Прохоров Ю. Г.

Лекции по алгебре

Семестр 1

Москва 2023

Разъяснение. Настоящие записки появились в процессе подготовки к лекциям. Они нисколько не претендуют на роль учебника.

Оглавление

Лекция 1

Системы линейных уравнений. Матрицы. Метод Гаусса. Решение систем линейных уравнений. Следствия.

Системы линейных уравнений

Системой линейных уравнений называется система вида

$$(*) \qquad \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,m}x_m &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,m}x_m &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,m}x_m &= b_n \end{cases}$$

где $a_{i,j}$ и b_i – известные константы, а x_i – неизвестные переменные. Краткая запись этой системы выглядит следующим образом:

$$\begin{cases} \sum\limits_{j=1}^{m} a_{1,j}x_j &= b_1 \\ \sum\limits_{j=1}^{m} a_{2,j}x_j &= b_2 \\ \cdots\cdots\cdots\cdots \\ \sum\limits_{j=1}^{m} a_{n,j}x_j &= b_n \end{cases}$$

Числа $a_{i,j}$ называются коэффициентами, а числа b_i – свободными членами системы. Система называется *квадратной*, если число ее уравнений равно числу неизвестных (т.е. $n = m$). Система называется *однородной*, если все ее свободные члены равны нулю: $b_j = 0$ для всех j.

Решением системы называется набор чисел $(c_1, \ldots, c_m)$, которые будучи подставлены в систему $(x_1 = c_1, \ldots, x_m = c_m)$ дают правильные равенства. Отметим, что решение системы – это *упорядоченный* набор чисел. Система называется *совместной*, если у нее имеется по крайней мере одно решение. Совместная система называется *определенной*, если ее решение единственно. Отметим, что однородная система обязательно совместна, поскольку она имеет нулевое решение $x_1 = \cdots = x_m = 0$.

Говорят, что две системы линейных уравнений от одинакового числа неизвестных *эквивалентны*, если множества их решений совпадают.

Далее мы изложим метод решения систем линейных уравнений, который называется *методом Гаусса*. Для этого удобно вместо системы $(*)$ ее коэффициенты и свободные члены собрать в таблицу, называемую *матрицей*.

Матрицы

Матрицей называется таблица

$$C = \begin{pmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,m} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n,1} & c_{n,2} & \cdots & c_{n,m} \end{pmatrix}$$

Размером этой матрицы считается $n \times m$ (число строк $\times$ число столбцов). Положение каждого элемента матрицы $c_{i,j}$ задается двумя числами (i, j) (номер строки, номер столбца). Краткая запись матрицы: $C = (c_{i,j})$. Мы рассматриваем также матрицы размера $1 \times m$ (строки) и размера $n \times 1$ (столбцы). Матрица размера $n \times n$ называется *квадратной* порядка n.

Система $(*)$ однозначно задается матрицей

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} & b_n \end{pmatrix}$$

которая называется *расширенной матрицей* системы. Матрица

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix}$$

получаемая из A отбрасыванием последнего столбца, называется *матрицей коэффициентов* или просто матрицей системы.

Элементарные преобразования

Определим преобразования типов $(\mathrm{I}_{\mathrm{i,j},\lambda})$, $(\mathrm{II}_{\mathrm{i},\lambda})$, $(\mathrm{III}_{\mathrm{i,j}})$ (здесь $1 \le i,\ j \le n$, а λ – некоторое число). Эти преобразования систему $(*)$ переводят в систему

$$(\dagger) \qquad \begin{cases} a'_{1,1}x_1 + a'_{1,2}x_2 + \cdots + a'_{1,m}x_m &= b'_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a'_{n,1}x_1 + a'_{n,2}x_2 + \cdots + a'_{n,m}x'_m &= b'_n \end{cases}$$

Тип (I)**.**

При преобразовании $(\mathrm{I}_{\mathrm{i,j},\lambda})$ все уравнения (†), кроме i-го имеют тот же вид, а i-е уравнение в (†) принимает вид

$$(a_{i,1}+\lambda a_{j,1})x_1+(a_{i,2}+\lambda a_{j,2})x_2+\cdots+(a_{i,m}+\lambda a_{j,m})x_m=b_i+\lambda b_j$$

Иначе говоря, к i-му уравнению прибавляется j-е, умноженное на λ. Здесь считается, что $i\neq j$.

Тип (II)**.**

При преобразовании $(\mathrm{II}_{\mathrm{i},\lambda})$ все уравнения (†), кроме i-го имеют тот же вид, а i-е уравнение в (†) принимает вид

$$\lambda a_{i,1}x_1+\lambda a_{i,2}x_2+\cdots+\lambda a_{i,m}x_m=\lambda b_i$$

Иначе говоря, к i-е уравнение умножается на λ. Считается, что $\lambda\neq 0$.

Тип (III)**.**

При преобразовании $(\mathrm{III}_{\mathrm{i,j}})$ все уравнения (†), кроме i-го и j-го имеют тот же вид, а i-е и j-е уравнения меняются местами. Считается, что $i\neq j$.

Преобразования $(\mathrm{I}_{\mathrm{i,j},\lambda})$, $(\mathrm{II}_{\mathrm{i},\lambda})$, $(\mathrm{III}_{\mathrm{i,j}})$ называются *элементарными преобразованиями* системы линейных уравнений.

Замечание. Все элементарные преобразования обратимы. Действительно, обратным к $(\mathrm{I}_{\mathrm{i,j},\lambda})$ является $(\mathrm{I}_{\mathrm{i,j},-\lambda})$, обратным к $(\mathrm{II}_{\mathrm{i},\lambda})$ является $(\mathrm{II}_{\mathrm{i},1/\lambda})$, а обратное преобразование к $(\mathrm{III}_{\mathrm{i,j}})$ совпадает с ним самим.

Замечание. Преобразования типа (III) могут быть получены последовательным применением преобразований типа (I) и (II). Действительно, $(\mathrm{III}_{\mathrm{i,j}})$ получается как композиция $(\mathrm{I}_{\mathrm{i,j},1})$, $(\mathrm{I}_{\mathrm{j,i},-1})$, $(\mathrm{I}_{\mathrm{i,j},1})$ и $(\mathrm{II}_{\mathrm{j},-1})$. (Проверьте!)

Задача. Можно ли преобразования типа (III) получить последовательным применением только преобразований типа (I)?

Предложение. *Если от системы* $(*)$ *можно перейти к системе* (†) *при помощи элементарных преобразований, то эти системы эквивалентны.*

Доказателъство. Достаточно доказать наше утверждение для одного преобразования типа (I) или (II). Для (II) утверждение очевидно. Пусть (†) получается из $(*)$ применением одного преобразования типа $(\mathrm{I}_{\mathrm{i,j},\lambda})$. Пусть M (соответственно, M') – множество решений системы $(*)$ (соответственно, (†)). Если $M=M'=\varnothing$, то доказывать нечего. Поэтому можно считать, что $M\neq\varnothing$. Пусть $(c_1,\ldots,c_m)\in M$. Это означает, что

$$a_{k,1}c_1+a_{k,2}c_2+\cdots+a_{k,m}c_m=b_k,\quad\forall k.$$

Так как $a'_{k,l} = a_{k,l}$ и $b'_k = b_k$ при $k \neq i$, то

$$a'_{k,1}c_1 + a'_{k,2}c_2 + \cdots + a'_{k,m}c_m = b'_k, \quad \forall k \neq i.$$

Для $k = i$ имеем $a'_{i,l} = a_{i,l} + \lambda a_{j,l}$ и $b'_i = b_i + \lambda b_j$. Отсюда

$$\begin{aligned} a'_{i,1}c_1 + a'_{i,2}c_2 + \cdots + a'_{i,m}c_m &= \\ (a_{i,1} + \lambda a_{j,1})c_1 + \cdots + (a_{i,m} + \lambda a_{j,m})c_m &= \\ a_{i,1}c_1 + \cdots + a_{i,m}c_m + \lambda(a_{j,1}c_1 + \cdots + a_{j,m}c_m) &= \\ b_i + \lambda b_j &= b'_i. \end{aligned}$$

Таким образом, $(c_1, \ldots, c_m) \in M'$ и поэтому $M' \subset M$. Обратное включение следует из обратимости элементарных преобразований. □

Элементарные преобразования матриц

Пусть $A_1, \ldots, A_n$ – строки матрицы A, т.е. A_i – это матрица вида

$$A_i = (a_{i1},\, a_{i,2}, \ldots, a_{i,m})\,.$$

Для строк определены операции сложения и умножения на число:

$$\begin{aligned} A_i + A_j &:= (a_{i1} + a_{j1},\, a_{i,2} + a_{j,2}, \ldots, a_{i,m} + a_{j,m}) \\ \lambda A_i &:= (\lambda a_{i1},\, \lambda a_{i,2}, \ldots, \lambda a_{i,m})\,. \end{aligned}$$

Иначе говоря, операции сложения и умножения строк на число выполняются *покомпонентно*.

Говорят, что матрица A' получена из матрицы A *элементарным преобразованием строк* типа $(\mathrm{I}_{\mathrm{i,j},\lambda})$, $(\mathrm{II}_{\mathrm{i},\lambda})$, $(\mathrm{III}_{\mathrm{i,j}})$ $(1 \leq i, j \leq n$, λ – некоторое число), если матрица A' тот же размер, что и A, а ее строки $A'_1, \ldots, A'_n$ имеют вид:

Преобразование $(\mathrm{I}_{\mathrm{i,j},\lambda})$**:** $A'_i = A_i + \lambda A_j$ и $A'_k = A_k$ при $k \neq i$ (считается, что $i \neq j$);

Преобразование $(\mathrm{II}_{\mathrm{i},\lambda})$**:** $A'_i = \lambda A_i$ и $A'_k = A_k$ при $k \neq i$ (считается, что $\lambda \neq 0$);

Преобразование $(\mathrm{III}_{\mathrm{i,j}})$**:** $A'_i = A_j$, $A'_j = A_i$ и $A'_k = A_k$ при $k \neq i, j$ (считается, что $i \neq j$).

Аналогично можно определить элементарные преобразования *столбцов*.

Определение. Говорят, что матрица

$$C = \begin{pmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,m} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n,1} & c_{n,2} & \cdots & c_{n,m} \end{pmatrix}$$

имеет *ступенчатый вид*, если она удовлетворяет следующим двум условиям:

(1) Ниже нулевой строки находятся только нулевые строки.

(2) Пусть $1 \le i < k \le n$ и пусть $c_{i,j}$ и $c_{k,l}$ – первые ненулевые элементы строк с номером i и k, соответственно. Тогда $l > j$. Иначе говоря, первый ненулевой элемент строки располагается строго правее первого ненулевого элемента любой более верхней строки.

Говорят, что матрица C имеет *улучшенный ступенчатый вид*, если она дополнительно к (1) и (2) удовлетворяет также следующему условию:

(3) Пусть $c_{i,j}$ – первый ненулевой элемент строки с номером i. Тогда $c_{i,j} = 1$ и $c_{r,j} = 0$ для всех r. Иначе говоря, первый ненулевой элемент строки равен 1 и все остальные элементы в столбце, содержащем данный элемент, равны 0.

Метод Гаусса

Теорема (Алгоритм Гаусса). (1) *Любая матрица C элементарными преобразованиями строк типа* (I) *приводится к матрице C', имеющей ступенчатый вид.*

(2) *Любая матрица C элементарными преобразованиями строк типа* (I) *и* (II) *приводится к матрице C', имеющей улучшенный ступенчатый вид.*

Доказателъство. (1) Индукция по количеству столбцов m. База индукции $m = 0$ очевидна. Предположим, что наше утверждение верно для всех $m' < m$. Можно считать, что первый столбец $\neq 0$, т.е. $c_{i,1} \neq 0$ для некоторого i. Если $c_{1,1} = 0$, то преобразованием $(\mathrm{I}_{1,i,1})$ добиваемся того, что $c'_{1,1} \neq 0$. Далее, если $c_{i,1} \neq 0$ для некоторого $i > 1$, то преобразованием $(\mathrm{I}_{i,1,\lambda})$, где $\lambda = -c_{i,1}/c_{1,1}$, добиваемся того, что $c'_{i,1} = c_{i,1} + \lambda c_{1,1} = 0$. Таким образом, мы можем считать, что $c_{1,1} \neq 0$ и все элементы в первом столбце кроме $c_{1,1}$ равны 0. Рассмотрим матрицу D, полученную из C вычеркиванием первого столбца и первой строки. По предположению индукции D элементарными преобразованиями строк типа (I) приводится к матрице D', имеющей ступенчатый вид. Проделаем над матрицей C те же элементарные преобразования, что и над матрицей D. Получим матрицу C', в которой все элементы в первом столбце кроме $c'_{1,1}$ равны 0, а матрица, полученная из C' вычеркиванием первого столбца и первой строки, совпадает с D' (и имеет ступенчатый вид). Поэтому и вся матрица C'имеет ступенчатый вид. Это доказывает (1).

(2) Согласно (1) мы можем считать, что матрица C уже имеет ступенчатый вид. Если $c_{i,j}$ – первый ненулевой элемент i-й строки, то преобразованием $(\mathrm{II}_{i,1/c_{i,j}})$ добиваемся того, что $c'_{i,j} = 1$. Далее снова применяем индукцию по числу столбцов. Пусть C_i – последняя ненулевая строка и пусть $c_{i,j} = 1$ – ее первый ненулевой элемент. Тогда $c_{k,j} = 0$ при $k > i$. При $k < i$ преобразованиями $(\mathrm{I}_{k,i,-c_{k,j}})$ добиваемся того, что $c_{k,j} = 0$. Далее используем предположение индукции. □

Определение. Система линейных уравнений называется *ступенчатой* (соответственно *улучшенной ступенчатой*), если таковой является ее расширенная матрица. В этом случае неизвестные, соответствующие столбцам матрицы в которых стоят первые ненулевые элементы строк называются *главными*, а остальные неизвестные – *свободными*. Уравнение вида $0 = b_i$, где $b_i \neq 0$ называется *противоречивым*.

Решение систем линейных уравнений.

Согласно теореме, каждая система линейных уравнений эквивалентна улучшенной ступенчатой системе. Более того, переход к улучшенному ступенчатому виду осуществляется элементарными преобразованиями строк. Таким образом, мы можем считать, что наша система имеет улучшенный ступенчатый вид. Если в системе имеется противоречивое уравнение, то она, очевидно, несовместна. Далее мы считаем, что система не содержит противоречивых уравнений. Пусть $x_{j_1}, \dots, x_{j_r}$ – все свободные неизвестные. В каждом нетривиальном уравнении участвует ровно одна главная неизвестная и каждая главная неизвестная x_j участвует ровно в одном нетривиальном уравнении. Таким образом, перенося члены со свободными неизвестными в правую часть, мы можем выразить главные неизвестные через свободные:

$$x_j = b_i - (a_{i,j_1}x_{j_1} + \dots + a_{i,j_r}x_{j_r}).$$

Совокупность этих уравнений эквивалентна нашей изначальной системе. Придадим свободным неизвестным произвольные значения $x_{j_1} = c_{j_1}, \dots, x_{j_r} = c_{j_r}$. Тогда им соответствуют (единственные) значения главных неизвестных, удовлетворяющие нашей ступенчатой системе $(*)$.

Следствие. *Ступенчатая система линейных уравнений совместна тогда и только тогда, когда она не содержит противоречивых уравнений.*

Следствие. *Для любых значений свободных неизвестных совместной системы линейных уравнений существует единственное решение, принимающее эти значения.*

Следствие. *Совместная система линейных уравнений определенна тогда и только тогда, когда все ее неизвестные – главные.*

Доказателъство. Если существует свободная неизвестная, то система линейных уравнений не может быть определенна по последнему следствию. Если неизвестные – главные, то в улучшенном ступенчатом виде каждое нетривиальное уравнение имеет вид $x_i = b_j$. □

Следствие. *Квадратная система линейных уравнений совместна и определенна тогда и только тогда, когда в ее ступенчатом виде нет противоречивых уравнений и соответствующая матрица имеет строго треугольный вид, т.е.* $a_{i,j} = 0$ *при* $i > j$ *и* $a_{i,i} \neq 0$.

Следствие. *Если у системы однородных линейных уравнений уравнений число уравнений меньше числа неизвестных, то система совместна и неопределенна. В частности, она имеет ненулевое решение.*

Лекция 2

Подстановки. Запись подстановок. Число подстановок. Произведение подстановок. Свойства.

Обратная и единичные подстановки. Понятие группы. Примеры. Транспозиции. Разложение подстановки в произведение транспозиций. Применение подстановки к перестановке. Четность. Корректность определения четности (как меняется число инверсий при применении транспозиции). Четность произведения подстановок. Четность обратной подстановки. Число четных и нечетных подстановок. Группа A_n. Понятие подгруппы. Неподвижные элементы. Орбиты. Независимые подстановки коммутируют. Циклы. Разложение подстановки в произведение независимых циклов.

Свойства отображений

Два отображения $f : X \to Y$ и $g : X \to Y$ считаются *равными*, если $f(x) = g(x)$ для всех $x \in X$. *Тождественное* отображение множества X в себя будет обозначаться ε_X или просто ε. Таким образом, $\varepsilon(x) = x$ для всех $x \in X$. Напомним, что *композицией* отображений $f : X \to Y$ и $g : Y \to Z$ называется отображение $g \circ f : X \to Z$ такое, что $g \circ f(x) = g(f(x))$.

Теорема. *Рассмотрим отображения $f : X \to Y$, $g : Y \to Z$ и $h : Z \to U$. Тогда*

$$(*) \qquad (h \circ g) \circ f = h \circ (g \circ f).$$

Свойство $(*)$ называется *ассоциативностью*.

Доказательство. Для любого $x \in X$ имеем

$$(h \circ g) \circ f(x) = (h \circ g)(f(x)) = h(g(f(x))),$$

$$h \circ (g \circ f)(x) = h((g \circ f)(x)) = h(g(f(x))).$$

Следовательно, $h(g(f(x))) = h(g(f(x)))$. □

Отображение $f : X \to Y$ называется *инъективным*, если из того, что $f(x_1) = f(x_2)$ следует $x_1 = x_2$. Оно называется *сюръективным*, если для любого $y \in Y$ существует $x \in X$ такое, что $f(x) = y$. Отображение $f : X \to Y$ называется *биективным* если оно инъективно и сюръективно.

Пусть дано отображение $f : X \to Y$. *Обратным* к нему называется отображение $f^{-1} : Y \to X$ такое, что $f \circ f^{-1} = \varepsilon_Y$ и $f^{-1} \circ f = \varepsilon_X$.

Подстановки

Рассмотрим множество $\Omega_n := \{i_1, \dots, i_n\}$ из n элементов любой природы.

Определение. Подстановкой множества Ω_n называется любое биективное отображение $\sigma : \Omega_n \to \Omega_n$.

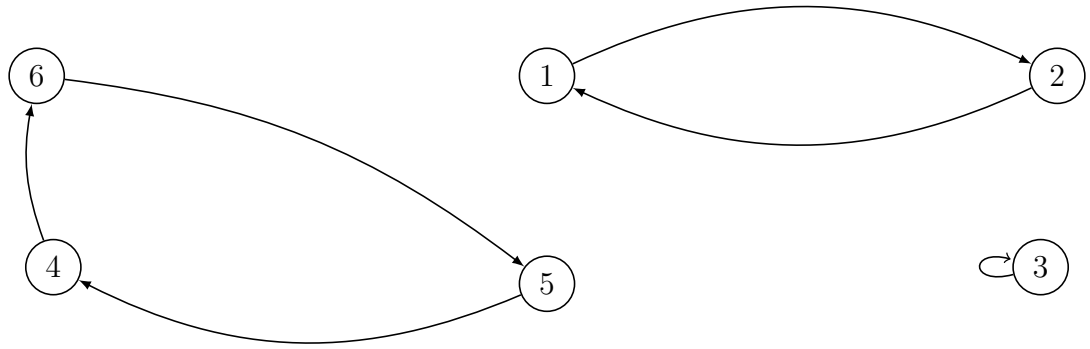
Далее мы считаем, что Ω_n – множество первых n натуральных чисел: $\Omega_n = \{1, 2, \dots, n\}$. Множество всех подстановок обозначим через S_n. Каждая подстановка однозначно $\sigma \in \mathrm{S}_n$ задается $2 \times n$-таблицей (матрицей)

$$(\dagger) \qquad \sigma = \begin{pmatrix} i_1 & i_2 & \cdots & i_{n-1} & i_n \\ j_1 & j_2 & \cdots & j_{n-1} & j_n \end{pmatrix}$$

Такая запись означает, что $\{i_1, \dots, i_n\} = \Omega_n$ и $j_k := \sigma(i_k)$. Схематично, это можно изобразить так:

$$\begin{array}{ccccc} i_1 & i_2 & \cdots & i_{n-1} & i_n \\ \downarrow & \downarrow & \cdots & \downarrow & \downarrow \\ j_1 & j_2 & \cdots & j_{n-1} & j_n \end{array}$$

Подстановку можно также задать направленным графом: вершины графа соответствуют элементам множества Ω_n и вершины i и j соединяются направленным ребром, если $\sigma(i) = j$. Например, подстановке $\left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 6 & 5 & 4 \end{smallmatrix}\right)$ соответствует граф



Отметим, что запись (†) не является единственной: при перестановке столбцов мы получаем ту же подстановку. Таким образом, каждая подстановка может быть записана в стандартном виде

$$\begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ j_1 & j_2 & \cdots & j_{n-1} & j_n \end{pmatrix}$$

Назовем *перестановкой* из n элементов упорядоченный набор $(i_1, \dots, i_n)$ элементов Ω_n. Таким образом, строки в (†) – перестановки. Ясно, что для перестановки $(i_1, \dots, i_n)$ всегда выполнено равенство множеств $\{i_1, \dots, i_n\} = \Omega_n$ и $i_k \neq i_l$ при $k \neq l$. Перестановку $(1, 2, \dots, n)$ назовем тривиальной.

Предложение. *Число число всех подстановок равно числу всех перестановок и равно* $n!$.

Произведением подстановок σ_1, $\sigma_2 \in \mathrm{S}_n$ назовем их композицию $\sigma_1 \circ \sigma_2 \in \mathrm{S}_n$. Тождественная подстановка обозначается через ε.

Свойства подстановок.

- $(\sigma \circ \varphi) \circ \delta = \sigma \circ (\varphi \circ \delta)$ для всех $\sigma, \varphi, \delta \in \mathrm{S}_n$ (ассоциативность);
- $\sigma \circ \varepsilon = \varepsilon \circ \sigma = \sigma$ для всех $\sigma \in \mathrm{S}_n$;
- для любой подстановки $\sigma \in \mathrm{S}_n$ существует подстановка $\sigma^{-1} \in \mathrm{S}_n$ такая, что

$$\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \varepsilon.$$

Подстановка σ^{-1} называется *обратной*.

Понятие группы

Определение. *Группой* называется множество G с операцией $(a, b) \mapsto a \circ b$. Такое, что выполняются свойства:

(1) $a \circ (b \circ c) = (a \circ b) \circ c$ для всех a, b, $c \in G$;

(2) существует элемент $1 \in G$ (*единичный* элемент или *нейтральный* элемент) такой, что $1 \circ a = a \circ 1 = a$ для всех $a \in G$;

(3) для любого $a \in G$ существует элемент $a^{-1} \in G$ (который называется *обратным* к a) такой, что $a \circ a^{-1} = a^{-1} \circ a = 1$.

Определение. Группа G называется *абелевой* (или *коммутативной*), если $a \circ b = b \circ a$ для любых a, $b \in G$.

В абелевой группе определено деление:

$$a/b := a \cdot b^{-1} = b^{-1} \cdot a.$$

В определениях выше операция $\circ$ “похожа” на операцию умножения. В этом случае говорят, что группа записана *мультипликативно*.

Иногда операция в группе записывается *аддитивно*. Тогда определение выше принимает вид:

Определение. *Аддитивной группой* называется множество G с операцией $(a, b) \mapsto a+b$. Такое, что выполняются свойства:

(1) $a + (b + c) = (a + b) + c$ для любых a, b, $c \in G$;

(2) существует элемент $0 \in G$ (*нулевой* элемент) такой, что $0 + a = a + 0 = a$ для любого $a \in G$;

(3) для любого $a \in G$ существует элемент $-a \in G$ (который называется *противоположным* к a) такой, что $a + (-a) = (-a) + a = 0$.

Обычно группа, записанная аддитивно, предполагается абелевой. В абелевой аддитивной группе определено вычитание:

$$a - b := a + (-b) = (-b) + a.$$

Примеры. (1) Множество $\{1\}$ из одного нейтрального элемента, очевидно, является группой. Множество чисел $\{\pm 1\}$ с операцией умножения является группой.

(2) Все подстановки из n элементов образуют группу S_n. Она называется *симметрической группой*. Эта группа неабелева при $n \geq 3$.

(3) Множество всех целых (соответственно, рациональных, действительных) чисел с операцией сложения образует (аддитивную) группу, которая обозначается $\mathbb{Z}^+$ (соответственно,$\mathbb{Q}^+$, $\mathbb{R}^+$). Эта группа является абелевой.

(4) Рассмотрим множество ненулевых рациональных (соответственно, действительных) чисел. Это множество с операцией умножения является (мультипликативной абелевой группой. Она обозначается $\mathbb{Q}^*$ (соответственно, $\mathbb{R}^*$). Эта группа также абелева.

Группа G, состоящая из конечного числа элементов, называется *конечной*. В этом случае число элементов G называется ее *порядком* и обозначается $|G|$. Таким образом,

$$|\,\mathrm{S}_n\,| = n!.$$

В мультипликативной группе имеет смысл понятие целой степени элемента: для $a \in G$ и $n \in \mathbb{Z}$ положим

$$a^n = \begin{cases} \underbrace{a \circ \cdots \circ a}_{n} & \text{если } n \in \mathbb{N}, \\ 1 & \text{если } n = 0, \\ (a^{-1})^{-n} & \text{если } -n \in \mathbb{N}. \end{cases}$$

Возведение в степень удовлетворяет стандартным свойствам:

$$a^n \circ a^m = a^{n+m} = a^m \circ a^n, \qquad (a^n)^m = (a^m)^n = a^{nm}.$$

Если группа G состоит из конечного числа элементов, то для любого $a \in G$ существует $n \in \mathbb{N}$ такое, что $a^n = 1$. Действительно, все элементы $1, a, a^2, a^3, \ldots$ не могут быть различны. Следовательно, $a^k = a^l$ для некоторых $k > l$. Тогда $a^{k-l} = a^k \circ (a^l)^{-1} = a^k \circ (a^k)^{-1} = 1$.

В аддитивной группе понятие степени заменяется на понятием умножения на целые числа: если $a \in G$ и $n \in \mathbb{Z}$, то

$$na = \begin{cases} \underbrace{a + \cdots + a}_{n} & \text{если } n \in \mathbb{N}, \\ 0 & \text{если } n = 0, \\ n(-a) & \text{если } -n \in \mathbb{N}. \end{cases}$$

Предложение. *Пусть G – группа. Тогда*

(1) *нейтральный элемент – единственный;*

(2) *для любого $a \in G$ обратный элемент a^{-1} – единственный;*

(3) *для любых $a, b \in G$ уравнение $a \circ x = b$ (соответственно, уравнение $x \circ a = b$) имеет единственное решение.*

(4) $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$.

Доказательство. (1) Предположим, что имеется два единичных элемента 1 и $1'$. Тогда $1 = 1 \circ 1' = 1'$.

(2) Предположим, что для $a \in G$ имеется два обратных элемента a^{-1} и $a^{-1\prime}$. Тогда $a = a^{-1} \circ a \circ a^{-1\prime} = a^{-1\prime}$.

(3) Для решения уравнения $a \circ x = b$ домножим обе части на a^{-1} *слева*. Получим $1 \circ x = a^{-1} \circ a \circ x = a^{-1} \circ b$. Уравнение $x \circ a = b$ решается аналогично.

(4) $(a \circ b) \circ \left(b^{-1} \circ a^{-1}\right) = a \circ (b \circ b^{-1}) \circ a^{-1} = a \circ 1 \circ a^{-1} = a \circ a^{-1} = 1$. □

Транспозиции

Определение. *Транспозицией* называется подстановка $\tau \in \mathrm{S}_n$ такая, что существуют элементы i, $j \in \Omega_n$, $i \neq j$ такие, что $\tau(i) = j$, $\tau(j) = i$ и $\tau(k) = k$ при $k \notin \{i, j\}$. Эта транспозиция обозначается $\sigma = [i, j]$.

Теорема. *Любая подстановка $\sigma \in \mathrm{S}_n$ представляется в виде произведения транспозиций, т.е. $\sigma = \tau_1 \circ \cdots \circ \tau_r$, где $\tau_1, \ldots, \tau_r$ – транспозиции.*

Доказательство. Индукция по n. Предположим, что утверждение верно для $n - 1$. Пусть

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ i_1 & i_2 & \cdots & i_{n-1} & i_n \end{pmatrix}$$

Если $i_n \neq n$, то рассмотрим транспозицию $\tau = [n, i_n]$. Если же $i_n = n$, то положим $\tau = \varepsilon$. В обоих случаях

$$\tau \circ \sigma = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ i'_1 & i'_2 & \cdots & i'_{n-1} & n \end{pmatrix}$$

Рассмотрим подстановку

$$\sigma' = \begin{pmatrix} 1 & 2 & \cdots & n-1 \\ i'_1 & i'_2 & \cdots & i'_{n-1} \end{pmatrix} \in \mathrm{S}_{n-1}$$

По предположению индукции она раскладывается в произведение транспозиций: $\sigma' = \tau'_1 \circ \cdots \tau'_m$, $\tau'_i = [k_i, l_i] \in \mathrm{S}_{n-1}$, $k_i\ l_i \in \{1, \ldots, n-1\}$, $k_i \neq l_i$. Рассмотрим транспозиции $\tau_i = [k_i, l_i] \in \mathrm{S}_n$. Очевидно, что $\tau \circ \sigma = \tau_1 \circ \cdots \circ \tau_m$. Поэтому $\sigma = \tau \circ \tau_1 \circ \cdots \circ \tau_m$. □

Замечание. На последнем шаге мы использовали то, что имеется вложение групп

$$\mathrm{S}_{n-1} \hookrightarrow \mathrm{S}_n, \qquad \begin{pmatrix} 1 & 2 & \cdots & n-1 \\ j_1 & j_2 & \cdots & j_{n-1} \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ j_1 & j_2 & \cdots & j_{n-1} & n \end{pmatrix}$$

т.е. вложение, при котором умножение в образе S_{n-1} совпадает с умножением в S_n. Таким образом, можно отождествить образ S_{n-1} с самой группой S_{n-1}. Отметим однако, что вложение $\mathrm{S}_{n-1} \hookrightarrow \mathrm{S}_n$ заведомо не является единственным: оно зависит от фиксации элемента $n \in \Omega_n$.

Заметим также, что разложение подстановки в произведение транспозиций не единственно.

Для перестановки $\Pi = (i_1, \ldots, i_n)$ и подстановки $\sigma \in \mathrm{S}_n$ положим

$$\sigma(\Pi) := (\sigma(i_1), \ldots, \sigma(i_n)).$$

Ясно, что $\sigma(\Pi)$ – перестановка и выполнено очевидное равенство

$$\delta \circ \sigma(\Pi) = \delta(\sigma(\Pi)).$$

Следствие. *Любые две перестановки из одинакового числа элементов могут быть получены друг из друга применением конечного числа транспозиций.*

Доказательство. Пусть $\Pi = (i_1, \ldots, i_n)$ и $\Pi' = (j_1, \ldots, j_n)$. Рассмотрим подстановку

$$\sigma = \begin{pmatrix} i_1 & i_2 & \cdots & i_{n-1} & i_n \\ j_1 & j_2 & \cdots & j_{n-1} & j_n \end{pmatrix}$$

По теореме $\sigma = \tau_1 \circ \cdots \circ \tau_r$, где τ_k – транспозиции. Тогда

$$\Pi' = \sigma(\Pi) = \tau_1 \circ \cdots \circ \tau_r(\Pi) = \tau_1(\tau_2(\cdots \tau_r(\Pi))). \qquad \square$$

Четность подстановок.

Определение. Пусть $\Pi = (i_1, \ldots, i_n)$ – перестановка. Зафиксируем два элемента i_k и i_l, где $k < l$:

$$\Pi = (i_1, \ldots, i_k, \ldots, i_l, \ldots, i_n).$$

Будем говорить, что элементы i_k и i_l образуют *инверсию*, если $i_k > i_l$. В противном случае (если $i_k < i_l$) будем говорить, что i_k и i_l образуют *порядок*.

Четностью перестановки Π называется четность общего числа инверсий в Π.

Пример. Пусть $1 \leq i < j \leq n$. Число инверсий в перестановке

$$(1, 2, \ldots i-1, j, i+1, \ldots, j-1, i, j+1, \ldots, n-1, n)$$

равно $j - i + \underbrace{1 + \cdots + 1}_{j-i-1} = 2(j-i) - 1$. Поэтому перестановка – нечетная.

Лемма. *Пусть* $\Pi = (i_1, \dots, i_n)$ *– перестановка и пусть* $\tau \in \mathrm{S}_n$ *– транспозиция. Тогда четности перестановок* Π *и* $\tau(\Pi)$ *противоположны. Иначе говоря, применение транспозиции меняет четность перестановки.*

Доказательство. Пусть $\Pi = (i_1, \dots, i_n)$, пусть $\tau = [a, b]$, $a \neq b$. Ясно, что $a = i_k$, $b = i_l$ для некоторых $i_k \neq i_l$, где мы можем считать, что $k < l$. Таким образом,

$$\Pi = (i_1, \dots, i_k, \dots, i_l, \dots, i_n)$$

и $\tau = [i_k, i_l]$. Тогда

$$\tau(\Pi) = (i_1, \dots, i_l, \dots, i_k, \dots, i_n).$$

Проведем доказательство индукцией по $l - k$.

Начнем с базы индукции,т.е. предположим, что $l = k+1$, т.е. $\tau = [i_k, i_{k+1}]$. Обозначим через s_r (соответственно, s'_r)) число инверсий, которые образует число, стоящее на месте r в Π (соответственно, $\tau(\Pi)$) со всеми последующими. Если $r < k$ или $r > k + 1$, то $s_r = s'_r$). Для $r = k$ и $r = k + 1$ имеем

$$s'_k = \begin{cases} s_{k+1} + 1 & \text{если } i_k < i_{k+1} \\ s_{k+1} & \text{если } i_k > i_{k+1} \end{cases}$$

$$s'_{k+1} = \begin{cases} s_k & \text{если } i_k < i_{k+1} \\ s_k - 1 & \text{если } i_k > i_{k+1} \end{cases}$$

В любом случае

$$s'_k + s'_{k+1} = s_k + s_{k+1} \pm 1.$$

В итоге получаем

$$\text{число инверсий в } \tau(\Pi) = \sum_{r=1}^{n} s'_r = \sum_{r=1}^{n} s_r \pm 1 = \big(\text{число инверсий в } \Pi\big) \pm 1.$$

Это доказывает наше утверждение в случае $l = k + 1$.

Поэтому далее считаем, что $l > k + 1$. Предположим, что утверждение верно для всех транспозиций вида $\tau = [i_{k'}, i_{l'}]$, где $0 < l' - k' < l - k$. Имеем

$$[i_k, i_l] = [i_k, i_{k+1}] \circ [i_{k+1}, i_l] \circ [i_{k+1}, i_k].$$

Согласно нашему предположению индукции применение транспозиций к любой перестановке меняет ее четность. Последовательно, применяя это соображение к $[i_k, i_{k+1}] \circ [i_{k+1}, i_l] \circ [i_{k+1}, i_k](\Pi)$ (три раза), получим, что четности Π и $\tau(\Pi)$ противоположны. $\square$

Определение. *Четностью подстановки*

$$\sigma = \begin{pmatrix} i_1 & i_2 & \cdots & i_n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix} \in \mathrm{S}_m \tag{‡}$$

назовем четность суммы четностей перестановок $(i_1, i_2, \dots, i_n)$ и $(j_1, j_2, \dots, j_n)$, т.е. четность числа инверсий в верхней и нижней строки в ее записи (‡).

Пример. Согласно предыдущему примеру транспозиция является нечетной подстановкой.

Лемма. *Четность подстановки не зависит от формы ее записи, т.е. четность подстановки не меняется при перестановке столбцов в* (‡).

Доказательство. Согласно предыдущей лемме при транспозиции столбцов меняется четность числа инверсий в обоих строках, значит четность суммарного числа инверсий не меняется. □

Таким образом, определение корректно и четность подстановки не зависит от вида записи. *Знаком* подстановки σ называется

$$\operatorname{sgn}(\sigma) := (-1)^{\text{четность }(\sigma)}.$$

Предложение. (1) $\operatorname{sgn}(\sigma_1 \circ \sigma_2) = \operatorname{sgn}(\sigma_1)\operatorname{sgn}(\sigma_2)$.

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

Доказательство. (1) Пусть $\sigma_1 = \tau_1 \circ \cdots \circ \tau_r$ и $\sigma_2 = \tau'_1 \circ \cdots \circ \tau'_r$ – разложения в произведение транспозиций. Тогда $\operatorname{sgn}(\sigma_1)\operatorname{sgn}(\sigma_2) = (-1)^r(-1)^k = (-1)^{r+k} = \operatorname{sgn}(\sigma_1 \circ \sigma_2)$.

(2) Аналогично, пусть $\sigma = \tau_1 \circ \cdots \circ \tau_r$ – разложение в произведение транспозиций. Тогда $\sigma^{-1} = \tau_r \circ \cdots \circ \tau_1$ – разложение в произведение транспозиций для σ^{-1}. Отсюда $\operatorname{sgn}(\sigma^{-1}) = r = \operatorname{sgn}(\sigma)$. □

Знакопеременная группа. Понятие подгруппы.

Предложение. *Число четных подстановок равно числу нечетных подстановок и равно* $n!/2$.

Доказательство. Зафиксируем некоторую нечетную подстановку τ (например, транспозицию). Согласно сказанному выше для любой подстановки $\sigma \in \mathrm{A}_n \quad \tau \circ \sigma \in \mathrm{S}_n \setminus \mathrm{A}_n$. Следовательно, имеется отображение $f : \mathrm{A}_n \to \mathrm{S}_n \setminus \mathrm{A}_n$, $f(\sigma) = \tau \circ \sigma$. Легко показать, что оно биективно. Следовательно, множества A_n и $\mathrm{S}_n \setminus \mathrm{A}_n$ равномощны. □

Множество всех четных подстановок мы обозначим через A_n. Поскольку произведение четных подстановок четно, то на A_n корректно определена операция умножения подстановок. Несложно видеть, что A_n с этой операцией удовлетворяет аксиомам в определении группы. Следовательно, A_n – группа. Она называется *знакопеременной группой*.

Определение. Подмножество H в группе G называется *подгруппой*, если оно является группой с той же операцией, что и в G.

Замечание. Для того, чтобы подмножество H в группе G было подгруппой необходимо и достаточно, чтобы выполнялись следующие условия:

(1) для любых элементов a, $b \in H$ их произведение $a \circ b$ принадлежит H;

(2) для любого элемента $a \in H$ обратный к нему a^{-1} принадлежит H.

Из последнего замечания легко видеть, что пересечение двух подгрупп является подгруппой. Более точно верно следующее утверждение:

Утверждение. *Пусть H_i, $i \in I$ – любое множество подгрупп группы G. Тогда их пересечение $\cap_{i\in I} H_i$ также является подгруппой.*

Примеры. (1) Симметрическая группа S_n содержит знакопеременную подгруппу $\mathrm{A}_n \subset \mathrm{S}_n$. Также мы уже обсуждали, что S_n содержит также подгруппы S_{n-1}.

(2) Группа $\mathbb{R}^+$ содержит подгруппы $\mathbb{Z}^+$ и $\mathbb{Q}^+$.

(3) Группа $\mathbb{R}^*$, содержит подгруппы $\mathbb{Q}^*$ и $\{\pm 1\}$.

Для элементов $a_1, \dots, a_m$ группы G пересечение всех подгрупп в G, содержащих $a_1, \dots, a_m$, является подгруппой. Она называется *подгруппой, порожденной множеством* $a_1, \dots, a_m$ и обозначается $\langle a_1, \dots, a_m\rangle$. Эта подгруппа состоит из всевозможных произведений степеней элементов $a_1, \dots, a_m$ (в произвольном порядке, с возможными повторениями). Очевидно, что также можно определить подгруппу порожденную бесконечным подмножеством элементов.

Например, мы доказали, что симметрическая группа S_n порождается транспозициями. Известно, что знакопеременная подгруппа $\mathrm{A}_n \subset \mathrm{S}_n$ порождается тройными циклами $[i, j, k]$, этот факт будет доказан позже.

Циклы

Зафиксируем подстановку $\sigma \in \mathrm{S}_n$. Два элемента i, $j \in \Omega_n$ назовем эквивалентными, если $j = \sigma^k(i)$ для некоторого $k \in \mathbb{N}$. Несложно проверить, что это отношение эквивалентности. Таким образом, множество Ω_n представляется в виде объединения непересекающихся классов эквивалентности:

$$(\S) \qquad \Omega_n = \Omega_n^{(1)} \cup \dots \cup \Omega_n^{(l)}, \qquad \Omega_n^{(k)} \cap \Omega_n^{(m)} = \varnothing \quad \text{при } k \neq m.$$

Эти классы называются *орбитами*.

Орбита состоит из одного элемента i тогда и только тогда, когда $\sigma(i) = i$. Такие элементы называются *неподвижными* относительно σ. В противном случае (если $\sigma(j) \neq j$) элемент j называется *подвижным*. Множество всех неподвижных (соответственно, подвижных) элементов мы обозначим через $\mathrm{Fix}(\sigma)$ (соответственно, $\mathrm{Mov}(\sigma)$). Ясно, что

$$\mathrm{Fix}(\sigma) \cup \mathrm{Mov}(\sigma) = \Omega_n \quad \text{и} \quad \mathrm{Fix}(\sigma) \cap \mathrm{Mov}(\sigma) = \varnothing.$$

Замечание. Несложно видеть, что

$$\mathrm{Mov}(\sigma_1 \circ \sigma_2) \subset \mathrm{Mov}(\sigma_1) \cup \mathrm{Mov}(\sigma_2) \quad \text{и} \quad \mathrm{Fix}(\sigma_1 \circ \sigma_2) \supset \mathrm{Fix}(\sigma_1) \cap \mathrm{Fix}(\sigma_2)$$

для любых подстановок σ_1, $\sigma_2 \in \mathrm{S}_n$. Следовательно, для любой подстановки $\sigma \in \mathrm{S}_n$ и любой ее степени σ^m имеют место включения

$$\mathrm{Fix}(\sigma) \subset \mathrm{Fix}(\sigma^m) \quad \text{и} \quad \mathrm{Mov}(\sigma) \supset \mathrm{Mov}(\sigma^m).$$

Лемма. *Если для подстановок* σ, $\varphi \in \mathrm{S}_n$ *выполнено* $\mathrm{Mov}(\sigma) \cap \mathrm{Mov}(\varphi) = \varnothing$, *то* $\sigma \circ \varphi = \varphi \circ \sigma$ (*т.е.* σ *и* φ *коммутируют*).

Доказательство. Возьмем элемент $i \in \Omega_n$. Если $i \notin \mathrm{Mov}(\sigma) \cup \mathrm{Mov}(\varphi)$, то $i \in \mathrm{Fix}(\sigma) \cap \mathrm{Fix}(\varphi)$, т.е. $\sigma(i) = i = \varphi(i)$. В этом случае $\sigma \circ \varphi(i) = i = \varphi \circ \sigma(i)$.

Пусть $i \in \mathrm{Mov}(\sigma)$. Тогда $i \notin \mathrm{Mov}(\varphi)$, $i \in \mathrm{Fix}(\varphi)$ и $\sigma(i) \in \mathrm{Mov}(\sigma)$. Значит $\varphi(i) = i$ и $\sigma(i) \notin \mathrm{Mov}(\varphi)$, т.е. $\sigma(i) \in \mathrm{Fix}(\varphi)$. Таким образом, $\sigma \circ \varphi(i) = \sigma(i) = \varphi \circ \sigma$. Случай $i \in \mathrm{Mov}(\varphi)$ разбирается аналогично. □

Определение. Подстановка $\sigma \in \mathrm{S}_n$ называется *циклом* (*циклической подстановкой*), если $\mathrm{Mov}(\sigma) = \{i_1, \ldots, i_m\}$ и

$$\sigma(i_k) = \begin{cases} i_{k+1} & \text{при } k = 1, \ldots, m-1, \\ i_1 & \text{при } k = m. \end{cases}$$

Такая подстановка обозначается $\sigma = [i_1, \ldots, i_m]$. Число m называется *длиной* цикла.

Цикл длины 2 – это транспозиция. Запись $\sigma = [i_1, \ldots, i_m]$ не единственна. Ясно, что $[i_1, \ldots, i_m] = [i_2, \ldots, i_m, i_1] = [i_3, \ldots, i_m, i_1, i_2]$ и т. д.

Циклы $\sigma = [i_1, \ldots, i_m]$ и $\varphi = [j_1, \ldots, j_l]$ называются *независимыми*, если $\mathrm{Mov}(\sigma) \cap \mathrm{Mov}(\varphi) = \varnothing$.

Теорема. *Любая подстановка* $\sigma \in \mathrm{S}_n$ *представляется в виде произведения независимых циклов* $\sigma = \sigma_1 \circ \cdots \circ \sigma_l$. *Это произведение единственно с точностью до порядка множителей.*

Пример.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 2 & 3 & 1 & 4 & 6 & 7 & 8 & 6 & 10 & 9 \end{pmatrix} = [1, 2, 3] \circ [5, 6, 7, 8] \circ [9, 10].$$

Основная идея доказательства состоит в том, что разложение соответствует разбиению (§) множества Ω_n на орбиты.

Доказательство. Докажем существование разложения индукцией по числу элементов в $\mathrm{Mov}(\sigma)$. Пусть $i_1 \in \mathrm{Mov}(\sigma)$ – подвижный элемент. Так как $\sigma^m = \varepsilon$ для некоторого m, то $\sigma^m(i_1) = i_1$. Возьмем m минимальным положительным таким, что $\sigma^m(i_1) = i_1$. Ясно, что $m > 1$. Положим $i_k := \sigma^{k-1}(i_1)$. Тогда для все числа

$$i_1, \ldots, i_{m_1}$$

различны.* Действительно, иначе $i_k = i_l$ для некоторых $1 \leq k < l < m$. Но тогда $\sigma^k = \sigma^l$ и $\sigma^{l-k} = 1$, где $0 < l - k < m$. Противоречие.

Более того, $\sigma(i_k) = i_{k+1}$ при $1 \leq k < m-1$ и $\sigma(i_{m-1}) = i_1$. Положим $\sigma_1 := [i_1, \ldots, i_{m-1}]$ и $\sigma' = \sigma \circ \sigma_1^{-1}$. Тогда $\sigma_1(i_k) = \sigma(i_k)$ при $k = 1, \ldots, m-1$ и $\sigma_1(j) = j$ при $j \notin \{i_1, \ldots, i_{m-1}\}$. Значит, $\sigma'(i_k) = i_k$ при $k = 1, \ldots, m-1$ и $\sigma'(j) = \sigma(j)$ при $j \notin \{i_1, \ldots, i_{m-1}\}$. Таким

*На самом деле, множество $\{i_1, \ldots, i_{m_1}\}$ – это одна из орбит в разложении (§).

образом, $\mathrm{Mov}(\sigma_1) = \{i_1, \ldots, i_{m-1}\} \subset \mathrm{Fix}(\sigma')$, $\mathrm{Fix}(\sigma') \supsetneq \mathrm{Fix}(\sigma)$ и $\mathrm{Mov}(\sigma') \subsetneq \mathrm{Mov}(\sigma)$. Более того, имеет место разбиение

$$\mathrm{Mov}(\sigma) = \mathrm{Mov}(\sigma') \cup \mathrm{Mov}(\sigma_1), \qquad \mathrm{Mov}(\sigma') \cap \mathrm{Mov}(\sigma_1) = \varnothing.$$

По предположению индукции мы можем записать

$$\sigma' = \sigma_2 \circ \cdots \circ \sigma_l,$$

где $\sigma_2, \ldots, \sigma_l$ – независимые циклы и $\mathrm{Mov}(\sigma_k) \subset \mathrm{Mov}(\sigma')$ при $k = 2, \ldots, l$. Тогда

$$\sigma = \sigma' \circ \sigma_1 = \sigma_1 \circ \sigma' = \sigma_1 \circ \sigma_2 \circ \cdots \circ \sigma_l.$$

Докажем единственность. Предположим, что имеется два разложения в произведение независимых циклов

$$\sigma = \sigma_1 \circ \cdots \circ \sigma_l = \sigma'_1 \circ \cdots \circ \sigma'_r.$$

Достаточно доказать, что для каждого цикла σ_k существует цикл σ'_s такой, что $\sigma_k = \sigma'_s$, т.е. $\sigma_k(i) = \sigma'_s(i)$ для всех $i \in \Omega_n$. По нашему предположению множество $\mathrm{Mov}(\sigma)$ двумя способами представляется в виде непересекающихся объединений

$$\mathrm{Mov}(\sigma) = \mathrm{Mov}(\sigma_1) \cup \cdots \cup \mathrm{Mov}(\sigma_l) = \mathrm{Mov}(\sigma'_1) \cup \cdots \cup \mathrm{Mov}(\sigma'_r).$$

Возьмем s так, что $\mathrm{Mov}(\sigma_k) \cap \mathrm{Mov}(\sigma_s) \neq \varnothing$. Пусть $i \in \mathrm{Mov}(\sigma_k) \cap \mathrm{Mov}(\sigma_s)$. Любой элемент $j \in \mathrm{Mov}(\sigma_k)$ представляется в виде $\sigma_k^a(i) = \sigma^a(i)$ для некоторого a. Аналогично, любой элемент $j \in \mathrm{Mov}(\sigma'_s)$ представляется в виде $\sigma_s'^a(i) = \sigma^a(i)$ для некоторого a. Поэтому

$$\mathrm{Mov}(\sigma_k) = \{\sigma_k^a(i) \mid a \in \mathbb{N}\} = \{\sigma^a(i) \mid a \in \mathbb{N}\} = \{\sigma_s'^b(i) \mid a \in \mathbb{N}\} = \mathrm{Mov}(\sigma'_s).$$

Значит для любого $j \in \mathrm{Mov}(\sigma_k) = \mathrm{Mov}(\sigma'_s)$ имеем $\sigma_k(j) = \sigma_k(j) = \sigma'_s(j)$.

Если же $j \notin \mathrm{Mov}(\sigma_k) = \mathrm{Mov}(\sigma'_s)$, то $\sigma_k(j) = j = \sigma'_s(j)$. Таким образом, $\sigma_k = \sigma'_s$. $\square$

Лекция 3

Операции сложения и умножения матриц на число. Свойства. Умножение матриц. Свойства.

Ассоциативность. Матричная запись систем линейных уравнений. Связь однородных и неоднородных систем линейных уравнений. Понятие кольца. Примеры. Умножение на диагональные матрицы. Умножение треугольных матриц.

Сложение матриц и умножение матриц на число

Множество матриц размера $n \times m$ с коэффициентами из $\mathbb{R}$ мы будем обозначать $\mathrm{Mat}_{n,m}(\mathbb{R})$. В случае $n = m$ мы будем сокращенно писать $\mathrm{Mat}_n(\mathbb{R})$ вместо $\mathrm{Mat}_{n,n}(\mathbb{R})$.

Пусть даны две матрицы $A, B \in \mathrm{Mat}_{n,m}(\mathbb{R})$ (одинакового размера)

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix} \qquad B = \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,m} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n,1} & b_{n,2} & \cdots & b_{n,m} \end{pmatrix}$$

Их суммой называется матрица $A + B \in \mathrm{Mat}_{n,m}(\mathbb{R})$, полученная покомпонентным сложением элементов:

$$A + B = \begin{pmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} & \cdots & a_{1,m} + b_{1,m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} + b_{n,1} & a_{n,2} + b_{n,2} & \cdots & a_{n,m} + b_{n,m} \end{pmatrix}$$

Если λ – произвольное число, то можно определить произведение $\lambda A \in \mathrm{Mat}_{n,m}(\mathbb{R})$:

$$\lambda A = \begin{pmatrix} \lambda a_{1,1} & \lambda a_{1,2} & \cdots & \lambda a_{1,m} \\ \cdots & \cdots & \cdots & \cdots \\ \lambda c_{n,1} & \lambda c_{n,2} & \cdots & \lambda a_{n,m} \end{pmatrix}$$

Нулевой матрицей будем называть матрицу из нулей:

$$O = \begin{pmatrix} 0 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 0 \end{pmatrix}$$

Размер нулевой матрицы всегда будет подходящим, т.е. таким, чтобы соответствующие операции имели смысл.

Следующие утверждения непосредственно вытекают из определений.

Предложение. *Операция сложения матриц удовлетворяет следующим свойствам:*

(1) $A+(B+C)=(A+B)+C$ *для любых матриц* $A,B,C\in \mathrm{Mat}_{n,m}(\mathbb{R})$ (*ассоциативность*);

(2) $O+A=A+O=A$ *для любой матрицы* $A\in \mathrm{Mat}_{n,m}(\mathbb{R})$, *где* $O\in \mathrm{Mat}_{n,m}(\mathbb{R})$ – *нулевая матрица;*

(3) *для любой матрицы* $A\in \mathrm{Mat}_{n,m}(\mathbb{R})$ *существует матрица* $-A\in \mathrm{Mat}_{n,m}(\mathbb{R})$ *такая, что* $A+(-A)=(-A)+A=O$;

(4) $A+B=B+A$ *для любых матриц* $A,B\in \mathrm{Mat}_{n,m}(\mathbb{R})$ (*коммутативность*).

Таким образом, $\mathrm{Mat}_{n,m}(\mathbb{R})$ – абелева группа по сложению.

Предложение. *Операция умножения матриц на число удовлетворяет следующим свойствам:*

(1) $\lambda(A+B)=\lambda A+\lambda B$ *для любых матриц* $A,B\in \mathrm{Mat}_{n,m}(\mathbb{R})$ *и любого числа* $\lambda\in\mathbb{R}$;

(2) $(\lambda+\mu)A=\lambda A+\mu A$ *для любой матрицы* $A\in \mathrm{Mat}_{n,m}(\mathbb{R})$ *и любых чисел* $\lambda,\mu\in\mathbb{R}$;

(3) $(\lambda\mu)A=\lambda(\mu A)$ *для любой матрицы* $A\in \mathrm{Mat}_{n,m}(\mathbb{R})$ *и любых чисел* $\lambda,\mu\in\mathbb{R}$;

(4) $1\,A=A$ *для любой матрицы* $A\in \mathrm{Mat}_{n,m}(\mathbb{R})$.

Умножение матриц.

Пусть даны две матрицы $A\in \mathrm{Mat}_{n,m}(\mathbb{R})$ и $B\in \mathrm{Mat}_{m,r}(\mathbb{R})$

$$(*)\qquad A=\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m}\\ \cdots & \cdots & \cdots & \cdots\\ a_{n,1} & a_{n,2} & \cdots & a_{n,m}\end{pmatrix}\qquad B=\begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,r}\\ \cdots & \cdots & \cdots & \cdots\\ b_{m,1} & b_{m,2} & \cdots & b_{m,r}\end{pmatrix}$$

Их произведением называется матрица $C=A\cdot B\in \mathrm{Mat}_{n,r}(\mathbb{R})$

$$C=\begin{pmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,r}\\ \cdots & \cdots & \cdots & \cdots\\ c_{n,1} & c_{n,2} & \cdots & c_{n,r}\end{pmatrix}$$

где элементы $c_{i,j}$ матрицы $C=A\cdot B$ вычисляются по правилу

$$c_{i,j}=\sum_{k=1}^{m} a_{i,k}b_{k,j}=a_{i,1}b_{1,j}+a_{i,2}b_{2,j}+\cdots+a_{i,m}b_{m,j}.$$

Пример. Обычно произведение матриц зависит от порядка сомножителей, т.е. умножение матриц некоммутативно:

$$\begin{pmatrix}0&1\\0&0\end{pmatrix}\cdot\begin{pmatrix}0&0\\1&0\end{pmatrix}=\begin{pmatrix}1&0\\0&0\end{pmatrix},\qquad \begin{pmatrix}0&0\\1&0\end{pmatrix}\cdot\begin{pmatrix}0&1\\0&0\end{pmatrix}=\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

Введем некоторые стандартные определения. Пусть

$$A = (a_{i,j}) = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}$$

– квадратная матрица порядка n. Ее *главной диагональю* называется диагональ, идущая из левого верхнего угла в правый нижний, т.е. диагональ из элементов $a_{i,i}$. Другая диагональ, идущая из правого верхнего угла в левый нижний, (диагональ из элементов $a_{i,n+1-i}$) называется *побочной*. Матрица A называется *верхнетреугольной*, если ниже ее главной диагонали стоят нули, т.е. $a_{i,j} = 0$ при $i > j$ и она называется *нижнетреугольной*, если выше ее главной диагонали стоят нули, т.е. $a_{i,j} = 0$ при $i < j$. Матрица A называется *диагональной*, если она и верхнетреугольная, и нижнетреугольная одновременно, т.е. $a_{i,j} = 0$ при $i \neq j$. *Единичной матрицей* называется диагональная матрица, у которой все элементы главной диагонали равны 1:

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

Единичную матрицу мы будем обозначать через E. Как и в случае нулевой матрицы, размер единичной матрицы всегда будет подходящим, т.е. таким, чтобы соответствующие операции имели смысл.

Умножение произвольной матриц B размера $n \times m$ на диагональную выглядит так:

$$\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d_n \end{pmatrix} \cdot \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,m} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n,1} & b_{n,2} & \cdots & b_{n,m} \end{pmatrix} = \begin{pmatrix} d_1 b_{1,1} & d_1 b_{1,2} & \cdots & d_1 b_{1,m} \\ d_2 b_{2,1} & d_2 b_{2,2} & \cdots & d_2 b_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ d_n b_{n,1} & d_n b_{n,2} & \cdots & d_n b_{n,m} \end{pmatrix}$$

$$\begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,m} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n,1} & b_{n,2} & \cdots & b_{n,m} \end{pmatrix} \cdot \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d_m \end{pmatrix} = \begin{pmatrix} b_{1,1} d_1 & b_{1,2} d_2 & \cdots & b_{1,m} d_m \\ b_{2,1} d_1 & b_{2,2} d_2 & \cdots & b_{2,m} d_m \\ \cdots & \cdots & \cdots & \cdots \\ b_{n,1} d_1 & b_{n,2} d_2 & \cdots & b_{n,m} d_m \end{pmatrix}$$

При умножении на единичную матрица не меняется:

$$E \cdot B = B, \qquad B \cdot E = B.$$

Здесь E – единичная матрица порядка n в первом равенстве и порядка m во втором. Следующую лемму предлагается доказать самостоятельно.

Лемма. *Произведение верхнетреугольных (соответственно, нижнетреугольных) матриц является верхнетреугольной (соответственно, нижнетреугольной).*

Лемма. (1) *Пусть A, B, C – матрицы размеров $n \times m$, $n \times m$ и $m \times r$, соответственно. Тогда*

$$(A + B) \cdot C = A \cdot C + B \cdot C.$$

(2) *Пусть A, B, C – матрицы размеров $n \times m$, $m \times r$ и $m \times r$, соответственно. Тогда*

$$A \cdot (B + C) = A \cdot B + A \cdot C.$$

(3) *Пусть A, B – матрицы размеров $n \times m$ и $n \times m$, соответственно. Тогда для любого числа λ выполнено*

$$(\lambda A) \cdot B = A \cdot (\lambda B) = \lambda (A \cdot B).$$

Доказательство. Докажем, например, (1). Пусть $A = (a_{i,j})$, $B = (b_{i,j})$, $C = (c_{i,j})$. Мы можем записать $A + B = (a_{i,j} + b_{i,j})$, $A \cdot C = (d_{i,j})$, $B \cdot C = (f_{i,j})$, $(A + B) \cdot C = g_{i,j}$, где

$$d_{i,j} = \sum_{k=1}^{m} a_{i,k} c_{k,j}, \qquad f_{i,j} = \sum_{k=1}^{m} b_{i,k} c_{k,j}, \qquad g_{i,j} = \sum_{k=1}^{m} (a_{i,k} + b_{i,k}) c_{k,j}.$$

Поэтому $d_{i,j} + f_{i,j} = g_{i,j}$. □

Теорема. *Пусть A, B, C – матрицы размеров $n \times m$, $m \times r$ и $r \times q$ соответственно. Тогда $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.*

Доказательство. Пусть $A = (a_{i,j})$, $B = (b_{i,j})$, $C = (c_{i,j})$. Мы можем записать $A \cdot B = (d_{i,l})$, $B \cdot C = (f_{k,j})$, где

$$d_{i,l} = \sum_{k=1}^{m} a_{i,k} b_{k,l}, \qquad f_{k,j} = \sum_{l=1}^{r} b_{k,l} c_{l,j}.$$

Поэтому $(A \cdot B) \cdot C = (g_{i,j})$, $A \cdot (B \cdot C) = (h_{i,j})$, где

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

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

Откуда видно, что $g_{i,j} = h_{i,j}$. □

Для матрицы $A = (a_{i,j}) \in \mathrm{Mat}_{n,m}(\mathbb{R})$ (см. $(*)$) определим *транспонированную матрицу* $A^{\mathrm{T}} \in \mathrm{Mat}_{m,n}(\mathbb{R})$ следующим образом $A^{\mathrm{T}} = (a'_{i,j})$, где $a'_{i,j} = a_{j,i}$. Иначе говоря,

$$A^{\mathrm{T}} = \begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1,n} & a_{2,n} & \cdots & a_{m,n} \end{pmatrix}$$

Таким образом, столбец из элементов $c_1, \ldots, c_n$ можно компактно записать так: $(c_1, \ldots, c_n)^{\mathrm{T}}$.

Предложение. $(A \cdot B)^{\mathrm{T}} = B^{\mathrm{T}} \cdot A^{\mathrm{T}}$.

Доказательство. Действительно, пусть $A = (a_{i,k})$, $B = (b_{k,j})$, $A \cdot B = (c_{i,j})$. Тогда $c_{i,j} = \sum_{k=1}^{m} a_{i,k} b_{k,j}$. С другой стороны, положим $A^{\mathrm{T}} = (a'_{j,k})$, $B^{\mathrm{T}} = (b'_{k,i})$, $B^{\mathrm{T}} \cdot A^{\mathrm{T}} = (d_{j,i})$. Тогда $a'_{j,k} = a_{i,k}$, $b'_{k,i} = b_{j,k}$,

$$d_{j,i} = \sum_{k=1}^{m} b'_{k,i} a'_{j,k} = \sum_{k=1}^{m} a_{i,k} b_{k,j} = c_{i,j}. \qquad \square$$

Матричная запись систем линейных уравнений

Рассмотрим систему линейных уравнений

$$(\dagger) \qquad \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,m}x_m &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,m}x_m &= b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,m}x_m &= b_n \end{cases}$$

Как уже обсуждалось, ей соответствует матрица коэффициентов и столбец свободных членов:

$$A := \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix} \qquad B := \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{pmatrix}$$

Тогда система (†) может быть записана в компактном виде:

$$A \cdot X = B$$

где X – столбец неизвестных

$$X := \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_m \end{pmatrix}$$

Связь однородных и неоднородных систем линейных уравнений.

Для системы линейных уравнений (†) *ассоциированная однородная* система – это система

$$(\ddagger) \qquad \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,m}x_m &= 0 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,m}x_m &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,m}x_m &= 0 \end{cases}$$

полученная из (†) занулением свободных членов b_i.

Теорема. *Пусть* (†) *– любая система линейных уравнений и пусть* (‡) *– ассоциированная с ней однородная система.*

(1) *Если* $C = (c_1, \ldots, c_m)$ *– решение неоднородной системы* (†), *а* $H = (h_1, \ldots, h_m)$ *– решение ассоциированной однородной системы* (‡), *то* $C+H = (c_1+h_1, \ldots, c_m+h_m)$ *– также решение неоднородной системы* (†).

(2) *Если* $C = (c_1, \ldots, c_m)$ *и* $C = (c_1, \ldots, c_m)$ *– решения неоднородной системы* (†), *то разность* $C - C' = (c_1 - c'_1, \ldots, c_m - c'_m)$ *является решением ассоциированной однородной системы* (‡).

(3) *Зафиксируем некоторое решение* $C^\bullet = (c^\bullet_1, \ldots, c^\bullet_m)$ *неоднородной системы* (†). *Тогда любое решение системы* (†) *является суммой* $C^\bullet$ *и некоторого решения ассоциированной однородной системы* (‡).

Доказательство. Будем записывать решения в виде столбцов. для доказательства (1) заметим, что $A \cdot C = B$ и $A \cdot H = O$. Отсюда $A \cdot (C + H) = A \cdot C + A \cdot H = B + O = B$. Утверждение (2) доказывается аналогично. В этом случае $A \cdot (C - C') = A \cdot C - A \cdot C' = B - B = O$. Наконец, в условиях (3), пусть $C = (c_1, \ldots, c_m)$ – любое решение системы (†). Тогда согласно (2) разность $H := C - C^\bullet$ является решением системы (‡). Следовательно, $C = C^\bullet + H$. □

Следствие. *Рассмотрим однородную систему линейных уравнений* (‡).

(1) *Если* $C = (c_1, \ldots, c_m)$ *и* $D = (d_1, \ldots, d_m)$ *– решения* (‡), *то* $C+D = (c_1+d_1, \ldots, c_m+d_m)$ *– также решение* (‡) .

(2) *Если* $C = (c_1, \ldots, c_m)$ *– решение* (‡), *а* λ *– произвольное число, то* $\lambda C = (\lambda c_1, \ldots, \lambda c_m)$ *– также решение* (‡) .

Понятие кольца

Определение. *Кольцом* называется непустое множество R с двумя операциями: сложением $(+)$ и умножением $(\cdot)$ такими, что

(1) (a) $a + (b + c) = (a + b) + c$ для любых a, b, $c \in R$;

(b) существует элемент $0 \in R$ (нулевой элемент) такой, что $0+a = a+0 = a$ для любого $a \in R$;

(c) для любого $a \in R$ существует элемент $-a \in R$ (который называется противоположным к a) такой, что $a + (-a) = (-a) + a = 0$;

(d) $a + b = b + a$ для любых a, $b \in R$;

(2) для любых a, b, $c \in R$ имеем

(a) $a \cdot (b + c) = a \cdot b + a \cdot c$;

(b) $(a + b) \cdot c = a \cdot c + b \cdot c$.

Замечание. Из определения кольца непосредственно выводится, что $a \cdot 0 = 0$ и $0 \cdot a = 0$ для любого $a \in R$. Читателю предлагается проделать это самостоятельно.

Кольцо называется *ассоциативным*, если выполнено свойство

$$(a \cdot b) \cdot c = a \cdot (b \cdot c) \qquad \forall a,\, b,\, c \in R.$$

Кольцо называется *коммутативным*, если выполнено свойство

$$a \cdot b = b \cdot a \qquad \forall a,\, b \in R.$$

Единицей кольца называется элемент $1 \in R$ такой, что

$$1 \neq 0 \quad \text{и} \quad 1 \cdot a = a \cdot 1 = a \qquad \forall a \in R.$$

(если такой существует).

Определение. Подмножество Q в кольце R называется *подкольцом*, если оно является кольцом с теми же операциями, что и в R.

Примеры. (1) Множество всех целых $\mathbb{Z}$, рациональных $\mathbb{Q}$ и действительных $\mathbb{R}$ чисел являются кольцам с обычными операциями сложения и умножения, причем $\mathbb{Z}$ и $\mathbb{Q}$ – подкольца в $\mathbb{R}$, а $\mathbb{Z}$ – подкольцо в $\mathbb{Q}$.

(2) Множество $n\mathbb{Z}$ целых чисел, делящихся на n, является подкольцом в $\mathbb{Z}$.

(3) Множество рациональных чисел вида $\{a/b \in \mathbb{Q} \mid b \equiv 0 \mod n\}$ является подкольцом в $\mathbb{Q}$.

(4) Множество действительных чисел вида $\{a+b\sqrt{2} \in \mathbb{R} \mid a,\, b \in \mathbb{Q}\}$ является подкольцом в $\mathbb{R}$.

(5) Множество $\{a_n\}$ всех последовательностей действительных чисел с операциями покомпонентного сложения и умножения является кольцам.

(6) Зафиксируем некоторое множество M. Множество $\{f \mid M \to \mathbb{R}\}$ всех числовых функций на M является кольцом.

(7) Множество $\mathrm{C}[a,b]$ всех непрерывных функций на отрезке является кольцом.

(8) Трехмерное векторное пространства с операцией векторного умножения является кольцом.

(9) Пусть M – любое множество и пусть 2^M – множество всех его подмножеств. Тогда 2^M с операциями симметрической разности $A \triangle B := (A \cup B) \setminus (A \cap B)$ и пересечения $A \cap B$ является кольцом.

Из доказанных нами свойств матриц вытекает следующая.

Теорема. $\mathrm{Mat}_n(\mathbb{R})$ – *ассоциативное кольцо с единицей.*

Замечание. Для того, чтобы подмножество Q в кольце R было подкольцом необходимо и достаточно, чтобы выполнялись следующие условия:

(1) для любых элементов a, $b \in Q$ их разность $a - b$ принадлежит Q;

(2) для любых элементов a, $b \in Q$ их произведение $a \cdot b$ принадлежит Q.

Из последнего замечания легко видеть, что пересечение двух подколец является подкольцом. Более точно верно следующее утверждение:

Утверждение. *Пусть Q_i, $i \in I$ – любое множество подколец кольца R. Тогда их пересечение $\cap_{i\in I} Q_i$ также является подкольцом.*

Лекция 4

Определители. Определитель треугольной матрицы. Определитель транспонированной матрицы. Полилинейные и кососимметрические функции. Полилинейность и кососимметричность определителя. Матричные единицы. Их произведения. Элементарные матрицы. Умножение произвольной матрицы на элементарную. Разложение матрицы в произведение элементарных. Вычисление определителя при помощи элементарных преобразований. Определитель с углом нулей. Определитель Вандермонда. Определитель произведения матриц.

Определители

Определение. Пусть A – квадратная матрица

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \cdots\cdots\cdots\cdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$$

Ее *определителем* называется число

$$|A| = \sum_{\sigma \in \mathrm{S}_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}.$$

Пример. Пусть $n = 2$. Тогда S_n состоит из двух подстановок: тождественной ε и транспозиции $\tau = [1, 2]$. Поэтому

$$\begin{aligned} |A| = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} &= \mathrm{sgn}(\varepsilon)\, a_{1,\varepsilon(1)} a_{2,\varepsilon(2)} + \mathrm{sgn}(\tau)\, a_{1,\tau(1)} a_{2,\tau(2)} \\ &= a_{1,1}a_{2,2} - a_{1,2}a_{2,1}. \end{aligned}$$

Пример. Пусть $n = 3$. Тогда S_n состоит из шести подстановок: тождественной ε, двух тройных циклов $[1, 2, 3]$, $[1, 3, 2]$ и трех транспозиций $[i, j]$. Поэтому

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

Пример. Покажем, что если $A = (a_{i,j})$ – верхнетреугольная (нижнетреугольная) матрица, то ее определитель равен произведению диагональных элементов:

$$|A| = a_{1,1}a_{2,2}\cdots a_{n,n}.$$

Действительно, в формуле для определителя член $a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{n,\sigma(n)}$ отличен от нуля только если $\sigma(i) \geq i$ для любого i. Отсюда $\sigma(n) = n$ и поэтому $\sigma(n-1) \neq n$. Тогда $\sigma(n-1) = n-1$ и т. д. Получим, что единственный ненулевой член соответствует единичной подстановке.

Теорема. $|A| = |A^{\mathrm{T}}|$.

Доказательство. Заметим, что

$$\begin{pmatrix} \sigma(1) & \sigma(2) & \cdots & \sigma(n) \\ 1 & 2 & \cdots & n \end{pmatrix} = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma^{-1}(1) & \sigma^{-1}(2) & \cdots & \sigma^{-1}(n) \end{pmatrix}$$

Поэтому

$$\begin{aligned} |A^{\mathrm{T}}| &= \sum_{\sigma\in \mathrm{S}_n} \mathrm{sgn}(\sigma) a^{\mathrm{T}}_{1,\sigma(1)}a^{\mathrm{T}}_{2,\sigma(2)}\cdots a^{\mathrm{T}}_{n,\sigma(n)} = \\ &= \sum_{\sigma\in \mathrm{S}_n} \mathrm{sgn}(\sigma) a_{\sigma(1),1}a_{\sigma(2),2}\cdots a_{\sigma(n),n} = \\ &= \sum_{\sigma\in \mathrm{S}_n} \mathrm{sgn}(\sigma^{-1}) a_{1,\sigma^{-1}(1)}a_{2,\sigma^{-1}(2)}\cdots a_{n,\sigma^{-1}(n)} = \\ &= \sum_{\tau\in \mathrm{S}_n} \mathrm{sgn}(\tau) a_{1,\tau(1)}a_{2,\tau(2)}\cdots a_{n,\tau(n)} = |A|. \end{aligned}$$

Это и доказывает утверждение. □

Полилинейность определителя

Функция $F(X_1, \ldots, X_N)$ от нескольких аргументов называется *полилинейной*, если при подстановке вместо любой переменной X_i значения $\lambda' X_i' + \lambda'' X_i''$, где λ' и λ'' – произвольные числа, мы имеем

$$F(X_1, \ldots, \lambda' X_i' + \lambda'' X_i'', \ldots, X_N) =$$

$$= \lambda' F(X_1, \ldots, X_i', \ldots, X_N) + \lambda'' F(X_1, \ldots, X_i'', \ldots, X_N)$$

Рассмотрим матрицу A как совокупность ее строк

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \ldots\ldots\ldots\ldots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix} = \begin{pmatrix} A_1 \\ \cdots \\ A_n \end{pmatrix}, \quad \text{где} \quad A_i = \begin{pmatrix} a_{i,1} & \cdots & a_{i,n} \end{pmatrix},$$

а определитель $|A|$ рассмотрим как числовую функцию строк

$$|A| = |A_1, \ldots, A_n| = F(A_1, \ldots, A_n).$$

Теорема (полилинейность определителя)**.** *Определитель является полилинейной функцией своих строк.*

Доказательство. Пусть $A_i = \lambda' A'_i + \lambda'' A''_i$ для некоторого (фиксированного) i, где

$$A'_i = \begin{pmatrix} a'_{i,1} & a'_{i,2} & \dots & a'_{i,n} \end{pmatrix}, \qquad A''_i = \begin{pmatrix} a''_{i,1} & a''_{i,2} & \dots & a''_{i,n} \end{pmatrix}.$$

Таким образом, $a_{i,j} = \lambda' a'_{i,j} + \lambda'' a''_{i,j}$ для любого j. Поэтому

$$\begin{aligned} |A| &= \sum_{\sigma\in \mathrm{S}_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{i,\sigma(i)} \cdots a_{n,\sigma(n)} = \\ &= \sum_{\sigma\in \mathrm{S}_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} \cdots (\lambda' a'_{i,\sigma(i)} + \lambda'' a''_{i,\sigma(i)}) \cdots a_{n,\sigma(n)} = \\ &= \lambda' \sum_{\sigma\in \mathrm{S}_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} \cdots a'_{i,\sigma(i)} \cdots a_{n,\sigma(n)} + \\ &+ \lambda'' \sum_{\sigma\in \mathrm{S}_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} \cdots a''_{i,\sigma(i)} \cdots a_{n,\sigma(n)} = \lambda' |A'| + \lambda'' |A''|, \end{aligned}$$

где A' (соответственно, A'') – матрица, составленная из строк $A_1, \dots, A'_i, \dots, A_n$ (соответственно, $A_1, \dots, A''_i, \dots, A_n$). □

Кососимметричность определителя

Функция $F(X_1, \dots, X_N)$ от нескольких аргументов называется *кососимметрической*, если при подстановке двух любых переменных X_i и X_i, $i \neq j$ функция меняет знак:

$$\begin{aligned} & F(X_1, \dots, X_i, \dots, X_j, \dots, X_N) = \\ = \; & -F(X_1, \dots, X_j, \dots, X_i, \dots, X_N). \end{aligned}$$

Лемма. *Если $F(X_1, \dots, X_N)$ – полилинейная функция такая, что*

$$F(X_1, \dots, X_i, \dots, X_j, \dots, X_N) = 0 \quad \textit{при } X_i = X_i, \quad \textit{для любых } i \neq j,$$

то эта функция является кососимметрической.

Доказательство. Докажем утверждение для $N = 2$. Общий случай ничем не отличается. Имеем

$$\begin{aligned} 0 &= F(X_1 + X_2, X_1 + X_2) = F(X_1 + X_2, X_1) + F(X_1 + X_2, X_2) = \\ &= F(X_1, X_1) + F(X_2, X_1) + F(X_1, X_2) + F(X_2, X_2) = F(X_2, X_1) + F(X_2, X_2). \end{aligned}$$

Следовательно, $F(X_2, X_1) = -F(X_1, X_2)$. □

Лемма. *Если в матрице две строки совпадают, то ее определитель равен нулю.*

Доказательство. Пусть в матрице A совпадают i-ая и j-ая строки, т.е. $a_{i,k} = a_{j,k}$ для любого k, где $i < j$. Докажем, что $|A| = 0$. Разобьём все подстановки из S_n на (непересекающиеся) пары

$$\{\sigma,\ \sigma' = \sigma \circ [i,j]\}.$$

Так как $a_{i,\sigma(i)} = a_{j,\sigma(i)}$ и $a_{j,\sigma(j)} = a_{i,\sigma(j)}$, то соответствующие члены

$$\begin{aligned}&\operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{i,\sigma(i)} \cdots a_{j,\sigma(j)} \cdots a_{n,\sigma(n)},\\ &\operatorname{sgn}(\sigma') a_{1,\sigma'(1)} \cdots a_{i,\sigma'(i)} \cdots a_{j,\sigma'(j)} \cdots a_{n,\sigma'(n)} \quad = \quad -\operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{i,\sigma(j)} \cdots a_{j,\sigma(i)} \cdots a_{n,\sigma(n)}\end{aligned}$$

в формуле для определителя сокращаются. □

Теорема (кососимметричность определителя)**.** *Определитель является кососимметрической функцией своих строк.*

Матричные единицы.

*Матричной единицей** назовем матрицу $E_{i,j} = (e_{k,l})$, в которой элемент, стоящий на месте (i,j) равен 1, а все остальные элементы равны 0, т.е.

$$e_{k,s} = \begin{cases} 1 & \text{если } k = i \text{ и } l = j, \\ 0 & \text{если } k \neq i \text{ или } l \neq j. \end{cases}$$

При умножении матрицы A на матричную единицу $E_{i,j}$ слева в матрице $E_{i,j} \cdot A$ месте i-ой строки будет стоять j-ая строка матрицы A, а все остальные строки в $E_{i,j} \cdot A$ будут нулевыми:

$$(*) \qquad E_{i,j} \cdot \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,m} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i,1} & a_{i,2} & a_{i,3} & \cdots & a_{i,m} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{j,1} & a_{j,2} & a_{j,3} & \cdots & a_{j,m} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & a_{n,m} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{j,1} & a_{j,2} & a_{j,3} & \cdots & a_{j,m} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \leftarrow i$$

Действительно, пусть $E_{i,j}A = B = (b_{k,l})$. Тогда

$$b_{k,l} = \sum_{s=1}^{n} e_{k,s} a_{s,l} = e_{k,j} a_{j,l} = \begin{cases} a_{j,l} & \text{если } k = i, \\ 0 & \text{если } k \neq i. \end{cases}$$

Аналогично, при умножении матрицы A на матричную единицу $E_{i,j}$ справа в матрице $A \cdot E_{i,j}$ месте j-ого столбца будет стоять i-ый столбец матрицы A, а все остальные столбцы

*Не нужно путать матричные единицы с единичной матрицей

в $A \cdot E_{i,j}$ будут нулевыми:

$$A \cdot E_{i,j} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix} \cdot E_{i,j} = \begin{pmatrix} 0 & 0 & \cdots & \overset{j}{\downarrow} & \cdots & 0 \\ 0 & 0 & \cdots & a_{1,i} & \cdots & 0 \\ 0 & 0 & \cdots & a_{2,i} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{n,i} & \cdots & 0 \end{pmatrix}$$

В частности, имеет место следующее правило умножения матричных единиц:

Лемма.

$$E_{i,j} \cdot E_{k,l} = \begin{cases} 0 & \text{если } j \neq k, \\ E_{i,l} & \text{если } j = k. \end{cases}$$

Элементарные матрицы:

Элементарные матрицей называется матрица $U_{i,j,\lambda}$ (соответственно, $U_{i,\lambda}, U_{i,j}$), полученная из единичной элементарным преобразованием ($\mathrm{I_{i,j,\lambda}}$) (соответственно, ($\mathrm{II_{i,\lambda}}$), ($\mathrm{III_{i,j}}$)). Таким образом,

$$U_{i,j,\lambda} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ \cdots & \lambda & \cdots & 1 & \cdots & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{matrix} \qquad U_{i,j} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 0 & \cdots & 1 & \\ & & \vdots & \ddots & \vdots & \\ & & 1 & \cdots & 0 & \\ & & & & & \ddots \\ & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{matrix} \qquad U_{i,\lambda} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ \cdots & & \lambda & & \cdots & \\ & & & \ddots & & \\ & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{matrix}$$

$$\begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ j \end{matrix} \qquad \begin{matrix} \\ \\ \\ \\ \\ \\ \\ i \quad j \end{matrix} \qquad \begin{matrix} \\ \\ \\ \\ \\ \\ \\ i \end{matrix}$$

Можно также записать

- $U_{i,j,\lambda} = E + \lambda E_{i,j}$,
- $U_{i,\lambda} = E + (\lambda - 1)E_{i,i}$,
- $U_{i,j} = E - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}$.

Предложение. *Пусть $A \in \mathrm{Mat}_{n,m}(\mathbb{R})$, пусть U – одна из элементарных матриц $U_{i,j,\lambda}$, $U_{i,j,\lambda}$, $U_{i,j,\lambda}$. Тогда матрица $U \cdot A$ получается из A соответствующим элементарным преобразованием* ($\mathrm{I_{i,j,\lambda}}$), ($\mathrm{II_{i,\lambda}}$), ($\mathrm{III_{i,j}}$).

Доказательство. Рассмотрим случай ($\mathrm{I_{i,j,\lambda}}$). Имеем

$$U_{i,j,\lambda} \cdot A = (E + \lambda E_{i,j}) \cdot A = A + \lambda E_{i,j} \cdot A.$$

Согласно $(*)$ все строки матрицы $E_{i,j} \cdot A$, кроме i-ой нулевые, а i-ая строка матрицы равна j-ой строке матрицы A. Следовательно, $U_{i,j,\lambda} \cdot A$ получается из A прибавлением к i-ой строке i-ой, умноженной на λ. Случаи ($\mathrm{II_{i,\lambda}}$) и ($\mathrm{III_{i,j}}$) рассматриваются аналогично. $\square$

Следствие. *Пусть A – квадратная матрица.*

(1) *Если $|A| \neq 0$, то $A = U_1 \cdots U_m$, где $U_1, \ldots U_m$ – элементарные матрицы.*

(2) *Если $|A| = 0$, то $A = U_1 \cdots U_m A'$, где A' – матрица с нулевой строкой, а $U_1, \ldots U_m$ – элементарные матрицы.*

Доказательство. Действительно, элементарными преобразованиями строк матрица A приводится к улучшенному ступенчатому виду A'. При этом обратными элементарными преобразованиями (выполненными в обратном порядке) из матрицы A' мы получим матрицу A. Следовательно, $A = U_1 \cdots U_m A'$, где $U_1, \ldots U_m$ – соответствующий элементарные матрицы. Если $|A| \neq 0$, то $|A'| \neq 0$ и тогда $A' = E$. Если же $|A| = 0$, тоA' – матрица с нулевой строкой. □

Вычисление определителя при помощи элементарных преобразований.

Теорема. *Пусть матрица A' получена из матрицы A одним элементарным преобразованием строк $A \longmapsto A'$. Тогда*

(1) *Если $A \longmapsto A'$ – преобразование типа $(\mathrm{I}_{\mathrm{i,j},\lambda})$, то $|A'| = |A|$;*

(2) *Если $A \longmapsto A'$ – преобразование типа $(\mathrm{II}_{\mathrm{i},\lambda})$, то $|A'| = \lambda|A|$;*

(3) *Если $A \longmapsto A'$ – преобразование типа $(\mathrm{III}_{\mathrm{i,j}})$, то $|A'| = -|A|$.*

Доказательство. (1) Пусть A' получается из A применением преобразования $(\mathrm{I}_{\mathrm{i,j},\lambda})$. Рассматривая матрицы A и A' как совокупности своих строк:

$$A = (A_1, \cdots, A_n), \qquad A' = (A'_1, \cdots, A'_n).$$

Тогда $A'_i = A_i + \lambda A_j$ и $A'_k = A_k$ при $k \neq i$. Из свойства полилинейности определителя получаем

$$|A'| = |A| + \lambda|A''| = |A|,$$

где A'' – матрица со строками $A''_i = A_j$ и $A''_k = A_k$ при $k \neq i$, т.е. матрица с двумя одинаковыми строками.

(2), (3) Для преобразований типов (II) и (III) утверждение непосредственно следует из свойств полилинейности и кососимметричности определителя, соответственно. □

Таким образом, при вычислении определителя матрицу A элементарными преобразованиями строк можно привести к ступенчатому виду A'. При этом определитель матрицы не изменится: $|A'| = |A|$. Поскольку матрица A' – ступенчатая и квадратная, то она треугольная. Следовательно, ее определитель равен произведению элементов главной диагонали.

Определитель с углом нулей

Говорят, что квадратная матрица

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \cdots & \cdots & \cdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$$

имеет нижний *угол нулей* если для некоторого $1 \le k \le n$ имеем $a_{i,j}$ при $i > k$, $j \le k$. Аналогично определяется матрица с верхним углом нулей. Таким образом, матрица с нижним углом нулей может быть записана в блочном виде

$$(\dagger) \qquad A = \begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix}$$

где A_1 и A_2 – квадратные матрицы размеров $k \times k$ и $(n-k) \times (n-k)$, соответственно, B – матрица размера $k \times (n-k)$, а 0 – нулевая матрица размера $(n-k) \times k$.

Теорема. *Пусть A – матрица с нижним углом нулей, записанная как выше. Тогда $|A| = |A_1||A_2|$.*

Доказательство. Приведем матрицу A_1 (соответственно, A_2) элементарными преобразованиями строк типа (I) к ступенчатой матрице A_1' (соответственно, A_2'). Тогда $|A_1| = |A_1'|$ и $|A_2| = |A_2'|$. Пусть матрица A' получена из A выполнением соответствующих элементарных преобразований строк (т.е. над первыми k строками матрицы A мы выполняем те же преобразования, что и над строками A_1, а над следующими $n-k$ строками матрицы A мы выполняем те же преобразования, что и над строками A_2). Тогда $|A'| = |A|$ и поэтому достаточно доказать, что $|A'| = |A_1'||A_2'|$. Ясно, что

$$A' = \begin{pmatrix} A_1' & B' \\ 0 & A_2' \end{pmatrix}$$

Матрицы A_1' и A_2' – треугольные (ниже главной диагонали стоят нули) и таковой же является A'. В этом случае равенство $|A'| = |A_1'||A_2'|$ следует из того, что определитель треугольной матрицы равен произведению элементов на ее диагонали. □

Следствие. *Пусть A – матрица с верхним углом нулей, записанная в виде*

$$A = \begin{pmatrix} A_1 & 0 \\ B & A_2 \end{pmatrix}$$

Тогда $|A| = |A_1| \cdot |A_2|$.

Определитель Вандермонда

Теорема.

$$\Delta_n = \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{1 \le j < i \le n} (a_i - a_j).$$

Доказательство. Индукция по n. $n = 2$: $\Delta_2 = a_2 - a_1$. Шаг индукции. Начиная с последней вычтем из каждого столбца предыдущий, умноженный на a_n. Получим

$$\Delta_n = \begin{vmatrix} 1 & a_1 - a_n & a_1^2 - a_1 a_n & \cdots & a_1^{n-1} - a_1^{n-2} a_n \\ 1 & a_2 - a_n & a_2^2 - a_2 a_n & \cdots & a_2^{n-1} - a_2^{n-2} a_n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & a_{n-1} - a_n & a_{n-1}^2 - a_{n-1} a_n & \cdots & a_{n-1}^{n-1} - a_{n-1}^{n-2} a_n \\ 1 & 0 & 0 & \cdots & 0 \end{vmatrix}$$

Далее переставим строки. Получим определитель с верхним углом нулей:

$$\Delta_n = (-1)^{n-1} \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & a_1 - a_n & a_1^2 - a_1 a_n & \cdots & a_1^{n-1} - a_1^{n-2} a_n \\ 1 & a_2 - a_n & a_2^2 - a_2 a_n & \cdots & a_2^{n-1} - a_2^{n-2} a_n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & a_{n-1} - a_n & a_{n-1}^2 - a_{n-1} a_n & \cdots & a_{n-1}^{n-1} - a_{n-1}^{n-2} a_n \end{vmatrix} =$$

$$(-1)^{n-1} \begin{vmatrix} a_1 - a_n & a_1^2 - a_1 a_n & \cdots & a_1^{n-1} - a_1^{n-2} a_n \\ a_2 - a_n & a_2^2 - a_2 a_n & \cdots & a_2^{n-1} - a_2^{n-2} a_n \\ \cdots & \cdots & \cdots & \cdots \\ a_{n-1} - a_n & a_{n-1}^2 - a_{n-1} a_n & \cdots & a_{n-1}^{n-1} - a_{n-1}^{n-2} a_n \end{vmatrix}$$

Вынесем из каждой строки множитель $a_i - a_n$

$$\Delta_n = (-1)^{n-1} (a_1 - a_n)(a_2 - a_n) \cdots (a_{n-1} - a_n) \begin{vmatrix} 1 & a_1 & \cdots & a_1^{n-2} \\ 1 & a_2 & \cdots & a_2^{n-2} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & a_{n-1} & \cdots & a_{n-1}^{n-2} \end{vmatrix}$$

и воспользуемся предположением индукции

$$\Delta_n = (a_n - a_1)(a_n - a_2) \cdots (a_n - a_{n-1}) \Delta_{n-1} = \prod_{j<n} (a_n - a_j) \prod_{1 \le j < i \le n-1} (a_i - a_j). \qquad \square$$

Определитель произведения матриц

Лемма. *Если U – элементарная матрица, то*

$$|U \cdot A| = |U| \cdot |A|.$$

Доказательство. Пусть, например, U – элементарная матрица типа (I). Тогда $|U| = 1$ и $U \cdot A$ получается из A применением преобразования (I). Следовательно, $|U \cdot A| = |A| = |U| \cdot |A|$. Случаи (II) и (III) разбираются аналогично. $\square$

Следствие. *Если $U_1, \ldots U_m$ – элементарные матрицы, то*

$$|U_1 \cdots U_m \cdot A| = |U_1 \cdots U_m| \cdot |A|.$$

Теорема. *Пусть A, $B \in \mathrm{Mat}_n(\mathbb{R})$. Тогда $|A| \cdot |B| = |A \cdot B|$.*

Доказательство. Предположим, что $|A| = 0$. Запишем $A = U_1 \cdots U_m A'$, где A' – матрица с нулевой строкой, а $U_1, \ldots U_m$ – элементарные матрицы. Тогда $A'B$ также имеет нулевую строку. Значит

$$|A \cdot B| = |(U_1 \cdots U_m \cdot A') \cdot B| = |U_1 \cdots U_m| \cdot |A' \cdot B| = 0 = |A| \cdot |B|.$$

Пусть $|A| \neq 0$. Тогда $A = U_1 \cdots U_m$, где $U_1, \ldots U_m$ – элементарные матрицы и

$$|A \cdot B| = |(U_1 \cdots U_m) \cdot B| = |U_1 \cdots U_m| \cdot |B| = |A| \cdot |B|. \qquad \square$$

Лекция 5

Эквивалентное определение определителя (как полилинейной кососимметрической формы). Разложение определителя по строке (и фальшивое разложение). Теорема Крамера.

Эквивалентное определение определителя (как полилинейной кососимметрической формы)

Теорема. *Пусть* $F(A_1, \ldots, A_n)$ *– функция*

$$F : \mathrm{Mat}_n(\mathbb{R}) \to \mathbb{R},$$

аргументами которой являются n *строк матрицы* $A = A_1, \ldots, A_n$. *Предположим, что она полилинейная и косо симметрична и пусть* $c := F(E)$. *Тогда*

$$F(A) = c|A|$$

для любой матрицы $A \in \mathrm{Mat}_n(\mathbb{R})$, *т.е. функция* F *пропорциональна определителю.*

Доказательство. Следующая лемма доказывается также как и соответствующий факт для определителей.

Лемма. *Пусть матрица* A' *получена из матрицы* A *одним элементарным преобразованием строк* $A \longmapsto A'$. *Тогда*

(1) *Если* $A \longmapsto A'$ *– преобразование типа* $(\mathrm{I}_{\mathrm{i,j},\lambda})$, *то* $F(A') = F(A)$;

(2) *Если* $A \longmapsto A'$ *– преобразование типа* $(\mathrm{II}_{\mathrm{i},\lambda})$, *то* $F(A') = \lambda F(A)$;

(3) *Если* $A \longmapsto A'$ *– преобразование типа* $(\mathrm{III}_{\mathrm{i,j}})$, *то* $F(A') = -F(A)$.

Приведем матрицу A элементарными преобразованиями строк к улучшенному ступенчатому виду

$$A = A^{(1)} \longmapsto A^{(2)} \longmapsto \cdots \longmapsto A^{(m)} =: B.$$

Предположим, что $|A| = 0$. Тогда матрица B имеет нулевую строку и $|B| = 0$. Согласно сказанному выше

$$F(A) = F(A^{(2)}) = \cdots = F(A^{(m-1)}) = F(B) = 0 = c|A|.$$

Пусть $|A| \neq 0$. Тогда $|A^{(i)}| \neq 0$ для всех $i = 1, \ldots, m$ и $B = E$. Снова, согласно сказанному выше, отношение $F(A^{(i)})/|A^{(i)}|$ сохраняется:

$$\frac{F(A)}{|A|} = \frac{F(A^{(2)})}{|A^{(2)}|} = \cdots = \frac{F(A^{(i)})}{|A^{(i)}|} = \frac{F(A^{(i+1)})}{|A^{(i+1)}|} = \cdots = \frac{F(E)}{|E|} = c. \qquad \square$$

Следствие. *Пусть* $\mathrm{Mat}_n(\mathbb{R}) \to \mathbb{R}$ *– полилинейная кососимметричная функция строк матриц такая, что* $F(E) = 1$. *Тогда* $F(A) = |A|$.

Разложение определителя по строке

Пусть $A = (a_{i,j})$ – $n \times m$-матрица. *Минором* порядка r (где $1 \leq r \leq n$ и $1 \leq r \leq m$) называется определитель матрицы из элементов, стоящих на пересечении некоторых r строк и r столбцов матрицы A. Если теперь $A = (a_{i,j})$ – $n \times n$-матрица, то через $M_{i,j}$ мы будем обозначать минор порядка $n-1$, полученный из A вычеркиванием i-ой строки и j-ого столбца. В этом случае *алгебраическим дополнением* к элементу $a_{i,j}$ называется число $A_{i,j} := (-1)^{i+j} M_{i,j}$.

Теорема (разложение определителя по строке)**.**

$$\sum_{j=1}^{n} a_{i,j} A_{k,j} = 0 \quad \textit{при } k \neq i.$$

Доказательство. Пусть $S = (a_{i,1}, \ldots, a_{i,n})$ – i-ая строка матрицы. Рассмотрим строки $S_j = (0, \ldots, a_{i,j}, \ldots, 0)$. Тогда $S = \sum S_i$. Пусть B_j – матрица, полученная из A заменой S на S_j. Согласно свойству полилинейности $|A| = \sum |B_j|$, где

$$|B_j| = \begin{vmatrix} a_{1,1} \cdots\cdots a_{1,j} \cdots\cdots a_{1,n} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{i-1,1} \cdots\cdots a_{i-1,j} \cdots\cdots a_{i-1,n} \\ 0 \cdots\cdots\cdots a_{i,j} \cdots\cdots\cdots 0 \\ a_{i+1,1} \cdots\cdots a_{i+1,j} \cdots\cdots a_{i+1,n} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n,1} \cdots\cdots a_{n,j} \cdots\cdots a_{n,n} \end{vmatrix} = (-1)^{i-1}(-1)^{j-1} \begin{vmatrix} a_{i,j} & 0 \cdots\cdots\cdots 0 \\ a_{1,1} & a_{1,2} \cdots\cdots a_{1,n} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{i-1,1} & a_{i-1,2} \cdots\cdots a_{i-1,n} \\ a_{i+1,1} & a_{i+1,2} \cdots\cdots a_{i+1,n} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n,1} & a_{n,2} \cdots\cdots a_{n,n} \end{vmatrix}$$

Здесь последний определитель получается из определителя матрицы B_i перестановками строк ($i-1$ раз) и перестановками столбцов ($j-1$ раз). Последний определитель имеет верхний угол нулей. Поэтому $|B_j| = (-1)^{i+j} M_{i,j}$. $\square$

Теорема (фальшивое разложение по строке)**.**

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

Доказательство. Рассмотрим матрицу

$$A' = (a'_{r,j}), \quad \text{где} \quad a'_{r,j} = \begin{cases} a_{r,j} & \text{при } r \neq k \\ a_{i,j} & \text{при } r = k \end{cases}$$

Мы в матрице A заменили k-ую строку на i-ю. Таким образом, в A' строки с номерами k и i совпадают. Поэтому $|A'| = 0$. Разлагая A' по k-й строке получим

$$0 = |A'| = \sum_{j=1}^{n} a'_{k,j} A'_{k,j} = \sum_{j=1}^{n} a_{i,j} A_{k,j}.$$

□

Теорема Крамера

Теорема. *Рассмотрим квадратную систему линейных уравнений*

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,n} = b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n,1}x_1 + \cdots + a_{n,n} = b_n \end{cases}$$

Положим

$$\Delta := \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \cdots\cdots & \cdots & \cdots\cdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix} \qquad \Delta_i := \begin{vmatrix} a_{1,1} & \cdots & b_1 & \cdots & a_{1,n} \\ \cdots\cdots & \cdots & \cdots & \cdots & \cdots\cdots \\ a_{n,1} & \cdots & b_n & \cdots & a_{n,n} \end{vmatrix}$$

Тогда

(1) *система совместна и определенна тогда и только тогда, когда* $\Delta \neq 0$,

(2) *если* $\Delta \neq 0$, *то решение системы имеет вид* $x_i = \Delta_i/\Delta$.

Доказательство. (1) Приведем матрицу системы элементарными преобразованиями строк к ступенчатому виду:

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} & b_1 \\ \cdots\cdots & \cdots & \cdots\cdots & \cdots \\ a_{n,1} & \cdots & a_{n,n} & b_n \end{pmatrix} \longmapsto \begin{pmatrix} a'_{1,1} & \cdots & a'_{1,n} & b'_1 \\ \cdots\cdots & \cdots & \cdots\cdots & \cdots \\ 0 & \cdots & a'_{n,n} & b'_n \end{pmatrix}$$

Пусть

$$\Delta' := \begin{vmatrix} a'_{1,1} & \cdots & a'_{1,n} \\ \cdots\cdots & \cdots & \cdots\cdots \\ 0 & \cdots & a'_{n,n} \end{vmatrix}$$

Определитель $\Delta \neq 0$ отличен от нуля тогда и только тогда, когда отличен от нуля $\Delta' \neq 0$. А это выполнено тогда и только тогда, когда $a'_{i,i} \neq 0$ для любого i. В свою очередь, это выполнено тогда и только тогда, когда система совместна и определенна.

(2) Положим $\alpha_j := \Delta_j/\Delta$. Тогда

$$\begin{aligned} \sum_{j=1}^{n} a_{i,j}\alpha_j &= \frac{1}{\Delta}\sum_{j=1}^{n} a_{i,j}\Delta_j = \frac{1}{\Delta}\sum_{j=1}^{n} a_{i,j}\sum_{k=1}^{n} b_k A_{k,j} = \\ &= \frac{1}{\Delta}\sum_{k=1}^{n} b_k \sum_{j=1}^{n} a_{i,j}A_{k,j} = \frac{b_i\Delta}{\Delta} = b_i \end{aligned}$$

поскольку

$$\sum_{j=1}^{n} a_{i,j}A_{k,j} = \begin{cases} \Delta & \text{если } k = i \\ 0 & \text{если } k \neq i \end{cases}$$

Таким образом, $\Delta \neq 0$, то система совместна и $x_i = \Delta_i/\Delta$ – решение. $\square$

Лекция 6

Единицы и обратные элементы в ассоциативном кольце. Делители нуля в кольце. Обратная матрица. Критерий существования обратной матрицы. Формула для обратной матрицы. Вычисление обратной матрицы при помощи элементарных преобразований. Делители нуля в кольце матриц. Вырожденные и невырожденные матрицы. Группы $\mathrm{GL}_n(\mathbb{R})$ и $\mathrm{SL}_n(\mathbb{R})$.

Пусть R – произвольное кольцо. Чтобы не рассматривать тривиальные случаи, мы будем считать, что R содержит по крайней мере один ненулевой элемент.

Элемент $a \in R$ называется *левым делителем нуля*, если существует ненулевой элемент $b \in R$. такой, что $a \cdot b = 0$. Соответственно, $b \in R$ называется *правым делителем нуля*, если существует $a \in R$, $a \neq 0$ такой, что $a \cdot b = 0$.

Примеры. (1) Если R – кольцо с нулевым умножением, то все его элементы – делители нуля.

(2) В кольце $\mathbb{Z}$ целых чисел 0 – единственный делитель нуля.

(3) Пусть R – кольцо всех вещественнозначных функций $f : M \to \mathbb{R}$ на множестве M. Если функция обращается в нуль на каком-то элементе $a \in M$, то она – делитель нуля в R.

Пусть R – кольцо с единицей 1 и пусть $a \in R$ – произвольный элемент. *Обратным* к a называется элемент $a^{-1} \in R$ такой, что $a \cdot a^{-1} = a^{-1} \cdot a = 1$. Если для $a \in R$ существует обратный элемент, то a называется *обратимым*.

Примеры. (1) В кольце целых чисел $\mathbb{Z}$ обратимыми являются только элементы ± 1.

(2) В кольце всех действительных чисел $\mathbb{R}$ обратимы все элементы, кроме 0.

Предложение. *Пусть R – ассоциативное кольцо.*

(1) *Если единичный элемент существует, то он единственный.*

(2) *Если для $a \in R$ существует обратный элемент, то он единственный.*

(3) *Если элемент $a \in R$ обратим, то он не может быть делителем нуля.*

(4) *Если элемент $a \in R$ обратим, то элемент a^{-1} обратим и $(a^{-1})^{-1} = a$.*

(5) *Если элементы a, $b \in R$ обратимы, то элемент $b \cdot a$ также обратим и $(b \cdot a)^{-1} = a^{-1} \cdot b^{-1}$.*

Обратная матрица

Обратной матрицей к квадратной матрице A называется квадратная матрица A^{-1} такая, что

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

где E – как обычно, единичная матрица соответствующего размера. Таким образом, Обратная матрица – это обратный элемент в кольце матриц $\mathrm{Mat}_n(\mathbb{R})$. Согласно предыдущему предложению, обратная матрица единственна, если она существует. Если у матрицы A существует обратная матрица, то мы будем говорить, что A *обратима*.

Заметим, что элементарные матрицы обратимы и матрица обратная к элементарной – также элементарная. Действительно,

$$U_{i,j,\lambda} \cdot U_{i,j,-\lambda} = E, \qquad U_{i,\lambda} \cdot U_{i,\lambda^{-1}} = E, \qquad U_{i,j} \cdot U_{i,j} = E.$$

Критерий обратимости матрицы

Присоединенной матрицей к квадратной матрице $A = (a_{i,j})$ называется матрица $\hat{A} = (\hat{a}_{i,j})$, где $\hat{a}_{i,j} = A_{j,i}$ – алгебраическое дополнение к транспонированному элементу.* Таким образом,

$$\hat{A} = \begin{pmatrix} A_{1,1} & \cdots & A_{n,1} \\ \cdots & \cdots & \cdots \\ A_{1,n} & \cdots & A_{n,n} \end{pmatrix}$$

Теорема (Формула для обратной матрицы)**.** *Пусть $A \in \mathrm{Mat}_n(\mathbb{R})$. Следующие условия эквивалентны:*

(1) $|A| \neq 0$*;*

(2) *существует обратная матрица A^{-1} (т.е. A обратима).*

Если эти условия выполнены, то

$$(*) \qquad A^{-1} = \frac{1}{|A|}\hat{A}.$$

Матрица, удовлетворяющая этим условиям, называются *невырожденной*, в противном случае она называется *вырожденной*.

Непосредственно из определения и свойств определителя следует, что для невырожденной матрицы имеет место равенство

$$|A^{-1}| = \frac{1}{|A|}.$$

*Это немного отличается от того, то было на лекции.

Доказательство. Запишем $A \cdot \hat{A} = (c_{i,j})$. Тогда используя формулы разложения определителя по строке и фальшивое разложение, мы можем записать

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k}\hat{a}_{k,j} = \sum_{k=1}^{n} a_{i,k}A_{j,k} = \begin{cases} |A| & \text{если } i = j, \\ 0 & \text{если } i \neq j. \end{cases}$$

Таким образом, $A \cdot \hat{A} = |A|E$. Если $|A| \neq 0$, то $A \cdot \frac{1}{|A|}\hat{A} = E$, т.е. A^{-1} существует и выполнено $(*)$.

Пусть для A существует обратная матрица. Тогда $1 = |E| = |A \cdot A^{-1}| = |A| \cdot |A^{-1}|$. Отсюда следует, что $|A| \neq 0$. □

Следствие. *Пусть для квадратной матрицы A существует квадратная матрица A' такая, что $A' \cdot A = E$. Тогда матрица A невырождена и $A^{-1} = A'$.*

Доказательство. Так как $1 = |E| = |A' \cdot A| = |A'| \cdot |A|$, то существует обратная матрица A^{-1}. Тогда $A^{-1} = E \cdot A^{-1} = A' \cdot A \cdot A^{-1} = A' \cdot E = A'$. □

Пример. Пусть $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Тогда

$$A^{-1} = \begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}.$$

Вычисление обратной матрицы при помощи элементарных преобразований

Пусть A – невырожденная матрица. Приведем ее элементарными преобразованиями строк к улучшенному ступенчатому виду

$$A = A^{(1)} \longmapsto A^{(2)} \longmapsto \cdots \longmapsto A^{(m)}.$$

Так как $|A| \neq 0$, то $|A^{(m)}| \neq 0$. Следовательно, $A^{(m)} = E$. Таким образом, последовательно умножая A на элементарные матрицы слева, мы получим единичную матрицу:

$$(U_1 \cdot U_2 \cdots U_m) \cdot A = U_1 \cdot U_2 \cdots U_m \cdot A = E.$$

Отсюда следует, что $A^{-1} = U_1 \cdot U_2 \cdots U_m = U_1 \cdot U_2 \cdots U_m \cdot E$. Это означает, что A^{-1} получается из E последовательным умножением на те же элементарные матрицы слева. Иначе говоря, A^{-1} получается из E последовательным применением тех же элементарных преобразований.

Практически, это можно проделать следующим образом. Рассмотрим матрицу размера $n \times 2n$, полученную из A приписыванием справа единичной матрицы:

$$B := \left(A \;\middle|\; E\right)$$

Последовательным применением элементарных преобразований строк приведем всю матрицу B к улучшенному ступенчатому виду. Тогда в левой части мы получим единичную матрицу, а в правой – обратную к A:

$$B' = \left(E \;\middle|\; A^{-1}\right)$$

Делители нуля в кольце матриц

Теорема. *Пусть* $A \in \mathrm{Mat}_n(\mathbb{R})$, $A \neq 0$. *Следующие условия эквивалентны:*

(1) $|A| = 0$,

(2) A – *левый делитель нуля в* $\mathrm{Mat}_n(\mathbb{R})$,

(3) A – *правый делитель нуля в* $\mathrm{Mat}_n(\mathbb{R})$,

(4) *не существует обратной матрицы к* A.

Напомним, что матрицы, удовлетворяющие этим условиям, называются *вырожденными*.

Доказательство. (1) $\Longrightarrow$ (2) Пусть $|A| = 0$. По теореме Крамера существует столбец X такой, что $A \cdot X = 0$. Пусть $B \in \mathrm{Mat}_n(\mathbb{R})$ – матрица у которой первый столбец совпадает с X, а остальные столбцы – нулевые. Тогда $A \cdot B = 0$.

(1) $\Longrightarrow$ (3) Пусть $|A| = 0$. Тогда $|A^{\mathrm{T}}| = 0$ и A^{T} – левый делитель нуля, т.е. $A^{\mathrm{T}} \cdot B = 0$. Отсюда $B^{\mathrm{T}} \cdot A = 0$.

(2)$\Longrightarrow$ (4) Иначе $A \cdot B = 0$, $0 = A^{-1} \cdot A \cdot B = B$.

(3) $\Longrightarrow$ (4) Аналогично.

(4) $\Longrightarrow$ (1) Иначе $|A| \neq 0$ и обратная матрица существует по формуле для обратной матрице. $\square$

Группы $\mathrm{GL}_n(\mathbb{R})$ и $\mathrm{SL}_n(\mathbb{R})$

Предложение. (1) *Множество* $\mathrm{GL}_n(\mathbb{R}) \subset \mathrm{Mat}_n(\mathbb{R})$ *всех невырожденных квадратных матриц порядка* n *образует группу с операцией умножения.*

(2) *Подмножество* $\mathrm{SL}_n(\mathbb{R}) \subset \mathrm{GL}_n(\mathbb{R})$, *состоящее из матриц с определителем* 1, – *подгруппа в* $\mathrm{GL}_n(\mathbb{R})$.

Группа $\mathrm{GL}_n(\mathbb{R})$ называется *полной линейной группой* степени n, а ее подгруппа $\mathrm{SL}_n(\mathbb{R})$ – *специальной линейной группой* степени n.

Доказательство. Обозначим через $\mathrm{GL}_n(\mathbb{R})$ множество невырожденных квадратных матриц порядка n. Несложно видеть, что оно замкнуто относительно умножения: если $A, B \in \mathrm{GL}_n(\mathbb{R})$, т.е. A и B – обратимые матрицы, то, как нами было показано, матрица $A \cdot B$ тоже обратима. Поэтому $A \cdot B \in \mathrm{GL}_n(\mathbb{R})$. Таким образом, умножение матриц – операция на $\mathrm{GL}_n(\mathbb{R})$. Ассоциативность умножения в $\mathrm{GL}_n(\mathbb{R})$ следует из того, что ассоциативность имеет место для любых матриц в $\mathrm{Mat}_n(\mathbb{R})$. Далее $E \in \mathrm{GL}_n(\mathbb{R})$ и для любой матрицы $A \in \mathrm{GL}_n(\mathbb{R})$ ее обратная также лежит в $\mathrm{GL}_n(\mathbb{R})$ поскольку обратная матрица A^{-1} также обратима.

Подмножество $\mathrm{SL}_n(\mathbb{R}) \subset \mathrm{GL}_n(\mathbb{R})$ является подгруппой поскольку для любых матриц $A, B \in \mathrm{SL}_n(\mathbb{R})$ мы имеем $|A| = |B| = 1$. Следовательно, $|A \cdot B| = |A| \cdot |B| = 1$ и $|A^{-1}| = 1/|A| = 1$. $\square$

Замечание. Нами было доказано, что полная линейная группа $\mathrm{GL}_n(\mathbb{R})$ порождается элементарными матрицами.

Определение. Квадратная матрица A называется *ортогональной*, если $A \cdot A^{\mathrm{T}} = E$.

Очевидно, что ортогональная матрица обратима и ее обратная совпадает с транспонированной: $A^{-1} = A^{\mathrm{T}}$.

Примеры. (1) Множество всех ортогональных $n \times n$-матриц является подгруппой в $\mathrm{GL}_n(\mathbb{R})$ и обозначается через $\mathrm{O}_n(\mathbb{R})$ (ортогональная группа).

(2) Множество всех ортогональных $n \times n$-матриц с определителем 1 является подгруппой в $\mathrm{O}_n(\mathbb{R})$ и в $\mathrm{SL}_n(\mathbb{R})$. Она обозначается через $\mathrm{SO}_n(\mathbb{R})$ (специальная ортогональная группа).

Следующий факт доказывается в точности также, как пункт (1) последнего предложения. Читателю предлагается привести доказательство самостоятельно.

Предложение. *Пусть R – ассоциативное кольцо с единицей и пусть R^* – множество всех его обратимых элементов. Тогда R^* – группа (с операцией умножения).*

Лекция 7

Векторные пространства. Примеры. Линейная зависимость. Критерий невырожденности матрицы. Базис. Координаты. Лемма о линейной зависимости. Следствия. Размерность и ранг. Ранг матрицы.

Векторные пространства.

Определение. *Векторным пространством* называется множество V, а котором задана операция сложения элементов V между собой

для любых $\mathbf{v}$, $\mathbf{w} \in V$, задан элемент $\mathbf{v} + \mathbf{w} \in V$

и операция умножения элементов $\mathbb{R}$ на элементы V

для любого $\lambda \in \mathbb{R}$ для любого $\mathbf{v} \in V$, задан элемент $\lambda\mathbf{v} \in V$,

такие, что выполняются следующие свойства:

(I) (1) $\mathbf{v} + (\mathbf{w} + \mathbf{u}) = (\mathbf{v} + \mathbf{w}) + \mathbf{u}$ для любых $\mathbf{v}$, $\mathbf{w}$, $\mathbf{u} \in V$;
(2) существует элемент $\mathbf{0} \in V$ $\mathbf{v} + \mathbf{0} = \mathbf{0} + \mathbf{v} = \mathbf{v}$ для любого $\mathbf{v} \in V$;
(3) для любого $\mathbf{v} \in V$ существует $-\mathbf{v} \in V$ $\mathbf{v} + (-\mathbf{v}) = (-\mathbf{v}) + \mathbf{v} = \mathbf{0}$;
(4) $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$ для любых $\mathbf{v}$, $\mathbf{w} \in V$.

(II) (1) $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$ для любого $\mathbf{v} \in V$, для любых α, $\beta \in \mathbb{R}$;
(2) $\alpha(\mathbf{v} + \mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w}$ для любых $\mathbf{v}$, $\mathbf{w} \in V$, для любого $\alpha \in \mathbb{R}$;
(3) $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v})$ для любого $\mathbf{v} \in V$, для любых α, $\beta \in \mathbb{R}$;
(4) $1\,\mathbf{v} = \mathbf{v}$ для любого $\mathbf{v} \in V$.

При этом элементы V называются *векторами*, а элементы $\mathbb{R}$ – *скалярами*.

Замечание. Непосредственно из определения получаем.

- Для любых $\mathbf{a}$, $\mathbf{b} \in V$ уравнение $\mathbf{a} + \mathbf{x} = \mathbf{b}$ имеет единственное решение.
- Нулевой элемент $\mathbf{0} \in V$ единственен.
- Для любого $\mathbf{a} \in V$ противоположный элемент $-\mathbf{a}$ единственен.

- Для любого $\mathbf{a} \in V$ имеем $0\,\mathbf{a} = \mathbf{0}$. Действительно, $\mathbf{a} + 0\,\mathbf{a} = 1\,\mathbf{a} + 0\,\mathbf{a} = (1+0)\mathbf{a} = 1\,\mathbf{a} = \mathbf{a}$.
- Для любого $\alpha \in \mathbb{R}$ имеем $\alpha\,\mathbf{0} = \mathbf{0}$. Действительно, $\alpha\,\mathbf{0} + \alpha\,\mathbf{a} = \alpha(\mathbf{0} + \mathbf{a}) = \alpha\,\mathbf{a}$.

Примеры. (1) $V = \{0\}$.

(2) Пусть $\mathbb{R}^n = \{(\alpha_1, \dots, \alpha_n)\}$ – множество строк длины n, где $\alpha_i \in \mathbb{R}$. Сложение и умножение на числа – покомпонентные. Аналогично определяется $\mathbb{Q}^n \dots$

(3) множество $\mathrm{Mat}_{n,m}(\mathbb{R})$ всех $n \times m$-матриц;

(4) геометрические векторы в двумерном (соответственно, трехмерном) пространстве;

(5) множество всех (бесконечных) последовательностей $(a_1, a_2, \dots, a_n, \dots)$;

(6) множество всех числовых функций $f : M \to \mathbb{R}$ на некотором множестве M.

Подмножество W векторного пространства V называется подпространством, если оно является пространством с теми же операциями (сложения и умножения на скаляры), что и в пространстве V. Несложно проверить, что для того, чтобы произвольное непустое подмножество $W \subset V$ было подпространством необходимо и достаточно, чтобы для любых векторов $\mathbf{a}$, $\mathbf{b} \in W$ и любого числа $\lambda \in \mathbb{R}$ было выполнено $\mathbf{a} + \mathbf{b} \in W$ и $\lambda\mathbf{a} \in W$.

Определение. Пусть V – векторное пространство и пусть $\mathbf{a}_1, \dots, \mathbf{a}_m \in V$ – конечный набор векторов. *Линейной комбинацией* векторов $\mathbf{a}_1, \dots, \mathbf{a}_m \in V$ с коэффициентами $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ называется выражение вида

$$\alpha_1\mathbf{a}_1 + \dots + \alpha_m\mathbf{a}_m \in V.$$

Линейная комбинация называется *тривиальной*, если $\alpha_1 = \dots = \alpha_m = 0$.

Определение. Векторы $\mathbf{a}_1, \dots, \mathbf{a}_m \in V$ называются *линейно зависимыми*, если их некоторая нетривиальная линейная комбинация равна нулю:

$$\alpha_1\mathbf{a}_1 + \dots + \alpha_m\mathbf{a}_m = \mathbf{0}, \quad \text{где } \alpha_i \neq 0 \text{ для некоторого } i.$$

Векторы не являющиеся линейно зависимыми, называются *линейно независимыми*. Это условие можно переформулировать следующим образом:

- $\alpha_1\mathbf{a}_1 + \dots + \alpha_m\mathbf{a}_m = \mathbf{0}$, то $\alpha_1 = \dots = \alpha_m = 0$ (если линейная комбинация равна нулю, то она тривиальна).

Примеры. (1) Один вектор $\mathbf{a}_1$ линейно зависим тогда и только тогда, когда $\mathbf{a}_1 = \mathbf{0}$.

(2) Если $\mathbf{a}_i = \mathbf{0}$, то система $\mathbf{a}_1, \dots, \mathbf{a}_m$ линейно зависима.

(3) Два вектора $\mathbf{a}_1$, $\mathbf{a}_2$ линейно зависимы тогда и только тогда, когда они пропорциональны, т.е. или $\mathbf{a}_1 = \lambda\mathbf{a}_2$ или $\mathbf{a}_2 = \lambda\mathbf{a}_1$ для некоторого $\lambda \in \mathbb{R}$.

Замечание. (1) Если подсистема $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_k}\} \subset \{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ линейно зависима, то и вся система $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ линейно зависима.

(2) Если система $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ линейно независима, то и любая ее подсистема $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_k}\}$ также линейно независима.

Предложение. *Система векторов $\mathbf{a}_1, \ldots, \mathbf{a}_m \in V$ линейно зависима тогда и только тогда, когда некоторый ее вектор $\mathbf{a}_j$ выражается как линейная комбинация остальных.*

Доказательство. Пусть система $\mathbf{a}_1, \ldots, \mathbf{a}_m \in V$ линейно зависима. Тогда существуют $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ такие, что имеет место равенство $\alpha_1\mathbf{a}_1 + \cdots + \alpha_m\mathbf{a}_m = \mathbf{0}$, где $\alpha_j \neq 0$ для некоторого j. Значит

$$\mathbf{a}_j = -\frac{\alpha_1}{\alpha_j}\mathbf{a}_1 - \cdots - \frac{\alpha_{j-1}}{\alpha_j}\mathbf{a}_{j-1} - \frac{\alpha_{j+1}}{\alpha_j}\mathbf{a}_{j+1} \cdots - \frac{\alpha_m}{\alpha_j}\mathbf{a}_m.$$

Наоборот, пусть

$$\mathbf{a}_j = \lambda_1\mathbf{a}_1 + \cdots + \lambda_{j-1}\mathbf{a}_{j-1} + \lambda_{j-1}\mathbf{a}_{j-1} \cdots + \lambda_m\mathbf{a}_m.$$

Тогда

$$\lambda_1\mathbf{a}_1 + \cdots + \lambda_{j-1}\mathbf{a}_{j-1} - \mathbf{a}_j + \lambda_{j+1}\mathbf{a}_{j+1} \cdots + \lambda_m\mathbf{a}_m = \mathbf{0}. \qquad \square$$

Критерий невырожденности матрицы

Теорема. *Пусть A – квадратная матрица. Следующие условия эквивалентны:*

(1) *A невырождена;*

(2) *столбцы A линейно независимы;*

(3) *строки A линейно независимы.*

Доказательство. Пусть $A_1, \ldots, A_n$ – столбцы матрицы. Запишем условие линейной зависимости $\lambda_1A_1 + \cdots + \lambda_nA_n = 0$, где $\lambda_i \in \mathbb{R}$, а 0 – нулевой столбец. Условие может быть переписано в виде

$$A \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Это однородная система линейных уравнений. По теореме Крамера $|A| \neq 0$ тогда и только тогда, когда система определенна (т.е. $\lambda_1 = \cdots = \lambda_n = 0$ – единственное решение), а это имеет место тогда и только тогда, когда столбцы $A_1, \ldots, A_n$ линейно независимы. Это доказывает эквивалентность (1) $\Longleftrightarrow$ (2). Так как $|A| = |A^{\mathrm{T}}|$, то аналогично получаем (1) $\Longleftrightarrow$ (3). $\square$

Базисы

Определение. Пусть V – векторное пространство и пусть $M \subset V$ – любое подмножество (система векторов). Говорят, что множество элементов $\{\mathbf{e}_i \in M \mid i \in I\}$ образует *базис* M, если

(1) для любого конечного набора индексов $\{i_1, \dots, i_k\} \subset I$ векторы $\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_k}$ линейно независимы,

(2) для любого $\mathbf{v} \in M$ найдется конечное подмножество $\{i_1, \dots, i_k\} \subset I$ такое, что векторы $\mathbf{v}, \mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_k}$ линейно зависимы.

Базисом пространства V называется базис $M = V$.

Пример. В пространстве $\mathbb{R}^n$ имеется стандартный базис

$$\mathbf{e}_1 = (1, \dots\dots, 0), \dots, \mathbf{e}_i = (0, \dots, \underset{\substack{\uparrow \\ i}}{1}, \dots, 0), \dots, \mathbf{e}_n = (0, \dots\dots, 1).$$

Если в пространстве имеется базис из конечного числа элементов, то пространство называется *конечномерным*. Мы будем рассматривать только конечномерные пространства. Отметим, однако, что не все векторные пространства конечномерны:

Пример. В пространстве всех бесконечных последовательностей $(a_1, \dots, a_n, \dots)$ не существует конечного базиса. Множество последовательностей из нулей и единиц (исключая нулевую последовательность) линейно независимо.

Предложение. *Пусть V – векторное пространство и пусть $M \subset V$ – произвольное подмножество.*

(1) *Если $\mathbf{e}_1, \dots, \mathbf{e}_m \in M$ – базис M, то любой вектор $\mathbf{v} \in M$ однозначно выражается через $\mathbf{e}_1, \dots, \mathbf{e}_m$.*

(2) *Наоборот, если $\mathbf{e}_1, \dots, \mathbf{e}_m \in M$ – векторы такие, что любой вектор $\mathbf{v} \in M$ однозначно выражается как линейная комбинация $\mathbf{e}_1, \dots, \mathbf{e}_m$ и $\mathbf{0} \in M$, то $\mathbf{e}_1, \dots, \mathbf{e}_m$ – базис M.*

Коэффициенты этого разложения называются координатами *вектора.*

Обычно координаты вектора записываются в виде строки. Поэтому базис следует рассматривать как упорядоченное множество.

Доказательство. (1) Так как $\mathbf{e}_1, \dots, \mathbf{e}_m$ – базис, то добавление к нему любого вектора $v \in M$ дает линейно зависимую систему. Таким образом,

$$\lambda \mathbf{v} + \sum \alpha_i \mathbf{e}_i = \mathbf{0}.$$

Если при этом $\lambda = 0$, то система $\mathbf{e}_1, \dots, \mathbf{e}_m$ окажется линейно зависимой. Значит $\lambda \neq 0$ и тогда

$$\mathbf{v} = -\sum \frac{\alpha_i}{\lambda} \mathbf{e}_i.$$

Это доказывает существование выражения.

Для доказательства единственности предположим, что

$$\mathbf{v} = \sum \lambda_i \mathbf{e}_i = \sum \lambda'_i \mathbf{e}_i.$$

Тогда

$$\sum (\lambda_i - \lambda'_i)\mathbf{e}_i = \mathbf{0}.$$

Так как векторы $\mathbf{e}_1, \dots, \mathbf{e}_m$ линейно независимы, то $\lambda_i = \lambda'_i$ для всех i.

(2) Если векторы $\mathbf{e}_1, \dots, \mathbf{e}_m$ линейно зависимы, то имеют место равенства

$$\lambda_1 \mathbf{e}_1 + \cdots \lambda_m \mathbf{e}_m = \mathbf{0} = 0\,\mathbf{e}_1 + \cdots 0\,\mathbf{e}_m,$$

где не все λ_i равны нулю. Это противоречит единственности разложения. Таким образом, $\mathbf{e}_1, \dots, \mathbf{e}_m$ линейно независимы. Если $\mathbf{v}$ – любой вектор, то $\mathbf{v} = \sum \alpha_i \mathbf{e}_i$ для некоторых α_i, значит $-\mathbf{v} + \sum \alpha_i \mathbf{e}_i = \mathbf{0}$ и векторы $\mathbf{v}, \mathbf{e}_1, \dots, \mathbf{e}_m$ линейно зависимы. □

Лемма о линейной зависимости

Теорема. *Пусть векторы $\mathbf{v}_1, \dots, \mathbf{v}_n$ линейно независимы и линейно выражаются через векторы $\mathbf{w}_1, \dots, \mathbf{w}_m$. Тогда $n \le m$.*

Доказательство. Предположим, что $n > m$. Запишем

$$\mathbf{v}_j = \alpha_{1,j}\mathbf{w}_1 + \cdots + \alpha_{m,j}\mathbf{w}_m.$$

Тогда для некоторых $\lambda_1, \dots, \lambda_n$ имеем

$$(*) \qquad \sum_{j=1}^{n} \lambda_j \mathbf{v}_j = \sum_{j=1}^{n} \lambda_j \left(\sum_{i=1}^{m} \alpha_{i,j} \mathbf{w}_i \right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \alpha_{i,j} \lambda_j \right) \mathbf{w}_i.$$

Рассмотрим систему однородных линейных уравнений

$$\sum_{j=1}^{n} \alpha_{i,j} \lambda_j = 0, \quad i = 1, \dots m.$$

относительно $\lambda_1, \dots, \lambda_n$. Так как число неизвестных меньше числа уравнений, то система имеет ненулевое решение. Для таких $\lambda_1, \dots, \lambda_n$ правая часть соотношения $(*)$ обращается в нуль. Значит, $\sum_{j=1}^{n} \lambda_j \mathbf{v}_j = \mathbf{0}$. □

Следствия

Пусть V – конечномерное векторное пространство, а $M \subset V$ – его подмножество.

Следствие. *Любой базис M содержит одинаковое количество элементов. Более точно, если $\mathbf{w}_1, \dots, \mathbf{w}_m$ – базис M, то любая система из n векторов $\mathbf{v}_1, \dots, \mathbf{v}_n \in M$ при $n > m$ линейно зависима. Если же при этом $\mathbf{v}_1, \dots, \mathbf{v}_n$ также образуют базис M, то $m = n$.*

Доказательство. Векторы $\mathbf{v}_1, \ldots, \mathbf{v}_n$ линейно выражаются через $\mathbf{w}_1, \ldots, \mathbf{w}_m$. Если векторы $\mathbf{v}_1, \ldots, \mathbf{v}_n$ линейно независимы, то по лемме о линейной зависимости $n \leq m$. Если же $\mathbf{v}_1, \ldots, \mathbf{v}_n$ образуют базис, то аналогично получаем $m \leq n$. □

Определение. *Рангом* подмножества M векторного пространства (обозначается $\mathrm{rk}(M)$) называется число элементов базиса (если конечный базис существует). Рангом столбцов матрицы называется ранг системы ее столбцов (как системы векторов $\mathbb{R}^n$). Рангом строк матрицы называется ранг системы ее строк*. *Размерностью*† пространства V (обозначается $\dim(V)$) называется ранг $M = V$.

Замечание. Если $M \subset V$, то $\mathrm{rk}(M)$ равен максимальному числу линейно независимых векторов $\mathbf{v}_1, \ldots, \mathbf{v}_n \in M$.

Следствие. *Пусть* $\dim(V) = n$. *Тогда система из* m *векторов в* V *линейно зависима при* $m > n$.

Следствие. *Если все векторы системы* $M \subset V$ *линейно выражаются через векторы системы* $M' \subset V$, *то* $\mathrm{rk}(M) \leq \mathrm{rk}(M')$.

Доказательство. Пусть $\mathbf{v}_1, \ldots, \mathbf{v}_n$ – базис M, а $\mathbf{v}'_1, \ldots, \mathbf{v}'_m$ – базис M'. По условию векторы $\mathbf{v}_1, \ldots, \mathbf{v}_n$ линейно выражаются через векторы M', а любой вектор M' линейно выражается через $\mathbf{v}'_1, \ldots, \mathbf{v}'_m$. Следовательно, $\mathbf{v}_1, \ldots, \mathbf{v}_n$ линейно выражаются через $\mathbf{v}'_1, \ldots, \mathbf{v}'_m$. По лемме о линейной зависимости $n \leq m$. □

Следствие. *Пусть* $\mathbf{e}_1, \ldots, \mathbf{e}_k \in M$ *– линейно независимые векторы. Тогда* $\mathbf{e}_1, \ldots, \mathbf{e}_k \in M$ *можно дополнить до базиса* M.

Доказательство. Если для любого $\mathbf{x} \in M$ векторы $\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{x}$ линейно зависимы, то $\mathbf{e}_1, \ldots, \mathbf{e}_k$ – базис M. В противном случае существует $\mathbf{x} \in M$ такой, что векторы $\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{x}$ линейно независимы. Полагаем $\mathbf{e}_{k+1} = \mathbf{x}$ и продолжаем процесс. Процесс оборвется поскольку число линейно независимых векторов не превосходит $\dim(V)$. □

Следствие. *Пусть* $W \subset V$ *– подпространство векторного пространства. Тогда*

$$\dim(W) \leq \dim(V)$$

Более того, если $\dim(W) = \dim(V)$, *то* $W = V$.

Доказательство. Система $\mathbf{w}_1, \ldots, \mathbf{w}_n \in W \subset V$ может быть дополнена до базиса V. Так как $\dim(V) = n$, то $\mathbf{w}_1, \ldots, \mathbf{w}_n$ – уже базис V. Следовательно, $V \subset W$. □

*Позже мы докажем, что для любой матрицы ранги строк и столбцов совпадают.

†Термин *ранг* употребляется в отношении подмножества (чаще всего конечного) $M \subset V$. Термин *размерность* употребляется в отношении векторного пространства.

Лекция 8

Теорема о ранге. Алгоритм нахождения базиса. Ранг суммы и произведения матриц. Критерий совместности системы линейных уравнений. Решения однородной системы линейных уравнений. Фундаментальная система решений. Задание подпространства системой линейных уравнений.

Теорема о ранге матрицы

Теорема. *Ранг системы строк матрицы равен рангу системы ее столбцов и равен наивысшему порядку отличного от нуля минора.*

Лемма. *Ранг (системы строк) матрицы не меняется при элементарных преобразованиях строк.*

Доказательство. Пусть матрица A' получена из матрицы A одним элементарным преобразованием строк. Тогда строка матрицы A' линейно выражаются через строки матрицы A. Следовательно, $\mathrm{rk}(A') \le \mathrm{rk}(A)$. Преобразование $A \mapsto A'$ обратимо, поэтому $\mathrm{rk}(A) \le \mathrm{rk}(A')$. □

Лемма. *При элементарных преобразованиях строк матрицы линейные зависимости (независимости) системы ее столбцов не меняются.*

Доказательство. Пусть $A = (a_{i,j})$ – $n \times m$-матрица и пусть $A_1, \dots, A_m$ – ее столбцы. Пусть матрица $A' = (a'_{i,j})$ получена из матрицы A одним элементарным преобразованием строк и пусть $A'_1, \dots, A'_m$ – столбцы A'. Предположим, что имеется линейная зависимость $\sum_{j=1}^{m} \lambda_j A_j = 0$. Это эквивалентно системе n равенств

$$(*) \qquad \sum_{j=1}^{m} \lambda_j a_{i,j} = 0, \quad i = 1, \dots, n.$$

Пусть, например, $A \mapsto A'$ – преобразование типа (I), т.е.

$$a'_{i,j} = \begin{cases} a_{i,j} & \text{при } i \ne i_0 \\ a_{i_0,j} + \mu a_{i_1,j} & \text{при } i = i_0 \end{cases}$$

для некоторых $i_0 \neq i_1$ и μ. Тогда равенство $(*)$ сохраняется для $a'_{i,j}$ при $i \neq i_0$, а для $i = i_0$ оно дает нам

$$0 = \sum_{j=1}^{m} \lambda_j a_{i_0,j} = \sum_{j=1}^{m} \lambda_j (a'_{i_0,j} - \mu a_{i_1,j}) = \sum_{j=1}^{m} \lambda_j a'_{i_0,j} - \mu \sum_{j=1}^{m} a_{i_1,j}.$$

Поскольку второй член равен нулю, то $\sum_{j=1}^{m} \lambda_j a'_{i_0,j} = 0$. Следовательно, $\sum_{j=1}^{m} \lambda_j A'_j = 0$.

Наоборот, предположим, что столбцы $A_{i_1}, \ldots, A_{i_r}$ линейно независимы. Если соответствующие столбцы $A_{i_1}, \ldots, A_{i_r}$ линейно зависимы, то из обратимости преобразования $A \mapsto A'$ мы получаем линейную зависимость столбцов $A_{i_1}, \ldots, A_{i_r}$. □

Следствие. *Ранг системы столбцов матрицы не меняется при элементарных преобразованиях строк.*

Лемма. *Пусть в матрице A минор, стоящий на пересечении строк $i_1, \ldots, i_r$ и столбцов $j_1, \ldots, j_r$, отличен от нуля. Тогда соответствующие строки (с номерами $i_1, \ldots, i_r$) линейно независимы. То же верно для столбцов (с номерами $j_1, \ldots, j_r$).*

Доказательство. Пусть M – соответствующая матрица. Согласно критерию невырожденности матрицы строки и столбцы M линейно независимы. Следовательно, линейно независимы также и удлинённые строки и столбцы. □

Лемма. *Ранг системы строк ступенчатой матрицы $A = (a_{i,j})$ равен рангу системы ее столбцов и равен числу ее ненулевых строк.*

Доказательство. Пусть $a_{1,j_1}, \ldots, a_{r,j_r}$ – первые ненулевые элементы строк, где $j_1 < j_2 < \cdots < j_r$ и r – числу ненулевых строк. Тогда минор, стоящий на пересечении строк $1, \ldots, r$ и столбцов $j_1, \ldots, j_r$, отличен от нуля. По предыдущей лемме ненулевые строки матрицы A линейно независимы и поэтому ранг системы строк A равен r. Далее, главные столбцы $A_{j_1}, \ldots, A_{j_r}$ также линейно независимы. Докажем, что они образуют базис системы столбцов. Пусть A_k – произвольный столбец матрицы A. Достаточно показать, что имеет место разложение $A_k = \lambda_1 A_{j_1} + \cdots + \lambda_r A_{j_r}$. Рассмотрим его как систему линейных уравнений относительно $\lambda_1, \ldots, \lambda_r$. Матрица этой системы треугольна и имеет определитель (равный минору) отличный от нуля. Следовательно, система совместна. □

Следствие. *Ранг системы строк матрицы равен рангу системы ее столбцов.*

Доказательство теоремы о ранге. Осталось доказать, что ранг матрицы не может превосходить наивысшего порядка отличного от нуля минора. Пусть в матрице строки с номерами $i_1, \ldots, i_r$ образуют базис системы строк. Докажем, что некоторый минор порядка r отличен от нуля. Пусть A' – матрица, полученная из A вычёркиванием всех строк кроме $i_1, \ldots, i_r$. Тогда $\mathrm{rk}(A') = r$ (поскольку строки A' линейно независимы). Пусть теперь столбцы матрицы A' с номерами $j_1, \ldots, j_r$ образуют базис системы ее столбцов и пусть A'' – матрица, полученная из A' вычёркиванием всех столбцов, кроме $j_1, \ldots, j_r$. Снова $\mathrm{rk}(A'') = r$ (поскольку столбцы A'' линейно независимы). Согласно критерию невырожденности матрицы $|A''| \neq 0$. □

Следствие. $\mathrm{rk}(A) = \mathrm{rk}(A^{\mathrm{T}})$.

Следствие. *Ранг матрицы равен числу ненулевых строк в ее ступенчатом виде.*

Алгоритм нахождения базиса

Пусть даны векторы $A_1, \ldots, A_m \in \mathbb{R}^n$. Найдем базис этой системы векторов $\{A_1, \ldots, A_m\}$. Для этого рассмотрим матрицу A, составленную из векторов $A_1, \ldots, A_m$, рассмотренных как столбцы:

$$A := \begin{pmatrix} A_1 & \cdots & A_m. \end{pmatrix}$$

Элементарными преобразованиями строк приведем матрицу A к ступенчатому виду A'. Пусть $A'_1, \ldots, A'_m$ – столбцы матрицы A'. Главные столбцы $A'_{i_1}, \ldots, A'_{i_r}$ образуют базис для $\{A'_1, \ldots, A'_m\}$. Тогда столбцы $A_{i_1}, \ldots, A_{i_r}$ матрицы A (с теми же номерами) образуют базис $\{A_1, \ldots, A_m\}$.

Ранг суммы и произведения матриц

Теорема. $\mathrm{rk}(A + B) \leq \mathrm{rk}(A) + \mathrm{rk}(B)$.

Доказательство. Пусть $A_1, \ldots, A_m$ – столбцы матрицы A, а $B_1, \ldots, B_m$ – столбцы матрицы B. Пусть $\{A_{i_1}, \ldots, A_{i_r}\}$ – базис для $\{A_1, \ldots, A_m\}$, а $\{B_{i_1}, \ldots, B_{i_s}\}$ – базис для $\{B_1, \ldots, B_m\}$. Таким образом, $\mathrm{rk}(A) = r$ и $\mathrm{rk}(B) = s$. Рассмотрим матрицу

$$C := \begin{pmatrix} A \mid B \end{pmatrix} = \begin{pmatrix} A_1 & \cdots & A_m & B_1 & \cdots & B_m \end{pmatrix}.$$

Ее столбцы выражаются как линейные комбинации векторов базисов $A_{i_1}, \ldots, A_{i_r}$ и $B_{i_1}, \ldots, B_{i_s}$. По лемме о линейной зависимости $\mathrm{rk}(C) \leq r + S$. С другой стороны, столбцы матрицы $A + B$ выражаются как линейные комбинации столбцов матрицы C. Следовательно, $\mathrm{rk}(A + B) \leq \mathrm{rk}(C) \leq r + S$. □

Теорема. $\mathrm{rk}(A \cdot B) \leq \mathrm{rk}(A),\ \mathrm{rk}(B)$.

Доказательство. Пусть $C := A \cdot B$. Мы считаем, что $A = (a_{i,j})$, $B = (b_{i,j})$, $C = (c_{i,j})$ и размеры матриц A, B, C – $n \times m$, $m \times q$ и $n \times q$, соответственно. Пусть $C_1, \ldots, C_q$ – столбцы C и пусть $A_1, \ldots A_m$ – столбцы A. На i-ом месте столбца C_j стоит элемент $c_{i,j} = \sum_{k=1}^{m} a_{i,k} b_{k,j}$. Поэтому $C_j = \sum_{k=1}^{m} b_{k,j} A_k$. Таким образом, столбцы матрицы C линейно выражаются через столбцы матрицы A. По следствию из леммы о линейной зависимости имеем $\mathrm{rk}(C) \leq \mathrm{rk}(A)$. Второе неравенство следует из того, что

$$\mathrm{rk}(C) = \mathrm{rk}(C^{\mathrm{T}}) = \mathrm{rk}(B^{\mathrm{T}} \cdot A^{\mathrm{T}}) \leq \mathrm{rk}(B^{\mathrm{T}}) = \mathrm{rk}(B).$$ □

Критерий совместности системы линейных уравнений

Теорема (Кронекера–Капелли). *Система линейных уравнений $A \cdot X = B$ совместна тогда и только тогда, когда* $\mathrm{rk}(A) = \mathrm{rk}(A \mid B)$.

Доказательство. Необходимость. Пусть $x_1 = \lambda_1, \ldots, x_n = \lambda_n$ – решение. Тогда $\lambda_1 A_1 + \cdots + \lambda_n A_n = B$, т.е. столбцы матрицы (A, B) выражаются через столбцы матрицы A. По лемме о линейной зависимости $\mathrm{rk}(A) \geq \mathrm{rk}(A \mid B)$. Очевидно также, что $\mathrm{rk}(A) \leq \mathrm{rk}(A \mid B)$.

Достаточность. Пусть $\mathrm{rk}(A) = \mathrm{rk}(A \mid B) = r$. Пусть $A_{i_1}, \ldots, A_{i_r}$ базис столбцов матрицы A. Тогда это и базис столбцов матрицы $(A \mid B)$. Следовательно, B выражается через этот базис: $\lambda_{i_1} A_{i_1} + \cdots + \lambda_{i_r} A_{i_r} = B$. Полагая $\lambda_j = 0$ при $j \notin \{i_1, \ldots, i_r\}$, мы можем записать $\lambda_1 A_1 + \cdots + \lambda_n A_n = B$, т.е. $(\lambda_1, \ldots, \lambda_n)$ – решение. □

Пространство решений однородной системы линейных уравнений

Все решения системы $A \cdot X = B$ от n неизвестных образуют подпространство в $\mathbb{R}^n$ тогда и только тогда, когда $B = 0$, т.е. система однородна.

Определение. *Фундаментальной системой решений* однородной системы называется любой базис пространства решений однородной системы.

Теорема (размерность пространства решений однородной системы уравнений)**.** *Пусть $A \cdot X = 0$ – однородная система линейных уравнений, где $X = (x_1, \ldots, x_n)^{\mathrm{T}}$ и пусть V – пространство решений. Тогда*

$$\dim(V) = n - \mathrm{rk}(A).$$

Доказательство. Пусть $r := \mathrm{rk}(A)$ и пусть $x_{i_1}, \ldots, x_{i_{n-r}}$ – свободные неизвестные. Все неизвестные выражаются через свободные:

$$x_j = \sum_{k=0}^{n-r} b_{j,k} x_{i_k}, \qquad j = 1, \ldots n.$$

Построим базис $\mathbf{e}_1, \ldots, \mathbf{e}_{n-r} \in V$. Например,

- $\mathbf{e}_s = (\alpha_1^s, \cdots, \alpha_n^s)$, где
$$\alpha_{i_k}^s = \delta_{k,s} = \begin{cases} 1 & \text{если } k = s, \\ 0 & \text{если } k \neq s. \end{cases}$$

Остальные α_j, $j \notin \{i_1, \ldots, i_{n-r}\}$ вычисляются по формулам выше. Тогда $\mathbf{e}_s \in V$. Векторы $\mathbf{e}_1, \ldots, \mathbf{e}_{n-r}$ линейно независимы. Действительно, укороченные векторы $\mathbf{e}'_1, \ldots, \mathbf{e}'_{n-r}$, полученные из $\mathbf{e}_1, \ldots, \mathbf{e}_{n-r}$ вычеркиванием всех координат, кроме координат с номерами $i_1, \ldots, i_{n-r}$, имеют вид $\mathbf{e}'_s = (0, \ldots, \underset{\uparrow}{1}, \ldots, 0)$. Эти векторы линейно независимы, а поэтому линейно независимы и векторы $\mathbf{e}_1, \ldots, \mathbf{e}_{n-r}$. Пусть $\mathbf{v} = (\alpha_1, \ldots, \alpha_n)$ – любое решение. Положим
$\quad\quad s$

$$\mathbf{w} = \mathbf{v} - (\alpha_{i_1} \mathbf{e}_1 + \cdots + \alpha_{i_{n-r}} \mathbf{e}_{n-r}).$$

Так как V – пространство, то $\mathbf{w}$ – решение. Значение свободных неизвестных для w равны нулю. Следовательно, $\mathbf{w} = \mathbf{0}$ и $\mathbf{v} = \alpha_{i_1} \mathbf{e}_1 + \cdots + \alpha_{i_{n-r}} \mathbf{e}_{n-r}$. □

Теорема. *Пусть $V \subset \mathbb{R}^n$ – m-мерное подпространство. Существует однородная система линейных уравнений $A \cdot X = 0$ такая что $X \in V$ тогда и только тогда, когда $A \cdot X = 0$.*

Доказательство. Если $V = \{0\}$, то утверждение очевидно. Пусть столбцы $B_1, \ldots, B_m$ составляют базис V. Составим $n \times m$-матрицу B, в которой $B_1, \ldots, B_m$ являются столбцами. Таким образом, $\mathrm{rk}(B) = m$. Рассмотрим систему линейных уравнений $B^{\mathrm{T}} \cdot Y = 0$, где Y – столбец n неизвестных. Пусть $Y = F_1, \ldots, F_{n-m}$ – фундаментальная система решений и пусть F – $n \times (n-m)$-матрица, в которой $F_1, \ldots, F_{n-m}$ являются столбцами. Тогда $B^{\mathrm{T}} \cdot F = 0$ и, следовательно, $F^{\mathrm{T}} \cdot B = 0$. Полагаем $A := F^{\mathrm{T}}$. Имеем $A \cdot B = 0$ и строки A линейно независимы. Пусть $W \subset \mathbb{R}^n$ – пространство решений однородной системы $A \cdot X = 0$. Далее, $A \cdot B_i = 0$. Поэтому $A \cdot X = 0$ для любого X, являющегося линейной комбинацией B_i, т.е. для любого X, принадлежащего V. Следовательно, $V \subset W$. Так как $\dim(W) = \dim(V) = m$, то $V = W$. $\square$

Лекция 9

Линейные отображения векторных пространств. Ядро и образ. Изоморфизмы. Изоморфизм векторных пространств одной размерности. Матрица линейного отображения. Гомоморфизмы групп. Примеры. Ядро и образ гомоморфизма групп. Изоморфизмы. Гомоморфизмы колец. Примеры. Ядро и образ гомоморфизма колец. Изоморфизмы.

Линейные отображения векторных пространств

Определение. Пусть V и W – векторные пространства. Отображение $f : V \to W$ называется *линейным*, если

$$f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2) \quad \text{и} \quad f(\lambda \mathbf{v}) = \lambda f(\mathbf{v}) \qquad \forall \mathbf{v}_1, \mathbf{v}_2, \mathbf{v} \in V. \qquad \forall \lambda \in \mathbb{R}.$$

Линейное отображение $f : V \to V$ векторного пространства в себя называется *линейным оператором*. *Изоморфизмом векторных пространств* называется линейное отображение $f : V \to W$, у которого существует обратное. Говорят, что векторные пространства V и W *изоморфны*, если существует изоморфизм $V \to W$.

Заметим, что обратное отображение к изоморфизму также является линейным отображением, и, следовательно, изоморфизмом. Таким образом, изоморфизм является отношением эквивалентности на множестве всех векторных пространств. Если векторные пространства V и W изоморфны, то мы будем это обозначать следующим образом $V \simeq W$.

Определение. Пусть $f : V \to W$ – линейное отображение. Подмножество

$$\mathrm{Ker}(f) := \{\mathbf{v} \in V \mid f(\mathbf{v}) = \mathbf{0}\}$$

называется *ядром* отображения f.

Лемма. *Пусть $f : V \to W$ – линейное отображение векторных пространств. Тогда верны следующие утверждения.*

(1) *Ядро* $\mathrm{Ker}(f)$ *является подпространством в* V.

(2) *Образ* $f(V)$ *является подпространством в* W.

(3) *Отображение f инъективно тогда и только тогда, когда* $\mathrm{Ker}(f) = \{\mathbf{0}\}$.

(4) *Отображение f является изоморфизмом тогда и только тогда, когда* $\mathrm{Ker}(f) = \{\mathbf{0}\}$ *и* $f(V) = W$.

Доказательство. (1) Пусть $\mathbf{v}_1, \mathbf{v}_2 \in \mathrm{Ker}(f)$ и пусть λ_1, $\lambda_2 \in \mathbb{R}$. Тогда $f(\mathbf{v}_1) = f(\mathbf{v}_2) = \mathbf{0}$. Значит $f(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2) = \lambda_1 f(\mathbf{v}_1) + \lambda_2 f(\mathbf{v}_2) = \mathbf{0}$. Следовательно, $\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 \in \mathrm{Ker}(f)$. Таким образом, $\mathrm{Ker}(f)$ – подпространство.

(2) Пусть $\mathbf{w}_1, \mathbf{w}_2 \in f(V)$ и пусть λ_1, $\lambda_2 \in \mathbb{R}$. Существуют векторы $\mathbf{v}_1, \mathbf{v}_2 \in V$ такие, что $f(\mathbf{v}_i) = \mathbf{w}_i$. Тогда $f(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2) = \lambda_1\mathbf{w}_1 + \lambda_2\mathbf{w}_2$. Следовательно, $\lambda_1\mathbf{w}_1 + \lambda_2\mathbf{w}_2 \in f(V)$ и $f(V)$ – подпространство.

(3) Если отображение f инъективно, то, очевидно, $\mathrm{Ker}(f) = \{\mathbf{0}\}$. Пусть $\mathrm{Ker}(f) = \{\mathbf{0}\}$. Предположим, что $f(\mathbf{v}_1) = f(\mathbf{v}_2)$. Тогда $f(\mathbf{v}_1 - \mathbf{v}_2) = f(\mathbf{v}_1) - f(\mathbf{v}_2) = \mathbf{0}$. Следовательно, $\mathbf{v}_1 - \mathbf{v}_2 \in \mathrm{Ker}(f) = \{\mathbf{0}\}$, т.е. $\mathbf{v}_1 = \mathbf{v}_2$.

(4) следует из (3). □

Теорема. *Конечномерные векторные пространства изоморфны тогда и только тогда, когда их размерности совпадают.*

Доказательство. Пусть $V \simeq V'$ и пусть $f : V \to V'$ – изоморфизм. Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ – базис V. Положим $\mathbf{e}'_i := f(\mathbf{e}_i)$. Докажем, что $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ – базис V'. Предположим, что $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ линейно зависимы. Тогда существуют $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ не все равные нулю, такие, что $\sum \lambda_i\mathbf{e}'_i = \mathbf{0}$. Так как f – линейное отображение, то $f(\sum \lambda_i\mathbf{e}_i) = \sum \lambda_i f(\mathbf{e}_i) = \sum \lambda_i\mathbf{e}'_i = \mathbf{0}$. Значит $\sum \lambda_i\mathbf{e}_i \in \mathrm{Ker}(f) = \{\mathbf{0}\}$, $\sum \lambda_i\mathbf{e}_i = \mathbf{0}$ и поэтому $\lambda_1 = \cdots = \lambda_n = 0$. Противоречие. Пусть $\mathbf{v}' \in V'$ – произвольный вектор. Существует вектор $\mathbf{v} \in V$ такой, что $f(\mathbf{v}) = \mathbf{v}'$. Так как $\mathbf{e}_1, \ldots, \mathbf{e}_n$ – базис, то $\mathbf{v} = \sum \mu_i\mathbf{e}_i$ для некоторых $\mu_1, \ldots, \mu_n \in \mathbb{R}$. Тогда $\mathbf{v}' = f(\mathbf{v}) = f(\sum \mu_i\mathbf{e}_i) = \sum \mu_i f(\mathbf{e}_i) = \sum \mu_i\mathbf{e}'_i$. Таким образом, $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ – базис V'. Значит, $\dim(V') = n = \dim(V)$.

Наоборот, пусть $\dim(V') = \dim(V)$. Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ – базисы V и V', соответственно. Для вектора $\mathbf{v} = \sum \mu_i\mathbf{e}_i \in V$ положим $f(\mathbf{v}) = \sum \mu_i\mathbf{e}'_i$. Тогда f – корректно определенное отображение, являющееся изоморфизмом. □

Определение. Пусть $f : V \to V'$ – линейное отображение векторных пространств. Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\mathbf{e}'_1, \ldots, \mathbf{e}'_m$ – базисы V и V', соответственно. Разложим элементы $f(\mathbf{e}_i)$ по базису $\mathbf{e}'_1, \ldots, \mathbf{e}'_m$:

$$(*) \qquad f(\mathbf{e}_i) = \sum_{j=1}^{m} a_{j,i}\mathbf{e}'_j.$$

Числа $a_{j,i}$ поместим в матрицу размера $m \times n$:

$$A := \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \ldots\ldots\ldots\ldots & & \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}$$

Эта матрица называется *матрицей линейного отображения.*

Таким образом, элементы матрицы A – это координаты векторов $f(\mathbf{e}_i)$ в базисе $\mathbf{e}'_1, \ldots, \mathbf{e}'_m$, записанные в столбцы. Конечно, матрица линейного отображения зависит от выбора базисов $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\mathbf{e}'_1, \ldots, \mathbf{e}'_m$.

Матрица A полностью определяет соответствующее линейное отображение:

Предложение. *Если $v_1, \dots, v_n$ – координаты вектора $\mathbf{v} \in V$ в базисе $\mathbf{e}_1, \dots, \mathbf{e}_n$, то $\mathbf{e}'_1, \dots, \mathbf{e}'_m$ координатами вектора $f(\mathbf{v}) \in V'$ в базисе $\mathbf{e}'_1, \dots, \mathbf{e}'_m$ вычисляются по формулам*

$$(\dagger) \qquad w_j = \sum_{i=1}^{n} a_{j,i} v_i.$$

Доказательство. Имеем $\mathbf{v} = \sum_{i=1}^{n} v_i \mathbf{e}_i$. Учитывая $(*)$ получим

$$f(\mathbf{v}) = f\left(\sum_{i=1}^{n} v_i \mathbf{e}_i\right) = \sum_{i=1}^{n} v_i f(\mathbf{e}_i) = \sum_{i=1}^{n} v_i \left(\sum_{j=1}^{m} a_{j,i} \mathbf{e}'_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} v_i a_{j,i} \mathbf{e}'_j =$$
$$= \sum_{j=1}^{m}\sum_{i=1}^{n} v_i a_{j,i} \mathbf{e}'_j = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_{j,i} v_i\right) \mathbf{e}'_j.$$

Это и доказывает $(\dagger)$. □

Следствие. *В обозначениях последнего предложения запишем координаты векторов $\mathbf{v}$ и $f(\mathbf{v})$ в виде столбцов:*

$$X := \begin{pmatrix} v_1 \\ \cdots \\ v_n \end{pmatrix} \qquad Y := \begin{pmatrix} w_1 \\ \cdots \\ w_m \end{pmatrix}$$

Тогда

$$Y = A \cdot X.$$

Гомоморфизмы групп

Определение. *Гомоморфизмом* групп называется отображение $f : G \to H$ группы G в группу H такое, что

$$f(a \circ b) = f(a) * f(b) \qquad \forall a, b \in G.$$

Предложение. *Пусть $f : G \to H$ – гомоморфизм групп. Тогда*

(1) $f(a^n) = f(a)^n$;

(2) $f(1_G) = 1_H$;

(3) $f(a^{-1}) = f(a)^{-1}$;

(4) *если у f существует обратное отображение, то оно тоже является гомоморфизмом.*

Примеры. (1) Определитель $\det : \mathrm{GL}_n(\Bbbk) \to \Bbbk^*$ является гомоморфизмом групп.

(2) Знак подстановки $\mathrm{sgn} : \mathrm{S}_n \to \{\pm 1\}$ является гомоморфизмом групп.

(3) Если A – абелева группа, то отображение $f(a) = a^n$ является гомоморфизмом. В абелевой аддитивной группе для любого $n \in \mathbb{Z}$ отображение $a \mapsto na$ является гомоморфизмом группы в себя.

(4) Экспонента $\exp : \mathbb{R} \to \mathbb{R}^*$, $a \mapsto e^a$ является гомоморфизмом групп.

(5) Взятие модуля $|\cdot| : \mathbb{R}^* \to \mathbb{R}_{>0}$, $z \mapsto |z|$ является гомоморфизмом групп.

(6) $\mathbb{Z} \to \mathbb{R}^*, \quad n \longmapsto a^n$.

Определение. Гомоморфизм групп $f : G \to H$ называется *изоморфизмом*, если существует обратное отображение. Изоморфизм $f : G \to G$ группы с собой называется *автоморфизмом.*

Примеры. (1) тождественное отображение $G \to G$ является автоморфизмом;

(2) $\mathbb{R}^+ \to \mathbb{R}_{>0}, \quad a \longmapsto e^a$ является изоморфизмом;

(3) если A – абелева группа, то отображение $f(a) = a^{-1}$ является автоморфизмом;

(4) $\mathrm{GL}_n(\Bbbk) \to \mathrm{GL}_n(\Bbbk), \qquad A \longmapsto (A^{-1})^{\mathrm{T}}$ является автоморфизмом;

(5) для любой группы G и элемента $a \in G$ отображение $f(x) = axa^{-1}$ является автоморфизмом. Он называется *внутренним* автоморфизмом.

Определение. Подмножество

$$\mathrm{Ker}(\varphi) := \{a \in G \mid \varphi(a) = 1_H\}$$

называется *ядром* гомоморфизма групп $\varphi : G \to H$.

Лемма. *Пусть $\varphi : G \to H$ – гомоморфизм групп. Тогда верны следующие утверждения.*

(1) *Ядро* $\mathrm{Ker}(\varphi)$ *является подгруппой в* G.

(2) *Образ* $\varphi(G)$ *является подгруппой в* H.

(3) *Гомоморфизм* φ *инъективен тогда и только тогда, когда* $\mathrm{Ker}(\varphi) = \{1\}$.

(4) *Гомоморфизм* φ *является изоморфизмом тогда и только тогда, когда* $\mathrm{Ker}(\varphi) = \{1\}$ *и* $\varphi(G) = H$.

Доказательство. Проверим, например, первое. Пусть $a_1, a_2 \in \mathrm{Ker}(\varphi)$. Тогда $\varphi(a_1) = \varphi(a_2) = 1$ и $\varphi(a_1 a_2^{-1}) = \varphi(a_1)\varphi(a_2)^{-1} = 1$. Отсюда $a_1 a_2^{-1} \in \mathrm{Ker}(\varphi)$. Поэтому $\mathrm{Ker}(\varphi)$ – подгруппа. □

Пример. Напомним, что для любой аддитивной абелевой группы A определено умножение его элемента $a \in A$ на целое число $n \in \mathbb{Z}$:

$$n \cdot a = \begin{cases} \underbrace{a + \cdots + a}_{n} & \text{если } n > 0, \\ 0 & \text{если } n = 0, \\ \underbrace{a + \cdots + a}_{-n} & \text{если } n < 0. \end{cases}$$

Это умножение удовлетворяет свойствам

$$(n_1 + n_2) \cdot a = n_1 \cdot a + n_2 \cdot a, \qquad n \cdot (a_1 + a_2) = n \cdot a_1 + n \cdot a_2.$$

Следовательно, для любого элемента $a \in A$ отображение

$$\phi_a : \mathbb{Z} \longrightarrow A$$

является гомоморфизмом групп.

Гомоморфизмы колец

Определение. *Гомоморфизмом колец* называется отображение $\varphi : R \to R_1$ такое, что $\varphi(a + b) = \varphi(a) + \varphi(b)$ и $\varphi(ab) = \varphi(a)\varphi(b)$ для любых элементов a, $b \in R$. Таким образом, гомоморфизм колец является гомоморфизмом их аддитивных групп. Как обычно в алгебре, биективный гомоморфизм называется *изоморфизмом*. Изоморфизм кольца на себя называется *автоморфизмом*.

Если R и R_1 – кольца с единицами 1 и $1'$, то обычно считается, что гомоморфизм колец $\varphi : R \to R_1$ единицу переводит в единицу, т.е. $\varphi(1) = 1'$. Заметим, что это свойство автоматически не выполняется для произвольного гомоморфизма. Например, пусть $R' := \mathbb{R}^n$, где сложение – это обычное сложение векторов, а умножение – покомпонентное: $(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = (a_1 b_1, \ldots, a_n b_n)$. Рассмотрим инъективный гомоморфизм

$$\varphi : \mathbb{R} \longrightarrow R' = \mathbb{R}^n, \qquad a \longmapsto (a, 0, \ldots, 0).$$

Оба кольца $\mathbb{R}$ и R' имеют единицы: 1 и $(1, \ldots, 1)$, соответственно. Однако, $\varphi(1) \neq (1, \ldots, 1)$.

Предложение. *Пусть $f : R \to S$ – гомоморфизм колец. Тогда*

(1) $f(n \cdot a) = n \cdot f(a)$;

(2) $f(a^n) = n \cdot f(a)^n$;

(3) $f(0_R) = 0_S$;

(4) $f(-a) = -f(a)$;

(5) *если у f существует обратное отображение, то оно тоже является гомоморфизмом.*

Примеры. (1) Если R – кольцо с единицей, то отображение

$$\phi : \mathbb{Z} \longrightarrow R, \qquad n \longmapsto n \cdot 1$$

является гомоморфизмом. Действительно, было показано, что ϕ – гомоморфизм аддитивных групп. С другой стороны, $\phi(nm) = (nm)\,1 = (n\,1)(m\,1) = \phi(n)\phi(m)$.

(2) Пусть A – кольцо числовых функций на некотором множестве M. Зафиксируем элемент $a \in M$. Отображение

$$\varphi : A \longrightarrow \mathbb{R}, \qquad f \longmapsto f(a).$$

Является гомоморфизмом.

Определение. Подмножество

$$\mathrm{Ker}(\varphi) := \{a \in G \mid \varphi(a) = 0\}$$

называется *ядром* гомоморфизма колец $\varphi : R \to S$.

Лемма. *Пусть $\varphi : R \to S$ – гомоморфизм колец. Тогда верны следующие утверждения.*

(1) *Ядро* $\mathrm{Ker}(\varphi)$ *является подкольцом в* R.

(2) *Образ* $\varphi(R)$ *является подкольцом в* S.

(3) *Гомоморфизм* φ *инъективен тогда и только тогда, когда* $\mathrm{Ker}(\varphi) = \{0\}$.

(4) *Гомоморфизм* φ *является изоморфизмом тогда и только тогда, когда* $\mathrm{Ker}(\varphi) = \{0\}$ *и* $\varphi(R) = S$.

Лекция 10

Поля. Определение, свойства, примеры. Изоморфизм полей. Поле комплексных чисел. Аксиоматическое определение, существование, единственность. Алгебраическая запись. Вещественная и мнимая части. Комплексное сопряжение. Тригонометрическая форма комплексного числа. Формула Муавра.

Определение. *Полем* называется ассоциативное коммутативное кольцо $\Bbbk$ такое, что любой ненулевой элемент $\Bbbk$ обратим.

Таким образом, поле это множество $\Bbbk$ с двумя операциями – сложения $a+b$ и умножения $a \cdot b$ такое, что выполнены следующие условия:

(I) $\Bbbk$ – группа по сложению;

(II) $\Bbbk^* := \Bbbk \setminus \{0\}$ – группа по умножению;

(III) для любых элементов $a,\ b,\ c \in \Bbbk$ имеет место равенство

$$(a+b) \cdot c = a \cdot c + b \cdot c.$$

Подмножество $\Bbbk_0$ поля $\Bbbk$ называется *подполем*, если оно является полем с теми же операциями, что и в $\Bbbk$. Если же сначала рассматривалось меньшее поле $\Bbbk_0$, а потом – большее $\Bbbk$, в котором $\Bbbk_0$ является подполем, то переход от $\Bbbk_0$ к $\Bbbk$ обычно называется *расширением полей*.

Примеры. (1) Стандартные примеры – это поле рациональных чисел $\mathbb{Q}$ и поле действительных чисел $\mathbb{R}$.

(2) $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a,\, b \in \mathbb{Q}\} \subset \mathbb{R}$.

(3) $\mathbb{Q}[\sqrt[3]{2}] = \{a + b\sqrt{2} + c\sqrt{4} \mid a,\, b,\, c \in \mathbb{Q}\} \subset \mathbb{R}$.

(4) Множество всех рациональных функций $\mathbb{R}(t)$ является полем.

Так как поле – это частный случай кольца, то для полей определены понятия гомоморфизмы и изоморфизма. Однако, следующее утверждение показывает, что гомоморфизм полей обязательно является вложением и поэтому этот термин в отношении полей не употребляется.

Утверждение. *Пусть $f : \mathbb{F} \to R$ – (ненулевой) гомоморфизм колец. Если $\mathbb{F}$ – поле, то f инъективен. Следовательно, f является изоморфизмом $\mathbb{F}$ с подполем $f(\mathbb{F}) \subset R$.*

Доказательство. Предположим, что не является инъективным. Тогда $\mathrm{Ker}(f) \neq \{0\}$, т.е. существует элемент такой, что $f(c) = 0$, $c \neq 0$. В этом случае

$$f(1) = f(c \cdot c^{-1}) = f(c) \cdot f(c^{-1}) = 0$$

и поэтому для любого элемента $x \in \mathbb{F}$ имеем $f(x) = f(x \cdot 1) = f(x) \cdot f(1) = 0$, т.е. $f(\mathbb{F}) = 0$. Противоречие. □

По определению, любое поле содержит по крайней мере два элемента: 0 и 1. Несложно построить поле из двух элементов. Зададим его таблицами сложения и умножения:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Аналогично можно построить поле из трех элементов:

$+$	0	1	a
0	0	1	a
1	1	a	0
a	a	0	1

$\cdot$	0	1	a
0	0	0	0
1	0	1	a
a	0	a	1

Отметим, что условия того, что уравнения $a + x = b$ и $a \cdot x = b$ и условия коммутативности позволяют заполнить эти таблицы однозначно. Однако, условия ассоциативности и дистрибутивности нужно проверять отдельно. Читателю предлагается самостоятельно проделать эту несложную проверку.

Поле комплексных чисел

Определение (аксиоматическое определение)**.** *Полем комплексных чисел* называется любое поле $\mathbb{C}$, которое обладает следующими свойствами:

(1) $\mathbb{C}$ содержит подполе $\mathbb{R}$ изоморфное полю действительных чисел;

(2) $\mathbb{C}$ содержит элемент $\boldsymbol{i}$ такой, что $\boldsymbol{i}^2 = -1$;

(3) если существует подполе $\mathbb{K}$ такое, что $\mathbb{R} \subset \mathbb{K} \subset \mathbb{C}$, то $\mathbb{K} = \mathbb{C}$ или $\mathbb{K} = \mathbb{R}$.

Элементы $\mathbb{C}$ называются *комплексными числами*.

Пока мы не знаем, что поле комплексных чисел существует, а также не знаем, что оно единственно. Однако, мы можем доказать следующий факт.

Утверждение. *Пусть $\mathbb{C}$ – поле комплексных чисел, определенное выше. Тогда любой элемент $z \in \mathbb{C}$ можно однозначно записать в виде*

$$(*) \qquad z = a + b\boldsymbol{i}, \qquad a,\, b \in \mathbb{R}.$$

Представление $(*)$ называется *алгебраической формой* комплексного числа. В этом представлении число a называется *действительной частью* комплексного числа z, а число b – его *мнимой частью*. Обычно используют обозначения

$$a =: \mathrm{Re}(z), \qquad b =: \mathrm{Im}(z).$$

Доказательство. Рассмотрим множество

$$\mathbb{K} := \{a + b\boldsymbol{i} \mid a,\, b \in \mathbb{R}\}.$$

Проверим, что $\mathbb{K}$ – подполе в $\mathbb{C}$. Действительно, если $z,\, z' \in \mathbb{K}$, то $z = a+b\boldsymbol{i}$ и $z = a'+b'\boldsymbol{i}$ для некоторых $a,\, b,\, a',\, b' \in \mathbb{R}$. Тогда

$$\begin{aligned} z \pm z' &= (a + b\boldsymbol{i}) + (a' + b'\boldsymbol{i}) = (a \pm a') + (b \pm b')\boldsymbol{i} \in \mathbb{K} \\ z \cdot z' &= (a + b\boldsymbol{i}) \cdot (a' + b'\boldsymbol{i}) = aa' + bb'\boldsymbol{i}^2 + ab'\boldsymbol{i} + a'b\boldsymbol{i} = (aa' - bb') + (ab' + a'b)\boldsymbol{i} \in \mathbb{K}. \end{aligned}$$

Если $z \neq 0$, то

$$z^{-1} = \frac{a}{a^2 + b^2} + \frac{b}{a^2 + b^2}\boldsymbol{i} \in \mathbb{K}.$$

Следовательно, $\mathbb{K} \subset \mathbb{C}$ – подполе. Так как $\boldsymbol{i} \in \mathbb{K}$, то согласно условию (3) имеем $\mathbb{K} = \mathbb{C}$. Таким образом, любой элемент $z \in \mathbb{C}$ представляется в виде $z = a + b\boldsymbol{i}$.

Если же имеется два представления $z = a + b\boldsymbol{i} = a' + b'\boldsymbol{i}$, то $(b - b')\boldsymbol{i} = a' - a \in \mathbb{R}$. Если $b \neq b'$, то $\boldsymbol{i} = \frac{a'-a}{b-b'} \in \mathbb{R}$. Противоречие показывает, что $b = b'$ и тогда $a = a'$. □

Следствие. *Поле* $\mathbb{C}$ *является также векторным пространством над* $\mathbb{R}$ *с базисом* $1,\, \boldsymbol{i}$.

Таким образом, комплексные числа $z \in \mathbb{C}$ можно изображать элементами двумерного векторного пространства $\mathbb{R}^2$. При этом сложение комплексных чисел соответствует сложению векторов. Действительная $\mathrm{Re}(z)$ и мнимая $\mathrm{Im}(z)$ части являются координатами в этом базисе.

Теорема. *Поле комплексных чисел* $\mathbb{C}$ *существует и единственно с точностью до изоморфизма. Более точно, если* $\mathbb{C}_0$ *и* $\mathbb{C}_1$ *– два поля, удовлетворяющих свойствам* (1), (2), (3), *то существует изоморфизм* $\varphi : \mathbb{C}_0 \to \mathbb{C}_1$, *переводящий действительные числа* $\mathbb{R}_0 \subset \mathbb{C}_0$ *в действительные* $\mathbb{R}_1 \subset \mathbb{C}_1$.

Доказательство. Существование. Рассмотрим множество квадратных матриц второго порядка

$$\mathbb{C}_0 = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \;\middle|\; a,\, b \in \mathbb{R} \right\} \subset \mathrm{Mat}_2(\mathbb{R}).$$

Легко видеть, что $\mathbb{C}_0$ – коммутативное подкольцо в $\mathrm{Mat}_2(\mathbb{R})$. Действительно,

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} + \begin{pmatrix} a' & b' \\ -b' & a' \end{pmatrix} = \begin{pmatrix} a + a' & b + b' \\ -(b + b') & a + a' \end{pmatrix},$$

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} a' & b' \\ -b' & a' \end{pmatrix} = \begin{pmatrix} aa' - bb' & ab' + a'b \\ -ab' - a'b & aa' - bb' \end{pmatrix} = \begin{pmatrix} a' & b' \\ -b' & a' \end{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Так как

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}^{-1} = \frac{1}{a^2+b^2}\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \mathbb{C},$$

то $\mathbb{C}_0$ – поле. Подмножество диагональных матриц

$$\mathbb{R}_0 := aE = \left\{ \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \;\middle|\; a \in \mathbb{R} \right\} \subset \mathbb{C}_0$$

является подполем, изоморфным $\mathbb{R}$. Изоморфизм $\mathbb{R}_0 \simeq \mathbb{R}$ задается формулой $aE \longmapsto a$. Ясно, что $\mathbb{C}_0$ является также векторным пространством над $\mathbb{R}_0$ с базисом E, I, где

$$I := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

и $I^2 = -E$. Наконец, любое промежуточное поле $\mathbb{R}_0 \subset \mathbb{K} \subset \mathbb{C}_0$ также векторным пространством над $\mathbb{R}$. Так как $\dim_{\mathbb{R}_0} \mathbb{C}_0 = 2$, то или $\mathbb{K} = \mathbb{R}_0$ или $\mathbb{K} = \mathbb{C}_0$.

Единственность. Пусть $\mathbb{C}_1$ – любое другое поле, удовлетворяющее свойствам (1), (2), (3). Построим изоморфизм φ между нашим полем $\mathbb{C}_0$, построенным выше, и $\mathbb{C}_1$:

$$\varphi : \mathbb{C}_0 \longrightarrow \mathbb{C}_1, \qquad \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \longmapsto a + b\boldsymbol{i}.$$

Несложно проверяется, что φ – гомоморфизм колец. Следовательно, φ – вложение и $\varphi(\mathbb{C}_0)$ – подполе в $\mathbb{C}_1$, содержащее $\mathbb{R}_1 = \varphi(\mathbb{R}_0)$ и элемент $\boldsymbol{i} := \varphi(I)$, удовлетворяющий $\boldsymbol{i}^2 = -1$. Значит $\varphi(\mathbb{C}_0) = \mathbb{C}_1$ согласно свойству (3). □

Следствие (алгебраическая запись комплексных чисел)**.** $\mathbb{C}$ *является двумерным векторным пространством над* $\mathbb{R}$ *и, таким образом, каждое комплексное число* $z \in \mathbb{C}$ *единственным образом представляется в виде* $z = a + b\boldsymbol{i}$, a, $b \in \mathbb{R}$.

Отображение

$$\mathbb{C} \longrightarrow \mathbb{C}, \qquad z = a + b\boldsymbol{i} \longmapsto \bar{z} := a - b\boldsymbol{i}$$

называется *комплексным сопряжением*. При геометрической реализации поля $\mathbb{C}$, как $\mathbb{R}^2$, комплексное сопряжение является отражением относительно действительной оси. При матричной реализации поля $\mathbb{C}$ комплексное сопряжение является транспонированием. Отсюда легко видеть, что комплексное сопряжение биективно и удовлетворяет следующим свойствам:

$$\overline{z+w} = \overline{z} + \overline{w}, \quad \overline{z \cdot w} = \overline{z} \cdot \overline{w}.$$

Следовательно, комплексное сопряжение – изоморфизм $\mathbb{C}$ на себя. Напомним, что изоморфизм некоторого объекта на себя называется *автоморфизмом*. Таким образом, комплексное сопряжение – автоморфизм поля $\mathbb{C}$.

Отметим, что $z = \bar{z}$ тогда и только тогда, когда $z \in \mathbb{R}$.

Модулем комплексного числа $z = a + b\boldsymbol{i}$ называется неотрицательное действительное число

$$|z| := \sqrt{a^2+b^2} = \sqrt{z \cdot \bar{z}}.$$

Пусть $|z| \neq 0$. Тогда $z/|z|$ – комплексное число, модуль которого равен 1, т.е. $z/|z| = a' + b'\boldsymbol{i}$, где $a' = a/|z|$, $b' = b/|z|$, $a'^2 + b'^2 = 1$. Следовательно, $a' = \cos(\varphi)$, $b' = \sin(\varphi)$ для некоторого $\varphi \in \mathbb{R}$. Это φ (определенное по модулю 2π) называется *аргументом* z. Таким образом, каждое ненулевое комплексное число z однозначно представляется в виде

$$z = r(\cos(\varphi) + \boldsymbol{i}\sin(\varphi)).$$

Это представление называется *тригонометрической* формой комплексного числа. Ясно, что $r'(\cos(\varphi)' + \boldsymbol{i}\sin(\varphi)') = r(\cos(\varphi) + \boldsymbol{i}\sin(\varphi))$ тогда и только тогда, когда $r' = r$ и $\varphi' = \varphi + 2\pi k$, $k \in \mathbb{Z}$.

Теорема. *Пусть*

$$z = r(\cos(\varphi) + \boldsymbol{i}\sin(\varphi)) \qquad u \qquad w = q(\cos(\psi) + \boldsymbol{i}\sin(\psi))$$

– тригонометрические формы двух комплексных чисел. Тогда

(1) $z \cdot w = rq(\cos(\varphi + \psi) + \boldsymbol{i}\sin(\varphi + \psi))$;

(2) $z/w = rq(\cos(\varphi - \psi) + \boldsymbol{i}\sin(\varphi - \psi))$.

Таким образом,

$$\begin{aligned} |z \cdot w| &= |z| \cdot |w|, \\ \arg(z \cdot w) &= \arg(z) + \arg(w), \\ \arg(z/w) &= \arg(z) - \arg(w) \qquad (\textit{если } z \neq 0 \ \textit{и} \ w \neq 0). \end{aligned}$$

Доказательство. Докажем, например, ((1)). Имеем

$$\begin{aligned} z \cdot w = r(\cos(\varphi) + \boldsymbol{i}\sin(\varphi)) \cdot q(\cos(\psi) + \boldsymbol{i}\sin(\psi)) = \\ = rq(\cos(\varphi)\cos(\psi) + \sin(\varphi)\sin(\psi)\boldsymbol{i}^2 + (\cos(\varphi)\sin(\psi)\sin(\varphi)\cos(\psi))) = \\ = rq(\cos(\varphi + \psi) + \boldsymbol{i}\sin(\varphi + \psi)). \quad \square \end{aligned}$$

Теорема (формула Муавра). *Пусть* $z = r(\cos(\varphi) + \boldsymbol{i}\sin(\varphi))$ *и* $n \in \mathbb{Z}$. *Тогда*

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

Доказательство. Пусть $n > 0$. Докажем утверждение индукцией по n. База индукции $n = 1$ очевидна. Предположим, что формула Муавра верна для n. Имеем

$$\begin{aligned} z^{n+1} = z^n \cdot z = r^n(\cos(n\varphi) + \boldsymbol{i}\sin(n\varphi)) \cdot r(\cos(\varphi) + \boldsymbol{i}\sin(\varphi)) = \\ = r^{n+1}(\cos((n+1)\varphi) + \boldsymbol{i}\sin((n+1)\varphi)). \end{aligned}$$

Если же $n < 0$, то мы можем применить уже доказанный случай для $m = -n > 0$:

$$\begin{aligned} z^n = \frac{1}{z^m} = \frac{1}{r^m(\cos(m\varphi) + \boldsymbol{i}\sin(m\varphi))} = \\ = \frac{1}{r^m}(\cos(-m\varphi) + \boldsymbol{i}\sin(-m\varphi)) = r^n(\cos(n\varphi) + \boldsymbol{i}\sin(n\varphi)). \quad \square \end{aligned}$$

Лекция 11

Решения уравнения $z^n = w$. Группа $\boldsymbol{\mu}_n$ корней из 1. Первообразные корни. Порядок элемента в группе. Циклические группы. Примеры. Подгруппа циклической группы. Кольца вычетов. Делители нуля и обратимые элементы в $\mathbb{Z}/n\mathbb{Z}$. Конечное ассоциативное кольцо без делителей нуля является телом.

Корни в поле комплексных чисел

Для натурального n *корнем n-ой степени* из комплексного числа w называется комплексный z корень уравнения $z^n = w$.

Напомним, что целые числа a и b называются *сравнимыми по модулю* натурального числа n если n делит $a - b$. В этом случае пишут

$$a \equiv b \mod n.$$

Теорема. *Пусть*

$$w = q\big(\cos(\psi) + \boldsymbol{i}\sin(\psi)\big)$$

– тригонометрическая форма ненулевого комплексного числа и пусть $n \in \mathbb{N}$. Тогда уравнение $z^n = w$ имеет ровно n корней. Они могут быть записаны в виде

$$z_k = \sqrt[n]{q}\left(\cos\frac{\psi + 2\pi k}{n} + \boldsymbol{i}\sin\frac{\psi + 2\pi k}{n}\right),$$

где $k \in \mathbb{Z}$, а $\sqrt[n]{q}$ – обычный вещественный корень n-ой степени. При этом $z_k = z_l$ тогда и только тогда, когда $k \equiv l \mod n$. Таким образом,

$$z_0, z_1, \ldots, z_{n-1}$$

– все различные корни уравнения $z^n = w$.

Доказательство. Ясно, что $z \neq 0$. Поэтому z можно представить в тригонометрическом виде

$$z = r\big(\cos(\varphi) + \boldsymbol{i}\sin(\varphi)\big).$$

По формуле Муавра

$$z^n = r^n\big(\cos(n\varphi) + \boldsymbol{i}\sin(n\varphi)\big) = w = q\big(\cos(\psi) + \boldsymbol{i}\sin(\psi)\big).$$

Отсюда получаем $r^n = q$ и $n\varphi = \psi + 2\pi k$. Отсюда $r = \sqrt[n]{q}$ и $\varphi = (\psi + 2\pi k)/n$. □

Обозначим через $\boldsymbol{\mu}_n$ множество всех корней степени n из 1. Таким образом,

$$\boldsymbol{\mu}_n = \{\boldsymbol{\epsilon} \in \mathbb{C} \mid \boldsymbol{\epsilon}^n = 1\}.$$

Согласно теореме, $\boldsymbol{\mu}_n$ содержит ровно n элементов $\boldsymbol{\epsilon}_0, \boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_{n-1}$, где

$$(*) \qquad \boldsymbol{\epsilon}_k = \sqrt[n]{q}\left(\cos\frac{2\pi k}{n} + \boldsymbol{i}\sin\frac{2\pi k}{n}\right).$$

Предложение. $\boldsymbol{\mu}_n$ – *подгруппа в* $\mathbb{C}^*$.

Доказательство. Действительно, если $\boldsymbol{\epsilon}, \boldsymbol{\epsilon}' \in \boldsymbol{\mu}_n$, то $\boldsymbol{\epsilon}^n = \boldsymbol{\epsilon}'^n = 1$. Отсюда $(\boldsymbol{\epsilon}\boldsymbol{\epsilon}')^n = 1$ и $(\boldsymbol{\epsilon}^{-1})^n = 1$. Следовательно, $\boldsymbol{\epsilon}\boldsymbol{\epsilon}' \in \boldsymbol{\mu}_n$ и $\boldsymbol{\epsilon}^{-1} \in \boldsymbol{\mu}_n$. □

Определение. Элемент $\boldsymbol{\epsilon} \in \boldsymbol{\mu}_n$ называется *первообра́зным* корнем степени n из 1, если он не является корнем меньшей степени из 1.

Предложение. *Для любого n первообразные корни степени n из 1 существуют. Все корни степени n из 1 являются степенями первообразного.*

Доказательство. Действительно, согласно $(*)$ и формуле Муавра имеем $(\boldsymbol{\epsilon}_1)^k = \boldsymbol{\epsilon}_k$ и $(\boldsymbol{\epsilon}_1)^k \neq 1$ при $k = 1, \ldots, n-1$. Следовательно, $\boldsymbol{\epsilon}_1$ – первообразный корень степени n из 1 и все остальные корни являются его степенями. Для любого другого первообразного корня $\boldsymbol{\epsilon} \in \boldsymbol{\mu}_n$ все элементы $\boldsymbol{\epsilon}^0, \boldsymbol{\epsilon}^1, \ldots, \boldsymbol{\epsilon}^{n-1}$ различны. Значит они заполняют $\boldsymbol{\mu}_n$. □

Порядок элемента в группе

Пусть G – группа. Напомним, что ее *порядком* называется число ее элементов. Порядок группы обозначается $|G|$. Назовем *порядком элемента* $a \in G$ число

$$|a| = \begin{cases} \min\{k \in \mathbb{N} : a^k = 1\} & \text{если существует } k \text{ такое, что } a^k = 1; \\ \infty & \text{если } a^k \neq 1 \text{ для всех } k. \end{cases}$$

Замечание. (1) В конечной группе порядок любого элемента конечен и не превосходит порядка группы. Действительно, пусть $|G| = N$. Тогда среди элементов $a, a^2, \ldots, a^{N+1}$ есть повторяющиеся, т.е. $a^k = a^l$ для некоторых $0 < k < l \leq N+1$. Значит $a^{l-k} = 1$, где $0 < l - k \leq N$.

(2) Те же соображения показывают, что если $|a| = \infty$, то все элементы a^k, $k \in \mathbb{Z}$ различны.

Примеры. (1) Группы $\mathbb{Z}^+$, $\mathbb{Q}^+$, $\mathbb{R}^+$, $\mathbb{C}^+$ не содержат нетривиальных элементов конечного порядка.

(2) Элементы конечного порядка в группе $\mathbb{C}^*$ – это корни из 1.

Предложение. *Пусть $|a| = n < \infty$. Тогда имеют место следующие утверждения:*

(1) $a^k = a^l$ *тогда и только тогда, когда* $k \equiv l \mod n$;

(2) $a^k = 1$ *тогда и только тогда, когда* k *делится на* n*;*

(3) *все элементы* $1, a, \ldots, a^{n-1}$ *различны;*

(4) *для любого* $k \in \mathbb{Z}$ *элемент* a^k *содержится в множестве* $\{1, a, \ldots, a^{n-1}\}$*;*

(5) $|a^m| = n/\text{НОД}(n, m)$.

Доказательство. Для того, чтобы доказать (1) разделим $k - l$ на n с остатком:

$$k - l = qn + r, \qquad 0 \leq r < n.$$

Тогда

$$a^{k-l} = a^{qn+r} = (a^n)q \cdot a^r.$$

Значит, $a^k = a^l$ тогда и только тогда, когда $r = 0$. Утверждения (2), (3) и (4) непосредственно следуют из (1).

Докажем (5). Положим $d := \text{НОД}(n, m)$ $n' = n/d$ и $m' = m/d$. Тогда $(a^m)^{n'} = a^{nm/d} = (a^n)^{m'} = 1$. Следовательно, $|a| \leq n'$. Обратно, пусть $(a^m)^k = 1$ для некоторого $k \in \mathbb{N}$. Тогда $a^{mk} = 1$ и, согласно пункту (2), имеем $n'd = n \mid mk = m'kd$. Следовательно, $n' \mid k$ и $n' \leq k$, поскольку $\text{НОД}(n', m') = 1$. □

Следствие. $|a| = |a^m|$ *тогда и только тогда, когда* $\text{НОД}(n, m) = 1$.

Предложение. *Пусть* G *– любая группа и пусть* $a, b \in G$ *– элементы такие, что* $a \cdot b = b \cdot a$. *Пусть* $|a| = n$, $|b| = m$ *и* $\text{НОД}(n, m) = 1$. *Тогда* $|a \cdot b| = nm$.

Доказательство. Имеем $(ab)^{nm} = 1$. С другой стороны, если $(ab)^k = 1$, то $a^k = b^{-k}$, $1 = a^{nk} = b^{-nk}$. Следовательно, $-nk \equiv 0 \mod m$ и $m \mid k$. □

Циклические группы

Определение. Группа G называется *циклической*, если существует элемент $a \in G$ такой, что любой элемент $g \in G$ является степенью a, т.е.

$$G = \{a^k \mid k \in \mathbb{Z}\}.$$

При этом элемент a называется *образующим* или *порождающим элементом* циклической группы G.

Если G – циклическая группа, порожденная элементом a, то мы будем писать $G = \langle a \rangle$. Заметим, что в любой группе G любой элемент $a \in G$ порождает циклическую подгруппу

$$\langle a \rangle = \{a^k \mid k \in \mathbb{Z}\}.$$

Она совпадает с пересечением всех подгрупп в G, содержащих a. Более того, любая группа является объединением своих циклических подгрупп:

$$G = \bigcup_{a \in G} \langle a \rangle.$$

Примеры. (1) Группа $\boldsymbol{\mu}_n$ – циклическая, она порождается любым первообразным корнем из 1.

(2) Группа целых чисел $\mathbb{Z}$ (по сложению) является циклической.

(3) Группа $\mathbb{R}^+$ – не является циклической, она несчетна.

Изоморфизм циклических групп одного порядка.

Теорема. (1) *Все циклические группы одного порядка изоморфны.*

(2) *Бесконечная циклическая группа изоморфна* $\mathbb{Z}$.

(3) *Конечная циклическая группа порядка* n *изоморфна* $\boldsymbol{\mu}_n$.

Доказательство. Пусть G – циклическая группа, порожденная элементом a, и H – циклическая группа (того же порядка), порожденная элементом b. Определим гомоморфизм $\varphi : G \to H$ формулой

$$(\dagger) \qquad \varphi(a^k) = b^k.$$

Во-первых, докажем корректность этого определения. Пусть $a_k = a^l$ и $k \neq l$. Тогда порядок группы G должен быть конечен. Обозначим $n := |G| = |H|$. Имеем $k = l + nq$ для некоторого $q \in \mathbb{Z}$. Следовательно,

$$\begin{aligned} \varphi(a^k) &= b^k = b^{l+nq}) = b^l b^{nq} = b^l, \\ \varphi(a^l) &= b^l. \end{aligned}$$

Это показывает, что $\varphi(a^k) = \varphi(a^l)$, т.е. доказывает корректность определения φ. Далее,

$$\varphi(a^k \cdot a^l) = \varphi(a^{k+l}) = b^{k+l} = b^k \cdot b^l = \varphi(a^k) \cdot \varphi(a^l).$$

Следовательно, φ – гомоморфизм групп. Согласно (†) он сюръективен. Предположим, что существует $a^k \in \mathrm{Ker}(\varphi)$, $a^k \neq 1$. Тогда $k \neq 0$ и $0 = \varphi(a^k) = b^k$. Это возможно только если группа H конечна и k делится на ее порядок. Но тогда $a^k = 1$, поскольку $|G| = |H|$. Противоречие показывает, что гомоморфизм φ также инъективен. Следовательно, он – изоморфизм. Это доказывает (1). Утверждения (2) и (3) следуют из (1). □

Следствие. *Порядок циклической группы равен порядку порождающего ее элемента.*

Теорема. (1) *Подгруппа циклической группы – циклическая.*

(2) *Порядок подгруппы в конечной циклической группе делит порядок группы.*

(3) $m \mid n$, *то в циклической группе порядка* n *существует подгруппа порядка* m *и эта подгруппа – единственная.*

Доказательство. Докажем (1). Пусть $G = \langle a \rangle$ и пусть $H \subset G$ – подгруппа. Возьмем

$$m := \min\{k \in \mathbb{N} \mid a^k \in H\}.$$

Пусть $b := a^m$ и пусть $c \in H$. Тогда $c = a^l$. Разделим l на m с остатком: $l = mq + r$, где $0 \leq r < m$. Получим $a^r = a^{l-mq} = a^l \cdot (a^m)^{-q} \in H$. Из минимальности выбора m получаем $r = 0$ и $c = a^l = a^{mq} = b^q$. Таким образом, $H = \langle b \rangle$. Это доказывает (1).

Теперь докажем (2). Пусть $G = \langle a \rangle$ – циклическая группа порядка n и пусть $H \subset G$ – ее подгруппа. Согласно утверждению (1) имеем $H = \langle a^m \rangle$ для некоторого m. Отсюда $|H| = |a^m| = n/\text{НОД}(n, m)$.

Наконец, докажем (3). Пусть $G = \langle a \rangle$ – циклическая группа порядка n и пусть $n = mq$. Тогда положим $H = \langle a^q \rangle$. Имеем $|H| = |a^q| = n/\text{НОД}(n, q) = m$. Единственность доказывается как и в (1). □

Подгруппы, порожденные множеством

Определение. Пусть G – группа и пусть $S \subset G$ – подмножество. Рассмотрим множество всевозможных произведений степеней элементов S:

$$\langle S \rangle := \{a_1^{n_1} \cdots a_m^{n_m} \mid a_i \in S, \quad n_i \in \mathbb{Z}\}.$$

Ясно, что $\langle S \rangle$ – подгруппа в G. Она называется группой, *порожденной* подмножеством S. Говорят, что группа G *порождается* множеством S, если $\langle S \rangle = G$. Иначе говоря, любой элемент $g \in G$ представляется в виде произведения степеней элементов из S.

Замечание. $\langle S \rangle$ – наименьшая подгруппа в G, содержащая S.

Примеры. (1) Симметрическая группа S_n порождается транспозициями.

(2) Полная линейная группа $\mathrm{GL}_n(\Bbbk)$ порождается элементарными матрицами.

(3) $\mathbb{Q}^*$ порождается простыми числами и -1.

Замечание. Не всякая группа порождается конечным числом элементов. Например, группа $\mathbb{R}^+$ несчётна, поэтому не может порождаться конечным числом элементов.

Кольца вычетов

Подмножество

$$\bar{a} := \{b \in \mathbb{Z} \mid b \equiv a \mod n\}$$

называется *классом вычетов* по модулю n. По определению $\bar{a} = \bar{a}'$ тогда и только тогда, когда $a \equiv a' \mod n$. Таким образом, для любого $\bar{a}$ мы можем записать $\bar{a} = \bar{a}'$, где $0 \leq a' < n$. Множество всех классов обозначается через $\mathbb{Z}/n\mathbb{Z}$ (или $\mathbb{Z}_n$). Согласно сказанному выше, $\mathbb{Z}/n\mathbb{Z} = \{\bar{0}, \bar{1}, \ldots, \overline{n-1}\}$.

Пример. Пусть $n = 2$. Тогда $\mathbb{Z}/2\mathbb{Z}$ состоит из двух элементов $\bar{0}$ (четные числа) и $\bar{1}$ (нечетные числа).

Определим сложение и умножение элементов $\mathbb{Z}/n\mathbb{Z}$ по правилам

$$(\ddagger) \qquad \bar{a} + \bar{b} = \overline{a+b}, \qquad \bar{a} \cdot \bar{b} = \overline{a \cdot b}.$$

Корректность определения: пусть $\bar{a} = \bar{a}'$ и $\bar{b} = \bar{b}'$. Тогда $a' = a + nk$ и $b' = b + nl$. Отсюда

$$a' + b' = (a + nk) + (b + nl) = a + b + n(k + l), \qquad \overline{a' + b'} = \overline{a + b},$$

$$a'b' = (a + nk)(b + nl) = ab + n(k + l + nkl), \qquad \overline{a'b'} = \overline{ab}.$$

Теорема. *$\mathbb{Z}/n\mathbb{Z}$ – коммутативное ассоциативное кольцо с 1.*

Доказательство. Следует из определения операций $(\ddagger)$ и соответствующих свойств кольца $\mathbb{Z}$. □

Предложение. *Пусть $n \in \mathbb{N}$, $a \in \mathbb{Z}$. Следующие условия эквивалентны:*

(1) *элемент $\bar{a} \in \mathbb{Z}/n\mathbb{Z}$ обратим;*

(2) *элемент $\bar{a} \in \mathbb{Z}/n\mathbb{Z}$ не является делителем нуля;*

(3) $\text{НОД}(a, n) = 1$.

Доказательство. Импликация (1) $\Longrightarrow$ (2) очевидна: если $\bar{a} \in \mathbb{Z}/n\mathbb{Z}$ обратим и $\bar{a}\bar{b} = \bar{0}$, то $\bar{b} = (\bar{a})^{-1}\bar{a}\bar{b} = \bar{0}$. Импликация (2) $\Longrightarrow$ (3) также очевидна: если $\text{НОД}(a, n) = 1$ и $\bar{a}\bar{b} = \bar{0}$, то $ab \equiv 0 \mod n$ и тогда $b \equiv 0 \mod n$, т.е. $\bar{b} = \bar{0}$. Наконец, докажем (3) $\Longrightarrow$ (1). Пусть $\text{НОД}(a, n) = 1$. Тогда существуют $b, c \in \mathbb{Z}$ такие, что $ab + nc = 1$. Значит, $\bar{1} = \overline{ab + nc} = \bar{a}\bar{b} + \bar{n}\bar{c} = \bar{a}\bar{b}$, т.е. $\bar{a}$ обратим. □

Эквивалентность (1) и (2) следует также из более общего факта:

Предложение. *Пусть R – конечное ассоциативное кольцо с 1 и пусть $a \in R$. Следующие условия эквивалентны:*

(1) *a обратим;*

(2) *a не является делителем нуля.*

Доказательство. Импликация (1) $\Longrightarrow$ (2) очевидна. Для доказательства (2) $\Longrightarrow$ (1) рассмотрим отображение $\varphi : R \to R$, $x \mapsto ax$. Если a не является делителем нули, то это отображение инъективно, а поскольку R – конечное множество, то φ и сюръективно. Поэтому существует $a' \in R$ такой, что $1 = \varphi(a') = aa'$. □

Лекция 12

Поля $\mathbb{F}_p$. Теорема Вильсона. Характеристика поля. Свойства полей характеристики p. Отображение Фробениуса. Алгебры над полем. Конечномерная ассоциативная алгебра без делителей нуля является алгеброй с делением. Кольцо многочленов. Универсальное свойство. Кольцо формальных степенных рядов. Подстановка элемента кольца в многочлен.

Поля $\mathbb{F}_p$

Следствие. *Пусть R – конечное коммутативное ассоциативное кольцо с 1 без делителей нуля. Тогда R – поле.*

Следствие. *Следующие условия эквивалентны:*

(1) $\mathbb{Z}/n\mathbb{Z}$ *– поле;*

(2) $\mathbb{Z}/n\mathbb{Z}$ *не имеет делителей нуля;*

(3) n *– простое число.*

Поле $\mathbb{Z}/p\mathbb{Z}$ (где p – простое число) обозначается через $\mathbb{F}_p$. Таким образом, $\mathbb{F}_2$ и $\mathbb{F}_3$ – поля из двух и трех элементов, соответственно, построенные ранее.

Следствие (теорема Вильсона)**.** *Натуральное число p является простым тогда и только тогда, когда*

$$(p-1)! \equiv -1 \mod p.$$

Доказательство. Действительно, в произведении $\bar{1}\cdot\bar{2}\cdots\overline{p-1}$ все элементы, кроме $\overline{p-1}$ разбиваются на пары взаимно обратных. □

Характеристика поля

Определение. Пусть $\Bbbk$ – поле. Если $\underbrace{1+\cdots+1}_{n} \neq 0$ для любого n, то говорят, что $\Bbbk$ – поле характеристики 0. Если $\underbrace{1+\cdots+1}_{n} = 0$ для некоторого n, то характеристикой $\Bbbk$ называется минимальное n такое, что $\underbrace{1+\cdots+1}_{n} = 0$.

Обычно характеристика поля обозначается $\mathrm{char}(\Bbbk)$.

Примеры. (1) Поля $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ имеют характеристику 0.

(2) Поля $\mathbb{F}_p$ имеют характеристику $p > 0$.

(3) Если поле $\mathbb{K}$ имеет характеристику n, а $\Bbbk \subset \mathbb{K}$ – подполе, то и характеристика $\Bbbk$ равна n.

Предложение. *Характеристика любого поля или равна* 0 *или является простым числом.*

Доказательство. Пусть n – характеристика поля $\Bbbk$. Предположим, что $n > 0$ и $n = n_1 n_2$, $n_i > 1$. Положим $\gamma := \underbrace{1 + \cdots + 1}_{n_1}$. Тогда

$$0 = \underbrace{1 + \cdots + 1}_{n} = \underbrace{\gamma + \cdots + \gamma}_{n_2} = \gamma \cdot (\underbrace{1 + \cdots + 1}_{n_2}).$$

Отсюда $\underbrace{1 + \cdots + 1}_{n_2} = 0$. Противоречие. □

Замечание. В любой аддитивной абелевой группы определено умножение ее элементов на целые числа. Для $n \in \mathbb{Z}$ и $a \in A$ положим

$$n \cdot a = \begin{cases} \underbrace{a + \cdots + a}_{n} & \text{если } n > 0, \\ 0 & \text{если } n = 0, \\ -((-n) \cdot a) & \text{если } n < 0. \end{cases}$$

Эта операция удовлетворяет стандартным свойствам (сравните с определением векторного пространства):

(1) $(n + m) \cdot a = n \cdot a + m \cdot a$ для любого $a \in A$, для любых n, $m \in \mathbb{Z}$;

(2) $n \cdot (a + b) = n \cdot a + n \cdot b$ для любых a, $b \in A$, для любого $n \in \mathbb{Z}$;

(3) $(nm) \cdot a = n \cdot (m \cdot a)$ для любого $a \in A$, для любых n, $m \in \mathbb{Z}$;

(4) $1 \cdot a = a$, для любого $a \in A$ (здесь $1 \in \mathbb{Z}$).

Например, свойство (3) для n, $m > 0$ доказывается следующей выкладкой:

$$(nm) \cdot a = \underbrace{a + \cdots + a}_{nm} = \underbrace{\underbrace{(a + \cdots + a)}_{m} + \cdots + \underbrace{(a + \cdots + a)}_{m}}_{n} = \underbrace{m \cdot a + \cdots + m \cdot a}_{n} = n \cdot (m \cdot a).$$

Свойство (3) очевидно, если одно из чисел n, m равно 0. Если оба числа n, m отрицательны, то

$$(nm) \cdot a = ((-n)(-m)) \cdot a = (-n) \cdot ((-m) \cdot a) = -(n \cdot (-(m \cdot a))) = n \cdot (m \cdot a).$$

Случаи $nm < 0$ аналогичны.

В частности, в любом коммутативном ассоциативном кольце определена операция умножения целых чисел на элементы кольца. Имеет место формула бинома Ньютона:

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

Предложение. *В поле $\Bbbk$ характеристики $p > 0$ для любых $\alpha, \beta \in \Bbbk$ выполнено равенство*

$$(\alpha + \beta)^p = \alpha^p + \beta^p.$$

Доказательство. Действительно, в формуле бинома все коэффициенты, $\frac{p!}{k!(p-k)!}$ при $k \neq 0$, p делятся на p. Значит, соответствующие члены равны нулю. □

Отображение

$$\varphi : \Bbbk \longrightarrow \Bbbk, \qquad \alpha \to \alpha^p.$$

φ называется *отображением Фробениуса.*

Предложение. *Пусть $\Bbbk$ – поле характеристики $p > 0$. Тогда отображение Фробениуса $\varphi : \Bbbk \to \Bbbk$ является изоморфизмом со своим подполем $\varphi(\Bbbk)$. Если $\Bbbk$ – конечное поле, то $\varphi(\Bbbk) = \Bbbk$.*

Доказательство. Имеем

$$\begin{aligned} \varphi(\alpha\beta) &= (\alpha\beta)^p = \alpha^p\beta^p = \varphi(\alpha)\varphi(\beta), \\ \varphi(\alpha+\beta) &= (\alpha+\beta)^p = \alpha^p + \beta^p = \varphi(\alpha) + \varphi(\beta). \end{aligned}$$

Следовательно, φ – гомоморфизм. Так как $\Bbbk$ – поле, то $\mathrm{Ker}(\varphi) = \{0\}$, φ инъективен и $\varphi(\Bbbk) \subset \Bbbk$ – подполе. Наконец, если поле $\Bbbk$ конечно, то $\varphi(\Bbbk) = \Bbbk$ столько же элементов. Отсюда $\varphi(\Bbbk) = \Bbbk$. □

Понятие алгебры над полем

Определение. *Алгеброй* A над полем $\Bbbk$ называется множество с тремя операциями: сложения элементов A между собой, умножения элементов A между собой и умножения элементов поля $\Bbbk$ на элементы A так, что выполнены следующие условия:

(1) A является кольцом;

(2) A является векторным пространством над $\Bbbk$;

(3) для любого $\alpha \in \Bbbk$ для любых $\mathbf{a}, \mathbf{b} \in A$ $(\alpha\mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\alpha\mathbf{b}) = \alpha(\mathbf{a} \cdot \mathbf{b})$.

Говорят, что алгебра A ассоциативна, коммутативна, с единицей и т. д. если таковым является соответствующее кольцо. Говорят, что (ассоциативная) алгебра A является *алгеброй с делением*, если в A имеется единица и для любого $\mathbf{a} \in A$, $\mathbf{a} \neq 0$ существует обратный элемент.

Примеры. (1) Любое векторное пространство с нулевым умножением является алгеброй.

(2) Если $\mathbb{K}$ – поле, а $\Bbbk$ его подполе, то $\mathbb{K}$ является алгеброй над $\Bbbk$.

(3) Множество $\{a_n\}$ всех последовательностей действительных чисел является алгеброй над $\mathbb{R}$. Множество всех вещественнозначных функций на некотором множестве является алгеброй над $\mathbb{R}$.

(4) Трехмерное векторное пространство с операцией векторного умножения является алгеброй над $\mathbb{R}$.

(5) Множество $\mathrm{Mat}_n(\Bbbk)$ квадратных матриц порядка n с элементами из поля $\Bbbk$ является алгеброй над этим полем.

Гомоморфизмом алгебр над полем $\Bbbk$ называется отображение $f : A \to A'$ такое, что это гомоморфизм колец и $\Bbbk$-линейное отображение векторных пространств. Таким образом, гомоморфизмом алгебр – это отображение $f : A \to A'$ такое, что

(1) $f(\mathbf{a}+\mathbf{b}) = f(\mathbf{a}) + f(\mathbf{b})$ для любых элементов $\mathbf{a}$, $\mathbf{b} \in A$,

(2) $f(\mathbf{a}\cdot\mathbf{b}) = f(\mathbf{a}) \cdot f(\mathbf{b})$ для любых элементов $\mathbf{a}$, $\mathbf{b} \in A$,

(3) $f(\lambda\mathbf{a}) = \lambda f(\mathbf{a})$ для любых элементов $\mathbf{a} \in A$ и $\lambda \in \Bbbk$.

Гомоморфизм алгебр называется *изоморфизмом*, если у него имеется обратное отображение.

Замечание. Если A – ассоциативная алгебра с единицей над полем $\Bbbk$, то отображение

$$f : \Bbbk \to A, \qquad f(\lambda) = \lambda\,\mathbf{1}$$

является (ненулевым) гомоморфизмом алгебр. Так как $\Bbbk$ – поле, то f инъективен. Значит, поле $\Bbbk$ можно отождествить с подалгеброй в A

$$f(\Bbbk) = \{\lambda\,\mathbf{1} \mid \lambda \in \Bbbk\} \subset A.$$

Теорема. *Пусть A – ассоциативная конечномерная алгебра с единицей над полем $\Bbbk$. Если $\mathbf{a} \in A$ не является делителем 0 ($и$ $\mathbf{a} \neq 0$), то $\mathbf{a}$ обратим.*

Доказательство. Для некоторого m элементы $\mathbf{1}, \mathbf{a}, \ldots, \mathbf{a}^m$ линейно зависимы. Выберем это m минимальным. Таким образом,

$$\lambda_m\mathbf{a}^m + s\lambda_1\mathbf{a} + \lambda_0\mathbf{1} = 0, \quad \lambda_i \in \Bbbk.$$

Если $\lambda_0 = 0$, то

$$\mathbf{a}(\lambda_m\mathbf{a}^{m-1} + \lambda_{m-1}\mathbf{a}^{m-2} + \lambda_1\,\mathbf{1}) = 0,$$

где выражение в скобках не равно нулю по нашему предположению о минимальности m. Но тогда $\mathbf{a}$ – делитель нуля, что противоречит нашему предположению. Значит, $\lambda_0 \neq 0$. Тогда мы можем положить

$$\mathbf{a}^{-1} = -\frac{1}{\lambda_0}(\lambda_m\mathbf{a}^{m-1} + s\lambda_2\mathbf{a} + \lambda_1\,\mathbf{1}). \qquad \square$$

Следствие. *Пусть A – ассоциативная конечномерная алгебра с единицей. Если в A нет делителей нуля, то A – алгебра с делением. Если дополнительно алгебра A коммутативна, то A – поле.*

Структурные константы алгебры

Пусть A – алгебра над полем $\Bbbk$. Для простоты предположим, что она конечномерна. Если $\mathbf{e}_1, \dots, \mathbf{e}_n$ – базис A, то мы можем произведения $\mathbf{e}_i \cdot \mathbf{e}_j$ по базису:

$$\mathbf{e}_i \cdot \mathbf{e}_j = \sum_{k=1}^{n} \gamma_{ijk} \mathbf{e}_k.$$

Константы γ_{ijk} называются *структурными константами алгебры* A. Они однозначно задают умножение: если $\mathbf{a} = \sum \alpha_i \mathbf{e}_i$ и $\mathbf{b} = \sum \beta_j \mathbf{e}_j$, то

$$\mathbf{a} \cdot \mathbf{b} = \Big(\sum_i \alpha_i \mathbf{e}_i \Big) \cdot \Big(\sum_j \beta_j \mathbf{e}_j \Big) = \sum_{i,j} \alpha_i \beta_j \mathbf{e}_i \cdot \mathbf{e}_j = \sum_{i,j,k} \alpha_i \beta_j \gamma_{ijk} \mathbf{e}_k.$$

Пример (Групповые алгебры)**.** Пусть G – конечная группа. Рассмотрим векторное пространство A над полем $\Bbbk$ с базисом be_g, занумерованным элементами группы $g \in G$. Таким образом, каждому элементу $g \in G$ соответствует свой вектор. Определим умножение элементов базиса по правилу

$$\mathbf{e}_g \cdot \mathbf{e}_g = \mathbf{e}_{g \cdot h}.$$

Согласно сказанному выше, это определяет структуру алгебры на A. Легко видеть, что эта алгебра ассоциативна. Вектор $\mathbf{e}_1$ является единичным элементом в A. Построенная алгебра называется *групповой алгеброй* группы G. Обычно она обозначается $\Bbbk[G]$.

Многочлены

Определение. Пусть R – коммутативное ассоциативное кольцо. Многочленом от одной переменной над R называется бесконечная последовательность

$$(*) \qquad f = (a_0, a_1, \dots, a_n, \dots), \qquad a_i \in R$$

такая, что только конечное число членов a_i отличны от 0. Элементы a_i называются коэффициентами многочлена.

Таким образом, два многочлена считаются равными, если равны все их соответствующие коэффициенты. Нулевой многочлен – это нулевая последовательность.

Определение. Пусть f – ненулевой многочлен $(*)$. Его *степенью* называется максимальное n такое, что $a_n \neq 0$. Степень многочлена f обозначается $\deg(f)$.

Отметим, что степень нулевого многочлена не определена. Однако, иногда удобно считать ее равной $-\infty$.

Обозначим через S множество всех многочленов над R. Определим на множестве S операции сложения и умножения. Пусть f, $g \in S$. Запишем

$$f = (a_0, a_1, \dots, a_n, \dots), \qquad g = (b_0, b_1, \dots, b_n, \dots).$$

Тогда

$$(\dagger) \qquad f + g = (a_0 + b_0, a_1 + b_1, \dots, a_n + b_n, \dots)$$

$$(\ddagger) \qquad f \cdot g = (c_0, c_1, \dots, c_n, \dots), \qquad c_k = \sum_{i+j=k} a_i b_j.$$

Ясно, что множество S с операцией сложения (покомпонентного) является абелевой группой.

В последней формуле мы считаем, что $a_i = 0$ при $i < 0$ и $b_j = 0$ при $j < 0$. Таким образом,

$$c_0 = a_0 b_0, \quad c_1 = a_0 b_1 + a_1 b_0, \quad c_2 = a_0 b_2 + a_1 b_1 + a_2 b_0, \quad c_3 = a_0 b_3 + a_1 b_2 + a_2 b_1 + a_3 b_0, \dots$$

Ясно, что $c_k = 0$ при $k > \deg(f) + \deg(g)$. Поэтому $f \cdot g$ – многочлен.

Предложение. *Пусть R – коммутативное ассоциативное кольцо. Тогда множество S всех многочленов над R – также коммутативное ассоциативное кольцо.*

Доказательство. Из (‡) следует, что умножение в S коммутативно. Пусть $f = f' + f''$, где

$$f = (a_0, a_1, \dots, a_n, \dots), \qquad f' = (a'_0, a'_1, \dots, a'_n, \dots), \qquad f'' = (a''_0, a''_1, \dots, a''_n, \dots).$$

Запишем

$$f' \cdot g = (c'_0, c'_1, \dots, c'_n, \dots), \qquad f'' \cdot g = (c''_0, c''_1, \dots, c''_n, \dots).$$

Тогда

$$c_k = \sum_{i+j=k} a_i b_j = \sum_{i+j=k} (a'_i + a''_i) b_j = \sum_{i+j=k} a'_i b_j + \sum_{i+j=k} a''_i b_j = c'_k + c''_k.$$

Следовательно,

$$(f' + f'') \cdot g = f' \cdot g + f'' \cdot g.$$

Остается проверить ассоциативность умножения. Пусть

$$f = (a_0, a_1, \dots, a_n, \dots), \qquad g = (b_0, b_1, \dots, b_n, \dots), \qquad h = (c_0, c_1, \dots, c_n, \dots).$$

Тогда

$$f \cdot g = (d_0, d_1, \dots, d_n, \dots), \qquad d_l = \sum_{i+j=l} a_i b_j,$$

$$g \cdot h = (d'_0, d'_1, \dots, d'_n, \dots), \qquad d'_r = \sum_{j+k=r} b_j c_k,$$

$$(f\cdot g)\cdot h=(e_0,e_1,\dots,e_n,\dots),\quad e_m=\sum_{l+k=m}d_lc_k=\sum_{l+k=m}\left(\sum_{i+j=k}a_ib_j\right)c_k=\sum_{i+j+k=m}a_ib_jc_k,$$

$$f\cdot(g\cdot h)=(e'_0,e'_1,\dots,e'_n,\dots),\quad e'_m=\sum_{i+r=m}a_id'_r=\sum_{i+r=m}a_i\left(\sum_{j+k=r}b_jc_k\right)=\sum_{i+j+k=m}a_ib_jc_k.$$

Мы получили, что $e_m=e'_m$. Следовательно, $(f\cdot g)\cdot h=f\cdot(g\cdot h)$. Таким образом, S является коммутативным ассоциативным кольцом □

Кольцо R вкладывается в $R[t]$ по правилу

$$a\longmapsto(a,0,0,0,\dots).$$

Поэтому всюду далее кольцо R будет отождествляться с подкольцом в $R[t]$. Положим

$$t:=(0,1,0,0,\dots).$$

Лемма.

$$t^k:=(0,0,\dots,\underset{\substack{\uparrow\\ k}}{1},\dots).$$

Доказательство. Доказывается индукцией по k с использованием формул (‡). □

Отсюда следует, что $(*)$ можно записать в виде конечной суммы

$$(\S)\qquad f=a_0+a_1t+\dots+a_nt^n$$

для некоторых $a_i\in R$, $n\in\mathbb{N}$. Такое представление многочлена *единственно*. Традиционно многочлены записываются в таком виде. Эта запись и будет использоваться в дальнейшем. Выделенный элемент t называется независимой переменной. Множество всех многочленов обозначается через $R[t]$.

Предложение. *Пусть R – коммутативное ассоциативное кольцо. Предположим, что существует другое коммутативное ассоциативное кольцо Q и элемент $s\in Q$ такие, что*

(1) *R является подкольцом в Q;*

(2) *любой элемент $f\in Q$* однозначно *представляется в виде*

$$(\P)\qquad f=a_0+a_1s+\dots+a_ns^n,\quad a_i\in R.$$

Тогда существует изоморфизм

$$\varphi:R[t]\longrightarrow Q$$

такой, что $\varphi(a)=a$ для всех $a\in R$ и $\varphi(t)=s$.

Доказательство. Рассмотрим отображение

$$\varphi:R[t]\longrightarrow Q,\qquad\varphi\left(\sum a_kt^k\right)=\sum a_ks^k.$$

Из единственности представлений (§) и (¶) следует, что оно корректно определено и является биекцией, а из определения операций (†) и (‡) следует, что φ – изоморфизм. □

Кольцо формальных степенных рядов

В представлении многочленов последовательностями $(*)$ требовалось, чтобы только конечное число членов последовательности было отлично от 0. Если отказаться от этого требования, то получится тоже очень важный в алгебре и анализе объект – кольцо формальных степенных рядов.

Определение. *Формальным степенным рядом* над коммутативным ассоциативным кольцом R с единицей называется любая бесконечная последовательность $(*)$. Обычно формальный степенной ряд $(*)$ записывается в виде бесконечной суммы

$$(\|) \qquad \sum_{k=0}^{\infty} a_k t^k = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n + \cdots .$$

Отметим, что бесконечные суммы не определены в кольцах. Так, что $(\|)$ – формальная запись бесконечной последовательности. Множество всех формальных степенных рядов обозначается через $R[[t]]$. На нем вводятся операции сложения (покомпонентно) и умножения (при помощи формул $(\ddagger)$). Как и в случае многочленов проверяется, что $R[[t]]$ – коммутативное ассоциативное кольцо с единицей.

Замечание. В случае, когда R – поле, кольца $R[t]$ и $R[[t]]$ являются бесконечномерными алгебрами над R.

Определение. *Формальным рядом Лорана* над коммутативным ассоциативным кольцом R с единицей называется любая бесконечная (вправо) последовательность

$$\sum_{k=-n}^{\infty} a_k t^k = a_{-n} t^{-n} + \cdots + a_{-1} t^{-1} + a_0 + a_1 t + a_2 t^2 + \cdots .$$

Отметим, что в этой формуле число n членов отрицательной степени ограничено, но для каждого ряда оно свое. Множество всех формальных степенных рядов над R обозначается через $R((t))$. Как и выше, на $R((t))$ вводятся операции сложения и умножения так, что $R((t))$ является коммутативным ассоциативным кольцом с единицей. Если же R – поле, то $R((t))$ также является полем.

Подстановка элементов кольца в многочлен.

Зафиксируем $c \in R$. Для многочлена

$$f = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n \in R[t]$$

положим

$$f(c) := a_0 + a_1 c + a_2 c^2 + \cdots + a_n c^n \in R[t].$$

Элемент $f(c)$ называется значением многочлена $f \in R[t]$ при $t = c$.

Таким образом, каждому $f \in R[t]$ сопоставляется функция

$$\bar{f} : R \to R, \qquad c \mapsto f(c).$$

Отметим, что в общем случае это сопоставление $f \longmapsto \bar{f}$ не является ни инъективным, ни сюръективным.

Пример. Предположим, что в кольце R конечное число элементов: $R = \{a_1, a_2, \ldots, a_n\}$. Рассмотрим многочлен

$$f := \prod_{i=1}^{n}(t - a_i).$$

Тогда f – ненулевой многочлен, но соответствующая функция $\bar{f} : R \to R$ тождественно равна нулю.

Говорят, что $c \in R$ – *корень* многочлена f, если $f(c) = 0$.

Замечание. Отметим также, что отображение

$$L_a : R[t] \longrightarrow R, \qquad f \longmapsto f(a)$$

является гомоморфизмом колец.

Лекция 13

Степень многочлена. Делители нуля в кольце многочленов. Деление многочленов над полем с остатком. Схема Горнера. Теорема Безу. Корни многочленов. Кратность корня. Функциональное равенство многочленов. Пример для конечных полей. Интерполяционная формула Лагранжа. Делимость в кольцах. Евклидовы кольца. Наибольший общий делитель. Алгоритм Евклида.

Степень многочлена

Напомним, что *степенью ненулевого многочлена* $f = \sum a_k t^k$ называется максимальное k такое, что $a_k \neq 0$.

Предложение. *Пусть* $f, g \in R[t]$ *– ненулевые многочлены. Тогда*

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

(2) *если* $\deg(f) \neq \deg(g)$, *то* $\deg(f+g) = \max\{\deg(f), \deg(g)\}$,

(3) $\deg(f \cdot g) \leq \deg(f) + \deg(g)$,

(4) *если кольцо* R *не имеет делителей нуля, то* $\deg(f \cdot g) = \deg(f) + \deg(g)$.

Доказательство. Докажем, например, (3) и (4). Запишем

$$f = \sum a_i t^i, \qquad g = \sum b_j t^j, \qquad a_n,\, b_m \neq 0.$$

Тогда

$$f \cdot g = \sum c_k t^k, \qquad c_k = \sum_{i+j=k} a_i b_j.$$

Пусть $\deg(f) = n$, $\deg(g) = m$. Тогда $a_i = 0$ при $i > n$ и $a_n \neq 0$. Аналогично, $b_j = 0$ при $j > m$ и $b_m \neq 0$. Поскольку $a_i = 0$ при $i < 0$ и $b_j = 0$ при $j < 0$, то $c_k = 0$ при $k > n+m$ и $c_{n+m} = a_n b_m$. Отсюда следует, что $\deg(f \cdot g) \leq n+m$. Более того, если a_n и b_m не являются делителями нуля, то $c_{n+m} = a_n b_m \neq 0$ и тогда $\deg(f \cdot g) = n+m$. □

Следствие. (1) *Если кольцо* R *не имеет делителей нуля, то и* $R[t]$ *не имеет делителей нуля.*

(2) *Пусть R не имеет делителей нуля. Если элемент $f \in R[t]$ обратим, то* $\deg(f) = 0$, $f \neq 0$.

(3) *Пусть $\Bbbk$ – поле. Тогда $f \in \Bbbk[t]$ обратим тогда и только тогда, когда* $\deg(f) = 0$, $f \neq 0$.

Деление многочленов над полем с остатком

Далее мы рассмотрим алгебру многочленов $\Bbbk[t]$, где $\Bbbk$ – поле.

Теорема (деление с остатком)**.** *Пусть $\Bbbk$ – поле. Для любых многочленов f, $g \in \Bbbk[t]$, $g \neq 0$ существуют единственные многочлены q, $r \in \Bbbk[t]$ такие, что*

$$(*) \qquad f = gq + r, \quad \text{где или } r = 0, \quad \text{или } \deg(r) < \deg(g).$$

Доказательство. Единственность. Предположим, что $f = gq + r = gq' + r'$. Тогда $g(q - q') = r - r'$. Отсюда $r = r'$ и $q = q'$.

Существование. Рассмотрим все представления многочлена f в виде $f = gq + r$, т.е. рассмотрим множество M всевозможных пар (q, r) многочленов $q, r \in \Bbbk[t]$, удовлетворяющих этому равенству. Это множество M непусто: $(0, f) \in M$.

Если в M существует элемент вида $(q, 0)$, то представление $(*)$, очевидно имеет место. Предположим, что M не содержит элементов вида $(q, 0)$.

Выберем элемент $(q, r) \in M$, у которого r имеет наименьшую степень. Предположим, что

$$k := \deg(r) \geq \deg(g) =: m.$$

Запишем

$$r = c_k t^k + \cdots + c_0, \qquad g = b_m t^m + \cdots + b_0$$

Положим,

$$r_1 := r - \frac{c_k}{b_m} t^{k-m} \cdot g, \qquad q_1 := q + \frac{c_k}{b_m} t^{k-m}.$$

Тогда коэффициент при t в r_1 равен нулю и поэтому $\deg(r_1) < \deg(r)$. С другой стороны,

$$f = g\left(q + \frac{c_k}{b_m} t^{k-m}\right) + r_1 = g \cdot q_1 + r_1.$$

Таким образом, $(q_1, r_1) \in M$. Противоречие показывает, что $k < m$. □

Практически, деление с остатком выполняется – столбиком, примерно также, как и деление многозначных чисел:

$$\begin{array}{rl|l}
x^3 - 2x^2 - 5x + 5 & & x - 1 \\
\cline{3-3}
-\,x^3 \;+ x^2 & & x^2 - x - 6 \\
\cline{1-1}
\quad - x^2 - 5x & & \\
\quad\; x^2 \;- x & & \\
\cline{1-1}
\quad\quad - 6x + 5 & & \\
\quad\quad\; 6x - 6 & & \\
\cline{1-1}
\quad\quad\quad - 1 & &
\end{array}$$

Очень важен частный случай теоремы о делении с остатком.

Следствие (схема Горнера)**.** *Для ненулевого многочлена $f \in \Bbbk[t]$ и элемента $\alpha \in \Bbbk$ имеет место представление*

$$(\dagger) \qquad f = (t-\alpha)q + c, \qquad \text{где } q \in \Bbbk[t] \text{ и } c = f(\alpha).$$

Практическая реализация алгоритма нахождения многочлена q и константы c называется *схемой Горнера*. Пусть

$$f = \sum_{i=1}^{n} a_i t^i.$$

Запишем многочлен q с неопределенными коэффициентами:

$$q = \sum_{k=0}^{n-1} b_k t^k.$$

Приравнивая коэффициенты в $f = (t-\alpha)q + c$ при одинаковых степенях, получим

$$\begin{array}{lcl}
a_n = b_{n-1} & \Longrightarrow & b_{n-1} = a_n \\
a_{n-1} = b_{n-2} - \alpha b_{n-1} & \Longrightarrow & b_{n-2} = a_{n-1} + \alpha b_{n-1} \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & \\
a_k = b_{k-1} - \alpha b_k & \Longrightarrow & b_{k-1} = a_k + \alpha b_k \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & \\
a_1 = b_0 - \alpha b_1 & \Longrightarrow & b_0 = a_1 + \alpha b_1 \\
a_0 = c - \alpha b_0 & \Longrightarrow & c = a_0 + \alpha b_0
\end{array}$$

Таким образом, коэффициенты $b_{n-1}, \dots, b_0$ и константа c получаются из таблицы

	a_n	a_{n-1}	a_{n-2}	$\dots$	a_1	a_0
α	b_{n-1}	b_{n-2}	b_{n-3}	$\dots$	b_0	c

которая заполняется по правилам $b_{n-1} = a_n$ и $b_k = a_{k+1} + \alpha b_{k+1}$ при $k < n-1$.

Следствие. *Для многочлена $f \in \Bbbk[t]$ и элемента $\alpha \in \Bbbk$ имеет место представление $f = c_n(t-\alpha)^n + \dots + c_1(t-\alpha) + c_0$, где $c_i \in \Bbbk$.*

Доказательство. Последовательно применим (†). □

Следствие (Теорема Безу)**.** *Элемент $\alpha \in \Bbbk$ является корнем ненулевого многочлена $f \in \Bbbk[t]$ тогда и только тогда, когда $f = (t-\alpha)g$ для некоторого $g \in \Bbbk[t]$.*

Следствие. *Для ненулевого многочлена $f \in \Bbbk[t]$ имеет место единственное представление*

$$(\ddagger) \qquad f = (t-\alpha_1)^{m_1} \cdots (t-\alpha_r)^{m_r} f_0,$$

где $\alpha_1, \dots, \alpha_r$ – все корни многочлена, а многочлен f_0 не имеет корней. В частности, многочлен степени n имеет не более n корней.

Доказательство. Последовательно применяя теорему Безу получим нужное представление. Для доказательства единственности предположим, что для некоторого многочлена $f \in \Bbbk[t]$ существуют два представления

$$f = (t-\alpha_1)^{m_1} \cdots (t-\alpha_r)^{m_r} f_0 = (t-\beta_1)^{m'_1} \cdots (t-\beta_l)^{m'_l} g_0.$$

Возьмем такой многочлен минимальной степени. Подставим β_1 в f:

$$f(\beta_1) = (\beta_1)-\alpha_1)^{m_1} \cdots (\beta_1)-\alpha_r)^{m_r} f_0(\beta_1) = 0.$$

Так как поле $\Bbbk$ не имеет делителей нуля, то один из множителей обращается в нуль. Так как f_0 не имеет корней, то $\beta_1 = \alpha_i$ для некоторого i. Поскольку $\Bbbk[t]$ не имеет делителей нуля, мы можем сократить на $t-\beta_1$ и получить аналогичное соотношение меньшей степени. Это противоречит нашему предположению. □

Кратность корня.

Определение. Пусть $f \in \Bbbk[t]$ – ненулевой многочлен. Кратностью его корня $\alpha \in \Bbbk$ называется m такое, что $(t-\alpha)^m$ делит f, а $(t-\alpha)^{m+1}$ не делит.

Иначе говоря, $\alpha \in \Bbbk$ – корень кратности m, если

$$f(t) = (t-\alpha)^m \cdot f_1(t),$$

где $f_1(\alpha) \neq 0$.

Замечание. В формуле (‡) числа $m_1, \ldots, m_r$ – это кратности корней $\alpha_1, \ldots, \alpha_r$, соответственно.

Следствие. *Ненулевой многочлен степени n имеет не более n корней, посчитанных с учетом кратностей.*

Многочлены и функции.

Теорема. *Пусть $f, g \in \Bbbk[t]$, $f \neq g$. Следующие условия эквивалентны:*

(1) $f(\alpha) = g(\alpha)$ *для любого* $\alpha \in \Bbbk$;

(2) *поле $\Bbbk$ конечно и все элементы поля $\Bbbk$ являются корнями многочлена $h := f - g$.*

Доказательство. Импликация (1) $\Longrightarrow$ (2) очевидна, поскольку ненулевой многочлен $h = f - g$ может иметь только конечное число корней.

Если $h(a) = 0$ для любого $a \in \Bbbk$, то $f(a) = g(a) + h(a) = g(a)$. □

Теорема (Интерполяционная формула Лагранжа). *Пусть $\alpha_1, \ldots, \alpha_{n+1}$, $\beta_1, \ldots, \beta_{n+1} \in \Bbbk$. Тогда существует единственный многочлен $f \in \Bbbk$ такой, что* $\deg(f) \leq n$ *и* $f(\alpha_i) = \beta_i$.

Доказательство. Существование. Положим

$$f = \sum_{i=1}^{n+1} \frac{(t-\alpha_1)\cdots\widehat{(t-\alpha_i)}\cdots(t-\alpha_{n+1})}{(\alpha_i-\alpha_1)\cdots\widehat{(\alpha_i-\alpha_i)}\cdots(\alpha_i-\alpha_{n+1})}\beta_i.$$

Несложно видеть, что $\deg(f) \le n$ и $f(\alpha_i) = \beta_i$.

Единственность. Предположим, что имеется два многочлена f и g таких, что $f(\alpha_i) = g(\alpha_i) = \beta_i$. Тогда элементы $\alpha_1, \ldots, \alpha_{n+1}$ являются корнями многочлена $f-g$ степени $\le n$. Противоречие. □

Делимость в кольцах

Определение. Ассоциативное коммутативное кольцо с единицей без делителей нуля называется *целостным* кольцом.

Пусть R – целостное кольцо. Говорят, что $f \in R$ *делится* на элемент $g \in R$, $g \neq 0$, если $f = gh$ для некоторого $h \in R$. В этом случае мы будем писать $g \mid f$ (читается "g делит f").

Следующие свойства легко выводятся из определения. Читателю предлагается это проделать самостоятельно.

Предложение. (1) *Если* $f \mid g_1$ *и* $f \mid g_2$, *то* $f \mid (g_1 \pm g_2)$;

(2) *Если* $f \mid g$, *то* $f \mid gh$, *для любого* $h \in R$;

(3) *если* $f \mid g$ *и* $g \mid h$, *то* $f \mid h$;

(4) $f \mid g$ *и* $g \mid f$ *тогда и только тогда, когда* $f = gu$, *где* $u \in R$ *– обратимый элемент.*

Определение. Пусть R – целостное кольцо. Для f, $g \in R$ таких, что f, $g \neq 0$ определим *наибольший общий делитель* $\text{НОД}(f, g) = h$ следующим образом

- $h \mid f$ и $h \mid g$;
- если $s \mid f$ и $s \mid g$, то $s \mid h$.

Отметим, однако, что наибольший общий делитель существует не всегда.

Замечание. Определение наибольшего общего делителя, данное выше, является наиболее правильным. Оно применимо к любым целостным кольцам. В случае кольца $\Bbbk[t]$ наибольший общий делитель многочленов $0 \neq f, g \in \Bbbk[t]$ совпадает с многочленом наибольшей степени, делящим f и g. Однако, такое определение не работает в других целостных кольцах.

Определение. Пусть R – целостное кольцо. Говорят, что оно *евклидово*, если существует функция $\delta : R \setminus \{0\} \to \mathbb{N} \cup \{0\}$ такая, что

(1) $a \mid b$, то $\delta(a) \le \delta(b)$;

(2) Для любых a, $b \in R$, $b \neq 0$ существует представление $a = bq + r$, где $r = 0$ или $\delta(r) < \delta(b)$.

Замечание. Из определения следует, что если $a \mid b$ и элемент b/a необратим, то $\delta(a) < \delta(b)$. Действительно, предположим, что $\delta(a) = \delta(b)$. Поделим a на b с остатком: $a = bq+r$, где $\delta(r) < \delta(b)$. Здесь $r \neq 0$ поскольку $c := b/a$ необратим. Тогда

$$\delta(b) = \delta(a) \leq \delta(a(1 - qc)) = \delta(a - acq) = \delta(a - bq) = \delta(r) < \delta(b).$$

Противоречие.

Пример. (1) Кольцо целых чисел $\mathbb{Z}$ евклидово, $\delta(a) = |a|$.

(2) Кольцо многочленов над полем $\Bbbk[t]$ евклидово, $\delta(f) = \deg(f)$.

Пример. Кольцо целых гуссовых чисел

$$R = \{a + b\boldsymbol{i} \mid a, b \in \mathbb{Z}\} \subset \mathbb{C}$$

является евклидовым, можно взять $\delta(a + b\boldsymbol{i}) = a^2 + b^2$.

Наибольший общий делитель. Алгоритм Евклида

Теорема. *Пусть $\Bbbk$ – поле и f, $g \in \Bbbk[t]$ – ненулевые многочлены. Тогда наибольший общий делитель f и g существует и единственен с точностью до пропорциональности. Более того, наибольший общий делитель $h = \text{НОД}(f, g)$ может быть представлен в виде*

$$h = fu + gv, \qquad \text{где } u, v \in \Bbbk[t]. \tag{§}$$

Замечание. Разложение (§) не единственно: $h = fu + gv = fu_1 + gv_1$ тогда и только тогда, когда $\frac{f}{h}(u_1-u)+\frac{g}{h}(v_1-v) = 0$ тогда и только тогда, когда существует r $u_1-u = r\frac{g}{h}$, $v_1 - v = -r\frac{f}{h}$.

Доказательство. Последовательно применим деление с остатком:

$$\begin{aligned}
f &= gq_1 + r_1 \\
g &= r_1q_2 + r_2 \\
r_1 &= r_2q_3 + r_3 \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
r_k &= r_{k+1}q_{k+2} + r_{k+2} \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
r_{n-1} &= r_nq_{n+1} + r_{n+1} \\
r_n &= r_{n+1}q_{n+2}.
\end{aligned}$$

Положим $h = r_{n+1}$ (последний ненулевой остаток). Убывающей индукцией по k доказываем, что h делит r_k для всех k. Действительно, если h делит r_l для $l > k$, то h делит $r_k = r_{k+1}q_{k+2} + r_{k+2}$. Отсюда получаем, что h делит $g = r_1q_2 + r_2$ и $f = gq_1 + r_1$.

Для доказательства (§) доказываем, что $r_k = fu_{k-1} + gu_{k-2}$ индукцией по k. □

Замечание. Многочлены u и v в (§) предыдущей теоремы можно выбрать так, что $\deg(u) < \deg(g)$, $\deg(v) < \deg(f)$.

Доказательство. Запишем $u = gu_1 + u_0$, $v = fv_1 + v_0$. Тогда $h = fu_0 + gv_0 + fg(u_1 + v_1)$. Если $u_1 + v_1 \neq 0$, то $\deg(fu_0 + gv_0 + fg(u_1 + v_1)) > \deg(f) + \deg(g)$. □

Вычисление наибольшего общего делителя примененное в доказательстве теоремы называется алгоритмом Евклида. Этот же метод работает в любом евклидовом кольце.

Теорема. *Пусть f, $g \in R$ – ненулевые элементы евклидова кольца R. Тогда наибольший общий делитель f и g существует и единственен с точностью до умножения на обратимый элемент. Более того, наибольший общий делитель $h = \text{НОД}(f, g)$ может быть представлен в виде*

$$h = fu + gv, \qquad \text{где } u, v \in R.$$

Лекция 14

Неприводимые многочлены. Факториальность кольца многочленов над полем. Факториальные кольца. Дифференцирования. Дифференцирования кольца многочленов над полем. Понижение кратности при дифференцировании. Формула Тейлора.

Неприводимые многочлены

Определение. Пусть $\Bbbk$ – поле. Говорят, что многочлен $f \in \Bbbk[t]$ положительной степени *неприводим*, если для любого разложения $f = f_1 f_2$, $f_i \in \Bbbk[t]$ имеем $\deg(f_1) = 0$ или $\deg(f_2) = 0$.

Замечание. Понятие неприводимости зависит от выбора поля: может случиться так, что многочлен f неприводим как элемент кольца $\Bbbk[t]$, но приводим как элемент кольца $\mathbb{K}[t]$, где $\mathbb{K}$ – большее поле, содержащее $\Bbbk$.

Примеры.

- Любой многочлен первой степени неприводим.
- Неприводимый многочлен f имеет корни тогда и только тогда, когда $\deg(f) = 1$.
- Многочлен, не имеющий корней может быть приводимым. Например, степень f^m неприводимого многочлена степени $\deg(f) \geq 2$ не имеет корней. Однако, f^m приводим при $m \geq 2$.
- По теореме Безу многочлен $f \in \Bbbk[t]$ степени 2 или 3 неприводим тогда и только тогда, когда он не имеет корней. Например, любой многочлен второй степени над $\mathbb{R}$ с отрицательным дискриминантом неприводим. Любой многочлен третьей степени над $\mathbb{R}$ приводим, так как имеет корень.

Факториальность кольца многочленов над полем

Теорема (факториальность кольца многочленов над полем)**.** *Пусть $f \in \Bbbk[t]$ – ненулевой многочлен положительной степени. Тогда имеет место разложение:*

$$f = f_1 \cdots f_m,$$

где все f_i – неприводимые многочлены. Это разложение единственно с точностью до порядка множителей и пропорциональности, т.е. если имеется два разложения

$$f = f_1 f_2 \dots f_m = g_1 g_2 \dots g_n$$

в произведение простых множителей, то $m = n$ и переставляя множители мы можем добиться того, что $f_i = g_i c_i$ для всех i, где $c_i \in \Bbbk^*$.

Лемма. *Пусть $f \in \Bbbk[t]$ неприводим и $f \mid gh$. Тогда $f \mid g$ или $f \mid h$.*

Доказательство. Пусть $f \nmid g$. Тогда НОД$(f, g) = 1$ и поэтому $fu+gv = 1$ для некоторых многочленов $u, v \in \Bbbk[t]$. Отсюда $f hu + (gh)v = h$. Следовательно, $f \mid h$. □

Доказательство. Существование. Индукция по степени. Действительно, для многочлена степени 1 разложение состоит из одного элемента. Предположим, что утверждение верно для всех многочленов степени $< d$. Пусть $\deg(f) = d$. Если f неприводим, то все доказано: опять разложение состоит из одного элемента. Если же f приводим, то $f = gh$, где $0 < \deg(g) < d$ и $0 < \deg(h) < d$. По предположению индукции многочлены g и h раскладываются в произведение неприводимых: $g = g_1 \cdots g_k$ и $h = h_1 \cdots h_l$. Тогда $f = g_1 \cdots g_k h_1 \cdots h_l$.

Единственность. Индукция по степени. Для многочлена степени 1 утверждение верно. Предположим, что утверждение верно для всех многочленов степени $< d$. Пусть $\deg(f) = d$. Предположим, что имеется два разложения $f = f_1 \cdots f_m = g_1 \cdots g_n$ в произведения неприводимых. Тогда неприводимый многочлен f_1 делит $g_1 \cdots g_n$. Значит он делит один из сомножителей, скажем g_1. Но многочлен g_1 также неприводим. Поэтому $f_1 = g_1 c_1$ для некоторого $c_1 \in \Bbbk^*$. Сокращая на f_1 получим $f_2 \cdots f_m = c_1 g_2 \cdots g_n$. По предположению индукции $m = n$ и $f_i = g_i c_i$ для всех i. □

Разложение на множители в целостных кольцах

Пусть R – целостное кольцо. Ненулевой необратимый элемент $f \in R$ называется *простым*, если для любого разложения $f = f_1 f_2$ в произведение элементов $f_i \in R$ один из этих элементов является обратимым.

Примеры.

- В поле нет простых элементов, поскольку в нем все ненулевые элементы обратимы.
- В кольце целых чисел $\mathbb{Z}$ простые элементы – это простые числа.
- В кольце $\Bbbk[t]$ многочленов над полем простые элементы – это неприводимые многочлены.

Кольцо R называется *факториальным*, если любой ненулевой необратимый элемент $f \in R$ может быть разложен в произведение простых множителей

$$f = p_1 p_2 \dots p_m$$

и это разложение единственно с точностью до порядка множителей и умножения множителей на обратимые элементы, т.е. если имеется два разложения

$$f = p_1 p_2 \dots p_m = p'_1 p'_2 \dots p'_n$$

в произведение простых множителей, то $m = n$ и переставляя множители мы можем добиться того, что $p_i = p'_i u_i$ для всех i, где u_i – обратимые элементы.

Пример.
- Из курса школьной арифметики мы знаем, что кольцо целых чисел $\mathbb{Z}$ факториально.
- Согласно последней теореме, кольцо многочленов над полем также факториально.

Пример. Кольцо E сходящихся степенных рядов на всей комплексной плоскости (кольцо целых функций) не является факториальным. В этом случае имеются функции, имеющие бесконечное множество простых делителей. Например, $\sin(z)$ представляется сходящимся степенным рядом, т.е. $\sin(z) \in E$. С другой стороны, имеет место формула

$$\sin(\pi z) = \pi z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right).$$

Пример. Кольцо чисел Эйлера

$$R = \{a + b\sqrt{-3} \mid a,\, b \in \mathbb{Z}\} \subset \mathbb{C}$$

не является факториальным: $4 = 2 \cdot 2 = (1 + \sqrt{-3})(1 - \sqrt{-3})$.

Факториальные и евклидовы кольца

Анализируя доказательство теоремы, можно заметить, что для того, чтобы целостное кольцо R было факториальным необходимо и достаточно выполнение следующих двух условий (сравните с последними двумя примерами).

(1) Процесс разложения элемента в произведения необратимых должен обрываться.

(2) Должна быть верна лемма о том, что если простой элемент $p \in R$ делит произведение, то он делит один из сомножителей.

Верна следующая теорема.

Теорема. *Если R – евклидово кольцо, то оно факториально.*

Доказательство следует той же схеме, что и доказательство факториальности кольца многочленов. Оба свойства (1) и (2) проверяются индукцией по $\delta(a)$, где δ – функция из определения евклидова кольца.

Пример. Кольцо целых гуссовых чисел является евклидовым, а следовательно, и факториальным.

Дифференцирования

Пусть $\Bbbk$ – поле и пусть A – коммутативная ассоциативная алгебра над $\Bbbk$. Отображение $\partial : A \to A$ называется *дифференцированием* если

- ∂ является $\Bbbk$-линейным;

- $\partial(ab) = a\partial(b) + \partial(a)b$.

Лемма. *Пусть $\partial : A \to A$ – дифференцирование. Тогда*

(1) *если $1 \in A$ – единица, то $\partial(1) = 0$;*

(2) $\partial(a^n) = na^{n-1}\partial(a)$.

Доказательство. (1) Имеем $\partial(1) = \partial(1\cdot 1) = 1\partial(1)+\partial(1)1 = \partial(1)+\partial(1)$. Отсюда $\partial(1) = 0$.
(2) Докажем по индукции:

$$\partial(a^{n+1}) = \partial(a^n a) = a^n\partial(a) + \partial(a^n)a = a^n\partial(a) + na^{n-1}\partial(a)a = (n+1)a^{n+1}\partial(a). \qquad \square$$

Заметим, что для любого дифференцирования $\partial : A \to A$ и любого элемента $\lambda \in \Bbbk$ отображение

$$\partial_\alpha : A \longrightarrow A, \qquad \partial_\lambda(a) = \lambda\partial(a)$$

также является дифференцированием.

Теорема. *Пусть $\Bbbk$ – поле. Существует единственное дифференцирование $\partial : \Bbbk[t] \to \Bbbk[t]$ такое, что $\partial(t) = 1$.*

Доказательство. Единственность. По лемме $\partial(t^k) = kt^{k-1}\partial(t) = kt^{k-1}$. Пусть $f = \sum_k a_k t^k$. Из линейности выводим

$$\partial(f) = \sum_k \partial(a_k t^k) = \sum_k k a_k t^{k-1}.$$

Существование. Определим ∂ формулой выше. Проверка линейности отображения ∂ оставляется читателю. Проверим равенство

$$(*) \qquad \partial(fg) = f\partial(g) + g\partial(f).$$

Во-первых, проверим эту формулу для одночленов. Пусть $f = at^n$ и $g = bt^m$. Тогда

$$\partial(fg) = \partial(abt^{n+m}) = ab(n+m)t^{n+m-1} = abnt^n t^{m-1} + abmt^m t^{n-1} = f\partial(g) + g\partial(f).$$

Теперь распишем многочлены в виде суммы одночленов $f = \sum_i f_i$, $g = \sum_j g_j$, $fg = \sum_{i,j} f_i g_j$. Из линейности выводим

$$\begin{aligned} \partial(fg) &= \sum_{i,j} \partial(f_i g_j), \\ f\partial(g) + g\partial(f) &= \left(\sum_i f_i\right)\sum_j \partial(g_j) + \left(\sum_j g_j\right)\sum_i \partial(f_i) = \\ &= \sum_{i,j}\left(f_i \sum_j \partial(g_j) + g_j \sum_i \partial(f_i)\right). \end{aligned}$$

Соответствующие члены в правых частях формул совпадают. Отсюда получаем $(*)$. $\square$

Следствие. *Любое дифференцирование кольца многочленов* $\Bbbk[t]$ *имеет вид*

$$\partial\left(\sum_k a_k t^k\right) = \lambda \sum_k k a_k t^{k-1}.$$

для некоторого $\lambda \in \Bbbk$.

Далее мы будем рассматривать дифференцирование $\partial : \Bbbk[t] \to \Bbbk[t]$, удовлетворяющее условию $\partial(t) = 1$. В этом случае обозначим $f' = \partial(f)$. Многочлен f' называется *производной* многочлена f.

Понижение кратности при дифференцировании.

Теорема. *Пусть* $\Bbbk$ *– поле. Пусть* $f \in \Bbbk[t]$ *– многочлен над* $\Bbbk$ *и пусть* g *– неприводимый множитель* f *кратности* m. *Предположим, что выполнено одно из следующих*

(1) $\mathrm{char}(\Bbbk) = 0$ *или*

(2) $\mathrm{char}(\Bbbk) = p > 0$, $p \nmid m$ *и* $p > \deg(g)$.

Тогда g *является множителем кратности* $m-1$ *для производной* f'.

Доказательство. Запишем $f = g^m h$, где $g \nmid h$. Тогда

$$f' = m g^{m-1} g' h + g^m h' = g^{m-1}(m g' h + g h').$$

Отсюда видно, что g^{m-1} делит f'. Предположим, что g^m делит f'. Тогда g делит $mg'h$. Следовательно, g делит mg'. Поскольку $\mathrm{char}(\Bbbk)$ равна 0 или не делит m, то $m \neq 0$ в $\Bbbk$ и поэтому g делит g'. Поскольку $\deg(g') < \deg(g)$, мы получаем, что $g' = 0$. Запишем $g = \sum b_k t^k$. Тогда $0 = g' = \sum k b_k t^{k-1}$ т.е. $k b_k = 0$ для любого k. Это возможно только если $\mathrm{char}(\Bbbk) = p > 0$ и $p \mid k$ как только $b_k \neq 0$. □

Следствие. *Пусть* $\Bbbk$ *– поле. Пусть* $f \in \Bbbk[t]$ *и пусть* α *– корень многочлена* f *кратности* m. *Предположим, что выполнено одно из следующих*

(1) $\mathrm{char}(\Bbbk) = 0$ *или*

(2) $\mathrm{char}(\Bbbk) = p > 0$, $p \nmid m$.

Тогда α *является корнем кратности* $m-1$ *для производной* f'.

Заметим, что дополнительные условия в случае положительной характеристики необходимы, что показывает следующие примеры.

Пример. Пусть $\Bbbk$ – поле характеристики $\mathrm{char}\,\Bbbk = p > 0$. Рассмотрим многочлен $f \in \Bbbk[t]$ вида

$$f(t) = (t^p - a)^m h(t),$$

где $a \in \Bbbk$, а $h(t)$ – многочлен положительной степени. Предположим, что множитель $t^p - a$ неприводим и не делит $h(t)$. Тогда

$$f'(t) = m(t^p - a)'h(t) + (t^p - a)^m h'(t) = mpt^{p-1}h(t) + (t^p - a)^m h'(t) = (t^p - a)^m h'(t).$$

Таким образом, $t^p - a$ – множитель кратности m, как для самого многочлена f, так и для его производной f'.

Пример. Пусть $\Bbbk$ – поле характеристики $\operatorname{char}\Bbbk = p > 0$. Рассмотрим многочлен $f \in \Bbbk[t]$ вида

$$f(t) = (t - \alpha)^p h(t),$$

где $\alpha \in \Bbbk$, а $h(t)$ – многочлен положительной степени такой, что $h(\alpha) \neq 0$. Тогда

$$f'(t) = p(t - \alpha)^{p-1}h(t) + (t - \alpha)^p h'(t) = (t - \alpha)^p h'(t).$$

Таким образом, α – корень кратности p, как для самого многочлена f, так и для его производной f'.

Следствие. *Пусть $\Bbbk$ – любое поле. Кратные корни многочлена $f \in \Bbbk[t]$ – это в точности общие корни f и f'.*

Доказательство. Запишем $f = (t - \alpha)^m h$, где $h(\alpha) \neq 0$. Тогда

$$f' = m(t - \alpha)^{m-1}h + (t - \alpha)^m h'.$$

Если α – кратный корень f, то $m > 1$ и $f'(\alpha) = 0$. Обратно, если $f(\alpha) = f'(\alpha) = 0$, то $m \geq 1$ и $m(t - \alpha)^{m-1}h(\alpha) = 0 \Longrightarrow m(t - \alpha)^{m-1} = 0$. Следовательно, или $m = 0$ или $(t - \alpha)^{m-1} = 0$. В обоих случаях $m > 1$. □

Отделение кратных множителей. Пусть $f \in \Bbbk[t]$, $\operatorname{char}\Bbbk = 0$ или $\operatorname{char}\Bbbk > \deg(f)$. Вычисляем f'. По алгоритму Евклида вычисляем $h := (f, f')$. Тогда многочлен f/h имеет те же неприводимые множители, что и f, но все – с кратностью 1. В частности, многочлен f/h имеет те же корни, что и f, но все – простые.

Формула Тейлора

Теорема. *Пусть $\Bbbk$ – поле и пусть $f \in \Bbbk[t]$, $\deg(f) = n$. Предположим, что* $\operatorname{char}\Bbbk = 0$ *или* $\operatorname{char}\Bbbk > n$. *Тогда f единственным образом представляется в виде*

$$(\dagger) \qquad f = b_0 + b_1(t - \alpha) + b_2(t - \alpha)^2 + \cdots + b_n(t - \alpha)^n,$$

где

$$(\ddagger) \qquad b_k = f^{(k)}(\alpha)/k!.$$

Доказательство. Существование разложения (†) доказывается индукцией по степени. База индукции очевидна. Предположим, что разложение (†) существует для всех многочленов степени $< n$. Имеем $f(t) = (t-\alpha)g(t) + b_0$, где $b_0 = f(\alpha)$. По предположению индукции

$$g = c_0 + c_1(t-\alpha) + c_2(t-\alpha)^2 + \cdots + c_{n-1}(t-\alpha)^{n-1}.$$

Тогда

$$\begin{aligned} f &= (t-\alpha) + b_0 = \big(c_0 + c_1(t-\alpha) + c_2(t-\alpha)^2 + \cdots + c_{n-1}(t-\alpha)^{n-1}\big)(t-\alpha) + b_0 \\ &= b_0 + c_0(t-\alpha) + c_1(t-\alpha)^2 + c_2(t-\alpha)^3 + \cdots + c_{n-1}(t-\alpha)^n. \end{aligned}$$

Для доказательства (‡) запишем

$$((t-\alpha)^l)^{(k)} = \begin{cases} l(l-1)\cdots(l-k+1)(t-\alpha)^{l-k} & \text{при } k \le l \\ 0 & \text{иначе.} \end{cases}$$

Это доказывается индукцией по k. Таким образом,

$$f^{(k)} = b_k k! + (t-\alpha)\cdot(\text{многочлен}).$$

Получаем $f^{(k)}(\alpha) = b_k k!$. □

Практически, разложение многочлена в формулу Тейлора (†) лучше всего находить при помощи схемы Горнера.

Лекция 15

Основная теорема алгебры. Сходимость последовательностей комплексных чисел. Лемма о возрастании модуля многочлена. Лемма Даламбера. Основная теорема алгебры (доказательство). Следствия. Неприводимые многочлены над $\mathbb{C}$ и $\mathbb{R}$. Поле частных целостного кольца.

Основная теорема алгебры комплексных чисел

Определение. Поле $\Bbbk$ называется *алгебраически замкнутым*, если любой многочлен $f \in \Bbbk[t]$ положительной степени имеет по крайней мере один корень в $\Bbbk$.

Лекция посвящена доказательству следующей теоремы, которая называется *основной теоремой алгебры комплексных чисел*.

Теорема. *Поле $\mathbb{C}$ алгебраически замкнуто.*

Сходимость последовательностей комплексных чисел

Напомним, что для любых комплексных чисел z_1, $z_2 \in \mathbb{C}$ выполнены следующие неравенства (неравенства треугольника):

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

Определение. *Пределом* последовательности комплексных чисел $\{z_n\}$ называется комплексной число z_0 такое, что $\lim_{n\to\infty} |z_n - z_0| = 0$. Если предел существует, то последовательность называется *сходящейся*.

Предел последовательности $\{z_n\}$ обозначается через $\lim_{n\to\infty} z_n$ или просто $\lim z_n$. Также, иногда пишут $z_n \longrightarrow z_0$.

Лемма. *Последовательность $\{z_n\}$ комплексных чисел сходится тогда и только тогда, когда сходятся последовательности действительных чисел $\{\mathrm{Re}(z_n)\}$ и $\{\mathrm{Im}(z_n)\}$. Более того, в этом случае*

$$\mathrm{Re}(\lim z_n) = \lim \big(\mathrm{Re}(z_n)\big), \qquad \mathrm{Im}(\lim z_n) = \lim \big(\mathrm{Im}(z_n)\big).$$

Доказательство. Запишем $z_n = x_n + iy_n$, $z_0 = x_0 + iy_0$. Тогда утверждение следует из неравенства

$$|z_n - z_0|^2 = (x_n - x_0)^2 + (y_n - y_0)^2,$$

где

$$|x_n - x_0| \le |z_n - z_0|, \quad |y_n - y_0| \le |z_n - z_0|. \qquad \square$$

Лемма. *Если последовательность комплексных чисел $\{z_n\}$ сходится, то сходится и последовательность действительных чисел $\{|z_n|\}$. Более того,*

$$\lim |z_n| = |\lim z_n|.$$

Доказательство. Пусть $\lim z_n = z_0$. Тогда утверждение следует из неравенства $||z_n| - |z_0|| \le |z_n - z_0|$. $\square$

Следствие. *Если последовательность $\{z_n\}$ сходится, то она ограничена, т.е. существует константа $c \in \mathbb{R}$ такая, что $|z_n| \le c$.*

Лемма. *Если последовательности $\{z_n\}$ и $\{w_n\}$ сходятся, то сходятся и последовательности $\{z_n + w_n\}$ и $\{z_n w_n\}$. Более того,*

$$\lim(z_n + w_n) = (\lim z_n) + (\lim w_n) \qquad \lim(z_n w_n) = (\lim z_n)(\lim w_n).$$

Доказательство. Пусть $\lim z_n = z_0$ и $\lim w_n = w_0$. Имеем

$$|z_n + w_n - (z_0 + w_0)| \le |z_n - z_0| + |w_n - w_0|.$$

Отсюда

$$\begin{aligned} |z_n w_n - z_0 w_0| &= |(z_n - z_0)w_n + (w_n - w_0)z_0| \le \\ &\le |(z_n - z_0)w_n| + |(w_n - w_0)z_0| \le c|z_n - z_0| + |(w_n - w_0)z_0|. \end{aligned} \quad \square$$

Следствие. *Зафиксируем многочлен $f \in \mathbb{C}[z]$. Если последовательность комплексных чисел $\{z_n\}$ сходится, то сходится и последовательность $\{f(z_n)\}$. Более того,*

$$\lim f(z_n) = f(\lim z_n).$$

Лемма о возрастании модуля многочлена

Лемма. *Пусть $f \in \mathbb{C}[z]$ – многочлен положительной степени. Тогда для любого $c > 0$ существует $R \in \mathbb{R}$ такое, что $|f(z)| \ge c$ при $|z| \ge R$.*

Доказательство. Запишем

$$f(z) = a_n z^n + \cdots + a_1 z + a_0, \quad a_n \ne 0, \quad n \ge 1.$$

Пусть $A := \max|a_i|$. Тогда

$$\begin{aligned}
|f(z)| &= |a_n z^n + \cdots + a_1 z + a_0| = \\
&= |z|^n \left| a_n + \frac{a_{n-1}}{z} + \cdots + \frac{a_1}{z^{n-1}} + \frac{a_0}{z^n} \right| \geq \\
&\geq |z|^n \left(|a_n| - \left| \frac{a_{n-1}}{z} + \cdots + \frac{a_1}{z^{n-1}} + \frac{a_0}{z^n} \right| \right) \geq \\
&\geq |z|^n \left(|a_n| - \frac{|a_{n-1}|}{|z|} - \cdots - \frac{|a_1|}{|z|^{n-1}} - \frac{|a_0|}{|z|^n} \right) \\
&\geq |z|^n \left(|a_n| - \frac{A}{|z|} - \cdots - \frac{A}{|z|^{n-1}} - \frac{A}{|z|^n} \right).
\end{aligned}$$

Мы можем считать, что

$$|z| \geq (nA+1)/|a_n| > 1.$$

Тогда $|z|^k > |z|$ и

$$|f(z)| \geq |z|^n \left(|a_n| - \frac{nA}{|z|} \right) = |z|^{n-1} \left(|a_n||z| - nA \right) \geq |z|^{n-1}.$$

Таким образом, неравенство $|f(z)| \geq c$ выполнено при

$$|z| \geq \max((nA+1)/|a_n|, \quad \sqrt[n-1]{c}) := R. \qquad \square$$

Лемма Даламбера

Лемма. *Пусть $f \in \mathbb{C}[z]$ – многочлен положительной степени такой, что $f(z_0) \neq 0$. Тогда для любого $\epsilon > 0$ существует $h \in \mathbb{C}$ такое, что $|h| < \epsilon$ и $|f(z_0+h)| < |f(z_0)|$.*

Доказательство. Запишем

$$f(z) = b_0 + b_k(z-z_0)^k + \cdots + b_n(z-z_0)^n, \quad b_0 \neq 0, \quad b_n \neq 0, \quad b_k \neq 0, \quad k \geq 1.$$

Пусть w_0 – один из корней многочлена $b_0 + b_k z^k$. Ясно, что $w_0 \neq 0$. Будем искать h в виде $h = tw_0$, $t \in \mathbb{R}$, $0 < t < 1$. Тогда

$$\begin{aligned}
f(z_0+h) &= b_0 + b_k w_0^k t^k + b_{k+1} w_0^{k+1} t^{k+1} + \cdots + b_n w_0^n t^n = \\
&= b_0(1-t^k) + \left(b_{k+1} w_0^{k+1} + \cdots + b_n w_0^n t^{n-k-1} \right) t^{k+1}
\end{aligned}$$

Пусть $C := |b_{k+1}||w_0|^{k+1} + \cdots + |b_n||w_0|^n$. При $t < |b_0|/C$ имеем

$$\begin{aligned}
|f(z_0+h)| &\leq |b_0|(1-t^k) + \left| b_{k+1} w_0^{k+1} + \cdots + b_n w_0^n t^{n-k-1} \right| \cdot t^{k+1} \leq \\
&\leq |b_0|(1-t^k) + \left(|b_{k+1}||w_0|^{k+1} + \cdots + |b_n||w_0|^n t^{n-k-1} \right) t^{k+1} \leq \\
&\leq |b_0|(1-t^k) + Ct^{k+1} = |b_0| + (Ct - |b_0|)t^k < |b_0| = |f(z_0)|.
\end{aligned}$$

Это выполнено при $h = tw_0$, $0 < t < \min(|b_0|/C,\ 1)$. Таким образом, в качестве ϵ можно взять любое число меньшее $|w_0| \min(|b_0|/C,\ 1)$. $\square$

Доказательство основной теоремы алгебры

Доказательство. Пусть $f \in \mathbb{C}[z]$, $\deg(f) > 1$. Предположим, что $|f(z)| > 0$, для любого $z \in \mathbb{C}$. Положим $M := \inf |f(z)|$. Существует последовательность $z_n \in \mathbb{C}$ такая, что $|f(z_n)|$ сходится к M. Имеются два случая.

Случай: последовательность $|z_n|$ не является ограниченной. Тогда для любого $R \in \mathbb{R}$ существует z_n такое, что $|z_n| > R$. Но по лемме о возрастании модуля для любого $c > 0$ существует R такое, что $|f(z)| > c$ при $|z| > R$. Противоречие.

Случай: последовательность $|z_n|$ ограничена. Запишем $z_n = x_n + iy_n$. Последовательности x_n и y_n ограничены. Выберем сходящиеся подпоследовательности x_{n_k} и y_{n_k}:

$$\lim x_{n_k} = x_0, \qquad \lim y_{n_k} = y_0.$$

Тогда последовательность $z_{n_k} = x_{n_k} + iy_{n_k}$ сходится к $z_0 = x_0 + iy_0$:

$$\lim z_{n_k} = z_0.$$

Следовательно, последовательность $f(z_{n_k})$ сходится к $f(z_0)$. Это противоречит лемме Даламбера. □

Неприводимые многочлены над $\mathbb{C}$ и $\mathbb{R}$

Следствие. *Неприводимые многочлены $f \in \mathbb{C}[z]$ – это только многочлены степени* 1.

Следствие. *Число корней многочлена $f \in \mathbb{C}[z]$, подсчитанных с учетом кратностей равно* $\deg(f)$.

Предложение. *Неприводимые многочлены $f \in \mathbb{R}[t]$ – это только многочлены степени* 1 *и многочлены степени* 2 *с отрицательным дискриминантом.*

В доказательстве используется следующий факт.

Лемма. *Пусть $f \in \mathbb{R}[t]$. Тогда для любого $w \in \mathbb{C}$ имеем $f(\overline{w}) = \overline{f(w)}$.*

Следствие. *Пусть $f \in \mathbb{R}[t]$ и пусть $w \in \mathbb{C}$ – корень f. Тогда и $\overline{w}$ – корень f.*

Доказательство предложения. Пусть $f \in \mathbb{R}[t]$ неприводим и пусть $\deg(f) > 1$. Тогда f не имеет действительных корней. Пусть $w \in \mathbb{C}$ – комплексный корень f. Тогда $\overline{w} \neq w$ и $\overline{w}$ – также корень f и $\overline{w} \neq w$. Следовательно, f делится на действительный многочлен второй степени

$$(t - w)(t - \overline{w}) = t^2 - 2\,\mathrm{Re}(w)t + |w|^2 \in \mathbb{R}[t]. \qquad \square$$

Поле частных целостного кольца

Определение. Пусть R – целостное кольцо. *Полем частных* кольца R называется поле $\mathbb{K}$ такое, что

(1) $\mathbb{K}$ содержит R как подкольцо;

(2) любой элемент $c \in \mathbb{K}$ представляется в виде $c = a/b$, a, $b \in R$, $b \neq 0$.

Полем частных кольца R обозначается через $\mathrm{Frac}(R)$.

Теорема. (1) *Для любого целостного кольца R поле частных $\mathrm{Frac}(R)$ существует.*

(2) *Поле частных $\mathrm{Frac}(R)$ целостного кольца R единственно с точностью до изоморфизма, т.е. если $\mathrm{Frac}(R)'$ – другое поле частных, то существует изоморфизм $\varphi : \mathrm{Frac}(R) \to \mathrm{Frac}(R)'$, который является тождественным на $R \subset \mathrm{Frac}(R)$.*

Доказательство. Существование. Рассмотрим множество

$$P := \{(a, b) \in R \times R \mid b \neq 0\}.$$

Зададим следующее отношение на этом множестве:

$$(a, b) \sim (a', b') \iff ab' = a'b.$$

Лемма. *Отношение $\sim$ является отношением эквивалентности.*

Доказательство. Ясно, что $\sim$ рефлексивно и симметрично. Докажем транзитивность. Пусть $(a, b) \sim (a', b')$ и $(a', b') \sim (a'', b'')$. Тогда $ab' = a'b$ и $a'b'' = a''b'$. Отсюда

$$ab'b'' = a'bb'' = a'b''b = a''b'b.$$

Сокращая на b', получим $ab'' = a''b$. Таким образом, $(a, b) \sim (a'', b'')$. □

Обозначим через $\mathrm{Frac}(R)$ множество классов эквивалентности $P/\sim$ и определим на этом множестве операции

$$(a_1, b_1) \cdot (a_2, b_2) = (a_1a_2, b_1b_2), \quad (a_1, b_1) + (a_2, b_2) = (a_1b_2 + a_2b_1, b_1b_2).$$

Лемма. *Формулы определяют корректные операции на $\mathrm{Frac}(R)$.*

Доказательство. Пусть $(a_1, b_1) \sim (a_1', b_1')$ и $(a_2, b_2) \sim (a_2', b_2')$. Тогда $a_1b_1' = a_1'b_1$ и $a_2b_2' = a_2'b_2$. Отсюда $a_1a_2b_1'b_2' = a_1'a_2'b_1b_2$, т.е. $(a_1a_2, b_1b_2) \sim (a_1'a_2', b_1'b_2')$. Следовательно,

$$(a_1, b_1) \cdot (a_2, b_2) = (a_1a_2, b_1b_2) \sim (a_1'a_2', b_1'b_2') = (a_1', b_1') \cdot (a_2', b_2').$$

Аналогично,

$$(a_1b_2 + a_2b_1)b_1'b_2' = (a_1b_1')b_2b_2' + (a_2b_2')b_1b_1' = (a_1'b_1)b_2b_2' + (a_2'b_2)b_1b_1' = (a_1'b_2' + a_2'b_1')b_1b_2.$$

Следовательно,

$$(a_1, b_1) + (a_2, b_2) = (a_1b_2 + a_2b_1, b_1b_2) \sim (a_1'b_2' + a_2'b_1', b_1'b_2') = (a_1', b_1') + (a_2', b_2').$$ □

Класс эквивалентности пары $(a, b) \in \mathrm{Frac}(R)$ традиционно обозначается через a/b (а отношение $\sim$ в $\mathrm{Frac}(R)$ считается равенством).

Докажем, что множество $\mathrm{Frac}(R)$ с операциями $+$ и $\cdot$ является полем. При сложении приводим к общему знаменателю и все сводится к умножению числителей:

$$\frac{a}{b} + \frac{a'}{b} = \frac{ab + a'b}{b^2} = \frac{(a + a')b}{b^2} = \frac{a + a'}{b}.$$

Элемент $0/1$ является нулевым, элемент $(-a)/1$ является противоположным к $a/1$. Умножение: коммутативность и ассоциативность очевидны, элемент $1/1$ является единичным, элемент b/a является обратным к a/b. Дистрибутивность (сначала приводим к общему знаменателю):

$$\left(\frac{a_1}{b} + \frac{a_2}{b}\right)\frac{a_3}{b_3} = \frac{a_1 + a_2}{b} \cdot \frac{a_3}{b_3} = \frac{(a_1 + a_2)a_3}{bb_3} = \frac{a_1}{b} \cdot \frac{a_3}{b_3} + \frac{a_2}{b} \cdot \frac{a_3}{b_3}.$$

Имеется естественное вложение $R \hookrightarrow \mathrm{Frac}(R)$, $a \longmapsto a/1$. Это инъективный гомоморфизм. Мы отождествим R с подкольцом $\mathrm{Frac}(R)$. По построению все элементы $\mathrm{Frac}(R)$ являются отношениями элементов R. Таким образом, построенное поле $\mathrm{Frac}(R)$ удовлетворяет всем условиям поля частных.

Единственность. Любой элемент $c \in \mathrm{Frac}(R)$ представляется в виде $c = a/b$, a, $b \in R$, $b \neq 0$. Определим отображение $\varphi : \mathrm{Frac}(R) \to \mathrm{Frac}(R)'$ по правилу

$$\varphi(a/b) = a /\!\!/ b,$$

где $/\!\!/$ – деление в поле $\mathrm{Frac}(R)'$. Во первых, нужно проверить корректность определения. Пусть $a/b = a'/b'$. Тогда $ab' = a'b$. Следовательно, $a' /\!\!/ b' = a /\!\!/ b$ и

$$\varphi(a/b) = a /\!\!/ b = a' /\!\!/ b' = \varphi(a'/b'),$$

т.е. $\varphi(a/b)$ не зависит от способа представления a/b в виде дроби. Далее

$$\varphi\left(\frac{a_1}{b_1} + \frac{a_2}{b_2}\right) = \varphi\left(\frac{a_1b_2 + a_2b_1}{b_1b_2}\right) = \frac{\overline{a_1b_2 + a_2b_1}}{\overline{b_1b_2}} = \frac{\overline{a_1}}{\overline{b_1}} + \frac{\overline{a_2}}{\overline{b_2}} = \varphi\left(\frac{a_1}{b_1}\right) + \varphi\left(\frac{a_2}{b_2}\right),$$

$$\varphi\left(\frac{a_1}{b_1} \cdot \frac{a_2}{b_2}\right) = \varphi\left(\frac{a_1a_2}{b_1b_2}\right) = \frac{\overline{a_1a_2}}{\overline{b_1b_2}} = \frac{\overline{a_1}}{\overline{b_1}} \cdot \frac{\overline{a_2}}{\overline{b_2}} = \varphi\left(\frac{a_1}{b_1}\right) \cdot \varphi\left(\frac{a_2}{b_2}\right).$$

Следовательно, $\varphi : \mathrm{Frac}(R) \to \mathrm{Frac}(R)'$ – гомоморфизм колец. По построению он сюръективен. Так как $\mathrm{Frac}(R)$ – поле, то φ также инъективен. Следовательно, φ – изоморфизм. □

Пример. Полем частных для кольца целых чисел $\mathbb{Z}$ является поле рациональных чисел $\mathbb{Q}$.

Пример. Пусть $\Bbbk$ – поле. Для кольца формальных степенных рядов $\Bbbk[[t]]$ над $\Bbbk$ поле частных – это поле формальных рядов Лорана $\Bbbk((t))$.

Лекция 16

Поле рациональных функций. Простейшие дроби. Разложение рациональной дроби в сумму простейших. Многочлены над факториальным кольцом. Лемма Гаусса. Факториальность кольца многочленов над факториальным кольцом.

Поле рациональных функций.

Определение. Поле частных $\mathrm{Frac}(R)$ кольца $R = \Bbbk[t]$ многочленов над полем называется *полем рациональных функций* и обозначается $\Bbbk(t)$.

Напомним, что $\Bbbk[t]$ – факториальное кольцо. Будем говорить, что дробь $f/g \in \Bbbk(t)$ имеет *несократимую* форму, если $\mathrm{НОД}(f, g) = 1$.

Утверждение. (1) *Для любой дроби $f/g \in \Bbbk(t)$ имеется несократимая запись.*

(2) *Любая другая запись $f'/g' \in \Bbbk(t)$ этой дроби получается из несократимой $f/g \in \Bbbk(t)$ умножением числителя и знаменателя на некоторый ненулевой элемент $h \in \Bbbk[t]$: $f' = fh$, $g' = gh$.*

(3) *Несократимая форма единственна с точностью до умножения числителя и знаменателя на константу $\lambda \in \Bbbk$.*

Доказательство. Для доказательства (1) просто поделим и числитель и знаменатель на $\mathrm{НОД}(f, g)$.

Докажем (2). Пусть $f'/g' = f/g$. Тогда $f'g = fg'$. Так как $(f, g) = 1$, то $f \mid f' = fh'$ и $g \mid g' = gh''$. Отсюда $fh'g = fgh''$. Следовательно, $h' = h''$.

Утверждение (3) следует из (2). □

Говорят, что дробь $f/g \in \Bbbk(t)$ – *правильная*, если $\deg(f) < \deg(g)$.

Лемма. *Сумма правильных дробей является правильной дробью.*

Утверждение. *Любая дробь $f/g \in \Bbbk(t)$ единственным образом разлагается в сумму*

$$\frac{f}{g} = q + \frac{f^*}{g}, \qquad q \in \Bbbk[t], \quad \frac{f^*}{g} \text{ – правильная дробь.}$$

Доказательство. Делим с остатком: $f = qg + f^*$. □

Замечание. Конструкция поля рациональных функций позволяет строить много новых примеров полей. Например, начиная с конечного поля $\mathbb{F}_p$ можно построить бесконечное поле $\mathbb{F}_p(t)$ (а также поля $\mathbb{F}_p(t_1,\dots,t_n)$) имеющие характеристику $p>0$.

Переход от поля $\Bbbk$ к полю $\Bbbk(t_1,\dots,t_n)$ называется *чисто трансцендентным расширением.*

Простейшие дроби

Определение. Дробь $f/g \in \Bbbk(t)$ называется *простейшей*, если $g=p^k$, где p – неприводимый многочлен и $\deg(f)<\deg(p)$.

Примеры.
- Дробь $c/(t-\alpha)^k$ является простейшей над любым полем $\Bbbk$.
- Если $\Bbbk=\mathbb{C}$, то любая простейшая дробь имеет вид $c/(t-\alpha)^k$.
- $\Bbbk=\mathbb{R}$, то любая простейшая дробь имеет вид $c/(t-\alpha)^k$ или $(at+b)/(t^2+pt+q)^k$, где знаменатель t^2+pt+q не имеет действительных корней.

Теорема. *Пусть $f/g\in\Bbbk(t)$ – рациональная дробь и пусть $g=p_1^{m_1}\cdots p_r^{m_r}$ – разложение знаменателя в произведение неприводимых. Тогда f/g является суммой многочлена и простейших дробей со знаменателями $p_1,p_1^2,\dots,p_1^{m_1},p_2,p_2^2,\dots,p_2^{m_2},\dots,p_r,p_r^2,\dots,p_r^{m_r}$:*

$$\frac{f}{g}=q+\sum_{i=1}^{r}\sum_{k=1}^{m_i}\frac{f_{i,k}}{p_i^k},\qquad \deg(f_{i,k})<\deg(p_i).$$

Это разложение единственно с точностью до порядка слагаемых.

Можно считать, что f/g – правильная дробь. Доказательство теоремы получается из следующих двух лемм.

Лемма. *Пусть f/g – правильная рациональная дробь и пусть $g=g_1g_2$ – разложение знаменателя в произведение взаимно простых многочленов отличных от констант. Тогда существует разложение*

$$f/g=f_1/g_1+f_2/g_2.$$

в сумму правильных дробей. Это разложение единственно.

Доказательство. Существуют многочлены u, v такие, что $g_1u+g_2v=1$. Отсюда

$$f/g=fv/g_1+fu/g_2.$$

Дроби в правой части можно разложить в суммы многочленов и правильных дробей. Получим

$$f/g=f_1/g_1+f_2/g_2+f^*,$$

где f_1/g_1 и f_2/g_2 – правильные дроби, а f^* – многочлен. Предположим, что $f^*\neq 0$. Тогда $\deg(gf^*)=\deg(f-f_1g_2-f_2g_1)<\deg(g)$. Противоречие.

Пусть имеются два различных разложения $f/g=f_1/g_1+f_2/g_2=f_1^*/g_1+f_2^*/g_2$. Тогда $(f_1-f_1^*)g_2=(f_2^*-f_2)g_1$. Отсюда получаем, что g_1 делит $f_1-f_1^*$. Следовательно, $\deg(f_1-f_1^*)\geq\deg(g_1)$. Противоречие. □

Следствие. *Всякая правильная рациональная дробь вида f/g, где $g = p_1^{m_1} \cdots p_r^{m_r}$ – разложение в произведение различных неприводимых, является суммой правильных дробей со знаменателями $p_1^{m_1}, p_2^{m_2}, \ldots, p_r^{m_r}$. Это разложение единственно с точностью до порядка.*

Доказательство. Индукция по степени знаменателя с использованием леммы. □

Лемма. *Всякая правильная рациональная дробь вида f/p^m, где p – неприводимый многочлен представляется в виде*

$$f/p^m = f_m/p^m + f_{m-1}/p^{m-1} + \cdots + f_1/p,$$

где $\deg(f_i) < \deg(p)$. Это разложение единственно.

Доказательство. Существование. Индукция по m. Делим f на p с остатком:

$$f = qp + f_m, \qquad \deg(f_m) < \deg(p), \quad \deg(q) < \deg(p^{m-1}).$$

Отсюда $f/p^m = f_m/p^m + q/p^{m-1}$.

Единственность. Предположим, что имеется другое разложение

$$f/p^m = f_m^*/p^m + f_{m-1}^*/p^{m-1} + \cdots + f_1^*/p$$

и пусть $h_i = f_i - f_i^*$. Тогда

$$0 = h_m/p^m + h_{m-1}/p^{m-1} + \cdots + h_1/p,$$

Пусть h_k – первый член отличный от нуля. Тогда

$$0 = h_k + h_{k-1}p + \cdots + h_1 p^{k-1}.$$

Отсюда $h_k = 0$. Противоречие. □

Факториальность кольца многочленов над факториальным кольцом

Теорема. *Пусть R – факториальное кольцо. Тогда кольцо $R[t]$ также факториально.*

Следствие. *Кольцо $\mathbb{Z}[t]$ факториально.*

Следствие. *Пусть $\Bbbk$ – поле. Тогда кольцо $\Bbbk[t_1][t_2]\cdots[t_n]$ факториально.*

Многочлен $f = a_n t^n + \cdots + a_0 \in R[t]$ положительной степени называется *примитивным*, если НОД$(a_n, \ldots, a_0) = 1$.

Обозначим через $\mathbb{K}$ поле частных кольца R, т.е. $\mathbb{K} := \mathrm{Frac}(R)$.

Лемма. *Любой многочлен $f \in \mathbb{K}[t]$ положительной степени представляется в виде $f = \frac{\alpha}{\beta} f^*$, где $\alpha, \beta \in R$, $\beta \neq 0$, $f^* \in R[t]$ – примитивный многочлен.*

Доказательство. Пусть $f = a_n t^n + \cdots + a_0$, где $a_i = b_i/c_i$, b_i, $c_i \in R$, $c_i \neq 0$. Положим $\beta = c_0 \cdots c_n$ и $a_i' = b_i\beta/c_i \in R$. Тогда $f = \frac{1}{\beta}\sum a_i' t^i$. Положим $\alpha = \text{НОД}(a_0, \ldots, a_n)$ и $a_i^* = a_i'/\alpha$. Тогда $f = \frac{\alpha}{\beta}\sum a_i^* t^i$. □

Лемма (лемма Гаусса)**.** *Пусть f, $g \in R[t]$ – примитивные многочлены. Тогда fg – примитивный многочлен.*

Доказательство. Запишем

$$f = \sum a_i t^i, \quad g = \sum b_j t^j, \quad fg = \sum c_k t^k,$$

где

$$c_k = \sum_{i+j=k} a_i b_j.$$

Предположим противное. Тогда существует простой элемент $p \in R$ такой, что

$$p \mid c_k, \quad \forall k.$$

По нашему предположению

$$\exists i \quad p \nmid a_i, \qquad \exists j \quad p \nmid b_j.$$

Выберем эти i и j минимальными. Тогда p делит все члены суммы $c_k = \sum_{i+j=k} a_i b_j$ кроме $a_i b_j$. Противоречие. □

Следствие. *Пусть f, $g \in R[t]$, причем g – примитивный многочлен. Если $g \mid f$ в кольце $\mathbb{K}[t]$, то $g \mid f$ в кольце $R[t]$.*

Доказательство. Пусть $f = gh$, где $h \in \mathbb{K}[t]$. Имеет место представление $h = \frac{\alpha}{\beta}h^*$, где $\alpha/\beta \in \mathbb{K}$ – несократимая дробь и $h^* \in R[t]$ – примитивный многочлен. Тогда $\beta f = \alpha g h^*$. Это противоречит лемме Гаусса. □

Следствие. *Пусть $f \in R[t]$ – примитивный многочлен. Тогда f – простой элемент в кольце $R[t]$ тогда и только тогда, когда многочлен f неприводим в кольце $\mathbb{K}[t]$.*

Таким образом простые элементы $f \in R[t]$ бывают двух типов:

- $\deg(f) = 0$, $f \in R$ – простой элемент,
- $\deg(f) > 0$, многочлен f примитивен в $R[t]$ и неприводимый в $\mathbb{K}[t]$.

Лемма. *Пусть $p \in R[t]$ – простой элемент в кольце $R[t]$ и пусть $p \mid fg$, f, $g \in R[t]$. Тогда $p \mid f$ или $p \mid g$.*

Доказательство. Случай $\deg(p) = 0$. Запишем $f = af^*$, $g = bg^*$, где f^*, $g^* \in R[t]$ – примитивные многочлены, а a, $b \in R$. По лемме Гаусса f^*g^* примитивный многочлен. Откуда получаем, что из $p \mid ab$ следует, что $p \mid a$ или $p \mid b$.

Случай $\deg(p) > 0$. Тогда p – примитивный многочлен. Пусть $p \nmid f$ в кольце $R[t]$. Тогда $p \nmid f$ в кольце $\mathbb{K}[t]$. По соответствующей лемме для многочленов над полем имеем $p \mid g$ в кольце $\mathbb{K}[t]$. По следствию $p \mid g$ в кольце $R[t]$. □

Доказательство теоремы. Существование. Применим индукцию по степени f. Если $\deg(f) = 0$, то $f \in R$ и утверждение выполнено, поскольку кольцо R факториально. Пусть $\deg(f) > 0$ и утверждение имеет место для все многочленов в $R[t]$ степени меньшей $\deg(f)$. Запишем $f = af^*$, где $f^* \in R[t]$ – примитивный многочлен, а $a \in R$. Если элемент f^* не является простым, то $f^* = f_1^* f_2^*$, где f_1^*, f_2^* – примитивные многочлены. По предположению индукции они допускают разложение в произведение простых элементов $\in R[t]$. Поскольку R факториально, то и a допускает разложение в произведение простых элементов.

Единственность. Пусть $f = p_1^{m_1} \cdots p_r^{m_r}$ – разложение в произведение простых с наименьшим $\sum m_i$. Докажем утверждение индукцией по $\sum m_i$. Предположим, что

$$f = p_1^{m_1} \cdots p_r^{m_r} = p_1'^{k_1} \cdots p_s'^{k_s}$$

По лемме $p_1 \mid p_j'$ для некоторого j. Отсюда элементы p_1 и p_j' отличаются обратимым множителем. Сокращая, получим два разложения f/p_1 с меньшим значением $\sum m_i$. По предположению индукции они совпадают. □

Лекция 17

Многочлены от нескольких переменных. Симметрические многочлены. Лексикографический порядок.

Многочлены от нескольких переменных

В этой лекции мы обсудим многочлены от нескольких переменных. Конечно, естественный способ определить кольцо многочленов от $t_1, \ldots, t_n$ над R – это задать его последовательным присоединением этих переменных: $R[t_1, \ldots, t_n] = R[t_1][t_2] \ldots [t_n]$. Однако, небольшим недостатком этого метода является то, что такое задание зависит от выбора порядка $t_1, \ldots, t_n$. Сформулируем универсальное определение.

Определение. Пусть R – коммутативное ассоциативное кольцо с 1. *Кольцом многочленов* от n переменных над R называется коммутативное ассоциативное кольцо S, содержащее выделенные элементы $t_1, \ldots, t_n$, такое, что

(1) R содержится в S как подкольцо и $S \neq R$,

(2) любой элемент $f \in S$ *однозначно* представляется в виде

$$f = \sum_{k_1,\ldots,k_n} a_{k_1,\ldots,k_n} t_1^{k_1} \cdots t_n^{k_n}, \quad a_{k_1,\ldots,k_n} \in R.$$

Элементы S называются *многочленами* от переменных $t_1, \ldots, t_n$. Выделенный элементы $t_1, \ldots, t_n$ называются независимыми переменными.

Лемма. *Пусть S – кольцо многочленов от переменных $t_1, \ldots, t_n$ и пусть $S_1 \subset S$ – подмножество элементов вида*

$$(*) \qquad \sum_{k_1,\ldots,k_{n-1}} a_{k_1,\ldots,k_{n-1}} t_1^{k_1} \cdots t_{n-1}^{k_{n-1}}, \quad a_{k_1,\ldots,k_{n-1}} \in R.$$

Тогда S_1 – подкольцо, являющееся кольцом многочленов от переменных $t_1, \ldots, t_{n-1}$ и $S = S_1[t_n]$.

Доказательство. Так как суммы, разности и произведения элементов вида $(*)$ снова являются элементами того же вида, то S_1 – подкольцо. Для него выполнены свойства

(1) и (2) из определения. Поэтому S_1 – кольцом многочленов от переменных $t_1, \dots, t_{n-1}$. Наконец, любой многочлен

$$f = \sum_{k_1,\dots,k_n} a_{k_1,\dots,k_n} t_1^{k_1} \cdots t_n^{k_n} \in S$$

можно записать в виде

$$f = \sum_{k_n} \left(\sum_{k_1,\dots,k_{n-1}} a_{k_1,\dots,k_{n-1},k_n}\, t_1^{k_1} \cdots t_{n-1}^{k_{n-1}} \right) t_n^{k_n} = \sum_{k_n} f_{k_n} t_n^{k_n},$$

где

$$f_{k_n} = \sum_{k_1,\dots,k_{n-1}} a_{k_1,\dots,k_{n-1},k_n}\, t_1^{k_1} \cdots t_{n-1}^{k_{n-1}} \in S_1.$$

Такое представление единственно. По определению кольца многочленов от одной переменной $S = S_1[t_n]$. □

Предложение. *Пусть R – коммутативное ассоциативное кольцо с 1. Кольцо многочленов S от n переменных над R существует. Если S' – другое кольцо многочленов от n переменных над R, то существует изоморфизм*

$$\varphi : S \longrightarrow S'$$

такой, что $\varphi(a) = a$ для всех $a \in R$ и $\varphi(t_i) = t'_i$, где $t_1, \dots, t_n$ (соответственно, $t'_1, \dots, t'_n$) – независимые переменные в S (соответственно, в S').

Доказательство. Существование. Определим кольцо S по индукции:

$$S = R[t_1][t_2] \dots [t_n].$$

Тогда S – коммутативное ассоциативное кольцо с единицей. Более того, S содержит R. Для доказательства свойства (2) из определения применим индукцию по n. База индукции очевидна. Пусть (2) выполнено для $n-1$. По определению $S = S_1[t_n]$, где $S_1 := R[t_1] \dots [t_{n-1}]$. Следовательно, любой элемент $f \in S$ единственным образом представляется в виде

$$f = \sum_{k_n} f_{k_n} t_n^{k_n}, \qquad f_{k_n} \in R[t_1, \dots, t_{n-1}].$$

По предположению индукции $f_{k_n} = \sum a_{k_1,\dots,k_{n-1},k_n} t_1^{k_1} \cdots t_{n-1}^{k_{n-1}}$ и это представление также единственно. Подставляя в формулу выше, получим

$$f = \sum_{k_n} \left(\sum_{k_1,\dots,k_{n-1}} a_{k_1,\dots,k_{n-1},k_n} t_1^{k_1} \cdots t_{n-1}^{k_{n-1}} \right) t_n^{k_n} = \sum_{k_1,\dots,k_n} a_{k_1,\dots,k_n} t_1^{k_1} \cdots t_n^{k_n}.$$

что доказывает (2).

Единственность. Снова применим индукцию по n. База индукции следует из соответствующего факта для многочленов от одной переменной. Предположим, что существование требуемого изоморфизма φ известно для колец многочленов от $n-1$ переменных. Пусть S и S' – два кольца многочленов от n переменных над R. По лемме $S = S_1[t_n]$ и $S' = S'_1[t'_n]$, где $S_1 \subset S$ и $S'_1 \subset S'$ – подкольца, являющиеся кольцами многочленов от переменных $t_1, \dots, t_{n-1}$ и $t'_1, \dots, t'_{n-1}$, соответственно. По предположению индукции имеется изоморфизм $\varphi : S_1 \to S'_1$ такой, что φ тождественен на R и $\varphi(t_i) = t'_i$ для $i = 1, \dots, n-1$. Из единственности кольца многочленов от одной переменной следует, что φ может быть продолжен до изоморфизма $\varphi : S \to S'$ такого, что $\varphi(t_n) = t'_n$. □

Кольцо многочленов от переменных $t_1, \dots, t_n$ обозначается через $R[t_1, \dots, t_n]$. Ясно, что умножение в $R[t_1, \dots, t_n]$ задается формулами

$$t_1^{k_1} \cdots t_n^{k_n} \cdot t_1^{l_1} \cdots t_n^{l_n} = t_1^{k_1+l_1} \cdots t_n^{k_n+l_n}.$$

Таким образом, для многочленов

$$f = \sum a_{k_1,\dots,k_n} t_1^{k_1} \cdots t_n^{k_n}, \qquad g = \sum b_{l_1,\dots,l_n} t_1^{l_1} \cdots t_n^{l_n}$$

их произведение выглядит следующим образом

$$fg = \sum c_{m_1,\dots,m_n} t_1^{m_1} \cdots t_n^{m_n}, \quad c_{m_1,\dots,m_n} = \sum a_{k_1,\dots,k_n} b_{l_1,\dots,l_n},$$

где суммирование ведётся по всем наборам $k_1, \dots, k_n$, $l_1, \dots, l_n$ таким, что $k_1 + l_1 = m_1$, $\dots$, $k_n + l_n = m_n$.

Как и в случае одной переменной, можно определить поле рациональных функций от нескольких переменных:

$$\Bbbk(t_1, \dots, t_n) := \mathrm{Frac}(\Bbbk[t_1, \dots, t_n]).$$

Здесь мы, конечно, предполагаем, что $\Bbbk$ – поле.

Следствие. $R[t_1, \dots, t_n] \simeq R[t_1][t_2] \dots [t_n]$.

Следствие. *Кольцо $R[t_1, \dots, t_n]$ не имеет делителей нуля тогда и только тогда, когда R не имеет делителей нуля.*

Следствие. $R[t_1, \dots, t_n] \simeq R[t_{\sigma(1)}, \dots, t_{\sigma(n)}]$ *для любой подстановки* $\sigma \in \mathrm{S}_n$.

Доказательство. Следует из того, что в определении кольца многочленов порядок переменных не имеет значения.. □

Используя результат предыдущей лекции, мы также получаем следующее.

Следствие. *Пусть R – факториальное кольцо. Тогда кольцо многочленов $R[t_1, \dots, t_n]$ также факториально. В частности, факториально кольцо многочленов $\Bbbk[t_1, \dots, t_n]$ над полем $\Bbbk$.*

Обозначение. $t_1^{k_1}\cdots t_n^{k_n} \in f$ если $at_1^{k_1}\cdots t_n^{k_n}$ присутствует в $f \in R[t_1,\dots,t_n]$ с ненулевым коэффициентом a.

Определение. *Степенью* ненулевого многочлена

$$f = \sum_{k_1,\dots,k_n} a_{k_1,\dots,k_n} t_1^{k_1}\cdots t_n^{k_n} \in R[t_1,\dots,t_n]$$

называется число

$$\deg(f) := \max\{k_1+\cdots+k_n \mid a_{k_1,\dots,k_n} \neq 0\}.$$

Степенью по переменной t_i этого многочлена называется число

$$\deg_{t_i}(f) := \max\{k_i \mid a_{k_1,\dots,k_n} \neq 0\}.$$

Определение. Ненулевой многочлен

$$f = \sum_{k_1,\dots,k_n} a_{k_1,\dots,k_n} t_1^{k_1}\cdots t_n^{k_n}$$

называется *однородным* степени d, если все коэффициенты $a_{k_1,\dots,k_n}$ обращаются в нуль при $\sum k_i \neq d$.

Утверждение. *Любой ненулевой многочлен $f \in R[t_1,\dots,t_n]$ представляется в виде суммы многочленов*

$$f = f_0 + f_1 + \cdots + f_d,$$

где каждый многочлен f_k является или однородным степени k или нулевым. Это разложение называется разложением в сумму однородных компонент.

Доказательство. Очевидно, нужно только сгруппировать все одночлены одной степени в соответствующее слагаемое f_k. □

Утверждение. *Пусть $f,\ g \in R[t_1,\dots,t_n]$ – однородные многочлены степеней d и e, соответственно. Тогда их произведение fg является или нулевым или однородным многочленом степени $d+e$.*

Если, кроме того, в кольце R нет делителей нуля, то $fg \neq 0$.

Доказательство. Пусть

$$f = \sum_{k_1,\dots,k_n} a_{k_1,\dots,k_n} t_1^{k_1}\cdots t_n^{k_n}, \qquad g = \sum_{l_1,\dots,l_n} b_{l_1,\dots,l_n} t_1^{l_1}\cdots t_n^{l_n},$$

где $\sum k_i = d$, $\sum l_j = e$. Тогда

$$fg = \sum a_{k_1,\dots,k_n} b_{l_1,\dots,l_n} t_1^{k_1+l_1}\cdots t_n^{k_n+l_n}.$$

$\sum(k_i+l_i) = d+e$.

Последнее утверждение следует из того, что в кольце $R[t_1,\dots,t_n]$ нет делителей нуля. □

Предложение. *Если в R нет делителей нуля, то* $\deg(fg) = \deg(f) + \deg(g)$.

Доказательство. Рассмотрим разложение в сумму однородных компонент

$$f = f_0 + \cdots f_d, \qquad g = g_0 + \cdots g_e.$$

Тогда разложение произведения $fg = \sum_{i,j} f_i g_j$ в сумму однородных компонент будет

$$fg = \sum h_k, \qquad h_k = \sum_{i+j=k} f_i g_j$$

(группируем однородные компоненты степени k). Компонента максимальной степени $f_d g_e$ отлична от нуля поскольку в R нет делителей нуля. □

Предложение. *Если в R нет делителей нуля, то* $\deg_{t_k}(fg) = \deg_{t_k} f + \deg_{t_k} g$.

Доказательство. Следует из соответствующего утверждения для многочленов от одной переменной поскольку $R_k[t_k]$, где $R_k := R[t_1, \ldots, t_{k-1}, t_{k+1}, \ldots, t_n]$ и в кольце R_k нет делителей нуля. □

Симметрические многочлены

Пусть R – коммутативное ассоциативное кольцо с единицей. Пусть $f \in R[t_1, \ldots, t_n]$ и пусть $\gamma \in \mathrm{S}_n$ – подстановка. Определим многочлен, полученный из f соответствующей перестановкой переменных:

$$f^\gamma(t_1, \ldots, t_n) := f(t_{\gamma(1)}, \ldots, t_{\gamma(n)}).$$

Таким образом, если $f = \sum_{k_1,\ldots,k_n} a_{k_1,\ldots,k_n} t_1^{k_1} \cdots t_n^{k_n}$, то

$$f^\gamma = \sum_{k_1,\ldots,k_n} a_{k_1,\ldots,k_n} t_{\gamma(1)}^{k_1} \cdots t_{\gamma(n)}^{k_n}.$$

Например, если $\gamma = [1, 2, 3]$ и $f = t_1 + t_2^2 + t_3^3 + t_4^4$, то $f^\gamma = t_2 + t_3^2 + t_1^3 + t_4^4$.

Утверждение. *Отображение*

$$\Phi_\gamma : R[t_1, \ldots, t_n] \longrightarrow R[t_1, \ldots, t_n], \qquad f \longmapsto f^\gamma$$

является изоморфизмом кольца многочленов на себя.

Доказательство. Действительно, Φ_γ взаимно однозначно и

$$\begin{aligned} (f+g)^\gamma(t_1, \ldots, t_n) &= (f+g)^\gamma(t_{\gamma(1)}, \ldots, t_{\gamma(n)}) = f^\gamma(t_{\gamma(1)}, \ldots, t_{\gamma(n)}) + g^\gamma(t_{\gamma(1)}, \ldots, t_{\gamma(n)}), \\ (fg)^\gamma(t_1, \ldots, t_n) &= (fg)^\gamma(t_{\gamma(1)}, \ldots, t_{\gamma(n)}) = f^\gamma(t_{\gamma(1)}, \ldots, t_{\gamma(n)}) g^\gamma(t_{\gamma(1)}, \ldots, t_{\gamma(n)}). \quad \square \end{aligned}$$

Определение. Многочлен называется *симметрическим*, если $f^\gamma = f$, для любой подстановки $\gamma \in \mathrm{S}_n$.

Поскольку S_n порождается транспозициями, то равенство $f^\gamma = f$ достаточно проверить для транспозиций.

Пример.

- Все многочлены от одной переменной – симметрические.
- *Степенные суммы* $s_m := t_1^m + \cdots + t_n^m$.
- *Элементарные симметрические многочлены*

$$\sigma_m := \sum_{i_1<\cdots<i_m} t_{i_1}\cdots t_{i_m} \quad 1 \le m \le n.$$

Таким образом,

$$\sigma_1 = \sum x_i, \quad \sigma_2 = \sum_{i<j} x_i x_j, \quad \sigma_3 = \sum_{i<j<k} x_i x_j x_k, \quad \ldots, \quad \sigma_n = \prod x_i.$$

- *Определитель Вандермонда*

$$\Delta := \begin{vmatrix} 1 & t_1 & \cdots & t_1^{n-1} \\ 1 & t_2 & \cdots & t_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & t_n & \cdots & t_n^{n-1} \end{vmatrix} = \prod_{i>j}(t_i - t_j)$$

Меняет знак при транспозициях. Поэтому Δ^2 – симметрический многочлен.

Предложение. *Все симметрические многочлены образуют подкольцо*

$$R[t_1, \ldots, t_n]^{\mathrm{S}_n} \subset R[t_1, \ldots, t_n].$$

Доказательство. Следует из того, что Φ_γ – изоморфизм для любой подстановки $\gamma \in \mathrm{S}_n$. □

Формулы Виета

Теорема. *Рассмотрим многочлен*

$$f = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0 \in \Bbbk[t]$$

положительной степени n, имеющий ровно n корней с учетом кратностей, и пусть

$$\alpha_1, \ldots, \alpha_n$$

– все эти корни. Тогда

$$\sigma_k(\alpha_1, \cdots, \alpha_n) = (-1)^k \frac{a_{n-k}}{a_n}$$

Таким образом,

$$\begin{aligned}
\alpha_1+\dots+\alpha_n &= -\tfrac{a_{n-1}}{a_n}\\
\alpha_1\alpha_2+\dots+\alpha_{n-1}\alpha_n &= \tfrac{a_{n-2}}{a_n}\\
\alpha_1\alpha_2\alpha_3+\dots+\alpha_{n-2}\alpha_{n-1}\alpha_n &= -\tfrac{a_{n-3}}{a_n}\\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
\alpha_1\cdots\alpha_n &= (-1)^n\tfrac{a_0}{a_n}.
\end{aligned}$$

Доказательство. Применим индукцию по n. Введем временное обозначение $\sigma_k^{n-1} = \sigma_k^{n-1}(t_1,\dots,t_{n-1})$ – элементарный симметрический многочлен от переменных $t_1,\dots,t_{n-1}$. Мы положим также $\sigma_n^{n-1}=0$. Тогда

$$\sigma_k = \sigma_k^{n-1} + t_n\sigma_{k-1}^{n-1}.$$

Запишем, $f(t)=f_1(t)(t-\alpha_n)$, где

$$f_1(t) = b_{n-1}t^{n-1}+\dots+b_1t+b_0$$

– многочлен степени $n-1$, $a_k=b_{k-1}-b_k\alpha_n$. По предположению индукции

$$\begin{aligned}
\sigma_k(\alpha_1,\dots,\alpha_n) &= \sigma_k^{n-1}(\alpha_1,\dots,\alpha_{n-1})+t_n\sigma_{k-1}^{n-1}(\alpha_1,\dots,\alpha_{n-1}) =\\
&= (-1)^k\,\frac{b_{n-1-k}}{b_{n-1}}+(-1)^{k-1}\,\frac{b_{n-1-k+1}}{b_{n-1}}\alpha_n\\
&= (-1)^k\,\frac{b_{n-1-k}-b_{n-k}\alpha_n}{a_n}\\
&= (-1)^k\,\frac{a_{n-k}}{a_n}. \quad \square
\end{aligned}$$

Лексикографический порядок

Будем говорить, что набор $(k_1,\dots,k_n)$ из n неотрицательных целых чисел меньше набора $(l_1,\dots,l_n)$ и писать $(k_1,\dots,k_n)\prec(l_1,\dots,l_n)$ если существует $1\le i\le n$ такое, что:

$$k_1=l_1,\ k_2=l_2,\dots,k_i=l_i,\qquad k_{i+1}<l_{i+1}.$$

Также мы можем определить порядок на ненулевых одночленах в $R[t_1,\dots,t_n]$:

$$a\,t_1^{k_1}\cdots t_n^{k_n}\prec b\,t_1^{l_1}\cdots t_n^{l_n}\quad \text{если}\quad (k_1,\dots,k_n)\prec(l_1,\dots,l_n).$$

Например, $2\,t_1^3t_2^4t_3^5t_4^3t_5^2\succ 3\,t_1^3t_2^4t_3^5t_4^2t_5^7$. Заметим, что этот порядок существенным образом зависит от порядка переменных $t_1,\dots,t_n$.

Лемма. *Для любых одночленов $u,v,w,u',v'\in R[t_1,\dots,t_n]$ с коэффициентом* 1 *имеем:*

(1) $u \succ v$, $v \succ w$, *то* $u \succ w$;

(2) $u \succ v$, *то* $uw \succ vw$;

(3) $u \succ v$, $u' \succ v'$, *то* $uu' \succ vv'$.

Доказательство. (1) Запишем $u = t_1^{k_1} \cdots t_n^{k_n}$, $v = t_1^{l_1} \cdots t_n^{l_n}$, $w = t_1^{m_1} \cdots t_n^{m_n}$. Пусть i – максимальное такое, что $k_1 = l_1 = m_1, \ldots, k_{i-1} = l_{i-1} = m_{i-1}$ (т.е. t_i – первая переменная, входящая в u, v, w в различных степенях). Отсюда $k_i \geq l_i \geq m_i$. Причем по крайней мере одно из этих неравенств – строгое.

(2) Запишем $u = t_1^{k_1} \cdots t_n^{k_n}$, $v = t_1^{l_1} \cdots t_n^{l_n}$, $w = t_1^{m_1} \cdots t_n^{m_n}$. Пусть i – максимальное такое, что $k_1 = l_1, \ldots, k_{i-1} = l_{i-1}$. Тогда $k_i > l_i$. Следовательно, $k_1 + m_1 = l_1 + m_1, \ldots, k_{i-1} + m_{i-1} = l_{i-1} + m_{i-1}, k_i + m_i > l_i + m_i$.

(3) $uu' \succ vu' \succ vv'$. □

Лекция 18

Симметрические многочлены. Основная теорема и симметрических многочленах. Формулы Виета. Дискриминант. Результант (определение и свойства). Связь результанта и дискриминанта.

Основная теорема о симметрических многочленах

Теорема (основная теорема о симметрических многочленах)**.** *Пусть $f \in R[t_1, \dots, t_n]^{\mathrm{S}_n}$ – симметрический многочлен. Тогда существует единственный многочлен $g \in R[x_1, \dots, x_n]$ такой, что*

$$f(t_1, \dots, t_n) = g(\sigma_1, \dots, \sigma_n).$$

Иначе говоря, любой симметрический многочлен единственным образом представляется в виде многочлена от элементарных симметрических.

Для простоты, мы докажем эту теорему в при одном дополнительном условии: мы будем предполагать, что кольцо R не имеет делителей нуля.

Определение. Пусть $f \in R[t_1, \dots, t_n]$. Одночлен $a\, t_1^{k_1} \cdots t_n^{k_n} \in f$ называется *старшим* в f, если $a\, t_1^{k_1} \cdots t_n^{k_n} \prec b\, t_1^{l_1} \cdots t_n^{l_n}$ для любого другого $b\, t_1^{l_1} \cdots t_n^{l_n} \in f$.

Введем временное обозначение $c(f)$ – старший член f и $o(f) = f - c(f)$

Пример. Пусть $f = t_1^2 t_2 + t_1 t_2 + t_1 t_2^2 + t_3^5$. Тогда $t_1^2 t_2$ – старший член.

Предложение. $c(fg) = c(f)c(g)$ (*старший член произведения равен произведению старших членов*).

Доказательство. Пусть $\tilde{f} \in o(f)$, $\tilde{g} \in o(g)$. Тогда $c(f) \succ \tilde{f}$ и $c(g) \succ \tilde{g}$. Следовательно, $c(f)c(g) \succ \tilde{f}\tilde{g}$, $c(f)c(g) \succ c(f)\tilde{g}$ и $c(f)c(g) \succ \tilde{f}c(g)$. Отсюда $c(f)c(g)$ – старший член. □

Лемма. *Пусть $f \in R[t_1, \dots, t_n]^{\mathrm{S}_n}$ и пусть $c(f) = a\, t_1^{k_1} \cdots t_n^{k_n}$. Тогда $k_1 \geq \cdots \geq k_n$.*

Доказательство. Предположим, что $k_i < k_{i+1}$. Возьмем это i минимальным. Рассмотрим транспозицию $\gamma = [i, i+1] \in \mathrm{S}_n$. Так как $c(f) = a t_1^{k_1} \cdots t_i^{k_i} t_{i+1}^{k_{i+1}} \cdots t_n^{k_n} \in f$, то $c(f)^\gamma = a t_1^{k_1} \cdots t_{i+1}^{k_{i+1}} t_i^{k_i} \cdots t_n^{k_n} \in f$. Но тогда $c(f)^\gamma \succ c(f)$. Противоречие. □

Лемма. *Для любого одночлена $u = t_1^{k_1} \cdots t_n^{k_n}$ такого, что $k_1 \geq \cdots \geq k_n$ существует единственный набор индексов $(l_1, \dots, l_n)$ такой, что $c(\sigma_1^{l_1} \cdots \sigma_n^{l_n}) = u$.*

Доказательство. Имеем $c(\sigma_r) = t_1 \cdots t_r$. Отсюда

$$c(\sigma_1^{l_1} \cdots \sigma_n^{l_n}) = t_1^{l_1}(t_1t_2)^{l_2} \cdots (t_1 \cdots t_n)^{l_n} = t_1^{l_1+\cdots+l_n} t_2^{l_2+\cdots+l_n} \cdots t_n^{l_n}.$$

Решаем следующую систему относительно $l_1, \ldots, l_n$

$$\begin{cases} l_1 + \quad \cdots + l_n &= k_1 \\ \quad l_2 + \cdots + l_n &= k_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \quad\quad\quad\quad l_n &= k_n \end{cases}$$

Получаем единственное решение

$$l_n = k_n, \qquad l_i = k_i - k_{i+1}, \quad i < n.$$

□

Доказательство теоремы. Существование. Предположим противное. Пусть M – множество всех $f \in R[t_1, \ldots, t_n]^{S_n}$, для которых теорема не верна. Выберем $f \in M$ с наименьшим $c(f)$. По лемме существует симметрический многочлен вида $\sigma_1^{l_1} \cdots \sigma_n^{l_n}$ такой, что $ac(\sigma_1^{l_1} \cdots \sigma_n^{l_n}) = c(f)$. Тогда $f_1 = f - a\sigma_1^{l_1} \cdots \sigma_n^{l_n}$ – симметрический многочлен, который, как и f, не выражается в виде многочлена от элементарных симметрических. Значит, $f_1 \in M$. С другой стороны, $c(f_1) \prec c(f)$. Противоречие.

Единственность. Предположим, что существуют два различных многочлена g_1, g_2 от n переменных такие, что $g_1(\sigma_1, \ldots, \sigma_n) = g_2(\sigma_1, \ldots, \sigma_n)$. Пусть $h := g_1 - g_2$. Разложим h в сумму одночленов

$$h = \sum h_i.$$

Ясно, что $h_i(\sigma_1, \ldots, \sigma_n) \neq 0$ (например, потому, что $R[t_1, \ldots, t_n]$ не имеет делителей нуля). Пусть $u_i := c(h_i(\sigma_1, \ldots, \sigma_n))$. По последней лемме среди u_i нет пропорциональных. Пусть u_* – старший среди них. Тогда u_* – единственный старший среди всех одночленов в $h(\sigma_1, \ldots, \sigma_n)$ и он не может сократиться. □

Практическая реализация.

Пусть $f \in R[t_1, \ldots, t_n]^{S_n}$. Можно считать, что f однороден степени d. Пусть

$$c(f) = at_1^{k_1} \cdots t_n^{k_n},$$

где $\sum k_i = d$. Выберем все наборы $(r_1, \ldots, r_n)$ неотрицательных целых чисел такие, что

- $r_1 \geq \cdots \geq r_n$;
- $\sum r_i = d$;
- $(r_1, \ldots, r_n) \prec (k_1, \ldots, k_n)$.

Тогда $g(\sigma_1, \ldots, \sigma_n)$ содержит все одночлены вида

$$\sigma_1^{r_1-r_2} \cdot \sigma_2^{r_2-r_3} \cdots \sigma_{n-1}^{r_{n-1}-r_n} \cdot \sigma_n^{r_n}.$$

Их коэффициенты находятся методом неопределённых коэффициентов.

Пример. Рассмотрим симметрический многочлен

$$(*) \qquad f = (t_1 - t_2)^2(t_2 - t_3)^2(t_1 - t_3)^2.$$

Его ставший член равен $t_1^4 t_2^2$. Применим описанный выше алгоритм. Для этого сначала перечислим все наборы (r_1, r_2, r_3), удовлетворяющие условиям

- $r_1 \geq r_2 \geq r_3$;
- $\sum r_i = 6$;
- $(r_1, r_2, r_3) \prec (4, 2, 0)$,

а также соответствующие одночлены $\sigma_1^{r_1-r_2} \cdot \sigma_2^{r_2-r_3} \cdot \sigma_3^{r_3}$. Получим

$(4,2,0)$	$\sigma_1^2 \cdot \sigma_2^2$
$(4,1,1)$	$\sigma_1^3 \cdot \sigma_3$
$(3,3,0)$	σ_2^3
$(3,2,1)$	$\sigma_1 \cdot \sigma_2 \cdot \sigma_3$
$(2,2,2)$	σ_3^2

Отсюда

$$(\dagger) \qquad f = c_0\sigma_1^2 \cdot \sigma_2^2 + c_1\sigma_1^3 \cdot \sigma_3 + c_2\sigma_2^3 + c_3\sigma_1 \cdot \sigma_2 \cdot \sigma_3 + c_4\sigma_3^2.$$

Заметим, что коэффициент при $\sigma_1^2 \cdot \sigma_2^2$ равен коэффициенту при старшем члене, т.е. $c_0 = 1$. Для нахождения остальных коэффициентов подставим в $(*)$ и $(\dagger)$ различные значения x_1, x_2, x_3 и сравним. Получим уравнения на c_i. Это удобно сделать при помощи таблицы:

t_1	t_2	t_3	σ_1	σ_2	σ_3	$(\dagger)$	$(*)$
1	1	0	2	1	0	$c_2 + 4$	0
2	-1	-1	0	-3	2	$-27c_2 + 4c_4$	0
1	1	1	3	3	1	$27c_1 + 27c_2 + 9c_3 + c_4 + 81$	0
2	1	-1	2	-1	-2	$-16c_1 - c_2 + 4c_3 + 4c_4 + 4$	36

Откуда находим последовательно $c_2 = -4$, $c_4 = -27$,

$$3c_1 + c_3 = 6, \quad -4c_1 + c_3 = 34, \quad c_1 = -4, \quad c_3 = 18.$$

Таким образом,

$$(\ddagger) \qquad f = \sigma_1^2 \cdot \sigma_2^2 - 4\sigma_1^3 \cdot \sigma_3 - 4\sigma_2^3 + 18\sigma_1 \cdot \sigma_2 \cdot \sigma_3 - 27\sigma_3^2.$$

Следствия из теоремы о симметрических многочленах

Ниже мы приведем несколько следствий из основной теоремы о симметрических многочленах. Для простоты мы рассмотрим случай многочленов над полем $\Bbbk$.

По-первых заметим, что кольцо симметрических многочленов изоморфно кольцу всех многочленов от того же числа переменных:

Следствие. $\Bbbk[t_1,\dots,t_n]^{\mathrm{S}_n} \simeq \Bbbk[x_1,\dots,x_n]$.

Для рациональной функции $f \in \Bbbk(t_1,\dots,t_n)$ и подстановки $\gamma \in \mathrm{S}_n$ определим рациональную функцию $f^\gamma \in \Bbbk(t_1,\dots,t_n)$ правилом

$$f^\gamma(t_1,\dots,t_n) = f(t_{\gamma(1)},\dots,t_{\gamma(n)}).$$

Несложно проверить, что это определение корректно, т.е. не зависит от представления f в виде дроби $f = g/h$. Действительно, пусть $f = g/h = g_1/h_1$. Тогда $gh_1 = hg_1$. Поскольку отображение $s \mapsto s^\gamma$ является изоморфизмом кольца $\Bbbk[t_1,\dots,t_n]$, то $g^\gamma h_1^\gamma = h^\gamma g_1^\gamma$. Отсюда $g^\gamma/h^\gamma = g_1^\gamma/h_1^\gamma$. Рациональная функция называется *симметрической*, если $f^\gamma = f$ для любой подстановки $\gamma \in \mathrm{S}_n$.

Следствие. *Пусть $f \in \Bbbk(t_1,\dots,t_n)$ – симметрическая рациональная функция. Тогда существует единственная рациональная функция $g \in R(x_1,\dots,x_n)$ такая, что*

$$f(t_1,\dots,t_n) = g(\sigma_1,\dots,\sigma_n).$$

т.е. любая симметрическая функция единственным образом представляется в виде рациональной функции от элементарных симметрических многочленов.

Доказательство. Представим f в виде дроби $f = g/h$, где $g,\, h \in \Bbbk[t_1,\dots,t_n]$, $h \neq 0$. Пусть $\gamma_1,\dots,\gamma_{n!}$ – все подстановки из S_n, причем γ_1 – тождественная. Тогда $h = h^{\gamma_1}$ и

$$f = \frac{g}{h} = \frac{g \cdot h^{\gamma_2} \cdots \gamma^{n!}}{h^{\gamma_1} \cdot h^{\gamma_2} \cdots h^{\gamma_{n!}}}$$

Ясно, что знаменатель последней дроби

$$H := h^{\gamma_1} \cdot h^{\gamma_2} \cdots h^{\gamma^{n!}}.$$

– симметрический многочлен. Так как $f^\gamma = f$ для любой подстановки $\gamma \in \mathrm{S}_n$, то и для ее числителя

$$G := g \cdot h^{\gamma_2} \cdots h^{\gamma^{n!}}$$

имеем $G^\gamma = G$, т.е. G – также симметрический многочлен. Далее применяем теорему о симметрических многочленах. □

Следующий факт получается применением формул Виета.

Следствие. *Пусть $f \in \Bbbk[t]$, $f = a_n t^n + \cdots a_1 t + a_0$ – многочлен положительной степени и пусть $\alpha_1,\dots,\alpha_n$ – корни, выписанные с учетом кратностей. Любой симметрический многочлен $\alpha_1,\dots,\alpha_n$ выражается в виде многочлена от $a_0/a_n,\dots,\, a_{n-1}/a_n$.*

Дискриминант

Пусть $\Bbbk$ – поле. Рассмотрим многочлен $f \in \Bbbk[t]$, $f = a_n t^n + \cdots a_1 t + a_0$ степени n над этим полем, имеющий ровно n корней (с учетом кратностей). Пусть $\alpha_1, \ldots, \alpha_n$ – все эти корни. Определим дискриминант многочлена f следующим образом

$$(\S) \qquad D_f = a_n^{2n-2} \prod_{i<j} (\alpha_i - \alpha_j)^2.$$

Дискриминант выражается через определитель Вандермонда:

$$D_f = a_n^{2n-2} \begin{vmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & \alpha_n & \alpha_n^2 & \cdots & \alpha_n^{n-1} \end{vmatrix}^2$$

Непосредственно из определения вытекает, что значение D_f характеризуем наличие кратных множителей у многочлена f: над полем $\Bbbk$ характеристики 0 многочлен f имеет кратные множители тогда и только тогда, когда D_f обращается в нуль. То же самое верно и в случае полей характеристики $p > 0$ при условии, что многочлен f не имеет неприводимых множителей вида $g(t^p)$.

Предложение. *D_f выражается как многочлен с целыми коэффициентами от a_i.*

Доказательство. D_f является многочленом (с целыми коэффициентами) от элементарных симметрических функций $\sigma_i(\alpha_1, \ldots, \alpha_n)$:

$$D_f = a_n^{2n-2} \cdot h(\sigma_1, \ldots, \sigma_n).$$

По формулам Виета

$$\sigma_1 = -a_{n-1}/a_n, \quad \sigma_2 = a_{n-2}/a_n, \ \ldots, \ \sigma_n = (-1)^n a_0/a_n.$$

Знаменатель каждого одночлена $\sigma_1^{l_1}\sigma_2^{l_2}\cdots\sigma_n^{l_n} \in h$ равен $a_n^{l_1+\cdots+l_n}$. С другой стороны,

$$c(\sigma_1^{l_1}\sigma_2^{l_2}\cdots\sigma_n^{l_n}) = \alpha_1^{l_1+\cdots+l_n}\alpha_2^{l_2+\cdots+l_n}\cdots\alpha_n^{l_n}.$$

Учитывая это и определитель Вандермонда, получаем

$$\sum l_i = \deg_{\alpha_1}(\sigma_1^{l_1}\sigma_2^{l_2}\cdots\sigma_n^{l_n}) = 2n-2.$$

Таким образом, знаменатель h равен a_n^{2n-2} и поэтому $a_n^{2n-2}\sigma_1^{l_1}\sigma_2^{l_2}\cdots\sigma_n^{l_n}$ – многочлен от $a_0, \ldots a_n$ с целыми коэффициентами. □

Пример. Пусть $n = 2$. Тогда $f = a_0 + a_1 t + a_2 t^2$,

$$D_f = a_2^2(\alpha_2 - \alpha_1)^2 = a_2^2(\alpha_1 + \alpha_2)^2 - 4a_2^2\alpha_1\alpha_2 = a_1^2 - 4a_0a_2.$$

Пример. Пусть $f = t^3 + pt + q$. Тогда из вычисления (‡) получаем

$$D_f = -4p^3 - 27q^2.$$

Действительные корни многочленов

Рассмотрим кубический многочлен $f \in \mathbb{R}[t]$ с действительными коэффициентами. Для простоты мы предположим, что его старший коэффициент равен 1. Имеются следующий возможности.

Многочлен f имеет кратный корень (который должен быть действительным). В этом случае $D_f = 0$.

Многочлен f имеет один действительный корень и два комплексно-сопряженных, т.е.

$$f = (t-\alpha)(t-\beta)(t-\bar{\beta}), \qquad \alpha \in \mathbb{R}, \quad \beta \notin \mathbb{R}, \quad \beta \neq \bar{\beta}.$$

Тогда

$$D_f = (\alpha-\beta)^2(\alpha-\bar{\beta})^2(\beta-\bar{\beta})^2 = \big((\alpha-\beta)\overline{(\alpha-\beta)}\big)^2(2\boldsymbol{i}\,\mathrm{Im}(\beta))^2 = -|\alpha-\beta|^4\,\mathrm{Im}(\beta)^2 < 0.$$

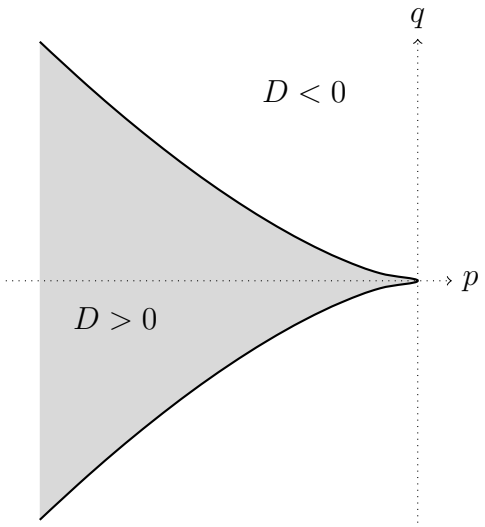
Многочлен f имеет три различных действительных корня. В этом случае $D_f > 0$ согласно (§).

Поскольку исчерпаны все возможности, то доказано следующее.

Утверждение. *Пусть $f \in \mathbb{R}[t]$ – кубический многочлен $f \in \mathbb{R}[t]$ с действительными коэффициентами. Тогда имеют место следующие утверждения:*

- *$D_f > 0$ тогда и только тогда f имеет три различных действительных корня;*
- *$D_f < 0$ тогда и только тогда f имеет один действительный корень и два (различных) комплексно-сопряженных;*
- *$D_f = 0$ тогда и только тогда f имеет кратные корни, в этом случае они действительны.*

Мы можем проиллюстрировать ситуацию на примере многочлена $f = t^3 + pt + q$. Каждому многочлену $f = t^3 + pt + q$ точку (p, q) на плоскости $\mathbb{R}^2$. Тогда многочленам с тремя действительными корнями соответствует серая область:



Аналогичное (но более слабое) утверждение имеет место для многочленов произвольной степени.

Решение уравнений третьей степени

Найдем корни многочлена

$$f = t^3 + pt + q.$$

Будем искать их в виде $t = u + v$. Тогда

$$u^3 + v^3 + (u+v)(3uv + p) + q = 0.$$

Потребуем $3uv + p = 0$. Тогда

$$\begin{cases} u^3 + v^3 = -q \\ uv = -p/3 \end{cases} \Longrightarrow \begin{cases} u^3 + v^3 = -q \\ u^3 v^3 = -p^3/27 \end{cases}$$

u^3 и v^3 находятся из решения $x^2 + qx - p^3/27 = 0$. Таким образом,

$$u^3,\, v^3 = -\frac{q}{2} \pm \sqrt{\frac{q^2}{4} + \frac{p^3}{27}} = -\frac{q}{2} \pm \sqrt{-D_f}.$$

С точностью до перестановки u и v, мы можем считать, что

$$\begin{aligned} u^3 &= -\frac{q}{2} + \sqrt{-D_f}, \\ v^3 &= -\frac{q}{2} - \sqrt{-D_f}. \end{aligned}$$

Отсюда получаем формулу для корней f:

$$\alpha_{1,2,3} = \underbrace{\sqrt[3]{-\tfrac{q}{2} + \sqrt{-D_f}}}_{u} + \underbrace{\sqrt[3]{-\tfrac{q}{2} - \sqrt{-D_f}}}_{v}$$

Она называется *формулой Кардано*. Отметим, что значения корней $\sqrt[3]{\ }$ в этой формуле должны быть выбраны так, чтобы выполнялось равенство $3uv + p = 0$.

Результант

Пусть $\Bbbk$ – поле. Рассмотрим многочлены

$$f = a_n t^n + \cdots a_1 t + a_0, \quad g = b_m t^m + \cdots b_1 t + b_0$$

степеней n и m над этим полем. Предположим, что они полностью разлагаются на линейные множители. Пусть

$$\alpha_1, \ldots, \alpha_n, \qquad \beta_1, \ldots, \beta_m$$

– корни этих многочленов, выписанные с учетом кратностей. Положим

$$R(f,g) := a_n^m b_m^n \prod_{i,j} (\alpha_i - \beta_j). \tag{¶}$$

Предложение. (1) $R(f,g) = (-1)^{nm} R(g,f)$;

(2) $R(f,g) = a_n^m \prod_i g(\alpha_i), \quad R(g,f) = b_m^n \prod_j f(\beta_j)$;

(3) (*основное свойство*) $R(f,g) = 0$ *тогда и только тогда, когда* f *и* g *имеют общий корень* (*мы считаем, что* $a_n b_m \neq 0$).

Доказательство. Свойства (1) и (3) следуют непосредственно из определения. Для доказательства (2) заметим, что

$$g(\alpha_i) = b_m \prod (\alpha_i - \beta_j), \qquad f(\beta_j) = a_n \prod (\beta_j - \alpha_i)$$

и сравним это с определением результанта (¶). □

Свойство (3) дает возможность выяснять имеют ли два многочлена общие множители. Напомним, что другой способ выяснить это основан на алгоритме Евклида.

Теорема. $R(f,f') = (-1)^{n(n-1)/2} a_n D_f$.

Доказательство. По определению результанта (¶) имеем

$$R(f,f') = a_n^{n-1} \prod_i f'(\alpha_i).$$

Пусть $\alpha_1, \ldots, \alpha_n$ – все корни многочлена f. Тогда

$$f = a_n \prod_j (t - \alpha_j).$$

Дифференцируя по правилу Лейбница получаем

$$f' = a_n \sum_{k=1}^{n} \prod_{j \neq k} (t - \alpha_j).$$

Подставим α_i:

$$f'(\alpha_i) = a_n \prod_{j \neq i} (\alpha_i - \alpha_j).$$

Отсюда

$$R(f,f') = a_n^{n-1} a_n^n \prod_i \prod_{j \neq i} (\alpha_i - \alpha_j) = (-1)^{n(n-1)/2} a_n \; a_n^{2n-2} \prod_{j>i} (\alpha_i - \alpha_j)^2.$$ □

Лекция 19

Вычисление результанта. Исключения неизвестных в системах алгебраических уравнений. Неприводимость дискриминанта. Кольцо многочленов инвариантных относительно знакопеременной группы

Вычисление результанта

Теорема. *Рассмотрим многочлены*

$$\begin{aligned} f(t) &= a_n t^n + \cdots + a_1 t + a_0, \\ g(t) &= b_m t^m + \cdots + b_1 t + b_0. \end{aligned}$$

Тогда

$$(*) \qquad R(f,g) = \begin{vmatrix} a_n & a_{n-1} & \cdots & a_0 & & & \\ & a_n & a_{n-1} & \cdots & a_0 & & \\ & & a_n & a_{n-1} & \cdots & a_0 & \\ & & & \ddots & \ddots & \ddots & \ddots \\ & & & & a_n & a_{n-1} & \cdots & a_0 \\ b_m & b_{m-1} & \cdots & b_0 & & & \\ & b_m & b_{m-1} & \cdots & b_0 & & \\ & & b_m & b_{m-1} & \cdots & b_0 & \\ & & & \ddots & \ddots & \ddots & \ddots \\ & & & & b_m & b_{m-1} & \cdots & b_0 \end{vmatrix} \begin{matrix} \\ \\ \\ \\ \\ \} m \\ \\ \\ \\ \\ \} n \end{matrix}$$

Здесь в правой части стоит определитель порядка $n+m$. Пустые места заполняются нулями.

Доказательство. Обозначим матрицу в правой части $(*)$ через A и положим $N := n+m$. Пусть

$$\alpha_1, \ldots, \alpha_n, \qquad \beta_1, \ldots, \beta_m$$

– все корни многочленов f и g, выписанные с учетом кратностей. Таким образом,

$$\begin{aligned} f(t) &= a_n(t-\alpha_1)\cdots(t-\alpha_n), \\ g(t) &= b_m(t-\beta_1)\cdots(t-\beta_m). \end{aligned}$$

Рассмотрим вспомогательный определитель порядка N:

$$|B| = \begin{vmatrix} \beta_1^{N-1} & \beta_2^{N-1} & \cdots & \beta_m^{N-1} & \alpha_1^{N-1} & \alpha_2^{N-1} & \cdots & \alpha_n^{N-1} \\ \beta_1^{N-2} & \beta_2^{N-2} & \cdots & \beta_m^{N-2} & \alpha_1^{N-2} & \alpha_2^{N-2} & \cdots & \alpha_n^{N-2} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \beta_1^2 & \beta_2^2 & \cdots & \beta_m^2 & \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \beta_1 & \beta_2 & \cdots & \beta_m & \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \end{vmatrix}$$

Вычислим $a_n^m b_m^n \, |A| \cdot |B|$ двумя способами.

Первое вычисление. Определитель B отличается от определителя Вандермонда только перестановками строк. Отсюда

$$|B| = \prod_{i<j} (\beta_i - \beta_j) \prod_{i,j} (\beta_j - \alpha_i) \prod_{i<j} (\alpha_i - \alpha_j).$$

Следовательно,

$$(\dagger) \qquad a_n^m b_m^n \, |A| \cdot |B| = |A| \cdot R(g,f) \prod_{i<j} (\beta_i - \beta_j) \prod_{i<j} (\alpha_i - \alpha_j).$$

Второе вычисление. Рассмотрим матрицу $C := A \cdot B$ размера $N \times N$ с элементами $c_{i.j}$. Запишем ее в блочном виде

$$C = A \cdot B = \left(\begin{array}{c|c} \mathrm{I} & \mathrm{II} \\ \hline \mathrm{III} & \mathrm{IV} \end{array} \right) \begin{matrix} m \\ n \end{matrix}$$
$$\qquad\quad m \quad n$$

Вычислим элементы $c_{i,j}$ матрицы C в каждом регионе.

$$\begin{aligned}
&(\mathrm{I}) && c_{i,j} &&= a_n \beta_j^{N-i} + a_{n-1} \beta_j^{N-i-1} + \cdots + a_0 \beta_j^{m-i} = \beta_j^{m-i} f(\beta_j), \\
&(\mathrm{II}) && c_{i,m+j} &&= a_n \alpha_j^{N-i} + a_{n-1} \alpha_j^{N-i-1} + \cdots + a_0 \alpha_j^{m-i} = \alpha_j^{m-i} f(\alpha_j) = 0, \\
&(\mathrm{III}) && c_{m+i,j} &&= b_m \beta_j^{N-i} + b_{m-1} \beta_j^{N-i-1} + \cdots + b_0 \beta_j^{n-i} = \beta_j^{n-i} g(\beta_j) = 0, \\
&(\mathrm{IV}) && c_{m+i,m+j} &&= b_m \alpha_j^{N-i} + b_{m-1} \alpha_j^{N-i-1} + \cdots + b_0 \alpha_j^{n-i} = \alpha_j^{n-i} g(\alpha_j).
\end{aligned}$$

Таким образом,

$$|A| \cdot |B| = |C| = \left| \begin{array}{c|c} C' & 0 \\ \hline 0 & C'' \end{array} \right| = |C'| \cdot |C''|,$$

где C' и C'' – квадратные матрицы размеров $m \times m$ и $n \times n$, соответственно. Согласно приведенным выше вычислениям элементы C' и C'' имеют вид

$$c'_{i,j} = \beta_j^{m-i} f(\beta_j), \qquad c''_{i,j} = \alpha_j^{n-i} g(\alpha_j).$$

Отсюда

$$|A|\cdot|B| = \prod_j f(\beta_j)\begin{vmatrix} \beta_1^{m-1} & \beta_2^{m-1} & \cdots & \beta_m^{m-1} \\ \vdots & \vdots & \cdots & \vdots \\ \beta_1^2 & \beta_2^2 & \cdots & \beta_m^2 \\ \beta_1 & \beta_2 & \cdots & \beta_m \\ 1 & 1 & \cdots & 1 \end{vmatrix} \prod_j g(\alpha_j) \begin{vmatrix} \alpha_1^{n-1} & \alpha_2^{n-1} & \cdots & \alpha_n^{n-1} \\ \vdots & \cdots & \vdots & \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ 1 & 1 & \cdots & 1 \end{vmatrix} =$$

$$= \prod_j f(\beta_j) \prod_{i<j}(\beta_i-\beta_j) \prod_j g(\alpha_j) \prod_{i<j}(\alpha_i-\alpha_j).$$

Окончательно получаем

$$(\ddagger) \qquad a_n^m b_m^n\,|A||B| = R(f,g)R(g,f)\prod_{i<j}(\beta_i-\beta_j)\prod_{i<j}(\alpha_i-\alpha_j).$$

Сравнивая (†) и (‡), получим

$$(\S) \qquad |A|\cdot R(g,f)\prod_{i<j}(\beta_i-\beta_j)\prod_{i<j}(\alpha_i-\alpha_j) = R(f,g)R(g,f)\prod_{i<j}(\beta_i-\beta_j)\prod_{i<j}(\alpha_i-\alpha_j).$$

Теперь, если все элементы α_i и β_j поля $\Bbbk$ различны, то множители $R(g,f)$, $\prod_{i<j}(\beta_i-\beta_j)$ и $\prod_{i<j}(\alpha_i-\alpha_j)$ отличны от нуля и мы можем сократить на них. Получим требуемое равенство $(*)$:

$$|A| = R(f,g).$$

Для того, чтобы доказать равенство $(*)$ в общем случае, можно воспользоваться одним из следующих двух способов.

Первый способ работает для поля комплексных чисел $\mathbb{C}$ (или любого его подполя). Заметим, что $|A|$ и $R(f,g)$ в доказываемом равенстве $(*)$ являются непрерывными функциями от корней $\alpha_1,\ldots,\alpha_n,\beta_1,\ldots,\beta_m \in \mathbb{C}^{n+m}$. Мы доказали, что эти функции равны на открытом плотном множестве $U \subset \mathbb{C}^{n+m}$, полученном из $\mathbb{C}^{n+m}$ удалением $(n+m)(n+m-1)/2$ плоскостей, заданных условиями равенства двух координат. Значит, функции $|A|$ и $R(f,g)$ равны всюду на $\mathbb{C}^{n+m}$.

Второй способ работает для любого поля. Мы рассмотрим α_i и β_j как независимые переменные. Тогда коэффициенты многочленов f и g, а также правая и левая части ($|A|$ и $R(f,g)$) доказываемого равенства $(*)$ являются многочленами от α_i и β_j. Мы можем сократить на общие множители. Получим требуемое равенство $|A| = R(f,g)$. Теперь, если нужно, мы можем подставить вместо α_i и β_j конкретные значения. □

Исключения неизвестных в системах алгебраических уравнений

В качестве приложения результанта рассмотрим следующую задачу. Дана система алгебраических уравнений

$$\begin{cases} f(x,y) = 0 \\ g(x,y) = 0. \end{cases}$$

где f и g – многочлены. Требуется исключить неизвестную y и свести решение системы к нахождению корней многочлена от одной переменной. Для этого рассмотрим f и g как многочлены от y с коэффициентами из $\Bbbk[x]$:

$$\begin{aligned} f &= a_n(x)y^n + a_{n-1}(x)y^{n-1} + \cdots + a_0(x), \\ g &= b_m(x)y^m + b_{m-1}(x)y^{m-1} + \cdots + b_0(x). \end{aligned}$$

Вычислить результант $R = R(f,g)$ по y, который также будет элементом $\Bbbk[x]$, т.е. многочленом от x. Результант $R = R(x)$ зануляется при $x = \alpha$ тогда и только тогда, когда или многочлены $f(\alpha, y)$ и $g(\alpha, y)$ имеют по крайней мере один общий корень или когда оба старших коэффициента $a_n(\alpha)$ и $b_m(\alpha)$ зануляется. Таким образом, решение системы из двух уравнений от двух неизвестных сводится к уравнений от одной переменной: сначала $R(x) = 0$, а потом $f(\alpha, y) = 0$ и $g(\alpha, y) = 0$.

Пример. Пусть дана система

$$\begin{cases} x^3y^2 + x^2y^2 + xy + x + 1 = 0 \\ xy^2 + x^3y + x^2y + y + 2 = 0. \end{cases}$$

Рассмотрим левые части как многочлены от y и запишем результант:

$$R = \begin{vmatrix} x^2+x^3 & x & x+1 & 0 \\ 0 & x^2+x^3 & x & x+1 \\ x & x^3+x^2+1 & 2 & 0 \\ 0 & x & x^3+x^2+1 & 2 \end{vmatrix}$$

Вычисляя, получим

$$R = x^{10} + 4x^9 + 6x^8 + 4x^7 + 6x^6 + 6x^5 - 3x^4 - x^3 + x^2.$$

Заметим, что корень $x = 0$ не дает решения: при $x = 0$ старшие коэффициенты в разложениях f и g по степеням y обращаются в нуль.

Неприводимость дискриминанта

Пусть характеристика основного поля $\Bbbk$ отлична от 2. Рассмотрим множество многочленов

$$\Bbbk[t_1, \ldots, t_n]^{\mathrm{A}_n} := \{f \in \Bbbk[t_1, \ldots, t_n] \mid f^\gamma = f \ \forall \gamma \in \mathrm{A}_n\}$$

не меняющихся при всех четных перестановках переменных. Ясно, что $\Bbbk[t_1, \ldots, t_n]^{\mathrm{A}_n}$ является подкольцом в $\Bbbk[t_1, \ldots, t_n]$ и содержит все симметрические многочлены. Назовем $\Bbbk[t_1, \ldots, t_n]^{\mathrm{A}_n}$ *кольцом многочленов инвариантных относительно знакопеременной группы*. С другой стороны, $\Bbbk[t_1, \ldots, t_n]^{\mathrm{A}_n}$ содержит многочлен Вандермонда

$$\Delta(t_1, \ldots, t_n) := \prod_{i>j}(t_i - t_j) = \begin{vmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^{n-1} \\ 1 & t_2 & t_2^2 & \cdots & t_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & t_n & t_n^2 & \cdots & t_n^{n-1} \end{vmatrix}$$

не являющийся симметрическим. Таким образом,

$$\Bbbk[t_1,\dots,t_n] \supsetneqq \Bbbk[t_1,\dots,t_n]^{\mathrm{A}_n} \supsetneqq \Bbbk[t_1,\dots,t_n]^{\mathrm{S}_n}.$$

При любой нечетной перестановке переменных многочлен $\Delta(t_1,\dots,t_n)$ меняет знак. Поэтому $\Delta^2(t_1,\dots,t_n)$ – симметрический многочлен и поэтому он выражается через элементарные симметрические: существует многочлен $D(x_1,\dots,x_n) \in \Bbbk[x_1,\dots,x_n]$ такой, что

$$\Delta^2(t_1,\dots,t_n) = D(\sigma_1,\dots,\sigma_n).$$

Предложение. *Многочлен $D(x_1,\dots,x_n)$ неприводим.*

Доказательство. Предположим противное, т.е. имеет место разложение в произведение непостоянных многочленов

$$D(x_1,\dots,x_n) = D_1(x_1,\dots,x_n)D_2(x_1,\dots,x_n).$$

Подставим $\sigma_1,\dots,\sigma_n$ в D_1. Получим

$$D_1(\sigma_1,\dots,\sigma_n) = c_1 \prod_{(i,j)\in M_1} (t_i - t_j),$$

где $c_1 \in \Bbbk$, а M_1, M_2 – подмножества в множестве всех пар $M := \{(i,j) \mid n \geq i > j \geq 1\}$ такие, что $M_1 \cup M_2 = M$ и $M_1 \cap M_2 = \varnothing$. Мы можем считать, что $(1,2) \in M_1$, т.е. правая часть последнего равенства содержит множитель $t_1 - t_2$. Для $i \neq j$ рассмотрим подстановку

$$\gamma := \begin{pmatrix} 1 & 2 & \cdots \\ i & j & \cdots \end{pmatrix}$$

Так как D_1 – симметрический многочлен, то $D_1^\gamma = D_1$, а значит правая часть последнего равенства содержит множитель $t_i - t_j$ или $t_j - t_i$, где (i,j) – любая пара различных индексов. Здесь мы пользуемся факториальностью кольца $\Bbbk[t_1,\dots,t_n]$. Мы получаем, что D_1 в кольце $\Bbbk[t_1,\dots,t_n]$ делится на

$$\prod_{i>j}(t_i - t_j) = \Delta.$$

Применяя те же рассуждения к D_2, получаем, что D_2 тоже делится на Δ. Но из соображений степени тогда получается, что $D_k = c_k\Delta$, $c_k \in \Bbbk$. С другой стороны, любая транспозиция $[i,j]$ меняет знак последнего многочлена. Противоречие. □

Следствие. *Дискриминант $D_f(a_n,\dots,a_0)$ многочлена*

$$f = a_n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

с неопределенными коэффициентами является неприводимым многочленом в $\Bbbk[a_n,\dots,a_0]$.

Доказательство. Следует из того, что дискриминант D_f получается из многочлена $D = \Delta^2$ подстановкой в него формул Виета. □

Следствие. *Любой многочлен $f \in \Bbbk[t, \dots, t_n]^{\mathrm{A}_n}$ единственным образом представляется в виде*

$$f = f_1 + \Delta f_2,$$

где f_1 и f_2 – симметрические многочлены.

Доказательство. Пусть γ – любая транспозиция. Положим

$$f_1 := \frac{1}{2}(f + f^\gamma), \qquad g := \frac{1}{2}(f - f^\gamma).$$

Тогда $f = f_1 + g$. Покажем, что многочлен f_1 – симметрический, а многочлен g – *кососимметрический*, т.е. $g^\nu = \mathrm{sgn}(\nu)g$ для любой подстановки $\nu \in \mathrm{S}_n$. Для любой четной подстановки ω подстановка $\omega' := \gamma \circ \omega \circ \gamma$ также четная и $(f^\gamma)^\omega = f^{\omega\circ\gamma}$. Отсюда

$$\begin{aligned} f_1^\omega &= \frac{1}{2}(f^\omega + f^{\omega\circ\gamma}) = \frac{1}{2}(f + f^{\gamma\circ\omega'}) = \frac{1}{2}(f + (f^{\omega'})^\gamma) = \frac{1}{2}(f + f^\gamma) = f_1, \\ g^\omega &= \frac{1}{2}(f^\omega - f^{\omega\circ\gamma}) = \frac{1}{2}(f - f^{\gamma\circ\omega'}) = \frac{1}{2}(f - (f^{\omega'})^\gamma) = \frac{1}{2}(f - f^\gamma) = g. \end{aligned}$$

Любая нечетная подстановка β представляется в виде произведения $\beta = \omega \circ \gamma$ транспозиции и четной: $\beta = \omega \circ \gamma$. Имеем $(f^\gamma)^\beta = f^\omega = f$. Отсюда

$$\begin{aligned} f_1^\beta &= \frac{1}{2}(f^{\omega\circ\gamma} + f) = \frac{1}{2}(f^{\gamma\circ\omega'} + f) = \frac{1}{2}((f^{\omega'})^\gamma + f) = \frac{1}{2}(f^\gamma + f) = f_1, \\ g^\beta &= \frac{1}{2}(f^{\omega\circ\gamma} - f) = \frac{1}{2}(f^{\gamma\circ\omega'} - f) = \frac{1}{2}((f^{\omega'})^\gamma - f) = \frac{1}{2}(f^\gamma - f) = -g. \end{aligned}$$

Таким образом, f_1 – симметрический многочлен и для любой подстановки $\nu \in \mathrm{S}_n$ имеем $g^\nu = \mathrm{sgn}(\nu)g$. Остается доказать, что g делится на Δ.

Так как $\Delta^\nu = \mathrm{sgn}(\nu)\Delta$, то $(g\Delta)^\nu = g\Delta$ для любой подстановки $\nu \in \mathrm{S}_n$, т.е. $g\Delta$ – симметрический многочлен. В кольце $\Bbbk[t, \dots, t_n]^{\mathrm{S}_n}$ многочлен $D = \Delta^2$ делит $(g\Delta)^2 = g^2\Delta^2$. Так как D неприводим и кольцо $\Bbbk[t, \dots, t_n]^{\mathrm{S}_n} \simeq \Bbbk[x_1, \dots, x_n]$ факториально, то он делит и $g\Delta$. Значит, существует симметрический многочлен f_2 такой, что $Df_2 = \Delta^2 f_2 = g\Delta$. Сокращая на Δ, получим равенство $g = \Delta f_2$ в кольце $\Bbbk[t, \dots, t_n]^{\mathrm{A}_n}$. Таким образом, $f = f_1 + \Delta f_2$. Ясно, что это представление единственно. □

Лекция 20

Смежные классы. Теорема Лагранжа. Малая теорема Ферма. Нормальные подгруппы. Свойства. Примеры. Факторгруппы. Теорема о гомоморфизме групп.

Смежные классы и теорема Лагранжа

Определение. Для любых подмножеств A, $B \subset G$ группы G положим

$$AB := \{ab \mid a \in A,\, b \in B\}.$$

В частности, для подгруппы $H \subset G$ и элемента $a \in H$ подмножество

$$aH := \{ah \mid h \in H\}$$

называется *левым смежным классом*. Аналогично определяются правые смежные классы:

$$Ha := \{ha \mid h \in H\}.$$

Множество всех левых смежных классов обозначается G/H. Мощность множества G/H обозначается $[G : H]$ и называется *индексом* подгруппы H.

Замечание. (1) Очевидно, что каждый элемент принадлежит своему смежному классу: $a \in aH$.

(2) Сама подгруппа H является смежным классом, причем как левым, так и правым: $H = 1\,H = H\,1$.

Заметим, что запись смежного класса в виде aH не является единственной:

Лемма. *$aH = a'H$ тогда и только тогда, когда существует элемент $h \in H$ $\quad a' = ah$.*

Доказательство. Если $aH = a'H$, то $a' \in aH$ и тогда $a' = ah$ для некоторого $h \in H$. Обратно, если $a' = ah$ для некоторого $h \in H$, то $a'h' = ahh' \in aH$ и поэтому $a'H \subset aH$. Аналогично доказывается обратное включение. □

Теорема. *Пусть G – группа и пусть $H \subset G$ – любая подгруппа.*

(1) *Группа G является объединением левых смежных классов $aH \in G/H$.*

(2) *Если два левых смежных класса a_1H и a_2H пересекаются, то они совпадают.*

(3) *Все левые смежные классы равномощны.*

Аналогичные утверждения верны для правых смежных классов.

Доказательство. (1) очевидно, поскольку $a \in aH$.

(2) Пусть $a \in a_1H \cap a_2H$. Тогда $a = a_1h_1 = a_2h_2$ для некоторых h_1, $h_2 \in H$. Отсюда $a_2 = a_1(h_1h_2^{-1})$ и $a_1H = a_2H$ по лемме.

(3) Отображение $H \to aH$, $h \mapsto ah$ является биекцией. □

Следствие (теорема Лагранжа)**.** *Если G – конечная группа, то*

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

В частности, порядок подгруппы делит порядок группы.

Доказательство. Следует из теоремы: $G = \cup_{a \in G} aH$, где все смежные классы не пересекаются и число элементов в каждом классе равно $|H|$. □

Следствие. *Порядок элемента делит порядок группы.*

Доказательство. Порядок элемента равен порядку порожденной им подгруппы. □

Следствие. *Если $|G| = n$, то $a^n = 1$ для всех $a \in G$.*

Доказательство. Если $a^m = 1$, то m делится на порядок элемента. □

Следствие. *Группа простого порядка – циклическая.*

Доказательство. Если $|G| = p$ – простое число, то то прядок любого неединичного элемента должен быть равен p. □

Следствие (Малая теорема Ферма)**.** *Пусть p – простое число. Тогда $m^p \equiv m \mod p$, для всех $m \in \mathbb{Z}$.*

Доказательство. Мы знаем, что $\mathbb{Z}/p\mathbb{Z}$ – поле. Поэтому множество $(\mathbb{Z}/p\mathbb{Z})^*$ ненулевых элементов – группа порядка $p-1$. Следовательно, $m^{p-1} \equiv 1 \mod p$, для всех целых чисел m, не делящихся на p. □

Примеры. (1) Пусть $G = H$. Тогда G/H состоит из одного элемента.

(2) Пусть $G = \mathrm{S}_n$ и $H = \mathrm{A}_n$. Тогда имеется ровно 2 смежных класса: четные и нечетные подстановки.

(3) Пусть $G = \mathrm{GL}_n(\Bbbk)$ и $H = \mathrm{SL}_n(\Bbbk)$. Тогда смежные классы G/H – матрицы с фиксированным определителем.

(4) Пусть $G = \mathbb{C}^*$ и $H = \{z \mid |z| = 1\}$. Тогда смежные классы G/H – комплексные числа с фиксированным модулем.

Нормальные подгруппы

Определение. Подгруппа $H \subset G$ группы G называется *нормальной* (обозначается $H \lhd G$), если $gH = Hg \quad$ для всех $g \in G$.

Примеры. (1) В абелевой группе любая подгруппа нормальна.

(2) В любой группе G имеются тривиальные нормальные подгруппы G и $\{1\}$.

(3) $\mathrm{SL}_n(\Bbbk) \lhd \mathrm{GL}_n(\Bbbk)$.

(4) $\mathrm{A}_n \lhd \mathrm{S}_n$.

Утверждение. *Подгруппа индекса* 2 *нормальна.*

Доказательство. В этом случае имеется ровно два левых смежных класса, причем один из них – сама подгруппа H:

$$G/H = \{H,\, aH\}, \quad \text{где } a \in G \setminus H.$$

Аналогично, имеется ровно два правых смежных класса:

$$\{H,\, Ha\}, \quad \text{где } a \in G \setminus H.$$

Но тогда $aH = G \setminus H = Ha$. □

Предложение. *Следующие условия эквивалентны:*

(1) $H \lhd G$;

(2) $gHg^{-1} \subset H \quad$ *для всех* $g \in G$.

Доказательство. Предположим, что $H \lhd G$, т.е. $gH = Hg \quad$ для всех $g \in G$ и пусть $h \in H$. Тогда $gh \in Hg$. Следовательно, $gh = h'g$ для некоторого $h' \in H$ и поэтому $ghg^{-1} = h' \in H$. Это означает, что $gHg^{-1} \subset H$.

Предположим, что $gHg^{-1} \subset H$. Пусть $g \in G$ и пусть $gh \in gH$, где $h \in H$. Тогда $ghg^{-1} = h' \in H$. Следовательно, $gh = h'g \in Hg$ и поэтому $gH \subset Hg$. Обратное включение доказывается аналогично. □

Замечание. На самом деле, в условиях выше верно равенство $gHg^{-1} = H$. (Докажите самостоятельно).

Определение. Говорят, что элементы a и a' группы G *сопряжены*, если существует $x \in G$ такой, что $a' = xax^{-1}$.

Не это сопряжение путайте с комплексным сопряжением!

Несложно проверить, что отношение сопряженности является отношением эквивалентности. Таким образом, группа G разбивается на непересекающееся объединение *классов сопряженных элементов*.

Следствие. *Подгруппа $H \subset G$ является нормальной тогда и только тогда, когда она составлена из классов сопряженных элементов.*

Зафиксируем элемент группы $g \in g$ и рассмотрим отображение

$$(*) \qquad \Phi_g : G \longrightarrow G, \qquad x \longmapsto gxg^{-1}.$$

Утверждение. *Отображение Φ_g, заданное формулой $(*)$, является автоморфизмом группы G.*

Доказательство. Для любых x, $y \in G$ имеем

$$\Phi_g(xy) = gxyg^{-1} = gxg^{-1}gyg^{-1} = \Phi_g(x)\Phi_g(y).$$

Следовательно, Φ_g является гомоморфизмом. Далее,

$$\Phi_g \circ \Phi_{g^{-1}}(x) = g(g^{-1}xg)g^{-1} = x.$$

Следовательно, $\Phi_{g^{-1}}$ – обратное отображение к Φ_g. Таким образом, Φ_g – изоморфизм. □

Автоморфизмы вида $(*)$ называются внутренними.

Следствие. *Подгруппа $H \subset G$ является нормальной тогда и только тогда, когда она переходит в себя при всех внутренних автоморфизмах:*

$$\Phi_g(H) \subset H$$

для любого элемента $g \in G$.

Замечание. (1) В абелевой группе все внутренние автоморфизмы являются тождественными. Обратное тоже верно: если в группе все внутренние автоморфизмы являются тождественными, то группа абелева.

(2) С другой стороны, в абелевой группе порядка порядка большего 2 имеются нетривиальные автоморфизмы. (например, отображение $a \longmapsto a^{-1}$). Это показывает, что далеко не все автоморфизмы группы являются внутренними.

Определение. *Центром* группы G называется подмножество

$$Z(G) := \{z \in G \mid gz = zg \quad \forall g \in G\}.$$

Утверждение. *Центр $Z(G)$ группы G является нормальной подгруппой в G. Более того, любая подгруппа $H \subset Z(G)$ нормальна в G.*

Доказательство. Пусть $z_1, z_2 \in Z(G)$. Тогда для любого элемента $g \in G$ имеем $z_i g = g z_i$. Следовательно,

$$(z_1z_2)g = z_1(z_2g) = z_1(gz_2) = (z_1g)z_2) = (gz_1)z_2) = g(z_1z_2),$$

Аналогично, для любых элементов $z \in Z(G)$ и $g \in G$ имеем $zg = gz$. Умножая обе части справа и слева на z^{-1} получим $gz_1^1 = z_1^1 g$. Таким образом, $Z(G)$ – подгруппа в G. Если $H \subset Z(G)$ – подгруппа, то для любых элементов $z \in H$ и $g \in G$ имеем

$$\Phi_g(z) = gzg^{-1} = zgg^{-1} = z \in H.$$

Следовательно, $H \lhd G$. □

Пример. (1) Центр $Z(G)$ группы G совпадает со всей группой G тогда и только тогда, когда она абелева.

(2) Центр $Z(\mathrm{S}_n)$ симметрической группы S_n тривиален.

(3) Центр $Z(\mathrm{GL}_n(\Bbbk))$ полной линейной группы $\mathrm{GL}_n(\Bbbk)$ состоит из скалярных матриц. Центр $Z(\mathrm{SL}_n(\Bbbk))$ специальной линейной группы $\mathrm{SL}_n(\Bbbk)$ состоит из скалярных матриц с определителем 1.

Утверждение. *Подгруппа $H \subset G$ порядка 2 нормальна тогда и только тогда, когда H содержится в центре.*

Доказательство. Пусть $H = \{1, h\}$. Для любого элемента $g \in G$ имеем $\Phi_g(h) \in H$ и $\Phi_g(h) \neq 1$. Следовательно, $\Phi_g(h) = h$. Тогда $ghg^{-1} = h$ и $gh = hg$. □

В частности, отсюда следует, что группа S_n не содержит нормальных подгрупп порядка 2. С другой стороны, любая транспозиция порождает подгруппу порядка 2 в S_n. Все это подгруппы не являются нормальными.

Факторгруппы

Пусть H – нормальная подгруппа группы G (в мультипликативной записи). Напомним, что через G/H мы обозначаем множество всех левых смежных классов. Определим умножение смежных классов следующим образом:

$$(aH) \cdot (bH) = (ab)H. \tag{†}$$

Лемма. *Определенное выше умножение не зависит от способа записи смежных классов.*

Доказательство. Пусть $aH = a'H$ и $bH = b'H$. Тогда $a' = ah_1$ и $b' = bh_2$ для некоторых $h_1, h_2 \in H$. Имеем

$$a'b' = (ah_1)(bh_2) = (ab)(b^{-1}h_1b)h_2,$$

где $b^{-1}h_1b \in H$ (поскольку $H \lhd G$). Следовательно, $a'b' = abh$ для $h := (b^{-1}h_1b)h_2 \in H$ и поэтому $(aH) \cdot (bH) = (ab)H$. □

Заметим, что доказательство леммы существенным образом опиралось на нормальность подгруппы H. На самом деле, верно и обратное: если формула (†) задает корректное умножение смежных классов, то подгруппа H является нормальной.

Предложение. *G/H является группой относительно определенного выше умножения.*

Доказательство. По определению имеем

$$(aH \cdot bH) \cdot cH = (ab)cH = a(bc)H = aH \cdot (bH \cdot cH).$$

Это доказывает ассоциативность операции. Нейтральным элементом в G/H является тривиальный смежный класс $1H = H$, а обратным к элементу aH является элемент $a^{-1}H$. □

Определение. Построенная группа G/H называется факторгруппой группы G по нормальной подгруппе H.

Непосредственно из определения умножения (†) следует, что отображение $\pi : G \to G/H$, $a \mapsto aH$ является гомоморфизмом групп.

Если G – конечная группа, то из теоремы Лагранжа следует, что

$$|G/H| = [G : H] = \frac{|G|}{|H|}.$$

Пример. Пусть $G := \mathbb{C}^*$ и пусть $H := \{z \mid |z| = 1\}$. Каждый элемент $z \in \mathbb{C}^*$ единственным образом записывается в виде $z = \alpha z_0$, где $\alpha \in \mathbb{R}_{>0}$, $z_0 \in H$. Поэтому каждый смежный класс G/H можно однозначно записать в виде αH, $\alpha \in \mathbb{R}_{>0}$. Следовательно, $G/H \simeq \mathbb{R}_{>0}$.

Теорема о гомоморфизме групп

Замечание. Пусть группа G порождается элементами $a_1, \ldots, a_n$. Если для двух гомоморфизмов $\varphi_1 : G \to G_1$ и $\varphi_2 : G \to G_1$ имеем $\varphi_1(a_i) = \varphi_2(a_i)$ для всех i, то $\varphi_1 = \varphi_2$.

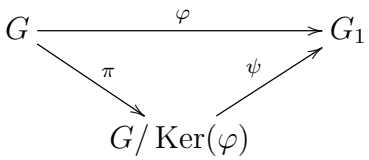
Теорема. *Пусть $\varphi : G \to G_1$ – гомоморфизм групп. Тогда*

(1) $\operatorname{Ker}(\varphi) \lhd G$.

(2) *Образ $\varphi(G)$ является подгруппой в G_1 и имеется естественный изоморфизм*

$$\psi : G/\operatorname{Ker}(\varphi) \to \varphi(G)$$

такой, что $\varphi = \psi \circ \pi$, где $\pi : G \to G/\operatorname{Ker}(\varphi)$ – естественный гомоморфизм на факторгруппу. В этом случае говорят, что диаграмма



коммутативна.

Доказательство. Положим $H := \operatorname{Ker}(\varphi)$. Пусть $a \in H$. Тогда $\varphi(a) = 1$. Для любого $b \in G$ имеем $\varphi(bab^{-1}) = \varphi(b)\varphi(a)\varphi(b)^{-1}) = \varphi(b)\varphi(b)^{-1}) = 1$. Следовательно, $bab^{-1} \in H$. Таким образом, H – нормальная подгруппа.

Определим ψ следующим образом: $\psi(aH) = \varphi(a)$. Во-первых проверяем, что это определение корректно. Пусть $aH = a'H$. Тогда $a' = ah$ для некоторого $h \in H$. Отсюда

$$\psi(a'H) = \varphi(ah) = \varphi(a)\varphi(h) = \varphi(a) = \psi(aH).$$

Далее проверяем, что ψ – гомоморфизм:

$$\psi(aH \cdot bH) = \psi\big((ab)H\big) = \varphi(ab) = \varphi(a)\varphi(b) = \psi(aH)\psi(bH).$$

Для проверки инъективности гомоморфизма достаточно проверить, что его ядро тривиально. Пусть $\psi(aH) = 1$. Это означает, что $\varphi(a) = 1$ и $a \in H$. Отсюда $aH = H$. Следовательно, гомоморфизм ψ инъективен. Наконец, ψ сюръективно по построению. $\square$

Примеры. Применяя теорему, получаем следующие изоморфизмы:

(1) $\mathrm{S}_n \,/\, \mathrm{A}_n \simeq \{\pm 1\}$;

(2) $\mathrm{GL}_n(\Bbbk)/\,\mathrm{SL}_n(\Bbbk) \simeq \Bbbk^*$;

(3) $\mathbb{R}^*/\{\pm 1\} \simeq \mathbb{R}_{>0}$;

(4) $\mathbb{C}^*/\boldsymbol{\mu}_n \simeq \mathbb{C}^*$;

(5) $\mathrm{O}_n(\mathbb{R})/\,\mathrm{SO}_n(\mathbb{R}) \simeq \{\pm 1\}$.

Лекция 21

Идеалы в кольцах. Примеры. Факторкольца. Теорема о гомоморфизме колец. Факторпространства. Факторалгебры.

Идеалы

Определение. Подгруппа $I \subset R$ аддитивной группы кольца называется (двусторонним) *идеалом*, если $aI \subset I$ и $Ia \subset I$ для любого $a \in R$.

Определение. Подгруппа $I \subset R$ аддитивной группы кольца называется *левым* (соотв., *правым*) *идеалом*, если $aI \subset I$ (соотв., $Ia \subset I$) для любого $a \in R$.

Примеры. (1) В каждом кольце имеется идеал $(0) := \{0\}$, который называется *нулевым*. Все кольцо R – также идеал.

(2) Множество всех четных чисел – идеал в кольце $\mathbb{Z}$. Более общий пример, множество всех целых чисел $(n) = n\mathbb{Z}$, делящихся на n – идеал в $\mathbb{Z}$.

(3) В кольце матриц $\mathrm{Mat}_n(\Bbbk)$ аддитивная подгруппа

$$I := \left\{ \begin{pmatrix} 0 & * & \cdots & * \\ \cdots & \cdots & \cdots & \cdots \\ 0 & * & \cdots & * \end{pmatrix} \right\}$$

(матриц с нулевым левым столбцом) идеалом не является. Однако она является левым идеалом: для всех $A \in I$, для всех $B \in \mathrm{Mat}_n(\Bbbk)$ $BA \in I$.

На самом деле, можно показать, что в кольце $\mathrm{Mat}_n(\Bbbk)$ нет идеалов, кроме нулевого и единичного. Кольца с таким свойством называются *простыми*.

Определение. Пусть R – коммутативное ассоциативное кольцо и пусть $a_1, \ldots, a_n \in R$. Множество

$$(a_1, \ldots, a_n) := \{a_1 b_1 + \cdots a_n b_n \mid b_i \in R\}$$

является идеалом в R. Он называется *идеалом, порожденным элементами* $a_1, \ldots, a_n$. Это наименьший идеал, содержащий $a_1, \ldots, a_n$. Идеал, порожденный одним элементом, т. е. идеал вида

$$(a) := \{ab \mid b \in R\}$$

называется *главным*. Говорят, что R – *кольцо главных идеалов*, если в нем каждый идеал является главным. Если R – кольцо с единицей, то $(1) = R$. Этот идеал называется *единичным*.

Пример. Пусть R – коммутативное ассоциативное кольцо с единицей. Если некоторый идеал I содержит обратимый элемент, то он является единичным: в этом случае из $a \in I$ следует, что $1 = aa^{-1} \in I$. Отсюда $I = (1) = R$.

В частности, любой идеал в поле является или нулевым или единичным.

Пример. (1) В кольце целых чисел $\mathbb{Z}$ любой идеал $I \subset \mathbb{Z}$ является главным. Если $I \neq (0)$, то он порождается наименьшим по модулю элементом $n \in I \setminus \{0\}$. Действительно, если $m \in I$ – другой элемент, то выполняя деление с остатком получим $m = qn + r$, где $0 \leq r < |n|$. Но тогда $r = n - mq \in I$ – меньший по модулю элемент идеала I. По нашему предположению $r = 0$ и, таким образом, $m = nq$. Следовательно, $I = (n)$. Мы доказали, что $\mathbb{Z}$ – кольцо главных идеалов.

(2) Аналогично, в кольце многочленов от одной переменной $\Bbbk[t]$ над полем любой идеал $I \subset \Bbbk[t]$ является главным. Если $I \neq (0)$ и $I \neq (1)$, то он порождается многочленом наименьшей степени $f \in I$. Действительно, как и выше, если $g \in I$ – другой элемент, то выполняя деление с остатком получим $g = qf + r$, где $\deg(r) < \deg(f)$ или $r = 0$. Но тогда $r = g - qf \in I$. По нашему предположению $r = 0$ и, таким образом, $g = qf$. Следовательно, $I = (f)$. Мы доказали, что $\Bbbk[t]$ – также кольцо главных идеалов.

Аналогичные соображения показывают, что любое евклидово кольцо является кольцом главных идеалов.

(3) В кольце формальных степенных рядов от одной переменной $\Bbbk[[t]]$ над полем любой идеал является главным. На самом деле, любой неединичный идеал в $\Bbbk[[t]]$ имеет вид (t^n) для некоторого n.

(4) В кольце $\mathscr{O}$ сходящихся в точке 0 степенных рядов от одной переменной любой идеал также является главным.

(5) Любой элемент $(\alpha_1, \ldots, \alpha_n)$ n-мерного пространства $\Bbbk^n$ определяет идеал

$$(t_1 - \alpha_1, \ldots, t - \alpha_n)$$

в кольце многочленов $\Bbbk[t_1, \ldots, t_n]$. При $n \geq 2$ такой идеал главным не является. Можно показать, что идеал $(t_1 - \alpha_1, \ldots, t - \alpha_n)$ совпадает с идеалом

$$\{f \in \Bbbk[t_1, \ldots, t_n] \mid f(\alpha_1, \ldots, \alpha_n) = 0\},$$

состоящим из многочленов, обращающихся в нуль в точке $(\alpha_1, \ldots, \alpha_n)$.

(6) Пусть $R = \mathrm{C}[a, b]$ – кольцо вещественнозначных непрерывных функций на отрезке. Функции, обращающиеся в нуль а некоторой точке образуют идеал

$$I_c := \{f \in \mathrm{C}[a, b] \mid f(c) = 0\}.$$

Этот идеал не является главным. Более того, он не может быть порожден никаким конечным множеством своих элементов.

Факторкольца

Пусть I – идеал кольца R. На факторгруппе R/I аддитивной группы определим умножение по правилу

$$(*) \qquad (a+I)(b+I) = ab + I.$$

Утверждение. *Определенное формулой* $(*)$ *умножение не зависит от вида записи смежных классов.*

Доказательство. Пусть $a + I = a' + I$ и $b + I = b' + I$. Тогда $a' = a + c$ и $b' = b + d$ для некоторых c, $d \in I$. Отсюда

$$a'b' - ab = ad + cb + cd \in I$$

и поэтому $a'b' + I = ab + I$. □

Таким образом, на факторгруппе R/I корректно определено умножение. Оно удовлетворяет свойствам дистрибутивности:

$$\big((a+I)+(b+I)\big)(c+I) = \big((a+b)+I\big)(c+I) = (a+b)c + I =$$
$$= (ac+bc) + I = (ac+I) + (bc+I) = (a+I)(c+I) + (bc+I)(c+I).$$

Следовательно, R/I является кольцом. По построению отображение

$$\pi : R \longrightarrow R/I, \qquad a \longmapsto a + I$$

является гомоморфизмом колец.

Определение. Кольцо R/I с умножением, заданным формулой $(*)$ называется *факторкольцом* кольца R по идеалу I.

С частным случаем факторколец мы уже сталкивались: кольцо вычетов $\mathbb{Z}/n\mathbb{Z}$ является факторкольцом кольца целых чисел $\mathbb{Z}$ по идеалу $n\mathbb{Z}$.

Замечание. Легко видеть, что

- кольцо R ассоциативно, то R/I ассоциативно,
- кольцо R коммутативно, то R/I коммутативно,
- кольцо R имеет единицу 1 и $1 \notin I$, то R/I имеет единицу.

Теорема о гомоморфизме колец

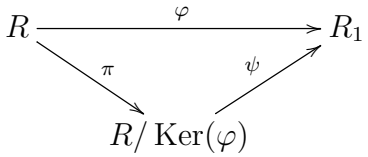
Теорема. *Пусть $\varphi : R \to R_1$ – гомоморфизм колец. Тогда*

(1) $\operatorname{Ker}(\varphi)$ – *идеал в* R.

(2) *Образ $\varphi(R)$ является подкольцом в R_1. Имеется естественный изоморфизм*

$$\psi : R/\operatorname{Ker}(\varphi) \longrightarrow \varphi(R)$$

такой, что $\varphi = \psi \circ \pi$, где $\pi : R \to R/\operatorname{Ker}(\varphi)$ – естественный гомоморфизм на факторкольцо. Таким образом, диаграмма



коммутативна.

Доказательство. Положим $I := \operatorname{Ker}(\varphi)$. Воспользуемся теоремой о гомоморфизме групп. Из нее следует, что I – подгруппа аддитивной группы R и имеется естественный изоморфизм групп $\psi : R/I \to \varphi(R)$, $\psi(a + I) = \varphi(a)$. Остается доказать, что ψ – гомоморфизм колец:

$$\psi((a + I)(b + I)) = \psi(ab + I) = \varphi(ab) = \varphi(a)\varphi(b) = \psi(a + I)\psi(b + I). \qquad \square$$

Примеры.

- Рассмотрим гомоморфизм

$$\varphi : \mathrm{C}[a, b] \longrightarrow \mathbb{R}, \qquad f \longmapsto f(c).$$

Он сюръективен и $\operatorname{Ker}\varphi = I_C$. Следовательно, $\mathrm{C}[a, b]/I_c \simeq \mathbb{R}$.

- Рассмотрим гомоморфизм

$$\varphi : \mathbb{R}[t] \longrightarrow \mathbb{C}, \qquad f \longmapsto f(\boldsymbol{i}).$$

Он сюръективен и $\operatorname{Ker}\varphi = (t^2 + 1)$. Следовательно, $\mathbb{R}[t]/(t^2 + 1) \simeq \mathbb{C}$.

Факторпространства

Пусть V – векторное пространство над полем $\Bbbk$ и $W \subset V$ – подпространство. На факторгруппе V/W аддитивной группы определим умножение на скаляры

$$(\dagger) \qquad \lambda(\mathbf{v} + W) = (\lambda\mathbf{v}) + W, \qquad \lambda \in \Bbbk.$$

Как и в конструкциях выше, определенное формулой (†) умножение не зависит от вида записи смежного класса. Действительно, пусть $\mathbf{v} + W = \mathbf{v}' + W$. Тогда $\mathbf{v}' = \mathbf{v} + \mathbf{w}$ для некоторого $\mathbf{w} \in W$. Отсюда

$$\lambda\mathbf{v}' - \lambda\mathbf{v} = \lambda(\mathbf{v}' - \mathbf{v}) = \lambda\mathbf{w} \in W$$

и поэтому $(\lambda\mathbf{v}') + W = (\lambda\mathbf{v}) + W$.

Таким образом, на факторгруппе V/W корректно определено умножение на скаляры. Оно удовлетворяет аксиомам векторного пространства:

$$(1) \qquad \begin{aligned}&(\lambda+\mu)(\mathbf{v}+W) = (\lambda+\mu)\mathbf{v} + W = \\ &= (\lambda\mathbf{v} + \mu\mathbf{v}) + W = \lambda\mathbf{v} + W + \mu\mathbf{v} + W = \lambda(\mathbf{v}+W) + \mu(\mathbf{v}+W).\end{aligned}$$

$$(2) \qquad \begin{aligned}&\lambda\big((\mathbf{v}_1+W) + (\mathbf{v}_2+W)\big) = \lambda\big((\mathbf{v}_1+\mathbf{v}_2)+W\big) = \\ &= (\lambda(\mathbf{v}_1+\mathbf{v}_2)) + W = (\lambda\mathbf{v}_1 + \lambda\mathbf{v}_2) + W = (\lambda\mathbf{v}_1 + W) + (\lambda\mathbf{v}_2 + W),\end{aligned}$$

$$(3) \qquad (\lambda\mu)(\mathbf{v}+W) = (\lambda\mu)\mathbf{v} + W = \lambda(\mu\mathbf{v}) + W = \lambda(\mu\mathbf{v}+W) = \lambda(\mu(\mathbf{v}+W)).$$

$$(4) \qquad 1\,(\mathbf{v}+W) = 1\,\mathbf{v} + W = \mathbf{v} + W.$$

Следовательно, V/W является пространством. По построению отображение

$$\pi : V \longrightarrow V/W, \qquad \mathbf{v} \longmapsto \mathbf{v} + W$$

является линейным отображением пространств.

Предположим, что наше пространство V конечномерно. Выберем базис $\mathbf{e}_1, \ldots, \mathbf{e}_m$ в W и дополним его до базиса $\mathbf{e}_1, \ldots, \mathbf{e}_n$ всего пространства V. Тогда для $\pi(\mathbf{e}_i)$ – нулевые векторы в V/W для $i = 1, \ldots, m$. С другой стороны, можно показать, что векторы $\pi(\mathbf{e}_{n-m}), \ldots, \pi(\mathbf{e}_n)$ линейно независимы, и, следовательно, образуют базис в V/W. В частности, отсюда получаем, что

$$\dim(V/W) = \dim(V) - \dim(W).$$

Определение. Пространство V/W с умножением, заданным формулой $(*)$ называется *факторпространством* пространства V по подпространству W.

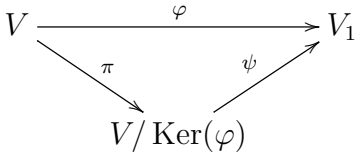
Теорема. *Пусть $\varphi : V \to V_1$ – линейное отображение пространств. Тогда*

(1) $\operatorname{Ker}(\varphi)$ – *подпространство в V.*

(2) *Образ $\varphi(V)$ является подпространством в V_1. Имеется естественный изоморфизм*

$$\psi : V/\operatorname{Ker}(\varphi) \longrightarrow \varphi(V)$$

такой, что диаграмма



коммутативна, где $\pi : V \to V/\operatorname{Ker}(\varphi)$ – естественный гомоморфизм на факторпространство.

Доказательство. Положим $W := \mathrm{Ker}(\varphi)$. Воспользуемся теоремой о гомоморфизме групп. Из нее следует, что W – подгруппа аддитивной группы V и имеется естественный изоморфизм групп $\psi : V/W \to \varphi(V)$, $\psi(\mathbf{v}+W) = \varphi(\mathbf{v})$. Остается доказать, что ψ – линейное отображение пространств:

$$\psi\big(\lambda(\mathbf{v}+W)\big) = \psi(\lambda\mathbf{v}+W) = \varphi(\lambda\mathbf{v}) = \lambda\varphi(\mathbf{v}) = \lambda\psi(\mathbf{v}+W). \qquad \square$$

Пример. Пусть V – конечномерное векторное пространство с базисом $\mathbf{e}_1, \ldots, be_n$ и пусть $U \subset V$ – подпространство, порожденное векторами $\mathbf{e}_1, \ldots, be_m$, $m \leq n$. Рассмотрим проекцию V на U, т.е. линейное отображение, заданное формулой

$$\varphi : (x_1, \ldots, x_m, \ldots, x_n) \longmapsto (x_1, \ldots, x_m, 0, \ldots, 0).$$

Ядром отображения будет подпространство W, порожденное векторами $\mathbf{e}_{m+1}, \ldots, be_n$. Следовательно, $V/W \simeq U$.

Факторалгебры

Пусть A – алгебра над полем $\Bbbk$. *Идеалом в алгебре* называется подмножество в $I \subset A$ такое, что

- I – идеал в A, где A рассматривается как кольцо,
- I – подпространство в A, где A рассматривается как векторное пространство.

Пусть I – идеал алгебры A. Согласно проверенным выше фактам на факторгруппе A/I корректно определено умножение элементов между собой и умножение элементов на скаляры, причем A/I является одновременно кольцом и векторным пространством. Для любых $\lambda \in \Bbbk$ и $\mathbf{a}, \mathbf{b} \in A$ имеем

$$\begin{aligned}\lambda\big((\mathbf{a}+I)\cdot(\mathbf{b}+I)\big) = \lambda\big((\mathbf{a}\cdot\mathbf{b})+I\big) = \lambda(\mathbf{a}\cdot\mathbf{b})+I &= (\lambda\mathbf{a})\cdot\mathbf{b}+I = (\lambda\mathbf{a}+I)\cdot(\mathbf{b}+I),\\ &= \lambda\mathbf{a}\cdot(\lambda\mathbf{b})+I = (\mathbf{a}+I)\cdot(\lambda\mathbf{b}+I).\end{aligned}$$

Следовательно, A/I является алгеброй. По построению отображение

$$\pi : A \longrightarrow A/I, \qquad \mathbf{a} \longmapsto \mathbf{a}+I$$

является гомоморфизмом алгебр.

Определение. Построенная выше алгебра A/I называется *факторалгеброй* алгебры A по идеалу I.

Замечание. Легко видеть, что

- если алгебра A ассоциативна, то и A/I факторалгебра ассоциативна,
- если алгебра A коммутативна, то и A/I факторалгебра коммутативна,
- если алгебра A имеет единицу 1 и $1 \notin I$, то A/I имеет единицу.

Следующая теоремы – непосредственное следствие теорем о гомоморфизмах колец и векторных пространств.

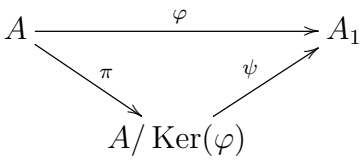
Теорема (теорема о гомоморфизме алгебр)**.** *Пусть* $\varphi : A \to A_1$ *– гомоморфизм алгебр. Тогда*

(1) $\operatorname{Ker}(\varphi)$ *– идеал в* A*.*

(2) *Образ* $\varphi(A)$ *является подалгеброй в* A_1*. Имеется естественный изоморфизм алгебр*

$$\psi : A/\operatorname{Ker}(\varphi) \longrightarrow \varphi(A)$$

такой, что диаграмма



коммутативна, где $\pi : A \to A/\operatorname{Ker}(\varphi)$ *– естественный гомоморфизм на факторалгебру.*

Лекция 22

Простые поля. Расширения полей.Теорема о башне полей. Присоединение к полю корня неприводимого многочлена. Поле разложения многочлена. Алгебраическое замыкание поля.

Лекция отменена. Отложено

Простые поля

Определение. Поле, не содержащее ни одного собственного подполя, называется *простым полем.*

Каждое поле $\Bbbk$ содержит единственное простое поле – пересечение всех подполей в $\Bbbk$. Примерами простых полей являются поле рациональных чисел $\mathbb{Q}$ и поля вычетов $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ по простому модулю p. Ниже мы докажем и обратное.

Напомним, что для любого кольца R с единицей корректно определено отображение

$$\phi : \mathbb{Z} \longrightarrow R, \qquad n \longmapsto n \cdot 1.$$

и это отображение является гомоморфизм колец.

Утверждение. *Пусть $\Bbbk$ – поле.*

(1) *Если* $\operatorname{char}(\Bbbk) = 0$, *то гомоморфизм ϕ инъективен.*

(2) *Если* $\operatorname{char}(\Bbbk) = p > 0$, *то* $\operatorname{Ker}(\phi) = p\,\mathbb{Z}$.

Доказательство. Если $\operatorname{char}(\Bbbk) = 0$, то $\phi(m) = m \cdot 1 \neq 0$ для любого $m \in \mathbb{Z}$. Следовательно, $\operatorname{Ker}(\phi) = (0)$.

Предположим, что $\operatorname{char}(\Bbbk) = p > 0$. Тогда p – порядок единицы в аддитивной группе поля. В частности, $p{\cdot}1 = 0$. Следовательно, $p \in \operatorname{Ker}(\phi)$. Обратно, пусть $m \in \operatorname{Ker}(\phi)$. Тогда $m \cdot 1 = 0$ и поэтому m делится на порядок единицы p. Таким образом, $\operatorname{Ker}(\phi) = p\mathbb{Z}$. $\square$

Теорема. *Любое простое поле $\Bbbk$ изоморфно $\mathbb{Q}$ или $\mathbb{F}_p$.*

Доказательство. Рассмотрим сначала случай $\operatorname{char}(\Bbbk) = p > 0$. Из утверждения и теоремы о гомоморфизме получаем, что $\phi(\mathbb{Z}) \simeq \mathbb{Z}/\operatorname{Ker}(\phi) = \mathbb{F}_p$ является подполем в $\Bbbk$, а так как $\Bbbk$ – простое, то $\phi(\mathbb{Z})$ совпадает с $\Bbbk$.

Теперь рассмотрим случай $\mathrm{char}(\Bbbk) = 0$. Тогда гомоморфизм ϕ инъективен и $Z := \phi(\mathbb{Z})$ – подкольцо в $\Bbbk$, изоморфное $\mathbb{Z}$. Элементы вида a/b, $a, b \in Z$, $b \neq 0$ образуют подполе $\Bbbk_0$ в $\Bbbk$. Так как поле $\Bbbk$ – простое, то $\Bbbk_0$ совпадает с $\Bbbk$. Тогда $\Bbbk$ должно совпадать с полем частных $\mathrm{Frac}(Z)$. Так как $Z \simeq \mathbb{Z}$, то $\Bbbk \simeq \mathrm{Frac}(Z) \simeq \mathrm{Frac}(\mathbb{Z}) \simeq \mathbb{Q}$. □

Таким образом, любое поле получается расширением поля $\mathbb{F}_p$ или $\mathbb{Q}$.

Расширения полей

Пусть $\mathbb{K}/\Bbbk$ – любое расширение полей. Заметим, что $\mathbb{K}$ является векторным пространством над $\Bbbk$. Расширение называется конечным, если $\mathbb{K}$ конечномерно как векторное пространство над $\Bbbk$. В этом случае размерность этого векторного пространства называется степенью расширения и обозначается $[K : \Bbbk]$.

Для элементов $\theta_1, \ldots, \theta_n \in \mathbb{K}$ обозначим через $\Bbbk[\theta_1, \ldots, \theta_n]$ (соответственно, через $\Bbbk(\theta_1, \ldots, \theta_n)$) – наименьшее подкольцо (соответственно подполе) в $\mathbb{K}$, содержащее $\Bbbk$ и все $\theta_1, \ldots, \theta_n$. Ясно, что

$$\Bbbk[\theta_1, \ldots, \theta_n] = \Big\{ \sum \alpha_{k_1,\ldots,k_n} \theta_1^{k_1} \cdots \theta_n^{k_n} \mid \alpha_{k_1,\ldots,k_n} \in \Bbbk,\ k_j \geq 0 \Big\},$$

$$\Bbbk(\theta_1, \ldots, \theta_n) = \Big\{ \tfrac{\beta}{\gamma} \mid \beta, \gamma \in \Bbbk[\theta_1, \ldots, \theta_n], \quad \gamma \neq 0 \Big\}.$$

Определение. Говорят, что элемент $\theta \in \mathbb{K}$ *алгебраичен* над $\Bbbk$, если существует ненулевой многочлен $f \in \Bbbk[t]$, для которого $f(\theta) = 0$. В противном случае элемент θ называется *трансцендентным* над $\Bbbk$. Расширение полей $\mathbb{K}/\Bbbk$ называется *алгебраическим*, если каждый элемент $\theta \in \mathbb{K}$ алгебраичен над $\Bbbk$.

Пример. Пусть $\mathbb{K}/\mathbb{R}$ – алгебраическое расширение. Тогда $\mathbb{K} = \mathbb{R}$ или $\mathbb{K} \simeq \mathbb{C}$ (как поле над $\mathbb{R}$).

Пример. (1) Пусть $\Bbbk = \mathbb{Q}$, а $\mathbb{K} = \mathbb{C}$. Хорошо известно, что множество $\mathbb{Q}[t]$ всех многочленов над $\mathbb{Q}$ счетно. Поэтому счетно также и множество всех алгебраических элементов $A \subset \mathbb{C}$. Однако множество $\mathbb{C}$ несчётно. Это показывает, что трансцендентных над $\mathbb{Q}$ элементов поля $\mathbb{C}$ "существенно больше" чем алгебраических. Однако, доказать трансцендентность *конкретных* действительных чисел довольно сложно.

(2) Для независимой переменной t, пусть $\Bbbk(t)$ – поле рациональных дробей над $\Bbbk$. Любой элемент $f \in \Bbbk(t) \setminus \Bbbk$ является трансцендентным.

Предложение. *Пусть $\mathbb{K}/\Bbbk$ – любое расширение полей. Следующие условия эквивалентны.*

(1) *элемент $\theta \in \mathbb{K}$ является алгебраическим над $\Bbbk$;*

(2) *кольцо $\Bbbk[\theta]$, как векторное пространство над $\Bbbk$, конечномерно;*

(3) *кольцо $\Bbbk[\theta]$ является полем.*

Доказательство. (1)$\Longrightarrow$(2) Пусть $\theta \in \mathbb{K}$ – алгебраический над $\Bbbk$ элемент и пусть $f \in \Bbbk[t]$ – ненулевой многочлен такой, что $f(\theta) = 0$. Мы можем считать, что старший коэффициент f равен 1, т.е.

$$f = t^n + \lambda_{n-1}t^{n-1} + \cdots \lambda_1 t + \lambda_0.$$

Тогда

$$\theta^n = -\lambda_{n-1}\theta^{n-1} - \cdots \lambda_1\theta - \lambda_0$$

и мы можем (по индукции) выразить все степени θ как линейные комбинации $1, \theta, \ldots, \theta^{n-1}$ с коэффициентами в $\Bbbk$. Таким образом, $\Bbbk[\theta]$ как векторное пространство над $\Bbbk$ порождается конечным числом элементов.

(2)$\Longrightarrow$(3) Достаточно показать, что любой ненулевой элемент $b \in \Bbbk[\theta]$ имеет обратный. Рассмотрим $\Bbbk[\theta]$ как векторное пространство над $\Bbbk$ и рассмотрим линейное отображение векторного пространства в себя

$$\varphi_b : \Bbbk[\theta] \longrightarrow \Bbbk[\theta], \qquad \varphi_b : \alpha \longmapsto b\alpha.$$

Так как в $\Bbbk[\theta]$ нет делителей нуля, то это отображение инъективно. Так как $\Bbbk[\theta]$ конечномерно, то оно и сюръективно. Значит, существует элемент $b^{-1} \in \Bbbk[\theta]$ такой, что $\varphi_b(b^{-1}) = b\,b^{-1} = 1$.

(3)$\Longrightarrow$(1) Мы можем записать $\theta^{-1} = \sum_{k=0}^{n} \lambda_k\theta^k$. Отсюда

$$\sum_{k=0}^{n} \lambda_k\theta^{k+1} - 1 = 0. \qquad \square$$

Следствие. *Если $\mathbb{K}/\Bbbk$ – конечное расширение, то любой элемент $\beta \in \mathbb{K}$ является алгебраическим над $\Bbbk$.*

Теорема о башне полей

Теорема. *Пусть $\mathbb{L}/\mathbb{K}$ и $\mathbb{K}/\Bbbk$ – конечные расширения полей, пусть $m := [\mathbb{L} : \mathbb{K}]$ и $n := [\mathbb{K} : \Bbbk]$. Тогда $\mathbb{L}/\Bbbk$ – конечное расширения полей и $[\mathbb{L} : \Bbbk] = nm$.*

Доказательство. Пусть $a_1, \ldots, a_n \in \mathbb{K}$ – базис $\mathbb{K}/\Bbbk$ и пусть $b_1, \ldots, b_m \in \mathbb{L}$ – базис $\mathbb{L}/\mathbb{K}$. Докажем, что элементы $a_ib_j \in \mathbb{L}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$ образуют базис $\mathbb{L}/\Bbbk$.

Предположим, что

$$\sum_{i,j} \lambda_{i,j}a_ib_j = 0$$

для $\lambda_{i,j} \in \Bbbk$. Тогда

$$0 = \sum_{i,j} \lambda_{i,j}a_ib_j = \sum_j \left(\sum_i \lambda_{i,j}a_i \right) b_j.$$

Так как $b_1, \ldots, b_m \in \mathbb{L}$ – базис $\mathbb{L}/\mathbb{K}$ и

$$\sum_i \lambda_{i,j}a_i \in \mathbb{K},$$

то $\sum_i \lambda_{i,j} a_i = 0$ для любого элемента j. Так как $a_1, \ldots, a_n \in \mathbb{K}$ – базис $\mathbb{K}/\Bbbk$, то $\lambda_{i,j} = 0$ для любых i, j. Следовательно, элементы $a_i b_j \in \mathbb{L}$ линейно независимы над $\Bbbk$.

Пусть $c \in \mathbb{L}$. Снова так как $b_1, \ldots, b_m \in \mathbb{L}$ – базис $\mathbb{L}/\mathbb{K}$, то имеет место разложение $c = \sum_j \mu_j b_j$ для некоторых $\mu_j \in \mathbb{K}$, а так как $a_1, \ldots, a_n \in \mathbb{K}$ – базис $\mathbb{K}/\Bbbk$, то $\mu_j = \sum_i \lambda_{i,j} a_i$ для некоторых $\lambda_{i,j} \in \Bbbk$. Таким образом,

$$c = \sum_j \left(\sum_i \lambda_{i,j} a_i \right) b_j = \sum_{i,j} \lambda_{i,j} a_i b_j,$$

т. е. элементы $a_i b_j \in \mathbb{L}$ порождают $\mathbb{L}$ как векторное пространство над $\Bbbk$. □

Предложение. *Если $\mathbb{K}/\Bbbk$ – любое расширение полей, то элементы поля $\mathbb{K}$, алгебраические над $\Bbbk$, также образуют поле.*

Доказательство. Достаточно доказать, что для любых двух алгебраических элементов α, $\beta \in \mathbb{K}$ элементы $\alpha \pm \beta$, $\alpha\beta$ и α/β также являются алгебраическими. Согласно следствию для этого достаточно доказать, что расширение $\Bbbk(\alpha, \beta)/\Bbbk$ конечно. Но $\Bbbk(\alpha, \beta) = \Bbbk(\alpha)(\beta)$. Расширения $\Bbbk(\alpha)/\Bbbk$ и $\Bbbk(\alpha)(\beta)/\Bbbk(\alpha)$ конечны. Требуемый факт теперь легко выводится из теоремы о башне полей. □

Пример. Пусть $\bar{\mathbb{Q}} \subset \mathbb{C}$ – множество всех алгебраических над $\mathbb{Q}$ элементов. Тогда $\bar{\mathbb{Q}}$ – поле. Оно называется *полем алгебраических чисел.*

Определение. Пусть $\mathbb{K}/\Bbbk$ – любое расширение полей и пусть $\theta \in \mathbb{K}$ – алгебраический над $\Bbbk$ элемент. Ненулевой многочлен $\mu_\theta^\Bbbk(t) \in \Bbbk[t]$ минимальной степени, для которого θ является корнем, называется *минимальным многочленом* элемента θ над $\Bbbk$. Если это не приводит к путанице, вместо $\mu_\theta^\Bbbk(t)$ мы будем писать просто μ_θ или даже μ.

Предложение. *Пусть $\mathbb{K}/\Bbbk$ – расширение полей и пусть $\theta \in \mathbb{K}$ – алгебраический над $\Bbbk$ элемент.*

(1) *Если $f \in \Bbbk[t]$ – многочлен такой, что $f(\theta) = 0$, то f делится на минимальный многочлен μ_θ. В частности, минимальный многочлен определен однозначно с точностью до постоянного множителя.*

(2) *Минимальный многочлен μ_θ неприводим в $\Bbbk[t]$.*

Доказательство. Разделим f на $\mu = \mu_\theta$ с остатком:

$$f = \mu g + r.$$

Тогда

$$0 = f(\theta) = \mu(\theta) g(\theta) + r(\theta) = r(\theta).$$

Так как $\deg r < \deg \mu$, то $r = 0$. Это доказывает (1).

Для доказательства второго утверждения предположим, что $\mu = \mu_1 \mu_2$. Тогда $\mu_1(\theta) = 0$ или $\mu_2(\theta) = 0$. Это противоречит минимальности многочлена μ. □

Присоединение к полю корня неприводимого многочлена

Напомним, что поле $\Bbbk$ называется *расширением поля* $\Bbbk_0$, если $\Bbbk_0$ является подполем в $\Bbbk$.

Пусть $\mathbb{L}$ – поле и пусть $\Bbbk \subset \mathbb{L}$ – подполе. Для элемента $\theta \in \mathbb{L}$ через $\Bbbk(\theta)$ мы обозначим подполе в $\mathbb{L}$, порожденное $\Bbbk$ и θ. Это наименьшее подполе в $\mathbb{L}$, содержащее $\Bbbk$ и θ.

Теорема. *Пусть $\Bbbk$ – поле и пусть $f \in \Bbbk[t]$ – многочлен положительной степени.*

(1) *Следующие три условия эквивалентны:*

   (a) *многочлен f неприводим,*

   (b) *факторкольцо $\Bbbk[t]/(f)$ является полем,*

   (c) *факторкольцо $\Bbbk[t]/(f)$ не имеет делителей нуля.*

(2) *Пусть многочлен f неприводим. Если $\mathbb{L}/\Bbbk$ – расширение полей такое, что f имеет корень $\theta \in \mathbb{L}$, то существует изоморфизм*

$$\varphi : \Bbbk[t]/(f) \to \Bbbk(\theta), \quad t \mapsto \theta,$$

являющийся тождественным отображением на $\Bbbk$.

Доказательство. (1) Докажем (1)(a) $\Longrightarrow$ (1)(b). Пусть $\pi : \Bbbk[t] \to \Bbbk[t]/(f)$ – естественный гомоморфизм. Рассмотрим ненулевой элемент $\bar{g} = g + (f) \in \Bbbk[t]/(f)$. Таким образом, $\bar{g} = \pi(g)$, где $g \in \Bbbk[t]$ – многочлен такой, что $g \notin (f)$. Последнее означает, что f и g взаимно просты (поскольку f неприводим). По теореме о наибольшем общем делителе существуют многочлены u, $v \in \Bbbk[t]$ такие, что $1 = fu + gv$. Отсюда

$$1 = \pi(1) = \pi(f)\pi(u) + \pi(g)\pi(v) = \bar{g}\pi(v).$$

Следовательно, любой ненулевой элемент $\bar{g} \in \Bbbk[t]/(f)$ обратим и поэтому $\Bbbk[t]/(f)$ – поле.

Импликация (1)(b) $\Longrightarrow$ (1)(c) очевидна. Для доказательства (1)(c) $\Longrightarrow$ (1)(a) предположим, что $f = f_1 f_2$, где $f_i \notin (f)$. Тогда в $\Bbbk[t]/(f)$ имеем

$$\pi(f_1)\pi(f_2) = \pi(f_1 f_2) = \pi(f) = 0,$$

т.е. $\pi(f_1)$, $\pi(f_2)$ – делители нуля. Противоречие.

(2) Рассмотрим отображение

$$\psi : \Bbbk[t] \longrightarrow \mathbb{L}, \qquad h \longmapsto h(\theta).$$

Ясно, что ψ – гомоморфизм. Его ядро является главным идеалом: $\mathrm{Ker}(\psi) = (h)$. С другой стороны, $f \in \mathrm{Ker}(\psi)$ и f неприводим. Поэтому $\mathrm{Ker}(\psi) = (f)$. По теореме о гомоморфизме $\psi(\Bbbk[t]) \simeq \Bbbk[t]/(f)$. □

Замечание. Построенное выше расширение называется *присоединением к полю корня неприводимого многочлена*. Действительно, поле $\Bbbk$ естественно вкладывается в $\Bbbk[t]/(f)$ (как композиция $\Bbbk \hookrightarrow \Bbbk[t] \stackrel{\pi}{\longrightarrow} \Bbbk[t]/(f)$), а согласно (2) образ $\theta := \pi(t)$ является корнем многочлена f.

Поле разложения многочлена

Определение. Пусть $\Bbbk$ – произвольное поле. *Полем разложения* многочлена $f \in \Bbbk[t]$ называется поле $\mathbb{K} \supset \Bbbk$ такое, что f над $\mathbb{K}$ разлагается на линейные множители:

$$f = c\prod_{i=1}^{n}(t-\alpha_i),$$

где $\alpha_i \in \mathbb{K}$ и элементы α_i порождают $\mathbb{K}$ над $\Bbbk$, т.е. $\mathbb{K} = \Bbbk(\alpha_1, \ldots, \alpha_n)$.

Таким образом, поле разложения многочлена $f \in \Bbbk[t]$ – это минимальное поле, содержащее $\Bbbk$, в котором f разлагается на линейные множители.

Теорема. *Пусть $\Bbbk$ – поле и $f \in \Bbbk[t]$ – некоторый многочлен. Существует поле $\mathbb{K} \supset \Bbbk$, являющееся полем разложения для f над $\Bbbk$.*

Замечание. На самом деле можно показать, что любые два поля разложения изоморфны над $\Bbbk$, т.е. существует изоморфизм тождественный на $\Bbbk$. Мы не будем доказывать этот факт в этом семестре.

Доказательство. Индукция по степени $n = \deg f$. База индукции очевидна. Предположим, что утверждение верно для всех многочленов степени $< n$. Пусть f_1 – неприводимый множитель f. Присоединим к $\Bbbk$ корень f_1, т.е. рассмотрим расширение $\mathbb{K}_1/\Bbbk$, где $\mathbb{K}_1 = \Bbbk[t]/(f_1)$. Пусть θ_1 – корень f_1 в $\mathbb{K}_1$. Запишем $f = (t-\theta_1)g$. Так как $\deg g < n$, то для g над $\mathbb{K}_1$ существует поле разложения $\mathbb{L}$. Таким образом, f разлагается на линейные множители в $\mathbb{L}$:

$$f = c(t-\theta_1)\cdots(t-\theta_n).$$

Положим $\mathbb{K} := \Bbbk(\theta_1, \ldots, \theta_n)$. □

Замечание. Пусть $f \in \Bbbk[t]$ – многочлен положительной степени и пусть $\mathbb{K}$ – его поле разложения f над $\Bbbk$. Мы считаем, что f неприводим. По конструкции и по теореме о башне полей $[\mathbb{K} : \Bbbk] \leq n!$, причем при $n = 2$ имеет место равенство. При $n = 3$ степень $[\mathbb{K} : \Bbbk]$ зависит от дискриминанта D многочлена f:

$$[\mathbb{K} : \Bbbk] = \begin{cases} 3 & \text{если } \sqrt{D} \in \Bbbk \\ 6 & \text{если } \sqrt{D} \notin \Bbbk \end{cases}$$

Действительно, пусть $\sqrt{D} \in \Bbbk$. По формулам Виета

$$\theta_2 + \theta_3,\ \theta_2\theta_3 \in \Bbbk(\theta_1).$$

С другой стороны, с точностью до знака имеем

$$\Bbbk \ni \sqrt{D} = (\theta_1-\theta_2)(\theta_1-\theta_3)(\theta_2-\theta_3) = (\theta_1^2 - (\theta_2+\theta_3)\theta_1 + \theta_2\theta_3)(\theta_2-\theta_3).$$

Поэтому $\theta_2 - \theta_3 \in \Bbbk(\theta_1)$, а отсюда и θ_2, $\theta_3 \in \Bbbk(\theta_1)$, т.е. $\Bbbk(\theta_1) = \mathbb{K}$. Следовательно, $[\mathbb{K} : \Bbbk] = 3$.

Пусть $\sqrt{D} \notin \Bbbk$. Предположим, что $[\mathbb{K} : \Bbbk] = 3$. Тогда $\Bbbk(\theta_1) = \mathbb{K}$. Следовательно, θ_2, $\theta_3 \in \Bbbk(\theta_1)$ и поэтому $\sqrt{D} \in \Bbbk(\theta_1)$. Таким образом, мы имеем башню полей

$$\Bbbk \subset \Bbbk(\sqrt{D}) \subset \Bbbk(\theta_1).$$

По теореме о башне полей

$$[\Bbbk(\theta_1) : \Bbbk] = [\Bbbk(\theta_1) : \Bbbk(\sqrt{D})] \cdot [\Bbbk(\sqrt{D}) : \Bbbk].$$

Так как $[\Bbbk(\theta_1) : \Bbbk] = 3$ и $[\Bbbk(\sqrt{D}) : \Bbbk] = 2$, то мы получаем противоречие.

Алгебраическое замыкание поля

Определение. Алгебраическим замыканием поля $\Bbbk$ называется алгебраическое расширение $\bar{\Bbbk} \supset \Bbbk$ такое, что поле $\bar{\Bbbk}$ алгебраически замкнуто.

Теорема. *Для любого поля $\Bbbk$ алгебраическое замыкание $\bar{\Bbbk}$. Любые два алгебраических замыкания одного поля $\Bbbk$ изоморфны над $\Bbbk$.*

Мы докажем только существование алгебраического замыкания поля из счетного числа элементов. Существование алгебраического замыкания в общем случае требует применения аксиомы выбора.

Лемма. *Пусть $\mathbb{K} \supset \Bbbk$ – алгебраическое расширение полей такое, что любой многочлен $f \in \Bbbk[t]$ имеет корень в $\mathbb{K}$. Тогда $\mathbb{K}$ – алгебраическое замыкание поля $\Bbbk$.*

Доказательство. Достаточно доказать, что $\mathbb{K}$ алгебраически замкнуто. Предположим противное. Пусть $g \in \mathbb{K}[t]$ – неприводимый многочлен степени > 1. Запишем

$$g(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0, \qquad a_i \in \mathbb{K}.$$

Элементы a_i являются алгебраическими над $\Bbbk$. По теореме о башне полей $\mathbb{L} := \Bbbk(a_0, \ldots, a_n)$ – конечное расширение поля $\Bbbk$. Пусть $\mathbb{L}_1 = \mathbb{L}(\theta) = \mathbb{L}/(g)$ – поле, полученное присоединением корня θ многочлена g к $\mathbb{L}$. Снова по теореме о башне полей $\mathbb{L}_1 \supset \Bbbk$ – также конечное расширение. Значит, элемент θ является алгебраическим над $\Bbbk$ и $\theta \in \mathbb{K}$. Следовательно, $\deg(g) = 1$. Противоречие. □

Доказательство теоремы. Пусть $\Bbbk$ – счетное поле. Множество всех неприводимых многочленов над ним счетно. Пронумеруем их:

$$f_1,\ f_2, \ldots, f_n, \ldots$$

По индукции построим башню полей

$$\Bbbk = \Bbbk_0 \subset \Bbbk_1 \subset \cdots \Bbbk_n \subset \cdots,$$

где $\Bbbk_i$ – поле разложение многочлена f_i над $\Bbbk_{i-1}$. Положим

$$\mathbb{K} := \bigcup_{i=1}^{\infty} \Bbbk_i.$$

Мы утверждаем, что $\mathbb{K}$ – поле. Действительно, для любых α, $\beta \in \mathbb{K}$ существует n такое, что α, $\beta \in \Bbbk_n$. Поскольку $\Bbbk_n$ – поле, то определены операции $\alpha \pm \beta$, $\alpha \cdot \beta$ и α/β (последнее, если $\beta \neq 0$). Так как $\Bbbk_n$ – подполе в $\Bbbk_m$ при $m > n$, то и, следовательно, эти операции определенные в $\Bbbk_n$ и $\Bbbk_m$ совпадают. Таким образом, корректно определены операции сложения и умножения на всем множестве $\mathbb{K}$. Очевидным образом для них проверяются аксиомы поля.

Далее мы утверждаем, что любой многочлен $f \in \Bbbk[t]$ имеет корень в $\mathbb{K}$. Действительно, f делится на какой-то многочлен f_i, который разлагается на множители в $\Bbbk_i \subset \mathbb{K}$. Наконец, по лемме $\mathbb{K}$ – алгебраическое замыкание поля $\Bbbk$. $\square$