线性代数讲义

王官杰

第 1-7 讲

目录

第一讲 高斯消元法及其矩阵表示

内容提要

(1) 线性方程组的高斯消元法
(2) 矩阵的概念
(3) 消元法的矩阵表示

本次课程, 我们将从大家熟悉的线性方程组出发, 引出求解线性方程组的高斯消元法. 之后, 引入矩阵, 并用矩阵的初等变换来实现高斯消元法.

1.1 线性方程组的消元法

考察二元一次线性方程组

$$\begin{cases} 2x - y = 1, & ① \\ x + y = 5. & ② \end{cases} \tag{1.1}$$

其中 x 和 y 是自变量.

在中学的时候, 我们已经学过, 可以使用消元法求解这个线性方程组. 我们可以首先做以下操作：$2 \times ② - ①$, 得

$$3y = 9. \quad ③$$

由③可得 $y = 3$. 将 $y = 3$ 代入①, 易得 $x = 2$. 因此方程组的解是

$$\begin{cases} x = 2, \\ y = 3. \end{cases}$$

容易看到, 上面的解法可以求解任意的二元线性方程组. 我们当然希望可以类似地求解有三个未知数的线性方程组, 有四个未知数的线性方程组, 以及更一般的有若干个未知数的线性方程组.

对于给定的 $n \in \mathbb{N}$, 我们称含有 n 个未知数的线性方程组为 n 元线性方程组. 使用 $x_1, x_2, \ldots, x_n$ 表示未知数, 则含有 m 个方程的线性方程组可以表示为

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

上式中 a_{ij}, b_i 是常数. 我们称 a_{ij} 是**系数**, b_i 是**常数项**. 注意系数 a_{ij} 的双下标的意义: 第一个下标 i 对应第 i 个方程, 第二个下标 j 对应第 j 个未知数 x_j. 我们还要强调的是, 线性方程组解的情况与未知数所用的符号无关, 而仅取决于系数与常数项.

如果一个线性方程组有解, 我们就称它是**相容的线性方程组**. 显然, 线性方程组的基本问题是如何求出解, 我们要讲述的就是线性方程组的求解故事.

问题 1.1 (线性方程组的三个基本问题)

我们可以将线性方程组的求解具体细化为下面的几个小问题:

(a) 线性方程组何时相容?
(b) 有解时, 有几个解?
(c) 给出具体的求解算法.

对于问题 (a), 我们希望找到判定是否有解的规则, 每个线性方程组不必求解, 只要去对照这个规则, 满足规则

则有解, 不满足则无解.

对于问题 (b), 我们希望找到计算解的“个数的公式”, 给定一个线性方程组, 不必求解, 代入这个公式就可以求出解的“个数”.

对于问题 (c), 我们希望有一个计算机看得懂的求解算法, 给定一个线性方程组, 输入它的系数和常数项后, 按照算法可以很快地得到解.

我们接下来要学习的内有, 有相当大一部分是围绕着以上三个问题展开的.

1.1.1 高斯消元法

解决问题的一个常用想法是先看简单的情形. 这里最简单的情形是: $m=n=1$, 此时方程组就是一个一元线性方程: $ax=b$. 它的解是非常明确的:

(1) 当 $a\neq 0$ 时有唯一解 $x=a^{-1}b$;

(2) 当 $a=0$ 且 $b=0$ 时, 任意数都是方程的解;

(3) 当 $a=0$ 且 $b\neq 0$ 时, 无解.

再看 $m=n=2$ 的情形, 此时方程组为

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

从几何上看, 这个方程组对应到求平面上直线 $a_{11}x_1+a_{12}x_2=b_1$ 和 $a_{21}x_1+a_{22}x_2=b_2$ 的交点问题. 于是我们得到

(1) 如果两直线相交, 那么有唯一解;

(2) 如果两直线重合, 那么任意的数对都是解;

(3) 如果两直线平行, 那么没有解.

在这两种简单情形下, 解都是三种情况. 于是我们有理由猜想一般的线性方程组的解的个数有三种可能性：一个解, 无穷多解, 以及无解. 后面我们将证明这个猜想成立. 另一方面, 我们看到, 未知数越少越容易求解. 所以一个自然的想法是, 是否可以把一般的线性方程组转化为未知数很少 (最好一个) 的“简单”线性方程来求解? 从很多未知数变到很少未知数, 就是消元, 而这正是中学求解线性方程组的方法.

消元通常有两种过程: 一种是逐步消去未知数, 另一种是一次消去其余未知数. 我们称前者为高斯消元法, 后者为加减消元法. 解线性方程组常用方法是高斯消元法, 于是线性方程组的一种解法是

$$\boxed{\text{线性方程组}} \xrightarrow{\text{高斯消元法}} \boxed{\text{简单线性方程组}}$$

下面我们举一个具体例子.

例题 1.1 在平面直角坐标系中, 作抛物线 $y=ax^2+bx+c$, 使其经过点 $(2,10),(1,0),(-3,20)$, 求抛物线方程.

解 由条件得到线性方程组

$$\begin{cases} 4a+2b+c=10 \\ \ \ a+\ \ b+c=0 \\ 9a-3b+c=20 \end{cases}$$

下面用高斯消元法解这个线性方程组:

$$
\text{原方程} \xrightarrow{(1)\leftrightarrow(2)} \begin{cases} a+\ \ b+c=0, \\ 4a+2b+c=10, \\ 9a-3b+c=20 \end{cases} \xrightarrow[(3)-9\times(1)]{(2)-4\times(1)} \begin{cases} a+b+c=0, \\ -2b-3c=10, \\ -12b-8c=20 \end{cases}
$$

$$
\xrightarrow{(3)-6\times(2)} \begin{cases} a+b+c=0, \\ -2b-3c=10, \\ 10c=-40 \end{cases} \xrightarrow{\frac{1}{10}\times(3)} \begin{cases} a+b+c=0, \\ -2b-3c=10, \\ c=-4 \end{cases}
$$

$$
\xrightarrow{(2)+3\times(3)} \begin{cases} a+b+c &=0, \\ -2b &=-2, \\ c &=-4 \end{cases} \xrightarrow{-\frac{1}{2}\times(2)} \begin{cases} a+b+c &=0, \\ b &=1, \\ c &=-4 \end{cases}
$$

$$
\xrightarrow[(1)-(3)]{(1)-(2)} \begin{cases} a=3 \\ b=1 \\ c=-4 \end{cases}
$$

于是所求的抛物线方程为 $y=3x^2+x-4$.

在上面求解线性方程组的过程中, 我们用 (1) 表示线性方程组的第一个方程, 其余类似. 记号 $(1)\leftrightarrow(2)$ 表示交换方程组中第一个和第二个方程的顺序, 记号 $(2)-4\times(1)$ 表示第二个方程的两边分别对应加上第一个方程两边的 -4 倍, 记号 $\frac{1}{10}\times(3)$ 表示第三个方程的两边同时乘 $\frac{1}{10}$, 其余类似.

1.1.2 初等变换

在上面的高斯消元法中, 我们将方程组变到另一个方程组. 一个自然的问题是: 在这个变换过程中, 方程组的解是否会改变? 现在我们回答这个问题.

定义 1.1

设 (A) 和 (B) 是两个 n 元线性方程组.

(1) 如果 (A) 和 (B) 有完全相同的解, 则称 (A) 和 (B) 同解.

(2) 设变换 $\mathcal{T}$ 将 (A) 变成 (B), 如果 (A) 和 (B) 同解, 则称 $\mathcal{T}$ 是一个同解变换.

注意到上面的高斯消元法中, 我们用到了三种变换. 于是有下面的定义.

定义 1.2

称下面三类线性方程组的变换为初等变换:

(1) 交换某两个方程的位置;

(2) 用非零常数乘某个方程;

(3) 用数乘某个方程再加到另一个方程上.

类似于上面的例题, 我们用 $(i)\leftrightarrow(j)$ 表示交换方程组中第 (i) 和第 (j) 两个方程的位置; 用 $a\times(i)$ 表示用非零数 a 乘第 (i) 个方程; 用 $(j)+a\times(i)$ 表示第 (i) 个方程乘数 a 后再加到第 (j) 个方程上. 容易证明这三个变换都是同解变换.

引理 1.1

线性方程组的初等变换是同解变换.

所以高斯消元法解线性方程组的过程可以描述如下:

$$\boxed{\text{线性方程组}} \xrightarrow{\text{初等变换}} \boxed{\text{简单线性方程组}}$$

在上面的求解过程中, 线性方程组中未知数所用符号并不是本质的, 本质的是系数和常数项. 为了抓住本质的东西, 记下来我们引入矩阵的概念, 用矩阵来记录线性方程组的系数和常数项.

1.2 矩阵的概念

1.2.1 矩阵

将线性方程组 $\begin{cases} 4a+2b+c=10 \\ a+\ \ b+c=0 \\ 9a-3b+c=20 \end{cases}$ 的未知数和等号去掉, 我们得到下面的数表

$$\begin{matrix} 4 & 2 & 1 & 10 \\ 1 & 1 & 1 & 0 \\ 9 & -3 & 1 & 20 \end{matrix}$$

注意为了能够恢复原来的线性方程组, 我们将这些数排得比较整齐: 每行对应一个方程, 前面三列是系数, 最后一列是常数项, 第一列是第一个未知数的系数, 等等. 这样整齐的一个数表就是矩阵. 为了告诉别人这个数表是一个整体, 我们用圆括号[1] 把它括起来, 得到

$$\begin{pmatrix} 4 & 2 & 1 & 10 \\ 1 & 1 & 1 & 0 \\ 9 & -3 & 1 & 20 \end{pmatrix}$$

一般地, 矩形 (长方形) 的数表称为矩阵. 于是矩阵的特点是外部框形, 内部成行成列. 一个 m 行 n 列的矩阵 ($m\times n$ 矩阵) 可以表示为

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

其中 a_{ij}, 是 (i,j) 位置的元素, i 表示所在的行, 而 j 表示所在的列. 常用 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 等大写的英文字母表示矩阵, 于是上面的矩阵可以表示为

$$\boldsymbol{A}=\boldsymbol{A}_{m\times n}=(a_{ij})=(a_{ij})_{m\times n}$$

行数和列数称为这个矩阵的型号, 两个矩阵如果型号相同, 则称它们同型矩阵.

注 数可以看成一个 1×1 矩阵, 所以矩阵是数的推广.

定义 1.3

设 $\boldsymbol{A}=(a_{ij})$ 和 $\boldsymbol{B}=(b_{ij})$ 是两个矩阵, 如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 同型, 且对任意的 i 和 j 有 $a_{ij}=b_{ij}$, 则称 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相等, 记为 $\boldsymbol{A}=\boldsymbol{B}$.

矩阵是本讲义的主角之一, 所以我们花费一些笔墨来介绍一些特殊的矩阵.

[1]有的书上用方括号

1.2.2 零矩阵

我们称所有元素为 0 的矩阵为零矩阵, 记为 $\boldsymbol{O}$, 即

$$\boldsymbol{O} = \boldsymbol{O}_{m\times n} = (0)_{m\times n}$$

是 $m \times n$ 的零矩阵. 根据定义, 零矩阵很多, 例如

$$\boldsymbol{O}_{2\times 2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq \boldsymbol{O}_{2\times 1} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

1.2.3 方阵

我们称 $n \times n$ 的矩阵为 n 阶方阵. 下面是一个一般的 n 阶方阵

$$\boldsymbol{A} = \boldsymbol{A}_n = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

这个方阵有两个对角线 (正方形的两个对角线), 其中一个对角线上的所有元素的行和列下标相等 (即 a_{ii} 所在的对角线), 称其为主对角线, 主对角线上的元素称为主对角元; 另一对角线上元素的行标和列标之和为 $n+1$, 称之为副对角线.

下面再介绍一些特殊的方阵. 如果一个方阵的主对角线下方的元素都等于零, 则称其为上三角阵, 于是一般形式是

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{pmatrix}, \quad (\text{空白位置为0}).$$

由定义, $\boldsymbol{A} = (a_{ij})_{n\times n}$ 是上三角阵的充分必要条件是, 对任意 $1 \leq j < i \leq n$, 有 $a_{ij} = 0$. 类似的, 如果一个方阵的主对角线上方的元素都等于零, 则称其为下三角阵, 其一般形式为

$$\begin{pmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad (\text{空白位置为0}).$$

如果一个方阵同时是上三角阵和下三角阵, 则称它是对角阵, 即主对角元外的元素都是零的方阵是对角阵. 对角阵常用下面的记号表示

$$\operatorname{diag}(a_1, a_2, \cdots, a_n) := \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}.$$

所有主对角元都相等的对角阵称为数量阵, 例如

$$\operatorname{diag}(\underbrace{a, \ldots, a}_{n}),$$

是 a 对应的 n 阶数量阵. 称数 1 对应的数量阵为**单位阵**, 记为 $\boldsymbol{E}$, 即 n 阶单位阵为

$$\boldsymbol{E}=\boldsymbol{E}_n:=\mathrm{diag}(\underbrace{1,\ldots,1}_{n})=\begin{pmatrix}1&&\\&\ddots&\\&&1\end{pmatrix}_{n\times n}.$$

1.2.4 向量

称 $n\times 1$ 的矩阵为 n 维列向量, 其一般形式为

$$\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}.$$

在不做特别说明的情况下, 本讲义中所提到的向量都指列向量. 列向量一般使用加粗的小写的字母表示, 如 $\boldsymbol{a},\boldsymbol{b},\ldots$. 称 $1\times n$ 的矩阵为 n 维行向量, 例如,

$$(a_1,a_2,\cdots,a_n),$$

行向量一般使用加粗的小写字母上加一个上标 T 表示[2], 如 $\boldsymbol{a}^{\mathrm{T}},\boldsymbol{b}^{\mathrm{T}},\ldots$.

我们称所有元素为 0 的向量为零向量, 记为 $\mathbf{0}$.

1.2.5 线性方程组与矩阵的对应关系

下面回到线性方程组. 设有 n 元线性方程组

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

我们用矩阵来记录这个方程组本质的东西. 令

$$\boldsymbol{A}=(a_{ij})_{m\times n}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix},$$

称其为**系数矩阵**. 记

$$\boldsymbol{b}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_m\end{pmatrix}$$

它是 m 维列向量, 称其为**常数项列向量**. 最后记

$$\widetilde{\boldsymbol{A}}=(\boldsymbol{A},\boldsymbol{b})=\left(\begin{array}{cccc|c}a_{11}&a_{12}&\cdots&a_{1n}&b_1\\a_{21}&a_{22}&\cdots&a_{2n}&b_2\\\vdots&\vdots&&\vdots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}&b_m\end{array}\right)$$

---

[1]有的书上记为 $\boldsymbol{I}$.

[2]该上标表示转置, 后面我们会讲到, 这里并不影响我们的讲述.

称其为**增广矩阵**. 增广矩阵记录了这个线性方程组的所有信息: 每一行对应一个方程, 最后一列对应常数项, 前面的列对应相应未知数的系数. 于是, 将每个线性方程组对应到它的增广矩阵, 我们得一个线性方程组与矩阵之间的对应关系

$$m\text{ 个方程的}n\text{ 元线性方程组} \xleftrightarrow{1-1\text{映射}} m\times(n+1)\text{ 矩阵}.$$

在这个对应关系下, 高斯消元法解线性方程组的过程可以描述如下:

$$\begin{tikzcd}
\boxed{\text{线性方程组}} \arrow[r, "\text{初等变换}"] \arrow[d, leftrightarrow, "\text{一一对应}"'] & \boxed{\text{简单线性方程组}} \arrow[d, leftrightarrow, "\text{一一对应}"] \\
\boxed{\text{增广矩阵}} \arrow[r, "??"] & \boxed{\text{简单矩阵}}
\end{tikzcd}$$

1.3 消元法的矩阵表示

我们的接下来的任务之一是将线性方程组的初等变换翻译到矩阵世界, 对应的变换我们称之为矩阵的初等(行) 变换.

1.3.1 矩阵的初等变换

线性方程组的方程对应到其增广矩阵的行, 对方程进行操作, 相当于对增广矩阵的行进行相应操作. 于是, 对应到线性方程组的初等变换, 我们有下面的定义.

定义 1.4

称下面三类变换为矩阵的初等行变换:

(a) 调行变换: $r_i \leftrightarrow r_j$ (交换第 i 行和第 j 行);

(b) 行数乘变换: $a \times r_i, a \neq 0$ (用 a 乘第 i 行所有元素);

(c) 行消去变换: $r_j + a \times r_i$ (用 a 乘第 i 行所有元素, 再加到第 j 行的对应位置元素上).

为了后面的应用, 我们还给出如下定义.

定义 1.5

称下面三类变换为矩阵的初等列变换:

(a) 调列变换: $c_i \leftrightarrow c_j$;

(b) 列数乘变换: $a \times c_i, a \neq 0$;

(c) 列消去变换: $c_j + a \times c_i$.

初等行变换和初等列变换统称为初等变换.

由于通常需要许多次初等行变换才能将一个矩阵化为“简单”矩阵, 所以我们引入下面概念.

定义 1.6

设 A, B 都是 $m \times n$ 阶矩阵,

(1) 如果 $\boldsymbol{A}$ 可经过有限次初等行变换化为 $\boldsymbol{B}$, 则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 行等价, 记为 $\boldsymbol{A} \overset{r}{\sim} \boldsymbol{B}$;

(2) 如果 $\boldsymbol{A}$ 可经过有限次初等列变换化为 $\boldsymbol{B}$, 则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 列等价, 记为 $\boldsymbol{A} \overset{c}{\sim} \boldsymbol{B}$;

(3) 如果 $\boldsymbol{A}$ 可经过有限次初等变换化为 $\boldsymbol{B}$, 则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价 , 记为 $\boldsymbol{A} \sim \boldsymbol{B}$.

行等价也称为行相抵, 列等价也称为列相抵, 等价也称为相抵.

行等价和列等价显然满足以下三个性质.

引理 1.2

矩阵的 (行, 列) 等价满足:

(1) (自反性) $\forall \boldsymbol{A}$, 有 $\boldsymbol{A} \sim \boldsymbol{A}$;

(2) (对称性) $\forall \boldsymbol{A}, \boldsymbol{B}$, 若 $\boldsymbol{A} \sim \boldsymbol{B}$, 则 $\boldsymbol{B} \sim \boldsymbol{A}$;

(3) (传递性) $\forall \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$, 若 $\boldsymbol{A} \sim \boldsymbol{B}, \boldsymbol{B} \sim \boldsymbol{C}$, 则 $\boldsymbol{A} \sim \boldsymbol{C}$.

证明 略 □

下面我们回到矩阵化简. 我们已经知道了化简的工具：初等 (行) 变换, 那么什么样的矩阵是“简单”的呢? 目前看来, 简单是为了可以直接写出对应的线性方程组的解. 我们来看一个例子.

例题 1.2 用矩阵的初等行变换, 表示例1.1中求解线性方程组

$$\begin{cases} 4a + 2b + c = 10 \\ a + b + c = 0 \\ 9a - 3b + c = 20 \end{cases} \tag{1.2}$$

的过程.

解 例1.1 中用高斯消元法解线性方程组(1.2)的过程, 用对应的增广矩阵表示为:

$$\begin{aligned}
\widetilde{\boldsymbol{A}} = \begin{pmatrix} 4 & 2 & 1 & 10 \\ 1 & 1 & 1 & 0 \\ 9 & -3 & 1 & 20 \end{pmatrix} &\xrightarrow{r_1 \leftrightarrow r_2} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 4 & 2 & 1 & 10 \\ 9 & -3 & 1 & 20 \end{pmatrix} \\
\xrightarrow[r_3 - 9r_1]{r_2 - 4r_1} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & -2 & -3 & 10 \\ 0 & -12 & -8 & 20 \end{pmatrix} &\xrightarrow{r_3 - 6r_2} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & -2 & -3 & 10 \\ 0 & 0 & 10 & -40 \end{pmatrix} \\
\xrightarrow{\frac{1}{10} r_3} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & -2 & -3 & 10 \\ 0 & 0 & 1 & -4 \end{pmatrix} &\xrightarrow{r_2 + 3r_3} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & -2 & 0 & -2 \\ 0 & 0 & 1 & -4 \end{pmatrix} \\
\xrightarrow{-\frac{1}{2} r_2} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -4 \end{pmatrix} &\xrightarrow[r_1 - r_2]{r_1 - r_3} \begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -4 \end{pmatrix}
\end{aligned}$$

1.3.2 阶梯型矩阵和简化阶梯型矩阵

观察上面例子中用初等行变换化简矩阵的过程, 我们主要做了两件事: 一个是利用行消去变换不断的产生零, 以至于最后得到的矩阵除最后一列外, 每一行至多只有一个数非零, 而且这些非零的数位于不同的列[3]; 另一个是把这些非零的数变成 1[4]. 这些非零的数我们称之为首元, 即从左往右看, 矩阵中非零行的第一个非零元是**首元**. 称首元所在的列为**首元列**.

例题 1.3 下面两个矩阵中加方框的数都是首元, 首元列都是第 1, 2, 3 列.

$$\begin{pmatrix} 0 & \boxed{3} & -1 & 0 & 2 \\ \boxed{-2} & 0 & 3 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{-1} & 5 & 1 \end{pmatrix}, \quad \begin{pmatrix} \boxed{1} & 0 & 0 & 3 \\ 0 & \boxed{1} & 0 & 1 \\ 0 & 0 & \boxed{1} & -4 \end{pmatrix}$$

需要注意的是, 并不是所有的线性方程组都可以像例题1.2中那样, 仅仅通过初等行变换就可以将增广矩阵中的系数矩阵转化为单位矩阵. 例如, 若某个线性方程组的增广矩阵是

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & -4 \end{pmatrix},$$

我们无法继续使用初等行变换将其系数矩阵变为初等矩阵.

但是, 我们可以仅仅通过初等行变换, 将线性方程组对应的增广矩阵转化为下面两种简单矩阵. 我们首先给出这两种简单矩阵的定义[5].

定义 1.7

(1) 称一个矩阵是阶梯形阵, 如果它满足

  (a) 全零行下方无非零行;

  (b) 各非零行的首元在上一行首元的右边.

(2) 称一个矩阵是简化阶梯形阵 (或者行等价标准形阵), 如果它满足

  (a) 是阶梯形阵;

  (b) 首元都是 1 ;

  (c) 各首元列只有一个非零元.

显然, 在阶梯形阵中

$$\text{非零行数} = \text{首元数} = \text{首元列数}.$$

且由定义, 在阶梯形阵中可以画一个台阶, 其一般形式是

$$\begin{pmatrix} 0 & \cdots & 0 & \triangle & * & 0 & * & 0 & & 0 & & 0 & \\ 0 & \cdots & 0 & & & \triangle & * & 0 & & 0 & & 0 & \\ 0 & \cdots & 0 & & & & & \triangle & & 0 & * & 0 & * \\ \vdots & \vdots & \vdots & & & & & & \ddots & \vdots & & \vdots & \\ 0 & \cdots & 0 & & & & & & & \triangle & & 0 & \\ 0 & \cdots & 0 & & & & & & & & & \triangle & \\ \vdots & \vdots & \vdots & & & & & & & & & & \\ 0 & \cdots & 0 & & & & & & & & & & \end{pmatrix}, \quad \triangle \neq 0.$$

---

[3] 如此就把未知数分离开了.

[4] 如此就可以直接写出解.

[5] 请读者仔细体会为什么我们觉得如此定义的 (简化) 阶梯形阵为简单矩阵.

而简化阶梯形阵的一般形式是

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & * & 0 & * & 0 & & 0 & & 0 & \\ 0 & \cdots & 0 & & & 1 & * & 0 & & 0 & & 0 & \\ 0 & \cdots & 0 & & & & & 1 & & 0 & * & 0 & * \\ \vdots & \vdots & \vdots & & & & & & \ddots & \vdots & & \vdots & \\ 0 & \cdots & 0 & & & & & & & 1 & & 0 & \\ 0 & \cdots & 0 & & & & & & & & & 1 & \\ \vdots & \vdots & \vdots & & & & & & & & & & \\ 0 & \cdots & 0 & & & & & & & & & & \end{pmatrix}$$

例题 1.4 判断下列矩阵是否为阶梯型矩阵、简化阶梯型矩阵.

(1) $\begin{pmatrix} 0 & 3 & -1 & 0 & 2 \\ -2 & 0 & 3 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 5 & 1 \end{pmatrix}$: 非阶梯形阵.

(2) $\begin{pmatrix} 0 & -1 & 2 & 3 & 1 \\ 0 & & 3 & 4 & 2 \\ 0 & & & & 2 \\ 0 & & & & \end{pmatrix}$ $\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -4 \end{pmatrix}$: 阶梯形阵, 但非简化阶梯形阵.

(3) $\begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -4 \end{pmatrix}$: 简化阶梯形阵.

通过前面的讨论我们知道, 对线性方程组的增广矩阵进行初等行变换等价于对线性方程组进行初等变换, 而线性方程组的初等变换并不改变线性方程组的解. 将增广矩阵转化为阶梯型矩阵或者简化阶梯型矩阵, 本质上是对线性方程组进行简化. 于是关于简化线性方程组的讨论可以转化为关于矩阵化为阶梯型矩阵或简化阶梯型矩阵的讨论. 下面我们进入本节的主题: 用初等行变换将一个矩阵化为阶梯形阵和简化阶梯形阵.

定理 1.1

设 $\boldsymbol{A}$ 是一个 $m \times n$ 的矩阵, 则

(1) $\boldsymbol{A} \overset{r}{\sim}$ 阶梯形阵;

(2) $\boldsymbol{A} \overset{r}{\sim}$ 简化阶梯形阵.

证明 这个证明就是给出化简的算法.

算法 1.1 (矩阵化为阶梯型和简化阶梯型矩阵)

(1) 算法一: $\boldsymbol{A} \overset{r}{\sim}$ 阶梯形阵.

第一步 若 $\boldsymbol{A} = \boldsymbol{O}$, 则结束; 否则, 可设

$$\boldsymbol{A} = \begin{pmatrix} 0 & \cdots & 0 & a_{1j} & \\ \vdots & & \vdots & \vdots & * \\ 0 & \cdots & 0 & a_{mj} & \end{pmatrix}, \quad \text{第} j \text{ 列是非零列}.$$

于是总可用调行变换, 使得 $\boldsymbol{A}$ 行等价于

$$\boldsymbol{A}_1=\begin{pmatrix} 0 & \cdots & 0 & \Delta & \\ \vdots & & \vdots & \vdots & * \\ 0 & \cdots & 0 & * & \end{pmatrix}, \quad \Delta\neq 0$$

再用行消去变换 $r_i-\dfrac{*}{\Delta}r_1$, 可得 $\boldsymbol{A}$ 行等价于

$$\boldsymbol{A}_2=\begin{pmatrix} 0 & \cdots & 0 & \Delta & * \\ 0 & \cdots & 0 & 0 & \\ \vdots & & \vdots & \vdots & B \\ 0 & \cdots & 0 & 0 & \end{pmatrix}$$

第二步 对 $\boldsymbol{B}$ 重复第一步. 由于 $\boldsymbol{B}$ 的左边都是零, 所以不论对 $\boldsymbol{B}$ 所在的行进行何种初等行变换, $\boldsymbol{B}$ 的左边还是零. 于是只要考虑 $\boldsymbol{B}$. 由于 $\boldsymbol{A}$ 的列数 n 是有限数, 所以有限步后可以将 $\boldsymbol{A}$ 行变为阶梯形阵.

(2) 算法二: $\boldsymbol{A}\overset{r}{\sim}$ 简化阶梯形阵.

$$\boldsymbol{A}\xrightarrow{\text{算法一}}\text{阶梯型矩阵}\xrightarrow{\text{行数乘}}\text{首元为 }1\xrightarrow{\text{行消去}}\text{首元列其余元为 }0$$

□

需要指出的是, 与 $\boldsymbol{A}$ 行等价的简化阶梯形阵是唯一的 (为什么?); 与 $\boldsymbol{A}$ 行等价的阶梯形阵虽然不是唯一的, 但是非零行数却是由 $\boldsymbol{A}$ 唯一确定的[6].

例题 1.5 求与矩阵 $\boldsymbol{A}=\begin{pmatrix} 0 & 0 & 0 & 2 & 2 \\ 0 & 2 & 4 & 3 & 1 \\ 0 & 1 & 2 & 4 & 3 \end{pmatrix}$ 行等价的简化阶梯形阵.

解 我们进行如下的初等行变换

$$\boldsymbol{A}\xrightarrow{r_1\leftrightarrow r_3}\begin{pmatrix} 0 & \boxed{1} & 2 & 4 & 3 \\ 0 & 2 & 4 & 3 & 1 \\ 0 & 0 & 0 & 2 & 2 \end{pmatrix}\xrightarrow{r_2-2r_1}\begin{pmatrix} 0 & 1 & 2 & 4 & 3 \\ 0 & 0 & 0 & -5 & -5 \\ 0 & 0 & 0 & 2 & 2 \end{pmatrix}$$
$$\xrightarrow[\frac{1}{2}r_3]{-\frac{1}{5}r_2}\begin{pmatrix} 0 & 1 & 2 & 4 & 3 \\ 0 & 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}\xrightarrow[r_3-r_2]{r_1-4r_2}\begin{pmatrix} 0 & \boxed{1} & 2 & 0 & -1 \\ 0 & 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

最后一个矩阵就是所求的简化阶梯形阵.

[6]称与 $\boldsymbol{A}$ 等价的阶梯型矩阵的非零行的数目为矩阵的秩, 矩阵的秩与问题1.1的答案密切相关, 后面我们还会做专门的介绍.

第二讲 矩阵的基本运算

内容提要

(1) 矩阵的加法
(2) 数量乘积
(3) 矩阵的乘法
(4) 矩阵的转置
(5) 分块矩阵及其运算

我们知道, 对于一个含有 m 个方程, n 个未知量的线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad\qquad \cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases} \tag{2.1}$$

记载其核心信息的是系数矩阵 $\boldsymbol{A}$ 和常数项 $\boldsymbol{b}$. 对于未知量 $x_1, x_2, \ldots, x_n$, 我们可以将他们看作是一个向量的各个分量[1], 记该向量为 $\boldsymbol{x}$. 在讨论线性方程组(2.1)时, 我们约定

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \ \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \ \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

为了更方便地讨论线性方程组, 我们在本节引入矩阵加法、乘法以及数量乘法.

2.1 矩阵的加法

定义 2.1

设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是两个同型矩阵, 称将对应位置元素相加得到的新的同型矩阵为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的和, 即若 $\boldsymbol{A} = (a_{ij})_{m\times n}$, $\boldsymbol{B} = (b_{ij})_{m\times n}$, 则

$$\boldsymbol{A} + \boldsymbol{B} = (a_{ij})_{m\times n} + (b_{ij})_{m\times n} = (a_{ij} + b_{ij})_{m\times n}.$$

例题 2.1

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} a+1 & b+2 \\ c+3 & d+4 \end{pmatrix}$$

为了定义矩阵的减法, 我们首先给出负矩阵的概念.

定义 2.2 (负矩阵)

设矩阵 $\boldsymbol{A} = (a_{ij})_{m\times n}$, 称矩阵 $(-a_{ij})_{m\times n}$ 为 $\boldsymbol{A}$ 的负矩阵, 记为 $-\boldsymbol{A}$.

有了负矩阵的概念, 我们就可以定义矩阵的减法.

定义 2.3

矩阵的减法定义为

$$\boldsymbol{A} - \boldsymbol{B} = \boldsymbol{A} + (-\boldsymbol{B}).$$

[1] 在本讲义中, 如果不做特别说明, 我们所称的向量都是列向量.

显然, 任何一个矩阵减去它本身, 得到的新矩阵的所有元素都为零. 我们把这种所有元素都为零的矩阵成为零矩阵, 记为 $\boldsymbol{O}$.

根据矩阵加法和减法的定义, 我们很容易得到矩阵加法和减法的性质

性质 2.1

设 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{O}$ 为同型矩阵, 其中 $\boldsymbol{O}$ 是零矩阵, 则

1. (交换律) $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$;
2. (结合律) $(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}=\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$;
3. (零元存在) $\boldsymbol{A}+\boldsymbol{O}=\boldsymbol{A}$;
4. (负元存在) $\boldsymbol{A}+(-\boldsymbol{A})=\boldsymbol{A}-\boldsymbol{A}=\boldsymbol{O}$.

2.2 数量乘积

定义 2.4

我们称矩阵 $\boldsymbol{A}$ 中的每个元素都乘以 k 得到的新矩阵为 k 与矩阵 $\boldsymbol{A}$ 的数量乘积(数乘), 记作 $k\boldsymbol{A}$. 即若 $\boldsymbol{A}=(a_{ij})_{m\times n}$, 则

$$k\boldsymbol{A}=k(a_{ij})_{m\times n}=(a_{ij})_{m\times n}.$$

例题 2.2

$$5\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}=\begin{pmatrix}5 & 10\\ 15 & 20\end{pmatrix}$$

显然, 数量乘积有以下性质

性质 2.2

设 $\boldsymbol{A},\boldsymbol{B}$ 为 $m\times n$ 的矩阵, k,k_1,k_2 为任意实数, 则

1. $k(\boldsymbol{A}+\boldsymbol{B})=k\boldsymbol{A}+k\boldsymbol{B}$;
2. $(k_1+k_2)\boldsymbol{A}=k_1\boldsymbol{A}+k_1\boldsymbol{B}$;
3. $(k_1k_2)\boldsymbol{A}=k_1(k_2\boldsymbol{A})$;
4. $1\cdot\boldsymbol{A}=\boldsymbol{A}$; $(-1)\cdot\boldsymbol{A}=-\boldsymbol{A}$;
5. $0\times\boldsymbol{A}=\boldsymbol{O}$; $a\boldsymbol{O}=\boldsymbol{O}$ 且
$$k\boldsymbol{A}=\boldsymbol{O}\Leftrightarrow k=0 \text{ 或} \boldsymbol{A}=\boldsymbol{O};$$
6. $\operatorname{diag}(k,k,\ldots,k)=k\boldsymbol{E}_n$.

有了矩阵的加法和数量乘法, 我们就可以将该线性方程组改写为

$$x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\ldots+x_n\boldsymbol{a}_n=\boldsymbol{b}, \tag{2.2}$$

其中

$$\boldsymbol{a}_i=\begin{pmatrix}a_{1i}\\ a_{2i}\\ \vdots\\ a_{mi}\end{pmatrix},\ \boldsymbol{b}=\begin{pmatrix}b_1\\ b_2\\ \vdots\\ b_m\end{pmatrix}\ (i=1,\ldots,n).$$

注意, 我们把列向量看作是一个 $m\times 1$ 的矩阵. 我们将等式(2.2)的左侧称作向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_n$ 的线性组合. 线性方程组有没有解这一问题的本质就是在问, 常数项 $\boldsymbol{b}$ 是否可以表示成向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_n$ 的线性组合; 线性方程组的解是不是唯一这一问题的本质上就是在问, 这种线性表示是否唯一. 于是, 关于线性方程组的讨论, 就可

以被转化为关于向量组的讨论；反之, 关于向量组的讨论, 也可以被转化为关于线性方程组的讨论. 我们将在后面的章节再对向量组进行详细的讨论.

2.3 矩阵的乘法

矩阵的乘法定义与矩阵加法的略微不同, 我们并不直接将矩阵乘法定义为对应位置元素相乘得到的新矩阵[2], 而是像下面这样定义矩阵的乘法.

定义 2.5

设 $\boldsymbol{A}$ 是 $m \times s$ 的矩阵, $\boldsymbol{B}$ 是 $s \times n$ 的矩阵, 则我们称由元素

$$\begin{aligned} c_{ij} &= a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{is}b_{sj} \\ &= \sum_{k}^{s} a_{ik}b_{kj}, \ (i = 1, \ldots, m; j = 1, \ldots, n) \end{aligned}$$

构成的矩阵 $\boldsymbol{C} = (c_{ij})_{m\times n}$ 为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的乘积, 记作

$$\boldsymbol{C} = \boldsymbol{AB}.$$

按照矩阵的定义, 易知

$$(a_{i1}, a_{i2}, \cdots, a_{is}) \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{sj} \end{pmatrix} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{is}b_{sj} = c_{ij},$$

由此可见 c_{ij} 即为 $\boldsymbol{A}$ 的第 i 行元素构成的行向量与 $\boldsymbol{B}$ 的第 j 列向量的乘积.

注 只有第一个矩阵的列数与第二个矩阵的行数相等时, 两个矩阵才可以做乘法.

由矩阵乘法的定义, 我们可以得到矩阵乘法的几个性质

性质 2.3

设 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 为矩阵, k 是一个数, 则有（假设运算可以进行）

(1) (结合律) $(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$;

(2) (分配律) $\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C}) = \boldsymbol{AB}+\boldsymbol{AC}$; $(\boldsymbol{B}+\boldsymbol{C})\boldsymbol{A} = \boldsymbol{BA}+\boldsymbol{CA}$;

(3) (单位元存在) $\boldsymbol{E}_m\boldsymbol{A}_{m\times n} = \boldsymbol{A}_{m\times n}$; $\boldsymbol{A}_{m\times n}\boldsymbol{E}_n = \boldsymbol{A}_{m\times n}$;

(4) $k(\boldsymbol{AB}) = (k\boldsymbol{A})\boldsymbol{B} = \boldsymbol{A}(k\boldsymbol{B}) =: k\boldsymbol{AB}$;

(5) $\boldsymbol{A}_{m\times s}\boldsymbol{O}_{s\times n} = \boldsymbol{O}_{m\times n}$; $\boldsymbol{O}_{m\times s}\boldsymbol{A}_{s\times n} = \boldsymbol{O}_{m\times n}$.

证明 这里仅证明结合律, 余下的性质都比较简单. 假设 $\boldsymbol{A} = (a_{ij})_{m\times s}$, $\boldsymbol{B} = (b_{ij})_{s\times t}$, $\boldsymbol{C} = (c_{ij})_{t\times n}$, 易知 $(\boldsymbol{AB})\boldsymbol{C}$ 和 $\boldsymbol{A}(\boldsymbol{BC})$ 都是 $m\times n$ 的矩阵, 因此我们只需要证明两个矩阵对应位置的元素相等即可. 记 $\boldsymbol{D} = (d_{ij})_{m\times t} = \boldsymbol{AB}$, $\boldsymbol{F} = (f_{ij})_{s\times n} = \boldsymbol{BC}$, 则 $(\boldsymbol{AB})\boldsymbol{C}$ 中第 i 行 j 列的元素为

$$\sum_{k=1}^{t} d_{ik}c_{kj} = \sum_{k=1}^{t}\left(\sum_{l=1}^{s} a_{il}b_{lk}\right)c_{kj} = \sum_{k=1}^{t}\sum_{l=1}^{s} a_{il}b_{lk}c_{kj}.$$

$\boldsymbol{A}(\boldsymbol{BC})$ 中第 i 行 j 列的元素为

$$\sum_{l=1}^{s} a_{il}f_{lj} = \sum_{l=1}^{s} a_{il}\left(\sum_{k=1}^{t} b_{lk}c_{kj}\right) = \sum_{l=1}^{s}\sum_{k=1}^{t} a_{il}b_{lk}c_{kj} = \sum_{k=1}^{t}\sum_{l=1}^{s} a_{il}b_{lk}c_{kj}.$$

[2] 事实上, 也有像矩阵加法那样定义的矩阵乘法, 假设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是同型矩阵, 称将对应位置元素相乘得到的新矩阵为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的 Hadamard 乘积, 记为 $\boldsymbol{A}\circ\boldsymbol{B}$, 但 Hadamard 乘积不是我们通常说的矩阵乘法. 除了 Hadamard 乘积, 还有其它的矩阵乘积, 这里不做介绍.

于是 $(\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{A}(\boldsymbol{BC})$. □

例题 2.3 $\boldsymbol{A}=\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}$, $\boldsymbol{B}=\begin{pmatrix}8&1\\7&0\\4&9\end{pmatrix}$, 求 $\boldsymbol{AB}$ 和 $\boldsymbol{BA}$.

解 由矩阵的乘法易知,

$$\boldsymbol{AB}=\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}\begin{pmatrix}8&1\\7&0\\4&9\end{pmatrix}=\begin{pmatrix}34&28\\91&58\end{pmatrix}$$

$$\boldsymbol{BA}=\begin{pmatrix}8&1\\7&0\\4&9\end{pmatrix}\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}=\begin{pmatrix}12&21&30\\7&14&21\\40&53&66\end{pmatrix}$$

这个例子说明, 乘法交换律 $\boldsymbol{AB}=\boldsymbol{BA}$ 对矩阵乘法一般不成立!

例题 2.4 设 $\boldsymbol{A}=\begin{pmatrix}1&0\\0&0\end{pmatrix}$, $\boldsymbol{B}=\begin{pmatrix}0&0\\0&1\end{pmatrix}$, 则 $\boldsymbol{AB}=\boldsymbol{O}$.

注 这个例子说明, 对于矩阵乘法, 从 $\boldsymbol{AB}=\boldsymbol{O}$ 不能推出 $\boldsymbol{A}=\boldsymbol{O}$ 或 $\boldsymbol{B}=\boldsymbol{O}$.
注 从 $\boldsymbol{AB}=\boldsymbol{AC}$ 和 $\boldsymbol{A}\neq\boldsymbol{O}$ 不能推出 $\boldsymbol{B}=\boldsymbol{C}$, 即乘法消去律一般不成立!

定义 2.6

设 $\boldsymbol{A}$ 是 n 阶方阵, k 是正整数, 称 k 个 $\boldsymbol{A}$ 连乘为 $\boldsymbol{A}$ 的 k 次幂, 记为 $\boldsymbol{A}^k$. 我们规定, $\boldsymbol{A}^0=\boldsymbol{E}$.

由定义可知

$$\boldsymbol{A}^k=\begin{cases}\boldsymbol{E}_n, & k=0,\\ \underbrace{\boldsymbol{A}\cdots\boldsymbol{A}}_{k\text{个相乘}}, & k\in\mathbb{N}^+\end{cases}$$

例题 2.5 设 $\boldsymbol{A}=(3,0,1,0)$, $\boldsymbol{B}=(-1,7,4,11)^{\mathrm{T}}$, 求 $(\boldsymbol{BA})^{2025}$.

解 由矩阵乘法的性质可知, $\boldsymbol{AB}=1$, 因此

$$\begin{aligned}(\boldsymbol{BA})^{2025}&=\boldsymbol{B}\left[(\boldsymbol{AB})^{2024}\right]\boldsymbol{A}=\boldsymbol{B}\left(1^{2024}\right)\boldsymbol{A}\\ &=\boldsymbol{BA}=\begin{pmatrix}-1\\7\\4\\11\end{pmatrix}(3,0,1,0)=\begin{pmatrix}-3&0&-1&0\\21&0&7&0\\12&0&4&0\\33&0&11&0\end{pmatrix}\end{aligned}$$

有了矩阵的乘法, 我们就可以将线性方程组(2.1)改写为

$$\boldsymbol{Ax}=\boldsymbol{b},$$

其中

$$\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix},\ \boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\ \boldsymbol{b}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_m\end{pmatrix}.$$

这种表示方法的好处是, 可以用矩阵来研究线性方程组解的结构; 反之, 也可以用线性方程组解的结构来讨论矩阵的性质.

2.4 矩阵的转置

定义 2.7

我们称将矩阵 $\boldsymbol{A}$ 的行和列交换后得到的矩阵为 $\boldsymbol{A}$ 的转置矩阵, 记为 $\boldsymbol{A}^{\mathrm{T}}$ 或 $\boldsymbol{A}'$, 即若 $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$, 则

$$\boldsymbol{A}^{\mathrm{T}} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

性质 2.4

设 $\boldsymbol{A}, \boldsymbol{B}$ 为矩阵, k 是一个数, 则有 (假设运算可以进行)

(1) $(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{A}$;

(2) $(\boldsymbol{A}+\boldsymbol{B})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}\mathrm{T}$;

(3) $(k\boldsymbol{A})^{\mathrm{T}} = k\boldsymbol{A}^{\mathrm{T}}$;

(4) $(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$, 进一步有 $(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k)^{\mathrm{T}} = \boldsymbol{A}_k^{\mathrm{T}}\cdots\boldsymbol{A}_2^{\mathrm{T}}\boldsymbol{A}_1^{\mathrm{T}}$.

证明 性质 (1)-(3) 显然成立, 下证性质 (4). 设 $\boldsymbol{A} = (a_{ij})_{m\times s}, \boldsymbol{B} = (B_{ij})_{s\times n}$, 则 $(\boldsymbol{AB})^{\mathrm{T}}$ 和 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$ 都是 $n \times m$ 矩阵, 我们只需证明它们对应位置的元素相同即可.

$(\boldsymbol{AB})^{\mathrm{T}}$ 第 i 行 j 列的元素与 $\boldsymbol{AB}$ 第 j 行第 i 列的元素相同, 也就是 $\boldsymbol{A}$ 的第 j 行与 $\boldsymbol{B}$ 的第 i 列的对应位置元素乘积的之和,

$$a_{j1}b_{1i} + a_{j2}b_{2i} + \ldots + a_{js}b_{si} = \sum_{k}^{s} a_{jk}b_{ki}.$$

$\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$ 位于第 i 行 j 的元素为 $\boldsymbol{B}^{\mathrm{T}}$ 的第 i 行与 $\boldsymbol{A}^{\mathrm{T}}$ 的第 j 列的对应位置元素乘积之和. 注意到 $\boldsymbol{B}^{\mathrm{T}}$ 的第 i 行是 $\boldsymbol{B}$ 的第 i 列, $\boldsymbol{A}^{\mathrm{T}}$ 的第 j 列是 $\boldsymbol{A}$ 的第 j 行, 因此该元素为

$$b_{1i}a_{j1} + b_{2i}a_{j2} + \ldots + b_{si}a_{js} = \sum_{k=1}^{s} b_{ki}a_{jk} \sum_{k}^{s} a_{jk}b_{ki}.$$

于是 $(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$. □

我们对转置等于本身的矩阵特别感兴趣, 这类矩阵被称为对称矩阵.

定义 2.8

设 $\boldsymbol{A}$ 是一个 n 阶方阵, 我们称 $\boldsymbol{A}$ 是对称矩阵, 若果 $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}$.

类似地, 转置等于本身与 -1 数量乘积的矩阵被称为反对称矩阵.

定义 2.9

设 $\boldsymbol{A}$ 是一个 n 阶方阵, 我们称 $\boldsymbol{A}$ 是反对称矩阵, 若果 $\boldsymbol{A}^{\mathrm{T}} = -\boldsymbol{A}$.

例题 2.6 试证明任意一个方阵都可以写成一个对称阵和反对称阵的和.

证明 易知

$$\boldsymbol{A}=\frac{1}{2}(\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}})+\frac{1}{2}(\boldsymbol{A}-\boldsymbol{A}^{\mathrm{T}}).$$

显然 $\frac{1}{2}(\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}})$ 是对称矩阵, $\frac{1}{2}(\boldsymbol{A}-\boldsymbol{A}^{\mathrm{T}})$ 是反对称矩阵. □

✍ **练习 2.1** 试证明

(1) 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为对称矩阵, 则 $\boldsymbol{A}\pm\boldsymbol{B}$ 也是对称矩阵.

(2) 设 $\boldsymbol{A}$ 为对称矩阵, 则 $k\boldsymbol{A}$ 也是对称矩阵.

(3) 设 $\boldsymbol{A}$ 为任何方阵, 则 $\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}$ 必是对称矩阵.

(4) 设 $\boldsymbol{A}$ 为任何矩阵, 则 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 和 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 都是对称矩阵.

✍ **练习 2.2** 试证明反对称矩阵的对角元全为 0.

2.5 分块矩阵及其运算

假设 $\boldsymbol{A}$ 是一个矩阵, 我们可以根据需要, 用若干条横线和竖线将 $\boldsymbol{A}$ 分割成许多个小矩阵, 每个小矩阵称为矩阵的子块, 以字块为元的形式上的矩阵称为分块矩阵.

例题 2.7 对于矩阵

$$\boldsymbol{A}=\begin{pmatrix}1&0&0&0\\0&1&0&0\\-1&2&1&0\\1&1&0&1\end{pmatrix}$$

分成子块的分法很多, 下面举出三种分块形式:

$$\left(\begin{array}{cc|cc}1&0&0&0\\0&1&0&0\\\hline -1&2&1&0\\1&1&0&1\end{array}\right),\quad \left(\begin{array}{cccc}1&0&0&0\\\hline 0&1&0&0\\\hline -1&2&1&0\\\hline 1&1&0&1\end{array}\right),\quad \left(\begin{array}{c|ccc}1&0&0&0\\0&1&0&0\\-1&2&1&0\\1&1&0&1\end{array}\right).$$

第一种分法可记为

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{E}_2&\boldsymbol{O}\\\boldsymbol{A}_1&\boldsymbol{E}_2\end{pmatrix},$$

其中

$$\boldsymbol{E}_2=\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad \boldsymbol{A}_1=\begin{pmatrix}-1&2\\1&1\end{pmatrix},\quad \boldsymbol{O}=\begin{pmatrix}0&0\\0&0\end{pmatrix},$$

即 $\boldsymbol{E}_2,\boldsymbol{O},\boldsymbol{A}_1,\boldsymbol{E}_2$ 为 $\boldsymbol{A}$ 的子块, 而 $\boldsymbol{A}$ 形式上成为以这些子块为元的分块矩阵. 第二种、第三种形式的分块矩阵的记法可类似给出.

例题 2.8

$$\boldsymbol{A}=\left(\begin{array}{cc|cc|cc}1&0&1&0&1&0\\0&1&0&1&0&1\\\hline 1&0&1&0&1&0\\0&1&0&1&0&1\end{array}\right)=\begin{pmatrix}\boldsymbol{E}_2&\boldsymbol{E}_2&\boldsymbol{E}_2\\\boldsymbol{E}_2&\boldsymbol{E}_2&\boldsymbol{E}_2\end{pmatrix}.$$

2.5.1 分块矩阵的运算

分块矩阵的运算法则与普通矩阵的运算法则类似.

(1) 分块矩阵的加法

设同型矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的分块法, 即

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1s}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2s}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{r1} & \boldsymbol{A}_{r2} & \cdots & \boldsymbol{A}_{rs}\end{pmatrix},\ \boldsymbol{B}=\begin{pmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \cdots & \boldsymbol{B}_{1s}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} & \cdots & \boldsymbol{B}_{2s}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{B}_{r1} & \boldsymbol{B}_{r2} & \cdots & \boldsymbol{B}_{rs}\end{pmatrix},$$

其中 $\boldsymbol{A}_{ij}$ 与 $\boldsymbol{B}_{ij}$ 也是同型矩阵 $(i=1,2,\cdots,r;j=1,2,\cdots,s)$, 则

$$\boldsymbol{A}\pm\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A}_{11}\pm\boldsymbol{B}_{11} & \boldsymbol{A}_{12}\pm\boldsymbol{B}_{12} & \cdots & \boldsymbol{A}_{1s}\pm\boldsymbol{B}_{1s}\\ \boldsymbol{A}_{21}\pm\boldsymbol{B}_{21} & \boldsymbol{A}_{22}\pm\boldsymbol{B}_{22} & \cdots & \boldsymbol{A}_{2s}\pm\boldsymbol{B}_{2s}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{r1}\pm\boldsymbol{B}_{r1} & \boldsymbol{A}_{r2}\pm\boldsymbol{B}_{r2} & \cdots & \boldsymbol{A}_{rs}\pm\boldsymbol{B}_{rs}\end{pmatrix}.$$

(2) 分块矩阵的数乘

用数 λ 乘一个分块矩阵时, 等于用 λ 去乘矩阵的每一个子块, 即

$$\lambda\boldsymbol{A}=\lambda\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1s}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2s}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{r1} & \boldsymbol{A}_{r2} & \cdots & \boldsymbol{A}_{rs}\end{pmatrix}=\begin{pmatrix}\lambda\boldsymbol{A}_{11} & \lambda\boldsymbol{A}_{12} & \cdots & \lambda\boldsymbol{A}_{1s}\\ \lambda\boldsymbol{A}_{21} & \lambda\boldsymbol{A}_{22} & \cdots & \lambda\boldsymbol{A}_{2s}\\ \vdots & \vdots & & \vdots\\ \lambda\boldsymbol{A}_{r1} & \lambda\boldsymbol{A}_{r2} & \cdots & \lambda\boldsymbol{A}_{rs}\end{pmatrix}.$$

(3) 分块矩阵的乘法

设 $\boldsymbol{A}$ 为 $m\times l$ 矩阵, $\boldsymbol{B}$ 为 $l\times n$ 矩阵, 分块成

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1t}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2t}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{r1} & \boldsymbol{A}_{r2} & \cdots & \boldsymbol{A}_{rt}\end{pmatrix},\ \boldsymbol{B}=\begin{pmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \cdots & \boldsymbol{B}_{1s}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} & \cdots & \boldsymbol{B}_{2s}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{B}_{t1} & \boldsymbol{B}_{t2} & \cdots & \boldsymbol{B}_{ts}\end{pmatrix},$$

其中 $\boldsymbol{A}_{i1},\boldsymbol{A}_{i2},\cdots,\boldsymbol{A}_{it}\ (i=1,2,\ldots,r)$ 的列数分别等于 $\boldsymbol{B}_{1j},\boldsymbol{B}_{2j},\ldots,\boldsymbol{B}_{tj}$ 的行数, 记

$$\boldsymbol{AB}=\begin{pmatrix}\boldsymbol{C}_{11} & \boldsymbol{C}_{12} & \cdots & \boldsymbol{C}_{1s}\\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} & \cdots & \boldsymbol{C}_{2s}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{C}_{r1} & \boldsymbol{C}_{r2} & \cdots & \boldsymbol{C}_{rs}\end{pmatrix},$$

则

$$\begin{aligned}\boldsymbol{C}_{ij}&=\boldsymbol{A}_{i1}\boldsymbol{B}_{1j}+\boldsymbol{A}_{i2}\boldsymbol{B}_{2j}+\ldots+\boldsymbol{A}_{it}\boldsymbol{B}_{tj}\\ &=\sum_{k=1}^{t}\boldsymbol{A}_{ik}\boldsymbol{B}_{kj}\ (i=1,2,\ldots,r;j=1,2,\ldots,s)\end{aligned}$$

(4) 分块矩阵的转置

设矩阵 A 可以写成分块矩阵

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1t}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2t}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{r1} & \boldsymbol{A}_{r2} & \cdots & \boldsymbol{A}_{rt}\end{pmatrix},$$

则 $\boldsymbol{A}$ 的转置矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 为

$$\boldsymbol{A}^{\mathrm{T}}=\begin{pmatrix}\boldsymbol{A}_{11}^{\mathrm{T}} & \boldsymbol{A}_{21}^{\mathrm{T}} & \cdots & \boldsymbol{A}_{r1}^{\mathrm{T}}\\ \boldsymbol{A}_{12}^{\mathrm{T}} & \boldsymbol{A}_{22}^{\mathrm{T}} & \cdots & \boldsymbol{A}_{r2}^{\mathrm{T}}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{1t}^{\mathrm{T}} & \boldsymbol{A}_{2t}^{\mathrm{T}} & \cdots & \boldsymbol{A}_{rt}^{\mathrm{T}}\end{pmatrix}$$

(5) 分块对角矩阵

设 $\boldsymbol{A}$ 为 n 阶方阵, 若 $\boldsymbol{A}$ 的分块矩阵的非零子块只在主对角线上出现, 其它子块都是零矩阵, 且非零子块都是方阵, 即

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_1 & \boldsymbol{0} & \cdots & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{A}_2 & \cdots & \boldsymbol{0}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{A}_s\end{pmatrix}$$

其中 $A_i\ (i=1,2,\ldots,s)$ 都是方阵, 则称 $\boldsymbol{A}$ 为分块对角矩阵.

(6) 分块对角矩阵的基本运算假设

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_1 & \boldsymbol{0} & \cdots & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{A}_2 & \cdots & \boldsymbol{0}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{A}_s\end{pmatrix},\ \boldsymbol{B}=\begin{pmatrix}\boldsymbol{B}_1 & \boldsymbol{0} & \cdots & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{B}_2 & \cdots & \boldsymbol{0}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{B}_s\end{pmatrix},$$

其中 $\boldsymbol{A}_i, \boldsymbol{B}_i, (i=1,2,\ldots,s)$ 是同阶子方块, 则

$$\boldsymbol{A}+\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A}_1+\boldsymbol{B}_1 & \boldsymbol{0} & \cdots & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{A}_2+\boldsymbol{B}_2 & \cdots & \boldsymbol{0}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{A}_s+\boldsymbol{B}_s\end{pmatrix},\ k\boldsymbol{A}=\begin{pmatrix}k\boldsymbol{A}_1 & \boldsymbol{0} & \cdots & \boldsymbol{0}\\ \boldsymbol{0} & k\boldsymbol{A}_2 & \cdots & \boldsymbol{0}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{0} & \boldsymbol{0} & \cdots & k\boldsymbol{A}_s\end{pmatrix}$$

$$\boldsymbol{A}\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A}_1\boldsymbol{B}_1 & \boldsymbol{0} & \cdots & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{A}_2\boldsymbol{B}_2 & \cdots & \boldsymbol{0}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{A}_s\boldsymbol{B}_s\end{pmatrix},\ \boldsymbol{A}^k=\begin{pmatrix}\boldsymbol{A}_1^k & \boldsymbol{0} & \cdots & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{A}_2^k & \cdots & \boldsymbol{0}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{A}_s^k\end{pmatrix}$$

2.5.2 矩阵按行分块和按列分块

矩阵按行分块和按列分块在矩阵的乘法和线性方程组的讨论中十分重要, 因此我们在这里特别讨论. 任意一个 $m\times n$ 的矩阵 A, 都有 m 行, 我们把每一行称作是矩阵的行向量, 若记第 i 行为

$$\boldsymbol{a}_i^{\mathrm{T}}=(a_{i1},a_{i2},\cdots,a_{in})\ (i=1,2,\ldots,m),$$

则矩阵 A 可以记为

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{a}_1^{\mathrm{T}}\\ \boldsymbol{a}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{a}_m^{\mathrm{T}}\end{pmatrix}$$

任意一个 $m\times n$ 的矩阵 A, 都有 n 列, 我们把每一列称作是矩阵的行向量, 若记第 j 列为

$$\boldsymbol{a}_j=(a_{1j},a_{2j},\cdots,a_{mj})^{\mathrm{T}}\ (j=1,2,\ldots,n),$$

则矩阵 A 可以记为

$$\boldsymbol{A}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n).$$

第三讲 行列式的公理化定义及性质

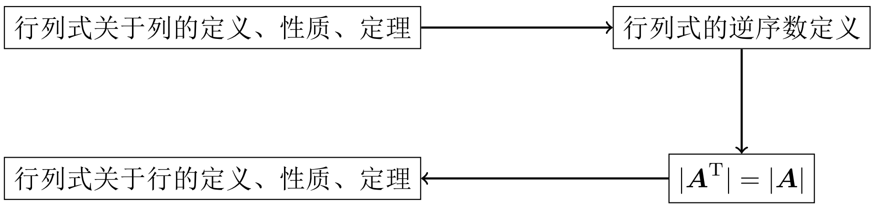
内容提要

(1) 二阶行列式、三阶行列式

(2) n 阶行列式的公理化定义

(3) 行列式的性质

行列式起源于线性方程组的研究, 接下来三节我们将介绍行列式的定义、性质以及应用. 为了方便大家理解, 这里将首先使用公理化的方式定义行列式, 然后再通过公理化定义给出行列式的显式表达式. 这样做的另外一个原因是想告诉大家, 为什么在行列式的显式定义中会出现逆序数这么奇怪的东西.



行列式思维导图

3.1 行列式的公理化定义

我们将以二阶行列式和三阶行列式作为切入点, 考察线性方程组解的性质并由此得出行列式的公理化定义.

3.1.1 二阶行列式

我们首先考察线性方程组

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

对于该线性方程组, 我们可以通过加减消元法得到它的解为（设分母不为零）:

$$\begin{cases} x = \dfrac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}} \\ y = \dfrac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}} \end{cases}. \tag{3.1}$$

为了更好地表示和记忆上面的公式, 我们引入二阶行列式这一记号 (概念).

定义 3.1

设 $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, 则称

$$a_{11}a_{22} - a_{12}a_{21}$$

为对应于二阶矩阵 $\boldsymbol{A}$ 的行列式, 记为 $\det(\boldsymbol{A})$ 或者 $|\boldsymbol{A}|$, 即

$$\det(\boldsymbol{A}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \tag{3.2}$$

我们一般使用对角线法则记忆二阶行列式的公式, 主对角元相乘取正号, 副对角元相乘取负号.

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

采用二阶行列式的记号, 线性方程组的求解公式(3.1)可以改写为:

$$x = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}. \tag{3.3}$$

3.1.2 三阶行列式

对于一般的三元线性方程组:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3, \end{cases}$$

利用消元法, 可求得它的解为（设分母不为零）:

$$x_1 = \frac{b_1a_{22}a_{33} + a_{12}a_{23}b_3 + a_{13}b_2a_{32} - b_1a_{23}a_{32} - a_{12}b_2a_{33} - a_{13}a_{22}b_3}{a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}},$$
$$x_2 = \frac{a_{11}b_2a_{33} + b_1a_{23}a_{31} + a_{13}a_{21}b_3 - a_{11}a_{23}b_3 - b_1a_{21}a_{33} - a_{13}b_2a_{31}}{a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}},$$
$$x_3 = \frac{a_{11}a_{22}b_3 + a_{12}b_2a_{31} + b_1a_{21}a_{32} - a_{11}b_2a_{32} - a_{12}a_{21}b_3 - b_1a_{22}a_{31}}{a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}}.$$

同样地, 为了记忆和书写的方便, 我们引入三阶行列式这一记号.

定义 3.2

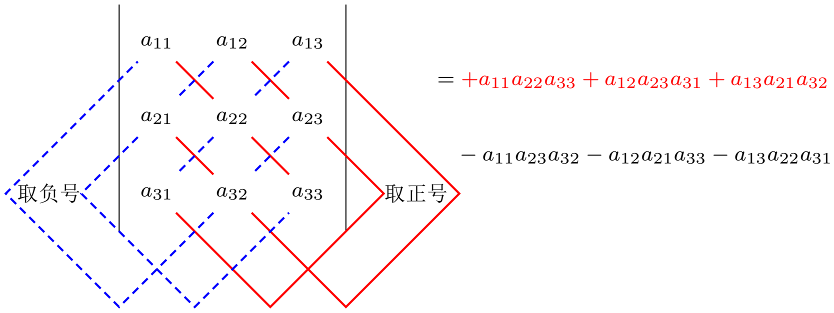
设 $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, 则称

$$\begin{aligned} & a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & -a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \end{aligned} \tag{3.4}$$

为对应于三阶矩阵 $\boldsymbol{A}$ 的行列式, 记为 $\det(\boldsymbol{A})$ 或者 $|\boldsymbol{A}|$, 即

$$\begin{aligned} \det(\boldsymbol{A}) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \end{aligned} \tag{3.5}$$

三阶行列式也有对角线法则, 平行于主对角的三个元素的乘积取正号, 平行于副对角的三个元素的乘积取负号.



采用三阶行列式的记号, 对上面的三元线性方程组, 我们得到: 当 $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \neq 0$ 时, 有

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

即 x_i 的分母是系数矩阵对应的三阶行列式, 而分子恰为将系数矩阵的第 i 列用常数项列向量替换后得到的方阵对应的三阶行列式.

3.1.3 n 阶行列式

继续上面的讨论, 我们希望能够给出一个由 n 个方程构成的 n 元线性方程组的解的表达式, 且该表达式类似于上面 $n = 2$ 和 $n = 3$, 即将解表示为

$$x_i = \frac{n\text{阶行列式}}{n\text{阶行列式}}$$

的形式.

直觉告诉我们, 直接给出 n 阶行列式的表达式是一个非常困难的任务, 因为三阶行列式的表达式就已经很复杂了. 我们可以反过来, 从行列式应该具有的性质入手, 得到行列式的表达式.

设 $\boldsymbol{A}$ 是一个 n 阶方阵, 它的列向量记为 $\boldsymbol{a}_1, \dots \boldsymbol{a}_n$, 即 $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n)$, 为了叙述方便, 我们将 $\boldsymbol{A}$ 的行列式记为

$$|\boldsymbol{A}| = \det(\boldsymbol{A}) = \det(\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n).$$

这里的行列式实际上是矩阵到实数的一个函数 ($R^{n\times n}$ 到 R 的一个函数), 我们希望通过该函数应该满足的法则来唯一确定该函数. 为此, 考察线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{3.6}$$

我们希望线性方程组的解可以表示成如下形式:

$$x_i = \frac{\det(\boldsymbol{A}_i)}{\det(\boldsymbol{A})} \tag{3.7}$$

期中 $\boldsymbol{A}_i$ 是系数矩阵 $\boldsymbol{A}$ 将第 i 列换成 $\boldsymbol{b}$ 之后得到的矩阵. 下面我们来研究一下 $\det(\cdot)$ 应该满足什么样的条件.

假设 $\boldsymbol{x}_1$ 是线性方程组(3.6)唯一的解, 则显然线性方程组 $\boldsymbol{A}\boldsymbol{x} = k\boldsymbol{b}$ 的解应该是 $k\boldsymbol{x}_1$. 这意味着, 如果线性方程组的系数矩阵保持不变, 但常数项变为原来的 k 倍, 那么每个未知数都会变成原先的 k 倍. 因此, 若线性方程组的解可以用(3.7)表示, $\det(\cdot)$ 至少应具有以下性质[1]:

[1] 分别考察 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 和 $\boldsymbol{A}\boldsymbol{x} = k\boldsymbol{b}$ 使用(3.7)求解即可得此结论. 注意到, 两方程组使用(3.7)求解时, 分母是相同的, 分子中的矩阵分别是将 $\boldsymbol{A}$ 的第 i 列变为 $\boldsymbol{b}$ 和 $k\boldsymbol{b}$. 注意, 后一方程组使用(3.7)求解时, 分子应该是前者的 k 倍.

性质 3.1

若矩阵的某一列乘以常数 k, 则新得到的矩阵的行列式是原先矩阵的行列式的 k 倍, 即

$$\det(\boldsymbol{a}_1,\ldots,k\boldsymbol{a}_i,\boldsymbol{a}_n)=k\det(\boldsymbol{a}_1,\ldots,\boldsymbol{a}_i,\boldsymbol{a}_n).$$

假设 $\boldsymbol{x}_1$ 是线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}_1$ 唯一的解, $\boldsymbol{x}_2$ 是线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}_2$ 唯一的解, 则线性方程组 $\boldsymbol{A}\boldsymbol{x}=(\boldsymbol{b}_1+\boldsymbol{b}_2)$ 的解应该是 $\boldsymbol{x}_1+\boldsymbol{x}_2$. 因此, 若线性方程组的解可以用(3.7)表示, $\det(\cdot)$ 至少应具有以下性质:

性质 3.2

若矩阵的某一列可以写成两个列向量的和的形式, 则该矩阵的行列式可以写成两个矩阵的行列式之和的形式, 即

$$\det(\boldsymbol{a}_1,\ldots,\boldsymbol{b}_1+\boldsymbol{b}_2,\boldsymbol{a}_n)=\det(\boldsymbol{a}_1,\ldots,\boldsymbol{b}_1,\boldsymbol{a}_n)+\det(\boldsymbol{a}_1,\ldots,\boldsymbol{b}_2,\boldsymbol{a}_n).$$

假设线性方程组的右端项与第系数矩阵的第 i 列相同, 即线性方程组是

$$\begin{cases} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=a_{1i},\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=a_{2i},\\ \qquad\qquad\ldots\\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n=a_{ni}.\end{cases}$$

若线性方程组存在唯一的解, 则该解必定是

$$\begin{cases} x_k=0,\ (k\neq i)\\ x_k=1\ (k=i).\end{cases}$$

因此, 若线性方程组的解可以用(3.7)表示, $\det(\cdot)$ 至少应具有以下性质:

性质 3.3

若矩阵的两列相同, 则其行列式为零, 即

$$\det(\boldsymbol{a}_1,\ldots\boldsymbol{a}_i,\ldots,\boldsymbol{a}_i,\cdots,\boldsymbol{a}_n)=0.$$

最后, 为了能够唯一确定 $\det(\cdot)$, 我们假设其还具有以下性质[2]

性质 3.4

单位矩阵的行列式为 1, 即 $\det(\boldsymbol{E}_n)=1$.

由3.2和性质3.3, 我们可以得到一个新的性质

性质 3.5

交换行列式的两列, 则行列式的值变为原来的相反数. 即

$$\det(\boldsymbol{a}_1,\ldots\boldsymbol{a}_i,\ldots,\boldsymbol{a}_j,\cdots,\boldsymbol{a}_n)=-\det(\boldsymbol{a}_1,\ldots\boldsymbol{a}_j,\ldots,\boldsymbol{a}_i,\cdots,\boldsymbol{a}_n).$$

证明 由性质3.3和性质3.2可得

$$\begin{aligned} 0&=\det(\boldsymbol{a}_1,\ldots\boldsymbol{a}_i+\boldsymbol{a}_j,\ldots,\boldsymbol{a}_i+\boldsymbol{a}_j,\cdots,\boldsymbol{a}_n)\\ &=\det(\boldsymbol{a}_1,\ldots\boldsymbol{a}_i,\ldots,\boldsymbol{a}_i,\cdots,\boldsymbol{a}_n)+\det(\boldsymbol{a}_1,\ldots\boldsymbol{a}_i,\ldots,\boldsymbol{a}_j,\cdots,\boldsymbol{a}_n)\\ &+\det(\boldsymbol{a}_1,\ldots\boldsymbol{a}_j,\ldots,\boldsymbol{a}_i,\cdots,\boldsymbol{a}_n)+\det(\boldsymbol{a}_1,\ldots\boldsymbol{a}_j,\ldots,\boldsymbol{a}_j,\cdots,\boldsymbol{a}_n)\\ &=\det(\boldsymbol{a}_1,\ldots\boldsymbol{a}_i,\ldots,\boldsymbol{a}_j,\cdots,\boldsymbol{a}_n)+\det(\boldsymbol{a}_1,\ldots\boldsymbol{a}_j,\ldots,\boldsymbol{a}_i,\cdots,\boldsymbol{a}_n)\end{aligned}$$

性质得证. □

[2]因为分子分母同时乘以一个倍数, 比值保持不变

定理 3.1

若 $\det(\cdot)$ 满足性质3.2, 则性质3.3和性质3.5是等价的.

证明 由性质3.5的证明, 可知, 性质3.2和性质3.3可以得到性质3.5, 下证性质3.2和性质3.5可以得到性质3.3.

由性质3.5可得,

$$\det(\boldsymbol{a}_1, \ldots \boldsymbol{a}_i, \ldots, \boldsymbol{a}_i, \cdots, \boldsymbol{a}_n) = -\det(\boldsymbol{a}_1, \ldots \boldsymbol{a}_i, \ldots, \boldsymbol{a}_i, \cdots, \boldsymbol{a}_n),$$

于是

$$\det(\boldsymbol{a}_1, \ldots \boldsymbol{a}_i, \ldots, \boldsymbol{a}_i, \cdots, \boldsymbol{a}_n) = 0.$$

定理得证. □

由于在计算行列式时, 性质3.5比性质3.3使用起来更为频繁和方便, 因此, 我们更愿意使用性质3.5来定义行列式.

定义 3.3 (行列式的公理化定义 1)

假设 $\boldsymbol{A}$ 是 n 阶方阵, 其列向量记为 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$, 则 n 阶行列是满足性质3.1、性质3.2、性质3.5和性质3.4的函数. 即行列式满足以下性质:

(1) 若某一列乘以 k, 则行列式变为原来的 k 倍.

(2) 若某一列的每个元素都可以写成两项之和的形式, 则该行列式可以写成两个行列式之和的形式.

(3) 交换行列式的两列, 其值反号.

(4) 单位矩阵的行列式为 1.

我们注意到, 线性方程组在求解的时候, 进行初等行变换不会改变线性方程组的解. 同样地, 如果我们将定义3.3中的列全部改成行, 那么按此定义的两行列式的比值也不会发生改变. 后面我们将证明, 行列式在列上所具有的性质在行上也同样具有 (见定理4.4). 定义3.3中的列全部改成行, 得到定义与定义3.3实际上是等价的.

定义 3.4 (行列式的公理化定义 2)

假设 $\boldsymbol{A}$ 是 n 阶方阵, 其行向量记为 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$, 则 n 阶行列是满足性质3.1、性质3.2、性质3.5和性质3.4的函数 (注意将上述提及的性质中的 “列” 全部换为 “行”). 即行列式满足以下性质:

(1) 若某一行乘以 k, 则行列式变为原来的 k 倍.

(2) 若某一行的每个元素都可以写成两项之和的形式, 则该行列式可以写成两个行列式之和的形式.

(3) 交换行列式的两行, 其值反号.

(4) 单位矩阵的行列式为 1.

从定义3.3和定义3.4, 我们不难看出, 行列式可以理解为对 n 维空间中, 列 (行) 向量所围成的几何体的体积.

例题 3.1 设 $\boldsymbol{A}$ 是一个 n 阶方阵, 则 $|a\boldsymbol{A}| = a^n|\boldsymbol{A}|$.

证明 每一列提取一个 a, 得 $|a\boldsymbol{A}| = a^n|\boldsymbol{A}|$. □

3.2 行列式的性质

接下来, 我们从定义3.3出发, 得到更多行列式的性质, 并最终推导出行列式的显示表达式. **注意, 下列所有关于行列式的性质、定理, 将 “行” 改成 “列”, 仍然成立. 后面, 我们将证明定义3.3和定义3.4是等价的 (见定理4.4), 而与下面对应的关于 “行” 的性质、定理可以类似地由定义3.4得到.**

性质 3.6

若行列式有一列 (行) 为零向量, 则行列式的值等于 0.

证明 由于行列式满足3.1, 设行列式第 i 列为零向量, 因此

$$\begin{aligned}\det(\boldsymbol{a}_1,\dots,0,\dots,\boldsymbol{a}_n)&=\det(\boldsymbol{a}_1,\dots,0\cdot\boldsymbol{a}_i,\dots,\boldsymbol{a}_n)\\&=0\cdot\det(\boldsymbol{a}_1,\dots,\boldsymbol{a}_i,\dots,\boldsymbol{a}_n)\\&=0\end{aligned}$$

结论得证. □

性质 3.7

若行列式有两列 (行) 元素成比例, 则行列式的值等于 0.

证明 由于行列式满足性质3.1, 设行列式第 i 列和第 j 列元素成比例, $\boldsymbol{a}_i=k\boldsymbol{a}_j$, 因此

$$\begin{aligned}\det(\boldsymbol{a}_1,\dots,\boldsymbol{a}_i,\dots,\boldsymbol{a}_j,\dots,\boldsymbol{a}_n)&=\det(\boldsymbol{a}_1,\dots,k\boldsymbol{a}_j,\dots,\boldsymbol{a}_j,\dots,\boldsymbol{a}_n)\\&=k\det(\boldsymbol{a}_1,\dots,\boldsymbol{a}_j,\dots,\boldsymbol{a}_j,\dots,\boldsymbol{a}_n)\\&=0\end{aligned}$$

其中最后一个等号用到了性质3.3. □

性质 3.8

对行列式做倍加列 (行) 变换, 行列式的值不变.

证明 根据性质3.2以及性质3.7, 我们可以得到

$$\begin{aligned}\det(\boldsymbol{a}_1,\dots,\boldsymbol{a}_i+k\boldsymbol{a}_j,\dots,\boldsymbol{a}_j,\dots,\boldsymbol{a}_n)&=\det(\boldsymbol{a}_1,\dots,\boldsymbol{a}_i,\dots,\boldsymbol{a}_j,\dots,\boldsymbol{a}_n)\\&\quad+\det(\boldsymbol{a}_1,\dots,k\boldsymbol{a}_j,\dots,\boldsymbol{a}_j,\dots,\boldsymbol{a}_n)\\&=\det(\boldsymbol{a}_1,\dots,\boldsymbol{a}_i,\dots,\boldsymbol{a}_j,\dots,\boldsymbol{a}_n)+0\\&=\det(\boldsymbol{a}_1,\dots,\boldsymbol{a}_i,\dots,\boldsymbol{a}_j,\dots,\boldsymbol{a}_n).\end{aligned}$$

第一个和第二个等号分别用到了性质3.2以及性质3.7. □

性质 3.9

对角矩阵的行列式为对角线元素的乘积, 即

$$\begin{vmatrix}a_1&&\\&\ddots&\\&&a_n\end{vmatrix}=a_1a_2\dots a_n.$$

证明 由性质3.1和性质3.4, 易得

$$\begin{vmatrix}a_1&&\\&\ddots&\\&&a_n\end{vmatrix}=a_1a_2\dots a_n\begin{vmatrix}1&&\\&\ddots&\\&&1\end{vmatrix}=a_1a_2\dots a_n.$$

□

性质 3.10

上三角形矩阵和下三角形矩阵的行列式等于对角线元素的乘积, 即

证明 我们以下三角形矩阵为例做证明, 上三角形矩阵的证明类似, 留作练习.

若矩阵的对角线元素全都不是 0, 则由性质3.8显然有

$$\begin{vmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} = \dots = \begin{vmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{vmatrix}.$$

由性质3.9可知结论成立.

若矩阵对角元元素至少有一个位 0, 不妨设对角线上最后一个为 0 的元素是 a_{ii}, 则由性质3.8有

$$\begin{vmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} = \dots = \begin{vmatrix} a_{11} & & & 0 & & & \\ a_{21} & a_{22} & & 0 & & & \\ \vdots & \vdots & \ddots & \vdots & & & \\ a_{i1} & \dots & 0 & 0 & & & \\ 0 & \dots & 0 & 0 & \ddots & & \\ \vdots & \vdots & \vdots & \vdots & & \ddots & \\ 0 & 0 & \dots & 0 & \dots & 0 & a_{nn} \end{vmatrix}.$$

注意到第 i 列的元素全部为 0, 由性质3.6易得行列式为 0, 而对角线元素乘积也是 0, 因此命题成立. □

由性质3.10, 可知, 我们在计算行列式时, 只需要对矩阵进行初等变换, 将矩阵转化为上三角形矩阵或者下三角形矩阵, 即可方便地计算任意矩阵的行列式.

例题 3.2 求 $\begin{vmatrix} 1 & 2 & 0 & 1 \\ 2 & 0 & -1 & 1 \\ -1 & 2 & 3 & 0 \\ 1 & 1 & -1 & 2 \end{vmatrix}$.

解 如只使用列变换

$$\begin{vmatrix} 1 & 2 & 0 & 1 \\ 2 & 0 & -1 & 1 \\ -1 & 2 & 3 & 0 \\ 1 & 1 & -1 & 2 \end{vmatrix} \xlongequal[c_4-c_1]{c_2-2c_1} \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & -4 & -1 & -1 \\ -1 & 4 & 3 & 1 \\ 1 & -1 & -1 & 1 \end{vmatrix} \xlongequal{c_2 \leftrightarrow c_4} - \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & -1 & -1 & -4 \\ -1 & 1 & 3 & 4 \\ 1 & 1 & -1 & -1 \end{vmatrix}$$

$$\xlongequal[c_4-4c_2]{c_3-c_2} - \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ -1 & 1 & 2 & 0 \\ 1 & 1 & -2 & -5 \end{vmatrix} = -10.$$

如只使用行变换

$$\begin{vmatrix} 1 & 2 & 0 & 1 \\ 2 & 0 & -1 & 1 \\ -1 & 2 & 3 & 0 \\ 1 & 1 & -1 & 2 \end{vmatrix} \xlongequal[r_4-r_1]{\substack{r_2-2r_1\\ r_3+r_1}} \begin{vmatrix} 1 & 2 & 0 & 1 \\ 0 & -4 & -1 & -1 \\ 0 & 4 & 3 & 1 \\ 0 & -1 & -1 & 1 \end{vmatrix} \xlongequal{r_2+r_3} \begin{vmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 4 & 3 & 1 \\ 0 & -1 & -1 & 1 \end{vmatrix}$$

$$\xlongequal{r_2 \leftrightarrow r_4} - \begin{vmatrix} 1 & 2 & 0 & 1 \\ 0 & -1 & -1 & 1 \\ 0 & 4 & 3 & 1 \\ 0 & 0 & 2 & 0 \end{vmatrix} \xlongequal{r_3+4r_2} - \begin{vmatrix} 1 & 2 & 0 & 1 \\ 0 & -1 & -1 & 1 \\ 0 & 0 & -1 & 5 \\ 0 & 0 & 2 & 0 \end{vmatrix}$$

$$\xlongequal{r_4+2r_3} - \begin{vmatrix} 1 & 2 & 0 & 1 \\ 0 & -1 & -1 & 1 \\ 0 & 0 & -1 & 5 \\ 0 & 0 & 0 & 10 \end{vmatrix} = -10.$$

两种方式所得结果相同. 事实上, 我们可以按照行列式的定义和性质, 自由的进行行变换和列变换.

✍ **练习 3.1** 已知 $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & 4 \\ 6 & 0 & 5 \end{pmatrix}$, 求 $|4A|$. (参考答案：$4^3 * 63$)

性质 3.11

若 $\boldsymbol{D} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{C} \end{pmatrix}$, 则 $\det(\boldsymbol{D}) = \det(\boldsymbol{A})\det(\boldsymbol{C})$, 其中 $\boldsymbol{O}$ 是零矩阵.

证明 假设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 分别是 p 阶和 q 阶方阵, 我们要计算 $\boldsymbol{D}$ 的行列式. 用行初等变换将 $\boldsymbol{D}$ 上三角化, 可以分别将 $\boldsymbol{A}$ 和 $\boldsymbol{C}$ 上三角化. 具体地, 设

$$\boldsymbol{A} \xrightarrow{\text{初等列变换}} \begin{pmatrix} a_{11} & \cdots & * \\ & \ddots & \vdots \\ & & a_{pp} \end{pmatrix}, \quad M = \text{初等变换中交换两列的次数},$$

$$\boldsymbol{B} \xrightarrow{\text{初等列变换}} \begin{pmatrix} c_{11} & \cdots & * \\ & \ddots & \vdots \\ & & c_{qq} \end{pmatrix}, \quad N = \text{初等变换中交换两列的次数}.$$

在 $\boldsymbol{D}$ 中, 分别对 $\boldsymbol{A}$ 和 $\boldsymbol{C}$ 所在的列做上面一样的行初等变换, 得到

$$\boldsymbol{D} \xrightarrow{\text{初等列变换}} \begin{pmatrix} a_{11} & \cdots & * & & & \\ & \ddots & \vdots & & & \\ & & a_{pp} & & & \\ & & & c_{11} & \cdots & * \\ & & & & \ddots & \vdots \\ & & & & & c_{qq} \end{pmatrix}$$

则在上述初等列变换中, 交换两列的次数是 $M+N$. 所以有

$$
\begin{aligned}
|\boldsymbol{D}| &= (-1)^{M+N}\begin{vmatrix} a_{11} & \cdots & * & & & \\ & \ddots & \vdots & & & \\ & & a_{pp} & & & \\ & & & c_{11} & \cdots & * \\ & & & & \ddots & \vdots \\ & & & & & c_{qq} \end{vmatrix} \\
&= (-1)^{M+N} a_{11}\cdots a_{pp} c_{11}\cdots c_{qq} \\
&= (-1)^{M} a_{11}\cdots a_{pp}\cdot(-1)^{N} c_{11}\cdots c_{qq} = |A||C|.
\end{aligned}
$$

□

同理可得

性质 3.12

若 $\boldsymbol{D}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{B} & \boldsymbol{C}\end{pmatrix}$, 则 $\det(\boldsymbol{D})=\det(\boldsymbol{A})\det(\boldsymbol{C})$, 其中 $\boldsymbol{O}$ 是零矩阵.

注 如果 $\boldsymbol{A}_1,\boldsymbol{A}_2,\ldots,\boldsymbol{A}_s$ 是方阵, 则有

$$
\begin{vmatrix} \boldsymbol{A}_1 & * & \cdots & * \\ & \boldsymbol{A}_2 & \cdots & * \\ & & \ddots & \vdots \\ & & & \boldsymbol{A}_s \end{vmatrix} = |\boldsymbol{A}_1||\boldsymbol{A}_2|\cdots|\boldsymbol{A}_s|, \begin{vmatrix} \boldsymbol{A}_1 & & & \\ * & \boldsymbol{A}_2 & & \\ \vdots & \vdots & \ddots & \\ * & * & \cdots & \boldsymbol{A}_s \end{vmatrix} = |\boldsymbol{A}_1||\boldsymbol{A}_2|\cdots|\boldsymbol{A}_s|,
$$

特别地,

$$
\begin{vmatrix} \boldsymbol{A}_1 & & \\ & \ddots & \\ & & \boldsymbol{A}_s \end{vmatrix} = |\boldsymbol{A}_1||\boldsymbol{A}_2|\cdots|\boldsymbol{A}_s|.
$$

证明 由性质3.11和性质3.12易得. □

性质 3.13

设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是 n 阶方阵, 则 $\det(\boldsymbol{AB})=\det(\boldsymbol{A})\det(\boldsymbol{B})$.

证明 设 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ -\boldsymbol{E} & \boldsymbol{B}\end{pmatrix}$, 则由性质3.12, $|\boldsymbol{C}|=|\boldsymbol{A}||\boldsymbol{B}|$. 另一方面, 对 $\boldsymbol{C}$ 进行如下初等变换

$$\boldsymbol{C}=\begin{pmatrix} a_{11} & \cdots & a_{1n} & & & \\ \vdots & & \vdots & & & \\ a_{n1} & \cdots & a_{nn} & & & \\ -1 & & & b_{11} & \cdots & b_{1n} \\ & \ddots & & \vdots & & \vdots \\ & & -1 & b_{n1} & \cdots & b_{nn} \end{pmatrix} \xrightarrow{c_{n+1}+b_{11}c_1+\cdots+b_{n1}c_n}$$

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} & \sum_{k=1}^n a_{1k}b_{k1} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} & \sum_{k=1}^n a_{nk}b_{k1} & 0 & \cdots & 0 \\ -1 & & & 0 & b_{12} & \cdots & b_{1n} \\ & \ddots & & \vdots & \vdots & & \vdots \\ & & -1 & 0 & b_{n2} & \cdots & b_{nn} \end{pmatrix} \xrightarrow[j=2,\ldots,n]{c_{n+j}+b_{1j}c_1+\cdots+b_{nj}c_n}$$

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} & \\ \vdots & & \vdots & \boldsymbol{AB} \\ a_{n1} & \cdots & a_{nn} & \\ -1 & & & \\ & \ddots & & \\ & & -1 & \end{pmatrix} \xrightarrow[i=1,\ldots,n]{r_i\leftrightarrow r_{n+i}} \begin{pmatrix} -1 & & & \\ & \ddots & & \\ & & -1 & \\ a_{11} & \cdots & a_{1n} & \\ \vdots & & \vdots & \boldsymbol{AB} \\ a_{n1} & \cdots & a_{nn} & \end{pmatrix}$$

由于上面共用了 n 个调行变换, 所以

$$|\boldsymbol{C}|=(-1)^n\begin{vmatrix}-\boldsymbol{E} & \boldsymbol{O}\\ \boldsymbol{A} & \boldsymbol{AB}\end{vmatrix}=(-1)^n|-\boldsymbol{E}||\boldsymbol{AB}|=|\boldsymbol{AB}|$$

最后就得到 $|\boldsymbol{AB}|=|\boldsymbol{A}||\boldsymbol{B}|$. □

性质 3.14

若 $\boldsymbol{A}$ 可逆, 则有 $|\boldsymbol{A}^{-1}|\neq 0$ 且 $|\boldsymbol{A}^{-1}|=1/|\boldsymbol{A}|$.

证明 由条件可知, $\boldsymbol{AA}^{-1}=\boldsymbol{E}$, 两边取行列式可得 $|\boldsymbol{AA}^{-1}|=\boldsymbol{E}$. 由性质3.13可知,

$$|\boldsymbol{A}||\boldsymbol{A}^{-1}|=1,$$

结论显然成立. □

注 实际上, 可以证明：$\boldsymbol{A}$ 可逆 $\Longleftrightarrow$ 有 $|\boldsymbol{A}|\neq 0$(见定理6.2).

第四讲 行列式的逆序数定义

内容提要

(1) 逆序数和行列式的逆序数定义

(2) 行列式的递归定义 (拉普拉斯展开)

在本节, 我们首先给出行列式的逆序数定义, 然后再给出行列式的拉普拉斯展开.

4.1 行列式的逆序数定义

我们以二阶行列式来引入行列式的逆序数定义. 对于二阶行列式, 我们有

$$\begin{aligned}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} &= \begin{vmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \\ &= \begin{vmatrix} a_{11} & 0 \\ 0 & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} \\ 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} \\ a_{21} & 0 \end{vmatrix}.\end{aligned}$$

注意上式中, 最右边的第二和第三个行列式中, 一行是另一行的倍数, 因此这两个式子为 0. 从而我们有

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 \\ 0 & a_{22} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} \\ a_{21} & 0 \end{vmatrix}$$

根据行列式的定义, 我们有

$$\begin{vmatrix} a_{11} & 0 \\ 0 & a_{22} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} \\ a_{21} & 0 \end{vmatrix} = a_{11}a_{22}\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} - a_{12}a_{21}\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

类似地, 对于三阶行列式, 每一列都可以拆分成

$$\begin{pmatrix} a_{1j} \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ a_{2j} \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ a_{3j} \end{pmatrix}$$

三种情况, 因此我们可以将 $\boldsymbol{A}$ 的行列式表示成 $3 \times 3 = 9$ 个行列式的和的形式. 注意到上述 9 个行列式中, 如果某个行列式的两列的非零元位于同一行 (因为两列成比例了), 则该行列式的值为 0, 因此, 只有非零元位于不同行不同列的行列式的值才有可能不为 0. 这样的行列式一共有 6 个, 即

$$\begin{aligned}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{32} \end{vmatrix} &= \begin{vmatrix} a_{11} & & \\ & a_{22} & \\ & & a_{33} \end{vmatrix} + \begin{vmatrix} & & a_{13} \\ a_{21} & & \\ & a_{32} & \end{vmatrix} + \begin{vmatrix} & a_{12} & \\ & & a_{23} \\ a_{31} & & \end{vmatrix} \\ &+ \begin{vmatrix} a_{11} & & \\ & & a_{23} \\ & a_{32} & \end{vmatrix} + \begin{vmatrix} & a_{12} & \\ a_{21} & & \\ & & a_{33} \end{vmatrix} + \begin{vmatrix} & & a_{13} \\ & a_{22} & \\ a_{31} & & \end{vmatrix}.\end{aligned} \tag{4.1}$$

注意到, 每一列的行指标实际上就是 $1, 2, 3$ 的一个排列, 因此我们可以使用排列来表示后面的 6 个行列式[1]. 如第一个行列式可以用 $[1, 2, 3]$ 表示, 第二个行列式可以用 $[2, 3, 1]$ 表示. 上式中, 所有的行指标为

$$\text{行指标} = [1, 2, 3], [2, 3, 1], [3, 1, 2], [1, 3, 2], [2, 1, 3], [3, 2, 1].$$

---

[1]注意, $3! = 6$.

公式(4.1)可以进一步改写为

$$\begin{aligned}|\boldsymbol{A}| = a_{11}a_{22}a_{33}\begin{vmatrix}1&&\\&1&\\&&1\end{vmatrix} + a_{21}a_{32}a_{13}\begin{vmatrix}&&1\\1&&\\&1&\end{vmatrix} + a_{31}a_{12}a_{23}\begin{vmatrix}&1&\\&&1\\1&&\end{vmatrix}\\ + a_{11}a_{32}a_{23}\begin{vmatrix}1&&\\&&1\\&1&\end{vmatrix} + a_{21}a_{12}a_{33}\begin{vmatrix}&1&\\1&&\\&&1\end{vmatrix} + a_{31}a_{22}a_{13}\begin{vmatrix}&&1\\&1&\\1&&\end{vmatrix}.\end{aligned} \tag{4.2}$$

上式中的每一个置换矩阵[2]的行列式都可以通过交换两列的操作变为单位矩阵的行列式. 后面三个行列式需要进行奇数次列交换（一次）, 前面三个行列式需要进行偶数次列交换（零次或者两次）. 例如, 若行指标是 $[3,1,2]$, 其对应的(4.2)中的置换矩阵进行两次行交换即可以变为单位矩阵, 因此其对应的行列式的值为 $a_{31}a_{12}a_{23}$ 乘以 $(-1)^2$. 其余的也类似.

注意到, 通过列交换将置换矩阵变为单位矩阵, 等价于将行指标变为从小到大的顺序排列. 我们在这里举例说明, 将 $[3,1,2]$ 变为从小到大排列的指标, 需要进行如下两次变换:

$$[3,1,2] \to [1,3,2] \to [1,2,3].$$

对于 n 阶方阵, 根据同样的分析, 我们有类似的表达式. 为了给出该表达式, 我们首先引入一些必要的术语.

定义 4.1

正整数数 $1,2,\dots,n$ 按照任意次序排成的有序数组称为一个 n 元排列, 记为 $[p_1,p_2,\dots,p_n]$. 所有 n 元排列做成的集合记为 S_n. 特别地, 称 $[1,2,\dots,n]$ 为自然排列.

容易知道 S_n 中元素的个数为 $n!$, 这意味着 n 元排列一共有 $n!$ 个. 为了研究需要经过奇数次还是偶数次行交换才可以将上面叙述中的置换矩阵变成单位矩阵, 我们引入逆序数和奇排列、偶排列的概念.

定义 4.2

在一个排列中, 如果两个数中前者大于后者, 则称这两个数构成一个逆序. 一个排列中所含逆序数的总数称为该排列的逆序数. 排列 $[p_1,p_2,\dots,p_n]$ 的逆序数记为 $\tau([p_1,p_2,\dots,p_n])$. 逆序数为偶数的排列称为偶排列, 逆序数为奇数的排列称为奇排列.

在一个排列中, 对调其中的两个数, 而其余的数保持不变, 就可以得到一个新的排列. 对排列所作的上述变换称为对换.

定理 4.1

对换改变排列的奇偶性, 即作一次对换, 奇排列变为偶排列, 偶排列变为奇排列.

证明 这里仅仅给出证明要点：分两步证明, 先证明相邻两个数交换顺序改变排列的奇偶性. 对换可以看做是相邻对换进行多次操作. □

我们前面提到过, 置换矩阵通过列变换变为单位矩阵需要进行的列变换的次数, 决定了展开式中某一项的符号. 置换矩阵与行指标构成的排列是一一对应的, 因此, 对置换矩阵进行一次列变换等价于对行指标构成的排列进行一次对换. 于是置换矩阵通过列变换变为单位矩阵需要进行的列变换的次数, 等于将行指标构成的排列通过对换变为自然排列所需要的对换的次数. 于是, 我们就得到了行列式展开的逆序数表达式：

$$|\boldsymbol{A}| = \sum_{[p1,\dots,p_n]\in S_n} (-1)^{\tau([p_1,\dots,p_n])} a_{p_1 1}a_{p_2 2}\dots a_{p_n n}.$$

事实上, 我们并不关心对换次数的具体数值, 我们只关心对换次数的奇偶性. 注意到自然排列是一个偶排列, 因此,

[2]单位经过调列 (行) 变换得到的矩阵称为置换矩阵

若一个排列是偶排列, 则通过对换将其变为自然排列所需要的对换的次数是偶数; 若一个排列是奇排列, 则通过对换将其变为自然排列所需要的对换的次数是奇数[3].

至此, 我们可以给出行列式的逆序数定义

定义 4.3 (行列式的逆序数定义)

设 $\boldsymbol{A}=(a_{ij})_{n\times n}$, 我们称

$$\sum_{[p_1,\ldots,p_n]\in S_n}(-1)^{\tau([p_1,\ldots,p_n])}a_{p_1 1}a_{p_2 2}\ldots a_{p_n n}$$

为矩阵 $\boldsymbol{A}$ 的行列式, 记为 $|\boldsymbol{A}|$, 或者 $\det(\boldsymbol{A})$.

通过行列式的逆序数定义, 我们可以得到行列式的更多性质.

定理 4.2

设 $\boldsymbol{A}=(a_{ij})_{n\times n}$, 则有

$$|\boldsymbol{A}|=\sum_{[q_1,\ldots,q_n]\in S_n}(-1)^{\tau([q_1,\ldots,q_n])}a_{1q_1}a_{2q_2}\ldots a_{nq_n}.$$

证明 仅给出证明要点, 考虑逆序数表达式中的项

$$(-1)^{\tau([p_1,\ldots,p_n])}a_{p_1 1}a_{p_2 2}\ldots a_{p_n n}=(-1)^{\tau([p_1,\ldots,p_n])+\tau([1,2,\ldots,n])}a_{p_1 1}a_{p_2 2}\ldots a_{p_n n},$$

若将上式中的两个元素 $a_{p_s s}$ 与 $a_{p_t t}$ 进行对换. 注意到, 新的下标的行排列和列排列是原下标的行排列和列排列各进行了一次对换, 因此, 两个下标行排列和列排列的逆序数之和的奇偶性保持不变. 通过对换, 可以将行排列变为自然排列, 假设此时行指标为 $[q_1,\ldots,q_n]$, 则有

$$(-1)^{\tau([p_1,\ldots,p_n])+\tau([1,2,\ldots,n])}a_{p_1 1}a_{p_2 2}\ldots a_{p_n n}=(-1)^{\tau([1,2,\ldots,n])+\tau([q_1,\ldots,q_n])}a_{1q_1}a_{2q_2}\ldots a_{nq_n}.$$

因此

$$|\boldsymbol{A}|=\sum_{[p_1,\ldots,p_n]\in S_n}(-1)^{\tau([p_1,\ldots,p_n])}a_{1p_1}a_{2p_2}\ldots a_{np_n}.$$

□

由定理4.2的讨论, 我们还可以得到下面的定理.

定理 4.3

设 $\boldsymbol{A}=(a_{ij})_{n\times n}$, 则 $|\boldsymbol{A}|$ 展开项中含有 $a_{p_1q_1}a_{p_2q_2}\ldots a_{p_nq_n}$ 的项的符号是

$$(-1)^{\tau([p_1,\ldots,p_n])+\tau([q_1,\ldots,q_n])},$$

其中 $a_{p_1q_1},a_{p_2q_2},\ldots,a_{p_nq_n}$ 是矩阵 $\boldsymbol{A}$ 位于不同行不同列的元素.

性质 4.1

矩阵的行列式与其转置的行列式相同, 即 $|\boldsymbol{A}|=|\boldsymbol{A}^{\mathrm{T}}|$.

证明 设 $\boldsymbol{A}^{\mathrm{T}}=(b_{ij})_{n\times n}$, 则 $b_{ij}=a_{ji}$. 于是由定理4.2

$$\begin{aligned}\det\left(\boldsymbol{A}^{\mathrm{T}}\right)&=\sum_{[p_1,p_2,\ldots,p_n]\in S_n}(-1)^{\tau([p_1,p_2,\ldots,p_n])}b_{1p_1}b_{2p_2}\cdots b_{np_n}\\&=\sum_{[p_1,p_2,\ldots,p_n]\in S_n}(-1)^{([p_1,p_2,\ldots,p_n])}a_{p_1 1}a_{p_2 2}\cdots a_{p_n n}\\&=\det(\boldsymbol{A}).\end{aligned}$$

□

[3]这段分析实质上给出了为什么按照定义3.3定义的行列式是唯一的

下面, 我们证明上一节遗留的一个问题.

定理 4.4

将行之前行列式中定义、性质中的所有行改成列, 仍然成立.

证明 因为 $|\boldsymbol{A}^{\mathrm{T}}|=|\boldsymbol{A}|$, 而对 $|\boldsymbol{A}|$ 作行变换, 相当于对 $\boldsymbol{A}^{\mathrm{T}}$ 作相应的列变换, 因此定理得证. □

4.2 行列式按行、按列展开

为了给出计算行列式的另外一种方法, 即行列式的递归算法, 我们首先需要引入余子式和代数余子式的概念:

定义 4.4

设 $\boldsymbol{A}$ 是一个 n 阶方阵, 将元素 a_{ij} 所在的第 i 行和第 j 列的所有元素划去后, 剩下的 $n-1$ 阶方阵的行列式称为元素 a_{ij} 的余子式, 记作 M_{ij}, 并把数 $A_{ij}=(-1)^{i+j}M_{ij}$ 称为元素 a_{ij} 的代数余子式.

注意, 虽然余子式和代数余子式在名称中含有式, 但实际上它们是一个数. 实际上行列式也称为"式", 但这些"式"只是形状上有个形式, 实际上只是一个数.

例题 4.1 根据代数余子式的定义, 计算行列式 $\begin{vmatrix} 2 & 1 & 3 \\ -1 & 0 & 2 \\ 1 & 5 & -2 \end{vmatrix}$ 每个元素的余子式和代数余子式.

解 我们只举一个例子, 第二行第一列元素 -1 的余子式和代数余子式. 根据定义, 它的余子式是去掉第二行和第一列所有元素剩余的二阶行列式

$$\begin{vmatrix} 1 & 3 \\ 5 & -2 \end{vmatrix} = -17,$$

因此它的代数余子式是 $A_{21}=(-1)^{2+1}(-17)=17$. 读者可以自行计算其他元素的余子式和代数余子式.

接下来我们便可以给行列式的拉普拉斯展开公式 (递归算法):

定理 4.5

设 $\boldsymbol{A}$ 是一个 n 阶方阵, 则

$$|\boldsymbol{A}|=\sum_{k=1}^{n}a_{kj}A_{kj}=a_{1j}A_{1j}+a_{2j}A_{2j}+\cdots+a_{nj}A_{nj},\qquad j=1,2,\ldots,n; \tag{4.3}$$

$$|\boldsymbol{A}|=\sum_{k=1}^{n}a_{ik}A_{ki}=a_{i1}A_{i1}+a_{i2}A_{i2}+\cdots+a_{in}A_{in},\qquad i=1,2,\ldots,n. \tag{4.4}$$

其中(4.3)是按照第 j 列展开, (4.4)是按照第 i 列展开.

证明 这里给出证明要点, 首先证明下面的引理

引理

如果矩阵 $\boldsymbol{A}$ 的第 j 列除了 a_{ij} 以外的其余元素全为 0, 则 $|\boldsymbol{A}|=a_{ij}A_{ij}$.

上述引理证明比较简单, 只需要将 a_{ij} 经过初等行变换和列变换变到第一行第一列即可. 引理证明结束后, 利用性质3.2和3.5即可得出结论. □

由于有两行相同, 或者两列相同的矩阵的行列式为 0. 因此我们有如下推论:

推论 4.1

设 $\boldsymbol{A}$ 是一个 n 阶方阵，则

$$\sum_{k=1}^{n} a_{kj}A_{ks} = a_{1j}A_{1s} + a_{2j}A_{2s} + \cdots + a_{nj}A_{ns} = 0,\ (j \neq s);$$

$$\sum_{k=1}^{n} a_{ik}A_{sk} = a_{i1}A_{s1} + a_{i2}A_{s2} + \cdots + a_{in}A_{sn} = 0,\ (i \neq s).$$

证明 我们以三阶行列式做说明，

$$\begin{aligned}
a_{21}A_{21} + a_{22}A_{22} + a_{23}A_{23} &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
\Longrightarrow uA_{21} \quad + \quad vA_{22} \quad + \quad wA_{23} &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ u & v & w \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
\Longrightarrow a_{11}A_{21} + a_{12}A_{22} + a_{13}A_{23} &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = 0
\end{aligned}$$

对于更高阶的行列式，证明方式类似. □

第五讲 行列式的应用及计算

内容提要

(1) 利用行列式求解线性方程组 (克莱姆法则)
(2) 利用行列式计算体积 (行列式的几何意义)
(3) 行列式计算举例

5.1 克莱姆法则

现在我们来回到我们最初的问题, 我们希望得到一个公式, 能够给出线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

的解. 对于系数矩阵为 n 阶方阵的问题, 我们这里可以给出答案.

例题 5.1 我们以线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3, \end{cases}$$

为例, 使用矩阵和行列式等工具, 给出线性方程组解的表达式.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 & 0 & 0 \\ x_2 & 1 & 0 \\ x_3 & 0 & 1 \end{pmatrix} = \begin{pmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{pmatrix}$$

两边取行列式我们有

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{vmatrix} x_1 & 0 & 0 \\ x_2 & 1 & 0 \\ x_3 & 0 & 1 \end{vmatrix} = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}$$

注意到

$$\begin{vmatrix} x_1 & 0 & 0 \\ x_2 & 1 & 0 \\ x_3 & 0 & 1 \end{vmatrix} = 1,$$

若系数矩阵的行列式不为 0, 我们有,

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}$$

同理可得

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

定理 5.1 (克莱姆法则)

设 $\boldsymbol{A}$ 是一个 n 阶方阵, 若 $|\boldsymbol{A}|\neq 0$, 则线性方程组

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},$$

有且仅有唯一解, 且其解为

$$x_1=\frac{|\boldsymbol{A}_1|}{|\boldsymbol{A}|},\dots,x_i=\frac{|\boldsymbol{A}_i|}{|\boldsymbol{A}|}\dots,x_n=\frac{|\boldsymbol{A}_n|}{|\boldsymbol{A}|}.$$

其中 $\boldsymbol{A}_i$ 是将 $\boldsymbol{A}$ 的第 i 列换为 $\boldsymbol{b}$ 得到的矩阵.

证明 我们首先证明当 $|\boldsymbol{A}|\neq 0$ 时,

$$x_1=\frac{|\boldsymbol{A}_1|}{|\boldsymbol{A}|},\dots,x_i=\frac{|\boldsymbol{A}_i|}{|\boldsymbol{A}|}\dots,x_n=\frac{|\boldsymbol{A}_n|}{|\boldsymbol{A}|}\tag{5.1}$$

是线性方程组的解.

考察线性方程组的第 i 个方程, 该方程为

$$a_{i1}x_1+\ldots+a_{in}x_n=b_i,\ i=1,2,\dots,n.$$

使用求和符号可以将上述方程表示为:

$$\sum_{j=1}^{n}a_{ij}x_j,\ i=1,2,\dots,n.\tag{5.2}$$

矩阵 $\boldsymbol{A}_j$ 的行列式按照第 j 列可以展开为

$$|\boldsymbol{A}_j|=b_1A_{1j}+\ldots+b_nA_{nj}=\sum_{k=1}^{n}b_kA_{kj}.$$

其中, A_{kj} 表示 $|\boldsymbol{A}|$ 中元素 a_{kj} 对应的代数余子式（也是 $\boldsymbol{A}_i$ 中元素 b_k 对应的代数余子式）. 将等式(5.1)代入方程(5.2), 可得

$$\begin{aligned}\sum_{j=1}^{n}a_{ij}x_j&=\sum_{j=1}^{n}a_{ij}\frac{|\boldsymbol{A}_j|}{|\boldsymbol{A}|}=\frac{1}{|\boldsymbol{A}|}\sum_{j=1}^{n}a_{ij}|\boldsymbol{A}_j|\\&=\frac{1}{|\boldsymbol{A}|}\sum_{j=1}^{n}a_{ij}(\sum_{k=1}^{n}b_kA_{kj})=\frac{1}{|\boldsymbol{A}|}\sum_{j=1}^{n}\sum_{k=1}^{n}a_{ij}b_kA_{kj}\\&=\frac{1}{|\boldsymbol{A}|}\sum_{k=1}^{n}b_k(\sum_{j=1}^{n}a_{ij}A_{kj})\\&=\frac{1}{|\boldsymbol{A}|}b_i|\boldsymbol{A}|,\qquad i=1,2,\dots,n.\end{aligned}$$

注意最后一个等式, 我们使用了推论4.1, 即

$$\sum_{j=1}^{n}a_{ij}A_{kj}=\begin{cases}|\boldsymbol{A}|,& k=i,\\0,& k\neq i.\end{cases}$$

然后证明唯一性, 该证明过程与例题5.1类似. 用向量 $\boldsymbol{x}$ 替换单位矩阵 $\boldsymbol{E}_n$ 的第 i 列, 将得到的新矩阵记为 $\boldsymbol{X}_i$,

易知 $|\boldsymbol{X}_i| = x_i$. 用 $\boldsymbol{A}$ 左乘 $\boldsymbol{X}_i$, 我们有

$$\boldsymbol{A}\boldsymbol{X}_i = \boldsymbol{A}_i,$$

其中 $\boldsymbol{A}_i$ 是将 $\boldsymbol{A}$ 的第一列换为 $\boldsymbol{b}$ 得到的矩阵. 两边取行列式我们有:

$$|\boldsymbol{A}||\boldsymbol{X}_i| = |\boldsymbol{A}_i|,$$

即

$$|\boldsymbol{A}|x_i = |\boldsymbol{X}_i|$$

若 $|\boldsymbol{A}| \neq 0$, 则必有

$$x_i = \frac{|\boldsymbol{A}_i|}{|\boldsymbol{A}|}.$$

命题得证. □

推论 5.1

设 $\boldsymbol{A}$ 是一个 n 阶方阵, 若 $|\boldsymbol{A}| \neq 0$, 则线性方程组

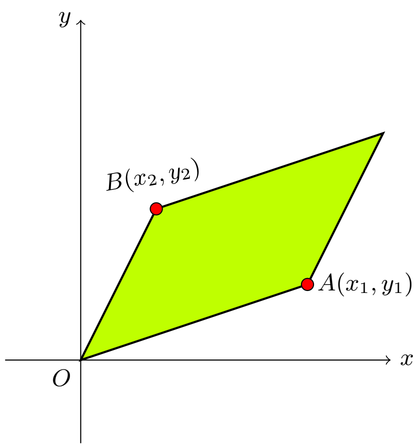
$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0},$$

只有零解.

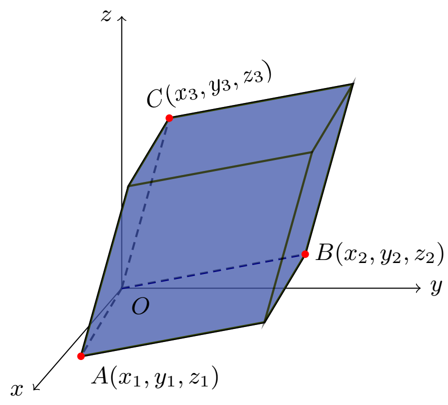
注 实际上, 若 $\boldsymbol{A}$ 是一个 n 阶方阵, 则 $|\boldsymbol{A}| \neq 0 \iff \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 只有零解.

5.2 行列式的几何意义

二阶行列式 $\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}$ 的绝对值等于平行四边形的面积.



三阶行列式 $\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$ 的绝对值等于平行六面体的体积.



在微积分中, 对二重积分进行坐标变换, 需要乘以一个雅克比行列式, 行列式的符号是有意义的. 此外, 在有限元方法中, 对区域进行三角形剖分, 可以使用行列式的符号判断剖分后每个三角形的顶点是否是按照逆时针排列的.

5.3 行列式的计算

例题 5.2 证明 n 阶范德蒙行列式

$$V_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{1\le i<j\le n}(x_j - x_i),$$

其中

$$\begin{aligned}\prod_{1\le i<j\le n}(x_j - x_i) = (x_2-x_1)(x_3-x_1)\cdots(x_n-x_1) \\ \cdot(x_3-x_2)\cdots(x_n-x_2) \\ \cdots\cdot \\ \cdot(x_{n-1}-x_{n-2})(x_n-x_{n-2}) \\ \cdot(x_n-x_{n-1}).\end{aligned}$$

证明 我们可以使用数学归纳法进行证明. 当 $n=2$ 时,

$$V_2 = \begin{vmatrix} 1 & 1 \\ x_1 & x_2 \end{vmatrix} = x_2 - x_1 = \prod_{1\le i<j\le n}(x_j - x_i),$$

结论成立. 假设结论对 $n-1$ 阶范德蒙行列式成立, 下面证明对 n 阶范德蒙行列式结论页成立. 对 n 阶范德蒙行列式, 从第 n 行开始, 每一行减去它的上面一行乘以 x_1, 可得

$$V_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & x_2-x_1 & x_3-x_1 & \cdots & x_n-x_1 \\ 0 & x_2(x_2-x_1) & x_3(x_3-x_1) & \cdots & x_n(x_n-x_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & x_2^{n-2}(x_2-x_1) & x_3^{n-2}(x_3-x_1) & \cdots & x_n^{n-2}(x_n-x_1) \end{vmatrix}$$

按照第一列展开, 可得

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

我们可以看到, 每一行都可以提取一个公因子 $x_i - x_1$ 可得

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

由归纳假设

$$V_n = \prod_{j=2}^{n}(x_j - x_1)\prod_{2\le i<j\le n}(x_j - x_i) = \prod_{1\le i<j\le n}(x_j - x_i).$$

综上, 结论对所有大于 1 的整数都成立. □

第六讲 逆矩阵和初等矩阵

内容提要

(1) 逆矩阵的定义和性质　　　　(2) 初等矩阵和初等变换

我们希望矩阵也能够有类似除法的运算. 这样, 在求解线性方程组时, 我们在两边同时“除以”系数矩阵, 就可以得到线性方程组的解. 为此, 我们需要先定义矩阵的“除法”. 本节我们将介绍逆矩阵 (“矩阵的除法”) 以及初等矩阵.

6.1 逆矩阵的定义和性质

在数字的四则运算法则中, 我们是通过乘法和倒数来定义除法的. 对于一个数字 a, 如果存在数字 b, 满足

$$ab = ba = 1,$$

则称 b 是 a 的倒数, 记为 a^{-1}. 类似地, 我们也可以定义矩阵的“倒数”.

定义 6.1

设 $\boldsymbol{A}$ 是一个 n 阶方阵, 我们称 $\boldsymbol{A}$ 是可逆的, 如果存在一个矩阵 $\boldsymbol{B}$ 使得

$$\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{E}_n,$$

矩阵 $\boldsymbol{B}$ 被称为 $\boldsymbol{A}$ 的逆矩阵, 记为 $\boldsymbol{A}^{-1}$.

根据逆矩阵的定义, 我们可以得到拟矩阵的一些性质.

性质 6.1

若矩阵 $\boldsymbol{A}$ 可逆, 则

(1) $\boldsymbol{AB} = \boldsymbol{C} \Rightarrow \boldsymbol{B} = \boldsymbol{A}^{-1}\boldsymbol{C}, \boldsymbol{BA} = \boldsymbol{C} \Rightarrow \boldsymbol{B} = \boldsymbol{CA}^{-1}$;

(2) $\boldsymbol{AB} = \boldsymbol{AC} \Rightarrow \boldsymbol{B} = \boldsymbol{C}, \boldsymbol{BA} = \boldsymbol{CA} \Rightarrow \boldsymbol{B} = \boldsymbol{C}$.

证明 易知

$$\boldsymbol{B} = \boldsymbol{EB} = (\boldsymbol{A}^{-1}\boldsymbol{A})\boldsymbol{B} = \boldsymbol{A}^{-1}(\boldsymbol{AB}) = \boldsymbol{A}^{-1}\boldsymbol{C}.$$

其余结论同理可证. □

性质 6.2

如果方阵 $\boldsymbol{A}$ 可逆, 则它的逆是唯一的.

证明 设 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 都是 $\boldsymbol{A}$ 的逆矩阵, 则有

$$\boldsymbol{B} = \boldsymbol{BE} = \boldsymbol{B}(\boldsymbol{AC}) = (\boldsymbol{BA})\boldsymbol{C} = \boldsymbol{EC} = \boldsymbol{C}$$

结论得证. □

性质6.1中的 (1) 和性质6.2告诉我们, 如果 $\boldsymbol{A}$ 可逆, 则线性方程组 $\boldsymbol{Ax} = \boldsymbol{b}$ 存在唯一解 $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$. 于是, 求解含有 n 个方程的 n 元线性方程组的问题转化为求可逆矩阵 $\boldsymbol{A}$ 的逆. 为了给出可逆矩阵求逆的公式, 我们引入伴随矩阵的概念.

定义 6.2

设 $\boldsymbol{A}$ 是一个 n 阶方阵, A_{ij} 是元素 a_{ij} 的代数余子式, 称矩阵 $\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}$ 为矩阵 $\boldsymbol{A}$ 的伴随矩阵, 记为 $\boldsymbol{A}^*$.

性质 6.3

设 $\boldsymbol{A}$ 是一个 n 阶方阵, 则 $\boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{A}^*\boldsymbol{A} = |\boldsymbol{A}|\boldsymbol{E}_n$.

证明 由定理4.5和推论4.1易得. □

定理 6.1

设 $\boldsymbol{A}$ 是一个 n 阶方阵, 若 $|\boldsymbol{A}| \neq 0$, 则 $|\boldsymbol{A}|$ 可逆, 且 $\boldsymbol{A}^{-1} = \dfrac{\boldsymbol{A}^*}{|\boldsymbol{A}|}$.

证明 由性质6.3易得, $\boldsymbol{A}^{-1} = \frac{\boldsymbol{A}^*}{|\boldsymbol{A}|}$. □

注 由定理6.1亦可得出 $|\boldsymbol{A}| \neq 0$, 克莱姆法则中解的公式.

定理 6.2

设 $\boldsymbol{A}$ 是一个 n 阶方阵, 则 $|\boldsymbol{A}| \neq 0 \Longleftrightarrow \boldsymbol{A}$ 可逆.

证明 若 $\boldsymbol{A}$ 可逆, 则由拟矩阵的定义可知, 存在一个矩阵 $\boldsymbol{B}$, 使得

$$\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A} = \boldsymbol{E},$$

两边取行列式可得

$$|\boldsymbol{A}\boldsymbol{B}| = |\boldsymbol{A}||\boldsymbol{B}| = |\boldsymbol{E}| = 1,$$

因此 $|\boldsymbol{A}| \neq 0$. 反之, 若 $|\boldsymbol{A}| \neq 0$, 则由定理6.1可得 $\boldsymbol{A}$ 可逆. □

定理 6.3

设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是 n 阶方阵, 如果 $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{E}_n$, 则 $\boldsymbol{B} = \boldsymbol{A}^{-1}$.

证明 由条件控制值,

$$|\boldsymbol{A}\boldsymbol{B}| = |\boldsymbol{A}||\boldsymbol{B}| = |\boldsymbol{E}_n| = 1,$$

因此 $|\boldsymbol{A}| \neq 0$, 由定理6.2可知, $\boldsymbol{A}$ 可逆. 于是我们有

$$\boldsymbol{B} = \boldsymbol{E}_n\boldsymbol{B} = (\boldsymbol{A}^{-1}\boldsymbol{A})\boldsymbol{B} = \boldsymbol{A}^{-1}(\boldsymbol{A}\boldsymbol{B}) = \boldsymbol{A}^{-1}\boldsymbol{E}_n = \boldsymbol{A}^{-1}.$$

定理得证. □

该定理告诉我们, 在判断方阵 $\boldsymbol{A}$ 是否可逆时, 只需要验证 $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{E}$ 或者 $\boldsymbol{B}\boldsymbol{A} = \boldsymbol{E}$ 即可, 这样将减少一半的工作量.

性质 6.4

设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是 n 阶可逆方阵, 则

(1) $\boldsymbol{A}^{-1}$ 可逆, 且 $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$;

(2) $\boldsymbol{A}^{\mathrm{T}}$ 可逆, 且 $(\boldsymbol{A}^{\mathrm{T}})^{-1} = (\boldsymbol{A}^{-1})^{\mathrm{T}}$;

(3) $(\boldsymbol{A}\boldsymbol{B})^{-1}$ 可逆, 且 $(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.

证明 按照定义6.1或定理6.3验证即可.

(1) 由条件可知

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

由定义6.1可得 A^{-1} 可逆, 且 $(A^{-1})^{-1} = A$;

(2) 由矩阵乘法的性质2.4, 易知

$$A^{\mathrm{T}}(A^{-1})^{\mathrm{T}} = (A^{-1}A)^{\mathrm{T}} = E,$$

由定理6.3可得 A^{T} 可逆, 且 $(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}$;

(3) 由矩阵乘法的性质2.3, 易知

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AEA^{-1} = E,$$

由定理6.3可得 $(AB)^{-1}$ 可逆, 且 $(AB)^{-1} = B^{-1}A^{-1}$.

□

注 设 A 是 n 阶方阵, 则 A 可逆 $\Longleftrightarrow A^{\mathrm{T}}$ 可逆.

性质 6.5

设 A, B 都是 n 阶可逆方阵, 则 $(AB)^* = B^*A^*$.

证明 由性质6.3和性质6.4可得,

$$(AB)^* = |AB|(AB)^{-1} = |A||B|(B^{-1}A^{-1}) = (|B|B^{-1})(|A|A^{-1}) = B^*A^*.$$

□

注 事实上, 不论 A, B 是否可逆, 都有 $(AB)^* = B^*A^*$. 这部分留作联系 (提示: 可以考虑当 t 足够大时有 $(A+tE)$ 和 $(B+tE)$ 都可逆, 此时有

$$((A+tE)(B+tE))^* = (B+tE)^*(A+tE)^*$$

注意上式两端实际上都是 t 的多项式, 因此对于任意的 t 上式成立, 取 $t=0$ 命题得证.)

性质 6.6

设 A 是 n 阶可逆方阵, 则 A^* 可逆, 且 $(A^*)^{-1} = (A^{-1})^*$.

证明 由性质6.3易知, 若 $|A| \neq 0$, 则 $|A^*| \neq 0$, 因此 A^* 可逆. 接下来只需证明

$$A^*(A^{-1})^* = E$$

即可. 由性质6.5可得

$$A^*(A^{-1})^* = (A^{-1}A)^* = E^* = E.$$

结论得证. □

注 事实上, 我们也可以直接给出 $(A^*)^{-1}$ 和 $(A^{-1})^*$. 由 $AA^* = |A|E$ 可得 $A^* = |A|A^{-1}$, 于是 $(A^*)^{-1} = A/|A|$; 同样地, 由 $(A^{-1})(A^{-1})^* = |A^{-1}|E$ 可得 $(A^{-1})^* = |A^{-1}|A = A/|A|$.

对于分块对角矩阵, 我们显然有以下结论.

性质 6.7

假设 $A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_s \end{pmatrix}$ 是可逆矩阵, 则有 $A^{-1} = \begin{pmatrix} A_1^{-1} & 0 & \cdots & 0 \\ 0 & A_2^{-1} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_s^{-1} \end{pmatrix}$.

6.2 初等矩阵和矩阵的初等变换

在本节, 我们将介绍初等矩阵以及初等矩阵和矩阵的初等变换之间的关系. 首先, 我们将给出初等矩阵的定义, 紧接着, 我们将介绍初等矩阵和矩阵的初等变换之间的关系.

定义 6.3

由 n 阶单位矩阵 $\boldsymbol{E}_n$ 经过一次以下变换得到的矩阵称为 n 阶初等矩阵.

由初等矩阵的定义可知, 有三种类型的初等矩阵

(1) **交换初等矩阵:** 交换矩阵的两行（或两列）, 即

$$\boldsymbol{E}_n(i,j)=\begin{pmatrix}1&&&&&&&&&&\\&\ddots&&&&&&&&&\\&&1&&&&&&&&\\&&&0&&\cdots&&1&&&\\&&&&1&&&&&&\\&&&\vdots&&\ddots&&\vdots&&&\\&&&&&&1&&&&\\&&&1&\cdots&&&0&&&\\&&&&&&&&1&&\\&&&&&&&&&\ddots&\\&&&&&&&&&&1\end{pmatrix}\begin{array}{l}\\ \\ \\ \leftarrow 第\ i\ 行\\ \\ \\ \\ \leftarrow 第\ j\ 行\\ \\ \\ \\ \end{array}$$

(2) **倍乘初等矩阵:** 将矩阵某行（或列）的所有元素乘以非零常数 k, 即

$$\boldsymbol{E}_n(i(k))=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&&&&\\&&&k&&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}\begin{array}{l}\\ \\ \\ \leftarrow 第\ i\ 行\\ \\ \\ \\ \end{array}$$

(3) **倍加初等矩阵:** 将矩阵某行（或列）的所有元素乘以常数 k 加到另一行（或列）上, 即

$$\boldsymbol{E}_n(i,j(k))=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&\cdots&k&&\\&&&\ddots&\vdots&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}\begin{array}{l}\\ \\ \leftarrow 第\ i\ 行\\ \\ \leftarrow 第\ j\ 行\\ \\ \\ \end{array}.$$

下面的定理, 揭示了初等矩阵与初等变换之间的关系.

定理 6.4

假设 $\boldsymbol{A}$ 是 $m\times n$ 的矩阵, 则

(1) 对 $\boldsymbol{A}$ 施行一次初等行变换, 相当于在矩阵 $\boldsymbol{A}$ 的左侧乘以对应的 m 阶初等矩阵;

(2) 对 $\boldsymbol{A}$ 施行一次初等变列换, 相当于在矩阵 $\boldsymbol{A}$ 的右侧乘以对应的 n 阶初等矩阵.

证明 直接验证即可. □

例题 6.1 观察下面的 3 阶矩阵乘法:

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \overset{(1)}{=\!=} \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

可以看出, 用这种矩阵左乘一个矩阵 A, 等同于将 A 的某两行交换了位置.

例题 6.2 观察下面的 3 阶矩阵乘法:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \overset{(2)}{=\!=} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ ka_{21} & ka_{22} & ka_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

可以看出, 用这种矩阵左乘一个矩阵 A, 等同于将 A 的某行变成原来的 k 倍

例题 6.3 观察下面的 3 阶矩阵乘法:

$$\begin{pmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \overset{(3)}{=\!=} \begin{pmatrix} a_{11}+ka_{31} & a_{12}+ka_{32} & a_{13}+ka_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

可以看出, 用这种矩阵左乘一个矩阵 A, 等同于将 A 的某行加上另一行的 k 倍.

定理 6.5

初等矩阵都是可逆的, 而且它们的逆仍然是初等矩阵. 事实上:

(1) $\boldsymbol{E}_n(i,j)^{-1} = \boldsymbol{E}_n(i,j)$;

(2) $\boldsymbol{E}(i(k))^{-1} = \boldsymbol{E}(i(k^{-1}))$;

(3) $\boldsymbol{E}(i,j(k))^{-1} = \boldsymbol{E}(i,j(-k))$.

证明 按照拟矩阵的定义直接验证即可. □

第七讲 利用初等矩阵求逆矩阵

内容提要

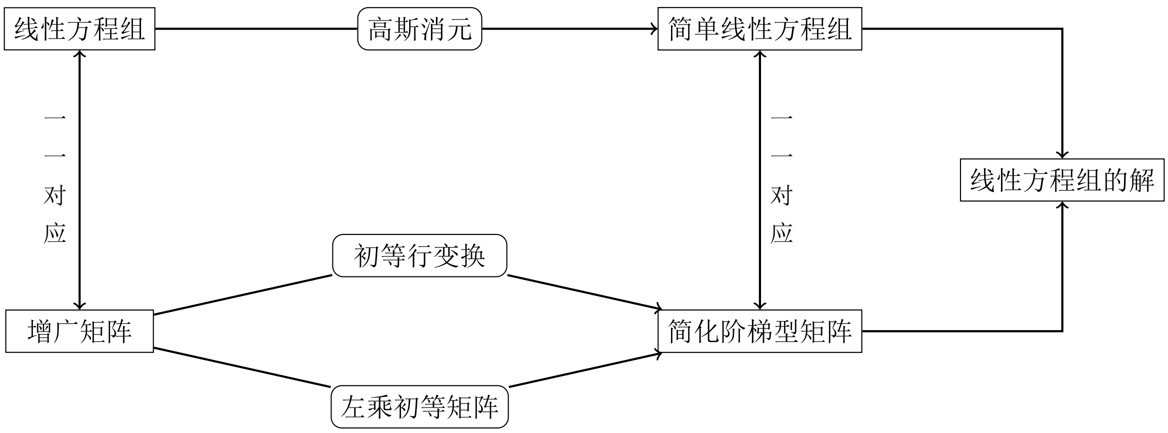
(1) 矩阵可逆的充分必要条件 　　(2) 利用初等行变换求逆矩阵

在上一节, 我们介绍了逆矩阵和初等矩阵, 虽然我们利用伴随矩阵, 给出了一个求逆矩阵的公式, 但是使用该公式需要计算 1 个 n 阶行列式, 和 n^2 个 $n-1$ 阶行列式, 通常情况下计算量非常庞大, 并不适合用来设计实用算法. 在实际计算中, 我们通过高斯消元法求解矩阵方程

$$\boldsymbol{A}\boldsymbol{X}=\boldsymbol{E}$$

来计算 $\boldsymbol{A}$ 的逆矩阵.

我们知道, 高斯消元法等价于对相应的矩阵矩阵进行一系列的初等行变换, 而初等行变换又可以通过对矩阵左乘相应的初等矩阵进行. 这个对应关系可以使用下图描述:



特别地, 系数矩阵 $\boldsymbol{A}$ 可逆时, 上图中的简化阶梯型矩阵的前 n 列实际上是一个单位矩阵. 在这种情况下, 我们可以直接写出线性方程组的解或者逆矩阵.

7.1 矩阵可逆的等价条件

我们知道, 对线性方程组对应的增广矩阵进行初等行变换, 不改变线性方程组（注意, 求解线性方程组时不可进行初等列变换）. 类似地, 对方阵进行初等行变换, 不改变方阵的可逆性.

引理 7.1

对方阵 $\boldsymbol{A}$ 进行一次初等行 (列) 变换得到矩阵 $\boldsymbol{B}$, 则 $\boldsymbol{A}$ 可逆 $\Longleftrightarrow \boldsymbol{B}$ 可逆.

证明 易知, 若 $\boldsymbol{A}$ 经过一次初等行变换得到矩阵 $\boldsymbol{B}$, 存在一个初等矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}\boldsymbol{A}=\boldsymbol{B}$, 两边同时取行列式可得

$$|\boldsymbol{P}||\boldsymbol{A}|=|\boldsymbol{B}|,$$

又因为 $\boldsymbol{A}$ 可逆, 故 $|\boldsymbol{P}|\neq 0$, 因此

$$|\boldsymbol{A}|\neq 0 \Longleftrightarrow |\boldsymbol{B}|\neq 0;$$

即 $\boldsymbol{A}$ 可逆 $\Longleftrightarrow \boldsymbol{B}$ 可逆.

如果是初等列变换, 同理可证. □

定理 7.1

任何 $m \times n$ 矩阵 $\boldsymbol{A}$ 经过若干次初等变换后, 可以化为如下形式的矩阵:

$$\boldsymbol{D} = \begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{O}_{r,n-r} \\ \boldsymbol{O}_{m-r,r} & \boldsymbol{O}_{m-r,n-r} \end{pmatrix}$$

即, 除了左上角为 r 阶单位阵, 其他元素都是 0. 该形式的矩阵称为矩阵 A 的等价标准形.

证明 使用归纳法, 容易证明. □

例题 7.1 3×4 矩阵的等价标准形有如下几种

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

由引理7.1和定理7.1, 我们可以得到如下定理

定理 7.2

若 $\boldsymbol{A}$ 可逆 $\Longleftrightarrow$ $\boldsymbol{A}$ 等价于单位矩阵.

证明 由定理7.1, 任意方阵都等价于以下形式

$$\begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{O}_{r,n-r} \\ \boldsymbol{O}_{n-r,r} & \boldsymbol{O}_{n-r,n-r} \end{pmatrix}.$$

由于初等变换不改变矩阵的可逆性, 因此, 若 $\boldsymbol{A}$ 可逆, 则必有 $n-r=0$, 即 $\boldsymbol{A}$ 等价于单位矩阵.

反之, 若 $\boldsymbol{A}$ 等价于单位矩阵, 则存在初等矩阵 $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_s$ 和 $\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_t$, 使得

$$\boldsymbol{P}_1, \ldots, \boldsymbol{P}_s \boldsymbol{A} \boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_t = \boldsymbol{E},$$

两边取行列式, 得

$$|\boldsymbol{P}_1| \cdots |\boldsymbol{P}_s||\boldsymbol{A}||\boldsymbol{Q}_1| \cdots |\boldsymbol{Q}_t| = 1,$$

因此 $|\boldsymbol{A}| \neq 0$, 从而 $\boldsymbol{A}$ 可逆. □

定理 7.3

若 $\boldsymbol{A}$ 可逆 $\Longleftrightarrow$ $\boldsymbol{A}$ 可以表示为一些列初等矩阵的乘积.

证明 若 $\boldsymbol{A}$ 可以表示为一些列初等矩阵的乘积, 则显然有 $\boldsymbol{A}$ 可逆. 反之, 若 $\boldsymbol{A}$ 可逆, 则由定理7.2, $\boldsymbol{A}$ 等价于单位矩阵. 故存在初等矩阵 $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_s$ 和 $\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_t$, 使得

$$\boldsymbol{P}_1, \ldots, \boldsymbol{P}_s \boldsymbol{A} \boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_t = \boldsymbol{E},$$

于是我们有

$$\boldsymbol{A} = \boldsymbol{P}_s^{-1} \cdots \boldsymbol{P}_1^{-1} \boldsymbol{Q}_t^{-1} \cdots \boldsymbol{Q}_1^{-1},$$

注意到初等矩阵的逆矩阵还是初等矩阵, 命题得证. □

定理 7.4

若 $\boldsymbol{A}$ 可逆 $\Longleftrightarrow$ $\boldsymbol{A}$ 可以只经过一系列初等行 (列) 变换转化为单位矩阵.

证明 由定理7.3证明中的讨论上面的讨论, 我们知道, $\boldsymbol{A}$ 可逆等价于

$$\boldsymbol{A} = \boldsymbol{P}_s^{-1} \cdots \boldsymbol{P}_1^{-1} \boldsymbol{Q}_t^{-1} \cdots \boldsymbol{Q}^{-1},$$

其中 $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_s$ 和 $\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_t$ 是初等矩阵. 于是我们有 $\boldsymbol{A}$ 可逆等价于

$$\boldsymbol{Q}_1 \cdots \boldsymbol{Q}_t \boldsymbol{P}_1, \ldots, \boldsymbol{P}_s \boldsymbol{A} = \boldsymbol{E}.$$

命题得证. □

7.2 利用初等变换求逆矩阵

定理7.4实质上给出了利用初等矩阵或者初等行变换求逆矩阵的方法. 假设 $\boldsymbol{A}$ 可逆, 则由定理7.4可知, 存在一系列的初等矩阵 $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_s$, 使得

$$\boldsymbol{P}_1, \ldots, \boldsymbol{P}_s \boldsymbol{A} = \boldsymbol{E},$$

即

$$\boldsymbol{P}_1, \ldots, \boldsymbol{P}_s = \boldsymbol{E}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}.$$

也就是说, 只要对单位矩阵进行与 $\boldsymbol{A}$ 相同的初等行变换, 则经过初等行变换之后, 若 $\boldsymbol{A}$ 化为单位矩阵, 则单位矩阵将化为 $\boldsymbol{A}^{-1}$.

为了同步进行这件事情, 我们可以将 $\boldsymbol{A}$ 和单位阵 $\boldsymbol{E}$ 放在一起, 构成一个 $n \times 2n$ 的矩阵

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{E} \end{pmatrix}.$$

对该矩阵进行初等行变换（注意不可以进行初等列变换）, 将 $\boldsymbol{A}$ 的位置化为单位阵, 则 $\boldsymbol{E}$ 所在的的位置将被化为 $\boldsymbol{A}^{-1}$.

注意, 该过程实质上是使用高斯消元法求矩阵方程 (可以看做是系数举着相同, 但常数向量不相同的 n 个联系方程组)

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{E}.$$

类似地, 若 $\boldsymbol{A}$ 可逆, 我们想求解矩阵方程

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{B},$$

只需要对矩阵

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \end{pmatrix}.$$

进行初等行变换, 将 $\boldsymbol{A}$ 的位置化为单位矩阵, $\boldsymbol{B}$ 的位置就会化为 $\boldsymbol{A}^{-1}\boldsymbol{B}$.

例题 7.2 用初等变换求矩阵的逆矩阵：

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & -2 & 1 \end{pmatrix}$$

解 对矩阵 $(\boldsymbol{A}, \boldsymbol{E})$ 作初等变换, 将其化为简化阶梯型矩阵,

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & -2 & 1 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{r_2 - r_1} \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & -2 & 1 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{r_2 \to (-1) r_2} \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 \\ 0 & -2 & 1 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[r_3 - 2r_2]{r_1 + r_2} \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 \\ 0 & 0 & -1 & 2 & -2 & 1 \end{pmatrix}$$

$$\xrightarrow{r_3 \to (-1) r_3} \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 \\ 0 & 0 & 1 & -2 & 2 & -1 \end{pmatrix} \xrightarrow[r_2 - r_3]{r_1 + r_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & -1 & 1 \\ 0 & 1 & 0 & -1 & 1 & -1 \\ 0 & 0 & 1 & -2 & 2 & -1 \end{array}\right)$$

因此

$$A^{-1}=\begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ -2 & 2 & -1 \end{pmatrix}$$

✍ **练习 7.1** 用初等变换求矩阵的逆矩阵.

$$\boldsymbol{A}=\begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ 2 & 0 & 3 \end{pmatrix};\quad \boldsymbol{B}=\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}.$$

解

$$\boldsymbol{A}^{-1}=\begin{pmatrix} -3 & 0 & 2 \\ 1 & 1 & -1 \\ 2 & 0 & -1 \end{pmatrix};\quad \boldsymbol{B}^{-1}=\frac{1}{5}\begin{pmatrix} -3 & 2 & 2 \\ 2 & -3 & 2 \\ 2 & 2 & -3 \end{pmatrix}.$$