Лекции по алгебре, I семестр, мех-мат МГУ

В. А. Артамонов

Содержание

ГЛАВА 1

Системы линейных уравнений и матрицы

1. Метод Гаусса

Рассмотрим прямоугольную систему линейных уравнений

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{1}$$

коэффициенты которой a_{ij}, b_j заданы. Наша цель – найти все решения и указать алгоритм для нахождения всех решений.

Определение 1.1. *Решением* системы (1) называется такой набор чисел $(\alpha_1, \ldots, \alpha_n)$, что для всех $i = 1, \ldots, m$ выполнены равенства

$$a_{i1}\alpha_1 + \cdots + a_{in}\alpha_n = b_i.$$

Мы будем различать следующие виды систем (1).

Определение 1.2. Система (1) *несовместна*, если она не имеет решения. Система (1) *совместна*, если она имеет решение. Совместная система (1) *неопределенна*, если она имеет более одного решения. Совместная система (1) *определенна*, если она имеет единственное решение.

Мы будем совершать ряд преобразований системы (1), не меняющие множества ее решений.

Определение 1.3. Две системы линейных уравнений вида (1) *эквивалентны*, если они имеют одинаковые множества решений.

Мы будем совершать ряд простейших преобразований системы (1), сохраняющих множества решений. Заметим, что все информация о системе (1) содержится в ее таблице ее коэффициентов.

Определение 1.4. *Матрицей* системы (1) называется прямоугольная таблица

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots\cdots & \cdots & \cdots\cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \tag{2}$$

Расширенной матрицей системы (1) называется прямоугольная таблица

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \cdots\cdots & \cdots & \cdots\cdots & \cdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix} \tag{3}$$

Замечание 1.5. Иногда расширенную матрицу системы (1) обозначают через

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right).$$

Определение 1.6. Следующие преобразования системы (1) (ее (расширенной) матрицы) называются *элементарными*:

- ♢ прибавление к одному уравнению (строке) другого уравнения (другой строки), умноженного(ой) на произвольное число;
- ♡ умножение уравнение (строки) на ненулевое число.

Теорема 1.7. *При элементарных преобразованиях переходим к эквивалентной системе.*

Доказательство. Предположим, что мы совершаем преобразование типа ♢, именно, к i-ому уравнению прибавляем j-ое, умноженное на α. Если $(\beta_1, \dots, \beta_n)$ – решение исходной системы (1). Все уравнения новой системы, кроме i-го, не изменились. Если мы подставим набор $(\beta_1, \dots, \beta_n)$ в i-ое уравнение новой системы, то получим

$$(a_{i1} + \alpha a_{j1})\beta_1 + \dots + (a_{in} + \alpha a_{jn})\beta_n =$$
$$(a_{i1}\beta_1 + \dots + a_{in}\beta_n) + \alpha(a_{j1}\beta_1 + \dots + a_{jn}\beta_n) = b_i + \alpha + b_j.$$

Таким образом, $(\beta_1, \dots, \beta_n)$ является решением новой системы. Поскольку исходная системы (1) получается из новой системы элементарным преобразованием прибалением к i-ому уравнению j-го, умноженного на $-\alpha$, то аналогично, каждое решение новой системы является решением исходной системы. □

Упражнение 1.8. Доказать, что совершая элементарные преобразования со строками матрицы можно в ней переставить любые две строки.

Будем приводить матрицу системы к наиболее простому – ступенчатому виду.

Определение 1.9. Матрица (3) называется *ступенчатой*, если

(1) ниже нулевой строки расположены только нулевые строки;
(2) первый ненулевой каждой строки равен 1;
(3) если первый ненулевой i-ой строки расположен на месте (i, k_i), то
  (a) $k_{i+1} > k_i$;
  (b) все элементы $a_{j,k_i} = 0$ для всех $j \neq i$.

Теорема 1.10. *Каждая матрица конечным числом элементарных преобразований строк приводится к ступенчатому виду.*

Доказательство. Пусть матрица A имеет вид (3). Если $A = 0$, то она уже имеет ступенчатый вид.

Пусть $A \neq 0$. Будем вести доказательство индукцией по числу строк m. Без ограничения общности можно считать, что в первом столбце есть ненулевой элемент a_{i1}. Если $i = 1$, то умножим 1-ую строку на a_{11}^{-1}. Итак, можно предполагать, что $a_{11} = 1$. Следовательно, если $m = 1$, то теорема доказана.

Пусть $m > 1$, и для $m - 1$ теорема доказана. Для каждого $i > 1$ вычтем из i-ой строки первую строку, умноженную на $a_{i1}a_{11}^{-1}$. В новой матрице все коэффициенты $a_{i1} = 0, i > 1$.

Рассмотрим подматрицу B в A, получающуюся отбрасыванием первой строки. По индукции можно считать, что матрица B имеет ступенчатый вид. Пусть в матрице B первые ненулевые элементы расположены в столбцах с номерами $1 < k_2 < k_3 < \dots$. Вычтем из первой строки 2-ую строку, умноженную на a_{1,k_2}, третью строку 3-ую строку, умноженную на a_{1,k_3}, и т. д. □

Определение 1.11. Пусть матрица системы (1) имеет ступенчатый вид. Назовем неизвестную x_i *главной*, если в некотором уравнении все коэффициенты при $x_1, \dots, x_{i-1}$ равны нуля, а коэффициент при x_i отличен от нуля (и потому равен 1). все остальные неизвестные назовем *свободными*.

Применим теоремы 1.7, 1.10 к исследованию системы (1). В силу указанных теорем можно считать, что расширенная матрица системы (1) имеет ступенчатый вид.

Пусть ее последняя ненулевая строка имеет вид

$$(0, \dots, 0, 1). \tag{4}$$

Это означает, что системы (1) содержит уравнение

$$0x_1 + \cdots + 0x_n = 1,$$

что невозможно. Следовательно, в этом случае система несовместна.

Пусть в A нет строки (4). Предположим для простоты, что переменные $x_1, \dots, x_r$ главные, а $x_{r+1}, \dots, x_n$ свободные. Тогда система имеет вид

$$\begin{cases} x_1 & & +a_{1,r+1}x_{r+1} & +\cdots+ & a_{1n}x_n & = b_1 \\ & x_2 & +a_{2,r+1}x_{r+1} & +\cdots+ & a_{2n}x_n & = b_2 \\ \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots \\ & x_r & +a_{r,r+1}x_{r+1} & +\cdots+ & a_{rn}x_n & = b_r \end{cases} \tag{5}$$

Перенося свободные переменные в правую часть, получаем выражение главных неизвестных через свободные

$$\begin{cases} x_1 &= b_1 - a_{1,r+1}x_{r+1} - \cdots - a_{1n}x_n \\ x_2 &= b_2 - a_{2,r+1}x_{r+1} - \cdots - a_{2n}x_n \\ \dots\dots & \dots\dots \\ x_r &= b_r - a_{r,r+1}x_{r+1} - \cdots - a_{rn}x_n. \end{cases} \tag{6}$$

Таким образом, придавая свободным неизвестным произвольные значения, мы однозначно находим значения главных неизвестных. Итак, система совместна, и, если есть свободные неизвестные, то система неопределенна. Если все неизвестные главные, то система определенна.

Определение 1.12. Система (1) *однородна*, если все ее свободные члены нулевые, т. е. $b_1 = \cdots = b_m = 0$.

Предложение 1.13. *Если в однородной системе число неизвестных n больше числа уравнений m, то система неопределенна.*

Доказательство. Приведем систему к ступенчатому виду. Ясно, что снова получим однородную систему, причем число главных неизвестных не превосходит числа ненулевых уравнений, т. е. не все неизвестные главные. □

2. Матрицы и операции над ними

Определение 1.14. $\mathrm{Mat}(n \times m)$ – всех матриц (прямоугольных таблиц) с n строками и m столбцами. Если $A \in \mathrm{Mat}(n \times m)$, то мы будем также писать $A = A_{n\times m}$. Если $A_{n\times m} = (a_{ij}), B_{n\times m} = (b_{ij})$, то полагаем $A + B = (a_{ij} + b_{ij})$. Кроме того, $\lambda A_{n\times m} = (\lambda a_{ij})$.

Предложение 1.15. *Пусть $A, B, C \in \mathrm{Mat}(n \times m)$ и λ, ν – числа. Тогда справедливы следующие 8 аксиом векторного пространства:*

(1) $A + B = B + A$;
(2) $A + (B + C) = (A + B) + C$;
(3) *если 0 – нулевая матрица (все ее коэффициенты равны нулю), то $A + 0 = A$ для любой матрицы A;*
(4) *для любой матрицы A существует такая матрица $-A$, что $A + (-A) = 0$;*
(5) $\lambda(A + B) = \lambda A + \lambda B$;
(6) $(\lambda + \nu)A = \lambda A + \nu A$;
(7) $(\lambda\nu)A = \lambda(\nu A)$;
(8) $1A = A$.

Доказательство. Приведем, например, доказательство первого утверждения, Если $A = (a_{ij}), B = (b_{ij})$, то $A + B = (a_{ij} + b_{ij}) = (b_{ij} + a_{ij}) = B + A$. Остальные утверждения доказываются аналогичною □

Определение 1.16. Пусть

$$A_{n\times m} = (a_{ij}), C_{m\times k} = (c_{st}).$$

Тогда $D = AC \in \mathrm{Mat}(n \times k) = (d_{is})$, где для всех $i = 1, \dots, n, \quad s = 1, \dots, k$

$$d_{is} = a_{i1}d_{1s} + \cdots + a_{in}d_{ns} \tag{7}$$

Предложение 1.17. *Умножение матриц ассоциативно, т.е.* $(AC)F = A(CF)$ *для любых матриц*

$$A \in \mathrm{Mat}(n \times m), C \in \mathrm{Mat}(m \times k), F \in \mathrm{Mat}(k \times l).$$

Доказательство. Пусть

$$A = A_{n\times m} = (a_{ij}), \quad C = C_{m\times k} = (c_{st}), \quad F = F_{k\times l} = (f_{tq}).$$

Если D из определения 1.16, то по (7) на месте (i, q) в матрице $(AC)F = DF$ стоит элемент

$$\sum_{=1}^{k} d_{i,}f_{,q} = \sum_{=1}^{k}\sum_{=1}^{n} a_{i,}d_{,}f_{,q}. \tag{8}$$

С другой стороны, если

$$CF = U = (u_{i,}) \in \mathrm{Mat}(m \times l),$$

то на месте (i, q) в матрице $A(CF) = AU$ стоит элемент

$$\sum_{=1}^{n} a_{i,}u_{,q} = \sum_{=1}^{n}\sum_{=1}^{k} a_{i,}d_{,}f_{,q}. \tag{9}$$

Из (8), (9) вытекает утверждение. □

Предложение 1.18. *Справедливы равенства:*

(1) $\lambda(AB) = (\lambda A)B = A(\lambda B)$.
(2) $A(B + C) = AB + AC, \quad (A + U)V = AV + UV$.

Доказательство. Докажем, например, второе утверждение. Пусть $A = (a_{ij}), B = (b_{ij}), C = (c_{ij})$. Тогда на месте (i, j) в матрице $A(B + C)$ строит элемент

$$\sum a_{ik}(b_{kj} + c_{kj}) = \sum a_{ik}b_{kj} + \sum a_{ik}c_{kj},$$

который равен элементу, стоящему на то же месте в матрице $AB + AC$. Так как размеры матриц $A(B+C)$ и $AB+AC$ совпадают, то они равны. Аналогично проверяются остальные утверждения. □

Определение 1.19. Пусть $A = (a_{ij}) \in \mathrm{Mat}(n)$. *Следом* $\mathrm{tr}\, A$ называется $a_{11} + \cdots + a_{nn}$.

Предложение 1.20. *Пусть* $A, B \in \mathrm{Mat}(n)$. *Тогда* $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Доказательство. Пусть $A = (a_{ij}), B = (b_{ij})$. Тогда на месте (i, i) в матрице AB стоит $\sum_{j=1}^{n} a_{ij}b_{ji}$, откуда

$$\mathrm{tr}(AB) = \sum_{i,j=1}^{n} a_{ij}b_{ji}.$$

Аналогично,

$$\mathrm{tr}(BA) = \sum_{s,t=1}^{n} b_{st}a_{ts} = \sum_{s,t=1}^{n} a_{ts}b_{st} = \mathrm{tr}(AB).$$

□

Определение 1.21. *Символ Кронекера* δ_{ij} равен 1, если $i = j$, и 0, если $i \neq j$. *Единичная* матрица $E = E_n \in \mathrm{Mat}(n)$ – это матрица, в которой на месте (i, j) стоит символ Кронекера δ_{ij}.

Предложение 1.22. *Пусть* $A \in \mathrm{Mat}(n \times m)$. *Тогда* $E_n A = A = AE_m$.

Доказательство. Пусть $A = (a_{ij})$. Тогда на месте (i, j) в матрице $E_n A$ стоит

$$\sum_{k=1}^{n} \delta_{ik} a_{kj} = \delta_{ii} a_{ij} = a_{ij},$$

т. е. $E_n A = A$. □

Определение 1.23. Пусть $A \in \mathrm{Mat}(n \times m)$. Тогда *транспонированная* матрица ${}^tA = A^* \in \mathrm{Mat}(m \times n)$ – это матрица, в которой на месте (i, j) стоит элемент a_{ji} матрицы A.

Предложение 1.24. ${}^t(A + B) = {}^tA + {}^tB$, $\quad {}^t(\lambda A) = \lambda {}^tA$, $\quad {}^t(AC) = {}^tC\,{}^tA$.

Доказательство. Докажем, например, последнее утверждение. В матрице ${}^t(AC)$ на месте (i, j) стоит $\sum_k a_{jk} c_{ki} = \sum_k c_{ki} a_{jk}$, т. е. элемент, стоящей на том же месте в матрице ${}^tC\,{}^tA$.

Аналогично доказываются остальные утверждения. □

Определение 1.25. *Матричные единицы* $E_{ij} \in \mathrm{Mat}(n \times m)$ – это матрицы E_{ij}, в которых на месте (s, t) стоит элемент $\delta_{si}\delta_{tj}$, т. е. на месте (i, j) стоит 1, и все остальные элементы равны 0.

Упражнение 1.26. Доказать, что

♠ ${}^tE_{ij} = E_{ji}$;

♣ если $A = (a_{ij})$, то $A = \sum_{i,j} a_{ij} E_{ij}$.

Предложение 1.27. $E_{ij} E_{rs} = \delta_{jr} E_{is}$.

Доказательство. На месте (u, v) в $E_{ij}E_{rs}$ стоит элемент

$$\sum_p (\delta_{ui}\delta pj)(\delta_{pi}\delta_{vs}) = \begin{cases} 1, & u = i = p = j, v = s; \\ 0 & \text{в противном случае.} \end{cases}$$

Отсюда вытекает утверждение. □

Следствие 1.28. *Пусть* $A = (a_{rs}) \in \mathrm{Mat}(n \times m)$. *Тогда*

$$E_{ij} A = a_{j1} E_{i1} + \cdots + a_{jm} E_{im},$$
$$A E_{ij} = a_{1i} E_{1j} + \cdots + a_{nj} E_{ni}.$$

Теорема 1.29. *Чтобы в матрице* $A \in \mathrm{Mat}(n \times m)$ *к* i*-ой строке прибавить* j*-ую, умноженную на* α *нужно рассмотреть матрицу* $(E_n + \alpha E_{ij})A$.

Доказательство. По предложениям 1.18, 1.27, упражнениям 1.22, ♣ и следствию 1.28

$$(E_n + \alpha E_{ij})A = E_n A + \alpha E_{ij} A =$$
$$A + \alpha(a_{j1} E_{i1} + \cdots + a_{jm} E_{im}) = \textstyle\sum_{rs} a_{rs} E_{rs} + (\alpha a_{j1}) E_{i1} + \cdots + (\alpha a_{jm}) E_{im} =$$
$$\textstyle\sum_{r \neq i, s} a_{rs} E_{rs} + \sum_{i,s} (a_{is} + \alpha a_{js}) E_{is}.$$

□

Следствие 1.30. *Чтобы в матрице* $A \in \mathrm{Mat}(n \times m)$ *к* i*-ому столбцу прибавить* j*-ый, умноженную на* α *нужно рассмотреть матрицу* $A(E_m + \alpha E_{ji})$.

Обозначение 1.31. Положим $D_i(\alpha) = E_n + (\alpha - 1)E_{ii} \in \mathrm{Mat}(n)$.

Теорема 1.32. *Чтобы в матрице $A \in \mathrm{Mat}(n \times m)$ i-ую строку (столбец) умножить на α нужно рассмотреть матрицу $D_i(\alpha)A$ ($AD_i(\alpha)$).*

Доказательство. Пусть $A = (a_{ij})$. Заметим, что в матрице $D_i(\alpha)$ на месте (s,t) стоит $\delta_{st} + (\alpha - 1)\delta_{is}\delta_{it}$. Поэтому в матрице $D_i(\alpha)A$ на месте (p,q) стоит

$$\textstyle\sum_r (\delta_{pr} + (\alpha-1)\delta_{ip}\delta_{ir})a_{rq} = \sum_r \delta_{pr}a_{rq} + (\alpha-1)\sum_r \delta_{ip}\delta_{ir}a_{rq} =$$

$$a_{pq} + (\alpha-1)\delta_{ip}a_{iq} = \begin{cases} a_{pq}, & i \neq q; \\ \alpha a_{iq}, & p = i. \end{cases}$$

□

Замечание 1.33. В терминах матричного умножения удобно записывать системы линейных уравнений. Именно, системы (1) имеет вид $AX = b$, где A – матрица (2) системы (1),

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

– столбец неизвестных,

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

– столбец свободных членов.

ГЛАВА 2

Множества и отображения. Перестановки

1. Множества и отображения.

Определение 2.1. *Отображение*

$$f : X \to Y. \tag{10}$$

Отображение (10)

(1) *инъективно*, если для любых $x, y \in X$ из $f(x) = f(y)$ следует $x = y$;
(2) *сюръективно*, если для любого $y \in Y$ существует такое $x \in X$, что $f(x) = y$;
(3) *биективно*, если оно инъективно и сюръективно.

Обозначение 2.2. Пусть задано отображение (10). Для любого $y \in Y$ положим

$$f^{-1}(y) = \{x \in X | f(x) = y\}.$$

Упражнение 2.3. Доказать, что отображение (10)

(1) *инъективно* тогда и только тогда, когда $|f^{-1}(y)| \leq 1$ для любого $y \in Y$;
(2) *сюръективно* тогда и только тогда, когда $|f^{-1}(y)| \geq 1$ для любого $y \in Y$;
(3) *биективно* тогда и только тогда, когда $|f^{-1}(y)| = 1$ для любого $y \in Y$.

Определение 2.4. Пусть $f : X \to Y, g : Y \to Z$. *Произведение (композиция)* отображений $gf : X \to Z$. *Тождественное* отображение $1_X : X \to X$. *Обратное* отображение $f^{-1} : Y \to X$.

Предложение 2.5. *Справедливы следующие утверждения:*

(1) *умножение отображений ассоциативно;*
(2) *произведение инъективных отображений инъективно;*
(3) *произведение сюръективных отображений сюръективно;*
(4) *если* f *из* (10), *то* $1_Y f = f 1_X = f$;
(5) *если* f *из* (10), *то* $f^{-1} f = 1_X$ *и* $f f^{-1} = 1_Y$;
(6) *обратное отображение к* f *из* (10) *существует тогда и только тогда, когда* f *биективно.*

Упражнение 2.6. Множество X конечно тогда и только тогда, когда любое инъективное (сюръективное) отображение $X \to X$ биективно.

2. Перестановки

Пусть $X_n = \{1, 2, \ldots, n\}$.

Определение 2.7. *Перестановкой (подстановкой) степени* n называется биективное отображение X_n в себя. Через S_n обозначается множество всех перестановок степени n.

Предложение 2.8. *Произведение перестановок и обратная и тождественная перестановки снова являются перестановками. Умножение перестановок ассоциативно.*

Доказательство. Нужно воспользоваться предложением 2.5. □

Обозначение 2.9. Пусть $\sigma \in S_n$. Тогда если $X_n = \{i_1, \dots, i_n\}$, то σ однозначно задается в виде двустрочной матрицы

$$\sigma = \begin{pmatrix} i_1 & \dots & i_n \\ \sigma(i_1) & \dots & \sigma(i_1) \end{pmatrix} \tag{11}$$

Упражнение 2.10. Пусть σ из (11), и

$$\tau = \begin{pmatrix} j_1 & \dots & j_n \\ i_1 & \dots & i_n \end{pmatrix}.$$

Тогда

$$\sigma\tau = \begin{pmatrix} j_1 & \dots & j_n \\ \sigma(i_1) & \dots & \sigma(i_n) \end{pmatrix}$$

и

$$\sigma^{-1} = \begin{pmatrix} \sigma(i_1) & \dots & \sigma(i_n) \\ i_1 & \dots & i_1 \end{pmatrix}$$

Определение 2.11. Пусть $i_1, \dots, i_k$ – различные числа из X_n. *Циклом* $(i_1, \dots, i_k) \in S_n$ длины k называется такая перестановка σ, что для $m \in X_n$

$$\sigma(m) = \begin{cases} i_{s+1}, & \text{если } m = i_s, s < k; \\ i_1, & \text{если } m = i_k; \\ m, , & \text{если } m \in X_n \setminus \{i_1, \dots, i_k\}. \end{cases}$$

Два цикла $(i_1, \dots, i_k), (j_1, \dots, j_s) \in S_n$ *независимы*, если все элементы $i_1, \dots, i_k, j_1, \dots, j_s$ различны.

Теорема 2.12. *Любая перестановка разлагается в произведение независимых циклов.*

Доказательство. Пусть $\sigma \in S_n$. Можно считать, что $\sigma \neq 1$. Возьмем произвольный элемент $k, 1 \leq k \leq n$, и предположим, что элементы $k_0 = k, k_1 = \sigma k, k_2 = \sigma^2 k, \dots, k_l = \sigma^l k$ различны, но $\sigma^{l+1} k = \sigma^s k$, где $0 \leq s \leq l$.

Лемма 2.13. $s = 0$.

Доказательство. Если $s > 0$, то $\sigma(k_{s-1}) = \sigma(k_l)$, что невозможно, ибо σ действует инъективно на $X = \{1, \dots, n\}$, но $k_{s-1} \neq k_l$ в силу выбора l. □

Итак, на множестве $\{k_0, k_1, \dots, k_l\}$ подстановка σ действует как

$$\begin{pmatrix} k_0 & k_1 & \dots & k_{l-1} & k_l \\ k_1 & k_2 & \dots & k_l & k_0 \end{pmatrix}$$

Выберем теперь произвольное число $j, 1 \leq j \leq n$, причем $j \notin \{k_0, k_1, \dots, k_l\}$. Как и выше строим множество $\{j_0, j_1, \dots, j_t\}$, на котором подстановка σ действует как цикл

$$\begin{pmatrix} j_0 & j_1 & \dots & j_{t-1} & j_t \\ j_1 & j_2 & \dots & j_t & j_0 \end{pmatrix}$$

Лемма 2.14. *Все элементы $k_0, k_1, \dots, k_l, j_0, j_1, \dots, j_t$ различны.*

Доказательство. Пусть $j_r = k_q$. Тогда

$$j_0 = \sigma^{-r} j_r \in \{k_0, k_1, \dots, k_l\},$$

что невозможно. □

Продолжая этот процесс, получаем подстановку

$$\tau = \begin{pmatrix} k_0 & k_1 & \dots & k_{l-1} & k_l \\ k_1 & k_2 & \dots & k_l & k_0 \end{pmatrix} \begin{pmatrix} j_0 & j_1 & \dots & j_{t-1} & j_t \\ j_1 & j_2 & \dots & j_t & j_0 \end{pmatrix} \cdots .$$

Непосредственная проверка показывает, что $\tau = \sigma$. □

Предложение 2.15. *Пусть $\pi \in S_n$ и $(i_1, \dots, i_k)$ – цикл из S_n. Тогда*

$$\pi(i_1, \dots, i_k)\pi^{-1} = (\pi(i_1), \dots, \pi(i_k)).$$

Доказательство. Непосредственная проверка. □

Определение 2.16. *Транспозицией* называется цикл длины 2.

Теорема 2.17. *Каждая перестановка является произведением транспозиций.*

Доказательство. $(i_1, \dots, i_k) = (i_1, i_2)(i_2, i_3) \cdots (i_{k-1}, i_k)$. □

Определение 2.18. Пусть $i_1, \dots, i_n$ – последовательность различных чисел из X_n. *Инверсией* в этой последовательности называется такая пара i_s, i_t, что $s < t$ и $i_s > i_t$. *Знаком* последовательности называется число $(-1)^M$, где M – число инверсий в последовательности. Если подстановка $\sigma \in S_n$ имеет двустрочную запись (11), где $i_1 = 1, \dots, i_n = n$, то *знак* $(-1)^\sigma$ перестановки σ равен знаку последовательности из второй строки.

Теорема 2.19. *Пусть задана перестановка* (11)*, где $i_k = k$ для всех k. Предположим, что заданы различные числа $y_1, \dots, y_n$. Тогда*

$$(-1)^{sigma} = \prod_{1 \le s < t \le n} \frac{y_{\sigma t} - y_{\sigma s}}{y_t - y_s}. \tag{12}$$

Доказательство. Если пара $\sigma s, \sigma t$ образует инверсию, то в числитель дроби (12) входит $y_{\sigma t} - y_{\sigma s}$, а в знаменателе встречается $y_{\sigma s} - y_{\sigma t}$. При делении возникает множитель -1. Если же эта пара не образует инверсии, то при делении возникает множитель 1. □

Теорема 2.20. *Пусть $\sigma, \tau \in S_n$. Тогда $(-1)^{\sigma\tau} = (-1)^\sigma (-1)^\tau$.*

Доказательство. Пусть $y_1, \dots, y_n$ – различные числа. Тогда и числа

$$z_1 = y_{\tau 1}, \dots, z_n = y_{\tau n}$$

различны. По теореме 2.19 имеем

$$(-1)^{\sigma\tau} = \prod_{1 \le s < t \le n} \frac{y_{\sigma\tau t} - y_{\sigma\tau s}}{y_t - y_s} =$$
$$\prod_{1 \le s < t \le n} \frac{y_{\sigma\tau t} - y_{\sigma\tau s}}{y_{\tau t} - y_{\tau s}} \prod_{1 \le s < t \le n} \frac{y_{\tau t} - y_{\sigma s}}{y_t - y_s} =$$
$$\prod_{1 \le s < t \le n} \frac{z_{\sigma t} - z_{\sigma s}}{z_t - z_s} \prod_{1 \le s < t \le n} \frac{y_{\tau t} - y_{\sigma s}}{y_t - y_s} =$$
$$(-1)^\sigma (-1)^\tau.$$

□

Следствие 2.21. $(-1)^{\sigma^{-1}} = (-1)^\sigma$.

Доказательство. Имеем

$$\sigma^{-1}\sigma = \varepsilon = \begin{pmatrix} 1 & \dots & n \\ 1 & \dots & n \end{pmatrix}.$$

Следовательно,

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

□

Следствие 2.22. *Если $\sigma \in S_n$ имеет произвольную двустрочную запись* (11), *то знак σ равен произведению знаков последовательностей из нижней и верхней строк.*

Доказательство. Перестановка (11) равна произведению перестановок $\psi\tau$, где

$$\psi = \begin{pmatrix} i_1 & \dots & i_n \\ 1 & \dots & n \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & \dots & n \\ \sigma 1 & \dots & \sigma n \end{pmatrix}.$$

При этом по следствию 2.21 знак ψ равен числу инверсий в верхней строке σ, а знак τ – числу инверсий в нижней строке σ. Остается воспользоваться теоремой 2.20. □

Теорема 2.23. *Знак транспозиции равен -1.*

Доказательство. По предложению 2.15 имеем

$$(i,j) = \begin{pmatrix} 1 & 2 & \dots \\ i & j & \dots \end{pmatrix} (1,2) \begin{pmatrix} 1 & 2 & \dots \\ i & j & \dots \end{pmatrix}^{-1}.$$

Следовательно, по теореме 2.20 и следствию 2.21 знаки (i,j) и (1,2) совпадают. Непосредственная проверка показывает, что $(-1)^{(1,2)} = -1$. □

Следствие 2.24. *Если подстановка разложена в произведение s транспозиций, то ее знак равен $(-1)^s$. В частности, знак цикла длины k равен $(-1)^{k-1}$.*

Доказательство. Нужно воспользоваться теоремой 2.17. □

Определение 2.25. Перестановка *четна*, если она имеет знак 1, в противном случае она *нечетна*. Через A_n обозначается множество всех четных перестановок из S_n.

Теорема 2.26. $|S_n| = n!, \quad |A_n| = \dfrac{n!}{2}$.

Доказательство. Отображение $\sigma \mapsto \sigma(1,2)$ переводит A_n в $S_n \setminus A_n$ и наоборот. □

ГЛАВА 3

Определители, обратная матрица

1. Определители

Определение 3.1. Пусть задана квадратная матрица

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots\cdots\cdots\cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \tag{13}$$

Определителем $\det A = |A|$ называется число

$$\det A = \sum_{\sigma \in S_n} (-1)^{\sigma} a_{1,\sigma 1} \cdots a_{n,\sigma n}. \tag{14}$$

Теорема 3.2. *Пусть A – верхнетреугольная матрица. Тогда*

$$\det A = a_{11} \cdots a_{nn}.$$

Доказательство. В матрице A элемент $a_{ij} = 0$, если $i > j$. Таким образом, если в определителе $\det A$ произведение

$$\sum_{\sigma \in S_n} (-1)^{\sigma} a_{1,\sigma 1} \cdots a_{n,\sigma n}$$

отлично от нуля, то $1 \le \sigma 1, 2 \le \sigma 2, \ldots, n \le \sigma n$. Учитывая, что индексы $\sigma 1, \ldots, \sigma n$ различны, получаем $n = \sigma n, \sigma(n-1) = n-1, \ldots, \sigma 1 = 1$. □

Теорема 3.3 (Простейшие свойства определителя) . *Если одна из строк A является линейно комбинацией двух строк, то $\det A$ является линейной комбинацией определителей соответствующих матриц.*

Доказательство. Пусть s-ая строка является указанной линейной комбинацией, т. е. $a_{sj} = \alpha a'_{sj} + \beta a''_{sj}$ для всех $j = 1, \ldots, n$. Тогда

$$\det A = \sum_{\sigma \in S_n} (-1)^{\sigma} a_{1,\sigma 1} \cdots a_{s-1,\sigma(s-1)} (\alpha a'_{s,\sigma s} + \beta a''_{s,\sigma s}) a_{s+1,\sigma(s+1)} \cdots a_{n,\sigma n} =$$

$$\alpha \sum_{\sigma \in S_n} (-1)^{\sigma)} a_{1,\sigma 1} \cdots a_{s-1,\sigma(s-1)} a'_{s,\sigma s} a_{s+1,\sigma(s+1)} \cdots a_{n,\sigma n} +$$

$$\beta \sum_{\sigma \in S_n} (-1)^{\sigma} a_{1,\sigma 1} \cdots a_{s-1,\sigma(s-1)} a''_{s,\sigma s} a_{s+1,\sigma(s+1)} \cdots a_{n,\sigma n}.$$

□

Следствие 3.4. *При элементарных преобразованиях строк типа ♡ из главы 1 определитель матрицы умножается на указанное число.*

Теорема 3.5. *Если в A две строки равны, то $\det A = 0$.*

Доказательство. Пусть в A i-ая и j-ая строки равны, т. е. $a_{ik} = a_{jk}$ для всех k, где $i < j$. В (14) входит произведение

$$a_{1,\sigma 1} \cdots a_{i,\sigma i} \cdots a_{j,\sigma j} \cdots a_{n,\sigma n} \tag{15}$$

со знаком $(-1)^\sigma$. Заметим, что в (14) входит также произведение

$$a_{1,\sigma 1} \cdots a_{i,\sigma j} \cdots a_{j,\sigma i} \cdots a_{n,\sigma n} \tag{16}$$

со знаком $(-1)^{\sigma(i,j)} = -(-1)^\sigma$ в силу теоремм 2.20 и 2.23. По условию произведения (15), (16) равны. □

Теорема 3.6. *При элементарных преобразованиях ♢ из главы 1 определитель не меняется.*

Доказательство. Выделим в матрице A строки A_i и A_j с номерами $i < j$,

$$A = \begin{pmatrix} \cdots \\ A_i \\ \cdots \\ A_j \\ \cdots \end{pmatrix}.$$

После преобравзования ♢ из главы 1 получаем матрицу

$$\begin{pmatrix} \cdots \\ A_i + \lambda A_j \\ \cdots \\ A_j \\ \cdots \end{pmatrix}.$$

Ее определитель по теореме 3.3 и теореме 3.5 равен

$$\begin{vmatrix} \cdots \\ A_i + \lambda A_j, \\ \cdots \\ A_j \\ \cdots \end{vmatrix} = \begin{vmatrix} \cdots \\ A_i \\ \cdots \\ A_j \\ \cdots \end{vmatrix} + \lambda \begin{vmatrix} \cdots \\ A_j \\ \cdots \\ A_j \\ \cdots \end{vmatrix} = \det A.$$

□

Следствие 3.7. *Если в матрице переставить две строки, то определитель изменит знак.*

Доказательство. Рассмотрим вспомогательную матрицу B, в которой i-ая и j-ая строки равны сумме $A_i + A_j$ i-ой и j-ой строк исходной матрницы A,

$$B = \begin{pmatrix} \cdots \\ A_i + A_j \\ \cdots \\ A_i + A_j \\ \cdots \end{pmatrix}.$$

Тогда по теореме 3.3

$$\det B = \begin{vmatrix} \cdots \\ A_i + A_j \\ \cdots \\ A_i + A_j \\ \cdots \end{vmatrix} = \begin{vmatrix} \cdots \\ A_i \\ \cdots \\ A_i + A_j \\ \cdots \end{vmatrix} + \begin{vmatrix} \cdots \\ A_j \\ \cdots \\ A_i + A_j \\ \cdots \end{vmatrix} = \begin{vmatrix} \cdots \\ A_i \\ \cdots \\ A_i \\ \cdots \end{vmatrix} + \begin{vmatrix} \cdots \\ A_i \\ \cdots \\ A_j \\ \cdots \end{vmatrix} + \begin{vmatrix} \cdots \\ A_j \\ \cdots \\ A_i \\ \cdots \end{vmatrix} + \begin{vmatrix} \cdots \\ A_j \\ \cdots \\ A_j \\ \cdots \end{vmatrix} = \begin{vmatrix} \cdots \\ A_i \\ \cdots \\ A_j \\ \cdots \end{vmatrix} + \begin{vmatrix} \cdots \\ A_j \\ \cdots \\ A_i \\ \cdots \end{vmatrix}$$

Отсюда вытекает утверждение. □

Теорема 3.8. $\det A = \det({}^tA)$.

Доказательство. В (14) входит произведение (15) со знаком

$$(-1)^{\text{четность } (i_1,\dots,i_n)}.$$

Оно входит в $\det({}^tA)$ со знаком

$$(-1)^{\text{четность } (j_1,\dots,j_n)}.$$

Нужно воспользоваться следствием 2.21. □

Следствие 3.9. *Все свойства строк в* $\det A$ *справедливы и для столбцов.*

Теорема 3.10 (Определитель с углом нулей). *Пусть*

$$A \in \mathrm{Mat}(n), \quad C \in \mathrm{Mat}(n \times m), \quad C \in \mathrm{Mat}(m).$$

Тогда

$$\left|\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right| = \det A \det B. \tag{17}$$

Доказательство. Приведем матрицы A и B к ступенчатому виду.

$$A \rightsquigarrow \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & \dots & a_{2,n-1} & a_{2,n} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix}, \quad B \rightsquigarrow \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1,m-1} & b_{1m} \\ 0 & b_{22} & \dots & b_{2,m-1} & b_{2,m} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & b_{mm} \end{pmatrix}.$$

Тогда по теоремам 3.6, 3.2 и следствию 3.4

$$\det A = \lambda a_{11} \cdots a_{nn}, \quad \det B = \mu b_{11} \cdots b_{mm},$$

где λ, μ возникают из-за применения преобразований типа ♡. Те же преобразования, примененные к матрице (17), приводят ее к верхнетреугольному виду с элементами

$$a_{11}, \dots, a_{nn}, b_{11}, \dots, b_{mm}$$

по главной диагонали. При этом определитель матрицы (17) равен

$$\lambda\mu a_{11} \cdots a_{nn} b_{11} \cdots b_{mm} = \det A \det B.$$

□

Следствие 3.11. *Определитель* , -¤¥à¬®-¤

$$W(x_1, \dots, x_n) = \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \dots & \dots & \dots & \dots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} = \prod_{1 \le i < j \le n} (x_j - x_i).$$

Доказательство. Случай $n = 2$ очевиден. Пусть для $n-1$ утверждение верно. Тогда вычтем из каждой строки, начиная снизу, предыдущую, умноженную на x_1. Получаем

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

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

$$(x_2 - x_1)\cdots(x_n - x_1)\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_2 & x_3 & \dots & x_n \\ x_2^2 & x_3^2 & \dots & x_n^2 \\ \dots & \dots & \dots & \dots \\ x_2^{n-2} & x_3^{n-2} & \dots & x_n^{n-2} \end{vmatrix} =$$

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

□

Теорема 3.12. *Пусть* $A, B \in \mathrm{Mat}(n)$. *Тогда* $\det(AB) = \det A \det B$.

Доказательство. Рассмотрим вспомогательную матрицу

$$\left|\begin{array}{c|c} A & 0 \\ \hline -E & B \end{array}\right|. \tag{18}$$

Транспонируя матрицу (18) получаем новую матрицу

$$\left|\begin{array}{c|c} {}^tA & -E \\ \hline 0 & {}^tB \end{array}\right|$$

с тем же определителем, равным $\det A \det B$. по теореме теореме 3.10. С другой стороны, прибавляя в каждому столбцу с номером $j > n$ линейную комбинацию первых n столбцов с коэффициентами, соответственно, $b_{1j}, b_{2j}, \cdots, b_{nj}$, получаем, что определитель (18) равен определителю матрицы

$$\left|\begin{array}{c|c} A & AB \\ \hline -E & 0 \end{array}\right|,$$

который с помощью n перестановок столбцов сводится к определителю

$$(-1)^n\left|\begin{array}{c|c} AB & A \\ \hline 0 & -E \end{array}\right| = \det(Ab)(-1)^{2n} = \det(AB).$$

□

Определение 3.13. *Минором* $M_{ij}, 1 \le i, j \le n$, матрицы (13) называется определитель матрицы размера $n-1$, получающейся вычеркиванием i-ой строки и j-го столбца. *Алгебраическим дополнением* A_{ij} называется $(-1)^{i+j}M_{ij}$.

Теорема 3.14 (Разложение определителя по строке (столбцу)) . *Для матрицы* A *из* (13) *и любого* $i = 1,\dots,n$ *имеем*

$$\det A = a_{i1}A_{i1} + \cdots + a_{in}A_{in}.$$

Доказательство. Имеем

$$(a_{i1}, \dots, a_{in}) = \sum_{j=1}^{n} (0, \dots, 0, a_{ij}, 0, \dots, 0)$$

Поэтому в силу теоремы 3.3 можно считать, что i-ая строка имеет вид

$$(0, \dots, 0, a_{ij}, 0, \dots, 0)$$

Переставляя эту строку на первой место со всеми предыдущими мы умножим определитель матрицы на $(-1)^{i-1}$. Затем переставляя столбцы мы умножим определитель матрицы на $(-1)^{j-1}$. Итак, определитель матрицы умножится на $(-1)^{i+j}$ и по теореме 3.10 он станет равным $a_{ij}M_{ij}$. □

Следствие 3.15 (Фальшивое разложение). *Если $i \neq j$, то*

$$a_{i1}A_{j1} + \dots + a_{in}A_{jn} = 0.$$

Доказательство. Воспользоваться теоремой 3.5 для вспомогательной матрицы, получающейся из A заменой j-ой строки на i-ую. □

Обозначение 3.16. Пусть задана квандратная матрица $A = (a_{ij}) \in \mathrm{Mat}(n)$. Через $\hat{A}\,\mathrm{Mat}(n)$ обозначим матрицу, в которой на месте (i, j) стоит алгебраическое дополнение A_{ji}.

Теорема 3.17. *Если $A \in \mathrm{Mat}(n)$, то $A\hat{A} = \hat{A}A = |A|E_n$.*

Доказательство. На месте (i, j) в $A\hat{A}$ стоит

$$a_{i1}A_{j1} + \dots + a_{in}A_{jn} = 0.$$

Остается воспользоваться теоремой 3.14 и следствием 3.15. □

Теорема 3.18. *Пусть $A \in \mathrm{Mat}(n)$. Следующие условия эквивалентны:*

(1) *матрица A элементарными преобразованиями строк приводится к единичной матрице E_n;*
(2) $|A| \neq 0$.

Доказательство. Заетим, что если матрица A приводится элементарными преобразованиями строк к ступенчатой матрице B, то определители $|A|, |B|$ отличаются на ненулевой множитель.

Пусть выполнено первое условие. Тогда $|A|$ получается из определителя единичной матрицы, равного 1, умножением на ненулевой число. Поэтому второе условие выполнено.

Обратно, пусть выполнено второе условие. Тогда матрица A приводится к ступенчатой квадратной матрице B, определитель которой отличен от нуля. Поэтому $B = E$. □

2. Обратная матрица. Матричные уравнения

Определение 3.19. Пусть $A \in \mathrm{Mat}(n)$. Матрица $A^{-1} \in \mathrm{Mat}(n)$ называется *обратной* к A, если $AA^{-1} = A^{-1}A = E_n$.

Предложение 3.20. *Если $i \neq j$, $\beta \neq 0$, то $(E + \alpha E_{ij})^{-1} = E - \lambda E_{ij}$, $D_i(\beta)^{-1} = D_i(\beta^{-1})$.*

Доказательство. Нужно воспользоваться предложением 1.27. □

Предложение 3.21. *Если A^{-1} существует, то она единственна.*

Доказательство. Пусть заданы две обратные B, C к A. Тогда $AC = E = BA$, откуда $B = B(AC) = (BA)C = .$ □

Теорема 3.22. *Пусть* $A \in \mathrm{Mat}(n)$. *Обратная матрица* A^{-1} *существует тогда и только тогда, когда* $\det A \neq 0$.

Доказательство. Пусть матрица A^{-1} существует. В силу определения 3.19 и теоремы 3.12 получаем $1 = \det A \det(A^{-1})$. Поэтому $\det A \neq 0$.

Обратно, пусть $\det A \neq 0$. Рассмотрим матрицу

$$B = \frac{1}{|A|}\hat{A} = (b_{ij}), \text{ где } b_{ij} = \frac{A_{ji}}{\det A}. \tag{19}$$

По теореме 3.17 $BA = AB = E$, т. е. $B = A^{-1}$. □

Теорема 3.23 (Теорема Крамера). *Квадратная система линейных уравнений* $AX = b$ *с матрицей* A *определенна тогда и только тогда, когда* $\det A \neq 0$. *Если* $\det A \neq 0$, *то решение находится по формуле*

$$x_i = \frac{\det A'_i}{\det A}, \tag{20}$$

где матрица A'_i *получается из* A *заменой* i*-го столбца на* b.

Доказательство. Если система $AX = b$ определенна, то все ее неизвестные главные. В этом случае ступенчатый вид A является единичной матрицей. Поэтому $\det A \neq 0$ в силу теореме 3.18.

Обратно, пусть $\det A \neq 0$. По теореме 3.18 ступенчатый вид матрицы A является единичной матрицей. Поэтому система определенна.

Пусть $\det A \neq 0$. По теореме 3.22 существует A^{-1}. Умножим уравнение на A^{-1} и получим $X = A^{-1}b$. Подставляя (19) и пользуясь разложением A'_i по i-му столбцу, завершаем доказательство. □

Формула (19) позволяет вычислять элементы A^{-1}. Укажем другой способ вычисления A^{-1}. Для этого нужно решить матричное уравнение $AX = E_n$. Рассмотрим более общий случай матричного уравнения $AX = B$, где $A \in \mathrm{Mat}(n)$, $\det A \neq 0$, и $X, B \in \mathrm{Mat}(n \times m)$.

Теорема 3.24. *Составим расширенную матрицу* $(A|B)$ *и приведем ее элементарными преобразованиями к ступенчатому виду* $(E|C)$. *Тогда* $C = X$.

Доказательство. По теоремам 1.29, 1.32 мы умножаем уравнение $AX = B$ на некоторые элементарные обратимые матрицы (см. предложение 3.20) $Z_1 \cdots Z_k$. Тем самым мы приходим к уравнению $Z_1 \cdots Z_k AX = Z_1 \cdots Z_k B$, причем по условию $Z_1 \cdots Z_k A = E$. Отсюда $X = Z_1 \cdots Z_k B = C$. □

Заметим, что в условии теоремы 3.24 решение уравнения $AX = B$ единственно, так как оно имеет вид $X = A^{-1}B$.

ГЛАВА 4

Линейные пространства. Ранг матрицы и его приложения

1. Линейные пространства

Определение 4.1. Множество V называется *линейным (векторным) пространством*, если в V определена операция сложения $x+y$ элементов из V, называемых *векторами*, и операция αx умножения вектора x на число α При этом выполнены следующие аксиомы:

(1) сложение ассоциативно, т. е. $(x+y)+z=x+(y+z)$ для всех $x,y,z\in V$;
(2) сложение коммутативно, т. е. $x+y=y+x$ для всех $x,y\in V$;
(3) существует такой элемент $0\in V$, называемый *нулем*, что $x+0=x$ для всех $x\in V$;
(4) для любого $x\in V$ существует такой элемент $-x\in V$, называемый *противоположным* к x, что $x+(-x)=0$;
(5) если α,β – числа и $x\in V$, то $(\alpha\beta)x=\alpha(\beta x)$;
(6) если α,β – числа и $x\in V$, то $(\alpha+\beta)x=\alpha x+\beta x$;
(7) если α – число и $x,y\in V$, то $\alpha(x+y)=\alpha x+\alpha y$;
(8) если $x\in V$, то $1x=x$.

Примеры 4.2. Примерами векторных пространств являются $\mathrm{Mat}(n\times m)$, векторы плоскости $\mathbb{R}^2$, пространства $\mathbb{R}^3$, все функции на отрезке $[a,b]$ и т. д.

Предложение 4.3. *Справедливы следующие утверждения:*

(1) *нулевой элемент единствен;*
(2) *противоположный элемент определен однозначно;*
(3) $0x=\alpha 0=0$*;*
(4) $(-1)x=-x$.

Доказательство. Проверим третье утверждение. Имеем $0x=(0+0)x=0x+0x$. Отсюда

$$0=0x+(-0x)=(0x+0x)+(-0x)=0x+(0x+(-0x))=0x+0=0x.$$

□

Следствие 4.4. $(\alpha-\beta)x=\alpha x-\beta x, \alpha(x-y)=\alpha x-\alpha y$.

Определение 4.5. Система векторов $x_1,\dots,x_n$ линейного пространства *линейно зависима*, если существует такой ненулевой набор чисел $\alpha_1,\dots,\alpha_n$ (т. е. не все эти числа равны нулю), что

$$\alpha_1x_1+\dots+\alpha_nx_n=0. \tag{21}$$

Система векторов $x_1,\dots,x_n$ линейного пространства *линейно независима*, если для любых чисел $\alpha_1,\dots,\alpha_n$ из (21) вытекает, что

$$\alpha_1=\dots=\alpha_n=0.$$

Замечание 4.6. Любая конечная система векторов линейного пространства либо линейно независима, либо линейно зависима.

Определение 4.7. Вектор x является *линейно комбинацией* векторов $x_1, \dots, x_n$ если существует такой набор чисел $\alpha_1, \dots, \alpha_n$, что

$$x = \alpha_1 x_1 + \cdots + \alpha_n x_n.$$

Линейной оболочкой $\langle x_1, \dots, x_n \rangle$ системы векторов $x_1, \dots, x_n$ называется множество всех линейных комбинаций этой системы векторов.

Предложение 4.8. *Если в системе векторов есть нулевой вектор, то система зависима. Если подсистема в системе линейно зависима, то и вся система зависима. Система зависима тогда и только тогда, когда один из ее векторов является линейно комбинацией остальных.*

Теорема 4.9 (Критерий равенства определителя нулю) . *Определитель квадратной матрицы A равен нулю тогда и только тогда, когда ее строки (столбцы) линейно зависимы.*

Доказательство. Пусть строки A зависимы, и, например, последняя строка A является линейной комбинацией предыдущих. Вычтем из последней строки эту линейную комбинацию. Тогда определитель не меняется, но, с другой стороны, он равен 0.

Обратно, пусть $\det A = \det {}^tA = 0$. Система линейных однородных уравнений с матрицей ${}^tA$ совместна и не имеет единственного решения по теореме Крамера 3.23. Пусть

$$\Lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

– ненулевое решение этой системы, т. е. $\Lambda A = 0$. Это дает зависимость строк A с коэффициентами $\lambda_1, \dots, \lambda_n$. □

Теорема 4.10 (Основная лемма о линейной зависимости) . *Пусть даны две системы векторов $a_1, \dots, a_n$ и $b_1, \dots, b_m$, причем каждый вектор $a_i \in \langle b_1, \dots, b_m \rangle$. Если $n > m$, то система $a_1, \dots, a_n$ линейно зависима.*

Доказательство. По условию

$$\begin{array}{rcccccc} a_1 & = & b_1\alpha_{11} & + & \cdots & + & b_m\alpha_{m1} \\ \dots\dots & & \dots\dots & & \dots\dots & & \dots\dots \\ a_n & = & b_1\alpha_{1n} & + & \cdots & + & b_m\alpha_{mn} \end{array} \tag{22}$$

Рассмотрим однородную систему линейных уравнений

$$\begin{cases} \alpha_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & 0 \\ \dots\dots & & \dots\dots & & \dots\dots & & \dots\dots \\ \alpha_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & 0 \end{cases} \tag{23}$$

По предложению 1.13 система (23) имеет ненулевое решение $c_1, \dots, c_n$. Из (22) и (23) следует, что

$$c_1a_1 + \cdots + c_na_n = 0.$$

□

Следствие 4.11. *Пусть даны две системы векторов $a_1, \dots, a_n$ и $b_1, \dots, b_m$, причем каждый вектор $a_i \in \langle b_1, \dots, b_m \rangle$. Если система $a_1, \dots, a_n$ независима, то $n \le m$.*

Следствие 4.12. *Пусть даны две независимые системы векторов*

$$a_1, \dots, a_n \text{ и } b_1, \dots, b_m,$$

причем каждый вектор

$$\langle a_1, \dots, a_n \rangle = \langle b_1, \dots, b_m \rangle.$$

Тогда $n = m$.

Определение 4.13. Система векторов $a_1, \dots, a_n$ в линейном пространстве L называется *базисом (базой)* L, если

(1) система $a_1, \dots, a_n$ независима;
(2) $L = \langle a_1, \dots, a_n \rangle$.

Размерностью $\dim V$ пространства V называется число векторов в базисе V. Векторное пространство конечномерно, если его размерность конечна.

Замечание 4.14. В силу следствия 4.12 число векторов в любом базисе постоянно, и поэтому размерность пространства определена однозначно.

Упражнение 4.15. Доказать, что матричные единицы E_{ij} образуют базис $\mathrm{Mat}(n \times m)$.

Теорема 4.16. *Любую независимую систему векторов конечномерного пространства можно дополнить до базиса.*

Доказательство. Пусть $\mathbf{e} = (e_1, \dots, e_n)$ – базис векторного пространства V, Предположим, что система векторов $a_1, \dots, a_k$ линейно независима. Рассмотрим система векторов $a_1, \dots, a_k, e_1$. Если она зависима, то $e_1 \in \langle a_1, \dots, a_n \rangle$. В этом случае переходим к рассмотрению вектора e_2. Если же эта система независима, рассматриваем систему векторов $a_1, \dots, a_k, e_1$, и т. д. В результате получаем такую независимую систему векторов

$$a_1, \dots, a_n, e_{i_1}, \dots, e_{i_s},$$

линейная оболочка L которой содержит базис $\mathbf{e}$. Отсюда $L = V$, и построенная система является базисом. □

Определение 4.17. Пусть $\mathbf{e} = (e_1, \dots, e_n)$ – базис векторного пространства V, и $x \in V$. Тогда

$$x = x_1 e_1 + \dots + x_n e_n = \mathbf{e}X, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

– разложение вектора x по базису $\mathbf{e}$. Набор X называется *системой (столбцом) координат* вектора x в базисе $\mathbf{e}$.

Предложение 4.18. *Система координат вектора в базисе определена однозначно.*

Предположим, что в векторном пространстве L выбраны два базиса $\mathbf{e} = (e_1, \dots, e_n)$ и $\mathbf{e}' = (e'_1, \dots, e'_n)$. Тогда для любых $i, j = 1, \dots, n$ имеем

$$e_i = e'_1 c_{1i} + \dots + e'_n c_{ni},$$
$$e'_j = e_1 c'_{1j} + \dots + e_n c'_{nj}.$$

или

$$\mathbf{e} = \mathbf{e}'C, \quad \mathbf{e}' = \mathbf{e}C',$$
$$\text{где } C = (c_{ij}), C' = (c'_{ij}) \in \mathrm{Mat}(n) \tag{24}$$

Определение 4.19. Матрица C (C') из (24) называется *матрицей перехода* от $\mathbf{e}$ к $\mathbf{e}'$ ($\mathbf{e}'$ к $\mathbf{e}$).

Из (24) вытекает, что $\mathbf{e} = \mathbf{e}CC'$, и поэтому $CC' = E_n$. Аналогично, $C'C = E_n$. Поэтому $C' = C^{-1}$. Обратно, если у матрицы C есть обратная, и $\mathbf{e}' = \mathbf{e}C$, то $\mathbf{e}'$ – базис пространства. Поэтому справедливо

Предложение 4.20. *Матрицами перехода от одного базиса к другому являются обратимые матрицы и только они.*

Предложение 4.21. *Пусть* $\mathbf{e}, \mathbf{e}'$ *– два базиса пространства* L*, и вектор* $x \in L$ *имеет в этих базисах столбцы координат* X, X'*, соответственно. Если* C *– матрица перехода от* $\mathbf{e}$ *к* $\mathbf{e}'$*, то* $X = CX'$.

Доказательство. В силу определения 4.17 и (24)

$$x = \mathbf{e}'X' = \mathbf{e}CX' = \mathbf{e}X,$$

откуда вытекает утверждение в силу единственности разложения по базису. □

Определение 4.22. Непустое подмножество L в линейном пространстве V называется *подпространством*, если из того, что $x, y \in L$ следует, что $x + y, \alpha x \in L$.

Упражнение 4.23. Если L – подпространство в V, то $0 \in L$. Если $x \in L$, то $-x \in L$.

Пример 4.24. Рассмотрим однородную систему линейных уравнений $AX = 0$ с матрицей $A \in \mathrm{Mat}(m \times n)$. Тогда все ее решения образуют подпространство в пространстве столбцов $\mathrm{Mat}(n \times 1)$.

Теорема 4.25. *Пусть* L *– подпространство конечномерного пространства* V*. Тогда* $\dim L \leq \dim V$*. Если* $\dim L = \dim V$*, то* $L = V$.

Доказательство. Пусть $\mathbf{e}$ – базис L, и $\mathbf{e}'$ – базис V. По следствию 4.11 число векторов в $\mathbf{e}$ $(= \dim L)$ не превосходит числа векторов в $\mathbf{e}'$ $(= \dim L)$.

Если $\dim L = \dim V$, то, присоединяя к $\mathbf{e}$ любой вектор из $\mathbf{e}'$, получаем зависимую систему. Поэтому любой вектор из $\mathbf{e}'$ лежит в линейной оболочке $\mathbf{e}$, т. е. в L. Отсюда $L = V$. □

Определение 4.26. Пусть заданы векторы $x_1, \ldots, x_k$. *Рангом* этой системы называется максимальное число линейно независимых векторов этой системы.

Упражнение 4.27. Ранг системы векторов $a_1, \ldots, a_k$ равен $\dim\langle a_1, \ldots, a_k\rangle$.

Предположим, что задана система векторов $a_1, \ldots, a_m$ в векторном пространстве V с базисом $\mathbf{e} = (e_1, \ldots, e_n)$. Пусть задано разложение каждого вектора a_i по базису $\mathbf{e}$,

$$a_i = e_1 a_{1i} + \cdots + e_n a_{ni}.$$

Положим $A = (a_{rs})\,\mathrm{Mat}(n \times m)$. Тогда

$$(a_1, \ldots, a_m) = \mathbf{e}A. \tag{25}$$

Опишем алгоритм решения следующей задачи:

- -¥®¡å®¤¨¬® - ©â¨ ¬ ªá¨¬ «ì-ãî -¥§ ¢¨á¨¬ãî ¯®¤á¨áâ¥¬ã á¨áâ¥¬ë ¢¥ªâ®à®¢

$$a_1, \ldots, a_m$$

¨ ¢ëà §¨âì ®áâ «ì-ë¥ ¢¥ªâ®àë á¨áâ¥¬ë ç¥à¥§ ¢ë¡à --ãî ¯®¤á¨áâ¥¬ã .

Для решения этой задачи приведем матрицу A элементарными преобразованиями к ступенчатому виду B. Пусть, например,

$$B = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & b_{1,r+1} & \ldots & b_{1m} \\ 0 & 1 & 0 & \ldots & 0 & 0 & b_{2,r+1} & \ldots & b_{2m} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 & b_{r,r+1} & \ldots & b_{rm} \end{pmatrix} \tag{26}$$

Заметим, что системы однородных уравнений с матрицами A и B эквивалентны. Решениями этих систем являются такие столбцы

$$\Lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix},$$

что $A\Lambda = B\Lambda = 0$. В силу (25) это эквивалентно тому, что

$$(a_1, \dots, a_n)\Lambda = 0. \tag{27}$$

Из вида B вытекает, что первые r векторов $a_1, \dots, a_r$ независимы. Кроме того, для любого $i = r+1, \dots, m$ имеем

$$B \begin{pmatrix} b_{1,i} \\ \vdots \\ b_{r,i} \\ 0 \\ \vdots \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0, \tag{28}$$

где -1 расположена на i-ом месте. Из (27) вытекает, что

$$a_i = a_1 b_{1,i} + \dots + a_r b_{r,i}.$$

Тем самым поставленная выше задача решена. Отметим, что в силу упражнения 4.27 приведенный алгоритм позволяет находить базис линейной оболочки системы векторов.

2. Ранг матрицы

Определение 4.28. *Рангом* матрицы A называется максимальное число линейно независимых строк A. Другими словами, ранг матрицы – это ранг ее системы строк или размерность линейной оболочки строк A.

Теорема 4.29. *Ранг матрицы не меняется при элементарных преобразованиях строк и столбцов.*

Доказательство. При элементарных преобразованиях строк линейная оболочка системы строк не меняется. Следовательно, не меняется и ее размерность.

Предположим, что мы совершаем элементарные преобразования столбцов матрицы A. По следствию 1.30 и теореме 1.32 матрица A заменяется на матрицу AM, где M – обратимая матрица. Если имеется линейное соотношение между строками A с коэффициентами $\lambda_1, \dots, \lambda_n$, то

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} A = 0, \quad \text{откуда} \quad \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} (AM) = 0,$$

и наоборот. Отсюда следует, что ранги системы строк A и AM совпадают. □

Теорема 4.30 (Теорема о ранге матрицы) . *Ранг системы строк матрицы совпадает с рангом системы столбцов.*

Доказательство. По теореме 4.29 можно считать, что матрица B имеет ступенчатый вид, например, (26). Тогда в ней первые r строк и столбцов образуют максимальные независимые системы. □

Определение 4.31. Пусть в матрице A выделены r строк и столбцов. На их пересечении строит квадратная матрица размера r. Ее определитель называется *минором* M порядка r матрицы A. Любой минор порядка $r+1$, получающийся присоединением еще одной строки и столбца A, называется *окаймляющим* для M.

Теорема 4.32 (Теорема об окаймляющем миноре) . *Ранг матрицы равен порядку ненулевого минора, все окаймляющие которого равны 0.*

Доказательство. Пусть для простоты указанный минор M размера r расположен в верхнем левом углу матрицы A. Присоединим к M i-ую строку и j-ый столбец. Получающийся окаймляющий минор по условию всегда равен 0 (включая случаи, когда либо $i < r$, либо $j < r$. Разложим этот минор по присоединенное столбцу

$$0 = a_{1j}A_{1,r+1} + \cdots + a_{rj}A_{r,r+1} + a_{ij}M.$$

Так как $M \neq 0$, то

$$a_{ij} = a_{1j}(-\frac{A_{1,r+1}}{M}) + \cdots + a_{rj}(-\frac{A_{r,r+1}}{M}). \tag{29}$$

В (29) коэффициенты

$$-\frac{A_{s,r+1}}{M}, \quad s = 1, \dots, r,$$

не зависят от j. Поэтому объединяя равенства (29) для всех j, получаем, что i-ая строка является линейной комбинацией первых r строк A. По теореме 4.9 первые r строк A независимы. □

В примере 4.24 отмечено, что все решения однородной системы $AX = 0$ образует подпространство в пространстве всех столбцов. Найдем его размерность.

Теорема 4.33. *Размерность пространства решений однородной системы $AX = 0$, где с n неизвестными равна $n - r(A)$, где $r(A)$ – ранг матрицы A.*

Доказательство. В силу теорем 1.7, 4.29 можно считать, что матрица A приведена в ступенчатому виду B из (26), где $r = r(A)$ – ранг матрицы A. Для $i = r+1$ положим

$$e_i = \begin{pmatrix} b_{1,i} \\ \vdots \\ b_{r,i} \\ 0 \\ \vdots \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0,$$

где -1 стоит на i-ом месте. Как уже отмечалось в (28) столбец e_i является решением системы $Ae_i = Be_i = 0$.

Пусть

$$a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

– произвольное решение системы. Тогда

$$b = a - a_{r+1}e_{r+1} - \cdots - a_n e_n$$

также является решением, у которого все значения свободных переменных нулевые. Из (26) видно, что вектор $b = 0$. Итак, любое решение системы является линейной комбинацией $e_{r+1}, \ldots, e_n$. Несложная проверка показывает, что они независимы. □

Теорема 4.34 (Теорема Кронекера-Капелли) . *Система линейных уравнений $AX = b$ совместна тогда и только тогда, когда ранги матриц A и $(A|b)$ совпадают.*

Доказательство. Пусть столбец Λ является решением системы. Тогда $A\Lambda = b$, т. е. линейные оболочки столбцов A и $(A|b)$ совпадают. Отсюда вытекает совпадение рангов этих матриц.

Обратно, пусть ранги матриц A и $(A|b)$ совпадают. Линейная оболочка столбцов A содержится в линейной оболочке $(A|b)$, причем их размерности совпадают. По теореме 4.25 эти оболочки совпадают. Поэтому b лежит в линейно оболочке столбцов A. □

Теорема 4.35. *Ранг произведения матриц не превосходит ранга множителей. Если один из множителей является обратимой матрицей, то ранг произведения равен другому множителю.*

Доказательство. Пусть $A, C, D = AC$ из определения 1.16. Тогда каждый столбец D является линейно комбинацией столбцов A, а каждая строка D является линейно комбинацией строк C. По теореме 4.10 (см. также теорему 4.25) получаем первое утверждение.

Пусть, например, матрица A обратима. по доказанному $r(D) \leq r(C)$. Кроме того, $r(C) = r(A^{-1}D) \leq r(D)$. Отсюда $r(C) = r(D)$. □

ГЛАВА 5

Комплексные числа

1. Действия с комплексными числами

Определение 5.1. *Комлексными числами* называются вещественные матрицы

$$z = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad a, b \in \mathbb{R}. \tag{30}$$

Через $\mathbb{C}$ обозначается множество всех комплексных чисел. *Модулем* числа z из (30) называется неотрицательное вещественное число $|z| = \sqrt{\det z}$.

Предложение 5.2. *Множество $\mathbb{C}$ замкнуто относительно сложения и умножения матриц. Модуль произведения комплексных чисел равен произведению модулей. Если $z \neq 0$, то $|z| > 0$.*

Доказательство. Пусть z из (30) и

$$z_1 = \begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix} \in \mathbb{C}.$$

Тогда

$$z + z_1 = \begin{pmatrix} a + a_1 & -(b + b_1) \\ b + b_1 & a + a_1 \end{pmatrix} \in \mathbb{C}$$

$$zz_1 = \begin{pmatrix} aa_1 - bb_1 & -(ab_1 + ba_1) \\ ab_1 + ba_1 & aa_1 - bb_1 \end{pmatrix} \in \mathbb{C}. \tag{31}$$

□

Предложение 5.3. *Умножение и сложение комплексных чисел коммутативно, ассоциативно, обладает свойствами дистрибутивности. В $\mathbb{C}$ содержатся нулевой и единичный элементы. В $\mathbb{C}$ для каждого элемента z имеется противоположный $-z$. Кроме того, у каждого ненулевого элемента имеется обратный.*

Доказательство. Достаточно доказать лишь последнее утверждение. Пусть z из (30). Тогда либо a, либо b отлично от нуля. Поэтому

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

□

Предложение 5.4. *$\mathbb{C}$ является векторным пространство над $\mathbb{R}$ с базой*

$$E = E_2, \quad i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

При этом $i^2 = -E$.

Всюду в дальнейшем мы будем отождествлять вещественное число a с матрицей aE. Ясно, что это отождествление согласовано с суммами и произведениями вещественных чисел. Тогда число z из (30) можно однозначно записать в виде $z = a + ib$ (*алгебраическая форма числа z*.

Определение 5.5. Пусть $z = a + bi$. Тогда комплексное число $\overline{z} = a - bi$ называется *комлексно сопряженным* к z.

Замечание 5.6. Если z имеет представление (30), то $\overline{z}$ – это транспонированная матрица. Поэтому $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$. Кроме того, $z\overline{z} = |z|^2$, и поэтому $z^{-1} = \dfrac{\overline{z}}{|z^2|}$.

2. Тригонометрическая форма комплексного числа

Представим комплексное число $z = a + bi$ в виде вектора на плоскости $\mathbb{R}^2$ с координатами a, b. Тогда $|z|$ – это длина этого вектора z.

Упражнение 5.7. Доказать, что $|z_1 + z_2| \le |z_1| + |z_2|$, и $|z_1 + z_2| \ge |z_1| - |z_2|$ для любых $z_1, z_2 \in \mathbb{C}$.

Определение 5.8. *Аргументом* ненулевого комплексного числа z называется угол ϕ между z и положительным направлением оси OX.

Аргумент ненулевого комплексного числа определен с точностью до слагаемого $2\pi n, n \in \mathbb{Z}$. Если $z = a + bi \neq 0$, то

$$\cos\phi = \frac{x}{|z|}, \quad \sin\phi = \frac{y}{|z|}.$$

Отсюда получаем *тригонометрическую форму комплексного числа*

$$z = |z|(\cos\phi + i\sin\phi) \tag{32}$$

Предложение 5.9. *Аргумент произведения комплексных чисел равен сумме аргументов множителей.*

Доказательство. Пусть z из (32) и $z_1 = |z_1|(\cos\psi + i\sin\psi)$. По формулам приведения

$$zz_1 = |z||z_1|(\cos(\phi + \psi) + i\sin(\phi + \psi)).$$

□

Обозначение 5.10. Предложение 5.9 позволяет ввести следующее обозначение. Если $a \in \mathbb{R}$, то

$$\exp(i\alpha) = \cos\alpha + i\sin\alpha.$$

Таким образом, тригонометрическая форма комплексного числа z имеет вид

$$z = |z|\exp(i\phi). \tag{33}$$

Следствие 5.11 (Формула Моавра). *Пусть z из (33) и $n \in \mathbb{Z}$. Тогда*

$$z^n = |z|^n \exp(ni\phi).$$

Рассмотрим вопрос об извлечении комплексных корней.

Определение 5.12. Пусть $z \in \mathbb{C}$ и $n \in \mathbb{Z}$. *Корнем* n-ой степени из z называется такое комплексное число t, что $t^n = z$.

Найдем все корни степени n из z. Пусть z из (33), и $t = |t|\exp(i\psi)$. Тогда

$$z = t^n = |t|^n \exp(ni\psi).$$

Отсюда $|z| = |t|^n$, т. е. $|t| = \sqrt[n]{|z|}$. Кроме того, $ni\psi \equiv \phi \pmod{2\pi m}, m \in \mathbb{Z}$. Следовательно,

$$\psi = \frac{\phi + 2\pi m}{n}, \quad m = 0, 1, \ldots, n-1.$$

Итак,

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

Поэтому, если $z \neq 0$, то имеется n различных корней степени n из 1. В частности, если $z = 1$, то m-ый корень из 1 имеет вид

$$\varepsilon_m = \left(\cos\frac{2\pi m}{n} + i\sin\frac{2\pi m}{n}\right).$$

Все эти корни расположены в вершинах правильного n-угольника, вписанного в единичную окружность с центром в 0.

ГЛАВА 6

Группы, кольца и поля

В этой главе изучаются основные понятия теории групп.

1. Группы, подгруппы, порядки элементов

Напомним некоторые необходимые определения.

Определение 6.1. Множество G с бинарной операцией умножения xy называется *группой*, если

(1) умножение ассоциативно, т. е. $(xy)z = x(yz)$ для всех $x, y, z \in G$;
(2) существует такой элемент $1 \in G$, называемый *единицей* G, что $x1 = 1x = x$ для всех $x \in G$;
(3) для любого элемента $x \in G$ найдется такой элемент x^{-1}, называемый *обратным* к x, что $xx^{-1} = x^{-1}x = 1$.

Определение 6.2. *Порядком* группы G называется число $|G|$ элементов в G.

Предложение 6.3. *Единичный элемент в группе единственен. Для каждого элемента $x \in G$ обратный элемент x^{-1} определен однозначно. Кроме того, если $x \in G$, то $(x^{-1})^{-1} = x$.*

Определение 6.4. Непустое подмножество H в группе G называется *подгруппой*, если вместе с любыми двумя его элементами оно содержит их произведение, и с каждым своим элементом H содержит его обратный.

Предложение 6.5. *Если H – подгруппа в группе G и 1 – единичный элемент G, то $1 \in H$.*

Упражнение 6.6. В произвольной группе произведение любого числа элементов не зависит от расстановки скобок.

Предложение 6.7. *Для непустого подмножества H в группе G следующие условия эквивалентны:*

(1) *H является подгруппой в G;*
(2) *если $x, y \in H$, то $xy^{-1} \in H$.*

Доказательство. Пусть выполнено условие (1), и $x, y \in H$. В силу определения 6.4 получаем $x, y^{-1} \in H$, откуда $xy^{-1} \in H$, т. е. выполнено условие (2).

Обратно, пусть выполнено условие (2), и $y \in H$. Тогда $y, y \in H$, откуда $1 = yy^{-1} \in H$ по (2). Далее $1, y \in H$, откуда $y^{-1} = 1y^{-1} \in H$ по (2). Наконец, если $x, y \in H$, то $x, y^{-1} \in H$ по доказанному выше. Отсюда $x(y^{-1})^{-1} = xy \in H$ по предложению 6.3. □

Примеры 6.8. Приведем примеры групп и их подгрупп:

(1) группа S_n содержит подгруппы A_n, S_{n-1};
(2) группа $GL(n, \mathbb{C})$ содержит подгруппы
$$GL(n, \mathbb{R}), \quad GL(n, \mathbb{Q}), \quad SL(n, \mathbb{C}), \quad SL(n, \mathbb{R});$$
(3) группа * содержит подгруппу $U_n = \{z \in \mathbb{C} | z^n = 1\}$.

Упражнение 6.9. Пусть

$$I = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, K = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \in SL(2, \mathbb{C}).$$

Доказать, что

(1) $I^2 = J^2 = K^2 = -E, \quad IJ = K, JK = I, KI = J, JI = -K, KJ = -I, IK = -J$;
(2) 8 матриц $\pm E, \pm I, \pm J \pm K$ образуют подгруппу кватернионов Q_8 в группе $SL(2, \mathbb{C})$.

Упражнение 6.10. Если $H_i, i \in I$ – подгруппы группы G, то $\cap_{i \in I} H_i$ – подгруппа группы G.

Определение 6.11. Пусть $\mathbb{Z}_n = \{0, 1, \dots, n-1\}$. Тогда $\mathbb{Z}_n$ – группа. Она называется *группой вычетов по модулю* n.

Предложение 6.12. *Определение группы* $\mathbb{Z}_n$ *корректно.*

Определение 6.13. Пусть a – элемент группы G. Для произвольного целого числа n положим

$$a^n = \begin{cases} 1, & \text{если } n = 0; \\ \underbrace{a \cdots a}_{n \text{ множителей}}, & \text{если } n > 0; \\ (a^{-n})^{-1}, & \text{если } n < 0. \end{cases}$$

Предложение 6.14. *Пусть* a *– элемент некоторой группы и* $n, m \in \mathbb{Z}$. *Тогда*

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

Определение 6.15. Пусть a – элемент некоторой группы. *Порядком* $|a|$ (или $o(a)$) элемента a называется такое наименьшее натуральное число n, что $a^n = 1$. Если такого числа n нет, то говорят, что порядок a равен бесконечности.

Предложение 6.16. *Пусть* $|a| = n < \infty$, *и* $m \in \mathbb{Z}$. *Следующие условия эквивалентны:*

(1) $n|m$ (n *делит* m);
(2) $a^m = 1$.

Определение 6.17. Пусть $a \in G$. Через $\langle a \rangle$ обозначим множество $\{a^n | n \in Z\}$ всех степеней элемента a.

Упражнение 6.18. $\langle a \rangle$ является подгруппой в G.

Определение 6.19. Пусть $a \in G$. Подгруппа $\langle a \rangle$ называется *циклической* подгруппой в группе G, порожденной элементом a. Группа G называется *циклической* с *порождающим (образующим)* элементом a, если $\langle a \rangle = G$.

Примеры 6.20. Доказать, что

(1) группа $\mathbb{Z}$ является циклическая с порождающим элементом 1 (или -1);
(2) группа U_n комплексных корней n-ой степени из 1 является циклической группой с порождающим элементом
$$\exp \frac{2\pi i}{n} = \cos \frac{2\pi i}{n} + i \sin \frac{2\pi i}{n};$$
(3) группа $\mathbb{Z}_n$ вычетов по модулю n является циклической группой с порождающим элементом 1.

Предложение 6.21. *Пусть* a *– элемент некоторой группы. Тогда* $|\langle a \rangle| = |a|$.

Доказательство. Если $a^r = a^m$ при некоторых $r < m$, то

$$a^{m-r} = 1, \text{ и } |a| = n < \infty.$$

В этом случае

$$\langle a \rangle = \{1, a, a^2, \dots, a^{n-1}\}.$$

□

Обозначение 6.22. Если $|a| = n$ в условии предложения 6.21, то циклическую группу, порожденную элементом a, мы будем обозначать $\langle a \rangle_n$.

Теорема 6.23. *Подгруппа циклической группы сама является циклической.*

Доказательство. Пусть H – подгруппа циклической группы $G = \langle a \rangle$. Если $H = 1$, то утверждение очевидно. Пусть H содержит неединичный элемент $a^m, m \neq 0$. Если $m < 0$, то H содержит и элемент $a^{-m}, -m > 0$. Выберем такое наименьшее натуральное число m, что $b = a^m \in H$. Если $a^r \in H, r \in \mathbb{Z}$, то, деля r с остатком на m, получаем $r = sm + q, 0 \leq q < m$. При этом по предложению 6.14

$$a^q = a^{r-sm} = a^r (a^m)^{-s} \in H,$$

что противоречит выбору m, если $q > 0$. □

Следствие 6.24. *Пусть $m_1, \dots, m_n \in Z$, и d – чисел $m_1, \dots, m_n$. Тогда существуют такие целые числа $u_1, \dots, u_n \in Z$, что $m_1 u_1 + \dots + m_n u_n = d$.*

Доказательство. Пусть $H = \mathbb{Z}m_1 + \dots + \mathbb{Z}m_n$. Тогда H – подгруппа в $\mathbb{Z}$, и, следовательно, $H = \mathbb{Z}d$. Остается убедиться, что $d = (m_1, \dots, m_n)$. □

Теорема 6.25. *Пусть $G = \langle a \rangle_n$ и H – подгруппа в G. Тогда существует и притом единственное такое число d, делящее n, что $H = \langle a^d \rangle_{\frac{n}{d}}$.*

Доказательство. По теореме 6.23 получаем $H = \langle a^k \rangle$ для некоторого $0 \leq k < n$. Положим $d = (n, k)$. Остается заметить, что $H = \langle a^d \rangle$. □

Следствие 6.26. *Пусть $a \in G$ имеет порядок n. Тогда $|a^k| = \dfrac{n}{(n,k)}$.*

Доказательство. Можно считать, что $G = \langle a \rangle$. Рассмотрим подгруппу $H = \langle a^k \rangle \subseteq G$. По теореме 6.25 $H = \langle a^d \rangle$, где $d = (n, k)$. Отсюда $|a^k| = |H| = \frac{n}{d}$. □

Следствие 6.27. *Пусть $G = \langle a \rangle_n$. Элемент a^k является порождающим в G тогда и только тогда, когда $(k, n) = 1$.*

Упражнение 6.28. Описать все подгруппы в $\langle a \rangle_{12}$.

Определение 6.29. Биективное отображение групп $f : G \to H$ называется *изоморфизмом*, если $f(xy) = f(x)f(y)$ для всех $x, y \in G$. Обозначение $\simeq$.

Пример 6.30. $(\mathbb{R}, +) \simeq (\mathbb{R}_{>0}, \cdot)$. В качестве f взять exp.

Предложение 6.31. *Циклическая группа порядка n изоморфна U_n. Бесконечная циклическая группа изоморфна $\mathbb{Z}$.*

Доказательство. Пусть $G = \langle a \rangle_n$. Зададим $f : G \to U_n$, полагая

$$f(a^k) = \exp(\frac{2\pi i}{n}).$$

Если $G = \langle a \rangle_\infty$, то определим $f : G \to \mathbb{Z}$, полагая $f(a^k) = k$. □

Следствие 6.32. $\mathbb{Z}_n \simeq U_n$.

2. Смежные классы и теорема Лагранжа

Определение 6.33. Пусть H – подгруппа в группе G, и $g \in G$. *Левым смежным классом* gH называется подмножество $\{gh|h \in H\}$ в G.

Упражнение 6.34. Найти

(1) левые смежные классы $GL(n, \mathbb{C})$ по $SL(n, \mathbb{C})$;
(2) левые смежные классы $\mathbb{Z}$ по $n\mathbb{Z}$;
(3) левые и правые смежные классы S_n по S_{n-1}.

Упражнение 6.35. Пусть H – подгруппа в группе G и $x, y \in G$. Доказать, что следующие условия эквивалентны:

(1) $xH = yH$;
(2) $x^{-1}y \in H$.

Предложение 6.36. *Пусть H – подгруппа в группе G и $x \in G$. Тогда $|H| = |xH|$.*

Предложение 6.37. *Пусть H – подгруппа в группе G и $x, y \in G$, причем $y \in xH$. Тогда $xH = yH$.*

Доказательство. Ясно, что $yH \subseteq xH$. По условию $y = xh$ для некоторого $h \in H$. Следовательно, для любого $u \in H$ получаем $xu = y(h^{-1}u)$, где $h^{-1}u \in H$. Отсюда $xH \subseteq yH$, т. е. $xH = yH$. □

Следствие 6.38. *Пусть H – подгруппа в группе G. Тогда два левых (правых) смежных класса G по H либо совпадают, либо не пересекаются.*

Доказательство. Воспользоваться предложением 6.37. □

Теорема 6.39 (Теорема Лагранжа). *Пусть H – подгруппа в конечной группе G. Тогда $|G| = |H|j$, где j – число левых (правых) смежных класса G по H.*

Доказательство. Разобъем G на левые смежные классы по H. Тогда каждый элемент $x \in G$ лежит в некотором классе, именно, в xH. Остается воспользоваться следствием 6.38 и предложением 6.36. □

Следствие 6.40. *Порядок элемента конечной группы делит порядок группы.*

Следствие 6.41. *Группа простого порядка является циклической.*

ГЛАВА 7

Кольца и поля

Определение 7.1. *Кольцо* (не обязательно ассоциативное). *Ассоциативные, коммутативные* кольца.

Предложение 7.2. *В любом кольцо имеем* $0x = x0 = 0$.

Примеры 7.3. Укажем ряд колец.

- ♡ Ассоциативные кольца – кольца матриц $\mathrm{Mat}(n, \mathbb{R})$.
- ♡ Ассоциативно-коммутативные кольца – кольца непрерывных функций на топологическом пространстве.

Определение 7.4. *Единичный* элемент, *делители нуля*, *обратимые* элементы алгебры.

Предложение 7.5. *Единичный элемент алгебры определен однозначно. Обратимые элементы ассоциативной алгебры образуют группу по умножению. Обратимый элемент ассоциативной алгебры не может быть делителем нуля.*

Следствие 7.6. *В поле нет делителей нуля.*

Теорема 7.7. *Группы обратимых элементов в* $\mathrm{Mat}(n, k)$ *– это* $GL(n, k)$*; делители нуля в* $\mathrm{Mat}(n, k)$ *– это вырожденные матрицы и только они.*

Определение 7.8. Структура кольца на $\mathbb{Z}_n$.

Предложение 7.9. *Структура кольца на* $\mathbb{Z}_n$ *определена корректно.*

Теорема 7.10. *Элемент* $k \in \mathbb{Z}_n$ *обратим тогда и только тогда, когда* $(k, n) = 1$. *Элемент* $k \in \mathbb{Z}_n$ *является делителем нуля тогда и только тогда, когда* $(k, n) > 1$.

Доказательство. Элемент k обратим тогда и только тогда, когда найдется такое число $u \in \mathbb{Z}_n$, что $ku = 1$, т. е. в $\mathbb{Z}$ $ku + nv = 1$. Это эквивалентно тому, что $(k, n) = 1$.

Пусть элемент $k \in \mathbb{Z}_n$ является делителем нуля. Тогда он не может иметь обратного по предложению 7.5. Следовательно, $(k, n) > 1$.

Обратно, пусть $d = (k, n) > 1$. Тогда $n > m = nd^{-1}$, т. е. $m \neq 0$ в $\mathbb{Z}_n$. При этом $km = k_1 dm = k_1 n = 0$ в $\mathbb{Z}_n$. □

Определение 7.11. Поле.

Примеры 7.12. Поля – $\mathbb{Q}, \mathbb{R}, \mathbb{C}$. Кроме того, $\mathbb{Q}[i], \mathbb{Q}[\sqrt{5}]$.

Определение 7.13. Характеристика поля char.

Теорема 7.14. *Характеристика поля либо равна нулю, либо простое число.*

Теорема 7.15. *Кольцо вычетов* $\mathbb{Z}_n$ *является полем тогда и только тогда, когда* n *– простое число.*

Доказательство. Нужно воспользоваться теоремой 7.10. □

Следствие 7.16. $\mathrm{char}\, \mathbb{Z}_p = p$.

Определение 7.17. Биективное отображение колец $f : R \to R'$ называется *изоморфизмом*, если $f(x+y) = f(x) + f(y), f(xy) = f(x)f(y)$. Обозначение $R \simeq R'$.

Пример 7.18. Доказать, что $\mathrm{Mat}(n, \mathrm{Mat}(m, \mathbb{R})) \simeq \mathrm{Mat}(nm, \mathbb{R})$.

Упражнение 7.19. Доказать, что при изоморфизме $f : R \to R'$

(1) $f(1) = 1$ и обратимые элементы переходят в обратимые;
(2) делители нуля переходят в делители нуля.

Определение 7.20. Подмножество K в кольце с 1 (поле) R называется *подкольцом (подполем)*, если K содержит 1 и из того, что $x, y \in K$ следует, что $x+y, xy \in K$ (в случае полей $x^{-1} \in K$, если $x \neq 0$.

Примеры 7.21. $\mathbb{Z}$ является подкольцом в $\mathbb{Q}$, $\mathbb{Q}$ является подполем в $\mathbb{R}, \mathbb{C}$. $\mathrm{Mat}(n, \mathbb{Z})$ является подкольцом в $\mathrm{Mat}(n, \mathbb{R})$.

Определение 7.22. Кольцо без делителей нуля называется *областью*. Другими словами, кольцо R является областью, если для любых $x, y \in R$ из того, что $xy = 0$, следует, что либо $x = 0$, либо $y = 0$.

Пример 7.23. Областями являются $\mathbb{Z}$, любое поле.

Упражнение 7.24. Пусть a, b, c – элементы из некоторой области, и $a \neq 0$. Если $ab = ac$, то $b = c$.

ГЛАВА 8

Многочлены и ряды от одной переменной

1. Кольцо многочленов от одной переменной

Пусть R – ассоциативное кольцо с 1. Рассмотрим множество $R[X]$ всех почти нулевых последовательностей

$$a = (a_0, a_1, \dots) \tag{34}$$

элементов из R. Определим в $R[X]$ операцию сложения по-координатно. Кроме того, если a из (34),

$$b = (b_0, b_1, \dots),$$

то

$$c = ab = (c_0, c_1, \dots),$$

где для любого $k \geq 0$

$$c_k = \sum_{i=0}^{k} a_i b_{k-i}. \tag{35}$$

Предложение 8.1. *$R[X]$ является ассоциативным кольцом с 1. Если R коммутативно, то и $R[X]$ коммутативно.*

Определение 8.2. $R[X]$ называется *кольцом многочленов*.

Предложение 8.3. *Все элементы $(a_0, 0, \dots) \in R[X]$ образуют подкольцо, изоморфное R.*

Доказательство. Изоморфизм задается по правилу $a \mapsto (a, 0, \dots) \in R[X]$. □

Обозначение 8.4. Всюду в дальнейшем мы будем отождествлять $a \in R$ с $(a, 0, \dots) \in R[X]$. Положим $X = (0, 1, 0, \dots)$.

Предложение 8.5. *Для любого $n \geq 1$*

$$X^n = (\underbrace{0, \dots, 0}_{n}, 1, 0, \dots)$$

Если a из (34), то

$$a = a_0 + a_1 X + \dots + a_n X^n, \quad a_i \in R. \tag{36}$$

Определение 8.6. Пусть $a \in R[X]$ из (36), причем $a_n \neq 0$. *Степенью* a называется $\deg a = n$. a_0 называется *свободным членом* a. a_n называется *старшим членом* a. Многочлены нулевой степени и нулевой многочлен называются *константами*. Многочлены со старшим коэффициентом 1 называются *унитарными*.

Предложение 8.7. *Старший (свободный) член произведения многочленов равен произведению старших (свободных) членов сомножителей. В частности, обратимыми элементами $R[X]$ являются обратимые в R константы.*

Следствие 8.8. *Если R – область, то и $R[X]$ – область. Кроме того, если $f, g \in r[X] \setminus 0$, то $\deg(fg) = \deg f + \deg g$. Если $f, g, f + g \neq 0$, то $\deg(f + g) \leq \max(\deg f, \deg g)$.*

2. Деление многочленов

Обозначение 8.9. Всюду в дальнейшем мы будем предполагать, что R – поле. В этом случае, группа обратимых элементов $R[X]$ совпадает с мультипликативной группой R^* всех ненулевых констант.

Теорема 8.10 (Деление с остатком). *Пусть $f, g \in R[X]$, причем $g \neq 0$. Тогда существуют и притом единственные такие $q, r \in R[X]$, что*

(1) $f = qg + r$;
(2) $r = 0$ *или* $\deg r < \deg g$.

Доказательство. **Существование.**

Если $f = 0$ или $\deg f < \deg g$, то полагаем $q = 0, r = g$. Пусть $\deg f = n$, и для меньших степеней теорема доказана. Предположим, что a_n, b_m – старшие коэффициенты f, g. Тогда степень $h = f - \frac{a_n}{b_m} X^{n-m} g$ меньше n. По индукции $h = q'g + r'$, откуда

$$f = h + \frac{a_n}{b_m} X^{n-m} g = (q' + \frac{a_n}{b_m} X^{n-m})g + r'.$$

Единственность.

Пусть

$$f = qg + r = q'g + r', \quad \deg r, \deg r' < \deg g, \text{ если они ненулевые.}$$

Тогда $g(q - q') = r' - r$. Если $q - q' \neq 0$, то $r - r' \neq 0$. Отсюда по следствию 8.8

$$\deg g > \deg(r - r') = \deg(g(q - q')) = \deg g + \deg(q - q') \geq \deg g.$$

□

Определение 8.11. Многочлен q называется *частным*, а многочлен r *остатком f деления на g*. Многочлен g *делит* f, если $r = 0$. Обозначение $g|f$.

Определение 8.12. *Наибольшим общим делителем* многочленов $f_1, \ldots, f_m$, не все из которых равны нулю, называется такой многочлен d, что

(1) $d|f_i, i = 1, \ldots, m$;
(2) если $d' \in R[X]$, и $d'|f_i, i = 1, \ldots, m$, то $d'|d$.

Наибольший общий делитель многочленов $f_1, \ldots, f_m$, обозначается либо $(f_1, \ldots, f_m)$, либо $(f_1, \ldots, f_m)$.

Предложение 8.13. *$(f_1, \ldots, f_m)$, определен однозначно, с точностью до множителя нулевой степени (ненулевой константы). Кроме того,*

$$(f_1, (f_2, \ldots, f_m)).$$

Изложим *алгоритм Эвклида* нахождения наибольшего общего делителя двух многочленов f, g, $g \neq 0$. Будем делить с остатком.

$$\begin{array}{ll} f = q_1 g + r_1, & \deg r_1 < \deg g; \\ g = q_2 r_1 + r_2, & \deg r_2 < \deg r_1; \\ r_1 = q_3 r_2 + r_3, & \deg r_3 < \deg r_2; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ r_k = q_{k+2} r_{k+1} + r_{k+2}, & \deg r_{k+2} < \deg r_{k+1}; \\ r_{k+1} = q_{k+3} r_{k+2}. & \end{array} \tag{37}$$

Отметим, что в (37) число k существует, поскольку степени остатков убывают.

Теорема 8.14. $(f, g) = r_{k+2}$.

Доказательство. Из предпоследнего равенства в (37) вытекает, что $r_{k+2}|r_{k+1}$, и т. д. Поднимаясь вверх получаем, что $r_{k+2}|g, r_{k+2}|f$.

Обратно, если $d'|f, d'|g$, то из первого равенства в (37) получаем, что $d'|r_1$, а из второго – $d'|g$. Двигаясь вниз, получаем, что $d'|r_{k+2}$. □

Следствие 8.15. *Пусть* $d = (f_1, \dots, f_m)$. *Тогда существуют такие многочлены* $u_1, \dots, u_m$, *что*

$$d = u_1 f_1 + \cdots + u_m f_m.$$

Доказательство. В силу предложения 8.13 можно считать, что $m = 2$, и $f_1 = f, f_2 = g$. Остается воспользоваться алгоритмом (37). □

Определение 8.16. Многочлены $f_1, \dots, f_m$ *взаимно просты*, если $(f_1, \dots, f_m) = 1$.

Предложение 8.17. *Многочлены* $f_1, \dots, f_m$ *взаимно просты, тогда и только тогда, когда существуют такие многочлены* $u_1, \dots, u_m$, *что*

$$1 = u_1 f_1 + \cdots + u_m f_m.$$

Определение 8.18. Многочлен p степени $\deg p \geq 1$ называется *неприводимым*, если p не разлагается в произведение многочленов меньшей степени.

Упражнение 8.19. Доказать, что

(1) любой многочлен первой степени,
(2) многочлен $X^2 + 1 \in \mathbb{R}[X]$

неприводим.

Предложение 8.20. *Пусть* $p \in \mathbb{R}[X]$ *неприводим и* $f \in R[X]$. *Тогда либо* $(p, f) = 1$, *либо* $p|f$.

Доказательство. Пусть $d = (p, f)$, и $f = f'd$. В силу определения 8.18 либо $\deg d = 0$, либо $\deg d = \deg p$.

В первом случае по предложению 8.7 элемент d обратим, т. е. в силу предложения 8.13 можно считать, что $d = 1$.

Пусть $\deg d = \deg p$. Тогда $p = dc$, где c – ненулевая константа, т. е. $d = pc^{-1}$. Отсюда $f = pc^{-1}f'$. □

Предложение 8.21. *Пусть* $p \in \mathbb{R}[X]$ *неприводим и* $p|(fg)$, *где* $f, g \in R[X]$. *Тогда либо* $p|f$, *либо* $p|g$.

Доказательство. По предложению 8.20 можно считать, что $(p, f) = 1$. Тогда $1 = fu + pv$ по предложению 8.17. Отсюда $g = 1g = fgu + pvg$. Поэтому $p|g$. □

Теорема 8.22. *Любой ненулевой многочлен разлагается в произведение неприводимых многочленов. Если*

$$f = p_1 \cdots p_n = q_1 \cdots q_m \tag{38}$$

– два разложения в произведения неприводимых многочленов $p_1, \dots, p_n, q_1, \dots, q_m$, *то* $n = m$, *и существует такая перестановка* $\sigma \in S_n$, *что* $p_i = c_i q_{\sigma i}$, *где* c_i *– ненулевая константа.*

Доказательство. **Существование.**

Индукция по $\deg f$. Случай $\deg f = 1$ очевиден, так как в этом случае f неприводим. Пусть для меньших степеней утверждение доказано. Если f неприводим, то утверждение

доказано. Если $f = gh$, где $\deg g, \deg g < \deg f$, то нужно воспользоваться индукцией для g, h.

Единственность. Индукция по $\deg f$. Случай $\deg f = 1$ очевиден. Пусть для меньших степеней утверждение доказано, и выполнено (38). По предложению 8.21 $p_1 | q_i$ для некоторого i. Перенумеровывая, можно считать, что $i = 1$. В силу неприводимости q_1 получаем $q_1 = c_1 p_1$, где c_1 – ненулевая константа. Тогда

$$p_1(p_2 \cdots p_n - (c_1 q_2) q_3 \cdots q_m) = 0.$$

По предложению 8.8 и упражнению 7.24 получаем, что

$$p_2 \cdots p_n - (c_1 q_2) q_3 \cdots q_m = 0,$$

или

$$p_2 \cdots p_n = (c_1 q_2) q_3 \cdots q_m.$$

по индукции $n - 1 = m - 1$ и $p_i = c_i q_i, i \geq 2, c_i \in R^*$. □

Определение 8.23. В силу теоремы 8.22 каждый ненулевой многочлен $f \in R[X]$ однозначно представляется в виде

$$f = p_1^{l_1} \cdots p_n^{l_n}, \quad c \in R^*, \quad l_i \in \mathbb{N} \cup 0, \tag{39}$$

где $p_1, \ldots, p_n$ – унитарные неприводимые многочлены. Число l_i называется *кратностью* p_i в f.

Теорема 8.24. *Пусть* $\operatorname{char} R = 0$, *и* $f \in R[X] \setminus 0$. *Если* k – *кратность неприводимого многочлена* p *в* f, *то кратность* p *в* f' *равна* $k - 1$.

Доказательство. Имеем $f = p^k g$, где $(p, g) = 1$. Тогда

$$f' = k p^{k-1} p' g + p^k g = p^{k-1}(k p' g + p g'). \tag{40}$$

Следовательно, $p^{k-1} | f'$. Если $p^k | f'$, то по $p | (k p' g + p g')$, и поэтому $p | k p' g$. По предложению 8.21 либо $p | k p'$, либо $p | g$, Но второе невозможно по предположению, а первое невозможно, ибо по предположению $k p' \neq 0$ и $\deg(k p') < \deg p$. □

Следствие 8.25. *Пусть* $\operatorname{char} R = 0$, *и* $f \in R[X] \setminus 0$ *имеет разложение* (39). *Тогда*

$$d = (f, f') = p_1^{l_1 - 1} \cdots p_n^{l_n - 1}, \quad \text{и} \quad \frac{f}{d} = p_1 \cdots p_n.$$

3. Корни многочленов

Определение 8.26. Пусть R – поле и $c \in R$. Для a из (36) положим $a(c) = a_0 + a_1 c + \cdots + a_n c^n$. Элемент является *корнем* a, если $a(c) = 0$.

Упражнение 8.27. Пусть $f, g \in R[X]$. Тогда

$$(f + g)(c) = f(c) + g(c), \quad (fg)(c) = f(c) g(c).$$

Теорема 8.28 (Теорема Безу). *Элемент* $c \in R$ *является корнем* $f \in R[X]$ *тогда и только тогда, когда* $(X - c) | f$.

Доказательство. Деля f с остатком на $X - c$ получаем, что $f = (X - c) q + r$, где $r \in R$. Отсюда $r = f(c)$. □

Отметим, что если $c \in R$, то деля любой многочлен последовательно с остатком на $X - c$, получаем для f разложение Тейлора

$$f = b_0 + b_1(X - c) + \cdots + b_n(X - c)^n. \tag{41}$$

Изложим *схему Горнера* для быстрого вычисления коэффициентов b_i в (41) по коэффициентам f. Для этого нужно уметь делать один шаг деления на $X - c$. Пусть

$$f = a_0 + a_1X + \cdots + a_nX^n = g(X - c) + r, \quad r \in R, \tag{42}$$

где

$$g = s_0 + s_1X + \cdots + s_{n-1}X^{n-1}.$$

Подставляя в (42) получаем

$$a_0 + a_1X + \cdots + a_nX^n = (X - c)(s_0 + s_1X + \cdots + s_{n-1}X^{n-1}) + r.$$

Приравнивая коэффициенты при одинаковых степенях X, получаем

$$\begin{array}{rcl} a_n & = & s_{n-1} \\ a_{n-1} & = & s_{n-2} - cs_{n-1} \\ \cdots\cdots & \cdots & \cdots\cdots \\ a_1 & = & s_0 - cs_1 \\ a_0 & = & r - cs_0. \end{array}$$

Отсюда

$$\begin{array}{rcl} s_{n-1} & = & a_n \\ s_{n-2} & = & a_{n-1} + cs_{n-1} \\ \cdots\cdots & \cdots & \cdots\cdots \\ s_0 & = & a_1 + cs_1 \\ r & = & a_0 + cs_0. \end{array} \tag{43}$$

Формулы (43) позволяют быстро за n умножений вычислить $r = f(c)$. Результаты этих вычислений обычно записываются в виде таблицы

	a_n	a_{n-1}	$\ldots$	a_1	a_0
c	$a_0 = s_{n-1}$	s_{n-2}	$\ldots$	s_1	r

(44)

Деля далее g с остатком на $X - c$ получаем, получаем формулу Тейлора. Результаты этих вычислений запишем в таблицу

	a_n	a_{n-1}	$\ldots$	a_1	a_0
c	$a_n = s_{n-1}$	s_{n-2}	$\ldots$	s_1	$r = b_0$
c	$a_n = t_{n-2}$	t_{n-3}	$\ldots$	$t_0 = b_1$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$		
c	$a_n = u_1$	$u_0 = b_{n-1}$			
c	$a_n = b_n$				

(45)

где b_i из (41).

Определение 8.29. *Кратностью корня* c многочлена f называется кратность неприводимого множителя $X - c$ в f. Другими словами, кратность равна k, если $(X - c)^k | f$, но $(X - c)^{k+1} \nmid f$.

Это определение согласовано с теоремой Безу 8.28

Упражнение 8.30. Следующие условия эквивалентны:

(1) кратность корня c многочлена f равна k;
(2) в формуле Тейлора

$$f = b_k(X - c)^k + b_{k+1}(X - c)^{k+1} + \cdots + b_n(X - c)^n, \quad b_k \neq 0.$$

Если $c_1, \dots, c_m$ – различные корни f с кратностями $k_1, \dots, k_m$, то

$$f = (X - c_1)^{k_1} \cdots (X - c_m)^{k_m} g.$$

Поэтому справедливо

Предложение 8.31. *Сумма числа корней (с кратностями) не превосходит степени многочлена.*

Упражнение 8.32. Пусть R – поле характеристики 0. Элемент $c \in R$ является кратным корнем (корнем кратности ≥ 2) тогда и только тогда, когда c общий корень многочлена и его производной.

4. Интерполяция

Пусть R – поле с различными элементами $a_0, a-1, \dots, a_n,$ и $b_0, b-1, \dots, b_n \in R$. Рассмотрим интерполяционный многочлен Лагранжа

$$F = \sum_{i=0}^{n} b_i \frac{(X - a_0) \cdots (X - a_{i-1})(X - a_{i+1}) \cdots (X - a_n)}{(a_i - a_0) \cdots (a_i - a_{i-1})(a_i - a_{i+1}) \cdots (a_i - a_n)}. \tag{46}$$

Теорема 8.33. $\deg F \leq n$ *и* $F(a_j) = b_j$ *для всех* $j = 0, \dots, n$.

Доказательство.

$$F(a_j) = \sum_{i=0}^{n} b_i \frac{(a_j - a_0) \cdots (a_j - a_{i-1})(a_j - a_{i+1}) \cdots (a_j - a_n)}{(a_i - a_0) \cdots (a_i - a_{i-1})(a_i - a_{i+1}) \cdots (a_i - a_n)} = b_j.$$

□

Теорема 8.34. *Существует в точности один многочлен* F *степени* $\leq n$ *с условием* $F(a_j) = b_j$ *для всех* $j = 0, \dots, n$.

Доказательство. Пусть существуют два многочлена F_1, F_2 с этим условием. Тогда $G = F_1 - F_2 \neq 0$, причем $\deg G \leq n$, и элементы $a_0, a-1, \dots, a_n$ является его корнями. Это противоречит теореме 8.31. □

Следствие 8.35. *Пусть* R *– бесконечное поле. Если два многочлена* f, g *задают одинаковую функцию на* R, *т. е.* $f(c) = g(c)$ *для всех* $c \in R$, *то* $f = g$.

Теорема 8.36. *Пусть* R *– поле из* q *элементов. Тогда ненулевой многочлен* $X^q - X$ *задает нулевую функцию на* R.

Доказательство. Все ненулевые элементы R^* поля R образуют мультипликативную группу порядка $q-1$. Если $c \in R^*$, то порядок c в этой группе делит $q-1$ по следствию 6.40. Поэтому по предложению 6.16 $c^{q-1} = 1$, откуда $c^q = c$. Это равенство верно и для нулевого элемента. □

5. Корни многочленов над $\mathbb{C}$ и $\mathbb{R}$

Теорема 8.37 (Основная теорема алгебры). *Пусть* $f \in \mathbb{C}[X]$ *имеет положительную степень. Тогда* f *имеет комплексный корень.*

Доказательство. Приведем два доказательства этой теоремы.

Первое доказательство

Пусть f не имеет комплексных корней. Тогда $F = \dfrac{1}{f}$ аналитична в $\mathbb{C}$ и ограничена. По теореме Лиувилля она постоянна. Но тогда и f постоянна, что неверно по следствию 8.35.

Второе доказательство

Лемма 8.38. *Пусть $f \in \mathbb{C}[X]$ имеет положительную степень. Если $z \to \infty$, то $|f(z)| \to \infty$.*

Доказательство. Если f имеет вид

$$f = a_0 + a_1 X + \cdots + a_n X^n, \quad a_n \neq 0. \tag{47}$$

Тогда

$$|f(z)| = |z|^n |a_n + \frac{a_{n-1}}{|z|} + \cdots + \frac{a_0}{|z|^n}| \geq |z|^n \left(|a_n| - \frac{|a_{n-1}|}{|z|} - \cdots - \frac{|a_0|}{|z|^n} \right).$$

С ростом $|z|$ число

$$\left(|a_n| - \frac{|a_{n-1}|}{|z|} - \cdots - \frac{|a_0|}{|z|^n} \right).$$

ограничено снизу некоторым M. Поэтому $|f(z)| \geq |z|^n M$. □

Лемма 8.39 (Лемма Даламбера). *Если f имеет вид (47) и $f(z_0) \neq 0$, то в любой окрестности z_0 найдется такое число z, что $|f(z)| < |f(z_0)|$.*

Доказательство. Разложим $f(z)$ в ряд Тейлора в окрестности z_0. Деля это разложение на $f(z_0) \neq 0$, получаем

$$\frac{f(z)}{f(z_0)} = 1 + c_p z^p + c_{p+1} z^{p+1} + \cdots + c_n z^n, \quad c_p, c_n \neq 0.$$

Пусть $z_1 \in \mathbb{C}$ и $z_1^p = -\dfrac{1}{c_p}$. Если $z = z_0 + t z_1, t \in (0,1) \subseteq \mathbb{R}$, то

$$\frac{f(z)}{f(z_0)} = 1 - t^p + t^{p+1} h_{n-p-1}(t),$$

где $h_{n-p-1}(t) \in \mathbb{C}[t]$ имеет степень $n-p-1$. Если C – максимум модулей коэффициентов $h_{n-p-1}(t)$, то при $t < \dfrac{1}{C(n-p)}$

$$\left| \frac{f(z)}{f(z_0)} \right| \leq 1 - t^p + |h_{n-p-1}(t)| \leq 1 - t^p + C(n-p)t^{p+1} = 1 - t^p(1 - C(n-p)t) < 1.$$

□

Завершим второе доказательство теоремы. Если $f \in \mathbb{C}[X]$ имеет положительную степень, и у f нет корней, то $\infty > M = \inf_{z \in \mathbb{C}} |f(z)| > 0$. Выберем последовательность таких комплексных чисел z_k, что $f(z_k) \to M$. Если эта последовательность неограничена, то из нее можно выбрать подпоследовательность $z_{i_k} \to \infty$. Получается противоречие с леммой 8.38.

Пусть последовательность z_k ограничена. В силу полноты $\mathbb{C}$ в ней можно выбрать сходящуюся подпоследовательность. Без ограничения общности можно считать, что это исходная последовательность. Итак, $z_k \to z_0$. Тогда $f(z_k) \to f(z_0) = M$. В силу леммы Даламбера получаем противоречие. □

Следствие 8.40. *Неприводимые многочлены над $\mathbb{C}$ имеют степень 1. В частности, каждый многочлен $f \in \mathbb{C}[X]$ однозначно представляется в виде*

$$f = a(X - c_1)^{k_1} \cdots (X - c_m)^{k_m}, \quad a \in \mathbb{C},$$

где $c_1, \ldots, c_m$ различные корни f кратностей $k_1, \ldots, k_m$.

Доказательство. Воспользоваться теоремами 8.37, 8.28. □

Предложение 8.41. *Пусть $f \in \mathbb{R}[X]$ и $c \in \mathbb{C}$ – корень f. Тогда $\overline{c}$ также корень f.*

Доказательство. Пусть f имеет вид 47. Тогда

$$0 = \overline{f(c)} = \overline{a_0 + a_1 c + \cdots + a_n c^n} = \overline{a_0} + \overline{a_1 c} + \cdots + \overline{a_n}\overline{c^n} = a_0 + a_1\overline{c} + \cdots + a_n\overline{c}^n = f(\overline{c}).$$

□

Упражнение 8.42. Пусть $f \in \mathbb{R}[X]$ и $c \in \mathbb{C}$ – корень f кратности k. Тогда $\overline{c}$ также корень f кратности k.

Теорема 8.43. *Неприводимые многочлены над $\mathbb{R}$ имеют степень ≤ 2. В частности, каждый многочлен $f \in \mathbb{R}[X]$ однозначно представляется в виде*

$$f = a(X - c_1)^{k_1} \cdots (X - c_m)^{k_m}(X^2 + p_1 X + q_1)^{l_1} \cdots (X^2 + p_s X + q_s)^{l_s}, \quad a \in \mathbb{R},$$

где $c_1, \ldots, c_m$ различные корни f кратностей $k_1, \ldots, k_m$, а многочлены $(X^2 + p_i X + q_i)^{l_i} \in \mathbb{R}[X]$ различны и не имеют вещественных корней.

Приведем оценку модулей корней комплексного многочлена.

Теорема 8.44. *Пусть $f \in \mathbb{C}[X]$ имеет вид (47). Предположим, что $a_{n-1} = \cdots = a_{k+1} = 0$, и $a_n, a_k \neq 0$, где $0 \leq k < n$. Положим*

$$B = \max(|a_k|, |a_{k-1}|, \ldots, |a_0|).$$

Если $z \in \mathbb{C}$ – корень f, то

$$|z| < 1 + \sqrt[n-k]{\frac{B}{|a_n|}}.$$

Доказательство. Пусть

$$|z| \geq 1 + \sqrt[n-k]{\frac{B}{|a_n|}}.$$

Достаточно показать, что $|f(z)| > 0$. Действительно,

$$|f(z)| = |a_n z^n + a_k z^k + \cdots + a_1 z + a_0| \geq |a_n z^n| - |a_k z^k + \cdots + a_1 z + a_0| \geq$$
$$|a_n||z|^n - \textstyle\sum_{j=0}^{k} |a_j||z|^j \geq |a_n||z|^n - B\sum_{j=0}^{k} |z|^j = |a_n||z|^n - B\frac{|z|^{k+1} - 1}{|z| - 1} =$$
$$\frac{|a_n|}{|z| - 1}\left[|z|^n(|z| - 1) - \frac{B}{|a_n|}|z|^{k+1} + \frac{B}{|a_n|}\right] > \frac{|a_n|}{|z| - 1}\left[|z|^n(|z| - 1) - \frac{B}{|a_n|}|z|^{k+1}\right] =$$
$$\frac{|a_n||z|^{k+1}}{|z| - 1}\left[|z|^{n-k-1}(|z| - 1) - \frac{B}{|a_n|}\right].$$

Таким образом, остается показать, что

$$|z|^{n-k-1}(|z| - 1) > \frac{B}{|a_n|}.$$

По условию

$$|z| - 1 \geq \alpha = \sqrt[n-k]{\frac{B}{|a_n|}} > 0.$$

Поэтому

$$|z|^{n-k-1}(|z| - 1) - \frac{B}{|a_n|} \geq (1 + \alpha)^{n-k-1}\alpha - \alpha^{n-k} = \alpha[(1 + \alpha)^{n-k-1} - \alpha^{n-k-1}] > 0,$$

ибо $1 + \alpha > \alpha$. □

Приведем алгоритм Штурма приближенного нахождения вещественных корней многочленов вещественными коэффициентами. Заметим, что если $f \in \mathbb{R}[X]$, то по следствию 8.25 многочлен $\dfrac{f}{(f,f')} \in \mathbb{R}[X]$ имеет те же корни, что и f, но их кратности становятся равными 1, т. е. $\dfrac{f}{(f,f')}$ не имеет кратных корней.

Теорема 8.45 (Штурм). *Пусть $f \in \mathbb{R}[X]$ не имеет кратных корней. Построим последовательность Штурма $f_0 = f$, и $f_1 = f'$. Далее каждое для любого $i \geq 2$ через f_i обозначим остаток со знаком - от деления f_{i-2} на f_{i-1}, т. е.*

$$f_{i-2} = g_{i-1}f_{i-1} - f_i. \tag{48}$$

Пусть $u < v$ – вещественные числа, не являющиеся корнями f. Для любого вещественного числа a через $W(a)$ обозначим число перемен знаков в ряду

$$f_0(a), f_1(a), \ldots, f_n(a), \quad n = \deg f. \tag{49}$$

Тогда число вещественных корней в $[u, v]$ равно $W(u) - W(v)$.

Доказательство. Нам потребуется ряд лемм.

Лемма 8.46. *Если $f_i(c) = 0$, то $f_{i+1}(c) \neq 0$. Кроме того, если $i > 0$, то $f_{i-1}(c) = -f_{i+1}(c)$.*

Доказательство. Если бы $f_i(c) = f_{i+1}(c) = 0$, то по (48) $f(c) = f'(c) = 0$, т. е. c – кратный корень f, что невозможно в силу (8.24) (См. также упражнение 8.32).

Второе утверждение вытекает из (48). □

Лемма 8.47. *Пусть $f_i(c) = 0$, где $i > 0$. Тогда существует такое $\varepsilon > 0$, что $f_{i-1}(x), f_{i+1}(x)$ имеют разные знаки при $x \in (c - \varepsilon, c + \varepsilon)$.*

Доказательство. Нужно воспользоваться непрерывностью f_j для всех j. □

Завершим доказательство теоремы. Пусть $\in (u, v)$ не является корнем f. Рассмотрим поведение функции $W(x)$ в окрестности c. Если все $f_j(c) \neq 0$, то в некоторой окрестности c все $f_j(c)$ имеют постоянный знак. Поэтому $W(x)$ постоянно в этой окрестности c. Если же некоторое $f_j(c) = 0, j \geq 1$, то в некоторой окрестности c по лемме 8.47 функции $f_{j-1}(x), f_{j+1}(x)$ имеют постоянный знак, и поэтому число перемен знаков в ряду

$$f_{j-1}(x), \quad f_j(x), \quad f_{j+1}(x) \tag{50}$$

постоянно. Беря пересечение этих окрестностей для всех j, получаем окрестность c, в которой $W(x)$ постоянно.

Пусть $f(c) = 0$. По лемме 8.46 имеем $f_1(c) \neq 0$. Пусть $f_1(c) > 0$. Тогда в некоторой окрестности U точки c функция $f(x)$ возрастает. Пусть $f(x) < 0, f(y) > 0$, при

$$x, y \in U, \quad x < c < y. \tag{51}$$

Если $f_j(c) \neq 0$, то можно считать, что $f_j(x)$ имеет постоянный знак в U. Если $f_j(c) = 0, j > 1$, то можно считать, что в ряду (50) число перемен знаков постоянно. Следовательно, для числа $W(a)$ перемен знаков в ряду (49) получаем $W(x) - W(y) = 1$, если x, y из (51).

Пусть $f_1(c) < 0$. Тогда в некоторой окрестности U точки c функция $f(x)$ убывает. Пусть $f(x) > 0, f(y) < 0$, при

$$x, y \in U, \quad x < c < y. \tag{52}$$

Если $f_j(c) \neq 0$, то можно считать, что $f_j(x)$ имеет постоянный знак в U. Если $f_j(c) = 0, j > 1$, то можно считать, что в ряду (50) число перемен знаков постоянно. Следовательно, для числа $W(a)$ перемен знаков в ряду (49) получаем $W(x) - W(y) = 1$, если x, y из (52). □

6. Неприводимые многочлены над $\mathbb{Z}$ и $\mathbb{Q}$

Определение 8.48. Многочлен из $\mathbb{Z}[X]$ называется *примитивным*, если причем наибольший общий делитель всех его коэффициентов равен 1.

Теорема 8.49 (Гаусс). *Произведение примитивных многочленов является примитивным.*

Доказательство. Пусть $f, g \in \mathbb{Z}[X]$ примитивны, но fg не примитивно. Тогда существует простое число p, делящее все коэффициенты fg. Перейдем к кольцу вычетов $\mathbb{Z}_p$. Если $\overline{f}, \overline{g}$ – образы f, g в $\mathbb{Z}_p[X]$, то в $\mathbb{Z}_p[X]$ получаем $\overline{f}\,\overline{g} = 0$, что невозможно, так как $\mathbb{Z}_p[X]$ – область. □

Если $f \in \mathbb{Q}[X] \setminus 0$, то $f = \dfrac{n}{m}\widetilde{f}$, где $n, m \in \mathbb{Z}$, причем $(n, m) = 1$, и $\widetilde{f} \in \mathbb{Z}[X]$, – примитивный многочлен.

Теорема 8.50. *Пусть $f \in \mathbb{Z}[X]$ примитивен, и $f = gh$, где $g, h \in \mathbb{Q}[X]$. Тогда $f = uv$, где $u, v \in \mathbb{Z}[X]$, и $u = \pm 1\widetilde{g}, v = \pm 1\widetilde{h}$.*

Доказательство. Имеем $g = \dfrac{r}{s}\widetilde{g}, \quad r, s \in \mathbb{Z}$, и $(r, s) = 1$. Аналогично, $h = \dfrac{c}{d}\widetilde{h}$, $c, d \in \mathbb{Z}$, и $(c, d) = 1$. Итак, $f = \dfrac{rc}{sd}\widetilde{g}\widetilde{h}$, или

$$(sd)f = (rc)\widetilde{g}\widetilde{h}, \tag{53}$$

По теореме 8.49 наибольший общий делитель коэффициентов правой части (53) равен rs, а с другой стороны, по (53) он равен sd. Отсюда $rs = \pm cd$. Отсюда $f = \pm\widetilde{g}\widetilde{h}$. □

Следствие 8.51. *Многочлен $f \in \mathbb{Z}[X]$ неразложим в $\mathbb{Z}[X]$ в произведение многочленов меньшей степени тогда и только тогда, когда он неприводим в $\mathbb{Q}[X]$.*

Теорема 8.52 (Критерий Эйзенштейна). *Пусть $f \in \mathbb{Z}[X]$ имеет вид (47), причем существует такое простое число p, что*

(1) $p \nmid a_n$;
(2) $p|a_{n-1}, \ldots, p|a_1, p|a_0$;
(3) $p^2 \nmid a_0$.

Тогда многочлен f неприводим в $\mathbb{Q}[X]$.

Доказательство. Пусть $d =$ всех коэффициентов f. Тогда $p \nmid d$. Поэтому f можно заменить на $d^{-1}f$ и считать, что f примитивен. Пусть $f = gh$ – нетривиальное разложение f в $\mathbb{Z}[X]$ (см. теорему 8.50). Факторизуя по модулю p получаем в $\mathbb{Z}_p[X]$ разложение $\overline{a_0}X^n = \overline{g}\,\overline{h}$, где черта сверху означает образ в $\mathbb{Z}_p[X]$ по модулю p. В силу теоремы 8.22 получаем, что

$$\overline{g} = \overline{b}X^r, \quad \overline{h} = \overline{c}X^{n-r},$$

где $b, c \in \mathbb{Z}$, и $0 < r < n$. Но тогда свободные члены g, h делятся на p, и поэтому свободный член $f = gh$ делится на p^2, что противоречит предположению. □

Пример 8.53. Пусть p – простое число. Тогда для любого $n \geq 1$ многочлен $X^n + p \in \mathbb{Z}[X]$ неприводим.

Следствие 8.54. *Пусть p – простое число. Тогда циклотомический многочлен*

$$\Phi_p(X) = X^{p-1} + x^{p-2} + \cdots + X + 1 = \frac{X^p - 1}{X - 1}$$

неприводим над $\mathbb{Q}$.

Доказательство. Положим $T = X + 1$. Тогда

$$\Phi_p(X) = \frac{(T+1)^p - 1}{T} = \frac{T^p + pX^{p-1} + \cdots + \binom{p}{k}T^k + \cdots + pT}{T} =$$
$$T^{p-1} + pX^{p-2} + \cdots + \binom{p}{k}T^{k-1} + \cdots + p.$$

Остается воспользоваться критерий Эйзенштейна. □

7. Рациональные дроби

Пусть K – коммутативная область целостности. Рассмотрим множество отношений $\frac{a}{b}$, где $a, b \in K$, причем $b \neq 0$.

Определение 8.55. Скажем, что два отношения $\frac{a}{b}$, $\frac{c}{d}$ равны, если $ad = bc$.

Предложение 8.56. *Это отношение является отношением эквивалентности.*

Упражнение 8.57. Следующие дроби равны $\frac{ab}{ac} = \frac{b}{c}$.

Обозначение 8.58. Через F обозначим множество классов равных дробей.

Введем в F сложение и умножение по правилу

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \quad \frac{a}{b}\frac{c}{d} = \frac{ac}{bd}.$$

Предложение 8.59. *Определение сложения и умножения корректно. F относительно этих операция является полем.*

Предложение 8.60. *Зададим отображение $\phi : K \to F$, полагая $\phi(a) = \frac{a}{1}$. Тогда ϕ инъективно, и $\phi(a + b) = \phi(a) + \phi(b)$, $\phi(ab) = \phi(a)\phi(b)$.*

Итак, можно считать, что $P \subseteq F$.

Пример 8.61. Если $K = \mathbb{Z}$, то $F = \mathbb{Q}$.

Определение 8.62. Если $K = R[X]$, где R – поле, то $F = R(X)$ называется *полем рациональных функций* над R. Элемент $\frac{f}{g} \in R(X)$ называется *правильной дробью*, если $\deg f < \deg g$. Элемент $\frac{f}{p^k} \in R(X)$ называется *простейшей дробью*, если $p \in R[X]$ – неприводимый многочлен, и $\deg f < \deg p$.

Теорема 8.63. *Любая дробь является суммой многочлена и простейших дробей.*

Доказательство. Нам потребуется

Лемма 8.64. *Пусть $\frac{f}{g} \in R(X)$, и $g = uv$, где $u, v \in R[X]$, причем $(u, v) = 1$. Тогда $\frac{f}{g} = \frac{h}{u} + \frac{t}{v}$.*

Доказательство. Имеем $ua + vb = 1$ для некоторых $a, b \in R[X]$. Отсюда $f = uaf + vbf$, и

$$\frac{f}{g} = \frac{uaf}{uv} + \frac{vbf}{uv} = \frac{af}{v} + \frac{bf}{u}.$$

□

В силу теоремы 8.22 и предыдущей леммы можно считать, что дробь имеет вид $\frac{f}{p^k}$, где p – неприводимый многочлен. Покажем утверждение теоремы индукцией по $\deg f$. Если $\deg f < \deg p$, то эта дробь простейшая. Пусть $\deg f \geq \deg p$. Разделим f с остатком на p,

$$f = qp + r, \quad r = 0 \text{ или } \deg r < \deg p.$$

Тогда

$$\frac{f}{p^k} = \frac{qp+r}{p^k} = \frac{q}{p^{k-1}} + \frac{r}{p^k}.$$

Второе слагаемое является правильной дробью, а в первом степень числителя уменьшилась. □

Следствие 8.65. *Каждая рациональная дробь из $\mathbb{C}(X)$ представляется в виде суммы многочлена и дробей вида $\frac{c}{(X-a)^k}$. Каждая рациональная дробь из $\mathbb{R}(X)$ представляется в виде суммы многочлена и дробей вида*

$$\frac{c}{(X-a)^k}, \quad \frac{aX+b}{(X^2+pA+q)^k},$$

где $X^2 + pX + q \in \mathbb{R}[X]$ не имеет вещественных корней.

8. Кольцо степенных рядов

Пусть R – ассоциативное кольцо с 1. Рассмотрим множество $R[[X]]$ всех последовательностей

$$a = (a_0, a_1, \dots) \tag{54}$$

элементов из R. Определим в $R[[X]]$ операцию сложения по-координатно. Кроме того, если a из (54),

$$b = (b_0, b_1, \dots),$$

то

$$c = ab = (c_0, c_1, \dots),$$

где для любого $k \geq 0$

$$c_k = \sum_{i=0}^{k} a_i b_{k-i}. \tag{55}$$

Предложение 8.66. *$R[[X]]$ является ассоциативным кольцом с 1. Если R коммутативно, то и $R[[X]]$ коммутативно.*

Определение 8.67. $R[[X]]$ называется *кольцом степенных рядов*.

Предложение 8.68. *Все элементы $(a_0, 0, \dots) \in R[[X]]$ образуют подкольцо, изоморфное R. Все почти нулевые последовательности в $R[[X]]$ образуют подкольцо, являющееся кольцом многочленов $R[X]$.*

Доказательство. Изоморфизм задается по правилу $a \mapsto (a, 0, \dots) \in R[[X]]$. □

Обозначение 8.69. Всюду в дальнейшем мы будем отождествлять $a \in R$ с $(a, 0, \dots) \in R[X]$. Положим $X = (0, 1, 0, \dots)$.

Предложение 8.70. *Для любого $n \geq 1$*

$$X^n = (0, \dots, 0, \overset{n}{1}, 0, \dots)$$

Если a из (54), то

$$f = a_0 + a_1 X + \cdots, \quad a_i \in R. \tag{56}$$

Определение 8.71. Пусть $a \in R[X]$ из (36), причем $a_0 = \ldots = a_{n-1} = 0, a_n \neq 0$. *Порядком* $o(f)$ называется число n. a_n называется *младшим членом* f.

Всюду в дальнейшем мы будем предполагать, R – область.

Предложение 8.72. *Порядок произведения рядов равен сумме их порядков. Младший член произведения рядов равен произведению младших членов сомножителей. В частности, $R[[X]]$ – область.*

Упражнение 8.73. $o(f+g) \geq o(f), o(g)$.

Теорема 8.74. *Элемент f из (56) обратим тогда и только тогда, когда a_0 обратим в R.*

Доказательство. Пусть

$$g = b_0 + b_1 X + \cdots = f^{-1}, \quad b_i \in R. \tag{57}$$

Тогда $1 = gf$, откуда $1 = a_0 b_0$. Аналогично, из $1 = fg$, откуда $1 = b_0 a_0$.

Обратно, пусть f из (56) и a_0 обратимо в R. Будем искать $g = f^{-1}$ в виде (57), где $b_0 = a_0^{-1}$. Для любого $j \geq 1$ из условия $fg = 1$ имеем

$$a_j b_0 + a_{j-1} b_1 + \cdots + a_0 b_j = 0.$$

Отсюда

$$b_j = -a_0^{-1} \left[a_j b_0 + a_{j-1} b_1 + \cdots + a_1 b_{j-1} \right].$$

Эта система равенств позволяет последовательно определить все коэффициенты $b_j, j \geq 1$. Итак, для f нашелся правый обратный элемент g. Так как свободный член g обратим, то для g найдется правый обратный h, т. е. $gh = 1$. Отсюда $f = f(gh) = (fg)h = h$. □

Всюду в дальнейшем мы будем предполагать, что R – поле.

Следствие 8.75. *Каждый ненулевой элемент f из $R[[X]]$ имеет вид $f = X^n u$, где $n = o(f) \geq 0$ и u – обратимый элемент из $R[[X]]$.*

Определение 8.76. Так как $R[[X]]$ – коммутативная область, то для нее существует поле дробей $R((X))$.Оно называется *полем лорановских рядов*.

Предложение 8.77. *Каждый ненулевой элемент из $R((X))$ однозначно представляется в виде $X^n u$, где $n \in \mathbb{Z}$ и u – обратимый элемент из $R[[X]]$.*

Доказательство. Пусть $a = X^m u, b = X^m v$, где $n, m \geq 0$, и u, v – обратимые элементы из $R[[X]]$. Тогда $\dfrac{a}{b} = X^{m-n} u v^{-1}$ имеет указанный вид.

Если $X^s u = X^l v \in R((X))$, и $s \geq l$, то $X^{s-l} = uv^{-1}$. Сравнивая свободные члены, получаем $s - l = 0$. тогда $u = v$. □

Следствие 8.78. *Если $f \in R((X)) \setminus 0$, то f однозначно представляется в виде лорановского ряда*

$$f = a_n X^n + a_{n+1} X^{n+1} + \cdots, \quad n \in \mathbb{Z}, \quad a_i \in R, \quad a_n \neq 0. \tag{58}$$

Определение 8.79. Пусть $f \in R((X))$ из (58). Положим

$$f' = n a_n X^{n-1} + (n+1) a_{n+1} X^n + \cdots \in \mathbb{R}((X)). \tag{59}$$

Предложение 8.80. *Если $f, g \in R((X))$, то $(f+g)' = f' + g'$, $(fg)' = f'g + fg'$.*

Доказательство. Непосредственная проверка с использованием (59) и (55). □

Следствие 8.81. *Если $f, g \in R((X))$, то $(g^{-1} f)' = g^{-2}(f'g - fg')$. В частности, $(g^n)' = n g^{n-1} g'$ для всех $n \in \mathbb{Z}$.*

Определение 8.82. Обозначим через $XR[[X]]$ множество всех $g \in R[[X]]$, имеющих порядок не меньше 1ю Пусть $f \in R[[X]]$, $g \in XR[[X]]$. *Подстановкой* ряда g в ряд f из (56) назовем ряд

$$f(g) = a_0 + a_1 g + \cdots + a_n g^n + \cdots \tag{60}$$

Определение 8.83. Предложение 8.84. *Определение подстановки ряда в ряд корректно.*

Доказательство. Действительно, так как $o(g^n) \geq n$, то в (60) коэффициент при X^m совпадает с коэффициентом при X^m у $a_0 + a_1 g + \cdots + a_{m-1} g^{m-1}$. □

Теорема 8.85. *Пусть* $f \in R[[X]]$, $g \in XR[[X]]$. *Тогда* $f(g)' = f'(g)g'$.

Доказательство. Коэффициент при X^m у $f(g)'$ совпадает с коэффициентом при X^m у $(a_0 + a_1 g + \cdots + a_m g^m)'$. Следовательно, по следствию 8.81 коэффициент при X^m у $f(g)'$ равен

$$a_1 g' + \cdots + m a_m g^{m-1} g' = (a_1 + 2a_2 g + \cdots + m a_m g^{m-1}) g'.$$

Отсюда следует утверждение. □

Всюду в дальнейшем мы будем предполагать, что поле R имеет нулевую характеристику.

Предложение 8.86. *Пусть* $f, g \in R[[X]]$, *причем* $f' = g'$. *Тогда* $f = g + c, c \in R$.

Упражнение 8.87. Пусть $f \in R[[X]]$ из (56). Тогда для любого $j \geq 0$

$$a_j = \frac{f^{(j)}(0)}{j!}.$$

Определение 8.88. Пусть $f \in XR[[X]]$. Положим

$$\exp(f) = 1 + \frac{f}{1!} + \cdots + \frac{f^n}{n!} + \cdots, \quad \ln(1+f) = \sum_{j \geq 1} \frac{(-1)^{i-1}}{i} f^i.$$

Теорема 8.89. *Если* $o(f), o(g) \geq 1$, *то* $\exp(f+g) = \exp f \exp g$, *и* $(\ln(1+f))' = \dfrac{f'}{1+f}$.

Доказательство. Имеем

$$\exp f \exp g = (1 + \frac{f}{1!} + \cdots + \frac{f^n}{n!} + \cdots)(1 + \frac{g}{1!} + \cdots + \frac{g^n}{n!} + \cdots) =$$

$$\textstyle\sum_{i,j \geq 0} \frac{f^i}{i!} \frac{f^j}{j!} = \sum_{t \geq 0} \frac{1}{t!} \sum_{i+j=t} \frac{t!}{i!j!} f^i g^j = \sum_{t \geq 0} \frac{1}{t!} (f+g)^t = \exp(f+g).$$

Для доказательства второго утверждения заметим, что по теореме 8.85

$$(\ln(1+f))' = \sum_{j \geq 1} (-1)^{j-1} f^{i-1} f' = (1+f)^{-1} f'.$$

□

Теорема 8.90. *Если* $o(f) \geq 1$, *то*

$$\ln(\exp f) = f, \quad \exp(\ln(1+f)) = f.$$

Доказательство. Достаточно показать утверждение при $f = X$. Вычислим производную, используя теорему 8.89,

$$(\ln(\exp X))' = (\ln(1 + (\exp X - 1)))' = \frac{(\exp X - 1)'}{1 + (\exp X - 1)} = \frac{\exp X}{\exp X} = 1 = X'.$$

Отсюда $\ln(\exp X) = X + c, c \in R$ по предложению предложению 8.86. Подставляя $X = 0$, получаем $\ln(\exp 0) = 0 = c$.

Рассмотрим теперь ряд

$$\frac{\exp(\ln(1+X))}{1+X}$$

и вычислим его производную, используя следствие 8.81 и теорему 8.89

$$\left[\frac{\exp(\ln(1+X))}{1+X}\right]' = \frac{[\exp(\ln(1+X))]'(1+X) - \exp(\ln(1+X))}{(1+X)^2} =$$
$$\frac{\frac{\exp(\ln(1+X))}{1+X}(1+X) - \exp(\ln(1+X))}{(1+X)^2} = 0$$

Отсюда

$$\frac{\exp(\ln(1+X))}{1+X} \in R.$$

Подставляя $X = 0$ получаем

$$\frac{\exp(\ln(1+X))}{1+X} = \exp(\ln 1)) = \exp 0 = 1.$$

□

Следствие 8.91. *Множество $XR[[X]]$ является группой относительно сложения. Пусть $1 + XR[[X]]$ – множество всех рядов из $R[[X]]$ со свободным членом 1. Тогда $1 + XR[[X]]$ – группа относительно умножения. Отображение $\exp : XR[[X]] \to 1 + XR[[X]]$ задает изоморфизм этих групп. Обратное отображение к нему имеет вид $\ln : 1 + XR[[X]] \to XR[[X]]$. В частности, $\ln((1+f)(1+g)) = \ln(1+f) + \ln(1+g)$ для всех $f, g \in XR[[X]]$.*

ГЛАВА 9

Многочлены от нескольких переменных

1. Кольцо многочленов от нескольких переменных

Определение 9.1. Пусть R – ассоциативное кольцо с 1. По индукции положим

$$R[X, \dots, X_n] = (R[X, \dots, X_{n-1}])[X_n].$$

Таким образом, каждый элемент из $R[X, \dots, X_n]$ однозначно представляется в виде суммы *одночленов*

$$a_{i_1,\dots,i_n} X_1^{i_1} \cdots X_n^{i_n}, \quad a_{i_1,\dots,i_n} \in R, \tag{61}$$

где $i_1, \dots, i_n \in \mathbb{N} \cup 0$. Пусть $(\mathbb{N} \cup 0)^n$ – множество *мультииндексов*, т. е. множество наборов $(i_1, \dots, i_n)$ элементов из $\mathbb{N} \cup 0$. Тогда одночлен (61) записывается в виде

$$a_i X^i, \quad i = (i_1, \dots, i_n) \in (\mathbb{N} \cup 0)^n, \quad X^i = X_1^{i_1} \cdots X_n^{i_n}.$$

Из следствия 8.8 вытекает

Теорема 9.2. *Если R – область, то и $R[X, \dots, X_n]$ – область.*

Для сравнения одночленов введем в $(\mathbb{N} \cup 0)^n$ отношение порядка.

Определение 9.3. Пусть

$$m = (m_1, \dots, m_n), r = (r_1 \dots, r_n) \in (\mathbb{N} \cup 0)^n.$$

Скажем, что $m > r$, если существует такое $1 \le j < n$, что

$$m_1 = r_1, \dots, m_{j-1} = r_{j-1}, m_j > r_j. \tag{62}$$

Предложение 9.4. *Если $m > m', m' > m''$, то $m > m''$. Кроме того, $m \ge m$, и если $m \ge m', m' \ge m$, то $m = m'$.*

Доказательство. Пусть

$$m = (m_1, \dots, m_n), m' = (m'_1 \dots, m'_n), m'' = (m''_1 \dots, m''_n). \tag{63}$$

Пусть

$$\begin{gathered} m_1 = m'_1, \dots, m_{j-1} = m'_{j-1}, m_j > m'_j; \\ m'_1 = m''_1, \dots, m'_{j'-1} = m''_{j'-1}, m'_{j'} > m''_{j'}. \end{gathered} \tag{64}$$

Если $s = \min(j, j')$, то

$$m_1 = m''_1, \dots, m_{s-1} = m''_{s-1}, m_s > m''_s.$$

Поэтому $m > m''$. Аналогично доказываются остальные утверждения. □

Теорема 9.5. *Любая последовательность $M_1 \ge M_2 \ge \dots$ в $(\mathbb{N} \cup 0)^n$ стабилизируется.*

Доказательство. Доказательство проводится индукцией по n. Если $n = 1$, то утверждение очевидно. Пусть для $n-1$ теорема доказана. Если $M_i = (m_{i1}, m_{i2}, \dots, m_{in})$, то по условию $m_{11} \geq m_{21} \geq \dots$. Следовательно, существует такое k, что $m_{k,1} = m_{k+1,1} = \dots$. По условию в этом случае в $(\mathbb{N} \cup 0)^{n-1}$ получаем

$$(m_{k,2}, \dots, m_{k,n}) \geq (m_{k+1,2}, \dots, m_{k+1,n}) \geq \dots$$

По индукции эта последовательность стабилизируется начиная с некоторого места t. Тогда $M_t = M_{t+1} = \dots$. □

Предложение 9.6. *Если $m, m', m'' \in (\mathbb{N} \cup 0)^n$ и $m > m'$, то $m + m'' > m' + m''$.*

Доказательство. Пусть m, m', m'' из (63) и выполнено (64). Тогда

$$m_1 + m_1'' = m_1' + m_1'', \dots, m_{j-1} + m_{j-1}'' = m_{j-1}' + m_{j-1}'', m_j + m_j'' > m_j' + m_j''.$$

□

Следствие 9.7. *Если $m, m', m'', m''' \in (\mathbb{N} \cup 0)^n$ и $m \geq m', m'' \geq m'''$, то $m + m'' > m' + m'''$.*

Определение 9.8. Пусть

$$f = \sum_{i \in (\mathbb{N} \cup 0)^n} a_i X^i \neq 0. \tag{65}$$

Одночлен $a_m X^m$ из этого представления называется *старшим* в f, если $m > j$ для всех таких $j \in (\mathbb{N} \cup 0)^n$, что $a_j \neq 0$.

Всюду в дальнейшем в этой главе мы будем предполагать, что R – ассоциативно-коммутативная область с 1. Из следствия 9.7 вытекает

Теорема 9.9. *Старший одночлен произведения многочленов равен произведению старших одночленов.*

Определение 9.10. Многочлен f из (9.7) *однороден степени d*, если в любом ненулевом одночлене $a_i X^i$ из (9.7), где $i = (i_1, \dots, i_n)$, выполнено условие $i_1 + \dots + i_n = d$.

Упражнение 9.11. Пусть f из (9.7) и f_d – *его однородная компонента степени d*, т. е. сумма всех его однородных одночленов степени d. Тогда $f = f_0 + f_1 + \dots$.

2. Симметричные многочлены

Определение 9.12. Многочлен f из (9.7) *симметричен*, если он не меняется при любой перестановке переменных. Примерами симметричных многочленов являются *элементарные симметричные* многочлены

$$\begin{array}{lllll} \sigma_1 & = & \sigma_1(X_1, \dots, X_n) & = & X_1 + \dots + X_n; \\ \sigma_2 & = & \sigma_2(X_1, \dots, X_n) & = & X_1X_2 + \dots + X_{n-1}X_n; \\ \dots & & \dots & & \dots \\ \sigma_n & = & \sigma_n(X_1, \dots, X_n) & = & X_1 \cdots X_n. \end{array} \tag{66}$$

Другими словами, σ_k – это сумма всех произведений по k неизвестных $X_{i_1} \cdots X_{i_k}$, $1 \geq i_1 < \dots < i_k \geq n$.

Упражнение 9.13. Все симметрические многочлены образуют подкольцо в кольце многочленов. Если симметрический многочлен f из (9.7) имеет представление $f = f_0 + f_1 + \dots$ в соответствии с упражнением 9.11, то все его однородные компоненты $f_0, f_1, \dots$ симметричны. Старший одночлен σ_k равен $X_1 \cdots X_k$.

Предложение 9.14. *Пусть*

$$aX_1^{m_1}\cdots,X_n^{m_n} \tag{67}$$

– старший член некоторого симметричного многочлена. Тогда

$$m_1 \geq m_2 \geq \cdots \geq m_n. \tag{68}$$

Доказательство. Пусть, например, $m_1 < m_2$. Переставляя X_1 и X_2, получаем, что в многочлен входит также одночлен $aX_1^{m_2}X_2m_1\cdots X_n^{m_n}$, который старше (67), что невозможно. □

Теорема 9.15. *Каждый симметричный многочлен представляется в виде многочлена от элементарных симметричных многочленов.*

Доказательство. Пусть задан симметричный многочлен f со страшим одночленом (67). По предложению 9.14 получаем (68). Рассмотрим симметричный многочлен

$$h = a\sigma_1^{m_1-m_2}\cdots\sigma_{n-1}^{m_n-m_{n-1}}\sigma_n^{m_n}$$

По теореме 9.9 и упражнению 9.13 старший член h равен

$$aX_1^{m_1-m_2}(X_1X_2)^{m_2-m_3}\cdots(X_1\cdots X_{n-1})^{m_n-m_{n-1}}(X_1\cdots X_n)^{m_n} = aX_1^{m_1}\cdots X_n^{m_n}$$

т. е. совпадает со старшим членом f. Поэтому старший член симметрического многочлена $f-h$ меньше старшего члена f. Продолжая этот процесс в силу теоремы 9.5 получаем утверждение теоремы. □

Приведем пример вычислений. Пусть задан однородный симметричный многочлен $f = X_1^3 + X_2^3 + X_3^3$ степени 3. Его старший член X_1^3 имеет набор показателей $(3,0,0)$. Выпишем все наборы $(m_1,m_2,m_3) \in (\mathbb{N}\cup 0)^3$ с условиями:

(1) $(m_1,m_2,m_3) < (3,0,0)$;
(2) $m_1 \geq m_2 \geq m_3$;
(3) $m_1+m_2+m+3 = 3$.

Эти наборы соответствуют наборам показателей старших членов, возникающих в доказательстве основной теоремы. В данном случае такими наборами являются $(2,1,0),(1,1,1)$. Каждому из получившихся наборов $(3,0,0),(2,1,0),(1,1,1)$ сопоставлен как и в теореме одночлен от элементарных симметричных многочленов. Тогда

$$f = \sigma_1^3 + a\sigma_1\sigma_2 + b\sigma_3, \tag{69}$$

где коэффициенты a,b нужно определить. Подставляя в (69) $X_1 = X_2 = 1, X_3 = 0$ получаем $2 = f = 8+2a$, откуда $a = -3$. Подставляя в (69) $X_1 = X_2 = X_3 = 1$ получаем $3 = f = 27-27+b$, откуда $b = 3$. Итак,

$$X_1^3 + X_2^3 + X_3^3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3.$$

Выведем теперь формулы Виета. Для этого в кольцо $R[X,X_1,\ldots,X_n]$ рассмотрим многочлен

$$f = a(X-X_1)\cdots(X-X_n), \quad a \in R.$$

Перемножая, получаем

$$f = aX^n + a(-1)\sigma_1X^{n-1} + \cdots + a(-1)^k\sigma_kX^{n-k} + \cdots + a(-1)^n\sigma_n. \tag{70}$$

Таким образом, справедливо

Предложение 9.16. *Пусть R – поле и*

$$f = a_nX^n + \cdots + a_0 \in R[X]$$

имеет корни $\alpha_1, \dots, \alpha_n$*, считая с их кратностями. Тогда для любого* $k = 1, \dots, n$ *справедливы формулы Виета*

$$\frac{a_{n-k}}{a_n} = (-1)^k \sigma_k. \tag{71}$$

Доказательство. Нужно подставить $X_1 = \alpha_1, \dots, X_n = \alpha_n, a = a_n$ в (70). □

Мы выражали симметричные многочлены через элементарные. Но в ряде случаев можно выражать и через другие симметричные многочлены. Именно, для любого натурального числа k положим

$$p_k = X_1^k + \dots + X_n^k \in \mathbb{Z}[X_1, \dots, X_n].$$

Теорема 9.17 (Формулы Ньютона). *В кольце* $\mathbb{Z}[X_1, \dots, X_n]$ *положим* $\sigma_k = 0$ *при* $k > n$ *и* $\sigma_0 = 1$*. Тогда для любого* $k \geq 1$ *имеем*

$$p_k - p_{k-1}\sigma_1 + p_{k-2}\sigma_2 + \dots + (-1)^{k-1} p_1 \sigma_{k-1} + (-1)^k \sigma_k = 0.$$

Доказательство. Заметим, что $\mathbb{Z}[X_1, \dots, X_n] \subseteq \mathbb{Q}[X, X_1, \dots, X_n]$. По (70)

$$F(X) = \sum_{k=0}^{n} \sigma_k X^k = \prod_{k=1}^{n} (1 + X_k X).$$

Тогда по (8.91)

$$\ln(F(X)) = \sum_{k=1}^{n} \ln(1 + X_k X).$$

Вычисляя производную по X по теореме 8.89 получаем, что

$$\frac{F(X)'}{F(X)} = \textstyle\sum_{k=1}^{n} \frac{X_k}{1 + X_k X} = \sum_{k=1}^{n} X_k (1 - X_k X + X_k^2 X^2 + \dots) =$$
$$p_1 - X p_2 + X^2 p_3 + \dots + (-1)^k T^k p_{k+1} + \dots$$

С другой стороны,

$$\frac{F(X)'}{F(X)} = \frac{\sum_{k=1}^{n} \sigma_k k X^{k-1}}{\sum_{k=1}^{n} \sigma_k X^k}.$$

Поэтому

$$\sum_{k=1}^{n} \sigma_k k X^{k-1} = (\sum_{k=1}^{n} \sigma_k X^k)(p_1 - X p_2 + X^2 p_3 + \dots + (-1)^k T^k p_{k+1} + \dots)$$

Итак, при $k = 1, \dots, n$

$$k\sigma_k = (-1)^k p_k + (-1)^{k-1} \sigma_1 + \dots + p_1 \sigma_{k-1}.$$

Отсюда находим p_k □

Следствие 9.18. *Если* R *– поле нулевой характеристики, то каждый симметричный многочлен является многочленом от* $p_1, \dots, p_n$*.*

Укажем явные выражения σ_k через $p_1, \dots, p_n$ и наоборот. В силу теоремы 9.17 имеем систему линейных уравнений

$$\begin{cases} \sigma_1 & & & & = & p_1 \\ \sigma_1 p_1 & -2\sigma_2 & & & = & p_2 \\ \sigma_1 p_2 & -\sigma_2 p_1 & 3\sigma_2 & & = & p_3 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \sigma_1 p_{k-1} & -\sigma_2 p_{k-2} & \sigma_3 p_{k-2} & \dots & (-1)^{k-1} k\sigma_k & = & p_k \end{cases} \tag{72}$$

Решая систему (72) относительно σ_k по правилу Крамера, получаем

$$\sigma_k = \frac{\begin{vmatrix} 1 & 0 & 0 & \dots & 0 & p_1 \\ p_1 & -2 & 0 & \dots & 0 & p_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ p_{k-1} & p_{k-2} & \dots & \dots & p_1 & p_k \end{vmatrix}}{\begin{vmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ p_1 & -2 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ p_{k-1} & p_{k-2} & \dots & \dots & p_1 & (-1)^{k-1}k \end{vmatrix}} = \frac{1}{k!}\begin{vmatrix} p_1 & 1 & 0 & \dots & 0 & 0 \\ p_2 & p_1 & 2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ p_k & p_{k-1} & \dots & \dots & p_2 & p_1 \end{vmatrix}.$$

Решая систему (72) относительно p_k по правилу Крамера, получаем, что

$$p_k = \begin{vmatrix} \sigma_1 & 1 & 0 & \dots & 0 & 0 \\ 2\sigma_2 & \sigma_1 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ (k-1)\sigma_{k-1} & \sigma_{k-2} & \dots & \dots & \sigma_1 & 1 \\ k\sigma_k & \sigma_{k-1} & \dots & \dots & \sigma_2 & \sigma - 1 \end{vmatrix} \tag{73}$$

В частности,

$$p_2 = \sigma_1^2 - 2\sigma_2, \quad p_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3.$$

3. Дискриминант и результант

Рассмотрим в $\mathbb{Z}[X_1, \dots, X_n]$ многочлен

$$W(X_1, \dots, X_n) = \prod_{1 \le i < j \le n} (X_j - X_i).$$

Многочлен $W(X_1, \dots, X_n)$ кососимметричен по $X_1, \dots, X_n$. Поэтому $W_n(X_1, \dots, X_n)^2$ является симметрическим многочленом в $\mathbb{Z}[X_1, \dots, X_n]$. Найдем его выражение через $p_1, \dots, p_n$ и тем самым через $\sigma_1, \dots, \sigma_n$ по формулам (73). В самом деле,

$$W(X_1, \dots, X_n)^2 = \begin{vmatrix} 1 & 1 & \dots & 1 \\ X_1 & X_2 & \dots & X_n \\ X_1^2 & X_2^2 & \dots & X_n^2 \\ \dots & \dots & \dots & \dots \\ X_1^{n-1} & X_2^{n-1} & \dots & X_n^{n-1} \end{vmatrix} \begin{vmatrix} 1 & X_1 & X_1^2 & \dots & X_1^{n-1} \\ 1 & X_2 & X_2^2 & \dots & X_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & X_n & X_n^2 & \dots & X_n^{n-1} \end{vmatrix} =$$
$$\begin{vmatrix} n & p_1 & p_2 & \dots & p_{n-1} \\ p_1 & p_2 & p_3 & \dots & p_n \\ \dots & \dots & \dots & \dots & \dots \\ p_{n-1} & p_n & p_{n+1} & \dots & p_{2n-2} \end{vmatrix}. \tag{74}$$

В частности, при $n = 2$

$$\mathrm{Disc}(x_1, x_2) = \begin{vmatrix} 2 & p_1 \\ p_1 & p_2 \end{vmatrix} = 2p_2 - p_1^2 = 2(\sigma_1^2 - 2\sigma_2) - \sigma_1^2 = \sigma_1^2 - 4\sigma_2.$$

Определение 9.19. Имеем

$$W(X_1, \dots, X_n)^2 = \mathrm{Disc}(\sigma_1, \dots, \sigma_n).$$

Многочлен

$$\mathrm{Disc}(X_1, \dots, X_n) \in \mathbb{Z}[X_1, \dots, X_n]$$

из (74) называется *дискриминантом*.

Из этого определения вытекает

Теорема 9.20. *Пусть*

$$f = X^n + a_{n-1}X^{n-1} + \cdots + a_0 \in R[X],$$

, R – поле, и $\alpha_1, \dots, \alpha_n$ – его корни. Тогда

$$\mathrm{Disc}(\sigma_1(\alpha_1, \dots, \alpha_n), \dots, \sigma_n(\alpha_1, \dots, \alpha_n))) = 0$$

в том и только в том случае, если f имеет кратный корень.

Определение 9.21. Пусть задан многочлен

$$f = a_n X^n + a_{n-1}X^{n-1} + \cdots + a_0, \quad a_n \neq 0.$$

Предположим, что $\alpha_1, \dots, \alpha_n$ – его корни. *Дискриминантом* этого многочлена называется

$$\mathrm{Disc}(f) = a_n^{2n-2} W(\alpha_1, \dots, \alpha_n)^2.$$

В силу (9.19) и формул Виета получаем

$$\mathrm{Disc}(f) = a_n^{2n-2}\,\mathrm{Disc}(-\frac{a_{n-1}}{a_n}, \dots, (-1)^n \frac{a_0}{a_n}).$$

Тем самым можно вычислить $\mathrm{Disc}(f)$ через коэффициенты $a_n, \dots, a_n$. Укажем явный способ этого вычисления. Для этого введем понятие результанта двух многочленов. Пусть даны два многочлена

$$f = a_n X^n + a_{n-1}X^{n-1} + \cdots + a_0, \quad g = b_m X^m + b_{m-1}X^{m-1} + \cdots + b_0$$

с коэффициентами в поле R. При этом мы предполагаем, что, по крайней мере, один из коэффициентов a_n, b_m отличен от нуля. Положим

$$\mathcal{R}(f,g) = \begin{pmatrix} a_n & a_{n-1} & a_{n-2} & \dots & a_0 & 0 & 0 & \dots & 0 \\ 0 & a_n & a_{n-1} & \dots & a_1 & a_0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_n & a_{n-1} & \dots & \dots & a_0 \\ b_m & b_{m-1} & b_{m-2} & \dots & b_1 & b_0 & 0 & \dots & 0 \\ 0 & b_m & b_{m-1} & \dots & b_2 & b_1 & b_0 & \dots & 0 \\ \dots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \dots & \dots & \dots & b_m & \dots & \dots & \dots & \dots & b_0 \end{pmatrix} \in \mathrm{Mat}(m+n, R).$$

В этой матрице первые n строк заполнены со сдвигом на один шаг коэффициентами многочлена f, а последние m строк – коэффициентами многочлена f. Из вида $\mathcal{R}(f,g)$ вытекает

Предложение 9.22. *Справедливо равенство*

$$\mathcal{R}(f,g) \begin{pmatrix} X^{n+m-1} \\ X^{n+m-2} \\ \vdots \\ X \\ 1 \end{pmatrix} = \begin{pmatrix} X^{m-1}f \\ X^{m-2}f \\ \vdots \\ f \\ X^{n-1}g \\ X^{n-2}g \\ \vdots \\ g \end{pmatrix} \tag{75}$$

Определение 9.23. *Результантом* многочленов f, g называется

$$R(f,g) = \det(\mathcal{R}(f,g)).$$

Теорема 9.24. *Существуют такие многочлены $u, v \in R[X]$, что $R(f,g) = fu + gv$.*

Доказательство. Если $R(f,g)=0$, то утверждение очевидно. Пусть $R(f,g)\neq 0$. Решая систему (75) относительно 1 по правилу Крамера, получаем

$$1=\frac{Uf+VG}{R(f,g)},\quad U,V\in R[X].$$

Остается положить $u=\dfrac{U}{R(f,g)}, v=\dfrac{V}{R(f,g)}$. □

Теорема 9.25. *Следующие условия эквивалентны:*

(1) $R(f,g)\neq 0$;
(2) $(f,g)=1$.

Доказательство. Пусть $(f,g)=1$ и $R(f,g)=0$. Тогда строки $\mathcal{R}(f,g)$ линейно зависимы. Поэтому найдется такая ненулевая строка

$$\Lambda=(\lambda_{n+m-1},\dots,\lambda_0),$$

что $\Lambda\mathcal{R}(f,g)=0$. В силу (75) получаем, что

$$\Lambda\begin{pmatrix} X^{m-1}f\\ X^{m-2}f\\ \vdots\\ f\\ X^{n-1}g\\ X^{n-2}g\\ \vdots\\ g\end{pmatrix}=(\lambda_{n+m-1}X^{m-1}+\dots+\lambda_n)f+(\lambda_{n-1}X^{n-1}+\dots+\lambda_0)g=0.$$

Но $(f,g)=1$. Поэтому из последнего равенства вытекает, что

$$f|(\lambda_{n-1}X^{n-1}+\dots+\lambda_0),\quad g|(\lambda_{n+m-1}X^{m-1}+\dots+\lambda_n).$$

Это возможно лишь в случае, когда $\Lambda=0$, так как либо $\deg f=n$, либо $\deg g=m$.

Обратно, пусть $R(f,g)\neq 0$. По теореме 9.24 существуют такие многочлены u,v, что $R(f,g)=fu+gv$. Тогда $1=f\frac{u}{R(f,g)}+g\frac{u}{R(f,g)}$, и поэтому $(f,g)=1$. □

Следствие 9.26. *Пусть $f,g\in\mathbb{C}[X]$. Следующие условия эквивалентны:*

(1) $R(f,g)=0$;
(2) *f,g имеют общий корень.*

Следствие 9.27. *Если R – произвольное поле, и $f,g\in R[X]$ имеют общий корень, то $R(f,g)=0$.*

Нам потребуется следующее

Предложение 9.28. *Пусть R – коммутативно-ассоциативная область с 1, и*

$$f(X_1,\dots,X_n)\in R[X_1,\dots,X_n].$$

Если $f(X_1,X_2,X_3,\dots,X_n)=0$, то

$$f(X_1,\dots,X_n)=(X_1-X_2)h(X_1,\dots,X_n),\quad h(X_1,\dots,X_n)\in R[X_1,\dots,X_n].$$

Доказательство. Заметим, что

$$R[X_1,\dots,X_n]=R[X_1,X_1-X_2,X_3,\dots,X_n].$$

Поэтому

$$f(X_1,\dots,X_n)=u(X_1,X_3,\dots,X_n)+(X_1-X_2)h(X_1,\dots,X_n).$$

Отсюда следует утверждение. □

Теорема 9.29. *Пусть K – коммутативно-ассоциативная область с 1,*

$$R = K[X_1, \dots, X_n, Y_1, \dots, Y_m].$$

В $R[X]$ положим

$$\begin{aligned} f &= a_n X^n + a_{n-1} X^{n-1} + \dots + a_0, \quad a_k = (-1)^k a_n \sigma(X_1, \dots, X_n), \\ g &= b_m X^m + b_{m-1} X^{m-1} + \dots + b_0, \quad b_k = (-1)^k b_m \sigma(Y_1, \dots, Y_m), \end{aligned} \tag{76}$$

где $a_n, b_n \in K$ не равны нулю. Тогда

$$R(f,g) = a_n^m b_m^n \prod_{i,j} (X_i - Y_j).$$

Доказательство. Без ограничения общности можно считать, что $a_n = b_m = 1$. Если X_i отождествить с некоторым Y_j, то $R(f,g)$ обратиться в 0 по следствию 9.27. Таким образом, по предложению 9.28

$$R(f,g) = A \prod_{i,j} (X_i - Y_j), \quad A \in R. \tag{77}$$

По формулам Виета

$$\prod_{i,j} (X_i - Y_j) = \prod_i g(X_i). \tag{78}$$

Следовательно, степень $\prod_{i,j}(X_i - Y_j)$ по любому X_i равна m. С другой стороны, степень $\sigma_j(X_1, \dots, X_n)$ по X_i равна 1. Поэтому степень $R(f,g)$ по X_i не меньше m. Таким образом, в (77) степень A по X_i равна 0. Аналогичные рассуждения применимы для всех переменных X_i, Y_j. Поэтому $A \in K$.

Положим

$$X_1 = \dots = X_n = 0, \quad Y_1 = \dots = Y_m = 1.$$

Тогда

$$\prod_{i,j} (X_i - Y_j) = (-1)^{mn}, \quad \text{и} \quad \sigma_i(0, \dots, 0) = 0, \quad \sigma_j(1, \dots, 1) = \binom{m}{j}.$$

Отсюда

$$R(f,g)(0,\dots,0,1,\dots,1) = \begin{vmatrix} 1 & 0 & 0 & \dots & & & & 0 \\ 0 & 1 & 0 & \dots & & & & 0 \\ \dots & & & & & & & \\ 0 & 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 1 & -\binom{m}{1} & \dots & & \dots & (-1)^m & 0 \dots & 0 \\ 0 & 1 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & (-1)^m & 0 \\ \dots & & & & & & & (-1)^m \end{vmatrix} = (-1)^{mn}.$$

Поэтому $A = 1$. □

Теорема 9.30. $a_n \operatorname{Disc}(f) = (-1)^{\frac{n(n-1)}{2}} R(f, f')$.

Доказательство. Пусть

$$f = a_n X^n + a_{n-1} X^{n-1} + \dots + a_0, \quad f' = n a_n X^{n-1} + \dots + a_1.$$

Если $x_1, \dots, x_n$ – корни f, то как и в предыдущей теореме (см. (78))

$$R(f,g) = a_n^{n-1} \prod_{i=1}^{n} f'(x_i).$$

Вычислим $f'(x_i)$. Имеем

$$f = a_n \prod_{i=1}^{n}(X - x_i), \quad f' = \sum_{i=1}^{n}\prod_{j\neq i}^{n}(X - x_j).$$

Отсюда

$$f'(x_i) = a_n \prod_{j\neq i}^{n}(x_i - x_j),$$

и, следовательно,

$$R(f, f') = a_n^{2n-1} \prod_{i=1}^{n}\prod_{j\neq i}(x_i - x_j) =$$
$$a_n^{2n-1}(-1)^{\frac{n(n-1)}{2}} \prod_{j<i}(x_i - x_j)^2 = a_n(-1)^{\frac{n(n-1)}{2}} \operatorname{Disc}(f).$$

□