m$, то $\sigma(m) = i \in \Omega_{m-1}$, а значит $m > \sigma(m)$. ■

Теперь докажем теорему 24.2:

Доказательство. Докажем теорему для нижнетреугольных матриц, к случаю верхнетреугольных матриц утверждение будет сводиться транспозицией. Итак, формула для определителя:

$$\det A = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma \cdot a_{1\sigma(1)} \dots a_{n\sigma(n)}.$$

Тождественная подстановка σ соответствует слагаемому $a_{11}\cdot\ldots\cdot a_{nn}$ (т. е. как раз произведению элементов на главной диагонали). А для любой другой подстановки найдётся i такое, что $i > \sigma(i)$, иными словами $a_{i\sigma(i)}$ будет находится выше главной диагонали, а потому $a_{i\sigma(i)} = 0$, а значит, занулится всё слагаемое. ■

Алгоритм вычисления определителя с помощью элементарных преобразований. Пусть мы хотим найти определитель матрицы A. Для этого будем приводить её к ступенчатому виду элементарными преобразованиями. При этом, следить за тем, преобразований каких типов мы применяем и домножать результат на c^{-1}, если применяем преобразование 3-го типа с константой c и на -1, если применяем преобразование 2-го типа (из линейности). А ступенчатый вид — это треугольная матрица, для неё определитель — это произведение диагональных элементов.

Теорема 24.3 (Эквивалентные условия невырожденности матрицы). Следующие условия эквивалентны:

1. К матрице A существует обратная.
2. $\operatorname{rk} A = n$.
3. $\det A \neq 0$.

Доказательство. Равносильность между первым и вторым пунктами уже была доказана. Докажем равносильность между вторым и третьим пунктами.

$\Rightarrow$. Если $\operatorname{rk} A = n$, то улучшенный ступенчатый вид матрицы A единичный, поэтому, находя определитель с помощью элементарных преобразований, не получим ни одного нулевого множителя. А значит, $\det A \neq 0$.

$\Leftarrow$. Находя $\det A$ с помощью элементарных преобразований, получаем $\det A = \lambda \det \widetilde{A}$, где $\widetilde{A}$ — улучшенный ступенчатый вид матрицы A, а $\lambda \neq 0$. $\det A \neq 0 \Rightarrow \det \widetilde{A} \neq 0$. А $\widetilde{A}$ — треугольная матрица, и её определитель равен произведению элементов на главной диагонали. Оно ненулевое, значит, $\widetilde{A} = E$. Отсюда следует невырожденность матрицы и $\operatorname{rk} A = n$. ■

Теорема 24.4 (Об определителе матрицы с углом нулей). Пусть матрица A имеет вид

$$A = \begin{pmatrix} B & D \\ 0 & C \end{pmatrix},$$

где B и C — квадратные матрицы. Тогда

$$\det A = \det B \cdot \det C.$$

Доказательство. При фиксированных B и D определитель матрицы A является полилинейной и кососимметричной её последних строк и, тем самым, кососимметричной и полилинейной функцией строк матрицы C. Согласно теореме 25.1, получаем

$$\det A = \det \begin{pmatrix} B & D \\ 0 & E \end{pmatrix} \cdot \det C.$$

Первый множитель, в свою очередь, является полилинейной и кососимметричной функцией первых столбцов матрицы, а потому (по той же теореме) получаем

$$\det A = \det \begin{pmatrix} B & D \\ 0 & E \end{pmatrix} \cdot \det C = \det \begin{pmatrix} E & D \\ 0 & E \end{pmatrix} \cdot \det B \cdot \det C,$$

но матрица $\begin{pmatrix} E & D \\ 0 & E \end{pmatrix}$ треугольная с единицами на главной диагонали, поэтому её определитель равен 1. ■

Утверждение (Задача Антона Александровича). Если матрицы A, B, C и D квадратные порядка n и при этом A и C коммутируют, то

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - CB).$$

Доказательство. Решить до конца не получилось, но вот наброски. Предположим, что матрица A невырождена. Тогда проводим элементарные преобразования:

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix} = \det A \cdot \det(D - CA^{-1}B) = \det(AD - ACA^{-1}B) = \det(AD - CB).$$

Однако непонятно, как избавиться от условия $\det A \neq 0$. В книге Р. Ф. Гантмахера «Теория матриц» я нашёл следующее утверждение (хотя и без доказательства): для любой вырожденной матрицы можно построить последовательность невырожденных, стремящихся к ней (покоординатно). А поэтому (в книге ровно так и написано) условие можно убрать из соображений непрерывности. ■

25 Единственность с точностью до пропорциональности линейной кососимметрической функции строк/столбцов. Определитель произведения матриц

Теорема 25.1. Всякая функция Φ на множестве квадратных матриц порядка n, являющаяся кососимметрической полилинейной функцией строк матрицы, имеет вид

$$\Phi(A) = \Phi(E) \det A.$$

Примечание. Во всех доказательствах этой теоремы, которые мне известны, без дополнительных объяснений используется утверждение «если у кососимметрической функции два аргумента равны, то она зануляется». Но я не понимаю, как его доказать, не упираясь при этом в проблему, описанную в примечании после доказательства теоремы 24.1.

Доказательство. Пусть Φ — полилинейная кососимметрическая функция строк матрицы. Пусть $e_1, e_2, \ldots, e_n$ — единичные строки. Тогда из линейности Φ

$$\Phi(A) = \Phi(A_{(1)}, A_{(2)}, \ldots, A_{(n)}) = \Phi\left(\sum_{i_1} a_{1i_1} e_{i_1}, \sum_{i_2} a_{2i_2} e_{i_2}, \ldots, \sum_{i_n} a_{ni_n} e_{i_n}\right) =$$
$$= \sum_{i_1,\ldots,i_n \in S_n} a_{1i_1} a_{2i_2} \ldots a_{ni_n} \Phi(e_{i_1}, e_{i_2}, \ldots, e_{i_n}).$$

При этом, если какие-то из $i_1, \ldots, i_n$ равны, то $\Phi(e_{i_1}, \ldots, e_{i_n}) = 0$ в силу кососимметричности функции. Поэтому можно считать только слагаемые, у которых $i_1, \ldots, i_n$ попарно различны. Тогда они однозначно задают перестановку $(i_1, \ldots, i_n)$, а эта перестановка в свою очередь сопоставляется подстановке $\sigma : j \in \Omega_n \mapsto i_j \in \Omega_n$. Получаем

$$\Phi(A) = \sum_{\sigma \in S_n} \Phi(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) a_{1\sigma(1)} \ldots a_{n\sigma(n)}.$$

Из кососимметричности Φ, количество операций, требуемых, чтобы переставить строки $e_{\sigma(i)}, \ldots, e_{\sigma(n)}$ в том же порядке, в котором они шли в единичной матрице, равно количеству инверсий в перестановке σ, а значит,

$$\Phi(e_{\sigma(1)}, e_{\sigma(2)}, \ldots, e_{\sigma(n)}) = \operatorname{sgn} \sigma \cdot \Phi(E).$$

Подставляя этот результат вы полученное ранее выражение, получим требуемое. ■

Теорема 25.2. $\det AB = \det A \cdot \det B$.

Доказательство. $\det AB$ является полилинейной кососимметрической функцией от строк матрицы AB. А эти строки, в свою очередь, являются линейными комбинациями строк A (по лемме 16.1). Поэтому $\det AB$ — полилинейная кососиммерическая функция от строк матрицы A. Отсюда

$$\det AB = \det EB \cdot \det A = \det B \cdot \det A.$$

■

26 Миноры. Алгебраические дополнения. Разложение определителя по строке/столбцу

Определение 26.1. Пусть A — произвольная матрица. Всякая матрица, составленная из элементов матрицы A, находящихся на пересечении каких-либо выбранных строк и каких-либо выбранных столбцов, называется **подматрицей** матрицы A.

Определение 26.2. Определитель квадратной подматрицы порядка k называется **минором** порядка k матрицы A. В частности, если A — квадратная матрицы порядка n, то минор поряка $n-1$, получаемый вычёркиванием i-ой строки и j-го столбца, называется **дополнительным минором** элемента a_{ij} и обозначается через M_{ij}.

Определение 26.3. Число

$$A_{ij} = (-1)^{i+j} M_{ij}$$

называется **алгебраическим дополнением** элемента a_{ij}.

Теорема 26.1 (Формула разложения определителя по строке/столбцу).

$$\det A = \sum_j a_{ij} A_{ij} = \sum_i a_{ij} A_{ij}.$$

Сначала докажем следующую лемму:

Лемма 26.1.

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & a_{ij} & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix} = a_{ij} A_{ij}.$$

Доказательство. Поменяем местами i-ую строку со всеми предыдущими, а затем j-ый столбец со всеми предыдущими. При этом всего мы произведём $i-1$ перестановку строк и $j-1$ перестановку столбцов. Поэтому результат домножится на

$$(-1)^{i-1+j-1} = (-1)^{i+j}.$$

В результате получится определитель вида

$$\det \begin{pmatrix} a_{1j} & 0 & \dots & 0 \\ a_{1j} & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{nj} & a_{n1} & \cdots & a_{nn} \end{pmatrix} = a_{ij} M_{ij}.$$

Последнее равенство выполняется в силу теоремы об определителе матрицы с углом нулей. С учётом вычисленного нами знака, отсюда и получается доказываемое равенство. ■

Теперь несложно доказать теорему 26.1:

Доказательство. Вспомним формулу для вычисления определителя:

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn} \sigma a_{1\sigma()} \dots a_{n\sigma(n)}.$$

Каждое слагаемое содержит ровно 1 элемент из i-ой строки, а предыдущая лемма означает, что сумма тех членов, которые содержат a_{ij} равна $a_{ij} A_{ij}$. Отсюда вытекает формула разложеиня по строке. Аналогично доказывается формула разложения по столбцу. ■

Примечание. Антон Александрович рассказывал, как быстро понимать знак у алгебраического дополнения A_{ij} (возведение -1 в степень — сложная и трудоёмкая операция). Он предложил следующую визуализацию:

$$\begin{pmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

27 Фальшивое разложение определителя по строке/столбцу. Явная формула для обратной матрицы

Теорема 27.1 (Фальшивое разложение определителя по строке/столбцу).

$$\sum_{j=1}^{n} a_{ij}A_{kj} = 0 \text{ при } k \neq i, \qquad \sum_{i=1}^{n} a_{ij}A_{ik} = 0 \text{ при } k \neq j.$$

Доказательство. Рассмотрим матрицу $A' = (a'_{ij})$, полученную из A заменой k-ой строки на i-ую. В матрице A' две одинаковые строки, следовательно, $\det A' = 0$. Разложим определитель матрицы A' по k-ой строке. Получим

$$0 = \det A' = \sum_{j=1}^{n} a'_{kj}A'_{kj} = \sum_{j=1}^{n} a_{ij}A_{kj}.$$

Аналогично получается фальшивое разложение по столбцу. ■

Теорема 27.2 (Явная формула для обратной матрицы). Пусть $A = (a_{ij})$ — невырожденная квадртная матрица. Тогда

$$A^{-1} = \frac{1}{\det A} \underbrace{\begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}}_{\widehat{A}^T}.$$

Доказательство. Докажем равносильное требуемому равенство:

$$A \cdot \widehat{A}^T = (\det A) \cdot E.$$

Перемножим по определению:

$$(A\widehat{A}^T)_{ii} = \sum_{t=1}^{n} a_{it}\widehat{a}_{ti} = \sum_{t=1}^{n} a_{it}A_{it} = \det A, \qquad (A\widehat{A}^T)_{ij} = \sum_{t=1}^{n} a_{ij}\widehat{a}_{tj} = \sum_{t=1}^{n} a_{it}A_{jt} = 0.$$

Первое равенство выполнено в силу теорему о разложении определителя по i-ой строке, а второе — по теореме о фальшивом разложении определителя по i-ой строке. ■

Примечание. Матрица $\widehat{A}$ в обозначениях предыдущей теоремы иногда называют **присоединённой матрицей**, но Сергей Александрович присоединённой матрицей называется $\widehat{A}^T$, а Винберг этого определения вообще не вводит.

Утверждение (Задача Антона Александровича). Пусть A — невырожденная целочисленная квадратная матрица. Матрица A^{-1} является целочисленной тогда и только тогда, когда $\det A = \pm 1$.

Доказательство. $\Rightarrow$. Заметим, что

$$\det A \cdot \det A^{-1} = \det AA^{-1} = \det E = 1.$$

Причём, матрица A^{-1} целочисленная, а поэтому и $\det A^{-1}$ (как и $\det A$) является целым числом (видно из формулы по определению). Поэтому числа $\det A$ и $\det A^{-1}$ целые и взаимно обратные. Значит, они равны ± 1.

$\Leftarrow$. Видно из явной формулы для обратной матрицы. ■

Утверждение (Задача Антона Александровича). Доказать, что в матрице A любой минор ранга $\operatorname{rk} A$, образуемый пересечением линейной независимых строк и линейно независимых столбцов, отличен от нуля.

Доказательство. Во-первых, можно перестановкой строк (элементарными преобразованиями 2 типа) передвинуть наш минор в левый верхний угол матрицы. Теперь, строки, входящие в минор, образуют базис

системы строк матрицы A, а поэтому можно, вычитая их линейные комбинации из остальных строк (элементраные преобразования 1 типа) занулить все строки, начиная с $\operatorname{rk} A + 1$. Ранее мы уже доказывали, что при этом линейные зависимости между столбцами не меняются (см. доказательство леммы 9.2), а потому первые $\operatorname{rk} A$ столбцов образуют базис системы столбцов матрицы A. Поступаем аналогично, вычитая эти столбцы из остальных. После этих действий, в нашей матрице останется только выбранный нами минор, а остальные элементы занулятся. При этом, мы совершали элементарные преобразования, поэтому ранг не изменился. Если предположить, что выбранный нами минор нулевой, но его ранг меньше $\operatorname{rk} A$, но тогда и ранг оставшейся матрицы меньше $\operatorname{rk} A$. Противоречие. ■

Примечание. Эти задачи есть также в Винберге, но впервые я их узнал от Антона Александровича.

28 Формулы Крамера

Теорема 28.1 (Крамер). Пусть дана система уравнений

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1, \\ \vdots \\ a_{n1}x_1 + \ldots + a_{nn}x_n = b_n \end{cases}$$

и $A = (a_{ij})$ — матрица её коэффициентов. Тогда эта система определена тогда и только тогда, когда $\det A \neq 0$. В этом случае решение находится по формулам

$$x_i = \frac{\det A_i}{\det A},$$

где A_i — матрица, полученная из A заменой её i-го столбца столбцом свободных членов.

Доказательство. Докажем первую часть теоремы. Условие $\det A \neq 0$ равносильно $\operatorname{rk} A = n$, что, в свою очередь, равносильно определённости системы (теорема 10.3).

Теперь перейдём ко второй части. При любом элементарном преобразовании системы в матрицах A и A_i одновременно происходит соответствующее элементарное преобразование строк и, следовательно, отношения, стоящие в правых частях формул Крамера, не изменяются. С помощью элементарных преобразований строк матрицу A можно привести к единичной матрице. Поэтому достаточно доказать теорему в том случае, когда $A = E$. Тогда система имеет вид

$$\begin{cases} x_1 & & & = b_1, \\ & x_2 & & = b_2, \\ & & \ddots & \vdots \\ & & x_n & = b_n. \end{cases}$$

Она, очевидно, имеет единственное решение $x_i = b_i$ $(i = 1, 2, \ldots, n)$. С другой стороны,

$$\det A = \det E = 1, \quad \det A_i = \det \begin{pmatrix} 1 & 0 & \cdots & b_1 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & b_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & b_i & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & b_{n-1} & \cdots & 1 & 0 \\ 0 & 0 & \cdots & b_n & \cdots & 0 & 1 \end{pmatrix} = b_i,$$

так что формулы Крамера в этом случае действительно верны. ■

29 Теорема о ранге матрицы. Метод окаймляющх миноров

Теорема 29.1 (О ранге матрицы). Пусть $A \in \underset{m\times n}{\mathrm{Mat}}$. Ранг матрицы A равен максимальному порядку ненулевого минора этой матрицы.

Доказательство. Для того, чтобы доказать заявленное равенство, докажем неравенства в одну и другую сторону. Сперва докажем, что ранг не меньше, чем порядок любого ненулевого минора данной матрицы. Пусть в A есть ненулевое минор порядка k, и пусть это $M_{i_1,\dots,i_k}^{j_1,\dots,j_k}$. Так как определитель этой $k \times k$ матрицы ненулевой, её строки линейно независимы. Однако строки данной матрицы — это строки $A_{(i_1)}, \dots, A_{(i_k)}$, в которых убраны некоторые координаты с одинаковыми номерами. Следовательно, строки $A_{(i_1)}, \dots, A_{(i_k)}$ линейно независимы, то есть, $\operatorname{rk} A \geqslant k$.

Пусть теперь $\operatorname{rk} A = k$. Докажем, что в A найдётся ненулевой минор порядка k. Выберем базисные строки $A_{(i_1)}, \dots, A_{(i_k)}$ и рассмотрим подматрицу P размера $k \times n$ в матрицы A состоящую из этих строк. Так как строки P линейно независимы, то $\operatorname{rk} P = k$. Выберем базисные столбцы матрицы P, их ровно k штук, они линейно независимы. Это в точности подматрицы порядка $k \times k$ данной матрицы и её ранг равен k, а значит, её определитель ненулевой. ■

Определение 29.1. Минор $(k+1) \times (k+1)$ называется **окаймляющим** минором для данного минора $k \times k$, если соответствующая подматрицы получена добавлением одной строки и одного столбца к подматрице минора $k \times k$.

Теорема 29.2 (Метод окаймляющих миноров). Если данный минор $k \times k$ матрицы A не равен нулю, а все его окаймляющие миноры равны нулю, то $\operatorname{rk} A = k$.

Доказательство. По теореме о ранге матрицы $\operatorname{rk} A \geqslant k$. Допустим, что $\operatorname{rk} A \geqslant k+1$. Пусть данный нам ненулевой минор стоит на пересечении строк с номерами $i_1, \dots, i_k$ и столбцов с номерами $j_1, \dots, j_k$. Тогда существует строка с номером $i_{k+1} \notin \{i_1, \dots, i_k\}$, такая что система строк с номерами $i_1, \dots, i_k, i_{k+1}$ линейно независима. Рассмотрим матрицу P размера $(k+1) \times n$, состоящую из этих строк. Ранг этой матрицы $k+1$. Тогда столбцы $P^{(j_1)}, \dots, P^{(j_k)}$ линейно независимы, т. к. даже есть убрать строчку i_{k+1}, то они будут таковыми. Дополним столбцы $P^{(j_1)}, \dots, P^{(j_k)}$ до базиса системы столбцов P некоторым столбцом $P^{(j_{k+1})}$. Матрицы, состоящая из столбцов $P^{(j_1)}, \dots, P^{(j_k)}, P^{(j_{k+1})}$ имеет ненулевой определитель. Но это подматрицы $(k+1) \times (k+1)$ в A, что противоречит условию. ■

30 Определитель Вандермонда. Задача интерполяции

Теорема 30.1 (Определитель Вандермонда).

$$\det \begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix} = \prod_{1 \leqslant j < i \leqslant n} (x_i - x_j).$$

Доказательство. Обозначим данный определитель через $V(x_2, \dots, x_n)$ и докажем утверждение индукцией по n:

База индукции ($n = 2$).

$$V(x_1, x_2) = \det \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \end{pmatrix} = x_2 - x_1.$$

Шаг индукции. Вычтем каждый столбец, умноженый на x_1 из следующего, при этом двигаясь справа налево по столбцам матрицы. Это элементарное преобразование, а потому определитель не изменится.

$$V(x_1, \dots, x_n) = \det \begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & x_2 - x_1 & x_2(x_2 - x_1) & \cdots & x_2^{n-2}(x_2 - x_1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n - x_1 & x_n(x_n - x_1) & \cdots & x_n^{n-2}(x_n - x_1) \end{pmatrix}.$$

По теореме об определителе матрицы с углом нулей, последний определитель равен

$$\det \begin{pmatrix} x_2 - x_1 & x_2(x_2 - x_1) & \cdots & x_2^{n-2}(x_2 - x_1) \\ \vdots & \ddots & \vdots & \vdots \\ x_n - x_1 & x_n(x_n - x_1) & \cdots & x_n^{n-2}(x_n - x_1) \end{pmatrix} = (x_2 - x_1) \dots (x_n - x_1) V(x_2, \dots, x_n).$$

А по предположению индукции $V(x_2,\dots,x_n)=\prod\limits_{2\leqslant j<i\leqslant n}(x_i-x_j)$, отсюда следует требуемое. ■

Задача интерполяции заключается в том, чтобы найти функцию $f(x)$ (обычно из некоторого заданного класса) такую, что в n попарно различных точках $x_1,\dots,x_n$ она принимает заданные значения $y_1,\dots,y_n$.

Теорема 30.2. Для любых различных $x_1,\dots,x_n$ и любых заданных $y_1,\dots,y_n$ существует единственный многочлен степени не более $n-1$, такой что $f(x_i)=y_i$ $(i=1,\dots,n)$.

Доказательство. Будем искать этот многочлен методом неопределённых коэффициентов. У многочлена не более чем $(n-1)$-ой степени их n (некоторые могут получиться нулями). Тогда условие $f(x_i)-y_i$ даёт линейное условие на эти коэффициенты, а совокупность таких условий является квадратной СЛУ. Матрица её коэффициентов выглядит так:

$$\begin{pmatrix}1 & x_1 & \cdots & x_1^{n-1}\\ 1 & x_2 & \cdots & x_2^{n-1}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & x_n & \cdots & x_n^{n-1}\end{pmatrix}$$

Её определитель равен $V(x_1,\dots,x_n)$, а т. к. все x_i попарно различны, то он ненулевой, и система определена. Так, коэффициенты нашего многочлена существуют и единственны.

Этот многочлен можно также выписать явно:

$$f(x)=\sum_{i=1}^{n}\left(\frac{\prod\limits_{j\neq i}(x-x_j)}{\prod\limits_{j\neq i}(x_i-x_j)}y_i\right).$$

Это действительно многочлен, т. к. знаменатели — константы, не равные нулю. Числители — многочлены степени $n-1$, поэтому их сумма — многочлен степени не более $n-1$. Если мы подставим в этот многочлен x_k, то все слагаемые, кроме k-го обратятся в ноль. А k-ое слагаемое становится равным y_k. Этот многочлен называется **интерполяционным многочленом Лагранжа**. ■

31 Понятие группы, абелевой группы. Примеры. Простейшие следствия из аксиом. Подгруппа. Критерий того, что подмножество является подгруппой. Порядок элемента, порядок подстановки. Циклическая группа и её порядок

Определение 31.1. Множество G с одной бинарной операцией $(x,y)\mapsto x*y$ называется **группой**, если выполнены следующие аксиомы:

1. $\forall x,y,z\in G$ выполнено $(x*y)*z=x*(y*z)$.
2. $\exists e\in G$ такой, что $\forall x\in G$ выполнено $e*x=x*e=x$.
3. $\forall x\in G$ существует $x^{-1}\in G$ такой, что $x*x^{-1}=x^{-1}*x=e$.

Определение 31.2. Группа $(G,*)$ называется **абелевой** (**коммутативной**), если дополнительно выполнена аксиома

4. $\forall x,y\in G$ выполнено $x*y=y*x$.

Определение 31.3. Если для множества с бинарной операцией выполнена только первая аксиома, то оно называется **полугруппой**, а если ещё и вторая, то **моноидом**.

Определение 31.4. Порядком группы G называется мощность множества различных её элементов $|G|$. Если количество элементов в группе G конечно, то группа называется **конечной**, а иначе — **бесконечной**.

Примечание. Если A — любое ассоциативное кольцо с единицей, то множество его обратимых элементов является группой по умножению и обозначается A^* (доказательство позднее).

Примеры групп:

1. Числовые группы (все они абелевы): $(\mathbb{Z},+)$, $(\mathbb{Q},+)$, $(\mathbb{R},+)$, $(\mathbb{Q}\setminus 0,\cdot)$, $(\mathbb{R}\setminus 0,\cdot)$.

2. Матрицы по сложению (все они абелевы): $\underset{m\times n}{\mathrm{Mat}}(\mathbb{R})$, $\underset{m\times n}{\mathrm{Mat}}(\mathbb{Q})$, $\underset{m\times n}{\mathrm{Mat}}(\mathbb{Z})$.
3. Вычеты $\bmod n$ (они абелевы) $(\mathbb{Z}_n, +)$, Z_n^*.
4. Группы подстановок по умножению: (при $n \geqslant 3$ не абелевы): $(S_n, \circ)$, $(A_n, \circ)$.
5. Невырожденные матрицы по умножению (при $n \geqslant 2$ не абелевы):

$$\mathrm{GL}_n(\mathbb{R}) := \Big(\big\{A \in \underset{n\times n}{\mathrm{Mat}}(\mathbb{R}) \ : \ \det A \neq 0\big\}, \cdot\Big), \quad \mathrm{GL}_n(\mathbb{Q}) := \Big(\big\{A \in \underset{n\times n}{\mathrm{Mat}}(\mathbb{Q}) \ : \ \det A \neq 0\big\}, \cdot\Big)$$

6. Преобразования: $(S(X), \circ)$, где $S(X)$ — множество биекций множества X.

Мультипликативная и аддитивная терминологии. Во многих примерах операцию $*$ мы будем называть умножением («Потому что нам так удобно и хочется, и позволяют региональные правила»). Такие группы будем называть **мультипликативными**, нейтральный элемент в них **единицей** (обозначается e или 1 или ε), а обратный элемент к x обозначать x^{-1}.

В других примерах операция — это сложение, причём обычно с ней получается абелева группа, поэтому операцию в произвольной абелевой группе принято называть сложением. При этом нейтральный элемент называется **нулём** (обозначается $\mathbf{0}$ или 0), а обратный элемент к x обозначается $-x$.

Теорема 31.1 (Простейшие следствия из аксиом). Пусть G — мультипликативная группа. Тогда

1. Нейтральный элемент в G единственный.
2. Обратный элемент к данному единственный.
3. $xy = xz \Rightarrow y = z$, $yx = zx \Rightarrow y = z$.
4. $xy = e \Rightarrow (x = y^{-1}) \wedge (y = x^{-1})$.
5. $(x^{-1})^{-1} = x$.
6. Выполнена обобщённая ассоциативность.
7. $(xy)^{-1} = (y^{-1}x^{-1})$.

Доказательство.

1. Пусть есть два нейтральных элемента e_1 и e_2. Тогда
$$e_1 = e_1 e_2 = e_2.$$
2. Пусть к элементу x есть два обратных: $(x^{-1})_1$ и $(x^{-1})_2$. Тогда
$$(x^{-1})_1 = (x^{-1})_1\big(x \cdot (x^{-1})_2\big) = \big((x^{-1})_1 \cdot x\big)(x^{-1})_2 = (x^{-1})_2.$$
3. Умножим обе части на x^{-1} слева или справа.
4. В самом деле, $g^{-1}g = e \Rightarrow (g^{-1})^{-1} = g$.
5. Следствие теоремы 15.1.
6. Проверим непосредственно:
$$y^{-1}x^{-1}xy = y^{-1}y = e.$$

■

Определение 31.5. Пусть $g \in G$, $k \in \mathbb{Z}$. Определим

$$g^k = \begin{cases} \underbrace{gg\ldots g}_{k \text{ раз}}, & \text{если } k > 0, \\ e, & \text{если } k = 0, \\ \underbrace{g^{-1}g^{-1}\ldots g^{-1}}_{-k \text{ раз}}, & \text{если } k < 0. \end{cases}$$

Лемма 31.1. $g^{(k+\ell)} = g^k g^\ell$, $(g^k)^\ell = g^{k\ell}$.

Доказательство. Заметим, что утверждение очевидно для $k, \ell > 0$. Рассмотрим случай, когда $k > 0$, $\ell < 0$, $k + \ell > 0$. Тогда

$$g^k g^\ell = \underbrace{gg\ldots g}_{k}\underbrace{g^{-1}g^{-1}\ldots g^{-1}}_{-\ell} = \underbrace{gg\ldots g}_{k+\ell} = g^{k+\ell}, \quad (g^k)^\ell = \underbrace{g^{-k}g^{-k}\ldots g^{-k}}_{-\ell} = g^{k\ell}.$$

Аналогично разбираются и другие случаи. ■

Определение 31.6. Пусть $(G, *)$ — группа. Подмножество $H \subseteq G$ называется **подгруппой** группы G, если $(H, *)$ является группой.

Теорема 31.2 (Критерий подгруппы). Пусть $H \subseteq G$. Тогда H — подгруппа G тогда и только тогда, когда

1. $H \neq \varnothing$
2. $h_1, h_2 \in H \Rightarrow h_1 h_2 \in H$
3. $h \in H \Rightarrow h^{-1} \in H$

Доказательство. Неободимость первых двух условий очевидна. H — подгруппа, значит, в ней для каждого элемента h есть обратный, при этом он также есть и в G (т. к. $H \subseteq G$). Однако обратный элемент к h в G только один, а значит, $h^{-1} \in H$.

Теперь докажем достаточность. Пусть условия выполнены. Тогда из условий 1 и 2, H — непустое множество с бинарной операцией (той же, что и в G, а значит, ассоциативной). По условию 1 найдётся $h \in H$. По условию 3, $h^{-1} \in H$. По условию 2, $hh^{-1} = e \in H$, то есть, в H есть нейтральный элемент. ■

Примечание. Из леммы 31.1 следует, что $(g^k)^{-1} = g^{-k}$. Кроме того, $e = g^0$. Таким образом, степени элемента g образуют подгруппу в группе G.

Определение 31.7. Группа G называется **циклической**, если она целиком состоит из степеней его элемента g. Тогда этот элемент g называется **порождающим**. Обозначается это так: $G = \langle g \rangle$.

Определение 31.8. Порядком элемента $g \in G$ называется минимальное натуральное k, такое что $g^k = e$. Если же такого числа k не существует, то говорят порядок k считается равным ∞. Обозначается порядок элемента g через $\operatorname{ord} g$.

Теорема 31.3. Порядок подстановки $\sigma \in S_n$ равен наименьшему общему кратному длин циклов в разложении σ в произведение независимых циклов.

Доказательство. Пусть разложение σ в произведение независимых циклов имеет вид $\sigma = \xi_1 \circ \ldots \circ \xi_N$, где ξ_i — цикл длины ℓ_i. Тогда, т. к. независимые циклы коммутируют, выполнено

$$\sigma^k = \xi_1^k \circ \ldots \circ \xi_N^k.$$

При этом подстановки $\xi_1^k, \ldots, \xi_N^k$ переставляют не пересекающиеся множества элементов. Из этого следует, что $\xi_1^k \circ \ldots \circ \xi_N^k = \varepsilon$ тогда и только тогда, когда $\xi_i^k = \varepsilon$ для каждого i. Но цикл в степени равен тождественной подстановке тогда и только тогда, когда степень делит длину данного цикла. Так, $\operatorname{ord}\sigma \mid \ell_i$ для каждого i и при этом, $\operatorname{ord}\sigma$ — минимальное число с таким свойством, то

$$\operatorname{ord}\sigma = \operatorname{НОК}(\ell_1, \ldots, \ell_N).$$

■

Утверждение (Задача из Винберга). Порядок любого элемента группы S_n не превосходит числа

$$e^{n/e} \approx 1.44^n.$$

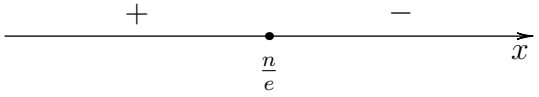
Доказательство. В обозначениях теоремы 31.3 максимальный порядок равен

$$\begin{aligned} \max_{k\in\mathbb{N}} \max\left\{\operatorname{НОК}(\ell_1,\ldots,\ell_k) \ :\ \ell_1,\ldots,\ell_k \in \mathbb{N}, \ell_1+\ldots+\ell_k = n\right\} \leqslant \\ \leqslant \max_{k\in\mathbb{N}} \max\left\{\ell_1\cdot\ldots\cdot\ell_k \ :\ \ell_1,\ldots,\ell_k \in \mathbb{N}, \ell_1+\ldots+\ell_k = n\right\} \leqslant \\ \leqslant \max_{k\in\mathbb{N}} \max\left\{\ell_1\cdot\ldots\cdot\ell_k \ :\ \ell_1,\ldots,\ell_k \in \mathbb{R}_+, \ell_1+\ldots+\ell_k = n\right\} \leqslant \max_{k\in\mathbb{N}} \left(\frac{n}{k}\right)^k \leqslant \max_{k\in\mathbb{R}_+} \left(\frac{n}{k}\right)^k. \end{aligned}$$

Осталось лишь найти максимум функции $f_n(x) = \left(\frac{n}{x}\right)^x$ на $\mathbb{R}_+$. Для этого продиффиренцируем её:

$$f_n(x) = e^{x\ln\frac{n}{x}}, \quad f'(x) = \left(e^{x\ln\frac{n}{x}}\right) = e^{x\ln\frac{n}{x}}\left(\ln\frac{n}{x} + x\cdot\frac{x}{n}\cdot\frac{-n}{x^2}\right) = e^{x\ln\frac{n}{x}}\left(\ln\frac{n}{x} - 1\right).$$

Отсюда $f_n'(x) = 0 \Leftrightarrow \ln\frac{n}{x} = 1$, отсюда $x = n/e$. Проверим, что нашли точку максимума:



Из правого луча можно взять число n, а из левого — $\frac{n}{2e}$. Итак, максимум функции f_n (а вместе с ним и верхняя оценка на порядок подстановки порядка n) равен $f_n(\frac{n}{e}) = n^{n/e}$. ■

Лемма 31.2. Если $\operatorname{ord} g = n$, то

1. $g^m = e \Leftrightarrow n \mid m$;
2. $g^k = g^\ell \Leftrightarrow k \equiv \ell \pmod n$.

Доказательство.

1. Разделим m на n с остатком:
$$m = qn + r, \quad 0 \leqslant r < n.$$
Тогда в силу определения порядка
$$g^m = (g^n)^q \cdot g^r = g^r = e \Leftrightarrow r = 0.$$
2. В силу предыдущего
$$g^k = g^\ell \Leftrightarrow g^{k-\ell} = e \Leftrightarrow n \mid (k - \ell) \Leftrightarrow k \equiv \ell \pmod n.$$

■

Теорема 31.4. $|\langle g\rangle| = \operatorname{ord} g$.

Доказательство. Пусть $\operatorname{ord} g = n$. Тогда
$$\langle g\rangle = \{e, g, g^2, \ldots, g^{n-1}\},$$
причём все перечисленные элементы различны. ■

Лемма 31.3. Если $\operatorname{ord} g = n$, то
$$\operatorname{ord} g^k = \frac{n}{\text{НОД}\,(n, k)}$$

Доказательство. Пусть $\text{НОД}\,(n, k) = d$, $n = n_1 d$, $k = k_1 d$, так что $\text{НОД}\,(n_1, k_1) = 1$. Имеем
$$(g^k)^m = e \Leftrightarrow n \mid km \Leftrightarrow n_1 \mid k_1 m \Leftrightarrow n_1 \mid m.$$
Следовательно, $\operatorname{ord} g^k = n_1$. ■

Определение 31.9. Группа G называется **циклической**, если существует такой элемент $g \in G$, что $G = \langle g\rangle$. Всякий такой элемент называется **порождающим элементом** группы G.

Утверждение. Элемент g^k циклической подгруппы $G = \langle g\rangle$ порядка n является порождающим тогда и только тогда, когда $\text{НОД}\,(n, k) = 1$.

Доказательство. Очевидное следствие леммы 31.1. ■

Теорема 31.5. Всякая бесконечная циклическая группа изоморфна группе $\mathbb{Z}$. Всякая конечная циклическая группа порядка n изоморфна группе $\mathbb{Z}_n$.

Доказательство. Если $G = \langle g\rangle$ — бесконечная циклическая группа, то в силу формул леммы 31.1 отображение $f : k \in \mathbb{Z} \mapsto g^k \in G$ есть изоморфизм.

Пусть теперь $G = \langle g\rangle$ — конечная группа порядка n. Рассмотрим отображение
$$f : [k] \in \mathbb{Z}_n \mapsto g^k \in G \quad (k \in \mathbb{Z}).$$
Так как
$$[k] = [\ell] \Leftrightarrow k \equiv \ell \pmod n \Leftrightarrow g^k = g^\ell,$$

то отображение f корректно определено и биективно. Свойство

$$f(k+\ell) = f(k) \cdot f(\ell)$$

напрямую вытекает из формул леммы 31.1. ■

Теорема 31.6.

1. Всякая подгруппа циклической группы является циклической
2. В циклической группе порядка n порядок любой подгруппы делит n и для любого делителя q числа n существует ровно одна группа порядка q.

Доказательство.

1. Пусть $G = \langle g \rangle$ — циклическая группа и H — её подгруппа, отличная от $\{e\}$ (единичная подгруппа, очевидно, является циклической). Заметим, что если $g^{-m} \in H$ для какого-либо $m \in \mathbb{N}$, то и $g^m \in H$. Пусть m — наименьшее из натуральных чисел, для которых $g^m \in H$. Докажем, что $H = \langle g^m \rangle$. Пусть $g^k \in H$. Разделим k на m с остатком:

   $$k = qm + r, \quad 0 \leqslant r < m.$$

   Имеем

   $$g^r = g^k (g^m)^{-q} \in H,$$

   откуда в силу определения числа m следует, что $r = 0$ и, значит, $g^k = (g^m)^q$.
2. Если $|G| = n$, то предыдущее рассуждение, применённое к $k = n$ (в этом случае $g^k = e \in H$), показывает, что $n = qm$. При этом

   $$H = \{e, g^m, g^{2m}, \ldots, g^{(q-1)m}\}, \qquad (*)$$

   и H является единственной подгруппой порядка q в группе G. Обратно, если q — любой делитель n и $n = qm$, то подмножество H, определяемое равенством $(*)$, является подгруппой порядка q.

■

32 Левые смежные классы по подгруппе. Индекс подгруппы. Теорема Лагранжа

Определение 32.1. Пусть G — группа и H — её подгруппа. Будем говорить, что элементы $g_1, g_2 \in G$ **сравнимы по модулю** H, и писать $g_1 \equiv g_2 \pmod{H}$, если

$$g_1^{-1} g_2 \in H, \qquad (*)$$

т. е. $g_2 = g_1 h$, где $h \in H$.

Примечание. Это определение обобщает определение сравнимости целых чисел по модулю n, которое получается в случае $G = \mathbb{Z}$, $H = n\mathbb{Z}$.

Утверждение. Отношение сравнимости по модулю H является отношением эквивалентности.

Доказательство.

1. $g \equiv g \pmod{H}$, т. к. $g^{-1} g = e \in H$;
2. если $g_1 \equiv g_2 \pmod{H}$, т. е. $g_1^{-1} g_2 \in H$, то $g_2 \equiv g_1 \pmod{H}$, т. к.

   $$g_2^{-1} g_1 = (g_1^{-1} g_2)^{-1} \in H;$$

3. если $g_1 \equiv g_2 \pmod{H}$ и $g_2 \equiv g_1 \pmod{H}$, т. е. $g_1^{-1} g_2, g_2^{-1} g_3 \in H$, то $g_1 \equiv g_3 \pmod{H}$, т. к.

   $$g_1^{-1} g_3 = (g_1^{-1} g_2)(g_2^{-1} g_3) \in H.$$

■

Определение 32.2. Классы этой эквивалентности называются **левыми смежными классами** группы G по подгруппе H. Ясно, что смежный класс, содержащий элемент g, имеет вид

$$gH := \{gh \ : \ h \in H\}.$$

Примечание. Умножение не обязано быть коммутативным, поэтому мы получим, вообще говоря, другое отношение эквивалентности (и другие классы), взяв вместо условия $(*)$ аналогичное ему $g_2 g_1^{-1} \in H$. Классы этой эквивалентности называются **правыми смежными классами** группы G по подгруппе H. Они имеют вид

$$Hg := \{hg \ : \ h \in H\}.$$

Заметим, что инверсия $g \mapsto g^{-1}$ устанавливает биекцию между множествами левых и правых смежных классов. А именно,

$$(gH)^{-1} = Hg^{-1}.$$

Определение 32.3. Множество левых смежных классов группы G по подгруппе H обозначается через G/H. Число смежных классов (левых или правых), если оно конечно, называется **индексом** подгруппы H и обозначается через $|G : H|$.

Теорема 32.1 (Лагранж). Если G — конечная грула и H — любая её подгруппа, то

$$|G| = |G : H| \cdot |H|.$$

Доказательство. Все смежные классы gH содержат одно и то же число элементов, равное $|H|$. Поскольку они образуют разбиение группы G (как классы эквивалентности), порядок группы G равен произведению их числа на $|H|$. ■

Отношения эквивалентности и кольца вычетов[1]

Определение 32.4. ПУсть M — какое-либо множество. Всякое подмножество $\mathbb{R} \subseteq M \times M$ называется **отношением** на множестве M. Если $(a, b) \in R$, то говорят, что элементы a и b находятся в отношении R и пишут aRb.

Определение 32.5. Отношение называется **отношение эквивалентности**, если оно обладает следующими свойствами:

1. aRa (рефлексивность)
2. $aRb \Rightarrow bRa$ (симметричность)
3. $(aRb \wedge bRc) \Rightarrow aRc$ (транзитивность)

Отношение эквивалентности обычно записывается как $a \underset{R}{\sim} b$ или просто $a \sim b$.

Определение 32.6. Пусть R — отношение эквивалентности на множестве M. Для каждого M положим

$$R(a) := \{b \in M \ : \ a \underset{R}{\sim} b\}.$$

Лемма 32.1. $R(a) \cap R(b) \neq \varnothing \Rightarrow R(a) = R(b)$.

Доказательство. Пусть $c \in R(a) \cap R(b)$. Тогда $a \underset{R}{\sim} c$ и $c \underset{R}{\sim} b$, отсюда $a \underset{R}{\sim} b$, а значит, они лежат в одном классе, поэтому $R(a) = R(b)$. ■

Определение 32.7. Таким образом, подмножества $R(a)$ образуют разбиение множества M (т. е. покрывают его и попарно не пересекаются). Они называются **классами эквивалентности отношения** R.

[1] Я не знал, куда ещё вставить...

Определение 32.8. Множество, элементами которого являются классы эквивалентности отношения R, называется **фактормножеством** множества M и обозначается через M/R (если R — отношение эквивалентности, иногда используется обозначение $M/\sim$). Отображение

$$a \in M \mapsto R(a) \in M/R$$

называется **отображением факторизации**.

Определение 32.9. Отношение эквивалентности R на множестве M называется **согласованным** с операцией $*$, если

$$a \underset{R}{\sim} a', b \underset{R}{\sim} b' \Rightarrow a * b \underset{R}{\sim} b * b'.$$

В этом случае на фактормножестве M/R также можно определить операцию $*$

Определение 32.10. Класс эквивалентности сравнимых по модулю n целых чисел, содержащий a, будем называть **вычетом числа a по модулю n** и обозначать через $[a]_n$ (или просто $[a]$, если понятно, какое n имеется в виду). Фактормножество множества $\mathbb{Z}$ по отношению сравнимости $\bmod n$ обозначается через $\mathbb{Z}_n$. Мы можем писать, что

$$\mathbb{Z}_n = \{[0]_n, [1]_n, \ldots, [n-1]_n\}.$$

Лемма 32.2. Отношение сравнимости по модулю n согласовано с операциями сложения и умножения в $\mathbb{Z}$.

Доказательство. Пусть

$$a \equiv q' \pmod n, \quad b \equiv b' \pmod n.$$

Тогда

$$a + b \equiv a' + b' \pmod n$$

и, аналогично,

$$ab = \equiv a'b \equiv a'b' \pmod n.$$

■

Определение 32.11. Таким образом, можем определить в множестве $\mathbb{Z}_n$ операции сложения и умножения по формулам

$$[a]_n + [b]_n := [a+b]_n, \quad [a]_n[b]_n := [ab]$$

33 Следствия из теоремы Лагранжа

Теорема 33.1. Порядок любой подгруппы конечной группы делит порядок группы.

Доказательство. По теореме Лагранжа $|G| = |G:H| \cdot |H|$, отсюда $|H| \mid |G|$. ■

Теорема 33.2. Порядок любого элемента конечной группы делит порядок группы.

Доказательство. Это вытекает из предыдущей теоремы и того, что порядок элемента равен порядку порождаемой им циклической подгруппы. ■

Теорема 33.3. Всякая конечная группа простого порядка является циклической.

Доказательство. В силу теоремы 33.1 такая группа должа совпадать с циклической подгруппой, порождённой любым элементом, отличным от 1. ■

Теорема 33.4. Если $|G| = n$, то $g^n = e$ для любого $g \in G$.

Доказательство. Пусть $\operatorname{ord} g = m$. В силу следствия 2 имеем $m \mid n$. Значит, $g^n = e$. ■

Лемма 33.1. Если p — простое число, то мультипликативная группа $\mathbb{Z}_p^*$ поля $\mathbb{Z}_p$ есть (абелева) группа порядка $p-1$.

Доказательство. Сначала докажем, что $\mathbb{Z}_p^* = \{1, 2, \ldots, p-1\}$ (т. е., обратимы все, кроме 0). Пусть m — число, не делящееся на p. Тогда

$$\text{НОД}\,(m, p) = 1 \Rightarrow \exists u, v \in \mathbb{Z} : um + vp = 1.$$

Возьмём $\bmod p$ от обеих частей. Получаем $um = 1 \pmod p$, отсюда $[u][m] = 1$, и мы нашли обратный. А утверждение следует напрямую из леммы 34.1. ■

Теорема 33.5 (Малая теорема Ферма). Пусть p — простое число и пусть $a \in \mathbb{Z}$, $p \nmid a$. Тогда

$$a^{p-1} \equiv 1 \pmod p.$$

Доказательство. Если p — простое число, то мультипликативная группа $\mathbb{Z}_p^*$ поля $\mathbb{Z}_p$ есть (абелева) группа порядка $p-1$ (из леммы выше). Следовательно, $g^{p-1} = 1$ для любого элемента $g \in \mathbb{Z}_p^*$. Это означает, что

$$a^{p-1} \equiv 1 \pmod p.$$

■

Определение 33.1 (Функция Эйлера). Для любого n порядок группы $\mathbb{Z}_n^*$ обратимых элементов кольца $\mathbb{Z}_n$, равный количеству чисел в ряде $1, 2, \ldots, n$, взаимно простых с n, обозначается через $\varphi(n)$.

Теорема 33.6 (Эйлер). Пусть $a, n \in \mathbb{N}$ и $\text{НОД}\,(a, n) = 1$. Тогда

$$a^{\varphi(n)} \equiv 1 \pmod n.$$

Доказательство. Применяя теорему 33.4 к группе $\mathbb{Z}_n^*$, получаем требуемое. ■

34 Определение кольца и поля. Примеры. Простейшие следствия из аксиом. Обратимые элементы, делители нуля, нильпотенты. Взаимное расположение множеств обратимых элементов, делителей нуля и нильпотентов. Критерий того, что кольцо $\mathbb{Z}_n$ является полем

Определение 34.1. Кольцом называется множество $\mathcal{R}$ с операциями сложения и умножения, для которого выполнены следующие аксиомы:

1. $\mathcal{R}$ есть абелева группа по сложению;
2. $\forall a, b, c \in \mathcal{R}$ выполнено $a(b+c) = ab + ac$ и $(a+b)c = ac + bc$.

Теорема 34.1 (Простешие следствия из аксиом кольца).

1. $0 \cdot a = a \cdot 0 = 0$;
2. $a(-b) = (-a)b = -ab$;
3. $a(b-c) = ab - ac$, $(a-b)c = ac - bc$;
4. В кольце не может быть двух различных единиц (но может не быть ни одной);
5. Если кольцо содержит более одного элемента, то $1 \neq 0$;

\+ Свойства аддитивной абелевой группы.

Доказательство.

1. В самом деле, пусть $a \cdot 0 = b$. Тогда

$$b + b = a \cdot 0 + a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 = b \Rightarrow b = 0.$$

Аналогично доказывается, что $0 \cdot a = 0$.

2. В самом деле,
$$ab + a(-b) = a(b + (-b)) = a \cdot 0 = 0.$$
3. В самом деле,
$$a(b - c) + ac = a(b - c + c) = ab$$
и, аналогично, $(a - b)c = ac - bc$.
4. Аналогично доказательству того же факта для группы
5. Если $1 = 0$, то для любого элемента a имеем
$$a = a1 = a0 = 0,$$
т. е. кольцо состоит из одного нуля.

■

Определение 34.2. Кольцо $\mathcal{R}$ называется **ассоциативным**, если $\forall\, a, b, c \in \mathcal{R}$ выполнено $a(bc) = (ab)c$, **коммутативным**, если $\forall\, a, b \in \mathcal{R}$ выполнено $ab = ba$, **кольцом с единицей**, если $\exists e \in \mathcal{R} : \forall r \in \mathcal{R}$ выполнено $re = er = r$ и **телом**, если оно ассоциативно и $\forall r \in \mathcal{R}\ \exists r^{-1} : rr^{-1} = r^{-1}r = e$.

Примеры колец:

1. Числовые множества $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ являются коммутативными ассоциативными кольцами с единицей относительно обычных операций сложения и умножения.
2. Вычеты $\bmod n$: $\mathbb{Z}_n$ является коммутативным ассоциативным кольцом с единицей относительно обычных операций сложения и умножения.
3. Множество $2\mathbb{Z}$ чётных чисел является коммутативным ассоциативным кольцом без единицы.
4. Множество всей функций, определённых на заданном подмножестве числовой прямой, является коммутативным ассоциативным кольцом с единицей относительно обычных операций сложения и умножения функций
5. Множество векторов пространства с операцией сложения и векторного умножения является некоммутативным и неассоциативным кольцом. Однако в нём выполняются следующие тождества, которые в некотором смысле заменяют коммутативность и ассоциативность:
$$\begin{array}{cc} [a, b] + [b, a] = 0 & \textit{антикоммутативность} \\ [[a, b], c] + [[c, a], b] + [[b, c], a] = 0 & \textit{тождество Якоби} \end{array}$$

Лемма 34.1. Если A — любое ассоциативное кольцо с единицей, то множество его обратимых элементов A^* является группой по умножению.

Доказательство. Множество A^* замкнуто относительно взятия обратного (по условию). Тогда если $a, b \in A^*$, то и $a^{-1}, b^{-1} \in A^*$, а отсюда следует замкнутность A^* относительно умножения — пусто не замкнуто, тогда ab необратим, а это неправда — $(ab)^{-1} = b^{-1}a^{-1}$. Ассоцитативность выполняется, потому что кольцо A ассоциативно. А единица есть, потому что она есть в A и из замкнутости: $A^* \ni aa^{-1} = 1$. ■

Примечание. Как следствие, $\mathbb{Z}_n^*$ является мультипликативной группой.

Определение 34.3. Полем называется коммутативное тело.

Примечание. Кольцо, состоящее из одного нуля, не считается полем.

Примеры полей:

1. $\mathbb{Q}$, $\mathbb{R}$.
2. $\mathbb{Z}$ не является полем, в нём обратимы только ± 1.
3. $\{0, 1\}$.

Определение 34.4. Подмножество $\mathcal{L}$ кольца $\mathcal{R}$ называется **подкольцом**, если

1. $\mathcal{L}$ является подгруппой аддитивной группы кольца $\mathcal{R}$
2. $\mathcal{L}$ замкнуто относительно умножения

Очевидно, что всякое подкольцо само является кольцом относительно тех же операций. При этом оно наследует такие свойства, как коммутативность и ассоциативность.

Определение 34.5. Подмножество $\mathcal{L}$ поля $\mathbb{F}$ называется **подполем**, если

1. $\mathcal{L}$ является подкольцом кольца $\mathbb{F}$.
2. $a \in L, a \neq 0 \Rightarrow a^{-1} \in L$
3. $1 \in L$

Очевидно, что всякое подполе является полем относительно тех же операций.

Утверждение (Задача из листочка кружка в Хамовниках). Любое подполе поля $\mathbb{R}$ содержит $\mathbb{Q}$.

Доказательство. Сначала заметим, что в любом подполе $\mathbb{R}$ есть 0 и 1 (потому что это поле), а значит, там есть и все целые числа. Действительно,

$$n \in \mathbb{Z} \Rightarrow n = \underbrace{1 + 1 + \ldots + 1}_{n \text{ раз}}.$$

А значит, есть и все рациональные:

$$q \in \mathbb{Q} \Rightarrow q = nm^{-1}, \quad n \in \mathbb{Z}, m \in \mathbb{Z} \setminus \{0\}.$$

■

Утверждение (Задача из Винберга). Поле $\mathbb{Q}$ не имеет нетривиальных отличных от него самого подполей.

Доказательство. Решение полностью аналогично решению предыдущей задачи — в любом подполе $\mathbb{Q}$ лежат все целые числа, а значит, и все рациональные, а значит, оно совпадает с $\mathbb{Q}$. ■

Определение 34.6. Элемент $r \in \mathcal{K}$ ($\mathcal{K}$ содержит единицу e) называется **обратимым**, если существует $r^{-1} \in \mathcal{K}$ такой, что $rr^{-1} = r^{-1}r = e$.

Определение 34.7. Если a и b — ненулевые элементы $\mathcal{K}$ и $ab = 0$, то a называется **левым делителем нуля**, а b — **правым делителем нуля**.

Примечание. Могут попросить привести пример коммутативного ассоциативного кольца с делителями нуля. Подойдёт кольцо функций на подмножестве X числовой прямой, содержащем больше одной точки. В самом деле, разобьём X на два непустых подмножества X_1 и X_2 и положим при $i = 1, 2$

$$f_i(x) = \begin{cases} 1 & \text{при } x \in X_i, \\ 0 & \text{при } x \notin X_i. \end{cases}$$

Тогда $f_1, f_2 \neq 0$, но $f_1 f_2 = 0$.

Лемма 34.2. В кольце делитель нуля не может быть обратим.

Доказательство. Пусть $ab = 0$ и a — обратимый элемент. Тогда

$$0 = a^{-1}0 = a^{-1}ab = b.$$

■

Примечание. А в поле все обратимы, поэтому в поле нет делителей нуля.

Определение 34.8. Элемент $x \in \mathcal{K}$, $x \neq 0$ называется **нильпотентом**, если $\exists n \in \mathbb{N}$ такое, что $x^n = 0$.

Утверждение (Задача Антона Александровича). Если матрица $A \in \underset{n \times n}{\mathrm{Mat}}$ нильпотента, то $A^n = 0$.

Доказательство. Пусть $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ — линейное преобразование, соответствующее матрице A. Тогда $\varphi(\mathbb{R}^n)$ — подпространство в $\mathbb{R}^n$ в силу линейности φ. Заметим, что образ базиса $\{e_1, \ldots, e_n\}$ является полной системой в $\mathbb{R}^n$ опять же в силу линейности φ. При этом,

$$(\varphi(e_1), \ldots, \varphi(e_n)) = (e_1, \ldots, e_n) \cdot A,$$

т. к. A — матрица линейного преобразования φ. А из верхней оценки на ранг произведения следует, что $\operatorname{rk}(\varphi(e_1), \ldots, \varphi(e_n)) \leqslant \operatorname{rk} A$ (а если вспомнить нижнюю оценку, то легко понять, что на самом деле всегда достигается равенство, но нам для решения задачи это не нужно). Поэтому образ базиса — полная линейно зависимая система. А значит, из неё можно выделить базис, в нём будет точно меньше n векторов, а значит, $\dim \varphi(\mathbb{R}^n) < \dim \mathbb{R}^n = n$. То есть, каждый раз размерность уменьшается хотя бы на один. Отсюда,

$$\varphi^n(\mathbb{R}^n) = \{\mathbf{0}\}.$$

А значит, $\forall v \in \mathbb{R}^n$ выполнено $A^n \cdot v = 0$. Подставив вместо v векторы из стандартного базиса в $\mathbb{R}^n$, легко убедиться, что все a_{ij} нулевые, т. е. $A = 0$. ■

Лемма 34.3. Нильпотент является делителем нуля.

Доказательство. Пусть x нильпотент. Тогда возьмём n наименьшим натуральным числом со свойством $x^n = 0$. Тогда

$$x \cdot x^{n-1} = 0, x^{n-1} \neq 0, x \neq 0.$$

■

Теорема 34.2. Кольцо $\mathbb{Z}_n$ является полем тогда и только тогда, когда n — простое число.

Теорема 34.3. $\Rightarrow$. Пусть n составное, т. е. $n = k \cdot \ell$, где $1 < k, \ell < n$. Тогда $[k]_n, [\ell]_n \neq 0$, но

$$[k]_n[\ell]_n = [k\ell]_n = [n]_n = 0.$$

Таким образом, в кольце $\mathbb{Z}_n$ имеются делители нуля и, значит, оно не является полем.

$\Leftarrow$. Уже доказывали в лемме 33.1.

Примечание. То, что написано в этом вопросе далее, не входит в программу экзамена и взято мной из лекций Е. Ю. Смирнова по алгебре в ВШЭ и А. Д. Елагина по алгебраической геометрии в НМУ.

Определение 34.9 (Прямое произведение колец). Пусть A и B — кольца. На прямом произведении множеств $A \times B$ можно ввести операции сложения и умножения:

$$(a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2), \quad (a_1, b_1) \cdot (a_2, b_2) = (a_1 a_2, b_1 b_2).$$

Очевидно, что множество $A \times B$ с введёнными таким образом операциями является кольцом.

Примечание. Нулём такого кольца является пара $(0, 0)$, единицей — пара $(1, 1)$. Прямое произвдеение колец всегда имеет делители нуля: $(a, 0) \cdot (0, b) = (0, 0)$.

Теорема 34.4 (Китайская теорема об остатках в слабой форме). Если числа $m_1, \ldots, m_k$ попарно взаимно просты, то отображение

$$[a] \in \mathbb{Z}_{m_1 \ldots m_k} \mapsto ([a], \ldots, [a]) \in \mathbb{Z}_{m_1} \times \ldots \times \mathbb{Z}_{m_k}$$

является изоморфизмом.

Доказательство. Докажем индукцией по количеству взаимно простых сомножителей k.

База индукции ($k = 2$). Рассмотрим прямое произведение колец $\mathbb{Z}_m \times \mathbb{Z}_n$ при взаимно простых n и m. Единицей является пара $([1]_m, [1]_n)$. Обозначим её для удобства за $\mathbf{1}$, а нулём — элемент $([0]_m, [0]_n)$, его обозначим за $\mathbf{0}$. Легко видеть, что в последовательности сумм

$$\begin{array}{l} \mathbf{0} \\ \mathbf{1} \\ \mathbf{1} + \mathbf{1} \\ \mathbf{1} + \mathbf{1} + \mathbf{1} \\ \ldots \\ \mathbf{1} + \mathbf{1} + \ldots + \mathbf{1} \\ \ldots \end{array} \qquad (*)$$

остатки в первой компоненте будут повторяться с периодом m, а во второй — с периодом n, так что первое повторение получится на mn-ом шаге (в силу взаимной простоты m и n). Следовательно, в $(*)$ мы перечислили ровно mn различных элементов, т. е. все элементы нашего прямого произведения. Но сложение умножение элементов вида $(*)$ определено однозначно, и, конечно, это в точности соответствует определению операций сложения и умножения в $\mathbb{Z}_{mn}$. Это значит, что отображение $Z_{mn} \to \mathbb{Z}_m \times \mathbb{Z}_n$, сопоставляющее остатку $[a] \in \mathbb{Z}_{mn}$ сумму

$$\underbrace{([1]_m, [1]_n) + \ldots + ([1]_m, [1]_n)}_{a \text{ раз}},$$

является изоморфизмом.

Шаг индукции. Пусть утверждение верно для всех $k < K + 1$. Тогда отображение

$$[a] \in \mathbb{Z}_{m_1 \ldots m_K m_{K+1}} \mapsto ([a], \ldots, [a], [a]) \in \mathbb{Z}_{m_1} \times \ldots \times \mathbb{Z}_{m_K} \times \mathbb{Z}_{m_{K+1}}$$

является изоморфизмом. Действительно, это утверждение для взаимно простых множителей $m = m_1 \ldots m_K$ и $n = m_{K+1}$ является базой индукции и было доказано. ■

35 Характеристика поля. Какие значения может принимать характеристика? Возведение суммы в степень, равную характеристике. Малая теорема Ферма

Определение 35.1. Характеристика $\operatorname{char} \mathbb{F}$ поля $\mathbb{F}$ равна наимньшему натуральному k такому, что сумма k единиц равна нулю, если такое натуральное k существует. А иначе говорят, что $\operatorname{char} \mathbb{F} = 0$.

Теорема 35.1. Характеристика поля либо равна нулю, либо является простым числом.

Доказательство. Допустим, что $\operatorname{char} \mathbb{F} = mn$. Тогда

$$\underbrace{1 + 1 + \ldots + 1}_{mn \text{ раз}} = \underbrace{1 + 1 + \ldots + 1}_{m \text{ раз}} + \underbrace{1 + 1 + \ldots + 1}_{n \text{ раз}}$$

Так как в поле нет делителей нуля, один из множителей равен 0. ■

Лемма 35.1. Пусть $\mathbb{F}$ — поле характеристики p. Если сложить элемент $a \in \mathbb{F}$ с собой pk раз, то получится 0.

Доказательство.

$$\underbrace{a + a + \ldots + a}_{pk \text{ раз}} = \underbrace{1 + 1 + \ldots + 1}_{p \text{ раз}} ka = 0 \cdot ka = 0.$$

■

Теорема 35.2. Пусть $\mathbb{F}$ — поле характеристики p. Тогда для $a, b \in \mathbb{F}$ выполнено

$$(a_1 + a_2 + \ldots + a_k)^p = a_1^p + a_2^p + \ldots + a_k^p.$$

Доказательство. Докажем индукцией по количеству слагаемых k.

База индукции $(k = 2)$. По формуле бинома Ньютона

$$(a + b)^p = \sum_{k=0}^{p} \binom{p}{k} a^k b^{p-k}.$$

При этом $p \mid \binom{p}{k} = \frac{p!}{k!(p-k!)}$ при $i = 1, 2, \ldots, p-1$. Таким образом, в поле $\mathbb{F}$ все слагаемые, кроме крайних, равны нулю.

Шаг индукции. Пусть утверждение верно для всех $k < K$. Тогда

$$(a_1 + \ldots + a_{K-1} + a_K)^p = ((a_1 + \ldots + a_{K-1}) + a_K)^p = (a_1 + \ldots + a_{K-1})^p + a_K^p = a_1^p + \ldots + a_{K-1}^p + a_K^p.$$

■

Теорема 35.3 (Малая теорема Ферма). Пусть p — простое число. Для любого целого числа n выполнено

$$n^p \equiv n \pmod p.$$

Доказательство. Утверждение теоремы равносильно тому, что в кольце $\mathbb{Z}_p$ выполнено $[n]^p = [n]$. Это следует из цепочки равенств

$$[n]^p = ([1] + \ldots + [1])^p = [1]^p + \ldots + [1]^p = [n].$$

■

36 Гомоморфизм и изоморфизм алгебраических структур. Комплексные числа. Доказательство того, что комплексные числа образуют поле

Определение 36.1. Алгебраическая структура — это множество X с несколькими операциями $X^n \to X$ (возможно, для различных n), удовлетворяющих некоторым аксиомам.

Определение 36.2. Пусть есть две алгебраичекие структуры с одинаковым количеством операций от одинакового количества переменных. **Гомоморфизмом** из одной в другую называется отображение множеств, переводящее операции в операции.

То есть, если имеем две алгебраические структуры A и B с операциями $\alpha_i : A^{n_i} \to A$ и $\beta_i : A^{n_i} \to B$ и $\varphi : A \to B$ — гомоморфизм, то

$$\varphi(\alpha_i(x_1, \ldots, x_{n_i})) = \beta_i(\varphi(x_1), \ldots, \varphi(x_{n_i})).$$

Изоморфные структуры одинаковы с алгебраической точки зрения, если какая-то алгебраическая аксиома выполнена для одной структуры, то она выполнена и для изоморфной ей. Изоморфные структуры обозначаются $A \cong B$.

Изоморфизм из множества на себя называется **автоморфизмом**.

Примеры гомоморфизмов:

1. $\varphi : (\mathbb{Z}, +) \to (\mathbb{Z}_n, +)$, $\varphi(k) = [k]_n$;
2. $(\mathrm{GL}_n(\mathbb{R}), \cdot) \to (\mathbb{R}^*, \cdot)$, $\varphi(M) = \det M$;
3. $\varphi : (\mathbb{Z}, +, \cdot) \to (\mathbb{Z}_n, +, \cdot)$, $\varphi(k) = [k]_n$;
4. $\varphi : (\mathbb{Z}, +, \cdot) \to (\mathbb{Q}, +, \cdot)$, $\varphi(k) = k$.

Определение 36.3. Комплексные числа — это множество $\mathbb{R}^2$ с операциями

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

Множество комплексных чисел обозначается $\mathbb{C}$.

Теорема 36.1. Комплексные числа образуют поле.

Чтобы доказать теорему, нам понадобится следующая лемма.

Лемма 36.1. Алгебраичекая структура комплексных чисел изоморфна алгебраической структуре $\mathbb{M}$ матриц вида $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, $a, b \in \mathbb{R}$ с операциями сложения и умножения.

Доказательство. Определим отображение

$$\varphi : (a, b) \in \mathbb{C} \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \mathbb{M}.$$

Очевидно, что φ — биекция. Кроме того,

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

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

Таким образом, φ — гомоморфизм. Однако, очевидно, что φ — биекция. То есть, φ — гомоморфизм. ■

Теперь докажем теорему:

Доказательство. Проверим непосредственно выполнение всех аксиом:

1. $\mathbb{C}$ коммутативно относительно умножения, это следует из формулы $(a,b)\cdot(c,d) = (ac-bd, ad+bc)$.
2. $\mathbb{C}$ является кольцом, т. к., очевидно, является абелевой группой относительно сложения и для него выполнена дистрибутивность, т. к. $\mathbb{C} \cong \mathbb{M}$, а для $\mathbb{M}$ она выполнена. Ассоциативность выполнена по той же причине. Далее, в этом кольце есть единица — $(1,0)$, и ко всему прочему, каждый ненулевой элемент обратим, т. к. $\mathbb{C} \cong \mathbb{M}$ и

   $$(a,b) \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \text{ причём } \det \begin{pmatrix} a & -b \\ b & a \end{pmatrix}^{-1} = a^2 + b^2.$$

   Таким образом, эта матрица вырожденна (или, что равносильно, элемент (a,b) обратим) тогда и только тогда, когда $(a,b) = (0,0)$.

■

Вложение вещественных чисел. Рассмотрим отображение

$$\psi : \mathbb{R} \to \big\{(x,0) \ : \ x \in \mathbb{R}\big\} \subset \mathbb{C}, \quad \psi(x) = (x,0).$$

Легко видеть, что ψ — изоморфизм. Таким образом, поле $(\mathbb{R}, +, \cdot)$ изоморфно подполю $\mathbb{C}$. В дальнейшем не будем различать эти поля.

Заметим, что $(a,0)\cdot(c,d) = (ac, ad)$. Получаем, что на $\mathbb{C}$ есть операции сложения и умножения на $\mathbb{R}$. С этими операциями $\mathbb{C}$ — векторное пространство над $\mathbb{R}$, изоморфное $\mathbb{R}^2$. Его базис — это $(1,0) =: 1$ и $(0,1) =: i$. Заметим, что $i^2 = -1$. Значит, комплексные числа можно записывать в виде $a + bi$.

Причём, операции записываются естественным образом:

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

Определение 36.4. Будем называть i **мнимой единицей**.

37 Модуль и аргумент комплексного числа. Сопряжение и его свойства. Вещественная и мнимая части комплексного числа. Алгебраическая и тригонометрическая форма записи комплексного числа, переход между ними. Деление чисел в алгебраической форме

Каждому комплексному числу $z = a + bi$ можно биективно сопоставить точку плоскости (a,b).

Определение 37.1. Модулем комплексного числа $z = a + bi$ называется неотрицательное вещественное число $|z| = \sqrt{a^2+b^2}$ (корень арифметический).

Определение 37.2. Аргументом ненулевого комплексного числа $z = a+bi \neq 0$ — это величина угла между положительным лучом оси абсцисс и радиус-вектором точки (a,b), причём этот угол откладывается против часовой стрелки. Обозначается аргумент $\arg z$.

Определение 37.3. Пусть $z = a + bi$. Тогда **комплексно сопряжённое** к z число — это $\overline{z} = a - bi$.

Теорема 37.1. Сопряжение является автоморфизмом поля $\mathbb{C}$.

Доказательство. Сначала докажем, что сопряжение является гомоморфизмом $\mathbb{C} \to \mathbb{C}$. Для этого нужно доказать, что оно сохраняет операции:

$$\overline{(a+bi)+(c+di)} = \overline{(a+c)+(b+d)i} = (a+c)-(b+d)i = \overline{a+bi} + \overline{c+di},$$

$$\overline{(a+bi)\cdot(c+di)} = \overline{(ac-bd)+(ad+bc)i} = (ac-bd)-(ad+bc)i = \overline{a-bi}\cdot\overline{c-di}.$$

Очевидно также, что сопряжение отображает $\mathbb{C}$ на $\mathbb{C}$ биективно. То есть, является автоморфизмом. ■

Определение 37.4. Вещественная часть числа $z = a + bi$ — это число $\operatorname{Re} z = a$. **Мнимая часть** числа z — это число $\operatorname{Im} z = b$.

Определение 37.5 (Алгебраическая запись комплексного числа). $z = \operatorname{Re} z + i \cdot \operatorname{Im} z$.

Определение 37.6 (Тригонометрическая запись комплексного числа). $z = |z|(\cos\arg z + i \cdot \sin\arg z)$.

При этом выполнено

$$\operatorname{Re} z = |z|\cos\arg z, \quad \operatorname{Im} z = |z|\sin\arg z.$$

Лемма 37.1. Тригонометрическая форма числа единственная.

Доказательство. Пусть

$$z = r(\cos\varphi + i\sin\varphi) = s(\cos\psi + i\sin\psi).$$

Тогда $|z| = \sqrt{r^2(\cos^2\varphi + \sin^2\varphi)} = r = \ldots = s$. Значит, модули равны. Приравнивая вещественные и мнимые части, получаем $\cos\varphi = \cos\psi$, $\sin\varphi = \sin\psi$, отсюда $\varphi = \psi$ с точностью до $2\pi k$. ■

Деление чисел в алгебраической форме. Домножаем на сопряжённое:

$$\frac{a+bi}{c+di} = \frac{(a+bi)(c-di)}{(c+di)(c-di)} = \frac{(a+bi)(c-di)}{c^2+d^2}.$$

38 Умножене и деление чисел в тригонометрической форме. Формула Муавра. Извлечение корней из комплексных чисел

Теорема 38.1. При умножении комплексных чисел их модули умножаются, а аргументы складываются.

Доказательство. Перемножим два числа в тригонометрической форме:

$$r(\cos\varphi + i\sin\varphi)\cdot s(\cos\psi + i\sin\psi) = rs((\cos\varphi\cos\psi - \sin\varphi\sin\psi) + i(\cos\varphi\sin\psi + \sin\varphi\cos\psi)) = \\ = rs(\cos(\varphi+\psi) + i\sin(\varphi+\psi)).$$

Как следствие, при делении модули делятся, а аргументы вычитаются. ■

Теорема 38.2 (Формула Муавра). Для $n \in \mathbb{Z}$ верна формула

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

Доказательство. Для натуральных n утверждение вытекает из кратного применения результата предыдущей теоремы, а для отрицательных n нужно применить результат для положительных и частный случай предыдущей теоремы:

$$(r(\cos\varphi + i\sin\varphi))^{-1} = r^{-1}(\cos(-\varphi) + i\sin(-\varphi)).$$

■

Определение 38.1. Пусть $n \in \mathbb{N}$, $z \in \mathbb{C}$. Комплексное число w является n-ым **корнем** из z, если $w^n = z$. Обозначается $w = \sqrt[n]{z}$.

Теорема 38.3. Число комплексных корней $\sqrt[n]{z}$ равно

$$\begin{cases} 1, & \text{если } z = 0, \\ n, & \text{если } z \neq 0. \end{cases}$$

Доказательство. Представим z и w в тригонометрическом виде.

$$z = r(\cos\varphi + i\sin\varphi), \quad w = s(\cos\psi + i\sin\psi).$$

Тогда

$$r(\cos\varphi + i\sin\varphi) = z = w^n = s^n(\cos n\psi + i\sin n\psi).$$

Мы доказывали, что тригонометрический вид единственный, значит, отсюда следует, что

$$\begin{cases} r = s^n, \\ \varphi + 2\pi k = n\psi, \ k \in \mathbb{Z}. \end{cases}$$

Таким образом, $s = \sqrt[n]{r}$ (корень арифметрический), $\psi = (\psi + 2\pi k)/n$, $k \in \mathbb{Z}$. Заметим, что при разных $k = 0, 1, \ldots, n-1$ получаются различные (не отлицающиеся на $2\pi k$) углы. Отсюда имеем явную формулу для n-го корня из z:

$$\sqrt[n]{z} = \sqrt[n]{z}\left(\cos\left(\frac{\varphi + 2\pi k}{n}\right) + i\sin\left(\frac{\varphi + 2\pi k}{n}\right)\right), \quad k \in \{0, 1, \ldots, n-1\}.$$

■

Геометрическое расположение корней. Как легко видеть из явной формулы, корни n-ые корни из z образуют вершины правильного n-угольника с центром в начале координат.

39 Целостное кольцо. Многочлены от одной переменной над целостным кольцом. Понятие степени многочлена и её свойства. Целостность кольца многочленов над целостным кольцом. Обратимые элементы в кольце многочленов над целостным кольцом. Разложение многочлена по степеням $x - x_0$. Теорема Безу

Определение 39.1. Коммутативное ассоциативное кольцо с единицей без делителей нуля называется **областью целостности** (**целостным кольцом**).

Определение 39.2. Пусть $\mathcal{R}$ — коммутативное ассциативное кольцо с единицей. **Многочлен** над $\mathcal{R}$ – это финитная (с конечным числом ненулевых элементов) последовательность $(a_0, a_1, a_2, \ldots)$, где $a_i \in \mathcal{R}$.

Определение 39.3 (Операции на многочленах).

$$(a_0, a_1, a_2, \ldots) + (b_0, b_1, b_2, \ldots) := (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots)$$

$$(a_0, a_1, a_2, \ldots) \cdot (b_0, b_1, b_2, \ldots) := (c_0, c_1, c_2, \ldots), \quad \text{где } c_k = \sum_{j=0}^{k} a_j b_{k-j}.$$

Определение 39.4. Будем обозначать множество многочленов над областью целостности $\mathcal{R}$ как $\mathcal{R}[x]$.

Примечание. Поле является областью целостности.

Определение 39.5. Алгеброй над полем $\mathcal{F}$ называется множество A с операциями сложения, умножения и умножения на элементы поля $\mathcal{F}$, для которого выполняются следующие аксиомы:

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

Лемма 39.1. $(\mathcal{R}[x], +, \cdot)$ — коммутативное ассоциативное кольцо с единицей.

Доказательство. Все аксиомы очевидны, кроме ассоциативности умножения. Пусть

$$((a_0, a_1, a_2, \ldots) \cdot (b_0, b_1, b_2, \ldots)) \cdot (c_0, c_1, c_2, \ldots) = (d_0, d_1, d_2, \ldots)$$

$$(a_0, a_1, a_2, \ldots) \cdot ((b_0, b_1, b_2, \ldots) \cdot (c_0, c_1, c_2, \ldots)) = (f_0, f_1, f_2, \ldots)$$

Обозначим $(a_0, a_1, a_2, \ldots) \cdot (b_0, b_1, b_2, \ldots) = (u_0, u_1, u_2, \ldots)$, $(b_0, b_1, b_2, \ldots) \cdot (c_0, c_1, c_2, \ldots) = (v_0, v_1, v_2, \ldots)$. Имеем

$$d_k = \sum_{j=0}^{k} u_j c_{k-j} = \sum_{j=0}^{k} \sum_{i=0}^{j} a_i b_{j-i} c_{k-j} = \sum_{p+q+r=k} a_p b_q c_r,$$

$$f_k = \sum_{k=0}^{k} a_i v_{k-i} = \sum_{i=0}^{k} \sum_{s=0}^{k-i} a_i b_s c_{k-i-s} = \sum_{p+q+r=k} a_p b_q c_r.$$

■

Заметим, что единицей кольца является элемент $(1, 0, 0, \ldots)$. При этом элементы $(r, 0, 0, \ldots)$ складываются и умножаются так же, как и элементы области целостности $\mathcal{R}$. Таким образом, отождествим $(r, 0, 0, \ldots) \mapsto r$ и получим вложение колец $\mathcal{R} \subset \mathcal{R}[x]$ (инъективный гомоморфизм). Обозначим $(0, 1, 0, \ldots) \in \mathcal{R}[x]$ через x.

Примечание. Кольцо многочленов над полем является алгеброй.

Лемма 39.2. x^n — это последовательность, в которой на n-ом месте стоит единица, а остальные элементы — нули.

Доказательство. Индукция по n.

Шаг индукции ($n = 1$). По обозначению.

База индукции.

$$x^n = x^{n-1} \cdot x = (0, 0, \ldots, 0, 1, 0, \ldots) \cdot (0, 1, 0, \ldots).$$

Из формулы умножения финитных последовательностей, получаем $c_n = 1$, а остальные 0. Действительно, ведь ненулевые только a_1 и b_{n-1}. ■

Таким образом, многочлен $f = (a + 0, a_1, \ldots, a_n, 0, 0, \ldots)$ может быть записан как

$$f(x) = a_0 + a_1 x + \ldots + a_n x^n.$$

Определение 39.6. Степень $\deg f$ многочлена $f \neq (0, 0, \ldots)$ равна максимальному n, такому что $a_n \neq 0$.

Теорема 39.1.

1. $\deg(f + g) \leqslant \max\{\deg f, \deg g\}$;
2. Если R — область целостности, то $\deg fg = \deg f + \deg g$.

Доказательство. Пусть $f = (a_0, a_1, \ldots, a_m, 0, 0, \ldots)$, $g = (b_0, b_1, \ldots, b_n, 0, 0, \ldots)$.

1. Тогда
$$f + g = (a_0 + b_0, a_1 + b_1, \ldots),$$
при $j > \max\{m, n\}$, то элемент с номером j будет нулевой.

2. Пусть $fg = (c_0, c_1, \ldots)$. Тогда $c_k = \sum\limits_{j=0}^{k} a_j b_{k-j}$. Если $k > m + n$, то $c_k = 0$. При этом $c_{m+n} = a_m b_n \neq 0$, т. к. $\mathcal{R}$ целостное. Значит, $\deg fg = m + n$.

■

Примечание. Если кольцо $\mathcal{R}$ не является целостным, то второй пункт теоремы не верен. Например, в кольце $\mathbb{Z}_4[x]$ выполнено $\deg(2x+1)=1$, но $(2x+1)^2=1$ и $\deg 1=0$.

Теорема 39.2. Кольцо многочленов над целостным кольцом целостное.

Доказательство. Если $fg=0$, то это противоречит $\deg fg=\deg f+\deg g$. ■

Теорема 39.3 (Безу). Пусть $\mathcal{R}$ — коммутативное ассоциативное кольцо с единицей, $c\in\mathcal{R}$ и $f\in\mathcal{R}[x]$. Тогда $f(x)=(x-c)q(x)+r$ ($r\in\mathcal{R}$), причём $q(x)$ — многочлен, а $r=f(c)$.

Доказательство. Пусть $\mathcal{R}$ — область целостности, $c\in\mathcal{R}$ и $f\in\mathcal{R}[x]$. Положим $\widetilde{x}=x-c$, тогда $x=\widetilde{x}+c$. Если подставить это в f и раскрыть кобки, получим многочлен $\widetilde{f}$, причём $\widetilde{f}(\widetilde{x})=f(x-c)$. Получаем разложение $f(x)$ по степеням $x-c$, т. е. выражения вида

$$f(x)=b_n(x-c)^n=b_{n-1}(x-c)^{n-1}+\ldots+b_0.$$

Причём, $f(c)=r$. ■

Деление с остатком на $x-c$ осуществляется по *схеме Горнера*. А именно, пусть

$$a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n=(x-c)(b_0x^{n-1}+b_1x^{n-2}+\ldots+b_{n-2}x+b_{n-1})+r.$$

Приравнивая коэффициенты при соответствующих степенях x, получаем цепочку равенств:

$$\begin{aligned}
&a_0=b_0,\\
&a_1=b_1-cb_0,\\
&a_2=b_2-cb_1,\\
&\vdots\\
&a_{n-1}=b_{n-1}-cb_{n-2},\\
&a_n=r-cb_{n-1}.
\end{aligned}$$

Отсюда находим следующие рекуррентные формулы для $b_0,b_1,\ldots,b_{n-1},r$:

$$\begin{aligned}
&b_0=a_0,\\
&b_1=a_1+cb_0,\\
&b_2=a_2+cb_1,\\
&\vdots\\
&b_{n-1}=a_{n-1}+cb_{n-2},\\
&r=a_n+cb_{n-1}.
\end{aligned}$$

Исходные данные и результаты вычислений удобно расположить в таблице:

	a_0	a_1	a_2	...	a_{n-1}	a_n
c	b_0	b_1	b_2	...	b_{n-1}	r

Чтобы разложить многочлен f по степеням $x-c$ можно последовательным делением многочлена f на $x-c$ с остатком. А именно, при первом делении получается остаток b_0 и неполное частное

$$f_1=b_1+b_2(x-c)+\ldots+b_n(x-c)^{n-1};$$

при делении f_1 на $x-c$ получается остаток b_1 и т. д. Записывать можно как последовательную схему Горнера, используя предыдущую строку в качестве входных данных для следующей.

40 Предел комплексных последовательностей и функций. Непрерывные функции комплексного аргумента. Непрерывная функция $f : \mathbb{C} \to \mathbb{R}$ достигает минимума на компакте. Лемма о возрастании модуля

Определение 40.1. Пусть $z_0 \in \mathbb{C}$. Тогда **ε-окрестность точки** z_0 — это

$$U_\varepsilon(z_0) = \{z \in \mathbb{C} \,:\, |z - z_0| < \varepsilon\}.$$

Определение 40.2. Пусть $z_1, z_2, \ldots, z_n, \ldots$ — последовательность комплексных чисел. Будем говорить, что она **имеет предел** $w \in \mathbb{C}$ при $n \to \infty$, если для любого $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$ существует $N \in \mathbb{N}$ такое, что для любого $n > N$ выполнено $z_n \in U_\varepsilon(w)$.

Лемма 40.1. Пусть $z_j = x_j + iy_j$ и $w = u + iv$. Тогда

$$\lim_{n\to\infty} z_n = w \Leftrightarrow \begin{cases} \lim\limits_{n\to\infty} x_n = u, \\ \lim\limits_{n\to\infty} y_n = v. \end{cases}$$

Доказательство.

$$\lim_{n\to\infty} z_n = w \Leftrightarrow \lim_{n\to\infty} |z_n - w| = 0 \Leftrightarrow \lim_{n\to\infty} \sqrt{(x_n-u)^2 + (y_n-v)^2} = 0 \Leftrightarrow \lim_{n\to\infty} \left((x_n-u)^2 + (y_n-v)^2\right) = 0 \Leftrightarrow$$
$$\Leftrightarrow \begin{cases} \lim\limits_{n\to\infty} x_n = u, \\ \lim\limits_{n\to\infty} y_n = v. \end{cases}$$

■

Теорема 40.1. Пусть $\lim\limits_{n\to\infty} z_n = w$ и $\lim\limits_{n\to\infty} z_n^* = w^*$. Тогда

$$\lim_{n\to\infty} (z_n + z_n^*) = w + w^*, \quad \lim_{n\to\infty} (z_n \cdot z_n^*) = w \cdot w^*.$$

Доказательство. По условию $x_n \to u$, $y_n \to v$, $x_n^* \to u^*$, $y_n^* \to v^*$. Тогда $z_n + z_n^* = (x_n + x_n^*) + i(y_n + y_n^*)$. Но $x_n + x_n^* \to u + u^*$, $y_n + y_n^* \to v + v^*$. Значит, $z_n + z_n^* \to w + w^*$. Для умножения аналогично. ■

Определение 40.3. Пусть $f : \mathbb{C} \to \mathbb{C}$ — функция. Тогда $\lim\limits_{z\to w} f(z) = c \in \mathbb{C}$, если для каждого $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$ найдётся такое $\delta \in \mathbb{R}$, $\delta > 0$ такое, что при $z \in U_\delta(w)$ выполнено $f(z) \in U_\varepsilon(c)$.

Для предела функций верны те же 2 утверждения (доказанных выше), что и для предела последовательностей (со схожим доказательством).

Определение 40.4. Функция $f : \mathbb{C} \to \mathbb{C}$ называется **непрерывной в точке** w, если $\lim\limits_{z\to w} f(z) = f(w)$.

Лемма 40.2. Сумма и произведение непрерывных функций — это непрерывная функция.

Как следствие, многочлен $f \in \mathbb{C}[z]$ задаёт непрерывную функцию $\mathbb{C} \to \mathbb{C}$.

Определение 40.5. Подмножество $L \subset \mathbb{C}$ называется **открытым**, если для любого $z \in L$ существует $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$ такое, что $U_\varepsilon(z) \subset L$. Подмножество $S \subset \mathbb{C}$ называется **замкнутым**, если $\mathbb{C} \setminus S$ открыто.

Лемма 40.3. Пусть $S \subset \mathbb{C}$ замкнуто. Тогда если $z_i \in S$ при всех i и существует предел $\lim\limits_{n\to\infty} z_n = w$, то $w \in S$.

Доказательство. Предположим $w \notin S$. Тогда найдётся $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$ такое, что $U_\varepsilon(w) \cap S = \varnothing$. Однако начиная с некоторого номера $z_n \in U_\varepsilon(w)$. Противоречие. ■

Определение 40.6. Подмножество $K \subset \mathbb{C}$ называется **компактом**, если K замкнуто и ограничено. То есть, существует $N \in \mathbb{R}$ такое, что $K \subset \{z : |z| < N\}$.

Лемма 40.4. Из любой последовательности в компакте K можно выбрать сходящуюся подпоследовательность.

Доказательство. Пусть есть последовательность $z_n = x_n + iy_n \in K$. Тогда последовательность x_n ограничена, а значит, можно найти такую подпоследовательность в z_n, что $\{x_n\}$ для неё сходится. Аналогично, последовательность y_n ограничена, а значит, мы можем перейти к последовательности, в которой $\{y_n\}$ сходится. Т. к. последовательности $\{x_n\}$ и $\{y_n\}$ для этой подпоследовательности имеют предел, то и сама последовательность имеет предел. В силу замкнутости K, предел лежит в K. ■

Теорема 40.2. Непрерывнаяфункция $f : \mathbb{C} \to \mathbb{R}$ достигает минимума на компакте.

Доказательство. Пусть $M = \inf\limits_{z \in K} f(z)$. Тогда существует последовательность $z_n \in K$ такая, что $\lim\limits_{n\to\infty} f(z_n) = M$. Выберем из этой последовательности сходящуюся подпоследовательность. Т. к. функция непрерывна, то её значение в предельной точке этой подпоследовательности равно M. ■

Лемма 40.5 (О возрастании модуля). Пусть $f(z) \in \mathbb{C}[z]$ — многочлен положительной степени. Тогда $\lim\limits_{|z|\to\infty} |f(z)| = \infty$. То есть, для каждого $C \in \mathbb{R}$ существует $D \in \mathbb{R}$ такое, что при $|z| > D$ выполнено $|f(z)| > C$.

Доказательство. Заметим, что $|z| \to \infty \Leftrightarrow z^{-1} \to 0$. Пусть

$$f(z) = a_0 + a_1 z + \ldots + a_n z^n = z^n \left(\frac{a_0}{z^n} + \frac{a_1}{z^{n-1}} + \ldots + \frac{a_{n-1}}{z} + a_n \right).$$

Тогда

$$|f(z)| = |z^n| \cdot \left| \frac{a_0}{z^n} + \frac{a_1}{z^{n-1}} + \ldots + \frac{a_{n-1}}{z} + a_n \right|.$$

Но при $|z| \to \infty$ выполнено $\dfrac{a_0}{z^n} + \dfrac{a_1}{z^{n-1}} + \ldots + \dfrac{a_{n-1}}{z} \to 0$. Значит, существует $P \in \mathbb{R}$ такое, что при $|z| > P$ выполнено

$$\left| \frac{a_0}{z^n} + \frac{a_1}{z^{n-1}} + \ldots + \frac{a_{n-1}}{z} \right| < \frac{a_n}{2}.$$

Для модулей комплексных чисел выполнено неравенство треугольника (модулю — длина вектора):

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

Отсюда

$$\left| \frac{a_0}{z^n} + \frac{a_1}{z^{n-1}} + \ldots + \frac{a_{n-1}}{z} + a_n \right| \geqslant |a_n| - \left| \frac{a_0}{z^n} + \frac{a_1}{z^{n-1}} + \ldots + \frac{a_{n-1}}{z} \right| > |a_n| - \frac{|a_n|}{2} = \frac{|a_n|}{2}.$$

Тогда $|f(z)| > |z^n| \cdot \frac{|a_n|}{2} > D^n \cdot \frac{|a_n|}{2}$. Если D таково, что $D^n \frac{|a_n|}{2} > C$, то $|f(z)| > C$. ■

41 Лемма Даламбера

Лемма 41.1 (Даламбер). Пусть $f \in \mathbb{C}[z]$ — многочлен положительной степени и $f(z_0) \neq 0$. Тогда сколь угодно близко к z_0 можно найти такое z, что $|f(z)| < |f(z_0)|$.

Доказательство. Разложим f по степеням $z - z_0$ и разделим на $f(z_0)$. Учитывая, что несколько первых коэффициентов разложения, следующих за свободным членом, могут оказаться равными нулю, запишем результат в виде

$$\frac{f(z)}{f(z_0)} = 1 + c_p(z - z_0)^p + c_{p+1}(z - z_0)^{p+1} + \ldots + c_n(z - z_0)^n \quad (c_p \neq 0).$$

Нам нужно доказать существование такого z, что

$$\left| \frac{f(z)}{f(z_0)} \right| < 1.$$

Идея доказательства состоит в том, что если выбирать z достаточно близким к z_0, выполнение этого неравенства будем зависеть только от суммы первых двух членов предыдущего разложения. Будем искать z в виде $z = z_0 + tz_1$, где $0 < t < 1$, а z_1 — комплексное число, удовлетворяющее услвоию $c_p z_1^p = -1$. Имеем тогда

$$\frac{f(z)}{f(z_0)} = 1 - t^p + t^{p+1}\varphi(t),$$

где φ — некоторый многочлен степени $n - p - 1$ (с комплексными коэффициентами). Если C — максимум модулей коэффициентов многочлена φ, то

$$|\varphi(t)| \leqslant A = (n-p)C$$

и, следовательно,

$$\left|\frac{f(z)}{f(z_0)}\right| \leqslant 1 - t^p + At^{p+1} = 1 - t^p(1 - At) < 1.$$

■

42 Основная теорема алгебры. Комплексные корни вещественных многочленов. Неприводимые многочлены над $\mathbb{C}$ и $\mathbb{R}$. Разложение комплексных и вещественных многочленов на неприводимые множители (существование)

Теорема 42.1. Всякий многочлен положительной степени над полем комплексных чисел имеет корень.

Доказательство. По лемме о возрастании модуля существует $D \in \mathbb{R}$ такое, что при $|z| > D$ выполнено $|f(z)| > |f(0)|$. Рассмотрим круг $K = \{|z| \leqslant 2D\}$. Это замкнутое и ограниченное множество, т. е. компакт. Значит, существует точка $w \in K$, в которой достигается минимум функции $|f(z)|$. Заметим, что w не лежит на границе K, т. к. на границе данная функция больше, чем в точке $0 \in K$. Значит, существует $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$ такое, что $U_\varepsilon(w) \subset K$. Если $f(w) = 0$, то корень найден. Допустим, что $f(w) \neq 0$. Тогда по лемме Даламбера существует $w' \in U_\varepsilon(w)$ такое, что $|f(w')| < |f(w)|$. Но $w' \in K$. Получаем противоречие с выбором w. ■

Лемма 42.1. Если $z \in \mathbb{C}$ — корень многочлена $f \in \mathbb{R}[x]$, то и $\overline{z} \in \mathbb{C}$ — тоже корень.

Доказательство. Сразу следует из того, что сопряжение является автоморфизмом на $\mathbb{C}$. ■

Утверждение. Любой многочлен с вещественными коэффициентами раскладывается в произведение линейных и квадратичных с отрицательным дискриминантом множителей.

Доказательство. Индукция по степени многочлена.

База индукции. $\deg f = 0$ и $\deg f = 1$ очевидна.

Шаг индукции. Пусть $f \in \mathbb{R}[x]$. Если f имеет вещественный корень c, то $f(x) = (x - c)g(x)$, при этом $\deg g < \deg f$ и к g можно применить предположение индукции.

Пусть теперь у $f(x)$ есть комплексный корень λ, тогда $\overline{\lambda}$ — тоже корень. Значит,

$$f(x) = (x - \lambda)(x + \lambda)h(x) = (x^2 - 2\mathrm{Re}\,\lambda + |\lambda|^2)h(x).$$

При этом дискриминант квадратного полученного трёхчлена отрицательный, и к h можно применить предположение индукции. ■

43 Кратные корни многочлена. Сумма кратностей не превышает степень многочлена. Формальное и функциональное равенство многочленов от одной переменной

Определение 43.1. Будем говорить, что **многочлен f имеет корень кратности k**, если он может быть представлен в виде $f(x) = (x - a)^k q(x)$ и не может быть представлен в виде $(x - a)^{k+1} r(x)$.

Утверждение. Любой многочлен $f \in \mathbb{C}[z]$ степени n раскладывается на линейные множители с коэффициентами из $\mathbb{C}$.

Доказательство. Следствие основой теоремы алгебры и теоремы Безу. ■

Утверждение. Сумма кратностей корней многочлена не превышает степень многочлена.

Доказательство. Из уже доказанного следует, что многочлен f (с коэффициентами из $\mathbb{R}$ или $\mathbb{C}$) степени n пишется как

$$f(x) = a(x - c_1)^{k_1} \dots (x - c_n)^{k_n}.$$

Отсюда сразу следует требуемое. ■

Формальное и функциональное равенство многочленов — не одно и то же. Например, многочлены $f(x) = x$ и $f(x) = x^2$ над полем $\mathbb{Z}_2$ задают одинаковые функции, но многочлены это, очевидно, разные (не совпадают задающие их финитные последовательности). Если $\mathcal{R}$ — конечная область целостности, то многочлен

$$f(x) = \prod_{r \in \mathcal{R}} (x - r)$$

задаёт тождественно нулевую функцию, хотя сам многочлен ненулевой. По малой теореме Ферма многочлены x^p и x задают одну и ту же функцию над полем $\mathbb{Z}_p$.

Теорема 43.1. Пусть $\mathcal{R}$ — бесконечная область целостности. Тогда из функционального равенства многочленов из $\mathcal{R}[x]$ следует их формальное равенство.

Доказательство. Пусть многочлены f и g определяют одну и ту же функцию. Тогда их разность $h = f - g$ определяет нулевую функцию, т. е. $h(c) = 0$ для всех $c \in \mathcal{R}$. Предположим, что $h \neq 0$, и пусть

$$h = a_0 + a_1 x + a_2 x^2 + \dots + a_{n-1} x^{n-1} \quad (a_{n-1} \neq 0).$$

возьмём различные $x_1, x_2, \dots, x_n \in \mathcal{R}$ (здесь используется бесконечность области целостности $\mathcal{R}$). Совокупность равенств

$$\begin{cases} a_0 + a_1 x_1 + a_2 x_1^2 + \dots + a_{n-1} x_1^{n-1} = 0, \\ a_0 + a_1 x_2 + a_2 x_2^2 + \dots + a_{n-1} x_2^{n-1} = 0, \\ \dots \\ a_0 + a_1 x_n + a_2 x_n^2 + \dots + a_{n-1} x_n^{n-1} = 0. \end{cases}$$

будем рассматривать как (квадратную) однородную СЛУ относительно $a_0, \dots, a_{n-1}$. Определитель матрицы коэффициентов этой системы есть определитель Вандермонда и потому отличен от нуля. Следовательно, система имеет только нулевое решение, что противоречит нашему предположению. ■

44 Деление многочленов от одной переменной над полем с остатком. Наибольший общий делитель. Алгоритм Евклида. Линейное выражение НОД. Доказательство того, что НОД делится на все общие делители

Теорема 44.1. Пусть $f, g \in \mathbb{F}[x]$, причём $g \neq 0$. Тогда существуют такие многочлены p и r, что $f = qg + r$ и $\deg r < \deg g$. Многочлены q и r определены этими условиями однозначно.

Доказательство. Докажем сначала существование, затем единственность указанного разложения. Если $\deg f < \deg g$, то можно взять $q = 0$, $r = f$. Если $\deg f \geqslant \deg g$, то q и r находятся процедурой «деления уголком». А именно, пусть

$$f = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n,$$
$$g = b_0 x^m + b_1 x^{m-1} + \dots + b_{m-1} x + b_n,$$

где $a_0, b_0 \neq 0$. Рассмотрим многочлен

$$f_1 = f - \frac{a_0}{b_0} x^{n-m} g.$$

Его степень меньше, чем степень многочлена f. Если $\deg f_1 < \deg g$, то мы можем взять

$$q = \frac{a_0}{b_0} x^{n-m}, \quad r = f_1.$$

В противном случае поступаем с многочленом f_1 так же, как с f.

Теперь единственность. Пусть

$$f = q_1 g + r_1 = q_2 g + r_2,$$

где $\deg r_1 < \deg g$ и $\deg r_2 < \deg g$. Тогда

$$r_1 - r_2 = (q_2 - q_1)g$$

и, если $q_1 \neq q_2$, то

$$\deg(r_1 - r_2) = \deg(q_2 - q_1) + \deg g \geqslant \deg g,$$

что, очевидно, неверно. Следовательно, $q_1 = q_2$ и $r_1 = r_2$. ■

Определение 44.1. Наибольшим общим делителем элементов a и b целостного кольца называется их общий делитель, делящийся на все их общие делители. Оно обозначается через НОД(a, b).

Определение 44.2. Целостное кольцо A, не являющееся полем, называется **евклидовым**, если существует функция

$$N : A \setminus \{0\} \to \mathbb{Z}_+$$

(называется **нормой**), удовлетворяющая следующим аксиомам:

1. $N(ab) \geqslant N(a)$, причём равенство достигается тогда и только тогда, когда элемент b обратим;
2. для любых $a, b \in A$, где $b \neq 0$, существуют такие $q, r \in A$, что $a = qb + r$ и либо $r = 0$, либо $N(r) < N(b)$;

Примечание. Условие 2 означает возможнось деления с остатком. Его единственности не требуется.

Вторая часть условия 1 на самом деле может быть выведена из остальных условий. В самом деле, пусть элемент b необратим. Тогда $ab \nmid a$. Разделим a на ab с остатком:

$$a = q(ab) + r.$$

Так как $r = a(1 - ab)$, то

$$N(a) \leqslant N(r) < N(ab).$$

Примеры евклидовых колец:

1. Целые числа: $\mathbb{Z}$ (в качестве нормы можно взять модуль).
2. Гауссовы целые числа: $\mathbb{Z}[i]$ (в качестве нормы можно опять же взять модуль).
3. Кольцо многочленов над полем: $\mathbb{F}[x]$ (в качестве нормы можно взять степень многочлена).

Теорема 44.2. В евклидовом кольце для любых элементов a и b существует наибольший общий делитель d, и он может быть представлен в виде $d = au + bv$, где u, v — какие-то элементы кольца.

Доказательство. Если $b = 0$, то $d = a = a \cdot 1 + b \cdot 0$. Если $b \mid a$, то $d = b = a \cdot 0 + b \cdot 1$. В противном случае разделим с остатком a на b, затем b на полученный остаток, затем первый остаток на второй остаток и т. д. Так как нормы остатков убывают, то в конце деление произойдёт без остатка. Получим цепочку равенств:

$$\begin{aligned}
&a = q_1 b + r_1,\\
&b = q_2 r_1 + r_2,\\
&r_1 = q_3 r_2 + r_2,\\
&\vdots\\
&r_{n-2} = q_n r_{n-1} + r_n,\\
&r_{n-1} = q_{n+1} r_n.
\end{aligned}$$

Докажем, что последний ненулевой остаток r_n и есть НОД(a, b). Двигаясь по выписанной цепочке равенств снизу вверх, получам последовательно

$$r_n \mid r_{n-1}, \quad r_n \mid r_{n-2}, \quad \ldots, \quad r_n \mid r_1, \quad r_n \mid b, \quad r_n \mid a.$$

Таким образом, r_n — общий делитель элементов a и b.

Двигаясь по той же цепочке равенств снизу вверх, получаем последовательно

$$\begin{aligned} r_1 &= au_1 + bv_1, \\ r_2 &= au_2 + bv_2, \\ &\vdots \\ r_n &= au_n = bv_n, \end{aligned}$$

где u_i, v_i $(i = 1, 2, \ldots, n)$ — какие-то элементы кольца. Таким образом, r_n можно представить в виде $au + bv$. Отсюда, в свою очередь, следует, что r_n делится на любой общий делитель a и b. ■

Определение 44.3. Необратимый нулевой элемент p целостного кольца называется **простым**, если он не может быть представлен в виде $p = ab$, где a и b — необратимые элементы. Простые элементы кольца $\mathbb{F}[x]$, где $\mathbb{F}$ — поле, называются **неприводимыми многочленами**.

Очевидно, что всякий многочлен первой степени неприводим. Из основной теоремы алгебры вытекает, что неприводимые многочлены над $\mathbb{C}$ — это только многочлены первой степени, а над $\mathbb{R}$ — это многочлены первой степени и многочлены второй степени с отрицательным дискриминантом.

Лемма 44.1. Если простой элемент p евклидова кольца A делит произведение $a_1 a_2 \ldots a_n$, то он делит хотя бы один из сомножителей $a_1, a_2, \ldots, a_n$.

Доказательство. Докажем это утверждение индукцией по n.

База индукции $(n = 2)$. Предположим, что $p \nmid a_1$. Тогда НОД$(p, a_1) = 1$ и, значит, существует такие $u, v \in A$, что $pu + a_1 v = 1$. Умножая это равенство на a_2, получаем

$$pua_2 + a_1 a_2 v = a_2,$$

откуда следует, что $p \mid a_2$.

Шаг индукции. Представим

$$a_1 a_2 \ldots a_n = a_1 \cdot (a_2 \ldots a_n).$$

Теперь, применяя базу, получаем $p \mid a_1$ или $p \mid a_2 \ldots a_n$. Если первое не выполняется, то применяем предположение индукции. ■

Определение 44.4. Два элемента a и b кольца $\mathcal{R}$ называются **ассоциированными**, если $a = bc$, где $c \in \mathcal{R}$ обратим.

Утверждение. Отношение ассоциированности является отношением эквивалентности.

Доказательство. Если $a = bc$, то $b = ac^{-1}$, поэтому отношение ассоциированности симметрично. Т. к. 1 — обратимый элемент $\mathcal{R}$, то отношение ассоциированности рефлексивно. И т. к. произведение двух обратимых элементов обратимо, данное отношение транзитивно. ■

Примечание. Получаем, что все элементы $\mathcal{R}$ распадаются на классы ассоциированности. В доказательстве следующей теоремы мы будем писать $p \sim q$, если элементы p и q ассоциированны.

Теорема 44.3. В евклидовом кольце всякий необратимый ненулевой элемент может быть разложен на простые множители, причём это разложение единственно с точностью до перестановки множителей и их ассоциированности.

Доказательство. Существование. Назовём необратимый ненулевой элемент $a \in A$ пацанским, если он может быть разложен на простые множители, остальные элементы будем называть чушпанскими. Предположим, что существуют чушпанские элементы. Выберем из них элемент с наименьшей нормой. Пусть это будет элемент a. Он не может быть простым. Следовательно, $a = bc$, гед b и c — необратимые элементы. Имеем $N(b) < N(a)$ и $N(c) < N(a)$ и, значит, b и c — пацанские элементы; но тогда, очевидно, и a — пацанский элемент, что противоречит нашему предположению. Таким образом, всякий необратимый ненулевой элемент кольца A может быть разложен на простые множители.

Единственость. Докажем индукцией по n, что если

$$a = p_1 p_2 \ldots p_n = q_1 q_2 \ldots q_m,$$

где p_i, q_j — простые элементы, то $m = n$ и, после подходящей перенумерации множителей, $p_i \sim q_i$ при $i = 1, 2, \ldots, n$.

База индукции ($n = 1$). $a = p_1$ — простой, поэтому и $m = 1$, причём $q_1 \sim p_1$ (из определения простого элемента).

Шаг индукции. При $n > 1$ имеем $p_1 \mid q_1 q_2 \ldots q_m$, а из предыдущей леммы это значит, что $p_1 \mid q_i$. Перенумеруем так, что $i = 1$. Тогда $p_1 \sim q_1$. Сокращая на p_1 и применяя предположение индукции, получаем требуемое. ■

Примечание. Это утверждение над евклидовым кольцом $\mathbb{Z}$ называется основной теоремой арифметики.

45 Факториальное кольцо. Факториальность кольца многочленов над полем

Определение 45.1. Целостное кольцо $\mathcal{R}$ называется **факториальным**, если любой необратимый элемент $r \in \mathcal{R}$ однозначто с точностью до перестановки множителей и ассоциированности множителей разлагается в произведение неприводимых элементов.

Примечание. С учётом этого определения, можно переформулировать теорему 44.3 как «любое евклидово кольцо факториально».

Кольцо многочленов над полем является евклидовым и, как следствие теоремы 44.3, факториальным.

46 Формальная производная многоченов. Связь значений кратных производных в даной точке с кратностью корня. Кратность корней $\text{НОД}(f, f')$. Избавление от кратных корней

Определение 46.1. Формальная производная — это отображение $\mathcal{D}$ алгебры $\mathbb{F}[x]$ в себя, для которого выполняются следующие аксиомы:

1. оно линейно;
2. $\mathcal{D}(fg) = (\mathcal{D}f)g + f(\mathcal{D}g)$ («правило Лейбница»);
3. $\mathcal{D}x = 1$.

Теорема 46.1. Отображение $\mathcal{D} : \mathbb{F}[x] \to \mathbb{F}[x]$, определённое нами выше, существует и единственно.

Доказательство. Существование. Построим линейное отображение $\mathcal{D} : \mathbb{F}[x] \to \mathbb{F}[x]$, задав его на базисных векторах формулами

$$\mathcal{D}1 = 0, \quad \mathcal{D}x^n = nx^{n-1}, \quad (n = 1, 2, \ldots),$$

и проверим, что оно обладает свойством 2 (остальные очевидно выполняются). В силу линейности достаточно проверить ео только на базисных векторах:

$$\mathcal{D}(x^m x^n) = \mathcal{D}x^{m+n} = (m+n)x^{m+n-1} = mx^{m+n-1} + mx^{m+n-1} = (\mathcal{D}x^m)x^n + (\mathcal{D}x^n)x^m.$$

Единственность. Заметим, что

$$\mathcal{D}1 = \mathcal{D}(1 \cdot 1) = \mathcal{D}1 + \mathcal{D}1 \Rightarrow \mathcal{D}1 = 0.$$

Докажем по индукции, что $\mathcal{D}(x^n) = n\mathcal{D}x^{n-1}$. При $n = 1$ это верно по аксиоме 3, а переход от $n-1$ к n делается выкладкой

$$\mathcal{D}x^n = \mathcal{D}(x^{n-1}x) = (\mathcal{D}x^{n-1})x + x^{n-1}(\mathcal{D}x) = (n-1)x^{n-2} \cdot x + x^{n-1} = nx^{n-1}.$$

■

Определение 46.2. Многочлен $\mathcal{D}f$ называется **производной** многочлена f и обозначается, как обычно, через f'.

Утверждение. Если $\operatorname{char}\mathbb{F} = 0$, то коэффициенты разложения многочлена $f \in \mathbb{F}[x]$ по степеням $x - c$

$$f = b_0 + b_1(x-c) + b_2(x-c)^2 + \ldots = b_n(x-c)^n \qquad (*)$$

могут быть найдены по формулам

$$b_k = \frac{f^{(k)}(c)}{k!}.$$

Доказательство. Продифференцируем равенство $(*)$ k раз и подставим $x = c$. ■

Определение 46.3. Из предыдущего утверждения можно сделать вывод, что

$$f = f(c) + \frac{f'(c)}{1!}(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \ldots + \frac{f^{(n)}(c)}{n!}(x-c)^n.$$

Эта формула называется **формулой Тейлора** для многочленов.

Теорема 46.2. При условии $\operatorname{char}\mathbb{F} = 0$ кратность корня c многочлена $f \in \mathbb{F}[x]$ равна наименьшему порядку производной многочлена f, не обращающейся в нуль в точке c.

Доказательство. Заметим, что кратность корня c равна номеру первого отличного от нуля коэффициента разложения $(*)$. Из формулы Тейлора сразу следует требуемое. ■

Далее идут следствия этой теоремы:

Утверждение. Пусть $c \in \mathbb{F}$ — корень кратности $k > 0$ многочлена $f \in \mathbb{F}[x]$. Тогда c — корень кратности $k-1$ многочлена f'.

Утверждение. Многочлен $\text{НОД}(f, f')$ своими корнями имеет только кратные корни $f(x)$. Причём, кратности всех корней в многочлене $\text{НОД}(f, f')$ на 1 меньше, чем в f.

Утверждение (Избавление от кратных корней). Многочлен

$$\frac{f(x)}{\text{НОД}(f, f')}$$

своими корнями имеет все корни $f(x)$ с кратностями 1.

47 Многочлены от нескольких переменных. Порядки на мономах. Лексикографический порядок и его свойства. Старший член и моном. Лемма о старшем члене

Определение 47.1. Кольцо многочленов от нескольких переменных $x_1, \ldots, x_n$ с коэффициентами из области целостности $\mathcal{R}$ определим рекурсивно:

$$\mathcal{R}[x1, \ldots, x_n] = \mathcal{R}[x_1, \ldots, x_{n-1}][x_n].$$

Исходя из теоремы 39.2, если $\mathcal{R}$ — область целостности, то и $\mathcal{R}[x_1, \ldots, x_n]$ — тоже область целостности.

Определение 47.2. Определим **лексикографический порядок на мономах**. Пусть

$$m_\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}, \quad m_\beta = x_1^{\beta_1} \ldots x_n^{\beta_n}$$

и $\alpha_1 = \beta_1, \ldots, \alpha_{k-1} = \beta_{k-1}\alpha_k \neq \beta_k$. Тогда если $\alpha_k > \beta_k$, то $m_\alpha \succ m_\beta$, иначе $m_\alpha \prec m_\beta$.

Теорема 47.1. Отношение лексикографического порядка обладает следующими свойствами:

1. для любых двух несовпадающих мономов либо $u \succ v$, либо $u \prec v$;
2. не может быть $u \succ v$ и $u \prec v$;

3. из того, что $u \succ v$ и $v \succ w$ следует $u \succ w$;
4. из того, что $u \succ v$ следует $uw \succ vw$ для любого монома w;
5. не существует бесконечных убывающих цепочек мономов $u_1 \succ u_2 \succ u_3 \succ \dots$

Доказательство. Свойства 1 и 2 очевидно следуют из определения. Докажем 3. Пусть первая переменная, которая не входит во все одночлены u, v, w с одним и тем же показателем, входит в них с показателями k, ℓ, m соответственно. Тогда

$$k \geqslant \ell \geqslant m,$$

причём хотя бы в одном из двух случаев имеет место строго неравенство. Следовательно, $k > m$ и $u \succ w$.

Докажем 4. При умножении на w к показателям, с которыми каждая из переменных входит в u и v, добавляется одно и то же число, и знак неравенства между этими показателями не меняется, а только эти неравенства и имеют значение при сравнении одночленов.

Докажем 5 индукцией по n.

База индукции ($n = 1$). Тогда $m_1 = x^k$ для некоторого k и в убывающей последовательности $m_1 \succ m_2 \succ m_3 \succ \dots$ встретятся только попарно различные мономы $m_t = x^t$, причём $t \leqslant k$, поэтому такая последовательность не может быть бесконечной.

Шаг индукции. Пусть свойство 5 доказано для всех $n < m$. Докажем для $n = m$. Допустим, что существует бесконечная убывающая последовательность $m_1 \succ m_2 \succ \dots$. При этом $m_1 = x_1^{\alpha_1} \dots x_m^{\alpha_m}$. При переходе от m_i к m_{i+1} показатель степени x_1 либо не меняется, либо убывает. Следовательно, убывать он может лишь конечное число раз. Значит, найдётся такое натуральное N, что начиная с m_N показатель степени x_1 не изменяется. Рассмотрим последовательность без x_1, начиная с N-го члена. Её длина $m-1 < m$ и она бесконечно убывает. Противоречие. ■

Определение 47.3. Среди ненулевых членов любого ненулевого многочлена $f \in \mathcal{R}[x_1, \dots, x_n]$ найдётся единственный, который лексикографически старше всех остальных. Он называется **старшим членом** многочлена f. **Старшим мономом** для удобства будем называть старший член с коэффициентом.

Лемма 47.1 (О старшем члене). Старший член произведения ненулевых многочленов равен произведению их старших членов.

Доказательство. Достаточно доказать это утверждение для двух многочленов (дальше индукцией). Пусть f_1, f_2 — ненулевые многочлены, u_1, u_2 — их старшие члены и v_1, v_2 — какие-то их члены. Если $v_1 \neq u_1$ и $v_2 \neq u_2$, то в силу теоремы 47.1

$$u_1 u_2 \succ v_1 v_2.$$

■

48 Симметрические многочлены. Основная теорема о симметрических многочленах

Определение 48.1. Многочлен $f \in \mathbb{F}[x_1, \dots, x_n]$ называется **симметрическим**, если он не изменяется ни при каких перестановках переменных.

Определение 48.2. Следующие симметрические многочлены называются **элементарными**:

$$\begin{aligned}
\sigma_1 &= x_1 + x_2 + \dots + x_n, \\
\sigma_2 &= x_1x_2 + x_2x_3 + \dots + x_{n-1}x_n, \\
&\vdots \\
\sigma_k &= \sum_{i_1 < \dots < i_k} x_{i_1} \dots x_{i_k}, \\
&\vdots \\
\sigma_n &= x_1x_2 \dots x_n.
\end{aligned}$$

Нетривиальный пример симметрического многочлена — это $V^2(x_1, \dots, x_n)$. Действительно, при перестановках $x_1, \dots, x_n$ определитель Вандермонда может лишь поменять знак, а значит, его квадрат при любых

перестановках переменных не меняется. То есть, они совпадают как функции, а из этого следует и совпадение многочленов (см. теорему 43.1).

Очевидно, что сумма и произведение симметрических многочленов, а также произведение симметрического многочлена на число являются симметрическими многочленами. Иными словами, симметрические многочлены образуют подалгебру в алгебре всех многочленов.

Теорема 48.1. Всякий симметрический многочлен единственным образом представляется в виде многочлена от элементарных симметрических многочленов.

Перед тем, как её доказывать, докажем два вспомогательных утверждения.

Лемма 48.1. Пусть $u = x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}$ — старший член симметрического многочлена f.

Доказательство. Предположим, что $k_i < k_{i+1}$ для некоторого i. В силу симметричности, многочлен f помимо члена u должен содержать и член

$$u^* = x_1^{k_1} x_2^{k_2} \dots x_i^{k_{i+1}} x_{i+1}^{k_i} \dots x_n^{k_n}, \qquad (\star)$$

ведь один получается из другого транспозицией $[i, i+1]$. Однако $u^* \succ u$, поэтому u не может быть старшим членом. Противоречие. ■

Лемма 48.2. Для любого одночлена $u = x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}$, показатели которого удовлетворяют неравенствам $(\star)$, существуют такие неотрицательные числа $\ell_1, \ell_2, \dots, \ell_n$, что старший член многочлена $\sigma_1^{\ell_1} \sigma_2^{\ell_2} \dots \sigma_n^{\ell_n}$ совпадает с u. Числа $\ell_1, \ell_2, \dots, \ell_n$ определены этим условием однозначно.

Доказательство. Старший член многочлена σ_k равен $x_1 x_2 \dots x_k$. В силу леммы о старшем члене старший член многочлена $\sigma_1^{\ell_1} \sigma_2^{\ell_2} \dots \sigma_n^{\ell_n}$ равен

$$x_1^{\ell_1} (x_1 x_2)^{\ell_2} \dots (x_1 x_2 \dots x_n)^{\ell_n} = x_1^{\ell_1 + \ell_2 + \dots + \ell_n} x_2^{\ell_2 + \dots + \ell_n} \dots x_n^{\ell_n}.$$

Приравнивая его одночлену u, получаем систему линейных уравнений

$$\begin{cases} \ell_1 + \ell_2 + \dots + \ell_n = k_1, \\ \ell_2 + \dots + \ell_n = k_2, \\ \dots \\ \ell_n = k_n, \end{cases}$$

которая, очевидно, имеет единственное решение

$$\ell_i = k_i - k_{i+1} \ (i = 1, 2, \dots, n-1), \quad \ell_n = k_n.$$

Из условия леммы следует, что определённые таким образом числа $\ell_1, \ell_2, \dots, \ell_n$ неотрицательны. ■

Теперь докажем теорему 48.1.

Доказательство. Пусть $f \in \mathbb{F}[x_1, \dots, x_n]$ — симметрический многочлен. Нам нужно найти такой многочлен $F \in \mathbb{F}[X_1, \dots, X_n]$, что

$$F(\sigma_1, \sigma_2, \dots, \sigma_n) = f.$$

Если $f = 0$, то можно взять $F = 0$. В противном случае пусть $u_1 = a x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}$ — старший член многочлена f. По лемме 48.1 выполняются неравенства $(\star)$, по лемме 48.2 существует такой одночлен $F_1 \in \mathbb{F}[X_1, \dots, X_n]$, что старший член многочлена $F_1(\sigma_1, \sigma_2, \sigma_n)$ был равен u_1. Рассмотрим симметрический многочлен

$$f_1 = f - F_1(\sigma_1, \sigma_2, \dots, \sigma_n).$$

Если $f_1 = 0$, то можно взять $F = F_1$. В противном случае повторяем процесс, строя многочлены $f_2, f_3, \dots$ Процесс обязательно конечный, т. к. степени многочленов f_k убывают. А выражение через элементарные симметрические многочлены будет выглядет как

$$F = F_1 + F_2 + \dots + F_N.$$

Теперь докажем, что многочлен F определён однозначно. Предположим, что F и G — такие многочлены, что

$$F(\sigma_1, \sigma_2, \dots, \sigma_n) = G(\sigma_1, \sigma_2, \dots, \sigma_n).$$

Рассмотрим разность $H = F - G$. Тогда

$$H(\sigma_1, \sigma_2, \ldots, \sigma_n) = 0.$$

Нам нужно доказать, что $H = 0$. Предположим, что это не так, и пусть $H_1, H_2, \ldots, H_s$ — все ненулевые члены многочлена H. Обозначим через w_i $(i = 1, 2, \ldots, s)$ старший член многочлена

$$H_i(\sigma_1, \sigma_2, \ldots, \sigma_n) \in \mathbb{F}[x_1, x_2, \ldots, x_n].$$

В силу леммы 2 среди одночленов $w_1, w_2, \ldots, w_s$ нет пропорциональных. Выберем из них старший. Пусть это будет w_1. По построению одночлен w_1 старше всех остальных членов многочлена $H_1(\sigma_1, \sigma_2, \ldots, \sigma_n)$ и всех членов многочленов $H_i(\sigma_1, \sigma_2, \ldots, \sigma_n)$ $(i = 2, \ldots, s)$. Поэтому после приведения подобных членов в сумме

$$H_1(\sigma_1, \sigma_2, \ldots, \sigma_n) + H_2(\sigma_1, \sigma_2, \ldots, \sigma_n) + \ldots + H_s(\sigma_1, \sigma_2, \ldots, \sigma_n) = H(\sigma_1, \sigma_2, \ldots, \sigma_n)$$

член w_1 не сократится, так что эта сумма не будет нулевой, что противоречит нашем предположению. ■

49 Теорема Виета. Дискриминант многочлена. Доказательство того, что дискриминант — многочлен от коэффициентов

Теорема 49.1 (Виет). Пусть многочлен $f \in \mathbb{F}[x]$ имеет $n = \deg f$ корней с учётом кратностей. То есть,

$$f = a_0 x^n + a_1 x^{n-1} + \ldots + a_n = a_0(x - \alpha_1)(x - \alpha_2)\ldots(x - \alpha_n).$$

Тогда

$$\sigma_k(\alpha_1, \alpha_2, \ldots, \alpha_n) = (-1)^k \frac{a_i}{a_0}, \quad (k = 1, 2, \ldots, n).$$

Доказательство.

$$f = a_0 x^n + a_1 x^{n-1} + \ldots + a_k x^{n-k} + \ldots + a_n = a_0(x - \alpha_1)(x - \alpha_2)\ldots(x - \alpha_n) =$$
$$= a_0 \sum_{k=0}^{n} x^{n-k} \underbrace{\sum_{i_1 < \ldots < i_k} (-1)^k \alpha_{i_1}\alpha_{i_2}\ldots\alpha_{i_k}}_{(-1)^k \sigma_k} = a_0 \sum_{k=0}^{n} x^{n-k} \underbrace{(-1)^k \sigma_k(\alpha_1, \ldots, \alpha_n)}_{a_k}.$$

■

Определение 49.1. Пусть многочлен $f = a_0 x^n + \ldots + a_n$ имеет n корней $\alpha_1, \alpha_2, \ldots, \alpha_n$ с учётом кратности. **Дискриминант** $D(f)$ многочлена $f \in \mathbb{F}[x]$ равен

$$D(f) = a_0^{2n-2} \prod_{i<j} (\alpha_i - \alpha_j)^2.$$

Теорема 49.2 (Основное свойство дискриминанта). Пусть многочлен $f = a_0 x^n + \ldots + a_n$ имеет n корней $\alpha_1, \alpha_2, \ldots, \alpha_n$ с учётом кратности. $D(f) = 0$ тогда и только тогда, когда у f есть кратные корни.

Доказательство. Если $D(f) = 0$, то существуют i и j такие, что $x_i = x_j$. ■

Теорема 49.3. Пусть многочлен $f = a_0 x^n + \ldots + a_n$ имеет n корней $\alpha_1, \alpha_2, \ldots, \alpha_n$ с учётом кратности. Тогда $D(f)$ — многочлен от коэффициентов a_i.

Доказательство. Из определения дискриминанта $D(f) = a_0^{2n-2} V^2(\alpha_1, \ldots, \alpha_n)$. Как уже доказывалось, этот многочлен симметрический. Применяя теорему Виета, получаем требуемое. ■

50 Поле частных целостного кольца. Вложение целостного кольца в своё поле частных Поле рациональных дробей. Формальное и функциональное равенство рациональных дробей

Таким же образом, как кольцо целых чисел расширяется до поля рациональных дробей, любое целостное кольцо можно расширить дополя.

Пусть A — целостное кольцо. Рассмотрим множество пар (a, b), гед $a, b \in A$, $b \neq 0$, и определим в нём отношение эквивалентности

$$(a_1, b_1) \sim (a_2, b_2) \Leftrightarrow a_1 b_2 = a_2 b_1.$$

Рефлексивность и симметричность этого отношения очевидны; докажем его транзитивность. Если $(a_1, b_1) \sim (a_2, b_2)$ и $(a_2, b_2) \sim (a_3, b_3)$, то

$$a_1 b_2 b_3 = a_2 b_1 b_3 = a_3 b_1 b_2,$$

откуда после сокращения на b_2 получаем

$$a_1 b_3 = a_3 b_1,$$

т. е. $(a_1, b_1) \sim (a_3, b_3)$.

Из данного определения следует, что

$$(a, b) \sim (ac, bc) \qquad (*)$$

для любого $c \neq 0$. С другой стороны, как показывает следующая ниже цепочка эквивалентностей, любая эквивалентность $(a_1, b_1) \sim (a_2, b_2)$ является следствием эквивалентностей типа $(*)$:

$$(a_1, b_1) \sim (a_1 b_2, b_1 b_2) = (a_2 b_1, b_1 b_2) \sim (a_2, b_2).$$

Определим теперь сложение и умножение пар по правилам

$$(a_1, b_1) + (a_2, b_2) := (a_1 b_2 + a_2 b_1, b_1 b_2), \quad (a_1, b_1) \cdot (a_2, b_2) := (a_1 a_2, b_1 b_2).$$

Докажем, что определённое выше отношение эквивалентности согласовано с этими операциями. В силу предыдущего достаточно показать, что при умножении обоих членов одной из пар (a_1, b_1) и (a_2, b_2) на элемент $c \neq 0$ сумма и произведение этих пар заменятся эквивалентными им парами; но очевидно, что при такой операции оба члена суммы и произведения умножатся на тот же элемент c.

Класс эквивалентности, содержащий пару (a, b), условимся записывать как дробь a/b. Ввиду доказанного выше операции сложения и умножения пар определяют операции сложения и умножения дробей, осуществляемые по обычным правилам:

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} = \frac{a_1 b_2 + a_2 b_1}{b_1 b_2}, \quad \frac{a_1}{b_1} \cdot \frac{a_2}{b_2} = \frac{a_1 a_2}{b_1 b_2}.$$

Докажем, что относительно этих операций дроби образуют поле.

Очевидно, что сложение дробей коммутативно и ассоциативно. Дробь $\frac{0}{1}$ служит нулём для операции сложения дробей, а дробь $-\dfrac{a}{b}$ противоположна дроби $\dfrac{a}{b}$. Таким образом, дроби образуют абелеву группу относительно сложения.

Коммутативность и ассоциативность умножения очевидны. Следующая цепочка равенств доказывает дистрибутивность умножения дробей относительно сложения (две дроби приведём к общему знаменателю):

$$\left(\frac{a_1}{b} + \frac{a_2}{b}\right) \frac{a_3}{b_3} = \frac{(a_1 + a_2) a_3}{b b_3} = \frac{a_1 a_3 + a_2 a_3}{b b_3} = \frac{a_1}{b} \frac{a_3}{b_3} + \frac{a_2}{b} \frac{a_3}{b_3}.$$

Дробь $1/1$ служит единицей для операции умножения дробей, а при $a \neq 0$ дробь $\dfrac{b}{a}$ обратна дроби $\dfrac{a}{b}$.

Определение 50.1. Построенно поле называется **полем частных целостного кольца** A и обозначается через $\operatorname{Quot} A$.

Сложение и умножения дробей вида $a/1$ сводятся к соответствующим операциям над их числителями. Кроме того, $a/1 = b/1$ только при $a = b$. Следовательно, дроби такого вида образуют подкольцо, изоморфное

A. Условившись отождествлять дробь вида $a/1$ с элементом a кольца A, мы получим вложение кольца A в поле $\operatorname{Quot} A$.

Далее, поскольку

$$\frac{a}{b}\frac{b}{1} = \frac{a}{1},$$

дробь a/b равна отношению элементов a и b кольца A в поле $\operatorname{Quot} A$.

В силу $(a, c) \sim (ac, bc)$ дробь не изменится, если её числитель и знаменатель умножить или разделить (если это возможно) на один и тот же элемент кольца A. Если A — евклидово кольцо, то путём сокращения числителя и знаменателя на их наибольший общий делитель любая дробь приводится к виду a/b, где $\text{НОД}\,(a, b) = 1$.

Определение 50.2. Такой вид дроби называется **несократимым**.

Утверждение. Любой вид дроби над евклидовым кольцом получается из любого её несократимого вида умножением числителя и знаменателя на один и тот же элемент.

Доказательство. Пусть $\dfrac{a}{b} = \dfrac{a_0}{b_0}$, причём $(a_0, b_0) = 1$. Из равенства $ab_0 = a_0 b$ следует, что $b_0 \mid a_0 b$ и, значит, $b_0 \mid b$. Пусть $b = cb_0$; ясно, что тогда $a = ca_0$. ■

Примечание. Как следствие, несократимый вид дроби над евклидовым кольцом определён однозначно с точностью до умножения числителя и знаменателя на один и тот же необратимый элемент.

Определение 50.3. Поле частных кольца $\mathbb{F}[x]$ многочленов над полем $\mathbb{F}$ называется **полем рациональных дробей** над полем $\mathbb{F}$ и обозначается $F(x)$.

Каждая рациональная дробь определяет функция на $\mathbb{F}$ со значениями в $\mathbb{F}$, определённую там, где её знаменатель (в несократмой записи) не обращается в ноль.

Определение 50.4. Назовём две рациональные функции равными, если на множестве, где оба заменателя не равны нулю, эти функции равны.

Теорема 50.1. Если две рациональные дроби формально равны, то они функционально равны. Если же поле $\mathbb{F}$ бесконечно, то верно и обратное.

Доказательство. Пусть $\frac{f}{g} = \frac{h}{s}$. Возьмём $a \in \mathbb{F}$ такое, что $g(a) \neq 0$ и $s(a) \neq 0$. Тогда $f(x)s(x) = g(x)h(x)$. Следовательно, $f(a)s(a) = g(a)h(a)$, что влечёт $\frac{f(a)}{g(a)} = \frac{h(a)}{s(a)}$.

Наоборот, пусть поле $\mathbb{F}$ бесконечно и две функции $\frac{f(x)}{g(x)}$ и $\frac{h(x)}{s(x)}$ совпадают везде, кроме корней g и s. Тогда их разность

$$\frac{f(x)}{g(x)} - \frac{h(x)}{s(x)} = \frac{f(x)s(x) - h(x)g(x)}{g(x)s(x)}$$

обращается в ноль везде, кроме корней g и s. То есть, многочлен $f(x)s(x) - h(x)g(x)$ имеет бесконечное количество корней. Значит, этот многочлен формально равен нулю. Следовательно, $\frac{f}{s} = \frac{h}{s}$. ■

51 Несократимые правильные и простейшие рациональные дроби. Разложение правильной дроби в сумму простейших

Определение 51.1. Рациональная дробь называется **правильной**, если $\deg f < \deg g$.

Утверждение. Любую дробь можно представить в виде суммы многочлена и правильной дроби, причём такое представление единственно.

Доказательство. Поделим числитель на заменатель с остатком: $f = gq + r$. Тогда

$$\frac{f}{g} = q + \frac{r}{g}.$$

Единственость такого представления следует из единственности деления с остатком. ■

Лемма 51.1. Всякая правильная рациональная дробь вида

$$\frac{f}{g_1 g_2 \dots g_s},$$

где $g_1, g_2, \dots, g_s$ попарно взаимно просты, разлагается в сумму правильных дробей со знаменателями $g_1, g_2, \dots, g_s$, причём единственным образом.

Доказательство. Докажем это утверждение индукцией по s. Положим $g = g_1 g_2 \dots g_s$. При $s = 2$, согласно теореме 44.2, существуют такие многочлены u и v, что $ug_1 + vg_2 = 1$. Домножив это равенство на f, получаем $f = u^* g_1 + v^* g_2$. Разделив на g, получим

$$\frac{f}{g} = \frac{v^*}{g_1} + \frac{u^*}{g_2}.$$

Так как дробь $\dfrac{f}{g}$ правильная, то сумма целых частей дробей v^*/g_1 и u^*/g_2 равна нулю. Выделив их, мы получим разложение дроби $\dfrac{f}{g}$ в сумму правильных дробей со знаменателями g_1 и g_2.

При $s > 2$ заметим, что многочлены g_1 и $g_2 \dots g_s$ взаимно просты, и по доказанному дробь $\dfrac{f}{g}$ разлагается в сумму правильных дроблей со знаменателями g_1 и $g_2 \dots g_s$. По предположению индукции правильная дробь со знаменателем $g_2 \dots g_s$ разлагается в сумму правильных со знаменателями $g_2, \dots, g_s$.

Докажем единственность. Пусть для данной дроби существует два разложения:

$$\frac{f_1}{g_1} + \dots + \frac{f_s}{g_s} = \frac{f_1^*}{g_1} + \dots + \frac{f_s^*}{g_s}.$$

Перенесём всё в одну сторону:

$$\frac{f_1 - f_1^*}{g_1} + \dots + \frac{f_s - f_s^*}{g_s} = 0.$$

Отсюда из взаимной простоты легко понять справедливость этого утверждения для $s = 2$, а для остальных добивается индукцией. ■

Определение 51.2. Рациональная дробь $\dfrac{f}{g}$ называется **простейшей**, если $g = p^k$, где $p \in \mathbb{F}[x]$ — неприводимый многочлен, и $\deg f < \deg p$.

В частности, всякая дробь вида

$$\frac{a}{(x-c)^k}, \quad (a, c \in \mathbb{F})$$

является простейшей. В случае $\mathbb{F} = \mathbb{C}$ дробями такого вида исчерпываюся все простейшие дроби. В случае $\mathbb{F} = \mathbb{R}$ имеются ещё простейшие дроби вида

$$\frac{ax + b}{(x^2 + px + q)^k}, \quad (a, b, p, q \in \mathbb{R}), \text{ где } p^2 - 4q < 0.$$

Теорема 51.1. Всякая правильная рациональная дробь fg разлагается в сумму простейших дробей.

Доказательство. Пусть $g = p_1^{k_1} p_2^{k_2} \dots p_s^{k_s}$. Ввиду леммы 51.1 дробь $\dfrac{f}{g}$ разлагается в сумму правильных дробей со знаменателями $p_1^{k_1}, p_2^{k_2}, \dots, p_s^{k_s}$. Поэтому нам достаточно доказать теорему в случае, когда $g = p^k$, где p — неприводимый многочлен. В этом случае, разделив f на p с остатком, получим

$$\frac{f}{p^k} = \frac{f_1}{p^{k-1}} + \frac{r}{p^k}, \quad \deg r < \deg p.$$

Второе слагаемое является простейшей дробью, а первое является правильной дробью как разность правильных дробей. Продолжая процесс, получим требуемое разложение.

Единственность следует из единственности разложения в лемме 51.1. ■

52 Примитивные многочлены над факториальным кольцом. Любой многочлен пропорционален примитивному. Лемма Гаусса

Определение 52.1. Многочлен $f(x) = a_n x^n + \ldots + a_1 x + a_0$ ($f \in A[x]$) называется примитивным, если

$$\text{НОД}\,(a_1, a_2, \ldots, a_n) = 1.$$

Напомним, что для целостного кольца A можно определить поле частных $\operatorname{Quot} A$. При этом существует вложение A в $\operatorname{Quot} A$, $a \mapsto a/1$. Тогда кольцо многочленов $A[x]$ вкладывается в кольцо многочленов $(\operatorname{Quot} A)[x]$.

Лемма 52.1. Любой многочлен $f \in (\operatorname{Quot} A)[x]$ можно представить в виде $r \cdot g$, где $r \in \operatorname{Quot} A$, $g \in A[x]$ — примитивный многочлен.

Доказательство. Представим каждый ненулевой коэффициент r_i в виде несократимой дроби $r_i = \frac{a_i}{b_i}$. Тогда $a = \text{НОД}\,(a_i)$, $b = \text{НОК}\,(b_i)$. Положим $r = \frac{a}{b}$. Тогда $s_i = \frac{r_i}{r} = \frac{a_i b}{a b_i} \in A$, т. к. $a \mid a_i$ и $b_i \mid b$. С другой стороны, допустим, что простой множитель p делит все s_i. Теперь рассмотрим случаи:

Случай 1: $p \mid b$. Тогда существует i такое, что p входит в b_i в той же степени, что и в b. Т. к. $p \mid b_i$, получаем $p \nmid a_i$ и, следовательно, $p \nmid a$. В итоге p не делит $s_i = \frac{a_i b}{a b_i}$. Противоречие.

Случай 2: $p \nmid b$. Т. к. для каждого i выполнено $p \mid s_i$, то $p \mid \text{НОД}\,(a_i, b)$. Поскольку $p \nmid b$, получаем, чо для каждого i верно $p \mid a_i$. Но существует такое i, такое что степень вхождения p в a_i такая же, как и в a. При этом i степень вхождения p в $s_i = \frac{a_i b}{a b_i}$ не может быть положительной. Противоречие. ■

Лемма 52.2 (Гаусс). Произведение двух примитивных многочленов есть примитивный многочлен.

Доказательство. Пусть $f = a_0 + a_1 x + \ldots + a_n x^n$, $g = b_0 + b_1 x + \ldots + b_m x^m$ — примитивные многочлены из $A[x]$. Пусть $fg = c_0 + c_1 x + \ldots + c_{m+n} x^{m+n}$. При этом

$$c_k = \sum_{j=0}^{k} a_j b_{k-j}.$$

Допустим, что $p \mid c_i$ для всех i. Пусть j — минимальное число, такое что $p \nmid a_j$, а ℓ — минимальное число, такое что $p \nmid b_\ell$. Тогда

$$c_{j+\ell} = a_0 b_{j+\ell} + \ldots + a_{j-1} b_{\ell+1} + a_j b_\ell + a_{j+1} b_{\ell-1} + \ldots + a_{j+\ell} b_0.$$

Все слагаемые делятся на p, кроме $a_j b_\ell$. Значит, и вся сумма не делится. Противоречие. ■

Примечание. Если $f, g \in A[x]$ примитивны и $f = rg$, где $r \in \operatorname{Quot} A$, то r обратим. В самом деле, если неприводимый элемент p входит в числитель несократимого вида r, то все коэффициенты f делятся на p. А если в знаменатель, то все коэффициенты g делятся на p.

53 Факториальность кольца многочленов над факториальным кольцом

Теорема 53.1. Пусть A — факториальное кольцо. Тогда кольцо $A[x]$ также факториально.

Доказательство. Докажем, что неприводимые элементы в $A[x]$ — это неприводимые элементы $p \in A$ и примитивные многочлены $f \in A[x]$, которые неприводимы в $(\operatorname{Quot} A)[x]$. В самом деле, если многочлен степени ноль, то он приводим тогда и только тогда, когда разлагается на 2 неприводимых многочлена степени ноль, то есть приводим в A. Многочлен положительной степени может разлагаться либо на произведение необратимой константы и многочлена, либо на произведение двух многочленов положительной степени. Первое возможно тогда и только тогда, когда многочлен не является примитивным. Второе даёт разложение f в произведение над $(\operatorname{Quot} A)[x]$. Таким образом, примитивный многочлен, который неприводим в $(\operatorname{Quot} A)[x]$ является неприводимым в $A[x]$. Осталось доказать, что примитивный многочлен f, приводимый

в $(\operatorname{Quot} A)[x]$ приводим и в $A[x]$. Пусть $f = gh$, где $g, h \in (\operatorname{Quot} A)[x]$. По лемме существуют элементы r_g и r_h из $(\operatorname{Quot} A)[x]$ такие, что $r_g g$ и $r_h h$ примитивны. Имеем $f = (r_g r_h)^{-1}(r_g g r_h h)$. По лемме Гаусса $r_g g r_h h$ примитивен. По замечанию $r_g r_h$ обратим в A. Значит, $f = ((r_g r_h)^{-1} r_g g)(r_h h)$ — разложение на необратимые множители в $A[x]$.

Существование. Пусть теперь $f \in A[x]$ — произвольный многочлен. Представим его в виде $f = rg$, где g — примитивный многочлен, и $r \in A$. Тогда r можно разложить на неприводимые в A, а g можно разложить на неприводимые в $(\operatorname{Quot} A)[x]$. При этом разложение g можно сделать разложением на примитивные неприводимые. Таким образом мы можем разложить любой элемента $f \in A[x]$ на неприводимые в $A[x]$.

Единственность. Пусть $f = p_1 \dots p_m g_1 \dots g_k = q_1 \dots q_u h_1 \dots h_v$, где p_i, q_j — неприводимые в A, а g_i, h_j — примитивные неприводимые в $\operatorname{Quot} A$. Заметим, что НОД всех коэффициентов f равен $p_1 \dots p_m$, т. к. $g_1 \dots g_k$ — примитивный многочлен. Аналогично, этот же НОД равен $q_1 \dots q_u$. Значит, произведения $p_1 \dots p_m$ и $q_1 \dots q_u$ ассоциированы. Т. к. A факториально, $u = m$ и p_i и q_j попарно ассоциированы. Кольцо $(\operatorname{Quot} A)[x]$ факториально, значит два разложения f совпадают с точностью до перестановки и ассоциированности множителей. Отсюда $k = v$ и $g_i = r_i h_i$ для некоторого $r_i \in (\operatorname{Quot} A)[x]$. Как доказано ранее из того, что g_i и h_i примитивны следует, что r_i — обратимый элемент A. ■

54 Результант. Свойства результанта. Связь результанта многочлена и его производной с дискриминантом многочлена. Выражение результанта через определитель (формулировка)

Пусть даны два многочлена из $\mathbb{F}[x]$:

$$f = a_0 x^n + a_1 x^{n-1} + \ldots + a_n, \quad g = b_0 x^m + b_1 x^{m-1} + \ldots + b_m.$$

Будем считать, что у них количество корней с учётом кратности равно степени:

$$f = a_0 \prod_{i=1}^{N} (x - x_i), \quad g = b_0 \prod_{j=1}^{m} (x - y_j).$$

Определение 54.1. Результантом $R(f, g)$ многочленов f и g называется число

$$a_0^m b_0^n \prod_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m} (x_i - y_j).$$

Теорема 54.1 (Свойства результанта).

1. $R(f, g) = 0$ тогда и только тогда, когда f и g имеют общий корень («основное свойство результанта»);
2. $R(g, f) = (-1)^{mn} R(f, g)$;
3. $R(f, g) = a_0^m \prod_{i=1}^{n} g(x_i) = (-1)^{mn} b_0^n \prod_{j=1}^{m} f(y_j)$.

Доказательство. Свойства 1 и 2 сразу следуют из определения. Второе равенство свойства 3 доказывать тоже не нужно — оно сразу следует из свойства 2. Итак, докажем первое равенство свойства 3:

$$a_0^m \prod_{i=1}^{n} g(x_i) = a_0^m \prod_{i=1}^{n} \left(b_0 \prod_{j=1}^{m} (x_i - y_j) \right) = a_0^m b_0^n \prod_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m} (x_i - y_j) = R(f, g).$$

■

Теорема 54.2.

$$R(f, f') = (-1)^{\frac{n(n-1)}{2}} a_0 D(f).$$

Доказательство. $f = a_0(x - x_1) \dots (x - x_n)$. Тогда

$$f'(x) = a_0 \sum_{i=1}^{n} \frac{\prod (x - x_j)}{x - x_i}.$$

Имеем

$$f'(x_i) = a_0 \prod i \neq j(x_i - x_j).$$

Теперь докажем равенство лобовым вычислением:

$$R(f, f') = a_0^{n-1} \prod_{i=1}^{n} f'(x_i) = a_0^{2n-1} \prod_{i=1}^{n} \prod_{i \neq j} (x_i - x_j) = a_0^{2n-1} \prod_{i<j} (x_i - x_j) \prod_{i>j} (x_i - x_j) =$$
$$a_0^{2n-1} \prod_{i<j} (x_i - x_j)(-1)^{\frac{n(n-1)}{2}} \prod_{i<j} (x_i - x_j) = (-1)^{\frac{n(n-1)}{2}} a_0^{2n-1} \prod_{i<j} (x_i - x_j)^2 = (-1)^{\frac{n(n-1)}{2}} a_0 D(f).$$

■

Теорема 54.3.

$$R(f, g) = \det \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_n & 0 & 0 & 0 & \cdots & 0 \\ 0 & a_0 & a_1 & a_2 & \cdots & a_n & 0 & 0 & \cdots & 0 \\ 0 & 0 & a_0 & a_1 & a_2 & \cdots & a_n & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & a_0 & a_1 & a_2 & \cdots & a_n \\ b_0 & b_1 & b_2 & \cdots & b_n & 0 & 0 & 0 & \cdots & 0 \\ 0 & b_0 & b_1 & b_2 & \cdots & b_n & 0 & 0 & \cdots & 0 \\ 0 & 0 & b_0 & b_1 & b_2 & \cdots & b_n & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & b_0 & b_1 & b_2 & \cdots & b_n \end{pmatrix}$$

Дополнительные билеты допишу потом.