Глава I. Матрицы и определители

§ 5. Приложения определителей

М.В.Волков

Уральский федеральный университет
Институт естественных наук и математики
кафедра алгебры и фундаментальной информатики

2021/2022 учебный год

Определение

Пусть A – квадратная матрица. Матрица B называется *обратной к* A, если $AB = BA = E$, где E – единичная матрица. Если матрица, обратная к A, существует, то матрица A называется *обратимой*.

Лемма (единственность обратной матрицы)

Если матрица A обратима, обратная к ней матрица определяется однозначно.

Доказательство. Пусть B и C – матрицы, обратные к A. Тогда

$$B = BE = B(AC) = (BA)C = EC = C. \qquad \square$$

Лемма оправдывает обозначение A^{-1} для матрицы, обратной к A.

Замечание

Доказательство дает немного больше, чем утверждается в лемме!
А именно, мы доказали, что если B – *левая обратная* к A (т.е. $BA = E$), а C – *правая обратная* к A (т.е. $AC = E$), то $B = C$.

Вспомним: систему n линейных уравнений с n неизвестными

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dotfill \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

можно записать в матричном виде

$$AX = B. \qquad (*)$$

Если матрица A обратима, умножив $(*)$ на A^{-1}, получим $X = A^{-1}B$. С другой стороны, подставив $X = A^{-1}B$ в $(*)$, получим $A(A^{-1}B) = B$.

Итак, *система линейных уравнений с обратимой матрицей имеет единственное решение*, которое можно найти, зная обратную матрицу.

Возникает два естественных вопроса:

1. Как узнать, имеет ли данная квадратная матрица обратную?
2. Если матрица A имеет обратную, то как вычислить A^{-1}?

Мы доказали теорему об определителе произведения матриц: *если* A *и* B – $n \times n$*-матрицы, то* $\det AB = \det A \cdot \det B$.

Следствие

Если матрица A обратима, то ее определитель отличен от 0.

Доказательство. Если матрица B обратна к матрице A, то $AB = E$. «Детерминируя» это равенство, т.е. записывая равенство определителей его левой и правой частей, получим $\det AB = \det E = 1$. По теореме $\det AB = \det A \cdot \det B$, откуда $\det A \cdot \det B = 1$, что влечет $\det A \neq 0$. □

Докажем, что верно и обратное: если $\det A \neq 0$, то матрица A обратима. Для этого понадобится следующая лемма:

Лемма «свой среди чужих»

Сумма произведений элементов некоторой строки матрицы на алгебраические дополнения соответствующих элементов *другой* строки равна нулю.

В символах: если $A = (a_{ij})_{n\times n}$, то $\sum\limits_{k=1}^{n} a_{ik}A_{jk} = 0$ при всех $i \neq j$.

Доказательство леммы. Рассмотрим матрицу A', которая получается из матрицы A заменой ее j-й строки на i-ю.

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1\,n-1} & a_{1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{i\,n-1} & a_{in} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{j1} & a_{j2} & \dots & a_{j\,n-1} & a_{jn} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n\,n-1} & a_{nn} \end{pmatrix} \quad \Rightarrow \quad A'=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1\,n-1} & a_{1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{i\,n-1} & a_{in} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{i\,n-1} & a_{in} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n\,n-1} & a_{nn} \end{pmatrix}.$$

Ясно, что $\det A'=0$, поскольку в матрице A' есть две равные строки. С другой стороны, разложение $\det A'$ по j-й строке дает $\det A'=\sum\limits_{k=1}^{n} a_{ik}A'_{jk}$.

Остается заметить, что $A'_{jk}=A_{jk}$ для всех $k=1,\dots,n$, так как матрицы A' и A различаются лишь j-й строкой. Поэтому

$$\sum_{k=1}^{n} a_{ik}A_{jk}=\sum_{k=1}^{n} a_{ik}A'_{jk}=\det A'=0. \qquad \square$$

Определение

Пусть $A = (a_{ij})_{n\times n}$ — матрица порядка $n > 1$. Матрица $\widehat{A} := (A_{ij})^T$ называется *присоединенной* к матрице A.

Таким образом, присоединенная матрица получается, если заменить каждый элемент матрицы его алгебраическим дополнением, а потом транспонировать получившуюся матрицу.

Предложение (основное свойство присоединенной матрицы)

$A\widehat{A} = \widehat{A}A = \det A \cdot E.$

Доказательство. Элемент произведения $A\widehat{A}$, стоящий на месте (i, j), равен сумме произведений элементов i-й строки матрицы A на элементы j-го столбца матрицы $\widehat{A}$. По определению, элементы j-го столбца матрицы $\widehat{A}$ – это алгебраические дополнения элементов j-й строки матрицы A. Сумма произведений элементов i-й строки на алгебраические дополнения элементов j-й строки равна $\det A$ при $j = i$ (формула разложения по строке) и равна 0 при $j \neq i$ (лемма «свой среди чужих»). Поэтому у $A\widehat{A}$ элементы на диагонали равны $\det A$, а вне диагонали равны нулю, т.е. $A\widehat{A} = \det A \cdot E$. Равноправие строк и столбцов дает $\widehat{A}A = \det A \cdot E$. □

Из основного свойства присоединенной матрицы немедленно вытекает, что в случае, когда $\det A \neq 0$, матрица A обратима и ее обратная матрица может быть вычислена по формуле

$$A^{-1} = \frac{1}{\det A}\widehat{A}. \qquad (*)$$

Формулой $(*)$ удобно пользоваться для 2×2-матриц, где она принимает вид

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

Для матриц более высокого порядка формула $(*)$ вычислительно неэффективна по сравнению с методом Гаусса, который мы опишем позже. Тем не менее, она полезна: например, из нее следует, что элементы матрицы A^{-1} *непрерывно* зависят от элементов матрицы A.

Матрицу с ненулевым определителем называют *невырожденной*.
Подытожим наши рассмотрения следующим утверждением:

Теорема (критерий невырожденности)

Для любой квадратной матрицы A следующие условия эквивалентны:

1. A обратима слева;
2. A обратима справа;
3. A обратима и слева, и справа;
4. $\det A \neq 0$.

Доказательство получается комбинированием фактов, доказанных в этой лекции. $\square$

Рассмотрим систему из n линейных уравнений с n неизвестными:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{cases}$$

Основную матрицу этой системы обозначим через A, ее определитель – через Δ, а столбец свободных членов – через $\mathbf{b}$.

Для $i = 1, 2, \dots, n$ обозначим через Δ_i определитель, полученной заменой i-го столбца определителя Δ на столбец $\mathbf{b}$. Итак,

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}, \quad \Delta_1 = \begin{vmatrix} b_1 & a_{12} & \dots & a_{1n} \\ b_2 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_n & a_{n2} & \dots & a_{nn} \end{vmatrix},$$

$$\Delta_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} & \dots & a_{1n} \\ a_{21} & b_2 & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \\ a_{n1} & b_n & a_{n3} & \dots & a_{nn} \end{vmatrix}, \dots, \Delta_n = \begin{vmatrix} a_{11} & \dots & a_{1\,n-1} & b_1 \\ a_{21} & \dots & a_{2\,n-1} & b_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{n\,n-1} & b_n \end{vmatrix}.$$

Теорема (правило Крамера, 1750)

Если $\Delta \neq 0$, то система

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{cases}$$

имеет единственное решение, которое вычисляется по формулам

$$x_1 = \frac{\Delta_1}{\Delta},\ x_2 = \frac{\Delta_2}{\Delta},\ \dots,\ x_n = \frac{\Delta_n}{\Delta}. \qquad (\diamondsuit)$$

Доказательство. Существование и единственность решения при $\Delta \neq 0$ обсуждались выше. Проверим, что решение дается формулами $(\diamondsuit)$, т.е. проверим равенство $A\mathbf{x} = \mathbf{b}$, где $\mathbf{x} = \left(\frac{\Delta_1}{\Delta}, \frac{\Delta_2}{\Delta}, \dots, \frac{\Delta_n}{\Delta}\right)^T$.

i-я координата столбца $A\mathbf{x}$ равна $\sum_{j=1}^n a_{ij}\frac{\Delta_j}{\Delta}$. В этом выражении разложим определитель Δ_1 по первому столбцу, определитель Δ_2 — по второму столбцу, …, определитель Δ_n — по n-му столбцу.

Учитывая, что в определителе Δ_j элемент b_k стоит на месте (k, j), получим

$$\sum_{j=1}^{n} a_{ij}\frac{\Delta_j}{\Delta} = \frac{1}{\Delta}\sum_{j=1}^{n} a_{ij}\sum_{k=1}^{n} b_k A_{kj} = \frac{1}{\Delta}\sum_{k=1}^{n} b_k\sum_{j=1}^{n} a_{ij} A_{kj}.$$

Имеем

$$\sum_{j=1}^{n} a_{ij}A_{kj} = \begin{cases} \Delta & \text{при } k = i \text{ по формуле разложения по строке,} \\ 0 & \text{при } k \neq i \text{ по лемме «свой среди чужих».} \end{cases}$$

Поэтому сумма $\frac{1}{\Delta}\sum_{k=1}^{n} b_k \sum_{j=1}^{n} a_{ij}A_{kj}$ сводится к $\frac{1}{\Delta}b_i\Delta = b_i$.
Итак, i-я координата столбца $A\mathbf{x}$ равна b_i, т.е. $A\mathbf{x} = \mathbf{b}$. $\square$

Здесь можно сделать замечание, аналогичное тому, которое делалось по поводу формулы для обратной матрицы: правило Крамера вычислительно неэффективно, но теорема Крамера имеет важное теоретическое значение.

В качестве приложения теоремы Крамера рассмотрим следующую важную задачу. Известны значения (вообще говоря, неизвестной) функции $f(x)$ в точках $x_0,\ x_1,\ \dots,\ x_n$: $f(x_0) = y_0,\ f(x_1) = y_1,\ \dots,\ f(x_n) = y_n$. Нужно *интерполировать* $f(x)$, т.е. построить *многочлен* $p(x)$ наименьшей возможной степени, принимающий в данных точках указанные значения.

Будем искать $p(x)$ в виде $a_0 + a_1x + a_2x^2 + \dots + a_nx^n$. Условия $p(x_0) = y_0,\ p(x_1) = y_1,\ \dots,\ p(x_n) = y_n$ дают следующую систему линейных уравнений относительно коэффициентов $a_0, a_1, a_2, \dots, a_n$:

$$\begin{cases} a_0 + a_1x_0 + a_2x_0^2 + \dots + a_nx_0^n = y_0 \\ a_0 + a_1x_1 + a_2x_1^2 + \dots + a_nx_1^n = y_1 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_0 + a_1x_n + a_2x_n^2 + \dots + a_nx_n^n = y_n \end{cases} \quad \text{или} \quad \begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{pmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Ее основная матрица – это матрица Вандермонда порядка $n+1$. Так как определитель Вандермонда от $x_0, x_1, \dots, x_n$ отличен от 0, по теореме Крамера эта система имеет единственное решение.

Мы доказали такой факт:

Теорема об интерполяционном многочлене

Пусть $x_0, x_1, \ldots, x_n$ – попарно различные, а $y_0, y_1, \ \ldots, \ y_n$ – произвольные действительные числа. Существует и притом только один многочлен $p(x)$ степени не выше n, такой, что $p(x_i) = y_i$ для всех $i = 0, 1, \ldots, n$.

Отметим одно простое, но важное следствие.

В алгебре многочлены рассматривают как *формальные выражения*, и равенство двух многочленов $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ и $q(x) = b_0 + b_1x + b_2x^2 + \cdots + b_kx^k$ означает, что эти выражения совпадают, т.е. $n = k$ и $a_0 = b_0$, $a_1 = b_1$, $\ldots$, $a_n = b_n$.
В анализе многочлены над полем $\mathbb{R}$ рассматривают как *функции* $\mathbb{R} \to \mathbb{R}$, и равенство двух многочленов $p(x) \in \mathbb{R}$ и $q(x) \in \mathbb{R}$ означает, что $p(\alpha) = q(\alpha)$ для любого $\alpha \in \mathbb{R}$. Понятно, что каждый многочлен $p(x)$ определяет функцию $\mathbb{R} \to \mathbb{R}$ по правилу $\alpha \mapsto p(\alpha)$.

Следствие о многочленах как функциях

Многочлены с действительными коэффициентами равны тогда и только тогда, когда они равны как функции.

Доказательство. Ясно, что одинаковые многочлены определяют одну и ту же функцию. Обратно, пусть $p(x)$ и $q(x)$ – многочлены степеней n и k соответственно, которые равны как функции.
Без ограничения общности можно считать, что $n \geq k$. Выберем попарно различные числа $x_0, x_1, \ldots, x_n \in \mathbb{R}$. По условию $p(x_i) = q(x_i)$ для всех $i = 0, 1, \ldots, n$. В силу теоремы многочлен степени не выше n однозначно определяется своими значениями в точках $x_0, x_1, \ldots, x_n$, откуда $p(x) = q(x)$. □

В приложениях бывает необходимо явно выписать многочлен $p(x)$ степени не выше n по заданным значениям $y_0, y_1, \ldots, y_n$ в попарно различных точках $x_0, x_1, \ldots, x_n$. Соответствующую процедуру (как и многие другие алгоритмы в алгебре и анализе) предложил Жозеф Луи Лагранж.

Для каждого $i = 0, 1, \ldots, n$ положим

$$p_i(x) := \prod_{\substack{j=0\\ j\neq i}}^{n} \frac{x - x_j}{x_i - x_j} = \frac{x - x_0}{x_i - x_0} \cdots \frac{x - x_{i-1}}{x_i - x_{i-1}} \cdot \frac{x - x_{i+1}}{x_i - x_{i+1}} \cdots \frac{x - x_n}{x_i - x_n}.$$

Видно, что $p_i(x)$ – многочлен степени n, $p_i(x_i) = 1$ и $p_i(x_j) = 0$ при $j \neq i$. *Интерполяционный многочлен Лагранжа* – это многочлен

$$p(x) := \sum_{i=1}^{n} y_i \prod_{\substack{j=0\\ j\neq i}}^{n} \frac{x - x_j}{x_i - x_j} = y_0 p_0(x) + y_1 p_1(x) + \cdots + y_n p_n(x).$$

По построению $p(x_i) = y_i$ для всякого $i = 0, 1, \ldots, n$, а степень многочлена $p(x)$ не превосходит n.