Версия от 27 января 2013 г.

Краткое содержание курса “Алгебра” (1-й семестр, 3-й поток) (лектор Марков В.Т.)

Предисловие

Этот текст не претендует ни на полноту изложения, ни на литературные достоинства — основной целью автора была краткость. В большинстве случаев приводятся только наброски доказательств (начало и конец доказательства отмечаются знаками ◀ и ▶, соответственно). Восстановление всех деталей всех доказательств — обязательное условие усвоения курса и хороший способ самостоятельной проверки понимания материала.

Автор приносит искреннюю и глубокую благодарность студентам 1 курса наборов 2009–2011 и особенно 2012 гг., нашедшим и указавшим ему многочисленные опечатки в тексте. На долю следующих групп опечаток тоже хватит.

Содержание

Лекция 1. Система линейных уравнений. Матрица коэффициентов и расширенная матрица системы. Приведение матриц и систем линейных уравнений к ступенчатому виду. Метод Гаусса. 4
Лекция 2. Линейная зависимость строк (столбцов). Основная лемма о линейной зависимости, база и ранг системы строк (столбцов). 8
Лекция 3. Ранг матрицы. Критерий совместности и определенности системы линейных уравнений в терминах рангов матриц. Фундаментальная система решений однородной системы линейных уравнений. 11
Лекция 4. Группа подстановок конечного множества, знак подстановки (четность), знакопеременная группа, разложение подстановки в произведение транспозиций и независимых циклов. 15
Лекция 5. Определитель квадратной матрицы, его основные свойства (линейность, кососимметричность, определитель транспонированной матрицы). Изменение определителя при элементарных преобразованиях строк (столбцов) матрицы. Определитель треугольной матрицы. Критерий равенства определителя нулю. 20
Лекция 6. Определитель матрицы с углом нулей. Определитель Вандермонда. Миноры и алгебраические дополнения элементов. Разложние определителя по строке (столбцу). Лемма о “фальшивом” разложении определителя. Формулы Крамера для решения определенных квадратных систем линейных уравнений. 26
Лекция 7. Операции над матрицами и их свойства. Обобщенная ассоциативность. Транспонирование произведения матриц. Умножение матрицы на диагональную матрицу слева и справа. Единичная матрица, ее единственность. Скалярные матрицы. Обратная матрица, ее единственность. 31
Лекция 8. Умножение треугольных матриц. Матричные единицы и их умножение. Элементарные матрицы и их связь с элементарными преобразованиями. Определитель произведения матриц. Критерий существования и способы нахождения обратной матрицы. 36
Лекция 9. Миноры прямоугольной матрицы. Вычисление ранга матрицы с помощью миноров (теорема о ранге матрицы). 40
Лекция 10. Ранг произведения матриц. Факторизационный ранг матрицы. Матричная запись системы линейных уравнений. Строение общего решения неоднородной системы уравнений, его геометрическая интерпретация. 42
Лекция 11. Основные алгебраические структуры: группы. 46
Лекция 12. Циклические группы. Порядок элемента. Подгруппы циклических групп. Изоморфизм циклических групп одного порядка. Теорема Кэли. Смежные классы, теорема Лагранжа и ее следствия. 49
Лекция 13. Основные алгебраические структуры: кольца, поля. 53
Лекция 14. Поле комплексных чисел. Комплексная плоскость. Модуль и аргумент комплексного числа. Алгебраическая и тригонометрическая форма записи комплексных чисел. Операция сопряжения комплексных чисел и ее свойства. Формула Муавра. Корни целой степени из комплексного числа. Группа комплексных корней из единицы. 57
Лекция 15. Кольцо многочленов от одной переменной над полем. Возможность и единственность деления на ненулевой многочлен с остатком. Наибольший общий делитель двух многочленов, его выражение через многочлены, алгоритм Евклида. . 62
Лекция 16. Неприводимые многочлены. Факториальность кольца многочленов и кольца целых чисел. Многочлен как функция. Схема Горнера. Корни многочлена, кратность корня. Понижение кратности корня при дифференцировании, избавление от кратных корней. 66

Лекция 17. Алгебраическая замкнутость поля комплексных чисел. Неприводимые многочлены над полями комплексных и действительных чисел. 71
Лекция 18. Интерполяционный многочлен, формула Лагранжа и метод Ньютона для его построения. Поле рациональных дробей. Простейшие дроби. Разложение правильной дроби в сумму простейших дробей, случай вещественного и комплексного полей. 74
Лекция 19. Границы корней многочлена. Теорема Декарта. 79
Лекция 20. Метод Штурма отделения вещественных корней многочлена. 82
Лекция 21. Кольцо многочленов от нескольких переменных. Лексикографический порядок на одночленах. Старший член произведения многочленов. Симметрические многочлены, их выражение через элементарные симметрические многочлены, формулы Виета. 85
Лекция 22. Результант двух многочленов, его выражение через корни многочленов. . . 89
Лекция 23. Дискриминант многочлена, выражение дискриминанта через корни многочлена. 92
Рекомендуемая литература. 96

Лекция 1. Система линейных уравнений. Матрица коэффициентов и расширенная матрица системы. Приведение матриц и систем линейных уравнений к ступенчатому виду. Метод Гаусса.

1°. Системы линейных алгебраических уравнений.

Определение 1.1. *Системой линейных (алгебраических) уравнений* называется система уравнений вида

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m, \end{array} \tag{1.1}$$

где *коэффициенты* a_{ij} и *правые части, или свободные члены* b_i системы (1.1) предполагаются (пока!) действительными числами. *Решением (или частным решением)* системы (1.1) называется такой набор чисел $x^0 = (x_1^0, x_2^0, \ldots, x_n^0)$, что при подстановке в (1.1) x_1^0 вместо x_1, x_2^0 вместо x_2 и т.д. каждое уравнение системы превращается в верное числовое равенство. *Решить систему уравнений* — значит найти все ее решения (множество всех решений системы (1.1) называется также ее *общим решением*).

Определение 1.2. Система (1.1) называется *совместной*, если она имеет хотя бы одно решение. В противном случае система называется *несовместной*. Совместная система называется *определенной*, если она имеет единственное решение. В противном случае совместная система называется *неопределенной*.

Определение 1.3. Две системы вида (1.1) называются *эквивалентными*, если множества их решений совпадают.

Определение 1.4. *Элементарным преобразованием системы* (1.1) называется преобразование одного из следующих трех видов:
ЭП1) к некоторому уравнению системы почленно прибавить другое уравнение той же системы, умноженное на произвольное число;
ЭП2) поменять местами два уравнения системы;
ЭП3) некоторое уравнение системы умножить на **ненулевое** число.

Предложение 1.5. При любом элементарном преобразовании получается система, эквивалентная исходной.

◄ Для преобразования ЭП2 утверждение очевидно. Для ЭП1: пусть к i-му уравнению прибавляется j-е, умноженное на λ. Тогда в новой системе i-е уравнение имеет вид

$$(a_{i1} + \lambda a_{j1})x_1 + (a_{i2} + \lambda a_{j2})x_2 + \ldots + (a_{in} + \lambda a_{jn})x_n = b_i + \lambda b_j. \tag{1.2}$$

Подставим в (1.2) любое решение $x^0 = (x_1^0, x_2^0, \ldots, x_n^0)$ исходной системы. Получим равенство

$$(a_{i1} + \lambda a_{j1})x_1^0 + (a_{i2} + \lambda a_{j2})x_2^0 + \ldots + (a_{in} + \lambda a_{jn})x_n^0 = \underbrace{a_{i1}x_1^0 + a_{i2}x_2^0 + \ldots + a_{in}x_n^0}_{b_i} + \lambda(\underbrace{a_{j1}x_1^0 + a_{j2}x_2^0 + \ldots + a_{jn}x_n^0}_{b_j}) = b_i + \lambda b_j.$$

Следовательно, любое решение исходной системы является решением преобразованной системы. Обратно, исходную систему можно получить из преобразованной системы, прибавляя к ее i-му уравнению j-е, умноженное на число $(-\lambda)$. Аналогичное доказательство — для ЭП3 — оставлено в качестве упражнения.►

2°. Прямоугольные матрицы. Сложение матриц. Умножение матрицы на число.

Определение 1.6. *Матрицей* размера $m \times n$ называется прямоугольная таблица из m строк и n столбцов, заполненная (пока — действительными) числами (*элементами* матрицы). Положение элемента матрицы определяется двумя индексами — номером строки и номером столбца. Таким образом, в общем виде можно записать

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

Принято элементы матрицы, обозначенной некоторой заглавной буквой, обозначать той же буквой в строчном написании. Например, если $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, то $a_{11} = 1,\ a_{12} = 2,\ a_{21} = 3,\ a_{22} = 4$. Сокращенно принято записывать $A = (a_{ij})$, $B = (b_{ij})$ и т.д.

Системе (1.1) соответствуют две матрицы: *матрица коэффициентов*, или просто *матрица системы*

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

и *расширенная матрица системы*

$$\overline{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}.$$

Определение 1.7. *Суммой* двух матриц $A = (a_{ij})$ и $B = (b_{ij})$ одного и того же размера $m \times n$ называется матрица того же размера, элементы которой — суммы соответствующих элементов матриц-слагаемых: $A + B = (a_{ij} + b_{ij})$. *Произведением* матрицы $A = (a_{ij})$ размера $m \times n$ на число λ называется матрица того же размера, элементы которой — произведения элементов исходной матрицы на одно и то же число λ: $\lambda A = (\lambda a_{ij})$.

Пример

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix},$$

$$10 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 10 & 20 \\ 30 & 40 \end{pmatrix}.$$

3°. Элементарные преобразования матриц. Преобразование матрицы к ступенчатому виду.

Элементарным преобразованиям системы линейных уравнений соответствуют элементарные преобразования ее расширенной матрицы (вспомним определение 1.4):

Определение 1.8. *Элементарным преобразованием матрицы* называется преобразование одного из следующих трех видов:
ЭП1) к некоторой строке матрицы прибавить другую строку той же матрицы, умноженную на произвольное число;
ЭП2) поменять местами две строки матрицы;
ЭП3) некоторую строку матрицы умножить на **ненулевое** число.
Операции над строками понимаются как операции над матрицами размера $1 \times n$.

Определение 1.9. *Главным элементом (или ведущим элементом, или лидером)* строки матрицы называется первый слева ненулевой элемент этой строки.

Определение 1.10. Матрица называется *ступенчатой*, если
1) все нулевые строки этой матрицы расположены ниже ее ненулевых строк;
2) лидер каждой следующей ненулевой строки расположен строго правее лидера предшествующей строки.
Говорят также, что такая матрица имеет ступенчатый вид

Теорема 1.11. Любую матрицу с помощью элементарных преобразований первого типа (ЭП1) можно привести к ступенчатому виду.

◄ Рассмотрим произвольную матрицу A размера $m \times n$. Если все элементы A равны 0, то матрица A уже ступенчатая. В противном случае среди лидеров ненулевых строк матрицы A выберем элемент a_{ij}, расположенный **не правее** всех остальных лидеров. Если $a_{1j} = 0$, прибавим i-ю строчку к первой. Значит, можно считать, что $i = 1$. Для каждой из строк с номерами $k = 2, \ldots, m$ положим $\lambda = -(a_{kj}/a_{1j})$ и применим ЭП1: к k-той строке прибавим 1-ю строку, умноженную на λ. Получится преобразование

$$\begin{pmatrix} 0 & \ldots & 0 & a_{1j} & \ldots & a_{1n} \\ 0 & \ldots & 0 & a_{2j} & \ldots & a_{2n} \\ \vdots & \ldots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & a_{mj} & \ldots & a_{mn} \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & \ldots & 0 & a_{1j} & \ldots & a_{1n} \\ 0 & \ldots & 0 & 0 & \ldots & a'_{2n} \\ \vdots & \ldots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & \ldots & a'_{mn} \end{pmatrix}$$

Повторим те же действия для подматрицы размера $(m-1) \times n$, образованной строками преобразованной матрицы с номерами 2, 3, ..., m. В итоге получим ступенчатую матрицу.►

4°. Метод Гаусса решения систем линейных уравнений. Из теоремы 1.11 можно, с учетом предложения 1.5, получаем

Следствие 1.12. Любая система линейных уравнений эквивалентна системе, расширенная матрица которой является ступенчатой.

Получение этой ступенчатой матрицы называют *прямым ходом* метода Гаусса. Остается описать множество решений системы, расширенная матрица которой является ступенчатой.

Если в ступенчатой матрице есть строка $(0, 0, \ldots, b)$, где $b \neq 0$, то соответствующее уравнение $0x_1 + 0x_2 + \ldots + 0x_3 = b$ не имеет решений, следовательно, исходная система несовместна.

Иначе система является совместной. Определим *главные* неизвестные, коэффициенты при которых являются лидерами некоторых строк, а остальные неизвестные назовем свободными.

Предложение 1.13. Для любого набора значений свободных неизвестных существует единственный набор значений главных неизвестных, который вместе с заданным набором значений образует решение системы.

◄ Рассмотрим последнее ненулевое уравнение совместной ступенчатой системы. В него с ненулевым коэффициентом входит единственное главное неизвестное, и его можно выразить через свободные неизвестные (и правую часть). Подставляя полученное выражение в остальные уравнения, получаем снова уравнение, содержащее одну главную неизвестную, и т.д. Окончательно получится набор выражений главных неизвестных через свободные.►

Процесс выражения главных неизвестных через свободные называют *обратным ходом* метода Гаусса.

Обратный ход метода Гаусса можно формализовать и другим способом.

Теорема 1.14. Любую ступенчатую матрицу с помощью элементарных преобразований строк первого и третьего типов (ЭП1 и ЭП3) можно привести к *улучшенному ступенчатому виду*, в котором лидер каждой ненулевой строки равен 1 и является единственным ненулевым элементом в своем столбце

◄ Деля каждую ненулевую строку на ее ведущий элемент, добиваемся выполнения первого условия. Преобразованиями первого типа, как в доказательстве теоремы 1.11, обращаем в 0 все остальные элементы столбца, содержащего ведущий элемент; при этом остальные ведущие элементы не меняются.►

Для улучшенной ступенчатой матрицы выражение главных неизвестных системы через свободные очевидно, так как каждое ненулевое уравнение содержит единственное главное неизвестное.

Следствие 1.15. Совместная система линейных уравнений является определенной тогда и только тогда, когда все неизвестные оказываются главными.

Следствие 1.16. Если число уравнений в совместной системе линейных уравнений меньше числа неизвестных, то система не является определенной.

Лекция 2. Линейная зависимость строк (столбцов). Основная лемма о линейной зависимости, база и ранг системы строк (столбцов).

1°. Линейная зависимость строк (столбцов).

В дальнейшем под *вектором* понимается строка (или столбец) действительных чисел. Говоря о *системе векторов*, мы имеем в виду, что все векторы этой системы имеют одну и ту же длину. Элементы строки (столбца) называют также *координатами* соответствующего вектора. Вектор, имеющий n координат, будем называть n-мерным вектором. Множество всех n-мерных векторов назовем n-мерным *арифметическим векторным пространством* и обозначим $\mathbb{R}^n$. Операции сложения векторов и умножения вектора на число определены как операции над матрицами. *Нулевым вектором* (обозначается 0) называется вектор, все координаты которого равны 0.

Определение 2.1. *Линейной комбинацией* системы векторов $a^1, a^2, \ldots, a^k$ с *коэффициентами* $\lambda_1, \lambda_2, ..., \lambda_k$ называется вектор

$$\lambda_1 a^1 + \lambda_2 a^2 + \ldots + \lambda_k a^k.$$

Линейная комбинация называется *тривиальной*, если все ее коэффициенты равны 0; если же хотя бы один коэффициент не равен 0, то такая линейная комбинация называется *нетривиальной*.

Определение 2.2. Система векторов $a^1, a^2, \ldots, a^k$ называется *линейно зависимой*, если существуют нетривиальная линейная комбинация этих векторов, равная нулевому вектору. Если такой линейной комбинации не существует, то система векторов $a^1, a^2, \ldots, a^k$ называется *линейно независимой*. Иными словами, система $a^1, a^2, \ldots, a^k$ линейно независима, если

$$\lambda_1 a^1 + \lambda_2 a^2 + \ldots + \lambda_k a^k = 0 \Rightarrow \lambda_1 = \lambda_2 = \ldots = \lambda_k = 0.$$

Предложение 2.3. Система векторов $a^1, a^2, \ldots, a^k$ является линейно зависимой тогда и только тогда, когда один из векторов — линейная комбинация остальных.

◀ Пусть система векторов $a^1, a^2, \ldots, a^k$ линейно зависима, и $\lambda_1 a^1 + \lambda_2 a^2 + \ldots + \lambda_k a^k = 0$ — нетривиальная линейная комбинация. Тогда $\lambda_i \neq 0$ для некоторого $i \in \{1, \ldots, k\}$, откуда

$$a^i = (-\frac{\lambda_1}{\lambda_i})a^1 + \ldots + (-\frac{\lambda_{i-1}}{\lambda_i})a^{i-1} + (-\frac{\lambda_{i+1}}{\lambda_i})a^{i+1} + \ldots + (-\frac{\lambda_k}{\lambda_i})a^k.$$

Обратно, если

$$a^i = \lambda_1 a^1 + \ldots + \lambda_{i-1} a^{i-1} + \lambda_{i+1} a^{i+1} + \ldots + \lambda_k a^k,$$

то

$$\lambda_1 a^1 + \ldots + \lambda_{i-1} a^{i-1} + (-1)a^i + \lambda_{i+1} a^{i+1} + \ldots + \lambda_k a^k = 0,$$

причем коэффициент $-1 \neq 0$.▶

Предложение 2.4. Если система векторов $a^1, a^2, \ldots, a^k$ линейно независима, а система $a^1, a^2, \ldots, a^k, b$ линейно зависима, то вектор b линейно выражается через $a^1, a^2, \ldots, a^k$.

◀ Запишем условие линейной зависимости:

$$\lambda_1 a^1 + \ldots + \lambda_k a^k + \lambda b = 0,$$

где хотя бы один из коэффициентов $\lambda_1, \ldots, \lambda_k, \lambda$ не равен 0. В случае $\lambda = 0$ получаем противоречие:

$$\lambda_1 a^1 + \lambda_2 a^2 + \ldots + \lambda_k a^k = 0 \Rightarrow \lambda_1 = \lambda_2 = \ldots = \lambda_k = 0 = \lambda.$$

Значит, $\lambda \neq 0$, и можно выразить вектор b, как в предыдущем доказательстве.►

2°. Основная лемма о линейной зависимости.

Мы говорим, что (конечная или бесконечная) система векторов T *линейно выражается* через систему векторов S, если каждый вектор системы T есть линейная комбинация некоторых векторов из системы S.

Замечание 2.5. Если система векторов T линейно выражается через систему векторов S, а система векторов S линейно выражается через систему векторов Q, то система векторов T линейно выражается через систему векторов Q.

Теорема 2.6 (Основная лемма о линейной зависимости). Пусть $a^1, a^2, \ldots, a^r$ и $b^1, b^2, \ldots, b^s$ — две системы векторов, причем вторая система линейно выражается через первую. Если $s > r$, то система $b^1, b^2, \ldots, b^s$ линейно зависима.

◄ По условию, можно записать

$$b^i = \alpha_1^i a^1 + \ldots + \alpha_r^i a^r, \quad i = 1, \ldots, s.$$

Для произвольных коэффициентов $\lambda_1, \ldots, \lambda_s$ запишем

$$\begin{aligned} \lambda_1 b^1 + \lambda_2 b^2 + \ldots + \lambda_s b^s = \\ \lambda_1(\alpha_1^1 a^1 + \ldots + \alpha_r^1 a^r) + \lambda_2(\alpha_1^2 a^1 + \ldots + \alpha_r^2 a^r) + \ldots + \lambda_s(\alpha_1^s a^1 + \ldots + \alpha_r^s a^r) = \\ (\alpha_1^1\lambda_1 + \ldots + \alpha_1^s\lambda_s)a^1 + (\alpha_2^1\lambda_1 + \ldots + \alpha_2^s\lambda_s)a^2 + \ldots + (\alpha_r^1\lambda_1 + \ldots + \alpha_r^s\lambda_s)a^r \end{aligned} \quad (2.1)$$

Рассмотрим систему линейных уравнений

$$\begin{array}{rcl} \alpha_1^1\lambda_1 + \ldots + \alpha_1^s\lambda_s & = & 0 \\ \alpha_2^1\lambda_1 + \ldots + \alpha_2^s\lambda_s & = & 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots & & \\ \alpha_r^1\lambda_1 + \ldots + \alpha_r^s\lambda_s & = & 0 \end{array} \quad (2.2)$$

относительно $\lambda_1, \ldots, \lambda_s$. Система (2.2) совместна, а поскольку число уравнений (r) меньше числа неизвестных (s), согласно следствию 1.16 теоремы 1.14, эта система не является определенной, значит, имеет ненулевое решение. Подставляя это решение вместо $\lambda_1, \ldots, \lambda_s$ в (2.1), получаем зависимость системы $b^1, b^2, \ldots, b^s$.►

Другая формулировка основной леммы о линейной зависимости: если система векторов $b^1, b^2, \ldots, b^s$ линейно независима и линейно выражается через систему векторов $a^1, a^2, \ldots, a^r$, то $s \leqslant r$.

Следствие 2.7. Любая система из $k > n$ векторов арифметического n-мерного векторного пространства является линейно зависимой.

◄ Любой вектор $x = (x_1, x_2, \ldots, x_n)$ линейно выражается через векторы

$$\begin{array}{rcl} e^1 & = & (1, 0, 0, \ldots, 0, 0) \\ e^2 & = & (0, 1, 0, \ldots, 0, 0), \\ \ldots\ldots\ldots\ldots\ldots\ldots & & \\ e^n & = & (0, 0, 0, \ldots, 0, 1) : \end{array} \quad (2.3)$$

$$x = x_1 e^1 + x_2 e^2 + \ldots + x_n e^n.$$

►

Запомним на будущее систему векторов $e^1, \ldots, e^n$ из (2.3).

3°. База и ранг системы строк (столбцов).

Определение 2.8. *Базой* (конечной или бесконечной) системы векторов S называется ее конечная линейно независимая подсистема, через которую линейно выражается любой вектор системы S. По определению, базой пустой системы, а также системы, состоящей из одного нулевого вектора, считается пустая подсистема.

Пример: система $e^1, \ldots, e^n$ из (2.3) является базой всего пространства $\mathbb{R}^n$ (проверьте, глядя на доказательство следствия 2.7). Эту базу мы будем называть *стандартной базой* в $\mathbb{R}^n$.

Теорема 2.9 (о дополнении до базы)**.** Любая линейно независимая подсистема системы векторов S в пространстве $\mathbb{R}^n$ может быть дополнена до базы системы S.

◀ Пусть S — система векторов в $\mathbb{R}^n$, S_0 — линейно независимая подсистема. Выберем линейно независимую подсистему S_1, в S, содержащую S_0 и состоящую из максимально возможного числа векторов (это можно сделать ввиду следствия 2.7). Согласно предложению 2.4, любой вектор из S линейно выражается через векторы из S_1. Значит, S_1 — база системы S.▶

Следствие 2.10. Любая система векторов в пространстве $\mathbb{R}^n$ обладает базой из не более чем n векторов.

◀ Применим теорему 2.9 к пустой подсистеме заданной системы векторов.▶

Теорема 2.11. Любые две базы одной и той же системы векторов содержат одинаковое число элементов.

◀ Пусть база S_0 системы S содержит r векторов, а база S_1 — s векторов. Поскольку система S_1 линейно выражается через систему S_0 и S_1 — линейно независимая система, ввиду основной леммы о линейной зависимости (теорема 2.6) имеем $s \leqslant r$. Но симметрично $r \leqslant s$, стало быть, $r = s$.▶

Определение 2.12. *Рангом* системы векторов S называется число элементов произвольной базы системы S. Это число будем обозначать $\mathrm{rk}(S)$.

Теорема 2.13. Если система векторов T линейно выражается через систему S, то $\mathrm{rk}(T) \leqslant \mathrm{rk}(S)$.

◀ Пусть S_0 — база системы S, содержащая $s = \mathrm{rk}(S)$ векторов, T_0 — база системы T, содержащая $t = \mathrm{rk}(T)$ векторов. Тогда система S линейно выражаются через S_0, значит, все векторы из T линейно выражаются через S_0. В частности, система T_0 линейно выражается через S_0, и по основной лемме о линейной зависимости (теорема 2.6), $t \leqslant s$.▶

Теорема 2.14. Если система векторов S имеет ранг r, то любая линейно независимая подсистема из r векторов системы S — база системы S.

◀ Пусть S_0 — линейно независимая подсистема системы S, содержащая $r = \mathrm{rk}(S)$ векторов. По теореме 2.9 подсистему S_0 можно дополнить до базы из r векторов. Но в S_0 уже есть r векторов, значит, добавление новых векторов невозможно, и S_0 — база системы S.▶

Лекция 3. Ранг матрицы. Критерий совместности и определенности системы линейных уравнений в терминах рангов матриц. Фундаментальная система решений однородной системы линейных уравнений.

1°. Определение ранга матрицы.

Теорема 3.1 (о ранге системы строк и столбцов матрицы). Ранг системы строк прямоугольной матрицы равен рангу системы ее столбцов и совпадает с числом ненулевых строк после приведения матрицы к ступенчатому виду.

Сначала докажем две вспомогательные леммы.

Лемма 3.2. При элементарных преобразованиях строк матрицы ранг системы ее строк не меняется.

◀ Пусть матрица A' получена из матрицы A одним элементарным преобразованием. Обозначим ранги систем строк матриц A и A' через r и r', соответственно. Система строк матрицы A' линейно выражается через систему строк матрицы A: в матрице A' может быть только одна "новая" строка, а она является линейной комбинацией двух строк (для ЭП1) или одной строки (для ЭП3) матрицы A. Следовательно, ввиду теоремы 2.13, $r' \leqslant r$. Ввиду обратимости элементарных преобразований (см. доказательство предложения 1.5) имеем также $r \leqslant r'$, т.е. $r' = r$.▶

Лемма 3.3. При элементарных преобразованиях строк матрицы ранг системы ее столбцов не меняется. Более того, если некоторая система столбцов является базой системы столбцов исходной матрицы, то та же система столбцов — база системы столбцов преобразованной матрицы.

◀ Пусть $j_1, \ldots, j_r$ — номера столбцов, образующих базу системы столбцов матрицы $A = a_{ij}$ размера $m \times n$. Это означает (см. определение базы), что система уравнений

$$\begin{array}{rcl} a_{1j_1}\lambda_1 + \ldots + a_{1j_r}\lambda_r & = & 0 \\ a_{2j_1}\lambda_1 + \ldots + a_{2j_r}\lambda_r & = & 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & & \\ a_{mj_1}\lambda_1 + \ldots + a_{mj_r}\lambda_r & = & 0 \end{array}$$

имеет только нулевое решение (линейная независимость) и что для любого $j = 1, 2, \ldots, n$ система уравнений

$$\begin{array}{rcl} a_{1j_1}\lambda_1 + \ldots + a_{1j_r}\lambda_r & = & a_{1j} \\ a_{2j_1}\lambda_1 + \ldots + a_{2j_r}\lambda_r & = & a_{2j} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & & \\ a_{mj_1}\lambda_1 + \ldots + a_{mj_r}\lambda_r & = & a_{mj} \end{array}$$

совместна (выражение столбцов через столбцы базы). Но элементарным преобразованиям системы строк матрицы соответствуют элементарные преобразования каждой из этих систем уравнений, следовательно, ввиду предложения 1.5, эти условия сохраняются для тех же столбцов преобразованной матрицы.▶

◀ **Доказательство теоремы 3.1.** Приведем произвольную матрицу A размера $m \times n$ к ступенчатому виду. Согласно доказанным леммам, ранги систем строк и столбцов матрицы A при этом не изменятся, т.е. **можно считать, что сама матрица A имеет ступенчатый вид**. Пусть в ней r ненулевых строк и $j_1, \ldots, j_r$ — номера столбцов, содержащих ведущие элементы этих строк. Покажем, что ненулевые строки ступенчатой матрицы линейно независимы. Действительно, рассмотрим их линейную комбинацию

$$
\begin{array}{rccccccc}
 & & j_1 & & j_2 & \dots & j_r & \\
\lambda_1(& 0\dots0 & a_{1j_1} & *\dots* & a_{1j_2} & *\dots* & a_{1j_r} & *\dots*)+ \\
\lambda_2(& 0\dots0 & 0 & 0\dots0 & a_{2j_2} & *\dots* & a_{2j_r} & *\dots*)+ \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
+\lambda_r(& 0\dots0 & 0 & 0\dots0 & 0 & 0\dots0 & a_{rj_r} & *\dots*) = \\
(& 0\dots0 & \lambda_1a_{1j_1} & *\dots* & \lambda_1a_{1j_2}+\lambda_2a_{2j_2} & *\dots* & \lambda_1a_{1j_r}+\dots+\lambda_ra_{rj_r} & *\dots*).
\end{array}
$$

Если последняя строка — нулевая, то получается ступенчатая система уравнений

$$
\begin{array}{rcl}
a_{rj_r}\lambda_r+\dots+a_{1j_r}\lambda_1 &=& 0 \\
\dots\dots\dots\dots\dots\dots\dots\dots & & \\
\lambda_1a_{1j_2}+\lambda_2a_{2j_2} &=& 0 \\
a_{1j_1}\lambda_1 &=& 0
\end{array}
$$

относительно неизвестных $\lambda_r,\dots,\lambda_1$, которая, по следствию 1.15, определена, т.е. имеет только нулевое решение. Второе условие из определения базы очевидно: все строки, кроме первых r, нулевые, и, значит, линейно выражаются через первые r строк. Итак, ранг системы строк матрицы A равен r.

Рассмотрим теперь линейную комбинацию столбцов с номерами $j_1,\dots,j_r$:

$$
\lambda_1\begin{pmatrix} a_{1j_1}\\0\\\vdots\\0\\\hline 0\\\vdots\\0\end{pmatrix}+\lambda_2\begin{pmatrix} a_{1j_2}\\a_{2j_2}\\\vdots\\0\\\hline 0\\\vdots\\0\end{pmatrix}+\dots+\lambda_r\begin{pmatrix} a_{1j_r}\\a_{2j_r}\\\vdots\\a_{rj_r}\\\hline 0\\\vdots\\0\end{pmatrix}=\begin{pmatrix} \lambda_1a_{1j_1}+\lambda_2a_{2j_2}+\dots+\lambda_ra_{1j_r}\\\lambda_2a_{2j_2}+\dots+\lambda_ra_{2j_r}\\\vdots\\\lambda_ra_{rj_r}\\\hline 0\\\vdots\\0\end{pmatrix}.
$$

Значит, для любых чисел $b_1,\dots,b_r$ система уравнений

$$
\begin{array}{rcl}
\lambda_1a_{1j_1}+\lambda_2a_{2j_2}+\dots+\lambda_ra_{1j_r} &=& b1 \\
\lambda_2a_{2j_2}+\dots+\lambda_ra_{2j_r} &=& b_2 \\
\dots\dots\dots\dots\dots\dots\dots\dots & & \\
\lambda_ra_{rj_r} &=& b_r \\
\hline 0 &=& 0 \\
\dots\dots\dots\dots\dots\dots\dots\dots & & \\
0 &=& 0
\end{array}
$$

является определенной. В частности, при $b_1=b_2=\dots=0$ получаем $\lambda_1=\dots=\lambda_r=0$, т.е. данная система столбцов линейно независима, а если в качестве $b_1,\dots,b_r$ взять первые r элементов любого столбца матрицы A, то получим выражение этого столбца через столбцы с номерами $j_1,\dots,j_r$. ▶

Заметим, что приведённое доказательство дает алгоритм нахождения базы любой системы столбцов: достаточно составить из них матрицу, привести ее к ступенчатому виду элементарными преобразованиями строк и выбрать столбцы, которые содержат ведущие элементы строк полученной ступенчатой матрицы. А для системы строк можно поступить так: записать их элементы в матрицу по столбцам и применить то же действие.

Определение 3.4. *Рангом матрицы* называется ранг системы ее строк (или, равносильно, ранг системы ее столбцов. Ранг матрицы A обозначим $\mathrm{rk}(A)$.

2°. Критерий совместности и определенности системы линейных уравнений.

Теорема 3.5. Система линейных уравнений совместна тогда и только тогда, когда ранги матрицы коэффициентов системы и расширенной матрицы системы равны.

◀ Рассмотрим систему уравнений (1.1) с матрицей коэффициентов A и расширенной матрицей $\overline{A}$. Пусть система совместна, и $x^0 = (x_1^0, x_2^0, \ldots, x_n^0)$ — некоторое ее решение. В векторной записи это означает, что

$$x_1^0 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2^0 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n^0 \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}. \tag{3.1}$$

Таким образом, система столбцов матрицы $\overline{A}$ линейно выражается через систему столбцов матрицы A, поэтому $\mathrm{rk}(A) \geqslant \mathrm{rk}(\overline{A})$. Обратное неравенство очевидно, значит $\mathrm{rk}(A) = \mathrm{rk}(\overline{A})$.

Обратно, допустим, что $\mathrm{rk}(A) = \mathrm{rk}(\overline{A}) = r$. Тогда можно выбрать линейно независимую подсистему из r столбцов матрицы A, и по теореме 2.14 эта подсистема — база системы столбцов расширенной матрицы $\overline{A}$. Но тогда столбец правых частей — линейная комбинация столбцов этой подсистемы и тем более — линейная комбинация столбцов матрицы A, что, как мы видели в записи (3.1), равносильно совместности системы уравнений.▶

Теорема 3.6. Совместная система линейных уравнений является определенной тогда и только тогда, когда ранг ее матрицы равен числу неизвестных.

◀ Условие определенности совместной системы уравнений равносильно тому, что все неизвестные — главные. Число главных неизвестных равно числу ненулевых строк после приведения матрицы системы к ступенчатому виду и, следовательно, равно рангу матрицы системы.▶

3°. Фундаментальная система решений однородной системы линейный уравнений.

Определение 3.7. Система линейных уравнений называется однородной, если правая часть каждого уравнения системы равна 0.

Рассмотрим однородную систему линейных уравнений

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & 0 \end{array}. \tag{3.2}$$

Очевидно: однородная система совместна (решение: $x_1 = x_2 = \ldots = x_n = 0$).

Заметим, что множество всех решений системы (3.2) замкнуто относительно сложения векторов и умножения вектора на число (позже мы проверим это средствами матричной алгебры). Следовательно, любая линейная комбинация решений однородной системы (3.2) снова является решением той же системы.

Иначе говоря, множество решений однородной линейной системы уравнений — *подпространство* арифметического векторного пространства $\mathbb{R}^n$.

Определение 3.8. *Фундаментальной системой решений* однородной системы линейных уравнений называется база множества всех решений этой системы.

Теорема 3.9. Число решений в фундаментальной системе решений однородной линейной системы из m уравнений от n переменных равно $n - r$, где r — ранг матрицы системы.

◄ Рассмотрим систему (3.2). Можно считать, что она приведена к ступенчатому виду. Предположим для упрощения записи, что первые r неизвестных являются главными. Тогда с помощью обратного хода метода Гаусса их можно выразить через свободные неизвестные:

$$\begin{array}{rcl} x_1 & = & \alpha_{11}x_{r+1} + \ldots + \alpha_{1,n-r}x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & & . \\ x_r & = & \alpha_{r1}x_{r+1} + \ldots + \alpha_{r,n-r}x_n \end{array} \tag{3.3}$$

Теперь можно построить $n - r$ решений системы (3.2) вида

$$\begin{array}{rcl} x^1 & = & (\underbrace{*, \ldots, *}_{r}, \ 1, \ 0, \ \ldots, \ 0), \\ x^2 & = & (\underbrace{*, \ldots, *}_{r}, \ 0, \ 1, \ \ldots, \ 0), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & & \\ x^{n-r} & = & (\underbrace{*, \ldots, *}_{r}, \ 0, \ 0, \ \ldots, \ 1), \end{array}$$

где первые r координат определяются в соответствии с (3.3). Заметим, что указанные решения линейно независимы (см. доказательство теоремы 3.1). Покажем, что произвольное решение $x^0 = (x_1^0, \ldots, x_n^0)$ системы (3.2) является линейной комбинацией указанных. Действительно, вектор $x^0 - (x_{r+1}^0 x^1 + \ldots + x_n^0 x^{n-r})$ удовлетворяет (3.3), и его последние $n - r$ координат равны 0. Значит, и первые r координат этого вектора равны 0, т.е.

$$x^0 = x_{r+1}^0 x^1 + \ldots + x_n^0 x^{n-r},$$

что и утверждалось.►

Утверждение теоремы 3.9 можно переформулировать так:
Общее решение однородной системы уравнений (3.2) *имеет вид*

$$x = C_1 x^1 + \ldots + C_{n-r} x^{n-r},$$

где $x^1, \ldots, x^{n-r}$ *— произвольная фундаментальная система решений этой системы уравнений, а* $C_1, \ldots, C_r$ *— произвольные числа.*

Лекция 4. Группа подстановок конечного множества, знак подстановки (четность), знакопеременная группа, разложение подстановки в произведение транспозиций и независимых циклов.

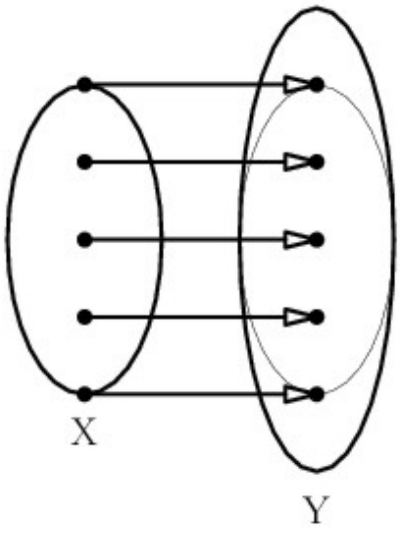
1°. Группа подстановок. Напомним (см. курс математического анализа), что отображение $f : X \to Y$ множества X в множество Y называется *инъективным*, если различным элементам множества X соответствуют различные элементы множества Y, т.е.

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

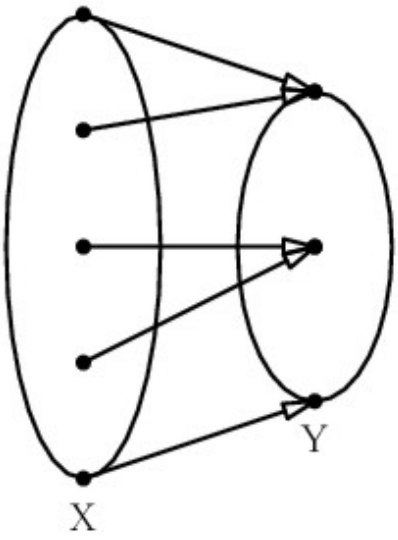
Соответственно, отображение $f : X \to Y$ множества X в множество Y называется *сюръективным*, если для любого элемента $y \in Y$ существует хотя бы один его *прообраз* в X, т.е. такой элемент $x \in X$, что $y = f(x)$:

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

Графически эти ситуации можно изобразить так:



Инъективное отображение



Сюръективное отображение

Отображение называется *биективным*, если оно инъективно и сюръективно одновременно. Биективное отображение часто также называют *взаимно-однозначным соответствием*.

Определение 4.1. *Группой подстановок* на конечном множестве X называется множество всех биективных отображений (*подстановок*) множества X в себя. Обозначается эта группа через S_X. Под *произведением* подстановок понимается их композиция как отображений:

$$\forall \sigma, \tau \in S_X : \forall x \in X, (\sigma\tau)(x) = \sigma(\tau(x)).$$

Если X содержит n элементов, то их можно занумеровать числами от 1 до n, т.е. считать, что $X = \{1, 2, \ldots, n\}$. В этом случае группа подстановок обозначается S_n и называется *симметрической группой степени* n.

Заметим, что любую подстановку из S_n можно задать таблицей вида

$$\sigma = \begin{pmatrix} i_1 & i_2 & \dots & i_n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}, \tag{4.1}$$

где $j_k = \sigma(i_k),\ k = 1, \dots, n$. Две такие таблицы определяют одну и ту же подстановку, если и только если они отличаются лишь порядком столбцов. Если верхняя строка в (4.1) имеет вид $(1\ 2\ \dots\ n)$, то говорят, что это *стандартная* таблица, задающая данную подстановку.

Предложение 4.2. Число элементов группы S_n равно $n! = 1 \cdot 2 \cdot \ldots \cdot n$.

◀ Возможное число образов элемента 1 равно n. Если образ 1 выбран, то для образа 2 остается $n-1$ возможность и т.д.▶

Отметим следующие свойства произведения подстановок:
1) произведение подстановок *ассоциативно*, т.е.

$$\forall \sigma, \tau, \rho \in S_n : \quad (\sigma\tau)\rho = \sigma(\tau\rho);$$

2) для *тождественной* подстановки $\varepsilon = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}$ и для любой подстановки $\sigma \in S_n$ выполняется равенство $\varepsilon\sigma = \sigma\varepsilon = \sigma$;
3) Для любой подстановки $\sigma \in S_n$ существует *обратная* подстановка σ^{-1}, для которой

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

Обратная подстановка к подстановке, заданной (4.1), определяется "перевернутой" таблицей $\begin{pmatrix} j_1 & j_2 & \dots & j_n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$.

Выполнение этих трех условий показывает, что S_n является группой относительно введенной операции (понятие *группы* мы определим позже).

2°. Четные и нечетные подстановки.

Определение 4.3. Рассмотрим строку $(i_1, i_2, ..., i_n)$, в которой каждое число от 1 до n встречается один раз (такую строку естественно назвать *перестановкой* множества $\{1, 2, \dots, n\}$). Определим *инверсию* в такой строке как пару индексов $k < l$, для которых $i_k > i_l$.

Предложение 4.4. Четность общего числа инверсий в верхней и нижней строках таблицы (4.1) не меняется при произвольной перестановке её столбцов.

◀ Заметим, что с помощью последовательности перестановок соседних столбцов можно любую таблицу вида (4.1) с заданным набором столбцов перевести в любую другую таблицу с тем же набором столбцов. Но при перестановке соседних столбцов $\begin{pmatrix} i_k \\ j_k \end{pmatrix}$ и $\begin{pmatrix} i_{k+1} \\ j_{k+1} \end{pmatrix}$ возможны 4 случая:
1) $i_k < i_{k+1},\ j_k < j_{k+1}$ — тогда при перестановке общее число инверсий увеличивается на 2;
2) $i_k < i_{k+1},\ j_k > j_{k+1}$ — тогда при перестановке общее число инверсий не меняется ;
3) $i_k > i_{k+1},\ j_k < j_{k+1}$ — тогда при перестановке общее число инверсий не меняется ;
4) $i_k > i_{k+1},\ j_k > j_{k+1}$ — тогда при перестановке общее число инверсий уменьшается на 2.
В любом случае четность общего числа инверсий не меняется.▶

Значит, можно дать такое

Определение 4.5. Подстановка σ, заданная таблицей (4.1), называется *четной*, если общее число инверсий в верхней и нижней строках таблицы — четное число, и *нечетной* в противном случае.

Если подстановка задана стандартной таблицей, то для определения четности достаточно найти число инверсий в нижней строке — ведь в этом случае в верхней строке инверсий нет.

Определение 4.6. *Знаком* подстановки σ называется число

$$\operatorname{sgn}(\sigma) = \begin{cases} 1, & \text{если } \sigma\text{— четная подстановка,} \\ -1, & \text{если } \sigma\text{— нечетная подстановка.} \end{cases} \tag{4.2}$$

Встречается также обозначение знака подстановки $\operatorname{sgn}(\sigma) = (-1)^\sigma$.

Определение 4.7. Транспозицией чисел i и j $(i \neq j)$ назвается подстановка (обозначается (i, j)), переставляющая эти числа и оставляющая остальные числа на месте, иначе говоря,

$$(i,j)(k) = \begin{cases} j, & \text{если } k = i, \\ i, & \text{если } k = j, \\ k, & \text{если } k \neq i,\ k \neq j. \end{cases}$$

Теорема 4.8. Любая подстановка — произведение некоторого числа транспозиций (тождественную подстановку считаем произведением пустого множества транспозиций).

◀ Рассмотрим стандартную таблицу, представляющую подстановку $\sigma \in S_n$:

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

Выберем первое число k, для которого $l = i_k \neq k$ (если такого числа нет, то $\sigma = \varepsilon$). Заметим, что при этом $l > k$, иначе получили бы противоречие: $\sigma(l) = l = \sigma(k)$. Положим $\sigma_1 = (k, l)\sigma$. Тогда уже $\sigma_1(k) = (k, l)(\sigma(k)) = (k, l)(l) = k$, причем $\sigma = (k, l)\sigma_1$, так как $(k, l)^2 = \varepsilon$. Применяя то же рассуждение к σ_1, получим следующую подстановку σ_2 и т.д., причем количество чисел, начиная с 1, переводимых в себя, будет расти, пока не останется тождественная подстановка.▶

Теорема 4.9. При умножении подстановки (слева или справа) на транспозицию четность подстановки меняется на противоположную.

◀ Рассмотрим сначала умножение на транспозицию слева. Пусть

$$\sigma = \begin{pmatrix} i_1 & i_2 & \dots & i_n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}.$$

Сначала предположим, что транспозиция $\tau = (j_k, j_{k+1})$ переставляет **соседние** числа из второй строки. Легко видеть, что

$$\tau\sigma = \begin{pmatrix} \dots & i_k & i_{k+1} \dots & \\ \dots & j_{k+1} & j_k & \dots \end{pmatrix}.$$

Поэтому число инверсий в таблице $\tau\sigma$ на 1 больше (при $j_k < j_{k+1}$) или на 1 меньше, чем в таблице подстановки σ, т.е. в этом случае $\operatorname{sgn}(\tau\sigma) = -\operatorname{sgn}(\sigma)$. Теперь рассмотрим общий случай: транспозиция τ переставляет числа j_k и j_l. Пусть для определенности $l > k$. Тогда можно заменить умножение на τ последовательностью перестановок соседних элементов: например, сначала передвинуть j_l на место j_{k+1} ($l - k - 1$ транспозиция), затем поменять местами j_k и j_{k+1}, и, наконец, вернуть j_l обратно (еще $l - k - 1$ транспозиция). Всего получилось $(l - k - 1) + 1 + (l - k - 1)$ транспозиций, т.е. нечетное число. Значит, для умножения слева теорема верна.

Доказательство для умножения справа аналогично, но использует первую строку таблицы, задающей подстановку.▶

Следствие 4.10. Произведение четного (нечетного) числа транспозиций — четная (нечетная) подстановка.

◀ Следует из теоремы 4.9 и того, что ε — четная подстановка:

$$\operatorname{sgn}(\tau_1\tau_2\dots\tau_r) = \operatorname{sgn}(\tau_1\tau_2\dots\tau_r\varepsilon) = (-1)^r\operatorname{sgn}(\varepsilon) = (-1)^r.$$

▶

Следствие 4.11. Для любых подстановок $\sigma, \rho \in S_n$,

$$\operatorname{sgn}(\sigma\rho) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\rho).$$

◀ Разложим, согласно 4.8, обе подстановки в произведение транспозиций:

$$\sigma = \tau_1\tau_2\dots\tau_k$$

и

$$\rho = \tau_{k+1}\tau_{k+2}\dots\tau_{k+l}.$$

Тогда

$$\operatorname{sgn}(\sigma\rho) = (-1)^{k+l} = (-1)^k(-1)^l = \operatorname{sgn}(\sigma)\operatorname{sgn}(\rho).$$

▶

Замечание. Из 4.11 вытекает, что множество всех четных подстановок в S_n образует группу относительно усмножения подстановок.

Определение 4.12. Множество всех четных подстановок степени n обозначается A_n и называется *знакопеременной* группой степени n.

Предложение 4.13. Количество четных подстановок в S_n при $n \geqslant 2$ равно числу нечетных подстановок и равно $n!/2$.

◀ Рассмотрим отображение $f : S_n \to S_n$, такое, что $f(\sigma) = (1,2)\sigma$. Оно биективно и переводит четные подстановки в нечетные и наоборот, нечетные — в четные, по теореме 4.9. Значит, количество тех и других одинаково. Осталось вспомнить, что всего подстановок данной степени $n!$.▶

3°. Разложение подстановок в произведение независимых циклов.[1]

Определение 4.14. Пусть $i_1, \dots, i_k$ — попарно различные числа из множества $\{1, \dots, n\}$. *Циклом* длины k на элементах $i_1, \dots, i_k$ называется подстановка степени n, переводящая i_1 в i_2, i_2 в i_3,..., i_{k-1} в i_k, i_k в i_1, а остальные элементы — в самих себя. Такой цикл обозначается так: $(i_1, \dots, i_k)$.

По этому определению, циклы длины 2 — то же самое, что транспозиции. Часто по формальным соображением удобно считать любой цикл длины 1 тождественной подстановкой.

Определение 4.15. *Носителем* подстановки $\sigma \in S_n$ называется множество

$$\operatorname{supp}(\sigma) = \{i \in \{1, \dots, n\} : \sigma(i) \neq i\}.$$

Две подстановки называются *независимыми*, если пересечение их носителей пусто.

[1] Предполагается рассмотреть эту тему на семинарах.

Предложение 4.16. Если подстановки $\sigma, \rho \in S_n$ независимы, то $\sigma\rho = \rho\sigma$.

◀ Пусть $i \in \{1, \ldots, n\}$ и $i \in \operatorname{supp}\sigma$. Тогда $\sigma(i) \in \operatorname{supp}(\sigma)$, иначе $\sigma(\sigma(i)) = \sigma(i)$, откуда, в силу биективности σ, получаем $\sigma(i) = i$, что невозможно. Следовательно, $i \notin \operatorname{supp}(\rho)$ и $\sigma(i) \notin \operatorname{supp}(\rho)$, поэтому имеем

$$\sigma\rho(i) = \sigma(i), \quad \rho\sigma(i) = \sigma(i).$$

Аналогично проверяется случай, когда $i \in \operatorname{supp}(\rho)$. Наконец, если $i \notin \operatorname{supp}(\sigma) \cup \operatorname{supp}(\rho)$, то

$$\sigma\rho(i) = \sigma(i) = i = \rho(i) = \rho\sigma(i).$$

▶

Предложение 4.17. Если τ — цикл длины k, то $\operatorname{sgn}(\tau) = (-1)^{k-1}$.

◀ Непосредственно проверяется соотношение

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

правая часть которого содержит $k - 1$ транспозиций.▶

Теорема 4.18. Любая подстановка может быть разложена в произведение незывисимых циклов, причем это разложение единственно, с точностью до порядка сомножителей.

◀ Пусть $\sigma \in S_n$. Выберем какой-то элемент $i \in \{1, \ldots, n\}$ и рассмотрим последовательность $i, \sigma(i), \sigma^2(i), \ldots$. Поскольку все ее элементы лежат в конечном множестве, найдутся степени k и $l > k$ такие, что $\sigma^k(i) = \sigma^l(i)$. Но из биективности σ получим, что $i = \sigma^r(i)$, где $r = l - k > 0$. Выбирая наименьшее положительное r с таким свойством, получим, что $(i, \sigma(i), \ldots, \sigma^{r-1}(i))$ — цикл, действие которого на его носителе совпадает с действием σ. Далее возьмем произвольный элемент, не лежащий в носителе данного цикла, и построим новый цикл. Будем повторять эти действия, пока не исчерпаем множество $\{1, \ldots, n\}$. Единственность разложения следует из того, что два числа i, j принадлежат носителю одного и того же цикла тогда и только тогда, когда существует показатель k, для которого $j = \sigma^k(i)$.▶

Лекция 5. Определитель квадратной матрицы, его основные свойства (линейность, кососимметричность, определитель транспонированной матрицы). Изменение определителя при элементарных преобразованиях строк (столбцов) матрицы. Определитель треугольной матрицы. Критерий равенства определителя нулю.

1°. Общее понятие определителя.

Определение 5.1. *Определителем* квадратной матрицы порядка n называется алгебраическая сумма всевозможных произведений элементов матрицы, взятых по одному из каждой строки и каждого столбца, причем каждое такое произведение умножается на знак подстановки, которая номеру каждой строки ставит в соответствие номер столбца элемента из этой строки, входящего в данное произведение.

То же самое можно описать формулой

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}. \tag{5.1}$$

Определитель квадратной матрицы $A = (a_{ij})$ часто обозначается $|A|$ или $\det(A)$. Чтобы разобраться в формуле (5.1), посмотрим, как она выглядит в случае $n = 2$ и $n = 3$.

При $n = 2$ в S_2 всего две подстановки: $\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ и $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$.

Первая — четная, соответствующее произведение $a_{11}a_{22}$ входит в сумму со знаком "+".

Вторая — нечетная, соответствующее произведение $a_{12}a_{21}$ входит в сумму со знаком "–".

Получаем

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

что совпадает с выражением, приведенным ранее.

При $n = 3$ имеем уже 6 подстановок и, соответственно, 6 произведений элементов матрицы

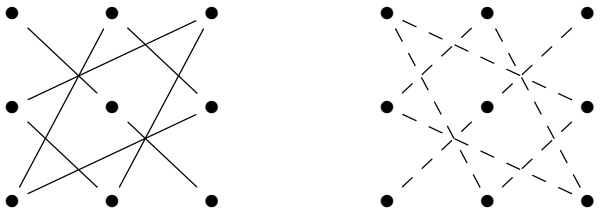
$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}:$$

Подстановка	Знак	Произведение
$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$	$+$	$a_{11}a_{22}a_{33}$
$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$	$-$	$a_{11}a_{23}a_{32}$
$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$	$-$	$a_{12}a_{21}a_{33}$
$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$	$+$	$a_{13}a_{21}a_{32}$
$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$	$+$	$a_{12}a_{23}a_{31}$
$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$	$-$	$a_{13}a_{22}a_{31}$

Таким образом,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \qquad (5.2)$$

Разумеется, запоминать формулу (5.2) не обязательно. Лучше запомнить следующие рисунки, выражающие *правило треугольников:*



$+$ $-$

На основании определения (5.1) легко вычислить определитель произвольной *треугольной матрицы*:

Определение 5.2. Квадратная матрица $A = (a_{ij})$ порядка n называется *верхнетреугольной*, если все элементы, расположенные ниже главной диагонали $a_{11}...a_{nn}$, равны нулю.

Иначе говоря, верхнетреугольная матрица имеет вид

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,n-1} & a_{nn} \\ 0 & a_{22} & \dots & a_{2,n-1} & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix}.$$

Можно также определить верхнетреугольную матрицу условием $a_{ij} = 0$ при $i > j$.
Аналогично определяется нижнетреугольная матрица.

Предложение 5.3. Определитель треугольной матрицы равен произведению ее диагональных элементов.

◀ Легко видеть, что единственное ненулевое произведение в сумме (5.1) — это произведение диагональных элементов, поскольку из первого столбца в такое произведение может входить только a_{11}, из второго — a_{22}, так как a_{12} вместе с a_{11} в произведение входить не может и т.д. А произведение элементов главной диагонали соответствует тождественной подстановке, которая является четной. Значит, оно входит в определитель со знаком "+".▶

2°. Основные свойства определителей. При описании основных свойств определителя мы будем часто обозначать строки квадратной матрицы $A = (a_{ij})$ порядка n символами $a^1, \dots, a^n$, считая $a^i = \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix}$. В этих обозначениях будем записывать матрицу A как

$$A = \begin{pmatrix} a^1 \\ a^2 \\ \vdots \\ a^n \end{pmatrix}.$$

Определение 5.4. Функция $f(x^1, x^2, \ldots, x^n)$ от n векторов со значениями в R называется *линейной по i-му аргументу*, если

$$\begin{aligned} &f(x^1, \ldots, x^{i-1}, y^i + z^i, x^{i+1}, \ldots, x^n) = \\ &\qquad f(x^1, \ldots, x^{i-1}, y^i, x^{i+1}, \ldots, x^n) + f(x^1, \ldots, x^{i-1}, z^i, x^{i+1}, \ldots, x^n), \\ &f(x^1, \ldots, x^{i-1}, \lambda x^i, x^{i+1}, \ldots, x^n) = \lambda f(x^1, \ldots, x^{i-1}, x^i, x^{i+1}, \ldots, x^n). \end{aligned} \tag{5.3}$$

Такая функция называется *полилинейной*, если она линейна по каждому аргументу.

Функция $f(x^1, x^2, \ldots, x^n)$ называется *симметричной*, если

$$f(x^1, \ldots, x^i, \ldots, x^j, \ldots, x^n) = f(x^1, \ldots, x^j, \ldots, x^i, \ldots, x^n)$$

для любых $i, j = 1, \ldots, n$.

Функция $f(x^1, x^2, \ldots, x^n)$ называется *кососимметричной*, если

$$f(x^1, \ldots, x^i, \ldots, x^j, \ldots, x^n) = -f(x^1, \ldots, x^j, \ldots, x^i, \ldots, x^n) \tag{5.4}$$

для любых $i, j = 1, \ldots, n,\ i \neq j$.

Теорема 5.5. Определитель квадратной матрицы — полилинейная и кососимметричная функция ее строк.

◀ Допустим, что i-я строка матрицы $A = (a_{ij})$ является суммой некоторых двух строк (с матрицей A никак не связанных!): $a^i = b^i + c^i$, т. е. $a_{ij} = b_{ij} + c_{ij}$, $j = 1, \ldots, n$. Тогда

$$\begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \ldots\ldots & \ldots & \ldots\ldots \\ a_{i1} & \ldots & a_{in} \\ \ldots\ldots & \ldots & \ldots\ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} \overset{(5.1)}{=} \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)} =$$

$$\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots (b_{i\sigma(i)} + c_{i\sigma(i)}) \cdots a_{n\sigma(n)} =$$

$$\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots b_{i\sigma(i)} \cdots a_{n\sigma(n)} + \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots c_{i\sigma(i)} \cdots a_{n\sigma(n)} =$$

$$\begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \ldots\ldots & \ldots & \ldots\ldots \\ b_{i1} & \ldots & b_{in} \\ \ldots\ldots & \ldots & \ldots\ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \ldots\ldots & \ldots & \ldots\ldots \\ c_{i1} & \ldots & c_{in} \\ \ldots\ldots & \ldots & \ldots\ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}.$$

Аналогично, если $a^i = \lambda b^i$, т. е. $a_{ij} = \lambda b_{ij}$, $j = 1, \ldots, n$, то

$$\begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \ldots\ldots & \ldots & \ldots\ldots \\ a_{i1} & \ldots & a_{in} \\ \ldots\ldots & \ldots & \ldots\ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} \overset{(5.1)}{=} \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)} =$$

$$\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots (\lambda b_{i\sigma(i)}) \cdots a_{n\sigma(n)} =$$

$$\lambda \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots b_{i\sigma(i)} \cdots a_{n\sigma(n)} = \lambda \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \ldots\ldots & \ldots & \ldots\ldots \\ b_{i1} & \ldots & b_{in} \\ \ldots\ldots & \ldots & \ldots\ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}.$$

Таким образом, доказана линейность определителя. Рассмотрим, далее, определитель

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots\dots\dots\dots \\ a_{j1} & \dots & a_{jn} \\ \dots\dots\dots\dots \\ a_{i1} & \dots & a_{in} \\ \dots\dots\dots\dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{j\sigma(i)} \cdots a_i\sigma(j) \cdots a_{n\sigma(n)} =$$

$$\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1\sigma\tau(1)} \cdots a_{i\sigma\tau(i)} \cdots a_{j\sigma\tau(j)} \cdots a_{n\sigma\tau(n)}, \quad (5.5)$$

где $\tau = (i, j)$ — транспозиция индексов i и j. Положим $\sigma' = \sigma\tau$. Тогда σ' тоже является произвольной подстановкой в S_n, причем из теоремы 4.9 следует, что $\operatorname{sgn}(\sigma) = -\operatorname{sgn}(\sigma')$. Следовательно, правую часть (5.5) можно записать так:

$$\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1\sigma\tau(1)} \cdots a_{i\sigma\tau(i)} \cdots a_{j\sigma\tau(j)} \cdots a_{n\sigma\tau(n)} =$$

$$- \sum_{\sigma'\in S_n} \operatorname{sgn}(\sigma') a_{1\sigma'(1)} \cdots a_{i\sigma'(i)} \cdots a_{j\sigma'(j)} \cdots a_{n\sigma'(n)} = - \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots\dots\dots\dots \\ a_{i1} & \dots & a_{jn} \\ \dots\dots\dots\dots \\ a_{j1} & \dots & a_{in} \\ \dots\dots\dots\dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix},$$

что и означает кососимметричность определителя.►

Следствие 5.6. Определитель, у которого одна строка нулевая или две строки совпадают, равен нулю.

◄ Первое утверждение следует из линейности определителя:

$$\begin{vmatrix} \dots\dots\dots \\ 0 \ \dots \ 0 \\ \dots\dots\dots \end{vmatrix} = \begin{vmatrix} \dots\dots\dots\dots\dots \\ 0\cdot 0 \ \dots \ 0\cdot 0 \\ \dots\dots\dots\dots\dots \end{vmatrix} = 0 \cdot \begin{vmatrix} \dots\dots\dots \\ 0 \ \dots \ 0 \\ \dots\dots\dots \end{vmatrix} = 0.$$

Второе утверждение вытекает из кососимметричности: если

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots\dots\dots\dots \\ a_{i1} & \dots & a_{in} \\ \dots\dots\dots\dots \\ a_{i1} & \dots & a_{in} \\ \dots\dots\dots\dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix},$$

то $|A| = -|A|$ (так как перестановка строк не меняет матрицы), откуда $|A| = 0$. ►

3°. Транспонирование матрицы. Определитель транспонированной матрицы.

Определение 5.7. Для произвольной матрицы $A = (a_{ij})$ размера $m \times n$ определим *транспонированную* матрицу $A^T = (a'_{i,j})$ размера $n \times m$, элементы которой определяются равенством $a'_{ij} = a_{ji}$, где $i = 1, \ldots, n$, $j = 1, \ldots, m$. Иными словами, строки матрицы A образуют столбцы матрицы A^T, и наоборот. Еще можно сказать, что матрица A^T получается зеркальным отражением матрицы A относительно главной диагонали.

Легко заметить, что $(A^T)^T = A$ для любой матрицы A и $(\alpha A + \beta B)^T = \alpha A^T + \beta B^T$ для любых матриц A и B одинакового размера и любых чисел α, β.

Теорема 5.8. Для любой квадратной матрицы A выполняется равенство $|A^T| = |A|$.

◀ Пусть $A = (a_{ij})$, $A^T = (a'_{i,j})$. Тогда

$$|A^T| = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a'_{1\sigma(1)} a'_{2\sigma(2)} \cdots a'_{n\sigma(n)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} =$$
$$\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma^{-1}(1)} a_{2\sigma^{-1}(2)} \cdots a_{n\sigma^{-1}(n)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) a_{1\sigma^{-1}(1)} a_{2\sigma^{-1}(2)} \cdots a_{n\sigma^{-1}(n)} = |A|,$$

поскольку для любой подстановки $\sigma \in S_n$ выполнено равенство $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$.▶

Из теоремы 5.8 следует такой общий вывод: любые утверждения об определителях остаются справедливыми, если заменить строки на столбцы. Например, определитель является полилинейной кососимметричной функцией столбцов матрицы.

4°. Изменение определителя матрицы при элементарных преобразованиях строк (столбцов).

Теорема 5.9. При элементарных преобразованиях типа 1 строк (столбцов) матрицы определитель не меняется. При элементарных преобразованиях типа 2 строк (столбцов) матрицы определитель умножается на (-1). При элементарном преобразовании третьего типа (умножение строки или столбца на $\lambda \neq 0$) определитель матрицы умножается на число λ.

◀ Второе утверждение — прямое следствие кососимметричности определителя как функции строк (столбцов), а третье — линейности этой функции. Рассмотрим преобразование ЭП1:

$$A = \begin{pmatrix} a_{11} & \ldots & a_{nn} \\ \ldots & \ldots & \ldots \\ a_{i1} & \ldots & a_{in} \\ \ldots & \ldots & \ldots \\ a_{j1} & \ldots & a_{jn} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \rightsquigarrow A' = \begin{pmatrix} a_{11} & \ldots & a_{nn} \\ \ldots & \ldots & \ldots \\ a_{i1} + \lambda a_{j1} & \ldots & a_{in} + \lambda a_{jn} \\ \ldots & \ldots & \ldots \\ a_{j1} & \ldots & a_{jn} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}.$$

Тогда в силу свойства линейности,

$$|A'| = \begin{vmatrix} a_{11} & \ldots & a_{nn} \\ \ldots & \ldots & \ldots \\ a_{i1} & \ldots & a_{in} \\ \ldots & \ldots & \ldots \\ a_{j1} & \ldots & a_{jn} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} + \lambda \begin{vmatrix} a_{11} & \ldots & a_{nn} \\ \ldots & \ldots & \ldots \\ a_{j1} & \ldots & a_{jn} \\ \ldots & \ldots & \ldots \\ a_{j1} & \ldots & a_{jn} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix},$$

причем второе слагаемое равно 0 в силу следствия 5.6. ▶

5°. Критерий равенства определителя нулю.

Теорема 5.10. Определитель квадратной матрицы равен 0 тогда и только тогда, когда ее строки (столбцы) линейно зависимы.

◄ Линейная зависимость строк (столбцов) квадратной матрицы A порядка n равносильна неравенству $\mathrm{rk}(A) < n$. Приводя матрицу A к ступенчатому виду элементарными преобразованиями строк типа 1, получим матрицу A', ранг и определитель которой равны, соответственно, рангу и определителю матрицы A. При этом, по теореме 3.1, $\mathrm{rk}(A')$ равен числу ненулевых строк матрицы A'. Следовательно, если $\mathrm{rk}(A') < n$, то в матрице A' имеется нулевая строка, и по следствию 5.6, $|A'| = 0$, откуда $|A| = 0$. Обратно, если в матрице A' нет нулевых строк, то лидеры всех строк расположены на главной диагонали, и $|A'| \neq 0$, так как равен произведению ненулевых диагональных элементов. Поэтому и $|A| \neq 0$. ►

Лекция 6. Определитель матрицы с углом нулей. Определитель Вандермонда. Миноры и алгебраические дополнения элементов. Разложние определителя по строке (столбцу). Лемма о "фальшивом" разложении определителя. Формулы Крамера для решения определенных квадратных систем линейных уравнений.

1°. Определитель матрицы с углом нулей. Рассмотрим *матрицу с углом нулей*, т.е. квадратную матрицу вида

$$A = \begin{pmatrix} b_{11} & \dots & b_{1k} & * & \dots & * \\ b_{21} & \dots & b_{2k} & * & \dots & * \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ b_{k1} & \dots & b_{kk} & * & \dots & * \\ 0 & \dots & 0 & c_{11} & \dots & c_{1l} \\ 0 & \dots & 0 & c_{21} & \dots & c_{2l} \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & \dots & 0 & c_{l1} & \dots & c_{ll} \end{pmatrix}. \tag{6.1}$$

Здесь символы $*$ обозначают элементы, значения которых для нас несущественны. Сокращенно будем записывать A как *блочную* матрицу

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

где $B = (b_{ij})$ и $C = (c_{ij})$ — квадратные матрицы размеров $k \times k$ и $l \times l$, соответственно.

Теорема 6.1. Если $A = \begin{pmatrix} B & * \\ 0 & C \end{pmatrix}$ — матрица с углом нулей, то

$$|A| = \begin{vmatrix} B & * \\ 0 & C \end{vmatrix} = |B||C|. \tag{6.2}$$

◀ Приведем матрицу A из (6.1) к треугольному виду элементарными преобразованиями строк первого типа следующим образом: сначала преобразуем первые k строк так, чтобы матрица B стала треугольной, затем, не меняя первых k строк, преобразуем последние l строк так, чтобы матрица стала треугольной (при этом угол нулей сохранится). Получится преобразование

$$A = \begin{pmatrix} B & * \\ 0 & C \end{pmatrix} \rightsquigarrow \begin{pmatrix} B' & * \\ 0 & C' \end{pmatrix} = A',$$

где матрицы B' и C' — треугольные. При этом $|A'| = |A|$, $|B'| = |B|$, $|C'| = |C|$, по теореме 5.9, а по предложению 5.3 $|A'| = |B'||C'|$, так как каждый из этих определителей — произведение соответствующих диагональных элементов. Стало быть, $|A| = |A'| = |B'||C'| = |B||C|$.▶

Из свойства определителя транспонированной матрицы видно, что если заменить левый нижний угол нулей на правый верхний, то равенство (6.2) даст равенство

$$\begin{vmatrix} B & 0 \\ * & C \end{vmatrix} = |B||C| \tag{6.3}$$

2°. Определитель Вандермонда. Определителем Вандермонда порядка n называется определитель вида

$$V(x_1,\dots,x_n)=\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \dots & \dots & \dots & \dots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix}. \tag{6.4}$$

где $x_1,\dots,x_n$ — произвольные числа, $n\geqslant 2$.

Теорема 6.2. Определитель Вандермонда равен произведению попарных разностей чисел $x_1,\dots,x_n$:

$$V(x_1,\dots,x_n)=\prod_{1\leqslant i<j\leqslant n}(x_j-x_i). \tag{6.5}$$

◄ Докажем (6.5) индукцией по n. При $n=2$ можно проверить непосредственно

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

Допустим справедливость (6.5) для определителей Вандермонда порядка $n-1$ и применим следующие элементарные преобразования строк к определителю 6.4: к каждой строке, начиная с последней, n-й, и кончая второй, прибавим предыдущую строку, умноженную на $(-x_1)$ (при этом значение определителя не изменится):

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \dots & \dots & \dots & \dots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix}=\begin{vmatrix} 1 & 1 & \dots & 1 \\ 0 & x_2-x_1 & \dots & x_n-x_1 \\ 0 & x_2^2-x_2x_1 & \dots & x_n^2-x_nx_1 \\ \dots & \dots & \dots & \dots \\ 0 & x_2^{n-1}-x_2^{n-2}x_1 & \dots & x_n^{n-1}-x_n^{n-2}x_1 \end{vmatrix}.$$

Получился определитель с углом нулей, значит, по теореме 6.1,

$$V(x_1,\dots,x_n)=1\cdot\begin{vmatrix} x_2-x_1 & \dots & x_n-x_1 \\ x_2(x_2-x_1) & \dots & x_n(x_n-x_1) \\ \dots & \dots & \dots \\ x_2^{n-2}(x_2-x_1) & \dots & x_n^{n-2}(x_n-x_1) \end{vmatrix}.$$

Из столбцов последнего определителя вынесем, соответственно, множители $(x_2-x_1),\dots(x_n-x_1)$:

$$V(x_1,\dots,x_n)=(x_2-x_1)\dots(x_n-x_1)\begin{vmatrix} 1 & \dots & 1 \\ x_2 & \dots & x_n \\ \dots & \dots & \dots \\ x_2^{n-2} & \dots & x_n^{n-2} \end{vmatrix}=(x_2-x_1)\dots(x_n-x_1)V(x_2,\dots,x_n).$$

Подставляя, по предположению индукции, значение $V(x_2,\dots,x_n)$, получаем

$$V(x_1,\dots,x_n)=(x_2-x_1)\dots(x_n-x_1)\prod_{2\leqslant i<j\leqslant n}(x_j-x_i)=\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).$$

►

3°. Миноры и алгебраические дополнения элементов определителя. Разложение определителя по строке (столбцу). Лемма о "фальшивом" разложении определителя.

Определение 6.3. *Минором* элемента a_{ij} определителя порядка n называется определитель порядка $n-1$, полученный вычеркиванием i-й строки и j-го столбца:

$$M_{ij} = \begin{vmatrix} a_{11} & \dots & a_{1,j-1} & a_{1,j} & a_{1,j+1} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i-1,1} & \dots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ \cancel{a_{i,1}} & \cancel{\dots} & \cancel{a_{i,j-1}} & \cancel{a_{ij}} & \cancel{a_{i,j+1}} & \cancel{\dots} & \cancel{a_{in}} \\ a_{i+1,1} & \dots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n,1} & \dots & a_{n,j-1} & a_{n,j} & a_{n,j+1} & \dots & a_{nn} \end{vmatrix}.$$

Алгебраическим дополнением элемента a_{ij} называется его минор, умноженный на $(-1)^{i+j}$:

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

Теорема 6.4 (Разложение определителя по строке (столбцу))**.** Определитель равен сумме произведений элементов любой его строки (любого столбца) на их алгебраические дополнения:

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{i1} & \dots & a_{in} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = a_{i1}A_{i1} + \dots + a_{in}A_{in}, \qquad i = 1, \dots, n. \tag{6.6}$$

$$\begin{vmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{vmatrix} = a_{1j}A_{1j} + \dots + a_{nj}A_{nj}, \qquad j = 1, \dots, n. \tag{6.7}$$

◄ Докажем равенство (6.6). Сначала представим i-ю строку определителя в виде суммы "простейших" строк:

$$(a_{i1}\ a_{i2}\ \dots\ a_{in}) = (a_{i1}\ 0\ \dots\ 0) + (0\ a_{i2}\ \dots\ 0) + \dots + (0\ 0\ \dots\ a_{in}).$$

В силу линейности получим

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{i1} & \dots & a_{in} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{i1} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ 0 & a_{i2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \dots + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}. \tag{6.8}$$

Теперь рассмотрим отдельно j-е слагаемое в (6.8). Переставляя соседние строки $i-1$ раз, начиная с i-й, поместим i-ю строку на место первой, не меняя порядка остальных строк:

$$\begin{vmatrix} a_{11} & \dots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & a_{ij} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \dots & a_{1n} \end{vmatrix} = (-1)^{i-1} \begin{vmatrix} 0 & \dots & 0 & a_{ij} & 0 & \dots & 0 \\ a_{11} & \dots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \dots & a_{1n} \end{vmatrix} \tag{6.9}$$

Аналогично, переставляя соседние столбцы последнего определителя $j-1$ раз, получим

$$\begin{vmatrix} 0 & \dots & 0 & a_{ij} & 0 & \dots & 0 \\ a_{11} & \dots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \dots & a_{nn} \end{vmatrix} = (-1)^{j-1} \begin{vmatrix} a_{ij} & 0 & \dots & 0 & 0 & \dots & 0 \\ a_{1j} & a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{nj} & a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{vmatrix}. \tag{6.10}$$

Комбинируя (6.9) и (6.10) и применяя теорему 6.1, получим

$$\begin{vmatrix} a_{11} & \dots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & a_{ij} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \dots & a_{1n} \end{vmatrix} = a_{ij} \underbrace{(-1)^{i-1}(-1)^{j-1}}_{=(-1)^{i+j}} M_{ij} = a_{ij}A_{ij}.$$

Суммируя по всем $j = 1, \dots, n$, получаем требуемое равенство (6.6). Равенство (6.7) доказывается аналогично. ►

Теорема 6.5 (Лемма о "фальшивом" разложении определителя)**.** Сумма произведений элементов некоторой строки (некоторого столбца) определителя на алгебраические дополнения соответствующих элементов **другой** строки (**другого** столбца) равна нулю. Иными словами, для определителя матрицы $A = (a_{ij})$ порядка n выполняются равенства

$$a_{i1}A_{k1} + a_{i2}A_{k2} + \dots + a_{in}A_{kn} = 0 \quad \text{при } i \neq k; \tag{6.11}$$

$$a_{1j}A_{1k} + a_{2j}A_{2k} + \dots + a_{nj}A_{nk} = 0 \quad \text{при } j \neq k; \tag{6.12}$$

Вспомним, что в "правильное" разложение (6.6) входят алгебраические дополнения элементов той же строки, по которой мы разлагаем определитель, в отличие от (6.11), с чем и связано название "фальшивое разложение".

◄ Заменим в определителе матрицы $A = (a_{ij})$ k-ю строку на i-ю. Получится определитель, равный нулю, так как две его строки совпадают. С другой стороны, вычислим его разложением по k-ой строке:

$$0 = \begin{matrix} \\ i \\ \\ k \\ \\ \end{matrix} \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{i1} & \dots & a_{in} \\ \dots & \dots & \dots \\ a_{i1} & \dots & a_{in} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = a_{i1}A_{k1} + \dots + a_{in}A_{kn},$$

причем алгебраические дополнения элементов k-й строки в исходном определителе и новом определителе совпадают, что и доказывает равенство (6.11). Для столбцов доказательство аналогично.►

4°. Формулы Крамера. Следующая теорема уточняет критерий совместности и определенности для *квадратных* систем линейных уравнений (т.е. систем, в которых число неизвестных равно числу уравнений). Пусть дана квадратная система уравнений

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n & = & b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots & & \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n & = & b_n \end{array}. \tag{6.13}$$

Определитель

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \tag{6.14}$$

называется определителем системы (6.13).

Теорема 6.6 (Теорема Крамера)**.** Квадратная система уравнений (6.13) является определенной тогда и только тогда, когда ее определитель Δ отличен от нуля.

◀ Мы знаем, что система (6.13) является определенной тогда и только тогда, когда $\mathrm{rk}(A) = \mathrm{rk}(\overline{A}) = n$, где A — матрица системы (6.13), а $\overline{A}$ — ее расширенная матрица (см. теоремы 3.5,3.6). Но для квадратной системы (6.13) всегда выполняется неравенство $\mathrm{rk}(A) \leqslant \mathrm{rk}(\overline{A}) \leqslant n$, значит, равенство $\mathrm{rk}(A) = n$ равносильно тому, что система (6.13) совместна и определена. Но условие $\mathrm{rk}(A) = n$ равносильно линейной независимости столбцов (или строк) матрицы A, что в силу критерия равенства определителя 0 (теорема 5.10) равносильно условию $\Delta = |A| \neq 0$.▶

Теорема 6.7 (Правило Крамера)**.** Пусть система уравнений (6.13) имеет определитель $\Delta \neq 0$. Положим

$$\Delta_i = \begin{vmatrix} a_{11} & \dots & a_{1,i-1} & b_1 & a_{1,i+1} & \dots & a_{1n} \\ a_{21} & \dots & a_{2,i-1} & b_2 & a_{2,i+1} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n,i-1} & b_n & a_{n,i+1} & \dots & a_{nn} \end{vmatrix}, \qquad i = 1, \dots, n. \tag{6.15}$$

Иначе говоря, определитель Δ_i получается из определителя Δ заменой его i-го столбца на столбец правых частей системы. Тогда единственное решение системы (6.13) выражается формулами Крамера

$$x_i = \frac{\Delta_i}{\Delta}, \qquad i = 1, \dots, n. \tag{6.16}$$

◀ Единственность решения при $\Delta \neq 0$ следует из теоремы Крамера 6.6. Остается проверить, что при подстановке в любое уравнение системы чисел, заданных формулами (6.16), получится верное равенство. Подставим эти числа в левую часть k-го уравнения, где $k = 1, \dots, n$. Получим

$$a_{k1}\frac{\Delta_1}{\Delta} + a_{k2}\frac{\Delta_2}{\Delta} + \dots + a_{kn}\frac{\Delta_n}{\Delta} =$$

$$\frac{1}{\Delta}(a_{k1}\Delta_1 + a_{k2}\Delta_2 + \dots + a_{kn}\Delta_n) = (\text{разложение } \Delta_i \text{ по } i\text{-му столбцу})$$

$$\frac{1}{\Delta}\begin{array}{l} (a_{k1}(b_1A_{11} + b_2A_{21} + \dots + b_nA_{n1}) + \\ a_{k2}(b_1A_{12} + b_2A_{22} + \dots + b_nA_{n2}) + \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ +a_{kn}(b_1A_{1n} + b_2A_{2n} + \dots + b_nA_{nn})) = \end{array}$$

$$\frac{1}{\Delta}\begin{array}{l} (b_1(a_{k1}A_{11} + a_{k2}A_{12} + \dots + a_{kn}A_{1n}) + \\ b_2(a_{k1}A_{21} + a_{k2}A_{22} + \dots + a_{kn}A_{2n}) + \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ +b_n(a_{k1}A_{n1} + a_{k2}A_{n2} + \dots + a_{kn}A_{nn})) \end{array}$$

Коэффициент при b_i в последней сумме — это "фальшивое" разложение определителя Δ по k-й строке при $i \neq k$ и "правильное" разложение определителя Δ по k-й строке при $i = k$. Поэтому

$$a_{k1}\frac{\Delta_1}{\Delta} + a_{k2}\frac{\Delta_2}{\Delta} + \dots + a_{kn}\frac{\Delta_n}{\Delta} = \frac{1}{\Delta}(b_1 \cdot 0 + \dots + b_k \cdot \Delta + \dots + b_n \cdot 0) = b_k,$$

что и доказывает справедливость правила Крамера. ▶

Лекция 7. Операции над матрицами и их свойства. Обобщенная ассоциативность. Транспонирование произведения матриц. Умножение матрицы на диагональную матрицу слева и справа. Единичная матрица, ее единственность. Скалярные матрицы. Обратная матрица, ее единственность.

1°. Определение и основные свойства операций над матрицами. Сначала вспомним определение суммы двух матриц и произведения матрицы на число (см определение 1.7).

Определение 7.1. *Суммой* двух матриц $A = (a_{ij})$ и $B = (b_{ij})$ одного и того же размера $m \times n$ называется матрица того же размера, элементы которой — суммы соответствующих элементов матриц-слагаемых: $A + B = (a_{ij} + b_{ij})$. *Произведением* матрицы $A = (a_{ij})$ размера $m \times n$ на число λ называется матрица того же размера, элементы которой — произведения элементов исходной матрицы на одно и то же число λ: $\lambda A = (\lambda a_{ij})$.

Определение 7.2. *Произведением* матрицы $A = (a_{ij})$ размера $m \times n$ на матрицу $B = (b_{ij})$ размера $n \times p$ называется матрица $C = AB = (c_{ij})$ размера $m \times p$, элемент которой, стоящий на пересечении i-й строки и j-го столбца равен сумме произведений элементов i-й строки матрицы A на соответствующие элементы j-го столбца матрицы B (часто коротко говорят "произведение i-й строки матрицы A на j-й столбец матрицы B):

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}, \quad i = 1, \ldots, m, \quad j = 1, \ldots p. \tag{7.1}$$

Следует обратить внимание, что произведение матрицы A на матрицу B определено **только в том случае, когда число столбцов первой матрицы равно числу строк второй.**

Теорема 7.3. Для любых матриц A, B, C допустимых размеров и любых чисел α и β выполнены равенства:
(1) $(A + B) + C = A + (B + C)$ (*ассоциативность сложения*);
(2) $A + B = B + A$ (*коммутативность сложения*);
(3) $A + 0 = 0 + A = A$, где 0 — нулевая матрица того же размера, что и матрица A (*существование нейтрального по сложению элемента*);
(4) $A + (-1)A = (-1)A + A = 0$ (*существование обратного по сложению, или противоположного элемента*);
(5) $\alpha(A + B) = \alpha A + \alpha B$ (*дистрибутивность умножения числа на матрицу*);
(6) $(\alpha + \beta)A = \alpha A + \beta A$ (*дистрибутивность умножения матрицы на число*);
(7) $\alpha(\beta A) = (\alpha\beta)A$ (*ассоциативность умножения матрицы на число*);
(8) $1A = A$ (*нейтральность числа 1 по умножению*);
(9) $A(BC) = (AB)C$ (*ассоциативность умножения*);
(10) $(A + B)C = AC + BC$ (*дистрибутивность умножения по первому сомножителю*;
(11) $A(B + C) = AB + AC$ (*дистрибутивность умножения по второму сомножителю*;
(12) $(\alpha A)B = A(\alpha B) = \alpha(AB)$.

◀ Свойства (1)-(8) очевидно следуют из определений. Проверим свойство (10). Пусть A, B — матрицы размера $m \times n$ и C — матрица размера $n \times p$. Тогда элемент i-й строки и j-го столбца матрицы, стоящей в левой части равенства (10), определяется выражением

$$\sum_{k=1}^{n}(a_{ik} + b_{ik})c_{kj} = \sum_{k=1}^{n}(a_{ik}c_{kj} + b_{ik}c_{kj}) = \sum_{k=1}^{n} a_{ik}c_{kj} + \sum_{k=1}^{n} b_{ik}c_{kj},$$

которое задает элемент i-й строки и j-го столбца матрицы, стоящей в правой части этого равенства.

Равенство (11) доказывается аналогично. Проверим, наконец, равенство (9). Заметим сначала, что для определенности произведений AB и BC необходимо и достаточно, чтобы матрицы A, B и C имели размеры $m \times n$, $n \times p$ и $p \times q$, соответственно. В этом случае матрица $U = AB$ имеет размер $m \times p$, а матрица $V = BC$ — размер $n \times q$, поэтому произведения, стоящие в левой и правой частях равенства (9), определены и являются матрицами размера $n \times q$. Осталось проверить, что соответствующие элементы матриц UC и AV равны. Запишем по определению выражение для элемента матрицы UC:

$$\sum_{l=1}^{p} u_{il}c_{lj} = \sum_{l=1}^{p} \left(\left(\sum_{k=1}^{n} a_{ik}b_{kl} \right) c_{lj} \right) = \sum_{l=1}^{p} \sum_{k=1}^{n} a_{ik}b_{kl}c_{lj}.$$

Аналогично, вычисляя элемент матрицы V с теми же индексами i, j, получим

$$\sum_{k=1}^{n} a_{ik}v_{kj} = \sum_{k=1}^{n} \left(a_{ik} \left(\sum_{l=1}^{p} b_{kl}c_{lj} \right) \right) = \sum_{k=1}^{n} \sum_{l=1}^{p} a_{ik}b_{kl}c_{lj}.$$

Последнее выражение, с точностью до порядка слагаемых, совпадает с полученным ранее.

Наконец, равенство (12) получается из (7.1) вынесением общего множителя λ из каждой суммы:

$$\sum_{k=1}^{n} (\lambda a_{ik})b_{kj} = \lambda(\sum_{k=1}^{n} a_{ik}b_{kj});$$
$$\sum_{k=1}^{n} a_{ik}(\lambda b_{kj}) = \lambda(\sum_{k=1}^{n} a_{ik}b_{kj}).$$

►

Замечание. Если A и B — квадратные матрицы одного и того же порядка n, то их сумма $A + B$ и произведение AB определены и снова являются квадратными матрицами порядка n. Значит, сложение и умножение квадратных матриц порядка n — операции, определенные на множестве этих матриц. Выполнение свойств (1)–(4) означает, что это множество является **абелевой группой** относительно операции сложения, а выполнение свойств (1)–(4) и свойств (9)–(11) — что это множество является **ассоциативным кольцом.** Общие понятия (абелевой) группы и кольца мы определим позже, при этом будем иметь в виду кольцо матриц как важнейший пример.

Заметим, кроме того, что умножение даже квадратных матриц, вообще говоря, некоммутативно. Например,

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

но

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

В то же время из ассоциативности произведения следует

Предложение 7.4. Произведение любого числа сомножителей не зависит от расстановки скобок между ними, в предположении, что это правильная расстановка скобок: каждое выражение в скобках — произведение двух сомножителей. Например, $(A(BC))D = (AB)(CD) = ((AB)C)D$ — правильные расстановки скобок в произведении четырех сомножителей (*Обобщенная ассоциативность*).

◀ Введем временно обозначение $A_1A_2\dots A_k = (\dots((A_1A_2)A_3)\dots)A_k$. Утверждение доказывается индукцией по числу сомножителей k. При $k = 3$ это верно в силу обычной ассоциативности. Для $k > 3$ рассмотрим последнюю в заданной расстановке операцию умножения: U_1U_2, где U_1, U_2 — произведения сомножителей $A_1, \dots, A_r$ и $A_{r+1}, \dots, A_k$, соответственно, с некоторой правильной расстановкой скобок в каждом произведении. Каждое из этих произведений содержит менее k сомножителей, поэтому из предположения индукции получаем:

$$U_1 = A_1 \cdot \dots \cdot A_r, \quad U_2 = A_{r+1} \cdot \dots \cdot A_k.$$

Если $k = r + 1$, то $U_1U_2 = A_1A_2\dots A_k$, по определению выражения в правой части. Если же $k > r + 1$, то

$$U_1U_2 = (A_1A_2\dots A_r)((A_{r+1}\dots A_{k-1})A_k) \overset{\text{(в силу ассоциативности)}}{=}$$
$$((A_1\dots A_r)(A_{r+1}\dots A_{k-1}))A_k \overset{\text{(в силу предположения индукции)}}{=} (A_1\dots A_{k-1})A_k = A_1\dots A_k,$$

что и требовалось.▶

2°. Транспонирование произведения матриц.

Теорема 7.5. Пусть $A = (a_{ij})$ — матрица размера $m \times n$, $B = (b_{ij})$ — матрица размера $n \times p$. Тогда $(AB)^T = B^TA^T$.

◀ Пусть $A = (a_{ij})$ — матрица размера $m \times n$, $B = (b_{ij})$ — матрица размера $n \times p$, $C = AB = (c_{ij})$. Тогда $B^T = (b'_{ij})$ — матрица размера $p \times n$, $A^T = (a'_{ij})$ — матрица размера $n \times m$, значит, произведение $C' = B^TA^T = (c'_{ij})$ определено и имеет размер $p \times m$, т.е. тот же размер, что и матрица C^T. Найдем элементы матрицы C':

$$c'_{ij} = \sum_{k=1}^{n} \underbrace{b'_{ik}}_{=b_{ki}} \underbrace{a'_{kj}}_{=a_{jk}} = \sum_{k=1}^{n} b_{ki}a_{jk} = \sum_{k=1}^{n} a_{jk}b_{ki} = c_{ji},$$

значит, $C' = C^T$. ▶

3°. Умножение матрицы на диагональную матрицу. Единичная матрица и ее свойства.

Определение 7.6. *Диагональной* матрицей называется квадратная матрица, все элементы которой, расположенные вне главной диагонали, равны нулю. Часто применяется обозначение

$$\operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}. \tag{7.2}$$

Предложение 7.7. Пусть $D = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ — диагональная матрица порядка n, $A = (a_{ij})$ — произвольная матрица размера $n \times p$. Тогда умножение матрицы A на матрицу D слева равносильно умножению каждой i-й строки матрицы A, $i = 1, \dots, n$, на число λ_i:

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} = \begin{pmatrix} \lambda_1a_{11} & \lambda_1a_{12} & \dots & \lambda_1a_{1p} \\ \lambda_2a_{21} & \lambda_2a_{22} & \dots & \lambda_2a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_na_{n1} & \lambda_na_{n2} & \dots & \lambda_na_{np} \end{pmatrix}. \tag{7.3}$$

Соответственно, если $A = (a_{ij})$ — матрица размера $m \times n$, то умножение матрицы A на матрицу D справа равносильно умножению каждого i-го столбца матрицы A, $i = 1, \ldots, n$, на число λ_i:

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 a_{11} & \lambda_2 a_{12} & \ldots & \lambda_n a_{1n} \\ \lambda_1 a_{21} & \lambda_2 a_{22} & \ldots & \lambda_n a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 a_{m1} & \lambda_2 a_{m2} & \ldots & \lambda_n a_{mn} \end{pmatrix}. \tag{7.4}$$

◄ Докажем (7.3). Положим $B = DA = (b_{ij})$ и вычислим элементы произведения:

$$b_{ij} = 0 \cdot a_{1j} + \ldots + 0 \cdot a_{i-1,j} + \lambda_i \cdot a_{ij} + 0 \cdot a_{i+1,j} + \ldots + 0 \cdot a_{nj} = \lambda_i a_{ij}, \quad j = 1, \ldots, p.$$

Аналогично доказывается (7.4). Можно также вывести (7.4) из (7.3) и формулы транспонирования произведения матриц. ►

Определение 7.8. *Единичной матрицей* порядка n называется матрица вида

$$E = \operatorname{diag}(1, 1, \ldots, 1) = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Иначе говоря, на главной диагонали единичной матрицы стоят единицы, а вне главной диагонали — нули.

Предложение 7.9. Пусть E — единичная матрица порядка n, A — произвольная матрица размера $n \times p$. Тогда $EA = A$. Соответственно, если A — матрица размера $m \times n$, то $AE = A$. Наконец, если A — квадратная матрица порядка n, то

$$AE = EA = A. \tag{7.5}$$

◄ Очевидное следствие предложения 7.7.►

Предложение 7.10. Квадратная матрица, удовлетворяющая условию (7.5) для любой квадратной матрицы того же порядка, определена однозначно.

◄ Допустим, что матрица E' обладает аналогичным свойством: $E'A = AE' = A$ для любой квадратной матрицы A. Подставляя E вместо A, получим $E'E = EE' = E$. Но в силу (7.5), $E'E = EE' = E'$. Значит, $E' = E$.►

Определение 7.11. Матрица вида $\lambda E = \operatorname{diag}(\lambda, \lambda, \ldots, \lambda)$ называется *скалярной* матрицей.

Умножение на скалярную матрицу λE (слева или справа) равносильно умножению всей матрицы на число λ (следует из (7.3) и (7.4)).

4°. Обратная матрица. Единственность обратной матрицы.

Определение 7.12. Квадратная матрица B называется *обратной* к квадратной матрице A, если

$$AB = BA = E. \tag{7.6}$$

Матрица называется *обратимой*, если обратная к ней матрица существует.

Теорема 7.13. Если обратная матрица к матрице A существует, то она единственна.

◀ Пусть матрицы B и B' удовлетворяют (7.6). Тогда

$$\begin{array}{ccccc} B(AB') & = & BE & = & B \\ || & & \text{(ассоциативность!)} & & \\ (BA)B' & = & EB' & = & B', \end{array}$$

т.е. $B = B'$. ▶

Доказанный факт позволяет ввести обозначение A^{-1} для обратной матрицы к матрице A. С учетом этого, получим более естественную запись для (7.6):

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

Лекция 8. Умножение треугольных матриц. Матричные единицы и их умножение. Элементарные матрицы и их связь с элементарными преобразованиями. Определитель произведения матриц. Критерий существования и способы нахождения обратной матрицы.

1°. Умножение треугольных матриц.

Предложение 8.1. Пусть $A = (a_{ij})$ и $B = (b_{ij})$ — две верхнетреугольные матрицы порядка n. Тогда их произведение — верхнетреугольная матрица, на главной диагонали которой расположены произведения соответствующих элементов главных диагоналей матриц A и B:

$$\begin{pmatrix} a_{11} & * & * & \dots & * \\ 0 & a_{22} & * & \dots & * \\ 0 & 0 & a_{33} & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & * & * & \dots & * \\ 0 & b_{22} & * & \dots & * \\ 0 & 0 & b_{33} & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & b_{nn} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} & * & * & \dots & * \\ 0 & a_{22}b_{22} & * & \dots & * \\ 0 & 0 & a_{33}b_{33} & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn}b_{nn} \end{pmatrix}.$$

◀ Пусть $i > j$. Рассмотрим элемент произведения $C = AB$, расположенный на пересечении i-й строки и j-го столбца:

$$c_{ij} = \underbrace{a_{i1}b_{1j} + \ldots + a_{i,i-1}b_{i-1,j}}_{=0,\ \text{т.к. } a_{ik}=0 \text{ при } i>k} + \overbrace{a_{ii}b_{ij} + \ldots + a_{in}b_{nj}}^{=0,\ \text{т.к. } b_{kj}=0 \text{ при } k \geqslant i > j} = 0.$$

Аналогично, на главной диагонали C расположены элементы

$$c_{ii} = \underbrace{a_{i1}b_{1i} + \ldots + a_{i,i-1}b_{i-1,i}}_{=0,\ \text{т.к. } a_{ik}=0 \text{ при } i>k} + a_{ii}b_{ii} + \overbrace{a_{i,i+1}b_{i+1,i} + \ldots + a_{in}b_{ni}}^{=0,\ \text{т.к. } b_{ki}=0 \text{ при } k>i} = a_{ii}b_{ii}.$$

►

2°. Матричные единицы. Правило умножения матричных единиц. Матричные единицы — база пространства матриц.

Определение 8.2. *Матричной единицей* E_{ij} называется матрица, у которой на пересечении i-й строки с j-м столбцом стоит 1, а все остальные элементы равны 0:

$$E_{ij} = \begin{array}{c} \\ \\ \\ \\ i \\ \\ \\ \end{array} \overset{\overset{\overset{\overset{\overset{j}{\downarrow}}{\downarrow}}{\downarrow}}{\downarrow}}{\begin{pmatrix} 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}}$$

Предложение 8.3. Произведение матричных единиц (допустимых размеров $m \times n$ и $n \times p$) определяется правилом

$$E_{ij}E_{kl} = \begin{cases} E_{il} \text{ при } j = k \\ 0 \text{ при } j \neq k \end{cases} \tag{8.1}$$

◄ Рассмотрим элементы произведения $A = E_{ij}E_{kl}$ двух матричных единиц указанных размеров. Очевидно, что все строки произведения, кроме i-й, нулевые, поскольку все строки матрицы E_{ij}, кроме i-й, нулевые. Аналогично, все столбцы произведения $E_{ij}E_{kl}$, кроме l-го, нулевые. Следовательно, достаточно рассмотреть элемент произведения на пересечении i-й строки и l-го столбца. При $j = k$ получим сумму

$$0 \cdot 0 + \ldots + \underbrace{1}_{\substack{j-\text{й элемент строки}}} \cdot \underbrace{1}_{\substack{j-\text{й элемент столбца}}} + \ldots + 0 \cdot 0 = 1,$$

а при $j \neq k$ — сумму

$$0 \cdot 0 + \ldots + \underbrace{1}_{\substack{j-\text{й элемент строки}}} \cdot \underbrace{0}_{\substack{j-\text{й элемент столбца}}} + \ldots + \underbrace{0}_{\substack{k-\text{й элемент строки}}} \cdot \underbrace{1}_{\substack{k-\text{й элемент столбца}}} + \ldots + 0 \cdot 0 = 0.$$

►

Очевидно, что любую матрицу размера $m \times n$ можно единственным образом представить как линейную комбинацию матричных единиц:

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

т.е. матричные единицы образуют базу пространства всех матриц заданного размера.

3°. Элементарные матрицы и элементарные преобразования.

Определение 8.4. Матрицы, полученные из единичной матрицы применением к её строкам (столбцам) одного элементарного преобразования любого из трех типов, называются *элементарными матрицами*. Легко проверить, что элементарное преобразование первого, второго и третьего типа дают, соответственно, матрицы вида $E+\lambda E_{ij}$, $E-E_{ii}-E_{jj}+E_{ij}+E_{ji}$ и $E+(c-1)E_{ii}$.

Теорема 8.5. Применение элементарного преобразования к строкам (столбцам) произвольной матрицы равносильно ее умножению слева (справа) на соответствующую элементарную матрицу.

◄ Рассмотрим произведение матрицы $A = (a_{ij})$ на элементарную матрицу $E + \lambda E_{ij}$ соответствующего размера слева:

$$(E + \lambda E_{ij})A = A + \lambda E_{ij}A = A + \lambda E_{ij} \left(\sum_{k=1}^{l} \sum_{l=1}^{n} a_{kl} E_{kl} \right) = (\text{только слагаемые с } k = j)$$

$$= \sum_{k=1}^{m} \sum_{l=1}^{n} a_{kl} E_{kl} + \sum_{l=1}^{n} \lambda a_{jl} E_{il} = (\text{соберем отдельно слагаемые с } k = i \text{ в первой сумме})$$

$$= \sum_{k \neq i} \sum_{l=1}^{n} a_{kl} E_{kl} + \sum_{l=1}^{n} (a_{il} + \lambda a_{jl}) E_{il}.$$

Таким образом, элементы всех строк произведения, кроме i-й, совпадают с элементами матрицы A, а элементы i-й строки — как в сумме i-й строки и j-й строки, умноженной на число λ.

Случай элементарного преобразования типа 2 рассматривается аналогично. Для элементарного преобразования типа 3 утверждение теоремы сразу следует из предложения 7.7.

Для преобразований столбцов и умножения справа можно либо применить аналогичные рассуждения, либо воспользоваться теоремой о транспонировании произведения.►

4°. Определитель произведения матриц.

Теорема 8.6. Пусть A и B — квадратные матрицы одного порядка. Тогда

$$|AB| = |A||B|. \tag{8.2}$$

◀ Известно, что матрицу A с помощью последовательности элементарных преобразований типа 1 можно привести к ступенчатому виду. Это означает, по предыдущей теореме, что существует такой набор элементарных матриц $P_1, P_2, \ldots, P_k$, что $A' = P_1P_2 \ldots P_kA$ — треугольная матрица. При этом $|A'| = |A|$. Но тогда $A'B = (P_1P_2 \ldots P_kA)B = P_1P_2 \ldots P_k(AB)$, значит, матрица $A'B$ получается из матрицы AB той же последовательностью элементарных преобразований, и поэтому $|A'B| = |AB|$. Если $|A| = 0$, то матрица A' содержит нулевую строку, но тогда и в матрице $A'B$ имеется нулевая строка. Значит, $0 = |A'B| = |AB| = \underbrace{|A|}_{=0}|B|$.

Если же $|A'| = |A| \neq 0$, то диагональные элементы матрицы A' отличны от 0, и можно провести еще несколько элементарных преобразований типа 1, в результате которых треугольная матрица превратится в диагональную: $A'' = Q_1Q_2 \ldots Q_lA' = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Получим $A''B = (Q_1Q_2 \ldots Q_lA')B = Q_1Q_2 \ldots Q_l(A'B)$, откуда

$$|A'B| = |A''B| = \underbrace{\lambda_1\lambda_2 \ldots \lambda_n}_{=|A''|=|A|}|B| = |A||B|.$$

▶

5°. Критерий существования и способы нахождения обратной матрицы.

Укажем следующий критерий обратимости матрицы.

Теорема 8.7. Квадратная матрица является обратимой тогда и только тогда, когда она *невырожденна*, т.е. когда ее определитель не равен нулю.

◀ Если A — обратимая квадратная матрица, то

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

что невозможно при $|A| = 0$.

Обратно, пусть $|A| \neq 0$. Тогда можно привести матрицу A элементарными преобразованиями строк типа 1 к диагональному виду, а затем — элементарными преобразованиями строк типа 3 сделать её единичной матрицей. Значит, $E = P_1P_2 \ldots P_kA$, т. е. существует матрица $B = P_1P_2 \ldots P_k$, такая, что $BA = E$. Применяя то же рассуждение к матрице A^T, получим, что

$$Q_1Q_2 \ldots Q_lA^T = E$$

для некоторых элементарных матриц $Q_1, Q_2, \ldots, Q_l$, т.е. $AB' = E$ для матрицы $B' = Q_l^TQ_{l-1}^T \ldots Q_1^T$. Но $B' = EB' = (BA)B' = B(AB') = B$ (см. доказательство теоремы 7.13), Но тогда $A^{-1} = B = B'$. ▶

Приведенное доказательство дает следующий метод нахождения обратной матрицы с помощью элементарных преобразований.

Пусть A — квадратная матрица порядка n. Запишем рядом матрицу A и единичную матрицу того же порядка: $(A|E)$. Приведем элементарными преобразованиями строк полученную матрицу размера $n \times 2n$ к виду $(E|A')$ (если в процессе приведения в левой части матрицы появится нулевая строка, то матрица A необратима, т.е. нет нужды предварительно проверять обратимость матрицы A). Тогда, по первой части доказательства теоремы 8.7, $A' = A^{-1}$.

Приведем пример вычисления обратной матрицы этим способом.

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 5 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -1 & -3 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 3 & -1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & -5 & 2 \\ 0 & 1 & 3 & -1 \end{array}\right),$$

значит,

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

Другой способ нахождения обратной матрицы — с помощью алгебраических дополнений.

Теорема 8.8. Пусть $A = (a_{ij})$ — невырожденная квадратная матрица порядка n. Положим

$$\tilde{A} = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}, \tag{8.3}$$

где A_{ij}, как обычно, алгебраическое дополнение элемента a_{ij}. Тогда

$$A^{-1} = \frac{1}{|A|}\tilde{A}. \tag{8.4}$$

◀ Найдем произведение $A\tilde{A}$:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix},$$

где

$$b_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \ldots + a_{in}A_{jn} = \begin{cases} |A|, & \text{если } i = j \text{ (“правильное” разложение по } i\text{-й строке)}, \\ 0, & \text{если } i \neq j \text{ (“фальшивое” разложение по } i\text{-й строке)}. \end{cases}$$

Значит, $A\tilde{A} = |A|E$. При $|A| \neq 0$ существует обратная матрица A^{-1}. Умножая обе части последнего равенства на $\dfrac{1}{|A|}A^{-1}$ слева, получаем (8.4). ▶

Матрицу $\tilde{A}$ называют *присоединенной* к матрице A. Отметим, что в матрице $\tilde{A}$ на месте каждого элемента a_{ij} матрицы A стоит алгебраическое дополнение к элементу, **симметричному** к a_{ij} относительно главной диагонали. Пример: если $A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$, то $|A| = -1$, $A_{11} = 5$, $A_{12} = -3$, $A_{21} = -2$, $A_{22} = 1$, значит, $A^{-1} = \dfrac{1}{-1}\begin{pmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{pmatrix} = -\begin{pmatrix} 5 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}$.

Лекция 9. Миноры прямоугольной матрицы. Вычисление ранга матрицы с помощью миноров (теорема о ранге матрицы).

1°. Миноры прямоугольной матрицы. Окаймляющие миноры.

Определение 9.1. *Минором порядка* k прямоугольной матрицы размера $m \times n$ называется определитель порядка k, составленный из элементов матрицы A, расположенных на пересечении некоторых k строк с некоторыми k столбцами. Если $i_1, \ldots, i_k$ — номера выбранных строк, а $j_1, \ldots, j_k$ — номера выбранных столбцов, то минор образуют отмеченные элементы матрицы:

$$\begin{array}{c} \begin{pmatrix} \ldots & j_1 & \ldots & j_2 & \ldots & \ldots & j_k & \ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots & a_{i_1,j_1} & \ldots & a_{i_1,j_2} & \ldots & \ldots & a_{i_1,j_k} & \ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots & a_{i_2,j_1} & \ldots & a_{i_2,j_2} & \ldots & \ldots & a_{i_2,j_k} & \ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots & a_{i_k,j_1} & \ldots & a_{i_k,j_2} & \ldots & \ldots & a_{i_k,j_k} & \ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{pmatrix} \\ \begin{matrix} \\ \\ i_1 \\ \\ i_2 \\ \\ \\ i_k \end{matrix} \end{array}$$

Этот минор имеет вид

$$\begin{vmatrix} a_{i_1,j_1} & \ldots & a_{i_1,j_k} \\ a_{i_2,j_1} & \ldots & a_{i_2,j_k} \\ \ldots\ldots\ldots\ldots & & \\ a_{i_k,j_1} & \ldots & a_{i_k,j_k} \end{vmatrix}.$$

Определение 9.2. *Окаймляющим минором* для минора порядка k матрицы A называется минор порядка $k+1$, полученный добавлением одной строки к набору строк и одного столбца к набору столбцов, задающих минор этот минор порядка k.

Теорема 9.3 (об окаймляющих минорах)**.** Если некоторый минор порядка k матрицы A, расположенный в строках с номерами $i_1, \ldots, i_k$ и столбцах с номерами $j_1, \ldots, j_k$, отличен от нуля, а все его окаймляющие миноры равны 0, то строки (столбцы) с номерами $i_1, \ldots, i_k$ ($j_1, \ldots, j_k$) образуют базу системы строк (столбцов) матрицы A.

◀ Для упрощения обозначений допустим, что указанный минор порядка k содержится в первых k строках и первых k столбцах матрицы $A = (a_{ij})$. Обозначим первые k столбцов матрицы A через $a^1, \ldots, a^k$ и покажем, что они образуют базу системы столбцов матрицы A. Ясно, что эти столбцы линейно независимы, иначе были бы линейно зависимы также "укороченные" столбцы

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{k1} \end{pmatrix}, \ldots, \begin{pmatrix} a_{1k} \\ \vdots \\ a_{kk} \end{pmatrix},$$

что противоречит условию

$$\Delta = \begin{vmatrix} a_{11} & \ldots & a_{k1} \\ \ldots\ldots\ldots\ldots & & \\ a_{1k} & \ldots & a_{kk} \end{vmatrix} \neq 0.$$

Остается проверить, что любой столбец матрицы A — линейная комбинация столбцов $a^1, \ldots, a^k$. Для этого выберем какой-нибудь столбец

$$a^r = \begin{pmatrix} a_{1r} \\ \vdots \\ a_{mr} \end{pmatrix}, \quad r > k,$$

и рассмотрим определитель

$$D = \begin{vmatrix} a_{11} & \dots & a_{1k} & a_{1r} \\ \dots & \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} & a_{kr} \\ a_{i1} & \dots & a_{ik} & a_{ir} \end{vmatrix}.$$

Заметим, что $D = 0$ при любом i: если $i \leqslant k$, то в нем две одинаковые строки, а если $i > k$, то D — окаймляющий минор для минора Δ и равен нулю по условию. Разлагая определитель D по последней строке, получим

$$D = a_{i1}A_{i1} + \ldots + a_{ik}A_{ik} + a_{ir}A_{ir} = 0. \tag{9.1}$$

При этом $A_{ir} = \Delta \neq 0$, а алгебраические дополнения $A_{i1}, \ldots, A_{ik}$ не зависят от i. Значит, можно обозначить

$$\lambda_1 = -\frac{A_{i1}}{\Delta}, \ldots, \lambda_k = -\frac{A_{ik}}{\Delta}$$

и переписать (9.1) в виде равенства

$$a_{ir} = \lambda_1 a_{i1} + \ldots + \lambda_k a_{ik}.$$

Поскольку это соотношение выполняется при всех $i = 1, \ldots, m$, получаем требуемое векторное равенство

$$a^r = \lambda_1 a^1 + \ldots + \lambda_k a^k.$$

Доказательство для строк аналогично.▶

2°. Теорема о ранге матрицы.

Теорема 9.4. Ранг прямоугольной матрицы A равен наибольшему порядку минора этой матрицы, не равного нулю.

◀ Пусть k — наибольший порядок минора матрицы A, не равного нулю. Пусть этот минор располагается в столбцах с номерами $j_1, \ldots, j_k$. По условию все миноры, окаймляющие данный минор, равны 0, поэтому, по теореме 9.3, столбцы с номерами $j_1, \ldots, j_k$ образуют базу столбцов матрицы A. Но число векторов в такой базе равно рангу матрицы A, по определению ранга матрицы.▶

Замечание. Если все элементы матрицы A равны нулю, то, разумеется, никаких отличных от нуля миноров у нее нет. Но в этом случае принято считать, что имеется “пустой” минор порядка 0, не равный нулю (и никакому, впрочем, другому числу), так что утверждение теоремы справедливо и в этом случае.

Доказанные теоремы позволяют вычислять ранг матриц, не прибегая к элементарным преобразованиям, а вычисляя только определители. Именно, выбираем ненулевой элемент матрицы, затем ищем окаймляющий минор порядка 2, не равный нулю, затем, если такой минор нашелся, ищем окаймляющий его минор порядка 3, не равный 0, и т.д. С вычислительной точки зрения этот процесс в общем случае менее эффективен, чем применение элементарных преобразований.

Лекция 10. Ранг произведения матриц. Факторизационный ранг матрицы. Матричная запись системы линейных уравнений. Строение общего решения неоднородной системы уравнений, его геометрическая интерпретация.

1°. Ранг произведения матриц.

Теорема 10.1. Если A — матрица размера $m \times n$, B — матрица размера $n \times p$, то

$$\mathrm{rk}(AB) \leqslant \min\{\mathrm{rk}(A), \mathrm{rk}(B)\}. \tag{10.1}$$

◀ Пусть, как обычно, $A = (a_{ij})$ и $B = (b_{ij})$. Заметим, что строки матрицы AB — линейные комбинации строк матрицы B. Действительно, i-я строка матрицы AB, $(i = 1, \ldots, m)$ — линейная комбинация строк матрицы B с коэффициентами $a_{i1}, \ldots, a_{in}$:

$$\begin{array}{rc} a_{i1}(b_{11}, b_{12}, \ldots, b_{1p}) & + \\ a_{i2}(b_{21}, b_{22}, \ldots, b_{2p}) & + \\ \ldots\ldots\ldots\ldots\ldots & \\ + \quad a_{in}(b_{n1}, b_{n2}, \ldots, b_{np}) & = \\ (\sum_{k=1}^{n} a_{ik}b_{k1}, \sum_{k=1}^{n} a_{ik}b_{k2}, \ldots, \sum_{k=1}^{n} a_{ik}b_{kn}), & \end{array}$$

а последняя строка и есть i-я строка произведения матриц. Но тогда из теоремы 2.13 следует, что ранг системы строк матрицы AB (т.е. ранг матрицы AB) не превосходит ранга системы строк матрицы B, т.е. $\mathrm{rk}(AB) \leqslant \mathrm{rk}(B)$. Неравенство $\mathrm{rk}(AB) \leqslant \mathrm{rk}(A)$ можно доказать аналогично, учитывая, что столбцы матрицы AB являются линейными комбинациями столбцов матрицы A, а ранг системы столбцов любой матрицы равен ее рангу. Другое рассуждение:
$\mathrm{rk}(AB) = \mathrm{rk}((AB)^T) = \mathrm{rk}(B^T A^T) \leqslant \mathrm{rk}(A^T) = \mathrm{rk}(A)$.▶
В случае умножения на обратимую матрицу можно сформулировать более точное утверждение.

Следствие 10.2. Если A — обратимая матрица размера $m \times m$, B — произвольная матрица размера $m \times n$, то $\mathrm{rk}(AB) = \mathrm{rk}(B)$.
Если A — обратимая матрица размера $n \times n$, B — произвольная матрица размера $m \times n$, то $\mathrm{rk}(BA) = \mathrm{rk}(B)$.

◀ Пусть A — обратимая матрица размера $m \times m$, B — произвольная матрица размера $m \times n$. Тогда, по теореме 10.1, выполнено неравенство $\mathrm{rk}(AB) \leqslant \mathrm{rk}(B)$. С другой стороны, $B = A^{-1}(AB)$, значит, $\mathrm{rk}(B) \leqslant \mathrm{rk}(AB)$. Равенство рангов доказано. Случай умножения на обратимую матрицу справа рассматривается аналогично.▶

2°. Факторизационный ранг матрицы.

Следующая теорема дает еще одно определение для ранга матрицы.

Теорема 10.3. Ранг ненулевой матрицы A размера $m \times n$ равен наименьшему натуральному числу k, такому, что существуют представление матрицы A в виде произведения матриц B, C размеров $m \times k$ и $k \times n$, соответственно (это число называется *факторизационным рангом* матрицы A.)

◀ Пусть ранг матрицы A равен r. Если $A = BC$, где B, C — матрицы размеров $m \times k$ и $k \times n$, соответственно, то $\mathrm{rk}(A) = \mathrm{rk}(BC) \leqslant \mathrm{rk}(B) \leqslant k$. Обратно, покажем, что существуют матрицы B, C размеров $m \times r$ и $r \times n$, для которых $A = BC$. Выберем r строк матрицы A, образующих базу системы ее строк, и образуем матрицу C из этих r строк. Для любой i-й строки a^i матрицы A существует выражение строки a^i через строки $c^1, \ldots, c^r$ матрицы C:

$$a^i = b_{i1}c^1 + \ldots b_{ir}c^r, \quad i = 1, \ldots, m$$

Тогда $A = BC$, где

$$B = \begin{pmatrix} b_{11} & \dots & b_{ir} \\ b_{21} & \dots & b_{2r} \\ \dots\dots\dots\dots \\ b_{m1} & \dots & b_{mr} \end{pmatrix}.$$

►

Следствие 10.4. Матрица $A = (a_{ij})$ размера $m \times n$ имеет ранг 1 тогда и только тогда, когда существуют числа $u_1, \dots, u_m$, не все равные 0, и $v_1, \dots, v_n$, не все равные 0, такие, что $a_{ij} = u_i v_j$, $i = 1, \dots, m$, $j = 1, \dots, n$.

◄ Условие $a_{ij} = u_i v_j$, $i = 1, \dots, m$, $j = 1, \dots, n$ равносильно тому, что $A = BC$, где B — столбец

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

а C — строка

$$C = \begin{pmatrix} v_1 & \dots & v_n \end{pmatrix}.$$

При этом A — ненулевая матрица тогда и только тогда, когда найдутся индексы i, j, для которых $u_i \neq 0$ и $v_j \neq 0$. В этом случае $\mathrm{rk}(A) \leqslant \mathrm{rk}(B) = 1$ и $\mathrm{rk}(A) \neq 0$, значит, $\mathrm{rk}(A) = 1$.

Обратно, пусть ранг матрицы A равен 1. Тогда, согласно предыдущей теореме, A можно представить в виде произведения столбца на строку. Если строка (столбец) содержит только нули, то и матрица A нулевая, что противоречит условию. ►

3°. Матричная запись системы линейных уравнений. Общее решение неоднородной системы линейных уравнений. Рассмотрим систему линейных уравнений

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n & = & b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n & = & b_m, \end{array} \tag{10.2}$$

Обозначим:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots\dots\dots\dots\dots\dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \text{ — матрица коэффициентов системы;}$$

$$\overline{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \text{ — столбец правых частей;}$$

$$\overline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ — столбец неизвестных.}$$

Тогда система (10.2) равносильна матричному уравнению

$$A\overline{x} = \overline{b}.$$

Действительно, $A\overline{x}$ — это столбец левых частей системы (10.2).

Из этого представления системы (10.2) легко получить следующие утверждения.

Предложение 10.5. Если система (10.2) — квадратная (т.е. $m = n$) и матрица A невырождена, то решение системы определяется равенством $\overline{x} = A^{-1}\overline{b}$.

◀ Поскольку $|A| \neq 0$, существует обратная матрица A^{-1}. Умножая обе части равенства $A\overline{x} = \overline{b}$ на A^{-1} слева, получаем требуемое равенство.▶

Использование данного способа решения эффективно, например, когда необходимо решить много систем с одинаковыми левыми частями и различными правыми частями.

Предложение 10.6. Если система (10.2) — однородная, т.е. $\overline{b} = 0$, то множество решений системы — подпространство линейного пространства $\mathbb{R}^n$, т.е. *множество всех решений этой системы замкнуто относительно сложения векторов и умножения вектора на число* (это уже упоминалось в лекции 3). Следовательно, любая линейная комбинация решений однородной системы снова является решением той же системы.

◀ Действительно, пусть $\overline{x}^1$ и $\overline{x}^2$ — два решения однородной системы $A\overline{x} = \overline{0}$. Тогда $A(\overline{x}^1 + \overline{x}^2) = A\overline{x}^1 + A\overline{x}^2 = \overline{0} + \overline{0} = \overline{0}$, и $A\lambda\overline{x}^1 = \lambda A\overline{x}^1 = \lambda\overline{0} = \overline{0}$.▶

Предложение 10.7. Пусть

$$A\overline{x} = \overline{b} \tag{10.3}$$

— произвольная совместная система линейных уравнений,

$$A\overline{x} = \overline{0} - \tag{10.4}$$

— система однородных линейных уравнений с той же матрицей коэффициентов, U — пространство решений однородной системы (10.4). Тогда множество всех решений системы (10.3) имеет вид

$$\overline{x}^0 + U = \{\overline{x}^0 + \overline{u} : \overline{u} \in U\},$$

где $\overline{x}^0$ — произвольное частное решение неоднородной системы (10.3). Часто этот факт выражают следующей фразой: **общее решение неоднородной системы уравнений есть сумма произвольного частного решения этой системы и общего решения соответствующей однородной системы**.

◀ Если $\overline{u} \in U$, а $\overline{x}^0$ — частное решение системы (10.3), то

$$A(\overline{x}^0 + \overline{u}) = A\overline{x}^0 + A\overline{u} = \overline{b} + \overline{0} = \overline{b}.$$

С другой стороны, если $\overline{x}^1$ — еще какое-то решение системы (10.3), то

$$\overline{x}^1 = \overline{x}^0 + (\overline{x}^1 - \overline{x}^0) \in \overline{x}^0 + U,$$

так как

$$A(\overline{x}^1 - \overline{x}^0) = A\overline{x}^1 - A\overline{x}^0 = \overline{b} - \overline{b} = \overline{0}.$$

▶

Докажем и обратное утверждение.

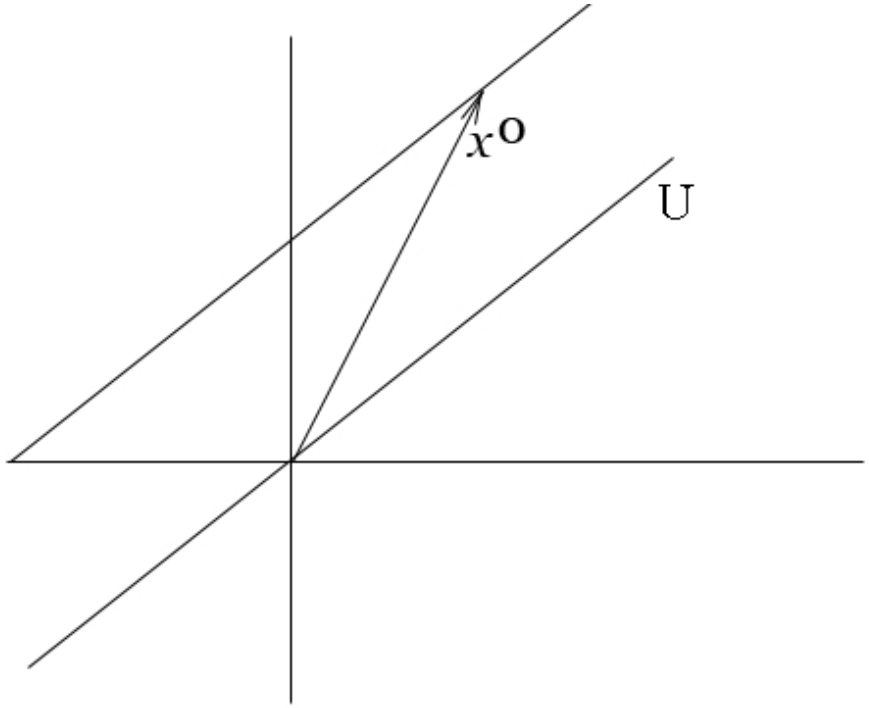
Предложение 10.8. Любое подпространство в $\mathbb{R}^n$ можно задать как множество решений некоторой системы однородных линейных уравнений.

◀ Пусть U — подпространство в $\mathbb{R}^n$ размерности k с базой $e^1, \ldots, e^k$. Векторы из $\mathbb{R}^n$ будем считать столбцами чисел. Рассмотрим множество V всех строк $(a_1, \ldots, a_n)$, для которых

$$(a_1 \ldots a_n)e^i = 0, \quad i = 1, \ldots, k. \tag{10.5}$$

Легко видеть, что V — подпространство линейного пространства $\mathbb{R}^n$. Поскольку равенства (10.5) можно рассматривать как систему уравнений относительно $a_1, \ldots, a_n$, с матрицей ранга k, то размерность пространства решений V равна $n - k$. Обозначим $m = n - k$ и выберем базу $a^1, \ldots, a^m$ подпространства V, где $a^i = (a_{i1}, \ldots, a_{in})$. Запишем матрицу $A = (a_{ij})$ размера $m \times n$. Пусть $\tilde{U}$ — пространство решений системы $Ax = 0$. Так как, по определению A, $Au = 0$ для любого вектора $u \in U$, имеем $U \subseteq \tilde{U}$. С другой стороны, ранг матрицы A равен m, в силу линейной независимости ее строк. Значит, $\tilde{U}$ имеет размерность $n - m = n - (n - k) = k$ — ту же размерность, что и U. Поэтому $U = \tilde{U}$.▶

В заключение отметим следующую геометрическую интерпретацию предложения 10.7: множество решений неоднородной системы есть *плоскость* в пространстве $\mathbb{R}^n$, т.е. множество точек, полученных сдвигом подпространства U на некоторый вектор $\overline{x}^0$ (см. рис.)



Лекция 11. Основные алгебраические структуры: группы.

1°. Бинарные операции на множествах.

Определение 11.1. Пусть X — произвольное непустое множество. *Бинарной операцией* на X называется произвольное отображение $\varphi : X \times X \to X$. Обычно вместо обозначения $\varphi(x, y)$ используют более привычную форму записи $a + b$ (операция "+") или $a * b$ (операция "$*$") или $a \circ b$ или $a \cdot b$ (последнее обозначение чаще всего сокращается до ab.

Наряду с бинарными можно рассматривать также n-арные операции при $n = 1, 2, 3, \ldots$ На множестве X можно, как правило, задать несколько различных операций.

Определение 11.2. Множество X вместе с одной или несколькими фиксированными операциями на нем называется "алгебраической структурой". Например, если на множестве X определена бинарная операция $*$, то соответствующая алгебраическая структура обозначается через $(X, *)$.

Как правило, исследуются не произвольные операции, а операции, удовлетворяющие определенным условиям (*аксиомам*).

Определение 11.3. Бинарная операция $*$ на множестве X (а также алгебраическая структура $(X, *)$) называется *ассоциативной*, если

$$\forall x, y, z \in X, \quad (x * y) * z = x * (y * z).$$

Определение 11.4. Бинарная операция $*$ на множестве X (а также алгебраическая структура $(X, *)$) называется *коммутативной*, если

$$\forall x, y \in X, \quad x * y = y * x.$$

Теорема 11.5. Результат последовательного применения ассоциативной бинарной операции к любому числу элементов не зависит от расстановки скобок между ними, в предположении, что это правильная расстановка скобок: каждое выражение в скобках — результат применения $*$ к двум выражениям.

◄ См. доказательство предложения 7.4 для операции умножения матриц.►

Определение 11.6. Множество с определенной на нем бинарной ассоциативной операцией называется *полугруппой*.

Определение 11.7. Элемент $e \in X$ называется *единичным* или *нейтральным* элементом относительно бинарной операции $*$ на множестве X, если

$$\forall x \in X, \quad x * e = e * x = x.$$

Предложение 11.8. Если нейтральный элемент существует, то он — единственный.

◄ См. доказательство предложения 7.10 для операции умножения матриц.►

Определение 11.9. Полугруппа, в которой есть единичный элемент, называется *моноидом*.

Определение 11.10. Пусть $(X, *)$ — моноид с нейтральным элементом e. Элемент $x \in X$ называется *обратимым*, если существует *обратный* элемент $y \in X$, для которого $x * y = y * x = e$.

Предложение 11.11. Если обратный элемент существует, то он — единственный.

◄ См. доказательство предложения 7.13 для операции умножения матриц.►
Последнее предложение позволяет ввести обозначение $y = x^{-1}$ для элемента, обратного к обратимому элементу x моноида X.

Определение 11.12. Алгебраическая структура (в частности, группа, кольцо, поле, которые мы определим ниже) называется *конечной*, если число ее элементов конечно, и *бесконечной* в противном случае. Число элементов алгебраической структуры называют ее *порядком*.

2°. Определение группы. Подгруппы.

Определение 11.13. Алгебраическая структура с одной бинарной операцией $(G, \cdot)$ называется *группой*, если выполняются следующие *аксиомы*:
1) операция ассоциативна: $\forall a, b, c \in G, \quad (ab)c = a(bc)$;
2) существует нейтральный элемент, (или *единица*) $e \in G$, для которого $ae = ea = a$ при любом $a \in G$.
3) любой элемент $a \in G$ обратим: существует элемент $a^{-1} \in G$, для которого $aa^{-1} = a^{-1}a = e$.
Иными словами, группа — это моноид, в котором каждый элемент обратим.

Определение 11.14. Группа G, в которой операция коммутативна, т.е.

$$\forall a, b \in G, \quad ab = ba,$$

называется *коммутативной*, или *абелевой* группой.

Замечание. Наряду с обозначением ab для результата бинарной операции (такое обозначение называют *мультипликативным*, а операцию в этом случае обычно называют умножением), часто используют, особенно для коммутативных групп, *аддитивное* обозначение операции $a + b$ (сложение). В этом случае нейтральный элемент называют нулевым и обозначают символом 0, а обратный к элементу a обозначают через $(-a)$, т.е. обозначения выбираются так, чтобы выполнялись привычные тождества $a + 0 = 0 + a = a$, $a + (-a) = (-a) + a = 0$.

Примеры групп:
$\mathbb{Z} = (\mathbb{Z}, +)$,
$\mathbb{R}^* = (\{x \in \mathbb{R} : x \neq 0\}, \cdot)$,
$\mathrm{GL}_n(\mathbb{R})$ — группа обратимых квадратных матриц порядка n относительно умножения матриц,
S_n — группа подстановок степени n относительно композиции.

Предложение 11.15. Если G — группа, и $a, b \in G$, то $(a^{-1})^{-1} = a$, $(ab)^{-1} = b^{-1}a^{-1}$.

◄ Первое равенство прямо следует из определения обратного элемента. Второе равенство проверяется:

$$ab(b^{-1}a^{-1}) = aea^{-1} = aa^{-1} = e,$$
$$(b^{-1}a^{-1})ab = beb^{-1} = bb^{-1} = e.$$

►

Определение 11.16. Подмножество H группы G называется подгруппой группы G, если
1) $e \in H$,
2) если $a, b \in H$, то $ab \in H$ (т. е. подмножество H замкнуто относительно операции),
3) если $a \in H$, то и $a^{-1} \in H$ (т. е. подмножество H замкнуто относительно взятия обратного элемента).

Иными словами, подгруппа группы G сама является группой относительно операции, заданной в G.

Примеры:
$n\mathbb{Z} = \{nz : z \in \mathbb{Z}\}$ — подгруппа в $\mathbb{Z}$,
$\mathbb{R}_+^* = \{x \in \mathbb{R} : x > 0\}$ — подгруппа в $\mathbb{R}^*$,
$\mathrm{SL}_n(\mathbb{R}) = \{A \in \mathrm{GL}_n(\mathbb{R}) : |A| = 1\}$ — подгруппа в $\mathrm{GL}_n(\mathbb{R})$,
A_n — подгруппа четных подстановок в S_n.

Определение 11.17. Пусть G, L — группы. Отображение $f : G \to L$ называется *гомоморфизмом*, если оно *сохраняет операцию*, т.е. если

$$\forall a, b \in G, \quad f(ab) = f(a)f(b).$$

Предложение 11.18. Пусть $f : G \to L$ — гомоморфизм групп. Тогда $f(e_G) = e_L$, $f(a^{-1}) = (f(a))^{-1}$ для любого $a \in G$.

◄ $f(e_G) = f(e_G e_G) = f(e_G)f(e_G)$. Умножая обе части на $(f(e_G))^{-1}$, получаем первое равенство. Из него выводим $f(a)f(a^{-1}) = f(e_G) = e_L$, откуда $f(a^{-1}) = (f(a))^{-1}$.►

Определение 11.19. Гомоморфизм групп, являющийся биективным (сюръективным, инъективным) отображением, называется *изоморфизмом* (эпиморфизмом, мономорфизмом). Группы, между которыми существует изоморфизм, называются *изоморфными* (обозначается $G \cong L$).

Упражнение. Пусть G, L, K — произвольные группы. Доказать, что
1) $G \cong G$;
2) если $G \cong L$, то $L \cong G$;
3) если $G \cong L$ и $L \cong K$, то $G \cong K$.

Лекция 12. Циклические группы. Порядок элемента. Подгруппы циклических групп. Изоморфизм циклических групп одного порядка. Теорема Кэли. Смежные классы, теорема Лагранжа и ее следствия.

1°. Циклические группы.

Определение 12.1. Пусть g — элемент группы G. Определим целые *степени* элемента g следующим образом:
1) $g^n = \underbrace{gg\cdots g}_{n \text{ раз}}$ при $n > 0$;
2) $g^0 = e$;
3) $g^n = (g^{-1})^{-n}$ при $n < 0$.

Теорема 12.2. Пусть g — элемент группы G. Тогда

$$g^{m+n} = g^m g^n, \quad g^{mn} = (g^m)^n \qquad \forall m, n \in \mathbb{Z}.$$

◀ Сначала заметим, что

$$g^{-n} = (g^{-1})^n = (g^n)^{-1} \text{ для любого } n \in \mathbb{Z}. \tag{12.1}$$

Затем докажем первое равенство. При $m \geqslant 0$ и $n \geqslant 0$, а также при $m \leqslant 0$ и $n \leqslant 0$ оно очевидно следует из определения степеней (и ассоциативности). Пусть теперь $m > 0, n < 0, m + n \geqslant 0$. Тогда $n' = -n > 0$, и

$$g^m g^n = \underbrace{gg\cdots g}_{m \text{ раз}}\underbrace{g^{-1}g^{-1}\cdots g^{-1}}_{n' \text{ раз}} = \underbrace{gg\cdots g}_{m-n' \text{ раз}}\underbrace{gg\cdots g}_{n' \text{ раз}}\underbrace{g^{-1}g^{-1}\cdots g^{-1}}_{n' \text{ раз}} = g^{m-n'} = g^{m+n}.$$

Теперь пусть $m > 0$, $n < 0$, $m + n < 0$. Тогда $n' = -n > 0$ и $n' - m > 0$, значит,

$$g^m g^n = \underbrace{gg\cdots g}_{m \text{ раз}}\underbrace{g^{-1}g^{-1}\cdots g^{-1}}_{n' \text{ раз}} = \underbrace{gg\cdots g}_{m \text{ раз}}\underbrace{g^{-1}g^{-1}\cdots g^{-1}}_{m \text{ раз}}\underbrace{g^{-1}g^{-1}\cdots g^{-1}}_{n'-m \text{ раз}} = (g^{-1})^{n'-m} = g^{m-n'} = g^{m+n}.$$

Остальные варианты знаков чисел $m, n, m + n$ рассматриваются аналогично.

Второе равенство вытекает из первого и соотношения (12.1) ▶

Следствие 12.3. Множество всех целых степеней элемента g группы G — подгруппа в группе G.

Определение 12.4. Подгруппа, состоящая из всех целых степеней элемента g группы G обозначается через $< g >$ и называется *циклической подгруппой, порожденной* элементом g. Группа G называется *циклической*, если $G =< g >$ для некоторого элемента g группы G.

2°. Порядок элемента.

Следствие 12.5. Любая циклическая группа коммутативна.

◀ Следует из тождества $g^m g^n = g^{m+n} = g^n g^m$.▶

Определение 12.6. *Порядком* элемента g группы G называется наименьшее положительное число n такое, что $g^n = e$. Обозначение $n = \operatorname{ord}(g)$. Если такого числа n не существует, говорят, что g — элемент *бесконечного порядка*, $\operatorname{ord}(g) = \infty$.

Теорема 12.7. Для любого элемента g группы G, $\mathrm{ord}(g) = |< g >|$.

◀ Достаточно убедиться, что различные степени элемента бесконечного порядка различны, а среди степеней элемента G конечного порядка n различны $e = g^0, g = g^1, \ldots, g^{n-1}$. Предположим противное, т.е. $g^k = g^l$ при $k > l$. Тогда $g^{k-l} = e$, значит, порядок g конечен. Но при $k < n = \mathrm{ord}(g)$ получаем противоречие. С другой стороны, если $n = \mathrm{ord}(g)$, то любая степень элемента g совпадает с одной из указанных. Действительно, если $k \geqslant n$, то $k = nq + r$, где r — остаток от деления k на n, значит, $0 \leqslant k < n$ и $g^k = (g^n)^q g^r = e^q g^r = g^r$. Если же $k < 0$, то $g^k = (g^{n-1})^{-k} = g^{-k(n-1)}$ и принадлежит данному множеству, по доказанному. ▶

Теорема 12.8. Любая подгруппа любой циклической группы — циклическая группа.

◀ Пусть $G =< g >$, H — подгруппа в G. Если $H = \{e\}$, то $H =< e >$. Иначе существует $k = \min\{n > 0 : g^n \in H\}$. Тогда если $g^m \in H$, разделим с остатком m на k: $m = kq + r$, $0 \leqslant r < k$ и получим $g^r = g^m (g^k)^{-q} \in H$, и в силу выбора k, $r = 0$, т.е. $g^m = (g^k)^q$. Значит, $H =< g^k >$.▶

Теорема 12.9. Две циклические группы изоморфны тогда и только тогда, когда их порядки (конечные или бесконечные) равны.

◀ Если две группы изоморфны, то их порядки, очевидно, равны, ведь изоморфизм — взаимно-однозначное отображение.

Обратно, пусть $G =< a >$, $H =< b >$ и $|G| = |H|$.

Если $|G| = |H| = \infty$, то отображение $a^k \mapsto b^k, \quad \forall k \in \mathbb{Z}$ определено корректно, так как различные степени элемента a различны, и из теоремы 12.2 следует, что это отображение — изоморфизм.

Если же $|G| = |H| = n < \infty$, то, как мы видели при доказательстве теоремы 12.7,

$$G = \{e = a^0, a = a^1, \ldots, a^{n-1}\},$$

$$H = \{e = b^0, b = b^1, \ldots, b^{n-1}\},$$

и снова легко проверить, что отображение $f : a^k \mapsto b^k, \quad 0 \leqslant k < n$ — изоморфизм. Действительно, для любого $m \in \mathbb{Z}$ можно записать $m = nq + r$, где $0 \leqslant r < n$, откуда $a^m = a^{nq} a^r = a^r$, $b^m = b^{nq} b^r = b^r$, значит, $f(a^m) = f(a^r) = b^r = b^m$, и снова можно воспользоваться теоремой 12.2.▶

Бесконечная циклическая группа — это, например, группа $(\mathbb{Z}, +)$. Построить конечную циклическую группу порядка n можно разными способами: как множество классов вычетов по модулю n (см. построение кольца вычетов ниже), как группу комплексных корней из 1 степени n (см. ниже), или как подгруппу в S_n, порожденную циклом длины n. Вообще, как показывает следующая теорема, из группы S_n можно "построить" любую конечную группу.

Теорема 12.10. (**Теорема Кэли**). Любая группа конечного порядка n изоморфна некоторой подгруппе группы подстановок S_n.

◀ Для любого элемента g группы G конечного порядка n определим отображение $\sigma_g : G \to G$ равенством

$$\sigma_g(x) = gx \ \ \forall x \in G.$$

Очевидно, что σ_g — биективное отображение: обратное к нему отображение имеет вид $\sigma_{g^{-1}}$. Для любых элементов $g, h \in G$ имеем

$$\sigma_{gh}(x) = (gh)x = g(hx) = g(\sigma_h(x)) = \sigma_g(\sigma_h(x)) \ \ \forall x \in G,$$

т.е. отображение $g \mapsto \sigma_g$ — гомоморфизм группы G в S_n. При этом при $g \neq h$ имеем $\sigma_g(e) = ge = g \neq h = he = \sigma_h(e)$, следовательно, образ G при этом гомоморфизме — подгруппа в S_n, изоморфная группе G.►

3°. Смежные классы. Теорема Лагранжа.

Определение 12.11. Пусть G — группа, H — подгруппа группы G, $g \in G$ — произвольный элемент группы G. Множество $gH = \{gh : h \in H\}$ называется *левым смежным классом элемента g по подгруппе H.*

Например, $H = eH$ — левый смежный класс.

Укажем основные свойства левых смежных классов.

Предложение 12.12. Пусть H — подгруппа группы G. Тогда
1) Если $g \in G$ и $g' \in gH$, то $g'H = gH$.
2) Если g_1H и g_2H — два левых смежных класса, то они либо равны, либо имеют пустое пересечение.
3) $\bigcup_{g\in G} gH = G$.
4) Если группа H конечна, то $|gH| = |H|$ для любого левого смежного класса gH.

◄
1). Пусть $g \in G$ и $g' \in gH$, тогда существует элемент $h' \in H$, для которого $g' = gh'$. Для любого $h \in H$ имеем $g'h = (gh')h = g(h'h) \in gH$, откуда $g'H \subseteq gH$. Обратно, $gh = (g'h'^{-1})h = g'(h'^{-1}h) \in g'H$, значит, $gH \subseteq g'H$.
2) Допустим, что $g_1H \cap g_2H \neq \emptyset$. Тогда существует элемент $g \in g_1H \cap g_2H$, при этом, согласно 1), $g_1H = gH = g_2H$.
3) По определению подгруппы, $e \in H$, значит, $g = ge \in gH$ для любого $g \in G$.
4) Заметим, что отображение $h \mapsto gh$ устанавливает взаимно-однозначное соответствие между элементами подгруппы H и элементами левого смежного класса gH: действительно, из $gh_1 = gh_2$ умножением слева на g^{-1} получаем $h_1 = h_2$. Значит, мощности множеств H и gH равны.►

Замечание 1. Левые смежные классы — это классы эквивалентности относительно отношения эквивалентности вида

$$g \sim f \overset{\text{def}}{\Leftrightarrow} f^{-1}g \in H.$$

Упражнение. Доказать, что определенное выше отношение на самом деле есть отношение эквивалентности, т.е. обладает свойствами:
1) $g \sim g \quad \forall g \in G$ (рефлексивность);
2) $g \sim f \Rightarrow f \sim g$ (симметричность); 3) $g \sim f, f \sim u \Rightarrow g \sim u$ (транзитивность).

Замечание 2. Аналогично определяются правые смежные классы Hg элементов группы G по ее подгруппе H. Их свойства аналогичны свойствам левых смежных классов.

Теорема 12.13. (**теорема Лагранжа**). Пусть G — конечная группа, H — подгруппа в G. Тогда порядок подгруппы H делит порядок группы G.

◄ Пусть k — число левых смежных классов элементов группы G по подгруппе H. Из предложения 12.12(2,4) следует, что всего в них сдержится $k|H|$ элементов, а из 12.12(3) — что $|G| = k|H|$.►

Следствие 12.14. Порядок любого элемента конечной группы делит порядок группы.

◄ Действительно, если $g \in G$, то $\mathrm{ord}(g) \overset{12.7}{=} |< g >| \,\big|\, |G|$.►

Следствие 12.15. Если порядок группы G — простое число, то группа G — циклическая.

◄ Пусть $|G| = p$ — простое число. Поскольку $p > 1$, в G есть элемент $g \neq e$. Положим $H =< g >$, тогда $|H| > 1$ и $|H| \,\big|\, p$, значит, $|H| = p$ и $H = G$.►

4°. Нормальные подгруппы и фактор-группы.

Определение 12.16. Подгруппа H группы G называется *нормальной*, если левые смежные классы по H совпадают с правыми:

$$gH = Hg \qquad \forall g \in G.$$

Обозначение: $H \lhd G$.

Предложение 12.17. Подгруппа H группы G нормальна тогда и только тогда, когда

$$\forall g \in G, h \in H, \quad ghg^{-1} \in H. \tag{12.2}$$

◄ Пусть $H \lhd G$, $g \in G, h \in H$. Тогда $gh \in Hg$, т.е. $gh = h'g$ для некоторого $h' \in H$. Но тогда $ghg^{-1} = h' \in H$. Обратно, пусть выполнено (12.2), и $g \in G$. Покажем, что $gH \subseteq Hg$. Для любого $h \in H$ положим $ghg^{-1} = h' \in H$. Получим $gh = h'g \in Hg$. Значит, $gH \subseteq Hg$. Аналогично, применим (12.2) к g^{-1} вместо g: $g^{-1}hg = h' \in H$, откуда $hg = gh'$ и, следовательно, $Hg \subseteq gH$. Итак, $gH = Hg$, и $H \lhd G$.►

Следствие 12.18. Если группа G коммутативна, то любая подгруппа в G нормальна.

◄ Используем 12.17 и учитывем, что в коммутативной группе $ghg^{-1} = h$.►

Еще примеры нормальных подгрупп (доказательство нормальности — **упражнение**):
$\{e\} \lhd G$ и $G \lhd G$ для любой группы G.
$A_n \lhd S_n$.
$SL_n(\mathbb{R}) \lhd GL_n(\mathbb{R})$.

На множестве смежных классов группы G по нормальной подгруппе H можно определить операцию следующим образом:

$$g_1H \cdot g_2H = (g_1g_2)H \qquad \forall g_1, g_2 \in G \tag{12.3}$$

Необходимо проверить корректность определения: пусть $g_1'H = g_1H$, $g_2'H = g_2H$. Это значит, что $g_1' = g_1h_1$ и $g_2' = g_2h_2$ для некоторых $h_1, h_2 \in H$, откуда

$$g_1'g_2' = g_1h_1g_2h_2 = g_1g_2 \underbrace{g_2^{-1}h_1g_2}_{\in H} h_2 \in g_1g_2H.$$

Следовательно, $g_1'g_2'H = g_1g_2H$ в силу 12.12(1). Легко проверить, что указанная операция задает структуру группы на множестве смежных классов (проверка оставляется читателю в качестве **упражнения**).

Определение 12.19. Множество всех смежных классов группы G по нормальной подгруппе H с операцией (12.3) называется *фактор-группой* группы G по H и обозначается: G/H.

Пример: $\mathbb{Z}/n\mathbb{Z}$ — циклическая группа из n элементов. Как мы увидим в следующей лекции, на этой группе имеется более богатая структура — структура кольца.

Лекция 13. Основные алгебраические структуры: кольца, поля.

1°. Понятие кольца.

Определение 13.1. Алгебраическая структура с двумя бинарными операциями $(R, +, \cdot)$ называется *кольцом*, если выполняются следующие аксиомы:
1) $(R, +)$ — абелева группа с нейтральным элементом 0;
2) $\forall a, b, c \in R, \quad (a + b)c = ac + bc$;
3) $\forall a, b, c \in R, \quad a(b + c) = ab + ac$.

Последние две аксиомы называют аксиомами, или законами, *дистрибутивности*.

Если выполняется аксиома ассоциативности умножения,
т.е. $\forall a, b, c \in R, \quad (ab)c = a(bc)$,
то кольцо R называется ассоциативным кольцом.

Если выполняется аксиома коммутативности умножения,
т.е. $\forall a, b \in R, \quad ab = ba$,
то кольцо R называется коммутативным кольцом.

Если в кольце R имеется элемент, нейтральный относительно умножения (обычно обозначается буквой e или цифрой 1), то говорят, что R — *кольцо с единицей*.

Группу $(R, +)$ из аксиомы 1) называют *аддитивной группой* кольца R.

Предложение 13.2. В любом кольце R выполняются соотношения:

$$a \cdot 0 = 0 \cdot a = 0 \qquad \forall a \in R;$$

$$a(-b) = (-a)b = -(ab), (-a)(-b) = ab \qquad \forall a, b \in R.$$

◀ Пусть a — элемент кольца R. Имеем $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$. Прибавим к обеим частям полученного равенства $a \cdot 0 = a \cdot 0 + a \cdot 0$ элемент $-(a \cdot 0)$:

$$a \cdot 0 + (-(a \cdot 0)) = 0 = a \cdot 0 + a \cdot 0 + (-(a \cdot 0)) = a \cdot 0 + 0 = a \cdot 0.$$

Равенство $0 \cdot a = 0$ доказывается аналогично. Теперь видно, что

$$a(-b) + ab = a(b + (-b)) = a0 = 0 \Rightarrow a(-b) = -(ab).$$

Соотношение $(-a)b = -(ab)$ доказывается аналогично. Наконец, $(-a)(-b) = -(a(-b)) = -(-ab) = ab$.▶

В нашем курсе, если не оговорено противное, все кольца предполагаются ассоциативными и имеющими единицу.

Примеры:
$\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ с обычными операциями сложения и умножения чисел (коммутативные кольца);
$M_n(\mathbb{R})$ — множество всех квадратных матриц порядка n относительно сложения и умножения матриц (некоммутативное кольцо при $n > 1$).

2°. Кольцо вычетов. Пусть n — натуральное число. Рассмотрим фактор-группу $(\mathbb{Z}/n\mathbb{Z}, +)$ группы $(\mathbb{Z}, +)$ по подгруппе $n\mathbb{Z}$. Смежные классы в этом случае имеют вид $a + n\mathbb{Z}$, $a \in \mathbb{Z}$. Введем более короткое обозначение: $a + n\mathbb{Z} = \overline{a}$. Операция сложения на фактор-группе тогда определена равенством

$$\overline{a} + \overline{b} = \overline{a + b}.$$

Определим еще операцию умножения на множестве $\mathbb{Z}/n\mathbb{Z}$ равенством

$$\overline{a} \cdot \overline{b} = \overline{ab} \quad \forall a, b \in \mathbb{Z}. \tag{13.1}$$

Проверим, что эта операция определена корректно. Пусть $\overline{a'} = \overline{a}$ и $\overline{b'} = \overline{b}$. Это значит, что существуют целые числа k, l, такие, что $a' = a+nk$, $b' = b+nl$. Тогда $a'b' = ab+anl+nkb+n^2kl = ab + n(al + bk + nkl)$, т.е. $a'b' \in ab + n\mathbb{Z}$, или $\overline{a'b'} = \overline{ab}$.

Легко убедиться, что указанные выше операции на множестве смежных классов $\mathbb{Z}/n\mathbb{Z}$ определяют структуру ассоциативного коммутативного кольца с единицей. Мы будем обозначать это кольцо через $\mathbb{Z}_n$.

Предложение 13.3. Кольцо $\mathbb{Z}_n$ содержит n элементов:

$$\overline{0}, \overline{1}, \ldots \overline{n-1}.$$

◀ Любое целое число m можно представить в виде $m = nq + r$, где $0 \leqslant r < n$ (деление с остатком). Действительно, пусть q — наибольшее целое число, для которого $r = m - nq \geqslant 0$. Если бы выполнялось неравенство $r \geqslant n$, то получилось бы, что $0 \leqslant r - n = m - (q+1)n$, что противоречит выбору q. Теперь видно, что $\overline{m} = \overline{r}$. Осталось заметить, что смежные классы $\overline{0}, \ldots, \overline{n-1}$ различны. В самом деле, если $0 \leqslant i < j < n$, и $\overline{i} = \overline{j}$, то получается противоречие: $0 < j - i = nk < n$, где k — целое число.▶

Для примера составим таблицы сложения и умножения в кольце $\mathbb{Z}_6$:

Сложение: $a + b$

$a \diagdown b$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{4}$	$\overline{4}$	$\overline{5}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{5}$	$\overline{5}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$

Умножение: ab

$a \diagdown b$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{4}$	$\overline{0}$	$\overline{2}$	$\overline{4}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{0}$	$\overline{3}$
$\overline{4}$	$\overline{0}$	$\overline{4}$	$\overline{2}$	$\overline{0}$	$\overline{4}$	$\overline{2}$
$\overline{5}$	$\overline{0}$	$\overline{5}$	$\overline{4}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

3°. Обратимые элементы и делители нуля в кольцах.

Определение 13.4. Пусть R — кольцо с единицей e. Элемент $a \in R$ называется *обратимым*, если существует элемент $b \in R$, такой, что $ab = ba = e$. Как и в случае групп, можно проверить единственность элемента b и ввести для него обозначение a^{-1}.

Предложение 13.5. Обратимые элементы кольца образуют группу относительно операции умножения в этом кольце.

◀ Ассоциативность умножения предполагается: мы рассматриваем только ассоциативные кольца. Единица кольца, очевидно, обратимый элемент: $ee = e$. Наконец, аксиома обратимости выполняется по определению обратимого элемента.▶

Группа, определенная в 13.5, называется *мультипликативной* группой кольца R и обозначается R^*.

Определение 13.6. Пусть R — кольцо. Элемент $a \in R$, отличный от 0, называется *левым делителем нуля*, если существует элемент $b \in R$, отличный от 0, для которого $ab = 0$. Аналогично определяются *правые делители нуля*. Элемент, являющийся левым или правым делителем нуля, называется *делителем нуля*.

Упражнение. Доказать, что для квадратной матрицы $A \neq 0$ порядка n равносильны условия:
(1) A — вырожденная матрица;
(2) A — левый делитель нуля в кольце $M_n(\mathbb{R})$;
(3) A — правый делитель нуля в кольце $M_n(\mathbb{R})$.

Предложение 13.7. Обратимый элемент кольца не может быть делителем нуля.

◀ Пусть $a \in R^*$ и $ab = 0$ для некоторого $b \in R$. Тогда $a^{-1}ab = b$ и $a^{-1}ab = a^{-1}0 = 0$, откуда $b = 0$. Значит, a не является левым делителем нуля. Аналогично проверяется, что a не является правым делителем нуля.▶

4°. Поля.

Определение 13.8. *Полем* называется ассоциативное коммутативное кольцо с единицей $e \neq 0$, в котором каждый ненулевой элемент обратим.

Иными словами, ассоциативное коммутативное кольцо R есть поле, если $R^* = R \setminus \{0\}$.

Примеры полей:
$\mathbb{Q}$, $\mathbb{R}$.

Подчеркнем особую важность понятия поля: все, что говорилось о линейных уравнениях, линейных зависимостях, матрицах, определителях и т.д. над полем $\mathbb{R}$, на самом деле верно для уравнений, векторов, матриц, определителей над произвольным полем. Надо только вспомнить, что в рассуждениях на эти темы мы пользовались только операциями сложения, вычитания и деления на число, не равное 0.

Теорема 13.9. Кольцо вычетов $\mathbb{Z}_n$ является полем тогда и только тогда, когда n — простое число.

◀ Пусть $\mathbb{Z}_n$ — поле. Если $n = ab$, где $1 < a < n$, то $1 < b < n$, поэтому $\overline{a} \neq 0$, $\overline{b} \neq 0$, но $\overline{a}\overline{b} = \overline{ab} = \overline{n} = \overline{0}$, т.е. $\overline{a}$ — делитель нуля в $\mathbb{Z}_n$, что противоречит 13.7.

Обратно, пусть n — простое число, и $\overline{a}$ — ненулевой элемент кольца $\mathbb{Z}_n$, т. е. a не делится на n. Рассмотрим классы $\overline{0}, \overline{a}, \overline{2a}, \ldots, \overline{(n-1)a}$ и покажем, что все эти классы различны. Действительно, если $\overline{ai} = \overline{aj}$ при $0 \leqslant i < j < n$, то $n \mid a(j-i)$, что невозможно, так как ни одно из чисел a и $j - i$ не делится на простое число n. Значит, среди этих классов есть и класс $\overline{1}$, т.е. $\overline{ai} = \overline{a}\overline{i} = \overline{1}$ для некоторого $i \in \{1, 2, \ldots, n-1\}$, т.е. $\overline{a}$ — обратимый элемент.▶

Заметим, что при доказательстве мы воспользовались известным еще Евклиду фактом: разложение натурального числа на простые множители единственно. В дальнейшем мы докажем аналогичное утверждение для многочленов. Наше доказательство можно будет легко переделать в доказательство для целых чисел.

Еще заметим, что фактически мы доказали следующий факт: конечное коммутативное ассоциативное кольцо без делителей нуля с единицей $e \neq 0$ — поле. Замечательная теорема, принадлежащая Веддербарну, утверждает, что любое ненулевое конечное ассоциативное кольцо без делителей нуля — поле, т.е. в случае конечного кольца и коммутативность, и наличие единицы следуют из ассоциативности.

Итак, для любого простого числа p мы построили поле $\mathbb{Z}_p$ из p элементов.

Определение 13.10. *Характеристикой* поля F (обозначение $\operatorname{char} F$) называется порядок единичного элемента e в аддитивной группе $(F, +)$. Если этот порядок бесконечен, то считают, что $\operatorname{char} F = 0$.

Примеры: $\operatorname{char}\mathbb{Q} = \operatorname{char}\mathbb{R} = 0$, $\operatorname{char}\mathbb{Z}_p = p$, если p — простое число.

Теорема 13.11. Характеристика любого поля F — либо 0, либо простое число.

◀ Фактически мы повторим часть доказательства теоремы 13.9. Пусть $\operatorname{char} F = n > 0$, т.е. n — наименьшее натуральное число, для которого $n \cdot e = 0$. Если $n = ab$, где $1 < a < n$, то $1 < b < n$, то $a \cdot e \neq 0$, $b \cdot e \neq 0$, но

$$(a \cdot e)(b \cdot e) = (\underbrace{e + e + \ldots + e}_{a \text{ раз}})(\underbrace{e + e + \ldots + e}_{b \text{ раз}}) = (\underbrace{e + e + \ldots + e}_{ab=n \text{ раз}}) = n \cdot e = 0.$$

Получается, что $a \cdot e$ — делитель нуля в поле, чего не может быть ввиду 13.7.▶

Отметим, что существуют конечные поля, не изоморфные $\mathbb{Z}_p$ для простых p. Именно, справедлива

Теорема 13.12. Для любого простого числа p и любого натурального числа k существует единственное, с точностью до изоморфизма, поле из p^k элементов (оно обозначается $GF(p^k)$, сокращение слов Galois field — поле Галуа.)

На доказательство этой теоремы у нас нет времени, но пример — поле $GF(4)$ — мы построим.

Заметим, что если F — поле из четырех элементов, то кроме 0 и 1 в нем есть еще 2 элемента. Обозначим их α и β. Положим $\operatorname{char} F = p$. Поскольку, по следствию из теоремы Лагранжа, $p \mid 4$, и p — простое число, получаем $p = 2$. Таким образом, в F 1+1=0, поэтому $\alpha + \alpha = \beta + \beta = 0$. Далее, $\alpha + 1 \neq 0$, $\alpha + 1 \neq 1$ и $\alpha + 1 \neq \alpha$, значит, $\alpha + 1 = \beta$. Это полностью определяет таблицу сложения в поле F:

Сложение: $a + b$

$a \diagdown b$	0	1	α	β
0	0	1	α	β
1	1	0	β	α
α	α	β	0	1
β	β	α	1	0

Перейдем к таблице умножения. Достаточно определить $\alpha\alpha$, поскольку тогда остальные произведения находятся однозначно. Но $\alpha\alpha \neq 0$, иначе α — делитель нуля. Аналогично, $\alpha\alpha \neq \alpha$, иначе $\alpha(\alpha + 1) = 0$ и опять α — делитель нуля. Наконец, если $\alpha\alpha = 1$, то $\beta\beta = (\alpha + 1)(\alpha + 1) = 1 + \alpha + \alpha + 1 = 0$, значит β — делитель нуля. Остается только одна возможность: $\alpha\alpha = \beta$. Теперь таблицу умножения в F легко достроить (например, $\alpha\beta = \alpha(\alpha + 1) = \alpha\alpha + \alpha = \alpha + 1 + \alpha = 1$):

Умножение: ab

$a \diagdown b$	0	1	α	β
0	0	0	0	0
1	0	1	α	β
α	0	α	β	1
β	0	β	1	α

Строго говоря, требуется еще проверка выполнения аксиом кольца для введенных операций, но это можно сделать либо простым, но длинным перебором, либо вывести из общих соображений, доказывающих теорему 13.12.

Упомянем без доказательства еще один важный факт:

Теорема 13.13. Мультипликативная группа конечного поля — циклическая группа.

Например, в построенном выше поле F из 4 элементов $F^* = \{1, \alpha, \beta\} = \{1, \alpha, \alpha^2\} = <\alpha>$.

Лекция 14. Поле комплексных чисел. Комплексная плоскость. Модуль и аргумент комплексного числа. Алгебраическая и тригонометрическая форма записи комплексных чисел. Операция сопряжения комплексных чисел и ее свойства. Формула Муавра. Корни целой степени из комплексного числа. Группа комплексных корней из единицы.

1°. Построение поля комплексных чисел.

На множестве $\mathbb{R}\times\mathbb{R}$ определим операции сложения и умножения формулами

$$\begin{array}{l}(a_1,b_1)+(a_2,b_2)=(a_1+a_2,b_1+b_2)\\(a_1,b_1)(a_2,b_2)=(a_1a_2-b_1b_2,a_1b_2+b_1a_2)\end{array}\qquad\forall(a_1,b_1),(a_2,b_2)\in\mathbb{R}\times\mathbb{R}.\tag{14.1}$$

Теорема 14.1. Множество $\mathbb{R}\times\mathbb{R}$ с операциями (14.1) является полем.

◀ Все аксиомы поля могут быть проверены непосредственно. Однако, проще заметить, что каждой паре $(a,b)\in\mathbb{R}\times\mathbb{R}$ можно сопоставить матрицу $\begin{pmatrix}a&b\\-b&a\end{pmatrix}$, и тогда операции (14.1) превращаются в обычные операции над матрицами: для сложения это очевидно, а для умножения вычислим:

$$\begin{pmatrix}a_1&b_1\\-b_1&a_1\end{pmatrix}\begin{pmatrix}a_2&b_2\\-b_2&a_2\end{pmatrix}=\begin{pmatrix}a_1a_2-b_1b_2&a_1b_2+b_1a_2\\-b_1a_2-a_1b_2&-b_1b_2+a_1a_2\end{pmatrix}\quad\forall(a_1,b_1),(a_2,b_2)\in\mathbb{R}\times\mathbb{R}.\tag{14.2}$$

Это сразу доказывает выполнение аксиом аддитивной группы, ассоциативности, дистрибутивности и существования единицы. Коммутативность умножения очевидна из симметричности выражения для $(a_1,b_1)(a_2,b_2)$. Осталось заметить, что если

$$0\neq A=\begin{pmatrix}a&b\\-b&a\end{pmatrix},$$

то

$$|A|=a^2+b^2\neq 0,$$

и существует матрица

$$A^{-1}=\frac{1}{|A|}\begin{pmatrix}a&-b\\b&a\end{pmatrix},$$

соответствующая паре действительных чисел $\left(\dfrac{a}{a^2+b^2},-\dfrac{b}{a^2+b^2}\right)$.▶

Определение 14.2. Поле $\mathbb{C}=\{(a,b):a,b\in\mathbb{R}\}$ с операциями (14.1) называется *полем комплексных чисел*. Плоскость с соответствующей системой координат называют комплексной плоскостью.

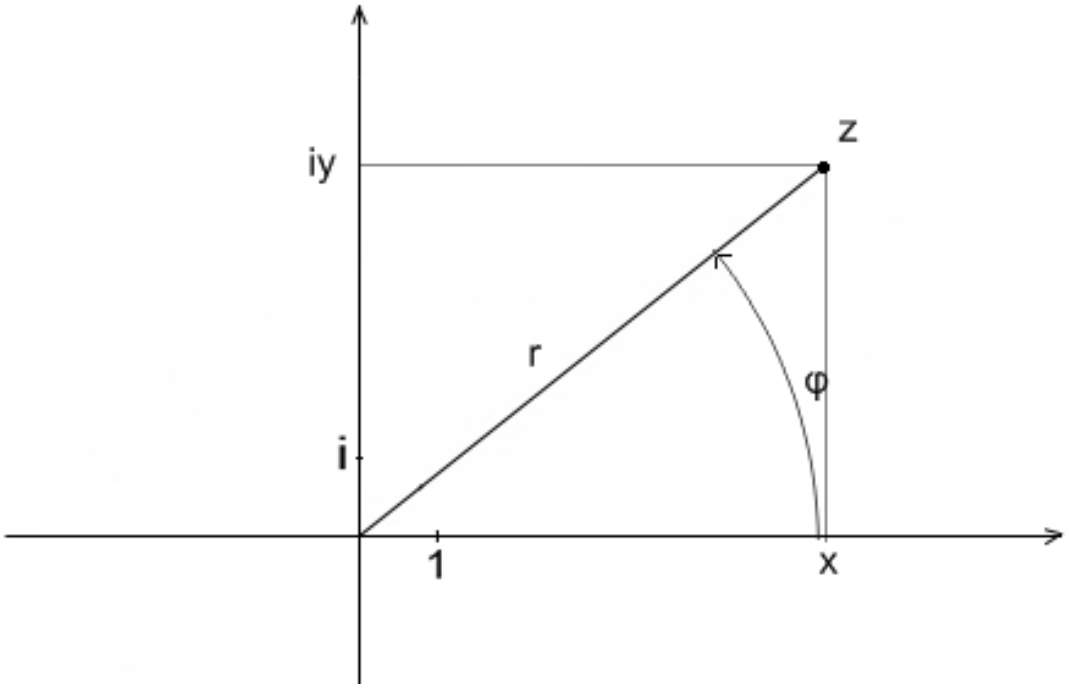
Легко видеть, что пары вида $(a,0), a\in\mathbb{R}$, соответствуют скалярным матрицам, и, следовательно, образуют подполе в $\mathbb{C}$, изоморфное полю $\mathbb{R}$. Поэтому можно отождествить пару $(a,0)$ с действительным числом a (точки оси абсцисс на комплексной плоскости рассматривать как точки действительной числовой оси), поэтому ось абсцисс также называют *действительной осью* комплексной плоскости. Далее, положим $i=(0,1)$. Тогда

$$i^2=(0,1)(0,1)\overset{(14.1)}{=}(0\cdot0-1\cdot1,1\cdot0+0\cdot1)=(-1,0)=-1.$$

Это объясняет название комплексного числа i — *мнимая единица*, а также чисел вида $bi=(b,0)(0,1)$, $b\in\mathbb{R}$ — *чисто мнимые числа*. На комплексной плоскости эти числа соответствуют точкам оси ординат, поэтому ось ординат называют *мнимой осью* комплексной

плоскости. Далее, любое комплексное число $z = (x, y)$ можно записать в виде $z = x + iy$. Соответственно, $x = \operatorname{Re} z$ и $y = \operatorname{Im} z$ называются *действительной* и *мнимой* частью числа $z = x + iy$.

2°. Модуль и аргумент комплексного числа. Введем на комплексной плоскости, наряду с декартовой, полярную систему координат (r, φ):



Тогда $x = r\cos\varphi$, $y = r\sin\varphi$, $r = \sqrt{x^2 + y^2}$.

Определение 14.3. Неотрицательное действительное число $|z| = r = \sqrt{x^2 + y^2}$ называется *модулем*, а угол $\arg z = \varphi$ — *аргументом* комплексного числа $z = x + yi$. Очевидно, из определения следует $z = r(\cos\varphi + i\sin\varphi)$. Последнее равенство называется *тригонометрической формой* записи комплексного числа. Следует заметимть, что аргумент ненулевого комплексного числа определен с точностью до слагаемого, кратного 2π, а аргумент числа $z = 0$ не определен.

Предложение 14.4. (**Неравенство треугольника.**) Для любых двух чисел $z, w \in \mathbb{C}$,

$$|z + w| \leqslant |z| + |w|.$$

◄ Комплексные числа складываются как обычные векторы, по “правилу параллелограмма”.►

Предложение 14.5. Для любых двух чисел $z, w \in \mathbb{C}$,

$$\begin{aligned} |zw| &= |z||w| \\ \arg(zw) &= \arg z + \arg w. \end{aligned} \tag{14.3}$$

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

$$\begin{aligned} zw &= rs(\cos\varphi + i\sin\varphi)(\cos\psi + i\sin\psi) \\ &= rs((\cos\varphi\cos\psi - \sin\varphi\sin\psi) + i(\cos\varphi\sin\psi + \sin\varphi\cos\psi)) \\ &= rs(\cos(\varphi + \psi) + i\sin(\varphi + \psi)). \end{aligned}$$

►

3°. Операция сопряжения.

Определение 14.6. Пусть $z = x + iy \in \mathbb{C}$. Число $\overline{z} = x - iy$ называется *сопряженным* к числу z.

Предложение 14.7. Для любых чисел $z, w \in \mathbb{C}$ выполнены равенства:
1) $\overline{\overline{z}} = z$;
2) $\overline{z + w} = \overline{z} + \overline{w}$;
3) $\overline{zw} = \overline{z}\,\overline{w}$;
4) $z + \overline{z} = 2\operatorname{Re} z \in \mathbb{R}, \quad z\overline{z} = |z|^2 \in \mathbb{R}$;
5) $|\overline{z}| = |z|, \quad \arg\overline{z} = -\arg z$.

◄ Равенства 1, 2, 4, 5 очевидны. Равенство 3 можно проверить непосредственно с помощью (14.1), а можно заметить, что сопряжению комплексного числа $a + bi$ соответствует транспонирование матрицы $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.►

4°. Формула Муавра. Корни из комплексного числа.

Теорема 14.8. Для любого комплексного числа $z = r(\cos\varphi + i\sin\varphi)$ и любого натурального числа n выполняется *формула Муавра*:

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

◄ Индукция по n. При $n = 1$ равенство выполняется по определению z. Рассмотрим z^{n+1}:

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

►

Выведем из формулы Муавра выражение для корней натуральной степени из комплексного числа $z \neq 0$.

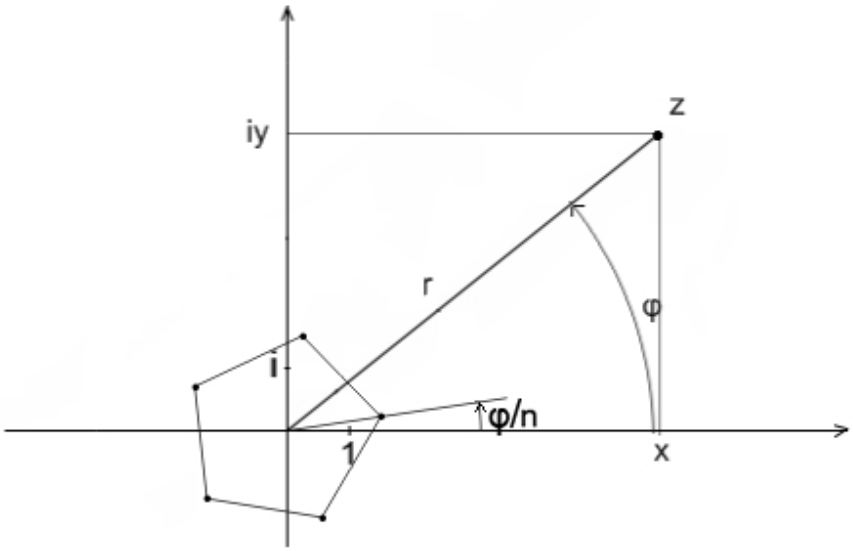
Определение 14.9. Обозначим через $\sqrt[n]{z}$ множество всех комплексных решений уравнения $w^n = z$.

Пусть $z = r(\cos\varphi + i\sin\varphi)$, $w = s(\cos\psi + i\sin\psi)$ и $w^n = z$. Тогда $s^n = r$, откуда $s = \sqrt[n]{r}$ — положительный действительный корень из положительного действительного числа, что определяет число $|w|$. Равенства $\cos(n\psi) = \cos\varphi$ и $\sin(n\psi) = \sin\varphi$ равносильны тому, что $n\psi = \varphi + 2\pi k$, где $k \in \mathbb{Z}$. Получаем $\psi = \dfrac{\varphi + 2\pi k}{n}, \quad k \in \mathbb{Z}$, причем числа $w_k = s\left(\cos\frac{\varphi+2\pi k}{n} + i\sin\frac{\varphi+2\pi k}{n}\right)$ и $w_{k'} = s(\cos\frac{\varphi+2\pi k'}{n} + i\sin\frac{\varphi+2\pi k'}{n})$ совпадают тогда и только тогда, когда $k \equiv k' \pmod n$. В результате получаем теорему.

Теорема 14.10. Для ненулевого комплексного числа $z = r(\cos\varphi + i\sin\varphi)$,

$$\sqrt[n]{z} = \left\{ \sqrt[n]{r}\left(\cos\frac{\varphi + 2\pi k}{n} + i\sin\frac{\varphi + 2\pi k}{n}\right) : k = 0, 1, \ldots, n-1 \right\}. \tag{14.4}$$

Ясно, что элементы указанного множества — вершины правильного n-угольника с центром в начале координат, причем одна из вершин n-угольника имеет полярные координаты $(\sqrt[n]{r}, \frac{\varphi}{n})$:



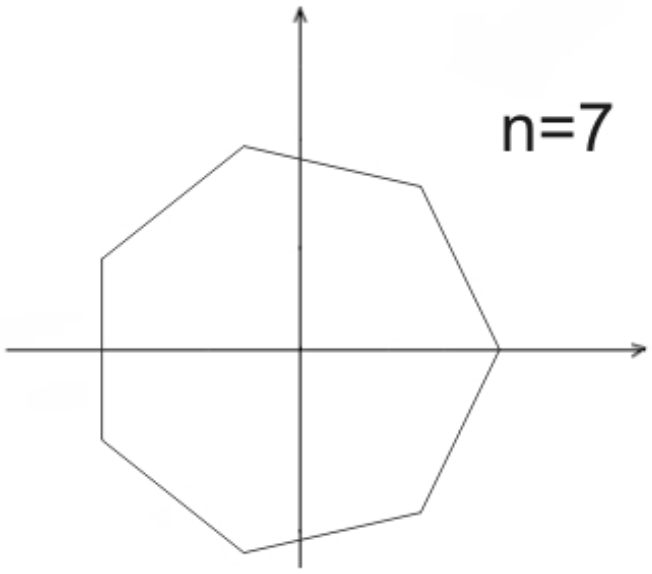
5°. Группа корней степени n из единицы. Рассмотрим случай $z = 1$ в (14.4). Поскольку $1 = 1(\cos 0 + i \sin 0)$, имеем

$$\sqrt[n]{1} = \left\{ \cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n} : k = 0, 1, \ldots, n-1 \right\}. \tag{14.5}$$

Предложение 14.11. Множество U_n корней из единицы степени n является циклической группой порядка n относительно умножения в $\mathbb{C}$.

◄ Сначала покажем, что U_n — подгруппа в $\mathbb{C}^*$. Достаточно проверить три факта: $1^n = 1 \Rightarrow 1 \in U_n$, $a, b \in U_n \Rightarrow (ab)^n = a^n b^n = 1 \cdot 1 = 1 \Rightarrow ab \in U_n$ и $a^n \in U_n \Rightarrow (a^{-1})^n = (a^n)^{-1} = 1^{-1} = 1$. Далее пусть $\varepsilon_k = \cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n}$, тогда по формуле Муавра $\varepsilon_k = \varepsilon_1^k$, т.е. $U_n =< \varepsilon_1 >$.► Порождающие элементы группы U_n называются *первообразными корнями* из единицы.

Геометрически корни степени n из единицы — это вершины правильного n-угольника с центром в начале координат, причем одна из вершин n-угольника расположена в точке 1 с координатами (1,0):



Предложение 14.12. Множество $U = \bigcup_{n>0} U_n$ всевозможных комплексных корней из 1 — подгруппа в $\mathbb{C}^*$.

◄ Достаточно показать, что если $a \in U_n$ и $b \in U_m$, то $ab \in U$. Но

$$(ab)^{mn} = a^{mn}b^{mn} = (a^n)^m(b^m)^n = 1^m 1^n = 1 \Rightarrow ab \in U_{mn} \subset U.$$

►

Лекция 15. Кольцо многочленов от одной переменной над полем. Возможность и единственность деления на ненулевой многочлен с остатком. Наибольший общий делитель двух многочленов, его выражение через многочлены, алгоритм Евклида.

1°. Кольцо многочленов от одной переменной над полем.

Определение 15.1. *Многочленом* над полем F от переменной x называется формальное выражение вида

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0, \tag{15.1}$$

где $a_0, a_1, \ldots, a_n$ произвольные элементы поля F, называемые *коэффициентами* многочлена, $n \geqslant 0$. Каждое слагаемое вида $a_i x^i$ называется *одночленом*. Многочлены, отличающиеся некоторым количеством слагаемых с нулевыми коэффициентами, считаются равными (более строго многочлен определяется как бесконечная сумма одночленов, в которых все коэффициенты, кроме конечного числа, равны нулю, и запись (15.1) — просто сокращение для этой бесконечной суммы). Операции над многочленами определяются следующим образом. Пусть

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0,$$

$$g(x) = b_m x^m + a_{m-1}x^{m-1} + \ldots + b_1 x + b_0.$$

Для определения суммы будем считать, при необходимости добавляя слагаемые с нулевыми коэффициентами к одному из многочленов, что $m = n$ и положим

$$f(x) + g(x) = (a_n + b_n)x^n + (a_{n-1} + b_{n-1})x^{n-1} + \ldots + (a_1 + b_1)x + (a_0 + b_0).$$

Произведение многочленов $f(x)$ и $g(x)$ определим так, чтобы выполнялись естественные равенства $x^i x^j = x^{i+j}$ и закон дистрибутивности:

$$f(x)g(x) = a_n b_m x^{n+m} + \ldots + (\sum_{i=0}^{k} a_i b_{k-i})x^k + \ldots + a_0 b_0.$$

(естественно, $a_i = 0$ при $i > n$ и $b_j = 0$ при $j > m$).

Предложение 15.2. Множество многочленов с указанными выше операциями — ассоциативное коммутативное кольцо с единицей.

◀ Проверим дистрибутивность. Пусть

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0,$$

$$g_1(x) = b_m x^m + b_{m-1}x^{m-1} + \ldots + b_1 x + b_0,$$

$$g_2(x) = b'_m x^m + b'_{m-1}x^{m-1} + \ldots + b'_1 x + b'_0.$$

Тогда

$$\begin{aligned} f(x)(g_1(x) + g_2(x)) &= a_n(b_m + b'_m)x^{n+m} + \ldots + \left(\textstyle\sum_{i=0}^{k} a_i(b_{k-i} + b'_{k-i})\right) x^k + \ldots + a_0(b_0 + b'_0) \\ &= (a_n b_m + a_n b'_m)x^{n+m} + \ldots + \\ &\quad \left((\textstyle\sum_{i=0}^{k} a_i b_{k-i}) + (\sum_{i=0}^{k} a_i b'_{k-i})\right) x^k + \ldots + \\ &\quad a_0 b_0 + a_0 b'_0 = f(x)g_1(x) + f(x)g_2(x). \end{aligned}$$

Ассоциативность и коммутативность кольца многочленов следует из дистрибутивности и того, что этими свойствами обладает умножение одночленов:

$$x^i(x^j x^k) = x^{i+j+k} = (x^i x^j)x^k; \qquad x^i x^j = x^{i+j} = x^j x^i.$$

►

Кольцо многочленов от переменной x над полем F обозначается $F[x]$. В данной лекции мы рассматриваем многочлены над произвольным фиксированным полем.

2°. Деление с остатком в кольце многочленов.

Определение 15.3. *Степенью* многочлена $f(x) = a_n x^n + \ldots + a_1 x + a_0$ называется число n, если $a_n \neq 0$. Обозначение: $\deg f(x)$.

Замечание. Согласно этому определению, степень нулевого многочлена не определена. Иногда удобно считать, что $\deg 0 = -\infty$.

Многочлены степени 0 (и нулевой многочлен) называют *константами*.

Предложение 15.4. Для любых двух ненулевых многочленов $f(x)$ и $g(x)$ над полем F, выполняются соотношения

$$\begin{array}{l} \deg(f(x)+g(x)) \leqslant \max\{\deg f(x), \deg g(x)\} \text{ (или } f(x)+g(x)=0); \\ \deg(f(x)g(x)) = \deg f(x) + \deg g(x). \end{array} \tag{15.2}$$

◄ Пусть $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$, $g(x) = b_m x^m + b_{m-1}x^{m-1} + \ldots + b_1 x + b_0$, причем $a_n \neq 0$, $b_m \neq 0$ и $n \geqslant m$. Полагая $b_{m+1} = \ldots = b_n = 0$, получим

$$f(x) + g(x) = (a_n + b_n)x^n + \ldots + (a_0 + b_0),$$

откуда $\deg(f(x)+g(x)) = n$ при $a_n + b_n \neq 0$ (в частности, при $m < n$) и $\deg(f(x)+g(x)) < n$ при $a_n + b_n = 0$.

Далее, $f(x)g(x) = a_n b_m x^{n+m} + \ldots + a_0 b_0$, и $a_n b_m \neq 0$, так как в поле нет делителей нуля, значит, $\deg(f(x)g(x)) = n + m$. ►

Следствие 15.5. В кольце многочленов над полем нет делителей нуля.

◄ Действительно, если $f(x) \neq 0$ и $g(x) \neq 0$, то $\deg(f(x)g(x)) = \deg f(x) + \deg g(x) \geqslant 0$.►

Теорема 15.6. Пусть F — поле. Для любого многочлена $f(x) \in F[x]$ и ненулевого многочлена $g(x)$ существуют единственные многочлены $q(x)$ и $r(x)$, удовлетворяющие условиям

$$\begin{array}{l} f(x) = q(x)g(x) + r(x); \\ r(x) = 0 \text{ либо } \deg r(x) < \deg g(x). \end{array} \tag{15.3}$$

◄ Докажем существование многочленов $q(x)$ и $r(x)$. Если $f(x) = 0$, то подходят многочлены $q(x) = r(x) = 0$. Поэтому будем считать, что $f(x) \neq 0$, $\deg f(x) = n$ и $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$. Обозначим $g(x) = b_m x^m + \ldots + b_0$, где $b_m \neq 0$, и докажем существование искомых многочленов индукцией по n. При $n < m$ можно положить $q(x) = 0$, $r(x) = f(x)$ — это дает базу индукции. Допустим, что для многочленов степени меньше n утверждение истинно, и докажем его для многочлена степени $n \geqslant m$. Рассмотрим многочлен

$$f_1(x) = f(x) - a_n b_m^{-1} g(x) x^{n-m} = a_n x^n + \ldots - a_n b_m^{-1} b_m x^m x^{n-m} - \ldots = a_n x^n + \ldots - a_n x^n - \ldots.$$

Видно, что $\deg f_1(x) < n$. По предположению индукции имеем $f_1(x) = q_1(x)g(x) + r(x)$, где $r(x) = 0$ либо $\deg r(x) < m$. Теперь

$$f(x) = f_1(x) + a_n b_m^{-1} g(x) x^{n-m} = \underbrace{\left(q_1(x) + a_n b_m^{-1} x^{n-m}\right)}_{=q(x)} g(x) + r(x).$$

Осталось проверить единственность. Предположим, что имеется два представления многочлена, удовлетворяющих (15.3):

$$f(x) = q_1(x)g(x) + r_1(x) = q_2(x)g(x) + r_2(x).$$

Тогда имеем

$$r_2(x) - r_1(x) = (q_1(x) - q_2(x))g(x).$$

Если $q_2(x) \neq q_1(x)$, то по (15.2) получим

$$\deg[(q_1(x) - q_2(x))g(x)] = m + \deg(q_1(x) - q_2(x)) \geqslant m,$$

в то время как

$$\deg(r_2(x) - r_1(x)) < m.$$

Получили противоречие. Значит, $q_2(x) = q_1(x)$, но тогда $r_2(x) = r_1(x)$. Единственность доказана.▶

Определение 15.7. Многочлены $q(x)$ и $r(x)$ из (15.3) называются, соответственно, *(неполным) частным* и *остатком* от деления многочлена $f(x)$ на многочлен $g(x)$.

Приведенное доказательство дает простой способ нахождения частного и остатка ("деление столбиком").

Пример:

$$\begin{array}{rl|l} & x^5 + 1 & x^2+1 \\ \cline{3-3} - & x^5 + x^3 & x^3 - x \text{ (частное)} \\ \cline{2-2} & -x^3 + 1 & \\ - & -x^3 - x & \\ \cline{2-2} & x + 1 \text{ (остаток)} & \end{array}$$

Определение 15.8. Говорят, что ненулевой многочлен $g(x)$ *делит* многочлен $f(x)$, и обозначают $g(x) \,\big|\, f(x)$, $f(x) = q(x)g(x)$ для некоторого многочлена $q(x)$. В этом случае остаток от деления многочлена $f(x)$ на многочлен $g(x)$ равен 0.

Определение 15.9. *Наибольшим общим делителем* двух многочленов $f(x)$ и $g(x)$ называется общий делитель этих двух многочленов, который делится на любой другой общий делитель многочленов $f(x)$ и $g(x)$. Обозначение: $(f(x), g(x))$.

Заметим, что существование наибольшего общего делителя требует доказательства. Мы дадим это доказательство и одновременно укажем способ нахождения наибольшего общего делителя.

Теорема 15.10. (**Алгоритм Евклида**). Пусть $f(x), g(x)$ — два многочлена над полем F. Тогда их наибольший общий делитель существует и может быть найден следующим образом.

Если один из многочленов равен нулю, то, очевидно, их наибольший общий делитель равен другому многочлену, поэтому далее предположим, что $f(x) \neq 0$ и $g(x) \neq 0$. Положим $r_0(x) = f(x)$, $r_1(x) = g(x)$ и выполним последовательно серию делений с остатком:

$$\begin{array}{rcl} r_0(x) & = & q_1(x)r_1(x) + r_2(x) \\ r_1(x) & = & q_2(x)r_2(x) + r_3(x) \\ \ldots\ldots & \ldots\ldots & \ldots\ldots\ldots\ldots \\ r_{k-1}(x) & = & q_k(x)r_k(x) + r_{k+1}(x) \end{array}$$

до тех пор, пока очередной остаток не окажется равным 0 (скажем, $r_{k+1}(x) = 0$). Тогда последний ненулевой остаток $r_k(x)$ — наибольший общий делитель многочленов $f(x)$ и $g(x)$.

◄ Положим $d(x) = r_k(x)$. Ясно, что $d(x) \,\big|\, r_k(x)$ и $d(x) \,\big|\, r_{k-1}(x)$, так как $r_{k+1}(x) = 0$. Но тогда

$$d(x) \,\big|\, r_{k-2}(x) = q_{k-1}(x)r_{k-1}(x) + r_k(x).$$

Повторяя это рассуждение $k-2$ раз, получим, что $d(x) \,\big|\, r_1(x) = g(x)$ и $d(x) \,\big|\, r_0(x) = f(x)$. Значит $d(x)$ — общий делитель указанных двух многочленов. Теперь пусть $h(x)$ — произвольный общий делитель $f(x)$ и $g(x)$. Из первого деления получаем, что

$$h(x) \,\big|\, r_2(x) = r_0(x) - q_1(x)r_1(x),$$

$$h(x) \,\big|\, r_3(x) = r_1(x) - q_2(x)r_2(x)$$

и т.д. В результате приходим к тому, что $h(x) \,\big|\, r_k(x) = d(x)$.►

Предложение 15.11. Наибольший общий делитель двух многочленов определен однозначно, с точностью до умножения на ненулевую константу.

◄ Если $f(x) = g(x) = 0$, то $(f(x), g(x)) = 0$, и доказывать нечего. Пусть $d_1(x)$ и $d_2(x)$ — два наибольших общих делителя многочленов $f(x)$ и $g(x)$. По определению, $d_1(x) \,\big|\, d_2(x)$ и $d_2(x) \,\big|\, d_1(x)$. Иначе говоря, $d_1(x) = q_1(x)d_2(x)$ и $d_2(x) = q_2(x)d_1(x)$, значит $d_1(x)d_2(x) = q_1(x)q_2(x)d_1(x)d_2(x)$, или $(q_1(x)q_2(x) - 1)d_1(x)d_2(x) = 0$, значит, по следствию 15.5, $q_1(x)q_2(x) = 1$. Следовательно, $\deg q_1(x) = \deg q_2(x) = 0$, т.е. $q_1(x)$ и $q_2(x)$ — ненулевые константы.►

Теорема 15.12. Для любых многочленов $f(x), g(x) \in F[x]$ существуют многочлены $u(x)$ и $v(x)$, для которых

$$(f(x), g(x)) = u(x)f(x) + v(x)g(x). \tag{15.4}$$

◄ Воспользуемся алгоритмом Евклида: построим последовательно многочлены $u_0(x), u_1(x), \ldots$ и $v_0(x), v_1(x), \ldots$, для которых выполнены равенства $r_i(x) = u_i(x)f(x) + v_i(x)g(x)$. Действительно, начнем с

$$u_0(x) = 1, v_0(x) = 0, u_1(x) = 0, v_1(x) = 1.$$

Далее воспользуемся индукцией: если $u_{i-2}(x), v_{i-2}(x), u_{i-1}(x), v_{i-1}(x)$ уже определены, то из равенства

$$\begin{aligned} r_i(x) &= r_{i-2} - q_{i-1}(x)r_{i-1}(x) \\ &= u_{i-2}(x)f(x) + v_{i-2}(x)g(x) - q_{i-1}(x)[u_{i-1}(x)f(x) + v_{i-1}(x)g(x)] \\ &= [u_{i-2}(x) - q_{i-1}(x)u_{i-1}(x)]f(x) + [v_{i-2}(x) - q_{i-1}(x)v_{i-1}(x)]g(x) \end{aligned}$$

видно, что можно взять $u_i = u_{i-2}(x) - q_{i-1}(x)u_{i-1}(x)$, $v_i(x) = v_{i-2}(x) - q_{i-1}(x)v_{i-1}(x)$. При $i = k$ получаем требуемое представление $r_k(x) = (f(x), g(x))$.►

Замечание. Для целых чисел также определена операция деления с остатком. Следовательно, наибольший общий делитель двух целых чисел можно найти с помощью алгоритма Евклида. Таким же способом можно доказать вариант теоремы 15.4 для целых чисел:

Теорема 15.13. Для любых двух целых чисел a, b существуют целые числа u и v, для которых

$$(a, b) = ua + vb.$$

Лекция 16. Неприводимые многочлены. Факториальность кольца многочленов и кольца целых чисел. Многочлен как функция. Схема Горнера. Корни многочлена, кратность корня. Понижение кратности корня при дифференцировании, избавление от кратных корней.

В этой лекции F — произвольное фиксированное поле, R — произвольное ассоциативное коммутативное кольцо с единицей.

1°. Простые элементы кольца. Неприводимые многочлены. Факториальные кольца. Факториальность кольца многочленов и кольца целых чисел.

Определение 16.1. Необратимый элемент $p \in R$ называется *простым*, если из равенства $p = ab$, где $a, b \in R$, следует, что один из элементов a, b обратим.

В кольце целых чисел простые элементы — это простые числа, а в кольце многочленов — неприводимые многочлены.

Определение 16.2. Многочлен $p(x) \in F[x]$ положительной степени n называется *неприводимым*, если он не разлагается в произведение двух многочленов степени меньшей, чем n.

Пример: любой многочлен степени 1 неприводим.

Определение 16.3. Многочлены $f(x)$ и $g(x)$ называются взаимно простыми, если $(f(x), g(x)) = 1$.

Определение 16.4. Два элемента a, b кольца R называются *ассоциированными* (обозначается $a \sim b$), если существует обратимый элемент $\lambda \in R$, такой, что $a = \lambda b$.

Определение 16.5. Кольцо R называется *факториальным*, если любой ненулевой необратимый элемент кольца R разлагается в произведение простых элементов, причем это разложение единственно, с точностью до перестановки сомножителей и их ассоциированности: если имеется два разложения

$$a = p_1 p_2 \dots p_r = q_1 q_2 \dots q_s,$$

где $p_1, \dots, p_r, q_1, \dots q_s$ — простые элементы, то $r = s$ и существует подстановка $\sigma \in S_r$ для которой $q_i \sim p_{\sigma(i)}$, $i = 1, 2, \dots, r$.

Отметим простейшие свойства неприводимых многочленов.

Предложение 16.6.
1) Если $p(x)$ — неприводимый многочлен, $f(x)$ — произвольный многочлен, то либо $p(x) \,\vdots\, f(x)$, либо $p(x)$ и $f(x)$ — взаимно простые.
2) Если $p(x)$ — неприводимый многочлен, и $p(x)$ делит произведение многочленов $f(x)$ и $g(x)$, то $p(x) \,\vdots\, f(x)$ или $p(x) \,\vdots\, g(x)$.

◄ Проверим 1). Если $d(x) = (p(x), f(x))$, то $d(x) \,\vdots\, p(x)$. Если $\deg d(x) = 0$, то $d(x)$ — ненулевая констата, и тогда 1 — тоже наибольший общий делитель многочленов $f(x)$ и $p(x)$. Если же $\deg d(x) = \deg p(x)$, то $p(x) = \lambda d(x)$, $\lambda \in F$, значит $p(x) \,\vdots\, f(x)$, т.к. $d(x) \,\vdots\, f(x)$.

Докажем 2). Если $p(x) \not\vdots f(x)$, то в силу 1) и теоремы 15.12 существуют многочлены $u(x)$ и $v(x)$, такие, что $u(x)p(x) + v(x)f(x) = 1$. Но тогда

$$g(x) = 1 \cdot g(x) = [u(x)p(x) + v(x)f(x)]g(x) = \underbrace{g(x)u(x)p(x)}_{p(x)\,\vdots} + \underbrace{v(x)(f(x)g(x))}_{p(x)\,\vdots}.$$

►

Теорема 16.7. Кольцо $F[x]$ факториально.

◄ Сначала докажем существование разложения ненулевого необратимого многочлена $f(x)$ в произведение неприводимых. Проведем индукцию по степени многочлена $f(x)$.

Если $n = \deg f(x) = 1$, то сам многочлен $f(x)$ неприводим.

Пусть теперь $n = \deg f(x) > 1$. Если $f(x)$ неприводим, то для этого многочлена утверждение теоремы выполнено. В противном случае $f(x) = g(x)h(x)$ для некоторых многочленов $g(x)$ и $h(x)$, степень каждого из которых меньше n. По предположению индукции

$$g(x) = p_1(x)p_2(x)\dots p_k(x),$$

$$h(x) = p_{k+1}(x)p_{k+2}(x)\dots p_{k+l}(x),$$

где $p_i(x)$ — неприводимые многочлены, $i = 1, 2, \dots, k+l$. Тогда

$$f(x) = p_1(x)p_2(x)\dots p_k(x)p_{k+1}(x)p_{k+2}(x)\dots p_{k+l}(x).$$

Теперь докажем единственность разложения, с точностью до перестановки сомножителей и их ассоциированности. Опять применим индукцию по степени многочлена $f(x)$. Если $n = \deg f(x) = 1$, то многочлен $f(x)$ неприводим, и из равенства $f(x) = q_1(x)\dots q_s(x)$ следует, что $s = 1$ и $q_1(x) = f(x)$. Допустим, что единственность разложения доказана для всех многочленов степени, меньшей чем $n = \deg f(x)$. Пусть

$$f(x) = p_1(x)p_2(x)\dots p_r(x) = q_1(x)q_2(x)\dots q_s(x).$$

Тогда в силу 16.6 можно считать, при необходимости переставляя сомножители, что $p_1(x) \mid q_1(x)$. Но тогда $q_1(x) = \lambda_1 p_1(x)$ для некоторого $\lambda_1 \in F$, т.е. $p_1(x) \sim q_1(x)$. Далее применяем предположение индукции к многочлену

$$g(x) = p_2(x)\dots p_r(x) = \lambda_1 q_2(x)\dots q_s(x),$$

степень которого меньше n и получаем, что $r - 1 = s - 1$, и после некоторой перестановки сомножителей $p_i(x) \sim q_i(x)$ при $i = 2, \dots r$. Вместе с ассоциированностью $p_1 \sim q_1$ это дает требуемое утверждение. ►

Аналогично доказывается факториальность кольца целых чисел.

Разложение на неприводимые множители, построенное в теореме 16.7, называется *каноническим разложением.*

Упражнение. Докажите, что над любым полем существует бесконечное множество попарно неассоциированных неприводимых многочленов.

2°. Многочлен как функция. Схема Горнера. Корень многочлена. Теорема Безу.

Определение 16.8. Пусть $f(x) \in F[x]$, $f(x) = a_n x^n + \dots + a_0$. *Значением* многочлена $f(x)$ при $x = a \in F$ называют элемент поля $f(a) = a_n a^n + \dots + a_0$. Таким образом, каждый многочлен задает отображение $F \to F : a \mapsto f(a)$.

Заметим, однако, что разные многочлены могут определять одну и ту же функцию.

Пример. Пусть $F = \mathrm{GF}(2)$ — поле из двух элементов $0, 1$. Тогда многочлены x и x^2 определяют одно и то же (тождественное) отображение F в F.

В то же время, как мы увидим в дальнейшем, если поле F бесконечно, то многочлен однозначно определен функцией, которую он задаёт.

Значение многочлена тесно связано с делением многочлена на двучлен.

Теорема 16.9. Пусть $f(x) \in F[x]$, $a \in F$. Тогда $f(a)$ — остаток от деления $f(x)$ на двучлен $x - a$.

◄ Запишем $f(x) = q(x)(x-a) + r$, где $r = 0$ либо $\deg r < \deg(x-a) = 1$, т.е. $\deg r = 0$. В обоих случаях $r \in F$. Подставляя a вместо x, получаем $f(a) = q(a)(a-a) + r = r$.►

Для одновременного вычисления остатка и частного от деления многочлена $f(x)$ на $x - a$ удобно использовать *схему Горнера*: введем обозначения $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$, $q(x) = b_{n-1}x^{n-1} + \ldots + b_0$. Тогда

$$q(x)(x-a) = b_{n-1}x^n + (b_{n-2} - ab_{n-1})x^{n-1} + \ldots + (b_0 - ab_1)x - b_0 a,$$

поэтому

$$\begin{aligned} b_{n-1} &= a_n, \\ b_{n-2} &= a_{n-1} + ab_{n-1} \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ b_{n-i-1} &= a_{n-i} + ab_{n-i} \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ b_0 &= a_1 + ab_1 \\ r &= a_0 + ab_0. \end{aligned}$$

Удобно представить вычисления в виде таблицы

a_n	a_{n-1}	$\ldots$	a_i	$\ldots$	a_1	a_0
$b_{n-1} = a_n$	$b_{n-2} = a_{n-1} + ab_{n-1}$	$\ldots$	$b_{i-1} = a_i + ab_i$	$\ldots$	$b_0 = a_1 + ab_1$	$r = a_0 + ab_0$

Определение 16.10. Корнем многочлена $f(x) \in F[x]$ называется элемент a поля F или его расширения, для которого $f(a) = 0$.

Теорема 16.11. (**теорема Безу.**) Элемент $a \in F$ является корнем многочлена $f(x)$ тогда и только тогда, когда $(x-a) \,\big|\, f(x)$.

◄ Условие $(x-a) \,\big|\, f(x)$ равносильно тому, что остаток r от деления $f(x)$ на двучлен $x - a$ равен нулю. Но в силу 16.9 $r = f(a)$.►

3°. Кратность неприводимого множителя (корня) многочлена. Формальная производная многочлена и её свойства. Понижение кратности неприводимого множителя (корня) при дифференцировании. "Освобождение" от кратных неприводимых множителей (корней).

Определение 16.12. Пусть $p(x)$ — неприводимый многочлен, $f(x)$ — некоторый ненулевой многочлен, и $p(x) \,\big|\, f(x)$. Наибольшее число k, для которого $p(x)^k \,\big|\, f(x)$, называется *кратностью* множителя $p(x)$ многочлена $f(x)$. Если $a \in F$ — корень многочлена $f(x)$, то кратность неприводимого множителя $x - a$ называется *кратностью* корня a.

Определение 16.13. *Формальной производной* многочлена

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in F[x]$$

называется многочлен

$$f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \ldots + a_1 \in F[x] \tag{16.1}$$

(вспомним, что $ka = \underbrace{a + \ldots + a}_{\text{k раз}}$ для любого $a \in F$ и натурального числа k).

Предложение 16.14. Для любых многочленов $f(x), g(x) \in F[x]$ и константы $\alpha \in F$ выполнены равенства:

$$\begin{array}{rcl} (f(x)+g(x))' & = & f'(x)+g'(x), \\ (\alpha f(x))' & = & \alpha f'(x), \\ (f(x)g(x))' & = & f'(x)g(x)+f(x)g'(x), \\ (f(x)^k)' & = & k(f(x))^{k-1}f'(x) \text{ для любого целого } k>0. \end{array} \tag{16.2}$$

◀ Первые два равенства из (16.2) очевидны. Проверим третье равенство для одночленов:

$$(x^m x^n)' = (x^{m+n})' = (m+n)x^{m+n-1} = mx^{m-1}x^n + nx^{n-1}x^m = (x^m)'x^n + x^m(x^n)',$$

как и утверждается. Для произвольных многочленов третье равенство из (16.2) следует из первых двух, дистрибутивности умножения и доказанного свойства производной произведения одночленов. Наконец, четвертое равенство выводится из третьего индукцией по k. Действительно, при $k=1$ оно очевидно: $f'(x)=f'(x)$. При $k>1$ имеем

$$(f(x)^k)' = (f(x)^{k-1}f(x))' = (f(x)^{k-1})'f(x) + f(x)^{k-1}f'(x) = \\ (k-1)f(x)^{k-2}f'(x)f(x) + f(x)^{k-1}f'(x) = kf(x)^{k-1}f'(x).$$

►

Теорема 16.15. Пусть $\operatorname{char} F = 0$. Если $p(x)$ — неприводимый множитель многочлена $f(x)$ кратности $k>1$, то $p(x)$ — неприводимый множитель производной $f'(x)$ кратности $k-1$. В частности, если $a \in F$ — корень многочлена $f(x)$ кратности $k>1$, то a — корень производной $f'(x)$ кратности $k-1$.

◀ По определению кратности неприводимого множителя, $f(x) = p(x)^k g(x)$, где $p(x) \nmid g(x)$. Значит,

$$f'(x) = ((p(x)^k)'g(x) + p(x)^k g'(x) = kp(x)^{k-1}p'(x)g(x) + p(x)^k g'(x) = p(x)^{k-1}[kp'(x)g(x) + p(x)g'(x)].$$

Итак, $p(x)^{k-1} \mid f'(x)$, но $p(x) \nmid p'(x)$, так как $\deg p'(x) < \deg p(x)$, и $p(x) \nmid g(x)$, по условию, следовательно, согласно предложению 16.6, $p(x) \nmid kp'(x)g(x)$. Поэтому $p(x) \nmid kp'(x)g(x) + p(x)g'(x)$, и $p(x)^k \nmid f'(x)$.

Утверждение о кратности корня получается, если положить $p(x) = x - a$. ►

Теорема 16.16. (Метод "избавления" от кратных множителей). Пусть $\operatorname{char} F = 0$, и $f(x) \in F[x]$ — многочлен положительной степени. Тогда многочлен

$$g(x) = \frac{f(x)}{(f(x), f'(x))}$$

имеет те же неприводимые множители, что и $f(x)$ (и, соответственно, те же корни), но кратность каждого неприводимого множителя (корня) равна 1.

◀ Пусть $d(x) = (f(x), f'(x))$. Поскольку $d(x) \mid f(x)$, все неприводимые множители многочлена $d(x)$ являются также неприводимыми множителями многочлена $f(x)$. Пусть $p(x)$ — неприводимый множитель многочлена $f(x)$ кратности $k \geqslant 1$, т.е. $f(x) = p(x)^k f_1(x)$, $p(x) \nmid f_1(x)$. По теореме 16.15, $p(x)^{k-1} \mid f'(x)$, значит, $p(x)^{k-1} \mid d(x)$. Но $p(x)^k \nmid f'(x)$, значит $p(x)^k \nmid d(x)$. Имеем $d(x) = p(x)^{k-1}h(x)$, $p(x) \nmid h(x)$. Поэтому

$$g(x) = p(x)\frac{f_1(x)}{h(x)},$$

откуда

$$g(x)h(x) = p(x)f_1(x).$$

Следовательно, по предложению 16.6, $p(x) \mid g(x)$, но $p(x)^2 \nmid g(x)$, иначе, сокращая на $p(x)$, мы получили бы $p(x) \mid f_1(x)$, что противоречит определению кратности неприводимого множителя. ►

Лекция 17. Алгебраическая замкнутость поля комплексных чисел. Неприводимые многочлены над полями комплексных и действительных чисел.

1°. Алгебраически замкнутые поля.

Определение 17.1. Поле F называется *алгебраически замкнутым*, если любой многочлен положительной степени над F имеет хотя бы один корень в F.

Теорема 17.2. Поле комплексных чисел алгебраически замкнуто.

Лемма 17.3. Пусть $f(x) \in \mathbb{C}[x]$ и $\deg f(x) > 0$. Если $\{z_n\}$ — последовательность комплексных чисел и $|z_n| \to \infty$ при $n \to \infty$, то $|f(z_n)| \to \infty$ при $n \to \infty$.

◄ Запишем $f(x) = a_k x^k + a_{k-1}x^{k-1} + \ldots + a_1 x + a_0$, где $k > 0$ и $a_k \neq 0$. Пусть

$$A = \max\left\{\frac{|a_{k-1}|}{|a_k|}, \ldots, \frac{|a_0|}{|a_k|}\right\}.$$

Согласно неравенству треугольника,

$$|a_k z_n^k| = |f(z_n) - (a_{k-1}z_n^{k-1} + \ldots + a_1 z_n + a_0)| \leqslant |f(z_n)| + |a_{k-1}z_n^{k-1} + \ldots + a_1 z_n + a_0|,$$

значит,

$$|f(z_n)| \geqslant |a_k z_n^k| - |a_{k-1}z_n^{k-1} + \ldots + a_1 z_n + a_0| \geqslant |a_k||z_n|^k - A|a_k|(|z_n|^{k-1} + \ldots + 1).$$

По условию, при достаточно больших n выполнено неравенство $|z_n| > 1$, значит, можно оценить

$$|f(z_n)| \geqslant |a_k||z_n|^k\left[1 - A\left(\frac{1}{|z_n|} + \frac{1}{|z_n|^2} + \ldots + \frac{1}{|z_n|^k}\right)\right] \geqslant |a_k||z_n|^k\left(1 - \underbrace{\frac{kA}{|z_n|}}_{\to 0}\right) \to +\infty \text{ при } n \to \infty.$$

►

Лемма 17.4. (Лемма Д'Аламбера). Пусть $f(x) \in \mathbb{C}[x]$ и $\deg f(x) > 0$. Если $u = f(z_0) \neq 0$ для некоторого числа $z_0 \in \mathbb{C}$, то существует число $z \in \mathbb{C}$, для которого $|f(z)| < |u|$.

◄ Пусть $g(x) = f(x + z_0)$. Тогда $u = g(0)$. Запишем $g(x) = a_0 + a_k x^k + \ldots + a_n x^n$, где $a_k \neq 0$, $a_n \neq 0$ и $a_0 = u$. Обозначим

$$A = \max\{|a_{k+1}|, \ldots, |a_n|\} \quad (\text{считая } A = 0 \text{ при } k = n).$$

Выберем число $r > 0$ так, чтобы выполнялись неравенства $r < 1$, $|a_k|r^k < |u|$, $(n-k)Ar < \frac{1}{2}|a_k|$. Пусть $\varphi = \arg u - \arg a_k$. Положим $\psi = \dfrac{\varphi + \pi}{k}$ и $z = r(\cos\psi + i\sin\psi)$. Тогда

$$|a_{k+1}z^{k+1} + \ldots + a_n z^n| \leqslant |a_{k+1}|r^{k+1} + \ldots + |a_n|r^n \leqslant Ar^{k+1}(n-k) < \frac{1}{2}|a_k|r^k,$$

и

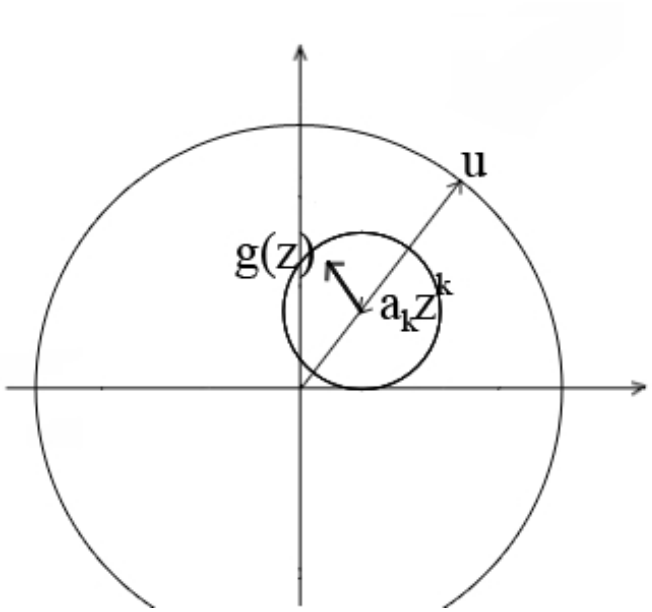
$$a_k z^k = a_k r^k(\cos(\varphi + \pi) + i\sin(\varphi + \pi)) = |a_k|r^k(\cos(\arg u + \pi) + i\sin(\arg u + \pi)).$$

Таким образом, векторы, соответствующие числам $a_k z^k$ и u, лежат на одной прямой и направлены противоположно, причем $|u| > |a_k z^k| = |a_k| r^k$. Следовательно, $u + a_k z^k = |u| - |a_k| r^k$, и

$$|g(z)| \leqslant |u + a_k z^k| + |a_{k+1} z^{k+1} + \ldots + a_n z^n|$$
$$= |u| - |a_k| r^k + |a_{k+1} z^{k+1} + \ldots + a_n z^n| < |u| - |a_k| r^k + \frac{1}{2}|a_k| r^k < |u|.$$

Осталось заметить, что $f(z + z_0) = g(z)$ и $|f(z + z_0)| < |f(z_0)|$. ►

Рассуждение из леммы Д'Аламбера иллюстрирует рисунок:



◄ **Доказательство теоремы 17.2.** Рассмотрим многочлен $f(x) \in \mathbb{C}$ положительной степени. Пусть $u = \inf\{|f(z)| : z \in \mathbb{C}\}$. По определению точной нижней грани, существует последовательность $\{z_n = x_n + iy_n\}$, для которой

$$\lim_{n\to\infty} |f(z_n)| = a.$$

В силу леммы 17.3, последовательность $\{|z_n|\}$ является ограниченной. Но тогда последовательность $\{x_n\}$ ограничена, следовательно, можно выбрать из нее сходящуюся подпоследовательность. Таким образом, будем считать, что

$$\lim_{n\to\infty} x_n = x_0.$$

Аналогично, можно считать, что

$$\lim_{n\to\infty} y_n = y_0.$$

Так же можно показать, что последовательности $\mathrm{Re}(f(z_n))$ и $\mathrm{Im}(f(z_n))$ ограничены, и, следовательно, их также можно считать сходящимися. Положим

$$u = \lim_{n\to\infty} \mathrm{Re}(f(z_n)) + i \lim_{n\to\infty} \mathrm{Im}(f(z_n)).$$

Поскольку значение многочлена $f(z_n)$ получается с помощью операций умножения и сложения из x_n и y_n, ясно, что $u = f(x_0 + iy_0)$ и $|u| = a$. Но лемма Д'Аламбера показывает, что при $u \neq 0$ число $|u|$ не является нижней гранью для множества модулей значений многочлена на $\mathbb{C}$. Значит, $u = 0$ и число $z_0 = x_0 + iy_0$ — искомый корень многочлена.►

Другое доказательство теоремы 17.2 использует следующее утверждение, точная формулировка и доказательство которого содержатся в курсе комплексного анализа.

Принцип максимума. *Если функция комплексного переменного $F(z)$ дифференцируема в каждой внутренней точке некоторой ограниченной области D и непрерывна на ее границе, то максимум её модуля* $\max\{|F(z)| : z \in D\}$ *достигается на границе области D.*

Действительно, если предположить, что моногочлен $f(x)$ положительной степени не имеет корней в $\mathbb{C}$, то для любого круга $|z| \leqslant R$ получаем, что максимум модуля функции $F(z) = \dfrac{1}{f(z)}$ достигается на границе этого круга. Однако, по лемме 17.3, $\max\{|F(z)| : |z| = R\} \to 0$ при $R \to \infty$. Противоречие.

Следствие 17.5. Неприводимыми многочленами над $\mathbb{C}$ являются линейные многочлены (т.е. многочлены степени 1) и только они.

◄ Если $p(x)$ — неприводимый многочлен, и z — его корень, то, по теореме Безу, $(x-z) \,\big|\, p(x)$, значит, $\deg p(x) = 1$. Обратное утверждение очевидно: любой многочлен степени 1 неприводим.►

Следствие 17.6. Любой многочлен положительной степени над $\mathbb{C}$ разлагается в произведение линейных множителей.

◄ Следует из предыдущего утверждения и тоеремы 16.7.►

Для описания неприводимых многочленов над $\mathbb{R}$ понадобится

Лемма 17.7. Если $f(x) \in \mathbb{R}[x]$ и $z \in \mathbb{C}$ — корень многочлена $f(x)$, то и $\overline{z}$ — корень многочлена $f(x)$.

◄ Если $f(x) = a_n x^n + \ldots + a_0$, где $a_i \in \mathbb{R}$, $i = 0, 1, \ldots, n$, то

$$f(\overline{z}) = a_n \overline{z}^n + \ldots + a_0 = \overline{a_n}\,\overline{z^n} + \ldots + \overline{a_0} = \overline{a_n z^n + \ldots + a_0} = \overline{f(z)} = \overline{0} = 0.$$

►

Следствие 17.8. Неприводимыми многочленами над $\mathbb{R}$ являются линейные многочлены и квадратичные многочлены с отрицательным дискриминантом, и только они.

◄ Линейные многочлены неприводимы над любым полем, в частности, над $\mathbb{R}$. Если квадратичный многочлен не является неприводимым, то он разлагается в произведение многочленов меньшей степени. Значит, эти множители имеют степень 1, и у многочлена имеется действительный корень. У квадратичного многочлена с отрицательным дискриминантом нет действительных корней, следовательно, такой многочлен неприводим.

Обратно, пусть $p(x) \in \mathbb{R}$ — неприводимый многочлен. Если он имеет действительный корень, то $\deg p(x) = 1$. В противном случае у него имеется комплексный корень z. По лемме 17.7, $\overline{z}$ также является корнем многочлена $p(x)$. При этом многочлены $x - z$ и $x - \overline{z}$ взаимно просты, так как $z \neq \overline{z}$, поэтому $(x - z)(x - \overline{z}) \,\big|\, p(x)$. Но $h(x) = (x - z)(x - \overline{z}) = x^2 - (z + \overline{z})x + z\overline{z} \in \mathbb{R}[x]$, значит, $q(x) = p(x)/h(x) \in \mathbb{R}[x]$ и из неприводимости многочлена $p(x)$ следует, что $p(x) \sim h(x)$, в частности, что $\deg p(x) = 2$. ►

Лекция 18. Интерполяционный многочлен, формула Лагранжа и метод Ньютона для его построения. Поле рациональных дробей. Простейшие дроби. Разложение правильной дроби в сумму простейших дробей, случай вещественного и комплексного полей.

1°. Задача интерполяции. В этой лекции F — произвольное поле.

Определение 18.1. *Задача (полиномиальной) интерполяции* ставится следующим образом. Для заданных различных элементов $x_0, x_1, \ldots, x_n \in F$ и элементов $y_0, y_1, \ldots, y_n \in F$ найти такой многочлен $f(x) \in F[x]$ наименьшей степени, что

$$f(x_0) = y_0,\ f(x_1) = y_1, \ldots, f(x_n) = y_n. \tag{18.1}$$

Такой многочлен называется *интерполяционным многочленом*. Элементы $x_0, \ldots, x_n$ называются *узлами интерполяции*.

Теорема 18.2. Для любых различных элементов $x_0, x_1, \ldots, x_n \in F$ и любых элементов $y_0, y_1, \ldots, y_n \in F$ существует единственный интерполяционный многочлен $f(x)$ степени не больше n (возможно, $f(x) = 0$), для которого выполнены равенства (18.1). Этот многочлен может быть задан *формулой Лагранжа*:

$$\begin{aligned} f(x) = y_0 \frac{(x-x_1)\ldots(x-x_n)}{(x_0-x_1)\ldots(x_0-x_n)} + \ldots \\ + y_i \frac{(x-x_0)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x-x_n)}{(x_i-x_0)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)} + \ldots \\ + y_n \frac{(x-x_0)\ldots(x-x_{n-1})}{(x_n-x_1)\ldots(x_n-x_{n-1})}. \end{aligned} \tag{18.2}$$

◀ Сначала заметим, что многочлен (18.2) является интерполяционным многочленом для поставленной задачи интерполяции. Действительно, пусть

$$\varphi_i(x) = \frac{(x-x_0)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x-x_n)}{(x_i-x_0)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)}, \quad i = 0, 1, \ldots, n.$$

Легко видеть, что $\deg \varphi_i(x) = n$, и что $\varphi_i(x_i) = 1$, $\varphi_i(x_j) = 0$ при $j \neq i$. Следовательно,

$$\deg f(x) \leqslant \max\{\deg(y_0\varphi_0(x)), \ldots, \deg(y_n\varphi_n(x))\} \leqslant n.$$

При этом

$$f(x_i) = y_0 \underbrace{\varphi_0(x_i)}_{=0} + \ldots + y_i \underbrace{\varphi_i(x_i)}_{=1} + \ldots + y_n \underbrace{\varphi_n(x_i)}_{=0} = y_i.$$

Далее, проверим единственность интерполяционного многочлена. Для этого запишем его в виде $f(x) = a_n x^n + \ldots + a_0$ и заметим, что в силу (18.1) его коэффициенты удовлетворяют системе линейных уравнений

$$\begin{aligned} a_n x_0^n + a_{n-1} x_0^{n-1} + \ldots + a_1 x_0 + a_0 &= y_0 \\ a_n x_1^n + a_{n-1} x_1^{n-1} + \ldots + a_1 x_1 + a_0 &= y_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_n x_n^n + a_{n-1} x_n^{n-1} + \ldots + a_1 x_n + a_0 &= y_n. \end{aligned}$$

Определитель этой системы с точностью до знака совпадает с определителем Вандермонда $\prod_{0 \leqslant i < j \leqslant n} (x_j - x_i) \neq 0$, так как по условию элементы $x_0, x_1, \ldots, x_n$ попарно различны. Следовательно, по теореме Крамера, ее решение единственно.▶

Недостатком формулы Лагранжа является то, что для нахождения коэффициентов интерполяционного многочлена по этой формуле необходимо знать все узлы $x_0, \ldots, x_n$ и все значения $y_0, \ldots, y_n$. Иногда удобнее строить многочлен, последовательно добавляя новые слагаемые, учитывающие новые узлы и значения в них.

Соответствующая формула называется *формулой Ньютона*:

$$f(x) = u_0 + u_1(x - x_0) + u_2(x - x_0)(x - x_1) + \ldots + u_n(x - x_0)\ldots(x - x_{n-1}).$$

Коэффициенты $u_0, u_1, \ldots, u_{n-1}$ определяются последовательно. Очевидно, что $u_0 = y_0$. Если $f_{n-1}(x)$ — интерполяционный многочлен для узлов $x_0, \ldots, x_{n-1}$, то из условия

$$y_n = f(x_n) = f_{n-1}(x_n) + u_n(x_n - x_0)\ldots(x_n - x_{n-1})$$

получаем

$$u_n = \frac{y_n - f_{n-1}(x_n)}{(x_n - x_0)\ldots(x_n - x_{n-1})}.$$

2°. Поле дробей коммутативного кольца без делителей нуля. В этом разделе R — коммутативное кольцо без делителей нуля с единицей 1. Рассмотрим множество пар

$$\hat{Q} = \{(a, b) : a, b \in R, b \neq 0\}.$$

Введем отношение эквивалентности на множестве $\hat{Q}$

$$(a_1, b_1) \sim (a_2, b_2) \Leftrightarrow a_1 b_2 = a_2 b_1 \tag{18.3}$$

Проверка того, что это отношение рефлексивно и симметрично, очевидна. Транзитивность отношения (18.3) можно установить так: если

$$(a_1, b_1) \sim (a_2, b_2) \text{ и } (a_2, b_2) \sim (a_3, b_3),$$

то

$$a_1 b_2 = a_2 b_1 \text{ и } a_2 b_3 = a_3 b_2,$$

откуда

$$a_1 b_2 b_3 = a_2 b_1 b_3 \Rightarrow a_1 b_3 b_2 = a_3 b_1 b_2 \Rightarrow (a_1 b_3 - a_3 b_1) b_2 = 0.$$

Поскольку $b_2 \neq 0$, и в R нет делителей нуля, получаем $a_1 b_3 - a_3 b_1 = 0$, т.е. $a_1 b_3 = a_3 b_1$ и $(a_1, b_1) \sim (a_3, b_3)$.

Определение 18.3. Класс эквивалентных пар называется *дробью*, класс пар, эквивалентных паре $(a, b) \in \hat{Q}$, обозначается $\dfrac{a}{b}$.

Из определения видно, что $\dfrac{a}{b} = \dfrac{ac}{bc}$ для любого ненулевого элемента $c \in R$. Определим операции над дробями:

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} = \frac{a_1 b_2 + a_2 b_1}{b_1 b_2}, \quad \frac{a_1}{b_1} \cdot \frac{a_2}{b_2} = \frac{a_1 a_2}{b_1 b_2}. \tag{18.4}$$

Предложение 18.4. Операции (18.4) определены корректно.

◄ Пусть $(a_1, b_1) \sim (a_1', b_1')$, $(a_2, b_2) \sim (a_2', b_2')$. Проверим, что $(a_1 b_2 + a_2 b_1, b_1 b_2) \sim (a_1' b_2' + a_2' b_1', b_1' b_2')$. Вычислим

$$(a_1 b_2 + a_2 b_1) b_1' b_2' = a_1 b_1' b_2 b_2' + a_2 b_2' b_1 b_1' = a_1' b_1 b_2 b_2' + a_2' b_2 b_1 b_1' = (a_1' b_2' + a_2' b_1') b_1 b_2.$$

Аналогично,

$$a_1 a_2 b_1' b_2' = a_1' b_1 a_2' b_2 = a_1' a_2' b_1 b_2,$$

т.е. $(a_1 a_2, b_1 b_2) \sim (a_1' a_2', b_1' b_2')$. ►

Предложение 18.5. Множество дробей является полем относительно операций (18.4).

◀ Коммутативность сложения и умножения, а также ассоциативность умножения — очевидны. Легко проверяется, что элемент $\frac{0}{1}$ является нейтральным относительно сложения, а элемент $\frac{1}{1}$ — нейтральным относительно умножения. Ясно, что $\frac{a}{b} + \frac{-a}{b} = \frac{0}{b} = \frac{0}{1}$, т.е для любой дроби существует противоположная. Проверим ассоциативность сложения:

$$\left(\frac{a_1}{b_1} + \frac{a_2}{b_2}\right) + \frac{a_3}{b_3} = \frac{a_1b_2 + a_2b_1}{b_1b_2} + \frac{a_3}{b_3} = \frac{(a_1b_2 + a_2b_1)b_3 + a_3b_1b_2}{b_1b_2b_3} =$$
$$\frac{a_1b_2b_3 + (a_2b_3 + a_3b_2)b_1}{b_1b_2b_3} = \frac{a_1}{b_1} + \frac{a_2b_3 + a_3b_2}{b_2b_3} = \frac{a_1}{b_1} + \left(\frac{a_2}{b_2} + \frac{a_3}{b_3}\right).$$

Дистрибутивность проверяется аналогично. Наконец, заметим, что если $\frac{a}{b} \neq 0$, то $a \neq 0$ и $\frac{a}{b} \cdot \frac{b}{a} = \frac{1}{1}$, значит, ненулевые дроби обратимы.▶

Определение 18.6. Поле $Q = Q(R)$ дробей вида $\frac{a}{b}, a, b \in R, b \neq 0$ с указанными операциями называется полем частных, или полем отношений кольца R. При этом множество дробей вида $\frac{a}{1}$ является подкольцом в Q, изоморфным кольцу R, соответственно, дробь $\frac{a}{1}$ отождествляется с элементом $a \in R$.

Пример: $\mathbb{Q} = Q(\mathbb{Z})$.
Упражнение: какое получится поле частных, $Q(R)$, если само кольцо R — поле?

3°. Поле рациональных дробей. В этом разделе F — произвольное поле. Применяя конструкцию поля частных к кольцу многочленов $F[x]$ над полем F, получим

Определение 18.7. Поле частных $Q(F[x])$ называется *полем рациональных дробей* и обозначается $F(x)$.

Определение 18.8. Ненулевая рациональная дробь $\frac{f(x)}{g(x)} \in F(x)$ называется *правильной*, если $\deg f(x) < \deg g(x)$.

Предложение 18.9. (выделение целой части.) Любая рациональная дробь может быть представлена в виде суммы

$$\frac{f(x)}{g(x)} = q(x) + \frac{r(x)}{g(x)},$$

где $q(x), r(x)$ — многочлены, причем либо $r(x) = 0$, либо $\frac{r(x)}{g(x)}$ — правильная рациональная дробь. Указанное представление единственно.

◀ Заметим, что данное представление равносильно равенству

$$f(x) = q(x)g(x) + r(x),$$

где либо $r(x) = 0$, либо $\deg r(x) = 0$. Значит, предложение следует из существования и единственности деления с остатком.▶

Определение 18.10. Ненулевая рациональная дробь $\frac{r(x)}{p(x)^k}$ называется *простейшей*, если $p(x)$ — неприводимый многочлен, и $\deg r(x) < \deg p(x)$.

Предложение 18.11. Сумма и произведение правильных рациональных дробей — правильная рациональная дробь.

◄ Пусть $\dfrac{f_1(x)}{g_1(x)}$ и $\dfrac{f_2(x)}{g_2(x)}$ — правильные рациональные дроби. Тогда $\dfrac{f_1(x)}{g_1(x)} \cdot \dfrac{f_2(x)}{g_2(x)} = \dfrac{f_1(x)f_2(x)}{g_1(x)g_2(x)}$, и

$$\deg(f_1(x)f_2(x)) = \deg f_1(x) + \deg f_2(x) < \deg g_1(x) + \deg g_2(x) = \deg(g_1(x)g_2(x)),$$

значит, для произведения утверждение верно. Аналогично,

$$\frac{f_1(x)}{g_1(x)} + \frac{f_2(x)}{g_2(x)} = \frac{f_1(x)g_2(x) + f_2(x)g_1(x)}{g_1(x)g_2(x)},$$

и

$$\deg(f_1(x)g_2(x) + f_2(x)g_1(x)) \leqslant \max\{\deg(f_1(x)g_2(x)), \deg(f_2(x)g_1(x))\} =$$
$$\max\{\deg f_1(x) + \deg g_2(x), \deg f_2(x) + \deg g_1(x)\} < \deg g_1(x) + \deg g_2(x) = \deg(g_1(x)g_2(x)),$$

значит, утверждение верно и для суммы. ►

Теорема 18.12. Любая правильная рациональная дробь может быть представлена в виде суммы простейших дробей:

$$\frac{f(x)}{g(x)} = \frac{r_{11}(x)}{p_1(x)} + \ldots + \frac{r_{1k_1}(x)}{p_1(x)^{k_1}} + \frac{r_{21}(x)}{p_2(x)} + \ldots + \frac{r_{2k_2}(x)}{p_2(x)^{k_2}} + \cdots + \frac{r_{s1}(x)}{p_s(x)} + \ldots + \frac{r_{sk_s}(x)}{p_s(x)^{k_s}}, \qquad (18.5)$$

где $p_1(x), \ldots, p_s(x)$ — попарно неассоциированные неприводимые многочлены. Представление (18.5) единственно, с точностью до перестановки слагаемых.

◄ Существование. Можно считать дробь $\dfrac{f(x)}{g(x)}$ несократимой. Пусть $p(x)$ — неприводимый многочлен, входящий в каноническое разложение многочлена $g(x)$ с кратностью $k > 0$. Запишем $g(x) = p(x)^k g_1(x)$, где $p(x) \nmid g_1(x)$. Из неприводимости $p(x)$ следует, что $p(x)^k$ и $g_1(x)$ взаимно просты. Следовательно, существуют такие многочлены $u(x)$ и $v(x)$, что

$$u(x)p(x)^k + v(x)g_1(x) = 1.$$

Получаем:

$$\frac{f(x)}{g(x)} = \frac{u(x)p(x)^k f(x) + v(x)g_1(x)f(x)}{p(x)^k g_1(x)} = \frac{u(x)f(x)}{g_1(x)} + \frac{v(x)f(x)}{p(x)^k}.$$

Пусть $q(x)$ и $r(x)$ — частное и остаток от деления многочлена $v(x)f(x)$ на $p(x)$.Тогда

$$\frac{v(x)f(x)}{p(x)^k} = \frac{q(x)p(x) + r(x)}{p(x)^k} = \frac{q(x)}{p(x)^{k-1}} + \frac{r(x)}{p(x)^k}.$$

Заметим, что второе слагаемое здесь — простейшая дробь. Рассмотрим разность

$$\frac{f(x)}{g(x)} - \frac{r(x)}{p(x)^k} = \frac{q(x)}{p(x)^{k-1}} + \frac{u(x)f(x)}{g_1(x)} = \frac{q(x)g_1(x) + p(x)^{k-1}u(x)f(x)}{p(x)^{k-1}g_1(x)}.$$

В левой части этого равенства — правильная дробь, согласно предложению 18.11, значит, и правая часть — правильная дробь. Применяем к ней то же преобразование многократно, пока не получится дробь, знаменатель которой равен $g_1(x)$. Тогда возьмем следующий неприводимый множитель многочлена $g(x)$ и т.д.

Единственность. Достаточно показать, что если левая часть (18.5) равна 0, то все слагаемые в правой части равны 0. Действительно, предположим противное, и выберем ненулевое слагаемое $\dfrac{r(x)}{p_1(x)^{k_1}}$ с наибольшим показателем k_1. Приводя соотношение (18.5) к общему знаменателю, получим $r(x)p_2(x)^{k_2}\dots p_s(x)_s^k + p_1(x)h(x) = 0$ для некоторого многочлена $h(x)$. Это невозможно, так как $p_1(x) \not| \; r(x)$ и $p_1(x) \not| \; p_i(x)$ при $i > 1$. ►

Следствие 18.13. Правильная рациональная дробь над полем $\mathbb{C}$ может быть представлена в виде

$$\frac{f(x)}{g(x)} = \frac{A_{11}}{x-a_1} + \ldots + \frac{A_{1k_1}}{(x-a_1)^{k_1}} + \frac{A_{21}}{x-a_2} + \ldots + \frac{A_{2k_2}}{(x-a_2)^{k_2}} + \cdots + \frac{A_{s1}}{x-a_s} + \ldots + \frac{A_{sk_s}}{(x-a_s)^{k_s}}, \quad (18.6)$$

где A_{ij} — некоторые константы, $a_1, \ldots, a_s$ — попарно различные комплексные корни многочлена $g(x)$, имеющие кратности $k_1, \ldots, k_s$, соответственно.

Следствие 18.14. Правильная рациональная дробь над полем $\mathbb{R}$ может быть представлена в виде

$$\frac{f(x)}{g(x)} = \frac{A_{11}}{x-a_1} + \ldots + \frac{A_{1k_1}}{(x-a_1)^{k_1}} + \ldots + \frac{A_{s1}}{x-a_s} + \ldots + \frac{A_{sk_s}}{(x-a_s)^{k_s}} + \\ \frac{B_{11}x + C_{11}}{x^2 + p_1x + q_1} + \ldots + \frac{B_{1m_1}x + C_{1m_1}}{(x^2 + p_1x + q_1)^{m_1}} + \ldots + \frac{B_{t1}x + C_{t1}}{x^2 + p_tx + q_t} + \ldots + \frac{B_{tm_t}x + C_{tm_t}}{(x^2 + p_tx + q_t)^{m_t}}, \quad (18.7)$$

где A_{ij}, B_{ij}, C_{ij} — некоторые константы, $a_1, \ldots, a_s$ — попарно различные действительные корни многочлена $g(x)$, имеющие кратности $k_1, \ldots, k_s$, соответственно, $x^2 + p_1x + q_1, \ldots, x^2 + p_tx + q_t$ — различные квадратные неприводимые делители многочлена $g(x)$, имеющие кратности $m_1, \ldots, m_t$.

Лекция 19. Границы корней многочлена. Теорема Декарта.

В этом разделе мы рассматриваем только многочлены с вещественными коэффициентами.

1°. Границы корней многочлена.

Предложение 19.1. Пусть $f(x) = x^n + a_1x^{n-1} + \ldots + a_n \in \mathbb{R}[x]$, c — положительный вещественный корень многочлена $f(x)$. Тогда $c \leqslant 1 + \sqrt[m]{B}$, где m — наименьшее число, для которого $a_m < 0$, B — максимальное значение абсолютных величин отрицательных коэффициентов многочлена $f(x)$ (заметим, что если все коэффициенты многочлена $f(x)$ неотрицательны, то он не имеет положительных вещественных корней).

◄ Пусть $c > 0$ — корень многочлена $f(x)$. Если $c \leqslant 1$, то доказывать нечего, поэтому пусть далее $c > 1$. Заметим, что $a_i \geqslant -B$ при $i = 1, \ldots n$ (для отрицательных коэффициентов по определению B, а для неотрицательных это неравенство очевидно). Тогда

$$0 = f(c) = c^n + \underbrace{a_1c^{n-1} \ldots a_{m-1}c^{n-m+1}}_{\geqslant 0} + a_mc^{n-m} + \ldots + a_n \geqslant$$

$$c^n - B(c^{n-m} + \ldots + 1) = c^n - B\frac{c^{n-m+1} - 1}{c-1} =$$

$$c^n - B\frac{c^{n-m+1}}{c-1} + \underbrace{\frac{B}{c-1}}_{>0} > \frac{c^{n-m+1}}{c-1}[c^{m-1}(c-1) - B].$$

Таким образом,

$$\underbrace{c^{m-1}}_{\geqslant (c-1)^{m-1}}(c-1) - B < 0 \Rightarrow (c-1)^m < B \Rightarrow c - 1 < \sqrt[m]{B}.$$

►

2°. Теорема Декарта. Теорема Декарта позволяет, не производя никаких вычислений, оценить число положительных корней многочлена.

Определение 19.2. Пусть $a_0, a_1, \ldots, a_n$ — произвольная последовательность вещественных чисел. Говорят, что на k-том месте в этой последовательности имеется *перемена знака*, если $a_k \neq 0$ и знак предшествующего ненулевого числа в последовательности противоположен знаку числа a_k (более формально: если существует число r, $0 \leqslant r < k$, такое, что $a_ra_k < 0$ и $a_{r+1} = \ldots = a_{k-1} = 0$).

Пример: в последовательности

$$1, 0, 0, \underline{-1}, 0, -2, \underline{4}, 5, \underline{-1}, 0, \underline{1}, 3, 0, \underline{-4}$$

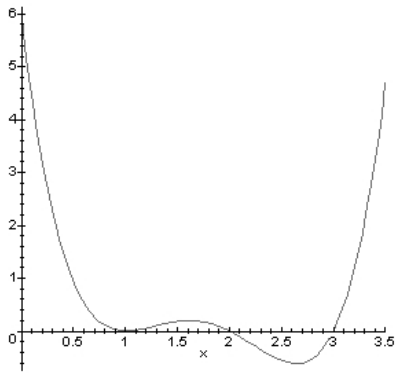
подчеркнуты те члены, на местах которых имеется перемена знака.

Теорема 19.3. (Теорема Декарта) Пусть $f(x) = a_0x^n + \ldots + a_n \in \mathbb{R}[x]$ — многочлен степени n. Тогда число положительных вещественных корней многочлена $f(x)$ не превосходит числа перемен знака в последовательности его коэффициентов и сравнимо с числом перемен знака по модулю 2, причем если все комплексные корни многочлена $f(x)$ вещественны, то эти два числа совпадают **(каждый корень считается столько раз, какова его кратность)**.

Обозначим через $L(f)$ число перемен знака в последовательности коэффициентов многочлена $f(x)$, а через $N(f)$ — число его положительных корней (с учетом кратностей). Без ограничения общности можно считать, что $a_0 > 0$ и $a_n \neq 0$ (потому что в противном случае можно умножить $f(x)$ на -1 и разделить на x, не меняя ни $L(f)$, ни $N(f)$).

Лемма 1. $N(f) \equiv L(f) \pmod 2$.

◄ $f(0) = a_n$, $f(x) > 0$ при всех достаточно больших x. Заметим, что при прохождениии (слева направо) корня четой кратности знак функции $f(x)$ не меняется, а при прохождении корня нечетной кратности меняется на противоположный. Значит, если $a_n < 0$, то знак должен измениться, т.е. $N(f)$ — нечетное число. В противном случае число $N(f)$ четно. Но то же верно и для $L(f)$: если знак в последовательности коэффициентов менялся нечетное число раз, то знак a_n противоположен знаку a_0, т.е. $a_n < 0$, и наоборот, при четном $L(f)$ имеем $a_n > 0$. Итак, оба числа $L(f)$ и $N(f)$ четные при $a_n > 0$ и нечетные при $a_n < 0$.►



На рисунке изображен график функции, заданной многочленом $f(x) = x^4-7x^3+17x^2-17x+6 = (x-1)^2(x-2)(x-3)$. Легко видеть, что $L(f) = N(f) = 4$.

Лемма 2. $N(f) \leqslant N(f') + 1$.

◄ По теореме Ролля, между любыми двумя корнями многочлена $f(x)$ лежит корень его производной. Каждый корень кратности $k > 1$ является корнем производной $f'(x)$ кратности $k-1$. Таким образом, если $c_1, \ldots, c_s$ — различные положительные корни многочлена $f(x)$ и их кратности равны $k_1, \ldots, k_s$, соответственно, то $N(f) = k_1 + \ldots + k_s$, а $N(f') \geqslant (s-1) + (k_1 - 1) + \ldots + (k_s - 1) = N(f) - 1$.►

Положим $\tilde{f}(x) = (-1)^n f(-x)$.

Лемма 3. $L(\tilde{f}) + L(f) \leqslant n = \deg f(x)$.

◄ Предположим сначала, что все коэффициенты многочлена $f(x)$ отличны от нуля. Сравним последовательности коэффициентов многочленов $f(x)$ и $\tilde{f}(x)$:

$$\begin{array}{lccccccc} f(x): & a_0 & a_1 & \ldots & a_k & \ldots & a_n \\ \tilde{f}(x): & a_0 & -a_1 & \ldots & (-1)^k a_k & \ldots & (-1)^n a_n \end{array}.$$

Видно, что в тех местах, где первая последовательность имеет перемену знака, вторая последовательность не имеет перемены знака, и наоборот (кроме, разумеется, места a_0). Значит, в этом случае $L(f) + L(\tilde{f}) = n$. Если же какие-то коэффициенты многочлена $f(x)$ равны нулю, то, заменяя их любыми ненулевыми числами, мы можем только увеличить число перемен знака в обеих последовательностях. ►

◄ **Доказательство теоремы Декарта.** Докажем неравенство $N(f) \leqslant L(f)$ индукцией по $n = \deg f(x)$. При $n = 0$ имеем $L(f) = N(f) = 0$, т.е. утверждение верно. Для $n > 0$ заметим, что $N(f') \leqslant L(f')$ по предположению индукции. Согласно лемме 2,

$$N(f) \leqslant N(f') + 1 \leqslant L(f') + 1 \leqslant L(f) + 1.$$

Но ввиду леммы 1, равенство $N(f) = L(f) + 1$ невозможно. Значит, $N(f) \leqslant L(f)$.

Последнее утверждение теоремы следует из леммы 3. Действительно, если все n корней многочлена $f(x)$ вещественные, и $f(0) \neq 0$, то $N(f) + N(\tilde{f}) = n$. Значит,

$$N(f) = n - N(\tilde{f}) \geqslant n - \underbrace{L(\tilde{f})}_{\leqslant n - L(f)} \geqslant n - (n - L(f)) = L(f) \Rightarrow N(f) = L(f).$$

►

Лекция 20. Метод Штурма отделения вещественных корней многочлена.

В этой лекции мы преполагаем, что многочлен $f(x) \in \mathbb{R}[x]$ не имеет кратных (комплексных) корней, т.е. к нему уже применили процедуру освобождения от кратных корней.

1°. Понятие системы Штурма. Теорема Штурма.

Определение 20.1. Конечная система многочленов $f_0(x) = f(x), f_1(x), \ldots, f_s(x)$ называется *системой Штурма* для многочлена $f(x)$ на отрезке $[a, b]$, если выполнены следующие условия:
(i) многочлен $f_s(x)$ не имеет корней на $[a, b]$;
(ii) $f_0(a)f_0(b) \neq 0$;
(iii) если $f_k(c) = 0$, $1 \leqslant k < s$, $a \leqslant c \leqslant b$, то $f_{k-1}(c)f_{k+1}(c) < 0$.
(iv) если $f(c) = 0$ для $c \in [a, b]$, то произведение $f_0(x)f_1(x)$ меняет знак с минуса на плюс, когда x проходит, возрастая, через точку c (более формально, если существует такое число $\delta > 0$, что $f_0(x)f_1(x) < 0$ при $x \in (c - \delta, c)$ и $f_0(x)f_1(x) > 0$ при $x \in (c, c + \delta)$).

Определение 20.2. Пусть $f_0(x) = f(x), f_1(x), \ldots, f_s(x)$ — система Штурма для многочлена $f(x)$ на отрезке $[a, b]$. Для любого $c \in [a, b]$ обозначим через $W(c) = W_f(c)$ количество перемен знака в числовой последовательности $f_0(c), f_1(c), \ldots, f_s(c)$.

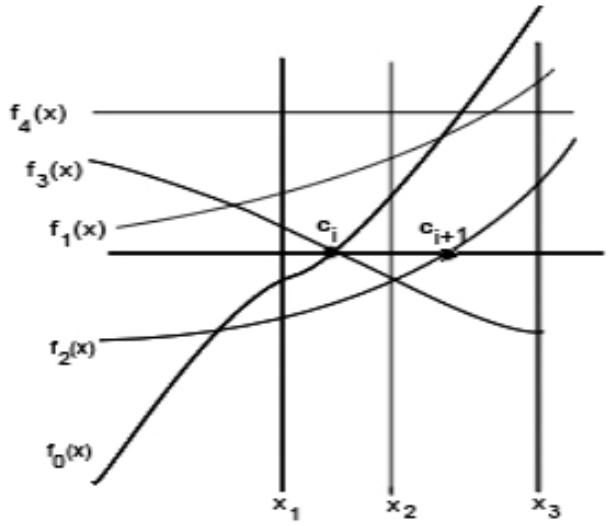
Теорема 20.3. (**Теорема Штурма**.) Пусть $f_0(x) = f(x), f_1(x), \ldots, f_s(x)$ — система Штурма для многочлена $f(x)$ на отрезке $[a, b]$. Тогда число вещественных корней многочлена $f(x)$ равно $W_f(a) - W_f(b)$.

◀ Рассмотрим разбиение $a = c_0 < c_1 < \ldots < c_{m-1} < c_m = b$ отрезка $[a, b]$ всеми различными корнями многочленов $f_0(x), f_1(x), \ldots, f_{s-1}(x)$, лежащими на отрезке $[a, b]$. Тогда на каждом интервале (c_{i-1}, c_i), $i = 1, \ldots, m$, ни один из многочленов системы Штурма не имеет корней, значит, на каждом таком интервале функция $W(c)$ постоянна. Предположим, что $f_r(c_i) \neq 0$, $r \geqslant 0$. Тогда число перемен знака в последовательности $f_r(c), \ldots, f_s(c)$ при $c \in (c_i, c_{i+1})$ (для $i < m$) и при $c \in (c_{i-1}, c_i)$ (для $i > 0$) равно числу перемен знака в последовательности $f_r(c_i), \ldots, f_s(c_i)$. Действительно, если $i < m$ и $f_k(c_i) \neq 0$ при всех $k = r + 1, \ldots, s$, то каждая из функций $f_r(c)$ сохраняет постоянный знак на полуинтервале $[c_i, c_{i+1})$, и в этом случае число перемен знака постоянно. В противном случае, если $f_k(c_i) = 0$, то $r < k < s$, в силу предположения и условия (i). В этом случае $f_{k-1}(c_i)$ и $f_{k+1}(c_i)$ имеют разные знаки в силу (iii), значит, на месте $k + 1$ имеется перемена знака в последовательности $f_0(c_i), \ldots, f_s(c_i)$. Теперь при $c \in (c_i, c_{i+1})$ независимо от знака числа $f_k(c)$ в тройке чисел $f_{k-1}(c), f_k(c), f_{k+1}(c)$ имеется также одна перемена знака. Опять получаем, что $W(c) = W(c_i)$. Аналогичное рассуждение применяется к полуинтервалу $(c_{i-1}, c_i]$ при $i > 0$.

В частности, если $r = 0$, т.е. $f(c_i) \neq 0$, то функция $W(c)$ постоянна на полуинтервале $[c_i, c_{i+1})$ (при $i < m$) и на полуинтервале $(c_{i-1}, c_i]$ (при $i > 0$).

Теперь пусть $f(c_i) = 0$. Значит, $0 < i < m$ в силу (ii). Тогда на месте $f_1(c)$ в последовательности $f_0(c), f_1(c), \ldots$, нет перемены знака при $c \in (c_i, c_{i+1})$, а при $c \in (c_{i-1}, c_i)$ перемена знака имеется. Число перемен знака в остальных местах последовательности значений сохраняется, поскольку из условия (ii) следует, что $f_1(c_i) \neq 0$. Итак, при переходе через каждый корень многочлена $f(x)$ значение функции $W(c)$ уменьшается на единицу, значит, если число корней равно u, то $W(b) = W(a) - u$. ▶

Поведение функций $f_0(x), \ldots, f_s(x)$ в окрестности корня $f(x)$ и в окрестности точки c_i, в которой $f(_i) \neq 0$, иллюстрируется следующим рисунком (при $s = 4$).



Из графиков видно, что

$$f_0(x_1) < 0, f_1(x_1) > 0, f_2(x_1) < 0, f_3(x_1) > 0, f_4(x_1) > 0,$$

значит, $W(x_1) = 3$;

$$f_0(x_2) > 0, f_1(x_2) > 0, f_2(x_2) < 0, f_3(x_2) < 0, f_4(x_2) > 0,$$

значит, $W(x_2) = 2$;

$$f_0(x_3) > 0, f_1(x_3) > 0, f_2(x_3) > 0, f_3(x_3) < 0, f_4(x_3) > 0,$$

значит, $W(x_3) = 2$.

2°. Построение стандартной системы Штурма. Пусть многочлен $f(x) \in \mathbb{R}[x]$ не имеет кратных (комплексных) корней и $f(a) \neq 0, f(b) \neq 0$ для некоторых $a, b \in \mathbb{R}$, $a < b$. Положим

$$\begin{array}{rcll} f_0(x) & = & f(x); & \\ f_1(x) & = & f'(x) & \\ f_0(x) & = & q_1(x)f_1(x) - f_2(x), & \deg f_2(x) < \deg f_1(x); \\ \ldots\ldots & & \ldots\ldots & \ldots\ldots \\ f_{k-1}(x) & = & q_k(x)f_k(x) - f_{k+1}(x), & \deg f_{k+1}(x) < \deg f_k(x); \\ \ldots\ldots & & \ldots\ldots & \ldots\ldots \\ f_{s-1}(x) & = & q_s(x)f_s(x). & \end{array} \tag{20.1}$$

Теорема 20.4. При сделанных выше предположениях, система (20.1), является системой Штурма для многочлена $f(x)$ на отрезке $[a, b]$.

◄ Поскольку $f(x)$ не имеет кратных корней, $(f, f') = 1$ и $f_s(x)$ — ненулевая константа, значит, условие 20.1(i) выполнено. Условие 20.1(ii) выполнено по предположению. Если $f_k(c) = 0$, то $f_{k-1}(c) = -f_{k+1}(c)$. Если бы при этом выполнялось равенство $f_{k-1}(c) = 0$, то мы получили бы из них последовательно $f_{k+1}(c) = f_{k+2}(c) = \ldots = f_s(c) = 0$, что невозможно. Значит, выполнено условие 20.1(iii). Наконец, проверим условие 20.1(iv). Если $f(c) = 0$, то $f'(c) \neq 0$, иначе число c было бы кратным корнем многочлена $f(x)$. Если $f'(c) > 0$, то и в некоторой окрестности точки c выполнено неравенство $f'(x) > 0$, функция $f(x)$ монотонно возрастает. Значит, в этой окрестности $f(x) < 0$ при $x < c$ и $f(x) > 0$ при $x > c$, поэтому $f(x)f'(x) > 0$ при $x \neq c$. Аналогично, при $f'(x) < 0$ в некоторой окрестности точки c функция $f(x)$ монотонно убывает, значит, в этой окрестности $f(x) > 0$ при $x < c$ и $f(x) < 0$ при $x > c$, поэтому $f(x)f'(x) > 0$ при $x \neq c$.►

Система (20.1) называется *стандартной системой Штурма*. Ясно, что если вместо системы Штурма $f_0(x), \ldots, f_s(x)$ взять систему $\lambda_0 f_0(x), \ldots, \lambda_s f_s(x)$, где $\lambda_k > 0$, $k = 0, \ldots, s$, то получится снова система Штурма. Это замечание позволяет в некоторых случаях упростить вычисления.

Для нахождения общего числа вещественных корней многочлена $f(x)$ заметим, что для достаточно большого положительного числа M все корни многочлена $f(x)$ лежат внутри интервала $(-M, M)$. Кроме того, можно считать, что для любого $k = 0, 1, \ldots, s$ знак числа $f_k(M)$ совпадает со знаком старшего коэффициента многочлена $f_k(x)$. В то же время знак числа $f_k(-M)$ совпадает со знаком старшего коэффициента многочлена $f_k(x)$, если $\deg f_k(x)$ — четное число, и противоположен знаку старшего коэффициента многочлена $f_k(x)$, если $\deg f_k(x)$ — нечетное число. Поэтому, даже не зная числа M, легко найти $W(M)$ и $W(-M)$.

Пример: $f_0(x) = f(x) = x^3 + 3x + 1$, $f_1(x) = \frac{1}{3}f'(x) = x^2 + 1$, деля $f(x)$ на $x^2 + 1$, получаем остаток $2x + 1$ и берем $f_2(x) = -2x - 1$. Делим $x^2 + 1$ на $-2x - 1$, получаем остаток $\frac{5}{4}$, берем $f_3(x) = -1$. Составляем таблицы знаков:

	$f_0(x) = x^3 + 3x + 1$	$f_1(x) = x^2 + 1$	$f_2(x) = -2x - 1$	$f_3(x) = -1$	W
$x = M$	$+$	$+$	$-$	$-$	1
$x = -M$	$-$	$+$	$+$	$-$	2

Итак, число вещественных корней данного многочлена равно $W(-M) - W(M) = 2 - 1 = 1$, т.е. данный многочлен имеет единственный вещественный корень.

Лекция 21. Кольцо многочленов от нескольких переменных. Лексикографический порядок на одночленах. Старший член произведения многочленов. Симметрические многочлены, их выражение через элементарные симметрические многочлены, формулы Виета.

1°. Определение и простейшие свойства кольца многочленов от нескольких переменных.

Определение 21.1. Кольцо многочленов от n переменных над полем F определяется по индукции. Кольцо $F[x_1]$ — это обычное кольцо многочленов от одной переменной x_1. Если кольцо многочленов $F[x_1, \ldots, x_{n-1}$ уже определено, положим

$$F[x_1, \ldots, x_n] = F[x_1, \ldots, x_{n-1}][x_n].$$

Теорема 21.2. Кольцо $F[x_1, \ldots, x_n]$ — ассоциативное коммутативное кольцо с единицей без делителей нуля.

◀ Доказательство аналогично доказательству предложения 15.2 и следствия 15.5.▶
Несколько более сложно доказывается

Теорема 21.3. Кольцо $F[x_1, \ldots, x_n]$ факториально.

Доказательство этого утверждения в наш курс не входит.

Итак, по определению, элементами кольца$F[x_1, \ldots, x_n]$ являются суммы вида

$$f(x_1, \ldots, x_n) = f_k(x_1, \ldots, x_{n-1})x_n^k + \ldots + f_0(x_1, \ldots, x_{n-1}).$$

В свою очередь, коэффициенты $f_k, \ldots, f_0$ разлагаются по степеням переменной x_{n-1}, и т.д. Таким образом, любой многочлен из $F[x_1, \ldots, x_n]$ единственным образом представляется в виде суммы ненулевых *одночленов* вида

$$\lambda_{i_1,\ldots,i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n},$$

где $0 \neq \lambda_{i_1,\ldots,i_n} \in F$ и $i_1, i_2, \ldots, i_n \geqslant 0$, для которых соответвующие векторы $(i_1, \ldots, i_n)$ различны. Мы будем также говорить, что данные одночлены $\lambda_{i_1,\ldots,i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$ входят в запись многочлена $f(x_1, \ldots, x_n)$. *Степенью* одночлена $\lambda_{i_1,\ldots,i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$ называется сумма показателей $i_1 + i_2 + \ldots i_n$. Многочлен $f(x_1, \ldots, x_n)$ называется *однородным*, если степени всех ненулевых одночленов, входящих в его запись, одинаковы (пример: $x_1^{100} + 2x_1^{34}x_2^{66} + x_2^{100}$).

2°. Лексикографический порядок на одночленах. Старший член многочлена.

Определение 21.4. Пусть $a = \alpha x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$ и $b = \beta x_1^{j_1} x_2^{j_2} \ldots x_n^{j_n}$ — два одночлена (с неулевыми коэффициентами). Будем говорить, что одночлен a *старше* одночлена b, и записывать $a \succ b$, если существует такое число $r > 0$, что $i_k = j_k$ при $k < r$ и $i_r > j_r$. Естественно, что в этом случае мы также будем говорить, что одночлен b *младше* одночлена a и обозначать $b \prec a$.

Иначе говоря, порядок на одночленах (без учета коэффициентов) такой же, как порядок слов в словаре (старшие слова предшествуют младшим), если представить произведения степеней переменных в виде длинных "слов":

$$a = \underbrace{x_1 \ldots x_1}_{i_1 \text{ раз}} \underbrace{x_2 \ldots x_2}_{i_2 \text{ раз}} \ldots \underbrace{x_n \ldots x_n}_{i_n \text{ раз}},$$

$$b = \underbrace{x_1 \ldots x_1}_{j_1 \text{ раз}} \underbrace{x_2 \ldots x_2}_{j_2 \text{ раз}} \ldots \underbrace{x_n \ldots x_n}_{j_n \text{ раз}}.$$

Предложение 21.5. Для любых двух **различных** одночленов $a = x_1^{i_1}x_2^{i_2}\dots x_n^{i_n}$ и $b = x_1^{j_1}x_2^{j_2}\dots x_n^{j_n}$ с коэффициентами 1 выполняется одно из отношений $a \succ b$ или $b \succ a$. Если $c = x_1^{k_1}x_2^{k_2}\dots x_n^{k_n}$ и выполнены отношения $a \succ b$ и $b \succ c$, то $a \succ c$.

◀ Пусть $a \neq b$. Это означает, что $i_r \neq j_r$ для некоторых $r = 1, 2, \dots, n$. Можно выбрать наименьшее число r, обладающее указанным свойством. Тогда $i_k = j_k$ при всех $k < r$ и либо $i_r > j_r$, и тогда $a \succ b$, либо $i_r < j_r$, и тогда $b \succ a$.

Теперь пусть $a \succ b$ и r — наименьшее число, для которого $i_r \neq j_r$ (и тогда $i_r > j_r$). Аналогично, если $b \succ c$, то можно выбрать наименьшее число s, для которого $j_s \neq k_s$ (и тогда $j_s > k_s$). Положим $m = \min(r, s)$. Тогда при $t < m$ имеем $i_t = j_t = k_t$, в то время как $i_m \geqslant j_m \geqslant k_m$, причем хотя бы одно из этих неравенств — строгое. Значит, $i_m > k_m$, и $a \succ c$.▶

Для примера покажем расположение в порядке убывания старшинства одночленов степени 4 от трех переменных:

$$x_1^4 \succ x_1^3x_2 \succ x_1^3x_3 \succ x_1^2x_2^2 \succ x_1^2x_2x_3 \succ x_1^2x_3^2 \succ x_1x_2^3 \succ x_1x_2^2x_3 \succ x_1x_2x_3^2 \succ x_1x_3^3 \succ$$
$$x_2^4 \succ x_2^3x_3 \succ x_2^2x_3^2 \succ x_2x_3^3 \succ x_3^4.$$

Выделяя слагаемое, соответствующее такому одночлену, приходим к следующему определению.

Определение 21.6. Заметим, что среди одночленов, входящих в запись ненулевого многочлена $f(x_1, \dots, x_n)$, всегда найдется такой, который старше всех остальных одночленов, входящих в запись этого многочлена. Этот одночлен называется *старшим членом* данного многочлена. Мы будем обозначать его $\mathrm{LT}(f)$.

Лемма 21.7. (**О старшем члене произведения многочленов.**) Старший член произведения многочленов равен произведению их старших членов.

◀ Ясно, что достаточно доказать утверждение леммы для двух сомножителей, а общий случай тогда легко вывести по индукции. Пусть $u = \alpha x_1^{i_1}x_2^{i_2}\dots x_n^{i_n} = \mathrm{LT}(f)$, $v = \beta x_1^{j_1}x_2^{j_2}\dots x_n^{j_n} = \mathrm{LT}(g)$. Покажем, что

$$\mathrm{LT}(fg) = \alpha\beta x_1^{i_1+j_1}x_2^{i_2+j_2}\dots x_n^{i_n+j_n} = uv = \mathrm{LT}(f)\,\mathrm{LT}(g).$$

Действительно, любой одночлен, входящий в запись произведения fg, пропорционален произведению одночленов $u' = x_1^{r_1}\dots x_n^{r_n}$ и $v' = x_1^{s_1}\dots x_n^{s_n}$, входящих с некоторыми ненулевыми коэффициентами в запись многочленов f и g, соответственно. При этом $u'v' = x_1^{r_1+s_1}x_2^{r_2+s_2}\dots x_n^{r_n+s_n}$. Выберем наименьшее число m, для которого выполнено хотя бы одно из неравенств $i_m > r_m$ или $j_m > s_m$ (такое число существует, поскольку $u \succ u'$ или $v \succ v'$). Пусть, для определенности, $i_m < r_m$. Поскольку, по выбору m, $i_k = r_k$ и $j_k = s_k$ при всех $k < m$, выполнено неравенство $j_m \geqslant s_m$ (иначе было бы $v' \succ v$). Значит, $i_k + j_k = r_k + s_k$ при всех $k < m$ и $i_m + j_m > r_m + s_m$, т.е. $uv \succ u'v'$.▶

Определение 21.8. Многочлен $f(x_1, \dots, x_n)$ называется *симметрическим*, если

$$f(x_1, \dots, x_n) = f(x_{\sigma(1)}, \dots, x_{\sigma(n)})$$

для любой подстановки $\sigma \in S_n$.

Определение 21.9. Одночлен $\alpha x_1^{i_1}\dots x_n^{i_n}$ назовем *монотонным*, если

$$i_1 \geqslant i_2 \geqslant \dots \geqslant i_n.$$

Лемма 21.10. (О старшем члене симметрического многочлена.) Старший член симметрического многочлена является монотонным одночленом.

◄ Пусть $f(x_1,\dots,x_n)$ — симметрический многочлен, и $u=\alpha x_1^{i_1}\dots x_n^{i_n}=\mathrm{LT}(f)$. Предположим, что он не является монотонным, т.е. существует число r, для которого $i_r<i_{r+1}$. Рассмотрим транспозицию $\sigma=(r,r+1)$. Заметим, что многочлен $f(x_{\sigma(1)},\dots,x_{\sigma(n)})$ содержит одночлен

$$v=\alpha x_1^{i_1}\dots x_{r-1}^{i_{r-1}}x_{r+1}^{i_r}x_r^{i_{r+1}}x_{r+2}^{i_{r+2}}\dots x_n^{i_n}=\alpha x_1^{i_1}\dots x_{r-1}^{i_{r-1}}x_r^{i_{r+1}}x_{r+1}^{i_r}x_{r+2}^{i_{r+2}}\dots x_n^{i_n}.$$

Видно, что $v\succ u$. Но $f(x_{\sigma(1)},\dots,x_{\sigma(n)})=f(x_1,\dots,x_n)$, так как $f(x_1,\dots,x_n)$ — симметрический многочлен, значит, одночлен v входит в запись многочлена $f(x_1,\dots,x_n)$, что противоречит тому, что u — старший член этого многочлена.►

Определение 21.11. Многочлены от n переменных $x_1,\dots,x_n$

$$\begin{aligned}
\sigma_1 &= x_1+\dots+x_n\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
\sigma_k &= \sum_{1\leqslant i_1<i_2<\dots<i_k\leqslant n} x_{i_1}\dots x_{i_k}\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
\sigma_n &= x_1\dots x_n
\end{aligned}$$

называются *элементарными симметрическими многочленами* от n переменных.

3°. Основная теорема о симметрических многочленах.

Теорема 21.12. Любой симметрический многочлен от переменных $x_1,\dots,x_n$ единственным образом представляется в виде многочлена от $\sigma_1,\dots,\sigma_n$.

◄ Нужно доказать существование и единственность такого многочлена $F(Y_1,\dots,Y_n)$, что

$$f(x_1,\dots,x_n)=F(\sigma_1,\dots,\sigma_n).$$

Существование. Без ограничения общности можно считать симметрический многочлен $f(x_1,\dots,x_n)$ однородным многочленом некоторой степени m, поскольку, очевидно, любой многочлен есть сумма однородных многочленов различных степеней. Рассмотрим $u=\mathrm{LT}(f)=\alpha x_1^{i_1}\dots x_n^{i_n}$. По лемме 21.10, $i_1\geqslant i_2\geqslant\dots\geqslant i_n$. Поскольку $\mathrm{LT}(\sigma_k)=x_1x_2\dots x_k$ при $k=1,\dots,n$, из леммы 21.7 вытекает, что

$$v=\mathrm{LT}(\sigma_1^{k_1}\dots\sigma_n^{k_n})=x_1^{k_1}(x_1x_2)^{k_2}\dots(x_1x_2\dots x_n)^{k_n}=x_1^{k_1+k_2+\dots+k_n}x_2^{k_2+\dots+k_n}\dots x_n^{k_n}.$$

Выберем степени $k_1,\dots,k_n$ так, чтобы $u=\alpha v$, что равносильно системе уравнений

$$\begin{aligned}
k_1+k_2+\dots+k_n &= i_1\\
k_2+\dots+k_n &= i_2\\
&\dots\dots\dots\dots\dots\dots\\
k_n &= i_n,
\end{aligned}\tag{21.1}$$

решение которой можно записать так:

$$\begin{aligned}
k_1 &= i_1-i_2\\
k_2 &= i_2-i_3\\
&\dots\dots\dots\dots\dots\\
k_{n-1} &= i_{n-1}-i_n\\
k_n &= i_n.
\end{aligned}$$

Следовательно, многочлен $f_1(x_1,\dots,x_n) = f(x_1,\dots,x_n) - \alpha\sigma_1^{k_1}\dots\sigma_n^{k_n}$ является однородным степени m многочленом, старший член которого младше чем u. Применяя то же преобразование к многочлену $f_1(x_1,\dots,x_n)$ и т.д., получим последовательность многочленов, старшие члены которых убывают относительно лексикографического порядка. Поскольку число одночленов степени m с коэффициентом 1 конечно, этот процесс закончится, когда очередная разность окажется равной нулю, т.е. когда будет получено искомое представление.

Единственность. Предположим, что $f(x_1,\dots,x_n) = F_1(\sigma_1,\dots,\sigma_n) = F_2(\sigma_1,\dots,\sigma_n)$, где $F_1(Y_1,\dots,Y_n)$ и $F_2(Y_1,\dots,Y_n)$ — различные многочлены. Положим $G(Y_1,\dots,Y_n) = F_1(Y_1,\dots,Y_n) - F_2(Y_1,\dots,Y_n)$. Тогда $G(Y_1,\dots,Y_n) \neq 0$, но $G(\sigma_1,\dots,\sigma_n) = 0$. Заметим, что если $u = \alpha Y_1^{r_1}Y_2^{r_2}\dots Y_n^{r_n}$ и $v = \beta Y_1^{s_1}Y_2^{s_2}\dots Y_n^{s_n}$ — различные (не ассоциированные) ненулевые одночлены, входящие в $G(Y_1,\dots,Y_n)$, то старшие члены многочленов $u(\sigma_1,\dots,\sigma_n)$ и $v(\sigma_1,\dots,\sigma_n)$ не ассоциированы, так как система уравнений (21.1) при фиксированных правых частях имеет единственное решение. Значит, самый старший среди этих старших членов не может сократиться ни с каким одночленом, входящим в $G(\sigma_1,\dots,\sigma_n)$. Противоречие.▶

4°. Формулы Виета. Рассмотрим многочлен от переменных $x_1,\dots,x_n,Y$:

$$F(x_1,\dots,x_n,Y) = (Y-x_1)(Y-x_2)\dots(Y-x_n). \tag{21.2}$$

Непосредственно проверяется, что

$$F(x_1,\dots,x_n,Y) = Y^n - \sigma_1 Y^{n-1} + \dots + (-1)^k\sigma_k Y^{n-k} + \dots + (-1)^n\sigma_n.$$

Теорема 21.13. Пусть $f(x) = a_0x^n + a_1x^{n-1} + \dots + a_n$ — многочлен над полем F степени $n > 0$, имеющий n корней $\alpha_1,\dots,\alpha_n$ в F (с учетом кратностей). Тогда

$$\sigma_k(\alpha_1,\dots,\alpha_n) = (-1)^k\frac{a_k}{a_0}, \quad k = 1,\dots,n. \tag{21.3}$$

◀ По теореме Безу, $f(x) = a_0(x-\alpha_1)\dots(x-\alpha_n)$. Подставляя x вместо Y и $\alpha_1\dots,\alpha_n$ вместо $x_1,\dots,x_n$ в многочлен $F(x_1,\dots,x_n,Y)$ из (21.2), получаем

$$f(x) = a_0\left(x^n - \sigma_1(\alpha_1,\dots,\alpha_n)x^{n-1} + \dots + (-1)^k\sigma_k(\alpha_1,\dots,\alpha_n)x^{n-k} + \dots + (-1)^n\sigma_n(\alpha_1,\dots,\alpha_n)\right).$$

Сравнивая коэффициенты при x^{n-k}, получаем равенства $a_0(-1)^k\sigma_k(\alpha_1,\dots,\alpha_n) = a_k$ при $k = 1,\dots,n$, что равносильно (21.3).▶

Лекция 22. Результант двух многочленов, его выражение через корни многочленов.

1°. Задачи, решаемые с помощью результанта. Определение результанта. Основная задача, решаемая введением результанта — дать явный критерий того, что два многочлена над полем имеют общий множитель положительной степени. Конечно, в этом можно убедиться, находя их наибольший общий делитель с помощью алгоритма Евклида. Однако, явная формула, включающая только коэффициенты многочлена, иногда имеет преимущество — будет приведен пример, показывающий, что применение позволяет исключать некоторые неизвестные из систем полиномиальных уравнений.

В этой лекции F — произвольное поле. Некоторые утверждения будут доказаны для случая $F = \mathbb{C}$.

Определение 22.1. *Результантом* многочленов $f(x) = a_0x^n + \ldots + a_n$ и $g(x) = b_0x^m + \ldots + b_m$, где $m, n > 0$, называется определитель вида

$$R(f,g) = \left|\begin{array}{cccccccc} a_0 & a_1 & \ldots & a_n & 0 & 0 & \ldots & 0 \\ 0 & a_0 & a_1 & \ldots & a_n & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & a_0 & a_1 & \ldots & a_{n-1} & a_n \\ b_0 & b_1 & \ldots & b_{m-1} & b_m & 0 & \ldots & 0 \\ 0 & b_0 & b_1 & \ldots & b_{m-1} & b_m & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & b_0 & b_1 & \ldots & b_{m-1} & b_m \end{array}\right| \begin{array}{l} \left.\vphantom{\begin{array}{c}a\\a\\a\\a\end{array}}\right\} m \text{ строк} \\ \left.\vphantom{\begin{array}{c}a\\a\\a\\a\end{array}}\right\} n \text{ строк} \end{array} \tag{22.1}$$

Лемма 22.2. В условиях определения 22.1,

$$R(g,f) = (-1)^{mn}R(f,g)$$

◀ Чтобы поменять местами нижние n строк с верхними m строками в определителе (22.1), можно сначала $m+1$-ю строку переставить последовательно снизу вверх m раз, затем то же самое повторить с $m+2$-й строкой и т.д. Всего получится mn перестановок соседних строк, при этом преобразованный определитель окажется равным $R(g,f)$▶

2°. Основное свойство результанта.

Теорема 22.3. Если $f(x) = a_0x^n + \ldots + a_0$ и $g(x) = b_0x^m + \ldots + b_m$ — два многочлена, $m, n > 0$, то $R(f,g) = 0$ тогда и только тогда, когда либо $a_0 = b_0 = 0$, либо многочлены $f(x)$ и $g(x)$ имеют общий множитель положительной степени.

◀ Ясно, что если $a_0 = b_0 = 0$, то $R(f,g) = 0$, так как в определителе первый столбец окажется нулевым. Значит, можно считать, что $a_0 \neq 0$ (в силу леммы 22.2 многочлены f и g можно поменять местами). Если при этом $g(x) = 0$, то, конечно, $R(f,g) = 0$ и многочлены $f(x)$ и $g(x)$ имеют общий множитель $f(x)$. Поэтому можно считать, что $g(x) \neq 0$. Докажем вспомогательное утверждение: многочлены $f(x)$ и $g(x)$ имеют общий множитель положительной степени тогда и только тогда, когда существуют такие ненулевые многочлены $u(x)$ и $v(x)$, что и выполняются условия

$$\deg u(x) < m, \quad \deg v(x) < n \quad \text{и} \quad f(x)u(x) + g(x)v(x) = 0. \tag{22.2}$$

Действительно, пусть у $f(x)$ и $g(x)$ имеется множитель $d(x)$ положительной степени. Тогда $f(x) = f_1(x)d(x)$, $g(x) = g_1(x)d(x)$, и можно положить $u(x) = g_1(x)$, $v(x) = -f_1(x)$:

$$\deg u(x) = \deg g(x) - \deg d(x) < m, \quad \deg v(x) = \deg f(x) - \deg d(x) < n,$$
$$f(x)u(x) + v(x)g(x) = d(x)f_1(x)g_1(x) + d(x)g_1(x)(-f_1(x)) = 0.$$

Обратно, пусть выполнены условия (22.2). Тогда $f(x) \,\big|\, g(x)v(x)$. Если бы многочлены $f(x)$ и $g(x)$ были взаимно простыми, то из этого бы следовало, что $f(x)|v(x)$, что невозможно, так как $\deg v(x) < \deg f(x)$. Вспомогательное утверждение доказано.

Теперь запишем $u(x)$ и $v(x)$ как многочлены с неопределенными коэффициентами:

$$u(x) = u_0x^{m-1} + u_1x^{m-2} + \dots u_{m-1}$$
$$v(x) = v_0x^{n-1} + v_1x^{n-2} + \dots v_{n-1}.$$

Вычисляя коэффициенты левой части равенства (22.2) при $x^{m+n-1}, x^{m+n-2} \dots, x, 1$, получим равносильную этому равенству систему линейных уравнений относительно $u_0, \dots, u_{m-1}$ и $v_0, \dots, v_{n-1}$:

$$\begin{array}{rcrcll}
a_0u_0 & + & b_0v_0 & = & 0 & \text{при } x^{m+n-1} \\
a_1u_0 + a_0u_1 & + & b_1v_0 + b_0v_1 & = & 0 & \text{при } x^{m+n-2} \\
a_2u_0 + a_1u_1 + a_0u_2 & + & b_2v_0 + b_1v_1 + b_0v_2 & = & 0 & \text{при } x^{m+n-3} \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & & & & \\
a_nu_{m-2} + a_{n-1}u_{m-1} & + & b_mv_{n-2} + b_{m-1}v_{n-1} & = & 0 & \text{при } x \\
a_nu_{m-1} & + & b_mv_{n-1} & = & 0 & \text{при } 1
\end{array} \tag{22.3}$$

По теореме Крамера 6.6, эта система имеет ненулевое решение в том и только в том случае, когда ее определитель равен нулю. Но матрица этой системы — транспонированная к матрице из (22.1), эначит, ее определитель равен $R(f, g)$.►

Следствие 22.4. При $F = \mathbb{C}$, в условиях теоремы 22.3 равенство $R(f, g) = 0$ равносильно тому, что либо $a_0 = b_0 = 0$, либо многочлены $f(x)$ и $g(x)$ имеют общий корень.

◄ Если x_0 — общий корень $f(x)$ и $g(x)$, то они имеют общий множитель $(x - x_0)$. Обратно, если у них есть общий множитель $d(x)$ положительной степени, то он имеет корень в $\mathbb{C}$, и этот корень — общий для $f(x)$ и $g(x)$. ►

3°. Выражение результанта через корни многочленов. Рассмотрим специальные многочлены от переменных $x_1, \dots, x_n, y_1, \dots, y_m, Z$ вида

$$F = a_0(Z - x_1)(Z - x_2)\dots(Z - x_n),$$
$$G = b_0(Z - y_1)(Z - y_2)\dots(Z - y_m),$$

где $a_0, b_0 \neq 0$.

Можно записать

$$F(Z) = a_0Z^n + a_1Z^{n-1} + \dots + a_n,$$
$$G(Z) = b_0Z^m + b_1Z^{m-1} + \dots + b_m,$$

где коэффициенты $a_1, \dots, a_n$, $b_1, \dots, b_m$ являются симметрическими многочленами от $x_1, \dots, x_n$ и $y_1, \dots, y_m$ и определяются формулами Виета. Определитель (22.1), элементами которого являются эти многочлены, также является многочленом от $x_1, \dots, x_n$, $y_1, \dots, y_m$.

Теорема 22.5. Для результанта многочленов $F(Z)$ и $G(Z)$ выполнены равенства:

$$R(F, G) = a_0^m \prod_{i=1}^{n} G(x_i) = (-1)^{mn} b_0^n \prod_{j=1}^{m} F(y_j) = a_0^m b_0^n \prod_{i,j} (x_i - y_j) \tag{22.4}$$

◄ Добавим к набору переменных, которые мы используем, новую переменную Y и рассмотрим результант $H(Y) = R(F(Z), G(Z) - Y)$. Ясно, что в матрице из 22.1 элементы b_m заменятся на $b_m - Y$, следовательно, $H(Y) = (-1)^n a_0^m Y^n + \ldots + R(F, G)$. Многочлены $F(Z)$ и $G(Z) - G(x_i)$ имеют общий корень x_i, значит, $H(G(x_i)) = 0$ по теореме 22.3. Поскольку значения $G(x_i)$ — различные элементы кольца $F[x_1, \ldots, x_n, y_1, \ldots, y_m]$, получаем равенство $H(Y) = (-1)^n a_0^m \prod_{i=1}^n (Y - G(x_i)) = a_0^m \prod_{i=1}^n (G(x_i) - Y)$. Подставляя $Y = 0$, получаем первое из равенств (22.4). Второе равенство вытекает из первого и леммы 22.2. Наконец, третье равенство вытекает из первого, так как $G(x_i) = b_0(x_i - y_1)\ldots(x_i - y_m)$ при $i = 1, \ldots n$ (достаточно перемножить почленно n последних соотношений).►

Следствие 22.6. (Выражение результанта через корни многочленов.) Пусть $f(x)$ и $g(x)$ — многочлены над полем комплексных чисел, причем $\deg f(x) = n > 0$, $\deg g(x) = m > 0$, a_0 и b_0 — старшие коэффициенты этих многочленов, $\alpha_1, \ldots, \alpha_n$ и $\beta_1, \ldots, \beta_m$ — комплексные корни многочленов $f(x)$ и $g(x)$ (с учетом кратностей). Тогда

$$R(f, g) = a_0^m \prod_{i=1}^{n} g(\alpha_i) = (-1)^{mn} b_0^n \prod_{j=1}^{m} f(\beta_j) = a_0^m b_0^n \prod_{i,j} (\alpha_i - \beta_j) \tag{22.5}$$

◄ По условию, $f(x) = a_0(x - \alpha_1)\ldots(x - \alpha_n)$ и $g(x) = a_0(x - \beta_1)\ldots(x - \beta_m)$. Подставим в (22.4) α_i вместо x_i, $i = 1, \ldots, n$ и β_j вместо вместо y_j, $j = 1, \ldots, m$. Получим (22.5). ►

4°. Исключение неизвестных из системы алгебраических уравнений. Ограничимся простейшим случаем двух уравнений с двумя неизвестными

$$\begin{array}{rcl} f(x, y) & = & 0 \\ g(x, y) & = & 0. \end{array} \tag{22.6}$$

Запишем оба многочлена по степеням переменного x:

$$f(x, y) = a_0(y)x^n + a_1(y)x^{n-1} + \ldots + a_n(y)$$
$$f(x, y) = b_0(y)x^n + b_1(y)x^{m-1} + \ldots + b_m(y)$$

Тогда $R(f(x), g(x)) = h(y)$ — многочлен от одного переменного y. Если (α, β) — некоторое решение системы (22.6), то α — общий корень многочленов $f(x, \beta)$ и $g(x, \beta)$, значит, $R(f(x, \beta), g(x, \beta)) = 0$, т.е. β — корень многочлена $h(y)$. Обратно, если β — корень многочлена $h(y)$, то либо $f(x, \beta)$ и $g(x, \beta)$ имеют общий корень (или несколько общих корней), либо $a_0(\beta) = b_0(\beta) = 0$, и в этом случае β может оказаться "посторонним" корнем. Приведенный метод в принципе сводит решение системы 22.6 к решению нескольких алгебраических уравнений с одним неизвестным (это рассуждение легко обобщить на случай систем из бóльшего числа уравнений с бóльшим числом неизвестных, однако практическая применимость метода ограничена тем, что степени получающихся уравнений растут очень быстро.

Пример. Рассмотрим систему уравнений

$$x^2 + xy^2 + y = 0$$
$$yx^2 + y^2 + 1 = 0$$

Запишем результант:

$$\begin{vmatrix} 1 & y^2 & y & 0 \\ 0 & 1 & y^2 & y \\ y & 0 & y^2+1 & 0 \\ 0 & y & 0 & y^2+1 \end{vmatrix} = \begin{vmatrix} 1 & y^2 & y \\ 0 & y^2+1 & 0 \\ y & 0 & y^2+1 \end{vmatrix} + y \begin{vmatrix} y^2 & y & 0 \\ 1 & y^2 & y \\ y & 0 & y^2+1 \end{vmatrix} = y^2+1+y(y^6+y^4-y) = y^7+y^5+1$$

Лекция 23. Дискриминант многочлена, выражение дискриминанта через корни многочлена.

1°. Определение дискриминанта с помощью результанта. В этой лекции мы рассматриваем многочлены с комплексными коэффициентами, хотя все результаты верны для многочленов над любым алгебраически замкнутым полем характеристики 0.

Поскольку кратные корни многочлена $f(x)$ — это общие корни многочлена $f(x)$ и его производной $f'(x)$, естественно использовать результант для выяснения того, имеет ли многочлен кратные корни.

Определение 23.1. *Дискриминантом* многочлена $f(x) = a_0x^n + \ldots + a_n$ положительной степени n называется число

$$D(f) = (-1)^{\frac{n(n-1)}{2}} a_0^{-1} R(f, f').$$

Из основной теоремы о результанте следует

Теорема 23.2. Многочлен $f(x)$ имеет кратные корни тогда и только тогда, когда $D(f) = 0$.

Теперь используем формулу, выражающую результант через корни многочлена (см. (22.5)).

Теорема 23.3. Для любого многочлена $f(x) = a_0x^n + \ldots + a_n$ положительной степени n выполняется равенство

$$D(f) = a_0^{2n-2} \prod_{1 \leqslant i < j \leqslant n} (\alpha_j - \alpha_i)^2, \tag{23.1}$$

где $\alpha_1, \ldots, \alpha_n$ — корни многочлена $f(x)$ (с учетом кратностей).

◀ По теореме Безу, можно записать $f(x) = a_0(x - \alpha_i)f_i(x)$ для любого $i = 1, \ldots, n$, где $f_i(x) = \prod\limits_{j \neq i}(x - \alpha_j)$ — некоторый многочлен степени $n - 1$ со старшим коэффициентом 1. Значит,

$$f'(x) = a_0((x - \alpha_i)f_i(x))' = a_0(f_i(x) + (x - \alpha_j)f_i'(x),$$

откуда $f'(\alpha_i) = a_0 f_i(\alpha_i) = a_0 \prod\limits_{j \neq i}(\alpha_i - \alpha_j)$. В силу (22.5), учитывая, что $\deg f'(x) = n - 1$, имеем

$$R(f, f') = a_0^{n-1} \prod_{i=1}^{n} f'(\alpha_i) = a_0^{n-1} \prod_{i=1}^{n} \underbrace{a_0 \prod_{j \neq i}(\alpha_i - \alpha_j)}_{f_i(\alpha_i)} = a_0^{2n-1} \prod_{\substack{1 \leqslant i, j \leqslant n \\ i \neq j}} (\alpha_i - \alpha_j)$$

Теперь заметим, что в произведение $\prod\limits_{\substack{1 \leqslant i, j \leqslant n \\ i \neq j}} (\alpha_i - \alpha_j)$ каждая разность $(\alpha_i - \alpha_j)$, где $i \neq j$, входит два раза с противоположным знаком. Заменяя разности с $i > j$ на противоположные, получим

$$\prod_{\substack{1 \leqslant i, j \leqslant n \\ i \neq j}} (\alpha_i - \alpha_j) = (-1)^{\frac{n(n-1)}{2}} \prod_{1 \leqslant i < j \leqslant n} (\alpha_j - \alpha_i)^2,$$

поскольку количество таких разностей равно $\binom{n}{2} = \dfrac{n(n-1)}{2}$. Значит,

$$D(f) = (-1)^{\frac{n(n-1)}{2}} a_0^{-1} R(f, f') = a_0^{2n-2} \prod_{1 \leqslant i < j \leqslant n} (\alpha_j - \alpha_i)^2.$$

►
Заметим, что дискриминант $D(f)$ — симметрический многочлен от корней многочлена $f(x)$.

2°. Дискриминант многочленов степени 2 и 3. Рассмотрим многочлен степени 2: $f(x) = ax^2 + bx + c$, где $a \neq 0$. Тогда $f'(x) = 2ax + b$, и по определению

$$D(f) = (-1)^1 a^{-1} \begin{vmatrix} a & b & c \\ 2a & b & 0 \\ 0 & 2a & b \end{vmatrix} = -\frac{1}{a}(ab^2 - 2ab^2 + 4a^2c) = b^2 - 4ac.$$

Теперь найдем дискриминант многочлена степени 3. Ограничимся случаем $f(x) = x^3 + px + q$, к которому, как мы увидим позже, легко сводится решение любого кубического уравнения. Пусть $\alpha_1, \alpha_2, \alpha_3$ — комплексные корни многочлена $f(x)$. Сначала выразим многочлен $g(x_1, x_2, x_3) = (x_2 - x_1)^2(x_3 - x_1)^2(x_3 - x_2)^2$ через элементарные симметрические многочлены. Видно, что $\mathrm{LT}(g) = x_1^4 x_2^2$. Составим таблицу монотонных векторов степеней по убыванию относительно лексикографического порядка:

$$\begin{array}{ccc} 4 & 2 & 0 \\ 4 & 1 & 1 \\ 3 & 3 & 0 \\ 3 & 2 & 1 \\ 2 & 2 & 2 \end{array}$$

Получим тождество с неопределенными коэффициентами

$$g(x_1, x_2, x_3) = A\sigma_1^2\sigma_2^2 + B\sigma_1^3\sigma_3 + C\sigma_2^3 + D\sigma_1\sigma_2\sigma_3 + E\sigma_3^2.$$

Учитывая, что $\sigma_1(\alpha_1, \alpha_2, \alpha_3) = 0$ в силу формул Виета, достаточно определить коэффициенты C и E. Для этого используем две подстановки:

x_1	x_2	x_3	σ_1	σ_2	σ_3	g	уравнение
1	-1	0	0	-1	0	4	$C(-1)^3 = 4 \Rightarrow C = -4$
1	1	-2	0	-3	-2	0	$C(-3)^3 + E(-2)^2 = 0 \Rightarrow E = -27$

По формулам Виета:

$$\sigma_2(\alpha_1, \alpha_2, \alpha_3) = p, \quad \sigma_3(\alpha_1, \alpha_2, \alpha_3) = -q.$$

Следовательно,

$$D(f) = g(\alpha_1, \alpha_2, \alpha_3) = -4p^3 - 27q^2.$$

3°. Решение уравнений степени 3 и 4. Рассмотрим кубическое уравнение

$$y^3 + ay^2 + by + c = 0.$$

С помощью подстановки $y = x - \dfrac{a}{3}$ получим уравнение

$$x^3 + px + q = 0, \tag{23.2}$$

где $p = b - \dfrac{a^2}{3}$ и $q = c - \dfrac{ab}{3} + \dfrac{2a^3}{27}$.

Далее, пусть x_0 — какой-то корень уравнения (23.2). Рассмотрим вспомогательный многочлен от нового переменного u:

$$h(u) = u^2 - x_0 u - \frac{p}{3}.$$

Корни многочлена $h(u)$ — числа α и β, причем

$$\alpha + \beta = x_0 \tag{23.3}$$

$$\alpha\beta = -\frac{p}{3} \tag{23.4}$$

Из (23.3) и (23.2) следует, что

$$(\alpha+\beta)^3 + p(\alpha + be) + q = 0,$$

т.е.

$$\alpha^3 + \beta^3 + \underbrace{(3\alpha\beta + p)}_{=0 \text{ из } (23.4)}(\alpha+\beta) + q = 0.$$

Итак,

$$\alpha^3 + \beta^3 = -q$$

$$\alpha^3\beta^3 = -\frac{p^3}{27}$$

Следовательно, α^3 и β^3 — корни квадратного уравнения

$$z^2 + qz - \frac{p^3}{27} = 0,$$

поэтому числа α и β можно найти с помощью следующих формул, известных как формулы **Кардано**:

$$\alpha = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}, \quad \beta = \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}.$$

Во-первых, заметим, что выражение под знаком квадратного корня пропорционально дискриминанту D многочлена $x^3 + px + q$:

$$\frac{q^2}{4} + \frac{p^3}{27} = -\frac{D}{128}.$$

Во-вторых, имеется 3 комплексных значения каждого из кубических корней в формулах Кардано. Корни исходного уравнения, однако, получаются только при таком выборе корней, когда выполняется равенство (23.4). В-третьих, числа α и β могут быть комплексными даже в том случае, когда уравнение имеет 3 действительных корня. В четвертых, для уравнения с вещественными коэффициентами важную роль играет знак дискриминанта:

— если $D < 0$, то уравнение (23.2) имеет 1 вещественный корень и два комплексных сопряженных корня,

— если $D = 0$, то уравнение (23.2) имеет 1 или 2 вещественных корня,

— если $D > 0$, то уравнение (23.2) имеет 3 различных вещественных корня.

Мы не будем доказывать эти утверждения.

Схематично опишем метод сведения уравнения степени 4 к уравнению степени 3, известный как метод **Феррари**. Для уравнения

$$y^4 + ay^3 + by^2 + cy + d = 0$$

сначала с помощью подстановки $y = x - \frac{a}{4}$ перейдем к "неполному" уравнению

$$x^4 + px^2 + qx + r = 0.$$

Далее введем вспомогательный параметр α и применим тождественное преобразование:

$$x^4 + px^2 + qx + r = \left(x^2 + \frac{p}{2} + \alpha\right)^2 - \left[2\alpha x^2 - qx + \left(\alpha^2 + p\alpha - r + \frac{p^2}{4}\right)\right] = 0.$$

Заметим, что если многочлен от x, стоящий в квадратных скобках, является полным квадратом линейного многочлена, то уравнение распадается на два квадратных уравнения. А для того, чтобы многочлен степени 2 был полным квадратом, необходимо и достаточно, чтобы его дискриминант был равен 0, т.е. чтобы выполнялось равенство

$$q^2 - 8\alpha\left(\alpha^2 + p\alpha - r + \frac{p^2}{4}\right) = 0.$$

Это — кубическое уравнение относительно α, найдя хотя бы один корень которого мы получим решение исходного уравнения степени 4.

Рекомендуемая литература.

Список учебников, приведенных здесь, далеко не полон. Кроме того, большинство из них многократно переиздавались. Мы даем ссылку лишь на одно издание каждой книги.

1. Винберг Э.Б. Курс алгебры. М., Факториал Пресс, 2002.
2. Кострикин А.И. Введение в алгебру. Т.1. Основы алгебры. М., Физико-математическая литература, 2002.
3. Курош А.Г. Курс высшей алгебры. М., Наука, 1975.
4. Михалев А.А., Михалев А.В. Начала алгебры, часть 1. М., ИУИТ, 2005.

Для дальнейшего изучения различных разделов алгебры можно рекомендовать книги

1. Бахтурин Ю.А. Основные структуры современной алгебры. М., Наука, 1990.
2. Ван дер Варден Б.Л. Алгебра. М., Наука, 1976.
3. Ленг С. Алгебра. М., Мир, 1968.