МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М. В. ЛОМОНОСОВА

Механико-математический факультет

Кафедра высшей алгебры

Курс лекций по алгебре

Лектор — Евгений Соломонович Голод

I курс, 1 семестр, отделение математики

Москва, 2006 г.

Предисловие

Внимание! Готовиться к экзамену по этому конспекту опасно для жизни!

Имеются противопоказания!

Если вам помог этот документ в сдаче экзамена, не поленитесь, подойдите и скажите "спасибо"... Знаете, как приятно? Любое другое вознаграждение приветствуется.

***Борис Агафонцев**, 102 группа*

Список литературы

[1] Конспекты лекций по алгебре. © МехМат, I курс, 1-й поток, 2006-2007 уч.г.

[2] Е.С. Голод. *Курс лекций по алгебре.* М., Издательство Центра прикладных исследований при механико-математическом факультете МГУ, 2004

[3] А.Г. Курош. *Курс высшей алгебры.* М.: Наука, 1971

[4] А.И. Кострикин. *Введение в алгебру. Основы алгебры.*. М.: Физматлит, 1994

[5] Пузыревский И.В. *Сказка про ε-окрестности.* М.: Hewlett-Packard, 2006

[6] Конспекты лекций И.Б. Кожухова в физико-математическом лицее №1557, © Первая группа, 2004-2006.

[7] ...

Последние изменения: 18 января 2007 г.

Об опечатках и неточностях пишите на agava@zelnet.ru

За информацией о последних изменениях и по другим вопросам обращайтесь по ICQ #216-059-136

Верстка в системе LaTeX 2_ε.

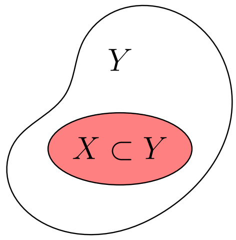
Содержание

1 Множества и отображения

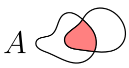
1.1 Множества

Определение. Под *множеством* принято понимать совокупность объектов некоторой природы (*множество* — неопределяемое понятие). Сами объекты называются *элементами множества.*

Над множествами определено несколько операций:



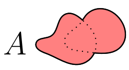
Пересечение



Обозначается «$\bigcap$».

$A \bigcap B \stackrel{\text{def}}{=} \{x | x \in A \land x \in B\}$

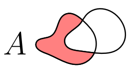
Объединение



Обозначается «$\bigcup$».

$A \bigcup B \stackrel{\text{def}}{=} \{x | x \in A \lor x \in B\}$

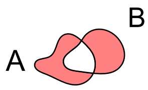
Разность



Обозначается «\».

$A \setminus B \stackrel{\text{def}}{=} \{x | x \in A \land x \notin B\}$

Симметрическая разность



Обозначается «Δ».

$A \Delta B \stackrel{\text{def}}{=} (A \setminus B) \cup (B \setminus A).$

Эти операции обладают следующими свойствами:

- Коммутативность пересечения и объединения:

$$A \cap B = B \cap A$$
$$A \cup B = B \cup A$$

- Ассоциативность пересечения и объединения:

$$(A \cap B) \cap C = A \cap (B \cap C)$$
$$(A \cup B) \cup C = A \cup (B \cup C)$$

- Дистрибутивность пересечения относительно объединения и наоборот.

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

- Относительно разности выполняются следующие соотношения:

$$M \setminus (A \cap B) = (M \setminus A) \cup (M \setminus B)$$
$$M \setminus (A \cup B) = (M \setminus A) \cap (M \setminus B)$$

- Относительно симметрической разности:

$$A\Delta B = (A \cup B) \setminus (A \cap B)$$

Доказательство.

1. Докажем включение $A\Delta B \subset (A \cup B) \setminus (A \cap B)$. Для этого возьмём произвольный элемент $a \in A\Delta B$, который по определению содержится в множестве $(A \setminus B) \cup (B \setminus A)$. Возможно две ситуации:
   (a) $a \in A \setminus B$
   (b) $a \in B \setminus A$

   Пусть имеет место пункт i. Тогда $A \in A$, $a \notin B \quad \Rightarrow \quad a \in A \cup B$ и $a \notin A \cap B$, откуда явно следует, что $a \in (A \cup B) \setminus (A \cap B)$. При предположении, что имеет место быть пункт ii рассуждения абсолютно аналогичны.
2. Докажем включение $A\Delta B \supset (A \cup B) \setminus (A \cap B)$. Возьмём произвольный элемент $a \in (A \cup B) \setminus (A \cap B)$, т.е.

$$\begin{cases} a \in A \cup B \\ a \notin A \cap B \end{cases}$$

   Возможно две ситуации:
   (a) $a \in A$
   (b) $a \in B$

   Пусть имеет место пункт i (для пункта ii рассуждения проводятся абсолютно аналогично). Так как $a \in A$ и $a \notin A \cap B$, то $a \notin B \quad \Rightarrow \quad a \in A \setminus B$. Включение доказано.

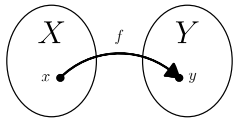
Утверждение доказано. □

1.2 Отображения

Определение. Множество $\mathbf{D} = \mathbf{A} \times \mathbf{B}$ называется декартовым (или прямым) произведением множеств, если оно состоит из всех возможнных пар $(a, b) \quad a \in A, b \in B$ и только из них.

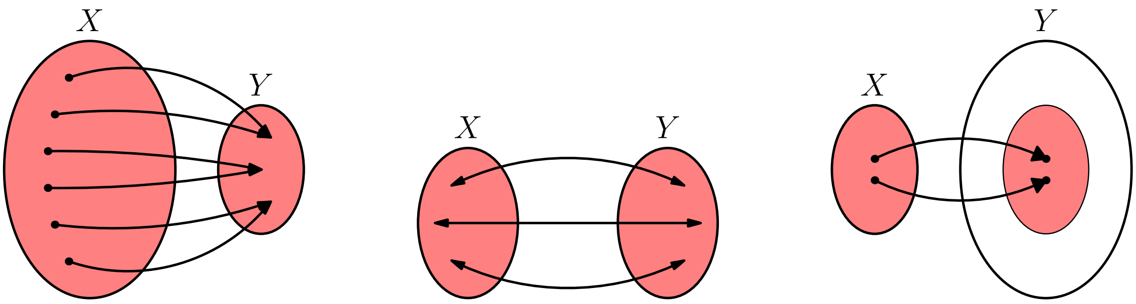
Определение. Подмножество $\mathbf{F}$ декартова произведения $\mathbf{D} = \mathbf{A} \times \mathbf{B}$ называется *отображением*, если $\forall x \in \mathbf{A} \quad \exists!(x, y) \in \mathbf{F}, \quad y \in B$. Тем самым отображение $\mathbf{F}$ определено на множестве $\mathbf{A}$. Обозначается следующим образом: $f\colon \mathbf{A} \to \mathbf{B}$ или $\mathbf{A} \xrightarrow{f} \mathbf{B}$. При этом пишут, что $y = f(x)$, $x \in \mathbf{A}$, $y \in \mathbf{B}$, а y называют *образом* x (соответственно, x — *прообразом* y). Множество всех элементов в $\mathbf{B}$, у которых есть прообраз в $\mathbf{A}$ назывют *полным проообразом*. Отображения f_1 и f_2 из $\mathbf{A}$ в $\mathbf{B}$ называют равными, если: $\forall x \in \mathbf{A} \quad f_1(x) = f_2(x)$.

Возможно и другое, менее строгое, определение функции. Пусть $\mathbf{F}$ — правило, согласно которому каждому $x \in \mathbf{A}$ ставится в соответствие не более одного элемента $y \in \mathbf{B}$. В таком случае правило $\mathbf{F}$ называется функцией. Если у x нет ни одного образа, то говорят, что функция $\mathbf{F}$ неопределена в точке x.



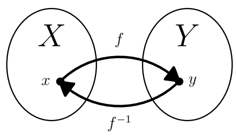
Существует классификация отображений, которая выделяет три типа:

- Отображение «на», или *сюръективное отображение*, или *сюръекция*, — это такое отображение, при котором каждый элемент из Y является образом хотя бы одного элемента из X (см. рис. 1a).
- Отображение «в», или *инъективное отображение*, или *инъекция* – это такое отображение, при котором двум разным элементам из X соответствуют разные элементы из Y: $x_1 \neq x_2 \quad \Rightarrow \quad f(x_1) \neq f(x_2)$ (см. рис. 1c).
- Отображение «между», или *биективное отображение*, или *биекция* — это такое отображение, которое является одновременно и сюръекцией, и инъекцией. То есть каждому элементу из X соответствует единственный элемент из Y и наоборот. Такое отображение также называется взаимнооднозначным (см. рис. 1b).



(a) Сюръекция (b) Биекция (c) Инъекция

Рис. 1: Классификация отображений



Обратное отображение Если задано биективное отображение $f\colon X \to Y$, то имеет смысл говорить об *обратном отображении*:

$$\forall y \in Y \exists !\, x \in X\colon y = f(x)$$

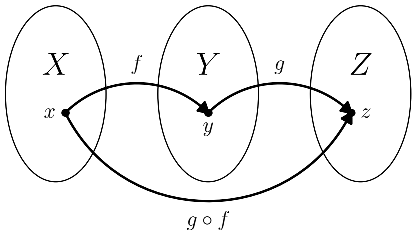
$$\exists g\colon Y \to X = f^{-1}\colon f^{-1}(y) = x$$

Очевидно, что $\left(f^{-1}\right)^{-1} \equiv f$.

Композиция отображений

Определение. Пусть заданы отображения $f\colon X \to Y$ и $g\colon Y \to Z$. Тогда композицией отображений f и g называют отображение $(g \circ f)$, которое действует по правилу:

$$(g \circ f)(x) \stackrel{\text{def}}{=} g(f(x))$$



Докажем два важных свойства композиций:

1. Композиция биекций является биекцией.
2. Ассоциативность композиций: $(g \circ f) \circ h = g \circ (f \circ h)$.

   Доказательство. Рассмотрим левую часть:

   $$((g \circ f) \circ h)(x) \overset{\text{def}}{=} (g \circ f)(h(x)) \overset{\text{def}}{=} g(f(h(x)))$$

   Рассмотрим правую часть:

   $$(g \circ (f \circ h))(x) \overset{\text{def}}{=} g((f \circ h)(x)) \overset{\text{def}}{=} g(f(h(x)))$$

   Очевидно, что они равны. □

Бином Ньютона

$$(a+b)^n = \underbrace{(a+b) \times \cdots \times (a+b)}_{n \text{ раз}} =$$
$$= \underbrace{a^n + na^{n-1}b + \cdots + \mathbf{C}_n^k a^{n-k} b^k + \cdots + nab^{n-1} + b^n}_{2^n \text{ неприведенных слагаемых}}$$

Таким образом,

$$(a+b)^n = \sum_{k=0}^{n} \mathbf{C}_n^k \times a^{n-k} \times b^k$$
$$T_{k+1} = \mathbf{C}_n^k \times a^{n-k} \times b^k$$

Полиномиальная формула

Полиномиальная формула является обобщением бинома Ньютона.

$$(a_1 + \cdots + a_k)^n = \sum_{i_1 + \cdots + i_k = n} \frac{n!}{i_1! \times \cdots \times i_k!} \times (a_1^{i_1} \times \cdots \times a_k^{i_k})$$

или

$$\left(\sum_{j=1}^{k} a_j\right)^n = \sum_{(\sum_{j=1}^{k} i_j) = n} \left(\frac{n!}{\prod_{j=1}^{k} i_j!} \times \prod_{j=1}^{k} a_j^{i_j}\right)$$

Число неподобных членов в разложении равно $\tilde{\mathbf{C}}_k^n = \mathbf{C}_{n+k-1}^k$ (получается, если представить n ячеек, в которые необходимо расставить с повторениями цифры от 1 до k).

Подсчёт числа отображений и подмножеств.

Используя принципы комбинаторики можно выяснить, что если $|X| = n, |Y| = m$, то:

1. $|P(X)|$ – количество всех подмножеств множества – равно 2^n.
2. Число всех m-элементных подмножеств в X равно $\mathbf{C}_n^m$.
3. Число всех отображений из X в Y равно m^n
4. Число всех сюръективных отображений из X в Y называется числом Стирлинга 2-го рода и вычисляется по формуле

$$S(n, m) = \frac{\sum\limits_{k=0}^{m} (-1)^k \mathbf{C}_m^k (m-k)^n}{m!}$$

   Прочитать об этом можно здесь:

   (a) http://mathworld.wolfram.com/StirlingNumberoftheSecondKind.html
   (b) http://ndp.jct.ac.il/tutorials/Discrete/node81.html

5. Число всех инъективных отображений из X в Y равно $m(m-1)\dots(m-n+1) = \mathbf{C}_n^m \cdot m! = \mathbf{A}_n^m$
6. Число всех биективных отображений из X в X равно $n!$

2 Перестановки

2.1 Группа перестановок

Под перестановками на множестве понимаются биективные отображения множества в себя. Рассмотрим множество, состоящее из конечного числа элементов. Занумеруем их от 1 до n. Множество всех биективных отображений из $(1, 2, \ldots n)$ в себя обозначим через S_n, $|S_n| = n!$. Всякую из таких перестановок можно записать в виде таблицы

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ i_1 & i_2 & \ldots & i_n \end{pmatrix},$$

которая означает, что при выполнении такого отображения элемент с номером 1 переходит в элемент с номером i_1 и т.п.

Перестановки не обладают свойством коммутативности, но обладают свойством ассоциативности. Существование единичной и обратной перестановки гарантируют нам то, что множество всех перестановок S_n является *группой*.

Условимся перемножать перестановки справа налево.

Определение. *Цикл* — перестановка, такая что $\alpha_1 \to \alpha_2 \to \cdots \to \alpha_k \to \alpha_1$, $\alpha_1, \alpha_2 \ldots \alpha_k \in \{1, 2, \ldots n\}$. Обозначается $(\alpha_1, \alpha_2, \ldots \alpha_k)$.

Если при выполнении какой-то циклической перестановки элементы множества переходят сами в себя, то они называются неподвижными, иначе – перемещаемыми. Два цикла называются независимыми, если у них нет общих перемещаемых элементов. Два независимых цикла коммутативны.

Теорема 2.1. *Всякая перестановка представима в виде произведения независимых циклов.*

Определение. *Транспозиция* – цикл длины 2, или изменение мест двух элементов между собой.

Утверждение. Любой цикл длины k представим в виде произведения $k - 1$ транспозиций:

$$(k_1, k_2 \ldots, k_n) = (k_1, k_n) \ldots (k_1, k_3)(k_1, k_2) = (k_1, k_2)(k_2, k_3) \ldots (k_{n-1}, k_n)$$

Утверждение. Число циклов длины k в S_n равно $\mathbf{C}_n^k \cdot (k-1)!$

Определение. *Декрементом* подстановки называется сумма длин независимых циклов в её разложении, уменьшенных на 1: $d = d(\sigma) = \sum\limits_{i=1}^{s} (k_i - 1)$. Легко понять, что $\operatorname{sgn} \sigma = (-1)^d$.

2.2 Чётность, знак перестановки

Определение. Назовём пару элементов (i, j) *правильной*, если $i < j \quad \Rightarrow \quad \sigma(i) < \sigma(j)$ и *неправильной* в обратном случае. Чётностью перестановки будет называть чётность количества неправильных пар k в этой перестановке. Знак перестановки $\operatorname{sgn}\sigma \overset{\text{def}}{=} (-1)^k$.

Утверждение. Транспозиция меняет знак перестановки. При умножении перестановок из знаки также перемножаются: $\operatorname{sgn}(\pi\sigma) = \operatorname{sgn}\pi \cdot \operatorname{sgn}\sigma$.

Утверждение. Число чётных перестановок равно числу нечётных и равно $n!/2$.

Определение. $\mathbf{A}_n \subset \mathbf{S}_n = \{\sigma | \operatorname{sgn}\sigma > 0\}$ — подгруппа, называемая *знакопеременной группой перестановок*.

3 Алгебраические структуры

3.1 Полугруппы

Определение. Пусть $\mathbf{G}$ – некоторое множество. Тогда *бинарной операцией* на этом множестве будет называться правило, согласно которому каждой упорядоченной паре (a, b), $a, b \in \mathbf{G}$, ставится в соответствие некоторый элемент $c \in \mathbf{G}$. Операция называется *частичной*, если она определена не для всех пар.

Определение. Операция называется *ассоциативной*, если $\forall a, b, c \in \mathbf{G}$ $(ab)c = a(bc)$. Множество, на котором задана ассоциативная операция, называется *полугруппой*. В случае ассоциативной операции значение выражения вообще не зависит от расстановки скобок:

Замечание. В случае частичной бинарной операции следует сказать, что закон ассоциативности состоит в следующем: если определено хотя бы одно из произведений $a(bc)$ или $(ab)c$, то второе также определено и они равны.

Теорема 3.1 (обобщённый закон ассоциативности). *Произведение n элементов в полугруппе не зависит от способа расстановки скобок.*

Доказательство. Проведём доказательство индукцией по n. База индукции – $n = 3$ – известно.

Теперь пусть $n \geqslant 4$ фиксированно и известно, что для произведения меньшего чем n числа элементов утверждение теоремы верно. Выделим скобки, над которыми производится последняя операция: $(a_1 \cdot \dots \cdot a_k) \cdot (a_{k+1} \cdot \dots \cdot a_n)$. Назовём такую расстановку скобок расстановкой типа k. Внутри каждой из них возможна любая расстановка скобок по предположению индукции.

Нам надо доказать, что для любой расстановки типа m, где $1 \leqslant m \leqslant n-1$ произведение определено и результат одинаков. Для этого нам дано, что для некоторой расстановки типа k это произведение определено.

Достаточно показать, что определены произведения для расстановок типа $k \pm 1$ тоже определены и результат одинаков во всех трёх случаях. Покажем, например, переход от k к $k-1$:

$$((a_1 \cdot \dots \cdot a_{k-1}) \cdot a_k) \cdot (a_{k+1} \cdot \dots \cdot a_n) = (a_1 \cdot \dots \cdot a_{k-1}) \cdot (a_k \cdot (a_{k+1} \cdot \dots \cdot a_n))$$

Нетрудно заметить, что доказательство на этом завершается. □

Определение. Пусть $\mathbf{G}$ – множество с бинарной операцией. Элемент $e \in \mathbf{G}$ называется *левой единицией*, если $ea = a \quad \forall a \in \mathbf{G}$. Определение *правой единицы* аналогично. Если в $\mathbf{G}$ имеется левая единица e_1 и правая единица e_2, то они совпадают и элемент $e = e_1 = e_2$ называется (двусторонней) единицей.

Определение. Пусть $(\mathbf{G}, \cdot)$ обладает единицей. Тогда для элемента a элемент a' называется *левым* (соответственно *правым*) *обратным*, если $a'a = e$ $(aa' = e)$. Также определяется двусторонний обратный элемент, который обозначается как a^{-1}. Элемент, имеющий обратный, называется *обратимым*.

Теорема 3.2. *Если элемент обладает левым обратным и правым обратным, то они равны между собой.*

Замечание. На экзамене необходимо проиллюстрировать примерами все эти свойства.

Определение. По определению $a^n = a \cdot \dots \cdot a$ (n раз), $a^0 = e$, $a^{-n} = (a^{-1})^n = (a^n)^{-1}$.

Определение. Элементы $a, b \in \mathbf{G}$ называются *коммутирующими*, если $ab = ba$. Если это выполнено для любых элементов, то операция (множество с этой операцией) называется коммутативной.

3.2 Группы

Определение. *Группа* — множество с операцией, удовлетворяющее условиям:

Интересно, какую хирургическую операцию надо провести каждому студенту из группы №102?

1. Ассоциативность
2. $\exists e: \quad ae = ea = a$
3. $\forall a \quad \exists b: \quad ab = ba = e$. Обозначается $b = a^{-1} \quad (a = b^{-1})$

Определение. Порядком элемента a в группе называется наименьшее $n \in \mathbb{N}$, такое что $a^n = e$. Обозначается $o(a) = n$. Если $\nexists o(a)$, то принято считать, что $o(a) = \infty$.

Порядок элемента группы отвечает следующему свойству:

Теорема 3.3. *Пусть* $o(a) = n$. *Тогда* $o(a^k) = \frac{n}{(k,n)}$

Доказательство. Пусть m – положительное число такое, что $(a^k)^m = e$. Значит $n \mid km$. Поэтому m – наименьший такой показатель, когда km — наименьшее общее кратное чисел k и n, т.е. $km = \frac{kn}{(k,n)} \quad \Rightarrow \quad m = \frac{n}{(k,n)}$. □

Теорема 3.4. *Порядок любого элемента конечной группы является делителем порядка группы (под порядком множества понимается количество его элементов).*

Доказательство. Пусть $|\mathbf{G}| = n$ и $a_1, \dots a_n$ – её элементы. Пусть порядок произвольного элемента $a \in \mathbf{G} \quad o(a) = k$. Рассмотрим такое отображение π_a: $\pi_a(a_i) = a a_i$. Это отображение биективно, следовательно может быть рассмотрено как перестановка на множестве $\mathbf{G}$. Разложим её в произведение независимых циклов. Каждый из циклов имеет вид $(a_i, a a_i, \dots a^{k-1} a_i)$ так как $a^l a_i \neq a_i$ при $1 \leqslant l < k$; таким образом, если их число равно m, то $km = n$. □

3.3 Кольца

Определение. Если на множестве K задано две операции (условно «+» и «·») и выполенены следующие условия:

1. $(K, +)$ – абелева (коммутативная) группа

2. $(K, \cdot)$ – полугруппа
3. $a(b+c) = ab + ac$ и $(a+b)c = ac + bc$,

то $(K, +, \cdot)$ называется *кольцом*.

Если в $(K, \cdot)$ существует единица $(1 \neq 0)$, то K называется кольцом с единицей. Кольцо называется коммутативным, если $\forall x, y \quad xy = yx$. Обратимость элементов в кольце зависит от свойств алгебраической структуры $(K, \cdot)$.

Определение. Элемент a называется левым (правым, двусторонним) *делителем нуля*, если $\exists b \neq 0$ такой, что $ab = 0$ ($ba = 0$, $ab = ba = 0$ соответственно).

Определение. *Поле* – коммутативное кольцо с $1 \neq 0$, в котором всякий ненулевой элемент обратим.

3.4 Кольца вычетов по модулю n

На множестве целых чисел введём отношение эквивалентности таким образом, что два числа a и b будут считаться эквивалентными, если $a \equiv b \mod n$. Все целые числа сгруппируются по классам эквивалентности, множество этих классов будем обозначать как $\mathbb{Z}/n$, а сами классы обозначим как $[a] = \{x \mid x \equiv a \mod n\}$. На множестве классов логично определить операции сложения и умножения как $[a] + [b] = [a+b]$ и $[a][b] = [ab]$, но требует проверки корректность такого определения, то есть чтобы результат не зависел от выбранных представителей в каждом из классов. В данном случае это не составит большого труда. Также тривиальны и остальные необходимые проверки, которые должны нам показать, что множество классов вычетов с такими операциями является коммутативным кольцом с единицей.

Также легко заметить, что элемент $[k]$ в $\mathbb{Z}/n$ обратим в том и только том случае, если $(k, n) = 1$. Так как кольцо вычетов конечно и коммутативно, то всякий неделитель нуля в нём обратим. Действительно, если a – неделитель нуля, то все элементы ax различны, значит один из них равен 1.

Далее, обозначим за $U(\mathbb{Z}/n)$ группу всех обратимых элементов кольца $\mathbb{Z}/n$. Порядок этой группы равен $\varphi(n)$. Кольцо $\mathbb{Z}/n$ является полем тогда и только тогда, когда n – простое.

Теорема 3.5 (обобщённая малая теорема Ферма). $(k, n) = 1 \quad \Rightarrow \quad k^{\varphi(n)} \equiv 1 \mod n$.

Доказательство. Если $(k, n) = 1$, то $[k] \subset U(\mathbb{Z}/n)$. По теореме 3.4 имеем, что $[k]^{\varphi(n)} = [1]$, что и требовалось доказать. $\square$

Замечание. В частном случае, когда n – простое, имеем, что почти все k таковы, что $(k, n) = 1$ и $\varphi(n) = n - 1$, то есть $k^n \equiv k \mod n$.

Определение. *Характеристикой поля* называется такое наименьшее неотрицательное число n, что $1 + \dots + 1$ (n раз) $= 0$. В случае, если такое не выполняется ни при каком натуральном n, принято считать, что характеристика поля $\operatorname{char} \mathbf{P} = 0$. Если $\operatorname{char} \mathbf{P} = n$, то n – простое.

3.5 Построение поля частных области целостности

Теорема 3.6. *Любое коммутативное кольцо с единицей и без делителей нуля вкладывается в поле.*

Доказательство. Пусть $\mathbf{K}$ – коммутативное кольцо без делителей нуля. Это означает, что

$$ab = 0 \quad \Rightarrow \quad (a = 0) \vee (b = 0)$$

Для доказательства построим некоторое множество, состоящее из элементов кольца $\mathbf{K}$, если данное множество с определёнными на нём операциями будет являться полем, и будет существовать вложение $\varphi: \quad \mathbf{K} \to \mathbf{F}$, то кольцо $\mathbf{K}$ будет вложено в это поле.

Рассмотрим множество пар $(a, b), \quad a, b \in \mathbf{K}, \quad b \neq 0$. Пусть пары (a, b) и (c, d) эквивалентны, если $ad = bc$. Эквивалентные пары будем обозначать таким образом: $(a, b) \sim (c, d)$. Пусть $\mathbf{A}$ — множество всех таких пар. Отождествим эквивалентные пары. Это значит, что

$$(a, b) \sim (c, d) \quad \& \quad (c, d) \sim (u, v) \quad \Rightarrow \quad (a, b) \sim (u, v)$$

▶

$$\begin{cases} ad &= bc \\ cv &= du \end{cases} \Leftrightarrow \begin{cases} adv &= bcv \\ bcv &= bdu \end{cases} \Rightarrow \quad adv = bdu \quad \Rightarrow \quad av = bu \blacktriangleleft$$

Обозначим за $[(a, b)]$ класс всех пар (a_i, b_i), эквивалентных между собой. Тогда пусть множество всех классов $\mathbf{F} = \{[(a, b)] \mid (a, b) \in \mathbf{A}\}$. Докажем, что это множество с указанными ниже операциями является полем, и в него будет вложено исходное кольцо $\mathbf{K}$.

На множестве $\mathbf{F}$ определим сложение двух классов:

$$[(a, b)] + [(c, d)] = [(ad + bc, bd)]$$

Далее будем обозначать класс или как $[(a, b)]$, или как $\frac{a}{b}$. Тогда определение сложения примет привычный вид $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$. Покажем, что класс суммы остаётся неизменным при различном выборе складываемых пар в классах слагаемых.

▶ Пусть $(a, b)+(c, d) = (ad+bc, bd), \quad (a', b')+(c', d') = (a'd'+b'c', b'd')$. Иначе,

$$\begin{cases} ab' &= ba' \\ cd' &= dc' \end{cases} \Leftrightarrow \begin{cases} ab'dd' &= dd'ba' \\ cd'bb' &= dc'bb' \end{cases}$$

Сложим оба равенства:

$$ab'dd' + cd'bb' = dd'ba' + dc'bb'$$

$$(ad + cb)b'd' = (d'a' + c'b')bd$$

Последнее означает не что иное, как $(ad + bc, bd) \sim (a'd' + b'c', b'd')$. ◀

Теперь определим операцию умножения двух классов как $\frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd}$. Необходимо теперь доказать корректность умножения, то есть что при различном выборе пар в классах $\frac{a}{b}$ и $\frac{c}{d}$ класс их произведения будет оставаться неизменным. Это доказательство аналогично вышеприведённому для сложения.

Для определённых нами операций сложения и умножения проверим аксиомы поля.

1. $\frac{a}{b} + \frac{c}{d} = \frac{c}{d} + \frac{a}{b}$ (по определению)
2. $\left(\frac{a}{b} + \frac{c}{d}\right) + \frac{s}{t} = \frac{a}{b} + \left(\frac{c}{d} + \frac{s}{t}\right)$

   ► Рассмотрим правую и левую части. Если мы их равносильными преобразованиям приведём к равным выражениям, то доказательство будет завершено.

$$\left(\frac{a}{b} + \frac{c}{d}\right) + \frac{s}{t} = \frac{ad + bc}{bd} + \frac{s}{t} = \frac{adt + bct + sbd}{bdt}$$

$$\frac{a}{b} + \left(\frac{c}{d} + \frac{s}{t}\right) = \frac{a}{b} + \frac{ct + ds}{dt} = \frac{adt + bct + sbd}{bdt} \blacktriangleleft$$

3. Класс $[(0; a)]$ отождествим с нулём.
4. $\forall \frac{a}{b} \quad \exists \frac{-a}{b}: \quad \frac{a}{b} + \frac{-a}{b} = 0$
5. $\left(\frac{a}{b} + \frac{c}{d}\right) \frac{s}{t} = \frac{a}{b}\frac{s}{t} + \frac{c}{d}\frac{s}{t}$
6. $\frac{a}{b} \cdot \frac{c}{d} = \frac{c}{d} \cdot \frac{a}{b}$
7. $\left(\frac{a}{b} \cdot \frac{c}{d}\right) \frac{s}{t} = \frac{a}{b} \left(\frac{c}{d} \cdot \frac{s}{t}\right)$
8. Класс $[(a, a)]$ отождествим с единицей.
9. Обратным к $\frac{a}{b}$ будет являться $\frac{b}{a}$.

Таким образом мы показали, что $\mathbf{F}$ с определёнными нами операциями является полем. $\mathbf{F}$ называется полем частных кольца $\mathbf{K}$.

Определим вложение $\varphi: \quad a \to \frac{ax}{x}, \quad a \in \mathbf{K}, \ \frac{ax}{x} \in \mathbf{F}$.

Это вложение сохраняет операцию: $a + b \xrightarrow{\varphi} [((a + b)x, x)] = [(ax + bx, x)] = [(ax, x)] + [(bx, x)]$ (аналогично проверяется для умножения) и разные элементы переходят в разные:

► Пусть $a \to \frac{ab}{b}, \quad a' \to \frac{a'b}{b}$. Необходимо доказать, что при различных a и a', $\frac{ab}{b}$ и $\frac{a'b}{b}$ также не совпадают. Докажем от противного. Пусть $a \neq a', \quad \frac{ab}{b} = \frac{a'b}{b}$. Последнее выражение означает, что $(ab, b) \sim (a'b, b)$ или $abb = a'bb$. Так как исходное кольцо $\mathbf{K}$ без делителей нуля, то или $a = a'$, или $b = 0$. Но мы определили классы в множестве $\mathbf{A}$ так, что $b \neq 0$. Следовательно $a = a'$. Противоречие. ◄

Таким образом, мы построили такое множество $\mathbf{F}$ из элементов $\mathbf{K}$ и определили на нём две операции таким образом, что множество с этими операциями является полем и показали такое отображение $\mathbf{K} \xrightarrow{\varphi} \mathbf{F}$, что оно является вложением. □

Пример. Кольцо целых чисел $\mathbb{Z}$ вложено в поле рациональных чисел $\mathbb{Q}$: $\mathbb{Z} \subset \mathbb{Q}$

3.6 Понятие гомоморфизма и изоморфизма

Определение. Отображение $\varphi\colon (G, \cdot) \to (K, *)$ называется *гомоморфизмом*, если оно «сохраняет операцию», т.е. $f(a \cdot b) = f(a) * f(b) \quad \forall a, b \in G$.

Утверждение. Образ группы при гомоморфизме является группой. При этом $\varphi(e) = e'$. Однако при гомоморфизме полугрупп образ единицы не обязательно является единицей.

Гомоморфизм колец уже обязан сохранять обе операции. В общем случае, опять же, единица не обязательно переходит в единицу.

Теорема 3.7. *Пусть $\mathbf{P}$ – поле и $\mathbf{K}$ – кольцо. В таком случае если существует гомоморфизм $\varphi\colon \mathbf{P} \to \mathbf{K}$, то φ либо тривиален, либо инъективен.*

Доказательство. Пусть $x \neq y$ и $f(x) = f(y) \quad \Rightarrow \quad f(x-y) = 0$. В силу существования $(x-y)^{-1}$ имеем $f(1) = f\left((x-y)(x-y)^{-1}\right) = 0$, то есть $\forall a \in \mathbf{P}\colon f(a) = f(a \cdot 1) = 0$. □

Определение. Отображение φ называется *изоморфизмом*, если φ является гомоморфизмом и φ – биективно. Два множества с алгебраическими операциями называются изоморфными, если существует изоморфизм, переводящий одно в другое.

Теорема 3.8 (Кейли). *Всякая группа порядка n изоморфна некоторой подгруппе симметрической группы S_n.*

Доказательство. Пусть $\mathbf{G} = \{a_1, a_2, \ldots, a_n\}$ – группа порядка n. Составим для группы $\mathbf{G}$ таблицу Кейли и сопоставим каждому элементу $a_i \in \mathbf{G}$ перестановку на множестве $\mathbf{G}$, задаваемую i-й строкой таблицы Кейли, то есть рассмотрим отображение $f\colon \mathbf{G} \to S_n$, при котором $f(a) = \pi_a, \quad \pi_a(a_j) = aa_j$. Разным элементам группы $\mathbf{G}$ соответствуют разные перестановки, таким образом отображение $\mathbf{G} \to \operatorname{Im} f$ – биективно.

Докажем, что оно является гомоморфизмом. $\forall a_j \in \mathbf{G}$ имеем $\pi_{ab}(a_j) = (ab)a_j = a(ba_j) = \pi_a(\pi_b(a_j)) = (\pi_a \circ \pi_b)(a_j)$. □

Теорема 3.9. *Пусть $\mathbf{P}$ – поле. Имеется единственный инъективный гомоморфизм*

1. $f\colon \mathbb{Q} \to \mathbf{P}$, *если* $\operatorname{char} \mathbf{P} = 0$;
2. $f\colon \mathbb{Z}/p \to \mathbf{P}$, *если* $\operatorname{char} \mathbf{P} = p$.

В обоих случаях $\operatorname{Im} f$ является единственным минимальным полем поля $\mathbf{P}$.

Доказательство.

1. $\operatorname{char} \mathbf{P} = 0$. Тогда $\forall n\colon n \cdot 1 \neq 0$. Зададим отображение f по правилу $f\left(\frac{m}{n}\right) = (m \cdot 1)(n \cdot 1)^{-1}$. Так как запись рационального числа в виде дроби неоднозначна, то необходимо проверить, что легко сделать, что подобное задание f корректно. Далее проверяем, что f – гомоморфизм, который является инъективным, так как $\mathbf{Q}$ – поле. Так как $f(1_{\mathbb{Q}}) = 1_{\mathbf{P}}$, то f определён однозначно.

2. $\operatorname{char} \mathbf{P} = p$. Определим $f([k]) = k \cdot 1$. Так как $o(1) = p$, то $k \cdot 1 = l \cdot 1 \quad \Leftrightarrow \quad k \equiv l \mod p$ и отображение определено корректно. Очевидно, что оно гомоморфизм, инъективно и однозначно.

В обоих случаях $\operatorname{Im} f$ содержится в любом подполе $\mathbf{L}$ поля $\mathbf{P}$ так как $1 \in \mathbf{L}$, а значит все элементы $n \cdot 1 \in \mathbf{L}$ и $(m \cdot 1)(n \cdot)^{-1} \in \mathbf{L} \quad (1 \cdot 1 \neq 0)$. $\operatorname{Im} f$ называется *простым подполем* поля $\mathbf{P}$. □

3.7 Циклические группы

Определение. Пусть $(\mathbf{G}, \cdot)$ – группа и $a \in \mathbf{G}$. Тогда множество $\langle a \rangle = \{a^n \mid n \in \mathbb{Z}\}$ называется *циклической подгруппой* с *порождающим элементом* a. Очевидно, что $|\langle a \rangle| = o(a)$. Может так оказаться, что $\langle a \rangle = \mathbf{G}$. В таком случае сама группа $\mathbf{G}$ называется циклической.

Если $\langle a \rangle$ – циклическая группа порядка n, то порядок элемента $o(a^k) = \frac{n}{(n,k)}$. Если $(n, k) = 1$, то элемент a^k является порождающим в этой группе. Легко понять, что число образующих в группе порядка n равно $\varphi(n)$. Очевидно, что если группа простого порядка, то каждый элемент этой группы будет являться порождающим. Если $|\langle a \rangle| = \infty$, то образующими являются только элементы a и a^{-1}.

Теорема 3.10. *Пусть $G = \langle a \rangle$ циклическая группа и $(K, \cdot)$ – произвольная группа. В таком случае существует единственный гомоморфизм $f\colon G \to K$, при котором $f(a) = b$ для любого b, если $|\langle a \rangle| = \infty$ и для $b\colon b^n = e$, если $|\langle a \rangle| = n$.*

Доказательство. Так как при гомоморфизме $f(a^k) = f^k(a)$, то f однозначно задаётся своим значением $f(a)$. В случае бесконечной группы G утверждение теоремы становится очевидным. В случае конечной, в общем-то, тоже. □

Теорема 3.11. *Две любые группы одного и того же простого порядка изоморфны между собой.*

3.8 Разложение группы на смежные классы. Теорема Лагранжа

Определение. Пусть H – подгруппа в $\mathbf{G}$. Множества вида $gH = \{gh \mid h \in H\}$ при фиксированном $g \in \mathbf{G}$ называются *левыми смежными классами* группы $\mathbf{G}$ по подгруппе H. Аналогично определяются правые смежные классы.

Теорема 3.12. *Каждый левый смежный класс определяется любым своим элементом, т.е. если $g' \in gH$, то $g'H = gH$. Два левых смежных класса либо не пересекаются, либо совпадают. Объединение всех левых смежных классов равно $\mathbf{G}$. Отображение $H \to gH$ биективно.*

Доказательство. Пусть $g' \in gH$, т.е. $g' = gh, \quad h \in H$. Тогда $g'H = ghH \subset gH$; так как $g = g'h^{-1}$, то $g \in g'H$ и $gH \subset g'H$, из чего следует $gH = g'H$. Остальные утверждения теоремы выводятся не менее просто. □

Разбиение множества $\mathbf{G}$ на попарно непересекающиеся классы равносильно введению отношения эквивалентности на этом множестве.

Теорема 3.13. *Два элемента* $g_1, g_2 \in \mathbf{G}$ *принадлежат одному левому смежному классу по подгруппе* H *тогда и только тогда, когда* $g_1^{-1}g_2 \in H$.

Доказательство. $g_2 \in g_1H \quad \Leftrightarrow \quad g_2 = g_1h$ для некоторого $h \in H$. Следовательно, $g^{-1}g_2 \in H$. □

Теорема 3.14. *Отображение* $x \to x^{-1}$ *группы* $\mathbf{G}$ *на себя задаёт биективное соответствие между множествами левых и правых смежных классов по подгруппе* H.

Доказательство.

$$(gH)^{-1} = \{(gh)^{-1} \mid g \in H\} = \{h^{-1}g^{-1} \mid h \in H\} = \{hg^{-1} \mid h \in H\} = Hg^{-1}$$

□

Определение. Число (левых или правых) смежных классов группы $\mathbf{G}$ по подгруппе H называется *индексом* подгруппы H в $\mathbf{G}$ и обозначается $(\mathbf{G} : H)$.

Теорема 3.15 (Лагранжа)**.** *Пусть* $\mathbf{G}$ *– конечная группа. Тогда* $|H| \cdot (\mathbf{G} : H) = |G|$.

Доказательство. Если $|H| = d$, то для каждого элемента $h \in H$ имеем $(\mathbf{G} : H)$, например, левых смежных классов, которые не пересекаются между собой и при объединении дают $\mathbf{G}$. Таким образом утверждение теоремы очевидно. □

Примеры разбиения множества на смежные классы:

1. $G = (\mathbb{Z}, +), \quad H = n\mathbb{Z}$. Получаемые таким образом классы называются *классами вычетов*.

2. $G = (\mathbb{R}, +), \quad H = 2\pi\mathbb{Z}$. Смежные классы $\alpha + 2\pi\mathbb{Z}$ находятся в биективном соответствии с с углами, которые образуют векторы на плоскости с положительным направление оси абсцисс.

3. ...

4 Векторная алгебра

4.1 Основные понятия. Линейная комбинация векторов

Определение. *Вектор* – упорядоченный набор из чисел, называемых координатами вектора. Вектор – элемент арифметического n-мерного пространства $\mathbb{R}^n$. Над векторами определены следующие операции:

1. Сложение (означает почленное сложение его координат)

2. Умножение на скаляр

Основные свойства операций:

1. $\forall \mathbf{a}, \mathbf{b} \in \mathbb{R}^n \quad \mathbf{a}+\mathbf{b} = \mathbf{b}+\mathbf{a}$

2. $\forall \mathbf{a}, \mathbf{b}, \mathbf{c} \quad (\mathbf{a}+\mathbf{b})+\mathbf{c} = \mathbf{a}+(\mathbf{b}+\mathbf{c})$

3. $\exists \mathbf{0} \colon \mathbf{a}+\mathbf{0} = \mathbf{a}$;

4. $\forall \mathbf{a} \quad \exists (-\mathbf{a}) \colon \mathbf{a}+(-\mathbf{a}) = 0$

5. $\forall \alpha, \beta \in \mathbb{R}, \mathbf{a} \colon (\alpha\beta)\mathbf{a} = \alpha(\beta\mathbf{a})$

6. $(\alpha+\beta)\mathbf{a} = \alpha\mathbf{a}+\beta\mathbf{a}$

7. $\alpha(\mathbf{a}+\mathbf{b}) = \alpha\mathbf{a}+\alpha\mathbf{b}$

8. $\forall \mathbf{a} \in \mathbb{R}^n \quad 1\cdot\mathbf{a} = \mathbf{a}, \quad 0\cdot\mathbf{a} = \mathbf{0}$

Определение. Пусть даны $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_m \in \mathbb{R}^n$ и $\alpha_1, \alpha_2, \dots \alpha_m \in \mathbb{R}$. Тогда выражение $\alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \dots + \alpha_m\mathbf{a}_m \in \mathbb{R}^n$ назовём *линейной комбинацией* векторов $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$. Если $\alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \dots + \alpha_m\mathbf{a}_m = 0$, то говорят, что задано *линейное соотношение* этих векторов. Соотношение $0\cdot\mathbf{a}_1 + \dots + 0\cdot\mathbf{a}_m = 0$ называется *тривиальным. Нетривиальным* называется соотношение, в котором хотя бы один из коэффициентов не равен 0.

Определение. Под *системой векторов* понимается индексированная совокупность векторов, т.е. в системе могут содержаться равные между собой вектора, но наделённые разными индексами.

Определение. Конечная система векторов называется *линейно зависимой*, если для неё существует нетривиальное линейное соотношение. Если же существует только тривиальное линейное соотношение, то такая система называется *линейно независимой.* Пустая система векторов линейно независима.

Утверждение. Любая подсистема в линейно независимой системе линейно независима.

Утверждение. Система линейно зависима тогда и только тогда, когда хотя бы один вектор этой системы выражается через остальные.

Утверждение. Если какая-то подсистема системы зависима, то и вся система линейно зависима. *Другими словами, любую линейно зависимую систему можно расширить.*

Утверждение. Если система векторов линейно зависима, то и система укороченных векторов линейно зависима.

Утверждение. В пространстве $\mathbb{R}^n$ всякая система, содержащая больше, чем n векторов, линейно зависима.

Определение. Говорят, что вектор $\mathbf{b}$ линейно выражается через систему векторов $\{\mathbf{a}_i\}$, если $\exists \alpha_1, \dots \alpha_n \colon \mathbf{b} = \sum\limits_i \alpha_i \mathbf{a}_i$.

Говорят, что система $\mathbf{S} \subset \mathbb{R}^n$ линейно выражается через систему $\mathbf{T} \subset \mathbb{R}^n$, если каждый вектор из $\mathbf{S}$ линейно выражается через конечную подсистему в $\mathbf{T}$. Две системы векторов называются *эквивалентными*, если каждая из них линейно выражается через другую.

Теорема 4.1 (Основная лемма о линейной зависимости). *Пусть даны системы векторов* $\mathbf{A} = \{\mathbf{a}_1, \dots, \mathbf{a}_m\}, \mathbf{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_k\} \in \mathbb{R}^n$ *и система* $\mathbf{B}$ *линейно выражается через* $\mathbf{A}$. *Тогда если* $k > m$, *то* $\mathbf{B}$ *линейно зависима.*

Доказательство. Имеем систему уравнений:

$$\begin{cases} \mathbf{b}_1 & = & \sum\limits_i^m \alpha_{1i} \mathbf{a}_i \\ & \dots & \\ \mathbf{b}_k & = & \sum\limits_i^m \alpha_{ki} \mathbf{a}_i \end{cases}$$

Рассмотрим систему векторов $\{\lambda_j\}$, в которой i-й вектор будет иметь координаты $(\alpha_{i_1}, \dots \alpha_{i_m})$. Система этих векторов обязательно линейно зависима, потому что количество векторов в ней больше размерности пространства, которому они принадлежат. Таким образом мы всегда можем выбрать коэффициенты $\{\mu_j\}$, так, чтобы $\sum_j^k \mu_j \lambda_j = 0$. Понятно, что если мы возьмём линейную комбинацию $\sum_j^k \mu_j \mathbf{b}_j$, то она тоже окажется равной 0. Теорема доказана. □

Замечание. Другая формулировка этой леммы такова: «Линейно независимую систему систему нельзя выразить через меньшее количество векторов, чем она содержит».

4.2 Базис системы векторов

Определение. Пусть $\mathbf{S} \subset \mathbb{R}^n$ – любая система векторов. Набор векторов $\mathbf{a}_1, \dots \mathbf{a}_r$ называется *базисом* системы $\mathbf{S}$, если

1. $\mathbf{a}_1, \dots \mathbf{a}_r$ линейно независимы.

2. $\forall \mathbf{b} \in \mathbf{S}$ линейно выражается через $\mathbf{a}_1, \dots \mathbf{a}_r$.

Утверждение. Базис – максимальная линейно независимая подсистема. Всякую линейно независимую систему в $\mathbf{S}$ можно дополнить до базиса $\mathbf{S}$.

Утверждение. Любая система $\mathbf{S} \subset \mathbb{R}^n$ имеет базис. Стандартным базисом для $\mathbb{R}^n$ называется система векторов вида $\mathbf{e}_i = (\underbrace{0,\ldots,0}_{i-1},1,\underbrace{0,\ldots,0}_{n-i})$

Теорема 4.2. *Любые 2 базиса* $\mathbf{A}$ *и* $\mathbf{B}$ *системы содержат равное количество векторов.*

Доказательство. Пусть подсистемы $\mathbf{A} = \{a_1 \ldots, a_n\}$ и $\mathbf{B} = \{b_1, \ldots, b_m\}$ являются базисами, тогда каждая из них линейно выражается через другую. По основной лемме о линейной зависимости (п. 4.1, стр. 21) имеем, что $m \leqslant n$ и $m \geqslant n$, т.е. $m = n$. □

Определение. *Рангом* системы векторов называется число векторов любом её базисе. Обозначается $\operatorname{rk}\mathbf{S} = r$. Например, $\operatorname{rk}\mathbb{R}^n = n$.

Теорема 4.3. *Конечная подсистема* $(a_1, \ldots, a_r)$ *системы векторов* $\mathbf{a}$ *является базисом её базисом в том и только том случае, если всякий вектор из* $\mathbf{a}$ *выражается через* $(a_1, \ldots, a_r)$ *единственным образом.*

Доказательство. Докажем необходимость и достаточность:

($\Leftarrow$) Пусть $(a_1, \ldots, a_r)$ – базис и допустим, что имеются два представления какого-то вектора $b \in \mathbf{a}$:

$$b = \lambda_1 a_1 + \lambda_r a_r$$
$$b = \lambda_1' a_1 + \lambda_r' a_r$$

Тогда $0 = b - b = (\lambda_1 - \lambda_1')a_1 + (\lambda_r - \lambda_r')a_r \quad \Rightarrow \quad \forall i\colon \lambda_i = \lambda_i'$ так как $(a_1, \ldots, a_r)$ линейно независима и для неё может существовать только тривиальная линейная комбинация.

($\Rightarrow$) Предположим, что векторы из $\mathbf{a}$ единственным образом выражаются через систему $(a_1, \ldots, a_r)$. Чтобы установить, что подсистема $(a_1, \ldots, a_r)$ является базисом, нужно показать, что $(a_1, \ldots, a_r)$ – линейно независима. Допустим обратное, т.е. имеется линейное соотношение

$$\alpha_1 a_1 + \cdots + \alpha_r a_r = 0.$$

Рассмотрим представление какого-то вектора $b \in \mathbf{a}$:

$$b = \lambda_1 a_1 + \cdots + \lambda_r a_r$$

В силу единственности представления вектора b в виде линейной комбинации векторов $a_1, \ldots, a_r$ получаем, что $\forall i\colon \lambda_i = \lambda_i + \alpha_i \quad \Rightarrow \quad \alpha_i = 0$, т.е. $(a_1, \ldots, a_r)$ линейно независима.

□

Утверждение. Если $\mathbf{B}$ линейно выражается через $\mathbf{A}$, то $\operatorname{rk}\mathbf{B} \leqslant \operatorname{rk}\mathbf{A}$. *(Доказательство проводится по основной лемме о линейной зависимости (п. 4.1, стр. 21))*

4.3 Подпространства в $\mathbb{R}^n$

Определение. Подмножество $\mathbf{L} \subset \mathbb{R}^n$ называется подпространством, если

1. $\mathbf{L} \neq \varnothing$
2. $a, b \in \mathbf{L} \quad \Rightarrow \quad a + b \in \mathbf{L}$
3. $a \in \mathbf{L} \quad \Rightarrow \quad \forall\lambda \in \mathbb{R} \quad \lambda a \in \mathbf{L}$

Определение. Плоскость в n-мерном пространстве есть множество векторов, полученное сдвигом какого-то подпространства на несущий вектор.

Определение. Пусть имеется какая-то система векторов $\mathbf{S} = (a_1 \dots a_s) \in \mathbb{R}^n$. Тогда множество всех линейных комбинаций векторов из $\mathbf{S}$ называется линейной оболочкой множества $\mathbf{S}$ и обозначается $<\mathbf{S}>$.

$$<\mathbf{S}>=< a_1, \dots, a_r >= \{\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_s a_s \mid \forall \lambda_i \in \mathbb{R}\}$$

Линейная оболочка пустого множества $<\varnothing>$ равна нулевому вектору.

Сформулируем несколько следствий из определения:

- Линейная оболочка любого множества является подпространством
- Всякое подпространство является линейной оболочкой своего базиса.
- Условие, что система векторов $\mathbf{b}$ линейно выражается через систему векторов $\mathbf{a}$, равносильно тому, что $<\mathbf{b}> \subseteq <\mathbf{a}>$.
- Две системы векторов эквивалентны тогда и только тогда, когда из линейные оболочки совпадают.

Задание подпространства линейной оболочкой какой-либо системы векторов являтся только одним из возможных вариантов. Вторым способом является задание подпространства множеством решений ОСЛУ (см. п. 7.2, стр. 41). В данном случае можно, конечно, считать, что подпространство задаётся линейной оболочкой её фундаментальной системы решений (далее: ФСР).

Определение. В случае разговора о подпространствах, вместо термина «ранг» подпространства употребляют термин «размерность» подпространства. Обозначается $\dim \mathbf{L}$. Размерность подпространства равна его рангу как системы векторов.

5 Матрицы

5.1 Основные понятия

Определение. *Матрица* – прямоугольная таблица из чисел или любых других объектов.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{mn} \end{pmatrix}$$

Над матрицами определены следующие операции:

1. Умножение на скаляр.
2. Транспонирование. Строки записываются в столбцы. Обозначается A^T
3. Сложение матриц. Определено только для матриц одинакового размера, производится почленное сложение всех элементов матрицы.
4. Умножение матриц. Определяется следующим способом: пусть даны матрицы A размером $m \times n$ и B размером $n \times k$. Тогда матрица $C = AB$ будет состоять из элементов $c_{i_j} = \sum\limits_{k=1}^{n} a_{i_k} b_{k_j}$.

Так как матрицы являются обощением векторов, то операциям сложения и умножения на скаляр соответствуют все заявленные для векторов 8 свойств, некоторые другие свойства теории линейной зависимости также сохраняются.

Можно также доказать следующие свойства, относящиеся к вновь определённым операциям:

1. $(A+B)^T = A^T + B^T$
2. $(A+B)C = AC + BC$
3. $A(BC) = (AB)C$

   Доказательство. Определим сначала базис пространства матриц $M_{m\times n}(\mathbb{R})$. Он будет состоять из *матричных единиц* E_{ij} *(на пересечении i-й строки и j-го столбца стоит 1, а все остальные элементы равны 0.* Всякая матрица будет представима в виде линейной комбинации матричных единиц.

   Теперь осталось проверить только ассоциативность умножения матричных единиц. В общем случае
   $$E_{ij}E_{kl} = \begin{cases} E_{il}, & k = j \\ 0, & k \neq j \end{cases}.$$

   Легко проверяется, что умножение матричных единиц ассоциативно. Из этого следует, что и умножение любых матриц ассоциативно, так как ранее мы доказали свойство дистрибутивности. □

4. $(AB)^T = B^T A^T$

Определение. *Следом* матрицы $\operatorname{tr} A$ называется сумма диагональных элементов матрицы.

Докажем следующее важное свойство: $\operatorname{tr}(AB) = \operatorname{tr}(BA)$:

Доказательство. Проведём доказательство одной строкой:

$$\operatorname{tr}(AB) = \sum_i \sum_k a_{ik} b_{ki} = \sum_k \sum_i a_{ik} b_{ki} = \sum_k \sum_i b_{ki} a_{ik} = \operatorname{tr}(BA)$$

$\square$

Определение. Единичной матрицей называется такая матрица E, что $AE = EA = A$. Такая матрица имеет следующий вид (все элементы, кроме элементов главной диагонали, равны нулю):

$$A = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Определение. Если в строке матрицы все элементы – нули, то такая строка называется *нулевой*. Первый ненулевой элемент в строке называется *главным* или *лидером*.

Определение. *Элементарными преобразованиями (далее: ЭП) строк матрицы* называются преобразования вида:

1. К i-й строке прибавить j-ю, умноженную на $\lambda \in \mathbb{R}$
2. Поменять местами i-ю и j-ю строки
3. Умножить i-ю строку на $\lambda \neq 0$

Элементарные преобразования строк или столбцов матрицы можно записать как произведение исходной матрицы на одну из *элементарных матриц* (прямых или транспонированных). Если матрицу умножить на одну из элементарных матриц слева ($A' = E_i A$), то соответствующее преобразование затронет строки, если справа – столбцы.

$$E_1 = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & \dots & \lambda & \dots \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \quad E_2 = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \quad E_3 = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \lambda & & \\ & & & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

Определение. Матрица называется ступенчатой, если она удовлетворяет следующим условиям:

1. Ниже нулевой строки (если она есть) находятся только нулевые строки
2. В каждой ненулевой строке её лидер стоит строго правее, чем в предыдущей.

$$\tilde{A}' = \left(\begin{array}{cccccc|c} * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & * & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & * & \cdot & \cdot & \cdot \\ 0 & & & & & & \cdot \end{array}\right)$$

Теорема 5.1. *Любую матрицу можно привести к ступенчатому виду используя только ЭП 1 типа.*

Доказательство. Докажем теорему при помощи метода математической индукции. Индукцию проведём по количеству строк матрицы.

Для $m = 1$ утверждение верно, так как матрица, состоящая из одной строки ступенчатая.

Для удобства будем рассматривать матрицу без нулевых столбцов. Рассмотрим матрицу из m строк и её первую строку. Возможно две ситации:

1. $a_{11} \neq 0$. Тогда с помощью ЭП 1 типа мы можем сделать все остальные элементы этого столбца нулевыми, добавив к каждой (i-й) строке этой матрицы первую, умноженную на $-\frac{a_{i1}}{a_{11}}$.
2. $a_{11} = 0$. Так как мы исключили из рассмотрения все нулевые столбцы, то существует ненулевой элемент в первом столбце. Добавим к первой строке строку, содержащую этот элемент и реализуем алгоритм, описанный в пункте 1.

По предположению индукции подматрицу этой матрицы из $m - 1$ строки (со 2-й по последнюю) можно привести к ступенчатому виду, одновременно с этим и вся матрица приведётся к ступечатому виду. Теорема доказана. □

Определение. Матрица называется *сильноступенчатой* (приведённой к улучшенному ступенчатому виду), если над лидерами её строк находятся только нулевые элементы, а сами лидеры равны 1. Сильноступенчатый вид матрицы единственен.

Приведение матрицы к сильноступенчатому виду включено в алгоритм нахождения базиса любой системы векторов, а также нахождения ФСР ОСЛУ.

5.2 Ранг матрицы

Определение. *Рангом матрицы* называется ранг системы её столбцов.

Теорема 5.2 (о ранге матрицы). *Ранг системы столбцов матрицы равен рангу системы её строк и равен числу ненулевых строк в её ступенчатом виде.*

Доказательство. Для доказательства этой теоремы докажем несколько вспомогательных утверждений.

Лемма 1. При ЭП строк матрицы ранг системы стоблцов не меняется.

►. Если какая-то система столбцов была линейно зависима, то и преобразованная система столбцов остаётся линейно зависимой. Тогда сохраняется и линейная независимость столбцов. В таком случае базис исходной системы эквивалентен базису новой системы, таким образом они имеют одинаковое число векторов. Ранг системы векторов не изменяется. ◄

Лемма 2. При ЭП строк матрицы ранг системы строк не меняется.

►. Если система строк имела базис, то и любая система, состоящая из линейных комбинаций этих строк будет иметь такой же базис (как один из возможных вариантов). ◄

Приведём матрицу к сильноступенчатому виду. Теперь очевидно, что ранг системы строк равен числу ненулевых строк в ступенчатом виде (т.е. числу «ступенек» или главных элементов) и равен рангу системы столбцов. □

Очевидно, что если $C = AB$, то столбцы (строки) C являются линейной комбинацией столбцов A (строк B). В таком случае понятно, что если система столбцов A была линейно независима, то и система столбцов C будет также линейно независимой и если система строк B была линейно независима, то и система строк C будет линейно независимой.

Таким образом получаем $\operatorname{rk}(AB) \leqslant \min(\operatorname{rk} A, \operatorname{rk} B)$. Если B – невырожденная матрица, то $\operatorname{rk} A = \operatorname{rk}(AB)$.

6 Определители

6.1 Определение

Путь задана некоторая квадратная матрица A, каждый элемент которой есть элемент коммутативного кольца с единицей.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Тогда определителем матрицы называют следующую сумму по всем перестановкам n символов:

$$\det A = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \sum_{\sigma} (\operatorname{sgn} \sigma \cdot a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}), \quad \text{Где} \quad \sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$$

Определение. Матрица называется *вырожденной*, если её определитель равен 0 и *невырожденной* в обратном случае.

6.2 Свойства определителей

- $\det A^{\mathbf{T}} = \det A$. Доказательство заключается в том, чтобы выразить элемент $A^{\mathbf{T}}$ через элементы A:

$$\det A^{\mathbf{T}} \overset{\text{def}}{=} \sum_{\sigma} \operatorname{sgn} \sigma \cdot a^{\mathbf{T}}_{1\sigma(1)} a^{\mathbf{T}}_{2\sigma(2)} \cdots a^{\mathbf{T}}_{n\sigma(n)} = \sum_{\sigma} \operatorname{sgn} \sigma \cdot a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} =$$

$$= \sum_{\sigma} \operatorname{sgn} \sigma \cdot a_{1\sigma^{-1}(1)} a_{2\sigma^{-1}(2)} \cdots a_{n\sigma^{-1}(n)} = \left\{ \tau = \sigma^{-1};\ |\tau| = |\sigma| \right\} =$$

$$= \sum_{\tau} (-1)^{|\tau|} a_{1\tau(1)} a_{2\tau(2)} \cdots a_{n\tau(n)} = \det A$$

- При домножении какой-либо строки/столбца на λ, определитель матрицы тоже увеличивается в λ раз.

$$\begin{vmatrix} a_{11} & a_{22} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{i1} & \lambda a_{i2} & \cdots & \lambda a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \lambda \begin{vmatrix} a_{11} & a_{22} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Доказательство тривиально.

- Если каждый элемент одной строки/столбца разбит на сумму, то определитель матрицы разбивается на сумму определителей.

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a'_{i1}+a''_{i1} & \cdots & a'_{in}+a''_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a'_{i1} & \cdots & a'_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a''_{i1} & \cdots & a''_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

Для доказательства рассмотрим левую часть:

$$\sum_\sigma \operatorname{sgn}\sigma \cdot a_{1\sigma(1)}a_{2\sigma(2)}\cdots\left(a'_{i\sigma(i)}+a''_{i\sigma(i)}\right)\cdots a_{n\sigma(n)} =$$

$$= \sum_\sigma \operatorname{sgn}\sigma \cdot \left(a_{1\sigma(1)}a_{2\sigma(2)}\cdots a'_{i\sigma(i)}\cdots a_{n\sigma(n)} + a_{1\sigma(1)}a_{2\sigma(2)}\cdots a''_{i\sigma(i)}\cdots a_{n\sigma(n)}\right) =$$

$$\sum_\sigma \operatorname{sgn}\sigma\cdot a_{1\sigma(1)}a_{2\sigma(2)}\cdots a'_{i\sigma(i)}\cdots a_{n\sigma(n)} + \sum_\sigma \operatorname{sgn}\sigma\cdot a_{1\sigma(1)}a_{2\sigma(2)}\cdots a''_{i\sigma(i)}\cdots a_{n\sigma(n)} \overset{\text{def}}{=}$$

$$\overset{\text{def}}{=} \det A_1 + \det A_2$$

(A_1, A_2 — матрицы правой части)

6.3 Частные случаи при вычислении определителя

Теорема 6.1. *Если в матрице поменять местами две строки, то определитель поменяет знак.*

Доказательство. Пусть начальная матрица — A, и в ней поменяли местами i-ую и j-ую строки и получили матрицу $\tilde{A}$. Тогда имеют место следующие соотношения;

$$\begin{cases} \det A & \overset{\text{def}}{=} \sum\limits_\sigma \operatorname{sgn}\sigma\cdot a_{1\sigma(1)}\cdots a_{i\sigma(i)}\cdots a_{j\sigma(j)}\cdots a_{n\sigma(n)} \\ \det\tilde{A} & \overset{\text{def}}{=} \sum\limits_\sigma \operatorname{sgn}\sigma\cdot a_{1\sigma(1)}\cdots a_{i\sigma(j)}\cdots a_{j\sigma(i)}\cdots a_{n\sigma(n)} \end{cases}$$

Пусть подстановка σ имеет такой вид:

$$\sigma = \begin{pmatrix} 1 & \cdots & i & \cdots & j & \cdots & n \\ p_1 & \cdots & p_i & \cdots & p_j & \cdots & p_n \end{pmatrix}$$

Тогда введем подстановку τ:

$$\tau = \begin{pmatrix} 1 & \cdots & i & \cdots & j & \cdots & n \\ p_1 & \cdots & p_j & \cdots & p_i & \cdots & p_n \end{pmatrix}$$

То есть подстановку, отличающуюся от σ транспозицией (ij). Из этого следует, что $\operatorname{sgn}\sigma = -\operatorname{sgn}\tau$. Ясно, что количество всевозможных σ и τ совпадает, так как существует взаимооднозначное соответствие между ними. То есть нет разницы, суммировать по всем σ или по всем τ. Заменим тогда в $\det\tilde{A}$ подстановку:

$$\det\tilde{A} = \sum_\sigma \operatorname{sgn}\sigma\cdot a_{1\sigma(1)}\cdots a_{i\sigma(j)}\cdots a_{j\sigma(i)}\cdots a_{n\sigma(n)} =$$

$$= \sum_{\tau} (-1) \operatorname{sgn} \tau a_{1\tau(1)} \cdots a_{i\tau(i)} \cdots a_{j\tau(j)} \cdots a_{n\tau(n)} = -\det A$$

□

Теорема 6.2. *Определитель с двумя одинаковыми строками равен нулю*

Доказательство. Доказательство тривиально и основано на предыдущей теореме. □

Теорема 6.3. *Определитель матрицы не изменится, если к какой-либо ее строке добавить другую строку, умноженную на λ.*

Доказательство.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (a_{i1}+\lambda a_{j1}) & (a_{i2}+\lambda a_{j2}) & (a_{i3}+\lambda a_{j3}) & \cdots & (a_{in}+\lambda a_{jn}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & a_{j3} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \lambda \underbrace{\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \hline \vdots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \hline \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}}_{= \quad 0 \text{ (п. 6.2, стр. 30)}} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

□

Из этих частных случаев следует, что матрица является невырожденной тогда и только тогда, когда система её строк линейно независима.

6.4 Аксиоматический подход

Стоит отметить, что определитель однозначно определяется своими свойствами, а именно можно ввести функцию определителя и аксиоматически:

Определителем называется функция $\det\colon M_{n\times n}(\mathbb{R}) \to \mathbb{R}$, удволетворяющая следующим условиям:

1. (линейность)

К i-ой строке можно прибавить j-ую, домноженную на некоторое число $\lambda \in \mathbb{R}$. Также i-ую строку можно домножить на ненулевое число $\lambda \in \mathbb{R}$. При этом определитель не изменяется.

$$\begin{vmatrix} \dots \\ i \\ \dots \\ j \\ \dots \end{vmatrix} = \begin{vmatrix} \dots \\ i + \lambda j \\ \dots \\ j \\ \dots \end{vmatrix}$$

$$\begin{vmatrix} \dots \\ i \\ \dots \end{vmatrix} = \begin{vmatrix} \dots \\ \lambda i \\ \dots \end{vmatrix}$$

2. (кососимметричность)

   Если i-ую и j-ую строки поменять местами, то определитель изменит знак.

$$\begin{vmatrix} \dots \\ i \\ \dots \\ j \\ \dots \end{vmatrix} = - \begin{vmatrix} \dots \\ j \\ \dots \\ i \\ \dots \end{vmatrix}$$

3. (равенство нулю)

   Если две строки матрицы совпадают, то ее определитель равен нулю.

$$\begin{vmatrix} \dots \\ \alpha \\ \dots \\ \alpha \\ \dots \end{vmatrix} = 0$$

4. (нормировка)

   Определитель единичной матрицы равен единице.

$$\det E = 1$$

Покажем вывод развёрнутой формулы определителя по заданным условиям:

$$\det A = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11}\varepsilon_1 + a_{12}\varepsilon_2 + \cdots + a_{1n}\varepsilon_n \\ \cdots \\ a_{n1}\varepsilon_1 + a_{n2}\varepsilon_2 + \cdots + a_{nn}\varepsilon_n \end{vmatrix} =$$

$$= \{\text{раскрывая по линейности определителя}\} =$$

$$= \sum_{i_1,\cdots,i_n \in \{1,\cdots,n\}} \begin{vmatrix} a_{1i_1}\varepsilon_{i_1} \\ a_{2i_2}\varepsilon_{i_2} \\ \cdots \\ a_{ni_n}\varepsilon_{i_n} \end{vmatrix} = \sum_{i_1,\cdots,i_n \in \{1,\cdots,n\}} a_{1i_1} a_{2i_2} \cdots a_{ni_n} \begin{vmatrix} \varepsilon_{i_1} \\ \varepsilon_{i_2} \\ \cdots \\ \varepsilon_{i_n} \end{vmatrix} =$$

$$= \{\text{убираем нулевые слагаемые}\} = \sum_{\substack{i_1,\cdots,i_n \in \{1,\cdots,n\} \\ \forall \alpha,\beta \in \{1,\cdots,n\}:\ i_\alpha \neq i_\beta}} a_{1i_1} a_{2i_2} \cdots a_{ni_n} \begin{vmatrix} \varepsilon_{i_1} \\ \varepsilon_{i_2} \\ \cdots \\ \varepsilon_{i_n} \end{vmatrix} =$$

$$= \{\text{введем подстановку } \sigma \in \mathbf{S}_n\text{: } \begin{pmatrix} 1 & \cdots & n \\ i_1 & \cdots & i_n \end{pmatrix}\} = \sum_{\sigma \in \mathbf{S}_n} a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \begin{vmatrix} \varepsilon_{\sigma(1)} \\ \varepsilon_{\sigma(2)} \\ \cdots \\ \varepsilon_{\sigma(n)} \end{vmatrix}$$

Теперь остается понять, что матрица, составленная из строк $\varepsilon_{\sigma(1)}, \varepsilon_{\sigma(2)}, \cdots, \varepsilon_{\sigma(n)}$, есть единичная матрица со строками, переставленными в соответствии с подстановкой σ. То есть определитель этой переставленной матрицы равен $\operatorname{sgn}\sigma \cdot \det E$.

Тогда:

$$\det A = \sum_{\sigma \in \mathbf{S}_n} \operatorname{sgn}\sigma a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \underbrace{\det E}_{=\ 1}$$

6.5 Треугольная матрица

Треугольной матрицей называется матрица, имеющая следующий вид:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{23} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & a_{nn} \end{pmatrix}$$

Отсюда видно, что $\det A = a_{11}a_{22}a_{33}\cdots a_{nn}$, так как существует единственная (диагональная) подстановка, которая не содержит нулевых элементов.

Любую матрицу можно привести к треугольному виду (п. 5.1, стр. 26).

6.6 Разложение определителя по строке или столбцу

Лемма 3. [Об определителе блочной матрицы] Пусть задана матрица размером $n \times n$ следующего вида:

$$M = \left(\begin{array}{c|c} \underset{k\times k}{A} & * \\ \hline 0 & \underset{(n-k)\times(n-k)}{B} \end{array}\right)$$

A, B — некоторые подматрицы; $*$ — произвольные элементы; 0 — нулевые элементы. Тогда выполняется следующее равенство:

$$\det M = \det A \cdot \det B$$

Доказательство. Будем приводить матрицы A и B к треугольному виду. Очевидно, что эти же преобразования, применённые к исходной матрице приведут и её к треугольному виду. Тогда её определитель будет равен произведению элементов главной диагонали, то есть произведению диагональных элементов первой матрицы, умноженному на произведение диагональных элементов второй матрицы. Доказательство завершено. □

Теперь рассмотрим определитель матрицы размером $n \times n$.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Такое разложение можно получить, представив i-ую строку в виде суммы:

$$\begin{aligned} (a_{i1}, a_{i2}, a_{i3}, \cdots, a_{in}) &= (a_{i1}, 0, 0, \cdots, 0) + (0, a_{i2}, a_{i3}, \cdots, a_{in}) = \\ &= (a_{i1}, 0, 0, \cdots, 0) + (0, a_{i2}, 0, \cdots, 0) + (0, 0, a_{i3}, \cdots, a_{in}) = \\ &= (a_{i1}, 0, 0, \cdots, 0) + \cdots + (0, 0, 0, \cdots, a_{in}) \end{aligned}$$

Теперь в правой части из каждого определителя вынесем элемент i-ой строки:

$$\cdots = a_{i1} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + a_{i2} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + a_{in} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Теперь будем менять местами столбцы, передвигая j-ый столбец (с единицей в i-ой строке) на первое место. При этом знак определителя изменится $(i-1)+(j-1)$ раз. Таким образом:

$$\dots = (-1)^{(i-1)+(1-1)} \cdot a_{i1} \begin{vmatrix} 1 & 0 & \cdots & 0 \\ a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{a_{i1}} & \widehat{a_{i2}} & \cdots & \widehat{a_{in}} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} +$$

$$+ (-1)^{(i-1)+(2-1)} \cdot a_{i2} \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{12} & a_{11} & a_{13} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widehat{a_{i2}} & \widehat{a_{i1}} & \widehat{a_{i3}} & \cdots & \widehat{a_{in}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n1} & a_{n3} & \cdots & a_{nn} \end{vmatrix} + \cdots$$

$$\cdots + (-1)^{(i-1)+(n-1)} \cdot a_{in} \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{1n} & a_{11} & a_{12} & \cdots & a_{1(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widehat{a_{in}} & \widehat{a_{i1}} & \widehat{a_{i2}} & \cdots & \widehat{a_{i(n-1)}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{nn} & a_{n1} & a_{n2} & \cdots & a_{n(n-1)} \end{vmatrix} = \cdots$$

Назовем *минором* i-ой строки и j-ого столбца определитель матрицы, полученный «вычеркиванием» i-ой строки и j-ого столбца (обозначим за M_{ij}). По лемме о блочной матрице (п. 6.6, стр. 33) разложим определители в полученном выражении. Также заметим, что если к степени (-1) прибавить двойку, то ничего не изменится. Таким образом, заменяя соответствующие определители минорами, получаем.

$$\cdots = a_{i1} \cdot (-1)^{i+1} \cdot M_{i1} + a_{i2} \cdot (-1)^{i+2} \cdot M_{i2} + \cdots + a_{in} \cdot (-1)^{i+n} \cdot M_{in} + = \sum_{j=1}^{n} a_{ij} \cdot (-1)^{i+j} \cdot M_{ij} = \cdots$$

Назовем выражение $A_{ij} = (-1)^{i+j} \cdot M_{ij}$ *алгебраическим дополнением* к элементу a_{ij}. Тогда окончательно получаем формулу разложения определителя матрицы по строке:

$$\Delta = \cdots = \sum_{j=1}^{n} a_{ij} \cdot A_{ij}$$

С разложением по строке связана следующая лемма:

Лемма 4. Если в разложении матрицы по i-ой строке вместо элементов i-ой строки взять элементы j-ой строки, то получится нуль.

$$a_{i1} \cdot A_{i1} + \cdots + a_{in} \cdot A_{in} = \Delta$$

$$a_{i1} \cdot A_{j1} + \cdots + a_{in} \cdot A_{jn} = 0$$

Доказательство.

$$\Delta = a_{j1} \cdot A_{j1} + \cdots + a_{jn} \cdot A_{jn}$$

Так как A_{ji} не зависит от элементов j-ой строки, то можно «подставить» на j-ую строку элементы i-ой:

$$\Delta' = a_{i1} \cdot A_{j1} + \cdots + a_{in} \cdot A_{jn}$$

А $\Delta' = 0$, так как в матрице получились две одинаковые строки (п. 6.2, стр. 30). Значит, равенство верно. □

6.7 Определитель произведения матриц

Теорема 6.4. *Определитель произведения двух квадратных матриц размера $n \times n$ равен произведению их определителей.*

$$C = AB \quad \Rightarrow \quad \det C = \det A \det B$$

Лемма 5. Всякая невырожденная матрица представляется в виде произведения элементарных матриц. Всякая вырожденная матрица представляется в виде произведения элементарных матриц и матрицы, имеющей нулевую строку.

Доказательство.

1. Если A – элементарная матрица, то равенство очевидно.

2. Если A – невырожденная матрица, она может быть представлена в виде произведения элементарных матриц на единичную:

   $$A = U_n U_{n-1} \ldots U_1 E$$

   Тогда применяя применяя пункт 1 получаем, что $\det(AB) = \det(U_n \ldots U_1 B) = \det U_n \det(U_{n-1} \ldots U_1 B) = \cdots = \det U_n \ldots \det U_1 \det B = \det A \cdot \det B$

3. Если A – вырожденная матрица, то используя пп. 1 и 2 получаем, что

   $$\det(AB) = \det U_n \ldots \det U_1 \det(A'B),$$

   где A' имеет нулевую строку, а следовательно и $A'B$ имеет нулевую строку, $\det(AB) = \det A \det B = 0$.

□

6.8 Определитель Вандермонда. Интерполяция

Определитель Вандермонда $V(a_1, \dots a_n)$ — определитель следующей матрицы:

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{vmatrix}$$

Сделаем все элементы первого столбца, кроме самого первого, нулями. Для этого вычтем из n-й строки a_1 раз $n-1$-ю, из $n-1$-й — a_1 раз $n-2$-ю и т.д. По схеме вычисления определителя блочной матрицы, определитель исходной матрицы равен определителю следующей матрицы

$$\begin{vmatrix} a_2 - a_1 & a_3 - a_1 & \cdots & a_n - a_1 \\ a_2^2 - a_2a_1 & a_3^2 - a_3a_1 & \cdots & a_n^2 - a_na_1 \\ \vdots & \vdots & \ddots & \vdots \\ a_2^{n-1} - a_1a_2^{n-2} & a_3^{n-1} - a_1a_3^{n-2} & \cdots & a_n^{n-1} - a_1a_n^{n-2} \end{vmatrix} =$$

$$= (a_2 - a_1)(a_3 - a_1)\dots(a_n - a_1) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_2 & a_3 & \cdots & a_n \\ a_2^2 & a_3^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_2^{n-2} & a_3^{n-2} & \cdots & a_n^{n-2} \end{vmatrix} = \cdots = \prod_{i>j}(a_i - a_j)$$

Этот определитель тесно связан с интерполяцией функций. Действительно, через любые $n+1$ точек с разными абсциссами обязательно проходит многочлен степени не больше n, причём только один (по двум точкам можно построить единственную прямую, проходящую через них, по трём — единственную параболу и т.п.)

Пусть у нас имеется набор из n точек (x_i, y_i). Пусть искомый многочлен имеет вид

$$P(x) = a_0 + a_1x + \cdots + a_nx^n.$$

Нам известно, что $\forall i \quad P(x_i) = y_i$. Запишем эти условия в виде системы (n уравнений, $a_0, \dots a_n$ — неизвестные):

$$\begin{cases} a_0 + a_1x_1 + \dots + a_nx_1^n = y_1 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_0 + a_1x_{n+1} + \dots + a_nx_{n+1}^n = y_{n+1} \end{cases}$$

Эта система имеет единственное решение (см. выше).

Теперь найдём саму интерполированную функцию: введём многочлен

$$L_i(x) = \frac{(x-x_1)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_{n+1})}{(x_i-x_1)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_{n+1})}$$

Этот многочлен называется интерполяционным многочленом Лагранжа. $L_i(x_j) = 0$, $L_i(x_i) = 1$. Тогда искомая функция будет иметь вид

$$P(x) = \sum_{i=1}^{n+1} y_iL_i(x)$$

6.9 Обратная матрица

Определение. *Обратной матрицей* A^{-1} к матрице A называется такая матрица, что выполнено следущее равенство:

$$A^{-1}A = AA^{-1} = E$$

Где E — единичная матрица.

Замечание. Понятие обратной матрицы вводится в том случае, если элементы A — элементы некоторого поля.

Теорема 6.5. *Если существует обратная матрица, то она единственна.*

Доказательство. Пусть существуют различные две матрицы B_1 и B_2, обратные к A. Рассмотрим произведение B_1AB_2. В силу ассоциативности умножения матриц, имеем:

$$(B_1A)B_2 = B_1(AB_2) \quad \Rightarrow \quad EB_2 = B_1E \quad \Rightarrow \quad B_2 = B_1$$

Значит, они совпадают. Противоречие с тем, что $B_1 \neq B_2$. □

Теорема 6.6. *Невырожденность матрицы является необходимым и достаточным условием для существования обратной матрицы.*

Необходимость. Доказательство проведем от обратного. Пусть B — обратная матрица к A и $\det A = 0$. Тогда, по определению, $AB = BA = E$. По теореме об определителе произведение (п. 6.7, стр. 35), имеем равенство:

$$\det AB = \det E \quad \Rightarrow \quad \det A \det B = \det E$$
$$0 \cdot \det B = 1 \quad \Rightarrow \quad \det B = \frac{1}{0}$$

Такого в поле быть не может, поэтому предположение неверно. □

Достаточность. Для доказательства достаточности воспользуемся формулой обратной матрицы:

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}^T = \frac{1}{\det A} \begin{pmatrix} A_{11} & \cdots & A_{n1} \\ \vdots & \ddots & \vdots \\ A_{1n} & \cdots & A_{nn} \end{pmatrix}$$

Где A_{ij} — алгебраическое дополнение к элементу a_{ij} (см. п. 6.6, стр. 33).

Остается проверить, что эта матрица действительно является обратной к A. Обозначим произведение $AA^{-1} = X$. Тогда:

$$X = \frac{1}{\det A} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} A_{11} & \cdots & A_{n1} \\ \vdots & \ddots & \vdots \\ A_{1n} & \cdots & A_{nn} \end{pmatrix} = \frac{1}{\det A} \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix}$$

Найдем элементы x_{ij}:

$$
\begin{array}{rcll}
x_{11} & = & a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n} & = \det A \\
x_{12} & = & a_{11}A_{21} + a_{12}A_{22} + \cdots + a_{1n}A_{2n} & = 0 \\
 & \ldots & & \ldots \\
x_{21} & = & a_{21}A_{11} + a_{22}A_{12} + \cdots + a_{2n}A_{1n} & = 0 \\
x_{22} & = & a_{21}A_{21} + a_{22}A_{22} + \cdots + a_{2n}A_{2n} & = \det A \\
 & \ldots & & \ldots \\
x_{n(n-1)} & = & a_{n1}A_{(n-1)1} + a_{n2}A_{(n-1)2} + \cdots + a_{nn}A_{(n-1)n} & = 0 \\
x_{nn} & = & a_{n1}A_{n1} + a_{n2}A_{n2} + \cdots + a_{nn}A_{nn} & = \det A
\end{array}
$$

Таким образом:

$$
X = \frac{1}{\det A} \begin{pmatrix} \det A & 0 & 0 & \cdots & 0 \\ 0 & \det A & 0 & \cdots & 0 \\ 0 & 0 & \det A & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \det A \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} = E
$$

Аналогично проверяется и произведение $A^{-1}A$. □

Обратная матрица легко вычисляется с помощью элементарных преобразований: будем приводить исходную матрицу к единичной, применяя одновременно те же преобразования к единичной матрице. В тот момент, когда из исходной мы получим единичную, из единичной мы получим обратную.

$$
(A \mid E) \longrightarrow \left(E \mid A^{-1}\right)
$$

6.10 Характеризация ранга матрицы в терминах миноров

Теорема 6.7. *Ранг матрицы равен наибольшему порядку её отличных от нуля миноров*

Доказательство. Пусть ранг матрицы равен r. Тогда всякая система столбцов или строк, содержащая в себе больше, чем r векторов, будет линейно зависима. Тогда и укороченные вектора будут линейно зависимы, то есть миноры порядков $r + 1$ и больших будут равны нулю.

Теперь покажем, что существует минор порядка r, отличный от нуля. Выберем в исходной матрице r линейно независимых строки, ранг этой матрицы по строкам, а следовательно и по столбцам равен r. То есть мы можем выбрать в ней r линейно независимых столбцов, тем самым получаем квадратную подматрицу порядка r, ранг которой равен r, то есть определитель которой не равен нулю.

Доказательство завершено. $\square$

7 Системы линейных уравнений

7.1 Основные понятия. Метод Гаусса.

Определение. Система линейных уравнений (далее: СЛУ) имеет вид

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

Для удобства имеет смысл рассматривать не саму СЛУ, а матрицу её коэффициентов A или расширенную матрицу коэффициентов $\tilde{A}$.

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{mn} \end{pmatrix} \quad \tilde{A} = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \\ a_{n1} & a_{n2} & \ldots & a_{mn} & b_m \end{array}\right)$$

Также иногда имеет смысл рассматривать всю систему в матричном виде $AX = B$, где A – матрица коэффициентов, X – вектор-столбец неизвестных, B – вектор-столбец свободных членов. В таком случае произведение $A^{(i)}$ на X будет равно b_i и соответствовать i-му уравнению системы. Можно рассматривать произведение двух матриц как систему из нескольких систем линейных уравнений.

Определение. Под решением системы будем понимать набор $x^0 = (x_1^0, \ldots, x_n^0)$, такой, что при подстановке уравнения обращаются в верные равенства. Под понятием «*решить СЛУ*» следует подразумевать найти все её решения. Если СЛУ не имеет решений, то она называется *несовместной*. Если СЛУ имеет решения, то она называется *совместной*. В таком случае возможно 2 варианта: если существует только единственное решение, то такая система называется *определённой*, если же решений бесконечно много, то *неопределённой*.

Определение. Две СЛУ называются *эквивалентными*, если они имеют одинаковое множество решений.

Легко понять, что если любую СЛУ рассматривать как расширенную матрицу её коэффициентов, то при ЭП строк этой матрицы множество решений исходной СЛУ не меняется. (*Это прямо вытекает из обратимости ЭП: на каждом шаге множество решений системы не уменьшается*)

Тогда понятно, что от одной из двух эквивалентных СЛУ мы можем перейти к другой при помощи ЭП. Для решения СЛУ мы должны понять, какой вид матрицы её коэффициентов нам больше всего удобен. Оказывается, что наиболее удобно для решения СЛУ приводить матрицу её коэффициентов к так называемому ступенчатому виду.

После приведения расширенной матрицы коэффициентов, ассоциированной со СЛУ, к ступенчатому виду возможно несколько ситуаций:

- Есть строка, в которой все элементы, кроме самого последнего (столбца свободных членов), равны нулю. Такую строку назовём «*экзотической*». В этом случае СЛУ несовместна.
- Если же экзотических строк нет, то система совместна. Далее приведём алгоритм нахождения решений.

Неизвестные в СЛУ можно разделить на 2 части: главные и свободные так, что если свободным неизвестным придать произвольные значения, то существует единственное решение данной системы, в котором свободные неизвестные принимают эти значения.

Пусть расширенная матрица коэффициентов СЛУ приведена к ступенчатому виду и в ней нет экзотических строк. Тогда главными неизвестными будем считать те, которым соответствуют столбцы, содержащие лидеров ненулевых строк этой матрицы. Остальные неизвестные будем считать свободными. Опять же, возможно 2 ситации:

1. Все неизвестные – главные. Тогда рассмотрим последнюю ненулевую строку расширенной матрицы коэффициентов. В терминах СЛУ оно имеет следующий вид: $a_{nn}x_n = b_n$. Отсюда единственным образом выражается x_n. Подставляя его значение в предыдущее уравнение, получаем единственное выражение для x_{n-1} и т.д. В таком случае система является определённой.
2. Не все неизвестные – главные. Придадим свободным неизвестным произвольные значения и выражаем главные неизвестные через свободные.

Утверждение. Если в расширенной матрице коэффициентов некоторой СЛУ, приведённой к ступенчатому виду нет экзотических строк, то СЛУ совместна. Если при этом нет свободных неизвестных, то она определена.

Описанный алгоритм решения СЛУ называется *методом Гаусса* или *методом последовательного исключения неизвестных*.

7.2 Однородные СЛУ

Определение. СЛУ называется *однородной*, если все её свободные члены равны 0. Очевидно, что однородная СЛУ всегда совместна.

Теорема 7.1. *Однородная СЛУ, в которой число уравнений меньше числа неизвестных является неопределённой.*

Доказательство. Пусть число уравнений равно m и оно меньше числа неизвестных n. Тогда в ступенчатом виде матрицы коэффициентов число ненулевых строк меньше n, оно же равно числу главных неизвестных, тогда есть и свободные неизвестные, система неопределена («экзотических» строк, понятно, нет) □

Утверждение. Все решения неоднородной системы уравнений получаются из её одного фиксированного решения прибавлением всевозможных решений ассоциированной с ней однородной системы.

Если придавать этому геометрический смысл, то решение системы уравнений есть плоскость, полученная сдвигом подпространства решений ассоциированной ОСЛУ на несущий вектор.

Фундаментальная система решений

Определение. Пусть есть следующая ОСЛУ:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= 0 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= 0 \end{cases} \qquad (*)$$

U — множество ее решений. Оно образует подпространство в $\mathbb{R}^n$ ($\dim U = r$). Тогда всякий базис U называется *фундаментальной системой решений* данной определенной системы линейных уравнений.

Теорема 7.2 (о ФСР ОСЛУ). *Пусть дана система* $(*)$ *и* $r < n$. *Тогда у нее существует фундаментальная система решений и она содержит* $(n - r)$ *элементов.*

Доказательство. Приведем ОСЛУ к ступенчатому виду. Матрица содержит r ненулевых строк, а значит есть r независимых и $(n - r)$ свободных неизвестных. Для удобства возьмем, что $x_1, \cdots, x_r$ — главные неизвестные, а $x_{r+1}, \cdots, x_n$ — свободные. Тогда возьмем, например, следующую систему решений:

	x_1	...	x_r	x_{r+1}	...	x_n	
α_1 :	$*$	...	$*$	1	0	0	$(n - r)$ решений
α_i :	$*$	...	$*$	0	...	0	
α_{n-r}:	$*$	...	$*$	0	...	1	

Надо доказать, что $(\alpha_1, \cdots, \alpha_{n-r})$ является фундаментальной системой решений. Тогда, так как любая линейно независимая подсистема векторов может быть дополнена до базиса системы, все другие ФСР будут иметь то же количество решений.

1. $(\alpha_1, \cdots, \alpha_{n-r})$ — линейно независимы. Вычеркнем первые r компонент из каждого решения: полученная укороченная система векторов — единичная, а значит линейно независимая. Таким образом, $(\alpha_1, \cdots, \alpha_{n-r})$ — линейно независимы.

2. Докажем, что любое решение выражается через полученную систему. Пусть есть некоторое решение $\beta(\beta_1, \cdots, \beta_n)$. Рассмотрим тогда следующий вектор:

   $$x = \beta - \beta_{r+1}\alpha_1 - \beta_{r+2}\alpha_2 - \dots - \beta_n\alpha_{n-r} = (\underbrace{*, *, \cdots, *}_{r}, \underbrace{0, 0, \cdots, 0}_{n-r})$$

   x является нулевым решением исходной ОСЛУ: значения свободных неизвестных равны нулю; главные неизвестные по методу Гаусса выражаются через свободные члены и свободные неизвестные, а и те, и другие равны нулю.

   Тогда:

$$\beta = \beta_{r+1}\alpha_1 + \beta_{r+2}\alpha_2 + \cdots + \beta_n\alpha_{n-r}$$

То есть получили, что любое решение ОСЛУ выражается через систему решений.

Таким образом, $(\alpha_1, \cdots, \alpha_{n-r})$ — фундаментальная система решений. □

7.3 Критерии совместности и определённости

Теорема Кронеккера-Капелли

Теорема 7.3. *СЛУ совместна тогда и только тогда, когда* $\operatorname{rk} A = \operatorname{rk} \tilde{A}$. *СЛУ определённа тогда и только тогда, когда* $\operatorname{rk} A = \operatorname{rk} \tilde{A}$ *и равен числу неизвестных.*

Доказательство. Основано на приведении к ступенчатому виду и на теореме о ранге матрицы. □

Теорема Крамера

Теорема 7.4. *Квадратная система линейных уравнений является определённой в том и только том случае, если определитель её матрицы коэффициентов отличен от 0.*

Формулы Крамера Пусть дана СЛУ:

$$AX = B, \quad A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

Можно считать, что $\det A \neq 0$. Нулевой определитель матрицы коэффициентов обозначает, что система либо не имеет решений, либо их бесконечно много. Этот случай нужно рассматривать отдельно каждый раз.

Если же $\Delta = \det A \neq 0$, то СЛУ имеет единственное решение:

$$X = \underbrace{\frac{\operatorname{adj} A}{\det A}}_{=A^{-1}} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

Замечание. $\operatorname{adj} A$ — дополнительная матрица к A; транспонированная матрица алгебраических дополнений к элементам матрицы A.

$$\operatorname{adj} A = \begin{pmatrix} A_{11} & \cdots & A_{n1} \\ \vdots & \ddots & \vdots \\ A_{1n} & \cdots & A_{nn} \end{pmatrix}$$

Найдем i-ый элемент этой матрицы:

$$x_i = \frac{1}{\det A}\left(A_{1i}b_1 + A_{2i}b_2 + \dots + A_{ni}b_n\right)$$

Выражение, стоящее в скобках, является разложением определителя матрицы A по i-ому столбцу, только вместо элементов этого столбца стоят элементы матрицы B. То есть, это определитель матрицы A с i-ым столбцом, замененным на матрицу B.

Правило Крамера формулируется следующим образом:

$$x_i = \frac{\Delta_i}{\Delta}, \quad \Delta = \begin{vmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nn} \end{vmatrix}, \quad \Delta_i = \begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix}$$

8 Комплексные числа

8.1 Построение поля комплексных чисел

Полем комплексных чисел $\mathbb{C}$ будем называть поле, которое удовлетворяет следующим свойствам:

1. $\mathbb{R} \subset \mathbb{C}$
2. $\exists i \in \mathbb{C}\colon i^2 = -1$
3. $\mathbb{R} \subset \mathbf{L} \subset \mathbb{C} \quad \Rightarrow \quad \mathbf{L} = \mathbb{R}$ или $\mathbf{L} = \mathbb{C}$

Существует несколько различных моделей (реализаций) поля комплексных чисел:

1. Евклидова плоскость. Элементы поля – векторы на этой плоскости с началом в точке $(0;0)$.
2. $\mathbb{R}^2$ – множество пар двух действительных чисел, записывается как $(a;b)$ или $a+bi$.
3. $\left\{\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \mathbf{M}_2(\mathbb{R})\right\}$ – множество матриц такого вида.

Во всех трёх случаях операция сложения над элементами поля нам уже понятна, операция умножения для случая 3 тоже определена нами, откуда следует её определение и для случая 2. Умножение в случае 1 будет определено позже.

Необходимо проверить, что построенное нами множество с двумя операциями является полем. Удобнее всего это делать в случае реализации поля комплексных чисел как множества квадратных матриц специального вида.

Также необходимо проверить выполнение дополнительных условий на поле $\mathbb{C}$, наложенных нами. Покажем только выполенение условия 3 в этом случае: пусть $\mathbf{L} \supset \mathbb{R} \quad \Rightarrow \quad \exists z \in \mathbf{L} \setminus \mathbb{R}$, но так как $\mathbf{L} \subset \mathbb{C}$, то z представимо в виде $z = a + bi$. Так как $z \notin \mathbb{R} \quad \Rightarrow \quad b \neq 0$, то мы имеем право переписать в виде $i = (z-a)b^{-1} \in \mathbf{L}$. Имеем, что $\forall x, y \quad z = x + iy$ является элементом $\mathbf{L} \quad \Rightarrow \quad \mathbf{L} = \mathbb{C}$.

8.2 Тригонометрическая форма. Формула Муавра

Рассматривая комплексное число как вектор с координатами (a,b), можно сказать, что образует некоторый угол с положительным направлением оси Ox.

Определение. *Аргументом* комплексного числа z, $\arg z$, называется число $\varphi \in$ $\in [0, 2\pi)$, что $z = x + iy = |z|\cos\varphi + i|z|\sin\varphi$.

Обычно $|z|$ обозначается как ρ, и комплексное число записывается в форме $z = \rho$ $\rho(\cos\varphi + i\sin\varphi)$

Пример. $z = 2 + 2i = \sqrt{2}(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4})$

С тригонометрической формой записи связана и показательная (экспоненциальная) форма записи комплексного числа: $z = \rho(\cos\varphi + i\sin\varphi) = \rho e^{i\varphi}$.

Пример. $e^{i\pi} = -1$

Рассмотрим произведение двух комплексных чисел, записанных в тригонометрической форме. Пусть

$$z_1 = \rho_1(\cos\varphi_1 + i\sin\varphi_1)$$

$$z_2 = \rho_2(\cos\varphi_2 + i\sin\varphi_2)$$

Тогда

$$z_1 z_2 = \rho_1\rho_2(\cos\varphi_1 + i\sin\varphi_1)(\cos\varphi_2 + i\sin\varphi_2) = \dots =$$
$$= \dots = \rho_1\rho_2\left(\cos(\varphi_1+\varphi_2) + i\sin(\varphi_1+\varphi_2)\right)$$

Из этого соотношения вытекает, что

$$|z_1z_2| = |z_1||z_2|, \qquad \arg(z_1z_2) = \arg z_1 + \arg z_2$$

Аналогично можно получить, что

$$\frac{z_1}{z_2} = \frac{\rho_1}{\rho_2}\left(\cos(\varphi_1-\varphi_2) + i\sin(\varphi_1-\varphi_2)\right)$$

Утверждение. Из формулы для умножения двух комплексных чисел в тригонометрической форме следует, что

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

Эта формула называется формулой Муавра.

Из формулы Муавра можно вывести формулу для корня n-ной степени из комплексного числа, также называемой формулой Муавра.

Пусть $\sqrt[n]{z} = \omega \quad \Leftrightarrow \quad \omega^n = z, \quad z = \rho(\cos\varphi + i\sin\varphi), \quad \omega = \sigma(\cos\alpha + i\sin\alpha).$

Тогда

$$\omega^n = \sigma^n(\cos n\alpha + i\sin n\alpha) = \rho(\cos\varphi + i\sin\varphi) = z$$

$$\Updownarrow$$

$$\begin{cases} \rho &= \sigma^n \\ \cos n\alpha &= \cos\varphi \\ \sin n\alpha &= \sin\varphi \end{cases}$$

$$\Downarrow$$

$$\begin{cases} \sigma &= \sqrt[n]{\rho} \\ n\alpha &= \varphi + 2\pi k, \quad k \in \mathbb{Z} \end{cases}$$

То есть существует $n-1$ различных корней n-ной степени из z.

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

8.3 Корни из единицы в поле комплексных чисел

Множество корней n-ной степени из 1 в поле комплексных чисел μ_n согласно формуле Муавра будет иметь вид $\{\cos(2\pi k/n)+i\sin(2\pi k/n) \mid k=0,1,\dots,n-1\}$. Устройство множества корней n-ной степени из 1 в поле комплексных чисел тесно связано со многими другими алгебраическими понятиями.

Определение. *Первообразным* корнем n-ной степени из 1 называется такое число, которое не является корнем из 1 никакой меньшей степени. Можно показать, что ε_k является первообразным корнем n-ной степени из 1 тогда и только тогда тогда, когда $(k,n)=1$.

Число первообразных корней определяется функцией Эйлера

$$\varphi(n)=|\{k \mid 1\leqslant k<n, (k,n)=1\}|.$$

Если $n=p_1^{k_1}\cdot\dots\cdot p_s^{k_s}$, где p_i – различные простые числа, то

$$\varphi(n)=(p_1-1)p_1^{k_1-1}\cdot\dots\cdot(p_s-1)p_s^{k_s-1}.$$

8.4 Единственность поля $\mathbb{C}$

Теорема 8.1. *Пусть* $\mathbf{P}$ *– поле, содержащее* $\mathbb{R}$ *и такой элемент* j*, что* $j^2=-1$*. Тогда отображение* $f\colon\mathbb{C}\to\mathbf{P}$*, при котором* $f(a+bi)=a+bj$ *является гомоморфизмом, и это единственный гомоморфизм, при котором поле* $\mathbb{R}$ *отображается тождественно, а* i *переходит в* j*.*

Доказательство. Очевидно, что $f(z_1+z_2)=f(z_1)+f(z_2)$, так же легко проверить, что $f(z_1z_2)=f(z_1)f(z_2)$. Единственность очевидна. $\square$

То есть из этого следует, что всякое поле, удовлетворяющее исходным условиям, изоморфно полю комплексных чисел, причём существует изоморфизм, тождественный на поле действительных чисел. Поле $\mathbb{C}$ имеет ровно два автоморфизма, тождественных на $\mathbb{R}$: один тождественный, а второй переводит i в $-i$ (комплексное сопряжение).

9 Кольцо многочленов

9.1 Построение кольца многочленов

Пусть $\mathbf{K}$ – коммутативное кольцо с единицей, тогда обозначим через $\mathbf{S} = \mathbf{K}[x]$ кольцо многочленов от одной переменной, если оно удовлетворяет следующим требованиям:

1. $\mathbf{K} \subset \mathbf{S}$
2. $\exists \mathbf{x} \in \mathbf{S} \quad \forall f \in \mathbf{S} \quad f$ однозначно представляется в виде многочлена от $\mathbf{x}$ с коэффициентами из $\mathbf{K}$.

Определим некоторые тривиальные понятия, а также проведём некоторые тривиальные проверки:

1. $a_i x^i \cdot b_j x^j = a_i b_j x^{i+j}$
2. $\sum\limits_i a_i x^i \cdot \sum\limits_j b_j x^j = \sum\limits_{i,j} a_i b_j x^{i+j} = \sum\limits_k c_k x^k$, где $c_k = \sum\limits_{i+j=k} a_i b_j$.
3. ...

Определение. *Степенью* многочлена f называется наибольшее из чисел k, что одночлен $a_k x^k$ входит в представление f. Обозначается $\deg f$. Принято считать, что $\deg 0 = -\infty$, $\deg a = 0$.

Отметим, что свойства кольца $\mathbf{K}[x]$ напрямую зависят от свойств кольца $\mathbf{K}$, то есть если $\mathbf{K}$ является областью целостности, то и $\mathbf{K}[x]$ является областью целостности; например: если a неделитель нуля, то и $g = ax^n + \dots$ не является делителем нуля. В случае, если g, h – неделители нуля, то $\deg(gh) = \deg g + \deg h$.

Если в $\mathbf{K}$ нет делителей нуля, то обратимых неконстант в $\mathbf{K}[x]$ тоже нет.

9.2 Функциональный взгляд

Несмотря на то, что под многочленом мы прежде всего понимаем формальную запись, можно также составить функцию $f\colon \mathbf{K} \to \mathbf{K}$, такую, что для

$$\forall f = a_0 + a_1 x + \dots \in \mathbf{K}[x] \quad \forall c \in \mathbf{K} \quad f(c) = a_0 + a_1 c + \dots \in \mathbf{K}.$$

Очевидно, что $f(c) + g(c) = (f+g)(c)$ и $f(c)g(c) = (fg)(c)$. Стоит отметить, что если $\mathbf{K}$ – бесконечная область целостности, то равенство многочленов в алгебраическом и функциональном смысле равносильны. Составить контрпример для случая конечного поля не составит труда: в $\quad 0 \neq \prod\limits_i^n (x - c_i)$, но $\forall i \quad f(c_i) = 0$. В случае $\mathbb{Z}/p$ нетрудно вывести теорему Вильсона с помощью похожей конструкции: $(p-1)! \equiv -1 \mod p$.

Теорема 9.1 (Безу). *Если* $f(x) = (x-c)h(x) + r$, *то* $r = f(c)$.

Определение. Элемент $c \in \mathbf{K}$ называется *корнем* f, если $f(c) = 0$. Как следствие из теоремы Безу также получаем, что c – корень, если $x - c \mid f$. Элемент c называется корнем кратности k, если $(x-c)^k \mid f$, но $(x-c)^{k+1} \nmid f$.

Теорема 9.2. *Пусть* $\mathbf{K}$ *– область целостности. Тогда любой многочлен единственным образом представим в виде* $f(x) = (x - c_1)^{k_1} \ldots (x - c_s)^{k_s} \cdot h(x) \in \mathbf{K}[x]$*, где* $h(x)$ *не имеет корней в* $\mathbf{K}$.

Доказательство. Доказательство существования такого представления проводится по индукции и тривиально. Теперь к доказательству единственности. Тут тоже всё достаточно просто: пусть не так, тогда существует два представления:

$$(x - c_1)^{k_1} \ldots (x - c_s)^{k_l} \cdot h_1(x) = (x - d_1)^{l_1} \ldots (x - d_t)^{l_l} \cdot h_2(x)$$

За несколько операций нетрудно убедиться, что число различных корней и их кратности совпадают, а, следовательно, и $h_1 = h_2$. Необходимо использовать тот факт, что в области целостности нет делителей нуля. □

Из этой теоремы имеется важное **следствие**: число корней с учётом кратности не превосходит степени многочлена.

9.3 Теория делимости в кольце многочленов от одной переменной над полем

Теорема 9.3. *Пусть* $\mathbf{P}$ *– поле. Тогда для любых двух многочленов* f *и* $g \neq 0$ *существует единственное представление* $f = gh + r$*, где* $\deg r < \deg g$.

Математики очень любят говорить, что тот или иной факт тривиален.

Доказательство. Пусть $\deg g = n$, $\deg f = m \geqslant n$ (иначе тривиально: $f = g \cdot 0 + f$). Доказательство проведём по индукции по степени многочлена f. Пусть утвереждение теоремы верно для всех многочленов степени меньшей m.

Пусть $g = bx^n + \ldots$, b – обратим в $\mathbf{P}$ и $f = ax^m + \ldots$. Тогда рассмотрим разность $f - b^{-1}ax^{m-n}g = f_1, \quad \deg f_1 < \deg f$, то есть по индуктивному предположению $f_1 = gh_1 + r$ и $f = f_1 + b^{-1}ax^{m-n}g + r = g(h_1 + b^{-1}ax^{m-n}) + r = gh + r$.

Осталось выяснить, почему такое представление единственно. Пусть $f = gu_1 + v_1 = gu_2 + v_2$. Тогда $g(u_1 - u_2) = r_1 - r_2$. Если $u_1 \neq u_2$, то $\deg(v_1 - v_2) \geqslant \deg g$, чего не может быть, откуда следует, что $u_1 = u_2$ и $v_1 = v_2$. □

Пусть $\mathbf{A}$ – область целостности. Говорят, что $b \mid a$, если $\exists u \in \mathbf{A} \quad a = bu$. Если деление возможно, то оно единственно в силу отсутствия делителей нуля. Два элемента называются *ассоциированными* друг другу, если $b \mid a$ и $a \mid b$. Например, в поле $\mathbb{Z}$ ассоциирванными друг другу являются только числа ± 1. В кольце $\mathbf{K}[x]$ многочлены f и g являются ассоциированными, если $\exists \alpha \in \mathbf{K}^* = \mathbf{K} \setminus \{0\}$: $f = \alpha g$.

Продолжая исследовать многочлены в кольце $\mathbf{K}[x]$, введём понятие *неприводимого* многочлена:

Определение. Многочлен $f, \quad \deg f > 0$, называется *неприводимым*, если он не представим в виде произведения двух многочленов строго меньшей степени. Очевидно, что любой многочлен первой степени неприводим над любым полем. В дальнейшем же, приводимость одного и того же многочлена, но над разными полями может меняться.

Понятие неприводимого многочлена в кольце многочленов от одной переменной в некоторой степени является аналогом простого числа в поле целых чисел.

Для области целостности $\mathbf{A}$ элемент $a \in \mathbf{A}$ называется неприводимым, если его нельзя разложить в произведение необратимых. Область целостности называется *факториальной*, если для неё справедливо утверждение об однозначном разложении на неприводимые элементы. Примерами нефакториальных областей целостности являются $\mathbb{Z}[i\sqrt{5}]$ и другие.

Неприводимых многочленов над любым полем существует бесконечное количество. В случае бесконечного поля утверждение очевидно, а в случае конечного поля мы можем гарантировать существование неприводимых многочленов сколь угодно большой степени.

Теорема 9.4. *Всякий ненулевой многочлен в кольце многочленов над полем однозначно (с точностью до ассоциированности) представим в виде произведения неприводимых.*

Доказательство. Существование такого представления докажем индукцией по степени многочлена. В качестве основания индукции можно взять $\deg f = 0$ или $\deg f = 1$. Индуктивный переход тривиален.

Единственность такого представления докажем тоже по индукции. Пусть для всех многочленов, степени меньшей n верно; проверим для $f, \quad \deg f = n$. Пусть не так, и имеется два представления, занумеруем все неприводимые многочлены в порядке возрастания степени:

$$f = ap_1p_2\dots p_s \qquad f = bq_1q_2\dots q_t$$

Возможно 2 взаимоисключающих случая:

1. Среди $q_j \quad \exists j\colon p_1 \mid q_j \quad \Rightarrow \quad q_j = p_1 \cdot u$. Так как q_j неприводим, то u – константа из поля, а значит мы можем сократить на p_1 и получим два разложения для многочлена строго меньшей степени, что невозможно по индуктивному предположению.

2. Все q_i не делятся на p_1. Тогда разделим $q_j, \quad \deg q_j \geqslant \deg p_1$, на p_1 с остатком. Далее считаем, что $q_j = q_1$ (перенумеруем для удобства). Тогда имеем следующее:

   $$f = ap_1p_1\dots p_s = b(p_1u + r)q_2\dots q_t = bp_1uq_2\dots q_t + brq_2\dots q_t = m + h$$

   Так как $p_1 \mid f$ и $p_1 \mid m$, то $p_1 \mid h$. Отметим, что $\deg h < \deg f$. Тогда по индуктивному предположению представление h в виде произведния неприводимых однозначно, а значит в его разложении обязательно присутствует p_1: $h = p_1g$. Но с другой стороны, так как $\deg r < \deg p_1$ и ни один из q_i не делится на p_1, p_1 не может присутствовать в этом разложении.

   Получили противоречие, теорема доказана.

□

Аналогично определению кратных корней для многочлена определятся понятие кратного вхождения неприводимых многочленов в разложение.

Стоит отметить, что кольцо многочленов факториально не только над полем, но и над любым факториальным кольцом. Об этом будет сказано позже.

9.4 Наибольший общий делитель

Определение. Наибольшим общим делителем элементов f и g одновременно не равных нулю называется такой элемент d, что

1. $d \mid f$ и $d \mid g$
2. $\forall d'\colon d' \mid f, d' \mid g \quad \Rightarrow \quad d' \mid d$

Определение. *Наименьшим общим кратным* двух элементов f и g называется такой элемент m, что

1. $f \mid m$ и $g \mid m$
2. $\forall m'\colon f \mid m', g \mid m' \quad \Rightarrow \quad m' \mid m$

Теорема 9.5 (Алгоритм Евклида). *Пусть* $\mathbf{P}$ *– поле. Тогда для* $\forall (f,g) \neq (0,0)$, $f,g \in \mathbf{P}[x]$ *наибольший общий делитель* f *и* g *определятся из системы равенств* $(f,g) = \dots = (r_{s-1}, r_s) = (r_s, 0) = r_s$*, где*

$$\begin{cases} f &= gq_0 + r_1 \\ g &= r_1 q_1 + r_2 \\ r_1 &= r_2 q_2 + r_3 \\ & \dots \\ r_{s-2} &= q_{s-1} r_{s-1} + r_s \\ r_{s-1} &= q_s r_s + 0 \end{cases}$$

Доказательство. Доказательство этой теоремы проводится в два этапа: «снизу вверх», когда проверяется, что r_s действительно является делителем f и g и «сверху вниз», когда проверяется, что r_s – наибольший из всех делителей. □

Стоит отметить, что НОД существует для любых элементов в факториальном кольце. Можно доказать факториальность евклидовых колец, то есть колец, в которых возможно задать функцию *норма*, отвечающую свойствам $N(ab) \geqslant N(a)$ и $\forall a, b \quad \exists q, r\colon a = bq + r$ и $N(r) < N(b)$ или $r = 0$.

Следствием из алгоритма Евклида является тот факт, что наибольший общий делитель d элементов f и g представим в виде $d = fu + gv$. Более сильное утверждение (для многочленов) сообщает нам о следующем:

Теорема 9.6. *Для любых многочленов* f, g *существует единственное представление* $(f,g) = d = fu + gv$*, где* $\deg u < \deg g/d$ *и* $\deg v < \deg f/d$.

Доказательство. То, что существуют произвольные $\tilde{u}, \tilde{v}$, для которых выполенено $d = f\tilde{u} + g\tilde{v}$, нам уже известно. Пусть $\tilde{f} = f/d$, $\tilde{g} = g/d$. Тогда $1 = \tilde{f}\tilde{u} + \tilde{g}\tilde{v}$. Разделим $\tilde{u}$ на $\tilde{g}$ с остатком: $\tilde{u} = \tilde{g}q + u, \quad \deg u < \deg \tilde{g}$ и подставим это выражение в вышенаписанное равенство: $1 = \tilde{f}(\tilde{g}q + u) + \tilde{g}\tilde{v} = \tilde{f}u + \tilde{g}(q\tilde{f} + \tilde{v})$.

Осталось проверить, что $\deg v$, где $v = q\tilde{f} + \tilde{v}$, меньше $\deg \tilde{f}$. Действительно, так как $\tilde{g}v = 1 - \tilde{f}u$, то

$$\deg \tilde{g} + \deg v = \deg(\tilde{g}v) = \deg(1 - \tilde{f}u) < \deg(\tilde{f}\tilde{g}) = \deg \tilde{g} + \deg \tilde{f} \quad \Rightarrow \quad \deg v < \deg \tilde{f}$$

□

Теорема 9.7 (об остатках). *Пусть* $f_1, f_2, \ldots, f_s$ *и* $g_1, g_2, \ldots, g_s$ *– произвольные наборы многочленов с тем лишь условием, что* $\forall i \neq j \quad (f_i, f_j) = 1$. *Тогда существует такой многочлен* h*, что остатки от деления* h *на* f_i *равны* g_i.

Доказательство. Доказательство проведём, естественно, по индукции по числу многочленов. Для $n = 2$ верно:

$$h = q_1 f_1 + g_1 = q_2 f_2 + g_2 \quad \Leftarrow \quad g_2 - g_1 = q_1 f_1 - q_2 f_2$$

Последнее представление возможно, так как $(f_1, f_2) = 1$.

Пусть для $n = s - 1$ такой многочлен h_{s-1} существует, тогда, используя технологию доказательства для $n = 2$ можем обобщить от $n = s - 1$ к $n = s$, введя многочлен $\tilde{f}_2 = f_2 \cdot \cdots \cdot f_s$ и получив многочлен h_s: для многочленов $f_2, \ldots, f_s$ и $f_2, \ldots, g_s$ требуемый условием теоремы многочлен h_{s-1} существует. Тогда для многочленов f_1, f_2 и g_1, h_{s-1} существует такой многочлен $h_s = f_1 q_1 + g_1 = \tilde{f}_2 \tilde{q}_2 + h_{s-1}$. Этот многочлен удовлетворяет нашим требованиям. □

9.5 Многочлены над факториальными кольцами

Определение. *Содержанием* многочлена $c(f)$ будем называть наибольший общий делитель его коэффицинтов. Многочлены над факториальным кольцом называются *примитивными*, если его содержание равно 1, что равносильно тому, что содержание принадлжеит классу обратимых элементов кольца $\mathbf{K}$.

В кольце $\mathbf{K}[x]$ неразложимыми многочленами являются неприводимые примитивные многочлены.

Теорема 9.8 (лемма Гаусса). *Произведение примитивных – примитивно.*

Доказательство. Пусть

$$P(x) = a_k x^k + a_{k-1} x^{k-1} + \cdots + a_1 x + a_0, \quad (a_k, \ldots, a_0) = 1$$

$$Q(x) = b_l x^l + b_{l-1} x^{l-1} + \cdots + b_1 x + b_0, \quad (b_l, \ldots, b_0) = 1$$

Рассмотрим их произведение

$$P(x) \cdot Q(x) = a_k b_l x^{k+l} + (a_k b_{l-1} + a_{k-1} b^l) x^{k+l-1} + \\ + (a_k b_{l-2} + a_{k-1} b_{l-1} + a_{k-2} b_l) x^{k+l-2} + \cdots + a_0 b_0 = \sum_i^{k+l} c_i x^i, \quad c_i = \sum_{s+t=i} a_s b_t$$

Предположим, что $P(x)Q(x)$ — не примитивный. Тогда все $c_i \,\vdots\, p$. Пусть a_i — первый, не делящийся на p, а b_j — последний, не делящийся на p. Тогда рассмотрим коэффициент при x^{i+j}:

$$c_{i+j} = a_0 b_{i+j} + a_1 b_{i+j-1} \cdots + a_i b_j + \ldots a_{i+j} b_0 \quad (\text{с поправками, если } i > l \text{ или } j > k)$$

Тогда все слагаемые кроме $a_i b_j$ делятся на p, а само это слагаемое не делится на p. Значит $c_{i+j} \not\vdots\, p$. Противоречие. □

Теорема 9.9. *Если многочлен $f(x)$ неприводим над $\mathbf{K}[x]$, то он неприводим и над $\mathbf{K}(x)$.*

Доказательство. Пусть $f(x)$ — примитивный (иначе вынесем наибольший общий делитель его коэффициентов и рассмотрим полученный многочлен; приводимость многочлена над $\mathbf{K}(x)$ не изменится, если его умножить или разделить на константу), и он приводим над $\mathbf{K}(x)$, но неприводим над $\mathbf{K}[x]$

$$f(x) = \underbrace{g(x) \cdot h(x)}_{\text{коэфф. из } \mathbf{K}}$$

Приведём все коэффициенты в $g(x)$ и $h(x)$ к общему знаменателю, вынесем его и вынесем наибольший общий множитель коэффициентов в числителе.

$$f(x) = \frac{a}{b}\varphi(x) \cdot \frac{c}{d}\psi(x) = \frac{ac}{bd}\varphi(x)\psi(x), \quad \varphi(x), \psi(x) \text{ — примитивные}$$

$$bd \cdot f(x) = ac \cdot \varphi(x)\psi(x)$$

По лемме Гаусса (п. 9.8, стр. 52) имеем, что $bd = ac$, $\quad f(x) = \varphi(x) \cdot \psi(x)$. То есть мы получили, что $f(x)$ приводим над $\mathbf{K}[x]$. Противоречие. □

Теорема 9.10. *Кольцо многочленов над любым факториальным кольцом факториально.*

Доказательство. Индукцией по степени многочлена получаем, что требуемое разложение всегда существует. Докажем единственность такого представления. Пусть существует два представления многочлена f:

$$f = a_1 a_2 \dots a_k p_1(x) p_2(x) \dots p_s(x) = b_1 b_2 \dots b_l q_1(x) q_2(x) \dots q_t(x)$$

В силу факториальности кольца $\mathbf{K}$ и того, что произведение примитивных примитивно, имеем тот факт, что константы из $\mathbf{K}$ определены в разложении однозначно (с точностью до ассоциированности). Покажем теперь однозначность разложения многочленов. В силу предыдущей теоремы приводимость каждого из многочленов в записи не изменится, если рассматривать многочлены не над кольцом $\mathbf{K}$, а над его полем частных $\mathbf{F}$. Нам известно, что $\mathbf{F}[x]$ – факториально как кольцо многочленов, над полем, но это означает, что $\exists r_i, \nu_i \in \mathbf{K}\colon p_i(x) = \frac{r_i}{\nu_i} q_i(x)$. Домножив последнее равенство на $\nu_i p_i(x) = r_i q_i(x)$. Так как $p_i(x)$ и $q_i(x)$ примитивны, то получаем, что они равны с точностью до ассоциированности уже в данном случае над $\mathbf{K}[x]$. Теорема доказана. □

9.6 Многочлены на полем комплексных чисел

Кольцо многочленов $\mathbb{C}[x]$ является наиболее важным случаем полей / колец многочленов от одной переменной. Введём следующее понятие:

Определение. Поле $\mathbf{P}$ называется алгебраически замкнутым, если $\forall f \in \mathbf{P}[x]$, $\deg f > 0$, имеет корень в $\mathbf{P}$.

Теорема 9.11 (основная теорема алгебры)**.** *Поле $\mathbb{C}$ алгебраически замкнуто.*

Доказательство этой теоремы мы проведём несколько позже, а пока будем пользоваться важным следствием: $\forall f \in \mathbb{C}[x]$ представим в в виде $a_n \cdot \prod\limits_{i=1}^{s}(x - z_i)_i^k$, где z_i – корни этого многочлена, k_i – их кратности, а a_n – старший коэффициент многочлена.

Одной из самых интересных задач для нас пока является исследование корней произвольного многочлена $f \in \mathbb{C}[x]$. Первым шагом к этому будет изучение границы для модуля корня многочлена в $\mathbb{C}[x]$. Утверждение состоит в следующем:

Теорема 9.12. *Пусть* $f = a_n x^n + a_{n-1}x^{n-1} + \dots + a_1 x + a_0$ *и* z: $f(z) = 0$. *Тогда* $|z| < 1 + A$, *где* $A = \max\limits_{0 \leqslant i \leqslant n-1} \left|\frac{a_i}{a_n}\right| > 0$

Доказательство. Покажем, что если $|z| \geqslant 1 + A$, то z не может быть корнем f. Рассмотрим выражение:

$$\left|\frac{f(z)}{a_n}\right| = \left|z^n + \frac{a_{n-1}}{a_n}z^{n-1} + \dots + \frac{a_0}{a_n}\right| \geqslant |z^n| - \left|\frac{a_{n-1}}{a_n}z^{n-1} + \dots + \frac{a_0}{a_n}\right| \geqslant$$
$$\geqslant |z^n| - A\left(|z|^{n-1} + \dots + 1\right) = |z^n| - A\frac{|z|^n - 1}{|z| - 1} = \frac{|z|^n(|z| - (1 + A))}{|z| - 1} > 0$$

□

Определим операцию комлексного сопряжения для многочленов: если $f = a_n x^n + \ldots + a_0$, то $\overline{f} = \overline{a_n} x_n + \dots + \overline{a_0}$. Очевидно, что $\overline{f + g} = \overline{f} + \overline{g}$ и т.д. Верны также утвереждения, что если z – корень кратности k для f, то $\overline{z}$ – корень кратности k для $\overline{f}$.

Теперь, если у нас есть $f \in \mathbb{C}[x]$, то тот факт, что $f \in \mathbb{R}[x] \subset \mathbb{C}[x]$ будет означать нам не что иное, что $f = \overline{f}$. Обобщая всё вышесказанное имеем следующее утверждение:

Теорема 9.13. *Всякий многочлен* $f \in \mathbb{R}[x]$ *однозначно представим в виде произведения линейных множителей и квадратных трёхчленов с отрицательным дискриминантом.*

Доказательство. $f \in \mathbb{R}[x] \quad \Leftrightarrow \quad f = \overline{f}$. Выделив сначала действительные корни многочлена, его комплексные корни мы можем сгруппировать по парам z и $\overline{z}$. Теперь, если раскрыть скобки $(x - z)(x - \overline{z})$, то получим квадратный трёхчлен с действительными коэффициентами. Легко проверить, что его дискриминант отрицателен, если только не $z = \overline{z}$. □

9.7 Основная теорема алгебры

Теорема 9.14. *Пусть дано поле* $\mathbf{P}$ *и неприводимый многочлен* $f \in \mathbf{P}[x]$. *Тогда существует расширение поля* $\mathbf{L} \supset \mathbf{P}$, *в котором* f *имеет корень.*

Доказательство. Имеет смысл рассматривать подстановку в многочлен $f = a_n x^n + \ldots + a_1 x + a_0$ над полем $\mathbf{P}$ не элемента из поля $\mathbf{P}$, а матрицы элементов из поля $\mathbf{P}$. В качестве матрицы, которую мы будем подставлять возьмём так называемую сопровождающую матрицу многочлена, изначально приняв $a_n = 1$:

$$A_f = \begin{pmatrix} & & & & -a_0 \\ 1 & & & & -a_1 \\ & 1 & & & -a_2 \\ & & \dots & & \\ & & & 1 & -a_{n-1} \end{pmatrix}$$

Рассмотрим строение выражения $B = f(A_f)$. В частности заметим, что

$$A_f e_1 = e_2,$$

$$A_f^2 e_1 = A_f e_2 = e_3,$$

$$\dots$$

$$A_f^n e_1 = A_f e_n = -a_0 e_1 - a_1 e_2 - \dots - a_{n-1} e_n.$$

Тогда

$$\begin{aligned} f(A_f)e_1 = A_f^n e_1 + a_{n-1}A_f^{n-1}e_1 + \dots + a_1 A_f e_1 + a_0 e_1 = \\ = -a_0 e_1 - a_1 e_2 - \dots - a_{n-1}e_n + a_{n-1}e_n + \dots + a_1 e_2 + a_0 e_1 = 0 \end{aligned}$$

Аналогично $f(A_f)e_i = f(A_f)A_f^{i-1}e_1 = A_f^{i-1} \cdot f(A_f)e_1 = 0$. Коммутативность была применена в силу того факта, что в выражении присутствуют степени одной и той же матрицы.

Рассмотрим теперь все множество всех матриц $\mathbf{A}_f = \{g(A_f) \mid \forall g \in \mathbf{P}[x]\}$. Это множество уже образует кольцо; корень многочлена f в этом кольце есть. Осталось показать, что оно образует поле. Заметим, что в качестве многочленов g из $\mathbf{P}[x]$ мы можем рассматривать только те, чья степень строго меньше степени f. Действительно, если разделить g на f с остатком, имеем $g(A_f) = q(A_f)f(A_f) + r(A_f) = r(A_f), \quad \deg r < \deg f$.

Любой элемент из $\mathbf{A}_f$ представим в виде $h = \alpha_0 E + \alpha_1 A_f + \dots + \alpha_{n-1}A_f^{n-1}$, что соответствует некоторому многочлену из $\mathbf{P}[x]$. Покажем, что если $h \neq 0$, то h обратим в $\mathbf{A}_f$. Так как f неприводим в $\mathbf{P}[x]$, то $\exists u, v\colon fu + hv = 1$. Подставляя в это равенство A_f получаем $f(A_f) \cdot u(A_f) + h(A_f) \cdot v(A_f) = E$. Так как $f(A_f) = 0$ имеем $v(A_f) = h^{-1}(A_f)$.

Доказано, что $\mathbf{A}_f$ – поле, в котором f имеет корень. □

В качестве примера мы можем взять $f = x^2 + 1 \in \mathbb{R}[x]$. Для него $A_f = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. Такая матрица отвечает одному из способов построения поля комплексных чисел $\mathbb{C}$.

Очевидно, что мы можем построить поле, в котором любой многочлен будет представим в виде представления линейных множителей, может быть придётся расширять поле не один раз. Доказательство по индукции.

Теорема 9.15 (Основная теорема алгебры)**.** *Любой многочлен $f \in \mathbb{C}[x]$ представим в виде произведения линейных множителей.*

Доказательство. Рассмотрим сначала вопрос существования корней у произвольного многочлена $f \in \mathbb{R}[x]$. Пусть его степень $\deg f = 2^k m$, где m – нечётное. Докажем существование у него хотя бы одного комплексного корня. По индукции: для $k = 0$ утверждение верно, так как всякий многочлен нечётной степени с действительными коэфициентами имеет даже хотя бы один действительный корень.

По доказанной выше теореме существует расширение поля $\mathbf{L} \supset \mathbb{C}$, в котором f раскладывается на линейные множители. Пусть $x_1, x_2, \ldots x_n \in \mathbf{L}$ – его корни. Для любого $t \in \mathbb{R}$ построим многочлен $g(x)$, корнями которого будут являться элементы вида $u_{ij} = x_i + x_j + tx_ix_j$, их число равно $\frac{n(n-1)}{2} = 2^{k-1}m'$. Очевидно, что коэффициенты многочлена g являются элементарными симметрическими выражениями от корней f, то есть лежат в поле действительных чисел. По предположению индукции многочлен g имеет комплексный корень.

Это означает, что при любом выборе числа t можно указать такую пару индексов i и j, что элемент $x_i + x_j + tx_ix_j$ будет являться комплексным числом. Верно даже более сильное утверждение: $\exists t_1 \neq t_2 \in \mathbb{R}$, что для одних и тех же индексов i и j

$$\begin{cases} x_i + x_j + t_1x_ix_j & = & a \\ x_i + x_j + t_2x_ix_j & = & b \end{cases}, \quad a, b \in \mathbb{C}$$

Легко показать, что из этого следует, что $x_i + x_j$ и x_ix_j по отдельности также принадлежат полю $\mathbb{C}$, а значит элементы x_i и x_j являются корнями квадратного уравнения $\mathbf{x}^2 - (x_i + x_j)\mathbf{x} + x_ix_j = 0$ с комплексными коэффициентами, а значит являются комплексными числами. Утверждение для $f \in \mathbb{R}[x]$ доказано.

Для произвольного многочлена $f \in \mathbb{C}[x]$ рассмотрим следующую конструкцию: $F(x) = f(x) \cdot \overline{f}(x)$, коэффициенты многочлена $F(x)$ $b_k = \sum\limits_{i+j=k} a_i\overline{a_j}$. Заметим, что $\overline{b}_k = b_k$, то есть коэффициенты многочлена F – действительные числа, а значит $\exists \beta \colon f(\beta)\overline{f}(\beta) = 0$. Это означает, что f имеет своим корнем или β, или $\overline{\beta}$. Теорема доказана. □

9.8 Формальная алгебраическая производная

Пусть $f = a_nx^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0$. Тогда

$$f' \stackrel{\text{def}}{=} na_nx^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + 2a_2x + a_1$$

Легко (а может быть и не очень) доказать следующие свойства (доказательства проводятся с помощью обобщения верности этих свойств для одночленов на произволные многочлены):

1. $(f + g)' = f' + g'$
2. $(fg)' = f'g + g'f$
3. $(f(g(x)))'_x = f'_g(g(x)) \cdot g'_x$

Рассмотрим следующее применение производной: пусть $f \in \mathbf{P}[x]$, где $\mathbf{P}$ – поле; $f = \prod p_i^{k-i}$, p_i – неприводимые, и $k_i > 0$ – их кратности. Наше утверждение гласит следующее:

Теорема 9.16.

1. *Кратность* p *в* f' *больше либо равна* $k-1$

2. *Если* $\operatorname{char}\mathbf{P}=0$*, то кратность* p *в* f' *в точности равна* $k-1$

Доказательство. Пусть $f=p^k g, \quad p\nmid g$. Вычислим f' по правилам дифференцирования, описанным выше:

$$f'=kp^{k-1}p'g+p_kg'=p^{k-1}(kp'g+pg')$$

В случае $\operatorname{char}\mathbf{P}=0 \quad\Rightarrow\quad p'\neq 0$ □

Следтвием этого утверждения является тот факт, что в случае поля характеристики 0 f не имеет кратных множителей тогда и только тогда, когда $(f,f')=1$. В случае поля характеристики p это верно только в обратную сторону. Также очевидно, что при дифференцировании кратность корня убывает не более чем на единицу (строго на единицу в случае $\operatorname{char}\mathbf{P}=0$).

И ещё одним следствием нашего утвереждения является тот факт, что в случае поля характеристики 0 можно решить задачу построения многочлена, имеющего такие же корни, но кратности 1, не находя самих корней: если $f=\prod\limits_i p_i^{k_i}\cdot h$, то $f'=\prod\limits_i p_i^{k_i-1}\cdot h$, а значит $g=f/(f,f')$ имеет те же корни, что и f, но первой кратности.

9.9 Теорема Штурма

Пусть дан многочлен $f\in\mathbb{R}[x]$. Нас будет интересовать задача нахождения таких интервалов на числовой прямой, в которых содержится ровно один корень этого многочлена.

Определение. *Системой Штурма* для многочлена f на отрезке $[a;b]$ называется последовательность многочленов $f_0,f_1,\ldots,f_s$, обладающая следующими свойствами:

1. f_0 и f имеют на $[a,b]$ одни и те же корни без учёта кратности. То есть имеем право взять в качестве f_0 многочлен $f/(f,f')$. Здесь и далее будем считать, что f не имеет кратных корней, то есть $f=f_0$.

2. f_s не имеет корней на $[a,b]$

3. f_i и f_{i+1} не имеют общих корней на $[a,b]$ для $0\leqslant i\leqslant s-1$

4. Если $f_i(c)=0 \quad c\in[a;b]$, то $f_{i-1}(c)f_{i+1}(c)<0$ для $1\leqslant i\leqslant s-1$

5. Если $f(c)=0 \quad c\in[a;b]$, то $f\cdot f_1$ меняет знак с «–» на «+» при прохождении точки c слева направо.

Функция $\omega(x)$ будет возвращать нам целое неотрицательное число – число перемен знаков в последовательности значений многочленов из последовательности Штурма в точке x.

Теорема 9.17 (Штурма).

1. *Число корней многочлена* f *в полуинтервале* $(a;b]$ *(без учёта кратности) равно* $\omega(a)-\omega(b)$.

2. *для любого многочлена система Штурма существует, и будет указан алгоритм её нахождения.*

Доказательство.

1. Будем идти по всем точкам из полуинтервала $(a;b]$ слева направо. Заметим, что если в точке x ни один из $f_i, \quad 0 \leqslant i \leqslant s-1$ не имеет корня, то значение $\omega(x)$ не меняется. Теперь, если в точке x какой-то из многочленов $f_i, \quad 1 \leqslant \leqslant i \leqslant s-1$ имеет корень, то, так как $f_{i-1}(x)f_{i+1} < 0$, число перемен знаков в последовательности Штурма тоже не меняется.

   Если же теперь $f(x)=0$, то так как f и f_1 не имеют общих корней и $f \cdot f_1$ меняет знак с «–» на «+», то число перемен знаков при прхождении через точку x, являющуюся корнем f убывает ровно на 1.

2. В качестве системы Штурма можно взять следующую последовательность многочленов $f_0, \ldots, f_s$, определяемую системой равенств:

$$
\begin{gathered}
f_0 = f \\
f_1 = f_0' \\
f_0 = g_1 f_1 - f_2 \\
\ldots \\
f_{i-1} = g_i f_i - f_{i+1} \\
\ldots \\
f_s = -(f, f')
\end{gathered}
$$

   Легко проверить, что данная последовательность многочленов удовлетворяет свойствам системы Штурма.

□

9.10 Кольцо многочленов от нескольких переменных

Пусть $\mathbf{K}$ – коммутативное кольцо с единицей. Кольцом многочленов от n переменных над $\mathbf{K}$ называется такое кольцо $\mathbf{S}$, что

1. $\mathbf{K} \subset \mathbf{S}$

2. $\exists x_1, x_2, \ldots, x_n \in \mathbf{S}$, что любой многочлен $f \in \mathbf{S}$ однозначно представим в виде $\sum\limits_i a_{i_1} \ldots a_{i_n} x_1^{i_1} \ldots x_n^{i_n}$, $a_{ij} \in \mathbf{K}$.

Существование такого кольца можно провести индукцией по числу переменных, определяя $\mathbf{K}[x_1,\ldots,x_n]$ как $\mathbf{K}[x_1,\ldots,x_{n-1}][x_n]$.

Выражение $x_1^{i_1}\cdot x_2^{i_2}\cdot\cdots\cdot x_n^{i_n}$ называется *мономом*. Степенью монома называется число $i_1+i_2+\cdots+i_n$. Многочлен f называется однородным степени m, если все мономы в его представлении имеют степень m. Сумма и произведение однородных многочленов однородны.

На множестве мономов нам необходимо ввести отношение порядка. Логично упорядочить все мономы лексикографически, то есть при сравнения мономов смотреть на показатели степени при переменных $x_1,x_2,\ldots,x_n$; моном с более высоким показателем степени при x_1 будет считаться старше. Вполне очевидно, что высший член произведения двух многочленов равен произведению их высших членов.

9.11 Симметрические многочлены. Формулы Виета

Определение. Многочлен $f(x_1,x_2,\ldots,x_n)$ называется *симметрическим*, если для любой перестановки π выполнено $(\pi\circ f)(x_1,x_2,\ldots,x_n)=f(x_{\pi(1)},x_{\pi(2)},\ldots,x_{\pi(n)})$. Любая комбинация симметрических многочленов является симметрическим многочленом.

Вводятся *элементарные* симметрические многочлены

$$\sigma_i(x_1,x_2,\ldots,x_n)=\sum_{1\leqslant k_i\leqslant n}(x_{k_1}x_{k_2}\ldots x_{k_i})\,.$$

Легко показать, что, если $f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ и $x_1,\ldots,x_n$ – его корни, то связь между коэффициентами многочлена и корнями выражается как $a_{n-i}=(-1)^i\cdot\sigma_i(x_1,x_2,\ldots x_n)$. Такое выражение называется *формулами Виета*. Стоит отметить, что хоть и x_i не обязательно принадлежат тому кольцу, над которым построено кольцо многочленов, в котором лежит многочлен f, но значение элементарных симметрических многочленов от этих корней лежит в этом кольце.

Оказывается, что наиболее общим способом получения симметрических многочленов является подстановка в многочлен $g\in\mathbf{K}[y_1,y_2,\ldots,y_n]$ в качестве переменных y_i элементарных симметрических многочленов $\sigma_i\in\mathbf{K}[x_1,x_2,\ldots,x_n]$. Но ещё более удивителен тот факт, что ...

Теорема 9.18. *Для любого симметрического многочлена $f\in\mathbf{K}[x_1,x_2,\ldots,x_n]$ существует и притом единственный многочлен $g(\sigma_1,\sigma_2,\ldots,\sigma_n)=f(x_1,x_2,\ldots,x_n)$. Коэффициенты многочлена g являются целочисленными линейными комбинациями коэффициентов многочлена f.*

Доказательство.

1. Можно считать многочлен f однородным, иначе его можно единственным образом представить в виде суммы однородных многочленов различных степеней.

2. Высший член любого симметрического многочлена монотонен, то есть в представлении $u=a\cdot x_1^{i_1}x_2^{i_2}\ldots x_n^{i_n}$ имеем $i_1\geqslant i_2\geqslant\cdots\geqslant i_n$.

3. Пусть старший член f равен $u = a \cdot x_1^{i_1} x_2^{i_2} \dots x_n^{i_n}$. Рассматривая многочлен $f_1 = f - a \cdot \sigma_1^{i_1 - i_2} \sigma_2^{i_2 - i_3} \dots \sigma_n^{i_n}$, имеем $\deg f_1 = \deg f$, в f_1 не входит старший член f и его коэффициенты линейным образом целочисленно выражаются через коэффициенты f. Далее проделав точно такую же схему для f_1, получаем многочлен f_2, для которого получаем многочлен f_3 и так далее конечное число раз. Таким образом было получено требуемое разложение f через элементарные симметрические.

4. Докажем единственность такого представления. В случае существования двух различных многочленов $g_1(\sigma_1, \sigma_2, \dots, \sigma_n) = g_2(\sigma_1, \sigma_2, \dots, \sigma_n) = f$ существовал бы отличный от нуля многочлен $g = g_1 - g_2$, для которого $g(\sigma_1, \sigma_2, \dots, \sigma_n) = 0$. Для каждого из одночленов в представлении g при подстановке в него элементарных симметрических получаются разные старшие члены от $x_1, x_2, \dots, x_n$, то есть среди них обязательно имеется самый высший и $g(\sigma_1, \sigma_2, \dots, \sigma_n) \neq 0$.

□

9.12 Результант пары многочленов

Пусть дано два многочлена $f, g \in \mathbf{K}[x]$

$$
\begin{aligned}
f(x) &= a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0, \quad a_n = 0 \\
g(x) &= b_m x^m + b_{m-1} x^{m-1} + \dots + b_1 x + b_0, \quad b_m = 0
\end{aligned}
$$

За определение результанта возьмём выражение

$$
\operatorname{Res}(f, g) = \begin{vmatrix}
a_n & a_{n-1} & & \dots & a_0 & & & \\
& a_n & a_{n-1} & \dots & & a_0 & & \\
& \dots & \ddots & \dots & \dots & & \ddots & \\
& & & a_n & & & & a_0 \\
b_m & b_{m-1} & & \dots & b_0 & & & \\
& b_m & b_{m-1} & \dots & & b_0 & & \\
& \dots & \ddots & \dots & \dots & & \ddots & \\
& & & b_m & & & & b_0
\end{vmatrix}
$$

Свойства:

1. $\operatorname{Res}(f, g) = 0 \quad \Leftrightarrow \quad \deg(f, g) > 0$ (имеются общие корни)

   Доказательство. Пусть $\operatorname{Res}(f, g) = 0$. Тогда существует нетривиальная линейная комбинация строк в определителе матрицы из коэффициентов многочленов. За $\alpha_1, \dots, \alpha_{m+n}$ обозначим строки этого определителя, имеем $c_1\alpha_1 + \dots + c_{m+n}\alpha_{m+n} = 0$. Умножая такую строку на столбец $(x^{m+n-1} // \dots // 1)$ получаем выражение $f(x)(c_1 x^{m-1} + \dots + c_m) + g(x)(c_{m+1} x^{n-1} + \dots + c_{m+n}) = 0$ или $fu + gv = 0, \quad \deg u < m, \deg v < n$. Очевидно, что $u, v \neq 0$. Если бы $(f, g) = 1$, то т.к. $f \mid gv$, то $f \mid v$. Противоречие с тем, что $\deg v < \deg f$.

Пусть $(f,g) = h$, $\deg h > 0$. Тогда $f = hf_1$, $g = -hg_1$; отсюда $fg_1 + gf_1 = 0$, значит существует набор коэффициентов, при котором линейная комбинация строк данного нам определителя равна нулю, а значит и $\mathrm{Res}(f,g) = 0$. $\square$

2. $\mathrm{Res}(f,g) = a_0^m \prod\limits_{i=1}^{n} g(\alpha_i) = (-1)^{mn} b_0^n \prod\limits_{j=1}^{m} f(\beta_j) = a_0^m b_0^n \prod\limits_{i,j} (\alpha_i - \beta_j)$.

Доказательство. Рассмотрим $\mathrm{Res}(f, g-y) = (-1)^n a_0^m y^n + \dots + \mathrm{Res}(f,g)$, где y – некоторая новая переменная. Рассмотретим это выражение как многочлен степени n от y. Придадим y значение $g(\alpha_i)$. Многочлены $f(x)$ и $g(x) - g(\alpha_i)$ имеют общий корень, а значит делятся на $x - \alpha_i$, или $\mathrm{Res}(f, g - g(\alpha_i)) = 0$. В таком случае многочлен $\mathrm{Res}(f, g-y)$ должен делиться на все $g(\alpha_i) - y$, или $\mathrm{Res}(f, g-y) = a_0^m \prod\limits_{i=1} (g(\alpha_i) - y)$. При $y = 0$ получаем требуемое равенство. $\square$

3. Результант применим для решения систем полиномиальных уравнений вида

$$\begin{cases} F(x,y) &= 0 \\ G(x,y) &= 0 \end{cases}$$

9.13 Дискриминант многочлена

Рассмотрим многочлен $f(t) = a_n t^n + a_{n-1} t^{n-1} + \dots + a_1 t + a_0$. Он имеет ровно n корней в некотором расширении кольца, из которого взяты его коэффициенты. Дискриминантом этого многочлена по определению называется элемент $D(f) \overset{\text{def}}{=} \prod\limits_{1 \leqslant j < i \leqslant n} (x_i - x_j)^2$, лежащий в том же кольце, что и коэффициенты f, так как является симметрической функцией от корней многочлена. Нетрудно заметить, что значение дискриминанта равно квадрату значения определителя Вандермонда, составленного из корней многочлена, в свою очередь равного определителю, составленному из степенных сумм $s_0, s_1, \dots s_{2n-2}$:

$$D = \begin{pmatrix} s_0 & s_1 & \dots & s_{n-1} \\ s_1 & s_2 & \dots & s_n \\ \vdots & \vdots & \ddots & \vdots \\ s_{n-1} & s_n & \dots & s_{2n-2} \end{pmatrix}$$

Как известно, степенные формулы выражаются через элементарные симметрические по формулам Ньютона: для $s^k = x_1^k + x_2^k + \dots + x_n^k$ определено рекуррентное соотношение

$$s_k - \sigma_1 s_{k-1} + \sigma_2 s_{k-2} - \dots + (-1)^i \sigma_i s_{k-i} + \dots + (-1)^{k-1} \sigma_{k-1} s_1 + (-1)^k k \sigma_k = 0$$

Получим ещё одно выражение дискриминанта. Рассмотрим $\mathrm{Res}(f, f')$. Вспомним, что если $f = a_n(t - x_1)\dots(t - x_n)$, то $f' = a_n \sum\limits_{i=1}^{n} (t - x_1) \dots \widehat{(t - x_i)} \dots (t - x_n)$; очевидно, что $f'(x_i) = a_n(t-x_1)\dots\widehat{(t-x_i)}\dots(t-x_n)$. Тогда

$$\operatorname{Res}(f,f') = a_n^{n-1}\prod_{i=1}^{n} f'(x_i) = a_n^{2n-1}\prod_{i=1}^{n}(t-x_1)\dots\widehat{(t-x_i)}\dots(t-x_n) =$$
$$= a_n^{2n-1}\cdot(-1)^{\frac{n(n-1)}{2}}\cdot\prod_{1\leqslant j<i\leqslant n}(x_i-x_j)^2 = a_n^{2n-1}\cdot(-1)^{\frac{n(n-1)}{2}}\cdot D(f)$$

9.14 Поле рациональных дробей

Руководствуясь теоремой 3.6, вложим кольцо многочленов от одной переменной в поле. В этом поле определим некоторые понятия:

Определение. Дробь $\frac{f}{g}$ называется правильной, если $\deg f < \deg g$

Теорема 9.19. *Всякую неправильную дробь можно представить в виде «многочлен + правильная дробь»*

Определение. Простейшая дробь – дробь вида $\frac{f}{p^m}$, где p — неприводимый и $\deg f < \deg p$.

Теорема 9.20. *Всякую правильную дробь можно представить в виде суммы простейших.*

Доказательство. Доказательство, разумеется, проведём по индукции. Индукцию проведём по степени g. Для $\deg g = 1$ верно. Пусть верно для всех правильных робей, степень знаменателя которых меньше n. Возможно два случая:

1. g раскладывается в произведение двух взаимнопростых многочленов. Тогда по следствию из алгоритма Евклида и исходя из индуктивного предположения получаем, что требуемое представление существует.

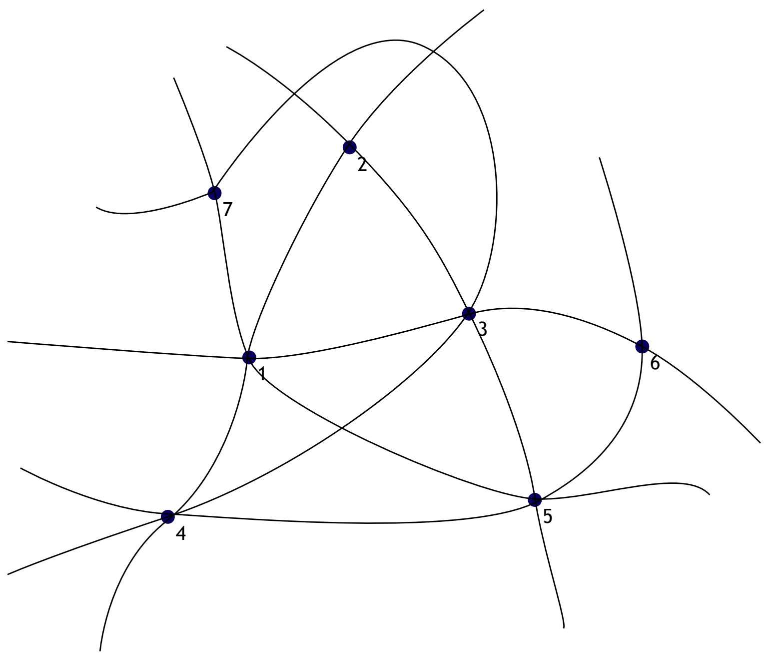
2. $g = p^k$. Всё равно разделим f на p с остатком, получим, что $f = qp + r$ или $\frac{f}{g} = \frac{q}{p^{k-1}} + \frac{r}{p^k}$, $\deg r < \deg p$. По индуктивному предположению опять получаем, что требуемое представление существует.

$\square$

Предметный указатель

FECI QUOD POTUI, FACIANT MELIORA POTENTES[1]

GAUDEAMUS IGITUR!



[1] Я сделал, что мог; кто может, пусть сделает лучше (*лат.*)