САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Факультет прикладной математики – процессов управления

Е. А. КАЛИНИНА, А. Ю. УТЕШЕВ

АЛГЕБРА И ГЕОМЕТРИЯ
ЧАСТЬ II. ВЫСШАЯ АЛГЕБРА

Учебное пособие

Санкт-Петербург
2006

УДК 512.8
Д32

Р е ц е н з е н т ы : д-р физ.-мат. наук, проф. *Е.И. Веремей* (С.-Петерб. гос. у-нт); д-р физ.-мат. наук, проф. *В.Ф. Зайцев* (Рос. гос. пед. ун-т им. А.И. Герцена)

Печатается по постановлению Редакционно-издательского совета факультета прикладной математики – процессов управления Санкт-Петербургского государственного университета

Д32 **Алгебра и геометрия. Часть 2. Высшая алгебра:** Учеб. пособие / Калинина Е.А., Утешев А.Ю. — СПб.: "СОЛО", 2006. — 172 с.

ISBN 5-98340-046-0

В настоящем пособии излагаются основные понятия, методы и алгоритмы классической высшей алгебры как науки о решении уравнений и систем уравнений.

Книга предназначена для студентов университетов, обучающихся по специальности “прикладная математика и информатика”.

Библиогр. 14 назв.

Работа выполнена при поддержке Федерального агентства по образованию в рамках Национального проекта “Образование”. Инновационный проект СПбГУ “Инновационная образовательная среда в классическом университете”, ИОП “Прикладная математика и информатика”

ISBN 5-98340-046-0

Перейти к оглавлению на странице: 460

ГЛАВА 4. СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ. МАТРИЦЫ И ОПРЕДЕЛИТЕЛИ

Алгебра — это наука о решении уравнений и систем уравнений. С алгебраическими уравнениями от одной переменной мы разобрались в главе 3. Теперь мы переходим к уравнениям от нескольких переменных и наше исследование начинаем с самого простого случая — линейного.

§1. Решение системы линейных уравнений: метод Гаусса

Определения

В настоящем параграфе $\mathbb{A}$ и $\mathbb{B}$ обозначают любые множества из $\mathbb{Q}, \mathbb{R}$ или $\mathbb{C}$.

Рассмотрим полином первой степени от переменных $x_1, \ldots, x_n$: $a_1x_1 + \cdots + a_nx_n - b$ с коэффициентами $\{a_1, \ldots, a_n, b\} \subset \mathbb{A}$.

Определение 1.1. Уравнение

$$a_1x_1 + \cdots + a_nx_n = b \tag{1.1}$$

называется **линейным уравнением**(**л.у.**) над $\mathbb{A}$ от **переменных** (**неизвестных**) $x_1, \ldots, x_n$. **Решением** уравнения (1.1) **во множестве** $\mathbb{B}$ называется любой набор значений для переменных $x_1 = \alpha_1, \ldots, x_n = \alpha_n$ из $\mathbb{B}$, обращающий уравнение в верное равенство:

$$a_1\alpha_1 + \cdots + a_n\alpha_n = b.$$

Пример 1.1. Найти все вещественные решения уравнения

$$x_1 - 2\,x_2 + x_3 = 1.$$

Решение. Из уравнения можем выразить любую из переменных через две оставшиеся, например: $x_1 = 1 + 2\,x_2 - x_3$. Придавая в этой формуле x_2 и x_3 произвольные значения из $\mathbb{R}$, можно получить значение для x_1. Формально решение можно записать в виде:

$$x_2 = t,\ x_3 = u,\ x_1 = 1 + 2\,t - u,$$

где t и u предполагаются произвольными вещественными числами. Других вещественных решений уравнение не имеет. △

Определение 1.2. **Системой линейных уравнений** (**л.у.**) над $\mathbb{A}$ называется совокупность из нескольких уравнений над $\mathbb{A}$ от одного и того же набора переменных:

$$\begin{cases} a_{11}x_1 & +a_{12}x_2 & +\ldots+a_{1n}x_n & = b_1, \\ a_{21}x_1 & +a_{22}x_2 & +\ldots+a_{2n}x_n & = b_2, \\ \ldots & & \ldots & \\ a_{m1}x_1 & +a_{m2}x_2 & +\ldots+a_{mn}x_n & = b_m. \end{cases} \tag{1.2}$$

или, в компактном виде[1]:

$$a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{jn}x_n = b_j \text{ для } j \in \{1, \ldots, m\}.$$

Относительно числа m уравнений не делается ни какого предположения: оно может быть меньше, больше или равно числу переменных n. Если $m > n$ то система (1.2) называется **переопределенной**.

Определение 1.3. **Решением** системы (1.2) называется решение одновременно всех уравнений системы. Система называется **совместной** если она имеет хотя бы одно решение, и **несовместной** в противном случае. Совместная система называется **определенной** если она имеет единственное решение и **неопределенной** если она имеет более одного решения. **Решить** систему (1.2) означает найти все ее решения или установить ее несовместность.

Определение 1.4. Две различные системы линейных уравнений называются **эквивалентными** или **равносильными** если их множества решений совпадают. Таким образом, любые две несовместные системы от одного набора переменных считаются эквивалентными.

Пример 1.2. Решить систему уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 & = b_1, \\ a_{21}x_1 + a_{22}x_2 & = b_2. \end{cases}$$

Решение. Если предположить, что $x_1 = \alpha_1$, $x_2 = \alpha_2$ — какое-то решение системы, то мы имеем два равенства:

$$a_{11}\alpha_1 + a_{12}\alpha_2 = b_1 \text{ и } a_{21}\alpha_1 + a_{22}\alpha_2 = b_2.$$

[1] Запомним, что первый индекс у коэффициента a_{jk} отвечает за номер уравнения, а второй — за номер переменной.

Умножим первое из этих равенств на a_{21}, а второе — на $(-a_{11})$ и сложим:

$$(a_{12}a_{21} - a_{11}a_{22})\,\alpha_2 = a_{21}b_1 - a_{11}b_2 .$$

Если же мы умножим первое равенство на a_{22}, а второе — на $(-a_{12})$ и снова сложим, то придем к

$$(a_{11}a_{22} - a_{12}a_{21})\,\alpha_1 = a_{22}b_1 - a_{12}b_2 .$$

Обозначим

$$\Delta \stackrel{\text{def}}{=} a_{11}a_{22} - a_{12}a_{21} .$$

Если $\Delta \neq 0$, то приходим к единственно возможному виду для α_j:

$$\alpha_1 = \frac{a_{22}b_1 - a_{12}b_2}{\Delta}\ , \ \alpha_2 = \frac{a_{11}b_2 - a_{21}b_1}{\Delta} . \tag{1.3}$$

Итак, если существует решение системы и $\Delta \neq 0$, то это решение должно иметь вид (1.3). Подстановка этих формул в исходную систему приводит к верным равенствам. Таким образом, условие $\Delta \neq 0$ является достаточным для единственности решения; позже мы увидим, что оно является и необходимым — если оно нарушается, то система либо несовместна как, например,

$$\left\{ \begin{array}{rl} x_1 + 2\,x_2 & = 1, \\ 2\,x_1 + 4\,x_2 & = b_2 \end{array} \right. \qquad \text{при } b_2 \neq 2;$$

либо неопределенна, а точнее — имеет бесконечное множество решений как, например,

$$\left\{ \begin{array}{rl} x_1 + 2\,x_2 & = 1, \\ 2\,x_1 + 4\,x_2 & = 2. \end{array} \right.$$

$\triangle$

Метод Гаусса

Определение 1.5. **Элементарными преобразованиями** системы **л.у.** называются преобразования следующих трех типов:

A) перестановка двух уравнений;

B) умножение обеих частей уравнения на любое отличное от нуля число из $\mathbb{A}$;

C) прибавление к одному уравнению любого другого, умноженного на произвольное число из $\mathbb{A}$: пара уравнений

$$\begin{aligned} a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{jn}x_n &= b_j, \\ a_{k1}x_1 + a_{k2}x_2 + \ldots + a_{kn}x_n &= b_k \end{aligned}$$

заменяется парой

$$\begin{array}{rcl} (a_{j1} + \lambda a_{k1})x_1 + (a_{j2} + \lambda a_{k2})x_2 + \ldots + (a_{jn} + \lambda a_{kn})x_n &=& b_j + \lambda b_k\,, \\ a_{k1}x_1 \qquad\qquad + a_{k2}x_2 + \ldots \qquad\qquad + a_{kn}x_n &=& b_k\,. \end{array} \tag{1.4}$$

Теорема 1.1. *Любое элементарное преобразование системы* **л.у.** *переводит эту систему в ей эквивалентную.*

Доказательство. Обозначим систему, получающуюся из (1.2) с помощью какого-то элементарного преобразования, через $(1.2)'$. Пусть таким преобразованием будет C). Если $x_1 = \alpha_1, \ldots, x_n = \alpha_n$ — решение (1.2), то оно обращает в верное равенство все уравнения системы и, в частности, j-е и k-е. Но тогда оно должно быть решением и их произвольной комбинации: например, первого из уравнений (1.4). Но все остальные уравнения системы $(1.2)'$ совпадают с уравнениями исходной системы (1.2). Следовательно $x_1 = \alpha_1, \ldots, x_n = \alpha_n$ — решение $(1.2)'$.

Обратно, пусть $x_1 = \beta_1, \ldots, x_n = \beta_n$ — решение $(1.2)'$. Система (1.2) получается из $(1.2)'$ с помощью обратного преобразования: когда к j-му уравнению прибавляется k-е, умноженное на $(-\lambda)$. На основании только что доказанного, преобразованная система, т.е. (1.2), должна иметь решением тот же самый набор $x_1 = \beta_1, \ldots, x_n = \beta_n$.

Таким образом, любое решение системы (1.2) должно быть решением системы $(1.2)'$ и обратно. Следовательно множества решений совпадают, а значит системы эквивалентны. ■

Задача. С помощью элементарных преобразований привести систему **л.у.** к наиболее простому виду: такому, из которого легко было бы установить множество решений.

Самым распространенным методом такого приведения является **метод исключения переменных** или **метод Гаусса**.

Предположим, что в системе (1.2) коэффициент $a_{11} \neq 0$. Вычтем из второго уравнения системы первое, умноженное на ${}^{a_{21}}/_{a_{11}}$. Получим новую систему **л.у.**, в которой все уравнения кроме второго совпадают с уравнениями исходной, а второе имеет вид:

$$\left(a_{22}-\frac{a_{21}}{a_{11}}a_{12}\right)x_2+\dots+\left(a_{2n}-\frac{a_{21}}{a_{11}}a_{1n}\right)x_n=b_2-\frac{a_{21}}{a_{11}}b_1,$$

т.е. из него пропала — *исключилась* — переменная x_1. По теореме 1.1 новая система эквивалентна системе (1.2). Аналогичное преобразование — вычитание из третьего уравнения системы первого, умноженного на ${}^{a_{31}}/_{a_{11}}$, позволяет исключить x_1 из этого уравнения, т.е. заменить его на

$$\left(a_{32}-\frac{a_{31}}{a_{11}}a_{12}\right)x_2+\dots+\left(a_{3n}-\frac{a_{31}}{a_{11}}a_{1n}\right)x_n=b_3-\frac{a_{31}}{a_{11}}b_1.$$

Продолжаем процесс далее. В конечном итоге исключаем x_1 из всех уравнений кроме первого:

$$\left\{\begin{array}{llll} a_{11}x_1+a_{12}x_2+\dots+a_{1n}x_n & = b_1, \\ \qquad\quad a_{22}^{[1]}x_2 \quad +\dots+a_{2n}^{[1]}x_n & = b_2^{[1]}, \\ \qquad\quad \dots & \dots \\ \qquad\quad a_{m2}^{[1]}x_2 \quad +\dots+a_{mn}^{[1]}x_n & = b_m^{[1]}. \end{array}\right. \quad \text{здесь} \quad \begin{array}{lcl} a_{jk}^{[1]} & \stackrel{\text{def}}{=} & a_{jk}-\dfrac{a_{j1}a_{1k}}{a_{11}}, \\ b_j^{[1]} & \stackrel{\text{def}}{=} & b_j-\dfrac{a_{j1}b_1}{a_{11}} \end{array}. \tag{1.5}$$

Система (1.5) эквивалентна системе (1.2), однако она имеет более простой вид: в ней выделилась *подсистема*

$$\left\{\begin{array}{ll} a_{22}^{[1]}x_2+\dots+a_{2n}^{[1]}x_n & = b_2^{[1]}, \\ \dots & \dots \\ a_{m2}^{[1]}x_2+\dots+a_{mn}^{[1]}x_n & = b_m^{[1]}, \end{array}\right. \tag{1.6}$$

которая не зависит от переменной x_1. К этой новой подсистеме можно применить те же рассуждения, что и к исходной системе (1.2), поставив теперь целью исключение переменной x_2. Предположим, что коэффициент $a_{22}^{[1]} \neq 0$. Тогда вычитая из второго уравнения системы (1.6) первое, домноженное на $a_{32}^{[1]}/a_{22}^{[1]}$, и т.д., из последнего уравнения системы (1.6) первое, домноженное на $a_{m2}^{[1]}/a_{22}^{[1]}$, добиваемся указанной цели: приходим к системе вида

$$\left\{\begin{array}{rcl} a_{22}^{[1]}x_2+a_{23}^{[1]}x_3+\ldots+a_{2n}^{[1]}x_n &=& b_2^{[1]},\\ a_{33}^{[2]}x_3+\ldots+a_{3n}^{[2]}x_n &=& b_3^{[2]},\\ \ldots & & \ldots\\ a_{m3}^{[2]}x_3+\ldots+a_{mn}^{[2]}x_n &=& b_m^{[2]}. \end{array}\right. \quad \text{здесь} \quad \begin{array}{rcl} a_{jk}^{[2]} &\stackrel{\mathrm{def}}{=}& a_{jk}^{[1]}-\dfrac{a_{j2}^{[1]}a_{2k}^{[1]}}{a_{22}^{[1]}},\\ b_j^{[2]} &\stackrel{\mathrm{def}}{=}& b_j^{[1]}-\dfrac{a_{j2}^{[1]}b_2^{[1]}}{a_{22}^{[1]}} \end{array}.$$

Понятно, что процесс исключения может быть продолжен и далее. Теперь посмотрим, где он может прерваться. Может так случиться, что очередная, ℓ-я подсистема имеет коэффициент $a_{\ell\ell}^{[\ell-1]}$ равным нулю, что не позволит алгоритму идти дальше — т.е. исключить переменную x_ℓ из оставшихся уравнений (в принципе, такое могло случиться уже на первом шаге, если бы коэффициент a_{11} был бы равен нулю). Возможные варианты дальнейших действий:

а) если хотя бы один коэффициент при x_ℓ в одном из оставшихся уравнений отличен от нуля: $a_{j\ell}^{[\ell-1]}\neq 0$, то это уравнение переставляется с ℓ-м;

б) если все коэффициенты $a_{j\ell}^{[\ell-1]}=0$, то переменная x_ℓ не входит ни в одно оставшееся уравнение, и можно перейти к исключению переменной $x_{\ell+1}$.

Поскольку число переменных конечно, то алгоритм исключения должен завершиться за конечное число шагов. Чем он может завершиться? Окончательная система должна иметь вид:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1\mathfrak{r}}x_{\mathfrak{r}} + a_{1,\mathfrak{r}+1}x_{\mathfrak{r}+1} + \ldots + a_{1n}x_n = b_1, \\ \qquad a_{22}^{[1]}x_2 + \ldots + a_{2\mathfrak{r}}^{[1]}x_{\mathfrak{r}} + a_{2,\mathfrak{r}+1}^{[1]}x_{\mathfrak{r}+1} + \ldots + a_{2n}^{[1]}x_n = b_2^{[1]}, \\ \qquad\qquad \ddots \qquad\qquad\qquad \ldots \\ \qquad\qquad\qquad a_{\mathfrak{r}\mathfrak{r}}^{[\mathfrak{r}-1]}x_{\mathfrak{r}} + a_{\mathfrak{r},\mathfrak{r}+1}^{[\mathfrak{r}-1]}x_{\mathfrak{r}+1} + \ldots + a_{\mathfrak{r},n}^{[\mathfrak{r}-1]}x_n = b_{\mathfrak{r}}^{[\mathfrak{r}-1]}, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad 0 = b_{\mathfrak{r}+1}^{[\mathfrak{r}-1]}, \\ \qquad\qquad\qquad\qquad\qquad\qquad \ldots \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad 0 = b_m^{[\mathfrak{r}-1]}, \end{cases} \tag{1.7}$$

при $\mathfrak{r} \le n$. Заметим, что все коэффициенты этой системы будут принадлежать тому же множеству $\mathbb{A}$, что и коэффициенты исходной системы (1.2).

Предположение. Будем считать, что каждое из первых $\mathfrak{r}$ уравнений системы (1.7) содержит в своей левой части хотя бы одну переменную с ненулевым коэффициентом.

Определение 1.6. Процесс приведения системы (1.2) к системе вида (1.7) называется **прямым ходом метода Гаусса**.

Очевидно, что в этом случае и эта окончательная система и, следовательно, исходная система (1.2) будет несовместной, если хотя бы одно из чисел $b_{\mathfrak{r}+1}^{[\mathfrak{r}-1]}, \ldots, b_m^{[\mathfrak{r}-1]}$ отлично от нуля.

Для простоты мы будем иллюстрировать наши рассуждения на системах **л.у.** над $\mathbb{R}$, в этом же множестве искать решения. Каждое из преобразований метода Гаусса будем обозначать $\to$.

Пример 1.3. Решить систему **л.у.**

$$\begin{cases} x_1 + x_2 - 3x_3 = -1 \\ 2x_1 + x_2 - 2x_3 = 1 \\ x_1 + x_2 + x_3 = 3 \\ x_1 + 2x_2 - 3x_3 = 1. \end{cases}$$

Решение.

$$\to\left\{\begin{array}{rl} x_1+x_2-3x_3 &= -1\\ -x_2+4x_3 &= 3\\ 4x_3 &= 4\\ x_2 \qquad &= 2\end{array}\right. \to\left\{\begin{array}{rl} x_1+x_2-3x_3 &= -1\\ -x_2+4x_3 &= 3\\ 4x_3 &= 4\\ 4x_3 &= 5\end{array}\right. \to\left\{\begin{array}{rl} x_1+x_2-3x_3 &= -1\\ -x_2+4x_3 &= 3\\ 4x_3 &= 4\\ 0 &= 1\end{array}\right.$$

Ответ. Система несовместна.

Пусть теперь $b_{\mathfrak{r}+1}^{[\mathfrak{r}-1]}=0,\dots,b_m^{[\mathfrak{r}-1]}=0$. Возможны два случая: $\mathfrak{r}=n$ и $\mathfrak{r}<n$. В случае $\mathfrak{r}=n$ перепишем систему (1.7):

$$\left\{\begin{array}{l} a_{11}x_1+a_{12}x_2+\dots+a_{1,n-1}x_{n-1}+a_{1,n}x_n=b_1,\\ \qquad a_{22}^{[1]}x_2+\dots+a_{2,n-1}^{[1]}x_{n-1}+a_{2,n}^{[1]}x_n=b_2^{[1]},\\ \qquad \dots \qquad\qquad \dots\\ \qquad\qquad a_{n-1,n-1}^{[n-2]}x_{n-1}+a_{n-1,n}^{[n-2]}x_n=b_{n-1}^{[n-2]},\\ \qquad\qquad\qquad a_{nn}^{[n-1]}x_n=b_n^{[n-1]}.\end{array}\right. \tag{1.8}$$

На основании предположения, $a_{nn}^{[n-1]}\neq 0$. Но тогда, поскольку система (1.7) является конечной стадией прямого хода метода Гаусса, то и все коэффициенты $a_{n-1,n-1}^{[n-2]},\dots,a_{22}^{[1]},a_{11}$ должны быть отличны от нуля — в противном случае метод Гаусса не остановился бы на системе такого вида (он называется **треугольным**).

Пример 1.4. При каких значениях коэффициентов система

$$\left\{\begin{array}{llll} a_{11}x_1+ & a_{12}x_2+ & a_{13}x_3= & b_1\\ a_{21}x_1+ & a_{22}x_2+ & a_{23}x_3= & b_2\\ a_{31}x_1+ & a_{32}x_2+ & a_{33}x_3= & b_3\end{array}\right. \tag{1.9}$$

приводится к треугольному виду?

Решение. Итак, если $a_{11}\neq 0$, то получаем:

$$a_{22}^{[1]}=\frac{a_{11}a_{22}-a_{12}a_{21}}{a_{11}}\ ,\ a_{23}^{[1]}=\frac{a_{11}a_{23}-a_{13}a_{21}}{a_{11}}.$$

Далее, если

$$\Delta_2\stackrel{\text{def}}{=}a_{11}a_{22}-a_{12}a_{21}\neq 0,$$

то

$$a_{33}^{[2]} = \frac{\Delta_3}{\Delta_2},$$

где

$$\begin{aligned} \Delta_3 &\stackrel{\text{def}}{=} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - \\ &- a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23}. \end{aligned}$$

Ответ. $a_{11} \neq 0, \Delta_2 \neq 0,\ \Delta_3 \neq 0$.

Из последнего уравнения системы (1.8) можно однозначно установить значение x_n:

$$x_n = b_n^{[n-1]}/a_{nn}^{[n-1]}.$$

Далее, подставляя это значение в $(n-1)$-е уравнение системы (1.8), выражаем x_{n-1}:

$$\begin{aligned} x_{n-1} &= \left(b_{n-1}^{[n-2]} - a_{n-1,n}^{[n-2]}x_n\right)/a_{n-1,n-1}^{[n-2]} = \\ &= \left(b_{n-1}^{[n-2]} - a_{n-1,n}^{[n-2]}b_n^{[n-1]}/a_{nn}^{[n-1]}\right)/a_{n-1,n-1}^{[n-2]}. \end{aligned}$$

Подставляем полученные значения для x_n и x_{n-1} в $(n-2)$-е уравнение системы (1.8), выражаем x_{n-2}, и т.д., в конце концов приходим к первому уравнению, из которого выражаем x_1 если ранее уже получены выражения для $x_2, \ldots, x_n$. Итак, в этом случае система (1.8), а вместе с ней и ей эквивалентная система (1.2), имеет единственное решение.

Пример 1.5. Решить систему **л.у.**

$$\begin{cases} x_1 + 3x_2 + x_3 = 5 \\ 2x_1 + x_2 + x_3 = 2 \\ x_1 + x_2 + 5x_3 = -7 \\ 2x_1 + 3x_2 - 3x_3 = 14. \end{cases}$$

Решение.

$$\to \begin{cases} x_1 + 3x_2 + x_3 = 5 \\ -5x_2 - x_3 = -8 \\ -2x_2 + 4x_3 = -12 \\ -3x_2 - 5x_3 = 4 \end{cases} \to \begin{cases} x_1 + 3x_2 + x_3 = 5 \\ -5x_2 - x_3 = -8 \\ {}^{22}\!/_5\, x_3 = -{}^{44}\!/_5 \\ -{}^{22}\!/_5\, x_3 = {}^{44}\!/_5 \end{cases} \to$$

$$\to \begin{cases} x_1 + 3\,x_2 \quad + x_3 = \quad 5 \\ \quad -5\,x_2 \quad - x_3 = \quad -8 \\ \qquad\qquad {}^{22}\!/\!{}_5\,x_3 = -{}^{44}\!/\!{}_5 \\ \qquad\qquad\qquad 0 = \quad 0 \end{cases}$$

Из третьего уравнения определяем $x_3 = -2$, подставляем во второе: $x_2 = 2$; оба значения подставляем в первое: $x_1 = 1$.

Ответ. $x_1 = 1$, $x_2 = 2$, $x_3 = -2$.

Исследуем теперь случай $\mathfrak{r} < n$ (такая форма системы (1.7) называется **трапециевидной**). На основании выдвинутого на с.191 предположения, в $\mathfrak{r}$-м уравнении системы (1.7) имеется хотя бы один ненулевой коэффициент в левой части, пусть $a_{\mathfrak{r}\mathfrak{s}}^{[\mathfrak{r}-1]} \ne 0$ — первый из них. Если $\mathfrak{s} = n$, то из этого уравнения однозначно определится x_n

$$x_n = \alpha_n \stackrel{\text{def}}{=} b_{\mathfrak{r}}^{[\mathfrak{r}-1]} / a_{\mathfrak{r}n}^{[\mathfrak{r}-1]}.$$

Если же $\mathfrak{s} < n$, то из того же уравнения можно выразить $x_{\mathfrak{s}}$ через переменные $x_{\mathfrak{s}+1}, \dots, x_n$:

$$x_{\mathfrak{s}} = \left(b_{\mathfrak{r}}^{[\mathfrak{r}-1]} - a_{\mathfrak{r},\mathfrak{s}+1}^{[\mathfrak{r}-1]} x_{\mathfrak{s}+1} - \dots - a_{\mathfrak{r},n}^{[\mathfrak{r}-1]} x_n \right) / a_{\mathfrak{r}\mathfrak{s}}^{[\mathfrak{r}-1]}. \tag{1.10}$$

Придавая в этой формуле переменным $x_{\mathfrak{s}+1}, \dots, x_n$ любой набор значений из $\mathbb{B}$:

$$x_{\mathfrak{s}+1} = \alpha_{\mathfrak{s}+1}, \dots, x_n = \alpha_n,$$

мы получим соответствующее значение для $x_{\mathfrak{s}}$:

$$x_{\mathfrak{s}} = \alpha_{\mathfrak{s}} \stackrel{\text{def}}{=} \left(b_{\mathfrak{r}}^{[\mathfrak{r}-1]} - a_{\mathfrak{r},\mathfrak{s}+1}^{[\mathfrak{r}-1]} \alpha_{\mathfrak{s}+1} - \dots - a_{\mathfrak{r},n}^{[\mathfrak{r}-1]} \alpha_n \right) / a_{\mathfrak{r}\mathfrak{s}}^{[\mathfrak{r}-1]}.$$

Рассмотрим теперь $(\mathfrak{r}-1)$-е уравнение системы (1.7). На основании все того же предположения со с.191, в этом уравнении имеется хотя бы один ненулевой коэффициент в левой части; пусть $a_{\mathfrak{r}-1,\mathfrak{k}}^{[\mathfrak{r}-2]} \ne 0$ — первый из них. Поскольку мы преположили, что система (1.7) является конечной стадией прямого хода метода Гаусса, то $\mathfrak{k} < \mathfrak{s}$, и переменная $x_{\mathfrak{k}}$ будет выражаться через переменные $x_{\mathfrak{k}+1}, \dots, x_n$. Снова различаются два случая. Если $\mathfrak{k} = \mathfrak{s} - 1$, то по фиксированным ранее значениям

$$x_{\mathfrak{s}} = \alpha_{\mathfrak{s}},\ x_{\mathfrak{s}+1} = \alpha_{\mathfrak{s}+1}, \dots, x_n = \alpha_n$$

значение переменной $x_{\mathfrak{k}}$ установится однозначно. Если же $\mathfrak{k} < \mathfrak{s}-1$, то переменным $x_{\mathfrak{k}+1},\dots,x_{\mathfrak{s}-1}$ могут быть приданы произвольные значения:

$$x_{\mathfrak{k}+1}=\alpha_{\mathfrak{k}+1},\dots,x_{\mathfrak{s}-1}=\alpha_{\mathfrak{s}-1},,$$

по которым величина $x_{\mathfrak{k}}$ установится однозначно. Произведем подстановку всех полученных значений переменных в $(\mathfrak{r}-2)$-е уравнение системы (1.7), и т.д. Во всей этой схеме нам на каком-то шаге обязательно встретится уравнение, в котором будут содержаться *по крайней мере две* переменные, значения которых еще не были зафиксированы на предыдущих шагах. Это следует из предположения, что число уравнений $\mathfrak{r}$ *меньше* числа неизвестных n. Такое уравнение допускает бесконечное число решений, любое из которых в ходе дальнейших шагов может быть "доделано" до решения системы (1.7). Таким образом, в этом случае система (1.7) и эквивалентная ей система (1.2) будет иметь бесконечное множество решений.

Определение 1.7. Процесс получения решения системы (1.2) из системы (1.7) называется **обратным ходом** метода Гаусса.

Пример 1.6. Решить систему **л.у.**

$$\textbf{а)}\ \left\{\begin{array}{rrrrl} x_1 & -2x_2 & +3x_3 & -4x_4 & = \ \ 4\\ & x_2 & -x_3 & +x_4 & =-3\\ x_1 & +3x_2 & & -3x_4 & = \ \ 1\\ & -7x_2 & +3x_3 & +x_4 & =-3\end{array}\right. ,\quad \textbf{б)}\ \left\{\begin{array}{rrrrl} x_1 & -2x_2 & -2x_3 & -x_4 & =-2\\ 2x_1 & -4x_2 & +3x_3 & -2x_4 & = \ \ 3\\ 3x_1 & -6x_2 & +5x_3 & -3x_4 & = \ \ 5\\ 4x_1 & -8x_2 & -3x_3 & -4x_4 & =-3.\end{array}\right.$$

Решение. Для **а)**:

$$\to \left\{\begin{array}{rrrrl} x_1 & -2x_2 & +3x_3 & -4x_4 & = \ \ 4\\ & x_2 & -x_3 & +x_4 & =-3\\ & 5x_2 & -3x_3 & +x_4 & =-3\\ & -7x_2 & +3x_3 & +x_4 & =-3\end{array}\right. \to \left\{\begin{array}{rrrrl} x_1 & -2x_2 & +3x_3 & -4x_4 & = \ \ 4\\ & x_2 & -x_3 & +x_4 & = -3\\ & & 2x_3 & -4x_4 & = \ \ 12\\ & & -4x_3 & +8x_4 & =-24\end{array}\right. \to$$

$$\to \left\{\begin{array}{rrrrl} x_1 & -2x_2 & +3x_3 & -4x_4 & = \ \ 4\\ & x_2 & -x_3 & +x_4 & =-3\\ & & 2x_3 & -4x_4 & = 12\\ & & & 0 & = \ \ 0\end{array}\right.$$

Придавая x_4 любые значения, из полученных уравнений последовательно выразим x_3, x_2 и x_1.

Перейти к оглавлению на странице: 460

Для **б)**:

$$\to \begin{cases} x_1 - 2\,x_2 - 2\,x_3 - x_4 = -2 \\ \qquad\qquad 7\,x_3 \qquad = \ \ 7 \\ \qquad\quad\ 11\,x_3 \qquad = 11 \\ \qquad\qquad 5\,x_3 \qquad = \ \ 5 \end{cases} \to \begin{cases} x_1 - 2\,x_2 - 2\,x_3 - x_4 = -2 \\ \qquad\qquad x_3 \qquad = \ \ 1 \\ \\ \\ \end{cases}$$

Второе уравнение дает единственное значение для x_3: $x_3 = 1$, подставив которое в первое, получим выражение для x_1 через переменные x_2 и x_4:

$$x_1 = 2\,x_2 + x_4.$$

Придавая x_2 и x_4 любые значения из $\mathbb{R}$, получим соответствующие значения для x_1.

Ответ. Система имеет бесконечное число решений, которые могут быть представлены формулами:

$$\textbf{а)}: \ x_1 = -8,\ x_2 = 3 + t,\ x_3 = 6 + 2\,t,\ x_4 = t \quad \text{для } \forall\, t \in \mathbb{R}.$$

$$\textbf{б)}: \ x_1 = 2\,t + u,\ x_2 = t,\ x_3 = 1,\ x_4 = u \quad \text{для } \forall\, \{t, u\} \subset \mathbb{R}.$$

Примеры показывают, что бесконечность множества решений может иметь структуру, различающуюся “размерностью”: она может быть “одномерной”, “двумерной” и т.д., в зависимости от количества параметров, вовлеченных в формулы, аналитически описывающие эту бесконечность.

Обобщим все полученные результаты.

Теорема 1.2. *Система уравнений* (1.2) *либо несовместна, либо имеет единственное решение, либо множество ее решений бесконечно.*

Задача. По коэффициентам системы (1.2) определить к какому из указанных в теореме случаев она относится и описать структуру ее множества решений.

Замечание 1.1. На примере системы

$$\begin{cases} 3\,x_1 + 2\,x_2 = 1, \\ 2\,x_1 + 3\,x_2 = 2 \end{cases}$$

проиллюстрируем типичную ошибку, допускаемую неопытными вычислителями. Вычтем из второго уравнения первое и, *одновременно*, из первого уравнения — второе:

$$\left\{\begin{array}{rl} x_1 - x_2 = & -1, \\ -x_1 + x_2 = & 1. \end{array}\right.$$

Последняя имеет бесконечное множество решений, но того же нельзя утверждать относительно исходной системы.

Замечание 1.2. В главе 1 мы интересовались существованием целочисленного решения уравнения $a_1x_1+a_2x_2=b$ при $\{a_1,a_2,b\}\subset\mathbb Z$. По своей постановке, эта задача представляет собой частный случай только что сформулированной. Однако методы ее решения существенно отличаются от тех, с которыми мы познакомимся в настоящей главе.

§2. Матрицы

Основные определения

Определение 2.1. **Матрица**[2] — это прямоугольная таблица, в каждой ячейке которой располагается некоторое число, называемое **элементом матрицы**. Будем обозначать матрицы прописными латинскими буквами $A,B,\dots,\mathfrak{A},\mathfrak{B},\dots$, а — при необходимости — их элементы буквами строчными $a,b,\dots,\mathfrak{a},\mathfrak{b},\dots$ Сами таблицы условимся ограничивать скобками — либо круглыми $(\ )$, либо квадратными $[\ ]$. Упорядоченная пара чисел

$$(\text{ количество строк, количество столбцов})$$

матрицы называется **порядком** (или **размерностью**) **матрицы**. Так, если матрица имеет m строк и n столбцов, то о ней говорят как о матрице “порядка m на n”, и записывают порядок в виде $m\times n$.

[2] matrix (*лат.*) — праматерь, первопричина, источник

Пример 2.1.

$$
\begin{array}{lll}
\textbf{а)}\ A = \begin{pmatrix} 2 & -1 & 7 \\ \boldsymbol{i} & 3 & 0 \end{pmatrix} & & \text{— матрица порядка } 2 \times 3; \\
\textbf{б)}\ B = \begin{pmatrix} 2 & \pi \\ -1 & \sqrt{3} \\ 7 & 0 \end{pmatrix} & & \text{— матрица порядка } 3 \times 2; \\
\textbf{в)}\ C = \begin{pmatrix} a, & 1, & b, & {}^3\!/_2 \end{pmatrix} & & \text{— матрица порядка } 1 \times 4; \\
\textbf{г)}\ D = [3] & & \text{— матрица порядка } 1 \times 1.
\end{array}
$$

Любая $m \times n$-матрица содержит всего mn элементов, и каждому из этих элементов можно поставить в соответствие "координаты его местоположения" в матрице, т.е. упорядоченную пару натуральных чисел (j, k), в которой первое число отвечает за номер строки элемента, а второе — за номер его столбца. При этом отсчет строк и столбцов будем вести, начиная от левого верхнего угла матрицы. Часто будет возникать необходимость записи матрицы общего вида, т.е. матрицы с элементами, числовые значения которых могут быть переменными (как это было в примере 2.1, **в**)). В самом общем случае — когда все элементы матрицы A могут быть произвольными — будем их записывать в виде a_{jk} или же, когда необходимо избежать недоразумения, в виде $a_{j,k}$. Так, a_{11} означает элемент матрицы A, стоящий в ее левом верхнем углу; $b_{10,3}$ — элемент матрицы B, стоящий в 10-й строке и 3-м столбце; $\mathfrak{c}_{m-3,2n-7}$ — элемент матрицы $\mathfrak{C}$, стоящий в $(m-3)$-й строке и $(2n-7)$-м столбце. Такая договоренность позволяет записывать компактно матрицы, для элементов которых имеется функциональная зависимость от местоположения в таблице. Так, к примеру, совершенно произвольную $m \times n$-матрицу мы можем записать в виде $A = [a_{jk}]_{\substack{j=1,\ldots,m \\ k=1,\ldots,n}}$.

Пример 2.2.

$$
A = \left[\frac{1}{j+k-1}\right]_{\substack{j=1,2,3,4 \\ k=1,2,3}} \quad \text{означает } 4\times 3 \text{ -матрицу} \quad \begin{bmatrix} 1 & {}^1\!/_2 & {}^1\!/_3 \\ {}^1\!/_2 & {}^1\!/_3 & {}^1\!/_4 \\ {}^1\!/_3 & {}^1\!/_4 & {}^1\!/_5 \\ {}^1\!/_4 & {}^1\!/_5 & {}^1\!/_6 \end{bmatrix}.
$$

Упражнение 1. *Найти развернутое выражение для матриц,* *представленных в компактной форме*

а) $A = [\max(j,k)]_{\substack{j=1,2,3,4\\k=1,2}}$; **б)** $B = [|j-k|]_{\substack{j=1,2,3\\k=1,2,3}}$; **в)** $C = [\delta_{jk}]_{\substack{j=1,\dots,5\\k=1,2,3}}$,

где δ_{jk} означает символ Кронекера[3].

Определение 2.2. Матрицы A и B называются **равными** если равны их порядки и совпадают элементы на соответствующих местах:

$$\text{если } A = [a_{jk}]_{\substack{j=1,\dots,m\\k=1,\dots,n}} \quad \text{и } B = [b_{jk}]_{\substack{j=1,\dots,m\\k=1,\dots,n}},$$

$$\text{то } A = B \iff a_{jk} = b_{jk} \; \forall j \in \{1,\dots,m\},\ k \in \{1,\dots,n\}.$$

Матрицы разных порядков не считаются равными.

Пример 2.3.

$$\begin{pmatrix} 2 & 1 & 11 & 4\\ 1 & -2 & 7 & 6\\ 11 & 7 & \sqrt{2} & 0 \end{pmatrix} \ne \begin{pmatrix} 2 & 1 & 11\\ 1 & -2 & 7\\ 11 & 7 & \sqrt{2}\\ 4 & 6 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0\\ 0 & 0 \end{pmatrix} \ne (0,0,0,0).$$

Еще одно упрощение записи — часто применяемое для матриц заранее не специфицированного порядка — заключается в том, что если некоторый участок матрицы занят равными нулю элементами, то они либо не указываются вовсе, либо вся их совокупность обозначается $\mathbb{O}$.

Пример 2.4.

$$\begin{pmatrix} 1 & 1\dots & 1 & 1 & \dots & 1\\ & 2\dots & 2 & 2 & \dots & 2\\ \mathbb{O} & \ddots & & & & \\ & & m-1 & m-1 & \dots & m-1\\ & & & m & \dots & m \end{pmatrix}_{m\times n} = \begin{pmatrix} 1 & 1\dots & 1 & 1 & \dots & 1\\ & 2\dots & 2 & 2 & \dots & 2\\ & \ddots & & & & \\ & & m-1 & m-1 & \dots & m-1\\ & & & m & \dots & m \end{pmatrix}_{m\times n}$$

Индексы внизу обозначают порядки матриц.

[3] См. раздел “Обозначения”.

Определение 2.3. **Произведением матрицы** A **на число** c называется матрица, каждый элемент которой равен произведению соответствующего элемента матрицы A на число c:

$$\text{если } A = [a_{jk}]_{\substack{j=1,\dots,m \\ k=1,\dots,n}}, \text{ то } c \cdot A \stackrel{\text{def}}{=} [ca_{jk}]_{\substack{j=1,\dots,m \\ k=1,\dots,n}}.$$

Пример 2.5.

$$3 \cdot \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 3 \\ -3 & 3 & 0 \\ 0 & 3 & 6 \end{pmatrix}, \quad 2 \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \qquad .$$

Определение 2.4. Для матриц A и B одного порядка их **суммой** называется матрица, каждый элемент которой равен сумме соответствующих элементов складываемых матриц:

$$\text{если } A = [a_{jk}]_{\substack{j=1,\dots,m \\ k=1,\dots,n}} \quad \text{и} \quad B = [b_{jk}]_{\substack{j=1,\dots,m \\ k=1,\dots,n}},$$
$$\text{то } A + B \stackrel{\text{def}}{=} [a_{jk} + b_{jk}]_{\substack{j=1,\dots,m \\ k=1,\dots,n}}.$$

Упражнение 2.

$$\begin{pmatrix} 1 & -2 & -1 \\ 2 & -1 & -2 \\ 1 & 2 & \sqrt{3} \\ 0 & 1 & 7 \end{pmatrix} + \begin{pmatrix} -1 & 3 & 2 \\ 1 & 0 & 1 \\ -1 & -2 & 3 \\ -1 & 1 & -5 \end{pmatrix} =$$

Свойства операций

Для произвольных матриц A, B, C одинакового порядка $m \times n$ и произвольных чисел c и d справедливо:

I. $(A + B) + C = A + (B + C)$;

II. $A + B = B + A$;

III. $\exists$ единственная матрица M такая, что $M + A = A$, именно $m \times n$-матрица, все элементы которой равны нулю:

$$M = \mathbb{O}_{m \times n} \stackrel{\text{def}}{=} \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}_{m \times n};$$

Перейти к оглавлению на странице: 460

будем называть эту матрицу **нулевой**;

IV. $\exists$ единственная матрица $\mathfrak{M}$ такая, что $\mathfrak{M}+A = \mathbb{O}_{m\times n}$, именно:

$$\mathfrak{M} = -A \overset{\text{def}}{=} -1 \cdot A = [-a_{jk}]_{\substack{j=1,\dots,m\\ k=1,\dots,n}};$$

V. $(c+d)A = cA + dA$;

VI. $c(A+B) = cA + cB$;

VII. $c(dA) = (cd)A$;

VIII. $\exists$ единственное число c, такое что $c \cdot A = A$, именно $c = 1$.

Определение 2.5. Если для элементов некоторого множества определены операции сложения и умножения на числа множества $\mathbb{A}$, и эти операции удовлетворяют аксиомам I – VIII, то это множество называется **линейным пространством**[4] **над** $\mathbb{A}$.

Теорема 2.1. *Множество $m\times n$-матриц образует линейное пространство над $\mathbb{A}$.*

Определение 2.6. Если $A_1,\dots,A_k$ — $m\times n$-матрицы, а $c_1,\dots,c_k$ — произвольные числа из $\mathbb{A}$, то матрица $c_1A_1+\dots+c_kA_k$ называется **линейной комбинацией** матриц $A_1,\dots,A_k$.

Пример 2.6. Задачу решения системы **л.у.** (1.2) можно интерпретировать следующим образом. Рассмотрим матрицы, состоящие из одного столбца

$$A_1 \overset{\text{def}}{=} \begin{pmatrix} a_{11}\\ a_{21}\\ \vdots\\ a_{m1}\end{pmatrix},\ A_2 \overset{\text{def}}{=} \begin{pmatrix} a_{12}\\ a_{22}\\ \vdots\\ a_{m2}\end{pmatrix},\dots,\ A_n \overset{\text{def}}{=} \begin{pmatrix} a_{1n}\\ a_{2n}\\ \vdots\\ a_{mn}\end{pmatrix},\ \mathcal{B} \overset{\text{def}}{=} \begin{pmatrix} b_1\\ b_2\\ \vdots\\ b_m\end{pmatrix},$$

Тогда систему (1.2) можно переписать в виде

$$x_1A_1 + x_2A_2 + \dots + x_nA_n = \mathcal{B},$$

[4]Общая теория линейных пространств будет излагаться в следующем семестре — в главе 1 раздела II.

и задача решения этой системы эквивалентна задаче подбора такой линейной комбинации столбцов $A_1, A_2, \dots, A_n$, которая бы давала столбец $\mathcal{B}$.

Определение 2.7. Для матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}_{m\times n}$$

ее **подматрицей** (**субматрицей**) будем называть матрицу

$$A\begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_\ell \\ \beta_1 & \beta_2 & \dots & \beta_s \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} a_{\alpha_1\beta_1} & a_{\alpha_1\beta_2} & \dots & a_{\alpha_1\beta_s} \\ a_{\alpha_2\beta_1} & a_{\alpha_2\beta_2} & \dots & a_{\alpha_2\beta_s} \\ \dots & & & \dots \\ a_{\alpha_\ell\beta_1} & a_{\alpha_\ell\beta_2} & \dots & a_{\alpha_\ell\beta_s} \end{pmatrix}_{\ell\times s}$$

где $1 \le \alpha_1 < \alpha_2 < \dots < \alpha_\ell \le m$, $1 \le \beta_1 < \beta_2 < \dots < \beta_s \le n$. Иначе говоря, это матрица, элементы которой стоят в матрице A на пересечении строк с номерами $\alpha_1, \alpha_2, \dots, \alpha_\ell$ и столбцов с номерами $\beta_1, \beta_2, \dots, \beta_s$. В частном случае, подматрицы

$$A_{[1]} \stackrel{\text{def}}{=} A\begin{pmatrix} 1\,2\dots\,m \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{pmatrix}, \dots,$$

$$A_{[n]} \stackrel{\text{def}}{=} A\begin{pmatrix} 1\,2\dots\,m \\ n \end{pmatrix} = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{pmatrix}$$

называются **столбцами** матрицы A (соответственно первым, и т.д., n-м), а матрицы

$$A^{[1]} \stackrel{\text{def}}{=} A\begin{pmatrix} 1 & & & \\ 1 & 2 & \dots & n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \end{pmatrix}, \dots,$$

$$A^{[m]} \stackrel{\text{def}}{=} A\begin{pmatrix} m & & & \\ 1 & 2 & \dots & n \end{pmatrix} = \begin{pmatrix} a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

— **строками** матрицы A (соответственно, первой, и т.д., m-й). Для избежания путаницы элементы строк иногда отделяют запятыми. Таким образом, можно образно сказать, что сама матрица A представляет

строку своих столбцов: $A = \left[A_{[1]}, A_{[2]}, \ldots, A_{[n]}\right]$; или же столбец своих строк: $A = \begin{pmatrix} A^{[1]} \\ \vdots \\ A^{[m]} \end{pmatrix}$.

Замечание 2.1. В современной литературе операция составления новой матрицы из двух (и более) меньшего порядка называется **конкатенацией**[5] и обозначается $\big|$. Так, например,

$$\text{если } A = \begin{pmatrix} 1 & 2 \\ -1 & -2 \\ 1 & 7 \end{pmatrix} \text{ и } B = \begin{pmatrix} 2 & 1 & 3 \\ -2 & 0 & 1 \\ 7 & -1 & 4 \end{pmatrix},$$

$$\text{то } A\big|B = \begin{pmatrix} 1 & 2 & 2 & 1 & 3 \\ -1 & -2 & -2 & 0 & 1 \\ 1 & 7 & 7 & -1 & 4 \end{pmatrix}.$$

Умножение матриц

Определение 2.8. Для матрицы-строки $U = (u_1, \ldots, u_n)$ и матрицы-столбца $V = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ определим произведение $U \cdot V$ как число

$$U \cdot V \stackrel{\text{def}}{=} u_1 v_1 + \cdots + u_n v_n \in \mathbb{A}.$$

Для произвольных матриц A и B произведение матрицы A на матрицу B определяется тогда и только тогда, когда их порядки связаны ограничением:

$$\text{количество столбцов матрицы } A = \text{ количество строк матрицы } B; \tag{2.1}$$

т.е. если матрица A имеет порядок $m \times n$, то матрица B может иметь порядок $n \times k$ при $\forall k \in \mathbb{N}$. В этом случае произведение

[5] concatenation (*англ.*) — сцепление.

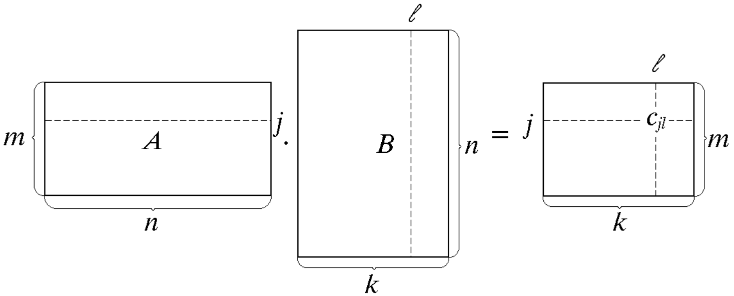
матрицы A на матрицу B обозначается $A \cdot B$ и представляет собой матрицу C порядка $m \times k$:

$$\underset{m\times k}{C} = \underset{m\times n}{A} \cdot \underset{n\times k}{B},$$

элементы которой вычисляются по следующему правилу

$$C = [c_{j\ell}]_{\substack{j=1,\ldots,m \\ \ell=1,\ldots,k}}, \text{ где } c_{j\ell} \overset{\text{def}}{=} A^{[j]}B_{[\ell]} = a_{j1}b_{1\ell} + a_{j2}b_{2\ell} + \cdots + a_{jn}b_{n\ell}. \tag{2.2}$$

Таким образом, элемент, стоящий в j-й строке и ℓ-м столбце матрицы C, равен произведению j-й строки матрицы A на ℓ-й столбец матрицы B.



Замечание 2.2. При этом произведение $B \cdot A$ может и не быть определено!

Пример 2.7.

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 0 \\ 3 & 7 \end{pmatrix}, B = \begin{pmatrix} \boldsymbol{i} & 0 & 0 & -1 \\ 4 & 2 & 0 & -2 \end{pmatrix} \implies A \cdot B = \begin{pmatrix} 8+\boldsymbol{i} & 4 & 0 & -5 \\ -\boldsymbol{i} & 0 & 0 & 1 \\ 28+3\boldsymbol{i} & 14 & 0 & -17 \end{pmatrix}$$

$$A = \begin{pmatrix} 3 & -1 & -1 \\ 2 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}, B = \begin{pmatrix} 2 & 1 \\ -1 & 0 \\ 0 & 1 \end{pmatrix} \implies A \cdot B =$$

$$A = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, B = (1, 2, -1, -2) \implies A \cdot B =$$

△

Упражнение 3. *Доказать, что*

$$AB = \left[AB_{[1]},\, AB_{[2]}, \dots, AB_{[k]}\right], \tag{2.3}$$

т.е. умножение матрицы A на матрицу B может быть произведено посредством умножения матрицы A на каждый из столбцов матрицы B.

Выясним теперь смысл операции умножения матриц с точки зрения системы **л.у.** (1.2). Эту систему можно переписать в матричном виде

$$AX{=}\mathcal{B}, \text{ где } X \stackrel{\text{def}}{=} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ — столбец неизвестных, а } \mathcal{B}{=}\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \text{ — столбец правых частей.} \tag{2.4}$$

Сделаем в системе (1.2) замену переменных (подстановку)

$$\begin{cases} x_1 = b_{11}y_1 + b_{12}y_2 + \dots + b_{1k}y_k, \\ \vdots \\ x_n = b_{n1}y_1 + b_{n2}y_2 + \dots + b_{nk}y_k, \end{cases} \tag{2.5}$$

которая в матричном виде может быть записана так

$$X = BY, \text{ где } B \stackrel{\text{def}}{=} \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1k} \\ \dots & & & \dots \\ b_{n1} & b_{n2} & \dots & b_{nk} \end{pmatrix}, \text{ а } Y \stackrel{\text{def}}{=} \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix}.$$

Подстановка (2.5) в j-е уравнение системы (1.2) приводит к следующему:

$$\begin{aligned} a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jn}x_n &= a_{j1}(b_{11}y_1 + b_{12}y_2 + \dots + b_{1k}y_k) &+ \\ &+ a_{j2}(b_{21}y_1 + b_{22}y_2 + \dots + b_{2k}y_k) &+ \\ &+ \qquad \dots &+ \\ &+ a_{jn}(b_{n1}y_1 + b_{n2}y_2 + \dots + b_{nk}y_k) &= \\ &= (a_{j1}b_{11} + a_{j2}b_{21} + \dots + a_{jn}b_{n1})y_1 &+ \\ &+ (a_{j1}b_{12} + a_{j2}b_{22} + \dots + a_{jn}b_{n2})y_2 &+ \\ &+ \qquad \dots &+ \\ &+ (a_{j1}b_{1k} + a_{j2}b_{2k} + \dots + a_{jn}b_{nk})y_k &= \\ = c_{j1}y_1 + c_{j2}y_2 + \dots + c_{jk}y_k, \end{aligned}$$

где коэффициенты $c_{j\ell}$ определяются формулой (2.2). Таким образом, правило умножения матрицы A на матрицу B, заданное формулой (2.2), соответствует правилу формирования новой матрицы системы **л.у.** при замене в этой системе переменных по формулам (2.5). Кроме того, в ходе вывода этого соответствия мы попутно доказали одно матричное равенство

$$A(BY) = (AB)Y \text{ для любого } Y = \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix} \tag{2.6}$$

Теорема 2.2. *Операция умножения матриц является ассоциативной: для любых матриц $A_{m\times n}, B_{n\times k}$ и $D_{k\times h}$ выполняется*

$$(AB)D = A(BD).$$

Доказательство. Действительно, формула (2.3) позволяет записать произведение матрицы AB на матрицу D посредством оперирования над столбцами матрицы D:

$$(AB)D = \left[(AB)D_{[1]}, \dots, (AB)D_{[h]}\right] =$$

На основании формулы (2.6) имеем право записать

$$(AB)D_{[1]} = A\left(BD_{[1]}\right), \ldots, (AB)D_{[h]} = A\left(BD_{[h]}\right).$$

Следовательно, наше произведение

$$= \left[A\left(BD_{[1]}\right), \ldots, A\left(BD_{[h]}\right)\right] =$$

Опять-таки, формула (2.3) позволяют переписать последнюю матрицу:

$$= A\left[BD_{[1]}, \ldots, BD_{[h]}\right] =$$

и завершит доказательство еще одно применение формулы (2.3):

$$= A(BD).$$

■

Операция умножения матриц, как правило, некоммутативна.

Если определено произведение A на B, то не обязательно определено произведение B на A. Действительно, правило умножения матриц (2.1) утверждает, что произведения A на B и B на A одновременно определены тогда и только тогда, когда при порядке матрицы A равном $m \times n$ порядок матрицы B равен $n \times m$.

Пример 2.8.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \end{pmatrix}, \ B = \begin{pmatrix} 7 & 0 \\ 1 & 0 \\ -1 & 1 \end{pmatrix} \Rightarrow$$

$$\Rightarrow \ AB = \begin{pmatrix} 6 & 3 \\ -8 & 1 \end{pmatrix}, \ BA = \begin{pmatrix} 7 & 14 & 21 \\ 1 & 2 & 3 \\ -2 & -2 & -2 \end{pmatrix}.$$

Видим, что матрицы AB и BA имеют разные порядки, т.е. $AB \neq BA$.

Матрицы AB и BA имеют одинаковый порядок тогда и только тогда, когда обе матрицы A и B имеют одинаковый порядок вида $n \times n$.

Определение 2.9. Матрица A называется **квадратной**, если количество ее строк равно количеству ее столбцов. Для квадратных $n \times n$-матриц будем их порядок обозначать одной буквой n, а записывать компактно — в виде $A = [a_{jk}]_{j,k=1}^{n}$.

Если матрицы A и B — квадратные порядка n, то обе матрицы AB и BA являются тоже квадратными порядка n. Тем не менее, и в этом случае, как правило $AB \neq BA$.

Пример 2.9.

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \ B = \begin{pmatrix} 1 & 2 \\ 5 & 3 \end{pmatrix} \ \Rightarrow \ AB = \begin{pmatrix} 11 & 8 \\ 23 & 18 \end{pmatrix}, \ BA = \begin{pmatrix} 7 & 10 \\ 14 & 22 \end{pmatrix}.$$

Определение 2.10. Говорят, что матрицы A и B **коммутируют** (или **перестановочны**), если $AB = BA$.

Существуют матрицы, коммутирующие с любой матрицей. Одна из таких матриц — нулевая: $\mathbb{O}_n \cdot A_n = A_n \cdot \mathbb{O}_n = \mathbb{O}_n$. Еще одна матрица, коммутирующая с любой матрицей A, имеет специальное название.

Определение 2.11. Матрица

$$E_n \stackrel{\text{def}}{=} \begin{pmatrix} 1 & & & \\ & 1 & & \mathbb{O} \\ \mathbb{O} & & \ddots & \\ & & & 1 \end{pmatrix}_n = [\delta_{jk}]_{j,k=1}^{n}$$

называется **единичной** (**порядка** n).

Упражнение 4. *Доказать, что единичная матрица играет роль единицы при умножении матриц:* $EA = AE = A$ *для любой* $n \times n$*-матрицы* A.

Интерпретация последнего результата в терминах линейной замены переменных (2.5) очевидна:

$$X = EY \iff x_1 = y_1, \ldots, x_n = y_n,$$

т.е. мы имеем дело с тождественной заменой переменных[6], которая, очевидно, не меняет матрицу линейной системы (1.2).

[6] В западной литературе единичная матрица обозначается I_n — по первой букве слова **identity** (*англ.*) — тождество.

Итак, умножение матриц ассоциативно, но не коммутативно. Будут ли выполняться другие свойства относительно операций сложения и умножения — из тех, что отмечены на с.58?

Упражнение 5. *Доказать, что для любых матриц, для которых определены указанные ниже операции, справедливы следующие свойства:*

$$\textbf{а)}(cA)B = A(cB) = cAB;\ \textbf{б)}(A_1 + A_2)B = A_1B + A_2B;$$
$$\textbf{в)}A(B_1 + B_2) = AB_1 + AB_2.$$

Определение 2.12. Матрицу

$$A^K \stackrel{\text{def}}{=} \begin{cases} \overbrace{A \cdot \ldots \cdot A}^{K \text{ раз}} & \text{при } K > 0, \\ E & \text{при } K = 0 \end{cases}$$

назовем K**-й степенью** матрицы A.

Наличие для матриц ассоциативного закона делает это определение корректным: перемножение матриц $A \cdot \ldots \cdot A$ можем производить, расставив скобки как пожелаем.

Упражнение 6. *Доказать, что степени одной и той же матрицы A коммутируют:*

$$A^K A^L = A^L A^K = A^{K+L} \quad \text{для } \{K, L\} \subset \mathbb{N}.$$

Матричные уравнения

Еще одним способом задания матрицы является задание соотношения, которому матрица должна удовлетворять.

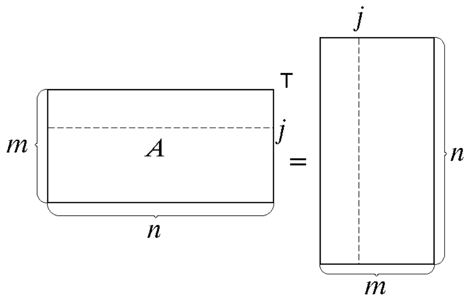
Определение 2.13. Преобразование матрицы, при котором ее строки становятся столбцами новой матрицы, называется **транспонированием**:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ \ldots & & & & \ldots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{pmatrix}^{\top} = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ a_{12} & a_{22} & \ldots & a_{m2} \\ a_{13} & a_{23} & \ldots & a_{m3} \\ \vdots & & & \vdots \\ a_{1n} & a_{2n} & \ldots & a_{mn} \end{pmatrix}$$

В компактном виде:

$$\left([a_{jk}]_{\substack{j=1,\ldots,m\\ k=1,\ldots,n}}\right)^{\top} = [a_{kj}]_{\substack{k=1,\ldots,n\\ j=1,\ldots,m}}\,,$$

а в графическом:



Замечание 2.3. В литературе встречаются следующие обозначения для операции транспонирования:

$$A^t = {}^tA = A' = A^*.$$

Упражнение 7. *Доказать следующие свойства операции транспонирования:*

$$\textbf{а)}\,\left(A^{\top}\right)^{\top} = A;\ \textbf{б)}(A+B)^{\top} = A^{\top} + B^{\top};\ \textbf{в)}(cA)^{\top} = cA^{\top};$$
$$\textbf{г)}(AB)^{\top} = B^{\top}A^{\top}$$

при условии, что все перации определены.

Определение 2.14. **Главной диагональю** квадратной матрицы A_n называется ее диагональ, идущая из левого верхнего в правый нижний угол, т.е. эта диагональ содержит элементы $a_{11}, \ldots, a_{jj}, \ldots, a_{nn}$.

Определение 2.15. Матрица A называется **симметричной** если она удовлетворяет соотношению

$$A = A^{\top}.$$

Из определения следует, что симметричная матрица может быть только квадратной, а ее элементы должны удовлетворять соотношению:

$$a_{jk} = a_{kj} \quad \text{для } \{j,k\} \subset \{1,\dots,n\}.$$

Иными словами, симметричная матрица — это такая матрица, которая симметрична относительно своей главной диагонали.

Упражнение 8. *Сколько элементов надо задать, чтобы однозначно определить симметричную матрицу порядка n?*

Частным случаем симметричной матрицы является **диагональная** матрица:

$$D = \begin{pmatrix} a_{11} & & & \\ & a_{22} & & \mathbb{O} \\ \mathbb{O} & & \ddots & \\ & & & a_{nn} \end{pmatrix}.$$

Определение 2.16. Матрица A называется **кососимметричной** если она удовлетворяет соотношениюindexМатрица!кососимметричная

$$A = -A^{\top}.$$

Из определения следует, что кососимметричная матрица может быть только квадратной, а ее элементы должны удовлетворять соотношению:

$$a_{jk} = -a_{kj} \quad \text{для } \{j,k\} \subset \{1,\dots,n\}.$$

Отсюда вытекает, что все элементы главной диагонали кососимметричной матрицы должны быть равны 0.

Упражнение 9. *Указать все элементы кососимметричной матрицы*

$$\begin{pmatrix} & 1 & \\ & & 3 \\ -2 & & \end{pmatrix}_{3\times 3}.$$

Упражнение 10. *Доказать, что при любой квадратной матрице* A

а) *матрица* $A + A^\top$ *будет симметричной;*

б) *матрица* $A - A^\top$ *будет кососимметричной.*

Перейдем теперь к более сложным соотношениям, задающим матрицу.

Матрицу можно задать матричным уравнением. Самыми простыми являются **линейные**

$$A \cdot X = B \quad \text{или} \quad Y \cdot A = C \tag{2.7}$$

при известных матрицах A — порядка $m \times n$, B — $m \times k$ и C — $\ell \times n$. Требуется определить неизвестные матрицы X или Y. Из двух видов матричных уравнений (1.2) задачу решения одного можно свести к задаче решения другого. Так, на основании результата упражнения 7 **г)**, имеем:

$$Y \cdot A = C \iff A^\top Y^\top = C^\top .$$

Получившееся матричное уравнение относится к первому типу уравнений (неизвестная матрица $Y^\top$ расположена справа от умножаемой на нее матрицы). Обычно рассматривают уравнения именно такого вида, и их частным случаем является система **л.у.** (1.2). С другой стороны, решение уравнения $AX = B$ сводится к решению подходящей системы **л.у.**

Пример 2.10. Решить матричное уравнение

$$\underbrace{\left(\begin{array}{rrr} 3 & 1 & 4 \\ -7 & -6 & 2 \\ 7 & 5 & 1 \\ 3 & 0 & 7 \end{array}\right)}_{A} X = \underbrace{\left(\begin{array}{rr} 1 & 0 \\ 0 & -1 \\ 0 & 1 \\ 1 & 0 \end{array}\right)}_{B} .$$

Решение. Матрица X должна иметь порядок 3×2. Подставим ее в уравнение в виде

$$X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{pmatrix} ,$$

домножим слева на матрицу A:

$$\begin{pmatrix} 3x_{11} & +x_{21} & +4x_{31}, & 3x_{12} & +x_{22} & +4x_{32} \\ -7x_{11} & -6x_{21} & +2x_{31}, & -7x_{12} & -6x_{22} & +2x_{32} \\ 7x_{11} & +5x_{21} & +x_{31}, & 7x_{12} & +5x_{22} & +x_{32} \\ 3x_{11} & & +7x_{31}, & 3x_{12} & & +7x_{32} \end{pmatrix}$$

и приравняем матрице B. Получим две системы **л.у.** относительно столбцов матрицы X:

$$\left\{ \begin{array}{rrrl} 3x_{11} & +x_{21} & +4x_{31} & = 1, \\ -7x_{11} & -6x_{21} & +2x_{31} & = 0, \\ 7x_{11} & +5x_{21} & +x_{31} & = 0, \\ 3x_{11} & & +7x_{31} & = 1; \end{array} \right. \quad \text{и} \quad \left\{ \begin{array}{rrrl} 3x_{12} & +x_{22} & +4x_{32} & = 0, \\ -7x_{12} & -6x_{22} & +2x_{32} & = -1, \\ 7x_{12} & +5x_{22} & +x_{32} & = 1, \\ 3x_{12} & & +7x_{32} & = 0. \end{array} \right.$$

Поскольку левые части этих систем идентичны, то прямой ход метода Гаусса будет давать одинаковые коэффициенты в левых частях. Опуская промежуточные вычисления, запишем окончательные варианты систем:

$$\left\{ \begin{array}{rrrl} 3x_{11} & +x_{21} & +4x_{31} & = 1, \\ & -x_{21} & +3x_{31} & = 0, \\ & & -{}^1\!/_3\, x_{31} & = -{}^7\!/_3; \end{array} \right. \quad \text{и} \quad \left\{ \begin{array}{rrrl} 3x_{12} & +x_{22} & +4x_{32} & = 0, \\ & -x_{22} & +3x_{32} & = 0, \\ & & -{}^1\!/_3\, x_{32} & = 1. \end{array} \right.$$

Ответ.

$$X = \begin{pmatrix} -16 & 7 \\ 21 & -9 \\ 7 & -3 \end{pmatrix}.$$

Определение 2.17. Решение уравнения $AX = E$ для квадратной матрицы A называется **правой обратной матрицей** для матрицы A, а решение уравнения $YA = E$ — **левой обратной матрицей** для матрицы A.

Вопрос о существовании правой и левой обратной матрицы пока открыт: примеры показывают, что для некоторых матриц A обратные существуют, а для других — нет.

Пример 2.11. Найти правую обратную матрицу для матрицы

$$\textbf{а)} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}; \ \textbf{б)} \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}.$$

Решение. Уравнение $AX = E$ запишем в привычном виде системы **л.у.** относительно элементов матрицы

$$X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}.$$

Эта система в случае **а)**

$$\left\{ \begin{array}{ll} x_{11} + x_{21} & = 1, \\ x_{11} + x_{21} & = 0; \end{array} \right. \qquad \left\{ \begin{array}{ll} x_{12} + x_{22} & = 0, \\ x_{12} + x_{22} & = 1 \end{array} \right.$$

оказывается несовместной, а в случае **б)**

$$\left\{ \begin{array}{rl} x_{11} + 2x_{21} & = 1, \\ x_{11} + x_{21} & = 0; \end{array} \right. \qquad \left\{ \begin{array}{rl} x_{12} + 2x_{22} & = 0, \\ x_{12} + x_{22} & = 1 \end{array} \right.$$

имеет единственное решение: $x_{11} = -1$, $x_{12} = 2$, $x_{21} = 1$, $x_{22} = -1$.

Ответ. **а)** Не $\exists$;

$$\textbf{б)} \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}.$$

В §10 будет установлено условие существования правой обратной матрицы, а также будет доказан следующий результат.

Теорема 2.3. *Если существует правая обратная матрица для матрицы A, то она единственна, и одновременно является левой обратной.*

Этот результат позволяет ввести следующее определение.

Определение 2.18. В случае ее существования, матрица, удовлетворяющая матричному равенству $AX = E$ называется **обратной матрицей** для матрицы A и обозначается A^{-1}.

Для задачи решения системы (1.2) при числе уравнений равном числу неизвестных, польза от знания матрицы A^{-1} может быть следующей: формула $X = A^{-1}\mathcal{B}$ определяет решение *любой* системы (1.2) при фиксированной матрице A и варьируемом столбце правых частей $\mathcal{B}$.

Упражнение 11. *Вычислить*

$$3\left[\begin{pmatrix} 2 & 3 \\ 2 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ -3 & -2 & -1 \end{pmatrix} + E\right]^{-1} - \left[\left([j-k]_{j,k=1}^{3}\right)^3 + [j+k \pmod 3]_{j,k=1}^{3}\right]^{\top}.$$

Из-за отсутствия коммутативности умножения матриц, встречаются линейные матричные уравнения более сложного вида.

Определение 2.19. Уравнение

$$A^\top X + XA = B$$

относительно неизвестной матрицы X называется **уравнением Ляпунова**; оно имеет важное значение в теории управления.

Можно поставить и задачу решения *нелинейных* матричных уравнений.

Упражнение 12. *Решить матричное уравнение*

$$X^2 = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Однако нашей основной задачей будет задача решения именно линейных матричных уравнений.

§3. Определение определителя

Определители малых порядков

В примерах 1.2 и 1.4 при решении (или преобразовании к эквивалентным) систем линейных уравнений общего вида, т.е. с буквенными коэффициентами, возникали некоторые выражения, зависящие от этих коэффициентов.

Определение 3.1. **Определитель** или **детерминант**[7] матрицы первого порядка считается равным единственному элементу этой матрицы:

$$\det(a_{11}) \stackrel{\text{def}}{=} a_{11}.$$

Определитель матрицы второго порядка определяется формулой

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \stackrel{\text{def}}{=} a_{11}a_{22} - a_{12}a_{21}. \tag{3.1}$$

[7] determinator (*лат.*) — определяющий.

Для этого определителя закреплено еще одно обозначение, когда скобки, ограничивающие матрицу, заменяются вертикальными чертами:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \stackrel{\text{def}}{=} a_{11}a_{22} - a_{12}a_{21}.$$

Это выражение встретилось нам при решении примера 1.2. Отличие его от нуля гарантирует существование и единственность решения системы линейных уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 & = b_1, \\ a_{21}x_1 + a_{22}x_2 & = b_2. \end{cases}$$

Более того, само это решение — т.е. формулы (1.3) — тоже можно записать с помощью определителей:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \ x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}. \tag{3.2}$$

Формально, определители в числителе получаются из определителя, стоящего в знаменателе, заменой столбцов на столбец правых частей системы. Для получения первой компоненты решения заменяется первый столбец, для получения второй компоненты — заменяется второй.

Пример 3.1. Решить систему уравнений

$$\begin{cases} 3\,x_1 & +4\,x_2 & = 1, \\ -x_1 & +2\,x_2 & = -5. \end{cases}$$

Решение.

$$\begin{vmatrix} 3 & 4 \\ -1 & 2 \end{vmatrix} = 3 \cdot 2 - 4 \cdot (-1) = 10 \neq 0.$$

Формулы (3.2) можно применять:

$$x_1 = \frac{\begin{vmatrix} 1 & 4 \\ -5 & 2 \end{vmatrix}}{10} = \frac{1 \cdot 2 - (-5) \cdot 4}{10} = \frac{11}{5}, \ x_2 = \frac{\begin{vmatrix} 3 & 1 \\ -1 & -5 \end{vmatrix}}{10} = -\frac{7}{5}.$$

$\triangle$

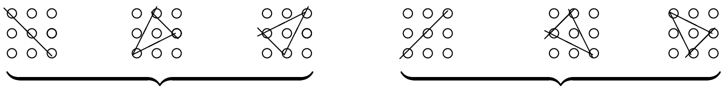
Перейти к оглавлению на странице: 460

Определитель матрицы третьего порядка определяется формулой

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \overset{\text{def}}{=} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \overset{\text{def}}{=}$$

$$\overset{\text{def}}{=} a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{21}a_{32}a_{13}-a_{31}a_{22}a_{13}-a_{21}a_{12}a_{33}-a_{11}a_{32}a_{23}. \tag{3.3}$$

Для запоминания этой формулы используются следующие диаграммы



со знаком + со знаком -

Выражение (3.3) также возникает при решении системы линейных уравнений. Рассмотрим систему (1.9).

$$\begin{array}{l|l} \quad\text{Уравнения} & \text{умножим на} \\ \left\{\begin{array}{llll} a_{11}x_1+ & a_{12}x_2+ & a_{13}x_3 = & b_1 \\ a_{21}x_1+ & a_{22}x_2+ & a_{23}x_3 = & b_2 \\ a_{31}x_1+ & a_{32}x_2+ & a_{33}x_3 = & b_3 \end{array}\right. & \begin{array}{l} \times(a_{22}a_{33}-a_{23}a_{32}) \\ \times-(a_{12}a_{33}-a_{13}a_{32}) \\ \times(a_{12}a_{23}-a_{13}a_{22}) \end{array} \end{array}$$

Сложив полученное, придем в левой части к

$$[a_{11}(a_{22}a_{33}-a_{23}a_{32})-a_{21}(a_{12}a_{33}-a_{13}a_{32})+a_{31}(a_{12}a_{23}-a_{13}a_{22})]x_1 +$$
$$+ [a_{12}(a_{22}a_{33}-a_{23}a_{32})-a_{22}(a_{12}a_{33}-a_{13}a_{32})+a_{32}(a_{12}a_{23}-a_{13}a_{22})]x_2 +$$
$$+ [a_{13}(a_{22}a_{33}-a_{23}a_{32})-a_{23}(a_{12}a_{33}-a_{13}a_{32})+a_{33}(a_{12}a_{23}-a_{13}a_{22})]x_3 =$$

$$= x_1\Delta + 0\cdot x_2 + 0\cdot x_3,$$

а в правой — к

$$= b_1(a_{22}a_{33}-a_{23}a_{32}) - b_2(a_{12}a_{33}-a_{13}a_{32}) + b_3(a_{12}a_{23}-a_{13}a_{22}).$$

Коэффициент Δ при x_1 в точности совпадает с определителем (3.3); это выражение уже встречалось нам на с. 192. Но и выражение в правой части тоже можно рассматривать как определитель

— оно равно
$$B_1 \stackrel{\text{def}}{=} \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}.$$

Снова произведем домножение уравнений системы (1.9):

$$\begin{array}{c|l} \text{Уравнения} & \text{умножим на} \\ \left\{\begin{array}{l} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{array}\right. & \begin{array}{l} \times -(a_{21}a_{33} - a_{31}a_{23}) \\ \times (a_{11}a_{33} - a_{13}a_{31}) \\ \times -(a_{11}a_{23} - a_{21}a_{13}) \end{array} \end{array}$$

Сложим полученное:

$$\begin{gathered} 0 \cdot x_1 + \\ +[-(a_{21}a_{33} - a_{31}a_{23})a_{12} + (a_{11}a_{33} - a_{13}a_{31})a_{22} - (a_{11}a_{23} - a_{21}a_{13})a_{32}]x_2 + \\ +0 \cdot x_3 = -b_1(a_{21}a_{33} - a_{31}a_{23}) + b_2(a_{11}a_{33} - a_{13}a_{31}) - b_3(a_{11}a_{23} - a_{21}a_{13}). \end{gathered}$$

Коэффициент при x_2 снова совпадает с определителем (3.3), а в правой части стоит определитель
$$B_2 \stackrel{\text{def}}{=} \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}.$$

Наконец, домножим первое из уравнений системы (1.9) на $(a_{21}a_{32} - a_{31}a_{22})$, второе — на $-(a_{11}a_{32} - a_{31}a_{12})$, третье — на $(a_{11}a_{22} - a_{21}a_{12})$ и сложим:

$$\begin{gathered} 0 \cdot x_1 + 0 \cdot x_2 + \\ +[(a_{21}a_{32} - a_{31}a_{22})a_{13} - (a_{11}a_{32} - a_{31}a_{12})a_{23} + (a_{11}a_{22} - a_{21}a_{12})a_{33}]x_3 = \\ = (a_{21}a_{32} - a_{31}a_{22})b_1 - (a_{11}a_{32} - a_{31}a_{12})b_2 + (a_{11}a_{22} - a_{21}a_{12})b_3. \end{gathered}$$

Коэффициент при x_3 снова совпадает с определителем (3.3), а в правой части стоит определитель
$$B_3 \stackrel{\text{def}}{=} \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}.$$

Итак, если решение системы (1.9) существует, то оно должно удовлетворять равенствам

$$x_1\Delta = B_1,\ x_2\Delta = B_2,\ x_3\Delta = B_3.$$

Теорема 3.1. *Если определитель (3.3) отличен от нуля, то система (1.9) имеет единственное решение, которое определяется по* **формулам Крамера**

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}},\ x_2 = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}},\ x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}. \quad (3.4)$$

Правило формирования дробей остается таким же, как и для случая системы из двух уравнений: j-я компонента решения в знаменателе имеет определитель матрицы A, а в числителе — определитель, получающийся из него заменой j-го столбца на столбец правых частей.

Доказательство того факта, что при условии существования решения, оно должно однозначно определяться формулами (3.4), уже проведено. Осталось показать, что формулы (3.4) действительно задают решение системы (1.9). Подставим эти значения в левую часть первого из уравнений:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 =$$

$$= \frac{a_{11}}{\Delta}[b_1(a_{22}a_{33}-a_{23}a_{32})-b_2(a_{12}a_{33}-a_{13}a_{32})+b_3(a_{12}a_{23}-a_{13}a_{22})] +$$
$$+ \frac{a_{12}}{\Delta}[-b_1(a_{21}a_{33}-a_{23}a_{32})+b_2(a_{11}a_{33}-a_{13}a_{31})-b_3(a_{11}a_{23}-a_{13}a_{21})] +$$
$$+ \frac{a_{13}}{\Delta}[b_1(a_{21}a_{32}-a_{22}a_{31})-b_2(a_{11}a_{32}-a_{12}a_{31})+b_3(a_{11}a_{22}-a_{12}a_{21})] =$$

$$\frac{1}{\Delta}[b_1\Delta + b_2 \cdot 0 + b_3 \cdot 0] = b_1.$$

Таким образом, определяемые формулами (3.4) значения переменных удовлетворяют первому из уравнений системы (1.9). Аналогично показывается справедливость оставшихся равенств. ■

Перейти к оглавлению на странице: 460

Выводы:

- Для системы линейных уравнений

$$A_{n\times n}X_{n\times 1}=\mathcal{B}_{n\times 1}\ \text{ при } n\in\{1,2,3\}$$

отличие от нуля определителя матрицы A является достаточным условием существования и единственности решения этой системы. Само это решение может быть записано также с помощью определителей.

- Определитель является функцией элементов матрицы A: его величина зависит от n^2 элементов этой матрицы. Эта зависимость полиномиальная: $\det(A)$ является однородным полиномом степени n относительно $a_{11}, a_{12}, \ldots, a_{nn}$ и линейным по каждой из этих переменных:

$$\deg(\det(A))=n,\ \deg_{a_{jk}}(\det(A))=1\ \text{ для } \{j,k\}\subset\{1,\ldots,n\}.$$

- Определитель порядка $n\geq 2$, у которого два столбца одинаковы, равен нулю.

Справедливость последнего вывода ранее не утверждалась явным образом. Тем не менее, легко проверить его справедливость для матрицы второго порядка. Для матрицы же третьего порядка частичная проверка справедливости этого вывода была проведена на с. 217: выражения

$$a_{12}(a_{22}a_{33}-a_{23}a_{32})-a_{22}(a_{12}a_{33}-a_{13}a_{32})+a_{32}(a_{12}a_{23}-a_{13}a_{22})$$

и

$$a_{13}(a_{22}a_{33}-a_{23}a_{32})-a_{23}(a_{12}a_{33}-a_{13}a_{32})+a_{33}(a_{12}a_{23}-a_{13}a_{22}),$$

с одной стороны равны 0, а, с другой стороны, совпадают, соответственно, с определителями

$$\begin{vmatrix} a_{12} & a_{12} & a_{13}\\ a_{22} & a_{22} & a_{23}\\ a_{32} & a_{32} & a_{33}\end{vmatrix}\quad \text{и} \quad \begin{vmatrix} a_{13} & a_{12} & a_{13}\\ a_{23} & a_{22} & a_{23}\\ a_{33} & a_{32} & a_{33}\end{vmatrix}.$$

Наша задача теперь заключается в распространении понятия определителя на случай матриц бо́льших порядков. Можно попытаться уловить закономерность в формировании определителей второго и третьего порядков и предложить затем гипотезу вычисления определителя четвертого порядка. С этой целью, проанализируем представление определителя третьего порядка, полученное на с. 217:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22}).$$

Его можно переписать в эквивалентном виде с помощью определителей второго порядка:

$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \tag{3.5}$$

Три определителя формально получаются вычеркиванием из определителя матрицы A первого столбца и, соответственно, первой, второй и третьей строк. Тем самым формула (3.5) позволяет свести вычисление определителя третьего порядка к вычислению трех определителей второго порядка. Она называется **формулой разложения определителя по первому столбцу**. Аналогичные формулы можно выписать и для второго и третьего столбцов:

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

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

Попробуем распространить такой подход на определитель

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}.$$

Разложим его по первому столбцу:

$$= \overset{?}{\pm} a_{11} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} \overset{?}{\pm} a_{21} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} \overset{?}{\pm} a_{31} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} \overset{?}{\pm}$$

$$\overset{?}{\pm} a_{41} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \end{vmatrix} .$$

Не очень понятно по какому правилу формировать знаки — но обратим внимание, что во всех трех формулах разложения определителя третьего порядка по столбцу эти знаки чередовались: либо в порядке $+,-,+$, либо — в порядке $-,+,-$. Примем в качестве гипотезы чередование $+,-,+,-$:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - \tag{3.6}$$

$$-a_{21} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{41} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \end{vmatrix} .$$

Теперь, поскольку у нас уже имеется способ вычисления определителя третьего порядка, мы можем выписать полное представление определителя четвертого порядка:

$$\begin{aligned} &= a_{11}a_{22}a_{33}a_{44} - a_{11}a_{22}a_{34}a_{43} - a_{11}a_{32}a_{23}a_{44} + a_{11}a_{32}a_{24}a_{43} + \\ &+ a_{11}a_{42}a_{23}a_{34} - a_{11}a_{42}a_{24}a_{33} - a_{21}a_{12}a_{33}a_{44} + a_{21}a_{12}a_{34}a_{43} + \\ &+ a_{21}a_{32}a_{13}a_{44} - a_{21}a_{32}a_{14}a_{43} - a_{21}a_{42}a_{13}a_{34} + a_{21}a_{42}a_{14}a_{33} + \\ &+ a_{31}a_{12}a_{23}a_{44} - a_{31}a_{12}a_{24}a_{43} - a_{31}a_{22}a_{13}a_{44} + a_{31}a_{22}a_{14}a_{43} + \\ &+ a_{31}a_{42}a_{13}a_{24} - a_{31}a_{42}a_{14}a_{23} - a_{41}a_{12}a_{23}a_{34} + a_{41}a_{12}a_{24}a_{33} + \\ &+ a_{41}a_{22}a_{13}a_{34} - a_{41}a_{22}a_{14}a_{33} - a_{41}a_{32}a_{13}a_{24} + a_{41}a_{32}a_{14}a_{23}. \end{aligned} \tag{3.7}$$

Всего в этой формуле 24 члена, половина из них имеет положительный знак, половина — отрицательный.

Теперь дело остается за проверкой: имеет ли какое-то отношение полученное выражение к задаче решения системы линейных уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 = b_3, \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + a_{44}x_4 = b_4 \ ? \end{cases} \tag{3.8}$$

Для этого применим тот же прием, что мы использовали на с. 217 при решении системы из трех уравнений: домножим

$$\left.\begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} \times \atop \times \begin{pmatrix} \text{1-е} \\ \text{уравнение} \end{pmatrix}\right| \left.-\begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} \times \atop \times \begin{pmatrix} \text{2-е} \\ \text{уравнение} \end{pmatrix}\right| \left.\begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} \times \atop \times \begin{pmatrix} \text{3-е} \\ \text{уравнение} \end{pmatrix}\right| -\begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \end{vmatrix} \times \atop \times \begin{pmatrix} \text{4-е} \\ \text{уравнение} \end{pmatrix}$$

и сложим:

$$A_1x_1 + A_2x_2 + A_3x_3 + A_4x_4 = B_1. \tag{3.9}$$

Коэффициент A_1 будет в точности совпадать с выражением (3.6), т.е. с определителем матрицы. А вот коэффициенты A_2, A_3 и A_4 обратятся в нуль. Этот факт, проверяемый непосредственными вычислениями, является проявлением того свойства, которое мы наблюдали для определителей меньших размерностей: выражение

$$A_2 = a_{12}\begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{22}\begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} + a_{32}\begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{42}\begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \end{vmatrix}$$

совпадает с определителем

$$\begin{vmatrix} a_{12} & a_{12} & a_{13} & a_{14} \\ a_{22} & a_{22} & a_{23} & a_{24} \\ a_{32} & a_{32} & a_{33} & a_{34} \\ a_{42} & a_{42} & a_{43} & a_{44} \end{vmatrix},$$

у которого два столбца одинаковы. Подобные определители второго и третьего порядков были равны нулю. По-видимому, это свойство сохраняется и для определителей бо́льших порядков.

Стоящее в правой части уравнения (3.9) выражение для B_1 представляет из себя определитель

$$\begin{vmatrix} b_1 & a_{12} & a_{13} & a_{14} \\ b_2 & a_{22} & a_{23} & a_{24} \\ b_3 & a_{32} & a_{33} & a_{34} \\ b_4 & a_{42} & a_{43} & a_{44} \end{vmatrix}.$$

Следовательно, компонента x_1 решения системы (3.8) может быть представлена по формуле, аналогичной первой из формул Крамера (3.4):

$$x_1 = B_1/A_1,$$

если только знаменатель — определитель (3.6) — отличен от нуля.

Замечание 3.1. Можно проверить, что если бы в формуле (3.6) мы приняли бы альтернативное чередование знаков у определителей третьего порядка $-, +, -, +$, окончательный результат не изменил бы своего вида: это равносильно смене знака у всех определителей.

Определитель n-го порядка

Понятие определителя вводится только для квадратных матриц.

В предыдущем пункте наметились два возможных подхода к обобщению понятия определителя матрицы n-го порядка. Первый заключался в сведéнии к случаю матрицы $(n-1)$-го порядка. Второй — в выписывании явного выражения определителя как функции своих элементов. Структуру этой функции можно предсказать: это — однородный полином, каждый член которого имеет вид $a_{1\alpha_1}a_{2\alpha_2} \times \cdots \times a_{n\alpha_n}$, с определенным знаком (см. формулы (3.1), (3.3) и (3.7)). В настоящем пункте нашей задачей является установить общее правило задания этого знака.

Определение 3.2. **Определителем** или **детерминантом** матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}_{n\times n}$$

называется сумма всевозможых произведений $a_{1\alpha_1}a_{2\alpha_2}\dots a_{n\alpha_n}$, взятых с определенными знаком $+$ или $-$. Здесь упорядоченный набор индексов $\alpha_1, \alpha_2, \dots, \alpha_n$ представляет перестановку[8] чисел $1, 2, \dots, n$.

$$\det A \stackrel{\text{def}}{=} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & \dots & \dots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum (-1)^{?} a_{1\alpha_1} a_{2\alpha_2} \times \dots \times a_{n\alpha_n}. \quad (3.10)$$

Определение 3.3. Упорядоченная пара (a, b) различных натуральных чисел образует **инверсию** (или нарушение порядка) если $a > b$. Можно ввести “счетчик” инверсий:

$$\mathfrak{inv}\,(a,b) \stackrel{\text{def}}{=} \begin{cases} 1 & \iff a > b; \\ 0 & \iff a < b \end{cases}$$

и с его помощью распространить определение на перестановку $(\alpha_1, \alpha_2, \dots, \alpha_n)$:

$$\mathfrak{inv}\,(\alpha_1, \alpha_2, \dots, \alpha_n) \stackrel{\text{def}}{=} \sum_{1 \le j < k \le n} \mathfrak{inv}\,(\alpha_j, \alpha_k). \quad (3.11)$$

Упражнение 13. *Показать, что* $\max \mathfrak{inv}\,(\alpha_1, \alpha_2, \dots, \alpha_n) = \mathsf{C}_n^2$.

Упражнение 14. *Показать, что*

$$\mathfrak{inv}\,(\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_\ell) = \\ = \mathfrak{inv}\,(\alpha_1, \dots, \alpha_k) + \mathfrak{inv}\,(\beta_1, \dots, \beta_\ell) + \sum_{\substack{1 \le i \le k \\ 1 \le j \le \ell}} \mathfrak{inv}\,(\alpha_i, \beta_j). \quad (3.12)$$

Теорема 3.2. *Если* $(\alpha_1, \dots, \alpha_k, \alpha_{k+1}, \dots, \alpha_n)$ *— перестановка чисел* $\{1, 2, \dots, n\}$, *то*

$$\mathfrak{inv}\,(\alpha_1, \dots, \alpha_k, \alpha_{k+1}, \dots, \alpha_n) = \\ = \mathfrak{inv}\,(\alpha_1, \dots, \alpha_k) + \mathfrak{inv}\,(\alpha_{k+1}, \dots, \alpha_n) + (\alpha_1 + \dots + \alpha_k) - \mathsf{C}_{k+1}^2. \quad (3.13)$$

---

[8] См. с. 11

Доказательство. Воспользуемся формулой (3.12). Преобразуем сумму:

$$\sum_{\substack{1\le i\le k\\ k+1\le j\le n}} \mathfrak{inv}\,(\alpha_i,\alpha_j) = [\mathfrak{inv}\,(\alpha_1,\alpha_{k+1}) + \cdots + \mathfrak{inv}\,(\alpha_1,\alpha_n)] +$$

$$+ [\mathfrak{inv}\,(\alpha_2,\alpha_{k+1}) + \cdots + \mathfrak{inv}\,(\alpha_2,\alpha_n)] +$$
$$+ \cdots +$$
$$+ [\mathfrak{inv}\,(\alpha_k,\alpha_{k+1}) + \cdots + \mathfrak{inv}\,(\alpha_k,\alpha_n)]\,.$$

Переставим скобки [ ], упорядочив числа $\alpha_1,\alpha_2,\ldots,\alpha_k$ по возрастанию. Иначе говоря, произведем перестановку

$$(\alpha_1,\ldots,\alpha_k) \longrightarrow (\alpha_{i_1},\ldots,\alpha_{i_k}) \quad \text{при} \quad \alpha_{i_1} < \cdots < \alpha_{i_k}\,.$$

Сумма не изменится

$$= [\mathfrak{inv}\,(\alpha_{i_1},\alpha_{k+1}) + \cdots + \mathfrak{inv}\,(\alpha_{i_1},\alpha_n)] +$$
$$+ [\mathfrak{inv}\,(\alpha_{i_2},\alpha_{k+1}) + \cdots + \mathfrak{inv}\,(\alpha_{i_2},\alpha_n)] +$$
$$+ \cdots +$$
$$+ [\mathfrak{inv}\,(\alpha_{i_k},\alpha_{k+1}) + \cdots + \mathfrak{inv}\,(\alpha_{i_k},\alpha_n)]\,.$$

Пусть сначала $\alpha_{i_1} > 1$. Среди чисел $\alpha_{k+1},\ldots,\alpha_n$ обязательно находятся $1,\ldots,\alpha_{i_1}-1$, так как их нет среди чисел $\alpha_1,\ldots,\alpha_k$. В результате все пары из первой скобки [ ]

$$(\alpha_{i_1},1),\ldots,(\alpha_{i_1},\alpha_{i_1}-1)$$

составляют инверсию, а все оставшиеся ее не составляют. Тогда первая скобка [ ] будет равна $\alpha_{i_1}-1$. Если $\alpha_{i_1}=1$, то все пары из первой скобки [ ] инверсии не составляют, и она будет равна 0, т.е. опять-таки $\alpha_{i_1}-1$. Среди чисел $\alpha_{k+1},\ldots,\alpha_n$ обязательно находятся $1,\ldots,\alpha_{i_2}-1$, за исключением α_{i_1}. Все последние составляют инверсию в паре (α_{i_2},α_j). Тогда вторая скобка [ ] будет равна $\alpha_{i_2}-2$. Продолжаем по индукции. Последняя скобка будет равна $\alpha_{i_k}-k$, так как среди чисел $\alpha_{k+1},\ldots,\alpha_n$ обязательно находятся $1,\ldots,\alpha_{i_k}-1$, за исключением $\alpha_{i_1},\ldots,\alpha_{i_{k-1}}$. В результате, вся сумма будет равна

$$[\alpha_{i_1}-1]+[\alpha_{i_2}-2]+\ldots+[\alpha_{i_k}-k] = (\alpha_{i_1}+\alpha_{i_2}+\ldots+\alpha_{i_k}) -$$

$$-(1+2+\ldots+k) = = (\alpha_1 + \cdots + \alpha_k) - k(k+1)/2,$$

поскольку числа $\alpha_{i_1}, \ldots, \alpha_{i_k}$ это переставленные $\alpha_1, \ldots, \alpha_k$. ■

Определение 3.4.

$$\det A \stackrel{\mathrm{def}}{=} \sum (-1)^{\mathfrak{inv}\,(\alpha_1,\alpha_2,\ldots,\alpha_n)} a_{1\alpha_1} a_{2\alpha_2} \ldots a_{n\alpha_n} \tag{3.14}$$

где сумма распространяется на всевозможные перестановки $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ элементов $\{1, 2, \ldots, n\}$.

Упражнение 15. *Показать, что в случае $n \in \{2, 3, 4\}$ формула (3.14) соответствует введенным в предыдущем пункте определениям определителей соответствующих размерностей.*

Свойства перестановок

Теорема 3.3. *Число всех перестановок из n элементов равно $n!$*

Доказательство приведено на с.11. ■

Итак, формула (3.14) заключает в себе $n!$ слагаемых. Покажем теперь, что в этой формуле число сложений совпадает с числом вычитаний.

Определение 3.5. Перестановки с четным числом инверсий называются **четными**, с нечетным — **нечетными**.

Рассмотрим две перестановки

$$\Theta = (a_1, \ldots, a_K, \alpha, b_1, \ldots, b_L, \beta, c_1, \ldots, c_M)$$

$$\Xi = (a_1, \ldots, a_K, \beta, b_1, \ldots, b_L, \alpha, c_1, \ldots, c_M)$$

Говорят, что перестановка Ξ получена из перестановки Θ в результате **транспозиции**.

Теорема 3.4. *Любая транспозиция меняет четность перестановки:*

$$\mathfrak{inv}\,\Theta - \mathfrak{inv}\,\Xi = \textit{ нечетное.}$$

Доказательство. **I.** Предположим сначала, что $L = 0$, т.е. производится транспозиция соседних элементов.

$$\begin{aligned}
\mathfrak{inv}\,\Theta &= \mathfrak{inv}\,(a_1, \dots, a_K, \alpha, \beta, c_1, \dots, c_M) = \\
&\overset{(3.11)}{=} \sum_{1\le i<j\le K} \mathfrak{inv}\,(a_i, a_j) + \sum_{1\le i\le K} \mathfrak{inv}\,(a_i, \alpha) + \sum_{1\le i\le K} \mathfrak{inv}\,(a_i, \beta) + \\
&+ \sum_{\substack{1\le i\le K\\ 1\le \ell\le M}} \mathfrak{inv}\,(a_i, c_\ell) + \mathfrak{inv}\,(\alpha, \beta) + \sum_{1\le \ell\le M} \mathfrak{inv}\,(\alpha, c_\ell) + \\
&+ \sum_{1\le \ell\le M} \mathfrak{inv}\,(\beta, c_\ell) + \sum_{1\le i<j\le M} \mathfrak{inv}\,(c_i, c_j)
\end{aligned}$$

Разложение для $\mathfrak{inv}\,\Xi$ отличается от этого только пятым слагаемым, которое меняется на $\mathfrak{inv}\,(\beta, \alpha)$. Поэтому

$$\mathfrak{inv}\,\Theta - \mathfrak{inv}\,\Xi = \mathfrak{inv}\,(\alpha, \beta) - \mathfrak{inv}\,(\beta, \alpha) = \pm 1.$$

II. Покажем, что сумма произвольного <u>нечетного</u> числа чисел, каждое из которых равно $+1$ или -1 есть число нечетное. Пусть среди заданных чисел $\omega_1, \dots, \omega_{2N+1}$ имеется p, равных 1, и q, равных -1. Тогда

$$2N+1 = p+q, \qquad \text{и}\ \omega_1 + \dots + \omega_{2N+1} = p - q = 2N+1-2q = 2(N-q)+1.$$

III. Покажем теперь справедливость утверждения теоремы для любого $L \ge 1$.

$$\begin{aligned}
&\mathfrak{inv}\,\Theta = \mathfrak{inv}\,(a_1, \dots, a_K, \alpha, b_1, \dots, b_L, \beta, c_1, \dots, c_M) = \\
&= \mathfrak{inv}\,(a_1, \dots, a_K, b_1, \alpha, \dots, b_L, \beta, c_1, \dots, c_M) + \omega_1 = \dots = \\
&= \mathfrak{inv}\,(a_1, \dots, a_K, b_1, \dots, b_L, \alpha, \beta, c_1, \dots, c_M) + \omega_1 + \dots + \omega_L = \\
&= \mathfrak{inv}\,(a_1, \dots, a_K, b_1, \dots, b_L, \beta, \alpha, c_1, \dots, c_M) + \omega_1 + \ldots + \omega_L + \omega_{L+1} = \\
&= \mathfrak{inv}\,(a_1, \dots, a_K, b_1, \dots, \beta, b_L, \alpha, c_1, \dots, c_M) + \\
&\qquad + \omega_1 + \dots + \omega_L + \omega_{L+1} + \omega_{L+2} = \dots = \\
&= \mathfrak{inv}\,(a_1, \dots, a_K, \beta, b_1, \dots, b_L, \alpha, c_1, \dots, c_M) + \omega_1 + \dots + \omega_{2L+1} = \\
&\overset{\text{п.II}}{=} \mathfrak{inv}\,\Xi + \text{ нечетное.}
\end{aligned}$$

■

Теорема 3.5. *Число четных перестановок из n элементов равно числу нечетных (и равно $n!/2$).*

Доказательство. Действительно, упорядочим все множество перестановок таким образом, чтобы каждая следующая получалась из предшествующей с помощью одной транспозиции. Возможность такого упорядочения доказывается индукцией по n. Для $n = 2$ это тривиально: $(1, 2)$ и $(2, 1)$. Предположим, что утверждение верно для перестановок из $(n-1)$−го элемента, т.е. их можно расположить в требуемом порядке.

Все множество перестановок из n элементов разобьем на подмножества по тем же правилам, что и в доказательстве теоремы 3.3. На основании индукционного предположения перестановки из первого подмножества $(\alpha_1, \dots, \alpha_{n-1}, n)$ можно упорядочить нужным нам образом. В получившейся последовательности перестановок возьмем последнюю и переставим местами n и $n-1$; в результате получим перестановку из второго подмножества $(\beta_1, \dots, \beta_{n-1}, n-1)$. По индуктивному предположению и эти перестановки можно упорядочить так, как требуется. В последней получившейся переставим местами $n-1$ и $n-2$, получим перестановку из третьего класса $(\gamma_1, \dots, \gamma_{n-1}, n-2)$ и т.д.

Таким способом можно упорядочить все $n!$ перестановок. По теореме 3.4 соседние перестановки должны иметь разные четности. ■

Теорема 3.6. *Если $(\alpha_1, \alpha_2, \dots, \alpha_n)$ и $(\beta_1, \beta_2, \dots, \beta_n)$ — произвольные перестановки элементов $\{1, 2, \dots, n\}$, то в разложении определителя по формуле (3.14) обязательно встретится слагаемое*

$$(-1)^{\mathfrak{inv}\,(\alpha_1,\alpha_2,\dots,\alpha_n)+\mathfrak{inv}\,(\beta_1,\beta_2,\dots,\beta_n)} a_{\alpha_1\beta_1} a_{\alpha_2\beta_2} \times \dots \times a_{\alpha_n\beta_n} \qquad (3.15)$$

Доказательство. Для последовательности пар индексов

$$\Theta = \left(\binom{\alpha_1}{\beta_1}, \binom{\alpha_2}{\beta_2}, \dots, \binom{\alpha_k}{\beta_k}, \dots, \binom{\alpha_\ell}{\beta_\ell}, \dots, \binom{\alpha_n}{\beta_n} \right) \qquad (3.16)$$

где $(\alpha_1, \alpha_2, \dots, \alpha_n)$ и $(\beta_1, \beta_2, \dots, \beta_n)$ — произвольные перестановки элементов $\{1, 2, \dots, n\}$, рассмотрим обобщение понятия числа инверсий:

$$\mathfrak{Inv}\,(\Theta) \stackrel{\text{def}}{=} \mathfrak{inv}\,(\alpha_1, \alpha_2, \dots, \alpha_n) + \mathfrak{inv}\,(\beta_1, \beta_2, \dots, \beta_n). \qquad (3.17)$$

Покажем, что при произвольной перестановке <u>любого числа</u> пар в последовательности (3.16) величина (3.17) меняется на четное число. Справедливость утверждения для перестановки <u>двух</u> пар следует из теоремы 3.4:

$$\begin{gathered}\mathfrak{Inv}\left(\begin{pmatrix}\alpha_1\\ \beta_1\end{pmatrix},\begin{pmatrix}\alpha_2\\ \beta_2\end{pmatrix},\ldots,\begin{pmatrix}\alpha_\ell\\ \beta_\ell\end{pmatrix},\ldots,\begin{pmatrix}\alpha_k\\ \beta_k\end{pmatrix},\ldots,\begin{pmatrix}\alpha_n\\ \beta_n\end{pmatrix}\right)=\\ =\mathfrak{inv}\,(\alpha_1,\alpha_2,\ldots,\alpha_\ell,\ldots,\alpha_k,\ldots,\alpha_n)+\\ +\mathfrak{inv}\,(\beta_1,\beta_2,\ldots,\beta_\ell,\ldots,\beta_k,\ldots,\beta_n)=\\ =\mathfrak{inv}\,(\alpha_1,\alpha_2,\ldots,\alpha_k,\ldots,\alpha_\ell,\ldots,\alpha_n)+\Omega_1+\\ +\mathfrak{inv}\,(\beta_1,\beta_2,\ldots,\beta_k,\ldots,\beta_\ell,\ldots,\beta_n)+\Omega_2=\\ =\mathfrak{Inv}\,(\Theta)+\Omega_1+\Omega_2.\end{gathered}$$

Поскольку Ω_1 и Ω_2 — нечетные, то $\Omega_1+\Omega_2$ — четное. Доказательство для произвольного числа перестановок пар проводится индукцией по числу пар.

Рассмотрим теперь сомножители произведения (3.15). Их можно переставить так, чтобы первые индексы шли по возрастанию:

$$a_{\alpha_1\beta_1}a_{\alpha_2\beta_2}\ldots a_{\alpha_n\beta_n}=a_{1\ \mathrm{в}_1}a_{2\ \mathrm{в}_2}\ldots a_{n\ \mathrm{в}_n}.$$

По доказанному, при такой перестановке величина $\mathfrak{Inv}$ изменится на четное число, поэтому знак этого члена в сумме (3.14) станет равным

$$(-1)^{\mathfrak{Inv}\,(\Theta)+\text{ четное}}=(-1)^{\mathfrak{Inv}\,(\Theta)}.$$

■

Определение 3.6. Определитель матрицы A есть сумма всевозможных произведений элементов матрицы, взятых по одному из каждой строки и по одному из каждого столбца, и знак произведения задается формулой (3.15).

Следствие 3.1. *Формулу (3.14) можно переписать, распространив суммирование на первые индексы:*

$$\det A\overset{\text{def}}{=}\sum(-1)^{\mathfrak{inv}\,(\alpha_1,\alpha_2,\ldots,\alpha_n)}a_{\alpha_1 1}a_{\alpha_2 2}\times\cdots\times a_{\alpha_n n}\tag{3.18}$$

Пример 3.2. С каким знаком войдет в разложение определителя матрицы 5-го порядка произведение $a_{43}a_{35}a_{52}a_{24}a_{11}$?

Решение. $\operatorname{sign} = (-1)^{\mathfrak{inv}\,(4,3,5,2,1)+\mathfrak{inv}\,(3,5,2,4,1)} = (-1)^{15} = -1$. △

Упражнение 16. *Вычислить*

$$\begin{vmatrix} 0 & a_{12} & 0 & a_{14} & 0 \\ a_{21} & 0 & 0 & 0 & a_{25} \\ 0 & 0 & a_{33} & 0 & 0 \\ a_{41} & 0 & 0 & 0 & a_{45} \\ 0 & a_{52} & 0 & a_{54} & 0 \end{vmatrix}.$$

Следствие 3.2. *В разложение* $\det A$ *всегда будет входить произведение элементов главной диагонали* $+a_{11}a_{22} \times \cdots \times a_{nn}$*, это выражение называется* **главным членом** *определителя.*

Следствие 3.3. *Определитель диагональной матрицы равен произведению ее диагональных элементов:*

$$\begin{vmatrix} a_{11} & & & \\ & a_{22} & & \mathbb{O} \\ \mathbb{O} & & \ddots & \\ & & & a_{nn} \end{vmatrix} = a_{11}a_{22} \times \cdots \times a_{nn}.$$

В частности,

$$\det E = 1. \tag{3.19}$$

Упражнение 17. *С каким знаком войдет в разложение* $\det A$ *произведение элементов другой диагонали матрицы, т.е.* $a_{n1}a_{n-1,2} \times \cdots \times a_{1n}$*?*

§4. Элементарные свойства определителя

Теорема 4.1. *Определитель матрицы не меняется при транспонировании:*

$$\det A = \det A^{\top}. \tag{4.1}$$

Доказательство. Обозначим $B \overset{\text{def}}{=} A^{\top}$, таким образом $B = [b_{jk}]_{j,k=1,\ldots,n}$ при $b_{jk} = a_{kj}$. На основании определения определителя (3.14):

$$\det B = \sum (-1)^{\mathfrak{inv}\,(\alpha_1,\alpha_2,\dots,\alpha_n)} b_{1\alpha_1} b_{2\alpha_2} \times \dots \times b_{n\alpha_n} = \\ = \sum (-1)^{\mathfrak{inv}\,(\alpha_1,\alpha_2,\dots,\alpha_n)} a_{\alpha_1 1} a_{\alpha_2 2} \times \dots \times a_{\alpha_n n}.$$

Однако последняя сумма совпадает с правой частью формулы (3.18), следовательно $\det B = \det A$. ■

Пример 4.1. Доказать, что при $\{u_{jk}, v_{jk}\}_{j,k=1}^{n} \subset \mathbb{R}$ определитель

$$\det A = \begin{vmatrix} u_{11} & u_{12} + \boldsymbol{i}\, v_{12} & u_{13} + \boldsymbol{i}\, v_{13} & \dots & u_{1n} + \boldsymbol{i}\, v_{1n} \\ u_{12} - \boldsymbol{i}\, v_{12} & u_{22} & u_{23} + \boldsymbol{i}\, v_{23} & \dots & u_{2n} + \boldsymbol{i}\, v_{2n} \\ u_{13} - \boldsymbol{i}\, v_{13} & u_{13} - \boldsymbol{i}\, v_{13} & u_{33} & \dots & u_{3n} + \boldsymbol{i}\, v_{3n} \\ \dots & & & & \dots \\ u_{1n} - \boldsymbol{i}\, v_{1n} & u_{2n} - \boldsymbol{i}\, v_{2n} & u_{3n} - \boldsymbol{i}\, v_{3n} & \dots & u_{nn} \end{vmatrix}$$

является числом вещественным.

Решение. Пусть $\det A = a + \boldsymbol{i}\, b$, где $\{a, b\} \subset \mathbb{R}$. Тогда по теореме 4.1 имеем: $a + \boldsymbol{i}\, b = \det A = \det A^{\top} =$

$$= \begin{vmatrix} u_{11} & u_{12} - \boldsymbol{i}\, v_{12} & u_{13} - \boldsymbol{i}\, v_{13} & \dots & u_{1n} - \boldsymbol{i}\, v_{1n} \\ u_{12} + \boldsymbol{i}\, v_{12} & u_{22} & u_{23} - \boldsymbol{i}\, v_{23} & \dots & u_{2n} - \boldsymbol{i}\, v_{2n} \\ u_{13} + \boldsymbol{i}\, v_{13} & u_{13} + \boldsymbol{i}\, v_{13} & u_{33} & \dots & u_{3n} - \boldsymbol{i}\, v_{3n} \\ \dots & & & & \dots \\ u_{1n} + \boldsymbol{i}\, v_{1n} & u_{2n} + \boldsymbol{i}\, v_{2n} & u_{3n} + \boldsymbol{i}\, v_{3n} & \dots & u_{nn} \end{vmatrix} = a - \boldsymbol{i}\, b.$$

Например, для $n = 3$:

$$\begin{vmatrix} u_{11} & u_{12} - \boldsymbol{i}\, v_{12} & u_{13} - \boldsymbol{i}\, v_{13} \\ u_{12} + \boldsymbol{i}\, v_{12} & u_{22} & u_{23} - \boldsymbol{i}\, v_{23} \\ u_{13} + \boldsymbol{i}\, v_{13} & u_{23} + \boldsymbol{i}\, v_{23} & u_{33} \end{vmatrix} =$$

$$= u_{11}u_{22}u_{33} - u_{11}\left(u_{23}^2 + v_{23}^2\right) - u_{22}\left(u_{13}^2 + v_{13}^2\right) - u_{33}\left(u_{12}^2 + v_{12}^2\right) + \\ + 2\left(u_{12}u_{13}u_{23} + u_{12}v_{13}v_{23} + v_{12}u_{13}v_{23} + v_{12}v_{13}u_{23}\right) \in \mathbb{R}.$$

△

Из теоремы 4.1 следует, что любое свойство, которое мы сможем доказать относительно строк определителя, будет иметь место и относительно его столбцов, и наоборот. Для удобства рассуждений объединим понятия строки и столбца под одним определением.

Перейти к оглавлению на странице: 460

Определение 4.1. Будем называть строку или столбец матрицы (или определителя) ее **рядом** — соответственно горизонтальным или вертикальным.

Теорема 4.2. *Общий множитель элементов любого ряда определителя можно вынести за знак определителя:*

$$\det\left[A_{[1]},\ldots,c\cdot A_{[j]},\ldots,A_{[n]}\right]=c\cdot\det\left[A_{[1]},\ldots,A_{[j]},\ldots,A_{[n]}\right],\tag{4.2}$$

$$\det\begin{bmatrix} A^{[1]} \\ \vdots \\ c\cdot A^{[k]} \\ \vdots \\ A^{[n]} \end{bmatrix}=c\cdot\det\begin{bmatrix} A^{[1]} \\ \vdots \\ A^{[k]} \\ \vdots \\ A^{[n]} \end{bmatrix}.\tag{4.3}$$

Доказательство. Докажем (4.3):

$$\sum(-1)^{\mathfrak{inv}\,(\alpha_1,\ldots,\alpha_n)}a_{1\alpha_1}\times\cdots\times(c\cdot a_{k\alpha_k})\times\cdots\times a_{n\alpha_n}=$$

$$=c\cdot\sum(-1)^{\mathfrak{inv}\,(\alpha_1,\ldots,\alpha_n)}a_{1\alpha_1}\times\cdots\times a_{k\alpha_k}\times\cdots\times a_{n\alpha_n}=c\det A.$$

■

Упражнение 18. *Доказать, что* **а)** $\det(-A)=(-1)^n\det A$;

б) *определитель кососимметричной матрицы нечетного порядка равен нулю;*

в) $\det\left[(-1)^{j+k}a_{jk}\right]_{j,k=1,\ldots,n}=\det\left[a_{jk}\right]_{j,k=1,\ldots,n}$.

Следующая теорема подтверждает одну гипотезу, выдвинутую в пункте 3 на основе анализа определителей малых порядков.

Теорема 4.3. *Определитель, имеющий два одинаковых ряда, равен нулю.*

Доказательство. Докажем утверждение теоремы для строк. Пусть $j\neq k$. В разложение (3.14) для $\det A$ обязательно войдут по одному элементу j-й и k-й строк и это разложение можно разбить на пары слагаемых вида:

$$(-1)^{\mathfrak{inv}\,(1,\ldots,j,\ldots,k,\ldots,n)+\mathfrak{inv}\,(\alpha_1,\ldots,\alpha_j,\ldots,\alpha_k,\ldots,\alpha_n)}\times$$
$$\times a_{1\alpha_1}\times\cdots\times a_{j\alpha_j}\times\cdots\times a_{k\alpha_k}\times\cdots\times a_{n\alpha_n}$$

и
$$(-1)^{\mathfrak{inv}\,(1,\dots,k,\dots,j,\dots,n)+\mathfrak{inv}\,(\alpha_1,\dots,\alpha_j,\dots,\alpha_k,\dots,\alpha_n)}\times$$
$$\times a_{1\alpha_1}\times\dots\times a_{k\alpha_j}\times\dots\times a_{j\alpha_k}\times\dots\times a_{n\alpha_n}\,,$$
различающихся лишь двумя сомножителями и знаками. По теореме 3.4 четность перестановки $(1,\dots,j,\dots,k,\dots,n)$ противоположна четности перестановки $(1,\dots,k,\dots,j,\dots,n)$, следовательно знаки противоположны. Если j-я и k-я строки матрицы A одинаковы:
$$\begin{matrix} a_{j1}\dots a_{jj}\dots a_{jk}\dots a_{jn}\\ \|\quad\dots\|\quad\dots\|\quad\dots\| \\ a_{k1}\dots a_{kj}\dots a_{kk}\dots a_{jn}\end{matrix}$$
то указанные слагаемые дают в сумме нуль. ■

Теорема 4.4. *Пусть имеются три определителя:* $\det A_1$, $\det A_2$ *и* $\det A$*, имеющие все ряды, кроме одного, одинаковыми. Исключительный ряд содержит:*

- *в* $\det A_1$ *— элементы* $\mathfrak{u}_1,\dots,\mathfrak{u}_n$*;*
- *в* $\det A_2$ *— элементы* $\mathfrak{v}_1,\dots,\mathfrak{v}_n$*;*
- *в* $\det A$ *— элементы* $\mathfrak{u}_1+\mathfrak{v}_1,\dots,\mathfrak{u}_n+\mathfrak{v}_n$*.*

Тогда
$$\det A=\det A_1+\det A_2. \tag{4.4}$$
Например,
$$\det\left[A_{[1]},\dots,\overbrace{\mathfrak{U}+\mathfrak{V}}^{A_{[j]}},\dots,A_{[n]}\right]=$$
$$=\det\left[A_{[1]},\dots,\mathfrak{U},\dots,A_{[n]}\right]+\det\left[A_{[1]},\dots,\mathfrak{V},\dots,A_{[n]}\right], \tag{4.5}$$
если $\mathfrak{U}$ *и* $\mathfrak{V}$ *означают столбцы, состоящие из элементов* $\mathfrak{u}_j$ *и* $\mathfrak{v}_j$ *соответственно.*

Доказательство.

$$\det A=\sum(-1)^{\mathfrak{inv}\,(\alpha_1,\dots,\alpha_n)}a_{1\alpha_1}\times\dots\times(\mathfrak{u}_j+\mathfrak{v}_j)\times\dots\times a_{n\alpha_n}=$$
$$=\sum(-1)^{\mathfrak{inv}\,(\alpha_1,\dots,\alpha_n)}a_{1\alpha_1}\times\dots\times\mathfrak{u}_j\times\dots\times a_{n\alpha_n}+$$
$$+\sum(-1)^{\mathfrak{inv}\,(\alpha_1,\dots,\alpha_n)}a_{1\alpha_1}\times\dots\times\mathfrak{v}_j\times\dots\times a_{n\alpha_n}=\det A_1+\det A_2.$$
■

Следствие 4.1. *Утверждение теоремы распространимо на любое количество слагаемых рядов.*

Следствие 4.2. *Определитель не изменится, если к любому его ряду прибавить любой другой ряд, умноженный на произвольное число из* $\mathbb{A}$ *.*

Доказательство. Для конкретности рассмотрим столбцы определителя $\det A$:

$$\det\left[\ldots, A_{[j]} + c \cdot A_{[k]}, \ldots, A_{[k]}, \ldots\right] =$$

$$\overset{(4.5)}{=} \det\left[\ldots, A_{[j]}, \ldots, A_{[k]}, \ldots\right] + \det\left[\ldots, c \cdot A_{[k]}, \ldots, A_{[k]}, \ldots\right]$$

$$\overset{(4.2)}{=} \det\left[\ldots, A_{[j]}, \ldots, A_{[k]}, \ldots\right] + c \cdot \det\left[\ldots, \cdot A_{[k]}, \ldots, A_{[k]}, \ldots\right] =$$

$$\overset{4.3}{=} \det\left[\ldots, A_{[j]}, \ldots, A_{[k]}, \ldots\right] = \det A.$$

■

Упражнение 19. *Верно ли равенство* $\det(A + B) = \det A + \det B$ *для любых квадратных матриц* A *и* B*?*

Теорема 4.5. *При перестановке местами его рядов определитель меняет знак.*

Доказательство. Для конкретности рассмотрим столбцы определителя. По теореме 4.3 следующий определитель с двумя одинаковыми j-ми и k-ми столбцами:

$$\det\left[\ldots, A_{[j]} + A_{[k]}, \ldots, A_{[j]} + A_{[k]}, \ldots\right],$$

равен нулю. Далее имеем цепочку равенств:

$$0 = \left[\ldots, A_{[j]} + A_{[k]}, \ldots, A_{[j]} + A_{[k]}, \ldots\right] =$$

$$\overset{(4.5)}{=} \left[\ldots, A_{[j]} + A_{[k]}, \ldots, A_{[j]}, \ldots\right] + \left[\ldots, A_{[j]} + A_{[k]}, \ldots, A_{[k]}, \ldots\right] =$$

$$\overset{(4.5)}{=} \left[\ldots, A_{[j]}, \ldots, A_{[j]}, \ldots\right] + \left[\ldots, A_{[k]}, \ldots, A_{[j]}, \ldots\right] +$$

$$+ \left[\ldots, A_{[j]}, \ldots, A_{[k]}, \ldots\right] + \left[\ldots, A_{[k]}, \ldots, A_{[k]}, \ldots\right] =$$

$$\overset{4.3}{=} \left[\ldots, A_{[k]}, \ldots, A_{[j]}, \ldots\right] + \left[\ldots, A_{[j]}, \ldots, A_{[k]}, \ldots\right]$$

■

Определение 4.2. Свойства определителя, выраженные теоремами 4.2 и 4.4 называются его **линейными**, а теоремой 4.5 — **кососимметрическим** свойствами относительно рядов этого определителя.

Эти свойства, вместе с (3.19), называются **определяющими свойствами определителя**: можно доказать, что любая функция от набора рядов матрицы A, обладающая этими свойствами, должна совпадать с $\det A$, вычисляемым по формуле (3.14).

§5. Миноры и алгебраические дополнения

В пункте 3 был предложен еще один подход к определению определителя: определитель n-го порядка вводился с помощью определителей $(n-1)$-го порядка. В настоящем параграфе такой подход будет строго обоснован.

Определение 5.1. Определитель $(n-1)$-го порядка, получающийся вычеркиванием из

$$\det A = \begin{vmatrix} a_{11} & \ldots a_{1,k-1} & a_{1k} & a_{1,k+1} & \ldots a_{1n} \\ \ldots & & \ldots & & \ldots \\ a_{j-1,1} & \ldots a_{j-1,k-1} & a_{j-1,k} & a_{j-1,k+1} & \ldots a_{j-1,n} \\ a_{j1} & \ldots a_{j,k-1} & a_{jk} & a_{j,k+1} & \ldots a_{jn} \\ a_{j+1,1} & \ldots a_{j+1,k-1} & a_{j+1,k} & a_{j+1,k+1} & \ldots a_{j+1,n} \\ \ldots & & \ldots & & \ldots \\ a_{n1} & \ldots a_{n,k-1} & a_{nk} & a_{n,k+1} & \ldots a_{nn} \end{vmatrix}$$

j-й строки и k-го столбца называется **минором**[9] ($(n-1)$-го порядка) этого определителя, **соответствующим элементу** a_{jk}. Будем обозначать его M_{jk}:

$$M_{jk} \stackrel{\text{def}}{=} \det A \begin{pmatrix} 1\,2 \ldots\, j-1,\, j+1 \ldots\, n \\ 1\,2 \ldots\, k-1,\, k+1 \ldots\, n \end{pmatrix} =$$

$$= \begin{vmatrix} a_{11} & \ldots a_{1,k-1} & a_{1,k+1} & \ldots a_{1n} \\ \ldots & & & \ldots \\ a_{j-1,1} & \ldots a_{j-1,k-1} & a_{j-1,k+1} & \ldots a_{j-1,n} \\ a_{j+1,1} & \ldots a_{j+1,k-1} & a_{j+1,k+1} & \ldots a_{j+1,n} \\ \ldots & & & \ldots \\ a_{n1} & \ldots a_{n,k-1} & a_{n,k+1} & \ldots a_{nn} \end{vmatrix}.$$

[9] minor (*лат.*) — меньший; minores — дети, молодежь.

Число
$$A_{jk} \stackrel{\text{def}}{=} (-1)^{j+k} M_{jk} \tag{5.1}$$
называется **алгебраическим дополнением элемента** a_{jk} в $\det A$.

Теорема 5.1. *Определитель матрицы равен сумме произведений элементов некоторого ряда определителя на их алгебраические дополнения. Иначе говоря, справедливы следующие формулы* **разложения определителя по** j**-й строке***:*

$$\det A = a_{j1}A_{j1} + a_{j2}A_{j2} + \cdots + a_{jn}A_{jn} = \sum_{\ell=1}^{n} a_{j\ell}A_{j\ell} \tag{5.2}$$

и **разложения определителя по** k**-му столбцу***:*

$$\det A = a_{1k}A_{1k} + a_{2k}A_{2k} + \cdots + a_{nk}A_{nk} = \sum_{\ell=1}^{n} a_{\ell k}A_{\ell k} \tag{5.3}$$

для любых $\{j,k\} \subset \{1,2,\ldots,n\}$.

Доказательство. Докажем сначала справедливость разложения (5.2) для $j=1$, т.е. для первой строки. Эту строку можно представить в виде суммы n строк:

$$(a_{11},a_{12},\ldots,a_{1n}) = (a_{11},0,\ldots,0)+(0,a_{12},\ldots,0)+\cdots+(0,0,\ldots,a_{1n}).$$

На основании теоремы 4.4, $\det A$ можно представить в виде суммы

$$\begin{vmatrix} a_{11} & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & & & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & \ldots & 0 \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & & & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} 0 & 0 & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & & & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}. \tag{5.4}$$

Рассмотрим первый из этих определителей:

$$\begin{vmatrix} a_{11} & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & & & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}. \tag{5.5}$$

На основании теоремы 4.2, из первой строки можно вынести множитель:

$$a_{11} \begin{vmatrix} 1 & 0 & \dots 0 \\ a_{21} & a_{22} & \dots a_{2n} \\ \dots & & \dots \\ a_{n1} & a_{n2} & \dots a_{nn} \end{vmatrix}.$$

Обозначим получившийся определитель через B_1 и разложим его по формуле определителя (3.14):

$$B_1 = \sum (-1)^{\mathfrak{inv}\,(1,\alpha_2,\dots,\alpha_n)} 1 \cdot a_{2\alpha_2} \dots a_{n\alpha_n}.$$

Здесь суммирование идет по всем перестановкам $(\alpha_2, \dots, \alpha_n)$ чисел $2, \dots, n$. Далее, на основании формулы (3.12):

$$\mathfrak{inv}\,(1,\alpha_2,\dots,\alpha_n) = \mathfrak{inv}\,(\alpha_2,\dots,\alpha_n) + \sum_{j=2}^{n} \mathfrak{inv}\,(1,\alpha_j) = \mathfrak{inv}\,(\alpha_2,\dots,\alpha_n).$$

Следовательно,

$$B_1 = \sum (-1)^{\mathfrak{inv}\,(\alpha_2,\dots,\alpha_n)} a_{2\alpha_2} \dots a_{n\alpha_n},$$

где суммирование идет по всем перестановкам $(\alpha_2, \dots, \alpha_n)$ чисел $2, \dots, n$. Однако последняя сумма — на основании определения — представляет следующий определитель $(n-1)$-го порядка

$$\begin{vmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ a_{32} & a_{33} & \dots & a_{3n} \\ \dots & & & \dots \\ a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix} = M_{11}.$$

Видим, что формула (5.2) будет справедлива по крайней мере для определителей вида (5.5).

Рассмотрим теперь второй определитель в (5.4). Переставим в нем местами первый и второй столбцы, на основании теоремы 4.5 получаем

$$\begin{vmatrix} 0 & a_{12} & 0 & \dots 0 \\ a_{21} & a_{22} & a_{23} & \dots a_{2n} \\ \dots & & & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{12} & 0 & 0 & \dots 0 \\ a_{22} & a_{21} & a_{23} & \dots a_{2n} \\ \dots & & & \dots \\ a_{n2} & a_{n1} & a_{n3} & \dots a_{nn} \end{vmatrix}.$$

На основании только что доказанного, получившийся определитель равен

$$= -a_{12} \begin{vmatrix} a_{21}\, a_{23} \dots a_{2n} \\ \dots \qquad \dots \\ a_{n1}\, a_{n3} \dots a_{nn} \end{vmatrix} = -a_{12}M_{12} = a_{12}A_{12}.$$

Теперь понятно как надо обходиться с остальными слагаемыми в формуле (5.4). Переставляем столбцы так, чтобы загнать единственный элемент первой строки в левый верхний угол. Каждая такая перестановка влечет за собой изменение знака, т.е. домножение нового определителя на (-1). Следовательно, последний определитель в (5.4):

$$\begin{vmatrix} 0 & 0 & \dots 0 & a_{1n} \\ a_{21} & a_{22} & \dots a_{2,n-1} & a_{2n} \\ \dots & & \dots & \\ a_{n1} & a_{n2} & \dots a_{n,n-1} & a_{nn} \end{vmatrix} = (-1)^{n-1} \begin{vmatrix} a_{1n} & 0 & 0 & \dots 0 \\ a_{2n} & a_{21} & a_{22} & \dots a_{2,n-1} \\ \dots & & & \dots \\ a_{nn} & a_{n1} & a_{n2} & \dots a_{n,n-1} \end{vmatrix} =$$

$$= (-1)^{n-1} a_{1n} \begin{vmatrix} a_{21}\, a_{22} \dots a_{2,n-1} \\ \dots \qquad \dots \\ a_{n1}\, a_{n2} \dots a_{n,n-1} \end{vmatrix} = (-1)^{n+1} a_{1n} M_{1n} = a_{1n}A_{1n}.$$

Итак, разложение (5.2) имеет место при $j = 1$. Если же $j > 1$, то рассуждения будут аналогичны предыдущему случаю. Для того, чтобы преобразовать определитель

$$\begin{vmatrix} a_{11} & \dots a_{1,k-1} & a_{1k} & a_{1,k+1} & \dots a_{1n} \\ \dots & & & & \dots \\ a_{j-1,1} & \dots a_{j-1,k-1} & a_{j-1,k} & a_{j-1,k+1} & \dots a_{j-1,n} \\ 0 & \dots 0 & a_{jk} & 0 & \dots 0 \\ a_{j+1,1} & \dots a_{j+1,k-1} & a_{j+1,k} & a_{j+1,k+1} & \dots a_{j+1,n} \\ \dots & & & & \dots \\ a_{n1} & \dots a_{n,k-1} & a_{nk} & a_{n,k+1} & \dots a_{nn} \end{vmatrix}$$

к виду (5.5), мы переставляем j-ю строку с предыдущими. Потребуется всего $(j-1)$ перестановка, чтобы эта строка оказалась на месте первой. А этот случай уже обсуждался ранее: потребуется еще $(k-1)$ перестановка, чтобы k-й столбец оказался на месте первого, и мы оказались бы в ситуации определителя вида (5.5).

Минор M_{jk} домножится на $(-1)^{j-1+k-1} = (-1)^{j+k}$. Следовательно, формула (5.2) справедлива для любого j. Справедливость же формулы (5.3) будет тогда следовать из теоремы 4.1. ■

Пример 5.1. Вычислить

$$\begin{vmatrix} 4 & 71 & 5 \\ 3 & 40 & -6 \\ -11 & 82 & 9 \\ -12 & -100 & 8 \end{vmatrix},$$

разложив определитель по третьей строке.

Решение. В формуле (5.2) берем $j = 3$:

$$-11(-1)^{3+1}\begin{vmatrix} 71 & 5 \\ 40 & -6 \\ -100 & 8 \end{vmatrix} + 8(-1)^{3+2}\begin{vmatrix} 41 & 5 \\ 30 & -6 \\ -120 & 8 \end{vmatrix} +$$

$$+2(-1)^{3+3}\begin{vmatrix} 4 & 7 & 5 \\ 3 & 4 & -6 \\ -12 & -10 & 8 \end{vmatrix} + 9(-1)^{3+4}\begin{vmatrix} 4 & 71 \\ 3 & 40 \\ -12 & -100 \end{vmatrix} =$$

и используем формулу вычисления определителя третьего порядка из §3:

$$= -11 \cdot 28 - 8 \cdot 48 + 3 \cdot 314 - 9 \cdot 18 = -226.$$

Заметим, что тот же самый результат можно было бы получить, сэкономив на вычислении определителей третьего порядка, если бы мы разложили исходный определитель по третьему *столбцу*:

$$1 \cdot \begin{vmatrix} 3 & 4 & -6 \\ -11 & 8 & 9 \\ -12 & -10 & 8 \end{vmatrix} - 0 \cdot (\dots) + 2 \cdot \begin{vmatrix} 4 & 7 & 5 \\ 3 & 4 & -6 \\ -12 & -10 & 8 \end{vmatrix} - 0 \cdot (\dots) =$$

$$= 1 \cdot (-854) + 2 \cdot 314 = -226.$$

Наличие нулевых элементов "облегчает жизнь" вычислителю ... △

А что делать, если среди элементов определителя нет нулевых?

Их надо “сделать”. Для этого в нашем распоряжении имеется такое средство, как элементарные преобразования строк или столбцов. В самом деле, на основании следствия 2 к теореме 4.4, к любому ряду определителя можно прибавить любой другой ряд, домноженный на произвольное число — определитель от этого не изменится. Но тогда мы можем эти множители подбирать так, чтобы добиться появления большого количества нулей в отдельном ряду.

Пример 5.2. Вычислить

$$\begin{vmatrix} 2 & 2 & 1 & 3 & 4 \\ 3 & 1 & 2 & 3 & 1 \\ 4 & -1 & 2 & 4 & -2 \\ 1 & -1 & 1 & 1 & 2 \\ 4 & -1 & 2 & 5 & 6 \end{vmatrix}.$$

Решение. Будем добиваться появления нулей во втором столбце. С этой целью прибавим вторую строку к третьей, четвертой и пятой, а также вычтем, домножив предварительно на 2, из первой:

$$= \begin{vmatrix} -4 & 0 & -3 & -3 & 2 \\ 3 & 1 & 2 & 3 & 1 \\ 7 & 0 & 4 & 7 & -1 \\ 4 & 0 & 3 & 4 & 3 \\ 7 & 0 & 4 & 8 & 7 \end{vmatrix} =$$

раскладываем по второму столбцу:

$$= \begin{vmatrix} -4 & -3 & -3 & 2 \\ 7 & 4 & 7 & -1 \\ 4 & 3 & 4 & 3 \\ 7 & 4 & 8 & 7 \end{vmatrix} =$$

и вот уже порядок понизился. Вычитаем из третьего столбца первый:

$$= \begin{vmatrix} -4 & -3 & 1 & 2 \\ 7 & 4 & 0 & -1 \\ 4 & 3 & 0 & 3 \\ 7 & 4 & 1 & 7 \end{vmatrix} =$$

теперь имеет смысл увеличить число нулевых элементов в третьем столбце — вычитаем из четвертой строки первую:

$$= \begin{vmatrix} -4 & -3 & 1 & 2 \\ 7 & 4 & 0 & -1 \\ 4 & 3 & 0 & 3 \\ 11 & 7 & 0 & 5 \end{vmatrix} =$$

Раскладываем по третьему столбцу:

$$= \begin{vmatrix} 7 & 4 & -1 \\ 4 & 3 & 3 \\ 11 & 7 & 5 \end{vmatrix} =$$

Можно было бы применить теперь формулу разложения определителя третьего порядка, но можно и продолжить упрощения — вычтем из третьей строки первую и вторую:

$$= \begin{vmatrix} 7 & 4 & -1 \\ 4 & 3 & 3 \\ 0 & 0 & 3 \end{vmatrix} =$$

и разложим по третьей строке:

$$= 3 \begin{vmatrix} 7 & 4 \\ 4 & 3 \end{vmatrix} = 3(21 - 16) = 15.$$

Ответ. 15.

Формализуем теперь способ вычисления любого определителя n-го порядка. Предположим, что $a_{11} \neq 0$. Тогда, вычитая из второй строки первую, домноженную на a_{21}/a_{11}, из третьей — первую, домноженную на a_{31}/a_{11} и т.д., из n-й — первую, домноженную на a_{n1}/a_{11}, мы добиваемся представления определителя в виде

$$= \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22}^{[1]} & \dots & a_{2n}^{[1]} \\ \dots & & & \dots \\ 0 & a_{m2}^{[1]} & \dots & a_{mn}^{[1]} \end{vmatrix} \quad \text{при } a_{jk}^{[1]} = a_{jk} - \frac{a_{j1}a_{1k}}{a_{11}}.$$

Тогда разложив полученный определитель по первому столбцу, мы понизим порядок определителя

$$= a_{11} \begin{vmatrix} a_{22}^{[1]} \dots & a_{2n}^{[1]} \\ \dots & \dots \\ a_{m2}^{[1]} \dots & a_{mn}^{[1]} \end{vmatrix}.$$

Легко видеть, что предлагаемый алгоритм преобразования строк определителя соответствует первому шагу изложенного в §1 метода Гаусса исключения переменных в системе линейных уравнений. Именно метод Гаусса является основным методом вычисления произвольного <u>числового</u> определителя. В самом деле, элементарными преобразованиями строк <u>квадратной</u> матрицы A ее можно привести к треугольному виду:

$$\begin{pmatrix} a_{11} & a_{12} \dots & a_{1n} \\ & a_{22}^{[1]} \dots & a_{2n}^{[1]} \\ & \mathbb{O} \ddots & \dots \\ & & a_{nn}^{[n-1]} \end{pmatrix}.$$

Если мы докажем, что эти элементарные преобразования не меняют величины определителя, то можно будет заключить, что

$$\det A = a_{11} a_{22}^{[1]} \times \dots \times a_{nn}^{[n-1]}.$$

На самом деле, из перечисленных на с. 187 преобразований только преобразования вида C) не меняют величины определителя. Преобразования вида A) меняют знак определителя, а вида B) — его величину. Однако, если откорректировать метод Гаусса с учетом этих особенностей, то он вполне пригоден для вычисления определителя.

Пример 5.3. Вычислить

$$\begin{vmatrix} 21 & 3 & 41 \\ 42 & 4 & 13 \\ 27 & 1 & 32 \\ -21 & -2 & -12 \\ 12 & 3 & 22 \end{vmatrix}$$

по методу Гаусса.

Решение. Вычитаем первую строку, умноженную на соответствующие числа из остальных строк, добиваясь появления нулей в первом столбце:

$$= \begin{vmatrix} 2 & 1 & 3 & 4 & 1 \\ 0 & 0 & -2 & -7 & 1 \\ 0 & 6 & -2 & -1 & 1 \\ 0 & 0 & 1 & 3 & 3 \\ 0 & 3/2 & 3/2 & 0 & 3/2 \end{vmatrix} =$$

Выносим общий множитель элементов последней строки:

$$= 3/2 \begin{vmatrix} 2 & 1 & 3 & 4 & 1 \\ 0 & 0 & -2 & -7 & 1 \\ 0 & 6 & -2 & -1 & 1 \\ 0 & 2 & 1 & 3 & 3 \\ 0 & 1 & 1 & 0 & 1 \end{vmatrix} =$$

Поскольку $a_{22}^{[1]} = 0$, то придется применять преобразование A): переставим вторую и пятую строки. При этом знак определителя поменяется:

$$= -\,3/2 \begin{vmatrix} 2 & 1 & 3 & 4 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 6 & -2 & -1 & 1 \\ 0 & 2 & 1 & 3 & 3 \\ 0 & 0 & -2 & -7 & 1 \end{vmatrix} =$$

Теперь можно воспользоваться второй строкой, чтобы обратить в нуль элементы второго столбца

$$= -\,3/2 \begin{vmatrix} 2 & 1 & 3 & 4 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & -8 & -1 & -5 \\ 0 & 0 & -1 & 3 & 1 \\ 0 & 0 & -2 & -7 & 1 \end{vmatrix} =$$

Элемент определителя, стоящий в третьей строке на главной диагонали, теперь отличен от нуля, и можно было бы продолжить прямой ход алгоритма Гаусса далее. Но, для того чтобы избежать появления дробных элементов, поставим на место третьей строки четвертую. Определитель снова поменяет знак

$$= {}^3/_2 \begin{vmatrix} 2 & 1 & 3 & 4 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & -1 & 3 & 1 \\ 0 & 0 & -8 & -1 & -5 \\ 0 & 0 & -2 & -7 & 1 \end{vmatrix} = {}^3/_2 \begin{vmatrix} 2 & 1 & 3 & 4 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & -1 & 3 & 1 \\ 0 & 0 & 0 & -25 & -13 \\ 0 & 0 & 0 & -13 & -1 \end{vmatrix} =$$

$$= {}^3/_2 \begin{vmatrix} 2 & 1 & 3 & 4 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & -1 & 3 & 1 \\ 0 & 0 & 0 & -25 & -13 \\ 0 & 0 & 0 & 0 & {}^{144}/_{25} \end{vmatrix} = \frac{3}{2} \cdot 2 \cdot 1 \cdot (-1) \cdot (-25) \cdot \frac{144}{25}.$$

Ответ. 432.

Вычисление определителя отличается от метода Гаусса еще и возможностью производить элементарные преобразования и над столбцами матрицы.

В заключение параграфа докажем один — необходимый для дальнейшего — теоретический результат, который мы наблюдали в пункте 3 на примерах определителей малых размерностей.

Теорема 5.2. *Сумма произведений элементов j-го ряда $\det A$ на алгебраические дополнения элементов k-го ряда равна 0 если $j \neq k$ и равна $\det A$ если $j = k$:*

$$\sum_{\ell=1}^{n} a_{\ell j} A_{\ell k} = \delta_{jk} \det A\,, \qquad \sum_{\ell=1}^{n} a_{j\ell} A_{k\ell} = \delta_{jk} \det A. \tag{5.6}$$

Доказательство. При $j = k$ утверждение следует из теоремы 5.1. Пусть $j \neq k$. Для определенности, докажем теорему для строк $\det A$. Составим новый определитель заменой k-й строки $\det A$ на j-ю строку того же определителя:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \ldots & & & \ldots \\ a_{j1} & a_{j2} & \ldots & a_{jn} \\ \ldots & & & \ldots \\ a_{k1} & a_{k2} & \ldots & a_{kn} \\ \ldots & & & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} \longrightarrow \det \widetilde{A} = \begin{array}{@{}c@{}l} \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \ldots & & & \ldots \\ a_{j1} & a_{j2} & \ldots & a_{jn} \\ \ldots & & & \ldots \\ a_{j1} & a_{j2} & \ldots & a_{jn} \\ \ldots & & & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} & \begin{matrix} \\ \\ \leftarrow j \\ \\ \leftarrow k \\ \\ \\ \end{matrix} \end{array}.$$

На основании теоремы 4.3, $\det \widetilde{A} = 0$. С другой стороны, разложим его по элементам k-й строки; очевидно, что алгебраические дополнения элементов этой строки будут совпадать с алгебраическими дополнениями элементов k-й строки $\det A$:

$$0 = a_{j1}A_{k1} + a_{j2}A_{k2} + \dots a_{jn}A_{kn},$$

т.е. мы получили вторую из формул (5.6). ■

§6. Формулы Крамера

В настоящем параграфе мы обобщим теорему 3.1.

Рассмотрим систему (1.2), в которой число уравнений равно числу неизвестных:

$$\left\{\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \qquad \dots \qquad\qquad\qquad \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{array}\right. \iff \qquad AX = \mathcal{B}. \qquad (6.1)$$

Здесь $A_{n\times n}$, $X_{n\times 1}$, $\mathcal{B}_{n\times 1}$. Рассмотрим столбцы матрицы A:

$$A = \left[A_{[1]}, \dots, A_{[k]}, \dots, A_{[n]}\right].$$

Теорема 6.1. *Если $\det A \ne 0$, то система (6.1) имеет единственное решение, которое определяется* **формулами Крамера**[10] *:*

$$x_k = \frac{\det\left[A_{[1]}, \dots, A_{[k-1]}, \mathcal{B}, A_{[k+1]}, \dots, A_{[n]}\right]}{\det A}, \qquad k \in \{1, \dots, n\} \qquad (6.2)$$

(k-й столбец матрицы заменяется на столбец правых частей).

Доказательство. Пусть решение системы (6.1) существует. Покажем, что оно должно иметь вид (6.2). Домножим первое уравнение системы на A_{11}, второе — на A_{21}, и т.д., n-е — на A_{n1} и просуммируем полученное:

$$x_1\sum_{j=1}^{n} a_{j1}A_{j1} + x_2\sum_{j=1}^{n} a_{j2}A_{j1} + \dots + x_n\sum_{j=1}^{n} a_{jn}A_{j1} = \sum_{j=1}^{n} b_jA_{j1}.$$

[10]**Крамер Габриель** (Cramer Gabriel, 1704–1752) — швейцарский математик.

На основании первого равенства (5.6) получаем: в левой части коэффициенты при $x_2, \dots, x_n$ пропадают, а коэффициент при x_1 равен $\det A$.

$$x_1 \det A = b_1 A_{11} + b_2 A_{21} + \cdots + b_n A_{n1} = \begin{vmatrix} b_1\, a_{12} \dots a_{1n} \\ b_2\, a_{22} \dots a_{2n} \\ \dots \quad \dots \\ b_n\, a_{n2} \dots a_{nn} \end{vmatrix} \quad \Rightarrow$$

$$\Rightarrow x_1 = \frac{\det \left[\mathcal{B}, A_{[2]}, \dots, A_{[n]}\right]}{\det A}.$$

Аналогично (т.е. домножением уравнений на A_{jk}) показывается справедливось и общей формулы (6.2).

Покажем теперь, что формулы (6.2) действительно дают решение, т.е. удовлетворяют системе (6.1). Подставим их в левую часть первого уравнения:

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n =$$

$$= a_{11}\frac{\sum\limits_{\ell=1}^{n} b_\ell A_{\ell 1}}{\det A} + a_{12}\frac{\sum\limits_{\ell=1}^{n} b_\ell A_{\ell 2}}{\det A} + \cdots + a_{1n}\frac{\sum\limits_{\ell=1}^{n} b_\ell A_{\ell n}}{\det A} =$$

перегруппируем слагаемые:

$$= b_1\frac{\sum\limits_{s=1}^{n} a_{1s}A_{1s}}{\det A} + b_2\frac{\sum\limits_{s=1}^{n} a_{1s}A_{2s}}{\det A} + \cdots + b_n\frac{\sum\limits_{s=1}^{n} a_{1s}A_{ns}}{\det A}.$$

на основании второго равенства (5.6) все слагаемые кроме первого пропадут, а первое превратится в $b_1 \det A \big/ \det A = b_1$. Аналогично проверяется истинность всех остальных равенств системы. ■

Пример 6.1. Решить систему уравнений

$$\begin{cases} 2x_1 + 3x_2 + 11x_3 + 5x_4 = 2, \\ x_1 + x_2 + 5x_3 + 2x_4 = 1, \\ 2x_1 + x_2 + 3x_3 + 2x_4 = -3, \\ x_1 + x_2 + 3x_3 + 4x_4 = -3. \end{cases}$$

Решение.

$$x_1 = \frac{\begin{vmatrix} 23115 \\ 11\ \ 52 \\ -31\ \ 32 \\ -31\ \ 34 \end{vmatrix}}{\begin{vmatrix} 23115 \\ 11\ \ 52 \\ 21\ \ 32 \\ 11\ \ 34 \end{vmatrix}} = \frac{-28}{14} = -2, x_2 = \frac{\begin{vmatrix} 2\ \ 2115 \\ 1\ \ 1\ \ 52 \\ 2-3\ \ 32 \\ 1-3\ \ 34 \end{vmatrix}}{\begin{vmatrix} 23115 \\ 11\ \ 52 \\ 21\ \ 32 \\ 11\ \ 34 \end{vmatrix}} = \frac{0}{14} = 0$$

Найдите оставшиеся компоненты решения. △

Замечание 6.1. В вычислительном отношении формулы Крамера существенно менее эффективны для решения системы (6.1), чем метод Гаусса (см. §1). Действительно, грубо говоря, метод Гаусса фактически эквивалентен вычислению одного определителя $\det A$, в то время как формулы (6.2) требуют вычисления $(n+1)$-го определителя. Тем не менее, эти формулы полезны с теоретической точки зрения, поскольку дают удобное для дальнейшего анализа явное аналитическое представление решения системы (6.1) через ее коэффициенты.

Следствие 6.1. *Если система* (6.1) *несовместна, то* $\det A = 0$.

Следствие 6.2. *Если система* (6.1) *имеет более одного решения, то* $\det A = 0$.

Определение 6.1. Система (6.1) называется **однородной** если $\mathcal{B} = \mathbb{O}$.

Однородная система всегда совместна, так как имеет тривиальное решение $X = \mathbb{O}$.

Следствие 6.3. *Условие* $\det A = 0$ *является необходимым для существования нетривиального решения у однородной системы.*

В параграфе 12 будет доказано, что это условие является и достаточным.

Перейти к оглавлению на странице: 460

§7. Теорема Лапласа

Определение 7.1. Выделим в $\det A$ строки с номерами $\alpha_1 < \dots < \alpha_k$ и столбцы с номерами $\beta_1 < \dots < \beta_k$. Элементы $a_{\alpha_j \beta_\ell}$, стоящие в этих строках и столбцах, образуют определитель k-го порядка:

$$M = A\begin{pmatrix}\alpha_1 \dots \alpha_k \\ \beta_1 \dots \beta_k\end{pmatrix} \overset{\text{def}}{=} \begin{vmatrix} a_{\alpha_1\beta_1} \dots a_{\alpha_1\beta_k} \\ \dots \qquad \dots \\ a_{\alpha_k\beta_1} \dots a_{\alpha_k\beta_k}\end{vmatrix}$$

Он называется **минором k-го порядка** определителя $\det A$. Если же из определителя $\det A$ вычеркиваются строки и столбцы с указанными номерами, то получившийся определитель $(n-k)$-го порядка

$$A\begin{pmatrix}\alpha_{k+1} \dots \alpha_n \\ \beta_{k+1} \dots \beta_n\end{pmatrix} \quad \text{здесь} \begin{matrix}\{\alpha_{k+1} < \dots < \alpha_n\} \subset \{1, 2, \dots, n\} \setminus \{\alpha_1, \dots \alpha_k\} \\ \{\beta_{k+1} < \dots < \beta_n\} \subset \{1, 2, \dots, n\} \setminus \{\beta_1, \dots \beta_k\}\end{matrix}$$

называется **минором, дополнительным минору** M в $\det A$. Число

$$\widetilde{M} \overset{\text{def}}{=} (-1)^{\alpha_1 + \dots + \alpha_k + \beta_1 + \dots + \beta_k} A\begin{pmatrix}\alpha_{k+1} \dots \alpha_n \\ \beta_{k+1} \dots \beta_n\end{pmatrix}$$

называется **алгебраическим дополнением** минора M в $\det A$.

Теорема 7.1. *[Лаплас] Выделим в* $\det A$ *произвольные строки с номерами* $\alpha_1 < \dots < \alpha_k$*. Образуем всевозможные миноры* k*-го порядка с элементами из этих строк:*

$$A\begin{pmatrix}\alpha_1 \dots \alpha_k \\ \beta_1 \dots \beta_k\end{pmatrix}, \quad \text{здесь } \{\beta_1 < \dots < \beta_k\} \subset \{1, 2, \dots, n\}.$$

Домножим эти миноры на их алгебраические дополнения в $\det A$*. Тогда величина* $\det A$ *равна сумме таких произведений по всем возможным выборкам* k *элементов* $(\beta_1, \dots, \beta_k)$ *из* $\{1, 2, \dots, n\}$*:*

$$\det A =$$

$$= \sum_{1 \le \beta_1 < \dots < \beta_k \le n} A\begin{pmatrix}\alpha_1 \dots \alpha_k \\ \beta_1 \dots \beta_k\end{pmatrix} A\begin{pmatrix}\alpha_{k+1} \dots \alpha_n \\ \beta_{k+1} \dots \beta_n\end{pmatrix} (-1)^{\left(\substack{\alpha_1 + \dots + \alpha_k + \\ + \beta_1 + \dots + \beta_k}\right)}. \tag{7.1}$$

Упражнение 20. *Доказать, что формула* (7.1) *содержит* C_n^k *слагаемых.*

Таким образом, формула (7.1) является естественным обобщением формулы (5.2) разложения определителя по элементам его строки. На основании свойства равноправия строк и столбцов определителя (теорема 4.1), в теореме Лапласа можно вместо фиксирования строк фиксировать столбцы:

$$\det A =$$

$$= \sum_{1\le\alpha_1<\dots<\alpha_k\le n} A\begin{pmatrix}\alpha_1\dots\alpha_k\\ \beta_1\dots\beta_k\end{pmatrix} A\begin{pmatrix}\alpha_{k+1}\dots\alpha_n\\ \beta_{k+1}\dots\beta_n\end{pmatrix}(-1)^{\left(\substack{\alpha_1+\dots+\alpha_k+\\+\beta_1+\dots+\beta_k}\right)}.$$

Доказательство. Пусть $M_1,\dots,M_N$ при $N \stackrel{\text{def}}{=} \mathsf{C}_n^k$ — все различные миноры k-го порядка, которые находятся в строках $\alpha_1,\dots,\alpha_k$, пусть $\widetilde{M}_1,\dots,\widetilde{M}_N$ — их алгебраические дополнения. Составим сумму

$$M_1\widetilde{M}_1+\dots+M_N\widetilde{M}_N. \tag{7.2}$$

Нам нужно доказать, что эта сумма совпадает с $\det A$. Сделаем это следующим образом. Будем считать все элементы матрицы A переменными. Тогда формально $\det A$ — это полиномиальная функция от своих элементов, множество одночленов которой имеет вид

$$\left\{(-1)^{\mathfrak{inv}\,(\alpha_1,\dots,\alpha_n)+\mathfrak{inv}\,(\beta_1,\dots,\beta_n)}a_{\alpha_1\beta_1}\times\dots\times a_{\alpha_n\beta_n}\right\} \tag{7.3}$$

где $(\alpha_1,\dots,\alpha_n)$ и $(\beta_1,\dots,\beta_n)$ пробегают все возможные перестановки элементов $1,\dots,n$. Сумма (7.2) тоже является полиномиальной функцией тех же переменных; она, в свою очередь, порождается множеством

$$\{M_1\widetilde{M}_1\}\cup\dots\cup\{M_N\widetilde{M}_N\} = \bigcup_{j=1}^{N}\{M_j\widetilde{M}_j\}, \tag{7.4}$$

т.е. объединением всех мономов, содержащихся в слагаемых (7.2). Покажем, что эти два полинома — $\det A$ и (7.2) — тождественно равны, т.е.[11] совпадают множества их одночленов.

[11] Глава 3, §11.1

I. Докажем, что любой элемент множества (7.4) входит в состав множества (7.3) (т.е. входит в состав суммы (3.14)). Выберем в разложениях миноров

$$A\begin{pmatrix}\alpha_1\ldots\alpha_k\\ \beta_1\ldots\beta_k\end{pmatrix} \quad \text{и} \quad A\begin{pmatrix}\alpha_{k+1}\ldots\alpha_n\\ \beta_{k+1}\ldots\beta_n\end{pmatrix}$$

произвольные мономы, пусть они имеют вид

$$a_{\alpha_1 j_1}\times\cdots\times a_{\alpha_k j_k} \quad \text{и} \quad a_{\alpha_{k+1} j_{k+1}}\times\cdots\times a_{\alpha_n j_n}$$

где $(j_1,\ldots,j_k)$ — перестановка чисел $\beta_1,\ldots,\beta_k$, а $(j_{k+1},\ldots,j_n)$ — перестановка чисел $\beta_{k+1},\ldots,\beta_n$ соответственно. Эти произведения будут входить в разложения миноров со знаками соответственно

$$(-1)^{\mathfrak{inv}\,(\alpha_1,\ldots,\alpha_k)+\mathfrak{inv}\,(j_1,\ldots,j_k)} \quad \text{и} \quad (-1)^{\mathfrak{inv}\,(\alpha_{k+1},\ldots,\alpha_n)+\mathfrak{inv}\,(j_{k+1},\ldots,j_n)}.$$

Поскольку $\alpha_1<\cdots<\alpha_k$ и $\alpha_{k+1}<\cdots<\alpha_n$, то $\mathfrak{inv}\,(\alpha_1,\ldots,\alpha_k)=0$ и $\mathfrak{inv}\,(\alpha_{k+1},\ldots,\alpha_n)=0$, следовательно знаки при произведениях равны

$$(-1)^{\mathfrak{inv}\,(j_1,\ldots,j_k)} \quad \text{и} \quad (-1)^{\mathfrak{inv}\,(j_{k+1},\ldots,j_n)}$$

соответственно. Если теперь перемножить эти выражения с учетом их знаков, а также знака, который приписывается алгебраическому дополнению минора M, то в произведении $M\widetilde{M}$ будет содержаться моном

$$a_{\alpha_1 j_1}\times\cdots\times a_{\alpha_k j_k}a_{\alpha_{k+1} j_{k+1}}\times\cdots\times a_{\alpha_n j_n} \tag{7.5}$$

со знаком

$$(-1)^{\mathfrak{inv}\,(j_1,\ldots,j_k)+\mathfrak{inv}\,(j_{k+1},\ldots,j_n)+\alpha_1+\cdots+\alpha_k+\beta_1+\cdots+\beta_k}. \tag{7.6}$$

То же произведение (7.5) войдет и в сумму (3.14), задающую $\det A$, поскольку индексы $\alpha_1,\ldots,\alpha_n$ все различны и $j_1,\ldots,j_n$ все различны. В этой сумме знак у произведения будет

$$(-1)^{\mathfrak{inv}\,(\alpha_1,\ldots,\alpha_k,\alpha_{k+1},\ldots,\alpha_n)+\mathfrak{inv}\,(j_1,\ldots,j_k,j_{k+1},\ldots,j_n)}. \tag{7.7}$$

Покажем, что величины (7.6) и (7.7) одинаковы. Для этого воспользуемся теоремой 3.2. Имеем:

$$\mathfrak{inv}\,(\alpha_1,\ldots,\alpha_k,\alpha_{k+1},\ldots,\alpha_n)=\mathfrak{inv}\,(\alpha_1,\ldots,\alpha_k)+\mathfrak{inv}\,(\alpha_{k+1},\ldots,\alpha_n)+$$

$$+(\alpha_1+\cdots+\alpha_k)-\mathsf{C}_{k+1}^2=(\alpha_1+\cdots+\alpha_k)-\mathsf{C}_{k+1}^2;$$

$$\mathfrak{inv}\,(j_1,\ldots,j_k,j_{k+1},\ldots,j_n)=\mathfrak{inv}\,(j_1,\ldots,j_k)+\mathfrak{inv}\,(j_{k+1},\ldots,j_n)+$$
$$+(j_1+\cdots+j_k)-\mathsf{C}_{k+1}^2=$$
$$=\mathfrak{inv}\,(j_1,\ldots,j_k)+\mathfrak{inv}\,(j_{k+1},\ldots,j_n)+(\beta_1+\cdots+\beta_k)-\mathsf{C}_{k+1}^2$$

(числа $j_1,\ldots,j_k$ представляют собой переставленные числа $\beta_1,\ldots,\beta_k$). Таким образом, величина (7.7) равна

$$(-1)^{\alpha_1+\cdots+\alpha_k+\beta_1+\cdots+\beta_k+\mathfrak{inv}\,(j_1,\ldots,j_k)+\mathfrak{inv}\,(j_{k+1},\ldots,j_n)-2\mathsf{C}_{k+1}^2}.$$

Число $2\mathsf{C}_{k+1}^2$ — четное, поэтому получившееся выражение совпадает с (7.6).

Таким образом, мы доказали, что множество (7.4) содержится во множестве (7.3).

II. Докажем теперь обратное, т.е. то, что множество (7.4) содержит множество (7.3). Произвольный одночлен в формуле $\det A$ можно представить в виде (7.3), где $\alpha_1,\ldots,\alpha_k$ — номера строк, упомянутых в формулировке теоремы. Этот же одночлен войдет и в состав произведения $M_j\widetilde{M}_j$, где M_j — минор, стоящий в строках $\alpha_1,\ldots,\alpha_k$ и в столбцах $\beta_1,\ldots,\beta_k$. В самом деле, в состав M_j войдет одночлен

$$(-1)^{\mathfrak{inv}\,(\alpha_1,\ldots,\alpha_k)+\mathfrak{inv}\,(\beta_1,\ldots,\beta_k)}a_{\alpha_1\beta_1}\times\cdots\times a_{\alpha_k\beta_k},$$

а в состав $\widetilde{M}_j$ — одночлен

$$(-1)^{\mathfrak{inv}\,(\alpha_{k+1},\ldots,\alpha_n)+\mathfrak{inv}\,(\beta_{k+1},\ldots,\beta_n)}a_{\alpha_{k+1}\beta_{k+1}}\times\cdots\times a_{\alpha_n\beta_n},$$

домноженный на $(-1)^{\alpha_1+\cdots+\alpha_k+\beta_1+\cdots+\beta_k}$. Произведение этих одночленов (с учетом формулы (3.13)) даст (7.3).

На основании пунктов **I** и **II** доказательства множества (7.4) и (7.3) совпадают.

III. Наконец, множества $\{M_j\widetilde{M}_j\}$ и $\{M_k\widetilde{M}_k\}$ не имеют общих элементов при $j\neq k$, поскольку миноры M_j и M_k различаются хотя бы одним столбцом. Следовательно, сумма (7.2) действительно совпадает с $\det A$. ■

Перейти к оглавлению на странице: 460

Пример 7.1. Вычислить определитель примера 5.2 с помощью теоремы Лапласа, взяв $\alpha_1 = 2, \alpha_2 = 5$.

Решение.

$$\begin{vmatrix} 2 & 2 & 1 & 3 & 4 \\ 3 & 1 & 2 & 3 & 1 \\ 4 & -1 & 2 & 4 & -2 \\ 1 & -1 & 1 & 1 & 2 \\ 4 & -1 & 2 & 5 & 6 \end{vmatrix} =$$

$$= (-1)^{2+5+1+2} \begin{vmatrix} 3 & 1 \\ 4 & -1 \end{vmatrix} \cdot \begin{vmatrix} 1 & 3 & 4 \\ 2 & 4 & -2 \\ 1 & 1 & 2 \end{vmatrix} + (-1)^{2+5+1+3} \begin{vmatrix} 3 & 2 \\ 4 & 2 \end{vmatrix} \cdot \begin{vmatrix} 2 & 3 & 4 \\ -1 & 4 & -2 \\ -1 & 1 & 2 \end{vmatrix} +$$

$$+(-1)^{2+5+1+4} \begin{vmatrix} 3 & 3 \\ 4 & 5 \end{vmatrix} \cdot \begin{vmatrix} 2 & 1 & 4 \\ -1 & 2 & -2 \\ -1 & 1 & 2 \end{vmatrix} + (-1)^{2+5+1+5} \begin{vmatrix} 3 & 1 \\ 4 & 6 \end{vmatrix} \cdot \begin{vmatrix} 2 & 1 & 3 \\ -1 & 2 & 4 \\ -1 & 1 & 1 \end{vmatrix}$$

$$+(-1)^{2+5+2+3} \begin{vmatrix} 1 & 2 \\ -1 & 2 \end{vmatrix} \cdot \begin{vmatrix} 2 & 3 & 4 \\ 4 & 4 & -2 \\ 1 & 1 & 2 \end{vmatrix} + (-1)^{2+5+2+4} \begin{vmatrix} 1 & 3 \\ -1 & 5 \end{vmatrix} \cdot \begin{vmatrix} 2 & 1 & 4 \\ 4 & 2 & -2 \\ 1 & 1 & 2 \end{vmatrix} +$$

$$+(-1)^{2+5+2+5} \begin{vmatrix} 1 & 1 \\ -1 & 6 \end{vmatrix} \cdot \begin{vmatrix} 2 & 1 & 3 \\ 4 & 2 & 4 \\ 1 & 1 & 1 \end{vmatrix} + (-1)^{2+5+3+4} \begin{vmatrix} 2 & 3 \\ 2 & 5 \end{vmatrix} \cdot \begin{vmatrix} 2 & 2 & 4 \\ 4 & -1 & -2 \\ 1 & -1 & 2 \end{vmatrix} +$$

$$+(-1)^{2+5+3+5} \begin{vmatrix} 2 & 1 \\ 2 & 6 \end{vmatrix} \cdot \begin{vmatrix} 2 & 2 & 3 \\ 4 & -1 & 4 \\ 1 & -1 & 1 \end{vmatrix} + (-1)^{2+5+4+5} \begin{vmatrix} 3 & 1 \\ 5 & 6 \end{vmatrix} \cdot \begin{vmatrix} 2 & 2 & 1 \\ 4 & -1 & 2 \\ 1 & -1 & 1 \end{vmatrix} =$$

$$= +(-7)(-16) - (-2)44 + 3 \cdot 20 - 14 \cdot (-4) + 4 \cdot (-10) - 8 \cdot 10 + 7 \cdot 2 +$$

$$+4 \cdot (-40) - 10 \cdot (-3) + 13 \cdot (-5) = 15.$$

$\triangle$

С точки зрения вычислительной эффективности теорема Лапласа не представляет интереса. Она, однако, полезна для доказательства некоторых теоретических результатов.

Следствие 7.1.

$$\begin{vmatrix} a_{11} & \dots a_{1k} & 0 & \dots 0 \\ \dots & \dots & \dots & \dots \\ a_{k1} & \dots a_{kk} & 0 & \dots 0 \\ a_{k+1,1} & \dots a_{k+1,k} & a_{k+1,k+1} & \dots a_{k+1,n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots a_{nk} & a_{n,k+1} & \dots a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} \dots a_{1k} \\ \dots \quad \dots \\ a_{k1} \dots a_{kk} \end{vmatrix} \cdot \begin{vmatrix} a_{k+1,k+1} \dots a_{k+1,n} \\ \vdots \qquad \vdots \\ a_{n,k+1} \quad \dots a_{nn} \end{vmatrix}.$$

Доказательство проводится применением формулы (7.1) для первых k строк. ■

Следствие 7.2. *Пусть* $\mathsf{A}_1, \dots, \mathsf{A}_k$ *— квадратные матрицы (не обязательно одинаковых порядков).*

$$\det \begin{pmatrix} \mathsf{A}_1 & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ * & \mathsf{A}_2 & \mathbb{O} & \mathbb{O} \\ * & * & \ddots & \\ * & * & * & \mathsf{A}_k \end{pmatrix} = \det \mathsf{A}_1 \times \dots \times \det \mathsf{A}_k.$$

Историческая справка

Лаплас Пьер Симон (Laplace Pierre Simon, 1749–1827) родился в семье крестьянина. С его рождением связана какая-то тайна. С ранних лет у мальчика были богатые покровители, которые заботились о его образовании и дальнейшей судьбе. Предполагают, что ученый был незаконнорожденным отпрыском местного синьора.

В 1766 он отправился в Париж. С рекомендательным письмом от одного влиятелього лица молодой человек пришел к Д'Аламберу, но тот не принял посетителя. Тогда Л. пришла в голову мысль самому изложить в письме великому ученому свои взгляды на механику. Д'Аламбер ответил незамедлительно:

> *"Милостивый государь! Вы имели случай убедиться, как мало обращаю я внимания на рекомендации, но вам они совершенно не нужны, вы зарекомендовали себя сами, и мне этого совершенно достаточно, моя помощь к вашим услугам. Приходите же, я жду вас."*

Через некоторое время Л. получил место преподавателя в Военной школе. Школа готовила, в-частности, артиллерийских офицеров. Одному из них Л. в характеристике написал что тот "*весьма способен к математике*", что и позволило кадету получить первый офицерский чин. Офицера звали Наполеон Бонапарт, и он не забыл своего учителя. После переворота 18 брюмера Наполеон назначил Л. министром внутренних дел, но через полгода заменил его на этой должности своим братом Люсьеном. Позже, на острове Св.Елены сосланный император писал:

"Первокласный геометр Л. вскоре заявил себя администратором менее чем посредственным; первые шаги его на этом поприще убедили нас в том, что мы в нем обманулись. Замечательно, что ни один из вопросов практической жизни не представлялся Лапласу в его истинном свете. Он везде искал чего-то, идеи его отличались загадочностью, и, наконец, он был насквозь проникнут духом бесконечно малых, который вносил в администрацию."

(Л. хотел реорганизовать работу судов на научной основе так, чтобы их приговоры с наибольшей вероятностью соответствовали сути дела.)

Почти все современники сходились в одном: Л. был великий ученый и беспринципный человек. При любой смене правительства он становился на сторону победивших, получая от них должности и награды. В 1814 г. он подал голос за низложение Наполеона. После реставрации Бурбонов получил пэрство и титул маркиза. Академик, вице-президент Сената, канцлер...

Был атеистом. Рассказывают, что когда Наполеон ознакомился с его "Трактатом по небесной механике", то воскликнул:

> *— Господин Лаплас! Ньютон в своей книге говорил о Боге, в вашей же книге, которую я уже просмотрел, я не встретил слова "Бог" ни разу!*

На что автор ответил:

> *— Гражданин Первый консул, в подобной гипотезе я не нуждался.*[12]

Научное наследие Л. огромно, у него есть труды в астрономии, математике, механике, мат.физике, теории вероятностей: Л-в азимут, уравнение Л., закон Био-Савара-Л. (о напряженности магнитного поля, создаваемого электрическим током) и пр. (в "Математической энциклопедии"с его именем связываются 12 разделов).

<u>Космогоническая гипотеза Лапласа</u>

Первоначально была раскаленная газовая туманность, которая медленно вращалась вследствие каких-то внутренних течений пара (газа) или в результате того, что близко от туманности когда-то прошло крупное космическое тело.

[12] Лагранж, которому Наполеон передал содержание этого разговора заметил: "Ах, но это — прекрасная гипотеза! Она так много объясняет."

Охлаждаясь, туманность сжималась и вращалась все быстрее и быстрее, подобно тому, как фигурист, прижимая руки к туловищу, ускоряет свое вращение. С ростом угловой скорости увеличивались центробежные силы, действующие на частицы экватора туманности и она все больше сплющивалась у полюсов. Наконец центробежные силы на экваторе стали равны силам притяжения к центру и от вращающейся туманности отделилось кольцо газа, которое вследствие взаимного притяжения частиц затем собралось в космическое тело — планету. Туманность продолжала сжиматься, ее скорость росла, отделилось второе кольцо, оно образовало другую планету и т.д. Так образовалась Солнечная система. Лаплас:

> *"Массы должны были принять сфероидальную форму с вращательным движением, направленным в сторону их обращения."*

Подтверждение гипотезы:

1) все планеты движутся по своим орбитам в прямом направлении (т.е. в том же, в каком вращается Солнце);
2) обращение спутников — тоже прямое;
3) орбиты планет и спутников почти круговые;
4) плоскости орбит всех планет и спутников почти совпадают с плоскостью экватора.

Книга Л. вышла в 1796 г., а спустя всего год В.Гершель открыл, что два спутника Урана движутся в обратном направлении. Однако гипотеза Л. просуществовала более ста лет. В ее рамках не удалось объяснить, почему 98% момента количества движения Солнечной системы приходится на орбитальное движение планет, хотя их суммарная масса в 750 раз меньше массы Солнца. Кроме того, развитие газодинамики показало, что вращающееся кольцо не может сгуститься в планету. Однако физические эффекты «остывания» и «гравитационного сжатия», которыми пользовался Л., являются главными и в современных моделях образования Солнечной системы.

Лаплас умер спустя 100 лет после смерти Ньютона и его последние слова фактически повторяют известное высказывание великого предшественника[13]:

> *"То, что мы знаем, так ничтожно по сравнению с тем, чего мы не знаем."*

§8. Теорема Бине–Коши

Пусть произведение двух матриц дает квадратную:

$$C_{m\times m} = A_{m\times n}\cdot B_{n\times m}.$$

Задача: выразить $\det C$ через миноры матриц A и B.

[13] См. с. 161.

Теорема 8.1. *[Бине, Коши]*

$$\det C = \begin{cases} 0 & \text{если } m > n; \\ \det A \det B & \text{если } m = n; \\ \displaystyle\sum_{1\le\beta_1<\dots<\beta_m\le n} A\begin{pmatrix} 1 & 2 & \dots m \\ \beta_1 & \beta_2 & \dots \beta_m \end{pmatrix} B\begin{pmatrix} \beta_1 & \beta_2 & \dots \beta_m \\ 1 & 2 & \dots m \end{pmatrix} \\ & \text{если } m < n. \end{cases} \tag{8.1}$$

Доказательство проведем только для второй из формул (8.1), для простоты рассуждений рассмотрев случай $m = n = 3$. Составим вспомогательный определитель порядка $m+n=6$:

$$\mathfrak{B} = \begin{vmatrix} A & \mathbb{O} \\ -E & B \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} & & & \\ a_{21} & a_{22} & a_{23} & & \mathbb{O} & \\ a_{31} & a_{32} & a_{33} & & & \\ -1 & 0 & 0 & b_{11} & b_{12} & b_{13} \\ 0 & -1 & 0 & b_{21} & b_{22} & b_{23} \\ 0 & 0 & -1 & b_{31} & b_{32} & b_{33} \end{vmatrix} \overset{\text{т. Лапласа}}{=} \det A \det B. \tag{8.2}$$

Действуя над его столбцами, добьемся, чтобы все элементы в правом нижнем углу обратились в нуль. Для этого прибавим к 4-му столбцу 1-й, умноженный на b_{11}, 2-й, умноженный на b_{21} и 3-й, умноженный на b_{31}. Величина определителя от этого не изменится:

$$\mathfrak{B} = \begin{vmatrix} a_{11} & a_{12} & a_{13} & d_{11} & 0 & 0 \\ a_{21} & a_{22} & a_{23} & d_{21} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & d_{31} & 0 & 0 \\ -1 & 0 & 0 & 0 & b_{12} & b_{13} \\ 0 & -1 & 0 & 0 & b_{22} & b_{23} \\ 0 & 0 & -1 & 0 & b_{32} & b_{33} \end{vmatrix}, \quad \text{где} \quad \begin{cases} d_{11} = a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31}, \\ d_{21} = a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31}, \\ d_{31} = a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31}. \end{cases}$$

Видим, что элементы d_{j1} формируются по закону умножения матриц A и B: $j-$я строчка матрицы A умножается на 1-й столбец матрицы B. Следовательно, $d_{j1} = c_{j1}$. Проделав аналогичные преобразования, ведущие к обнулению оставшихся элементов матрицы B, получим в результате:

$$\mathfrak{B} = \begin{vmatrix} A & C \\ -E & \mathbb{O} \end{vmatrix}.$$

Снова применяем теорему Лапласа, теперь уже для миноров последних трех строк (т.е. $\alpha_1 = 4, \alpha_2 = 5, \alpha_3 = 6$):

$$\mathfrak{B} = (-1)^{4+5+6+1+2+3} \det(-E) \cdot \det C = \det C \tag{8.3}$$

Сравнивая (8.2) и (8.3), получаем справедливость второй из формул (8.1). ■

Упражнение 21. *Воспользовавшись идеей доказательства, показать справедливость остальных равенств теоремы.*

Теорема 8.2. *Для произвольных* $\{a_j, b_j\}_{j=1}^n \subset \mathbb{C}$ *справедливо* **неравенство Коши***:*

$$\sum_{j=1}^n |a_j|^2 \sum_{j=1}^n |b_j|^2 \geq \left| \sum_{j=1}^n a_j \overline{b_j} \right|^2 . \tag{8.4}$$

Доказательство. Применим теорему 8.1 к произведению следующих матриц:

$$A = \begin{pmatrix} a_1\, a_2 \dots a_n \\ b_1\, b_2 \dots b_n \end{pmatrix}, \qquad B = \begin{pmatrix} \overline{a_1}\, \overline{b_1} \\ \overline{a_2}\, \overline{b_2} \\ \vdots \;\; \vdots \\ \overline{a_n}\, \overline{b_n} \end{pmatrix}.$$

С одной стороны:

$$\det(A\,B) = \begin{vmatrix} a_1\overline{a_1} + a_2\overline{a_2} + \dots + a_n\overline{a_n} & a_1\overline{b_1} + a_2\overline{b_2} + \dots + a_n\overline{b_n} \\ \overline{a_1}b_1 + \overline{a_2}b_2 + \dots + \overline{a_n}b_n & b_1\overline{b_1} + b_2\overline{b_2} + \dots + b_n\overline{b_n} \end{vmatrix} =$$

$$= (|a_1|^2 + |a_2|^2 + \dots + |a_n|^2)(|b_1|^2 + |b_2|^2 + \dots + |b_n|^2) -$$

$$- \left| a_1\overline{b_1} + a_2\overline{b_2} + \dots + a_n\overline{b_n} \right|^2 .$$

С другой стороны, применяя к тому же произведению третью из формул (8.1), получим:

$$\det(A\,B) = \begin{vmatrix} a_1\, a_2 \\ b_1\, b_2 \end{vmatrix} \cdot \begin{vmatrix} \overline{a_1}\, \overline{b_1} \\ \overline{a_2}\, \overline{b_2} \end{vmatrix} + \begin{vmatrix} a_1\, a_3 \\ b_1\, b_3 \end{vmatrix} \cdot \begin{vmatrix} \overline{a_1}\, \overline{b_1} \\ \overline{a_3}\, \overline{b_3} \end{vmatrix} + \dots +$$

$$+\begin{vmatrix} a_{n-1} & a_n \\ b_{n-1} & b_n \end{vmatrix} \cdot \begin{vmatrix} \overline{a_{n-1}} & \overline{b_{n-1}} \\ \overline{a_n} & \overline{b_n} \end{vmatrix} = \sum_{1\le j<k\le n} |a_j b_k - a_k b_j|^2.$$

Сравнивая два представления, получаем:

$$\sum_{j=1}^{n} |a_j|^2 \sum_{j=1}^{n} |b_j|^2 - \left|\sum_{j=1}^{n} a_j \overline{b_j}\right|^2 = \sum_{1\le j<k\le n} |a_j b_k - a_k b_j|^2,$$

и, следовательно, левая часть должна быть неотрицательной. ■

<u>*Историческая справка*</u>

Коши Огюстен Луи (Cauchy, Augustin Louis, 1789–1857) родился 21 августа, месяц спустя штурма Бастилии. Отец его был глубоко верующим католиком, и таким же стал сам К. Первоначальный интерес к гуманитарным дисциплинам (участвовал в конкурсах на написание стихов на латинском и французком языках) к 14 годам сменился страстью к математике. По протекции отца, получил поддержку от Лапласа и Лагранжа. Последний как-то сказал:

> *Вы видите этого маленького молодого человека, он заменит нас всех как геометров.*

Такая поддержка уже была необходима для молодого таланта: эпоха быстрых карьерных взлетов, характерная для революции[14] и начала наполеоновской империи, сменялась эпохой застоя и протекций. Двадцатилетнему К. надо было бороться за место под солнцем.

На первых порах ему повезло. После реставрации Бурбонов была произведена чистка Академии Наук, и из нее были исключены республиканец (“организатор победы”) Карно и верный сторонник Бонапарта Монж. На одно из освободившихся мест был назначен К. Эта гнусная чистка была плохо встречеа в Академии и вообще в ученом мире: оба исключенных были уважаемыми учеными. После этой истории К. нажил много недоброжелателей.

С 1817 года К. стал преподавать в Политехническом институте — самом престижном вузе Франции, готовившем инженеров. Он читал лекции по математическому анализу — и сразу же стал составлять учебник, который послужил многим поколениям студентов. Именно в учебнике К. впервые появилось привычное для современных курсов анализа определение непрерывной функции, теория сходимости рядов.

К. было 40 лет, когда разразилась июльская революция 1830 года. В его жизни произошел резкий перелом. По своим политическим взглядам К. был ультрароялистом — сторонником “альянса трона и алтаря”. К. поддерживал все начинания католической церкви, а в Академии выступал против исследований, противоречащих — по его мнению — религии.

[14]“Что такое революция?” — “Это десять тысяч новых вакансий.”

Давайте будем развивать с упоением математические знания без того, чтобы распространять их на чужие области, давайте не будем воображать себе, что мы можем применить формулы к истории или же, что можем вывести законы морали из теорем алгебры или интегрального исчисления.

К 1830 г. клерикализм вызывал растущее сопротивление французского общества. Пойдет ли духовенство против науки и образования, возьмет ли под контроль прессу и законодательство, идеологию? Либералы популяризовали тему клерикального заговора, требовали разгона околоцерковных светских организаций, запрещения деятельности иезуитов. После июльской революции им удалось осуществить свои планы. К. воспринял это как личную трагедию. Почти все его коллеги, студенты, все члены его семьи присягнули на верность новому королю — герцогу Орлеанскому. Сам К. остался верен Карлу X, отправился за ним в эмиграцию, и в течение 8 лет жил в Италии и Австрии, следуя за двором изгнанного монарха. Лучший — к тому времени — математик Франции был лишен всех должностей и возможности преподавания.

Коши принимал участие в многочисленных академических комиссиях, занимавшихся рассмотрением научных исследований, представленных молодыми учеными. В исследованиях по истории математики его принято упрекать за недостаточное внимание к работам Абеля и Галуа, рецензентом которых он был назначен. Несмотря на формальную справедливость этих упреков, в оправдание К. следует отметить исключительную загруженность ученого как научными трудами так и общественными обязанностями. С целью "частичной реабилитации" К. приведем и противоположные свидетельства.

По признанию великого французского математика Шарля Эрмита (1822–1901) он стал убежденным католиком именно под влиянием Коши (и после того как К. оказал ему поддержку во время тяжелой болезни).

Из заметки академика А.Н.Крылова "Памяти М.В.Остроградского":

"Остроградский... просил отца отправить его в Париж учиться у знаменитых математиков: Лапласа, Пуассона, Фурье, Ампера. Отец М.В. на это согласился, и в мае 1822 г. Остроградский выехал за границу с каким-то попутчиком, но, доехав до Чернигова, был вынужден вернуться обратно, ибо был своим попутчиком, "как рассказывают", обокраден.

В августе, т.е. после уборки хлеба, отец ему вновь дал на поездку денег, и на этот раз Остроградский доехал до Парижа благополучно и стал слушать лекции в Сорбонне и Collédge de France. Своими дарованиями он обратил на себя внимание французских математиков, в особенности, Коши, который в начале своего мемуара: "Sur les intégrales prises entre des limites imagineures", изданного отдельно и в извлечении помещенного в "Bulletin de Ferussas" за апрель 1825 г., упомянув о Лапласе и Бриссоне, занимавшихся подобными вопросами, говорит:

Наконец, молодой русский, одаренный большой проницательностью и весьма сведущий в анализе бесконечно малых, г.Остроградский, воспользовавшись этими интегралами и их преобразованием в обыкновенные, дал новые доказательства формул, о которых я упоминаю, а также обобщил другие формулы, находящиеся в моей статье, помещенной в 19-й тетради журнала политехнической школы.

Г.Остроградский любезно сообщил мне главные результаты своей работы.

В 1826 г. с Остроградским случилось происшествие, о котором рассказывали старики, его знавшие, но о котором умалчивает Сомов в очерках его жизни и его трудов.

В 1821 г. началась война греков за освобождение от турецкого владычества, приведшая к Наваринскому сражению в 1827 г. За все эти семь лет в Архипелаге, постепенно усиливаясь, развилось такое пиратство, что коммерческое мореплавание стало почти невозможным без военного конвоя, — грабили и алжирцы и тунисцы, и турки, и левантинцы, и египтяне, и греки.

В те времена банковские сношения еще не были развиты, и для пересылки денег в Париж отцу Остроградского приходилось ... покупать у экспортеров хлеба в Ростове, Херсоне или Одессе вексель какого-нибудь марсельского купца, который затем поручал своему дебитору или кредитору оплатить этот вексель в Париже.

По какой-то причине в 1826 г. Остроградский денег своевременно от отца не получил, задолжал в гостинице за "харч и постой" и по жалобе хозяина был посажен в Clichy, т.е. долговую тюрьму в Париже. Здесь он, видимо, особенно усердно занимался математикой и написал свою знаменитую работу: "Mémoire sur la propagation des ondes dans un bassin cylindrique" и послал эту работу Коши.

Коши в ноябре 1826 г. представил этот мемуар с самым лестным отзывом Парижской Академии, которая удостоила эту работу высшего отличия – напечатания в "Mémoires des Savants étrangers à l'Académie"... Более того, Коши сам, не будучи богатым человеком, выкупил Остроградского из "долгового" и вместе со своими коллегами рекомендовал его на должность преподавателя математики в Collége Henri IX, в каковой он и оставался до конца 1827 г..."

§9. Определители специального вида

В предыдущих параграфах мы иллюстрировали теоретические результаты на примерах числовых определителей. Довольно часто, однако, на практике возникает необходимость вычислять определители, элементы которых зависят от параметров. Метод Гаусса оказывается не слишком приспособленным для такой задачи.

Пример 9.1. Вычислить

$$\left|\begin{array}{rrrr} \alpha+1 & \alpha^2+1 & \alpha^2-1 & \alpha \\ \alpha^2+\alpha+1 & \alpha^2-\alpha+1 & \alpha^2 & 1 \\ 2\,\alpha+1 & \alpha^2+2 & \alpha & \alpha^2-1 \\ 2\,\alpha & 2\,\alpha^2+2\,\alpha+1 & \alpha^2-\alpha-1 & \alpha+1 \end{array}\right|.$$

Решение. Разложение по формуле (3.14) даст величину этого определителя в виде полинома от α. С другой стороны, если для его вычисления мы попытаемся применить метод Гаусса, то на первом же шаге элементы преобразованного определителя окажутся дробно–рациональными функциями от параметра α. Понятно, что после приведения определителя к треугольному виду и перемножения стоящих на диагонали дробей мы, в конце концов, получим тот же ответ полиномиального вида, но сам факт, что для его получения потребовалось "выйти за пределы" множества полиномиальных функций не свидетельствует в пользу метода Гаусса ... $\triangle$

Что делать? Универсальных методов вычисления подобных определителей (отличных, естественно, от определения (3.14)) нет. Успех во многом будет зависеть от искусства вычислителя. В этом параграфе мы проиллюстрируем несколько полезных приемов на примерах вычисления часто встречающихся определителей.

Определитель Вандермонда

$$V(x_1,\dots,x_n) \stackrel{\text{def}}{=} \det\left[x_j^{k-1}\right]_{j,k=1}^{n} \stackrel{\text{def}}{=} \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \dots & & & & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix}_{n\times n} =$$

Вычтем последовательно из n-го, $(n-1)$-го, ..., второго столбца предыдущий, домноженный на x_1:

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

разложим по первой строке, и вынесем общие множители элементов строк получившегося определителя $(n-1)$-го порядка:

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

$$= (x_2 - x_1) \times \cdots \times (x_n - x_1) V(x_2, \ldots, x_n).$$

Определитель $V(x_2, \ldots, x_n)$ имеет тот же вид, что и исходный, но на единицу меньший порядок. Его можно преобразовать аналогично:

$$V(x_2, \ldots, x_n) = (x_3 - x_2) \times \cdots \times (x_n - x_2) V(x_3, \ldots, x_n).$$

Продолжая процесс далее, приходим к окончательному ответу

$$V(x_1, \ldots, x_n) = \prod_{1 \le j < k \le n} (x_k - x_j). \tag{9.1}$$

Упражнение 22. *Сколько сомножителей содержится в этом произведении?*

Упражнение 23. *Вычислить*

$$\begin{vmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ x_1 & x_2 & x_3 & x_4 & 1 & 0 & 0 \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 & 2\,x_1 & 1 & 0 \\ x_1^3 & x_2^3 & x_3^3 & x_4^3 & 3\,x_1^2 & 3\,x_1 & 1 \\ x_1^4 & x_2^4 & x_3^4 & x_4^4 & 4\,x_1^3 & 6\,x_1^2 & 4\,x_1 \\ x_1^5 & x_2^5 & x_3^5 & x_4^5 & 5\,x_1^4 & 10\,x_1^3 & 10\,x_1^2 \\ x_1^6 & x_2^6 & x_3^6 & x_4^6 & 6\,x_1^5 & 15\,x_1^4 & 20\,x_1^3 \end{vmatrix}.$$

Ганкелев определитель

Определение 9.1. **Ганкелевой**[15] матрицей называется симметричная матрица следущего вида:

$$H = \begin{bmatrix} h_0 & h_1 & h_2 & \ldots h_{n-2} & h_{n-1} \\ h_1 & h_2 & h_3 & \ldots h_{n-1} & h_n \\ h_2 & h_3 & h_4 & \ldots h_n & h_{n+1} \\ \ldots & & & & \ldots \\ h_{n-1} & h_n & h_{n+1} & \ldots h_{2n-3} & h_{2n-2} \end{bmatrix}_{n \times n} = [h_{j+k-2}]_{j,k=1}^n .$$

Элементы $h_0, \ldots, h_{n-1}, h_n, \ldots, h_{2n-2}$ — **образующие** ганкелевой матрицы.

[15] **Ганкель** (Ханкель) **Герман** (Hankel Hermann, 1839–1873) — немецкий математик.

Перейти к оглавлению на странице: 460

Теорема 9.1. *Если* $h_j = x_1^j + \cdots + x_n^j$ *при* $j \in \{0, \ldots, 2n-2\}$, *то*

$$\det H = \prod_{1 \le j < k \le n} (x_k - x_j)^2 \tag{9.2}$$

Доказательство. Матрицу H можно представить в виде произведения:

$$H = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ x_1 & x_2 & x_3 & \ldots & x_n \\ \ldots & & & & \ldots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \ldots & x_n^{n-1} \end{bmatrix} \cdot \begin{bmatrix} 1 & x_1 & x_1^2 \ldots x_1^{n-1} \\ 1 & x_2 & x_2^2 \ldots x_2^{n-1} \\ \ldots & & \ldots \\ 1 & x_n & x_n^2 \ldots x_n^{n-1} \end{bmatrix}$$

На основании теоремы 8.1 (Бине–Коши), $\det H$ равен тогда произведению двух определителей Вандермонда:

$$\det H = V(x_1, \ldots, x_n)^2 = \prod_{1 \le j < k \le n} (x_k - x_j)^2.$$

■

Теорема 9.2. *Если* $h_{j+k-2} = \frac{1}{j+k-1}$ *при* $\{j, k\} \in \{1, \ldots, n\}$, *то определитель* **матрицы Гильберта**

$$\mathfrak{H}_n = \begin{bmatrix} 1 & 1/2 & 1/3 & \ldots 1/n \\ 1/2 & 1/3 & 1/4 & \ldots 1/(n+1) \\ 1/3 & 1/4 & 1/5 & \ldots 1/(n+2) \\ \vdots & & & \vdots \\ 1/n & 1/(n+1) & 1/(n+2) & \ldots 1/(2n-1) \end{bmatrix}_{n \times n} \tag{9.3}$$

равен

$$\frac{[1!\,2!\,3!\ldots(n-1)!]^3}{n!\,(n+1)!\,(n+2)!\,\ldots(2n-1)!}.$$

Доказательство. Вычислим сначала более общий определитель

Коши:

$$\det\left[\frac{1}{a_j+b_k}\right]_{j,k=1}^{n} = \begin{vmatrix} \frac{1}{a_1+b_1} & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_n} \\ \frac{1}{a_2+b_1} & \frac{1}{a_2+b_2} & \cdots & \frac{1}{a_2+b_n} \\ & \cdots & & \cdots \\ \frac{1}{a_n+b_1} & \frac{1}{a_n+b_2} & \cdots & \frac{1}{a_n+b_n} \end{vmatrix}_{n\times n},$$

из которого определитель матрицы Гильберта получится при $a_j = = j-1, b_j = j$. Вычтем из второго, третьего и т.д., n-го столбца первый:

$$\begin{vmatrix} \frac{1}{a_1+b_1} & \frac{b_1-b_2}{(a_1+b_2)(a_1+b_1)} & \cdots & \frac{b_1-b_n}{(a_1+b_n)(a_1+b_1)} \\ \frac{1}{a_2+b_1} & \frac{b_1-b_2}{(a_2+b_2)(a_2+b_1)} & \cdots & \frac{b_1-b_n}{(a_2+b_n)(a_2+b_1)} \\ \cdots & & & \cdots \\ \frac{1}{a_n+b_1} & \frac{b_1-b_2}{(a_n+b_2)(a_n+b_1)} & \cdots & \frac{b_1-b_n}{(a_n+b_n)(a_n+b_1)} \end{vmatrix}_{n\times n} =$$

и вынесем общие множители из числителей и знаменателей строк и столбцов:

$$= \frac{(b_1-b_2)\times\ldots\times(b_1-b_n)}{(a_1+b_1)(a_2+b_1)\times\ldots\times(a_n+b_1)} \begin{vmatrix} 1 & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_n} \\ 1 & \frac{1}{a_2+b_2} & \cdots & \frac{1}{a_2+b_n} \\ \cdots & & & \cdots \\ 1 & \frac{1}{a_n+b_2} & \cdots & \frac{1}{a_n+b_n} \end{vmatrix}_{n\times n}.$$

Вычтем первую строку полученного определителя из второй, третьей и т.д., n-й:

$$= \frac{(b_1-b_2)\ldots(b_1-b_n)}{(a_1+b_1)(a_2+b_1)\ldots(a_n+b_1)} \begin{vmatrix} 1 & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_n} \\ 0 & \frac{a_1-a_2}{(a_2+b_2)(a_1+b_2)} & \cdots & \frac{a_1-a_2}{(a_2+b_n)(a_1+b_n)} \\ \cdots & & & \cdots \\ 0 & \frac{a_1-a_n}{(a_n+b_2)(a_1+b_2)} & \cdots & \frac{a_1-a_n}{(a_n+b_n)(a_1+b_n)} \end{vmatrix}_{n\times n},$$

Перейти к оглавлению на странице: 460

разложим по первому столбцу и вынесем общие множители из числителей и знаменателей строк и столбцов:

$$=\frac{(-1)^{2(n-1)}(b_2-b_1)\dots(b_n-b_1)(a_2-a_1)\dots(a_n-a_1)}{(a_1+b_1)(a_2+b_1)\dots(a_n+b_1)(a_1+b_2)\dots(a_1+b_n)}\times$$

$$\times\begin{vmatrix} \frac{1}{a_2+b_2} & \dots & \frac{1}{a_2+b_n} \\ \dots & & \dots \\ \frac{1}{a_n+b_2} & \dots & \frac{1}{a_n+b_n} \end{vmatrix}_{(n-1)\times(n-1)}.$$

В результате получили определитель той же структуры, что и исходный, но на единицу меньшего порядка. Продолжая процесс по аналогии, получим окончательно:

$$\det\left[\frac{1}{a_j+b_k}\right]_{j,k=1}^{n}=\frac{\prod\limits_{1\le j<k\le n}[(a_j-a_k)(b_j-b_k)]}{\prod\limits_{j,k=1}^{n}(a_j+b_k)}.$$

■

Пример 9.2.

$$\det\mathfrak{H}_3=\frac{1}{2160},\ \det\mathfrak{H}_4=\frac{1}{6\,048\,000},\ \det\mathfrak{H}_6=\frac{1}{186313420339200000}.$$

Замечание 9.1. Матрицы Вандермонда и Гильберта возникают в задачах аппроксимации (приближения) функций.

Упражнение 24. *Построить полином $p(x)$ второй степени,*

а) *принимающий значение $\sqrt{x}$ в точках $x=\{1/9,\ 1/4, 1\}$;*

б) *такой, чтобы величина интеграла*

$$\int_0^1\left[p(x)-\sqrt{x}\right]^2 d\,x$$

была минимальной.

Ленточный определитель

Определитель Якоби:

$$\mathfrak{J}_n \stackrel{\text{def}}{=} \begin{vmatrix} a_1 & b_1 & 0 & 0 \dots & 0 & 0 \\ -c_2 & a_2 & b_2 & 0 \dots & 0 & 0 \\ 0 & -c_3 & a_3 & b_3 \dots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 \dots & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & 0 \dots & -c_n & a_n \end{vmatrix}_{n\times n} \tag{9.4}$$

после разложения по общей формуле (3.14) разложения определителя будет представлять из себя полином по $a_1, \dots, a_n, b_1, \dots, b_{n-1}, c_2, \dots, c_n$, линейный по каждой переменной. Если разложить $\mathfrak{J}_n$ по последней строке, то получим:

$$\begin{aligned} \mathfrak{J}_n &= a_n\mathfrak{J}_{n-1} + b_{n-1}c_n\mathfrak{J}_{n-2} = \\ &= a_n(a_{n-1}\mathfrak{J}_{n-2} + b_{n-2}c_{n-1}\mathfrak{J}_{n-3}) + b_{n-1}c_n\mathfrak{J}_{n-2} = \\ &= (a_na_{n-1} + b_{n-1}c_n)\mathfrak{J}_{n-2} + a_nb_{n-2}c_{n-1}\mathfrak{J}_{n-3} = \dots \end{aligned} \tag{9.5}$$

Пример 9.3. $\mathfrak{J}_2 = a_1a_2 + b_1c_2$, $\mathfrak{J}_3 = a_1a_2a_3 + a_1b_2c_3 + b_1c_2a_3$,

$$\begin{aligned} \mathfrak{J}_5 &= a_1a_2a_3a_4a_5 + b_1c_2a_3a_4a_5 + a_1b_2c_3a_4a_5 + a_1a_2b_3c_4a_5 + \\ &+ a_1a_2a_3b_4c_5 + b_1c_2b_3c_4a_5 + b_1c_2a_3b_4c_5 + a_1b_2c_3b_4c_5. \end{aligned}$$

Теорема 9.3. *Значение $\mathfrak{J}_n$ равно сумме главного члена $a_1a_2\times\dots\times a_n$ и всевозможных произведений, получающихся из него заменой одной или нескольких пар соседних множителей a_ja_{j+1} на b_jc_{j+1}.*

Частный случай определителя Якоби — **континуант**:

$$\mathsf{K}_n(a_1, a_2, \dots, a_n) \stackrel{\text{def}}{=} \begin{vmatrix} a_1 & 1 & 0 & 0 \dots & 0 & 0 \\ -1 & a_2 & 1 & 0 \dots & 0 & 0 \\ 0 & -1 & a_3 & 1 \dots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 \dots & a_{n-1} & 1 \\ 0 & 0 & 0 & 0 \dots & -1 & a_n \end{vmatrix}_{n\times n}$$

Его величина совпадает с континуантой[16].

Исследуем еще один частный случай определителя Якоби (9.4) — при одинаковых элементах на диагоналях

$$a_1 = \dots = a_n \stackrel{\text{def}}{=} a,\ b_1 = \dots = b_{n-1} \stackrel{\text{def}}{=} b,\ c_2 = \dots = c_n \stackrel{\text{def}}{=} c;$$

таким образом:

$$\mathfrak{I}_n = \left| \begin{array}{l} a\,b\,0\,0\dots 0\,0 \\ c\,a\,b\,0\dots 0\,0 \\ 0\,c\,a\,b\dots 0\,0 \\ \vdots \qquad \ddots \quad \vdots \\ 0\,0\,0\,0\dots a\,b \\ 0\,0\,0\,0\dots c\,a \end{array} \right|_{n\times n} .$$

В этом случае уравнение (9.5) перейдет в $\mathfrak{I}_n = a\mathfrak{I}_{n-1} - bc\mathfrak{I}_{n-2}$.

Определение 9.2. Уравнение

$$x_n = p\,x_{n-1} + q\,x_{n-2}, \quad \{p, q, x_n\} \subset \mathbb{A} \tag{9.6}$$

называется **линейным разностным** (или **возвратным**) уравнением (второго порядка). Пусть числа x_0, x_1 заданы. Тогда уравнение (9.6) определяет **линейную рекуррентную**[17] (или **возвратную**) последовательность (второго порядка): каждый элемент этой последовательности определяется через два предшествующих.

Пример 9.4. Уравнение второго порядка

$$x_n = x_{n-1} + x_{n-2}$$

определяет при $x_0 = x_1 = 1$ последовательность чисел **Фибоначчи**

$$\{x_n\}_{n=0}^{\infty} = \{1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34, \dots\}.$$

Определение 9.3. **Решить уравнение** (9.6) означает найти выражение для x_n в виде явной функции от номера n и “начальных условий” x_0, x_1.

[16] Глава 1, §2
[17] recurrere (*лат.*) — возвращаться.

Перейти к оглавлению на странице: 460

Алгоритм решения разностного уравнения второго порядка

$$x_n = p\,x_{n-1} + q\,x_{n-2}$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

$$x^2 = p\cdot x \quad + q\cdot 1$$

$$\Longleftrightarrow \quad x^2 - p\,x - q = 0$$

$$\Downarrow$$

$$\boxed{\mathcal{D} = p^2 + 4\,q = 0?}$$

$$\swarrow \qquad\qquad \searrow$$

$$\boxed{\lambda_1 \neq \lambda_2} \qquad\qquad \boxed{\lambda_1 = \lambda_2 \stackrel{def}{=} \lambda}$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$x_n = C_1\lambda_1^n + C_2\lambda_2^n \qquad\qquad x_n = (C_1 + C_2 n)\lambda^n$$

Упражнение 25. *Показать, что найденные формулы дают решения уравнения* (9.6)*.*

Неопределенные коэффициенты C_1 и C_2 ищутся из “начальных условий”: формулы должны оставаться справедливыми при $n=0$ и $n=1$.

$$\begin{cases} x_0 = C_1 + C_2, \\ x_1 = C_1\lambda_1 + C_2\lambda_2 \end{cases} \qquad \text{или} \qquad \begin{cases} x_0 = C_1, \\ x_1 = (C_1 + C_2)\lambda \end{cases}$$

Упражнение 26. *Найти выражение для n-го числа Фибоначчи (формулу Бине).*

Пример 9.5. Вычислить

$$)\ \begin{vmatrix} 2 & 2 & 0 & \dots & 0 & 0 \\ 1 & 2 & 2 & \dots & 0 & 0 \\ 0 & 1 & 2 & \dots & 0 & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & 2 & 2 \\ 0 & 0 & 0 & \dots & 1 & 2 \end{vmatrix}_{n\times n}; \quad)\ \begin{vmatrix} 2\boldsymbol{i} & -1 & 0 & \dots & 0 & 0 \\ 1 & 2\boldsymbol{i} & -1 & \dots & 0 & 0 \\ 0 & 1 & 2\boldsymbol{i} & \dots & 0 & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & 2\boldsymbol{i} & -1 \\ 0 & 0 & 0 & \dots & 1 & 2\boldsymbol{i} \end{vmatrix}_{n\times n}.$$

Решение. Для случая **а)** разностное уравнение имеет вид $\mathfrak{J}_n = 2\mathfrak{J}_{n-1} - 2\mathfrak{J}_{n-2}$. Корни соответствующего квадратного уравнения $\lambda_{1,2} = 1 \pm \boldsymbol{i}$ различны. Согласно приведенному выше алгоритму, решение разностного уравнения имеет вид $C_1(1+\boldsymbol{i})^n + C_2(1-\boldsymbol{i})^n$. Для определения констант C_1 и C_2 вычислим определители первого и второго порядков: $\mathfrak{J}_1 = 2, \mathfrak{J}_2 = 2$.

$$\begin{cases} 2 = C_1(1+\boldsymbol{i}) + C_2(1+\boldsymbol{i}), \\ 2 = C_1(1+\boldsymbol{i})^2 + C_2(1+\boldsymbol{i})^2 \end{cases} \Rightarrow \quad C_1 = \frac{1-\boldsymbol{i}}{2}, \; C_2 = \frac{1+\boldsymbol{i}}{2}$$

Окончательный ответ:

$$\mathfrak{J}_n = \frac{1-\boldsymbol{i}}{2}(1+\boldsymbol{i})^n + \frac{1+\boldsymbol{i}}{2}(1-\boldsymbol{i})^n = (1+\boldsymbol{i})^{n-1} + (1-\boldsymbol{i})^{n-1}. \quad (9.7)$$

Для случая **б)**: $\mathfrak{J}_n = 2\,\boldsymbol{i}\,\mathfrak{J}_{n-1} + \mathfrak{J}_{n-2}$. Корни соответствующего квадратного уравнения $\lambda_{1,2} = \boldsymbol{i}$ одинаковы. Выражение для $\mathfrak{J}_n$ следует искать в виде $\mathfrak{J}_n = (C_1 + C_2 n)\,\boldsymbol{i}^{\,n}$.

$$\mathfrak{J}_1 = 2\boldsymbol{i}\,, \; \mathfrak{J}_2 = -3 \; \Rightarrow \begin{cases} 2\boldsymbol{i} = \boldsymbol{i}\,C_1 + \boldsymbol{i}\,C_2, \\ -3 = -C_1 - 2\,C_2 \end{cases} \Rightarrow \quad C_1 = 1, \; C_2 = 1.$$

Окончательно:

$$\mathfrak{J}_n = (1+n)\,\boldsymbol{i}^{\,n}.$$

Отметим одно интересное обстоятельство. Определитель **а)**, элементами которого являются вещественные числа, очевидно и сам должен быть вещественным — в соответствии с определением (3.10). Более того, он должен быть целым. Между тем в его представлении (9.7) участвует мнимая единица. Как разрешить противоречие? Подстановка в формулу тестовых значений n и использование формулы бинома Ньютона позволяет заметить, что мнимая часть получившегося выражения каждый раз обращается в нуль. Для обоснования этой гипотезы, удобно перейти к тригонометрической форме записи комплексных чисел. Имеем:

$$1+\boldsymbol{i} = \sqrt{2}\left(\cos\frac{\pi}{4} + \boldsymbol{i}\,\sin\frac{\pi}{4}\right), \quad 1-\boldsymbol{i} = \sqrt{2}\left(\cos\left(-\frac{\pi}{4}\right) + \boldsymbol{i}\,\sin\left(-\frac{\pi}{4}\right)\right).$$

Для возведения в степень применяем формулу Муавра[18]:

$$
\begin{aligned}
\mathfrak{I}_n = & 2^{(n-1)/2}\left(\cos\frac{(n-1)\pi}{4} + \boldsymbol{i}\sin\frac{(n-1)\pi}{4}\right) + \\
& + 2^{(n-1)/2}\left(\cos\left(-\frac{(n-1)\pi}{4}\right) + \boldsymbol{i}\sin\left(-\frac{(n-1)\pi}{4}\right)\right) = \\
& = 2^{(n+1)/2}\cos\frac{\pi(n-1)}{4},
\end{aligned}
$$

т.е. $\mathfrak{I}_n$ оказывается действительно действительным. Тот факт, что на самом деле $\mathfrak{I}_n$ является числом целым подтверждается тем, что последовательность

$$
\{\cos\pi(n-1)/4\}_{n=1}^{\infty} = \left\{1, 2^{-1/2}, 0, -2^{-1/2}, -1, -2^{-1/2}, 0, 2^{-1/2}, 1, \dots\right\}
$$

является циклической и выражения $2^{-1/2}$ возникают только при четных n. При домножении на $2^{(n+1)/2}$ дробные степени двойки пропадают. Резюмируем приведенные рассуждения: присутствие в формуле (9.7) мнимой единицы, образно говоря, само является *мнимым*, иллюзорным; в вычислениях она участвует, но в ответе пропадает!

Ответ. **а)** $2^{(n+1)/2}\cos\dfrac{\pi(n-1)}{4}$; **б)** $(1+n)\boldsymbol{i}^{\,n}$.

Принцип решения разностных уравнений более высоких порядков — такой же, как и уравнения второго порядка[19]. Это позволяет вычислять ленточные определители, у которых заполнены больше трех диагоналей.

⊖ **Упражнение 27.** *Вычислить определитель*

$$
\begin{vmatrix}
6 & 11 & 6 & 0 & \dots & 0 & 0 \\
1 & 6 & 11 & 6 & \dots & 0 & 0 \\
0 & 1 & 6 & 11 & \dots & 0 & 0 \\
\vdots & & & \ddots & \ddots & & \vdots \\
0 & 0 & 0 & 0 & \dots & 6 & 11 \\
0 & 0 & 0 & 0 & \dots & 1 & 6
\end{vmatrix}_{n\times n}.
$$

[18] Глава 2, формула (2.8).

[19] Подробное изложение теории разностных уравнений будет произведено в 3-м семестре.

Историческая справка

Фибоначчи (filius Bonacij , в современном написании Fibonacci) или Леонардо Пизанский (?1180–?1250)

Итальянский математик. В своей книге “Об абаке” (“Liber Abaci”, 1202) рассмотрел следующую задачу о числе кроликов. Требуется определить число пар зрелых кроликов, образовавшихся от одной пары в течение года, если известно, что каждая зрелая пара кроликов ежемесячно рождает новую пару, причем новорожденные достигают полной зрелости в течение месяца. Итак в начальный момент имеется $x_0 = 1$ пара, через месяц прибавится еще одна пара, но число зрелых пар останется прежнее $x_1 = 1$. Через два месяца крольчата достигнут зрелости и общее число зрелых пар будет равно $x_2 = 2$. Пусть мы вычислили уже количества x_{n-2} зрелых пар через $n-2$ месяцев и x_{n-1} — через $n-1$ месяцев. Тогда через n месяцев общее число зрелых пар будет $x_{n-2} + x_{n-1}$.

Характеристический полином

Определение 9.4. Характеристическим полиномом матрицы A называется

$$\det(A - \lambda E).$$

Пример 9.6. а) Для $n = 2$:

$$\det(A-\lambda E) = \begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} = \lambda^2 - (a_{11}+a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21})$$

б) Для $n = 3$:

$$\det(A - \lambda E) = \begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix} =$$

$$= -\lambda^3 + (a_{11}+a_{22}+a_{33})\lambda^2 - \left\{ \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} \right\} \lambda + \det A.$$

⊖ Характеристические полиномы матриц возникают в задачах о классификации типов кривых или поверхностей второго порядка. Общее уравнение кривой второго порядка можно представить в виде:

$$a_{11}x^2 + 2\,a_{12}xy + a_{22}y^2 + 2a_1x + 2\,a_2y + a = 0. \qquad (9.8)$$

Уравнение (9.8) всегда является уравнением одной из кривых: эллипса, гиперболы или параболы (не считая случаев вырождения — пары прямых, точки или пустого множества). Поворотом системы координат на некоторый угол уравнение (9.8) всегда можно преобразовать к виду:

$$\lambda_1 X^2 + \lambda_2 Y^2 + 2b_1 X + 2b_2 Y + b = 0. \tag{9.9}$$

При этом коэффициенты λ_1, λ_2 будут корнями уравнения

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{12} & a_{22} - \lambda \end{vmatrix} = 0.$$

Легко доказать (**сделайте это!**), что при любых наборах a_{jk} эти корни будут вещественными. Преобразуем уравнение (9.9) следующим образом:

$$\lambda_1 \left(X + \frac{b_1}{\lambda_1} \right)^2 + \lambda_2 \left(Y + \frac{b_2}{\lambda_2} \right)^2 + c = 0, \quad \text{при} \quad c = b - \frac{b_1^2}{\lambda_1} - \frac{b_2^2}{\lambda_2}.$$

Если $\lambda_1 \lambda_2 > 0$, то при $\operatorname{sign} c = -\operatorname{sign} \lambda_1$, то уравнение (9.8) представляет эллипс. При $c = 0$ получим одну точку, а при $\operatorname{sign} c = \operatorname{sign} \lambda_1$ — пустое множество.

⊖ **Упражнение 28.** *Исследуйте случай* $\lambda_1 \lambda_2 < 0$.

Аналогично исследуются и поверхности второго порядка.

Теорема 9.4. $\det(A - \lambda E) = (-1)^n (\lambda^n + a_1 \lambda^{n-1} + \dots + a_n)$, *где*

$$a_1 = -\sum_{j=1}^{n} a_{jj} = -(a_{11} + a_{22} + \dots + a_{nn}),$$

$$a_2 = \sum_{1 \le j_1 < j_2 \le n} \begin{vmatrix} a_{j_1 j_1} & a_{j_1 j_2} \\ a_{j_2 j_1} & a_{j_2 j_2} \end{vmatrix},$$

$$a_3 = -\sum_{1 \le j_1 < j_2 < j_3 \le n} \begin{vmatrix} a_{j_1 j_1} & a_{j_1 j_2} & a_{j_1 j_3} \\ a_{j_2 j_1} & a_{j_2 j_2} & a_{j_2 j_3} \\ a_{j_3 j_1} & a_{j_3 j_2} & a_{j_3 j_3} \end{vmatrix},$$

$$\dots$$

$$a_n = (-1)^n \det A$$

Упражнение 29. *Показать, что характеристический полином матрицы* **Фробениуса**

$$\mathfrak{F} \stackrel{\text{def}}{=} \begin{pmatrix} 0 & 1 & 0 & 0 \dots 0 & 0 \\ 0 & 0 & 1 & 0 \dots 0 & 0 \\ 0 & 0 & 0 & 1 \dots 0 & 0 \\ \dots & & & & \dots \\ 0 & 0 & 0 & 0 \dots 0 & 1 \\ a_n & a_{n-1} & a_{n-2} & \dots a_2 & a_1 \end{pmatrix}_{n\times n}$$

равен $(-1)^n(\lambda^n - a_1\lambda^{n-1} - \dots - a_n)$.

⊖ **Упражнение 30.** *Показать, что характеристический полином кососимметричной матрицы четвертого порядка не имеет вещественных корней.*

Подробное исследование характеристического полинома матрицы и его различных приложений будет произведено в 3-м семестре.

Целочисленный определитель

Пример 9.7. Верно ли равенство

$$\begin{vmatrix} 51239 & 79922 & 55538 & 29177 \\ 46152 & 16596 & 37189 & 82561 \\ 71489 & 23165 & 26563 & 61372 \\ 44350 & 42391 & 91185 & 64809 \end{vmatrix} = 0?$$

Решение. Фактическое вычисление подобного определителя — каким бы методом мы не воспользовались — задача довольно трудоемкая. Однако вопрос ставится не о фактическом значении, а о равенстве его нулю. Это обстоятельство может упростить вычисления. Обозначим неизвестное значение определителя через x; очевидно это число целое. Если $x = 0$, то и его остаток при делении на любое число $M \in \mathbb{Z}$ тоже должен быть равен нулю. Если же хоть для одного $M \in \mathbb{Z}$ выполнится условие $x \not\equiv 0 \pmod{M}$, то и $x \neq 0$. Вычисление определителя фактически (см. формулу (3.14)) сводится к умножению элементов определителя. Если же мы ставим задачу определения остатка от деления этого выражения на M, то[20] имеет смысл сразу же "сократить" каждый элемент определителя до его остатка от деления на M.

Возьмем сначала $M=10$; от каждого элемента определителя оставляем только последнюю цифру:

$$\left|\begin{array}{cccc} 9&2&8&7\\ 2&6&9&1\\ 9&5&3&2\\ 0&1&5&9 \end{array}\right| \equiv_{10} \left|\begin{array}{rrrr} -1&2&-2&-3\\ 2&-4&-1&1\\ -1&5&3&2\\ 0&1&5&-1 \end{array}\right| = \left|\begin{array}{rrrr} -1&2&-2&-3\\ 0&0&-5&-5\\ 0&3&5&-5\\ 0&1&5&-1 \end{array}\right| =$$

$$= - \left|\begin{array}{rrr} 0&-5&-5\\ 3&5&-5\\ 1&5&-1 \end{array}\right| \equiv_{10} 0.$$

Итак, полученный ответ является необходимым, но не достаточным условием равенства определителя нулю. Сделаем еще одну проверку: возьмем $M=7$.

$$\left|\begin{array}{cccc} 6&3&0&1\\ 1&6&5&3\\ 5&2&5&3\\ 5&6&3&3 \end{array}\right| \equiv_7 3 \ne 0.$$

Ответ. Равенство неверно.

Понятно, что если бы определитель был равен нулю, то каждое вычисление по модулю только "увеличивало бы достоверность" этого события.

Можно ли на основе серии модулярных вычислений установить истинное значение определителя?

Попробуем это сделать для определителя примера 9.7. Прежде всего, оценим сверху абсолютную величину определителя. Каждое слагаемое в разложении определителя по формуле (3.14) представляет собой произведение четырех пятизначных чисел, следовательно, оно не превосходит 10^{24}. Всего таких слагаемых 24, из них половина — с отрицательным знаком. Поэтому $|x| < 12 \cdot 10^{24}$. Берем

[20] Глава 1, §5.1

теперь все простые числа p, так чтобы их произведение превысило эту оценку:

$$L \stackrel{\text{def}}{=} \underbrace{2 \cdot 3 \cdot 5 \times \cdots \times 67 \cdot 71}_{20} > 12 \cdot 10^{24}$$

и вычисляем остатки $B_p \stackrel{\text{def}}{=} x \pmod{p}$:

$$B_2 = 0, B_3 = 0, B_5 = 0, B_7 = 3, B_{11} = 7, \ldots, B_{67} = 64, B_{71} = 39.$$

Китайская теорема об остатках[21] позволяет однозначно определить величину x если она находится в пределах $0 < x < L$:

$$x = +557940821520864633874788300.$$

Однако, поскольку мы не знаем знака определителя, то должны предложить еще один вариант ответа — вычитая из полученного положительного значения величину L:

$$x = -8605834327092627090.$$

Из двух получившихся вариантов только один — именно последний — удовлетворяет оценке $|x| < 12 \cdot 10^{24}$. Это и есть величина нашего определителя.

§10. Обратная матрица

Определение 10.1. Для квадратной матрицы A ее **правой обратной** называется матрица X такая, что $AX = E$. В случае существования такой матрицы, она обозначается A^{-1}.

Теорема 10.1. *Для матрицы A правая обратная матрица A^{-1} существует тогда и только тогда, когда* $\det A \neq 0$.

Доказательство необходимости. Если A^{-1} существует, то $AA^{-1} = E$, следовательно, на основании формулы (8.1), $1 = \det A \det A^{-1}$. Это равенство гарантирует, что $\det A \neq 0$. ■

[21] Глава 1, теорема 6.3.

Доказательство достаточности. Пусть $\det A \neq 0$. Вычислим алгебраические дополнения ко всем элементам матрицы A и составим из них матрицу

$$\widetilde{A} \stackrel{\text{def}}{=} \begin{pmatrix} A_{11}\, A_{12} \dots A_{1n} \\ A_{21}\, A_{22} \dots A_{2n} \\ \dots \qquad\quad \dots \\ A_{n1}\, A_{n2} \dots A_{nn} \end{pmatrix}^{\top}, \tag{10.1}$$

которая называется матрицей **союзной** матрице A. Докажем, что

$$A\widetilde{A} = E \det A.$$

Действительно,

$$\begin{pmatrix} a_{11}\, a_{12} \dots a_{1n} \\ a_{21}\, a_{22} \dots a_{2n} \\ \dots \qquad\quad \dots \\ a_{n1}\, a_{n2} \dots a_{nn} \end{pmatrix} \begin{pmatrix} A_{11}\, A_{21} \dots A_{n1} \\ A_{12}\, A_{22} \dots A_{n2} \\ \dots \qquad\quad \dots \\ A_{1n}\, A_{2n} \dots A_{nn} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11}A_{11}+a_{12}A_{12}+\dots+a_{1n}A_{1n} & \dots & a_{11}A_{n1}+a_{12}A_{n2}+\dots+a_{1n}A_{nn} \\ a_{21}A_{11}+a_{22}A_{12}+\dots+a_{2n}A_{1n} & \dots & a_{21}A_{n1}+a_{22}A_{n2}+\dots+a_{2n}A_{nn} \\ \dots & & \dots \\ a_{n1}A_{11}+a_{n2}A_{12}+\dots+a_{nn}A_{1n} & \dots & a_{n1}A_{n1}+a_{n2}A_{n2}+\dots+a_{nn}A_{nn} \end{pmatrix}.$$

Воспользуемся теперь результатом теоремы 5.2. Согласно второй из формул (5.6), только диагональные элементы последней матрицы могут быть отличными от нуля, и они равны как раз $\det A$.

Поскольку, по предположению, $\det A \neq 0$, то матрица

$$\frac{1}{\det A}\widetilde{A} = \begin{pmatrix} \dfrac{A_{11}}{\det A} & \dfrac{A_{21}}{\det A} & \dots & \dfrac{A_{n1}}{\det A} \\ \dfrac{A_{12}}{\det A} & \dfrac{A_{22}}{\det A} & \dots & \dfrac{A_{n2}}{\det A} \\ \dots & & & \dots \\ \dfrac{A_{1n}}{\det A} & \dfrac{A_{2n}}{\det A} & \dots & \dfrac{A_{nn}}{\det A} \end{pmatrix} \tag{10.2}$$

как раз и будет правой обратной матрицей для матрицы A. ■

Теорема 10.2. *Если существует правая обратная матрица для матрицы A, то* **а)** *она единственна и* **б)** *она является и левой обратной.*

Доказательство. Утверждение **б)** проверяется непосредственной проверкой (с применением теоремы 5.2). Если же предположить наличие еще одной левой обратной матрицы Y:

$$YA = E \ \Rightarrow \ (YA)A^{-1} = A^{-1} \ \Rightarrow \ Y(AA^{-1}) = A^{-1} \ \Rightarrow \ Y = A^{-1},$$

т.е. левая обратная должна совпадать с A^{-1}.

а) Если $AX = E$, то с использованием только что доказанного имеем:

$$A^{-1}(AX) = A^{-1} \ \Rightarrow \ (A^{-1}A)X = A^{-1} \ \Rightarrow \ X = A^{-1}.$$

■

Пример 10.1. Вычислить

$$\begin{pmatrix} 4 & 71 & 5 \\ 3 & 40 & -6 \\ -11 & 82 & 9 \\ -12 & -100 & 8 \end{pmatrix}^{-1}.$$

Решение. Определитель этой матрицы был вычислен в примере 5.1: -226. Обратная матрица существует. Вычисляя алгебраические дополнения элементов, строим союзную матрицу:

$$A_{11} = \begin{vmatrix} 40 & -6 \\ 82 & 9 \\ -100 & 8 \end{vmatrix} = -56, \ A_{12} = -\begin{vmatrix} 30 & -6 \\ -112 & 9 \\ -120 & 8 \end{vmatrix} = 96,$$

$$A_{13} = \begin{vmatrix} 3 & 4 & -6 \\ -11 & 8 & 9 \\ -12 & -10 & 8 \end{vmatrix} = -854, \ A_{14} = -\begin{vmatrix} 3 & 40 \\ -11 & 82 \\ -12 & -100 \end{vmatrix} = 36,$$

$$A_{21} = -58, \ A_{22} = 164, \ A_{23} = -1506, \ A_{24} = 118, \dots, A_{44} = 58.$$

Ответ.

$$\begin{pmatrix} {}^{28}/_{113} & {}^{29}/_{113} & -{}^{14}/_{113} & {}^{20}/_{113} \\ -{}^{48}/_{113} & -{}^{82}/_{113} & {}^{24}/_{113} & -{}^{117}/_{226} \\ {}^{427}/_{113} & {}^{753}/_{113} & -{}^{157}/_{113} & {}^{949}/_{226} \\ -{}^{18}/_{113} & -{}^{59}/_{113} & {}^{9}/_{113} & -{}^{29}/_{113} \end{pmatrix}$$

Определение 10.2. Матрица A с ненулевым определителем называется **неособенной** или **невырожденной** или **обратимой**. Операция нахождения обратной матрицы для матрицы A называется **обращением** матрицы.

Упражнение 31. *Доказать следующие свойства операции обращения:*

$$\textbf{а)}\ \left(A^{-1}\right)^{-1}=A;\ \textbf{б)}(AB)^{-1}=B^{-1}A^{-1};\ \textbf{в)}(A^{\top})^{-1}=(A^{-1})^{\top};$$
$$\textbf{г)}\ \det A^{-1}=(\det A)^{-1}.$$

при условии, что все действия определены.

Еще один способ вычисления обратной матрицы заключается в модернизации метода решения матричного уравнения $AX=E$, который мы применяли в §2. Условно он называется "метод приписывания единичной матрицы" и заключается в следующем:

1) формируем расширенную $n\times 2n$-матрицу $(A\mid E)$, приписывая к матрице A справа единичную матрицу E того же порядка;

2) элементарными преобразованиями над строками расширенной матрицы, добиваемся, чтобы в левой ее половине получилась единичная матрица;

3) если это удается сделать, то матрица, получившаяся в правой половине и будет A^{-1}.

Пример 10.2. Вычислить

$$\left(\begin{array}{rrr} 4 & 5 & 1\\ 1 & 3 & -2\\ 3 & 1 & 2\end{array}\right)^{-1}$$

приписыванием единичной матрицы.

$$\left(\begin{array}{rrr|rrr} 4 & 5 & 1 & 1 & 0 & 0\\ 1 & 3 & -2 & 0 & 1 & 0\\ 3 & 1 & 2 & 0 & 0 & 1\end{array}\right)\to\left(\begin{array}{rrr|rrr} 1 & 3 & -2 & 0 & 1 & 0\\ 4 & 5 & 1 & 1 & 0 & 0\\ 3 & 1 & 2 & 0 & 0 & 1\end{array}\right)\to$$

$$\to \left(\begin{array}{rrr|rrr} 1 & 3 & -2 & 0 & 1 & 0 \\ 0 & -7 & 9 & 1 & -4 & 0 \\ 0 & -8 & 8 & 0 & -3 & 1 \end{array}\right) \to \left(\begin{array}{rrr|rrr} 1 & 3 & -2 & 0 & 1 & 0 \\ 0 & -8 & 8 & 0 & -3 & 1 \\ 0 & -7 & 9 & 1 & -4 & 0 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|rrr} 1 & 3 & -2 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & -3/8 & 1/8 \\ 0 & -7 & 9 & 1 & -4 & 0 \end{array}\right) \to \left(\begin{array}{rrr|rrr} 1 & 3 & -2 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & -3/8 & 1/8 \\ 0 & 0 & 2 & 1 & -11/8 & -7/8 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|rrr} 1 & 3 & -2 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & -3/8 & 1/8 \\ 0 & 0 & 1 & 1/2 & -11/16 & -7/16 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|rrr} 1 & 3 & -2 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1/2 & 5/16 & 9/16 \\ 0 & 0 & 1 & 1/2 & -11/16 & -7/16 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & -1/2 & 9/16 & 13/16 \\ 0 & 1 & 0 & 1/2 & -5/16 & -9/16 \\ 0 & 0 & 1 & 1/2 & -11/16 & -7/16 \end{array}\right).$$

Ответ.

$$\begin{pmatrix} -1/2 & 9/16 & 13/16 \\ 1/2 & -5/16 & -9/16 \\ 1/2 & -11/16 & -7/16 \end{pmatrix}$$

Упражнение 32. *Найти неизвестную матрицу X из уравнения*

$$(X - E)^{-1}(X + E) = B \ , \quad \textit{здесь } E_{n\times n}, B_{n\times n}.$$

Упражнение 33. *Доказать, что при условии* $a_{11} \neq 0, a_{22} \neq 0, \ldots,$ $a_{nn} \neq 0$

$$\begin{pmatrix} a_{11} & & & \\ & a_{22} & \mathbb{O} & \\ & \mathbb{O} & \ddots & \\ & & & a_{nn} \end{pmatrix}^{-1} = \begin{pmatrix} 1/a_{11} & & & \\ & 1/a_{22} & \mathbb{O} & \\ & \mathbb{O} & \ddots & \\ & & & 1/a_{nn} \end{pmatrix}$$

Теорема 10.3. *[Обращение блочной матрицы] Пусть*

$$\mathfrak{B} = \begin{pmatrix} A & \mathbb{O} \\ C & D \end{pmatrix} \quad ,$$

где $k \times k$-матрица A и $\ell \times \ell$-матрица D — неособенные. Тогда

$$\mathfrak{B}^{-1} = \begin{pmatrix} A^{-1} & \mathbb{O} \\ -D^{-1}CA^{-1} & D^{-1} \end{pmatrix}_{n\times n}.$$

Доказательство. Будем искать $\mathfrak{B}^{-1}$ в виде:

$$\mathfrak{B}^{-1} = \begin{pmatrix} X & Y \\ U & V \end{pmatrix}_{n\times n}$$

при $k \times k$-матрице X и $\ell \times \ell$-матрице V. Разбиваем матричное равенство

$$\begin{pmatrix} X & Y \\ U & V \end{pmatrix}\begin{pmatrix} A & \mathbb{O} \\ C & D \end{pmatrix} = \begin{pmatrix} E_k & \mathbb{O} \\ \mathbb{O} & E_\ell \end{pmatrix}.$$

на четыре отдельных

$$\begin{matrix} XA + YC = E_k, YD = \mathbb{O}, \\ UA + VC = \mathbb{O},\ VD = E_\ell \end{matrix} \quad \Rightarrow \quad \begin{matrix} Y = \mathbb{O}, \\ V = D^{-1}. \end{matrix}$$

Подставляем полученное в два оставшихся равенства: $X = A^{-1}$, $U = -D^{-1}CA^{-1}$. ■

Упражнение 34. *Алгоритм шифрования* **Rijndael**, *используемый в мобильной телефонии, имеет в одной из стадий следующее преобразование байтов*

$$\begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{pmatrix} = \begin{pmatrix} 10001111 \\ 11000111 \\ 11100011 \\ 11110001 \\ 11111000 \\ 01111100 \\ 00111110 \\ 00011111 \end{pmatrix}\begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \pmod 2$$

Найти обратное преобразование.

§11. Ранг

Ранг системы рядов

Определение 11.1. Система (совокупность, набор) n рядов (строк или столбцов)

$$\{A_1, \dots, A_n\} \tag{11.1}$$

называется **линейно зависимой** (**л.з.**) если существуют числа $\alpha_1, \dots, \alpha_n$, хотя бы одно из которых отлично от нуля, такие что

$$\alpha_1 A_1 + \dots + \alpha_n A_n = \mathbb{O}. \tag{11.2}$$

Если же равенство (11.2) возможно только при $\alpha_1 = 0, \dots, \alpha_n = 0$, то система (11.1) называется линейно независимой (**л.н.з.**).

Теорема 11.1. **а)** *Если система* (11.1) *содержит хотя бы один нулевой ряд, то она* **л.з.**

б) *Если система* (11.1) **л.н.з.**, *то и любая ее подсистема* **л.н.з.**

в) *При $n > 1$ система* (11.1) **л.з.** *тогда и только тогда, когда по меньшей мере один ее ряд линейно выражается через остальные, т.е. существуют $j \in \mathbb{N}$ и константы $\gamma_1, \dots, \gamma_{j-1}, \gamma_{j+1}, \dots, \gamma_n$ такие, что*

$$A_j = \gamma_1 A_1 + \dots + \gamma_{j-1} A_{j-1} + \gamma_{j+1} A_{j+1} + \dots + \gamma_n A_n.$$

Пример 11.1. Найти все значения параметра λ, при которых строка $B = (7, -2, \lambda)$ выражается через строки

$$A_1 = (2, 3, 5),\ A_2 = (3, 7, 8),\ A_3 = (1, -6, 1).$$

Решение. Составим уравнение $B = \gamma_1 A_1 + \gamma_2 A_2 + \gamma_3 A_3$ и попытаемся подобрать неопределенные параметры γ_j ему удовлетворяющие.

$$\begin{cases} 2\gamma_1 + 3\gamma_2 + \gamma_3 = 7 \\ 3\gamma_1 + 7\gamma_2 - 6\gamma_3 = -2 \\ 5\gamma_1 + 8\gamma_2 + \gamma_3 = \lambda \end{cases} \iff \begin{cases} \gamma_1 + 4\gamma_2 - 7\gamma_3 = -9 \\ \gamma_2 - 3\gamma_3 = -5 \\ 0 = \lambda - 15 \end{cases}.$$

Ответ. $\lambda = 15$.

Теорема 11.2. *Если каждый из рядов системы* (11.1) *линейно выражается через ряды $B_1, \dots, B_k$ и при этом $k < n$, то система* (11.1) *будет* **л.з.**

Перейти к оглавлению на странице: 460

Доказательство проводится индукцией по k. Если $k = 1$, то $A_j = \gamma_j B_1$ при $j \in \{1, \ldots, n\}$. Если хоть один из A_j — нулевой ряд, то утверждение следует из пункта а) теоремы 11.1. Пусть все $A_j \neq \mathbb{O}$, тогда и $\gamma_1 \neq 0$. Имеем:

$$B_1 = \frac{1}{\gamma_1} A_1 \implies A_n = \gamma_n B_1 = \frac{\gamma_n}{\gamma_1} A_1,$$

т.е. A_n линейно выражается через A_1. Из пункта в) теоремы 11.1 следует справедливость утверждения теоремы 11.2.

Пусть утверждение справедливо для $(k-1)$-го рядов. Покажем справедливость для k рядов. Пусть

$$\begin{cases} A_1 = \gamma_{11} B_1 + \gamma_{12} B_2 + \cdots + \gamma_{1k} B_k, \\ \ldots \qquad\qquad \ldots \\ A_n = \gamma_{n1} B_1 + \gamma_{n2} B_2 + \cdots + \gamma_{nk} B_k. \end{cases} \tag{11.3}$$

Пусть в (11.3) хоть один коэффициент γ_{ij} отличен от нуля. Например $\gamma_{nk} \neq 0$ (если это не так, перенумеруем ряды $B_1, \ldots, B_k$). Тогда

$$B_k = \frac{1}{\gamma_{nk}} \left(A_n - \gamma_{n1} B_1 - \cdots - \gamma_{n,k-1} B_{k-1} \right).$$

Подставляем это выражение в первые $n-1$ из уравнений (11.3):

$$\underbrace{A_j - \frac{\gamma_{jk}}{\gamma_{nk}} A_n}_{\stackrel{\text{def}}{=} A'_j} = \left(\gamma_{j1} - \frac{\gamma_{n1}\gamma_{jk}}{\gamma_{nk}} \right) B_1 + \cdots + \left(\gamma_{j,k-1} - \frac{\gamma_{n,k-1}\gamma_{jk}}{\gamma_{nk}} \right) B_{k-1}$$

при $j \in \{1, \ldots, n-1\}$. Ряды A'_j линейно выражаются через $k-1$ ряд $B_1, \ldots, B_{k-1}$. По индуктивному предположению система $A'_1, \ldots, A'_{n-1}$ **л.з.**, т.е. существуют не все равные нулю константы $\alpha_1, \ldots, \alpha_{n-1}$ такие, что

$$\alpha_1 A'_1 + \cdots + \alpha_{n-1} A'_{n-1} = \mathbb{O} \implies$$

$$\alpha_1 A_1 + \ldots + \alpha_{n-1} A_{n-1} + \alpha_n A_n = \mathbb{O}$$

$$\text{при } \alpha_n \stackrel{\text{def}}{=} -\alpha_1 \frac{\gamma_{1k}}{\gamma_{nk}} - \ldots - \alpha_{n-1} \frac{\gamma_{n-1,k}}{\gamma_{nk}}.$$

Следовательно, система $A_1, \ldots, A_n$ **л.з.** ■

Определение 11.2. **Рангом системы рядов** (11.1) называется число рядов в ее максимальной линейно независимой подсистеме:

$$\mathfrak{r} = \operatorname{rank}\{A_1, \ldots, A_n\}, \qquad \operatorname{rank}\{\mathbb{O}, \ldots, \mathbb{O}\} \stackrel{\text{def}}{=} 0.$$

Теорема 11.3. *Ранг системы рядов* (11.1) *равен* $\mathfrak{r} \geq 1$ *тогда и только тогда, когда в этой системе существует* $\mathfrak{r}$ *линейно независимых рядов, через которые выражается каждый ряд системы.*

Доказательство необходимости. Пусть ранг системы рядов (11.1) равен $\mathfrak{r} \geq 1$. Тогда в этой системе имеются $\mathfrak{r}$ линейно независимых рядов. Без ограничения общности можно предположить, что такими рядами являются $A_1, \ldots, A_\mathfrak{r}$; в противном случае можно перенумеровать ряды системы. Если ряд A_j есть один из рядов $A_1, \ldots, A_\mathfrak{r}$, то A_j линейно выражается через $A_1, \ldots, A_\mathfrak{r}$. Например, если $j = 1$, то $A_1 = 1 \cdot A_1 + 0 \cdot A_2 + \cdots + 0 \cdot A_\mathfrak{r}$. Если $j > \mathfrak{r}$, то $\mathfrak{r} + 1$ рядов $A_1, \ldots, A_\mathfrak{r}, A_j$ будут **л.з.** по предположению:

$$\alpha_1 A_1 + \cdots + \alpha_\mathfrak{r} A_\mathfrak{r} + \alpha_j A_j = \mathbb{O} \tag{11.4}$$

при хотя бы одном из чисел $\alpha_1, \ldots, \alpha_\mathfrak{r}, \alpha_j$ отличном от нуля. Легко видеть, что $\alpha_j \neq 0$; в противном случае имели бы равенство

$$\alpha_1 A_1 + \cdots + \alpha_\mathfrak{r} A_\mathfrak{r} = \mathbb{O},$$

свидетельствующее о линейной зависимости рядов $A_1, \ldots, A_\mathfrak{r}$, что противоречит условию. Следовательно $\alpha_j \neq 0$ и из (11.4) можно выразить A_j через $A_1, \ldots, A_\mathfrak{r}$. ■

Доказательство достаточности. Обратно, пусть в системе (11.1) существует $\mathfrak{r}$ линейно независимых рядов, через которые линейно выражается каждый ряд системы. Пусть для определенности это будут ряды $A_1, \ldots, A_\mathfrak{r}$. Возьмем подсистему из $(\mathfrak{r}+1)$-го произвольных рядов

$$A_{i_1}, A_{i_2}, \ldots, A_{i_{\mathfrak{r}+1}}. \tag{11.5}$$

По предположению, каждый ряд этой подсистемы линейно выражается через $A_1, \ldots, A_\mathfrak{r}$. На основании теоремы 11.2 система (11.5) линейно зависима. Таким образом, в системе (11.1) существует $\mathfrak{r}$ линейно независимых рядов, но всякие $\mathfrak{r} + 1$ рядов уже **л.з.**. Согласно определению $\operatorname{rank}\{A_1, \ldots, A_n\} = \mathfrak{r}$. ■

Только что доказанная теорема позволяет дать другое определение ранга.

Определение 11.3. **Рангом системы рядов** (11.1) называется число $\mathfrak{r}$ таких линейно независимых рядов системы, через которые линейно выражается каждый ряд системы. Любая совокупность $\mathfrak{r}$ линейно независимых рядов системы (11.1) ранга $\mathfrak{r} \neq 0$ называется **базисом** этой системы.

Теорема 11.4. *Любую линейно независимую подсистему*

$$\{A_{i_1}, A_{i_2}, \dots, A_{i_k}\} \tag{11.6}$$

системы рядов (11.1) *можно дополнить до базиса этой системы.*

Доказательство. Если $k = \mathfrak{r}$, то подсистема (11.6) уже является базисом. Пусть $k < \mathfrak{r}$. Выберем произвольный ряд A_j системы (11.1), не совпадающий с (11.6). Хотя бы один из таких рядов должен быть линейно независим с (11.6). В противном случае, по теореме 11.3, система (11.6) была бы базисом для (11.1), что противоречило бы условию $\mathfrak{r} > k$. К системе

$$\{A_{i_1}, A_{i_2}, \dots, A_{i_k}, A_j\}$$

применим те же рассуждения. Продолжая процесс далее, дойдем до линейно независимой подсистемы, содержащей ровно $\mathfrak{r}$ рядов. ■

Пример 11.2. Найти какой-нибудь базис системы строк

$$A_1{=}(5,2,-3,1),\ A_2{=}(4,1,-2,3),\ A_3{=}(1,1,-1,-2),\ A_4{=}(3,4,-1,2)$$

и все строки системы, не входящие в этот базис, выразить через базисные строки.

Решение. Воспользуемся теоремой 11.4. Строка $A_1 \neq \mathbb{O}$, возьмем ее в качестве первой строки искомого базиса и попытаемся дополнить оставшимися строками до базиса. Если $\operatorname{rank} = 1$, то все оставшиеся строки должны линейно выражаться через A_1, в частности $A_2 = \gamma A_1$ при подходящем числе γ. Однако система линейных уравнений

$$4 = 5\gamma,\ 1 = 2\gamma,\ -2 = -3\gamma,\ 3 = \gamma$$

несовместна, так что предположение неверно. Итак, rank > 1 и строки A_1, A_2 **л.н.з.**

Если rank $= 2$, то каждая строка A_3, A_4 должна линейно выражаться через A_1, A_2. Из двух составленных систем уравнений:

$$\left\{\begin{array}{rcr} 5\gamma_1 + 4\gamma_2 &=& 1 \\ 2\gamma_1 + \gamma_2 &=& 1 \\ -3\gamma_1 - 2\gamma_2 &=& -1 \\ \gamma_1 + 3\gamma_2 &=& -2 \end{array}\right. \quad \text{и} \quad \left\{\begin{array}{rcr} 5\gamma_1 + 4\gamma_2 &=& 3 \\ 2\gamma_1 + \gamma_2 &=& 4 \\ -3\gamma_1 - 2\gamma_2 &=& -1 \\ \gamma_1 + 3\gamma_2 &=& 2 \end{array}\right.$$

первая совместна и имеет решение $\gamma_1 = 1, \gamma_2 = -1$; вторая же несовместна. Итак, rank > 2 и строки A_1, A_2, A_4 **л.н.з.**

В системе осталась только одна строка A_3, но поскольку уже установлена ее линейная зависимость с A_1, A_2, то ее включение в уже построенную подсистему не может увеличить ранга.

Ответ. Базис системы составляют, например, строки A_1, A_2, A_4; при этом $A_3 = A_1 - A_2$.

В следующем пункте будет дан еще один алгоритм вычисления ранга системы рядов, а теперь установим, какие действия над системой не изменяют ее ранга.

Теорема 11.5. *Если систему* (11.1) *дополнить рядом, линейно выражающимся через ряды системы, то ранг системы не изменится. Точно так же ранг системы не меняется при удалении ряда, линейно выражающегося через остальные ряды системы.*

Доказательство. **а)** Дополним систему (11.1) рядом $\mathfrak{A}$, линейно выражающимся через $A_1, \dots, A_n$. Если rank$\{A_1, \dots, A_n\} = \mathfrak{r}$, то в системе, согласно теореме 11.3, существует $\mathfrak{r}$ базисных рядов, например $A_1, \dots, A_{\mathfrak{r}}$. Поскольку $\mathfrak{A}$ линейно выражается через $A_1, \dots, A_n$, то он будет выражаться и через $A_1, \dots, A_{\mathfrak{r}}$. Таким образом, в расширенной системе

$$A_1, \dots, A_n, \mathfrak{A}$$

существует $\mathfrak{r}$ **л.н.з.** рядов, через которые выражаются все ряды этой системы. Поэтому

$$\operatorname{rank}\{A_1, \dots, A_n, \mathfrak{A}\} = \operatorname{rank}\{A_1, \dots, A_n\}.$$

б) Пусть для определенности в системе (11.1) ряд A_n линейно выражается через остальные ряды системы. Удалим его. Пусть ранг системы

$$\{A_1, \dots, A_{n-1}\} \tag{11.7}$$

равен $\mathfrak{r}$. Тогда, на основании доказанного в пункте **а)**, и ранг исходной системы (11.1) должен равняться $\mathfrak{r}$, так как она получается путем присоединения к (11.7) ряда A_n, линейно выражающегося через ряды (11.7). ■

Этот результат можно усилить, введя следующее

Определение 11.4. **Элементарными преобразованиями системы рядов** (11.1) называются следующие

A) перестановка двух любых рядов;

B) умножение ряда на число $c \ne 0$;

C) прибавление к одному ряду любого другого ряда.

Теорема 11.6. *Элементарные преобразования не меняют ранга системы рядов.*

Доказательство. Пусть $\operatorname{rank}\{A_1, \dots, A_i, \dots, A_j, \dots, A_n\} = \mathfrak{r}$. Дополним эту систему рядом $Y \overset{\text{def}}{=} cA_i$, где $c \ne 0$:

$$A_1, \dots, A_i, \dots, A_j, \dots, A_n, Y. \tag{11.8}$$

Новая система, на основании теоремы 11.5, будет иметь тот же ранг, так как Y линейно выражается через A_i. Но, в свою очередь, A_i можно выразить через Y: $A_i = {}^1\!/_c Y$. Следовательно, удаляя из системы (11.8) ряд A_i, мы придем к системе

$$A_1, \dots, A_i, \dots, A_j, \dots, A_n, Y$$

того же ранга $\mathfrak{r}$. Но эта система как раз и получается из первоначальной с помощью второго элементарного преобразования.

Рассмотрим теперь третье элементарное преобразование. Дополним систему рядом $Z \overset{\text{def}}{=} A_i + A_j$:

$$A_1, \dots, A_i, \dots, A_j, \dots, A_n, Z. \tag{11.9}$$

Ранг новой системы остается равным $\mathfrak{r}$ так как Z линейно выражается через ряды первоначальной системы. Но, в свою очередь,

A_i можно выразить через Z и A_j: $A_i = Z - A_j$. Следовательно, удаляя из системы (11.9) ряд A_i, мы придем к системе

$$A_1, \dots, A_i, \dots, A_j, \dots, A_n, Z$$

того же ранга $\mathfrak{r}$. Но эта система как раз и получается из первоначальной с помощью третьего элементарного преобразования. ■

Следствие 11.1. *Если каждый из рядов системы* (11.1) *линейно выражается через ряды* $B_1, \dots, B_k$, *то*

$$\operatorname{rank}\{A_1, \dots, A_n\} \le \operatorname{rank}\{B_1, \dots, B_k\}.$$

Доказательство. Составим объединенную систему

$$\{A_1, \dots, A_n, B_1, \dots, B_k\}.$$

Ранг любой подсистемы не может превосходить ранга самой системы, в частности,

$$\operatorname{rank}\{A_1, \dots, A_n\} \le \operatorname{rank}\{A_1, \dots, A_n, B_1, \dots, B_k\}.$$

С другой стороны, на основании теоремы 11.5 число справа равно $\operatorname{rank}\{B_1, \dots, B_k\}$. ■

Ранг матрицы

Определение 11.5. **Рангом матрицы** $A_{m\times n}$ называется наибольший порядок ее отличных от нуля миноров. Иначе говоря: $\operatorname{rank} A = \mathfrak{r} \in \mathbb{N}$ тогда и только тогда, когда существует ее минор порядка $\mathfrak{r}$, отличный от нуля, а все миноры более высокого порядка равны нулю. Кроме того, полагают:

$$\operatorname{rank} \mathbb{O}_{m\times n} \stackrel{\text{def}}{=} 0.$$

Очевидно, что если все миноры матрицы A порядка $\mathfrak{r}+1$ равны нулю, то и все миноры бо́льшего порядка должны быть равны нулю. Но, оказывается для проверки условия $\operatorname{rank} A = \mathfrak{r}$ нет необходимости вычислять и <u>все</u> миноры порядка $\mathfrak{r}+1$.

Определение 11.6. Для произвольного минора матрицы A порядка $k < \min(m,n)$

$$A\begin{pmatrix}\alpha_1\,\alpha_2\ldots\,\alpha_k\\ \beta_1\,\beta_2\ldots\,\beta_k\end{pmatrix}, \qquad \begin{cases}1 \le \alpha_1 < \alpha_2 < \cdots < \alpha_k \le m\\ 1 \le \beta_1 < \beta_2 < \cdots < \beta_k \le n\end{cases}$$

минор $(k+1)$-го порядка, получающийся добавлением еще одной строки и одного столбца:

$$A\begin{pmatrix}\alpha_1\,\alpha_2\ldots\,\alpha_k\,\alpha_{k+1}\\ \beta_1\,\beta_2\ldots\,\beta_k\,\beta_{k+1}\end{pmatrix}, \qquad \alpha_{k+1} \ne \alpha_j, \beta_{k+1} \ne \beta_j$$

называется его **окаймляющим**.

Следующий результат связывает понятие ранга матрицы с введенным в предыдущем пункте понятием ранга системы рядов. Рассмотрим системы строк и столбцов матрицы A:

$$\{A^{[1]},\ldots,A^{[m]}\} \quad \text{и} \quad \{A_{[1]},\ldots,A_{[n]}\}.$$

Теорема 11.7. *Если матрица A имеет минор порядка $\mathfrak{r}$, отличный от нуля, для которого все окаймляющие его миноры порядка $\mathfrak{r}+1$ равны нулю, то ранг системы ее столбцов равен $\mathfrak{r}$, и ранг системы ее строк равен $\mathfrak{r}$.*

Доказательство. Для определенности предположим, что отличен от нуля минор, стоящий в левом верхнем углу матрицы[22]:

$$\Delta \stackrel{\text{def}}{=} A\begin{pmatrix}1\,2\ldots\,\mathfrak{r}\\ 1\,2\ldots\,\mathfrak{r}\end{pmatrix}.$$

При таком расположении минора первые $\mathfrak{r}$ строк матрицы A будут **л.н.з.** Действительно, если бы первые $\mathfrak{r}$ строк были линейно зависимыми, то по теореме 11.1 одна из них выражалась бы через остальные. Вычитая из этой строки соответствующую линейную комбинацию остальных, мы все ее элементы превратили бы в нули. Но тогда одна из строк минора Δ стала бы нулевой. Сам минор при таких преобразованиях не должен менять своего значения (по

[22] Если бы ненулевой минор находился в другом месте матрицы, то рассуждения были бы аналогичными.

свойству линейности определителя). Получили: $\Delta = 0$, что противоречит предположению.

Покажем теперь, что все строки матрицы A линейно выражаются через первые $\mathfrak{r}$ строк.

Выберем какой-то ненулевой элемент первой строки минора Δ: пусть $a_{1k} \neq 0,\ k \leq \mathfrak{r}$. Такой элемент всегда существует, в противном случае $\Delta = 0$. Вычтем из второй, третьей, $\dots$, $m-$й строки матрицы A ее первую строку, умноженную соответственно на $a_{2k}/a_{1k}, a_{3k}/a_{1k}, \dots, a_{mk}/a_{1k}$. Тогда все элементы k-го столбца матрицы, кроме a_{1k}, обратятся в нули. Очевидно, что при всех этих преобразованиях как сам минор Δ, так и любой минор, его окаймляющий, не изменят своих значений. После таких преобразований вторая строка минора Δ изменится, однако она будет содержать по крайней мере один элемент, отличный от нуля; в противном случае $\Delta = 0$. Пусть таким элементом будет $a'_{2\ell}$. Очевидно, $\ell \neq k$, так как все элементы k-го столбца, кроме a_{1k} стали равными нулю. Вычтем из третьей, $\dots$, $m-$й строки преобразованной матрицы ее вторую строку, умноженную соответственно на $a'_{3\ell}/a'_{2\ell}, \dots, a'_{m\ell}/a'_{2\ell}$. Тогда все элементы $\ell-$го столбца матрицы, кроме первых двух, обратятся в нули. Затем обращаемся к третьей строке и т.д., пока не дойдем до $\mathfrak{r}-$й строки. В результате мы добьемся того, что в каждой из строк матрицы, лежащих под минором Δ, будут равны нулю первые $\mathfrak{r}$ элементов[23]. Оказывается, что будут равны нулю и все элементы каждой такой строки. В самом деле, пусть b_{ij} — элемент с индексами $i > \mathfrak{r},\ j > \mathfrak{r}$ преобразованной матрицы. Возьмем минор $(\mathfrak{r}+1)$-го порядка этой матрицы, окаймляющий Δ и имеющий в правом нижнем углу элемент b_{ij}:

$$D = \begin{vmatrix} & & b_{1j} \\ & & \vdots \\ & \Delta & \\ & & b_{\mathfrak{r}j} \\ 0 & \dots\ 0 & b_{ij} \end{vmatrix}.$$

Разлагая D по элементам последней строки, получаем

$$D = b_{ij}\Delta.$$

[23]Здесь предполагается, что $\mathfrak{r} < m$. В случае $\mathfrak{r} = m$ все строки матрицы линейно независимы и $\operatorname{rank}\{A^{[1]}, \dots, A^{[m]}\} = \mathfrak{r}$.

Но по условию $D = 0$ и $\Delta \neq 0$. Следовательно, $b_{ij} = 0$, что и требовалось доказать.

Таким образом, после всех указанных преобразований элементы $(\mathfrak{r}+1)$-й, $(\mathfrak{r}+2)$-й, $\ldots$, m-й строк матрицы A стали равными нулю. Это означает, что $(\mathfrak{r}+1)$-я, $(\mathfrak{r}+2)$-я, $\ldots$, m-я строки линейно выражаются через первые $\mathfrak{r}$ строк.

Итак, первые $\mathfrak{r}$ строк матрицы A **л.н.з.**, а все остальные строки линейно через них выражаются. По теореме 11.3

$$\operatorname{rank}\{A^{[1]}, \ldots, A^{[m]}\} = \mathfrak{r}.$$

Наконец, если транспонировать матрицу A, то в $A^\top$ минор Δ и его окаймляющие $(\mathfrak{r}+1)-$го порядка сохраняют свои значения на основании известного свойства определителей. Поскольку строками $A^\top$ являются транспонированные столбцы $A_{[1]}, \ldots, A_{[n]}$ матрицы A, то отсюда вытекает справедливость теоремы и для столбцов:

$$\operatorname{rank}\{A_{[1]}, \ldots, A_{[n]}\} = \mathfrak{r}.$$

■

Следствие 11.2. *Ранг матрицы равен рангу системы ее строк (столбцов).*

Доказательство. Пусть $\operatorname{rank} A = \mathfrak{r}$. Тогда, по определению, существует по крайней мере один минор $\mathfrak{r}-$го порядка, отличный от нуля, а все миноры более высокого порядка (если такие имеются) равны нулю. Отсюда, в частности, следует, что все миноры $(\mathfrak{r}+1)-$го порядка, окаймляющие данный, равны нулю. По теореме 11.7 ранг системы строк (столбцов) равен $\mathfrak{r}$, т.е. равен $\operatorname{rank} A$. ■

Следствие 11.3. $\det A_{n\times n} = 0 \iff$ *его строки (столбцы) линейно зависимы.*

Доказательство. Если строки $\det A$ линейно зависимы, то ранг системы строк меньше n. По следствию 11.2 и $\operatorname{rank} A < n$, но тогда дожно быть $\det A = 0$. Обратно, если $\det A = 0$, то $\operatorname{rank} A < n$ и потому система строк матрицы A линейно зависима. ■

Алгоритм вычисления ранга по методу окаймляющих миноров:

A) ищем минор первого порядка (т.е. элемент матрицы a_{jk}), отличный от нуля (если такого нет, то $\operatorname{rank} A = 0$);

B) ищем минор второго порядка, содержащий a_{jk} и отличный от нуля:

$$A\begin{pmatrix} j\, j_2 \\ k\, k_2 \end{pmatrix} = \begin{vmatrix} a_{jk} & a_{jk_2} \\ a_{j_2k} & a_{j_2k_2} \end{vmatrix}$$

(если такого нет, то $\operatorname{rank} A = 1$);

C) продолжаем процесс окаймления до тех пор, пока не найдем такой минор $\mathfrak{r}-$го порядка, который сам отличен от нуля, а все окаймляющие его $(\mathfrak{r}+1)-$го порядка равны нулю. Тогда $\operatorname{rank} A = \mathfrak{r}$.

Пример 11.3. Вычислить ранг матрицы

$$\begin{pmatrix} 3 & 4 & -1 & 5 & -2 \\ 1 & 5 & -2 & 3 & 4 \\ 2 & -1 & 1 & 2 & 3 \\ 3 & -7 & 4 & 1 & -7 \\ 0 & 11 & -5 & 4 & -4 \end{pmatrix}.$$

Решение. Для проверки условия $\mathfrak{r} > 0$ достаточно найти хотя бы один ненулевой элемент матрицы.

$$\begin{pmatrix} \underline{3|} & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix}.$$

Далее, для проверки условия $\mathfrak{r} > 1$ пытаемся подобрать ненулевой минор второго порядка, окаймляя выбранный минор первого порядка:

$$\begin{pmatrix} 3 & 4| & * & * & * \\ \underline{1} & \underline{5|} & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix}.$$

Из всех окаймляющих этот минор миноров третьего порядка:

$$
\left(\begin{array}{ccc|cc}
3 & 4 & -1 & * & * \\
1 & 5 & -2 & * & * \\
2 & -1 & 1 & * & * \\
\hline
* & * & * & * & * \\
* & * & * & * & *
\end{array}\right),\quad
\left(\begin{array}{ccc|cc}
3 & 4 & -1 & * & * \\
1 & 5 & -2 & * & * \\
\hline
* & * & * & * & * \\
3 & 7 & 4 & * & * \\
\hline
* & * & * & * & *
\end{array}\right),\quad
\left(\begin{array}{ccc|cc}
3 & 4 & -1 & * & * \\
1 & 5 & -2 & * & * \\
\hline
* & * & * & * & * \\
* & * & * & * & * \\
0 & 11 & -5 & * & *
\end{array}\right),
$$

$$
\left(\begin{array}{cc|c|c|c}
3 & 4 & * & 5 & * \\
1 & 5 & * & 3 & * \\
2 & -1 & * & 2 & * \\
\hline
* & * & * & * & * \\
* & * & * & * & *
\end{array}\right),\quad
\left(\begin{array}{cc|c|c|c}
3 & 4 & * & 5 & * \\
1 & 5 & * & 3 & * \\
\hline
* & * & * & * & * \\
3 & -7 & * & 1 & * \\
\hline
* & * & * & * & *
\end{array}\right),\quad
\left(\begin{array}{cc|c|c|c}
3 & 4 & * & 5 & * \\
1 & 5 & * & 3 & * \\
\hline
* & * & * & * & * \\
* & * & * & * & * \\
0 & 11 & * & 4 & *
\end{array}\right),
$$

$$
\left(\begin{array}{cc|cc|c|}
3 & 4 & * & * & -2 \\
1 & 5 & * & * & 4 \\
2 & -1 & * & * & 3 \\
\hline
* & * & * & * & * \\
* & * & * & * & *
\end{array}\right),\quad
\left(\begin{array}{cc|cc|c|}
3 & 4 & * & * & -2 \\
1 & 5 & * & * & 4 \\
\hline
* & * & * & * & * \\
3 & -7 & * & * & -7 \\
\hline
* & * & * & * & *
\end{array}\right),\quad
\left(\begin{array}{cc|cc|c|}
3 & 4 & * & * & -2 \\
1 & 5 & * & * & 4 \\
\hline
* & * & * & * & * \\
* & * & * & * & * \\
0 & 11 & * & * & -4
\end{array}\right)
$$

два последних отличны от нуля[24]. Это означает, что $\mathfrak{r} \geq 3$. Окаймляем какой-нибудь из ненулевых миноров третьего порядка. Оказывается, что все 4 минора четвертого порядка равны нулю.

Ответ. Ранг матрицы равен 3.

Еще один способ вычисления ранга матрицы был фактически нами использован при доказательстве теоремы 11.7. Теорема 11.6 утверждала, что элементарные преобразования не меняют ранга системы рядов. Отсюда сразу вытекает, что ранг матрицы A не изменяется при элементарных преобразованиях ее строк и столбцов. Простейший метод вычисления ранга матрицы с числовыми элементами состоит в приведении ее элементарными преобразованиями к ступенчатой матрице.

Алгоритм вычисления ранга по методу элементарных преобразований:

А) ищем ненулевой элемент матрицы a_{jk} (если такого нет, то $\operatorname{rank} A = 0$);

[24] Подчеркнем: нас интересовал хотя бы один отличный от нуля минор; если бы нам повезло его угадать сходу, то не нужно было бы вычислять все остальные миноры третьего порядка.

B) перестановкой строк и столбцов матрицы, добиваемся, чтобы ненулевой элемент a_{jk} попал в левый верхний угол;

C) применяя метод исключения Гаусса, добиваемся, чтобы все элементы первого столбца полученной матрицы, кроме верхнего, обратились в нуль;

D) к полученной в результате исключения подматрице порядка $(m-1)\times(n-1)$ применяем процедуры пунктов A)–C).

Процесс заканчивается, когда матрица оказывается приведенной к виду:

$$\begin{pmatrix} b_{11} & b_{12} & \dots & b_{1\mathfrak{r}} & b_{1,\mathfrak{r}+1} & \dots & b_{1n} \\ 0 & b_{22} & \dots & b_{2\mathfrak{r}} & b_{2,\mathfrak{r}+1} & \dots & b_{2n} \\ \vdots & & \ddots & & & & \vdots \\ 0 & 0 & \dots & b_{\mathfrak{r}\mathfrak{r}} & b_{\mathfrak{r},\mathfrak{r}+1} & \dots & b_{\mathfrak{r}n} \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & & & & & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

Здесь $b_{11}\neq 0, b_{22}\neq 0,\dots, b_{\mathfrak{r}\mathfrak{r}}\neq 0$. Тогда $\operatorname{rank} A = \mathfrak{r}$.

Пример 11.4. Вычислить ранг матрицы

$$\begin{pmatrix} 2 & 3 & 5 & -3 & -2 \\ 3 & 4 & 3 & -1 & -3 \\ 5 & 6 & -1 & 3 & -5 \end{pmatrix}.$$

Решение. Элементарными преобразованиями строк матрицы, приводим ее к ступенчатой:

$$\to \begin{pmatrix} 2 & 3 & 5 & -3 & -2 \\ 1 & 1 & -2 & 2 & -1 \\ 5 & 6 & -1 & 3 & -5 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & -2 & 2 & -1 \\ 2 & 3 & 5 & -3 & -2 \\ 5 & 6 & -1 & 3 & -5 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & -2 & 2 & -1 \\ 0 & 1 & 9 & -7 & 0 \\ 0 & 1 & 9 & -7 & 0 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & 1 & -2 & 2 & -1 \\ 0 & 1 & 9 & -7 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Ответ. Ранг матрицы равен 2.

Перейти к оглавлению на странице: 460

Упражнение 35. *Найти ранги матриц по методу элементарных преобразований:*

$$\text{а)}\ \begin{pmatrix} 17 & 51 & 27 & 31 \\ 93 & 25 & 14 & 121 \\ 94 & 27 & 15 & 120 \\ 18 & 53 & 28 & 30 \end{pmatrix}, \qquad \text{б)}\ \begin{pmatrix} 8 & 5 & 3 & 7 & 1 & 2 \\ 4 & 6 & 7 & 9 & 11 & 12 \\ 3 & 0 & -1 & 2 & 7 & 1 \\ 15 & 11 & 9 & 13 & 19 & 15 \\ 17 & 10 & 2 & -5 & 6 & 3 \end{pmatrix}.$$

Выясним теперь какие действия над матрицей не изменят ее ранга.

Теорема 11.8. *[Сильвестр] Для любых матриц $A_{m\times n}$ и $B_{n\times \ell}$ имеет место* **неравенство Сильвестра***:*

$$\operatorname{rank} A + \operatorname{rank} B - n \le \operatorname{rank}(AB) \le \min(\operatorname{rank} A, \operatorname{rank} B) \qquad (11.10)$$

Доказательство проведем только для правого неравенства. На основании следствия 11.2 к теореме 11.7, имеем:

$$\operatorname{rank} A = \operatorname{rank}\{A_{[1]},\dots,A_{[n]}\}$$

т.е. рангу системы столбцов. Аналогично:

$$\operatorname{rank}(AB) = \operatorname{rank}\left(\left[A_{[1]},\dots,A_{[n]}\right] B\right) =$$

$$= \operatorname{rank}\{b_{11}A_{[1]} + \dots + b_{n1}A_{[n]},\dots, b_{1\ell}A_{[1]} + \dots + b_{n\ell}A_{[n]}\},$$

т.е. равен рангу системы линейных комбинаций столбцов $\{A_{[1]},\dots,A_{[n]}\}$. По следствию 11.1 к теореме 11.6, ранг этой комбинации не может превосходить ранга комбинируемых столбцов: $\operatorname{rank}(AB) \le \operatorname{rank} A$. Неравенство $\operatorname{rank}(AB) \le \operatorname{rank} B$ доказывается аналогично. ■

Следствие 11.4. *Если A и B — квадратные матрицы $n-$го порядка и $\det B \ne 0$, то $\operatorname{rank}(AB) = \operatorname{rank} A$.*

Доказательство. Обозначим $C = AB$. По доказанному, $\operatorname{rank} C \le \le \operatorname{rank} A$. С другой стороны, $A = CB^{-1}$ и на основании теоремы 11.8 должно быть: $\operatorname{rank} A \le \operatorname{rank} C$. Из двух неравенств следует требуемое равенство. ■

Перейти к оглавлению на странице: 460

Упражнение 36. *Найти ранги матриц*

$$
\text{а)} \begin{pmatrix} \lambda & 1 & 1 & 1 & 1 \\ 1 & \lambda & 1 & 1 & 1 \\ 1 & 1 & \lambda & 1 & 1 \\ 1 & 1 & 1 & \lambda & 1 \end{pmatrix}, \quad \text{б)} \begin{pmatrix} \lambda & 1 & 3 & 4 & 1 \\ 1 & \lambda & -1 & 1 & 5 \\ \lambda & \lambda & 4 & 3 & -4 \\ 1 & 1 & 7 & 7 & -3 \end{pmatrix}
$$

в зависимости от значений параметра λ.

§12. Условия совместности линейной системы

Теорема Кронекера — Капелли

Рассмотрим систему линейных уравнений

$$
\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1, \\ \dots \qquad \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{cases} \qquad \Longleftrightarrow \qquad AX = \mathcal{B} \tag{12.1}
$$

Наряду с матрицей A этой системы рассмотрим еще **расширенную матрицу**:

$$
C \stackrel{\text{def}}{=} \left[A \middle| \mathcal{B}\right] = \left[A_{[1]}, \dots, A_{[n]}, \mathcal{B}\right] = \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \dots & & \dots & \dots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix},
$$

получаемую приписыванием к матрице A столбца правых частей.

Теорема 12.1. *[Кронекер, Капелли] Система* (12.1) *совместна тогда и только тогда когда ранг матрицы* A *совпадает с рангом расширенной матрицы* C:

$$
\operatorname{rank} A = \operatorname{rank} C. \tag{12.2}
$$

При выполнении этого условия система (12.1) *имеет единственное решение если* $\operatorname{rank} A = n$ *(т.е. числу неизвестных) и бесконечно много решений если* $\operatorname{rank} A < n$.

Доказательство необходимости. Пусть существует решение $x_1 = \alpha_1, \dots, x_n = \alpha_n$ системы, тогда

$$
\alpha_1 A_{[1]} + \dots + \alpha_n A_{[n]} = \mathcal{B},
$$

т.е. столбец $\mathcal{B}$ линейно выражается через столбцы $A_{[1]},\dots,A_{[n]}$. Но тогда

$$\operatorname{rank}\{A_{[1]},\dots,A_{[n]}\}=\operatorname{rank}\{A_{[1]},\dots,A_{[n]},\mathcal{B}\}$$

по теореме 11.5. На основании следствия 11.2 к теореме 11.7, $\operatorname{rank} A=\operatorname{rank} C$. ■

Доказательство достаточности. Обратно, пусть $\operatorname{rank} A = \operatorname{rank} C = \mathfrak{r}$. По определению ранга матрицы, в матрице A существует минор порядка $\mathfrak{r}$, отличный от нуля; этот же минор останется и минором матрицы C. Пусть, для определенности, этот минор находится в левом верхнем углу матрицы[25]:

$$\Delta \overset{\mathrm{def}}{=} A\begin{pmatrix}1\,2\dots\mathfrak{r}\\ 1\,2\dots\mathfrak{r}\end{pmatrix}\neq 0. \tag{12.3}$$

Тогда первые $\mathfrak{r}$ строк матрицы A линейно независимы, а остальные будут линейно выражаться через них. Это же утверждение будет справедливо и для строк матрицы C. Умножая первые $\mathfrak{r}$ уравнений системы (12.1) на соответствующие числа и складывая их, получим любое оставшееся уравнение. Таким образом, система (12.1) может быть заменена эквивалентной ей системой из первых $\mathfrak{r}$ уравнений:

$$\left\{\begin{array}{l} a_{11}x_1+\dots+a_{1\mathfrak{r}}x_{\mathfrak{r}}+a_{1,\mathfrak{r}+1}x_{\mathfrak{r}+1}+\dots+a_{1n}x_n=b_1,\\ \qquad\qquad\dots\\ a_{\mathfrak{r}1}x_1+\dots+a_{\mathfrak{r}\mathfrak{r}}x_{\mathfrak{r}}+a_{\mathfrak{r},\mathfrak{r}+1}x_{\mathfrak{r}+1}+\dots+a_{\mathfrak{r}n}x_n=b_{\mathfrak{r}}\end{array}\right. \iff A'X=\mathcal{B}'. \tag{12.4}$$

Если $\mathfrak{r}=n$, то матрица A' квадратная. По предположению $\det A'\neq 0$. По теореме 6.1 (Крамера) решение такой системы единственно.

Пусть теперь $\mathfrak{r}<n$. Перепишем уравнения (12.4) в виде

$$\left\{\begin{array}{l} a_{11}x_1+\dots+a_{1\mathfrak{r}}x_{\mathfrak{r}}=b_1-(a_{1,\mathfrak{r}+1}x_{\mathfrak{r}+1}+\dots+a_{1n}x_n),\\ \qquad\qquad\dots\\ a_{\mathfrak{r}1}x_1+\dots+a_{\mathfrak{r}\mathfrak{r}}x_{\mathfrak{r}}=b_{\mathfrak{r}}-(a_{\mathfrak{r},\mathfrak{r}+1}x_{\mathfrak{r}+1}+\dots+a_{\mathfrak{r}n}x_n)\end{array}\right. \tag{12.5}$$

По предположению определитель матрицы при $x_1,\dots,x_{\mathfrak{r}}$ (т.е. $A_{\mathfrak{r}}$) отличен от нуля. По теореме Крамера у этой системы существует единственное решение относительно неизвестных $x_1,\dots,x_{\mathfrak{r}}$ при

[25] В противном случае, можно добиться этого перенумеровкой переменных и перестановкой уравнений.

произвольных фиксированных значениях $x_{\mathfrak{r}+1},\dots,x_n$:

$$x_j=\frac{\begin{vmatrix} a_{11}\dots a_{1,j-1}\,[b_1-(a_{1,\mathfrak{r}+1}x_{\mathfrak{r}+1}+\dots+a_{1n}x_n)]\,a_{1,j+1}\dots a_{1\mathfrak{r}}\\ \dots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \dots \qquad\qquad \dots\\ a_{\mathfrak{r}1}\dots a_{\mathfrak{r},j-1}\,[b_{\mathfrak{r}}-(a_{\mathfrak{r},\mathfrak{r}+1}x_{\mathfrak{r}+1}+\dots+a_{\mathfrak{r}n}x_n)]\,a_{\mathfrak{r},j+1}\dots a_{\mathfrak{r}\mathfrak{r}}\end{vmatrix}}{\Delta} \tag{12.6}$$

при $j\in\{1,\dots,\mathfrak{r}\}$. Таким образом, в этом случае система (12.1) имеет бесконечное множество решений. ■

Используя свойство линейности определителя по столбцу, формулы (12.6) можно переписать в виде

$$x_j=\beta_j+\gamma_{j,\mathfrak{r}+1}x_{\mathfrak{r}+1}+\dots+\gamma_{jn}x_n\ (j\in\{1,\dots,\mathfrak{r}\}). \tag{12.7}$$

Здесь

$$\beta_j\stackrel{\text{def}}{=}\frac{\begin{vmatrix} a_{11}\dots a_{1,j-1} & b_1 & a_{1,j+1}\dots a_{1\mathfrak{r}}\\ \dots & \dots & \dots\\ a_{\mathfrak{r}1}\dots a_{\mathfrak{r},j-1} & b_{\mathfrak{r}} & a_{\mathfrak{r},j+1}\dots a_{\mathfrak{r}\mathfrak{r}}\end{vmatrix}}{\Delta},$$

$$\gamma_{jk}\stackrel{\text{def}}{=}-\frac{\begin{vmatrix} a_{11}\dots a_{1,j-1} & a_{1k} & a_{1,j+1}\dots a_{1\mathfrak{r}}\\ \dots & \dots & \dots\\ a_{\mathfrak{r}1}\dots a_{\mathfrak{r},j-1} & a_{\mathfrak{r}k} & a_{\mathfrak{r},j+1}\dots a_{\mathfrak{r}\mathfrak{r}}\end{vmatrix}}{\Delta}.$$

Определение 12.1. Формулы (12.7) называются **общим решением** системы (12.1). В этих формулах переменные $x_{\mathfrak{r}+1},\dots,x_n$ называются **основными** (или свободными), а $x_1,\dots,x_{\mathfrak{r}}$ — **зависимыми**. Решение, получающееся из общего решения фиксированием значений основных переменных, называется **частным решением** системы (12.1).

Пример 12.1. Исследовать совместность и найти общее решение системы уравнений:

$$\begin{cases} 2x_1-x_2+x_3+2x_4+3x_5=2,\\ 6x_1-3x_2+2x_3+4x_4+5x_5=3,\\ 6x_1-3x_2+4x_3+8x_4+13x_5=9,\\ 4x_1-2x_2+x_3+x_4+2x_5=1\end{cases}$$

Решение. Приведем здесь два алгоритма решения этой задачи, соответствующие двум способам вычисления ранга матрицы. Вычисляем сначала ранг матрицы A по методу окаймляющих миноров:

$$|2| \neq 0, \qquad \begin{vmatrix} 2 & 1 \\ 6 & 2 \end{vmatrix} \neq 0, \qquad \begin{vmatrix} 2 & 1 & 2 \\ 6 & 2 & 4 \\ 4 & 1 & 1 \end{vmatrix} = 2 \neq 0,$$

а все миноры, окаймляющие последний, равны нулю. Итак, $\operatorname{rank} A = 3$. Для нахождения ранга расширенной матрицы C достаточно проверить окаймление найденного ненулевого минора третьего порядка с помощью элементов взятых из столбца правых частей. Имеется всего один такой минор, и он равен нулю. Следовательно $\operatorname{rank} C = 3$, система совместна, и имеет бесконечное множество решений.

Ненулевой минор третьего порядка находится в первой, второй и четвертых строках, что означает линейную независимость соответствующих уравнений. Третье уравнение линейно зависит от остальных, и может быть отброшено. Далее, указанный ненулевой минор третьего порядка образован коэффициентами при x_1, x_3 и x_4. Следовательно оставшиеся уравнения могут быть разрешены относительно этих переменных, т.е. они — зависимые, а x_2 и x_5 — основные. Использование формулы (12.6) дает общее решение

$$x_1 = \frac{\begin{vmatrix} 2 & 1 & 2 \\ 3 & 2 & 4 \\ 1 & 1 & 1 \end{vmatrix}}{2} - x_2 \frac{\begin{vmatrix} -1 & 1 & 2 \\ -3 & 2 & 4 \\ -2 & 1 & 1 \end{vmatrix}}{2} - x_5 \frac{\begin{vmatrix} 3 & 1 & 2 \\ 5 & 2 & 4 \\ 2 & 1 & 1 \end{vmatrix}}{2} = -\frac{1}{2} + \frac{x_2}{2} + \frac{x_5}{2},$$

$$x_3 = \frac{\begin{vmatrix} 2 & 2 & 2 \\ 6 & 3 & 4 \\ 4 & 1 & 1 \end{vmatrix}}{2} - x_2 \frac{\begin{vmatrix} 2 & -1 & 2 \\ 6 & -3 & 4 \\ 4 & -2 & 1 \end{vmatrix}}{2} - x_5 \frac{\begin{vmatrix} 2 & 3 & 2 \\ 6 & 5 & 4 \\ 4 & 2 & 1 \end{vmatrix}}{2} = 3 - 4x_5,$$

$$x_4 = \frac{\begin{vmatrix} 2 & 1 & 2 \\ 6 & 2 & 3 \\ 4 & 1 & 1 \end{vmatrix}}{2} - x_2 \frac{\begin{vmatrix} 2 & 1 & -1 \\ 6 & 2 & -3 \\ 4 & 1 & -2 \end{vmatrix}}{2} - x_5 \frac{\begin{vmatrix} 2 & 1 & 3 \\ 6 & 2 & 5 \\ 4 & 1 & 2 \end{vmatrix}}{2} = 0.$$

Решим теперь ту же задачу, воспользовавшись гауссовским методом исключения переменных в системе линейных уравнений (см.

§1):

$$\begin{cases} 2x_1 - x_2 + x_3 + 2x_4 + 3x_5 = 2, \\ x_3 + 2x_4 + 4x_5 = 3, \\ x_4 = 0 \end{cases}$$

Используя обратный ход метода Гаусса, снова приходим к полученным формулам.

Ответ. Общее решение системы: $x_1 = {}^1/_2(x_2 + x_5 - 1)$, $x_3 = 3 - 4\,x_5$, $x_4 = 0$.

Пример 12.2. Исследовать совместность и найти общее решение системы уравнений

$$\begin{cases} \lambda x_1 + x_2 + x_3 + x_4 = 1, \\ x_1 + \lambda x_2 + x_3 + x_4 = 1, \\ x_1 + x_2 + \lambda x_3 + x_4 = 1, \\ x_1 + x_2 + x_3 + \lambda x_4 = 1, \end{cases}$$

в зависимости от значения параметра λ.

Решение. В этом примере число уравнений совпадает с числом неизвестных. Это обстоятельство несколько облегчает рассуждения. Прежде всего, исследуем случай единственности решения. Для этого достаточно вычислить $\det A =$

$$= \begin{vmatrix} \lambda & 1 & 1 & 1 \\ 1 & \lambda & 1 & 1 \\ 1 & 1 & \lambda & 1 \\ 1 & 1 & 1 & \lambda \end{vmatrix} = \begin{vmatrix} (\lambda-1) & (1-\lambda) & 0 & 0 \\ 0 & (\lambda-1) & (1-\lambda) & 0 \\ 0 & 0 & (\lambda-1) & (1-\lambda) \\ 1 & 1 & 1 & \lambda \end{vmatrix} = (\lambda-1)^3 \begin{vmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 1 & 1 & 1 & \lambda \end{vmatrix} =$$

$= (\lambda - 1)^3(\lambda + 3)$. При $\lambda \neq 1$ и при $\lambda \neq -3$ решение системы единственно:

$$x_1 = x_2 = x_3 = x_4 = 1/(\lambda + 3).$$

Осталось исследовать критические случаи. При $\lambda = 1$ имеем

$$\operatorname{rank} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \operatorname{rank} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} = 1,$$

и система совместна. Она эквивалентна единственному уравнению

$$x_1 + x_2 + x_3 + x_4 = 1,$$

из которого получаем общее решение

$$x_1 = 1 - x_2 - x_3 - x_4.$$

При $\lambda = -3$

$$\operatorname{rank}\begin{pmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{pmatrix} = 3, \quad \operatorname{rank}\begin{pmatrix} -3 & 1 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 & 1 \\ 1 & 1 & -3 & 1 & 1 \\ 1 & 1 & 1 & -3 & 1 \end{pmatrix} = 4$$

и система несовместна.

Ответ. Система несовместна при $\lambda = -3$; она имеет единственное решение $x_1 = x_2 = x_3 = x_4 = 1/(\lambda+3)$ при $\lambda \neq 1$ и общее решение $x_1 = 1 - x_2 - x_3 - x_4$ при $\lambda = 1$.

Система однородных уравнений

В частном случае обращения в нуль правых частей уравнений (12.1) эта система называется **однородной**:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0, \\ \qquad \cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = 0 \end{cases} \iff AX = \mathbb{O}. \tag{12.8}$$

Такая система всегда совместна, так как имеет **тривиальное** решение $x_1 = 0, \ldots, x_n = 0$. Из теоремы 12.1 следует

Теорема 12.2. *Система (12.8) имеет только тривиальное решение, если* $\operatorname{rank} A = n$ *и бесконечно много решений если* $\operatorname{rank} A < n$.

Следствие 12.1. *Если матрица системы (12.8) квадратная, то для существования нетривиального решения* **Н.** *и* **Д.** *чтобы* $\det A = 0$.

Определение 12.2. Для системы (12.1) система (12.8) называется **соответствующей** (или приведенной) однородной системой.

Теорема 12.3. *Общее решение неоднородной системы (12.1) можно представить в виде суммы какого-либо частного решения этой системы и общего решения соответствующей однородной системы (12.8).*

Доказательство фактически следует из формул (12.7). В этих формулах β_j представляет решение системы (12.1), получаемое при $x_{\mathfrak{r}+1} = 0, \ldots, x_n = 0$. Величины же коэффициентов γ_{jk} вовсе не зависят от правых частей системы (12.1) и будут одинаковыми при любых значениях $b_1, \ldots, b_m$. В частности, если $b_1{=}0, \ldots, b_m{=}0$, то в формулах (12.7) величины β_j обращаются в нуль и эти формулы превращаются в

$$x_j = \gamma_{j,\mathfrak{r}+1}x_{\mathfrak{r}+1} + \cdots + \gamma_{jn}x_n \ (j \in \{1, \ldots, \mathfrak{r}\}). \tag{12.9}$$

Но это как раз и есть общее решение однородной системы (12.8), соответствующей системе (12.1). ■

Определение 12.3. Совокупность $\{X_1, \ldots, X_k\}$ решений системы (12.8) называется **фундаментальной системой решений** (сокращение— **ф.с.р.**) если она линейно независима и любое другое решение системы линейно выражается через $X_1, \ldots, X_k$. **Ф.с.р.** не определяется для случая $\operatorname{rank} A = n$.

Теорема 12.4. *Если* $\operatorname{rank} A = \mathfrak{r} < n$, *то система* (12.8) *имеет* **ф.с.р.**, *состоящую из* $n - \mathfrak{r}$ *элементов. Число элементов в* **ф.с.р.** *не зависит от способа ее построения.*

Доказательство. Пусть, для определенности, отличен от нуля минор (12.3). Повторяя все рассуждения доказательства теоремы 12.1, получим общее решение системы (12.8) в виде (12.9). Число основных переменных равно $n - \mathfrak{r}$. Придадим этим переменным системы значений

$$\begin{array}{ll} x_{\mathfrak{r}+1}{=}\mathbf{x}_{1,\mathfrak{r}+1}, & \ldots, x_n{=}\mathbf{x}_{1n} \\ \ldots & \ldots \\ x_{\mathfrak{r}+1}{=}\mathbf{x}_{n-\mathfrak{r},\mathfrak{r}+1}, & \ldots, x_n{=}\mathbf{x}_{n-\mathfrak{r},n} \end{array} \quad \text{так, чтобы} \quad \mathbf{D} \overset{\text{def}}{=} \begin{vmatrix} \mathbf{x}_{1,\mathfrak{r}+1} & \ldots & \mathbf{x}_{1n} \\ \ldots & & \ldots \\ \mathbf{x}_{n-\mathfrak{r},\mathfrak{r}+1} & \ldots & \mathbf{x}_{n-\mathfrak{r},n} \end{vmatrix} \neq 0.$$

Этого всегда можно добиться, взяв в качестве $\mathbf{D}$ единичную матрицу. Восстановив по формулам (12.9) значения зависимых переменных, получим некоторую систему из $n - \mathfrak{r}$ решений:

$$\left\{ \begin{array}{llll} (\mathbf{x}_{11}, & \ldots \mathbf{x}_{1\mathfrak{r}}, & \mathbf{x}_{1,\mathfrak{r}+1}, & \ldots \mathbf{x}_{1,n}) \\ \ldots & & \ldots & \ldots \\ (\mathbf{x}_{n-\mathfrak{r},1}, & \ldots \mathbf{x}_{n-\mathfrak{r},\mathfrak{r}}, & \mathbf{x}_{n-\mathfrak{r},\mathfrak{r}+1}, & \ldots \mathbf{x}_{n-\mathfrak{r},n}) \end{array} \right\} \tag{12.10}$$

Покажем, что (12.10) образуют **ф.с.р.** системы (12.8). Предположим, что имеется еще одно решение этой системы

$$(\mathbf{x}_1,\dots,\mathbf{x}_{\mathfrak{r}},\mathbf{x}_{\mathfrak{r}+1},\dots,\mathbf{x}_n).$$

Составим матрицу

$$M=\begin{pmatrix} \mathbf{x}_{11} & \dots \mathbf{x}_{1\mathfrak{r}} & \mathbf{x}_{1,\mathfrak{r}+1} & \dots \mathbf{x}_{1,n} \\ \dots & & & \mathbf{D}\ \dots \\ \mathbf{x}_{n-\mathfrak{r},1} & \dots \mathbf{x}_{n-\mathfrak{r},\mathfrak{r}} & \mathbf{x}_{n-\mathfrak{r},\mathfrak{r}+1} & \dots \mathbf{x}_{n-\mathfrak{r},n} \\ \mathbf{x}_1 & \dots \mathbf{x}_{\mathfrak{r}} & \mathbf{x}_{\mathfrak{r}+1} & \dots \mathbf{x}_n \end{pmatrix}$$

Покажем, что $\operatorname{rank} M = n-\mathfrak{r}$. Действительно, имеется минор этой матрицы порядка $n-\mathfrak{r}$ отличный от нуля, это — $\mathbf{D}$. Поэтому первые $n-\mathfrak{r}$ строк матрицы **л.н.з.** и $\operatorname{rank} M \ge n-\mathfrak{r}$. С другой стороны, на основании формул (12.9) первые $\mathfrak{r}$ столбцов матрицы линейно выражаются через последние $n-\mathfrak{r}$, т.е. $\operatorname{rank} M \le n-\mathfrak{r}$. Из двух неравенств для $\operatorname{rank} M$ получаем искомое равенство. Но оно означает, что последняя строка матрицы линейно выражается через предыдущие, и, следовательно, произвольное решение системы (12.8) может быть представлено в виде линейной комбинации решений (12.10).

Покажем теперь, что число элементов в **ф.с.р.** системы (12.8) не зависит от способа ее построения. Действительно, пусть

$$\{X_1,\dots,X_k\} \quad \text{и} \quad \{Y_1,\dots,Y_\ell\}$$

— две различные **ф.с.р.** системы (12.8). По определению **ф.с.р.** любое решение Y_j представимо в виде линейной комбинации $\{X_1,\dots,X_k\}$. Поскольку все $Y_1,\dots,Y_\ell$ **л.н.з.** то по теореме 11.2 должно быть $\ell \le k$. Но, с другой стороны, любое решение X_i представимо в виде линейной комбинации $\{Y_1,\dots,Y_\ell\}$ и на том же основании должно быть $k \le \ell$. Следовательно $\ell = k$. ■

Пример 12.3. Найти **ф.с.р.** для системы уравнений

$$\begin{cases} x_1 - x_2 + x_3 - x_4 = 0, \\ x_1 - x_2 + 2x_3 + 3x_4 = 0, \\ x_1 - x_2 - x_3 - 9x_4 = 0 \end{cases}$$

Решение. Приводим систему к ступенчатому виду:

$$\begin{cases} x_1 - x_2 + x_3 - x_4 = 0, \\ x_3 + 4x_4 = 0 \end{cases}$$

В качестве зависимых переменных можно взять, например, x_1 и x_3.

$$\begin{array}{cc|cc} \multicolumn{2}{c}{\overbrace{\hphantom{x_1 \quad x_3}}^{\text{зависимые}}} & \multicolumn{2}{c}{\overbrace{\hphantom{x_2 \quad x_4}}^{\text{основные}}} \\ x_1 & x_3 & x_2 & x_4 \\ \hline 1 & 0 & 1 & 0 \\ 5 & -4 & 0 & 1 \\ & & \multicolumn{2}{c}{\underbrace{\hphantom{0 \quad 1}}_{E}} \end{array}$$

Ответ. **Ф.с.р.** состоит, например, из решений

$$x_1 = 1, x_2 = 1, x_3 = 0, x_4 = 0 \quad \text{и} \quad x_1 = 5, x_2 = 0, x_3 = -4, x_4 = 1.$$

Упражнение 37. *Пусть матрица системы* (12.8) *квадратная и* $\operatorname{rank} A_{n\times n} = n - 1$. *Доказать, что если ненулевой минор матрицы порядка* $n-1$ *соответствует какому-нибудь элементу* i*-й строки, то система алгебраических дополнений к элементам этой строки составляет* **ф.с.р.** *для* (12.8). *Например, для системы*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = 0, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_2 = 0 \end{cases}$$

ф.с.р. *состоит из решения*

$$x_1 = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}, \; x_2 = -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}, \; x_3 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},$$

если только хотя бы один из миноров отличен от нуля.

Упражнение 38. *Найти* **ф.с.р.** *для систем уравнений*

а) $$\begin{cases} 5x_1 + 6x_2 - 2x_3 + 7x_4 + 4x_5 = 0, \\ 2x_1 + 3x_2 - x_3 + 4x_4 + 2x_5 = 0, \\ 7x_1 + 9x_2 - 3x_3 + 5x_4 + 6x_5 = 0, \\ 5x_1 + 9x_2 - 3x_3 + x_4 + 6x_5 = 0; \end{cases}$$

б) $$\begin{cases} x_1 - x_3 + x_5 = 0, \\ x_2 - x_4 + x_6 = 0, \\ x_1 - x_2 + x_5 - x_6 = 0, \\ x_2 - x_3 + x_6 = 0, \\ x_1 - x_4 + x_5 = 0. \end{cases}$$

Историческая справка

Понятие ранга матрицы и "теорема Кронекера–Капелли" были открыты несколькими независимыми исследователями. Первое доказательство этой теоремы принадлежит Ч.Л.Додсону, оно было напечатано им в 1867 г. в книге

An elementary treatise on determinants

в следующей формулировке.

Теорема. *Для того чтобы система n неоднородных уравнений была совместна, необходимо и достаточно, чтобы порядок наибольшего отличного от нуля минора был одинаков в расширенной и нерасширенной матрице системы.*

Дод(ж)сон Чарльз Лютвидж (Dod(g)son Charles Lutwidge, 1832–1898)

Один из самых знаменитых математиков мира работал магистром (master of house) Тринити–колледжа в Оксфорде, и имел сан дьякона. Д. был болезненно застенчив, глух на одно ухо, заикался. Был крайне педантичен: записывал в особую тетрадь содержание всех отправляемых и получаемых им писем — последнее зарегистрировано номером 98721.

Его трудами зачитывались (и зачитываются до сих пор) все — от школьника до королевы Виктории. Самые популярные его книги подписаны псевдонимом, полученным двойным переводом личных имен:

$$\underbrace{\text{Charles Lutwidge}}_{\text{английский}} \longrightarrow \underbrace{\text{Carolus Ludovicus}}_{\text{латынь}} \longrightarrow \underbrace{\text{Louis (Lewis) Carrol}}_{\text{английский}}$$

Сказка, начатая 4 июля 1862 г. на лодочной прогулке и рассказанная им Алисе Лиддел и двум ее сестрам переросла в книгу

Alice's adventures in wonderland (1865, "Алиса в стране чудес"),

а затем в ее продолжение

Through the looking glass (1872, "Алиса в Зазеркалье").

Д. вывел сам себя в качестве персонажа — птицы Додо (сам он, заикаясь, представлялся "До-до-дсон"). Книги при жизни автора разошлись в невероятном по тем временам количестве — 180 тыс. экземпляров. Это требовало труда, деловитости, организованности. Сказочник входил во все детали издательского дела, вплоть до упаковки тиража. В экспедиции издательства "Макмиллан" долгие годы в качестве рабочей инструкции висело его письмо с точными указаниями, как надо укладывать в пачки и перевязывать книги.

У Додсона было еще одно увлечение: фотография. В середине XIX века фотографирование требовало невероятной пунктуальности, терпения и чистоты. Зарегистрированное количество сделанных им снимков — 2700. Д. был исключительно силен как портретист. Наиболее известная фотография Фарадея принадлежит именно ему.

Служебная записка Д. руководству колледжа:

“*Поскольку фотография ныне широко используется для регистрации выражений человеческого лица, и, вероятно, может быть применена к алгебраическим выражениям, было бы желательно устроить небольшую фотографическую лабораторию — как для обычного применения, так и для регистрации тяжести, нарушений равновесия, решимости и пр., что проявляются на лице при сложных математических вычислениях*”.

Описание Додсоном одного семейного фотопортрета:

“*Предполагалось, что он представит младенца, которого венчают цветами соединенные усилия детей под руководством отца и личным наблюдением мамаши. Одновременно сцена должна была изображать*

«Невинность, на которую Победа возлагает лавровый венец при благосклонном содействии Решимости, Независимости, Веры, Надежды и Милосердия, в то время как Мудрость взирает на них с благосклонной улыбкой».

Результат для всякого непредубежденного наблюдателя обнаруживал, что младенец был в обмороке. Мамаша (несомненно вследствие какого-то ошибочного представления об анатомии человека) приводит его в чувство, сворачивая ему голову. Два юнца, в предвидении неизбежного исхода, хотят вырвать по локону с его головы в память о трагическом событии. Две девицы, ожидая своей очереди за локонами, пока что пытаются удавить третью. Папаша, в отчаянии от необычного поведения своего семейства, закололся, но еще тянется за карандашом, чтобы составить соответствующую записку”.

В 1867 г. вместе со своим коллегой преподобным Лидделлом (отцом той самой Алисы, которой посвящены знаменитые сказки) Додсон отправился в Россию. Кэролловеды долгое время пытались разведать, что побудило двух англичан отправиться за тридевять земель. И только недавно удалось установить: они ездили на празднование юбилея митрополита московского Леонида. Результатом путешествия явились заметки

“Дневник путешествия в Россию в 1867 году”

Маршрут пролегал через Германию. Посещение Берлина с его знаменитыми парками и скверами навело Д. на следующие размышления:

“*Мне кажется, что архитектура Берлина основана на двух принципах. Если на крыше найдется удобное местечко, туда необходимо поставить фигуру человека. Лучше всего, если он будет стоять на одной ноге. Если местечко найдется на земле, то на нем следует расставить по кругу бюсты на пъедесталах так, чтобы лицом они были обращены внутрь и как бы совещались о чем-то между собой, или воздвигнуть гигантскую фигуру человека, убивающего, намеревающегося убить или убившего (предпочтение отдается настоящему времени) какое-нибудь живое существо.*

Чем больше шипов у этого существа, тем лучше. Наиболее подходящим считается дракон, но если изобразить его художнику не под силу, то можно ограничиться львом или свиньей. «Принцип умерщвления живых тварей» проведен всюду с такой неукоснительной последовательностью, что некоторые районы Берлина выглядят как гигантская бойня доисторических животных."

В Петербурге Д. открыл много чудес:

"*Неподалеку от Адмиралтейства стоит великолепная конная статуя Петра Великого . . . Лошадь поднялась на дыбы, а у ее задней ноги извивается змея, на которую, как мне кажется, лошадь наступила. Если бы это памятник был воздвигнут в Берлине, то Петр, несомненно, был бы самым деятельным образом вовлечен в убийство чудовища. Здесь же он не обращает на змею никакого внимания: «теория умерщвления» в России не признана. Мы обнаружили также две гигантские фигуры львов, бывших до такой степени трогательно ручными, что каждый из них, как котенок, катил перед собой большой шар.*"

— А долго у вас шли занятия? — спросила Алиса, торопясь перевести разговор.
— Это зависело от нас, — отвечал Черепаха Квази. — Как все займем, так и кончим.
— Займете? — удивилась Алиса.
— Занятия почему так называются? — пояснил Грифон. — Потому что на занятиях мы у нашего учителя ум занимаем... А как все займем и ничего ему не оставим, тут же и кончим. В таких случаях говорят: «Ему ума не занимать»... Поняла?

Перейти к оглавлению на странице: 460

ГЛАВА 5. ИНТЕРПОЛЯЦИЯ

Определение 1. **Интерполяция** или интерполирование — приближенное или точное нахождение какой-либо величины по известным отдельным значениям этой же величины, или других величин, с ней связанных.

Историческая справка

Interpolatio (*лат.*) — изменение, искажение; **interpolo, interpolare** (*лат.*) — 1) разглаживать, подновлять, 2) искажать, подделывать, фальсифицировать.

Интерполяциями в римском праве назывались изменения, внесенные компиляторами при составлении Corpus juris civilis (530 г.) в тексты различных авторов и конституций. Юстинианов сборник представлял собою цельный законодательный акт, и в нем необходимо было устранить противоречия и устарелые положения. С этой целью составители одни тексты пропускали, другие исправляли, нередко приписывая цитируемому юристу слова, которых он не мог сказать, а подчас даже искажая тексты до неузнаваемости... Кстати, одна из статей кодекса называлась

> *De maleficiis et mathematicis et ceteris similibus*
> (О злодеях, математиках и прочих подобных)

и предписывала:

> *Ars autem mathematica damnabilis interdicta est omnino*
> (Преступное же искусство математики безусловно воспрещается).

§1. Интерполяционный полином

Будем обозначать через $\mathbb{A}$ какое-либо из множеств $\mathbb{Q}, \mathbb{R}$ или $\mathbb{C}$.

Постановка задачи

Задача. Построить полином $y = f(x)$, принимающий значения согласно следующей таблице:

$$\begin{array}{c|c} x & x_1 \dots x_n \\ \hline y & y_1 \dots y_n \end{array} \quad \{x_1, \dots, x_n, y_1, \dots, y_n\} \subset \mathbb{A}. \tag{1.1}$$

Определение 1.1. Искомый полином называется **интерполяционным**, а значения $x_1, \dots, x_n$ — **узлами** интерполяции.

Теорема 1.1. *Если все узлы x_j различны, то существует единственный полином $f(x) \in \mathbb{A}[x]$, $\deg f \le n-1$ такой, что $f(x_1) = y_1, \dots, f(x_n) = y_n$.*

Доказательство. Воспользуемся методом неопределенных коэффициентов[1]:

$$f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_{n-1} x^{n-1}.$$

Условия $f(x_j) = y_j$ позволяют выписать систему уравнений для определения $a_0, \dots, a_{n-1}$:

$$\left\{\begin{array}{l} a_0 + a_1 x_1 + \ldots + a_{n-1} x_1^{n-1} = y_1, \\ a_0 + a_1 x_2 + \ldots + a_{n-1} x_2^{n-1} = y_2, \\ \qquad\qquad \dots \qquad\qquad \dots \\ a_0 + a_1 x_n + \ldots + a_{n-1} x_n^{n-1} = y_n \end{array}\right. \iff$$

$$\iff \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ & \dots & & & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \dots \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}. \qquad (1.2)$$

Определитель матрицы системы является определителем Вандермонда[2] :

$$V(x_1, x_2, \dots, x_n) = \prod_{1 \le j < k \le n} (x_k - x_j),$$

он отличен от нуля ввиду условия теоремы. Поэтому существует единственное решение системы (1), которое может быть получено, например, по формулам Крамера[3]. Поскольку эти формулы представляют решение в виде рациональных функций от $x_1, \dots, x_n, y_1, \dots, y_n$, то $f(x) \in \mathbb{A}[x]$. ■

[1] Мы изменили здесь порядок нумерации коэффициентов полинома по сравнению с принятым в главе 3.

[2] Глава 4, §9.1.

[3] Глава 4, §6.

Пример 1.1. Построить интерп. полином по таблице

x	-3	-1	2	4
y	70	22	-5	7

и с его помощью интерполировать значение неизвестной функции при $x = 1$.

Решение. Система (1) имеет вид

$$\begin{pmatrix} 1 & -3 & 9 & -27 \\ 1 & -1 & 1 & -1 \\ 1 & 2 & 4 & 8 \\ 1 & 4 & 16 & 64 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 70 \\ 22 \\ -5 \\ 7 \end{pmatrix}.$$

Решив ее, получим $a_0 = 7$, $a_1 = -12$, $a_2 = 3$, $a_3 = 0$.

Ответ. $f(x) = 3\,x^2 - 12\,x + 7$, $f(1) = -2$.

Упражнение 39. *При каком условии на x_j, y_j степень построенного по теореме 1.1 полинома будет меньше $n-1$?*

Упражнение 40. *Пусть $f(x)$ является интерп. полиномом, построенным по теореме 1.1 (т.е. $\deg f \le n-1$). Укажите общий вид интерп. полинома произвольной степени, принимающего значения по той же таблице.*

Упражнение 41. *Пусть в таблице (1.1): $y_j = g(x_j)$, где $g(x)$ некоторый полином $\deg g = m > n-1$ (т.е. ставится задача приближения полинома большей степени полиномом меньшей степени). Доказать, что интерполяционный полином совпадает с остатком от деления $g(x)$ на $(x-x_1)\times\dots\times(x-x_n)$.*

Теорема 1.1 утверждает, что интерп. полином степени $\le n-1$ для таблицы (1.1) единствен. Но вот строить его можно разными способами.

Интерполяционный полином в форме Лагранжа

Поясним сначала идею на примере $n = 3$. Перепишем уравнения системы (1), дополнив их формальным тождеством $a_0 + a_1 x +$

$+a_2x^2 = f(x)$:

$$\begin{cases} a_0 + a_1x_1 + a_2x_1^2 & -y_1 & = 0, \\ a_0 + a_1x_2 + a_2x_2^2 & -y_2 & = 0, \\ a_0 + a_1x_3 + a_2x_3^2 & -y_3 & = 0, \\ a_0 + a_1x + a_2x^2 & -f(x) & = 0 \end{cases} \iff \begin{pmatrix} 1 & x_1 & x_1^2 & -y_1 \\ 1 & x_2 & x_2^2 & -y_2 \\ 1 & x_3 & x_3^2 & -y_3 \\ 1 & x & x^2 & -f(x) \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ 1 \end{pmatrix} = \mathbb{O}_{4\times 1}$$

Теорема 1.1 утверждает, что при условии различности x_j существует интерполяционный полином, принимающий указанные значения y_j. Следовательно, последняя система совместна: существует хотя бы один столбец $[a_0, a_1, a_2, 1]^\top$, ей удовлетворяющий. Поскольку этот набор не тривиальный, а система однородна, то по следствию 1 к теореме 12.2 главы 4, определитель этой системы должен быть равен нулю:

$$0 = \begin{vmatrix} 1 & x_1 & x_1^2 & -y_1 \\ 1 & x_2 & x_2^2 & -y_2 \\ 1 & x_3 & x_3^2 & -y_3 \\ 1 & x & x^2 & -f(x) \end{vmatrix} = \begin{vmatrix} 1 & x_1 & x_1^2 & 0 \\ 1 & x_2 & x_2^2 & 0 \\ 1 & x_3 & x_3^2 & 0 \\ 1 & x & x^2 & -f(x) \end{vmatrix} + \begin{vmatrix} 1 & x_1 & x_1^2 & -y_1 \\ 1 & x_2 & x_2^2 & -y_2 \\ 1 & x_3 & x_3^2 & -y_3 \\ 1 & x & x^2 & 0 \end{vmatrix}.$$

Отсюда получаем детерминантное представление интерп. полинома

$$f(x) = \begin{vmatrix} 1 & x_1 & x_1^2 & -y_1 \\ 1 & x_2 & x_2^2 & -y_2 \\ 1 & x_3 & x_3^2 & -y_3 \\ 1 & x & x^2 & 0 \end{vmatrix} \Bigg/ \begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix}. \tag{1.3}$$

При разложении по последнему столбцу определителя, стоящего в числителе, очевидным образом возникают определители Вандермонда:

$$f(x) = y_1 \frac{V(x_2, x_3, x)}{V(x_1, x_2, x_3)} - y_2 \frac{V(x_1, x_3, x)}{V(x_1, x_2, x_3)} + y_3 \frac{V(x_1, x_2, x)}{V(x_1, x_2, x_3)} =$$
$$= y_1 \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)} + y_2 \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)} + y_3 \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)},$$

и мы получаем представление интерп. полинома в форме, которая и называется **формой Лагранжа**.

Идея может быть развита и на общий случай, но мы приведем здесь другой вывод. Обозначим пока нам неизвестный интерп. полином через $f(x)$, на основании теоремы 1.1, $\deg f \le n-1$. Введем в рассмотрение новый полином

$$W(x) \stackrel{\text{def}}{=} (x-x_1)\times\cdots\times(x-x_n),\ \deg W = n. \tag{1.4}$$

Поскольку $\deg f \le n-1$, то дробь $f(x)/W(x)$ будет правильной. Разложим ее на простейшие дроби над $\mathbb{C}$, воспользовавшись формулой Лагранжа[4]:

$$\frac{f(x)}{W(x)} \equiv \sum_{j=1}^{n} \frac{f(x_j)}{W'(x_j)(x-x_j)} = \sum_{j=1}^{n} \frac{y_j}{W'(x_j)(x-x_j)}. \tag{1.5}$$

Обозначим

$$W_j(x) \stackrel{\text{def}}{=} \frac{w(x)}{x-x_j} = (x-x_1)\times\cdots\times(x-x_{j-1})(x-x_{j+1})\times\cdots\times(x-x_n).$$

Очевидно

$$W'(x_j) = W_j(x_j) = (x_j-x_1)\times\cdots\times(x_j-x_{j-1})(x_j-x_{j+1})\times\cdots\times(x_j-x_n).$$

Тогда, домножая (1.5) на $W(x)$, получаем интерп. полином в форме Лагранжа:

$$f(x) \equiv \sum_{j=1}^{n} \frac{y_j W_j(x)}{W_j(x_j)} = \tag{1.6}$$

$$= y_1 \frac{(x-x_2)\dots(x-x_n)}{(x_1-x_2)\dots(x_1-x_n)} + y_2 \frac{(x-x_1)(x-x_3)\dots(x-x_n)}{(x_2-x_1)(x_2-x_3)\dots(x_2-x_n)} + \cdots + \\ + y_n \frac{(x-x_1)\dots(x-x_{n-1})}{(x_n-x_1)\dots(x_n-x_{n-1})}.$$

Пример 1.2. Построить интерп. полином по таблице

x	-2	-1	1	2
y	-29	-8	-2	7

и с его помощью интерполировать значение неизвестной функции при $x=0$.

[4] Глава 3, теорема 10.7.

Решение. Имеем: $W(x) = (x+2)(x+1)(x-1)(x-2)$. По формуле (1.6):

$$f(x) = \frac{-29}{-12}(x+1)(x-1)(x-2) + \frac{-8}{6}(x+2)(x-1)(x-2) + \\ + \frac{-2}{-6}(x+2)(x+1)(x-2) + \frac{7}{12}(x+2)(x+1)(x-1).$$

Подставляем сюда $x = 0$: $f(0) = -3$. △

В предыдущем примере мы решили поставленную задачу, не преобразовывая форму Лагранжа интерп. полинома к канонической форме: после получения представления (1.6) не раскладывали правую часть по степеням переменной. Иногда такая необходимость может возникнуть, например, в случае, когда нам нужно найти корни интерп. полинома. Попробуем получить явное выражение коэффициентов интерп. полинома через $x_1, \dots, x_n, y_1, \dots, y_n$.

Теорема 1.2. *Пусть числа $x_1, \dots, x_n$ все различны. Для полинома $W(x)$, задаваемого формулой (1.4), справедливы следующие* **равенства Эйлера–Лагранжа***:*

$$\sum_{j=1}^{n} \frac{x_j^k}{W'(x_j)} = \begin{cases} 0 & \text{если } k < n-1; \\ 1 & \text{если } k = n-1. \end{cases} \tag{1.7}$$

Доказательство. Построим интерп. полином по следующей таблице:

x	$x_1 \dots x_n$
y	$x_1^k \dots x_n^k$

С одной стороны, ответ известен заранее: $f(x) \equiv x^k$. С другой стороны, формула (1.6) дает его же в виде суммы:

$$x^k \equiv \sum_{j=1}^{n} x_j^k \frac{W_j(x)}{W_j(x_j)} \equiv \underbrace{x^{n-1} \sum_{j=1}^{n} \frac{x_j^k}{W'(x_j)} + \dots}_{\text{разложение по степеням } x} .$$

В этом тождестве степени полиномов слева и справа должны быть одинаковыми. Если $k < n-1$, то старший коэффициент правого полинома должен обратиться в нуль. Если же $k = n-1$, то должны совпасть старшие коэффициенты обоих полиномов. ■

Следствие 1.1. *Обозначим*

$$\sigma_k \stackrel{\text{def}}{=} \sum_{j=1}^{n} \frac{x_j^{n+k-1}}{W'(x_j)} \quad , \quad \tau_k \stackrel{\text{def}}{=} \sum_{j=1}^{n} \frac{x_j^k y_j}{W'(x_j)}. \tag{1.8}$$

Имеют место равенства, связывающие коэффициенты интерп. полинома[5] $f(x) = A_0 x^{n-1} + \dots + A_{n-1}$ *с величинами* σ *и* τ*:*

$$\tau_0 = A_0,\ \tau_k = A_0\sigma_k + A_1\sigma_{k-1} + \dots + A_{k-1}\sigma_1 + A_k \ \text{ при } k \in \{1, \dots, n-1\}. \tag{1.9}$$

Доказательство. В самом деле, на основании формул (1.7) имеем:

$$\sigma_0 = 1,\ \sigma_{-1} = \dots = \sigma_{-n} = 0.$$

Тогда

$$\tau_k = \sum_{j=1}^{n} \frac{x_j^k y_j}{W'(x_j)} =$$

$$= \sum_{j=1}^{n} \frac{(A_0 x_j^{n-1} + A_1 x_j^{n-2} + \dots + A_k x_j^{n-k-1} + \dots + A_{n-1}) x_j^k}{W'(x_j)} =$$

$$= A_0\sigma_k + A_1\sigma_{k-1} + \dots + A_k\sigma_0.$$

■

Формулы (1.9) позволяют рекурсивно, начиная со старшего, вычислить коэффициенты интерп. полинома по величинам (1.8).

Пример 1.3. Найти корни интерп. полинома, заданного таблицей

x	1	2	3	4	5
y	1	−2	33	166	481

Решение. $W(x) = (x-1)(x-2)(x-3)(x-4)(x-5)$,

$$W'(1) = 24,\ W'(2) = -6,\ W'(3) = 4,\ W'(4) = -6,\ W'(5) = 24.$$

$$\sigma_1 = \frac{1^5}{W'(1)} + \frac{2^5}{W'(2)} + \frac{3^5}{W'(3)} + \frac{4^5}{W'(4)} + \frac{5^5}{W'(5)} = 15,$$

[5] Здесь мы снова вернулись к традиционной, т.е. канонической форме записи полинома.

$$\sigma_2 = \frac{1^6}{W'(1)} + \frac{2^6}{W'(2)} + \frac{3^6}{W'(3)} + \frac{4^6}{W'(4)} + \frac{5^6}{W'(5)} = 140,$$

$$\sigma_3 = 1050,\ \sigma_4 = 6951.$$

$$\tau_0 = \frac{1^0\cdot 1}{W'(1)} + \frac{2^0\cdot(-2)}{W'(2)} + \frac{3^0\cdot 33}{W'(3)} + \frac{4^0\cdot 166}{W'(4)} + \frac{5^0\cdot 481}{W'(5)} = 1,$$

$$\tau_1 = \frac{1^1\cdot 1}{W'(1)} + \frac{2^1\cdot(-2)}{W'(2)} + \frac{3^1\cdot 33}{W'(3)} + \frac{4^1\cdot 166}{W'(4)} + \frac{5^1\cdot 481}{W'(5)} = 15,$$

$$\tau_2 = 134,\ \tau_3 = 960,\ \tau_4 = 6117.$$

Формулы (1.9):

$$\begin{aligned}
1 &= A_0 \ \Rightarrow\ A_0 = 1,\\
15 &= 15\,A_0 + A_1 \ \Rightarrow\ A_1 = 0,\\
134 &= 140\,A_0 + 15\,A_1 + A_2 \ \Rightarrow\ A_2 = -6,\\
960 &= 1050\,A_0 + 140\,A_1 + 15\,A_2 + A_3 \ \Rightarrow\ A_3 = 0,\\
6117 &= 6951\,A_0 + 1050\,A_1 + 140\,A_2 + 15\,A_3 + A_4 \ \Rightarrow\ A_4 = 6.
\end{aligned}$$

Уравнение $x^4 - 6\,x^2 + 6 = 0$ легко решается подстановкой $X \stackrel{\text{def}}{=} x^2$.

Ответ. $\sqrt{3\pm\sqrt{3}},\ -\sqrt{3\pm\sqrt{3}}$.

Упражнение 42. *Рассмотрим задачу двумерной интерполяции. Пусть известны значения функции в узлах прямоугольной сетки:*

$$f(x_j, y_k) = z_{jk} \qquad \{j,k\} \subset \{1,\dots,n\}.$$

Доказать, что существует единственный полином $f(x,y)$, степень которого по каждой переменной не превосходит $n-1$, принимающий значения по этой таблице. Построить его выражение обобщением формы Лагранжа.

Интерполяционный полином в форме Ньютона

Основной недостаток построения интерп. полинома по методу (в форме) Лагранжа заключается в том, что при добавлении в таблицу (1.1) нового узла (новых результатов измерений), в формуле приходится пересчитывать все слагаемые. От этого недостатка свободен метод Ньютона, в котором добавление нового узла ведет к добавлению лишь одного слагаемого к построенному ранее полиному.

Для пояснения идеи вновь обратимся к случаю $n=3$ и в качестве стартового представления интерп. полинома воспользуемся его детерминантной формой (1.3). Преобразуем определитель, стоящий в числителе; с этой целью вычтем из третьего столбца второй, домноженный на x, а потом из второго столбца — первый, домноженный на x, получим

$$\begin{vmatrix} 1 & x_1-x & x_1^2-x_1x & -y_1 \\ 1 & x_2-x & x_2^2-x_2x & -y_2 \\ 1 & x_3-x & x_3^2-x_3x & -y_3 \\ 1 & 0 & 0 & 0 \end{vmatrix} = \begin{vmatrix} x_1-x & x_1^2-x_1x & y_1 \\ x_2-x & x_2^2-x_2x & y_2 \\ x_3-x & x_3^2-x_3x & y_3 \end{vmatrix} =$$

Далее, вычтем из второго столбца первый, домноженный на x_1:

$$= \begin{vmatrix} x_1-x & 0 & y_1 \\ x_2-x & (x_2-x)(x_2-x_1) & y_2 \\ x_3-x & (x_3-x)(x_3-x_1) & y_3 \end{vmatrix} =$$

Теперь вычтем первую строчку из второй и третьей:

$$= \begin{vmatrix} x_1-x & 0 & y_1 \\ x_2-x_1 & (x_2-x)(x_2-x_1) & y_2-y_1 \\ x_3-x_1 & (x_3-x)(x_3-x_1) & y_3-y_1 \end{vmatrix} =$$

и вынесем множители из строк:

$$= (x_2-x_1)(x_3-x_1)\times$$

$$\times \begin{vmatrix} x_1-x & 0 & y_1 \\ 1 & x_2-x & (y_2-y_1)/(x_2-x_1) \\ 1 & x_3-x & (y_3-y_1)/(x_3-x_1) \end{vmatrix} =$$

Из третьей строчки вычтем вторую:

$$= (x_2-x_1)(x_3-x_1)\times$$

$$\times \begin{vmatrix} x_1-x & 0 & y_1 \\ 1 & x_2-x & (y_2-y_1)/(x_2-x_1) \\ 0 & x_3-x_2 & (y_3-y_1)/(x_3-x_1)-(y_2-y_1)/(x_2-x_1) \end{vmatrix} =$$

Перейти к оглавлению на странице: 460

и снова вынесем множитель:

$$= (x_2 - x_1)(x_3 - x_1)(x_3 - x_2)\times$$

$$\times \begin{vmatrix} x_1 - x & 0 & y_1 \\ 1 & x_2 - x & (y_2 - y_1)/(x_2 - x_1) \\ 0 & 1 & \left[(y_3 - y_1)/(x_3 - x_1) - (y_2 - y_1)/(x_2 - x_1)\right]/(x_3 - x_2) \end{vmatrix}.$$

В результате интерп. полином получается в виде определителя, имеющего структуру

$$f(x) = \begin{vmatrix} x_1 - x & 0 & B_1 \\ 1 & x_2 - x & B_2 \\ 0 & 1 & B_3 \end{vmatrix},$$

а в случае произвольного n:

$$f(x) = \begin{vmatrix} x_1 - x & 0 & 0 & \dots 0 & 0 & B_1 \\ 1 & x_2 - x & 0 & \dots 0 & 0 & B_2 \\ & 1 & x_3 - x & \dots 0 & 0 & B_3 \\ \dots & & & \ddots & & \\ 0 & 0 & 0 & \dots 1 & x_{n-1} - x & B_{n-1} \\ 0 & 0 & 0 & \dots 0 & 1 & B_n \end{vmatrix},$$

который напоминает характеристический полином матрицы Фробениуса[6]. Раскладывая его рекурсивно по строкам, получаем для $n = 3$ интерп. полином в форме

$$f(x) = B_1 + (x - x_1)B_2 + (x - x_1)(x - x_2)B_3,$$

а в случае произвольного n — в форме

$$f(x) = B_1 + (x - x_1)B_2 + (x - x_1)(x - x_2)B_3 + \dots + \\ + (x - x_1)\times \dots \times (x - x_{n-1})B_n,$$

которая и называется **формой Ньютона**.

Приведем другой вывод этого представления, который даст нам алгоритм вычисления коэффициентов B_j. Рассмотрим последовательность интерполяционных таблиц

$$\begin{array}{c|c} x & x_1 \\ \hline y & y_1 \end{array}, \quad \begin{array}{c|c} x & x_1\; x_2 \\ \hline y & y_1\; y_2 \end{array}, \quad \dots, \quad \begin{array}{c|c} x & x_1 \dots x_{k-1} \\ \hline y & y_1 \dots y_{k-1} \end{array}, \quad \begin{array}{c|c} x & x_1 \dots x_{k-1}\; x_k \\ \hline y & y_1 \dots y_{k-1}\; y_k \end{array}$$

[6] Глава 4, упражнение 9.9.

Соответствующие интерп. полиномы обозначим

$$f_1(x), f_2(x), \dots, f_{k-1}(x), f_k(x), \quad \deg f_j \le j-1.$$

Очевидно $f_n(x)$ равен искомому интерп. полиному $f(x)$. Поскольку

$$f_{k-1}(x_1) = y_1 = f_k(x_1), \dots, f_{k-1}(x_{k-1}) = y_{k-1} = f_k(x_{k-1})$$

то разность $f_k(x) - f_{k-1}(x)$ должна делиться на $(x-x_1)$, ..., $(x-$ $-x_{k-1})$, а, следовательно, и на произведение

$$f_k(x) - f_{k-1}(x) \equiv q(x)(x-x_1)\dots(x-x_{k-1}).$$

Но $\deg(f_k(x) - f_{k-1}(x)) \le k-1$ и поэтому $q(x) = const \stackrel{\text{def}}{=} A_k$:

$$f_k(x) - f_{k-1}(x) \equiv A_k(x-x_1)\dots(x-x_{k-1}). \tag{1.10}$$

Для определения константы A_k используем последний узел k-й таблицы:

$$A_k = \frac{f_k(x_k) - f_{k-1}(x_k)}{(x_k-x_1)\dots(x_k-x_{k-1})} = \frac{y_k - f_{k-1}(x_k)}{(x_k-x_1)\dots(x_k-x_{k-1})}.$$

Таким образом, если полином $f_{k-1}(x)$ уже построен, то можно вычислить A_k, а тогда из (1.10) получить выражение для $f_k(x)$. Имеем представление интерп. полинома в форме Ньютона:

$$\begin{aligned} f(x) \;&\equiv\; f_n(x) = (f_n - f_{n-1}) + (f_{n-1} - f_{n-2}) + \dots + (f_2 - f_1) + f_1 = \\ &\stackrel{(1.10)}{=} A_n(x-x_1)\dots(x-x_{n-1}) + \dots + A_2(x-x_1) + A_1. \end{aligned} \tag{1.11}$$

Здесь величины $A_1, A_2, \dots, A_n$ удобно находить по методу неопределенных коэффициентов, последовательно подставляя $x = x_1$, x_2, ..., x_n:

$$\begin{aligned} &f(x_1) = A_1 = y_1, \\ &f(x_2) = A_1 + A_2(x_2-x_1) \Longrightarrow A_2 = \frac{y_2-y_1}{x_2-x_1}, \\ &f(x_3) = A_1 + A_2(x_3-x_1) + A_3(x_3-x_1)(x_3-x_2) \Longrightarrow \\ &\qquad\qquad \Longrightarrow A_3 = \frac{y_3-y_1}{(x_3-x_1)(x_3-x_2)} + \frac{y_2-y_1}{(x_2-x_1)(x_3-x_2)}, \\ &\quad \dots \end{aligned}$$

Перейти к оглавлению на странице: 460

Упражнение 43. *Построить интерполяционный полином наименьшей степени по данной таблице*

$$\text{а)}\ \begin{array}{c|cccc} x & 1 & 9/4 & 4 & 25/4 \\ \hline y & 1 & 3/2 & 2 & 5/2 \end{array}\ ,\ \textit{найти}\ f(2); \qquad \text{б)}\ \begin{array}{c|ccccc} x & 0 & 1 & 2 & \dots & n \\ \hline y & 1 & 2 & 4 & \dots & 2^n \end{array}$$

Замечание 1.1. Материалы настоящего параграфа позволяют заполнить последнюю строчку в “таблице соответствий” из упражнения 11.3 главы 3: найти целое число по его остаткам при делении на числа $M_1, \dots, M_k \longleftrightarrow$ найти полином по его остаткам от деления на линейные $x - x_1, \dots, x - x_k$.

§2. Приближенная интерполяция

В этом параграфе все числа предполагаются вещественными. Нам потребуются некоторые результаты из §11 главы 3.

Метод наименьших квадратов

Пусть из физических соображений можно считать (предполагать), что величины x и y связаны линейной зависимостью вида $y = kx + b$, а неизвестные коэффициенты k и b должны быть оценены экспериментально. Экспериментальные данные представляют собой $m > 1$ точек на координатной плоскости $(x_1, y_1), \dots, (x_m, y_m)$. Если бы эти опыты производились без погрешностей, то подстановка данных в уравнение приводила бы нас к системе из m линейных уравнений для двух неизвестных k и b:

$$y_1 = k\,x_1 + b, \dots, y_m = k\,x_m + b. \tag{2.1}$$

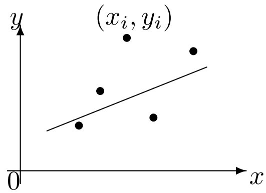
Из любой пары уравнений этой системы можно было бы однозначно определить коэффициенты k и b.

Однако, в реальной жизни опытов без погрешностей не бывает[7]. Для простоты будем предполагать, что величины $x_1, \dots, x_m$

[7]Письмо в редакцию: Дорогая редакция! Формулировку закона Ома следует уточнить следующим образом: “Если использовать тщательно отобранные и безупречно подготовленные исходные материалы, то при наличии некоторого навыка из них можно сконструировать электрическую цепь, для которой измерения отношения тока к напряжению, даже если они проводятся в течение ограниченного времени, дают значения, которые после введения соответствующих поправок оказываются равными постоянной величине”. — А.М.Б.Розен (Из книги “Физики шутят”М.Мир.1993).

известны точно, а им соответствующие $y_1, \dots, y_m$ — приближенно. Если $m > 2$, то при любых различных x_i и x_j пара точек (x_i, y_i) и (x_j, y_j) определяет прямую. Но другая пара точек определяет другую прямую, и у нас нет оснований выбрать какую-нибудь одну из всех прямых.

В этом случае то, что в действительности нужно, — это найти такую прямую на координатной плоскости, которая, может быть, не проходит ни через одну пару экспериментальных точек, или даже ни через одну из точек, но в каком-то смысле возможно более близко расположена ко всем точкам.



Обычно в этой задаче удаленность точки от прямой измеряют не расстоянием, а разностью ординат $k\,x_i + b - y_i$, и выбирают прямую так, чтобы сумма **квадратов** всех таких разностей была минимальна. Коэффициенты k_0 и b_0 уравнения этой прямой дают некоторое решение стоящей перед нами задачи, которое отнюдь не является решением системы линейных уравнений (2.1) (вообще говоря, несовместной).

Рассмотрим теперь обобщение предложенной задачи. Пусть искомая зависимость между величинами y и x полиномиальная:

$$y_1 = f(x_1), \dots, y_m = f(x_m), \qquad \text{где} \quad f(x) = a_0 + a_1 x + \dots + a_{n-1} x^{n-1} \tag{2.2}$$

Определение 2.1. Величина $\varepsilon_i = f(x_i) - y_i$ называется i-й **невязкой**. Уравнения (2.2) называются **условными**.

Задача. Подобрать неопределенные коэффициенты $a_0, \dots, a_{n-1}$ так, чтобы величина

$$\sum_{i=1}^{m} \varepsilon_i^2 = \sum_{i=1}^{m} [f(x_i) - y_i]^2 \tag{2.3}$$

была наименьшей.

Теорема 2.1. *Если $m \geq n$ и $x_1, \dots, x_m$ все различны, то существует единственный набор коэффициентов $a_0, \dots, a_{n-1}$, обеспе-*

чивающий минимальное значение для (2.3)*. Этот набор определяется из системы уравнений*

$$\underbrace{\begin{pmatrix} s_0 & s_1 & s_2 & \dots s_{n-2} & s_{n-1} \\ s_1 & s_2 & s_3 & \dots s_{n-1} & s_n \\ s_2 & s_3 & s_4 & \dots s_n & s_{n+1} \\ & \dots & & \dots & \\ s_{n-1} & s_n & s_{n+1} & \dots s_{2n-3} & s_{2n-2} \end{pmatrix}}_{S_{n\times n}} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \dots \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} t_0 \\ t_1 \\ t_2 \\ \dots \\ t_{n-1} \end{pmatrix} \qquad (2.4)$$

где $s_k \stackrel{\text{def}}{=} x_1^k + \dots + x_m^k$, $t_k \stackrel{\text{def}}{=} x_1^k y_1 + \dots + x_m^k y_m$*. (Здесь считаем* $0^0 \stackrel{\text{def}}{=} 1$*, так что* $s_0 = m$ *при любых* $x_1, \dots, x_m$*).*

Определение 2.2. Система (2.4) называется системой **нормальных** уравнений или просто **нормальной** системой.

Доказательство. Рассмотрим сумму (2.3) как полином от неопределенных коэффициентов a_j:

$$F(a_0, \dots, a_{n-1}) \stackrel{\text{def}}{=} \sum_{i=1}^m [f(x_i) - y_i]^2 = \sum_{i=1}^m [(a_0 + a_1 x_i + \ldots + a_{n-1} x_i^{n-1}) - y_i]^2.$$

На основании теоремы 11.8 главы 3 такая функция может принимать экстремальные значения только на вещественных решениях системы уравнений

$$\partial F/\partial a_0 = 0, \dots, \partial F/\partial a_{n-1} = 0. \qquad (2.5)$$

Распишем выражение для $\partial F/\partial a_j =$

$$= \sum_{i=1}^m 2\,[f(x_i) - y_i] \frac{\partial (a_0 + a_1 x_i + \dots + a_{n-1} x_i^{n-1})}{\partial a_j} = 2 \sum_{i=1}^m [f(x_i) - y_i] x_i^j =$$

$$= 2 \sum_{i=1}^m \left[\left(a_0 x_i^j + a_1 x_i^{j+1} + \dots + a_{n-1} x_i^{j+n-1} \right) - y_i x_i^j \right] =$$

$$= 2 \left[a_0 \sum_{i=1}^m x_i^j + a_1 \sum_{i=1}^m x_i^{j+1} + \dots + a_{n-1} \sum_{i=1}^m x_i^{j+n-1} - \sum_{i=1}^m y_i x_i^j \right] =$$

$$= 2[a_0\, s_j + a_1\, s_{j+1} + \dots + a_{n-1}\, s_{j+n-1} - t_j]$$

Таким образом, условия (2.5) можно переписать в виде системы (2.4).

Покажем теперь, что матрица этой системы имеет ненулевой определитель. Действительно, матрица S — ганкелева. При $m=n$ ее определитель был нами вычислен в §9.2 главы 4 с помощью теоремы Бине—Коши:

$$\det S = \prod_{1\le j<k\le n} (x_k - x_j)^2.$$

Воспользуемся той же теоремой и для случая $m > n$:

$$S = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ x_1 & x_2 & x_3 & \dots & x_m \\ \dots & & & & \dots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \dots & x_m^{n-1} \end{pmatrix} \cdot \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \dots & & & & \dots \\ 1 & x_m & x_m^2 & \dots & x_m^{n-1} \end{pmatrix}$$

$$\det S = \sum_{1\le j_1<j_2<\dots<j_n\le m} \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_{j_1} & x_{j_2} & \dots & x_{j_n} \\ \dots & & & \dots \\ x_{j_1}^{n-1} & x_{j_2}^{n-1} & \dots & x_{j_n}^{n-1} \end{vmatrix} \cdot \begin{vmatrix} 1 & x_{j_1} & \dots & x_{j_1}^{n-1} \\ 1 & x_{j_2} & \dots & x_{j_2}^{n-1} \\ \dots & & & \dots \\ 1 & x_{j_n} & \dots & x_{j_n}^{n-1} \end{vmatrix} =$$

$$= \sum_{1\le j_1<j_2<\dots<j_n\le m} \ \prod_{1\le L<K\le n} (x_{j_K} - x_{j_L})^2.$$

Каждое слагаемое неотрицательно и отлично от нуля поскольку, по предположению, все x_j различны. Поэтому $\det S > 0$. По теореме Крамера система (2.4) имеет единственное решение.

Мы оставляем без доказательства тот факт, что это решение доставляет функции F именно значение минимума, причем глобального. ■

Упражнение 44. *Доказать, что при $m = n$ полином, построенный по методу наименьших квадратов, совпадает с обычным интерполяционным.*

Пример 2.1. Определить по методу наименьших квадратов полином второй степени, принимающий таблицу значений

x	0	1	2	3	4
y	1	1.4	2	2.7	3.6

Решение. Имеем: $n = 3$

$$s_0 = 5,\ s_1 = 0+1+2+3+4 = 10,\ s_2 = 0+1^2+2^2+3^2+4^2 = 30,$$
$$s_3 = 100,\ s_4 = 354,$$

$$t_0 = 1+1.4+2+2.7+3.6 = 10.7,\ t_1 = 0\cdot1+1\cdot1.4+2\cdot2+3\cdot2.7+4\cdot3.6 = 27.9,$$
$$t_2 = 0\cdot1+1^2\cdot1.4+2^2\cdot2+3^2\cdot2.7+4^2\cdot3.6 = 91.3.$$

Решаем нормальную систему:

$$\begin{pmatrix} 5 & 10 & 30 \\ 10 & 30 & 100 \\ 30 & 100 & 354 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 10.7 \\ 27.9 \\ 91.3 \end{pmatrix}$$

Ответ. $f(x) = {}^1/_7 \left(6.98 + 2.35\,x + 0.55\,x^2\right)$.

Упражнение 45. *Определить по методу наименьших квадратов полином третьей степени, принимающий таблицу значений*

x	-2	-1	0	1	2
y	16	1	0	1	16

Упражнение 46. *Определить главную квадратичную тенденцию (тренд) изменения курса доллара по отношению к рублю по временному ряду среднемесячных курсов*

Месяц	*ноябрь*	*декабрь*	*январь*	*февраль*	*март*
руб. за доллар	26.7	26.4	26.4	26.3	26.4

и спрогнозировать его курс на апрель[8].

Замечание 2.1. Чем оправдан выбор критерия (2.3) в виде минимума суммы **квадратов** невязок, а не суммы , например, их первых или третьих или четвертых степеней? С одной стороны, это диктуется соображениями простоты вычислений. (Так, например,

[8]Подсказка. Вычисления упрощаются при грамотной пронумеровке месяцев: см. решения двух предыдущих упражнений.

$|x|$ является недифференцируемой функцией при $x = 0$.) Однако более глубинной причиной является следующая, которую мы снова проиллюстрируем на примере из начала параграфа. Пусть физический процесс описывается линейным законом $y = k\,x + b$ с неизвестными нам параметрами k и b, и мы пытаемся определить эти параметры с помощью серии из m опытов, испытаний. Несовершенство экспериментальной аппаратуры приводит к тому, что вместо точных значений y_j нам приходится иметь дело с их приближенными значениями. Точки (x_j, y_j) оказываются не на искомой прямой, а располагаются случайным образом в некоторой ее окрестности. Результат напоминает стрельбу по мишени с изображенной на ней целью — нашей прямой. Теперь представим себе, что после произведения m выстрелов мишень снимается и предъявляется наблюдателю *тыльной* стороной, на которой видны только пулевые отверстия, но не видна сама цель. И наблюдателю нужно по расположению этих отверстий угадать эту цель. Наблюдатель может даже не знать того факта, что стрельба производилась именно по прямой, а не по более сложной цели. Но, предположим, что эта информация ему известна, или же составляет содержание его *гипотезы*. Какими словами описать тот результат, что должен быть им получен? Как **максимально правдоподобный** при данном наборе испытаний и данной гипотезе. Из всех возможных прямых мы должны выбрать максимально правдоподобную. Оказывается, что именно выбор **квадратов** (а не первых, или третьих или четвертых и т.д степеней) невязок гарантирует получение такого результата. Этот факт был установлен великим немецким математиком Гауссом.

Историческая справка

Гаусс Карл Фридрих (Gauß, или Gauss Carl Friedrich, 1777–1855) родился в Брауншвейге в семье чернорабочего. Ему было 12 лет, когда разразилась Великая французская революция, 29, когда была распущена казавшаяся вечной Священная Римская империя, 38, когда был разгромлен Наполеон, и за 70, когда в самой Германии произошла революция 1848 года.

Математические способности проявились еще в школе, и, к счастью, получили поддержку учителей. Помощник учителя Мартин Бартельс, который впоследствии стал в Казани университетским профессором Н.И.Лобачевского, добился того, что герцог Брауншвейгский оказал Гауссу материальную поддержку для продолжения образования. Первый успех пришел к Г. в 19 лет: он доказал возможность построения правильного 17-угольника циркулем и линейкой. Для $\cos 2\pi/17$ он приводит следующее явное выражение:

$$-\frac{1}{16}+\frac{\sqrt{17}}{16}+\frac{1}{16}\sqrt{34-2\sqrt{17}}+\frac{1}{8}\sqrt{17+3\sqrt{17}-\sqrt{34-2\sqrt{17}}-2\sqrt{34+2\sqrt{17}}}.$$

В 1799 г. он публикует (также на средства герцога) свою диссертацию: она посвящена доказательству основной теоремы высшей алгебры. Примерно с этого же года его интересы устремляются одновременно и в теорию чисел и в астрономию. Мода на астрономию в те времена была огромной — каждое государство считало долгом престижа построить обсерваторию. Гаусс получил европейскую известность после того как в 1801 г. подтвердился его прогноз орбиты Цереры. Его приглашают в Санкт–Петербург на должность директора обсерватории. Брауншвейг не мог предложить Г. много в этом отношении, но в 1802–1803 годах были начаты серьезные переговоры о строительстве для него маленькой обсерватории. Благодаря улучшению условий, Г. решил остаться на родине.

Цитата из гауссовской "Вводной лекции по астрономии", в которой он полемизирует с утилитарным взглядом на науку, осуждая тех, кто требует от нее немедленной выгоды:

> *"Подобные суждения свидетельствуют не только о нашем убожестве, но в равной мере и о мелком, узком и ленивом мышлении; они выдают тенденцию всегда трусливо вычислять наперед плату за каждое усилие, холодное сердце и нечувствительность ко всему великому и почетному для человека. К сожалению, нельзя закрывать глаза на то, что такой образ мыслей очень распространен в наш век, и совершенно несомненно, что этот образ мыслей находится в очень тесной связи с тем несчастьем, которое поразило многие страны в последнее время; поймите меня правильно, я говорю не о столь частом недостатке, как простое непонимание науки, а об источнике, из которого все это вытекает, о тенденции всегда сначала спрашивать о выгоде и все соотносить с физическим благополучием, о безразличии к великим идеям, об отвращении ко всякому усилию, порожденному чистым энтузиазмом: я думаю, что такая тенденция, если она одерживает верх, может играть решающую роль в катастрофах того рода, что мы пережили."*

В 1818–1832 годах большое место в жизни Г. занимал проект геодезического исследования Ганноверского государства. Гаусс лично руководил работами. Метод наименьших квадратов был главным средством обработки наблюдений. Из письма к астроному Ольберсу от 1822 г.:

> *"Я на днях, после трех месяцев упорной работы, закончил по методу наименьших квадратов уравнительные вычисления, заключавшие около 300 условных уравнений с 55 неизвестными."*

Трудно представить себе сегодня тот объем вычислений, который приходилось производить Г. для своей научной работы. По его собственной оценке, ему пришлось иметь дело более чем с миллионом чисел в одном только геодезическом исследовании. Вследствие непрестанных и неутомимых вычислений, многие числа обладали для Г. индивидуальностью, почти как живые существа.

Г. никогда не был политически активен и даже воздерживался от открытых политических высказываний. Он никогда не расставался со своими юношескими впечатлениями того, как великодушие государя позволило ему стать ученым и помогло оставить позади узкий и безнадежный мир, в котором он родился. Г. не был атеистом, но, судя по его письмам, не верил в Бога как в личность. Принципиальной частью его веры была убежденность в гармонии и целостности общего проекта творения. Математика была ключом к усилиям человека получить хотя бы приблизительное представление о плане Бога.

Псевдорешения линейной системы

Рассмотрим теперь обобщение задачи предыдущего пункта. В практических задачах часто бывает нужно найти решение, удовлетворяющее большому числу возможно противоречивых требований. Если такая задача сводится к системе линейных уравнений

$$\left\{\begin{array}{ll} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = b_2 \\ \ldots & \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = b_m \end{array}\right. \iff AX = \mathcal{B} \qquad (2.6)$$

при числе уравнений m большем числа неизвестных n, то такая **переопределенная** система, как правило, несовместна. В этом случае задача может быть решена только путем выбора некоторого компромисса — все требования могут быть удовлетворены не полностью, а лишь до некоторой степени.

Определение 2.3. **Псевдорешением** системы (2.6) называется столбец X, обеспечивающий минимум величины

$$\sum_{i=1}^{m} [b_i - (a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n)]^2. \qquad (2.7)$$

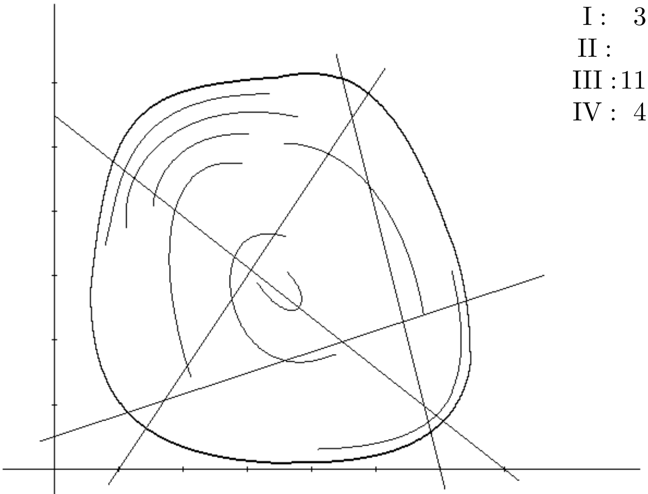
Замечание 2.2. Такому определению можно также соотнести вероятностную интерпретацию. Пусть для определения неизвестных величин $x_1, \ldots, x_n$ проводятся m экспериментов, описываемых линейными уравнениями:

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i \quad \text{при } i \in \{1, \ldots, m\}.$$

При этом величины a_{ij} — известные постоянные, не подверженные сопутствующим экспериментам (наблюдениям) погрешностям, а вот величины b_i этим погрешностям подвержены. Формально каждое из равенств следует рассматривать как приближенное. Понятно, что при таких обстоятельствах не имеет смысла гоняться за *точным* решением системы (2.6) (его может и не существовать вовсе!). Искать следует *приближенное* решение, оптимальное в некотором смысле. Оказывается, что именно выбор критерия в виде (2.7) обеспечивает то, что псевдорешение дает **максимально правдоподобные** значения неизвестных величин $x_1,\dots,x_n$.

Упражнение 47. *На дубовой колоде лежит мелкая монетка. К колоде по очереди подходят 4 рыцаря и каждый наносит удар мечом, стараясь попасть по монетке. Все промахиваются. Расстроенные, рыцари уходят в харчевню пропивать злосчастную монетку. Укажите максимально правдоподобное ее расположение, имея перед глазами зарубки:*

$$
\begin{array}{rrcrcr}
\mathrm{I}: & 3x- & 2y & = & 6, \\
\mathrm{II}: & x- & 3y & = & -3, \\
\mathrm{III}: & 11x+ & 14y & = & 154, \\
\mathrm{IV}: & 4x+ & y & = & 48.
\end{array}
$$



Какой из ударов можно считать наиболее удачным?

Теорема 2.2. *Существует псевдорешение системы* (2.6) *и оно* *является решением системы*

$$\left[A^\top A\right] X = A^\top \mathcal{B}. \tag{2.8}$$

Это решение будет единственным тогда и только тогда, когда $\operatorname{rank} A = n$.

Определение 2.4. Система (2.8) называется **нормальной** системой по отношению к системе (2.6). Формально она получается домножением системы (2.6) слева на матрицу $A^\top$.

В смысле этого определения система (2.4) была нормальной по отношению к системе условий (2.2), рассматриваемых относительно неизвестных $a_0, \ldots, a_{n-1}$. Заметим также, что если $m = n$ и $\det A \neq 0$, то псевдорешение системы (2.6) совпадает с ее обычным решением.

Доказательство. Как обычно, $A^{[i]}$ обозначает i-ю строчку матрицы A, а $A_{[j]}$ — ее j-й столбец. На основании теоремы 11.8 главы 3 точка экстремума функции

$$F(X) \stackrel{\text{def}}{=} \sum_{i=1}^{m} [(a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n) - b_i]^2 = \sum_{i=1}^{m} [A^{[i]}X - b_i]^2$$

ищется из условий

$$\partial F/\partial x_1 = 0, \ldots, \partial F/\partial x_n = 0. \tag{2.9}$$

Распишем выражение для $\partial F/\partial x_j =$

$$= \sum_{i=1}^{m} 2[A^{[i]}X - b_i] \frac{\partial(a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n)}{\partial x_j} = 2\sum_{i=1}^{m} [A^{[i]}X - b_i]a_{ij} =$$

$$= 2\left[\left(a_{1j}A^{[1]} + \cdots + a_{mj}A^{[m]}\right) X - (a_{1j}b_1 + \cdots + a_{mj}b_m)\right] =$$

$$= 2\left[A_{[j]}^\top AX - A_{[j]}^\top \mathcal{B}\right].$$

Таким образом, условия (2.9) эквивалентны следующим

$$A_{[1]}^\top AX = A_{[1]}^\top \mathcal{B}, \ldots, A_{[n]}^\top AX = A_{[n]}^\top \mathcal{B} \iff A^\top AX = A^\top \mathcal{B}.$$

Итак, если существует псевдорешение системы (2.6), то оно обязательно должно быть (обычным) решением системы (2.8).

Покажем теперь, что система (2.8) всегда совместна. Предположим сначала, что $\operatorname{rank} A = n$. Для доказательства единственности решения системы (2.8) в этом случае, вычислим определитель матрицы $A^\top A$ с помощью теоремы Бине–Коши[9]. Если $m = n$, то

$$\det(A^\top A) = \det A^\top \det A = (\det A)^2.$$

Если же $m > n$, то $\det(A^\top A) =$

$$= \sum_{1 \le j_1 < \dots < j_n \le m} \begin{vmatrix} a_{j_1 1} & a_{j_2 1} & \dots & a_{j_n 1} \\ a_{j_1 2} & a_{j_2 2} & \dots & a_{j_n 2} \\ \dots & & & \dots \\ a_{j_1 n} & a_{j_2 n} & \dots & a_{j_n n} \end{vmatrix} \cdot \begin{vmatrix} a_{j_1 1} & a_{j_1 2} & \dots & a_{j_1 n} \\ a_{j_2 1} & a_{j_2 2} & \dots & a_{j_2 n} \\ \dots & & & \dots \\ a_{j_n 1} & a_{j_n 2} & \dots & a_{j_n n} \end{vmatrix} =$$

$$= \sum_{1 \le j_1 < \dots < j_n \le m} \begin{vmatrix} a_{j_1 1} & a_{j_1 2} & \dots & a_{j_1 n} \\ a_{j_2 1} & a_{j_2 2} & \dots & a_{j_2 n} \\ \dots & & \dots & \\ a_{j_n 1} & a_{j_n 2} & \dots & a_{j_n n} \end{vmatrix}^2 \ge 0. \qquad (2.10)$$

По предположению $\operatorname{rank} A = n$. Для случая $m = n$ это означает, что $\det A \ne 0$, но тогда и $\det(A^\top A) > 0$. Для случая $m > n$ то же условие означает существование у матрицы A хотя бы одного минора порядка n отличного от нуля. Соответствующее слагаемое в (2.10) строго положительно, и снова $\det(A^\top A) > 0$. По теореме Крамера, система (2.8) имеет единственное решение.

Пусть теперь $\operatorname{rank} A = \mathfrak{r} < n$. Такое может произойти, например, когда $m < n$. Докажем, сначала, что $\operatorname{rank}(A^\top A) = \mathfrak{r}$. Теорема 11.8 (Сильвестра) из главы 4 утверждает, что $\operatorname{rank}(A^\top A) \le$ $\le \operatorname{rank} A = \mathfrak{r}$. Установим теперь существование ненулевого минора порядка $\mathfrak{r}$ для матрицы $C \stackrel{\text{def}}{=} A^\top A$. Предположим для определенности, что ненулевой минор порядка $\mathfrak{r}$ матрицы A находится в ее столбцах с номерами $1, \dots, \mathfrak{r}$. Тогда

$$C \begin{pmatrix} 1 \dots \mathfrak{r} \\ 1 \dots \mathfrak{r} \end{pmatrix} = \det \left(\begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & & & \dots \\ a_{1\mathfrak{r}} & a_{2\mathfrak{r}} & \dots & a_{m\mathfrak{r}} \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1\mathfrak{r}} \\ a_{21} & a_{22} & \dots & a_{2\mathfrak{r}} \\ \dots & & & \dots \\ a_{m1} & a_{m2} & \dots & a_{m\mathfrak{r}} \end{bmatrix} \right).$$

[9] Глава 4, §7

Применяем теорему Бине–Коши, и, рассуждая по аналогии с предыдущей частью доказательства, придем к заключению, что данный минор матрицы $A^\top A$ отличен от нуля. Итак, $\operatorname{rank}(A^\top A)=\mathfrak{r}$.

Вычислим теперь ранг расширенной матрицы системы (2.8). С одной стороны:

$$\mathfrak{r} = \operatorname{rank}\left(A^\top A\right) \le \operatorname{rank}\left[A^\top A \mid A^\top \mathcal{B}\right].$$

С другой стороны, на основании теоремы 11.8 (Сильвестра) из главы 4, имеем:

$$\operatorname{rank}\left[A^\top A \mid A^\top \mathcal{B}\right] = \operatorname{rank} A^\top \left[A \mid \mathcal{B}\right] \le \operatorname{rank} A^\top = \mathfrak{r}.$$

Два неравенства приводят к заключению: расширенная матрица системы (2.8) имеет ранг $\mathfrak{r}$. На основании теоремы Кронекера — Капелли делаем вывод: в этом случае система (2.8) совместна и имеет бесконечное множество решений.

Мы оставляем без доказательства тот факт, что любое решение системы (2.8) доставляет функции F именно значение минимума, причем глобального. ■

Замечание 2.3. Если система (2.8) имеет бесконечное количество решений, то обычно в качестве псевдорешения берут какое-то одно из них — как правило то, у которого минимальна сумма квадратов компонент (“длина”).

Пример 2.2. Найти псевдорешение системы

$$x_1 + x_2 = 2,\ x_1 - x_2 = 0,\ 2x_1 + x_2 = 2.$$

Решение. Имеем:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 1 \end{pmatrix},\ \operatorname{rank} A = 2,\ \mathcal{B} = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix},\ A^\top A = \begin{pmatrix} 6 & 2 \\ 2 & 3 \end{pmatrix},\ A^\top \mathcal{B} = \begin{pmatrix} 6 \\ 4 \end{pmatrix}.$$

Упражнение 48. *Показать, что матрица $A^\top A$ всегда симметрична.*

Упражнение 49. *Найти псевдорешения систем линейных уравнений:*

а) $$\begin{cases} 27\,x_1 - 55\,x_2 = 1, \\ -13\,x_1 + 27\,x_2 = 1, \\ -14\,x_1 + 28\,x_2 = 1; \end{cases}$$ **б)** $$\begin{cases} x_1 + x_2 + x_3 + x_4 = 2, \\ x_1 + x_2 + x_3 + x_4 = 3, \\ x_1 + x_2 + x_3 + x_4 = 4; \end{cases}$$

в) $$\begin{cases} -x_1 - 2x_2 = 1, \\ 2\,x_1 + 4\,x_2 = 0, \\ x_1 + 2\,x_2 = 0, \\ 3\,x_1 + 6\,x_2 = 0. \end{cases}$$

Перейти к оглавлению на странице: 460

ГЛАВА 6. ТЕОРИЯ ИСКЛЮЧЕНИЯ

Мы вновь возвращаемся к основной задаче раздела: решению уравнений и систем уравнений. Задача решения системы нелинейных алгебраических уравнений от нескольких переменных будет рассмотрена в настоящей главе на примере системы из двух уравнений от двух переменных. Собственно эта задача будет решена в §5, а в предваряющих его параграфах будет разработана технология — в приложении к одной, уже решенной в §4 главы 3, задаче для полиномов от одной переменной.

§1. Результант

Определение

Пример 1.1. Найти условие, при котором полиномы

$$f(x) = a_0x^5+a_1x^4+a_2x^3+a_3x^2+a_4x+a_5 \text{ и } g(x) = b_0x^3+b_1x^2+b_2x+b_3$$

$(a_0 \neq 0, b_0 \neq 0)$ из $\mathbb{A}[x]$ имеют общий корень.

Решение. Пусть f и g имеют корень $x=c \in \mathbb{C}$: $f(c)=0$, $g(c)=0$. Тогда

$$\begin{array}{rlr} c^2f(c) = & a_0c^7+a_1c^6+a_2c^5+a_3c^4+a_4c^3+a_5c^2 & = 0, \\ cf(c) = & a_0c^6+a_1c^5+a_2c^4+a_3c^3+a_4c^2+a_5c & = 0, \\ f(c) = & a_0c^5+a_1c^4+a_2c^3+a_3c^2+a_4c+a_5 & = 0, \\ g(c) = & b_0c^3+b_1c^2+b_2c+b_3 & = 0, \\ cg(c) = & b_0c^4+b_1c^3+b_2c^2+b_3c & = 0, \\ c^2g(c) = & b_0c^5+b_1c^4+b_2c^3+b_3c^2 & = 0, \\ c^3g(c) = & b_0c^6+b_1c^5+b_2c^4+b_3c^3 & = 0, \\ c^4g(c) = & b_0c^7+b_1c^6+b_2c^5+b_3c^4 & = 0. \end{array} \qquad (1.1)$$

Запишем эти равенства как систему линейных уравнений

$$\begin{array}{r} 3\left\{\begin{array}{l} \\ \\ \\ \end{array}\right. \\ 5\left\{\begin{array}{l} \\ \\ \\ \\ \\ \end{array}\right. \end{array} \underbrace{\begin{pmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & & \\ & a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & \\ & & a_0 & a_1 & a_2 & a_3 & a_4 & a_5 \\ & & & & b_0 & b_1 & b_2 & b_3 \\ & & & b_0 & b_1 & b_2 & b_3 & \\ & & b_0 & b_1 & b_2 & b_3 & & \\ & b_0 & b_1 & b_2 & b_3 & & & \\ b_0 & b_1 & b_2 & b_3 & & & & \end{pmatrix}}_{M_{8\times 8}} \underbrace{\begin{pmatrix} c^7 \\ c^6 \\ c^5 \\ c^4 \\ c^3 \\ c^2 \\ c \\ 1 \end{pmatrix}}_{X} = \mathbb{O}_{8\times 1} \tag{1.2}$$

относительно столбца неизвестных $X = \left[c^7, c^6, c^5, c^4, \ldots, 1\right]^\top$ (неуказанные элементы матрицы M считаются равными нулю). Эта система однородная и имеет нетривиальное решение (последняя компонента вектора X равна единице). Следовательно[1], определитель ее матрицы равен нулю: $\det M = 0$. Это условие является необходимым для существования общего корня у полиномов f и g. $\triangle$

Для случая произвольных полиномов

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n \quad \text{и} \quad g(x) = b_0x^m + b_1x^{m-1} + \ldots + b_m$$

из $\mathbb{A}[x]$ $(a_0 \neq 0,\ b_0 \neq 0)$ составим квадратную матрицу порядка $m+n$:

$$M = \begin{pmatrix} a_0 & a_1 & a_2 & \ldots & \ldots & a_n & 0 & \ldots & 0 & 0 \\ 0 & a_0 & a_1 & \ldots & \ldots & a_{n-1} & a_n & \ldots & 0 & 0 \\ \vdots & & \ddots & & & & & & \ddots & \\ 0 & 0 & \ldots & a_0 & \ldots & \ldots & & \ldots & a_{n-1} & a_n \\ 0 & 0 & \ldots & & b_0 & b_1 & \ldots & \ldots & b_{m-1} & b_m \\ 0 & 0 & \ldots & b_0 & b_1 & \ldots & & \ldots & b_m & 0 \\ \vdots & & \iddots & \ldots & & & & \iddots & & \vdots \\ 0 & b_0 & \ldots & \ldots & & b_m & \ldots & & \ldots & 0 \\ b_0 & \ldots & \ldots & & b_m & 0 & \ldots & & & 0 \end{pmatrix} \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} m \\ \left.\begin{array}{l} \\ \\ \\ \\ \\ \end{array}\right\} n \end{array}, \tag{1.3}$$

элементы выше a_n и b_0, и ниже a_0 и b_m все равны нулю.

[1] Следствие 1 к теореме 12.2 главы 4

Определение 1.1. Выражение

$$\mathcal{R}(f,g) \stackrel{\mathrm{def}}{=} (-1)^{n(n-1)/2} \det M \tag{1.4}$$

называется **результантом** (**в форме Сильвестра**) полиномов f и g.

По построению, результант является полиномом относительно коэффициентов $a_0,\dots,a_n,b_0,\dots,b_m$:

$$\mathcal{R}(a_0x^n+\dots+a_n,\ b_0x^m+\dots+b_m) \in \mathbb{Z}[a_0,\dots,a_n,b_0,\dots,b_m].$$

Упражнение 50. *Доказать равенство*

$$\begin{aligned}\mathcal{R}(a_0x^2+a_1x+a_2, b_0x^2+b_1x+b_2) =\\ =(a_0b_2-a_2b_0)^2-(a_0b_1-a_1b_0)(a_1b_2-a_2b_1).\end{aligned} \tag{1.5}$$

Теорема 1.1. *Для того чтобы f и g имели общий корень, необходимо и достаточно выполнение условия $\mathcal{R}(f,g)=0$.*

Доказательство. Нам осталось показать достаточность. Идею доказательства проиллюстрируем на примере 1.1. Пусть

$$f(x) \equiv a_0(x-\lambda_1)\times\dots\times(x-\lambda_5),\ g(x)\equiv b_0(x-\mu_1)(x-\mu_2)(x-\mu_3).$$

Дополнительно предположим, что все корни $\lambda_1,\dots,\lambda_5$ полинома $f(x)$ различны и все корни μ_1,μ_2,μ_3 полинома $g(x)$ различны.

Домножим матрицу M справа на матрицу

$$V=\begin{pmatrix} \mu_1^7 & \mu_2^7 & \mu_3^7 & \lambda_1^7 & \dots & \lambda_5^7 \\ \mu_1^6 & \mu_2^6 & \mu_3^6 & \lambda_1^6 & \dots & \lambda_5^6 \\ \dots & & & & & \dots \\ \mu_1 & \mu_2 & \mu_3 & \lambda_1 & \dots & \lambda_5 \\ 1 & 1 & 1 & 1 & \dots & 1 \end{pmatrix}_{8\times 8}.$$

Результатом умножения будет следующая матрица:

$$L=\begin{pmatrix} \mu_1^2 f(\mu_1) & \mu_2^2 f(\mu_2) & \mu_3^2 f(\mu_3) & 0 & \dots & 0 \\ \mu_1 f(\mu_1) & \mu_2 f(\mu_2) & \mu_3 f(\mu_3) & 0 & \dots & 0 \\ f(\mu_1) & f(\mu_2) & f(\mu_3) & 0 & \dots & 0 \\ 0 & 0 & 0 & g(\lambda_1) & \dots & g(\lambda_5) \\ 0 & 0 & 0 & \lambda_1 g(\lambda_1) & \dots & \lambda_5 g(\lambda_5) \\ 0 & 0 & 0 & \lambda_1^2 g(\lambda_1) & \dots & \lambda_5^2 g(\lambda_5) \\ 0 & 0 & 0 & \lambda_1^3 g(\lambda_1) & \dots & \lambda_5^3 g(\lambda_5) \\ 0 & 0 & 0 & \lambda_1^4 g(\lambda_1) & \dots & \lambda_5^4 g(\lambda_5) \end{pmatrix}.$$

В получившемся матричном равенстве перейдем к определителям:

$$M \cdot V = L \;\Rightarrow\; \det M \det V = \det L. \tag{1.6}$$

Найдем выражения для $\det V$ и для $\det L$. Матрица V представляет собой матрицу Вандермонда (с точностью до перестановки строк). Как известно, ее определитель равен произведению всевозможных разностей, составленных из чисел $\lambda_1, \dots, \lambda_5, \mu_1, \mu_2, \mu_3$:

$$\det V = \left[\prod_{1\le L<K\le 3} (\mu_K - \mu_L)\right] \left[\prod_{1\le \ell<k\le 5} (\lambda_k - \lambda_\ell)\right] \left[\prod_{k=1}^{3}\prod_{j=1}^{5} (\mu_k - \lambda_j)\right].$$

Далее, матрица L блочно-диагональная, и

$$\det L = \begin{vmatrix} \mu_1^2 f(\mu_1) & \mu_2^2 f(\mu_2) & \mu_3^2 f(\mu_3) \\ \mu_1 f(\mu_1) & \mu_2 f(\mu_2) & \mu_3 f(\mu_3) \\ f(\mu_1) & f(\mu_2) & f(\mu_3) \end{vmatrix} \cdot \begin{vmatrix} g(\lambda_1) \dots & g(\lambda_5) \\ \lambda_1 g(\lambda_1) \dots & \lambda_5 g(\lambda_5) \\ \lambda_1^2 g(\lambda_1) \dots & \lambda_5^2 g(\lambda_5) \\ \lambda_1^3 g(\lambda_1) \dots & \lambda_5^3 g(\lambda_5) \\ \lambda_1^4 g(\lambda_1) \dots & \lambda_5^4 g(\lambda_5) \end{vmatrix} =$$

$$= f(\mu_1) f(\mu_2) f(\mu_3) \begin{vmatrix} \mu_1^2 & \mu_2^2 & \mu_3^2 \\ \mu_1 & \mu_2 & \mu_3 \\ 1 & 1 & 1 \end{vmatrix} g(\lambda_1) \times \dots \times g(\lambda_5) \begin{vmatrix} 1 \dots & 1 \\ \lambda_1 \dots & \lambda_5 \\ \lambda_1^2 \dots & \lambda_5^2 \\ \lambda_1^3 \dots & \lambda_5^3 \\ \lambda_1^4 \dots & \lambda_5^4 \end{vmatrix} =$$

(обращаем внимание на то, что определители — снова типа Вандермонда)

$$= -\left[\prod_{1\le L<K\le 3} (\mu_K - \mu_L)\right] \cdot \left[\prod_{1\le \ell<k\le 5} (\lambda_k - \lambda_\ell)\right] \times$$
$$\times f(\mu_1) f(\mu_2) f(\mu_3) g(\lambda_1) \times \dots \times g(\lambda_5).$$

Итак, это — величина правой части равенства (1.6), а величина левой, с точностью до знака, равна $\mathcal{R}(f,g) \cdot \det V$. Если $\mathcal{R}(f,g) = 0$, то хотя бы одно из чисел

$$f(\mu_1),\ f(\mu_2),\ f(\mu_3),\ g(\lambda_1), \dots, g(\lambda_5)$$

должно обратиться в нуль (поскольку, по предположению, $\lambda_k \ne \lambda_\ell, \mu_K \ne \mu_L$). ■

Замечание 1.1. Используя выведенное в ходе доказательства выражение для $\det V$, можем получить из равенства (1.6) явное представление результанта $\mathcal{R}(f,g)$ через корни обоих полиномов. Действительно, сократив[2] общий множитель, получим:

$$\mathcal{R}(f,g) = a_0^m b_0^n \prod_{k=1}^{m} \prod_{j=1}^{n} (\lambda_j - \mu_k), \tag{1.7}$$

где $\lambda_1, \ldots, \lambda_n$ — корни полинома $f(x)$, а $\mu_1, \ldots, \mu_m$ — корни полинома $g(x)$. Таким образом, величина $\mathcal{R}(f,g)$ характеризует расстояние между множествами корней полинома $f(x)$ и полинома $g(x)$. В свою очередь, формулу (1.7) можно переписать в эквивалентных видах:

$$\mathcal{R}(f,g) = a_0^m \prod_{j=1}^{n} g(\lambda_j) = \tag{1.8}$$

$$= (-1)^{mn} b_0^n \prod_{k=1}^{m} f(\mu_k). \tag{1.9}$$

Иногда в литературе равенство (1.8) берут в качестве формального определения результанта. В частном случае $g(x) \equiv const$ из нее следует:

$$\mathcal{R}(f, const) = (const)^n. \tag{1.10}$$

Упражнение 51. *Доказать, что* $\mathcal{R}(f,g) = (-1)^{nm}\mathcal{R}(g,f)$.

Упражнение 52. *Доказать, что*

$$\mathcal{R}\,(f_1 \cdot f_2,\, g) = \mathcal{R}(f_1,\, g) \cdot \mathcal{R}(f_2,\, g).$$

Пример 1.2. Пусть

$$f^*(x) \stackrel{\text{def}}{=} x^n f(1/x) = a_n x^n + \ldots + a_0, \quad g^*(x) \stackrel{\text{def}}{=} x^m g(1/x) = b_m x^m + \ldots + b_0,$$

и числа a_0, a_n, b_0, b_m отличны от нуля. Доказать, что

$$\mathcal{R}(f^*, g^*) = (-1)^{mn}\mathcal{R}(f,g).$$

[2] Что, строго говоря, не всегда допустимо... Тем не менее, мы просим принять на веру, что следующая формула будет верна всегда!

Решение. Корнями полинома f^* являются числа $1/\lambda_j$. Используя равенство (1.8) и формулу Виета, получаем

$$\mathcal{R}(f^*,g^*) = a_n^m \prod_{j=1}^{n} g^*\left(\frac{1}{\lambda_j}\right) = a_n^m \prod_{j=1}^{n}\left(\frac{1}{\lambda_j}\right)^m \prod_{j=1}^{n} g(\lambda_j) =$$

$$= (-1)^{mn} a_0^m \prod_{j=1}^{n} g(\lambda_j) = (-1)^{mn}\mathcal{R}(f,g).$$

$\triangle$

Упражнение 53. *Вычислить результант полиномов:*
а) $3\,x^2+x-2$ *и* $x^2-2\,x-2$;
б) $x^3-3\,x+6$ *и* x^3+x^2-x-1.

Упражнение 54. *Найти все значения параметра p, при которых полиномы*
а) $x^3+p\,x+1$ *и* $x^2+p\,x+1$;
б) $x^3+p\,x^2-14$ *и* $x^3+p\,x-14$
имеют общий корень.

Результант и алгоритм Евклида

Итак, согласно теореме 1.1, условие $\mathcal{R}(f,g)=0$ равносильно существованию общего корня полиномов $f(x)$ и $g(x)$, или же, иными словами, *нетривиального* наибольшего общего делителя НОД (f,g) этих полиномов. Обратно, условие $\mathcal{R}(f,g)\neq 0$ гарантирует взаимную простоту полиномов $f(x)$ и $g(x)$, т.е. то, что НОД $(f,g) = const \neq 0$. Поскольку конструктивное вычисление НОД (f,g) может быть организовано по алгоритму Евклида[3], следует ожидать, что в этом алгоритме выражение для $\mathcal{R}(f,g)$ должно явно проявиться на каком-то этапе. Восстановим здесь решение примера 4.5 из §4.2 главы 3.

Пример 1.3. Найти условия взаимной простоты полиномов

$$f(x)=a_0x^2+a_1x+a_2 \quad g(x)=b_0x^2+b_1x+b_2,\ a_0\neq 0, b_0\neq 0.$$

[3]Глава 3, §4.1

Решение. Первым шагом алгоритма Евклида будет деление (с остатком) $f(x)$ на $g(x)$:

$$f(x)=\frac{a_0}{b_0}g(x)+\underbrace{\left(a_1-\frac{a_0b_1}{b_0}\right)x+\left(a_2-\frac{a_0b_2}{b_0}\right)}_{\stackrel{\text{def}}{=}r_1(x)}.$$

Предположим сначала, что $a_1b_0-a_0b_1\ne 0$. Тогда второй шаг алгоритма Евклида заключается в делении $g(x)$ на $r_1(x)$. В ходе длинных выкладок получим:

$$g(x)=r_1(x)\frac{b_0}{a_1b_0-a_0b_1}\left[b_0x+\left(b_1-b_0\frac{a_2b_0-a_0b_2}{a_1b_0-a_0b_1}\right)\right]+\underbrace{\frac{b_0\mathcal{R}(f,g)}{(a_1b_0-a_0b_1)^2}}_{\stackrel{\text{def}}{=}r_2(x)},$$

где $\mathcal{R}(f,g)$ в точности совпадает с результантом (1.5).

Теперь проанализируем $\mathsf{НОД}\,(f,g)$ в зависимости от величины $\mathcal{R}(f,g)$. Если $\mathcal{R}(f,g)\ne 0$, то очевидно, что на следующем, т.е. третьем, шаге алгоритм Евклида остановится и $\mathsf{НОД}\,(f,g)=r_2(x)=$ $=const\ne 0$. В этом случае полиномы $f(x)$ и $g(x)$ взаимно просты. Если же $\mathcal{R}(f,g)=0$, то $\mathsf{НОД}\,(f,g)=r_1(x)$, т.е. НОД является линейным по x полиномом.

Предположим теперь, что $a_1b_0-a_0b_1=0$. Алгоритм Евклида остановится уже на втором шаге. Полиномы $f(x)$ и $g(x)$ будут взаимно простыми при $a_2b_0-a_0b_2\ne 0$. Одновременное выполнение условий

$$a_1b_0-a_0b_1=0 \quad \text{и} \quad a_2b_0-a_0b_2=0$$

равносильно тому, что коэффициенты полиномов $f(x)$ и $g(x)$ пропорциональны, т.е. $f(x)=Cg(x)$ при некоторой константе C. Тогда $\mathsf{НОД}\,(f,g)$ совпадает с любым из этих полиномов.

Обобщая предшествующие рассуждения, можем выписать условие взаимной простоты $f(x)$ и $g(x)$.

Ответ. $\mathsf{НОД}\,(f,g)=1$ тогда и только тогда, когда $\mathcal{R}(f,g)\ne 0$, где число $\mathcal{R}(f,g)$ определяется формулой (1.5).

Рассмотрим теперь общую схему алгоритма Евклида. Пусть схема последовательного деления имеет вид:

$$
\begin{aligned}
f(x) &= g(x)q_1(x) + r_1(x)\ , \quad 0 \le \deg r_1(x) < \deg g(x)\ ,\\
g(x) &= r_1(x)q_2(x) + r_2(x)\ , \quad 0 \le \deg r_2(x) < \deg r_1(x),\\
r_1(x) &= r_2(x)q_3(x) + r_3(x)\ , \quad 0 \le \deg r_3(x) < \deg r_2(x),\\
&\dots \quad \dots\\
r_{k-2}(x) &= r_{k-1}(x)q_k(x) + r_k(x)\ , \quad 0 \le \deg r_k(x) < \deg r_{k-1}(x),\\
r_{k-1}(x) &= r_k(x)q_{k+1}(x);
\end{aligned}
$$

т.е. $r_k(x) = \text{НОД}\,(f(x), g(x))$. Рассмотрим первую строчку схемы. Очевидно, что на корнях полинома $g(x)$ значения $f(x)$ и $r_1(x)$ совпадают: $f(\mu_k) = r_1(\mu_k)$. Тогда, на основании формулы (1.9), имеем

$$
\begin{aligned}
\mathcal{R}(f,g) &= (-1)^{mn} b_0^n \prod_{k=1}^{m} f(\mu_k) = (-1)^{mn} b_0^n \prod_{k=1}^{m} r_1(\mu_k) =\\
&= (-1)^{mn} b_0^{n-n_1} \mathcal{R}(g, r_1),
\end{aligned}
$$

здесь $n_1 \stackrel{\text{def}}{=} \deg r_1(x)$. Итак, с точностью до степени старшего коэффициента полинома g, результант полиномов f и g совпадает с результантом полиномов g и r_1. Теперь перейдем ко второй строчке схемы. На основании тех же рассуждений можем утверждать, что $\mathcal{R}(g, r_1) = (-1)^{n_1 m} \widetilde{b}_0^{n_1-n_2} \mathcal{R}(r_1, r_2)$, где $n_2 \stackrel{\text{def}}{=} \deg r_2(x)$, а $\widetilde{b}_0$ — старший коэффициент полинома $r_1(x)$. Продолжаем процесс далее. На каждом шаге степени остатков понижаются, т.е. вычисление результанта все упрощается. И если предположить, что предпоследний остаток, т.е. $r_{k-1}(x)$, является линейным полиномом, то $r_k(x)$ должен быть константой: $r_k(x) \equiv const$. Согласно формуле (1.10), эта константа совпадает с $\mathcal{R}(r_{k-1},\, r_k)$ и, следовательно, будет обращаться в нуль тогда и только тогда, когда $\mathcal{R}(f,g) = 0$. Если взять коэффициенты полиномов f и g *символьными* (буквенными) — как это было сделано в примере 1.3, то результант этих полиномов, рассматриваемый как полином относительно их коэффициентов, будет совпадать с $\mathcal{R}(r_{k-1},\, r_k)$, с точностью до сомножителя, являющегося рациональной функцией коэффициентов. Таким образом, вычисление $\mathcal{R}(f,g)$ может быть произведено с помощью вычисления $\text{НОД}\,(f,g)$ по алгоритму Евклида; именно эта схема вычисления реализована во всех пакетах компьютерной алгебры.

Линейное представление результанта

Покажем теперь отношение результанта к еще одной задаче, которую мы решали в §4.1 главы 3.

Задача. Найти полиномы $u(x)$ и $v(x)$ из $\mathbb{A}[x]$, обеспечивающие справедливость тождества

$$f(x)v(x)+g(x)u(x)\equiv \text{НОД}\,(f,g), \qquad (1.11)$$

которое называется **линейным представлением наибольшего общего делителя**.

Теорема 1.2. *Существуют полиномы $\tilde{u}(x)$ и $\tilde{v}(x)$ из $\mathbb{A}[x]$ со степенями $\deg \tilde{u} \le (\deg f)-1$, $\deg \tilde{v} \le (\deg g)-1$, удовлетворяющие тождеству*

$$\mathcal{R}(f,g)\equiv f(x)\tilde{v}(x)+g(x)\tilde{u}(x). \qquad (1.12)$$

Если, вдобавок, полиномы $f(x)$ и $g(x)$ взаимно просты, то полиномы $\tilde{u}(x)$ и $\tilde{v}(x)$ будут определяться единственным образом.

Доказательство проведем снова для случая $n=5$ и $m=3$. Прибавим к последнему столбцу матрицы M ее первый столбец, домноженный на x^7, второй, домноженный на x^6, и т.д., предпоследний, домноженный на x. Величина определителя не изменится. С другой стороны,

$$\mathcal{R}(f,g)=\left|\begin{array}{cccccccc}
a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & 0 & x^2f(x)\\
0 & a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & xf(x)\\
0 & 0 & a_0 & a_1 & a_2 & a_3 & a_4 & f(x)\\
0 & 0 & 0 & 0 & b_0 & b_1 & b_2 & g(x)\\
0 & 0 & 0 & b_0 & b_1 & b_2 & b_3 & xg(x)\\
0 & 0 & b_0 & b_1 & b_2 & b_3 & 0 & x^2g(x)\\
0 & b_0 & b_1 & b_2 & b_3 & 0 & 0 & x^3g(x)\\
b_0 & b_1 & b_2 & b_3 & 0 & 0 & 0 & x^4g(x)
\end{array}\right| .$$

Представим последний столбец определителя в виде суммы двух:

$$[x^2f(x),xf(x),f(x),0,0,0,0,0]^{\top} \quad \text{и}$$

$$[0,0,0,g(x),xg(x),x^2g(x),x^3g(x),x^4g(x)]^{\top};$$

тогда определитель можно также представить в виде суммы двух слагаемых. Следовательно, полином $\tilde{v}(x)$ (или $\tilde{u}(x)$) равен определителю, получающемуся из результанта заменой в нем последнего столбца на $[x^2, x, 1, 0, 0, 0, 0, 0]^\top$ (или, соответственно, на $[0, 0, 0, 1, x, x^2, x^3, x^4]^\top$).

Пусть теперь полиномы $f(x)$ и $g(x)$ взаимно просты. Тогда тождество (1.11) имеет вид

$$f(x)\tilde{v}(x) + g(x)\tilde{u}(x) \equiv 1. \tag{1.13}$$

Будем искать полиномы $\tilde{u}(x)$ и $\tilde{v}(x)$, ему удовлетворяющие, методом неопределенных коэффициентов:

$$\tilde{v}(x) \stackrel{\text{def}}{=} v_0x^2 + v_1x + v_2,\ \ \tilde{u}(x) \stackrel{\text{def}}{=} u_0x^4 + u_1x^3 + u_2x^2 + u_3x + u_4.$$

Подставим эти выражения в тождество (1.13)

$$(v_0a_0 + u_0b_0)x^7 + (v_0a_1 + v_1a_0 + u_0b_1 + u_1b_0)x^6 + \dots + (v_2a_5 + u_4b_3) \equiv 1$$

и приравняем коэффициенты при одинаковых степенях x:

$$\begin{array}{lllllllll}
v_0a_0 & & & +u_0b_0 & & & & & = 0 \\
v_0a_1 & +v_1a_0 & & +u_0b_1 & +u_1b_0 & & & & = 0 \\
v_0a_2 & +v_1a_1 & +v_2a_0 & +u_0b_2 & +u_1b_1 & +u_2b_0 & & & = 0 \\
 & & \dots & & & & \dots & & \dots \\
 & & v_2a_5 & & & & & +u_4b_3 & = 1
\end{array}$$

Получим систему из 8 линейных уравнений для определения 8 коэффициентов $v_0, v_1, v_2, u_0, u_1, u_2, u_3, u_4$. Определитель этой системы (с точностью до транспонирования и перестановки столбцов) совпадает с результантом. По предположению, $\mathcal{R}(f, g) \neq 0$. Следовательно, система имеет единственное решение, которое может быть определено по формулам Крамера. Легко показать тождественность этого решения тому, что получено в доказательстве первой части теоремы. ■

Пример 1.4. Найти полиномы $\tilde{u}(x)$ и $\tilde{v}(x)$, удовлетворяющие тождеству (1.12) для $f(x) = x^5 - 4\,x - 2,\ g(x) = x^3 - 1$.

Решение. Разложим по последнему столбцу определитель

$$\tilde{v}(x) = \begin{vmatrix} 1 & 0 & 0 & 0 & -4 & -2 & 0 & x^2 \\ 0 & 1 & 0 & 0 & 0 & -4 & -2 & x \\ 0 & 0 & 1 & 0 & 0 & 0 & -4 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \end{vmatrix} = -18\,x^2 + 7\,x - 8.$$

Аналогично находим $\tilde{u}(x)=18x^4-7x^3+8x^2+18x-79$, а $\mathcal{R}(f,g)=95$. $\triangle$

Упражнение 55. *Найти полиномы $\tilde{u}(x)$ и $\tilde{v}(x)$, удовлетворяющие тождеству* (1.12) *для*

а) $f(x) = x^3 + 3\,x + 3,\ g(x) = x^2 - x - 2$;

б) $f(x) = x^4,\ g(x) = x^3 - 3\,x^2 + 4$;

в) $f(x) = x^5 - x^3 + 2\,x^2 - 2\,x + 2,\ g(x) = x^4 + 2\,x^3 + 7\,x^2 + 2\,x + 6$.

Упражнение 56. *Зная, что полиномы $f(x)$ и $g(x)$ взаимно просты, методом неопределенных коэффициентов подобрать полиномы $\tilde{u}(x)$ и $\tilde{v}(x)$ наименьшей степени, удовлетворяющие тождеству* (1.13) *для*

а) $f(x) = x^5,\ g(x) = (x-1)^3$;

б) $f(x) = x^3 + x^2 - x - 2,\ g(x) = x^3 - 4\,x^2 + 5\,x - 2$.

§2. Дискриминант

Для того чтобы полином $f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n \in \mathbb{A}[x]$ имел кратный корень, необходимо и достаточно[4], чтобы он имел общий корень со своей производной $f'(x)$. По теореме 1.1, для этого

[4] Глава 3, следствие 1 к теореме 6.1.

Перейти к оглавлению на странице: 460

Н. и Д., чтобы $\mathcal{R}(f,f')=0$. Соответствующий определитель

$$D=\left|\begin{array}{ccccccccc}
a_0 & a_1 & \dots & a_{n-2} & a_{n-1} & a_n & & & \\
 & a_0 & \dots & a_{n-3} & a_{n-2} & a_{n-1} & a_n & \mathbb{O} & \\
 & & \ddots & & & & \ddots & \ddots & \\
 & & & a_0 & a_1 & \dots & & a_{n-1} & a_n \\
 & \mathbb{O} & & & na_0 & & \dots & 2a_{n-2} & a_{n-1} \\
 & & & na_0 & (n-1)a_1 & & \dots & a_{n-1} & \\
 & & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & & & \cdot^{\cdot^{\cdot}} & & \\
 & na_0 & (n-1)a_1 & \dots & & a_{n-1} & & \mathbb{O} & \\
na_0 & (n-1)a_1 & \dots & & a_{n-1} & & & &
\end{array}\right|$$

будет делиться на a_0 (общий множитель элементов первого столбца).

Определение 2.1. Выражение D/a_0 называется **дискриминантом** полинома $f(x)$ и обозначается $\mathcal{D}(f)$:

$$\mathcal{D}(f)\stackrel{\text{def}}{=}D/a_0=(-1)^{n(n-1)/2}\mathcal{R}(f,f')/a_0. \qquad (2.1)$$

По построению, дискриминант является полиномом относительно коэффициентов $a_0,\dots,a_n$:

$$\mathcal{D}(a_0x^n+\dots+a_n)\in\mathbb{Z}[a_0,\dots,a_n];$$

степень этого полинома равна $2n-2$, и в своем разложении по степеням $a_0,\dots,a_n$ он будет содержать слагаемое $(-1)^{n(n-1)/2}n^na_0^{n-1}a_n^{n-1}$.

Пример 2.1.

$$\mathcal{D}(a_0x^2+a_1x+a_2)=\frac{1}{a_0}\left|\begin{array}{ccc} a_0 & a_1 & a_2\\ 0 & 2a_0 & a_1\\ 2a_0 & a_1 & 0\end{array}\right|=a_1^2-4\,a_0a_2.$$

Упражнение 57. *Доказать, что*

а) $\mathcal{D}(x^3+p\,x+q)=-108\left(\dfrac{q^2}{4}+\dfrac{p^3}{27}\right)$;

б) $\mathcal{D}(a_0x^3+a_1x^2+a_2x+a_3)=a_1^2a_2^2-4a_1^3a_3-4a_0a_2^3+18a_0a_1a_2a_3-27a_0^2a_3^2$;

в) $\mathcal{D}(x^4+p\,x+q)=6912\left(\dfrac{q^3}{27}-\dfrac{p^4}{256}\right)$.

Следующий результат является очевидным следствием теоремы 1.1.

Теорема 2.1. *Полином $f(x)$ имеет кратный корень тогда и только тогда, когда $\mathcal{D}(f)=0$.*

Теорема 2.2. *Имеет место равенство*

$$\mathcal{D}(f)=a_0^{2n-2}\prod_{1\le j<k\le n}(\lambda_k-\lambda_j)^2. \tag{2.2}$$

Здесь $\lambda_1,\dots,\lambda_n$ — корни $f(x)$.

Доказательство. Воспользуемся равенством (1.8) в применении к случаю $g(x)\equiv f'(x)$:

$$\mathcal{R}(f,f')=a_0^{n-1}\prod_{j=1}^n f'(\lambda_j)=$$

$$=a_0^{n-1}\prod_{j=1}^n\left[a_0(\lambda_j-\lambda_1)(\lambda_j-\lambda_2)\dots(\lambda_j-\lambda_{j-1})(\lambda_j-\lambda_{j+1})\dots(\lambda_j-\lambda_n)\right].$$

Последнее произведение содержит $n(n-1)$ сомножителей, причем наряду с $(\lambda_j-\lambda_k)$ включает и $(\lambda_k-\lambda_j)$. Поменяем у половины сомножителей знаки:

$$\mathcal{R}(f,f')=(-1)^{n(n-1)/2}a_0^{2n-1}\prod_{0\le j<k\le n}(\lambda_k-\lambda_j)^2.$$

Из (2.1) тогда следует (2.2). ■

Таким образом, величина $\mathcal{D}(f)$ характеризует расстояние между корнями полинома $f(x)$.

Упражнение 58. *Доказать, что для вещественности всех корней полинома $f(x)\in\mathbb{R}[x]$ необходимо выполнение условия $\mathcal{D}(f)\ge 0$.*

Упражнение 59. *Доказать следующие формулы:*
а) $\mathcal{D}(f\cdot g)=\mathcal{D}(f)\mathcal{D}(g)[\mathcal{R}(f,g)]^2$;
б) $\mathcal{D}(f(x)(x-a))=\mathcal{D}(f(x))[f(a)]^2$.

Упражнение 60. *Пусть* $f^* \overset{\text{def}}{=} x^n f(1/x) = a_n x^n + \ldots + a_0$ *и* $a_0 \neq 0$, $a_n \neq 0$. *Доказать, что* $\mathcal{D}(f) = \mathcal{D}(f^*)$.

§3. Субрезультанты

Рассмотрим снова пример 1.1. Пусть $\mathcal{R}(f,g) = 0$. Тогда, по теореме 1.1, существует общий корень полиномов $f(x)$ и $g(x)$. Выразим его через коэффициенты полиномов.

Неизвестное значение $x = c$ этого корня должно удовлетворять равенствам (1.1). Отбросим первое и последнее из них:

$$\begin{aligned}
a_0c^6+a_1c^5+a_2c^4+a_3c^3+a_4c^2+a_5c &= 0,\\
a_0c^5+a_1c^4+a_2c^3+a_3c^2+a_4c+a_5 &= 0,\\
b_0c^3+b_1c^2+b_2c+b_3 &= 0,\\
b_0c^4+b_1c^3+b_2c^2+b_3c &= 0,\\
b_0c^5+b_1c^4+b_2c^3+b_3c^2 &= 0,\\
b_0c^6+b_1c^5+b_2c^4+b_3c^3 &= 0.
\end{aligned} \tag{3.1}$$

Перепишем в матричном виде

$$\underbrace{\begin{pmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & a_5 \\ 0 & a_0 & a_1 & a_2 & a_3 & a_4 \\ 0 & 0 & 0 & b_0 & b_1 & b_2 \\ 0 & 0 & b_0 & b_1 & b_2 & b_3 \\ 0 & b_0 & b_1 & b_2 & b_3 & 0 \\ b_0 & b_1 & b_2 & b_3 & 0 & 0 \end{pmatrix}}_{M_1} \begin{pmatrix} c^6 \\ c^5 \\ c^4 \\ c^3 \\ c^2 \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ -a_5 \\ -b_3 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Пусть $\det M_1 \neq 0$. Тогда, по теореме Крамера, существует единственное решение системы (3.1), и корень c представи́м в виде

$$c = -\frac{\overbrace{\begin{vmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & \\ & a_0 & a_1 & a_2 & a_3 & a_5 \\ & & & b_0 & b_1 & b_3 \\ & & b_0 & b_1 & b_2 & \\ & b_0 & b_1 & b_2 & b_3 & \\ b_0 & b_1 & b_2 & b_3 & & \end{vmatrix}}^{\det M_1^{(1)}}}{\det M_1}.$$

Определение 3.1. Определитель матрицы M_1, получаемой из матрицы (1.3) вычеркиванием первого и последнего столбцов, первой и последней строк, называется **первым субрезультантом** полиномов f и g; будем обозначать его $\mathcal{R}^{(1)}(f,g)$.

Теорема 3.1. *Пусть $\mathcal{R}(f,g)=0$ и $x=c$ — общий корень $f(x)$ и $g(x)$. Обозначим $f_1(x)=f(x)/(x-c)$, $g_1(x)=g(x)/(x-c)$. Справедливо равенство*

$$\mathcal{R}^{(1)}(f,g)=\mathcal{R}(f_1,g_1). \tag{3.2}$$

Доказательство. Для случая

$$f(x)=a_0x^5+a_1x^4+a_2x^3+a_3x^2+a_4x+a_5,\ g(x)=b_0x^3+b_1x^2+b_2x+b_3$$

рассуждения будут следующими. Поскольку $f(x)\equiv(x-c)f_1(x)$, $g(x)\equiv(x-c)g_1(x)$, то коэффициенты полиномов

$$f_1(x)=a_0'x^4+a_1'x^3+a_2'x^2+a_3'x+a_4',\ g_1(x)=b_0'x^2+b_1'x+b_2'$$

можно найти из равенств

$$a_0=a_0', a_j=a_j'-a_{j-1}'c,\ (j=1,2,3,4)\ , a_5=-a_4'c,$$

$$b_0=b_0', b_j=b_j'-b_{j-1}'c,\ (j=1,2)\ , b_3=-b_2'c.$$

Если из каждого столбца результанта

$$\mathcal{R}(f_1,g_1)=\begin{vmatrix} a_0' & a_1' & a_2' & a_3' & a_4' & 0\\ 0 & a_0' & a_1' & a_2' & a_3' & a_4'\\ 0 & 0 & 0 & b_0' & b_1' & b_2'\\ 0 & 0 & b_0' & b_1' & b_2' & 0\\ 0 & b_0' & b_1' & b_2' & 0 & 0\\ b_0' & b_1' & b_2' & 0 & 0 & 0 \end{vmatrix}$$

(кроме первого) вычесть предшествующий, домноженный на c, то получим выражение для $\mathcal{R}^{(1)}(f,g)$. ■

Теорема 3.2. *Для того чтобы $f(x)$ и $g(x)$ имели только один общий корень, необходимо и достаточно, чтобы*

$$\mathcal{R}(f,g)=0,\ \mathcal{R}^{(1)}(f,g)\neq 0.$$

Доказательство следует из теорем 1.1 и 3.1. ■

Следствие 3.1. *При выполнении условия теоремы 3.2 единственный общий корень рационально выражается через коэффициенты полиномов $f(x)$ и $g(x)$:*

$$c = -\frac{\det M_1^{(1)}}{\mathcal{R}^{(1)}(f,g)}. \tag{3.3}$$

Здесь матрица $M_1^{(1)}$ получается из $M^{(1)}$ заменой ее последнего столбца на

$$[\underbrace{0,\dots,0,}_{m-2} a_n, b_m, \underbrace{0,\dots,0}_{n-2}]^\top.$$

Пример 3.1. Найти общий корень полиномов

$$f(x) = x^2 - 4x - 5 \quad \text{и} \quad g(x) = x^2 - 7x + 10.$$

Решение. Общий корень $f(x)$ и $g(x)$ существует, поскольку

$$\mathcal{R}(f,g) = \begin{vmatrix} 1 & -4 & -5 & 0 \\ 0 & 1 & -4 & -5 \\ 0 & 1 & -7 & 10 \\ 1 & -7 & 10 & 0 \end{vmatrix} = 0.$$

Имеем далее:

$$\mathcal{R}^{(1)}(f,g) = \begin{vmatrix} 1 & -4 \\ 1 & -7 \end{vmatrix} = -3 \ne 0, \ \det M_1^{(1)} = \begin{vmatrix} 1 & -5 \\ 1 & 10 \end{vmatrix} = 15.$$

По формуле (3.3), получаем: $c = -\dfrac{15}{-3} = 5$. △

Пусть теперь $\mathcal{R}(f,g) = 0, \mathcal{R}^{(1)}(f,g) = 0$. Тогда у полиномов $f(x)$ и $g(x)$ имеется, по крайней мере, два общих корня c_1 и c_2.

Оба эти корня должны удовлетворять уравнениям (3.1). Рассмотрим подсистему (субсистему) системы (3.1), вычеркнув первое и последнее уравнения и считая c равным c_1 или c_2:

$$\left\{\begin{array}{rl} a_0c^5+a_1c^4+a_2c^3+a_3c^2+a_4c+a_5 & =0, \\ b_0c^3+b_1c^2+b_2c+b_3 & =0, \\ b_0c^4+b_1c^3+b_2c^2+b_3c & =0, \\ b_0c^5+b_1c^4+b_2c^3+b_3c^2 & =0. \end{array}\right.$$

Перейти к оглавлению на странице: 460

Запишем эти уравнения в матричной форме:

$$\underbrace{\begin{pmatrix} a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & b_0 & b_1 \\ 0 & b_0 & b_1 & b_2 \\ b_0 & b_1 & b_2 & b_3 \end{pmatrix}}_{M_2} \begin{pmatrix} c^5 \\ c^4 \\ c^3 \\ c^2 \end{pmatrix} = \begin{pmatrix} -a_4c - a_5 \\ -b_2c - b_3 \\ -b_3c \\ 0 \end{pmatrix}.$$

Пусть $\det M_2 \neq 0$. Тогда, по теореме Крамера, существует единственное решение этой системы, и c^2 должно удовлетворять уравнению

$$c^2 = \frac{\begin{vmatrix} a_0 & a_1 & a_2 & -a_4c - a_5 \\ & & b_0 & -b_2c - b_3 \\ & b_0 & b_1 & -b_3c \\ b_0 & b_1 & b_2 & \end{vmatrix}}{\det M_2},$$

откуда получаем следующее квадратное уравнение, которому должны удовлетворять c_1 и c_2:

$$\begin{vmatrix} a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & b_0 & b_1 \\ 0 & b_0 & b_1 & b_2 \\ b_0 & b_1 & b_2 & b_3 \end{vmatrix} c^2 + \begin{vmatrix} a_0 & a_1 & a_2 & a_4 \\ 0 & 0 & b_0 & b_2 \\ 0 & b_0 & b_1 & b_3 \\ b_0 & b_1 & b_2 & 0 \end{vmatrix} c + \begin{vmatrix} a_0 & a_1 & a_2 & a_5 \\ 0 & 0 & b_0 & b_3 \\ 0 & b_0 & b_1 & 0 \\ b_0 & b_1 & b_2 & 0 \end{vmatrix} = 0. \qquad (3.4)$$

Определение 3.2. Определитель матрицы M_2, получаемой из матрицы M вычеркиванием двух первых и двух последних столбцов, двух первых и двух последних строк, называется **вторым субрезультантом** полиномов f и g и обозначается $\mathcal{R}^{(2)}(f,g)$.

Теорема 3.3. *Для того чтобы $f(x)$ и $g(x)$ имели в точности два общих корня, необходимо и достаточно, чтобы выполнялись условия*

$$\mathcal{R}(f,g) = \mathcal{R}^{(1)}(f,g) = 0, \ \mathcal{R}^{(2)}(f,g) \neq 0.$$

Доказательство следует из теорем 1.1 и 3.1. ∎

Следствие 3.2. *При условии теоремы 3.3 оба корня должны удовлетворять квадратному уравнению*

$$x^2\mathcal{R}^{(2)}(f,g) + x \det M_2^{(1)} + \det M_2^{(2)} = 0. \qquad (3.5)$$

Здесь матрицы $M_2^{(1)}$ *и* $M_2^{(2)}$ *получаются из* M_2 *заменой последнего ее столбца на*

$$[\underbrace{0,\dots,0}_{m-4}, a_n, a_{n-1}, b_{m-1}, b_m, \underbrace{0,\dots,0}_{n-4}]^\top \quad \textit{и} \quad [\underbrace{0,\dots,0}_{m-3}, a_n, b_m, \underbrace{0,\dots,0}_{n-3}]^\top$$

соответственно. Полином, стоящий в левой части уравнения (3.5)*, является* НОД (f,g).

Определение 3.3. Определитель матрицы M_k, получаемой из матрицы M вычеркиванием k первых и k последних столбцов, k первых и k последних строк, называется k**-м субрезультантом** полиномов f и g и обозначается $\mathcal{R}^{(k)}(f,g)$. Для однообразия будем считать нулевым субрезультантом определитель матрицы M: $\mathcal{R}^{(0)}(f,g) \stackrel{\text{def}}{=} \det M = (-1)^{n(n-1)/2}\mathcal{R}(f,g)$.

Пример 3.2.

$$\mathcal{R}^{(3)}(a_0x^5+a_1x^4+a_2x^3+a_3x^2+a_4x+a_5,\ b_0x^3+b_1x^2+b_2x+b_3) =$$

$$= \begin{vmatrix} 0 & b_0 \\ b_0 & b_1 \end{vmatrix} = -\,b_0^2.$$

Обобщая результаты теорем 3.2 и 3.3, приходим к следующему результату.

Теорема 3.4. **а)** *Для существования* d *общих корней у* $f(x)$ *и* $g(x)$ *(т.е. для того, чтобы* $\deg(\text{НОД}\,(f,g)) = d$*), необходимо и достаточно, чтобы*

$$\mathcal{R}(f,g)=0, \mathcal{R}^{(1)}(f,g)=0,\dots,\mathcal{R}^{(d-1)}(f,g)=0,\ \mathcal{R}^{(d)}(f,g)\neq 0.$$

б) *В этом случае* НОД (f,g) *можно представить в виде*

$$x^d\mathcal{R}^{(d)}(f,g) + x^{d-1}\det M_d^{(1)} + \ldots + \det M_d^{(d)}.$$

Здесь $M_d^{(j)}$ *— матрица, получаемая из* M_d *заменой последнего столбца на столбец*

$$[a_{m+n-2d+j-1}, a_{m+n-2d+j-2}, \dots, a_{n-d+j}, b_{m-d+j}, b_{m-d+j+1},\dots, b_{m+n-2d+j-1}]^\top$$

(здесь полагаем $a_K = 0$ при $K > n$ и $b_L = 0$ при $L > m$).

в) *Полиномы $v(x)$ и $u(x)$ из $\mathbb{A}[x]$, дающие линейное представление* НОД (f,g) *по формуле* (1.11), *получаются из $\mathcal{R}^{(d)}$ заменой в нем последнего столбца на*

$$\left[x^{m-d-1}, x^{m-d-2}, \dots, x, 1, 0, 0, \dots, 0\right]^{\top} \quad \text{и}$$
$$\left[0, 0, \dots, 0, 0, 1, x, \dots, x^{n-d-1}\right]^{\top}$$

соответственно. Эти полиномы будут единственными при ограничениях на степени:

$$\deg v \le m - d - 1, \ \deg u \le n - d - 1.$$

Доказательство. Для пункта **в)** в случае

$$f(x) = a_0x^5 + a_1x^4 + a_2x^3 + a_3x^2 + a_4x + a_5, \ g(x) = b_0x^3 + b_1x^2 + b_2x + b_3$$

и $d = 2$ рассуждения будут следующими. Домножим $\mathcal{R}^{(2)}$ на x^2:

$$x^2\mathcal{R}^{(2)} = x^2 \begin{vmatrix} a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & b_0 & b_1 \\ 0 & b_0 & b_1 & b_2 \\ b_0 & b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} a_0 & a_1 & a_2 & a_3x^2 \\ 0 & 0 & b_0 & b_1x^2 \\ 0 & b_0 & b_1 & b_2x^2 \\ b_0 & b_1 & b_2 & b_3x^2 \end{vmatrix} =$$

теперь к последнему столбцу прибавим первый, умноженный на x^5, второй, умноженный на x^4, и третий, умноженный на x^3:

$$= \begin{vmatrix} a_0 & a_1 & a_2 & a_0x^5+a_1x^4+a_2x^3+a_3x^2 \\ 0 & 0 & b_0 & b_0x^3+b_1x^2 \\ 0 & b_0 & b_1 & b_0x^4+b_1x^3+b_2x^2 \\ b_0 & b_1 & b_2 & b_0x^5+b_1x^4+b_2x^3+b_3x^2 \end{vmatrix} = \begin{vmatrix} a_0 & a_1 & a_2 & f(x)-(a_4x+a_5) \\ 0 & 0 & b_0 & g(x)-(b_2x+b_3) \\ 0 & b_0 & b_1 & xg(x)-b_3x \\ b_0 & b_1 & b_2 & x^2g(x) \end{vmatrix} =$$

последний столбец представляем в виде линейной комбинации остальных:

$$= f(x) \begin{vmatrix} a_0 & a_1 & a_2 & 1 \\ 0 & 0 & b_0 & 0 \\ 0 & b_0 & b_1 & 0 \\ b_0 & b_1 & b_2 & 0 \end{vmatrix} + g(x) \begin{vmatrix} a_0 & a_1 & a_2 & 0 \\ 0 & 0 & b_0 & 1 \\ 0 & b_0 & b_1 & x \\ b_0 & b_1 & b_2 & x^2 \end{vmatrix} - \begin{vmatrix} a_0 & a_1 & a_2 & a_4x+a_5 \\ 0 & 0 & b_0 & b_2x+b_3 \\ 0 & b_0 & b_1 & b_3x \\ b_0 & b_1 & b_2 & 0 \end{vmatrix}.$$

Если перенести последний определитель в левую часть (т.е. к $x^2\mathcal{R}^{(2)}$), то слева получим выражение для $\mathsf{НОД}\,(f,g)$ (см. формулу (3.4)); в правой же части коэффициенты при $f(x)$ и $g(x)$ будут равны соответственно полиномам $v(x)$ и $u(x)$ из тождества (1.11). ■

Упражнение 61. *Найти* $\mathsf{НОД}\,(f,g)$ *для*

а) $f(x)=x^4-x^3-x^2+1,\ g(x)=x^4-2\,x^3+2\,x^2-2\,x+1$ *;*
б) $f(x)=x^3+5\,x^2+5\,x+4,\ g(x)=x^3-x^2-x-2$ *;*
в) $f(x)=x^3-4\,x^2+4\,x-3,\ g(x)=3\,x^4-2\,x^3+4\,x^2-x+2$ *;*
г) $f(x)=2\,x^3+5\,x^2-6\,x-9,\ g(x)=3\,x^3+7\,x^2-11\,x-5$ *.*

Рассмотрим теперь частный случай $g(x)\equiv f'(x)$. Теорема 3.4 позволяет установить наличие d общих корней у полинома и его производной, т.е. наличие d кратных корней у $f(x)$ — с учетом их кратностей. Двойные корни полинома $f(x)$ оказываются простыми корнями $D(x)=\mathsf{НОД}\,(f,f')$, корни же кратностей больших двух становятся кратными корнями $D(x)$, но кратность их понижается на единицу, и для их поиска можно снова применить теорему 3.4 — теперь уже для поиска $\mathsf{НОД}\,(D,D')$.

Упражнение 62. *Имеет ли полином кратные корни? Если да, найти их кратности*

а) $f(x)=x^4-x^3-104x^2+514x-720$ *;*
б) $f(x)=x^4-x^3-30x^2-76x-56$ *;*
в) $f(x)=x^4+x^3-2x^2+4x-24$ *;*
г) $f(x)=x^5+x^4-2x^3-2x^2+x+1$ *;*
д) $f(x)=2x^6+6x^5+6x^4+x^3-3x^2-3x-1$ *.*

Упражнение 63. *При каких значениях параметров полиномы имеют*

а) *двойной корень:* $px^3+p^2x^2+x+p$ *;*
б) *двойной корень:* x^5+px^4+q *;*
в) *тройной корень:* x^4+px^3+2x+q *;*
г) *только два различных корня:* $x^4+4\,x^3-2\,x^2+px+q$ *?*

Перейти к оглавлению на странице: 460

§4. Приложения

Уничтожение иррациональности в знаменателе

Пусть $f(x), g(x), g_1(x)$ — полиномы из $\mathbb{A}[x]$, и $\lambda_1,\dots,\lambda_n$ — корни $f(x)$.

Задача. Для рациональной дроби $g_1(x)/g(x)$ найти такой полином $G(x)\in\mathbb{A}[x]$, чтобы

$$G(\lambda_1)=g_1(\lambda_1)/g(\lambda_1),\ \dots,\ G(\lambda_n)=g_1(\lambda_n)/g(\lambda_n).$$

Главным образом, нас будет интересовать случай, когда коэффициенты $f(x), g(x)$ и $g_1(x)$ являются рациональными числами, но среди корней $\lambda_1,\dots,\lambda_n$ полинома $f(x)$ имеются иррациональные. Именно в этом случае поставленную задачу и называют задачей об уничтожении иррациональности в знаменателе выражения $g_1(\lambda)/g(\lambda)$. Здесь предполагается, что $\lambda\in\{\lambda_1,\dots,\lambda_n\}$.

Теорема 4.1. *Если $\mathcal{R}(f,g)\neq 0$, то всегда существует полином $G(x)$, решающий поставленную задачу. При условии $\deg G<n$ такой полином определяется единственным образом, и его коэффициенты будут рационально зависеть от коэффициентов $f(x), g(x)$ и $g_1(x)$.*

Доказательство. По теореме 1.2, существуют полиномы $\tilde v(x)$ и $\tilde u(x)$, удовлетворяющие тождеству (1.12):

$$\tilde v(x)f(x)+\tilde u(x)g(x)\equiv\mathcal{R}(f,g),$$

а если их строить согласно алгоритму теоремы, то легко видеть, что их коэффициенты — как и коэффициенты $\mathcal{R}(f,g)$ — принадлежат $\mathbb{A}$. Подставляем в тождество $x=\lambda_j$:

$$\tilde u(\lambda_j)g(\lambda_j)=\mathcal{R}(f,g)\ \Rightarrow\ \frac{1}{g(\lambda_j)}=\frac{\tilde u(\lambda_j)}{\mathcal{R}(f,g)}\ \text{ и }\ \frac{g_1(\lambda_j)}{g(\lambda_j)}=\frac{g_1(\lambda_j)\tilde u(\lambda_j)}{\mathcal{R}(f,g)}.$$

Следовательно, в качестве искомого полинома $G(x)$ можно взять полином $g_1(x)\tilde u(x)/\mathcal{R}(f,g)$. Легко видеть, что наряду с полиномом $G(x)$ полином $G(x)+q(x)f(x)$, где $q(x)$ — произвольный полином из $\mathbb{A}[x]$, тоже дает решение поставленной задачи. Взяв в качестве

$G(x)$ остаток от деления $g_1(x)\tilde{u}(x)$ на $f(x)$ и поделив его на $\mathcal{R}(f,g)$, получим полином с указанным в теореме ограничением на степень. Результант и полином $\tilde{u}(x)$ можно искать с помощью теоремы 1.2, а также методом неопределенных коэффициентов. ■

Упражнение 64. *Доказать единственность полинома* $G(x)$ *при выполнении условия* $\deg G < n$.

Пример 4.1. Уничтожить иррациональность в знаменателе выражения $\lambda/(\lambda^3-1)$, где λ — корень полинома $x^5-4\,x-2$.

Решение. Из решения примера 1.4 имеем:

$$\tilde{u}(x) = 18\,x^4 - 7\,x^3 + 8\,x^2 + 18\,x - 79,\ \mathcal{R}(f,g) = 95.$$

Воспользуемся теперь алгоритмом из доказательства теоремы 4.1:

$$G(x) = (18\,x^5 - 7\,x^4 + 8\,x^3 + 18\,x^2 - 79\,x)/95.$$

Если поделить $G(x)$ на $f(x)$, то остаток от деления

$$G_1(x) = (-7\,x^4 + 8\,x^3 + 18\,x^2 - 7\,x + 36)/95$$

также является решением задачи — причем единственным среди полиномов степеней меньших 5. △

Упражнение 65. *Уничтожить иррациональность в знаменателе выражений*

а) $\lambda/(\lambda-1)$, *где* λ *— корень полинома* x^3-2x-2;

б) $1/(\lambda^3+3\lambda^2+3\lambda+2)$, *где* λ *— корень полинома* x^4+x^3- $-4x^2-3x+2$.

Преобразование Чирнгауза

Задача. Пусть даны два полинома из $\mathbb{A}[x]$:

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n,\ g(x) = b_0x^m + b_1x^{m-1} + \ldots + b_m$$

$(a_0 \ne 0, b_0 \ne 0)$; обозначим $\lambda_1,\ldots,\lambda_n$ (неизвестные нам) корни $f(x)$. Построить полином $F(y)$, имеющий корни $g(\lambda_1),\ldots,g(\lambda_n)$. Нахождение этого полинома называется **преобразованием Чирнгауза** $y=g(x)$ полинома $f(x)$.

Теорема 4.2. *Существует единственный нормализованный полином $F(y)$ степени n, решающий задачу. При этом его коэффициенты будут рационально зависеть от коэффициентов $f(x)$ и $g(x)$.*

Доказательство. Требуемый полином

$$F(y) = (y - g(\lambda_1)) \times \cdots \times (y - g(\lambda_n)).$$

По формуле (1.8):

$$F(y) \equiv \mathcal{R}(f(x), y - g(x))/a_0^m$$

(результант вычисляется для полиномов относительно переменной x, а y считается числовым параметром). ■

Пример 4.2. Найти преобразование Чирнгауза $y = x^2 + x - 1$ полинома $f(x) = x^3 - 2x + 3$.

Решение.

$$F(y) = \mathcal{R}(x^3 - 2x + 3,\ -x^2 - x + (1+y)) =$$
$$= -\begin{vmatrix} 1 & 0 & -2 & 3 & 0 \\ 0 & 1 & 0 & -2 & 3 \\ 0 & 0 & -1 & -1 & 1+y \\ 0 & -1 & -1 & 1+y & 0 \\ -1 & -1 & 1+y & 0 & 0 \end{vmatrix} = y^3 - y^2 + 6y - 4.$$

△

Пример 4.3. Для полинома $f(x) = x^3 + a_1x^2 + a_2x + a_3 \in \mathbb{C}[x]$ подобрать преобразование Чирнгауза вида $y = x^2 + b_1x + b_2 \in \mathbb{C}[x]$ так, чтобы получившийся в результате преобразования полином имел вид $F(y) = y^3 + c_3$.

Решение. Преобразование Чирнгауза при первоначально неопределенных b_1 и b_2 дает[5]:

$$F(y) = y^3 + c_1y^2 + c_2y + c_3$$

[5] Для вычислений использовался пакет MAPLE.

при

$$
\begin{aligned}
c_1 &= -a_1^2 + 2\,a_2 + a_1 b_1 - 3\,b_2,\\
c_2 &= -a_1a_2b_1 + a_2b_1^2 - 2a_1a_3 + a_2^2 - 4a_2b_2 + 3a_3b_1 + 2a_1^2b_2 - 2a_1b_1b_2 + 3b_2^2,\\
c_3 &= a_1b_1b_2^2 - a_1^2b_2^2 + a_1a_2b_1b_2 + a_2a_3b_1 - a_3^2 + 2\,a_2b_2^2 +\\
&\quad + a_3b_1^3 - a_2b_1^2b_2 - a_1a_3b_1^2 + 2\,a_1a_3b_2 - b_2^3 - 3\,a_3b_1b_2 - a_2^2b_2.
\end{aligned}
$$

Требуется подобрать b_1 и b_2 так, чтобы коэффициенты c_1 и c_2 обратились в нуль. Получаем систему из двух уравнений относительно b_1 и b_2: первое из них — линейное, второе — квадратное. Эта система разрешима в радикалах:

$$
b_1 = \frac{2\,a_1^3 - 7\,a_1a_2 + 9\,a_3 \pm \sqrt{-3\mathcal{D}(f)}}{2(a_1^2 - 3\,a_2)}, \tag{4.1}
$$

а b_2 выражается через b_1 по формуле

$$
b_2 = \frac{a_1b_1 - a_1^2 + 2\,a_2}{3}. \tag{4.2}
$$

Здесь $\mathcal{D}(f) = -27\,a_3^2 + 18\,a_1a_2a_3 - 4\,a_1^3a_3 + a_1^2a_2^2 - 4\,a_2^3$, т.е. является дискриминантом кубического полинома (см. упражнение 57, б)); предполагается также, что $a_1^2 - 3\,a_2 \neq 0$. Итак, преобразование Чирнгауза при указанных значениях параметров b_1 и b_2 дает полином $y^3 + c_3$ (здесь в приведенное выше выражение для c_3 также следует подставить полученные выражения для b_1 и b_2). $\triangle$

Только что решенный пример позволяет ответить на вопрос: *зачем нужно преобразование Чирнгауза*? Именно, это преобразование *иногда* позволяет решать алгебраические уравнения в радикалах. В самом деле, уравнение $y^3 + c_3 = 0$ можно решить в радикалах; если обозначить μ_1, μ_2 и μ_3 его корни, то корни исходного кубического уравнения $x^3 + a_1x^2 + a_2x + a_3 = 0$ получатся как решения квадратных уравнений $\lambda^2 + b_1\lambda + b_2 = \mu_j$ при b_1 и b_2 заданных формулами (4.1) и (4.2).

Упражнение 66. *Найти преобразование Чирнгауза, позволяющее решить в радикалах уравнение* $x^3 + a_1x^2 + {}^1/_3\, a_1^2x + a_3 = 0$.

Упражнение 67. *Считая теперь доказанным факт разрешимости кубического уравнения в радикалах, придумать способ решения в радикалах уравнения* $x^4 + a_1x^3 + a_2x^2 + a_3x + a_4 = 0$ *с помощью преобразования Чирнгауза вида* $y = x^2 + b_1x + b_2$.

Перейти к оглавлению на странице: 460

Пример 4.4. Пусть $f(x), g(x), g_1(x)$ — полиномы с рациональными коэффициентами. Построить полином $F(y)$, имеющий корни

$$g_1(\lambda_1)/g(\lambda_1),\ \dots,\ g_1(\lambda_n)/g(\lambda_n).$$

Решение. Можно воспользоваться результатами §4. Построив полином $G(x)$ такой, что

$$G(\lambda_1) = g_1(\lambda_1)/g(\lambda_1), \dots, G(\lambda_n) = g_1(\lambda_n)/g(\lambda_n),$$

сведем задачу к уже решенной. Можно действовать и напрямую: полином

$$F(y) \stackrel{\text{def}}{=} \mathcal{R}(f(x), yg(x) - g_1(x))$$

решает задачу при условии $\mathcal{R}(f(x), g(x)) \neq 0$. $\triangle$

Историческая справка

Граф **Эренфрид Вальтер фон Чирнгауз(ен) (Ehrenfried Walter von Tschirnhaus(en), 1651–1708).**

Предки графа жили в северных областях Чехии и носили фамилию Черноус. Математик и естествоиспытатель — в те времена ученые были универсалистами, и их способности использовались на разных поприщах. Строил заводы по производству оптических линз и зеркал. В 1704 г. саксонский курфюрст и польский король Август Сильный поручил Ч. контролировать деятельность молодого Иоганна Фридриха Бётгера (1662—1719). Бётгер начал свои научные занятия с того что во время работы в Берлине учеником аптекаря посвятил себя алхимическим опытам, имея целью получение “философского” камня, позволяющего превращать любой металл в золото. Слухи об этом дошли до прусского короля Фридриха I и тот решил “приватизировать” потенциальный источник золота. Вовремя предупрежденный, Бётгер успел сбежать в 1701 г. в Виттенберг, где, однако, попал в лапы Августа. Последний распорядился доставить умельца в Дрезден и там посадить под надзор, потребовав от него того же: наполнить его казну золотом. Для Бётгера все могло кончиться очень плохо — трансмутация элементов дело неприбыльное и в наши времена. Но Чирнгауз сумел спасти его от гнева короля одной только фразой:

“*Ваше величество! Золото может быть и белым!*”

Он имел в виду фарфор. Приблизительно в VI веке состав фарфора изобрели китайцы, однако производственный секрет (know-how) хранился в строгой тайне. Эпизодически попадая в Европу начиная с XIII века, китайский фарфор вставлялся европейскими ювелирами в оправу и наряду с другими драгоценными предметами хранился в церковных и дворянских сокровищницах. Тот же Август Саксонский обменял однажды на фарфоровое блюдо роту своих солдат-гвардейцев, правда “без мундиров и оружия”, как было особо отмечено в договоре. Попытки раскрытия секрета предпринимались неоднократно — и итальянцами (фарфор Медичи), и французами и англичанами. Легенды утверждали, что фарфор

— это смесь яичной скорлупы и морских раковин (отсюда и европейское название фарфора — порцелин, итальянские рыбаки морские раковины называли “порчелло” — свинушка, поросенок).

И только двум немцам — молодому талантливому авантюристу и старому опытному ученому — фортуна улыбнулась. И то не сразу. В конце XVII века Ч. проводит обширные геологические исследования в Саксонии. Его целью было найти сырье — огнеупорный материал для плавильных печей. Экспериментируя с обжигом, он имел дело с глиной из Кольдица — которая и стала впоследствии одной из главных составных частей мейсенского фарфора. В 1709 г. — уже после смерти Ч. — Бётгер добился окончательного успеха, и в 1710 г. в мейсенской крепости Альбрехтсбург была пущена первая европейская мануфактура твердого фарфора. Она работала день и ночь, принося Августу огромный доход. Всякий работающий на мануфактуре давал клятву в том, что под страхом смерти не раскроет тайны “порцелина”. Но однажды один из шпионов, подсматривавших за Бётгером, донес, что будто бы мастер собирается перебежать к прусскому королю. Этого доноса оказалось достаточно для того чтобы Бётгер кончил свою жизнь в тюрьме.

Десятилетия спустя в его память в Мейсене была построена церковь с 37 колоколами — фарфоровыми.

§5. Исключение переменных в системе уравнений

Рассмотрим теперь систему уравнений

$$f(x,y)=0, g(x,y)=0. \tag{5.1}$$

Здесь $\{f,g\}\subset\mathbb C[x,y]$. Нас интересуют все решения этой системы во множестве $\mathbb C$.

В §11.3 главы 3 мы упоминали, что в случае вещественности коэффициентов полинома $f(x,y)$ уравнение $f(x,y)=0$ определяет на плоскости (x,y) некоторую кривую, которая может состоять из нескольких ветвей. В случае, когда оба полинома f и g имеют вещественные коэффициенты, задачу поиска вещественных решений системы (5.1) можно считать задачей нахождения точек пересечения двух алгебраических кривых. Частным случаем пересечения кривых является их касание (рис. 2).

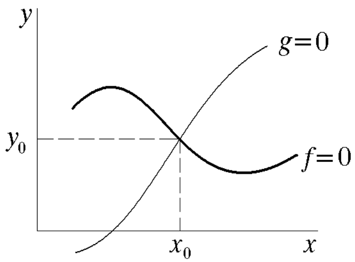


Рис. 1

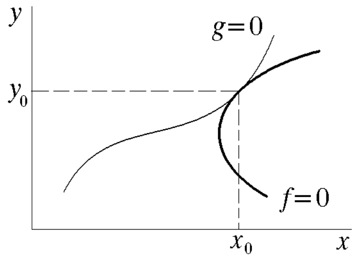


Рис. 2

Для того чтобы точка (x_0, y_0) была точкой касания необходимо, чтобы в ней выполнялись условия

$$f(x,y)=0,\ \ g(x,y)=0,\ \ \frac{\partial f}{\partial x}\Big/\frac{\partial f}{\partial y}=\frac{\partial g}{\partial x}\Big/\frac{\partial g}{\partial y},$$

где последнее характеризует равенство угловых коэффициентов касательных к алгебраическим кривым $f(x,y)=0$ и $g(x,y)=0$.

Определение 5.1. Выражение

$$\mathcal{J}(x,y)=\frac{\partial f}{\partial x}\frac{\partial g}{\partial y}-\frac{\partial g}{\partial x}\frac{\partial f}{\partial y} \tag{5.2}$$

называется **якобианом** полиномов $f(x,y)$ и $g(x,y)$.

Якобиан является двумерным аналогом производной полинома от одной переменной. В частности, обращение его в нуль на некотором решении системы (5.1) означает, что решение — **кратное**. Такое решение при бесконечно-малом возмущении коэффициентов полиномов распадается на несколько простых. Геометрически это означает, что при небольшой деформации этих кривых такая точка неустойчива: она либо распадается на две (или более) точки обычного пересечения, либо порождает пару мнимых.

Представим $f(x,y)$ и $g(x,y)$ в виде сумм их форм:

$$\begin{aligned}f(x,y)&=f_n(x,y)+f_{n-1}(x,y)+\ldots+f_0(x,y),\\ g(x,y)&=g_m(x,y)+g_{m-1}(x,y)+\ldots+g_0(x,y).\end{aligned}$$

Относительно коэффициентов старших форм

$$\begin{aligned}f_n(x,y)&=a_{n0}x^n+a_{n-1,1}x^{n-1}y+\ldots+a_{0n}y^n,\\ g_m(x,y)&=b_{m0}x^m+b_{m-1,1}x^{m-1}y+\ldots+b_{0m}y^m\end{aligned} \tag{5.3}$$

сделаем следующее предположение:

$$a_{n0} \neq 0, a_{0n} \neq 0, b_{m0} \neq 0, b_{0m} \neq 0. \tag{5.4}$$

Пара $(\alpha, \beta) \in \mathbb{C}^2$ будет решением (5.1) тогда и только тогда, когда полиномы $f(\alpha, y)$ и $g(\alpha, y)$ имеют общий корень $y = \beta$, а, следовательно, на основании теоремы 1.1,

$$\mathcal{R}(f(\alpha, y), g(\alpha, y)) = 0. \tag{5.5}$$

Запишем этот результант в форме Сильвестра (1.4). Для этого разложим $f(\alpha, y)$ и $g(\alpha, y)$ по убывающим степеням y

$$\begin{aligned} f(\alpha, y) &= A_0 y^n + A_1(\alpha) y^{n-1} + \ldots + A_n(\alpha), \\ g(\alpha, y) &= B_0 y^m + B_1(\alpha) y^{m-1} + \ldots + B_m(\alpha) \end{aligned}$$

(здесь $A_0 = a_{0n} \neq 0, B_0 = b_{0m} \neq 0$ по (5.4), $\deg A_j(\alpha) \leq j$, $\deg B_j(\alpha) \leq j$) и вычислим определитель Сильвестра

$$\begin{array}{l} m\left\{\begin{array}{l}\\ \\ \\ \\ \end{array}\right. \\ n\left\{\begin{array}{l}\\ \\ \\ \\ \\ \end{array}\right. \end{array}
\left|\begin{array}{ccccccc}
A_0 & A_1(\alpha) & \ldots & A_n(\alpha) & & \mathbb{O} & \\
 & A_0 & \ldots & & A_n(\alpha) & & \\
 & & \ddots & & \ldots & \ddots & \\
 & & & A_0 & A_1(\alpha) & \ldots & A_n(\alpha) \\
 & \mathbb{O} & & & B_0 & \ldots & B_m(\alpha) \\
 & & & B_0 & B_1(\alpha) & \ldots & \\
 & & \cdot^{\cdot^{\cdot}} & & \ldots & \cdot^{\cdot^{\cdot}} & \\
 & B_0 & B_1(\alpha) \ldots & & B_m(\alpha) & & \\
B_0 & B_1(\alpha) & \ldots & B_m(\alpha) & & \mathbb{O} &
\end{array}\right| = \mathcal{X}(\alpha). \tag{5.6}$$

$\mathcal{X}(\alpha)$ — полином по α. Для выполнения условия (5.5) необходимо и достаточно, чтобы значение α удовлетворяло уравнению $\mathcal{X}(x) = 0$. Если $\{f, g\} \subset \mathbb{A}[x, y]$, то и $\mathcal{X}(x) \in \mathbb{A}[x]$.

Определение 5.2. Полином $\mathcal{X}(x)$ — т.е. результант $f(x, y)$ и $g(x, y)$, рассматриваемых как полиномы по переменной y

$$\mathcal{X}(x) \stackrel{\mathrm{def}}{=} \mathcal{R}_y\left(f(x, y), g(x, y)\right),$$

называется **элиминантой системы** (5.1) по x. Аналогично определяется и вторая элиминанта системы

$$\mathcal{Y}(y) \stackrel{\mathrm{def}}{=} \mathcal{R}_x(f(x, y), g(x, y)).$$

Для простоты, мы не учитывали здесь (и не будем учитывать в последующих примерах) знак $(-1)^{n(n-1)/2}$ в выражениях обеих элиминант.

Теорема 5.1. *Пусть выполнено условие* (5.4). *Если пара* (α,β) *является решением* (5.1), *то необходимо, чтобы* $\mathcal{X}(\alpha)=0$ *и* $\mathcal{Y}(\beta)=0$.

Таким образом, решение системы (5.1) сводится к решению одного уравнения от одной переменной: $\mathcal{X}(x)=0$ (или $\mathcal{Y}(y)=0$). Говорят, что другая переменная исключена. Поэтому и соответствующий раздел алгебры называется **теорией исключения**.

Пусть теперь $x{=}\alpha$ — произвольный корень $\mathcal{X}(x)$: $\mathcal{X}(\alpha){=}0$. Тогда выполнено условие (5.5), а значит, $f(\alpha,y)$ и $g(\alpha,y)$ как полиномы от y имеют хотя бы один общий корень β.

Теорема 5.2. *Пусть выполнено условие* (5.4). *Тогда для любого корня* α *элиминанты* $\mathcal{X}(x)$ *существует хотя бы одно значение* $y=\beta$ *такое, что пара* (α,β) *оказывается решением* (5.1).

Пример 5.1. Решить систему уравнений

$$\left\{\begin{array}{l} f(x,y)=4x^2-7xy+y^2+13x-2y-3=0,\\ g(x,y)=9x^2-14xy+y^2+28x-4y-5=0.\end{array}\right.$$

Решение. Составим элиминанту $\mathcal{X}(x)$:

$$\begin{array}{l} f(x,y)=y^2+(-7\,x-2)y+(4\,x^2+13\,x-3),\\ g(x,y)=y^2+(-14\,x-4)+(9\,x^2+28\,x-5),\end{array}$$

$$\mathcal{X}(x)=\left|\begin{array}{cccc} 1 & -7x-2 & 4x^2+13x-3 & 0\\ 0 & 1 & -7x-2 & 4x^2+13x-3\\ 0 & 1 & -14x-4 & 9x^2+28x-5\\ 1 & -14x-4 & 9x^2+28x-5 & 0\end{array}\right|=$$

$$=24(x^4-x^3-4x^2+4x).$$

Найдем ее корни: $\alpha_1=0,\alpha_2=1,\alpha_3=2,\alpha_4=-2$.

Итак, найдены x-компоненты решений системы. Как найти их y-компоненты? Можно построить вторую элиминанту $\mathcal{Y}(y)$, отыскать ее корни, составить всевозможные пары из корней $\mathcal{X}(x)$ и $\mathcal{Y}(y)$, подставить их в $f(x,y)$ и $g(x,y)$ и проверить на равенство нулю. Либо же найденный корень $x=\alpha$ подставить в одно из уравнений: $f(\alpha,y)=0$, решить его относительно y, и каждую полученную таким образом пару подставить в $g(x,y)$; хотя бы одна из них должна удовлетворить уравнению $g(x,y)=0$. На этих путях нас ожидает следующее препятствие: *как правило*, корни элиминант невозможно установить **точно**; погрешность же вычислений может повлечь ошибку при выявлении истинной пары (α,β).

Хотелось бы минимизировать использование численных методов. Удовлетворить это желание поможет один из результатов первой части. Вспомним, что при $x=\alpha$ полиномы $f(\alpha,y)$ и $g(\alpha,y)$ имеют общий корень, а, следовательно, существует нетривиальный $\operatorname{НОД}(f(\alpha,y),g(\alpha,y))$. Степень этого НОД и его аналитическое выражение через коэффициенты полиномов $f(\alpha,y)$ и $g(\alpha,y)$ можно найти с помощью теории субрезультантов.

Так, если первый субрезультант полиномов $f(\alpha,y)$ и $g(\alpha,y)$ не обращается в нуль: $\mathcal{R}^{(1)}(\alpha)\neq 0$, то НОД будет первой степени и его выражение[6]:

$$\mathcal{R}^{(1)}(\alpha)y+\det M_1^{(1)}(\alpha). \tag{5.7}$$

Для нашего примера

$$\left|\begin{array}{cc} 1 & -7x-2 \\ 1 & -14x-4 \end{array}\right| y+\left|\begin{array}{cc} 1 & 4x^2+13x-3 \\ 1 & 9x^2+28x-5 \end{array}\right| = (-7x-2)y+(5x^2+15x-2),$$

$\mathcal{R}^{(1)}(x)=-7x-2\neq 0$ при $x=0,1,2,-2$. Следовательно, соответствующие этим корням значения y находятся по формуле (3.3):

$$y=-\det M_1^{(1)}(x)/\mathcal{R}^{(1)}(x). \tag{5.8}$$

Ответ. Решения системы: $(1,2)$; $(2,3)$; $(0,-1)$; $(-2,1)$.

Теорема 5.3. *При выполнении условий (5.4) и*

$$\mathcal{R}(\mathcal{X}(x),\mathcal{R}^{(1)}(x))\neq 0 \tag{5.9}$$

система (5.1) может быть сведена к эквивалентной ей (т.е. имеющей такое же множество решений) системе

$$\mathcal{X}(x)=0,\quad \mathcal{R}^{(1)}(x)y+\det M_1^{(1)}(x)=0. \tag{5.10}$$

[6] См. теорему 3.4.

Действительно, условие (5.9) эквивалентно тому, что $\mathcal{X}(x)$ и $\mathcal{R}^{(1)}(x)$ не имеют общих корней и, следовательно, $\mathcal{R}^{(1)}(\alpha) \neq 0$.

Что происходит при нарушении условия (5.9)?

Пример 5.2. Решить систему уравнений

$$\left\{ \begin{array}{l} f(x,y) = x^3 - 2x^2y - 4xy^2 + 2y^3 + 6x^2 + 12xy - 16x - 8y = 0, \\ g(x,y) = -3x^3 - 4x^2y - 3xy^2 + 4y^3 + 2x^2 + 24xy - 10y^2 - 12x - 16y + 40 = 0. \end{array} \right.$$

Решение. Найдем элиминанту $\mathcal{X}(x)$, первый субрезультант и определитель матрицы $M_1^{(1)}$:

$$\begin{gathered} \mathcal{X}(x) = -750x^2(3x^7 + 2x^6 - 120x^5 + 112x^4 + 1136x^3 - 2400x^2 + 256x + 1536) \\ = -750x^2(x + {}^2\!/_3)(x+4)(x-4)(x+6)(x-2)^3, \\ \mathcal{R}^{(1)}(x) = -1000x(x-2)^2, \ \det M_1^{(1)}(x) = 50x(x-2)^2(3x^2+10x+32). \end{gathered}$$

На корнях $-{}^2\!/_3$, -4, 4, -6 элиминанты $\mathcal{X}(x)$ формула (5.8) дает истинные значения y-компоненты: ${}^4\!/_3$, 2, 6, 4 соответственно.

Однако на корнях $x = 0$ и $x = 2$ имеем $\mathcal{R}^{(1)}(x) = 0$, и формула (5.8) неприменима — хотя формула (5.10) продолжает оставаться справедливой, обращаясь в тождество вида $0 \equiv 0$. (Попытка использования формулы (5.8) после удаления общего множителя $x(x-2)^2$ оканчивается неудачей: так, пара $(90, {}^8\!/_5)$ не является решением системы.)

По-видимому, при этих значениях α полиномы $f(\alpha, y)$ и $g(\alpha, y)$ имеют более одного общего корня, и для нахождения НОД $(f(\alpha,y), g(\alpha,y))$ воспользуемся вторым субрезультантом:

$$\begin{aligned} &\mathcal{R}^{(2)}(x) = 10(x-2), \ \det M_2^{(1)}(x) \equiv 0, \\ &\det M_2^{(2)}(x) = -10(x^3 + 2x^2 - 4x - 8) = -10(x-2)(x+2)^2. \end{aligned}$$

Для соответствующих y-компонент получаем уравнение

$$\mathcal{R}^{(2)}(\alpha)y^2 + \det M_2^{(1)}(\alpha)y + \det M_2^{(2)}(\alpha) = 0. \qquad (5.11)$$

Для нашего примера $(\alpha-2)y^2 - (\alpha-2)(\alpha+2)^2 = 0$, и $\mathcal{R}^{(2)}(\alpha) \neq 0$ при $\alpha = 0$; так что для этого значения x уравнение (5.11) дает два верных значения y-компоненты: $y^2 - 4 = 0$. В самом деле, $(0,2)$ и $(0,-2)$ — решения системы. Более того, уравнение (5.11) остается справедливым и на группе решений, полученной на предыдущем этапе. Однако при $\alpha = 2$ оно обращается в тождество вида $0 \equiv 0$ (не помогает даже удаление общего множителя $(x-2)$: точки $(2,4)$ и $(2,-4)$ решениями системы не являются).

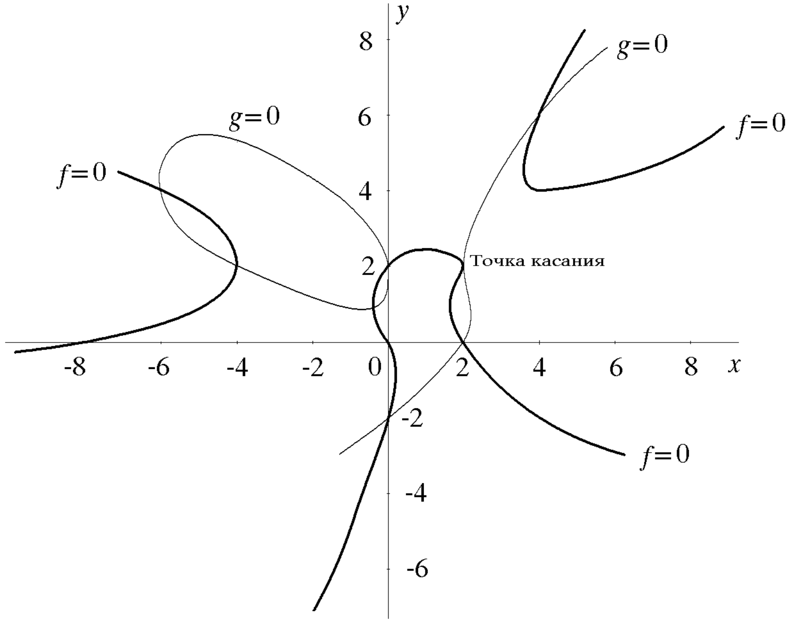


Рис. 3

По-видимому, при этом α полиномы $f(\alpha,y)$ и $g(\alpha,y)$ имеют более двух общих корней, и для нахождения НОД $(f(\alpha,y), g(\alpha,y))$ следует воспользоваться третьим субрезультантом. Но он не существует (в матрице M вычеркивать нечего). Это неудивительно — НОД $(f(2,y), g(2,y))$ должен иметь степень не меньшую трех, но сами исходные полиномы $f(2,y)$ и $g(2,y)$ — как раз третьей степени

$$f(2,y)=2(y^3-4y^2+4y)=2y(y-2)^2, \ g(2,y)=4(y^3-4y^2+4y)=4y(y-2)^2,$$

и их НОД совпадает с любым из них. Следовательно, третью (и последнюю) группу решений составляют пары $(2,0)$ и $(2,2)$, причем $(2,2)$ оказывается *кратным*, так как якобиан (5.2) обращается на нем в нуль. Такое решение будем считать за два.

Геометрический смысл: абсциссе $x=0$ соответствуют две ординаты точек пересечения кривых $f(x,y)=0$ и $g(x,y)=0$: $y=-2$

и $y = 2$. Абсциссе $x = 2$ также соответствуют два значения y: $y = 0$ и $y = 2$. Оказывается, что в точке $(2, 2)$ кривые не пересекаются, но соприкасаются (рис. 3).

Ответ. Система имеет 9 решений (с учетом кратностей):

$$(-{}^2/_3, {}^4/_3),\ (-4, 2),\ (4, 6),\ (-6, 4),\ (0, 2),\ (0, -2),\ (2, 0),\ (2, 2),\ (2, 2).$$

Замечание 5.1. В отличие от теоремы 5.3, исходную систему удалось свести к эквивалентной системе, имеющей, однако, бóльшее число уравнений

$$\begin{gathered} \mathcal{X}(x) = 0;\ \mathcal{R}^{(1)}(x)y + \det M_1^{(1)}(x) = 0; \\ \mathcal{R}^{(2)}(x)y^2 + \det M_2^{(1)}(x)y + \det M_2^{(2)}(x) = 0; \dots \end{gathered} \tag{5.12}$$

Структура системы (5.12): первое уравнение от y не зависит; второе — зависит линейно; третье — квадратично и т.д.

Упражнение 68. *Решить системы уравнений*

а) $$\begin{aligned} &x^2 + y^2 - 3x - y = 0, \\ &-x^2 - 6xy + y^2 + 7x + 11y - 12 = 0; \end{aligned}$$

б) $$\begin{aligned} &5x^2 - 6xy + 5y^2 - 16 = 0, \\ &2x^2 - xy + y^2 - x - y - 4 = 0; \end{aligned}$$

в) $$\begin{aligned} &x^2 + xy + y^2 - 2x - 4y + 3 = 0, \\ &x^3 + y^3 - x^2 + xy - 5y^2 - 5x + 7y - 3 = 0. \end{aligned}$$

§6. Число решений системы уравнений

Каково общее число решений системы (5.1)? На основании теорем 5.1 и 5.2 видим, что оно совпадает с $\deg \mathcal{X}(x)$ (если принимать во внимание все корни последнего, включая невещественные и кратные с учетом кратности).

Теорема Безу

Теорема 6.1. *[Безу] Пусть выполнено условие (5.4). Тогда, как правило*

$$\deg \mathcal{X}(x) = \deg f(x, y) \cdot \deg g(x, y) = nm. \tag{6.1}$$

Перейти к оглавлению на странице: 460

Доказательство приведем для случая $n = 3$ и $m = 2$.

$$f(x,y) = A_0 y^3 + A_1(x)y^2 + A_2(x)y + A_3(x),$$
$$g(x,y) = B_0 y^2 + B_1(x)y + B_2(x),$$
$$\mathcal{X}(x) = \begin{vmatrix} A_0 & A_1(x) & A_2(x) & A_3(x) & \\ & A_0 & A_1(x) & A_2(x) & A_3(x) \\ & & B_0 & B_1(x) & B_2(x) \\ & B_0 & B_1(x) & B_2(x) & \\ B_0 & B_1(x) & B_2(x) & & \end{vmatrix}.$$

Здесь $A_0 = a_{03}, B_0 = b_{02}; \deg A_j(x) \le j;\ \deg B_j(x) \le j\ (j = 1,2,3)$. Старший моном $\mathcal{X}(x)$ образуется из старших мономов элементов определителя. Выделим их

$$\begin{array}{ll} A_0 = a_{03}; & B_0 = b_{02}; \\ A_1(x) = a_{12}x + \dots; & B_1(x) = b_{11}x + \dots; \\ A_2(x) = a_{21}x^2 + \dots; & B_2(x) = b_{20}x^2 + \dots; \\ A_3(x) = a_{30}x^3 + \dots & \end{array}$$

Следовательно,

$$\mathcal{X}(x) = \begin{vmatrix} a_{03} & a_{12}x & a_{21}x^2 & a_{30}x^3 & \\ & a_{03} & a_{12}x & a_{21}x^2 & a_{30}x^3 \\ & & b_{02} & b_{11}x & b_{20}x^2 \\ & b_{02} & b_{11}x & b_{20}x^2 & \\ b_{02} & b_{11}x & b_{20}x^2 & & \end{vmatrix} + \dots,$$

и нам осталось извлечь степень x из первого определителя. Проделаем это с помощью процедуры, которую можно обобщить на случай произвольных полиномов $f(x,y)$ и $g(x,y)$: домножим вторую и четвертую строки на x, третью — на x^2:

$$= \frac{1}{x^4} \begin{vmatrix} a_{03} & a_{12}x & a_{21}x^2 & a_{30}x^3 & \\ & a_{03}x & a_{12}x^2 & a_{21}x^3 & a_{30}x^4 \\ & & b_{02}x^2 & b_{11}x^3 & b_{20}x^4 \\ & b_{02}x & b_{11}x^2 & b_{20}x^3 & \\ b_{02} & b_{11}x & b_{20}x^2 & & \end{vmatrix} + \dots$$

Столбцы делятся на $\quad \overset{\uparrow}{x} \quad \overset{\uparrow}{x^2} \quad \overset{\uparrow}{x^3} \quad \overset{\uparrow}{x^4}$

Выносим указанные множители из определителя

$$= \frac{x^{1+2+3+4}}{x^4} \begin{vmatrix} a_{03} & a_{12} & a_{21} & a_{30} & \\ & a_{03} & a_{12} & a_{21} & a_{30} \\ & & b_{02} & b_{11} & b_{20} \\ & b_{02} & b_{11} & b_{20} & \\ b_{02} & b_{11} & b_{20} & & \end{vmatrix} + \ldots =$$

$$= x^6 \mathcal{R}(a_{03}y^3 + a_{12}y^2 + a_{21}y + a_{30}, b_{02}y^2 + b_{11}y + b_{20}) + \ldots$$

А для произвольных m и n получим

$$\mathcal{X}(x) = \mathcal{A}_0 x^{nm} + \mathcal{A}_1 x^{nm-1} + \cdots + \mathcal{A}_{nm} \qquad (6.2)$$

и аналогично

$$\mathcal{Y}(y) = \mathcal{B}_0 y^{nm} + \mathcal{B}_1 y^{nm-1} + \ldots + \mathcal{B}_{nm}$$

при

$$\mathcal{A}_0 \stackrel{\text{def}}{=} \mathcal{R}(f_n(1,y), g_m(1,y)) \text{ и } \mathcal{B}_0 \stackrel{\text{def}}{=} \mathcal{R}(f_n(x,1), g_m(x,1)). \qquad (6.3)$$

Упражнение 69. *Доказать, что старшие коэффициенты* $\mathcal{X}(x)$ *и* $\mathcal{Y}(y)$ *совпадают с точностью до знака.*

Упражнение 70. *Доказать, что*

$$\mathcal{A}_{nm} = \mathcal{R}(f(0,y), g(0,y)) \text{ и } \mathcal{B}_{nm} = \mathcal{R}(f(x,0), g(x,0)).$$

Замечание 6.1. Итак, мы выяснили смысл выражения “*как правило*” из формулировки теоремы Безу: если число $\mathcal{A}_0$, определяемое формулой (6.3), отлично от нуля. Подчеркнем, что это число зависит только от старших форм (5.3) в разложениях $f(x,y)$ и $g(x,y)$.

Упражнение 71. *Установить структуру множества решений системы* (5.1) *при*

$$f(x,y) \stackrel{\text{def}}{=} \prod_{j=1}^{n} (a_j x + b_j y + c_j), \ g(x,y) \stackrel{\text{def}}{=} \prod_{k=1}^{m} (A_k x + B_k y + C_k).$$

При каком условии система имеет бесконечное множество решений?

Перейти к оглавлению на странице: 460

Замечание 6.2. Упражнение 71 и теорема 11.11 главы 3 позволяют дать следующую геометрическую интерпретацию теоремы Безу. Алгебраические кривые $f(x,y)=0$ и $g(x,y)=0$ можно рассматривать как сильно деформированные семейства прямых, причем если в произвольной ограниченной области плоскости деформация еще заметна, то на бесконечности кривые продолжают вести себя "почти как прямые". Теорема Безу утверждает, что число точек пересечения алгебраических кривых остается инвариантным при деформации этих кривых — по крайней мере, до тех пор, пока асимптоты не станут параллельными.

Исключительные случаи теории исключения

Нас будут интересовать случаи, когда либо обращается в нуль числа (6.3), либо нарушаются условия (5.4).

I. Пусть $\mathcal{A}_0 = \mathcal{R}(f_n(1,y), g_m(1,y)) = 0$. Возможны следующие случаи:

а) Число решений системы понижается.

Пример 6.1. Решить систему уравнений

$$\left\{ \begin{array}{l} f(x,y) = x^3 + 3x^2y + 3xy^2 + y^3 - yx - y^2 + x + 2y = 0, \\ g(x,y) = x^2 + 2xy + y^2 - y = 0. \end{array} \right.$$

Решение. Элиминанты этой системы имеют степени меньшие, чем оценка Безу:

$$\mathcal{X}(x) = x(x+2),\ \mathcal{Y}(y) = y(y-1), \quad \deg \mathcal{X} = \deg \mathcal{Y} = 2 < mn = 6.$$

Действуя согласно алгоритму §5 , получаем систему в виде двух уравнений

$$\mathcal{X}(x) = 0, \quad -2y - x = 0,$$

эквивалентную исходной. Последняя имеет два решения: $(0,0), (-2,1)$. $\triangle$

Теорема 6.2. *Пусть* $\deg(\text{НОД}\,(f_n(1,y), g_m(1,y))) = d$. *Тогда число решений системы уменьшается по крайней мере на* d.

Теорема дает лишь достаточное условие понижения степени, как видно из примера 6.1 или 6.2.

Перейти к оглавлению на странице: 460

Пример 6.2. Решить систему уравнений

$$\left\{\begin{array}{l} f(x,y)=x^2-4xy+3y^2-\dfrac{3}{13}x+\dfrac{45}{13}y+\dfrac{126}{169}=0,\\ g(x,y)=x^2-xy-6y^2+\dfrac{16}{13}x+\dfrac{42}{13}y+\dfrac{1266}{169}=0. \end{array}\right.$$

Решение. Здесь

$$\mathcal{X}(x)=\frac{378}{169}\left(-6\,x^2+\frac{8701}{13}x+\frac{726}{169}\right),$$

и хотя $\deg\mathcal{X}$ понижается на 2 по отношению к оценке Безу, при

$$\mathcal{A}_0=\mathcal{R}(f_n(1,y),g_m(1,y))=\mathcal{R}(1-4y+3y^2,1-y-6y^2)=0,$$

тем не менее, $\mathcal{R}^{(1)}(f_n(1,y),g_m(1,y))\neq 0$. На понижение степени элиминанты здесь влияют младшие формы разложений $f(x,y)$ и $g(x,y)$.

Ответ. Система имеет 2 решения:

$$\left(\frac{11}{156}(791\pm 5\sqrt{25033}),\ \frac{1}{156}(2869\pm 19\sqrt{25033})\right).$$

Заметим, что для этой системы свободные члены в разложениях $f(x,y)$ и $g(x,y)$ не влияют на число решений, но влияют на их вид.

Пример 6.3. Решить систему уравнений

$$\left\{\begin{array}{l} f(x,y)=x^2-4\,xy+3\,y^2-\dfrac{3}{13}x+\dfrac{45}{13}y=0,\\ g(x,y)=x^2-xy-6y^2+\dfrac{16}{13}x+\dfrac{42}{13}y=0. \end{array}\right.$$

Решение. Здесь

$$\mathcal{X}(x)=\frac{81}{169}\left(606x^2-\frac{4136}{13}x\right).$$

Ответ. Система имеет 2 решения:

$$(0,0),\ \left(\frac{2068}{3939},-\frac{893}{3939}\right).$$

б) Число решений становится бесконечным.

Пример 6.4. Решить систему уравнений

$$\left\{ \begin{array}{l} f(x,y) = x^2 - 2xy + y^2 - 1 = 0, \\ g(x,y) = x^2 - y^2 + 2x + 1 = 0. \end{array} \right.$$

Решение. Здесь $\mathcal{X}(x) \equiv 0$, $\mathcal{Y}(y) \equiv 0$. Алгоритм из §8 дает систему

$$0 = 0,\ (-x)y + (x^2 + x) = 0,\ g(x,y) = 0,$$

которая имеет бесконечное множество решений: $(\alpha, \alpha + 1)$ при любом $\alpha \in \mathbb{C}$ и дополнительно точку $(0, -1)$.

Геометрический смысл: алгебраические кривые $f(x,y)=0$ и $g(x,y)=0$ имеют общую ветвь — прямую $y = x + 1$. $\triangle$

в) Система решений не имеет (несовместна).

Пример 6.5. Решить систему уравнений

$$f(x,y) = x^2 - y^2 + 1 = 0,\ g(x,y) = x - y = 0.$$

Решение. Здесь $\mathcal{X}(x) \equiv 3 \neq 0$ ни при одном $x \in \mathbb{C}$.

Ответ. Система несовместна.

Рассмотренными примерами исчерпываются все возможные случаи для числа решений системы (5.1): оно либо бесконечно, либо не превосходит mn.

II. Пусть нарушено условие (5.4).

Чисто формально это может привести к невозможности конструкции $\mathcal{X}(x)$ по формуле (5.6): если $a_{0n} = b_{0m} = 0$, то $\mathcal{X}(x) \equiv 0$. В этом случае степени полиномов $f(x,y)$ и $g(x,y)$, рассматриваемых как полиномы от y, становятся меньшими, чем n и m соответственно. Это обстоятельство необходимо учитывать при построении результанта в виде определителя матрицы M, так как ее порядок должен понизиться.

Пример 6.6. Решить систему уравнений

$$f(x,y) = xy - 1 = 0,\ g(x,y) = x^2 + 2xy - 1 = 0.$$

Решение. Вычисляя $\mathcal{R}_y(f,g)$ как результант полиномов первой степени, запишем:

$$\mathcal{X}(x) = x(x^2+1) \ \Rightarrow \ \alpha_1 = 0, \alpha_{2,3} = \pm \boldsymbol{i} \, .$$

Однако $\alpha_1 = 0$ не соответствует ни одному решению: $f(0,y) = g(0,y) = -1 \ne 0$. Только значения $\alpha_{2,3}$ определяют решения системы: им соответствуют $\beta_{2,3} = \mp \boldsymbol{i}$. В объяснение факта появления "лишнего" корня у элиминанты обратим внимание на то, что при $x = 0$ степени полиномов f и g понижаются, и этот эффект проявляется при построении элиминанты в виде определителя матрицы M. Заметим, что вторая элиминанта $\mathcal{Y}(y) = y^2+1$ не имеет "лишнего" корня.

Ответ. Система имеет 2 решения: $(\pm \boldsymbol{i}\, , \mp \boldsymbol{i}\,)$.

Пример 6.7. Решить систему уравнений

$$f(x,y) = xy - 1 = 0, \ \ g(x,y) = x^2y + x - 2 = 0.$$

Решение. Каждая из элиминант этой системы

$$\mathcal{Y}(y) = -2y(y-1), \ \ \mathcal{X}(x) = 2x(x-1)$$

имеет "лишний" корень, в результате порождается "ложное" решение $(0,0)$. Причина эффекта та же, что и в предыдущем примере. Единственным решением системы будет $(1,1)$. $\triangle$

Для контроля подобных случаев необходимо проверять подозрительные значения переменных, т.е. те из них, которые понижают степени исходных уравнений.

Пример 6.8. Решить систему уравнений

$$\left\{ \begin{array}{l} f(x,y) = y^2x^2 - xy^2 - 2y^2 + yx^3 + 2x^2y + 3xy + 2x - 4y + 10 = 0, \\ g(x,y) = x^2y - 3xy + 2y + 2x^3 - 6 = 0. \end{array} \right.$$

Решение. Раскладываем полиномы системы по степеням y

$$\begin{array}{rcl} f(x,y) &=& (x^2-x-2)y^2 + (x^3+2\,x^2+3\,x-4)y + 2\,x + 10 \, , \\ g(x,y) &=& (x^2-3\,x+2)y + 2\,x^3 - 6 \end{array}$$

и составляем элиминанту по x:

$$\begin{array}{rcl} \mathcal{X}(x) &=& 2x^8 - 2x^7 - 6x^6 + 2x^5 - 20x^4 + 24x^3 + 88x^2 - 40x - 80 = \\ &=& 2(x-2)(x+1)(x^2-2)(x^4+x^2-10). \end{array}$$

Ищем подозрительные корни $\mathcal{X}(x)$:

$$\text{НОД}\,\big(\mathcal{X}(x),(x^2-x-2)\big)=x^2-x-2=(x-2)(x+1).$$

Итак, при $x=-1$ и $x=2$ степень полинома $f(x,y)$ по y понижается. Поскольку $f(-1,y)=-6\,y+8$, $g(-1,y)=6\,y-8$, то корню $x=-1$ соответствует решение системы $(-1,\ {}^4\!/_3)$. Что же касается корня $x=2$, то здесь ситуация иная: $f(2,y)=18\,y+14$, $g(2,y)=10$ и система несовместна.

Ответ. Решения системы:

$$(-1,\ {}^4\!/_3);\ (\pm\sqrt{2},\mp\sqrt{2});$$

$$\left(\pm\,{}^1\!/_2\sqrt{-2+2\sqrt{41}};\ {}^3\!/_4\sqrt{41}+\ {}^9\!/_4\pm\ {}^1\!/_4\sqrt{242+38\sqrt{41}}\right);$$

$$\left(\pm\,{}^1\!/_2\,\boldsymbol{i}\ \sqrt{2+2\sqrt{41}},\ -\,{}^3\!/_4\ \sqrt{41}+\ {}^9\!/_4\pm\boldsymbol{i}\ {}^1\!/_4\sqrt{-242+38\sqrt{41}}\right).$$

Упражнение 72. *Решить системы уравнений*

а) $$x^2+y^2-9=0,\quad x^2+y^2+2\,x-y-3=0;$$

б) $$5x^2-5y^2-3x+9y=0,\quad 5x^3+5y^3-15x^2-13xy-y^2=0;$$

в) $$3x^3+9x^2y+9xy^2+3y^3+2x^2-4xy+2y^2-5=0,\quad 4x^3+12x^2y+12xy^2+4y^3-x^2+2xy-y^2-3=0.$$

§7. Замечания

Об эквивалентных системах

В §5 мы строили системы, эквивалентные исходной системе (5.1). Эти эквивалентные системы имели специальный вид (5.12): первое уравнение зависело только от x, второе зависело от y линейно, возможное третье — квадратично, и т.д. Нас теперь интересует вопрос о единственности системы такого вида, эквивалентной исходной.

Обратимся к примеру 5.1. Эквивалентная система была найдена в следующем виде:

$$\mathcal{X}(x)/24 = x^4 - x^3 - 4x^2 + 4x = 0,\ (-7x-2)y + (5x^2+15x-2) = 0,$$

и поскольку выполнено условие (5.9), то можно воспользоваться формулой (5.8) для получения y-компоненты:

$$y = \frac{5x^2+15x-2}{7x+2}. \tag{7.1}$$

Воспользуемся теперь результатами §4. По теореме 4.1, на корнях $\mathcal{X}(x)$ дробь (7.1) будет принимать такие же значения, что и некоторый полином $G(x)$, при этом $G(x)$ можно подобрать степени меньшей $\deg \mathcal{X}$. Построим такой полином по алгоритму теоремы 4.1:

$$\mathcal{R}(x^4 - x^3 - 4\,x^2 + 4\,x,\ 7\,x+2) = -3456,$$

$$\tilde{u}(x) = \left| \begin{array}{ccccc} 1 & -1 & -4 & 4 & 0 \\ 0 & 0 & 0 & 7 & 1 \\ 0 & 0 & 7 & 2 & x \\ 0 & 7 & 2 & 0 & x^2 \\ 7 & 2 & 0 & 0 & x^3 \end{array} \right| = -343\,x^3 + 441\,x^2 + 1246\,x - 1728.$$

Остаток от деления $\tilde{u}(x)(5x^2+15x-2)$ на $x^4-x^3-4x^2+4x$ равен

$$2016x^3 - 2592x^2 - 9792x + 3456$$

и $G(x)$ отличается от него только на множитель $-1/3456$. Итак, еще одной системой, эквивалентной исходной, будет

$$x^4 - x^3 - 4x^2 + 4x = 0,\ y + (7x^3 - 9x^2 - 34x + 12)/12 = 0.$$

Теорема 7.1. *При выполнении условия* (5.9) *существует полиномиальная система уравнений вида*

$$\mathcal{X}(x) = 0,\ y - G(x) = 0 \quad (\deg G(x) < \deg \mathcal{X}), \tag{7.2}$$

эквивалентная исходной системе (5.1).

Перейти к оглавлению на странице: 460

Замечание 7.1. Условие (5.9) будет выполнено, если $\mathcal{X}(x)$ не имеет кратных корней, т.е.

$$\mathcal{D}(\mathcal{X}) \neq 0 \ \Rightarrow \ \mathcal{R}(\mathcal{X}(x), \mathcal{R}^{(1)}(x)) \neq 0$$

(обратное, вообще говоря, неверно). При этом условии все решения системы различны и имеют различные x-компоненты.

Об эквидистанте

Определение 7.1. Рассмотрим гладкую кривую K, в каждой ее точке A проведем перпендикуляр и возьмем на этом перпендикуляре точки, находящиеся на некотором фиксированном расстоянии h от точки A. Полученные точки формируют две кривые, каждую из которых назовем **эквидистантой** кривой K и будем обозначать K_{+h} и K_{-h}.

Эквидистанты имеют очевидный физический смысл. Если предположить, что каждая точка кривой является источником излучения, то эквидистанта представляет собой волновой фронт[7].

Задача. Построить уравнение эквидистант графика $y = f(x)$, где $f(x) \in \mathbb{R}[x]$.

Теорема 7.2. *Эквидистанты K_h и K_{-h} задаются уравнением*

$$\Phi(x,y) = 0, \quad \Phi(x,y) \stackrel{\text{def}}{=} \mathcal{D}_X \left([X-x]^2 + [f(X)-y]^2 - h^2 \right) .$$

Здесь дискриминант берется по переменной X, в то время как остальные переменные считаются параметрами.

Доказательство. Пусть точка (x,y) эквидистанты находится на расстоянии h от ближайшей к ней точки $(X, Y = f(X))$ кривой K. Тангенс угла наклона касательной к графику $Y = f(X)$ равен $f'(X)$. Следовательно, по определению эквидистанты, вектор $(X - x, Y-y)$ должен быть перпендикулярен направляющему вектору касательной. Таким образом, получаем систему из трех уравнений

$$(x-X)^2 + (y-Y)^2 = h^2, \ Y = f(X), \ (x-X) + f'(X)(y-Y) = 0 \, .$$

[7] В предположении изотропности физической среды, из которой состоит плоскость.

Среднее из этих условий позволяет исключить переменную Y из первого и третьего уравнения:

$$(x-X)^2+(y-f(X))^2-h^2=0,\ (x-X)+f'(X)(y-f(X))=0\ .$$

Теперь для исключения переменной X мы должны были бы составить элиминанту по переменной X. Очевидно, однако, что второе уравнение является (с точностью до множителя) производной первого по X. ■

Пример 7.1. Найти уравнение эквидистант параболы $y=x^2$.

Решение. После вычисления дискриминанта, отбросим общий множитель его коэффициентов и сгруппируем по степеням h:

$$\begin{aligned}\Phi(x,y) &= \mathcal{D}_X\left(X^4+(1-2y)X^2-2\,xX+x^2+y^2-h^2\right)=\\ &= (16\,y^2+16\,x^2-8\,y+1)(y-x^2)^2+\\ &+ \left[8(-4y^2-8yx^2-y+1-8x^4)(y-x^2)-(4x^2+1)^3\right]h^2+\\ &+ 8\,(2\,y^2+4\,y+6\,x^2-1)h^4-16\,h^6.\end{aligned}$$

Уравнение $\Phi(x,y)=0$ и дает искомые эквидистанты K_h и K_{-h} для параболы $y=x^2$. На рис. 4 показаны эквидистанты параболы для $h=1$

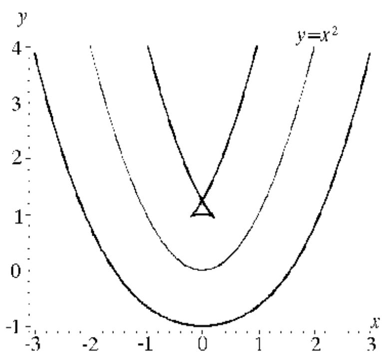


Рис. 4

(чтобы не загромождать график, мы сдвинули координатные оси на несколько единиц). △

Перейти к оглавлению на странице: 460

Упражнение 73. *Рассмотрим снова гладкую кривую* K, *которую теперь — в отличие от только что рассмотренной задачи об эквидистанте — будем считать не источником излучения, а зеркалом.* **Кривой зеркального изображения** *назовем геометрическое место точек, являющихся зеркальными изображениями источника излучения* (x_0, y_0) *относительно касательных к точкам кривой* K. *Доказать, что в случае, когда кривая* K *задана уравнением* $y = f(x)$ *при* $f(x) \in \mathbb{R}[x], \deg f > 1$, *кривая зеркального изображения определяется уравнением* $\Phi(x, y) = 0$, *где*

$$\Phi(x,y) \stackrel{\text{def}}{=} \mathcal{D}_X \left(2X(x - x_0) + 2f(X)(y - y_0) - (x^2 - x_0^2) - (y^2 - y_0^2)\right).$$

Построить это уравнение для $(x_0, y_0) = (0, 0)$, $f(x) = x^2 + 1$.

§8. Основная теорема высшей алгебры

Вводные замечания

Теорема 8.1. *Полином* $f(z) = a_0z^n + a_1z^{n-1} + \dots + a_n$ $(a_0 \ne 0,\ n \ge \ge 1)$ *с комплексными коэффициентами всегда имеет хотя бы один комплексный корень.*

В таком виде теорема была сформулирована и доказана в 1746 году независимо французом Даламбером и швейцарцем Эйлером. Предложенный ими доказательства были подвергнуты критике со стороны Гаусса, который, в свою очередь, предложил четыре новых доказательства. В XIX веке были предложены еще несколько доказательств (Кэли, Вейерштрасс), различающиеся между собой “степенью алгебраичности”. Дело в том, что основная теорема высшей алгебры не может быть доказана средствами только одной алгебры. Все известные доказательства основываются на свойствах полинома, которые не могут быть доказаны с помощью алгоритмов, использующих конечное число элементарных операций над коэффициентами полинома. Приходится использовать предельные переходы. И проблема сводится к тому, чтобы свести число таких “неалгебраических” мест до минимума.

В настоящем параграфе приводится доказательство, использующее лишь единственный подобный факт — следствие 3 к теореме

Больцано[8].

Лемма 8.1. *Полином $f(z)$ нечетной степени с вещественными коэффициентами имеет по крайней мере один вещественный корень.*

Еще один результат мы позаимствуем из §1.2.

Лемма 8.2. *Полиномы $f(z)$ и $g(z)$ с комплексными коэффициентами имеют нетривиальный* НОД *тогда и только тогда, когда их результант $\mathcal{R}(f,g)$ равен нулю.*

Доказательство основной теоремы

Очевидно теорема справедлива для $n \le 2$.

I. Сведение к системе уравнений. Пусть $z = x + \boldsymbol{i}\, y$, причем x и y в общем случае *комплексные*[9]. Воспользуемся формулой Тейлора:

$$f(z)=f(x+\boldsymbol{i}\, y)=f(x)+\frac{f'(x)}{1!}\boldsymbol{i}\, y+\frac{f''(x)}{2!}(\boldsymbol{i}\, y)^2+\dots+\frac{f^{(n)}(x)}{n!}(\boldsymbol{i}\, y)^n=$$

$$=\left[f(x)-\frac{f''(x)}{2!}y^2+\frac{f^{(4)}(x)}{4!}y^4-\dots\right]+$$

$$+\boldsymbol{i}\, y\left[\frac{f'(x)}{1!}-\frac{f'''(x)}{3!}y^2+\frac{f^{(5)}(x)}{5!}y^4-\dots\right]=$$

$$=F_1(x,Y)+\boldsymbol{i}\, yF_2(x,Y) \quad \text{где} \quad Y \stackrel{\text{def}}{=} -y^2, \tag{8.1}$$

$$\begin{cases} F_1(x,Y) \stackrel{\text{def}}{=} f(x)+\dfrac{f''(x)}{2!}Y+\dfrac{f^{(4)}(x)}{4!}Y^2+\dots, \\ F_2(x,Y) \stackrel{\text{def}}{=} \dfrac{f'(x)}{1!}+\dfrac{f'''(x)}{3!}Y+\dfrac{f^{(5)}(x)}{5!}Y^2+\dots \end{cases} \tag{8.2}$$

Полиномы F_1 и F_2 имеют комплексные коэффициенты.

Исключим из системы

$$F_1(x,Y)=0,\ F_2(x,Y)=0 \tag{8.3}$$

[8] Глава 3, §7.3.

[9] Вовсе не обязательно, чтобы $x = \mathrm{Re}\, z,\ y = \mathrm{Im}\, z$.

переменную Y, составив элиминанту по x:

$$\mathcal{X}(x) \stackrel{\text{def}}{=} \mathcal{R}_Y(F_1, F_2). \qquad (8.4)$$

Лемма 8.3. $\deg \mathcal{X} = N \stackrel{\text{def}}{=} n(n-1)/2$.

Действительно, по аналогии с доказательством теоремы 6.1 (Безу), можно установить старший член разложения $\mathcal{X}$ по степеням x:

$$\mathcal{X}(x) = \mathcal{A}_0 x^N + \dots\ , \qquad \text{где } \mathcal{A}_0 = (-1)^K 2^N a_0^{n-1},\ K = \left\lfloor \frac{n}{4} \right\rfloor. \qquad (8.5)$$

Здесь $\lfloor\ \rfloor$ означает целую часть числа.

Пример 8.1. Для $n=6$ и $a_0=1$ имеем:

$$\mathcal{X}(x) = -32768\, x^{15} - 81920\, a_1 x^{14} - (32768 a_2 + 81920\, a_1^2) x^{13} + \dots$$

II. Пусть коэффициенты $f(z)$ вещественны и $n = \deg f = 2\eta$ при η нечетном.

В рассуждениях пункта **I** доказательства возьмем x вещественным. Тогда построенные полиномы $F_1(x,Y)$ и $F_2(x,Y)$ будут иметь вещественные коэффициенты, равно как и элиминанта $\mathcal{X}(x)$. Далее, число $N = \deg \mathcal{X}(x) = \eta(2\eta-1)$ нечетно. По лемме 8.1, у полинома $\mathcal{X}(x)$ существует вещественный корень $x_0 : \mathcal{X}(x_0) = 0$.

Согласно лемме 8.2 это условие является необходимым и достаточным для существования нетривиального общего делителя у $F_1(x_0,Y)$ и $F_2(x_0,Y)$:

$$\mathfrak{D}(Y) \stackrel{\text{def}}{=} \text{НОД}\,(F_1(x_0,Y), F_2(x_0,Y)),\ \deg \mathfrak{D} \ge 1. \qquad (8.6)$$

При этом $\mathfrak{D} \in \mathbb{R}[Y]$. Если обозначить

$$\Phi_j(Y) \stackrel{\text{def}}{=} F_j(x_0,Y)/\mathfrak{D}(Y) \quad (j=1,2)$$

то и $\Phi_j \in \mathbb{R}[Y]$. Из формулы (8.1) следует

$$f(z) = f(x_0 + \boldsymbol{i}\, y) = F_1(x_0,Y) + \boldsymbol{i}\, y F_2(x_0,Y) = \mathfrak{D}(Y)[\Phi_1(Y) + \boldsymbol{i}\, y \Phi_2(Y)] =$$

Вспоминая, что $\boldsymbol{i}\, y = z - x_0$, получаем разложение

$$f(z) = \underbrace{\mathfrak{D}((z-x_0)^2)}_{f_1(z)} \underbrace{[\Phi_1((z-x_0)^2) + (z-x_0)\Phi_2((z-x_0)^2)]}_{f_2(z)}$$

полинома $f(z)$ на нетривиальные множители

$$f(z) = f_1(z)f_2(z) \quad \text{при } 1 \le n_j = \deg f_j < n,\ f_j(z) \in \mathbb{R}[z]. \tag{8.7}$$

Здесь оба числа n_1, n_2 четные, при этом хотя бы одно представимо в виде $n_j = 2\eta_j$ где η_j нечетное. (В противном случае $n = n_1 + n_2 = 2(\eta_1 + \eta_2)$ и $n/2$ было бы четно, что противоречит предположению этого пункта доказательства.) Тогда относительно соответствующего полинома $f_j(z)$ выполнены те же предположения, что и относительно $f(z)$. Применяем к $f_j(z)$ алгоритм пункта **а)** и раскладываем его на нетривиальные множители

$$f_j(z) = f_{j1}(z)f_{j2}(z) \quad \text{при } 1 \le n_{jk} = \deg f_{jk} < n_j,\ f_{jk}(z) \in \mathbb{R}[z]. \tag{8.8}$$

Продолжаем процесс далее. Получаем последовательность полиномов

$$f(z),\ f_j(z),\ f_{jk}(z), \dots, \tag{8.9}$$

каждый из которых является делителем предыдущего и степени их убывают:

$$n > n_j > n_{jk} > \dots$$

Это убывание завершится за конечное число шагов, и степень последнего полинома будет ≤ 2. Поскольку полином $\deg \le 2$ всегда имеет хотя бы один комплексный корень, то этот же корень будет корнем и любого полинома последовательности (8.9), в том числе и $f(z)$.

Во всех предшествующих рассуждениях неявно предполагалось, что $F_2(x_0, Y) \not\equiv 0$. Если это условие не выполняется, то $f(z) = F_1(x_0, Y) \in \mathbb{R}[Y]$. По предположению $\deg F_1(x_0, Y) = \eta$ — нечетна и на основании леммы 8.1 такой полином имеет вещественный корень.

III. Пусть коэффициенты $f(z)$ <u>комплексны</u> и $n = \deg f$ <u>нечетно</u>.

Определение 8.1. Для полинома $f(z) = a_0z^n + a_1z^{n-1} + \dots + a_n$ ему **сопряженным** назовем полином $f_*(z) = \overline{a_0}z^n + \overline{a_1}z^{n-1} + \ldots + \overline{a_n}$.

Справедливо очевидное равенство

$$f(\overline{z}) = \overline{f_*(z)} \tag{8.10}$$

Утверждается, что произведение $\mathfrak{F}(z) = f(z)f_*(z)$ будет полиномом по z с вещественными коэффициентами. Действительно, если выделить в коэффициентах полинома $f(z)$ вещественные и мнимые части, то его можно представить в виде $f(z) \equiv u(z) + \boldsymbol{i}\ v(z)$, где полиномы

$$u(z) = \operatorname{Re} e(a_0)z^n + \dots + \operatorname{Re} e(a_n) \text{ и } v(z) = \operatorname{Im} m(a_0)z^n + \dots + \operatorname{Im} m(a_n)$$

принадлежат $\mathbb{R}[z]$. Соответственно $f_*(z) \equiv u(z) - \boldsymbol{i}\ v(z)$. Но тогда

$$\mathfrak{F}(z) = f(z)f_*(z) = u(z)^2 + v(z)^2 \in \mathbb{R}[z].$$

Далее, $\deg \mathfrak{F}(z) = 2n$, и, по условию настоящего пункта, n нечетное. По доказанному в пункте **II** у $\mathfrak{F}(z)$ существует корень:

$$\exists z = z_0 \in \mathbb{C}: \ \mathfrak{F}(z_0) = 0 \Longrightarrow \left\{ \begin{array}{l} f(z_0) = 0 \\ \text{или} \\ f_*(z_0) = 0 \stackrel{(8.10)}{\Longrightarrow} f(\overline{z_0}) = 0 \end{array} \right.$$

т.е. какое-то из чисел $z_0, \overline{z_0}$ будет корнем $f(z)$. Теорема доказана и для этого случая.

IV. Пусть, наконец, коэффициенты $f(z)$ комплексны и $n = \deg f$ четно: $n = 2\eta$ при $\eta \in \mathbb{N}$.

Определение 8.2. Четное число $n \in \mathbb{N}$ называется числом **четности** $\mathfrak{r}$:

$$\operatorname{parity}(n) = \mathfrak{r} \quad \Longleftrightarrow \quad n \text{ делится на } 2^{\mathfrak{r}}, \ \text{ но не делится на } 2^{\mathfrak{r}+1}.$$

Будем полагать $\operatorname{parity}(n) = 0$ если n нечетно.

Пример 8.2. $\operatorname{parity}(6) = 1, \ \operatorname{parity}(8) = 3, \ \operatorname{parity}(12) = 2$.

Утверждается, что для элиминанты (8.4)

$$\operatorname{parity}(\deg \mathcal{X}(x)) = \mathfrak{r} - 1 \quad \text{если} \quad \operatorname{parity}(n) = \mathfrak{r}.$$

Действительно, $N = n(n-1)/2$ делится на $2^{\mathfrak{r}-1}$, но не делится на $2^{\mathfrak{r}}$ (поскольку $n-1$ нечетно).

Докажем теорему индукцией по $\mathfrak{r} = \operatorname{parity}(n)$. Для $\mathfrak{r} = 0$ утверждение доказано в пункте **III**. Пусть оно справедливо для любого полинома $\operatorname{parity}(\deg) \leq \mathfrak{r}-1$. Докажем для полинома $\operatorname{parity}(\deg f)=\mathfrak{r}$.

Действуем сначала по алгоритму пункта **I**. Получаем $\mathcal{X}(x) \in \mathbb{C}[x]$ и $\operatorname{parity}(\deg \mathcal{X})=\mathfrak{r}-1$. По индуктивному предположению, существует $x=x_0 \in \mathbb{C}: \mathcal{X}(x_0)=0$. Далее повторяем аргументы пункта **II** доказательства с тем отличием, что теперь все полиномы будут иметь комплексные коэффициенты. По лемме 2 существует нетривиальный $\text{НОД}(F_1(x_0,Y),F_2(x_0,Y))$, и, следовательно, возможно разложение (8.7), где теперь $f_j \in \mathbb{C}[z]$.

Хотя бы одно из чисел n_1 или n_2 имеет четность $\leq \mathfrak{r}$ (оба не могут быть четности большей $\mathfrak{r}$, так как тогда $\operatorname{parity}(n) = \operatorname{parity}(n_1+n_2) > \mathfrak{r}$). Если хоть одно из чисел n_j имеет четность меньшую $\mathfrak{r}$, то можно применить индуктивное предположение. Соответствующий полином имеет тогда корень $z=z_0 \in \mathbb{C}$, который и будет корнем $f(z)$.

Если хоть одно из чисел n_1, n_2 имеет четность равную $\mathfrak{r}$, то к этому полиному $f_j(z)$ применяем снова рассуждения пункта **I**, выписываем для него разложение (8.8), где $f_{jk}(z) \in \mathbb{C}[z]$, и т.д. Получаем последовательность полиномов (8.9) с убывающими степенями. Через конечное число шагов эти степени станут меньшими $2^{\mathfrak{r}}$. Следовательно и соответствующие четности степеней станут меньшими $\mathfrak{r}$ (если только нам не встретится понижение четности степени на более раннем шаге). По индуктивному предположению у такого полинома существует корень $z=z_0 \in \mathbb{C}$, который и будет корнем $f(z)$. ■

Нахождение комплексного корня

Идею, использованную для доказательства существования корня полинома, можно развить в алгоритм его нахождения. Ограничимся случаем $f(z) \in \mathbb{R}[z]$. Пусть теперь $x=\operatorname{Re} z$, $y=\operatorname{Im} z$. Тогда элиминанта $\mathcal{X}(x)$, определенная формулой (8.5), имеет вещественные коэффициенты. Если мы найдем (точно или приближенно) хотя бы один вещественный корень $x=x_0$ этого полинома, то мы сможем выделить нетривиальный множитель полинома $f(z)$:

$$f_1(z) \stackrel{\text{def}}{=} \mathfrak{D}((z-x_0)^2) \in \mathbb{R}[z].$$

Выражение для $\mathfrak{D}(Y)$ дается формулой (8.6) и может быть установлено с помощью субрезультантов. Действительно, предположим, например, что $\mathcal{X}(x)$ не имеет кратных корней $(\mathcal{D}(\mathcal{X}(x)) \neq 0)$. Тогда $\deg \mathfrak{D} = 1$ и

$$\mathfrak{D}(Y) = \mathcal{R}^{(1)}(x_0)Y + \mathcal{R}_1^{(1)}(x_0),$$

где $\mathcal{R}^{(1)}$ — первый субрезультант результанта (8.4), а $\mathcal{R}_1^{(1)}$ получается из последнего заменой одного столбца. Следовательно,

$$f_1(z) = \mathfrak{D}((z-x_0)^2) = \mathcal{R}^{(1)}(x_0)(z-x_0)^2 + \mathcal{R}_1^{(1)}(x_0) \ \in \mathbb{R}[z]$$

— квадратичный множитель полинома $f(z)$. Корни полинома $f(z)$:

$$\lambda_{1,2} = \begin{cases} x_0 \pm \sqrt{-\mathcal{R}_1^{(1)}(x_0)/\mathcal{R}^{(1)}(x_0)} & \text{если } \ \mathcal{R}_1^{(1)}(x_0)/\mathcal{R}^{(1)}(x_0) < 0, \\ x_0 \pm i\sqrt{\mathcal{R}_1^{(1)}(x_0)/\mathcal{R}^{(1)}(x_0)} & \text{если } \ \mathcal{R}_1^{(1)}(x_0)/\mathcal{R}^{(1)}(x_0) > 0. \end{cases} \tag{8.11}$$

В первом случае оба корня вещественны, во втором — мнимы (комплексно–сопряженны).

Пример 8.3. Найти корни полинома

$$f(z) = z^6 + 8\,z^5 + 24\,z^4 + 51\,z^3 + 146\,z^2 + 298\,z + 312.$$

Решение.

$$\begin{aligned}\mathcal{X}(x) = &-32768x^{15} - 655360x^{14} - 6029312x^{13} - 33972224x^{12} - \\ &-131440640x^{11} - 369635328x^{10} - 773799936x^9 - 1195447296x^8 - \\ &-1272087808x^7 - 672626816x^6 + 450670208x^5 + 1362662720x^4 + \\ &+1430858024x^3 + 853936272x^2 + 300635412x + 56964852.\end{aligned}$$

Ищем вещественные корни $\mathcal{X}(x)$: $x_1 = 1$, $x_2 = -{}^7/_2$, $x_3 = -{}^3/_2$. Теперь осталось подставить эти значения в формулы (8.11), в которых

$$\begin{aligned}\mathcal{R}^{(1)} &= 896x^6 + 7168x^5 + 23552x^4 + 39944x^3 + 34704x^2 + 12620x + 231, \\ \mathcal{R}_1^{(1)} &= 384x^8 + 4096x^7 + 18432x^6 + 48504x^5 + 92560x^4 + 144452x^3 + \\ &\quad +163285x^2 + 101812x + 22050\end{aligned}$$

x_0	$\mathcal{R}_1^{(1)}/\mathcal{R}^{(1)}$	λ
1	5	$1 \pm \boldsymbol{i}\sqrt{5}$
$-{}^7\!/_2$	${}^3\!/_4$	$-{}^7\!/_2 \pm \boldsymbol{i}\,{}^{\sqrt{3}}\!/_2$
$-{}^3\!/_2$	${}^7\!/_4$	$-{}^3\!/_2 \pm \boldsymbol{i}\,{}^{\sqrt{7}}\!/_2$

Ответ. $\lambda_{1,2}{=}1 \pm \boldsymbol{i}\sqrt{5}$, $\lambda_{3,4}{=}-{}^7\!/_2 \pm \boldsymbol{i}\ {}^{\sqrt{3}}\!/_2$, $\lambda_{5,6}{=}-{}^3\!/_2 \pm \boldsymbol{i}\ {}^{\sqrt{7}}\!/_2$.

Дальнейшие свойства системы (8.3)

В этом пункте[10] предполагаем, что $f(z) \in \mathbb{R}[z]$. Наша задача теперь установить связь между корнями $f(z)$ и элиминанты (8.4), а также элиминанты

$$\mathcal{Y}(Y) \stackrel{\rm def}{=} \mathcal{R}_x(F_1, F_2)/a_0. \qquad (8.12)$$

В разложении обеих элиминант по степеням переменных

$$\mathcal{X}(x) = \mathcal{A}_0 x^N + \mathcal{A}_1 x^{N-1} + \dots + \mathcal{A}_N\,,\ \mathcal{Y}(Y) = \mathcal{B}_0 Y^N + \mathcal{B}_1 Y^{N-1} + \dots + \mathcal{B}_N$$

коэффициенты можно выразить через $a_0, \dots, a_n$, например

$$\mathcal{A}_1 = \frac{(n-1)a_1}{2a_0}\mathcal{A}_0,\ \mathcal{A}_2 = \frac{n-2}{4a_0^2}\left(\frac{n-1}{2}a_1^2 + a_0 a_2\right)\mathcal{A}_0, \dots, \qquad (8.13)$$

$$\mathcal{B}_0 = 2^{2N} a_0^{2n-2},\ \mathcal{B}_1 = \frac{1}{4a_0^2}\left(2n\,a_0 a_2 - (n-1)a_1^2\right)\mathcal{B}_0, \dots,$$

$$\mathcal{A}_N = \mathcal{X}(0) = \mathcal{R}(F_1(0,Y), F_2(0,Y)), \qquad (8.14)$$

$$\mathcal{B}_N = \mathcal{Y}(0) = \frac{\mathcal{R}(F_1(x,0), F_2(x,0))}{a_0} = \frac{\mathcal{R}(f(x), f'(x))}{a_0} = \mathcal{D}(f(x)). \qquad (8.15)$$

Теорема 8.2. *[Лагранж] Множество решений системы* (8.3) *имеет вид*

$$\left\{\left(x_{jk} = \frac{\lambda_j + \lambda_k}{2}, Y_{jk} = \left[\frac{\lambda_j - \lambda_k}{2}\right]^2\right)\right\}_{1 \le j < k \le n}. \qquad (8.16)$$

[10] Материал настоящего пункта относится к главе 8.

Доказательство. Ограничимся случаем когда все корни полинома различны: $\mathcal{D}(f)\neq 0$. Введем в рассмотрение

$$y_{jk}\stackrel{\text{def}}{=}\frac{\lambda_j-\lambda_k}{2\boldsymbol{i}};\ y_{kj}\stackrel{\text{def}}{=}-y_{jk}\ \Longrightarrow\ Y_{jk}=-y_{jk}^2=-y_{kj}^2.$$

Подставляя в равенство (8.1) пары (x_{jk},y_{jk}) и (x_{jk},y_{kj}), получим соотношения

$$F_1(x_{jk},Y_{jk})+\boldsymbol{i}\,y_{jk}F_2(x_{jk},Y_{jk})=f\left(\frac{\lambda_j+\lambda_k}{2}+\boldsymbol{i}\,\frac{\lambda_j-\lambda_k}{2\boldsymbol{i}}\right)=f(\lambda_j)=0;$$
$$F_1(x_{jk},Y_{jk})+\boldsymbol{i}\,y_{kj}F_2(x_{jk},Y_{jk})=f\left(\frac{\lambda_j+\lambda_k}{2}+\boldsymbol{i}\,\frac{\lambda_k-\lambda_j}{2\boldsymbol{i}}\right)=f(\lambda_k)=0.$$

которые будем рассматривать как линейную систему относительно $F_1(x_{jk},Y_{jk})$ и $y_{jk}F_2(x_{jk},Y_{jk})$. Поскольку определитель этой системы отличен от нуля, должно иметь место:

$$F_1(x_{jk},Y_{jk})=0,\quad y_{jk}F_2(x_{jk},Y_{jk})=0.$$

Но $y_{jk}\neq 0$ по предположению различности корней. Поэтому любая пара из множества (8.16) — решение системы (8.3).

Обратно, рассмотрим произвольное решение $x=\alpha,\ Y=-\beta^2$ системы (8.3):

$$\begin{cases}F_1(\alpha,-\beta^2)=0,\\ F_2(\alpha,-\beta^2)=0\end{cases}\Longleftrightarrow F_1(\alpha,-\beta^2)\pm\boldsymbol{i}\,\beta F_2(\alpha,-\beta^2)=0\Longleftrightarrow f(\alpha\pm\boldsymbol{i}\,\beta)=0.$$

Таким образом, числа $\alpha\pm\boldsymbol{i}\,\beta$ должны быть корнями полинома $f(z)$. Если $\beta\neq 0$, то существуют индексы j,k такие, что

$$\alpha+\boldsymbol{i}\,\beta=\lambda_j,\quad \alpha-\boldsymbol{i}\,\beta=\lambda_k\ \Longrightarrow\ \alpha=\frac{\lambda_j+\lambda_k}{2},\ \beta=\frac{\lambda_j-\lambda_k}{2\boldsymbol{i}}.$$

Если $\beta=0$, то $F_1(\alpha,0)=F_2(\alpha,0)=0$. Из формул (8.2) тогда следует, что $f(\alpha)=f'(\alpha)=0$, что противоречит предположению о простоте корней $f(z)$. ■

Следствие 8.1. *Корнями $\mathcal{X}(x)$ являются числа $(\lambda_j+\lambda_k)/2$, среди которых, очевидно, будут находиться и* $\operatorname{Re}\lambda_j$ *при* $\lambda_j\notin\mathbb{R}$. *В частности, если у $f(z)\in\mathbb{Q}[z]$ имеется делитель вида $z^2+pz+q\in\in\mathbb{Q}[z]$, $p^2-4q<0$ то у $\mathcal{X}(x)$ имеется корень $-p/2\in\mathbb{Q}$.*

Теорема 8.3. *Все корни полинома $f(z) \in \mathbb{R}[z]$ будут вещественными и различными тогда и только тогда, когда у полинома $\mathcal{Y}(Y)$ все коэффициенты ненулевые, а их знаки чередуются.*

Доказательство необходимости. Если все корни λ_j вещественны и различны, то и все числа Y_{jk}, определяемые формулой (8.16) тоже вещественны и положительны. Тогда по следствию 2 к теореме Декарта[11] необходимо

$$\operatorname{nrr}\{\mathcal{Y}(Y)=0 \ | \ Y>0\} = \mathcal{V}(\mathcal{B}_0, \mathcal{B}_1, \dots, \mathcal{B}_N) = N.$$

Последнее возможно тогда и только тогда, когда знаки коэффициентов $\mathcal{Y}(Y)$ чередуются. ■

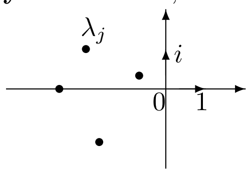
Доказательство достаточности. Пусть теперь выполнено условие теоремы, но у полинома $f(z)$ существует пара комплексно-сопряженных корней $\lambda_{1,2} = \alpha \pm \boldsymbol{i}\ \beta$, $\beta \ne 0$. Этой паре соответствует по формуле (8.16) корень $Y_{1,2} = -\beta^2 < 0$ полинома $\mathcal{Y}(Y)$. Но тогда, из следствия 1 к теореме Декарта следует, что в последовательности коэффициентов $\mathcal{Y}(Y)$ должно иметься по крайней мере одно знакопостоянство. Последнее противоречит предположению.

Отсутствие у полинома $f(z)$ кратных корней следует из того, что $\mathcal{Y}(0) = \mathcal{B}_N \ne 0$ (см. формулу (8.15)). Однако утверждение теоремы можно обобщить и на случай их наличия. Если $\mathcal{Y}(Y) = \mathcal{B}_0 Y^N + \mathcal{B}_1 Y^{N-1} + \dots + \mathcal{B}_{N-k} Y^k$ и $\mathcal{B}_{N-k} \ne 0$ то для вещественности корней $f(z)$ Н. и Д. чтобы все числа $\mathcal{B}_0, \mathcal{B}_1, \dots, \mathcal{B}_{N-k}$ были ненулевыми, а их знаки чередовались. ■

Теперь проанализируем свойства корней $f(z)$ в зависимости от коэффициентов $\mathcal{X}(x)$.

Определение 8.3. Полином $f(z)$ называется **устойчивым**, если все его корни лежат в левой полуплоскости комплексной плоскости:

$$\operatorname{Re}\lambda_1 < 0, \dots, \operatorname{Re}\lambda_n < 0.$$



[11] Глава 3, §7.4.

Теорема 8.4. *[Раус] Для устойчивости полинома $f(z) \in \mathbb{R}[z]$* **Н.** *и* **Д.** *чтобы*

а) *все коэффициенты $f(z)$ были одного знака;*

б) *все коэффициенты $\mathcal{X}(x)$ были одного знака.*

Доказательство. Для определенности будем считать числа a_0 и $\mathcal{A}_0$ положительными. Покажем, что для устойчивости $f(z)$ Н. и Д. чтобы все коэффициенты $f(z)$ и $\mathcal{X}(x)$ были положительными. Пусть $\lambda_1, \lambda_2, \dots, \lambda_k$ — вещественные, а

$$\lambda_{k+1} = \beta_1 + \boldsymbol{i}\ \gamma_1,\ \lambda_{k+2} = \beta_2 + \boldsymbol{i}\ \gamma_2,\ \dots,\ \lambda_n = \beta_{n-k} + \boldsymbol{i}\ \gamma_{n-k}$$

— комплексно–сопряженные корни $f(z)$. Разложение полинома $f(z) \in \mathbb{R}[z]$ на вещественные множители будет иметь вид:

$$\begin{aligned} f(z) &= a_0(z-\lambda_1)\times\dots\times(z-\lambda_k)\times \\ &\times \left(z^2-2\beta_1 z+(\beta_1^2+\gamma_1^2)\right)\times\ldots\times\left(z^2-2\beta_{n-k}z+(\beta_{n-k}^2+\gamma_{n-k}^2)\right). \end{aligned}$$

Н. Предположим, что $\mathrm{Re}\, e\, \lambda_j < 0$, т.е.

$$\lambda_1 < 0, \dots, \lambda_k < 0, \beta_1 < 0, \dots, \beta_{n-k} < 0.$$

Это означает, что все сомножители в произведении — как линейные, так и квадратичные — имеют только положительные коэффициенты. Следовательно, и все коэффициенты $f(z)$ тоже положительны.

По следствию 1 к теореме 8.2 корнями $\mathcal{X}(x)$ являются числа $(\lambda_j + \lambda_k)/2$. Имеем:

$$\mathrm{Re}\, e(\lambda_j + \lambda_k)/2 = (\mathrm{Re}\, e\, \lambda_j + \mathrm{Re}\, e\, \lambda_k)/2 < 0.$$

Таким образом корни полинома $\mathcal{X}(x)$ удовлетворяют тому же условию, что корни $f(z)$. Той же последовательностью рассуждений приходим к выводу, что и все коэффициенты $\mathcal{X}(x)$ должны быть положительными.

Д. Пусть $a_0 > 0, \dots, a_n > 0$. Правило знаков Декарта утверждает, что полином $f(z)$ не имеет положительных корней; он не имеет и нулевых, так как $a_n \neq 0$. Следовательно, все вещественные корни $f(z)$ (если таковые имеются) должны быть отрицательными. Аналогично доказывается, что и все вещественные корни $\mathcal{X}(x)$ должны быть отрицательными. По следствию 1 к теореме 8.2 среди вещественных корней $\mathcal{X}(x)$ должны находиться и $\mathrm{Re}\, e\, \lambda_{k+1}, \dots, \mathrm{Re}\, e\, \lambda_n$. Поэтому все корни $f(z)$ удовлетворяют условию $\mathrm{Re}\, e\, z < 0$. ∎

Теорема 8.4 накладывает $n+N = n(n+1)/2$ условий на коэффициенты полинома $f(z)$ необходимых для установления его устойчивости. Не все эти условия независимы: так, например, из формул (8.13) очевидно, что первые коэффициенты в разложении $\mathcal{X}(x)$ будут одинакового знака если все $a_0, \ldots, a_n$ одного знака. Выделим важнейшие условия. Поскольку корни полинома являются непрерывными функциями его коэффициентов, то при непрерывном изменении последних свойство устойчивости может быть нарушено только когда какой-то из корней выйдет из левой полуплоскости на ее границу — т.е. на мнимую ось. Произойти это может двумя путями:

1) один из корней λ_j станет равным 0: тогда $a_n = 0$;

2) у пары комплексно-сопряженных корней $\lambda_j, \lambda_k = \overline{\lambda_j}$ вещественная часть станет нулевой: $\lambda_{j,k} = \pm i\,\beta$; тогда система (8.3) будет иметь решение $x = 0, Y = -\beta^2$, и, следовательно, $\mathcal{X}(0) = 0$, т.е. $\mathcal{A}_N = 0$, где $\mathcal{A}_N$ определяется по формуле (8.14).

Итак, два условия из теоремы 8.4 являются существенными, в том смысле, что потеря полиномом $f(z)$ свойства устойчивости может произойти исключительно за счет обращения в нуль одного из чисел a_n или $\mathcal{A}_N$. Для последнего числа формула (8.14) дает представление в виде результанта.

Пример 8.4. Найти все значения параметра A, при которых полином

$$z^5 + 3\,z^4 + A\,z^3 + 5\,z^2 + z + 3 - 2\,A$$

будет устойчив.

Решение. Условие **а)** теоремы 8.4 будет выполнено при $A \in]0, {}^3/_2[$. Построим теперь полином (8.4):

$$\begin{aligned}
\mathcal{X}(x) &= x^{10} + 6x^9 + \left({}^3/_4 A + {}^{27}/_2\right) x^8 + \left({}^{27}/_8 A + {}^{113}/_8\right) x^7 + \\
&+ \left({}^3/_{16} A^2 + {}^{81}/_{16} A + {}^{69}/_8\right) x^6 + \left({}^9/_{16} A^2 + {}^{113}/_{32} A + {}^{177}/_{32}\right) x^5 + \\
&+ \left({}^1/_{64} A^3 + {}^{27}/_{64} A^2 + {}^{55}/_{16} A + {}^{29}/_{64}\right) x^4 + \\
&+ \left({}^3/_{128} A^3 + {}^{13}/_{128} A^2 + {}^{57}/_{16} A - {}^{113}/_{32}\right) x^3 +
\end{aligned}$$

$$
\begin{aligned}
+ & \left({}^{85}\!/_{256}A^2 + {}^{49}\!/_{256}A - {}^{43}\!/_{256}\right) x^2 + \\
+ & \left({}^{1}\!/_{256}A^3 + {}^{9}\!/_{64}A^2 - {}^{161}\!/_{512}A + {}^{55}\!/_{512}\right) x + \\
+ & \ {}^{3}\!/_{512}A^3 - {}^{23}\!/_{1024}A^2 + {}^{15}\!/_{512}A - {}^{25}\!/_{1024}
\end{aligned}
$$

Условие **б)** теоремы 8.4 порождает несколько неравенств относительно параметра A. Решать эту систему в общем случае было бы довольно затруднительно. Вспомним, однако, что для устойчивости полинома наиболее существенным из этих неравенств является условие на свободный член:

$$\frac{3}{512}A^3-\frac{23}{1024}A^2+\frac{15}{512}A-\frac{25}{1024}\equiv\frac{1}{1024}\,(2A-5)\left(3A^2-4A+5\right)>0.$$

Решением его является интервал $]\,{}^{5}\!/_{2},+\infty[$, который не пересекается с полученным из условия **а)**.

Ответ. $A\in\varnothing$.

Можно, оказывается, уменьшить число условий теоремы 8.4 существенных для установления устойчивости за счет рассмотрения субрезультантов результанта $\mathcal{R}(F_1(0,Y),F_2(0,Y))$.

Теорема 8.5. *[Льенар, Шипар]***.** *Для устойчивости полинома* $f(z)=a_0z^n+\dots+a_n\in\mathbb{R}[z],\quad a_0>0$ **Н.** *и* **Д.** *чтобы*

а) *все коэффициенты* $f(z)$ *были положительными;*

б) *все субрезультанты результанта*

$$
\begin{aligned}
\mathcal{A}_N &= \mathcal{R}(F_1(0,Y),F_2(0,Y)) = \\
&= \mathcal{R}\left(a_n+a_{n-2}Y+a_{n-4}Y^2+\dots,\ a_{n-1}+a_{n-3}Y+a_{n-5}Y^2+\dots\right)
\end{aligned}
$$

были положительными.

Пример 8.5. Условие **б)** для $n=7$:

$$
\mathcal{R}^{(0)}=\left|\begin{array}{cccccc}
a_1 & a_3 & a_5 & a_7 & & \\
 & a_1 & a_3 & a_5 & a_7 & \\
 & & a_1 & a_3 & a_5 & a_7 \\
 & & a_0 & a_2 & a_4 & a_6 \\
 & a_0 & a_2 & a_4 & a_6 & \\
a_0 & a_2 & a_4 & a_6 & &
\end{array}\right|>0,\ \
\mathcal{R}^{(1)}=\left|\begin{array}{cccc}
a_1 & a_3 & a_5 & a_7 \\
 & a_1 & a_3 & a_5 \\
 & a_0 & a_2 & a_4 \\
a_0 & a_2 & a_4 & a_6
\end{array}\right|>0,\ \
\mathcal{R}^{(2)}=\left|\begin{array}{cc}
a_1 & a_3 \\
a_0 & a_2
\end{array}\right|>0;
$$

и для $n = 8$:

$$\mathcal{R}^{(0)} = \begin{vmatrix} a_0\,a_2\,a_4\,a_6\,a_8 & & & \\ & a_0\,a_2\,a_4\,a_6\,a_8 & & \\ & & a_0\,a_2\,a_4\,a_6\,a_8 & \\ & & & a_1\,a_3\,a_5\,a_7 \\ & & a_1\,a_3\,a_5\,a_7 & \\ & a_1\,a_3\,a_5\,a_7 & & \\ a_1\,a_3\,a_5\,a_7 & & & \end{vmatrix} > 0, \ \mathcal{R}^{(1)} = \begin{vmatrix} a_0\,a_2\,a_4\,a_6\,a_8 & & \\ & a_0\,a_2\,a_4\,a_6 & \\ & & a_1\,a_3\,a_5 \\ & a_1\,a_3\,a_5\,a_7 & \\ a_1\,a_3\,a_5\,a_7 & & \end{vmatrix} > 0,$$

$$\mathcal{R}^{(2)} = \begin{vmatrix} a_0 & a_2 & a_4 \\ & a_1 & a_3 \\ a_1 & a_3 & a_5 \end{vmatrix} > 0,$$

и $\mathcal{R}^{(3)} = a_1 > 0$. Последнее условие будет выполнено автоматически при выполнении условия **а)**.

Теорема 8.5 требует проверки $n + \lfloor (n-1)/2 \rfloor$ условий на коэффициенты полинома для проверки его устойчивости, что существенно меньше числа условий из теоремы 8.4.

Упражнение 74. *Выписать явно условия устойчивости полиномов* **а)** *третьей и* **б)** *четвертой степеней.*

Перейти к оглавлению на странице: 460

ГЛАВА 7. КВАДРАТИЧНЫЕ ФОРМЫ

§1. Приведение квадратичной формы к каноническому виду по методу Лагранжа

Обозначения. $X \stackrel{\text{def}}{=} (x_1, \dots, x_n)^\top$ — столбец переменных, $\mathbb{A}$ — множество $\mathbb{Q}$ (рациональных), или $\mathbb{R}$ (вещественных), или $\mathbb{C}$ (комплексных) чисел.

Определение

Определение 1.1. Квадратичной формой над $\mathbb{A}$ называют однородный полином второй степени с коэффициентами из $\mathbb{A}$:

$$f(X) = \sum_{1 \le j \le k \le n} f_{jk} x_j x_k = \begin{array}{llll} f_{11}x_1^2 & +f_{12}x_1x_2 & +\dots & +f_{1n}x_1x_n+ \\ & +f_{22}x_2^2 & +\dots & +f_{2n}x_2x_n+ \\ & +\dots & +f_{jk}x_jx_k & +\dots+ \\ & & & +f_{nn}x_n^2 \end{array}$$

Пример 1.1. Функции

$$x_1^2 - x_1x_2 + x_3^2\,, \quad \sqrt{3}\,x_2^2 - \pi\,x_3^2\,, \quad -x_1x_2\,, \quad \boldsymbol{i}\ x_1^2$$

являются кв. формами. Функции

$$x_1^2 - 3\,x_1 + 1\,, \quad 5\,x_1^2x_2^2\,, \quad \frac{x_1x_3^2}{x_2}\,, \quad \sqrt{x_1x_2x_3x_4}$$

не являются кв. формами.

Если определить верхнетреугольную матрицу $\mathbf{F}$ равенством:

$$\mathbf{F} \stackrel{\text{def}}{=} \left(\begin{array}{cccc} f_{11} & f_{12} & \dots & f_{1n} \\ & f_{22} & \dots & f_{2n} \\ \mathbb{O} & & \ddots & \vdots \\ & & & f_{nn} \end{array} \right),$$

то кв. форму можно записать в виде $f(X) = X^\top \mathbf{F} X$. Более удобно, однако, другое представление. Рассмотрим матрицу

$$\mathbf{A} \overset{\text{def}}{=} \frac{\mathbf{F} + \mathbf{F}^\top}{2} = \begin{pmatrix} f_{11} & {}^1\!/\!{}_2 f_{12} & \dots & {}^1\!/\!{}_2 f_{1n} \\ {}^1\!/\!{}_2 f_{12} & f_{22} & \dots & {}^1\!/\!{}_2 f_{2n} \\ \dots & & & \dots \\ {}^1\!/\!{}_2 f_{1n} & {}^1\!/\!{}_2 f_{2n} & \dots & f_{nn} \end{pmatrix},$$

которая, очевидно, симметрична: $\mathbf{A}^\top = \mathbf{A}$. Тогда

$$f(X) = \sum_{1 \le j,k \le n} a_{jk} x_j x_k = X^\top \mathbf{A} X. \tag{1.1}$$

Определение 1.2. Представление (1.1) называют **правильной записью** кв. формы, матрицу $\mathbf{A}$ — **матрицей** кв. формы f, а $\det \mathbf{A}$ — ее **дискриминантом**.

Канонический вид

Произведем невырожденную замену (преобразование) переменных в кв. форме:

$$\left\{ \begin{array}{lcl} x_1 & = & c_{11} y_1 + c_{12} y_2 + \dots + c_{1n} y_n, \\ x_2 & = & c_{21} y_1 + c_{22} y_2 + \dots + c_{2n} y_n, \\ \dots & & \dots \\ x_n & = & c_{n1} y_1 + c_{n2} y_2 + \dots + c_{nn} y_n, \end{array} \right. \quad \Longleftrightarrow \quad \begin{array}{c} X = CY \\ \text{при} \\ \det C \ne 0. \end{array} \tag{1.2}$$

Кв. форма перейдет в новую относительно новых переменных:

$$f(X) = X^\top \mathbf{A} X = (CY)^\top \mathbf{A} C Y = Y^\top \underbrace{C^\top \mathbf{A} C}_{\mathbf{B}} Y = Y^\top \mathbf{B} Y \overset{\text{def}}{=} \widetilde{f}(Y).$$

Определение 1.3. Говорят, что кв.форма $f(X) = X^\top \mathbf{A} X$ имеет **канонический вид**, если ее матрица диагональна, или, что то же:

$$f(X) = a_{11} x_1^2 + a_{22} x_2^2 + \dots + a_{nn} x_n^2$$

("представлена в виде суммы квадратов").

Теорема 1.1. *Для любой кв. формы над* $\mathbb{A}$ *существует линейная замена переменных* (1.2) *такая, что преобразованная кв. форма* $\widetilde{f}(Y)$ *имеет канонический вид.*

Доказательство (**метод Лагранжа**). Пусть $a_{11} \neq 0$. Выделим в $f(X)$ все слагаемые, содержащие x_1:

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n + \sum_{2\le j,k\le n} a_{jk}x_jx_k =$$

$$= a_{11}\left(x_1^2 + 2\frac{a_{12}}{a_{11}}x_1x_2 + \cdots + 2\frac{a_{1n}}{a_{11}}x_1x_n\right) + \cdots =$$

$$= a_{11}\left[\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \ldots + \frac{a_{1n}}{a_{11}}x_n\right)^2 - \left(\frac{a_{12}}{a_{11}}x_2 + \ldots + \frac{a_{1n}}{a_{11}}x_n\right)^2\right] + \ldots =$$

$$= a_{11}\underbrace{\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \ldots + \frac{a_{1n}}{a_{11}}x_n\right)}_{y_1}^2 - \underbrace{a_{11}\left(\frac{a_{12}}{a_{11}}x_2 + \ldots + \frac{a_{1n}}{a_{11}}x_n\right)^2 + \ldots}_{\text{нет } x_1} =$$

$$= a_{11}y_1^2 + f_2(x_2, \ldots, x_n). \tag{1.3}$$

Итак, подстановка $y_1 \overset{\text{def}}{=} x_1 + {}^{a_{12}}/_{a_{11}}\, x_2 + \cdots + {}^{a_{1n}}/_{a_{11}}\, x_n$ позволяет свести задачу к форме $f_2(x_2, \ldots, x_n)$ от $(n-1)$-й переменной. Предположив, что эта форма приводится к каноническому виду, получаем доказательство по индукции.

Если $a_{11} = 0$, но $\exists k: \ a_{kk} \neq 0$. Замена переменных $x_k = \mathfrak{x}_1$, $x_1 = \mathfrak{x}_k$ сводит этот случай к предыдущему.

Если $a_{11} = \cdots = a_{nn} = 0$, но $\exists a_{jk} \neq 0$. Предварительная подстановка $x_j = \mathfrak{x}_j + \mathfrak{x}_k$, $x_k = \mathfrak{x}_k$ сводит этот случай к предыдущему. ■

Пример 1.2. Методом Лагранжа привести кв. форму

$$f = 4x_1^2 + 2x_2^2 + x_3^2 + x_4^2 - 4x_1x_2 - 4x_1x_3 + 4x_1x_4 + 4x_2x_3 - 4x_3x_4$$

к каноническому виду.

Решение.

$$
\begin{aligned}
f &= 4\left(x_1^2 - x_1x_2 - x_1x_3 + x_1x_4\right) + 2x_2^2 + x_3^2 + x_4^2 + 4x_2x_3 - 4x_3x_4 = \\
&= 4\Bigl[\bigl(\underbrace{x_1 - {}^1\!/_2\,x_2 - {}^1\!/_2\,x_3 + {}^1\!/_2\,x_4}_{y_1}\bigr)^2 - \left(-{}^1\!/_2x_2 - {}^1\!/_2x_3 + {}^1\!/_2x_4\right)^2\Bigr] + \\
&+ 2x_2^2 + x_3^2 + x_4^2 + 4x_2x_3 - 4x_3x_4 = 4y_1^2 + x_2^2 + 2x_2x_3 + 2x_2x_4 - 2x_3x_4 = \\
&= 4y_1^2 + \Bigl[\bigl(\underbrace{x_2 + x_3 + x_4}_{y_2}\bigr)^2 - (x_3 + x_4)^2\Bigr] - 2x_3x_4 = \\
&= 4y_1^2 + y_2^2 - x_3^2 - 4x_3x_4 - x_4^2 = 4y_1^2 + y_2^2 - \Bigl[\bigl(\underbrace{x_3 + 2\,x_4}_{y_3}\bigr)^2 - 4x_4^2\Bigr] - x_4^2 = \\
&= 4y_1^2 + y_2^2 - y_3^2 + 3x_4^2.
\end{aligned}
$$

Ответ. Подстановка

$$
\left\{
\begin{array}{rcrrrr}
y_1 & = & x_1 & -{}^1\!/_2\,x_2 & -{}^1\!/_2\,x_3 & +{}^1\!/_2\,x_4, \\
y_2 & = & & x_2 & +x_3 & +x_4, \\
y_3 & = & & & x_3 & +2x_4, \\
y_4 & = & & & & x_4
\end{array}
\right.
$$

приводит кв. форму f к каноническому виду $4\,y_1^2 + y_2^2 - y_3^2 + 3\,y_4^2$.

§2. Формула Якоби

Метод Лагранжа и метод Гаусса

Рассмотрим матрицу кв.формы последнего примера, и, временно выходя из круга поставленных в предыдущем параграфе проблем, побробуем применить к ней метод Гаусса приведения к треугольному виду[1]:

$$
\begin{pmatrix} 4 & -2 & -2 & 2 \\ -2 & 2 & 2 & 0 \\ -2 & 2 & 1 & -2 \\ 2 & 0 & -2 & 1 \end{pmatrix}
\to
\begin{pmatrix} 4 & -2 & -2 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix}
\to
\begin{pmatrix} 4 & -2 & -2 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & -2 & -1 \end{pmatrix}
\to
\begin{pmatrix} 4 & -2 & -2 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & 3 \end{pmatrix}.
$$

[1] Глава 4, §1.

Обратим внимание на два обстоятельства: диагональные элементы матрицы совпадают с коэффициентами канонического вида кв.формы, а коэффициенты замены переменных, приводящей к этому каноническому виду, совпадают с элементами строк последней матрицы, если их разделить на диагональный элемент. Возникает подозрение, что метод Лагранжа является “замаскированной” версией метода Гаусса.

Для того, чтобы выяснить аналитический смысл преобразований по методу Лагранжа найдем правило формирования коэффициентов формы $f_2(x_2, \dots, x_n)$ в (1.3).

$$
\begin{aligned}
f_2 &= \sum_{2\le j,k\le n} a_{jk}x_jx_k - a_{11}\left(\frac{a_{12}}{a_{11}}x_2 + \dots + \frac{a_{1n}}{a_{11}}x_n\right)^2 = \\
&= \sum_{2\le j,k\le n} a_{jk}x_jx_k - a_{11}\sum_{2\le j,k\le n}\frac{a_{1j}a_{1k}}{a_{11}^2}x_jx_k = \\
&= \sum_{2\le j,k\le n}\left(a_{jk} - \frac{a_{1j}}{a_{11}}a_{1k}\right)x_jx_k. \qquad (2.1)
\end{aligned}
$$

Теорема 2.1. *Метод Лагранжа приведения кв.формы $X^\top \mathbf{A} X$ к каноническому виду эквивалентен методу Гаусса приведения матрицы $\mathbf{A}$ к треугольному виду.*

Доказательство. Действительно, первый шаг прямого хода метода исключения переменных Гаусса преобразует матрицу $\mathbf{A}$ следующим образом:

$$
\begin{pmatrix}
a_{11}\, a_{12}\, a_{13} \dots a_{1n} \\
a_{12}\, a_{22}\, a_{23} \dots a_{2n} \\
\dots \quad \dots \\
a_{1n}\, a_{2n}\, a_{3n} \dots a_{nn}
\end{pmatrix}
\to
\begin{pmatrix}
a_{11}\, a_{12} \dots a_{1n} \\
0 \quad a_{22}^{[1]} \dots a_{2n}^{[1]} \\
\dots \quad \dots \\
0 \quad a_{n2}^{[1]} \dots a_{nn}^{[1]}
\end{pmatrix};
$$

здесь

$$
a_{jk}^{[1]} \stackrel{\text{def}}{=} a_{jk} - \frac{a_{j1}a_{1k}}{a_{11}},
$$

и предполагается, что $a_{11} \neq 0$. Видим, что формула формирования элементов матрицы

$$\begin{pmatrix} a_{22}^{[1]} & \dots & a_{2n}^{[1]} \\ \dots & & \dots \\ a_{n2}^{[1]} & \dots & a_{nn}^{[1]} \end{pmatrix}_{(n-1)\times(n-1)}$$

точно такая же, как и матрицы кв.формы (2.1). Более того, поскольку матрица $\mathbf{A}$ *симметрична* ($a_{jk} = a_{kj}$), то и только что полученная матрица оказывается симметричной. Если $a_{22}^{[1]} \neq 0$, то к этой новой матрице можно снова применить ту же процедуру, и т.д., и в конце концов придем к матрице первого порядка. Собирая все промежуточные результаты в одну матрицу, получим ее в треугольном виде

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,n-1} & a_{1n} \\ 0 & a_{22}^{[1]} & \dots & a_{2,n-1}^{[1]} & a_{2n}^{[1]} \\ & & \ddots & & \dots \\ 0 & 0 & & a_{n-1,n-1}^{[n-2]} & a_{n-1,n}^{[n-2]} \\ 0 & 0 & \dots & 0 & a_{nn}^{[n-1]} \end{pmatrix} \tag{2.2}$$

при условии, что ни одно из чисел на диагонали не обратилось в нуль:

$$a_{11} \neq 0,\ a_{22}^{[1]} \neq 0, \dots,\ a_{n-1,n-1}^{[n-2]} \neq 0,\ a_{nn}^{[n-1]} \neq 0.$$

Если теперь обратиться к методу Лагранжа (доказательству теоремы 1.1), то увидим, что полученная матрица как раз и определяет замену переменных

$$\left\{\begin{array}{lllll} y_1 = & x_1 + \frac{a_{12}}{a_{11}} x_2 + & \dots + & \frac{a_{1,n-1}}{a_{11}} x_{n-1} + & \frac{a_{1n}}{a_{11}} x_n \\ y_2 = & x_2 + & \dots + & \frac{a_{2,n-1}^{[1]}}{a_{22}^{[1]}} x_{n-1} + & \frac{a_{2n}^{[1]}}{a_{22}^{[1]}} x_n \\ \vdots & & \ddots & & \dots \\ y_{n-1} = & & & x_{n-1} + & \frac{a_{n-1,n}^{[n-2]}}{a_{n-1,n-1}^{[n-2]}} x_n \\ y_n = & & & & x_n \end{array}\right. , \tag{2.3}$$

приводящую кв. форму к каноническому виду:

$$a_{11} y_1^2 + a_{22}^{[1]} y_2^2 + \dots + a_{n-1,n-1}^{[n-2]} y_{n-1}^2 + a_{nn}^{[n-1]} y_n^2. \tag{2.4}$$

■

Определение 2.1. Замена переменных вида (2.3) называется **треугольной**. Различают два типа треугольных замен переменных — правую и левую — по типу матриц, их задающих:

$$R \stackrel{\text{def}}{=} \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ & r_{22} & \dots & r_{2n} \\ \mathbb{O} & & \ddots & \vdots \\ & & & r_{nn} \end{pmatrix} \quad \text{и} \quad L \stackrel{\text{def}}{=} \begin{pmatrix} \ell_{11} & & & \\ \ell_{21} & \ell_{22} & & \mathbb{O} \\ \vdots & & \ddots & \\ \ell_{n1} & \ell_{n2} & \dots & \ell_{nn} \end{pmatrix}.$$

Теорема 2.2. *Матрица, обратная неособенной треугольной матрице, также является треугольной, причем того же типа.*

Доказательство проводится фактическим построением обратной матрицы. ■

Итак замена переменных вида (2.3) может быть записана в матричной форме:

$$Y = RX \quad \text{и если} \ \det R \neq 0, \quad \text{то } X = R^{-1}Y$$

и при этом R^{-1} остается правой треугольной. Этой заменой переменных кв. форма f преобразуется к каноническому виду (2.4):

$$\left(R^{-1}Y\right)^{\top} \mathbf{A} R^{-1} Y = Y^{\top} \left(R^{-1}\right)^{\top} \mathbf{A} R^{-1} Y = Y^{\top} \begin{pmatrix} a_{11} & & & \\ & a_{22}^{[1]} & & \mathbb{O} \\ \mathbb{O} & & \ddots & \\ & & & a_{nn}^{[n-1]} \end{pmatrix} Y.$$

Задача. Найти необходимые и достаточные условия приводимости кв. формы к каноническому виду с помощью треугольной замены переменных.

Матричный формализм метода Гаусса

Поставим более общую задачу: для произвольной (не обязательно симметричной) матрицы $\mathbf{A}$ найти условия существования неособенных треугольных матриц L и R таких, чтобы матрица

$L\mathbf{A}R$ была диагональной. Имея в виду результат теоремы 2.2, можем переформулировать эту задачу: установить возможность представления матрицы $\mathbf{A}$ в виде $L\mathbf{D}R$, где $\mathbf{D}$ — диагональная матрица.

Определение 2.2. Минор матрицы $\mathbf{A}$, стоящий в ее левом верхнем углу называется **главным**:

$$\det \mathbf{A}_k \stackrel{\text{def}}{=} \mathbf{A}\begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}.$$

Теорема 2.3. *Неособенная матрица* $\mathbf{A}$ *представима в виде произведения*

$$\underbrace{\begin{pmatrix} 1 & & & \\ \ell_{21} & 1 & & \mathbb{O} \\ \vdots & & \ddots & \\ \ell_{n1} & \ell_{n2} & \dots & 1 \end{pmatrix}}_{L} \underbrace{\begin{pmatrix} d_{11} & & & \\ & d_{22} & & \mathbb{O} \\ \mathbb{O} & & \ddots & \\ & & & d_{nn} \end{pmatrix}}_{\mathbf{D}} \underbrace{\begin{pmatrix} 1 & r_{12} & \dots & r_{1n} \\ & 1 & \dots & r_{2n} \\ \mathbb{O} & & \ddots & \vdots \\ & & & 1 \end{pmatrix}}_{R} \tag{2.5}$$

тогда и только тогда, когда

$$\det \mathbf{A}_1 \neq 0, \dots, \det \mathbf{A}_{n-1} \neq 0, \det \mathbf{A}_n \neq 0. \tag{2.6}$$

При выполнении условия (2.6) *представление* (2.5) *однозначно.*

Доказательство необходимости. Обозначим через L_k, $\mathbf{D}_k$, R_k главные подматрицы матриц L, $\mathbf{D}$, R:

$$L_k = \begin{pmatrix} 1 & & & \\ \ell_{21} & 1 & & \mathbb{O} \\ \vdots & & \ddots & \\ \ell_{k1} & \ell_{k2} & \dots & 1 \end{pmatrix}, \ \mathbf{D}_k = \begin{pmatrix} d_{11} & & & \\ & d_{22} & & \mathbb{O} \\ \mathbb{O} & & \ddots & \\ & & & d_{kk} \end{pmatrix}, \ R_k = \begin{pmatrix} 1 & r_{12} & \dots & r_{1k} \\ & 1 & \dots & r_{2k} \\ \mathbb{O} & & \ddots & \vdots \\ & & & 1 \end{pmatrix}.$$

Если справедливо представление (2.5), то справедливо и

$$\mathbf{A}_k = L_k \, \mathbf{D}_k \, R_k, \tag{2.7}$$

переходя к определителям, получим

$$\det \mathbf{A}_k = \det \mathbf{D}_k = d_{11} d_{22} \times \dots \times d_{kk}. \tag{2.8}$$

Последнее равенство при $k = n$ дает:

$$\det \mathbf{A} = \det \mathbf{A}_n = d_{11} d_{22} \times \cdots \times d_{nn} \neq 0,$$

поскольку, по предположению, $\det \mathbf{A} \neq 0$. Следовательно, все числа $d_{11}, \ldots, d_{nn}$ отличны от нуля, а тогда на основании (2.8) и все главные миноры отличны от нуля.

Д. Предположим теперь, что условия (2.6) выполнены. Будем строить матрицы из представления (2.5) построчечно. Для элемента a_{11} это представление дает равенство

$$a_{11} = 1 \cdot d_{11} \cdot 1 \;\Rightarrow\; d_{11} = a_{11},$$

т.е. определяется однозначно. Предположим, что нам удалось построить $(k-1)$ первых строк и столбцов матриц L, $\mathbf{D}$, R. Построим k-ые:

$$L_k = \begin{pmatrix} L_{k-1} & \mathbb{O} \\ Y & 1 \end{pmatrix}, \; \mathbf{D}_k = \begin{pmatrix} \mathbf{D}_{k-1} & \mathbb{O} \\ \mathbb{O} & d_{kk} \end{pmatrix}, \; R_k = \begin{pmatrix} R_{k-1} & X \\ \mathbb{O} & 1 \end{pmatrix};$$

здесь элементы столбца $X_{(k-1)\times 1}$, строки $Y_{1\times(k-1)}$ и d_{kk} пока не определены. Для их определения обратимся к равенству (2.8), выделив для удобства в матрице $\mathbf{A}_k$ элементы последней строки и последнего столбца:

$$\mathbf{A}_k = \begin{pmatrix} \mathbf{A}_{k-1} & U \\ V & a_{kk} \end{pmatrix} \quad , \; U \stackrel{\text{def}}{=} \begin{pmatrix} a_{1k} \\ \vdots \\ a_{k-1,k} \end{pmatrix}, \; V \stackrel{\text{def}}{=} (a_{k1}, \ldots, a_{k,k-1}) \, .$$

В этих обозначениях равенство (2.8) переписывается в виде системы матричных уравнений

$$\left\{ \begin{array}{ll} L_{k-1}\mathbf{D}_{k-1}R_{k-1} = \mathbf{A}_{k-1} \; , & L_{k-1}\mathbf{D}_{k-1}X = U \\ Y\mathbf{D}_{k-1}R_{k-1} = V \; , & Y\mathbf{D}_{k-1}X + d_{kk} = a_{kk} \end{array} \right.$$

Поскольку по индуктивному предположению матрицы $L_{k-1}, \mathbf{D}_{k-1}$ и R_{k-1} определяются однозначно и при этом $\det \mathbf{D}_{k-1} = \det A_{k-1} \neq 0$, то однозначно определяются и ряды X и Y:

$$X = (L_{k-1}\mathbf{D}_{k-1})^{-1} U, \quad Y = V (\mathbf{D}_{k-1}R_{k-1})^{-1},$$

тогда однозначно определится и элемент матрицы $\mathbf{D}_k$:

$$d_{kk} = a_{kk} - Y\mathbf{D}_{k-1}X = a_{kk} - V\mathbf{A}_{k-1}^{-1}U.$$

На основании формулы (2.7) этот элемент отличен от нуля. Итак, элементы k-х рядов матриц L, $\mathbf{D}$, R определяются однозначно. На основании индукции, получаем справедливость утверждения для любых рядов искомых матриц. ■

Следствие 2.1. *При выполнении условий* (2.6)*, элементы матрицы* $\mathbf{D}$ *определяются по формулам*

$$d_{11} = \det \mathbf{A}_1 = a_{11}, \; d_{22} = \frac{\det \mathbf{A}_2}{\det \mathbf{A}_1} = \frac{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}{a_{11}},$$

$$d_{33} = \frac{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \dots, d_{nn} = \frac{\det \mathbf{A}_n}{\det \mathbf{A}_{n-1}}. \tag{2.9}$$

Доказательство следует из формул (2.8). ■

Следствие 2.2. *При выполнении условий* (2.6)*, матрица* (2.2) *из метода Гаусса получается в виде произведения*

$$L^{-1}\mathbf{A},$$

т.е. домножением слева матрицы $\mathbf{A}$ *на неособенную левую треугольную.*

Формула Якоби

Теперь посмотрим как результаты предыдущего пункта проявляются для симметричных матриц.

Теорема 2.4. *Кв. форма* $f(X) = X^\top \mathbf{A} X$ *при симметричной неособенной матрице* $\mathbf{A}$ *приводится к каноническому виду заменой переменных*

$$X = CY \quad \textit{при } C = R = \begin{pmatrix} 1 & r_{12} & \dots & r_{1n} \\ & 1 & \dots & r_{2n} \\ \mathbb{O} & & \ddots & \vdots \\ & & & 1 \end{pmatrix}$$

тогда и только тогда, когда все главные миноры матрицы $\mathbf{A}$ *отличны от нуля.*

Доказательство необходимости. Если $C^\top \mathbf{A} C = \mathbf{D}$, то $\mathbf{A} = \left(C^\top\right)^{-1} \mathbf{D} C^{-1}$, где по теореме 2.2 матрица $C^{-1} = R^{-1}$ — правая треугольная, а матрица $\left(C^\top\right)^{-1} = \left(C^{-1}\right)^\top$ — левая треугольная. На основании теоремы 2.3, необходимым условием такого представления матрицы $\mathbf{A}$ является условие (2.6). ■

Доказательство достаточности. Пусть теперь выполнены условия (2.6). Теорема 2.3 утверждает, что возможно <u>единственное</u> представление матрицы $\mathbf{A}$ в виде произведения (2.5). В нашей теореме $\mathbf{A} = \mathbf{A}^\top$, таким образом

$$\mathbf{A}^\top = (L\mathbf{D}R)^\top = R^\top \mathbf{D}^\top L^\top = R^\top \mathbf{D} L^\top = \mathbf{A},$$

иначе говоря, для матрицы $\mathbf{A}$ существует еще одно представление в виде произведения (2.5). Противоречие может быть устранено только условием $R^\top = L$, но тогда матрица $C \overset{\text{def}}{=} R$ — та, что требуется. ■

Следствие 2.3. *Если все главные миноры матрицы* $\mathbf{A}$ *отличны от нуля, то кв.форма* $f(X) = X^\top \mathbf{A} X$ *приводится к следующему каноническому виду (***формула Якоби***):*

$$\det \mathbf{A}_1 y_1^2 + \frac{\det \mathbf{A}_2}{\det \mathbf{A}_1} y_2^2 + \ldots + \frac{\det \mathbf{A}_n}{\det \mathbf{A}_{n-1}} y_n^2 = \sum_{k=1}^{n} \frac{\det \mathbf{A}_k}{\det \mathbf{A}_{k-1}} y_k^2 \ (\det \mathbf{A}_0 \overset{\text{def}}{=} 1) \tag{2.10}$$

"треугольной" заменой переменных (2.3)*.*

Упражнение 75. *Найти треугольное преобразование, приводящее кв. форму* $\sum\limits_{i,j=1}^{n} \min\{i,j\}\, x_i x_j$ *к каноническому виду.*

Упражнение 76. *Найти канонический вид кв.формы* $\sum\limits_{i,j=1}^{n} \max\{i,j\}\, x_i x_j\,.$

§3. Закон инерции

Разумеется, для заданной кв.формы канонические виды, т.е. представления в виде сумм квадратов, можно построить разными способами. Выясним, какие характеристики являются общими (инвариантными) для этих представлений.

Ранг квадратичной формы

Предположим, что с помощью какой-либо замены переменных (1.2) мы привели кв.форму к каноническому виду:

$$\widetilde{f}(Y) = \alpha_1 y_1^2 + \cdots + \alpha_n y_n^2.$$

Но может так случиться, что часть коэффициентов обратится в нуль.

Определение 3.1. **Рангом** кв.формы называется ранг ее матрицы:

$$\operatorname{rank} f \stackrel{\text{def}}{=} \operatorname{rank} \mathbf{A}.$$

Теорема 3.1. *Ранг кв.формы не меняется при неособенных заменах переменных:*

$$\operatorname{rank} f = \operatorname{rank} C^{\top} \mathbf{A} C \quad \textit{при} \quad \forall C,\ \det C \neq 0.$$

Доказательство основано на следствии 11.4 к теореме 11.8 главы 4: ранг матрицы не меняется при домножении ее на произвольную неособенную. ■

Следствие 3.1. *Ранг кв.формы равен числу ненулевых коэффициентов в ее каноническом виде.*

Закон инерции

Начиная с этого момента рассматриваем только вещественные кв.формы: $\mathbb{A} = \mathbb{R}$. Алгоритмы § 1 приводят к вещественному каноническому виду:

$$\widetilde{f}(Y) = \alpha_1 y_1^2 + \dots + \alpha_{\mathfrak{r}} y_{\mathfrak{r}}^2, \quad , \alpha_j \neq 0, \alpha_j \in \mathbb{R} \ (\mathfrak{r} \overset{\text{def}}{=} \operatorname{rank} f).$$

Так, например, если все главные миноры $\det \mathbf{A}_1, \dots, \det \mathbf{A}_{\mathfrak{r}}$ отличны от нуля, то справедлив аналог формулы Якоби (2.10):

$$X^\top \mathbf{A} X = \det \mathbf{A}_1 \, y_1^2 + \frac{\det \mathbf{A}_2}{\det \mathbf{A}_1} y_2^2 + \dots + \frac{\det \mathbf{A}_{\mathfrak{r}}}{\det \mathbf{A}_{\mathfrak{r}-1}} y_{\mathfrak{r}}^2 = \sum_{k=1}^{\mathfrak{r}} \frac{\det \mathbf{A}_k}{\det \mathbf{A}_{k-1}} y_k^2. \tag{3.1}$$

Здесь $\det \mathbf{A}_0 \overset{\text{def}}{=} 1$.

Определение 3.2. Число положительных (или отрицательных) коэффициентов в каноническом виде кв.формы называется ее **положительным** (или **отрицательным**) **индексом инерции** $n_+(f)$ (или $n_-(f)$). Разность

$$\sigma(f) \overset{\text{def}}{=} n_+(f) - n_-(f)$$

называется **сигнатурой** кв.формы.

Теорема 3.2. *[закон инерции] Индексы инерции не зависят от способа приведения кв.формы к каноническому виду.*

Доказательство. Предположим, что в результате двух подстановок кв. форма приведена к различным каноническим видам:

$$X^\top \mathbf{A} X = \begin{cases} \alpha_1 y_1^2 + \ldots + \alpha_p y_p^2 - \alpha_{p+1} y_{p+1}^2 - \ldots - \alpha_{\mathfrak{r}} y_{\mathfrak{r}}^2 & \text{при } Y = C_1 X; \\ \beta_1 z_1^2 + \ldots + \beta_q z_q^2 - \beta_{q+1} z_{q+1}^2 - \ldots - \beta_{\mathfrak{r}} z_{\mathfrak{r}}^2 & \text{при } Z = C_2 X. \end{cases}$$

Здесь все α_j, β_k положительны, но $q > p$. Можно подобрать вектор $X = X_* \neq \mathbb{O}$ так, чтобы у вектора $Y_* = C_1 X_*$ первые p компонент, а у вектора $Z_* = C_2 X_*$ последние $\mathfrak{r} - q$ компонент обращались в нуль. Действительно, число уравнений для определения n компонент нужного нам вектора равно $p + \mathfrak{r} - q = \mathfrak{r} - (q-p) < \mathfrak{r} \le n$. Такая система всегда имеет нетривиальное решение X_*, при этом соответствующий ему вектор Y_* будет ненулевым[2]. Подстановка Y_* и Z_* в канонические виды дает:

$$X_*^\top \mathbf{A} X_* = \begin{cases} -\alpha_{p+1} y_{p+1,*}^2 - \cdots - \alpha_\mathfrak{r} y_{\mathfrak{r},*}^2 & < 0; \\ \beta_1 z_{1,*}^2 + \cdots + \beta_q z_{q,*}^2 & \ge 0. \end{cases}$$

Противоречие доказывает ошибочность предположения $q > p$. ■

Пример 3.1. Найти ранг и сигнатуру кв.формы $f(x_1, x_2, x_3) = x_1 x_2 - x_2 x_3$.

Решение. Приводим кв.форму к каноническому виду по методу Лагранжа:

$$f = {}^1\!/_4\,(x_1 + x_2 - x_3)^2 - {}^1\!/_4\,(x_1 - x_2 - x_3)^2.$$

Ответ. $\operatorname{rank} f = 2$, $\sigma(f) = 0$.

Следствие 3.2. *В предположении, что* $\det \mathbf{A}_1 \ne 0, \dots, \det \mathbf{A}_\mathfrak{r} \ne 0$ ($\mathfrak{r} \stackrel{\text{def}}{=} \operatorname{rank} \mathbf{A}$) *имеем:*

$$n_+(f) = \mathcal{P}(1, \det \mathbf{A}_1, \dots, \det \mathbf{A}_\mathfrak{r}), \ \ n_-(f) = \mathcal{V}(1, \det \mathbf{A}_1, \dots, \det \mathbf{A}_\mathfrak{r}). \tag{3.2}$$

Здесь $\mathcal{P}$ *— число постоянств, а* $\mathcal{V}$ *— число перемен знака в числовой последовательности.*

Доказательство фактически следует из формулы Якоби (3.1). ■

Пример 3.2. Найти ранг и сигнатуру кв.формы

$$f_\alpha(x_1, x_2, x_3) = 3\,x_1^2 - 4\,x_1 x_2 - 2\,x_1 x_3 + \alpha\,x_2^2 + 6\,x_2 x_3$$

в зависимости от значений параметра α.

Решение. Сначала пробуем применить формулу (3.2):

$$\det \mathbf{A}_1 = 3, \ \det \mathbf{A}_2 = 3\,\alpha - 4, \ \det \mathbf{A}_3 = \det \mathbf{A} = -\alpha - 15.$$

[2] Глава 4, теорема 12.2.

При $\alpha \notin \{\, ^4\!/_3,\, -15\}$ формула (3.2) применима при $\mathfrak{r} = 3$:

$$n_+(f) = \begin{cases} 2 & \text{при} \quad \alpha > {}^4\!/_3; \\ 2 & \text{при} \quad -15 < \alpha < {}^4\!/_3; \\ 1 & \text{при} \quad \alpha < -15. \end{cases}$$

При $\alpha = {}^4\!/_3$, по-прежнему, $\mathfrak{r} = 3$, но формула (3.2) неприменима. В этом случае приходится действовать по методу Лагранжа:

$$f_{4/3}(x_1,x_2,x_3) = 3\,(x_1 - {}^2\!/_3\, x_2 - {}^1\!/_3\, x_3)^2 - {}^1\!/_3\,(x_3 - 7\,x_2)^2 + {}^{49}\!/_3 x_2^2.$$

Следовательно, $n_+(f) = 2$. Осталось рассмотреть случай $\alpha = -15$, когда $\mathfrak{r} = 2$. Поскольку условие следствия 1 к теореме 3.2 выполняются, то формула (3.2) применима: $n_+(f) = 1$.

Во всех случаях отрицательный индекс инерции вычисляется по формуле $n_-(f) = \mathfrak{r} - n_+(f)$.

Ответ. $\operatorname{rank} f = 3, \sigma(f) = 1$ при $\alpha > -15$; $\operatorname{rank} f = 3, \sigma(f) = -1$ при $\alpha < -15$; $\operatorname{rank} f = 2, \sigma(f) = 0$ при $\alpha = -15$.

Заметим, что при непрерывном изменении параметра сигнатура кв.формы может измениться только когда параметр проходит через значение, при котором происходит изменение ранга.

Конгруэнтность кв.форм

Определение 3.3. Матрицы $\mathbf{A}$ и $\mathbf{B}$, связанные соотношением $\mathbf{B} = C^\top \mathbf{A} C$ при некоторой неособенной матрице C, называются **конгруэнтными**: $\mathbf{A} \cong \mathbf{B}$. Если, вдобавок, матрицы $\mathbf{A}$ и $\mathbf{B}$ симметричны, то конгруэнтными называются и соответствующие им кв. формы $X^\top \mathbf{A} X$ и $X^\top \mathbf{B} X$.

Теорема 3.3. *Справедливы следующие свойства отношения конгруэнтности:*

а) *симметричность: если* $\mathbf{A}_1 \cong \mathbf{A}_2$, *то* $\mathbf{A}_2 \cong \mathbf{A}_1$;

б) *транзитивность: если* $\mathbf{A}_1 \cong \mathbf{A}_2$ *и если* $\mathbf{A}_2 \cong \mathbf{A}_3$, *то* $\mathbf{A}_1 \cong \mathbf{A}_3$;

в) *рефлексивность:* $\mathbf{A} \cong \mathbf{A}$.

Иными словами, отношение конгруэнтности является отношением **эквивалентности**.

Доказательство. **а)** Если $\mathbf{A}_2 = C_1^\top \mathbf{A}_1 C_1$ при $\det C_1 \neq 0$, то $\mathbf{A}_1 = C_2^\top \mathbf{A}_2 C_2$ при $C_2 = C_1^{-1}$.

б) Если $\mathbf{A}_2 = C_1^\top \mathbf{A}_1 C_1$ и $\mathbf{A}_3 = C_2^\top \mathbf{A}_2 C_2$ то $\mathbf{A}_3 = C_3^\top \mathbf{A}_1 C_3$ при $C_3 \stackrel{\text{def}}{=} C_1 C_2$. Матрица C_3 неособенная, если C_1 и C_2 неособенные.

Пункт **в)** теоремы доказывается выбором $C_1 \stackrel{\text{def}}{=} E$. ■

Теорема 3.4. *Кв.формы $X^\top \mathbf{A} X$ и $X^\top \mathbf{B} X$ конгруэнтны тогда и только тогда, когда совпадают их индексы инерции, или, что то же, равны их ранги и сигнатуры.*

Доказательство необходимости. следует из теоремы 3.2 (закона инерции). Если $\mathbf{A} = C^\top \mathbf{B} C$ при $\det C \neq 0$ и кв. форма $X^\top \mathbf{A} X$ приводится к каноническому виду преобразованием $X = C_1 Y$ ($\det C_1 \neq 0$), то к тому же каноническому виду приводится и кв.форма $X^\top \mathbf{B} X$ преобразованием $X = C C_1 Y$. ■

Доказательство достаточности. Рассмотрим канонические виды кв.форм $X^\top \mathbf{A} X$ и $X^\top \mathbf{B} X$:

$$\begin{aligned} X^\top \mathbf{A} X &= \alpha_1 y_1^2 + \cdots + \alpha_{n_+(A)} y_{n_+(A)}^2 - \\ &\qquad -\alpha_{n_+(A)+1} y_{n_+(A)+1}^2 - \cdots - \alpha_{\mathfrak{r}(A)} y_{\mathfrak{r}(A)}^2 \end{aligned} \tag{3.3}$$

$$\begin{aligned} X^\top \mathbf{B} X &= \widetilde{\alpha}_1 \widetilde{y}_1^2 + \cdots + \widetilde{\alpha}_{n_+(B)} \widetilde{y}_{n_+(B)}^2 - \\ &\qquad -\widetilde{\alpha}_{n_+(B)+1} \widetilde{y}_{n_+(B)+1}^2 - \cdots - \tilde{\alpha}_{\mathfrak{r}(B)} \widetilde{y}_{\mathfrak{r}(B)}^2. \end{aligned} \tag{3.4}$$

Здесь все числа $\alpha, \widetilde{\alpha}$ предполагаются положительными. Если индексы инерции кв. форм одинаковы, то количества положительных коэффициентов в канонических видах одинаковы, и количества отрицательных коэффициентов одинаковы. Замена переменных

$$\widetilde{y}_j = y_j \sqrt{\frac{\alpha_j}{\tilde{\alpha}_j}} \quad \text{при} \quad j \in \{1, \ldots, \mathfrak{r}\},$$
$$\widetilde{y}_j = y_j \quad \text{при} \quad j \in \{\mathfrak{r}+1, \ldots, n\}$$

будет неособенной и переведет один канонический вид в другой. Таким образом канонические виды кв.форм будут конгруэнтными, но тогда, на основании пунктов **а)** и **б)** теоремы 3.3, будут конгруэнтны и сами формы $X^\top \mathbf{A} X$ и $X^\top \mathbf{B} X$. ■

Из всего разнообразия канонических видов, которые можно построить для кв.формы, выберем самый простой, именно тот, коэффициенты которого равны $+1$ или -1. Например, если кв. форма $f(X)$ имеет канонический вид (3.3), то преобразование

$$y_j = \frac{z_j}{\sqrt{\alpha_j}} \text{ при } j \in \{1,\dots,\mathfrak{r}\}\ ,\ y_j = z_j \text{ при } j \in \{\mathfrak{r}+1,\dots,n\}$$

приводит кв. форму к виду

$$z_1^2 + \dots + z_{n_+(A)}^2 - z_{n_+(A)+1}^2 - \dots - z_{\mathfrak{r}}^2,$$

который называется **нормальным** видом кв.формы.

Множество всех кв.форм над $\mathbb{R}$ можно разбить на **классы** эквивалентности, в каждом из которых будут находиться только конгруэнтные между собой формы. Каждый из классов полностью описывается каким-то из своих представителей. Таким представителем можно взять нормальный вид.

Упражнение 77. *Сколько существует классов эквивалентности на множестве форм от n переменных?*

Упражнение 78. *Какие из следующих форм принадлежат одинаковым классам эквивалентности:*

$$\begin{array}{rcl} f_1 &=& x_1^2 + 4\,x_2^2 + x_3^2 + 4\,x_1x_2 - 2\,x_1x_3\ , \\ f_2 &=& x_1^2 + 2\,x_2^2 - x_3^2 + 4\,x_1x_2 - 2\,x_1x_3 - 4\,x_2x_3\ , \\ f_3 &=& -4\,x_1^2 - x_2^2 - x_3^2 - 4\,x_1x_2 - 4\,x_1x_3 + 18\,x_2x_3\ ? \end{array}$$

Упражнение 79. *Доказать, что для того чтобы кв.форма $f(X) \in \mathbb{R}[X]$ могла быть представлена в виде произведения двух линейных форм над $\mathbb{R}$ необходимо и достаточно, чтобы $\operatorname{rank} f \le$ ≤ 2, а при $\operatorname{rank} f = 2$ еще и выполнения условия $\sigma(f) = 0$. Удовлетворяет ли этим условиям кв.форма $\displaystyle\sum_{i,j=1}^n (i+j)x_ix_j$?*

§4. Положительная определенность

Здесь рассматриваются только вещественные кв.формы.

Перейти к оглавлению на странице: 460

Определение 4.1. Кв.форма $f(X) = X^{\top}\mathbf{A}X$ называется

а) неотрицательной если $f(X) \ge 0$ для любого $X \in \mathbb{R}^n$;

б) положительно определенной, если она неотрицательна и $f(X) = 0 \iff X = \mathbb{O}$;

в) неопределенной, если существуют $X_1, X_2 \in \mathbb{R}^n$ такие что $f(X_1)$ и $f(X_2)$ имеют разные знаки: $f(X_1)f(X_2) < 0$.

По аналогии с пунктами а) и б) определяются неположительные и отрицательно определенные кв.формы. Иногда неотрицательные или неположительные формы называют **полуопределенными**.

Пример 4.1. При $n = 3$ кв.форма

а) $x_1^2 + 3x_2^2 + 4x_3^2$ будет полож.определенной;

б) $x_1^2 + x_3^2$ (или $(x_1 + \sqrt{2}x_2 - x_3)^2$) будет неотрицательной, но не полож.определенной;

в) $-x_1^2$ будет неположительной, но не отрицательно определенной;

г) $-x_1^2 - 2x_2^2 - 4\,x_3^2$ будет отрицательно определенной;

д) $x_1x_2 + x_2x_3$ будет неопределенной.

Очевидны необходимые условия неотрицательности кв.формы: все коэффициенты при квадратах переменных должны быть неотрицательными: $a_{11} \ge 0, \dots, a_{nn} \ge 0$. Эти же условия, очевидно, будут и достаточными, если все остальные коэффициенты a_{jk} при $j \ne k$ обратятся в нуль. Если же последнее не выполняется, то имеет смысл предварительно преобразовать кв.форму к сумме квадратов, т.е. исследовать ее канонический вид.

Теорема 4.1. *Ненулевая квадратичная форма $f(X) = X^{\top}\mathbf{A}X$ будет неотрицательной тогда и только тогда, когда*

$$n_-(\mathbf{A}) = 0 \qquad \text{или, что то же} \qquad \sigma(\mathbf{A}) = \operatorname{rank}\mathbf{A}.$$

Если это условие выполнено, то для полож.определенности необходимо и достаточно чтобы $\det\mathbf{A} \ne 0$.

Доказательство. По теореме 1.1 всегда найдется замена переменных $X = CY$, приводящая кв.форму к каноническому виду:

$$f(X) = \alpha_1 y_1^2 + \dots + \alpha_{\mathfrak{r}} y_{\mathfrak{r}}^2\ ,\ \mathfrak{r} = \operatorname{rank} f.$$

Для записи обратной замены переменных $Y = C^{-1}X$ обозначим $\widetilde{C} \overset{\text{def}}{=} C^{-1}$: $\widetilde{C} = [\widetilde{c}_{jk}]_{j,k=1}^{n}$.

Д. Если $n_-(\mathbf{A}) = 0$, то $\alpha_1 > 0, \ldots, \alpha_{\mathfrak{r}} > 0$, и тогда $f(X) \geq 0$.

Н. Пусть теперь $f(X) \geq 0$, но $n_-(\mathbf{A}) > 0$, т.е. в каноническом виде какой-то коэффициент отрицателен: $\alpha_j < 0$. Система линейных уравнений

$$\begin{cases} \widetilde{c}_{11}x_1 & + \cdots + & \widetilde{c}_{1n}x_n & = 0 \\ \ldots & & & \ldots \\ \widetilde{c}_{j-1,1}x_1 & + \cdots + & \widetilde{c}_{j-1,n}x_n & = 0 \\ \widetilde{c}_{j+1,1}x_1 & + \cdots + & \widetilde{c}_{j+1,n}x_n & = 0 \\ \ldots & & & \ldots \\ \widetilde{c}_{n1}x_1 & + \cdots + & \widetilde{c}_{nn}x_n & = 0 \end{cases}$$

является однородной и имеет нетривиальное решение $X = X_*$ (поскольку число уравнений меньше числа переменных). Соответствующие значения переменных $Y = Y_* = \widetilde{C}X_*$ обратятся в нуль:

$$y_{1*} = 0, \ldots, y_{j-1,*} = 0, y_{j+1,*} = 0, \ldots, y_{n*} = 0.$$

Однако же $y_{j,*}$ должно быть отличным от нуля. В противном случае $Y_* = \mathbb{O}$ и система $\widetilde{C}X = \mathbb{O}$ имеет нетривиальное решение $X = X_*$. Тогда на основании следствия 1 к теореме 12.2 главы 4, должно выполняться $\det \widetilde{C} = 0$, что невозможно, поскольку $\det C \neq 0$. Поэтому $f(X_*) = \alpha_j y_{j,*}^2 < 0$, но это противоречит предположению о неотрицательности $f(X)$. Итак, предположение $n_-(\mathbf{A}) > 0$ ошибочно, т.е. $n_-(\mathbf{A}) = 0$.

Осталось доказать, что для полож. определенности Н. и Д., чтобы $\mathfrak{r} = n$, т.е. $\det \mathbf{A} \neq 0$. Предположив, что $\mathfrak{r} < n$, мы приходим к существованию $X = X_{**} \neq \mathbb{O}$ такого, что первые $\mathfrak{r}$ соответствующих значений переменных $Y = Y_{**} = \widetilde{C}X_{**}$ обратятся в нуль: в системе

$$\begin{cases} \widetilde{c}_{11}x_1 & + \cdots + & \widetilde{c}_{1n}x_n & = 0, \\ \ldots & & & \ldots \\ \widetilde{c}_{\mathfrak{r}1}x_1 & + \cdots + & \widetilde{c}_{\mathfrak{r}n}x_n & = 0 \end{cases}$$

число уравнений меньше числа переменных. Это значит, что $f(X_{**}) = 0$ при $X_{**} \neq \mathbb{O}$, что противоречит полож. определенности. Следовательно, $\mathfrak{r} = n$.

Обратно, пусть $\mathfrak{r} = n = n_+(f)$, т.е.

$$f(X) = \alpha_1 y_1^2 + \cdots + \alpha_n y_n^2 \quad \alpha_1 > 0, \ldots, \alpha_n > 0.$$

Очевидно, что $f(X) = 0$ тогда и только тогда, когда $Y = \mathbb{O}$, но это возможно только при условии $X = \mathbb{O}$ поскольку $X = CY$ при $\det C \neq 0$. ■

Теорема 4.2. *[Сильвестр] Для полож.определенности кв.формы $f(X) = X^\top \mathbf{A} X$ необходимо и достаточно чтобы все главные миноры матрицы $\mathbf{A}$ были положительны:*

$$\det \mathbf{A}_1 = a_{11} > 0,\ \det \mathbf{A}_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} > 0,\ \ldots, \det \mathbf{A}_n = \det \mathbf{A} > 0. \tag{4.1}$$

Доказательство достаточности. фактически следует из формулы (2.10) (Якоби) и теоремы 4.1. ■

Доказательство необходимости. Пусть $f(X) = X^\top \mathbf{A} X$ полож.определена. Тогда любая форма

$$f_j(x_1, \ldots, x_j) \stackrel{\text{def}}{=} f(x_1, \ldots, x_j, 0, \ldots, 0) = X_j^\top \mathbf{A}_j X_j,$$

рассматриваемая как форма относительно переменных $X_j \stackrel{\text{def}}{=} (x_1, \ldots, x_j)^\top$. также будет полож.определенной. В самом деле, если $f(X) \geq 0$, то и $f_j(X_j) \geq 0$. Если предположить, что существует $X_j = X_{j*} \neq \mathbb{O}$ такой, что $f_j(X_j) = 0$, то при $X = X_* = (x_{1*}, \ldots, x_{j*}, 0, \ldots, 0)^\top$ обратится в нуль и сама форма $f(X)$. Это противоречит предположенной полож. определенности формы. Следовательно, форма $f_j(X_j)$ является полож. определенной относительно своих переменных. Но тогда, согласно теореме 4.1, ее дискриминант должен быть отличным от нуля: $\det \mathbf{A}_j \neq 0$. Поскольку предшествующие рассуждения справедливы для любого индекса j, то получаем возможность воспользоваться формулой (3.2) для вычисления индекса инерции. Теорема 4.1 требует, чтобы $n_+(f) = n$, но это возможно только когда все главные миноры положительны. ■

Примером положительно определенной матрицы может служить матрица Гильберта[3].

[3] Глава 4, §9.2.

Пример 4.2. Найти все значения параметра α, при которых кв. форма

$$2\,x_1^2+2\,x_2^2+x_3^2+2\,\alpha\,x_1x_2+6\,x_1x_3+2\,x_2x_3$$

будет полож.определенной.

Решение. Имеем:

$$\det \mathbf{A}_1=2,\ \det \mathbf{A}_2=4-\alpha^2,\ \det \mathbf{A}_3=-\alpha^2+6\,\alpha-16.$$

Последнее выражение будет отрицательно при любых $\alpha\in\mathbb{R}$.

Ответ. $\alpha\in\varnothing$.

Следствие 4.1. *Для отрицательной определенности* $f(X) = X^\top\mathbf{A}X$ *необходимо и достаточно чтобы знаки главных миноров чередовались:*

$$\det \mathbf{A}_1<0,\ \det \mathbf{A}_2>0,\ \det \mathbf{A}_3<0,\ \dots,\ (-1)^n\det \mathbf{A}_n>0.$$

Замечание 4.1. Можно ли получить условия неотрицательности кв. формы, преобразовав неравенства Сильвестра (4.1) в нестрогие:

$$\det \mathbf{A}_1\ge 0,\ \det \mathbf{A}_2\ge 0,\ \dots,\det \mathbf{A}_n\ge 0 \quad ?$$

Оказывается, что такие условия являются необходимыми, но не достаточными для неотрицательности: для кв. формы

$$f(x_1,x_2,x_3,x_4)=x_1^2+2\,x_1x_3+2\,x_2x_4+x_4^2=X^\top\begin{pmatrix}1&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&1\end{pmatrix}X$$

они выполняются, но она — неопределенная, т.к. $f(1,0,-1,0) = -1<0$, а $f(1,0,0,0)=1>0$.

Замечание 4.2. Где используются квадратичные формы? Прежде всего, в математическом анализе. В §11 главы 3 мы привели теорему 11.9 о достаточном условии наличия экстремума полинома нескольких переменных в некоторой точке: для того, чтобы полином $f(x_1,x_2,\dots,x_\ell)$ имел в точке $(c_1,c_2,\dots,c_\ell)$ локальный минимум (максимум) достаточно, чтобы квадратичная форма

$$\left(\frac{\partial\Box}{\partial x_1}X_1+\frac{\partial\Box}{\partial x_2}X_2+\dots+\frac{\partial\Box}{\partial x_\ell}X_\ell\right)^2 f$$

Перейти к оглавлению на странице: 460

от переменных $X_1, \dots, X_\ell$ была положительной (отрицательной) при всех вещественных значениях этих переменных, среди которых хотя бы одно было отлично от нуля. Последнее условие как раз и является определением знакоопределенной квадратичной формы.

Перейти к оглавлению на странице: 460

ГЛАВА 8. ЛОКАЛИЗАЦИЯ КОРНЕЙ ПОЛИНОМА

Если мы хотим найти приближенное значение корня полинома применением какого-либо численного метода, то мы должны предварительно дать оценку точности вычислений: сколько значащих цифр мы должны сохранять в промежуточных выкладках, чтобы гарантировать достоверность получаемых результатов?

Это порождает более общую задачу оценки *чувствительности* корней, т.е. оценки их изменения при некотором возмущении коэффициентов полинома. Принципиальным результатом здесь является теорема 7.8 главы 3 о непрерывной зависимости корней полинома от его коэффициентов. Выясним на примере, насколько малым может быть возмущение коэффициентов полинома $f(x) = x^n + a_1 x^{n-1} + \dots + a_n \in \mathbb{R}[x]$, чтобы сохранилось неизменным хотя бы число его вещественных корней. Для определенности, рассмотрим случай, когда полином $f(x)$ не имеет кратных корней (т.е. $\mathcal{D}(f) \neq 0$). Тогда число вещественных корней не изменится при малых (вещественных) вариациях его коэффициентов. В самом деле, вещественный корень полинома с вещественными коэффициентами не может "сойти" с вещественной оси, т.к. мнимые корни полинома над $\mathbb{R}$ образуют комплексно-сопряженные пары. Согласно теореме о непрерывной зависимости корней полинома от его коэффициентов один корень не может "расщепиться" на два. Осталось только оценить малость допустимого возмущения.

Пример. [Уилкинсон]. *Вычислить корни полинома*

$$
\begin{aligned}
f(x)={}&x^{20}+210x^{19}+20615x^{18}+1256850x^{17}+53327946x^{16}+1672280820x^{15}+\\
&+40171771630x^{14}+756111184500x^{13}+11310276995381x^{12}+\\
&+135585182899530x^{11}+1307535010540395x^{10}+10142299865511450x^{9}+\\
&+63030812099294896x^{8}+311333643161390640x^{7}+1206647803780373360x^{6}+\\
&+3599979517947607200x^{5}+8037811822645051776x^{4}+12870931245150988800x^{3}+\\
&+13803759753640704000x^{2}+8752948036761600000x+2432902008176640000
\end{aligned}
$$

по методу Ньютона.

Решение. Забегая вперед, укажем ответ: данный полином имеет каноническое разложение

$$f(x) = \prod_{j=1}^{20} (x+j),$$

и таким образом его корнями являются числа $-1, -2, \ldots, -20$, достаточно хорошо разнесенные. Пусть, однако же, эта информация нам изначально недоступна, и мы для поиска корней применяем метод Ньютона[1], задав точность вычислений в 10^{-5}. Является ли эта точность достаточной для нахождения приближенных значений корней? Оказывается, нет. Рассмотрим полином

$$\widetilde{f}_\varepsilon(x) \stackrel{\text{def}}{=} f(x) + \varepsilon\, x^{19} \qquad \varepsilon = 2^{-23} \approx 1.192092896 \times 10^{-7}.$$

Уже при таком малом возмущении в одном-единственном коэффициенте происходит потеря **десяти** вещественных корней! Корнями $\widetilde{f}_\varepsilon(x)$ являются

$$\lambda_1 = -1.000000,\ \lambda_2 = -2.000000,\ \lambda_3 = -3.000000,\ \lambda_4 = -4.000000,$$

$$\lambda_5 = -4.999999,\ \lambda_6 = -6.000007,\ \lambda_7 = -6.999697,\ \lambda_8 = -8.007268,$$

$$\lambda_9 = -8.917250$$

$$\lambda_{10,11} = -10.095266 \pm 0.643501\, \boldsymbol{i}\ ,\ \lambda_{12,13} = -11.793634 \pm 1.652330\, \boldsymbol{i}\ ,$$

$$\lambda_{14,15} = -13.992358 \pm 2.518830\, \boldsymbol{i}\ ,\ \lambda_{16,17} = -16.730737 \pm 2.812625\, \boldsymbol{i}\ ,$$

$$\lambda_{18,19} = -19.502439 \pm 1.940330\, \boldsymbol{i}\ ,$$

$$\lambda_{20} = -20.846908.$$

В данном примере допустимые возмущения для ε, т.е. такие, при которых сохранится свойство вещественности всех корней $\widetilde{f}_\varepsilon(x)$, находятся в пределах

$$-1.3508 \times 10^{-10} \ < \ \varepsilon \ < +1.4213 \times 10^{-10}.$$

$\triangle$

[1] Глава 3, §9.

Задача. Для полинома $f(z)\in\mathbb{C}[z]$ получить точную информацию о числе его корней в заданной области $\mathbb{S}\subset\mathbb{C}$.

Оказывается, для достаточно широкого класса областей $\mathbb{S}$ эту информацию можно получить без применения численных, т.е. приближенных методов. Существуют алгоритмы, позволяющие за конечное число элементарных алгебраических операций $(+,-,\times,\div)$ над коэффициентами $f(z)$ установить количество корней этого полинома в таких областях, как, к примеру,

$$
\begin{aligned}
\mathbb{S} &= \{z\in\mathbb{R}\big|a<z<b\} \text{ при } \{a,b\}\subset\mathbb{R},\\
\mathbb{S} &= \{z\in\mathbb{C}\big|\operatorname{Re} e(z)<0\},\\
\mathbb{S} &= \{z\in\mathbb{C}\big||z|<1\}.
\end{aligned}
$$

§1. Теорема Штурма

Задача. Для полинома $f(x)\in\mathbb{R}[x]$ установить точное число его корней на заданном интервале $]a,b[$: $\operatorname{nrr}\{f(x)=0\mid a<x<b\}$.

Система полиномов Штурма

Определение 1.1. Для полинома $f(x)$ система полиномов

$$
f_0(x)\overset{\text{def}}{=}f(x),f_1(x),\dots,f_K(x) \tag{1.1}
$$

называется **системой полиномов Штурма**[2] (сокращение **с.п.Шт.**) на заданном интервале $]a,b[$ если на этом интервале

А) соседние полиномы $f_j(x)$ и $f_{j+1}(x)$ не имеют общих корней;

Б) $f_K(x)\neq 0$;

В) если $f_j(x_0)=0$ при $x_0\in]a,b[$ и $j=1,\dots,k-1$, то числа $f_{j-1}(x_0)$ и $f_{j+1}(x_0)$ имеют разные знаки: $f_{j-1}(x_0)f_{j+1}(x_0)<0$;

[2]**Штурм** (Стюрм) ***Жак Шарль Франсуа*** (Sturm Jaques Charles François, 1803–1855). Уроженец Швейцарии, работал и жил во Франции. Как иностранец и протестант, он долгое время занимал сравнительно скромное положение, пока в 1836 г. благодаря самоотверженной поддержке своих друзей Лиувилля и Дюамеля, которые в последний момент сняли свои кандидатуры, не был избран членом Парижской академии наук. Работы по алгебре и дифференциальным уравнениям.

Г) произведение $f_0(x)f_1(x)$ меняет знак с отрицательного на положительный когда x, возрастая, проходит корень $\lambda \in]a, b[$ полинома $f_0(x) \equiv f(x)$.

Число перемен знака

$$\mathcal{V}_x \stackrel{\text{def}}{=} \mathcal{V}(f_0(x), f_1(x), \ldots, f_K(x)) \tag{1.2}$$

при x возрастающем от a к b, будет меняться когда x проходит через корень какого-либо полинома системы. Мы покажем, что это число может разве лишь уменьшаться, и уменьшается на единицу тогда и только тогда, когда x проходит через корень начального полинома системы (1.1), т.е. $f(x)$.

Лемма 1.1. *[о стабилизации знака непрерывной функции] Пусть $F(x)$ — функция непрерывная в некоторой окрестности точки x_0. Если $F(x_0) \neq 0$, то существует такая окрестность точки x_0, в которой*

$$\operatorname{sign} F(x) = \operatorname{sign} F(x_0).$$

Теорема 1.1. *Если $f(a) \neq 0$, $f(b) \neq 0$, и система (1.1) является системой полиномов Штурма для $f(x)$, то*

$$\begin{gathered}\operatorname{nrr}\{f(x){=}0 \mid a < x < b\} = \mathcal{V}_a - \mathcal{V}_b = \\ = \mathcal{V}(f_0(a), f_1(a), \ldots, f_K(a)) - \mathcal{V}(f_0(b), f_1(b), \ldots, f_K(b)).\end{gathered} \tag{1.3}$$

Доказательство. Начнем движение аргумента от a к b. Пусть x_0 — первый корень какого-то из полиномов $f_j(x)$ системы (1.1), встретившийся нам при этом движении. Предположим, что $f_\ell(x_0) \neq 0$ при $\ell \neq j$. Проанализируем различные возможные ситуации для индекса j.

I. Пусть $j \neq 0, j \neq K$. По свойству А числа $f_{j-1}(x_0)$ и $f_{j+1}(x_0)$ не равны нулю, а по свойству В они должны иметь разные знаки. Следовательно

$$\mathcal{V}(f_{j-1}(x_0), 0, f_{j+1}(x_0)) = 1.$$

По лемме 1.1 эти знаки сохранаются и в некоторой достаточно малой окрестности x_0

$$\mathcal{V}(f_{j-1}(x_0 - \Delta x), f_j(x_0 - \Delta x), f_{j+1}(x_0 - \Delta x)) =$$

$$= \mathcal{V}(f_{j-1}(x_0+\Delta x), f_j(x_0+\Delta x), f_{j+1}(x_0+\Delta x)) = 1$$

при $\Delta x < \delta$. Но тогда можно утверждать, что число $\mathcal{V}(f_{j-1}(x),\ f_j(x),\ f_{j+1}(x))$ не меняется при прохождении аргументом интервала $]x_0-\delta\ ,\ x_0+\delta[$ ($\operatorname{sign} f_j(x)$ не влияет на число знакоперемен). Можно выбрать δ настолько малым, чтобы ни один другой полином системы (1.1) не обратился в нуль в указанном интервале (на основании леммы 1.1). Тогда можно утверждать, что и число (1.2) не изменится при прохождении интервала.

II. Случай $j = K$ невозможен на основании свойства Б.

III. Пусть $j = 0$, т.е. x_0 совпадает с каким-то корнем λ полинома $f(x)$. По свойству А $f_1(x_0) \neq 0$, а по свойству Г произведение $f_0(x)f_1(x)$ меняет знак с $-$ на $+$ когда x, возрастая, проходит x_0:

$$f_0(x_0-\Delta x)f_1(x_0-\Delta x) < 0, \ \text{т.е. эти числа разных знаков,}$$

$$f_0(x_0+\Delta x)f_1(x_0+\Delta x) > 0, \ \text{т.е. эти числа одного знака}$$

для $\Delta\, x < \delta$ при достаточно малых δ. Следовательно,

$$\mathcal{V}(f_0(x_0-\Delta x),\ f_1(x_0-\Delta x)) = 1, \ \text{ а } \ \mathcal{V}(f_0(x_0+\Delta x),\ f_1(x_0+\Delta x)) = 0.$$

Можно выбрать δ настолько малым, чтобы ни один другой полином системы (1.1) не обратился в нуль в указанном интервале. Тогда число (1.2) должно уменьшаться на единицу при возрастании x от $x_0-\Delta x$ до $x_0+\Delta x$.

Осталось рассмотреть последний случай, отсеченный нами в самом начале доказательства. Предположим, что $f_j(x_0) = 0$ и $f_\ell(x_0) = 0$ при $j < \ell$. По свойству А эти полиномы не могут быть соседними: $\ell \neq j+1$. Тогда для тройки полиномов $f_{\ell-1}(x), f_\ell(x), f_{\ell+1}(x)$ повторяются рассуждения пункта **I** доказательства. Изменение знаков этих полиномов при прохождении x_0 не повлияет на число (1.2).

Обобщая все собранные факты, приходим к выводу: число (1.2) будет меняться, уменьшаясь на 1, только при прохождении аргументом корней полинома $f(x)$. Мы получили “счетчик” корней полинома на интервале $]a,b[$: разность между величинами $\mathcal{V}_a$ и $\mathcal{V}_b$ даст нам точное значение числа корней полинома на этом интервале. ■

Построение системы полиномов Штурма с помощью алгоритма Евклида

Лемма 1.2. *Пусть на интервале* $]a,b[$ *полином* $f(x)$ *имеет кратный корень четной кратности. Тогда построение* **с.п.Шт.** *невозможно.*

Доказательство. Пусть $f(x)=(x-\lambda)^{2m}\varphi(x)$, и $\varphi(\lambda)\ne 0$. Имеем $\operatorname{sign} f(x)=\operatorname{sign}\varphi(x)=\operatorname{sign}\varphi(\lambda)$ в достаточно малой окрестности λ. Но тогда при произвольном выборе полинома $f_1(x)$ такого, чтобы $f_1(\lambda)\ne 0$, произведение $f(x)f_1(x)$ будет сохранять знак в этой окрестности, что противоречит свойству Г) для **с.п.Шт.** ■

Предположим теперь, что $f(x)$ не имеет вовсе кратных корней. Это равносильно тому, что НОД$(f(x),f'(x)) = const \ne 0$. Установить этот факт можно по алгоритму Евклида нахождения НОД[3]. Оказывается, что в качестве полиномов системы Штурма (1.1) можно взять последовательность остатков из алгоритма Евклида, если только домножить некоторые из них на -1. Именно, возьмем

$$f_1(x)\stackrel{\mathrm{def}}{=} f'(x).$$

Поделим $f_0(x)\equiv f(x)$ на $f_1(x)$ и обозначим через $f_2(x)$ остаток, домноженный на -1:

$$f_0(x)\equiv q_1(x)f_1(x)-f_2(x),\quad \deg f_2<n-1.$$

Поделим $f_1(x)$ на $f_2(x)$ и обозначим через $f_3(x)$ остаток, домноженный на -1:

$$f_1(x)\equiv q_2(x)f_2(x)-f_3(x),\quad \deg f_3<\deg f_2.$$

Продолжаем алгоритм далее, в конце концов дойдем до последнего ненулевого остатка $f_K(x)$, который совпадает с НОД$(f(x),f'(x))$. По предположению, этот последний $f_K(x)\equiv const\ne 0$.

Теорема 1.2. *Последовательность полиномов, построенная с помощью алгоритма Евклида будет* **с.п.Шт.** *для* $f(x)$ *на произвольном интервале* $]a,b[\subset\mathbb{R}$.

[3] Глава 3, §4.

Доказательство. **1.** Проверим выполнение свойства Г). Пусть $f(x) = (x - \lambda)\varphi(x)$ при $\varphi(\lambda) \neq 0$. Тогда $f'(x) = \varphi(x) + (x - \lambda)\varphi'(x)$ и

$$f_0(x)f_1(x) = (x-\lambda)\varphi(x)^2 + (x-\lambda)^2\varphi(x)\varphi'(x) = \\ = (x-\lambda)\underbrace{\left[\varphi(x)^2 + (x-\lambda)\varphi(x)\varphi'(x)\right]}_{\stackrel{\text{def}}{=}\Phi(x)}.$$

Очевидно $\Phi(\lambda) = \varphi(\lambda)^2 > 0$. По лемме 1.1, полином $\Phi(x)$ положителен в достаточно малой окрестности λ. В этой же окрестности будет выполнено

$$\operatorname{sign}\left[f_0(x)f_1(x)\right] = \operatorname{sign}\left[(x-\lambda)\Phi(x)\right] = \operatorname{sign}(x-\lambda) = \begin{cases} -1 & \text{если } x<\lambda; \\ +1 & \text{если } x>\lambda. \end{cases}$$

2. Свойство Б) выполняется поскольку последний полином $f_K(x)$, построенный по алгоритму Евклида, будет константой, отличной от 0.

3. Установим, наконец, справедливость свойств А) и В). Тройка полиномов $f_{j-1}(x), f_j(x), f_{j+1}(x)$ связана соотношением:

$$f_{j-1}(x) = q_j(x)f_j(x) - f_{j+1}(x).$$

Если $f_j(x_0) = 0$ то $f_{j-1}(x_0) = -f_{j+1}(x_0)$, что подтверждает свойство В, если только не выполняется $f_{j-1}(x_0) = f_{j+1}(x_0) = 0$. Но в последнем случае и все последующие остатки в алгоритме Евклида обратятся в нуль при $x = x_0$. В частности, $f_K(x_0) = 0$, но, по доказанному ранее, $f_K(x) \equiv const \neq 0$. ∎

Пример 1.1. Отделить корни полинома $f(x) = x^4 - x - 1$.

Решение. $f_1 = f'(x) = 4x^3 - 1$.

$$\begin{array}{r|l} x^4 + 0x^3 + 0x^2 \quad -x - \quad 1 & 4x^3 - 1 \\ \cline{2-2} x^4 + \qquad\qquad -{}^1\!/\!_4 x \quad & {}^1\!/\!_4 x \\ \hline -{}^3\!/\!_4 x \quad -1 & \end{array}$$

Перейти к оглавлению на странице: 460

$f_2(x) = {}^3/_4\, x + 1\,.$

$$
\begin{array}{rrrr|l}
4x^3+ & 0x^2 & +0x & -1 & {}^3/_4\, x + 1 \\ \cline{5-5}
4x^3+ & {}^{16}/_3\, x^2 & & & {}^{16}/_3\, x^2 - {}^{64}/_9\, x + {}^{256}/_{27} \\ \cline{1-4}
 & -{}^{16}/_3\, x^2 & & -1 & \\
 & -{}^{16}/_3\, x^2 & -{}^{64}/_9\, x & & \\ \cline{2-4}
 & & {}^{64}/_9\, x & -1 & \\
 & & {}^{64}/_9\, x & +{}^{256}/_{27} & \\ \cline{3-4}
 & & & -{}^{283}/_{27} &
\end{array}
$$

$f_3(x) = {}^{283}/_{27}\,.$

x	$f(x)$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$\mathcal{V}_x$	Комментарии
$-\infty$	$+$	$-$	$-$	$+$	2	сначала устанавливаем
$+\infty$	$+$	$+$	$+$	$+$	0	число вещественных корней,
0	$-$	$-$	$+$	$+$	1	затем положит. и отрицат.,
-1	$+$	$-$	$+$	$+$	2	затем просто дробим про-
1	$-$	$+$	$+$	$+$	1	межутки, отыскивая такие,
2	$+$	$+$	$+$	$+$	0	чтобы на каждом $\mathcal{V}_a - \mathcal{V}_b = 1$

Ответ. Полином $f(x)$ имеет два различных вещественных корня, один на интервале $]-1, 0[$, другой — на $]1, 2[$.

Упрощения, допустимые при построении с.п.Шт.:

1) любой полином можно домножать на положительную константу (это позволяет сделать коэффициенты полиномов возможно более простыми — например, целыми);

2) нет необходимости проделывать все деления до последнего, т.е. до тех пор, пока не получится функция $f_K(x) \equiv const$. Деление можно прекратить на том полиноме $f_j(x)$, про который известно, что он не имеет вещественных корней на $]a, b[$;

3) если $f(x)$ имеет кратные корни, то последняя функция, полученная применением алгоритма Евклида, т.е. $f_K(x)$ будет отличной от константы. Тогда разность $\mathcal{V}_a - \mathcal{V}_b$ дает число *различных* корней полинома на $]a, b[$. Каждый корень учитывается один раз (вне зависимости от его кратности).

Перейти к оглавлению на странице: 460

Пример 1.2. Отделить корни полинома $f(x) = x^5 + 2\,x^4 + x^3 - x^2 - 2\,x - 1$.

Решение. $f_1 = f'(x) = 5\,x^4 + 8\,x^3 + 3\,x^2 - 2\,x - 2$

$$
\begin{array}{rrrrrr|l}
x^5+ & 2\,x^4+ & x^3- & x^2- & 2\,x & -\,1 & 5x^4+8x^3+3x^2-2x-2 \\ \cline{7-7}
x^5+ & {}^{8}\!/_{5}\,x^4+ & {}^{3}\!/_{5}\,x^3- & {}^{2}\!/_{5}\,x^2- & {}^{2}\!/_{5}\,x & & {}^{1}\!/_{5}\,x + {}^{2}\!/_{25} \\ \cline{1-5}
 & {}^{2}\!/_{5}\,x^4+ & {}^{2}\!/_{5}\,x^3- & {}^{3}\!/_{5}\,x^2- & {}^{8}\!/_{5}\,x & -\,1 & \\
 & {}^{2}\!/_{5}\,x^4+ & {}^{16}\!/_{25}x^3+ & {}^{6}\!/_{25}x^2- & {}^{4}\!/_{25}x & -\,{}^{4}\!/_{25} & \\ \cline{2-6}
 & & -\,{}^{6}\!/_{25}x^3- & {}^{21}\!/_{25}x^2- & {}^{36}\!/_{25}x & -\,{}^{21}\!/_{25} &
\end{array}
$$

$f_2 = 2\,x^3 + 7\,x^2 + 12\,x + 7$;

$$
\begin{array}{rrrrr|l}
5\,x^4+ & 8\,x^3+ & 3\,x^2- & 2\,x & -\,2 & 2\,x^3 + 7\,x^2 + 12\,x + 7 \\ \cline{6-6}
5x^4+ & {}^{35}\!/_{2}\,x^3+ & 30\,x^2+ & {}^{35}\!/_{2}\,x & & {}^{5}\!/_{2}\,x - {}^{19}\!/_{4} \\ \cline{1-4}
 & -\,{}^{19}\!/_{2}\,x^3- & 27\,x^2- & {}^{39}\!/_{2}\,x & -\,2 & \\
 & -\,{}^{19}\!/_{2}\,x^3- & {}^{133}\!/_{4}\,x^2- & 57\,x & -\,{}^{133}\!/_{4} & \\ \cline{2-5}
 & & {}^{25}\!/_{4}\,x^2+ & {}^{75}\!/_{2}\,x & +\,{}^{125}\!/_{4} &
\end{array}
$$

$f_3 = -x^2 - 6\,x - 5$;

$$
\begin{array}{rrrr|l}
2\,x^3+ & 7\,x^2+ & 12\,x & +\,7 & -x^2 - 6\,x - 5 \\ \cline{5-5}
2\,x^3+ & 12\,x^2+ & 10\,x & & -2x + 5 \\ \cline{1-3}
 & -\,5\,x^2+ & 2\,x & +\,7 & \\
 & -\,5\,x^2- & 30\,x & -\,25 & \\ \cline{2-4}
 & & 32\,x & +\,32 &
\end{array}
$$

$f_4 = -x - 1$. Следующее деление приводит к нулевому остатку. Вывод: $\text{НОД}\,(f, f') = f_4$, и, следовательно, $x = -1$ является кратным корнем второй кратности для $f(x)$.

x	$f(x)$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$	$\mathcal{V}$
$-\infty$	$-$	$+$	$-$	$-$	$+$	3
$+\infty$	$+$	$+$	$+$	$-$	$-$	1
0	$-$	$-$	$+$	$-$	$-$	2
-2	$-$	$+$	$-$	$+$	$+$	3
2	$+$	$+$	$+$	$-$	$-$	1

Ответ. Полином $f(x)$ имеет два различных вещественных корня, один на интервале $]-2, 0[$, другой — на $]0, 2[$.

Перейти к оглавлению на странице: 460

Условие вещественности всех корней полинома

В некоторых прикладных задачах требуется установить условия, при которых полином $f(x) \in \mathbb{R}[x]$ имеет определенное число вещественных корней.

Теорема 1.3. *Полином $f(x) \in \mathbb{R}[x]$ имеет только вещественные и простые корни тогда и только тогда, когда в его системе полиномов Штурма (1.1) число $K = \deg f$ и все старшие коэффициенты этих полиномов имеют одинаковый знак.*

Доказательство. Для простоты всех корней $f(x)$ **Н.** и **Д.** чтобы $\mathsf{НОД}(f, f') \equiv const \ne 0$. При построении **с.п.Шт** по алгоритму Евклида степени полиномов $f_j(x)$ будут убывать до $\deg f_K(x) = 0$. Следовательно $K \le n$. Все корни $f(x)$ будут вещественными тогда и только тогда, когда

$$\operatorname{nrr}\{f(x) = 0 \mid -\infty < x < +\infty\} = n.$$

Но тогда из формулы (1.3) следует, что для этого **Н.** и **Д.** чтобы

$$\mathcal{V}(f_0(-\infty), f_1(-\infty), \dots, f_K(-\infty)) = n,$$
$$\mathcal{V}(f_0(+\infty), f_1(+\infty), \dots, f_K(+\infty)) = 0.$$

Первое из этих равенств влечет за собой то, что $K = n$ т.е. $\deg f_j(x) = \deg f_{j-1} - 1$; из второго же следует, что знаки полиномов $f_0(x), \dots, f_n(x)$ должны быть одинаковыми при $x = +\infty$. Но эти знаки определяются только старшими членами полиномов. ■

Пример 1.3. Установить условия вещественности всех корней полинома $x^3 + p\,x + q$.

Решение. Строим **с.п.Шт.** по алгоритму Евклида

$$f_1(x) = 3\,x^2 + p,\ \ f_2(x) = -{}^2\!/_3\,p\,x + q,$$
$$f_3(x) = -(4p^3 + 27q^2)/(4p^2) = \mathcal{D}(f)/(4p^2),$$

если $p \ne 0$. Здесь $\mathcal{D}(f)$ — дискриминант[4] полинома $f(x)$. При этом предположении для вещественности всех корней **Н.** и **Д.** чтобы

$$p < 0,\ \ \mathcal{D}(f) > 0\ ,$$

[4]Глава 6, §2.

и, можно проверить, что первое из этих условий вытекает из второго. Если же $p=0$, то **с.п.Шт.** превращается в

$$f_0(x)=x^3+q,\ f_1(x)=3\,x^2,\ f_2(x)=q,$$

и при любом значении $q\ne 0$ полином имеет лишь один вещественный корень. При $p=0, q=0$ полином x^3 имеет единственный вещественный корень третьей кратности.

Ответ. Полином имеет все корни вещественными тогда и только тогда, когда либо $4p^3+27q^2<0$, либо $p=0,\ q=0$.

Упражнение 80. *Найти все значения* p *и* q*, при которых полином* $x^4+p\,x+q$ *не имеет вещественных корней.*

§2. Ганкелевы матрицы в задаче локализации корней

В настоящем параграфе мы рассмотрим еще один подход к задаче локализации корней полинома, основанный на теории квадратичных форм.

Формулы Ньютона

Пусть $f(x)=a_0x^n+a_1x^{n-1}+\dots+a_n\in\mathbb{C}[x],\ a_0\ne 0$. Обозначим $\lambda_1,\dots,\lambda_n$ его (нам неизвестные) корни.

Определение 2.1. Выражение

$$s_k\stackrel{\rm def}{=}\lambda_1^k+\dots+\lambda_n^k$$

называется k-й **суммой Ньютона** полинома $f(x)$.

Задача. Найти выражение для s_k через коэффициенты $a_0,\dots,a_n$.

Согласно теореме 2.6 главы 3, эта задача разрешима в следующем смысле: существует некоторый полином $\mathfrak{F}(x_1,\dots,x_n)$ с целыми коэффициентами такой, что

$$\lambda_1^k+\dots+\lambda_n^k\equiv\mathfrak{F}(a_1/a_0,\dots,a_n/a_0).$$

Для малых показателей k функцию $\mathfrak{F}$ удается подобрать с помощью формул Виета.

Пример 2.1. Имеем:

$$s_0=\lambda_1^0+\dots+\lambda_n^0=n\ ,s_1=\lambda_1^1+\dots+\lambda_n^1=-a_1/a_0,$$

$$s_2=\lambda_1^2+\dots+\lambda_n^2=(\lambda_1+\dots+\lambda_n)^2-2\sum_{1\le j<k\le n}\lambda_j\lambda_k=$$

$$=s_1^2-2\,a_2/a_0=(a_1^2-2\,a_0a_2)/a_0^2.$$

Для вывода общей формулы воспользуемся следующим приемом, предположив, для упрощения рассуждений, что $f(x)$ не имеет кратных корней. Разложим дробь $f'(x)/f(x)$ на простейшие с помощью формулы Лагранжа[5]

$$\frac{f'(x)}{f(x)}=\sum_{j=1}^n\frac{f'(\lambda_j)}{f'(\lambda_j)(x-\lambda_j)}=\sum_{j=1}^n\frac{1}{x-\lambda_j}. \qquad (2.1)$$

Разложим теперь каждую дробь $1/(x-\lambda_j)$ в ряд по степеням ${}^1/_x$

$$\frac{1}{x-\lambda_j}=\frac{1}{x}\left(1+\frac{\lambda_j}{x}+\frac{\lambda_j^2}{x^2}+\dots\right)=\frac{1}{x}+\frac{\lambda_j}{x^2}+\dots+\frac{\lambda_j^k}{x^{k+1}}+\dots$$

и подставим в (2.1):

$$\begin{array}{rcccccc}
\dfrac{f'(x)}{f(x)} & = & \dfrac{1}{x} & +\dfrac{\lambda_1}{x^2} & +\dots & +\dfrac{\lambda_1^k}{x^{k+1}}+ & \dots+\\
 & + & \dfrac{1}{x} & +\dfrac{\lambda_2}{x^2} & +\dots & +\dfrac{\lambda_2^k}{x^{k+1}}+ & \dots+\\
 & + & & & & & \dots+\\
 & + & \dfrac{1}{x} & +\dfrac{\lambda_n}{x^2} & +\dots & +\dfrac{\lambda_n^k}{x^{k+1}}+ & \dots=\\
 & = & \dfrac{s_0}{x} & +\dfrac{s_1}{x^2} & +\dots & +\dfrac{s_k}{x^{k+1}}+ & \dots
\end{array}$$

Домножим обе части этого равенства на $f(x)$, получим тождество:

$$f'(x)\equiv f(x)\left[\frac{s_0}{x}+\frac{s_1}{x^2}+\dots+\frac{s_k}{x^{k+1}}+\dots\right].$$

[5] Глава 3, §10.

Сравнивая коэффициенты при одинаковых степенях x в этом тождестве получаем равенства:

$$\begin{array}{lllllll}
x^{n-1}: & n\ a_0 & = a_0\ s_0 & & & & \\
x^{n-2}: & (n-1)a_1 & = a_1\ s_0 & +a_0\ s_1 & & & \\
x^{n-3}: & (n-2)a_2 & = a_2\ s_0 & +a_1\ s_1 + a_0 s_2 & & & \\
\dots & & \dots & & & & \\
1\ : & a_{n-1} & = a_{n-1}s_0 & +a_{n-2}s_1 + \dots + a_0 s_{n-1} & & & \\
x^{-1}\ : & 0 & = a_n\ s_0 & +a_{n-1}s_1 + \dots + a_1 s_{n-1} + a_0 s_n & & & \\
x^{-2}\ : & 0 & = & a_n\ s_1 + \dots + a_2 s_{n-1} + a_1 s_n + a_0 s_{n+1} & & & \\
\dots & & \dots & & & &
\end{array}$$

Разрешая их последовательно, получаем рекурсивные **формулы Ньютона** для вычисления s_k:

$$s_0 = n,\ s_1 = -a_1/a_0, \dots,$$

$$s_k = \begin{cases} -(a_1 s_{k-1} + a_2 s_{k-2} + \dots + a_{k-1}s_1 + a_k k)/a_0 & \text{если } k < n; \\ -(a_1 s_{k-1} + a_2 s_{k-2} + \dots + a_n s_{k-n})/a_0 & \text{если } k \ge n. \end{cases} \tag{2.2}$$

Пример 2.2. Выписать выражения для s_3, s_4 и s_5 через коэффициенты полинома.

Решение. Используя результаты примера 2.1, имеем

$$s_3 = -(a_1 s_2 + a_2 s_1 + 3\,a_3)/a_0 =$$

$$= -\left(a_1(a_1^2 - 2\,a_0 a_2)/a_0^2 + a_2(-a_1/a_0) + 3\,a_3\right)/a_0 =$$

$$= \left(-a_1^3 + 3\,a_1 a_2 - 3\,a_3\right)/a_0^3\ ,$$

$$s_4 = -\left(a_1 s_3 + a_2 s_2 + a_3 s_1 + 4\,a_4\right)/a_0 =$$

$$= \left(a_1^4 - 4\,a_1^2 a_2 + 4\,a_1 a_3 - 4\,a_4 + 2\,a_2^2\right)/a_0^4\ ,$$

$$s_5 = \left(-a_1^5 + 5(a_1^3 a_2 + a_1 a_4 - a_1^2 a_3 - a_1 a_2^2 + a_2 a_3 - a_5)\right)/a_0^5.$$

Полученные формулы универсальны: они не зависят от степени полинома (при $j > n$ коэффициент a_j полагается равным нулю). $\triangle$

Упражнение 81. *Получить обращения формул Ньютона: по известным величинам $s_0, \dots, s_n$ восстановить коэффициенты полинома $f(x)$ (считая $a_0 = 1$).*

Упражнение 82. *Найти аналоги формул* (2.2) *для вычисления* s_k *при отрицательных индексах* k.

Теорема Якоби

Рассмотрим теперь полином

$$f(x) = a_0x^n + a_1x^{n-1} + \dots + a_n, \ (a_0 \ne 0)$$

с вещественными коэффициентами. Вычислим его суммы Ньютона $s_1, \dots, s_{2n-2}$ и составим ганкелеву матрицу

$$S = [s_{j+k}]_{j,k=0}^{n-1} . \tag{2.3}$$

Теорема 2.1. *[Якоби] Число различных корней* $f(x)$ *равно* $\operatorname{rank} S$, *а число различных вещественных корней* $f(x)$ *равно* $\sigma(S)$:

$$\operatorname{nrr}\{f(x) = 0\} = \sigma(S) = n_+(S) - n_-(S). \tag{2.4}$$

Доказательство. Рассмотрим квадратичную форму с матрицей S относительно вещественных переменных $x_0, x_1, \dots, x_{n-1}$:

$$X^{\top} S\, X = (x_0, x_1, \dots, x_{n-1}) \left(\begin{array}{llll} s_0 & s_1 & \dots & s_{n-1} \\ s_1 & s_2 & \dots & s_n \\ \dots & & & \dots \\ s_{n-1} & s_n & \dots & s_{2n-2} \end{array} \right) \left(\begin{array}{l} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{array} \right) =$$

$$= (x_0, x_1, \dots, x_{n-1}) \left(\begin{array}{llll} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \dots & & & \dots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \end{array} \right) \left(\begin{array}{llll} 1 & \lambda_1 & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{n-1} \\ \dots & & & \dots \\ 1 & \lambda_n & \dots & \lambda_n^{n-1} \end{array} \right) \left(\begin{array}{l} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{array} \right) =$$

$$= \sum_{j=1}^{n} \left[x_0 + x_1\lambda_j + x_2\lambda_j^2 + \dots + x_{n-1}\lambda_j^{n-1} \right]^2 .$$

Итак, мы фактически представили квадратичную форму в виде суммы квадратов. Число различных квадратов совпадает с числом различных λ_j: если

$$f(x) = a_0(x-\lambda_1)^{\mathfrak{m}_1}(x-\lambda_2)^{\mathfrak{m}_2} \times \dots \times (x-\lambda_p)^{\mathfrak{m}_p} \ , \text{ где } \lambda_j \ne \lambda_k$$

то

$$X^{\top} S\, X = \sum_{j=1}^{p} \mathfrak{m}_j \left[x_0 + x_1\lambda_j + x_2\lambda_j^2 + \dots + x_{n-1}\lambda_j^{n-1} \right]^2 = \sum_{j=1}^{p} \mathfrak{m}_j z_j^2, \tag{2.5}$$

где подстановка

$$z_j = \begin{cases} x_0 + x_1\lambda_j + x_2\lambda_j^2 + \dots + x_{n-1}\lambda_j^{n-1} & \text{если } j \in \{1,\dots,p\}; \\ x_{j-1} & \text{если } j \in \{p+1,\dots,n\}. \end{cases} \tag{2.6}$$

или в матричной форме

$$\left(\begin{array}{c} z_1 \\ z_2 \\ \vdots \\ z_p \\ z_{p+1} \\ \vdots \\ z_n \end{array} \right) = \left(\begin{array}{ccccccc} 1 & \lambda_1 & \dots & \lambda_1^{p-1} & \lambda_1^p & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{p-1} & \lambda_2^p & \dots & \lambda_2^{n-1} \\ \dots & & & & & \dots & \\ 1 & \lambda_p & \dots & \lambda_p^{p-1} & \lambda_p^p & \dots & \lambda_p^{n-1} \\ & & & & 1 & & \mathbb{O} \\ & & \mathbb{O} & & & \ddots & \\ & & & & & & 1 \end{array} \right) \left(\begin{array}{c} x_0 \\ x_1 \\ \vdots \\ x_{p-1} \\ x_p \\ \vdots \\ x_{n-1} \end{array} \right) \tag{2.7}$$

очевидно является неособенной (**почему?**). Таким образом, в правой части (2.5) стоит канонический вид квадратичной формы $X^{\top} S\, X$ и поэтому

$$\operatorname{rank} S = p = \text{ число различных корней } f(x).$$

Докажем теперь равенство (2.4). Оно очевидно в том случае, когда все $\lambda_1, \dots, \lambda_p$ вещественны, т.к. тогда вещественна и матрица (2.7) и новые переменные z_j. Пусть имеется корень $\lambda_j = \alpha + \boldsymbol{i}\ \beta$ при $\beta \neq 0$. Тогда, поскольку коэффициенты полинома вещественны, корнем будет и $\lambda_{j+1} = \overline{\lambda_j} = \alpha - \boldsymbol{i}\ \beta$ причем той же кратности, что и λ_j. Введем новые переменные u_j и u_{j+1} по формулам:

$$u_j = \frac{z_j + z_{j+1}}{2}\ , \quad u_{j+1} = \frac{z_j - z_{j+1}}{2\,\boldsymbol{i}}. \tag{2.8}$$

Оказывается, что переменные связаны с исходными $x_0, \dots, x_{n-1}$ линейными зависимостями с *вещественными* коэффициентами.

Действительно, если выделить в представлении (2.6) для z_j вещественные и мнимые части

$$z_j = x_0 + x_1\lambda_j + \cdots + x_{n-1}\lambda_j^{n-1} = X(x_0, \dots, x_{n-1}) + \boldsymbol{i}\ Y(x_0, \dots, x_{n-1}) ,$$

то

$$z_{j+1} = x_0 + x_1\lambda_{j+1} + \cdots + x_{n-1}\lambda_{j+1}^{n-1} =$$

$$= x_0 + x_1\overline{\lambda_j} + \cdots + x_{n-1}\overline{\lambda_j^{n-1}} = X(x_0, \dots, x_{n-1}) - \boldsymbol{i}\ Y(x_0, \dots, x_{n-1}) .$$

Из формул (2.8) следует, что

$$\begin{aligned} u_j = X(x_0, \dots, x_{n-1}) &\stackrel{\text{def}}{=} x_0 + x_1 \mathrm{Re}\, e(\lambda_j) + \cdots + x_{n-1}\mathrm{Re}\, e(\lambda_j^{n-1}) \\ u_{j+1} = Y(x_0, \dots, x_{n-1}) &\stackrel{\text{def}}{=} x_1 \mathrm{Im}\, m(\lambda_j) + \cdots + x_{n-1}\mathrm{Im}\, m(\lambda_j^{n-1}) \end{aligned}$$

имеют вещественные коэффициенты.

В матричной форме записи подстановка (2.8) эквивалентна домножению формул (2.7) слева на матрицу

$$\begin{array}{r} \\ \\ \\ \mathrm{j} \longrightarrow \\ \mathrm{j+1} \longrightarrow \\ \\ \\ \\ \end{array} \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & \mathbb{O} & & \\ & & 1 & & & & & \\ & & & 1/2 & 1/2 & & & \\ & & & 1/2i & -1/2i & & & \\ & & & & & 1 & & \\ & & \mathbb{O} & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}$$

Матрица подстановки остается неособенной, а ее результатом будет то, что пара слагаемых в каноническом виде (2.5) приведется к виду:

$$\mathfrak{m}_j z_j^2 + \mathfrak{m}_j z_{j+1}^2 = 2\mathfrak{m}_j u_j^2 - 2\mathfrak{m}_j u_{j+1}^2 .$$

Таким образом, каждая пара комплексно-сопряженных корней $f(x)$ порождает два вещественных квадрата в каноническом виде квадратичной формы $X^{\top} S\, X$: один с положительным коэффициентом, второй — с отрицательным. Тогда

$$\begin{array}{rcl} n_+(S) & = & \text{число вещественных } + \text{ число пар комплексно-} \\ & & \text{сопряженных корней,} \\ n_-(S) & = & \text{число пар комплексно-сопряженных корней.} \end{array}$$

Отсюда и следует формула (2.4). ■

Согласно результатам §3 главы 7, конструктивное вычисление индексов инерции симметричной матрицы S возможно посредством определения знаков ее главных миноров $S_1,\dots,S_n$.

Следствие 2.1. *Пусть*

$$S_n=0,\dots,S_{\mathfrak{r}+1}=0, S_{\mathfrak{r}}\ne 0,\dots,S_1\ne 0.$$

Тогда $\operatorname{rank} S=\mathfrak{r}$ *и*

$$\operatorname{nrr}\{f(x)=0\}=\mathcal{P}(1,S_1,\dots,S_{\mathfrak{r}})-\mathcal{V}(1,S_1,\dots,S_{\mathfrak{r}}).$$

Пример 2.3. Определить число вещественных корней полинома $x^5-3\,x^3-x-1$.

Решение. Суммы Ньютона:

$$\{s_j\}_{j=0}^{8}=\{5,\ 0,\ 6,\ 0,\ 22,\ 5,\ 72,\ 21,\ 238\}.$$

$$S=\left[\begin{array}{rrrrr} 5 & 0 & 6 & 0 & 22 \\ 0 & 6 & 0 & 22 & 5 \\ 6 & 0 & 22 & 5 & 72 \\ 0 & 22 & 5 & 72 & 21 \\ 22 & 5 & 72 & 21 & 238 \end{array}\right].$$

Главные миноры:

$$S_1=5,\ S_2=30,\ S_3=444,\ S_4=-4598,\ S_5=-56\,123.$$

Поскольку $S_5\ne 0$, все корни $f(x)$ различны.

$$\mathcal{P}(1,5,30,444,-4598,-56\,123)=4,\ \mathcal{V}(1,5,30,444,-4598,-56\,123)=1.$$

Ответ. Три вещественных корня.

Следствие 2.2. *Определитель матрицы S фактически совпадает с дискриминантом полинома $f(x)$:*

$$S_n = \det S = \mathcal{D}(f)/a_0^{2n-2}.$$

Это следует из формулы (2.2) главы **??** и формулы (9.2) из главы 4.

Следствие 2.3. *Пусть* $\operatorname{rank} S = \mathfrak{r}$. *Для вещественности всех корней* $f(x)$

а) *достаточно, чтобы* $S_1 > 0, \dots, S_{\mathfrak{r}} > 0$,

б) *и необходимо, чтобы* $S_1 \ge 0, \dots, S_{\mathfrak{r}} \ge 0$.

Отделение корней

Идею, использованную при доказательстве теоремы 2.1 можно развить и для задачи отделения корней $f(x)$. Для этого вычислим дополнительно еще одну сумму Ньютона s_{2n-1} и составим следующую ганкелеву матрицу, зависящую от параметра t:

$$H(t) = [s_{j+k}t - s_{j+k+1}]_{j,k=0}^{n-1} =$$

$$= \left[\begin{array}{llll} s_0t - s_1 & s_1t - s_2 & \dots & s_{n-1}t - s_n \\ s_1t - s_2 & s_2t - s_3 & \dots & s_nt - s_{n+1} \\ \dots & & & \dots \\ s_{n-1}t - s_n & s_nt - s_{n+1} & \dots & s_{2n-2}t - s_{2n-1} \end{array} \right]_{n\times n} . \qquad (2.9)$$

Легко видеть, что матрица, составленная из коэффициентов при t совпадает с (2.3). Обозначим $\ell-$й главный минор матрицы (2.9) через $H_{\ell}(t)$:

$$H_{\ell}(t) = \det [s_{j+k}t - s_{j+k+1}]_{j,k=0}^{\ell-1} .$$

Можно получить иное детерминантное представление для $H_{\ell}(t)$ — в виде определителя порядка $\ell+1$, в котором параметр t оказывается "выметенным" в последний столбец:

$$H_{\ell}(t) \equiv \left| \begin{array}{lllll} s_0 & s_1 & \dots & s_{\ell-1} & 1 \\ s_1 & s_2 & \dots & s_{\ell} & t \\ \vdots & & & \vdots & \vdots \\ s_{\ell} & s_{\ell+1} & \dots & s_{2\ell-1} & t^{\ell} \end{array} \right|_{(\ell+1)\times(\ell+1)} .$$

Разложение по этому столбцу позволит получить каноническое представление полинома $H_{\ell}(t)$.

Теорема 2.2. *[Йоахимшталь] Пусть* $\operatorname{rank} S = \mathfrak{r}$. *Тогда*

$$\operatorname{nrr}\{f(x)=0 \mid a<x<b\} =$$
$$= \mathcal{V}(1,H_1(a),\dots,H_{\mathfrak{r}}(a)) - \mathcal{V}(1,H_1(b),\dots,H_{\mathfrak{r}}(b)), \qquad (2.10)$$

при условии, что в ряду

$$1, H_1(t),\dots,H_{\mathfrak{r}}(t) \qquad (2.11)$$

нет двух последовательных нулей при $t=a$ *и* $t=b$.

Идея доказательства основывается на представимости квадратичной формы $X^{\top}H(t)X$ в виде

$$\sum_{j=1}^{n}(t-\lambda_j)\left[x_0+x_1\lambda_j+x_2\lambda_j^2+\dots+x_{n-1}\lambda_j^{n-1}\right]^2 .$$

Следствие 2.4. $H_n(t)=\det H(t)\equiv S_n f(t)/a_0$.

Таким образом система полиномов (2.11) играет роль системы полиномов Штурма для полинома $f(x)$. В отличие от классической **с.п.Шт.**, построенной в §1 с помощью алгоритма Евклида, здесь степени полиномов $H_\ell(t)$ возрастают при возрастании ℓ.

Пример 2.4. Локализовать вещественные корни полинома

$$f(x)=x^4-x^3-9\,x^2+14\,x-4.$$

Решение. Вычисляем суммы Ньютона:

$$\{s_k\}_{k=0}^7=\{4,\ 1,\ 19,\ -14,\ 159,\ -229,\ 1474,\ -2869\}$$

и строим последовательность миноров (2.11):

$$H_1(t)=4\,t-1,\ H_2(t)=75(t^2+t-5),$$
$$H_3(t)=1250(3\,t^3-26\,t+17),\ H_4(t)=62500\,f(t).$$

Вычисляем числа знакоперемен в нескольких узлах:

t	1	$H_1(t)$	$H_2(t)$	$H_3(t)$	$H_4(t)$	$\mathcal{V}_t$
$-\infty$	+	−	+	−	+	4
$+\infty$	+	+	+	+	+	0
0	+	−	−	+	−	3
−4	+	−	+	−	+	4
−3	+	−	+	+	−	3
1	+	+	−	−	+	2
2	+	+	+	−	−	1
3	+	+	+	+	+	0

Ответ. Полином $f(x)$ имеет 4 различных вещественных корня, лежащие на интервалах $]-4,-3[,\ ]0,1[,\ ]1,2[,\ ]2,3[$.

Упражнение 83. *Как теорема 2.2 связана с $?*

§3. Корни полинома в областях комплексной плоскости

Левая полуплоскость

Материал этого пункта изложен в §8 главы 6.

Единичный круг

Определение 3.1. **Единичным кругом** на комплексной плоскости назовем круг $|z| \le 1$.

Задача. Найти **Н.** и **Д.** условия на коэффициенты полинома $f(z) = a_0 z^n + \dots + a_n$, при которых все корни $f(z)$ находятся внутри единичного круга, т.е. удовлетворяют условию $|z| < 1$.

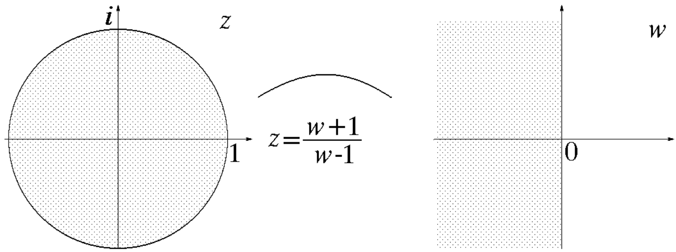
Решить эту задачу можно сведением ее к задаче установления критерия устойчивости некоторого вспомогательного полинома.

Лемма 3.1. *Замена переменной*

$$z = \frac{w+1}{w-1}$$

производит взаимно-однозначное отображение внутренности единичного круга плоскости z в левую полуплоскость плоскости w.

Перейти к оглавлению на странице: 460



Доказательство. Покажем, что каждому значению $z \neq 1$ соответствует единственное значение w. В самом деле,

$$z = 1+\frac{2}{w-1} \iff z-1 = \frac{2}{w-1} \iff w-1 = \frac{2}{z-1} \iff w = \frac{z+1}{z-1}.$$

Далее, покажем, что условие $|z| < 1$ эквивалентно $\operatorname{Re} e(w) < 0$.

$$|z| < 1 \iff \left|\frac{w+1}{w-1}\right| < 1 \iff |w+1| < |w-1| \iff$$

$$\iff \sqrt{(\operatorname{Re}(w)+1)^2+(\operatorname{Im} m(w))^2} < \sqrt{(\operatorname{Re}(w)-1)^2+(\operatorname{Im}(w))^2} \iff$$

$$\iff 4\operatorname{Re}(w) < 0.$$

■

Теорема 3.1. *Полином $f(z) = a_0 z^n + \cdots + a_n$ имеет все свои корни лежащими внутри единичного круга тогда и только тогда, когда полином*

$$F(w) \overset{\text{def}}{=} (w-1)^n f\left(\frac{w+1}{w-1}\right) =$$

$$= a_0(w+1)^n + a_1(w+1)^{n-1}(w-1) + \cdots + a_n(w-1)^n \qquad (3.1)$$

будет устойчив.

Доказательство тривиально ввиду леммы 1. ■

Пример 3.1. Определить все значения параметра $\alpha \in \mathbb{R}$, при которых полином

$$f(z) = 3\,z^3 + \alpha\,z^2 + z + 2$$

будет иметь все корни лежащими внутри единичного круга.

Решение. Строим полином (3.1):

$$F(w) = \underbrace{(6+\alpha)}_{A_0}w^3 + \underbrace{(2+\alpha)}_{A_1}w^2 + \underbrace{(14-\alpha)}_{A_2}w + \underbrace{2-\alpha}_{A_3}.$$

Теорема 7.5 (Льенара, Шипара) главы 6 дает условия устойчивости $F(w)$ в виде

$$A_0 > 0,\ A_1 > 0,\ A_2 > 0,\ A_3 > 0,\ A_1A_2 - A_0A_3 > 0;$$

и

$$A_0 < 0,\ A_1 < 0,\ A_2 < 0,\ A_3 < 0,\ A_1A_2 - A_0A_3 > 0.$$

Подставляя сюда выражения для коэффициентов, получим, что первая система ограничений имеет решение $-1 < \alpha < 2$, а вторая — несовместна.

Косвенной проверкой истинности полученного интервала могут служить его границы:

$$f(z) \equiv \left\{ \begin{array}{ll} (3z+2)(z^2-z+1) & \text{при } \alpha = -1; \\ (z+1)(3z^2-z+2) & \text{при } \ \alpha = 2. \end{array} \right.$$

В обоих случаях имеются корни, удовлетворяющие условию $|z| = 1$: в первом случае это будет комплексно-сопряженная пара ${}^1\!/_2 \pm \pm \boldsymbol{i}\ {}^{\sqrt{3}}\!/_2$, во втором — корень (-1).

Ответ. $-1 < \alpha < 2$.

Известен еще один результат, позволяющий решить поставленную задачу.

Теорема 3.2. *[Шур, Кон] Полином $f(z) = a_0 z^n + \dots + a_n$ имеет все свои корни лежащими внутри единичного круга тогда и только тогда, когда*

$$|\text{старший коэффициент } f(z)| > |\text{свободный член } f(z)|\ , \qquad (3.2)$$

т.е. $|a_0| > |a_n|$, *и полином*

$$f_1(z) \stackrel{\text{def}}{=} \frac{a_0 f(z) - a_n f^*(z)}{z} \tag{3.3}$$

имеет все свои корни лежащими внутри единичного круга. Здесь $f^*(z) \stackrel{\text{def}}{=} z^n f(1/z) = a_0 + a_1 z + \dots + a_n z^n$.

На первый взгляд, конструктивность этого результата не очень очевидна: исходная задача для полинома $f(z)$ сводится к аналогичной задаче для полинома $f_1(z)$. Обратим, однако, внимание на то, что полином

$$\begin{aligned} f_1(z) &= [a_0(a_0 z^n + \dots + a_n) - a_n(a_0 + a_1 z + \dots + a_n z^n)]\,/z = \\ &= \left[(a_0^2 - a_n^2)z^n + (a_0 a_1 - a_{n-1}a_n)z^{n-1} + \ldots + (a_0 a_{n-1} - a_1 a_n)z\right]/z = \\ &= (a_0^2 - a_n^2)z^{n-1} + (a_0 a_1 - a_{n-1}a_n)z^{n-2} + \dots + (a_0 a_{n-1} - a_1 a_n) \end{aligned}$$

имеет степень меньшую, чем $\deg f$. Таким образом, алгоритм конструктивен в том смысле, что он сводит исходную задачу к более простой. Применяя к полиному $f_1(z)$ снова критерий Шура–Кона, получим следующее необходимое условие

$$\begin{array}{rcl} |\text{старший коэффициент } f_1(z)| & > & |\text{свободный член } f_1(z)| \\ \Longleftrightarrow \qquad |a_0^2 - a_n^2| & > & |a_0 a_{n-1} - a_1 a_n|, \end{array} \tag{3.4}$$

при выполнении которого исследованию подлежит полином

$$f_2(z) \stackrel{\text{def}}{=} \frac{(a_0^2 - a_n^2) f_1(z) - (a_0 a_{n-1} - a_1 a_n) f_1^*(z)}{z}.$$

Продолжая процедуру, за конечное число шагов мы дойдем до полинома первой степени. Окончательно, необходимые и достаточные условия нахождения всех корней полинома $f(z)$ степени n внутри единичного круга получаются объединением n условий

$$\begin{array}{rcl} |\text{старший коэффициент } f(z)| & > & |\text{свободный член } f(z)|, \\ |\text{старший коэффициент } f_1(z)| & > & |\text{свободный член } f_1(z)|, \\ \dots & & \dots \\ |\text{старший коэффициент } f_{n-1}(z)| & > & |\text{свободный член } f_{n-1}(z)|. \end{array}$$

Пример 3.2. Определить все значения параметра $\alpha \in \mathbb{R}$, при которых полином

$$5\,z^4 + 4\,z^3 + \alpha\,z^2 + 2\,z + 1$$

будет иметь все корни лежащими внутри единичного круга.

Решение. Условие (3.2) выполнено.

$$\begin{array}{rcrrrrr|l} f(z) & = & 5\,z^4 & +4\,z^3 & +\alpha\,z^2 & +2\,z & +1 & \times 5 \\ f^*(z) & = & z^4 & +2\,z^3 & +\alpha\,z^2 & +4\,z & +5 & \times 1 \\ \hline & & 24\,z^4 & +18\,z^3 & +4\alpha\,z^2 & +6\,z & & \end{array}$$

Для $f_1(z) \stackrel{\text{def}}{=} 24\,z^3 + 18\,z^2 + 4\alpha\,z + 6$ условие (3.4) выполнено[6].

$$\begin{array}{rcrrrr|l} f_1(z) & = & 12\,z^3 & +9\,z^2 & +2\alpha\,z & +3 & \times 12 \\ f_1^*(z) & = & 3\,z^3 & +2\alpha\,z^2 & +9\,z & +12 & \times 3 \\ \hline & & 135\,z^3 & +(108-6\alpha)\,z^2 & +(24\alpha-27)\,z & & \end{array}$$

Для $f_2(z) \stackrel{\text{def}}{=} 135\,z^2 + (108-6\alpha)\,z + (24\alpha-27)$ необходимое условие расположения всех его корней внутри единичного круга выполнено при $|8\alpha - 9| < 45$.

$$\begin{array}{rrrr|l} f_2(z) = & 45z^2 & +(36-2\alpha)z & +(8\alpha-9) & \times 45 \\ f_2^*(z) = & (8\alpha-9)z^2 & +(36-2\alpha)z & +45 & \times(8\alpha-9) \\ \hline & [45^2-(8\alpha-9)^2]z^2 & +(36-2\alpha)(45-(8\alpha-9))z & & \end{array}$$

Единственный корень полинома

$$f_3(z) \stackrel{\text{def}}{=} [45^2 - (8\alpha-9)^2]\,z + (36-2\alpha)(45-(8\alpha-9))$$

удовлетворяет неравенству $|z| < 1$ тогда и только тогда, когда

$$|45 - (8\alpha-9)| \cdot |45 + (8\alpha-9)| > |45 - (8\alpha-9)| \cdot |36 - 2\alpha|.$$

Собираем полученные условия на α в систему неравенств:

$$|8\alpha - 9| < 45, \quad 2\,|9 + 2\alpha| > |18 - \alpha|,$$

[6] В дальнейшем для упрощения вычислений каждый полином $f_j(x)$ делим на положительный НОД его коэффициентов.

и решаем ее: $0 < \alpha < {}^{27}/_{4}$. Это условие является необходимым и достаточным для того чтобы все корни $f(z)$ удовлетворяли условию $|z| < 1$. На границах получившегося интервала лежат значения параметра, при которых полином $f(z)$ будет обладать корнями с модулями равными 1:

$$\begin{array}{lll} \text{при } \alpha=0: & f(z)\equiv(z+1)(5z^3-z^2+z+1) & \Rightarrow \lambda=-1 \\ \text{при } \alpha={}^{27}/_{4}: & f(z)\equiv{}^{1}/_{4}(2z^2+z+2)(10z^2+3z+2) & \Rightarrow \lambda=-{}^{1}/_{4}\pm \boldsymbol{i}\,{}^{\sqrt{15}}/_{4} \end{array}$$

Ответ. $0 < \alpha < {}^{27}/_{4}$.

Перейти к оглавлению на странице: 460

ГЛАВА 9. НЕКОТОРЫЕ АЛГЕБРАИЧЕСКИЕ СТРУКТУРЫ

§1. Группа

Бинарная операция

Что такое алгебраическая операция? Рассмотрим произвольное непустое множество $\mathbb{S}$, его элементы будем обозначать готическими буквами:

$$\mathbb{S} = \{\mathfrak{A}, \mathfrak{B}, \dots, \mathfrak{a}, \mathfrak{b}, \dots\}.$$

Определение 1.1. **Бинарной операцией**, определенной на $\mathbb{S}$, называется соответствие, при котором каждой упорядоченной паре $\mathfrak{a}$ и $\mathfrak{b}$ элементов из $\mathbb{S}$ отвечает определенный элемент $\mathfrak{c}$ того же множества. Записывать этот факт будем в виде

$$\mathfrak{a} * \mathfrak{b} = \mathfrak{c}.$$

Пример 1.1. Если $\mathbb{S} = \mathbb{Z}$, то *сложение* является бинарной операцией; если $\mathbb{S} = \mathbb{Q}_+$ (положительных рациональных чисел), то *деление* тоже является бинарной операцией.

Подчеркнем, что понятие бинарной операции неразрывно связано с множеством, на котором она определена: результат операции не должен выводить за пределы множества. Примеры показывают, что бинарная операция, определенная на множестве $\mathbb{S}$, не обязательно “наследуется” при переходе к подмножеству.

Пример 1.2. На подмножестве четных чисел множества $\mathbb{Z}$ операция сложения продолжает оставаться бинарной: сумма любых четных чисел остается четным числом. Напротив, на подмножестве нечетных чисел та же операция не является бинарной.

Определение 1.2. Говорят, что подмножество $\mathbb{S}_1$ множества $\mathbb{S}$ **замкнуто** относительно бинарной операции $*$, определенной на $\mathbb{S}$, если выполнено $\mathfrak{a} * \mathfrak{b} \in \mathbb{S}_1$ для любой пары элементов $\mathfrak{a}$ и $\mathfrak{b}$ из $\mathbb{S}_1$. Если это условие не выполняется хотя бы для одной пары элементов $\mathfrak{a}$ и $\mathfrak{b}$ из $\mathbb{S}_1$, то говорят, что $\mathbb{S}_1$ **не замкнуто** относительно $*$.

Упражнение 84. *Пусть* $\mathbb{S} \stackrel{\text{def}}{=} \mathbb{C}$, $\mathbb{S}_1 \stackrel{\text{def}}{=} \{1, -1, \boldsymbol{i}, -\boldsymbol{i}\}$. *Является ли* $\mathbb{S}_1$ *замкнутым относительно сложения? А умножения?*

Два произвольных элемента $\mathfrak{a}$ и $\mathfrak{b}$ множества $\mathbb{S}$, на котором определена бинарная операция $*$, определяют четыре возможные результата действия этой операции:

$$\mathfrak{a} * \mathfrak{b}, \ \mathfrak{a} * \mathfrak{a}, \ \mathfrak{b} * \mathfrak{a}, \ \mathfrak{b} * \mathfrak{b},$$

не все из которых обязательно различны.

Определение 1.3. Если $\mathfrak{a} * \mathfrak{b} = \mathfrak{b} * \mathfrak{a}$, то говорят, что элементы $\mathfrak{a}$ и $\mathfrak{b}$ **перестановочны** или, что они **коммутируют относительно** операции $*$, если же $\mathfrak{a} * \mathfrak{b} \neq \mathfrak{b} * \mathfrak{a}$, то эти элементы **не перестановочны** (**не коммутируют**) **относительно** $*$.

Пример 1.3. Приведем несколько случаев некоммутативности:

а) множество $\mathbb{Q}_+$ с операцией деления: ${}^1/_2 : {}^1/_3 \neq {}^1/_3 : {}^1/_2$;

б) множество квадратных матриц n-го порядка с операцией умножения:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 5 & 3 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 \\ 5 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Рассмотрим теперь последовательное применение операции $*$ к трем элементам $\mathfrak{a}$, $\mathfrak{b}$ и $\mathfrak{c}$ множества $\mathbb{S}$. Возможные варианты исчерпываются

$$(\mathfrak{a} * \mathfrak{b}) * \mathfrak{c} \quad \text{и} \quad \mathfrak{a} * (\mathfrak{b} * \mathfrak{c}).$$

Определение 1.4. Говорят, что для упорядоченной тройки элементов $\mathfrak{a}$, $\mathfrak{b}$ и $\mathfrak{c}$ выполняется свойство **ассоциативности** операции $*$ если $(\mathfrak{a} * \mathfrak{b}) * \mathfrak{c} = \mathfrak{a} * (\mathfrak{b} * \mathfrak{c})$.

Пример 1.4. Для чисел ${}^1/_2$, ${}^1/_3$, ${}^1/_4$ не выполняется свойство ассоциативности деления: $({}^1/_2 : {}^1/_3) : {}^1/_4 \neq {}^1/_2 : ({}^1/_3 : {}^1/_4)$.

Определение 1.5. Будем говорить, что операция $*$ **коммутативна** (**ассоциативна**) на множестве $\mathbb{S}$ если для любых двух (трех) элементов множества выполняется свойство коммутативности (ассоциативности).

Упражнение 85. *Пусть операция* $*$ *подчиняется двум законам:*

$$\mathfrak{a} * \mathfrak{a} = \mathfrak{a}, \quad (\mathfrak{a} * \mathfrak{b}) * \mathfrak{c} = (\mathfrak{b} * \mathfrak{c}) * \mathfrak{a}$$

для любых $\mathfrak{a}, \mathfrak{b}$ *и* $\mathfrak{c}$ *из* $\mathbb{S}$. *Доказать, что* $*$ *ассоциативна и коммутативна.*

Определение группы

Определение 1.6. Множество $\mathbb{S}$ с определенной на нем бинарной операцией $*$ называется **полугруппой** если $*$ — ассоциативна.

Определение 1.7. **Нейтральным** элементом множества $\mathbb{S}$ относительно операции $*$ называется такой элемент $\mathfrak{e}$ этого множества, что для любого элемента $\mathfrak{a} \in \mathbb{S}$ выполняются соотношения

$$\mathfrak{a} * \mathfrak{e} = \mathfrak{e} * \mathfrak{a} = \mathfrak{a}.$$

Пример 1.5. На множестве четных чисел не существует нейтрального элемента относительно умножения.

Определение 1.8. Пусть множество $\mathbb{S}$ содержит нейтральный элемент относительно операции $*$. Элемент $\mathfrak{A}$ называется **обратным** элементу $\mathfrak{a} \in \mathbb{S}$ относительно операции $*$ если выполняются соотношения

$$\mathfrak{a} * \mathfrak{A} = \mathfrak{A} * \mathfrak{a} = \mathfrak{e}.$$

Определение 1.9. Полугруппа называется **группой**, если в ней существует нейтральный элемент и для любого ее элемента существует обратный относительно $*$.

Обозначения и терминология. Традиционно группа обозначается буквой $\mathbb{G}$; операция $*$ называется **умножением**; результат бинарной операции $\mathfrak{a} * \mathfrak{b}$ записывается тогда как $\mathfrak{a} \cdot \mathfrak{b}$ или просто $\mathfrak{a}\mathfrak{b}$; нейтральный элемент $\mathfrak{e}$ называется **единичным** или просто **единицей**; элемент, обратный $\mathfrak{a}$ обозначается $\mathfrak{a}^{-1}$. Результат умножения элемента на самого себя называется его **степенью**:

$$\underbrace{\mathfrak{a} * \mathfrak{a} * \cdots * \mathfrak{a}}_{k} \stackrel{\text{def}}{=} \mathfrak{a}^k, \ \mathfrak{a}^0 \stackrel{\text{def}}{=} \mathfrak{e}.$$

Перейти к оглавлению на странице: 460

Примеры групп

Аддитивная группа целых чисел. Множество $\mathbb{Z}$ является группой относительно операции сложения, единичным элементом которой является 0, а элемент $(-\mathfrak{a})$ будет обратным элементу $\mathfrak{a}$. То же самое множество не является группой относительно умножения: хотя единичный элемент и существует, но обратного не существует ни для какого другого. *Является ли это множество полугруппой?*

Классы вычетов. Рассмотрим полную систему классов вычетов[1] по модулю M, т.е. множеств целых чисел имеющих одинаковый остаток r при делении на M:

$$\overline{r} = \left\{ r + Mt \;\middle|\; t \in \mathbb{Z} \right\}$$

при $r \in \{0, 1, \dots, M-1\}$. Это множество — обозначаемое $\mathbb{Z}_M$ — является группой относительно сложения классов:

$$\overline{a} + \overline{b} = \overline{c} \quad \text{где } c = a + b \pmod{M}.$$

Единичным элементом является класс $\overline{0}$ (множество чисел, делящихся на M нацело); элементом, обратным $\overline{a}$, очевидно является $\overline{M-a}$.

То же множество $\mathbb{Z}_M$ относительно умножения классов является полугруппой но не группой: для класса $\overline{0}$ не существует обратного. Но даже если выбросить этот класс и рассмотреть множество $\mathbb{Z}_M \setminus \overline{0} = \{\overline{1}, \dots, \overline{M-1}\}$ то и оно не для всяких M будет группой : хотя единичный элемент и существует, но нет, например, класса, обратного $\overline{2}$ по модулю 6. Тем не менее, если дополнительно потребовать, чтобы модуль M был числом *простым*: $M = p$, то множество $\{\overline{1}, \dots, \overline{p-1}\}$ становится группой относительно умножения. Элементом, обратным $\overline{a}$ будет класс $\overline{x}$, где x является решением сравнения $ax \equiv 1 \pmod{p}$. Это решение единственно среди чисел множества $\{1, \dots, p-1\}$ и может быть найдено по алгоритмам §6.1 главы 1. Например, используя теорему Ферма[2], его можно представить в виде $x = a^{p-2} \pmod{p}$.

[1] Глава 1, §5.2.

[2] Глава 1, теорема 5.6.

Корни из единицы. Во множестве $\mathbb{C}$ рассмотрим подмножество корней n-й степени из единицы:

$$\mathbb{G} = \left\{ \varepsilon_k = \cos\frac{2\pi k}{n} + \boldsymbol{i}\ \sin\frac{2\pi k}{n} \right\}_{k=0}^{n-1}.$$

Это подмножество является группой относительно умножения:

$$\mathfrak{e} = \varepsilon_0 = 1;\ \varepsilon_j\varepsilon_k = \varepsilon_{j+k \pmod n} = \begin{cases} \varepsilon_{j+k} \in \mathbb{G} & \text{если } j+k<n, \\ \varepsilon_{j+k-n} \in \mathbb{G} & \text{если } j+k \ge n; \end{cases}$$

$$\varepsilon_j^{-1} = \varepsilon_{-j} = \varepsilon_{n-j} \in \mathbb{G}.$$

Полиномы. Множества полиномов $\mathbb{Z}[x], \mathbb{Q}[x], \mathbb{R}[x], \mathbb{C}[x]$ с коэффициентами из соответственно $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ или $\mathbb{C}$ являются группами относительно сложения. Единичным элементом является тождественно нулевой полином, а полином $(-f(x))$ — обратным полиному $f(x)$. Те же множества не являются группами относительно умножения.

Полная линейная группа. Множество квадратных $n \times n$ матриц над $\mathbb{R}$ является полугруппой, но не группой относительно операции умножения, т.к. не для всякой матрицы A существует обратная. Но вот подмножество невырожденных матриц образует группу, называемую **полной линейной группой степени n над $\mathbb{R}$** и обозначаемую[3] $\mathbf{GL}(n, \mathbb{R})$.

Определение 1.10. Группа называется **абелевой**[4] если операция $*$ коммутативна. В противном случае группа **неабелева**.

Группы из первых четырех примеров настоящего пункта являются абелевыми, но $\mathbf{GL}(n, \mathbb{R})$ — неабелева.

Замечание 1.1. В литературе иногда операцию $*$ в абелевой группе обозначают $+$; тогда нейтральный элемент $\mathfrak{e}$ обозначают 0; а $\mathfrak{a}^{-1}$ — через $-a$ (и называют элементом, противоположным $\mathfrak{a}$).

[3] ganze lineare (*нем.*) — полная линейная

[4] **Абель Нильс Хенрик** (Abel Niels Henrik, 1802–1829) — норвежский математик. Доказал (теорема 3.1 главы 3) неразрешимость в радикалах общего алгебраического уравнения пятой степени. При изучении алгебраических уравнений широко использовал понятие коммутативной группы. Создал теорию эллиптических функций.

Перейти к оглавлению на странице: 460

Монотонные функции. Рассмотрим множество функций $\{\varphi(x)\}$, непрерывных, (строго) возрастающих на $[0,1]$ и таких, что $\varphi(0)=0, \varphi(1)=1$. В качестве операции $*$ возьмем суперпозицию функций, т.е. подстановку одной функции в другую :

$$\varphi(x) * \psi(x) \stackrel{\text{def}}{=} \varphi\left(\psi(x)\right).$$

Множество с введенной операцией образует группу: ассоциативность имеется; $\mathfrak{e}=x$; $\varphi^{-1}(x)$ — функция обратная к $\varphi(x)$ (существует и принадлежит рассматриваемому множеству на основании известных результатов из мат.анализа). Группа неабелева, т.к., например, для

$$\varphi=x^2,\ \psi=\sin\frac{\pi x}{2} \quad \text{имеем } \varphi*\psi=\left(\sin\frac{\pi x}{2}\right)^2,\ \text{а } \psi*\varphi=\sin\frac{\pi x^2}{2}.$$

Определение 1.11. Полугруппа (группа) называется **конечной** если она состоит из конечного числа элементов; это число называется тогда **порядком** полугруппы (группы): $\mathbf{Card}(\mathbb{G})$ или $|\mathbb{G}|$ или $\#\mathbb{G}$.

Отображения плоских фигур. Рассмотрим группу, элементами которой являются отображения плоскости: именно, вращения плоских фигур на углы, кратные $\varphi>0$, вокруг начала координат. Эти отображения можно задать в матричной форме: если (x,y) — координаты точки до поворота, а (X,Y) — ее координаты после поворота на угол $k\,\varphi$ $(k\in\mathbb{Z})$, то связь между ними дается формулой

$$\left(\begin{array}{c} X \\ Y \end{array}\right)=P_k\left(\begin{array}{c} x \\ y \end{array}\right) \quad \text{при} \quad P_k \stackrel{\text{def}}{=} \left(\begin{array}{cc} \cos k\varphi & -\sin k\varphi \\ \sin k\varphi & \cos k\varphi \end{array}\right).$$

Бинарной операцией над элементами этого множества возьмем комбинацию поворотов. Легко видеть, что эта операция коммутативна: результат поворота на угол $k\varphi$, а затем — на угол $\ell\varphi$ будет таким же и при изменении последовательности поворотов и совпадать с поворотом на угол $(k+\ell)\varphi$. Аналитика подтверждает геометрию:

$$P_kP_\ell=P_\ell P_k=P_{k+\ell}.$$

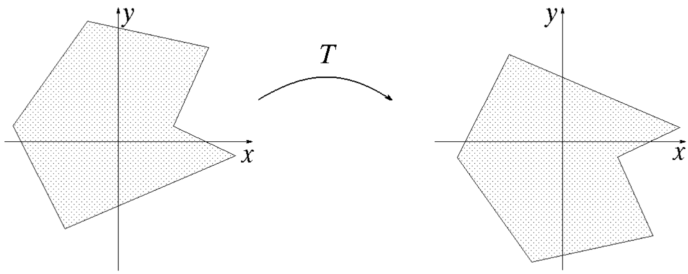
Далее, за единичный элемент относительно комбинации поворотов возьмем поворот, при котором все точки остаются на месте: $P_0=E$.

Перейти к оглавлению на странице: 460

Наконец, поворотом, обратным повороту на угол $k\varphi$, будет поворот на угол $(-k\varphi)$: $P_k^{-1} = P_{-k}$. Таким образом, все свойства группы выполнены, и наше множество является группой, причем абелевой.

Будет ли эта группа конечной? Ответ на этот вопрос зависит от угла φ. Так, если $\varphi = 2\pi/m$, $m \in \mathbb{Z}$, то $P_m = E = P_0, P_{m+1} = P_1, \ldots, P_{m+k} = P_k$, при этом матрицы $P_0, P_1, \ldots, P_{m-1}$ все различны. Следовательно, $\mathbf{Card}(\mathbb{G}) = m$. Если $\varphi = 2\pi p/q$ и дробь p/q несократима, то $\mathbf{Card}(\mathbb{G}) = q$. Наконец, при $\varphi = 2\pi\alpha$ и α — иррациональном имеем: $P_k \neq P_\ell$ ни при каких целых индексах $k, \ell, k \neq \ell$. В этом случае группа бесконечна.

Определим теперь на той же самой плоскости еще одно отображение: зеркальное отражение. Плоская фигура отображается в зеркально симметричную ей относительно абсолютно плоского зеркала, проходящего через ось абсцисс и перпендикулярного плоскости:



Это отображение также можно задать в матричной форме:

$$\begin{pmatrix} X \\ Y \end{pmatrix} = T \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{при } T \stackrel{\text{def}}{=} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Очевидно, что два последовательных отражения возвращают фигуру в исходное состояние; действительно:

$$T^2 = E.$$

Таким образом, множество из двух отображений — тождественного и зеркального — образует группу второго порядка, причем абелеву.

Попробуем теперь составить объединение множеств рассмотренных выше отображений, одновременно допустив к рассмотрению и повороты и отражения, и их всевозможные комбинации. Для простоты, рассмотрим повороты на угол $\varphi = 2\pi/3$. Комбинируя теперь все возможные последовательные отображения, приходим ко множеству:

$$\Big\{\textit{тождественное},\ \textit{поворот на угол}\ 2\pi/3,\ \textit{поворот на угол}\ 4\pi/3,$$

$$\textit{отражение},\ \textit{поворот на угол}\ 2\pi/3\ \textit{и отражение},$$

$$\textit{отражение и поворот на угол}\ 2\pi/3,$$

$$\textit{поворот на угол}\ 4\pi/3\ \textit{и отражение},$$

$$\textit{отражение и поворот на угол}\ 4\pi/3, \dots\Big\}$$

Проанализируем полученное с помощью матричного аппарата; соответствующая последовательность матриц имеет вид:

$$E,\ P_1 = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix},\ P_2 = \begin{pmatrix} -1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & -1/2 \end{pmatrix},\ T,$$

$$TP_1 = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{pmatrix},\ P_1T = \begin{pmatrix} -1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix},$$

$$TP_2 = \begin{pmatrix} -1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix},\ P_2T = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{pmatrix},\ \dots$$

Видим, что имеется всего 6 различных отображений; кроме того $P_1T \ne TP_1$. Итак, получившаяся группа порядка 6 является неабелевой.

Образующие элементы группы

Пусть $\mathfrak{a}$ и $\mathfrak{b}$ — элементы группы $\mathbb{G}$. Тогда, по определению группы, $\mathfrak{a}^{-1}$ и $\mathfrak{b}^{-1}$ также должны быть элементами $\mathbb{G}$ наряду с $\mathfrak{a}\mathfrak{b}^{-1}\mathfrak{a}$, $\mathfrak{a}\mathfrak{b}\mathfrak{a}^{-1}\mathfrak{b}$ и т.д. Любое произведение, которое можно записать, используя в качестве сомножителей $\mathfrak{a}, \mathfrak{b}, \mathfrak{a}^{-1}, \mathfrak{b}^{-1}$ в любом порядке и в любом конечном числе, является элементом $\mathbb{G}$. Если все элементы группы можно записать в виде произведений, включающих лишь $\mathfrak{a}$ и $\mathfrak{b}$ (и их обратные), то мы назовем $\mathfrak{a}$ и $\mathfrak{b}$ **образующими** (или **образующими элементами**) группы $\mathbb{G}$. Это понятие обобщается очевидным образом.

Определение 1.12. Если все элементы группы $\mathbb{G}$ могут быть выражены в виде произведений элементов из некоторого множества $\mathbb{S}$ (и их обратных), то элементы множества $\mathbb{S}$ называются **образующими** группы $\mathbb{G}$, или говорят, что они **порождают** группу $\mathbb{G}$.

Вернемся к примерам предыдущего пункта. Аддитивная группа целых чисел имеет единственную образующую, именно 1. Для группы корней n-й степени из 1 образующим можно выбрать ε_1, поскольку $\varepsilon_k = \varepsilon_1^k$ для любого $k \in \{0, 1, \dots, n-1\}$. С тем же успехом можно было бы взять и некоторые другие корни — но не любые! Так, для $n = 6$, корни ε_2, ε_3 или ε_4 образующими группы не будут:

$$\varepsilon_2^2 = \varepsilon_4,\ \varepsilon_2^3 = 1 = \varepsilon_0, \dots$$

Образующим можно взять любой <u>первообразный</u> корень ε_k, т.е. тот, для которого $\mathsf{НОД}\,(k, n) = 1$.

Группа отображений плоской фигуры — если допускаются к рассмотрению и повороты и отражения — имеет две образующие: поворот на угол φ и отражение.

Простейший случай — это группа с одной образующей.

Определение 1.13. Группа $\mathbb{G}$ называется **циклической** если она порождается единственным своим элементом. Если этот элементом является $\mathfrak{a}$, то циклическую группу обозначают $\langle \mathfrak{a} \rangle$:

$$\langle \mathfrak{a} \rangle \stackrel{\text{def}}{=} \{\mathfrak{a}^k\}_{k \in \mathbb{Z}}.$$

Упражнение 86. *Доказать, что циклическая группа всегда абелева.*

Вычисляя положительные степени элемента $\mathfrak{a}$ мы можем никогда не встретить повторений: все элементы бесконечной последовательности

$$\mathfrak{a},\, \mathfrak{a}^2, \dots, \mathfrak{a}^k, \dots$$

будут различными. В этом случае говорят, что элемент $\mathfrak{a}$ имеет **бесконечный порядок**. Если же $\mathfrak{a}^k = \mathfrak{a}^\ell$ при некоторой паре показателей k и ℓ, $k < \ell$, то, домножая обе части равенства на $\mathfrak{a}^{-k}$, получаем: $\mathfrak{a}^{\ell-k} = \mathfrak{e}$. Наименьшее целое число n, для которого $\mathfrak{a}^n = \mathfrak{e}$ называется **порядком** элемента $\mathfrak{a}$.

Упражнение 87. *Рассматривается мультипликативная группа классов вычетов по модулю* 17*. Найти порядок классов* $\overline{2}, \overline{4}, \overline{7}$*.*

Подгруппа

Определение 1.14. Подмножество $\mathbb{H}$ группы $\mathbb{G}$ называется ее **подгруппой**, если оно само является группой (относительно бинарной операции $*$). Иначе говоря, $\mathfrak{e} \in \mathbb{H}$, и при произвольных элементах $\mathfrak{a}$ и $\mathfrak{b}$ из $\mathbb{H}$ должны выполняться условия $\mathfrak{a} * \mathfrak{b} \in \mathbb{H}$, $\mathfrak{a}^{-1} \in \mathbb{H}$, $\mathfrak{b}^{-1} \in \mathbb{H}$.

Любая группа $\mathbb{G}$ имеет очевидные подгруппы: саму себя и подгруппу, состоящую из единственного элемента $\mathfrak{e}$. Эти две подгруппы называются **несобственными**. Нас интересует случай существования **собственных** подгрупп.

Такие подгруппы очевидно имеет аддитивная группа целых чисел: например, множество четных чисел является замкнутым относительно операции сложения. Группа корней n-й степени из единицы при $n = 6$ имеет собственные подгруппы:

$$\mathbb{H}_1 = \{1,\, \varepsilon_2,\, \varepsilon_4\} \quad \text{и} \quad \mathbb{H}_2 = \{1,\, \varepsilon_3\},$$

а при $n = 7$ их не имеет. В самом деле, если предположить, что подгруппа $\mathbb{H}$ содержит элемент ε_k при $k \in \{1, \dots, 6\}$, то, по свойству замкнутости относительно умножения, она должна содержать и любую степень этого элемента: ε_k^1, $\varepsilon_k^2 = \varepsilon_{2k \pmod 7}$, $\varepsilon_k^3 = \varepsilon_{3k \pmod 7}, \dots$ Можно доказать, что первые 6 членов этой последовательности будут различными. Но тогда они обязаны пробегать все множество корней седьмой степени из 1.

Упражнение 88. *Показать, что множество* $\mathbf{SL}(n, \mathbb{R})$ *матриц порядка* n *с определителями, равными* 1 *образует подгруппу* $\mathbf{GL}(n, \mathbb{R})$ *(называется* **специальной линейной группой степени** n **над** $\mathbb{R}$*). Будет ли* $\mathbf{SL}(n, \mathbb{R})$ *абелевой?*

Упражнение 89. *Образует ли множество матриц вида*

$$\left(\begin{array}{cc} a & b \\ b & a \end{array} \right) \quad \text{при } a^2 \neq b^2$$

подгруппу группы $\mathbf{GL}(n, \mathbb{R})$*? Будет ли эта подгруппа абелевой?*

Обратим внимание, что во всех приведенных примерах подгруппы строились по одному приему: брался произвольный элемент $\mathfrak{a}$ группы $\mathbb{G}$ и генерировались его степени (положительные, отрицательные и нулевая). Таким образом, эти подгруппы оказывались *циклическими*, порожденными элементом $\mathfrak{a}$. Разумеется, подгруппа может порождаться и не одним образующим.

Упражнение 90. *Доказать, что если* $\mathbb{H}_1$ *и* $\mathbb{H}_2$ *— подгруппы группы* $\mathbb{G}$, *то и* $\mathbb{H}_1 \bigcap \mathbb{H}_2$ *— подгруппа группы* $\mathbb{G}$.

Предположим теперь, что группа $\mathbb{G}$ — конечная, тогда и любая ее подгруппа тоже должна быть конечной. *Каким может тогда быть порядок подгруппы?*

Теорема 1.1. *[Лагранж] Порядок конечной группы кратен порядку любой из ее подгрупп.*

Для доказательства теоремы введем новое понятие.

Определение 1.15. Пусть $\mathbb{H}$ — собственная подгруппа группы $\mathbb{G}$ и пусть $\mathfrak{b}$ — элемент группы $\mathbb{G}$. Рассмотрим множество всех элементов вида $\mathfrak{b}\mathfrak{h}$ при $\mathfrak{h} \in \mathbb{H}$. Это множество называется **левым смежным классом** группы $\mathbb{G}$ по подгруппе $\mathbb{H}$ и обозначается $\mathfrak{b}\mathbb{H}$:

$$\mathfrak{b}\mathbb{H} \stackrel{\text{def}}{=} \left\{ \mathfrak{b}\mathfrak{h} \middle| \mathfrak{h} \in \mathbb{H} \right\}.$$

Аналогично определяется **правый** смежный класс $\mathbb{H}\mathfrak{b}$.

Замечание 1.2. В литературе встречаются и противоположные определения: когда левый смежный класс (в смысле только что приведенного определения) называют правым и наоборот.

Пример 1.6. Рассмотрим аддитивную группу целых чисел $\mathbb{G} = \mathbb{Z}$, в качестве ее подгруппы возьмем множество чисел кратных числу $M \in \mathbb{N}$: $\mathbb{H} = \left\{ tM \middle| t \in \mathbb{Z} \right\}$. Тогда при b не кратном M левый смежный класс

$$b + \mathbb{H} = \left\{ b + tM \middle| t \in \mathbb{Z} \right\}$$

очевидно совпадает с правым смежным классом, и представляет собой класс вычетов $\overline{b}$ по модулю M.

Доказательство теоремы Лагранжа. Сначала покажем, что все элементы смежного класса $\mathfrak{b}\mathbb{H}$ различны. Предположим, что $\mathfrak{b}\mathfrak{h}_1 = \mathfrak{b}\mathfrak{h}_2$ при $\{\mathfrak{h}_1, \mathfrak{h}_2\} \subset \mathbb{H}$. Тогда, домножая это равенство *слева* на

$\mathfrak{b}^{-1}$, необходимо приходим к $\mathfrak{h}_1 = \mathfrak{h}_2$. Следовательно, при $\mathfrak{h}_1 \neq \mathfrak{h}_2$ должно выполняться и $\mathfrak{b}\mathfrak{h}_1 \neq \mathfrak{b}\mathfrak{h}_2$. Итак, все элементы класса $\mathfrak{b}\mathbb{H}$ действительно различны; как следствие получаем, что $\mathbf{Card}(\mathbb{H}) = \mathbf{Card}(\mathfrak{b}\mathbb{H})$.

Теперь покажем, что множества $\mathbb{H}$ и $\mathfrak{b}\mathbb{H}$ не содержат общих элементов: $\mathbb{H} \bigcap \mathfrak{b}\mathbb{H} = \varnothing$. Допустим, что общие элементы имеются: $\mathfrak{h}_j = \mathfrak{b}\mathfrak{h}_k$. Тогда, домножая это равенство *справа* на $\mathfrak{h}_k^{-1}$, получаем $\mathfrak{b} = \mathfrak{h}_j\mathfrak{h}_k^{-1}$, и, поскольку $\mathfrak{h}_j\mathfrak{h}_k^{-1}$ принадлежит $\mathbb{H}$, то и $\mathfrak{b} \in \mathbb{H}$. Однако это противоречит предположению.

Итак, если $\mathbb{H}$ — несобственная подгруппа группы $\mathbb{G}$, то доказательство теоремы тривиально. Если же она собственная, то существует элемент $\mathfrak{b}$ группы $\mathbb{G}$, не входящий в $\mathbb{H}$. Тогда смежный класс $\mathfrak{b}\mathbb{H}$ содержит ровно $\mathbf{Card}(\mathbb{H})$ различных элементов, ни один из которых не содержится в $\mathbb{H}$. Следовательно множество $\mathbb{H} \bigcup \mathfrak{b}\mathbb{H}$ содержит ровно $2\mathbf{Card}(\mathbb{H})$ различных элементов. Если $\mathbb{H} \bigcup \mathfrak{b}\mathbb{H} = \mathbb{G}$, то теорема доказана. Если же существует элемент $\mathfrak{f}$ группы $\mathbb{G}$, не принадлежащий $\mathbb{H} \bigcup \mathfrak{b}\mathbb{H}$, то мы образовываем новый смежный класс $\mathfrak{f}\mathbb{H}$, который — по доказанному выше — имеет все элементы различными и не входящими в $\mathbb{H}$. Покажем, что $\mathfrak{f}\mathbb{H} \bigcap \mathfrak{b}\mathbb{H} = \varnothing$. В самом деле, если $\mathfrak{f}\mathfrak{h}_j = \mathfrak{b}\mathfrak{h}_k$, то $\mathfrak{f} = \mathfrak{b}\mathfrak{h}_k\mathfrak{h}_j^{-1} \in \mathfrak{b}\mathbb{H}$, что противоречит предположению о том, что $\mathfrak{f} \notin \mathfrak{b}\mathbb{H}$.

Следовательно, множество $\mathbb{H} \bigcup \mathfrak{b}\mathbb{H} \bigcup \mathfrak{f}\mathbb{H}$ содержит ровно $3\mathbf{Card}(\mathbb{H})$ различных элементов. Если это множество совпадает с $\mathbb{G}$, то теорема доказана. В противном случае, продолжаем процедуру выделения новых смежных классов. Эта процедура конечна, поскольку сама группа $\mathbb{G}$ конечна. Но тогда использование индукции позволит утверждать, что группа $\mathbb{G}$ раскладывается в объединение конечного числа попарно непересекающихся множеств: смежных классов по подгруппе $\mathbb{H}$. Каждое из подмножеств имеет мощность, равную $\mathbf{Card}(\mathbb{H})$. ■

Следствие 1.1. *Если* $\mathbf{Card}(\mathbb{G})$ *— простое число, то группа* $\mathbb{G}$ *не имеет собственных подгрупп и является циклической.*

Упражнение 91. *Пусть* p *— простое число, а число* $k \in \mathbb{N}$ *— минимальное, удовлетворяющее условию* $A^k \equiv 1 \pmod{p}$. *Доказать, что* k *является делителем* $(p-1)$.

Изоморфизм групп

В предыдущих пунктах мы видели, что группы, образованные элементами различной природы часто имели сходные свойства: такими, например были аддитивная группа классов вычетов по модулю n, т.е. $\mathbb{Z}_n$, и мультипликативная группа комплексных корней n-й степени из 1. Формализовать подобное сходство можно с помощью установления соответствия между элементами этих групп: чтобы описание свойств одной группы свободно переводилось в описание свойств другой.

Определение 1.16. Рассмотрим две группы: $\mathbb{G}$ с операцией $*$ и $\widetilde{\mathbb{G}}$ с операцией $\circledast$. Отображение $F : \mathbb{G} \mapsto \widetilde{\mathbb{G}}$, отображающее каждый элемент группы $\mathbb{G}$ на элемент группы $\widetilde{\mathbb{G}}$, называется **гомоморфизмом**, если

$$F(\mathfrak{a} * \mathfrak{b}) = F(\mathfrak{a}) \circledast F(\mathfrak{b}) \quad \text{д } \forall \mathfrak{a} \in \mathbb{G}, \forall \mathfrak{b} \in \mathbb{G}.$$

Если, вдобавок, гомоморфизм задает взаимно-одозначное соответствие между $\mathbb{G}$ и $\widetilde{\mathbb{G}}$, то он называется **изоморфизмом**. В случае существования изоморфизма между группами, сами группы называются **изоморфными** или **абстрактно равными**. Легко показать, что отношение изоморфизма групп является отношением эквивалентности.

Пример 1.7. Рассмотрим аддитивную группу вещественных чисел и мультипликативную группу положительных вещественных чисел. Эти группы изоморфны, поскольку функция $\log x$ (здесь логарифм рассматривается по любому основанию бо́льшему 1) взаимно-однозначно отображает множество $\mathbb{R}_+$ на $\mathbb{R}$, и, вдобавок, $\log(xy) = \log x + \log y$.

Упражнение 92. *Может ли группа быть изоморфна собственной подгруппе?*

Упражнение 93. *Доказать, что любые две циклические группы одинакового порядка (конечного или бесконечного) изоморфны.*

Можно доказать, что существует лишь конечное число абстрактно различных групп порядка n. В самом деле, с точностью до обозначения элементов для множества, состоящего из n различных символов, существует лишь конечное число таблиц умножения, имеющих n^2 клеток. Так, например, при $n = 6$ существуют всего *две* абстрактно различные группы: циклическая, изоморфная группе корней 6-й степени из единицы, и неабелева группа, изоморфная группе отображений плоских фигур (см. с. 443). Известно, что существует 267 абстрактно различных групп порядка 64.

§2. Кольцо, поле, алгебра

В предыдущем параграфе мы имели дело со множествами, в каждом из которых была определена единственная операция. Но те объекты, с которыми нам довелось встречаться в курсе, часто допускали еще одну (или несколько) бинарных операций. Оформим это наблюдение новым определением.

Определение 2.1. Пусть $\mathbb{K}$ — некоторое множество с двумя определенными в нем операциями $\circ$ и $*$. Это множество называется **кольцом** относительно указанных операций если выполнено

I. $\circ$ — коммутативна: $\mathfrak{a} \circ \mathfrak{b} = \mathfrak{b} \circ \mathfrak{a}$;

II. $\circ$ — ассоциативна: $(\mathfrak{a} \circ \mathfrak{b}) \circ \mathfrak{c} = \mathfrak{a} \circ (\mathfrak{b} \circ \mathfrak{c})$;

III. $\circ$ — обратима, т.е. для любых двух элементов $\mathfrak{a}$ и $\mathfrak{b}$ из $\mathbb{K}$ существует решение уравнения $\mathfrak{a} \circ x = \mathfrak{b}$;

IV. операции подчиняются дистрибутивному (распределительному) закону:

$$\mathfrak{a} * (\mathfrak{b} \circ \mathfrak{c}) = \mathfrak{a} * \mathfrak{b} \circ \mathfrak{a} * \mathfrak{c} \quad \text{и} \quad (\mathfrak{b} \circ \mathfrak{c}) * \mathfrak{a} = \mathfrak{b} * \mathfrak{a} \circ \mathfrak{c} * \mathfrak{a}.$$

Первые три условия определяют кольцо $\mathbb{K}$ как абелеву группу относительно операции $\circ$.

Замечание 2.1. Традиционно принято называть операцию $\circ$ **суммой**, а $*$ — **произведением**; нейтральный элемент относительно суммы называют **нулем**: $\mathfrak{o} + \mathfrak{a} = \mathfrak{a}$.

На произведение не накладывается практически никаких ограничений. Однако в приложениях, как правило, возникают кольца,

Перейти к оглавлению на странице: 460

в которых умножение может удовлетворять одному или нескольким из следующих условий (справедливых для любых элементов $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}$):

V. $(\mathfrak{a} * \mathfrak{b}) * \mathfrak{c} = \mathfrak{a} * (\mathfrak{b} * \mathfrak{c})$ (ассоциативное кольцо);

VI. $\mathfrak{a} * \mathfrak{b} = \mathfrak{b} * \mathfrak{a}$ (коммутативное кольцо);

VII. существование **единицы**, т.е. элемента $\mathfrak{e}$ такого, что $\mathfrak{a} * \mathfrak{e} = \mathfrak{e} * \mathfrak{a} = \mathfrak{a}$ (кольцо с единицей);

VIII. существование обратного элемента относительно умножения для любого элемента $\mathfrak{a} \neq \mathfrak{o}$: $\mathfrak{a} * \mathfrak{a}^{-1} = \mathfrak{e}$.

Пример 2.1. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ и $\mathbb{C}$ — коммутативные и ассоциативные кольца относительно сложения и умножения. $\mathbb{N}$ не является кольцом, т.к. не существует натурального решения уравнения $3+x=2$.

Пример 2.2. $\mathbb{Z}[x], \mathbb{Q}[x], \mathbb{R}[x]$ и $\mathbb{C}[x]$ (и вообще $\mathbb{Z}[x_1, \ldots, x_n]$, $\mathbb{Q}[x_1, \ldots, x_n]$, $\mathbb{R}[x_1, \ldots, x_n]$, $\mathbb{C}[x_1, \ldots, x_n]$) — коммутативные и ассоциативные кольца относительно сложения и умножения.

Пример 2.3. Множество квадратных $n \times n$ матриц с элементами из $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ или $\mathbb{C}$ — некоммутативное и ассоциативное кольцо относительно сложения и умножения.

Пример 2.4. $\mathbb{Z}$ — кольцо с единицей, а его подмножество четных чисел — кольцо без единицы.

Определение 2.2. **Полем** называется коммутативное и ассоциативное кольцо $\mathbb{P}$ с единицей, в котором выполняется условие VIII. Иначе говоря, в $\mathbb{P}$ уравнение $\mathfrak{a} * x = \mathfrak{b}$ разрешимо при любых $\mathfrak{a}$ и $\mathfrak{b}$ из $\mathbb{P}$, $\mathfrak{a} \neq \mathfrak{o}$.

Пример 2.5. $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ — поля, а $\mathbb{Z}$ — не поле. Множества дробно-рациональных функций над каждым из этих множеств — поля. Множество квадратных $n \times n$ матриц не является полем, т.к. не при всех A и B уравнение $AX = B$ разрешимо.

Определение 2.3. **Алгеброй** $\mathbb{A}$ над полем $\mathbb{P}$ называется такое кольцо, в котором определено умножение $\star$ элементов на элементы из $\mathbb{P}$, удовлетворяющее естественным аксиомам:

1) $(\mathbf{A} \circ \mathbf{B}) \star \mathfrak{a} = \mathbf{A} \star \mathfrak{a} \circ \mathbf{B} \star \mathfrak{a}, \quad \mathbf{A} \star \mathfrak{e} = \mathbf{A}$;

2) $(\mathbf{A} * \mathbf{B}) \star \mathfrak{a} = (\mathbf{A} \star \mathfrak{a}) * \mathbf{B} = \mathbf{A} * (\mathbf{B} \star \mathfrak{a})$

при любых $\mathbf{A}$ и $\mathbf{B}$ из $\mathbb{A}$, $\mathfrak{a}$ из $\mathbb{P}$.

Первым примером алгебр над полем $\mathbb{R}$ явились **гиперкомплексные** числа. В 1843 Уильям Гамильтон году придумал теорию **кватернионов**, т.е. чисел, являющихся n-мерными (при $n > 2$) аналогами комплексных чисел. Произвольный кватернион записывается в виде линейной комбинации

$$X = x_0 1 + x_1 \boldsymbol{i} + x_2 \boldsymbol{j} + x_3 \boldsymbol{k}, \quad \{x_0, x_1, x_2, x_3\} \subset \mathbb{R}$$

элементов базиса $1, \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$. Здесь 1 — обычная единица, а умножение остальных элементов базиса задается таблицей

	$\boldsymbol{i}$	$\boldsymbol{j}$	$\boldsymbol{k}$
$\boldsymbol{i}$	-1	$\boldsymbol{k}$	$-\boldsymbol{j}$
$\boldsymbol{j}$	$-\boldsymbol{k}$	-1	$\boldsymbol{i}$
$\boldsymbol{k}$	$\boldsymbol{j}$	$-\boldsymbol{i}$	-1

Часто в записи кватерниона 1 опускается:

$$X = x_0 + \underbrace{x_1 \boldsymbol{i} + x_2 \boldsymbol{j} + x_3 \boldsymbol{k}}_{\mathbf{V}}$$

здесь x_0 называется скалярной, а $\mathbf{V}$ — векторной частями кватерниона; при $x_0 = 0$ кватернион называется вектором и может быть отождествлен с обычным вектором из $\mathbb{R}^3$. Произведение двух таких векторов $\mathbf{V}_1$ и $\mathbf{V}_2$ (в соответствии с приведенной выше таблицей) выражается через скалярное и векторное произведения векторов:

$$\mathbf{V}_1 \cdot \mathbf{V}_2 = -(\mathbf{V}_1, \mathbf{V}_2) + [\mathbf{V}_1, \mathbf{V}_2]$$

что показывает тесную связь кватернионов с векторным исчислением (последнее и возникло из теории кватернионов).

Кватернионы обладают всеми свойствами поля, кроме коммутативности умножения, и включают в себя комплексные числа.

*　　*　　*

Перейти к оглавлению на странице: 460

Усвоение преподанного зависит от привычек слушателя; какие у нас сложились привычки, такого изложения мы и требуем, и то, что говорят против обыкновения, кажется неподходящим, и из-за непривычности — более непонятным и чуждым, ибо привычное более понятно. А какую силу имеет привычное, показывают законы, в которых то, что выражено в форме мифов и по-детски просто, благодаря привычке имеет бóльшую силу, нежели знание самих законов. Одни не воспринимают преподанного, если излагают математически, другие — если не приводят примеров, третьи требуют, чтобы приводилось свидетельство поэта. И одни хотят, чтобы все излагалось точно, а других точность тяготит или потому, что они не в состоянии связать [одно с другим], или потому, что считают точность мелочностью. В самом деле, есть у точности что-то такое, из-за чего она как в делах, так и в рассуждениях некоторым кажется низменной. Поэтому надо приучиться к тому, как воспринимать каждый предмет, ибо нелепо в одно и то же время искать и знание и способ его усвоения. Между тем нелегко достигнуть даже одного из них.

А математической точности нужно требовать не для всех предметов, а лишь для нематериальных. Вот почему этот способ не подходит для рассуждающего о природе, ибо вся природа, можно сказать, материальна ...

Аристотель “Метафизика”.

Перейти к оглавлению на странице: 460

Подсказки и ответы к упражнениям

Глава 4

2.11

$$\begin{pmatrix} 11 & -1 & -15 \\ -20 & 11 & 20 \\ 0 & -13 & -17 \end{pmatrix}$$

9.2 $(x_1 - x_4)^4(x_1 - x_3)^4(x_1 - x_2)^4(x_3 - x_4)(x_2 - x_4)(x_2 - x_3)$.
9.3 а) $-\,{}^{3}/_{5}x^2 + {}^{17}/_{12}x + {}^{11}/_{60}$; **б)** ${}^{6}/_{35} + {}^{48}/_{35}x - {}^{4}/_{7}x^2$
10.2 $(E + B)(B - E)^{-1} = E + 2(B - E)^{-1}$
10.4

$$\begin{pmatrix} 0&0&1&0&0&1&0&1 \\ 1&0&0&1&0&0&1&0 \\ 0&1&0&0&1&0&0&1 \\ 1&0&1&0&0&1&0&0 \\ 0&1&0&1&0&0&1&0 \\ 0&0&1&0&1&0&0&1 \\ 1&0&0&1&0&1&0&0 \\ 0&1&0&0&1&0&1&0 \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \pmod 2$$

11.1 а) 3 **б** 5
11.2 а) $\operatorname{rank} = 1$ при $\lambda = 1$, и $\operatorname{rank} = 4$ при $\lambda \neq 1$.
б) $\operatorname{rank} = 3$ при $\lambda \in \{\,{}^{1}/_{4}, 1\}$, и $\operatorname{rank} = 4$ при $\lambda \notin \{\,{}^{1}/_{4}, 1\}$.
12.2 Ф.с.р. состоит, например, из решений

а) $x_1{=}0, x_2{=}1, x_3{=}3, x_4{=}0, x_5{=}0$ и $x_1{=}0, x_2{=}0, x_3{=}2, x_4{=}0, x_5{=}1$;

б) $x_1{=}x_2{=}x_3{=}x_4{=}1, x_5{=}x_6{=}0$, $x_1{=}-1, x_2{=}x_3{=}x_4{=}0, x_5{=}1, x_6{=}0$ и

$$x_1 = 0, x_2 = -1, x_3 = x_4 = x_5 = 0, x_6 = 1.$$

Глава 5

1.2 а) $1 + {}^{2}/_{5}(x-1) - {}^{1}/_{105}(x-1)(4x-9) + {}^{1}/_{945}(x-1)(4x-9)(x-4)$, $f(2) = {}^{1334}/_{945} \approx 1.41164$;
б) $1 + x/1! + x(x-1)/2! + \cdots + x(x-1)\ldots(x-n+1)/n!$

1.3 $f(x) = {}^{1}/_{7}(31\,x^2 - 14.4)$.
2.1 $x = {}^{71261}/_{8669} \approx 8.22$, $y = {}^{40298}/_{8669} \approx 4.65$. К этой точке ближе всего расположена зарубка III.
2.5 $x_1 = {}^{5}/_{7}, x_2 = {}^{6}/_{7}$
2.6 а) $x_1{=}x_2{=}0$; **б)** $x_1{=}x_2{=}x_3{=}x_4{=}{}^{3}/_{4}$; **в)** $x_1{=}-{}^{1}/_{75}, x_2{=}-{}^{2}/_{75}$.

Глава 6

1.4 а) -26, **б)** -256. **1.5 а)** -2, **б)** $\{0, 13\}$.
1.6 а) $-6x+11,\ 6x^2-5x+25$; **б)** $16x^2-96x+144,\ -16x^3+48x^2+64$;
в) $0,\ 0$
1.7 а) $15\,x^2 - 35\,x + 21,\ -15\,x^4 - 10\,x^3 - 6\,x^2 - 3\,x - 1$;
б) ${}^{41}/_{8}\,x^2 - {}^{57}/_{4}\,x + {}^{65}/_{8},\ -{}^{41}/_{8}\,x^2 - {}^{91}/_{8}\,x - {}^{69}/_{8}$
3.3 а) $p \in \{0, 1, {}^{1}/_{2} \pm {}^{\sqrt{3}}/_{2}\,\boldsymbol{i}\}$; **б)** $p = 0$ или $3125\,q + 256\,p^5 = 0$;
в) $p = -2, q \in \{-1, {}^{11}/_{16}\}$ или $p = 1 \pm \sqrt{3}\boldsymbol{i}, q \in \{{}^{1}/_{2}(1 \pm \pm\sqrt{3}\boldsymbol{i}), -{}^{11}/_{32}(1 \pm \sqrt{3}\boldsymbol{i})\}$;
г) $p = -12, q = 9$.
4.2 а) ${}^{1}/_{3}\,x^2 + {}^{1}/_{3}\,x + {}^{2}/_{3}$ **б)** ${}^{9}/_{91}\,x^2 - {}^{34}/_{91}\,x + {}^{36}/_{91}$ при $x \neq -2$.
5.1 а) $(3, 0)$; $(2, -1)$; $(0, 1)$; $(2, 2)$. **б)** $(-1, 1)$; $(2, 2)$; $(1, -1)$ (кратное)
в) $(0, 3)$; $(0, 1)$; $(2, 1 \pm \sqrt{2}\boldsymbol{i})$; $(-1, 2)$; $(-1, 3)$
6.1 См. пример 1.2.
8.1 а) При $a_0 > 0$: $a_1 > 0, a_2 > 0, a_3 > 0,\ a_1a_2 - a_0a_3 > 0$.
б) При $a_0 > 0$: $a_1 > 0, a_2 > 0, a_3 > 0, a_4 > 0,\ a_1a_2a_3 - a_1^2a_4 - -a_0a_3^2 > 0$.

Глава 7

2.1 $\sum\limits_{i=1}^{n} y_i^2$, где $y_i = x_i + x_{i+1} + \dots + x_n$.
2.2 $y_1^2 - 2\,y_2^2 - {}^{3}/_{2}\,y_3^2 - {}^{4}/_{3}\,y_4^2 - \dots - {}^{n}/(n-1)\,y_n^2$.
3.1 C_{n+2}^2. **3.2** f_2 и f_3.

Перейти к оглавлению на странице: 460

ЛИТЕРАТУРА

1. *Гантмахер Ф.Р.* Теория матриц. М.: Наука. $\leq$ 1975
2. *Окунев Л.Я.* Высшая алгебра. М.: Учпедгиз. $\leq$ 1958
3. *Фаддеев Д.К.* Лекции по алгебре. М.: Наука. 1984

ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

Глава 5

4. *Беклемишев Д.В.* Дополнительные главы линейной алгебры. Гл. VI. М.: Наука. 1983
5. *Крылов А.Н.* Лекции о приближенных вычислениях. Гл. VIII. М.: ГИТТЛ. 1954
6. *Линник Ю.В.* Метод наименьших квадратов и основы математико–статистической теории обработки наблюдений. М.: ГИФМЛ. 1958

Глава 6

7. *Калинина Е.А., Утешев А.Ю.* Теория исключения. СПб. НИИ Химии СПбГУ. 2002
8. *Сушкевич А.К.* Основы высшей алгебры. М.-Л.ОНТИ. 1937

Глава 9

9. *Гроссман И., Магнус В.* Группы и их графы.М.Мир. 1971.

ИСТОРИЧЕСКИЕ СПРАВКИ

составлены на основании следующих источников

10. *История математики с древнейших времен до начала XIX столетия. Т.3. Под ред. Юшкевича А.П. М. Наука. 1972.*
11. *Белхост Б.* Огюстен Коши. М.Наука. 1997
12. *Бюлер В.* Гаусс. М.Наука. 1989
13. *Вендровский К.В.* Фотограф из Зазеркалья.// Химия и жизнь. 1983, №8, сс. 62–67
14. *Данилов Ю.А.* Льюис Кэрролл как нелинейное явление.// Известия вузов. Прикладная нелинейная динамика. 1996, т.4, №1, сс. 119–125

15. *Крылов А.Н.* Воспоминания и очерки. М. АН СССР. 1956
16. *Фаццари Г.* Краткая история математики. М.: Колос. 1923
17. *O'Connor J.J., Robertson E.F.* The MacTutor History of Mathematics archive
http://www-history.mcs.st-andrews.ac.uk/history/index.html

Предметный указатель

ОГЛАВЛЕНИЕ

Перейти к оглавлению на странице: 460