

第 2 章 矩阵及其运算

在上一章的最后一节，我们以行列式为工具讨论了线性方程组，并且在当未知数的个数与方程的个数相等且系数行列式不等于零时，给出了解的计算公式（克拉默法则）。但现实生活中所产生的数学模型往往都不满足克拉默法则的条件，为了求解一般的线性方程组，我们将在本章引入矩阵理论和矩阵计算方面的一些基本内容。

2.1 矩阵与线性变换初步

定义 2-1 (矩阵) 由 $m \times n$ 个数 a_{ij} ($i=1,2,\dots,m; j=1,2,\dots,n$) 排成的 m 行 n 列的数表

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

称为 m 行 n 列矩阵，为了表示它是一个整体，我们往往对数表加上一个括号，并用大写加粗的英文字母来表示它，例如：

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad (2-1)$$

这 $m \times n$ 个数称为矩阵 A 的**元素**，简称**元**，数 a_{ij} 位于矩阵的第 i 行第 j 列，称为矩阵的 **(i, j) 元**，以数 a_{ij} 为 (i, j) 元的 m 行 n 列矩阵可简记作 $(a_{ij})_{m \times n}$ 或 (a_{ij}) ，也可记作 $A_{m \times n}$ ，读作 m 行 n 列矩阵 A 。

例如，某厂三种产品在一年中四个季度的产量可以列成一个 3 行 4 列的矩阵：

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix},$$

其中 a_{ij} 为第 i 种产品在第 j 季度的产量。这三种产品的单价及单件重量则可以列成一个 3 行 2 列矩阵：

$$P = \begin{pmatrix} p_1 & w_1 \\ p_2 & w_2 \\ p_3 & w_3 \end{pmatrix},$$

其中 p_i ($i=1,2,3$) 为第 i 种产品的单价, w_i ($i=1,2,3$) 为第 i 种产品的单件重量.

又如, 四个城市间的单向航线如图 2-1 所示. 若令:

$$a_{ij} = \begin{cases} 1, & \text{从 } i \text{ 市到 } j \text{ 市有 1 条单向航线} \\ 0, & \text{从 } i \text{ 市到 } j \text{ 市无单向航线} \end{cases},$$

则图 2-1 可用矩阵表示为:

$$A = (a_{ij}) = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

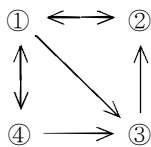


图 2-1

一般地, 若干个点之间的单向通道都可用这样的矩阵表示.

如果多个矩阵的行数和列数都对应相等, 则称它们为同型矩阵. 若同型矩阵 $A = (a_{ij})_{m \times n}$ 与 $B = (b_{ij})_{m \times n}$ 的对应元素相等, 即:

$$a_{ij} = b_{ij} \quad (i=1,2,\dots,m; j=1,2,\dots,n),$$

则称矩阵 A 与矩阵 B 相等, 记作:

$$A = B.$$

本章重点讨论实矩阵 (即元素为实数的矩阵). 下面给出几个特殊矩阵的概念.

行数和列数都相同的矩阵称为方阵. 如 $A_{m \times m}$ 叫做 m 阶矩阵或 m 阶方阵, 简记为 A_m , $B_{n \times n}$ 叫做 n 阶矩阵或 n 阶方阵, 简记为 B_n , $C_{r \times r}$ 叫做 r 阶矩阵或 r 阶方阵, 简记为 C_r 等.

只有一行的矩阵 $A = (a_1 \ a_2 \ \cdots \ a_n)$ 称为行矩阵或行向量, 通常写作:

$$A = (a_1, a_2, \dots, a_n),$$

其对应的元素 a_j ($j=1,2,\dots,n$) 又称该向量的第 j 个分量.

只有一列的矩阵

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

称为列矩阵或列向量. 同样, 其对应的元素 b_i ($i=1,2,\dots,n$) 也称该向量的第 i 个分量.

例如, 可以把矩阵

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

的每一行元素构成的矩阵都看成是一个行向量, 把它的每一列元素构成的矩阵都看成一个列向量.

元素都为零的矩阵称为**零矩阵**, 记作 $\mathbf{O}$. 注意不同型的零矩阵是不相等的. 例如:

$$\mathbf{O}_{2 \times 2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{O}_{3 \times 4} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{O}_{1 \times 4} = (0 \ 0 \ 0 \ 0), \quad \mathbf{O}_{3 \times 1} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

都是零矩阵, 但它们不同型, 因此是互不相等的.

通常将只有一列元素的零向量(即列向量的每个元素都为 0)记为 $\mathbf{0}$.

在本章及后续章节中, 我们还会陆续引入其他特殊矩阵的概念.

定义 2-2 (线性变换) 若 m 个变量 $y_1, y_2, \dots, y_m$ 的取值可分别由 n 个变量 $x_1, x_2, \dots, x_n$ 通过如下关系式

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \quad \quad \quad \cdots \cdots \cdots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{cases} \quad (2-2)$$

得到, 则称该关系式为从变量 $x_1, x_2, \dots, x_n$ 到变量 $y_1, y_2, \dots, y_m$ 的**线性变换**, 其中常数 a_{ij} 称为**变换系数**, 所有变换系数构成一个 m 行 n 列矩阵 $A = (a_{ij})_{m \times n}$, 称为该线性变换的**系数矩阵**(也称为**变换矩阵**).

由定义 2-2 知, 当给出了线性变换 (2-2) 后, 它的系数矩阵也就唯一确定了; 反过来, 如果给出一个矩阵作为线性变换的系数矩阵, 线性变换也就唯一确定了. 从这个意义上讲, 线性变换和矩阵之间存在着——对应的关系.

例如线性变换

$$\begin{cases} y_1 = x_1 \\ y_2 = x_2 \\ \quad \quad \quad \cdots \cdots \cdots \\ y_n = x_n \end{cases}$$

叫做**恒等变换**, 它的变换矩阵

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

是一个 n 阶方阵, 叫做 n 阶单位矩阵, 简称单位阵, 其特点是: 从左上角到右下角的直线 (叫做(主) 对角线) 上的元素都是 1, 其余元素都为 0, 即 n 阶单位阵 I 的 (i, j) 元为

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (i, j = 1, 2, \dots, n).$$

又如线性变换

$$\begin{cases} y_1 = \lambda_1 x_1 \\ y_2 = \lambda_2 x_2 \\ \dots\dots\dots \\ y_n = \lambda_n x_n \end{cases}$$

的变换矩阵

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

同样是一个 n 阶方阵, 叫做 n 阶对角矩阵, 简称对角阵, 其特点是: 除 (主) 对角线上的元素以外的所有元素都为 0. 上述对角阵也可简写作:

$$A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n).$$

从几何上看, 矩阵 $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ 所对应的线性变换 $\begin{cases} x_1 = x \\ y_1 = y \\ z_1 = 0 \end{cases}$ 可以看作是三维空

间直角坐标系 $oxyz$ 中的向量 $\overline{OP} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ 变为向量 $\overline{OP_1} = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$ 的变换 (或看作把

点 $P(x, y, z)$ 的坐标变为点 $P_1(x, y, 0)$ 的变换, 如图 2-2 所示), 由于向量 $\overline{OP_1}$ 是向量 $\overline{OP}$ 在 xoy 坐标面上的投影向量 (即点 P_1 是点 P 在 xoy 坐标面上的投影), 因此这样的变换叫做投影变换. 读者自己可以写出点 P 在 yoz 和 zox 坐标面上的投影变换及其对应的变换矩阵和点 P 在 x 轴、 y 轴、 z 轴上的投影变换及其对应的变换矩阵.

又如, 矩阵 $\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$ 所对应的线性变换 $\begin{cases} x_1 = x \cos \alpha - y \sin \alpha \\ y_1 = x \sin \alpha + y \cos \alpha \end{cases}$ 可以看作

是把 xoy 平面上的向量 $\overline{OP} = \begin{pmatrix} x \\ y \end{pmatrix}$ 绕原点 O 逆时针旋转 α 角后变为向量 $\overline{OP_1} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$

(或把点 $P(x, y)$ 绕原点 O 逆时针旋转 α 角后变为点 $P_1(x_1, y_1)$) 的变换, 叫做旋转变换. 事实上, 如图 2-3 所示, 若设 $\overline{OP}$ 的长度为 r , 辐角为 θ , 即 $x = r \cos \theta$, $y = r \sin \theta$, 则:

$$x_1 = x \cos \alpha - y \sin \alpha = r \cos \theta \cos \alpha - r \sin \theta \sin \alpha = r \cos(\theta + \alpha),$$

$$y_1 = x \sin \alpha + y \cos \alpha = r \cos \theta \sin \alpha + r \sin \theta \cos \alpha = r \sin(\theta + \alpha),$$

这表明 $\overline{OP_1}$ 的长度也为 r ，但辐角变为 $\theta + \alpha$ ，即辐角在原来的基础上增加了 α 。

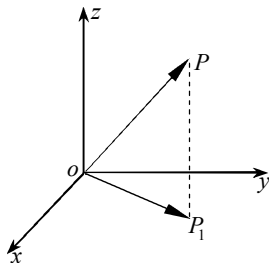


图 2-2

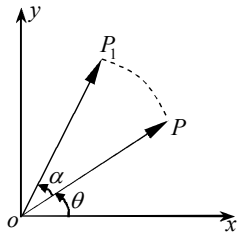


图 2-3

习题 2.1

1. 写出下列线性变换的变换矩阵：

$$(1) \begin{cases} y_1 = 5x_1 - x_3 \\ y_2 = x_1 + 2x_2 + 3x_3 \\ y_3 = -2x_2 \end{cases}$$

(2) 三维空间直角坐标系 $oxyz$ 中把点 $P(x, y, z)$ 的坐标变为其关于原点 O 的对称点 $P_1(-x, -y, -z)$ ；

(3) 三维空间直角坐标系 $oxyz$ 中把点 $P(x, y, z)$ 的坐标变为其在 zox 平面上的投影点 $P_1(x, 0, z)$ ；

(4) xoy 平面上把点 $P(x, y)$ 的 x 、 y 坐标分别变为其原来的 λ 、 μ 倍（这里 $\lambda\mu \neq 0$ ）。

2.2 矩阵的运算

2.2.1 矩阵的四则运算

1. 矩阵的加法

定义 2-3 (矩阵的和) 设 $A = (a_{ij})$ 与 $B = (b_{ij})$ 是两个同型的 $m \times n$ 矩阵，则将 A 与 B 对应位置元素相加后所得的矩阵 $(a_{ij} + b_{ij})$ 称为矩阵 A 与矩阵 B 的和，记作 $A + B$ ，即：

$$A+B=(a_{ij}+b_{ij})=\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix},$$

而把求同型矩阵的和的运算称为**矩阵的加法** (运算).

特别地, 若矩阵 O 表示与矩阵 A 同型的零矩阵, 则:

$$A+O=A.$$

根据定义 2-3, 矩阵的加法显然满足下列运算规律 (设 A 、 B 、 C 都是 $m \times n$ 矩阵):

- ① 交换律 $A+B=B+A$;
- ② 结合律 $(A+B)+C=A+(B+C)$.

设矩阵 $A=(a_{ij})$, 记

$$-A=(-a_{ij})$$

为将 A 的各元素取其相反数后得到的矩阵, 称 $-A$ 为 A 的**负矩阵**, 根据矩阵的加法, 显然有:

$$A+(-A)=O.$$

这里矩阵 O 表示与矩阵 A 同型的零矩阵, 由此规定同型矩阵 A 与 B 之差为:

$$A-B=A+(-B),$$

这就是**矩阵的减法** (运算).

矩阵 $A-B$ 实际上就是将 A 与 B 对应位置的元素相减后所得的矩阵. 显然若 $A=B$, 则 $A-B=O$, 反之亦然.

2. 数乘矩阵

定义 2-4 (数与矩阵的积) 设 λ 是一个常数, $A=(a_{ij})$ 是一个 $m \times n$ 矩阵, 则将 λ 乘以 A 的所有元素后所得的矩阵 (λa_{ij}) 称为**数 λ 与矩阵 A 的积**, 记作 λA , 即:

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

而把求数与矩阵相乘的运算称为**数乘矩阵**. 在不引起混淆的情况下, 有时也把 λA 简称为数乘矩阵.

根据定义 2-3 和定义 2-4, 数乘矩阵满足下列运算规律 (设 A 、 B 为 $m \times n$ 矩阵, λ 、 μ 为常数):

- ① 关于数的结合律 $(\lambda\mu)A=\lambda(\mu A)$;
- ② 关于数的分配律 $(\lambda+\mu)A=\lambda A+\mu A$;
- ③ 关于矩阵的分配律 $\lambda(A+B)=\lambda A+\lambda B$.

我们将矩阵的加法 (运算) 及数乘矩阵统称为矩阵的**线性运算**, 而把 $\lambda A+\mu B$

称为同型矩阵 A 、 B 的**线性组合**(这里 λ 、 μ 为常数), 它表示数乘矩阵 λA 与 μB 的和. 一般地, 同型矩阵 A_i ($i=1, 2, \dots, n$) 的线性组合写作:

$$\lambda_1 A_1 + \lambda_2 A_2 + \dots + \lambda_n A_n = \sum_{k=1}^n \lambda_k A_k,$$

它表示数乘矩阵 $\lambda_1 A_1, \lambda_2 A_2, \dots, \lambda_n A_n$ 的和, 这里 λ_i ($i=1, 2, \dots, n$) 是常数.

例 2-1 设:

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

求 $3A - 2B$.

$$\begin{aligned} \text{解} \quad 3A - 2B &= 3 \begin{pmatrix} 1 & 0 & -5 \\ -2 & 3 & 1 \end{pmatrix} - 2 \begin{pmatrix} 5 & 1 & 2 \\ 3 & -4 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 0 & -15 \\ -6 & 9 & 3 \end{pmatrix} - \begin{pmatrix} 10 & 2 & 4 \\ 6 & -8 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 3-10 & 0-2 & -15-4 \\ -6-6 & 9-(-8) & 3-0 \end{pmatrix} = \begin{pmatrix} -7 & -2 & -19 \\ -12 & 17 & 3 \end{pmatrix}. \end{aligned}$$

3. 矩阵的乘法

将变量 x_1, x_2 变为 y_1, y_2, y_3 的线性变换为:

$$(i) \quad \begin{cases} y_1 = b_{11}x_1 + b_{12}x_2 \\ y_2 = b_{21}x_1 + b_{22}x_2 \\ y_3 = b_{31}x_1 + b_{32}x_2 \end{cases},$$

再将 y_1, y_2, y_3 变为 z_1, z_2 的线性变换为:

$$(ii) \quad \begin{cases} z_1 = a_{11}y_1 + a_{12}y_2 + a_{13}y_3 \\ z_2 = a_{21}y_1 + a_{22}y_2 + a_{23}y_3 \end{cases},$$

要求将变量 x_1, x_2 变为 z_1, z_2 的线性变换及其变换矩阵.

将 (ii) 右端的 y_1, y_2, y_3 分别用 (i) 右端的关于 x_1, x_2 的表达式代替, 得:

$$\begin{cases} z_1 = a_{11}(b_{11}x_1 + b_{12}x_2) + a_{12}(b_{21}x_1 + b_{22}x_2) + a_{13}(b_{31}x_1 + b_{32}x_2) \\ z_2 = a_{21}(b_{11}x_1 + b_{12}x_2) + a_{22}(b_{21}x_1 + b_{22}x_2) + a_{23}(b_{31}x_1 + b_{32}x_2) \end{cases},$$

整理得:

$$(iii) \quad \begin{cases} z_1 = (a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31})x_1 + (a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32})x_2 \\ z_2 = (a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31})x_1 + (a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32})x_2 \end{cases}.$$

线性变换 (iii) 可看成是先作线性变换 (i) 再作线性变换 (ii) 的结果, 我们把线性变换 (iii) 称为线性变换 (ii) 与线性变换 (i) 的**乘积**, 相应地把线性变换 (iii) 的变换矩阵定义为线性变换 (ii) 与线性变换 (i) 的变换矩阵的**乘积**, 即:

$$\begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}.$$

一般地, 我们有:

定义 2-5 (矩阵的积) 设 $A = (a_{ij})$ 是一个 $m \times p$ 矩阵, $B = (b_{ij})$ 是一个 $p \times n$ 矩阵, 则矩阵 $C = (c_{ij})$ 是将 A 与 B 的元素作下列运算所得到的 $m \times n$ 矩阵:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^p a_{ik}b_{kj} \quad (i=1, 2, \cdots, m; j=1, 2, \cdots, n). \quad (2-3)$$

矩阵 C 称为矩阵 A 与矩阵 B 的积, 记作:

$$C = AB.$$

把求矩阵的积的运算称为**矩阵的乘法**(运算).

由定义 2-5 可知, 求两个矩阵的积, 第一个矩阵(左矩阵)的列数必须等于第二个矩阵(右矩阵)的行数, 且积的第 i 行第 j 列的元素为第一个矩阵(左矩阵)的第 i 行元素构成的行向量(共有 p 个分量)与第二个矩阵(右矩阵)的第 j 列元素构成的列向量(也有 p 个分量)对应分量相乘的和.

特别地, 任何矩阵与零矩阵的乘积(如果可以相乘的话)都是零矩阵. 一个 $1 \times p$ 的矩阵(行向量)与一个 $p \times 1$ 矩阵(列向量)的积是一个 1 阶方阵, 它是一个数, 即:

$$(a_1, a_2, \cdots, a_p) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix} = a_1b_1 + a_2b_2 + \cdots + a_pb_p = \sum_{k=1}^p a_kb_k.$$

例 2-2 求矩阵

$$A = \begin{pmatrix} 1 & 0 & -3 & 1 \\ 2 & 3 & 1 & 0 \end{pmatrix} \text{ 与 } B = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & 3 \\ 1 & 3 & 4 \\ 3 & 1 & 0 \end{pmatrix}$$

的乘积 AB .

解 A 为 2×4 矩阵, B 为 4×3 矩阵, A 的列数等于 B 的行数, 故 A 与 B 可以相乘, 其积 $C = AB$ 是一个 2×3 矩阵, 据式 (2-3), 有:

$$C = AB = \begin{pmatrix} 1 & 0 & -3 & 1 \\ 2 & 3 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & 3 \\ 1 & 3 & 4 \\ 3 & 1 & 0 \end{pmatrix}$$

$$\begin{aligned}
 &= \begin{pmatrix} 1 \times 1 + 0 \times 2 & 1 \times (-1) + 0 \times 0 & 1 \times 2 + 0 \times 3 \\ +(-3) \times 1 & +(-3) \times 3 & +(-3) \times 4 \\ +1 \times 3 & +1 \times 1 & +1 \times 0 \\ 2 \times 1 + 3 \times 2 & 2 \times (-1) + 3 \times 0 & 2 \times 2 + 3 \times 3 \\ +1 \times 1 & +1 \times 3 & +1 \times 4 \\ +0 \times 3 & +0 \times 1 & +0 \times 0 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & -9 & -10 \\ 9 & 1 & 17 \end{pmatrix}.
 \end{aligned}$$

由定义 2-5 可以看出, 本题中乘积 BA 是没有意义的.

例 2-3 求矩阵

$$A = \begin{pmatrix} 6 & 3 \\ -8 & -4 \end{pmatrix} \text{ 与 } B = \begin{pmatrix} -2 & 1 \\ 4 & -2 \end{pmatrix}$$

的乘积 AB 及 BA .

解 由定义 2-5 和 AB 与 BA 都有意义, 根据式 (2-3), 有:

$$\begin{aligned}
 AB &= \begin{pmatrix} 6 & 3 \\ -8 & -4 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ 4 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \\
 BA &= \begin{pmatrix} -2 & 1 \\ 4 & -2 \end{pmatrix} \begin{pmatrix} 6 & 3 \\ -8 & -4 \end{pmatrix} = \begin{pmatrix} -20 & -10 \\ 40 & 20 \end{pmatrix}.
 \end{aligned}$$

可见, 本题中虽然乘积 AB 和 BA 都是有意义的, 但 $AB \neq BA$.

综合以上两例可知, 在矩阵的乘法中必须注意矩阵相乘的顺序, AB 是 A 左乘 B (或 B 被 A 左乘) 的乘积, 而 BA 是 A 右乘 B (或 B 被 A 右乘) 的乘积.

AB 有意义时, BA 可以没有意义, 反之亦然; 若 AB 和 BA 都有意义, 如 A 是 $m \times n$ 矩阵, B 是 $n \times m$ 矩阵, 则 AB 是 m 阶方阵, BA 是 n 阶方阵, 当 $m \neq n$ 时 $AB \neq BA$, 即使 $m = n$, 如例 2-3, $m = n = 2$, AB 和 BA 都是二阶方阵, 但 AB 与 BA 仍然可以不相等.

总之, 矩阵的乘法不满足交换律, 即在一般情况下, $AB \neq BA$.

例 2-3 还表明, 两个都不为零(矩阵)的矩阵的乘积可以是一个零矩阵, 即 $AB = O$, 但 A 与 B 都不为 O , 这就进一步提醒读者要特别注意: 由 $AB = O$ 不能得出 $A = O$ 或 $B = O$ 的结论; 同理, 若 $A \neq O$, 而 $A(X - Y) = O$ (这里同型矩阵 X 、 Y 右乘 A 都有意义), 也不能得出 $X - Y = O$ 的结论.

矩阵的乘法虽然不满足交换律, 但仍然满足下列结合律和分配律 (假设运算都是有意义的, A 、 B 、 C 都为矩阵, λ 为常数), 证明从略:

- ① 结合律 $(AB)C = A(BC)$, $\lambda(AB) = (\lambda A)B = A(\lambda B)$;
- ② 分配律 $A(B + C) = AB + AC$, $(A + B)C = AC + BC$.

对于单位矩阵 I , 容易验证:

$$I_m A_{m \times n} = A_{m \times n}, \quad A_{m \times n} I_n = A_{m \times n},$$

或简写成:

$$IA = AI = A.$$

可见单位矩阵 I 在矩阵的乘法运算中的作用相当于数 1 在数的乘法运算中的作用.

矩阵 (λ 为常数)

$$\lambda I = \begin{pmatrix} \lambda & & & \\ & \lambda & & \\ & & \ddots & \\ & & & \lambda \end{pmatrix}$$

称为纯量阵, 也称数量矩阵, 由 $(\lambda I_n)A_n = I_n(\lambda A_n) = \lambda A_n$, $A_n(\lambda I_n) = (\lambda A_n)I_n = \lambda A_n$ 可知, 纯量阵 λI_n 与 n 阶方阵 A_n 的乘积等于数 λ 与 A_n 的乘积, 并且:

$$(\lambda I_n)A_n = \lambda A_n = A_n(\lambda I_n),$$

上式表示纯量阵 λI 与任何同阶方阵的乘积满足交换律.

有了矩阵的乘法 (运算), 就可以定义方阵的幂 (运算). 设 A 是 n 阶方阵, 定义

$$\begin{aligned} A^0 &= I, \quad A^1 = A^0 A = A, \quad A^2 = A^1 A = AA, \quad \dots, \\ A^k &= A^{k-1} A = \underbrace{AA \cdots A}_{k \text{ 个 } A} \quad (k=1, 2, 3, \dots) \end{aligned}$$

即 A 的 k 次幂 A^k 就是 k 个 A 连乘. 显然只有方阵的幂才有意义.

读者可以证明, 对角阵的 k 次幂同样为对角阵, 且其对角线上的元素为对应的原矩阵对角线上元素的 k 次幂, 即对角阵

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

的 k 次幂为

$$A^k = \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{pmatrix}.$$

满足 $A^2 = A$ 的方阵 A 称为幂等矩阵, 满足 $A^2 = I$ 的方阵 A 称为对合矩阵.

对于幂等矩阵和对合矩阵, 有如下定理:

定理 2-1 设 A 、 B 为 n 阶方阵, I 为 n 阶单位阵, 则:

- (1) A 为对合矩阵的充分必要条件是 $(I-A)(I+A) = O$;
- (2) 若 A 、 B 都是对合矩阵, 则 AB 为对合矩阵的充分必要条件是 $AB = BA$;
- (3) 若 A 、 B 都是幂等矩阵, 则 $A+B$ 为幂等矩阵的充分必要条件是 $AB = -BA$.

证 (1) 必要性. 若 A 为对合矩阵, 则 $A^2 = I$, 从而:

$$(I - A)(I + A) = I + A - A - A^2 = I - A^2 = O.$$

充分性. 若 $(I - A)(I + A) = O$, 即 $I - A^2 = O$, 则 $A^2 = I$.

(2) 必要性. 若 AB 为对合矩阵, 即 $(AB)^2 = I$. 又因为 A 、 B 都是对合矩阵, 所以 $A^2 = I, B^2 = I$, 从而:

$$A(AB - BA)B = A^2B^2 - (AB)^2 = I - (AB)^2 = O.$$

上式左乘矩阵 A 后再右乘矩阵 B , 得:

$$A^2(AB - BA)B^2 = AOB = O,$$

即 $AB - BA = O$, 从而 $AB = BA$.

充分性. 若 $AB = BA$, 则:

$$(AB)^2 = (AB)(AB) = A(BA)B = A(AB)B = A^2B^2 = II = I,$$

即 AB 为对合矩阵.

(3) 必要性. 若 $A + B$ 为幂等矩阵, 则 $(A + B)^2 = A + B$; 另一方面, 因为 A 、 B 都是幂等矩阵, 所以 $(A + B)^2 = A^2 + AB + BA + B^2 = A + B + AB + BA$. 从而有 $AB + BA = O$, 即 $AB = -BA$.

充分性. 若 $AB = -BA$, 则 $AB + BA = O$, 又 A 、 B 都是幂等矩阵, 从而有:

$$(A + B)^2 = A^2 + AB + BA + B^2 = A + B + AB + BA = A + B + O = A + B,$$

即 $A + B$ 为幂等矩阵. ■

由矩阵乘法的结合律可知, 矩阵的幂满足以下运算规律 (设 k, l 为正整数, λ 为常数):

$$(\lambda A)^k = \lambda^k A^k, \quad A^k A^l = A^{k+l}, \quad (A^k)^l = A^{kl}.$$

又因为矩阵乘法一般不满足交换律, 所以对两个 n 阶方阵 A 与 B , 一般说来, $(AB)^k \neq A^k B^k$, 只有当 A 与 B 可交换时才有 $(AB)^k = A^k B^k$; 类似地, 像完全平方公式 $(A \pm B)^2 = A^2 \pm 2AB + B^2$ 及平方差公式 $(A + B)(A - B) = A^2 - B^2$ 等也只有当乘积中的 A 与 B 可交换时才成立.

图 2-1 所示的四个城市间的单向航线的例子中, 由

$$A = (a_{ij}) = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} \text{ 有 } A^2 = (b_{ij}) = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 1 \end{pmatrix},$$

则 $b_{ij} (i \neq j)$ 表示从 i 市经一次中转到 j 市的单向航线条数, 而 $b_{ij} (i = j)$ 表示过 i (或 j) 市的双向航线条数. 例如:

$b_{11} = 2$ 表示过 1 市的双向航线有 2 条 ($\textcircled{1} \rightarrow \textcircled{2} \rightarrow \textcircled{1}$, $\textcircled{1} \rightarrow \textcircled{4} \rightarrow \textcircled{1}$, 参看图 2-1, 下同);

$b_{23}=1$ 表示从 2 市经一次中转到 3 市的单向航线有 1 条 ($② \rightarrow ① \rightarrow ③$);

$b_{33}=0$ 表示过 3 市没有双向航线;

$b_{42}=2$ 表示从 4 市经一次中转到 2 市的单向航线有 2 条 ($④ \rightarrow ① \rightarrow ②$, $④ \rightarrow ③ \rightarrow ②$).

利用矩阵的乘法, 可将线性变换 (2-2) 记作矩阵相乘的形式:

$$\mathbf{y} = \mathbf{A}\mathbf{x}.$$

其中:

$$\mathbf{A} = (a_{ij}), \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

这里, 列向量 (列矩阵) $\mathbf{x}$ 表示 n 个变量 $x_1, x_2, \dots, x_n$, 列向量 $\mathbf{y}$ 表示 m 个变量 $y_1, y_2, \dots, y_m$, 线性变换 (2-2) 把 $\mathbf{x}$ 变成 $\mathbf{y}$, 相当于用矩阵 $\mathbf{A}$ 左乘 $\mathbf{x}$ 得到 $\mathbf{y}$.

用矩阵 $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ 左乘向量 $\overrightarrow{OP} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, 相当于把向量 $\overrightarrow{OP}$ 投影到 xoy 坐标

面上 (参看图 2-2).

用矩阵 $\mathbf{A} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$ 左乘向量 $\overrightarrow{OP} = \begin{pmatrix} x \\ y \end{pmatrix}$, 相当于把 xoy 坐标平面中的向量

$\overrightarrow{OP}$ 绕原点 O 逆时针旋转 α 角 (参看图 2-3). 进一步还可推知, 以 $\mathbf{A}^n = \begin{pmatrix} \cos n\alpha & -\sin n\alpha \\ \sin n\alpha & \cos n\alpha \end{pmatrix}$

左乘向量 $\overrightarrow{OP}$, 相当于把 xoy 坐标平面中的向量 $\overrightarrow{OP}$ 绕原点 O 逆时针旋转 n 个 α 角, 即

旋转 $n\alpha$ 角, 这时所对应的矩阵应为 $\begin{pmatrix} \cos n\alpha & -\sin n\alpha \\ \sin n\alpha & \cos n\alpha \end{pmatrix}$, 即有:

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}^n = \begin{pmatrix} \cos n\alpha & -\sin n\alpha \\ \sin n\alpha & \cos n\alpha \end{pmatrix}.$$

用数学归纳法也可以证明上述结论. 事实上, 当 $n=1$ 时, 等式显然成立. 设 $n=k$ 时等式成立, 即设:

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}^k = \begin{pmatrix} \cos k\alpha & -\sin k\alpha \\ \sin k\alpha & \cos k\alpha \end{pmatrix},$$

则当 $n=k+1$ 时, 有:

$$\begin{aligned} & \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}^{k+1} \\ &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}^k \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
&= \begin{pmatrix} \cos k\alpha & -\sin k\alpha \\ \sin k\alpha & \cos k\alpha \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \\
&= \begin{pmatrix} \cos k\alpha \cos \alpha - \sin k\alpha \sin \alpha & -\cos k\alpha \sin \alpha - \sin k\alpha \cos \alpha \\ \sin k\alpha \cos \alpha + \cos k\alpha \sin \alpha & -\sin k\alpha \sin \alpha + \cos k\alpha \cos \alpha \end{pmatrix} \\
&= \begin{pmatrix} \cos(k+1)\alpha & -\sin(k+1)\alpha \\ \sin(k+1)\alpha & \cos(k+1)\alpha \end{pmatrix}.
\end{aligned}$$

即当 $n = k + 1$ 时等式成立.

2.2.2 矩阵的其他运算

1. 矩阵的转置

定义 2-6 矩阵 A 的行换成同序号的列后所得的新矩阵称为 A 的**转置矩阵**, 记作 A^T . 例如矩阵

$$A = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 0 & 4 \end{pmatrix}$$

的转置矩阵为:

$$A^T = \begin{pmatrix} 1 & 2 \\ 3 & 0 \\ -2 & 4 \end{pmatrix}.$$

显然, 行向量的转置为列向量, 列向量的转置为行向量.

满足 $A^T = A$ 的方阵称为**对称矩阵**, 简称**对称阵**. 对称阵的特点是: 它的元素以主对角线为对称轴对应相等. 例如单位阵 I 是对称阵, 矩阵

$$A = \begin{pmatrix} 1 & 3 & -4 \\ 3 & 2 & 0 \\ -4 & 0 & -3 \end{pmatrix}$$

也是对称阵.

若设列向量 $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, 则矩阵

$$X^T X = x_1^2 + x_2^2 + \cdots + x_n^2$$

是一个一阶方阵, 也就是一个数. 而

$$XX^T = \begin{pmatrix} x_1x_1 & x_1x_2 & \cdots & x_1x_n \\ x_2x_1 & x_2x_2 & \cdots & x_2x_n \\ \vdots & \vdots & & \vdots \\ x_nx_1 & x_nx_2 & \cdots & x_nx_n \end{pmatrix} = (x_ix_j)$$

是一个 n 阶方阵, 并且是对称阵.

矩阵的转置满足下列运算规律（假设运算都是有意义的）：

- ① $(A^T)^T = A$;
- ② $(A+B)^T = A^T + B^T$;
- ③ $(\lambda A) = \lambda A^T$ (λ 为常数);
- ④ $(AB)^T = B^T A^T$.

我们仅证明④，其余的证明留给读者．由定义 2-6，只需证明 AB 的 (j,i) 元即为 $B^T A^T$ 的 (i,j) 元即可．事实上，设 $A = (a_{ij})_{m \times p}$ ， $B = (b_{ij})_{p \times n}$ ，并记 AB 的 (j,i) 元为 $(AB)_{ji}$ ， $B^T A^T$ 的 (i,j) 元为 $(B^T A^T)_{ij}$ ．根据式 (2-3)，有：

$$(AB)_{ji} = \sum_{k=1}^p a_{jk} b_{ki},$$

而 B^T 的第 i 行为行向量 $(b_{1i}, b_{2i}, \dots, b_{pi})$ ， A^T 的第 j 列为列向量 $(a_{j1}, a_{j2}, \dots, a_{jp})^T$ ，因此：

$$(B^T A^T)_{ij} = \sum_{k=1}^p b_{ki} a_{jk} = \sum_{k=1}^p a_{jk} b_{ki},$$

即：

$$(AB)_{ji} = (B^T A^T)_{ij} \quad (j=1, 2, \dots, m; \quad i=1, 2, \dots, n),$$

从而：

$$(AB)^T = B^T A^T.$$

例 2-4 已知：

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

求 $(AB)^T$ ．

解 方法一： $AB = \begin{pmatrix} 1 & -3 & 0 \\ 2 & 1 & -1 \end{pmatrix} \begin{pmatrix} -4 & 1 & 2 \\ 3 & 4 & 5 \\ 1 & 0 & 3 \end{pmatrix} = \begin{pmatrix} -13 & -11 & -13 \\ -6 & 6 & 6 \end{pmatrix}$ ，所以：

$$(AB)^T = \begin{pmatrix} -13 & -6 \\ -11 & 6 \\ -13 & 6 \end{pmatrix}.$$

$$\text{方法二：} (AB)^T = B^T A^T = \begin{pmatrix} -4 & 3 & 1 \\ 1 & 4 & 0 \\ 2 & 5 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -3 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -13 & -6 \\ -11 & 6 \\ -13 & 6 \end{pmatrix}.$$

例 2-5 设列矩阵 $X = (x_1, x_2, \dots, x_n)^T$ 满足 $X^T X = 1$ ， I 为 n 阶单位阵， $P = I - 2XX^T$ ．证明 P 是对称阵，且 $PP^T = I$ ．

证 因为:

$$\begin{aligned} P^T &= (I - 2XX^T)^T \\ &= I^T - 2(XX^T)^T \\ &= I - 2XX^T = P, \end{aligned}$$

所以 P 是对称阵, 且:

$$\begin{aligned} PP^T &= P^2 = (I - 2XX^T)^2 \\ &= I^2 - 4XX^T + (2XX^T)^2 \\ &= I - 4XX^T + 4(XX^T)(XX^T) \\ &= I - 4XX^T + 4X(X^T X)X^T \\ &= I - 4XX^T + 4XX^T = I. \end{aligned}$$

2. 方阵的行列式

定义 2-7 设 A 为 n 阶方阵, 由 A 的元素所构成的行列式 (各元素的位置不变) 称为 A 的行列式, 记作 $|A|$ 或 $\det A$.

注意, 只有方阵才会有对应的行列式, 但方阵与行列式是两个完全不同的概念: n 阶方阵是由 n^2 个数按一定的方式 (长方形) 排成的数表或集合 (有时也称为阵列), 而 n 阶行列式则是这个阵列中的数按一定的运算法则所确定的一个数.

方阵行列式满足下列运算规律 (设 A 、 B 为 n 阶方阵, λ 为常数):

- ① $|A^T| = |A|$;
- ② $|\lambda A| = \lambda^n |A|$;
- ③ $|AB| = |A||B|$.

我们仅证明③, 其余的证明留给读者. 事实上, 设 $A = (a_{ij})$, $B = (b_{ij})$, 并记 $2n$ 阶行列式

$$D = \begin{vmatrix} a_{11} & \cdots & a_{1n} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} & 0 & \cdots & 0 \\ -1 & & & b_{11} & \cdots & b_{1n} \\ & \ddots & & \vdots & & \vdots \\ & & -1 & b_{n1} & \cdots & b_{nn} \end{vmatrix} = \begin{pmatrix} A & O \\ -I & B \end{pmatrix}.$$

由例 1-11 知, $D = |A||B|$, 而在 D 中以 b_{1j} 乘第 1 列, 以 b_{2j} 乘第 2 列, ..., 以 b_{nj} 乘第 n 列, 都加到第 $n+j$ 列上 ($j=1, 2, \dots, n$), 有:

$$D = \begin{vmatrix} A & C \\ -I & O \end{vmatrix},$$

其中 C 的 (i, j) 元为:

$$b_{1j}a_{i1} + b_{2j}a_{i2} + \cdots + b_{nj}a_{in} = \sum_{k=1}^n a_{ik}b_{kj},$$

即 $C = AB$. 再对 D 作第一初等行变换 $r_i \leftrightarrow r_{n+i} (i=1, 2, \dots, n)$, 有:

$$D = (-1)^n \begin{vmatrix} -I & O \\ A & C \end{vmatrix},$$

按例 1-11 的证法, 有:

$$D = (-1)^n |-I||C| = (-1)^n (-1)^n |I||C| = |C| = |AB|,$$

于是:

$$|AB| = |A||B|.$$

由③可知, 即使对于 n 阶方阵 A, B , 一般说来 $AB \neq BA$, 但总有 $|AB| = |BA|$.

3. 伴随矩阵

定义 2-8 设 A 为 n 阶方阵, 由 A 的行列式 $|A|$ 的各个元素的代数余子式 A_{ij} 为元素所构成的 n 阶方阵

$$A^* = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

称为矩阵 A 的**伴随矩阵**, 简称**伴随阵**.

注意, 伴随矩阵 A^* 的 (i, j) 元是行列式 $|A|$ 的 (j, i) 元的代数余子式.

方阵 A 与其伴随矩阵 A^* 满足如下的等量关系 (I 是与 A 同阶的单位矩阵):

$$AA^* = A^*A = |A|I. \quad (2-4)$$

事实上, 设 $A = (a_{ij})$, $(AA^*)_{ij}$ 、 $(A^*A)_{ij}$ 分别表示矩阵 AA^* 、 A^*A 的 (i, j) 元, 则由定理 1-4 及其推论, 有:

$$(AA^*)_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = |A|\delta_{ij},$$

$$(A^*A)_{ij} = a_{1j}A_{i1} + a_{2j}A_{i2} + \cdots + a_{nj}A_{in} = |A|\delta_{ij},$$

故:

$$AA^* = A^*A = (|A|\delta_{ij}) = |A|(\delta_{ij}) = |A|I.$$

*4. 复矩阵的一些运算

在实际问题中, 我们还可能遇到与**复矩阵** (即元素为复数的矩阵) 相关的一些运算.

定义 2-9 设 $A = (a_{ij})$ 为复矩阵, 若用 $\bar{a}_{ij}$ 表示 a_{ij} 的共轭复数, 并记

$$\bar{A} = (\bar{a}_{ij}), \quad A^H = \bar{A}^T,$$

$\bar{A}$ 称为 A 的**共轭矩阵**, A^H 称为 A 的**共轭转置矩阵**, 满足 $A^H = A$ 的复方阵称为**埃尔米特 (Hermitian) 矩阵**.

复矩阵的运算满足下列运算规律 (设 A, B 为复矩阵, λ 为复常数, 且运算都是有意义的), 证明从略:

- ① $\overline{A^T} = \overline{A}^T$, $(A^H)^H = A$;
 ② $\overline{A+B} = \overline{A} + \overline{B}$, $(A+B)^H = A^H + B^H$;
 ③ $\overline{\lambda A} = \overline{\lambda} \overline{A}$, $(\lambda A)^H = \overline{\lambda} A^H$;
 ④ $\overline{AB} = \overline{A} \overline{B}$, $(AB)^H = B^H A^H$.

由性质①和性质④可知, 若 A 是一个复矩阵, 则 AA^H 与 $A^H A$ 都是埃尔米特矩阵.

习题 2.2

1. 用矩阵相乘的形式表示习题 2.1 中题 1 的线性变换.
 2. 计算下列乘积:

$$(1) (1, 2, 3) \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}; \quad (2) \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} (1, -2);$$

$$(3) \begin{pmatrix} 4 & 3 & 1 \\ 1 & -2 & 3 \\ 5 & 7 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 7 \end{pmatrix}; \quad (4) \begin{pmatrix} 2 & 1 & 4 & 0 \\ 1 & -1 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 3 & 1 \\ 0 & -1 & 2 \\ 1 & -3 & 1 \\ 4 & 0 & -2 \end{pmatrix};$$

$$(5) (x_1, x_2, x_3) \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

3. 设 $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 & 3 \\ -1 & -2 & 4 \\ 0 & 5 & 1 \end{pmatrix}$, 求 $3AB - 2A$ 和 $A^T B$.

4. 已知两个线性变换:

$$\begin{cases} x_1 = 2y_1 + y_3 \\ x_2 = -2y_1 + 3y_2 + 2y_3 \\ x_3 = 4y_1 + y_2 + 5y_3 \end{cases}, \quad \begin{cases} y_1 = -3z_1 + z_2 \\ y_2 = 2z_1 + z_3 \\ y_3 = -z_2 + 3z_3 \end{cases},$$

求从 z_1, z_2, z_3 到 x_1, x_2, x_3 的线性变换.

5. 设 $A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$, 问:

- (1) $AB = BA$ 成立吗?
 (2) $(A+B)^2 = A^2 + 2AB + B^2$ 成立吗?
 (3) $(A+B)(A-B) = A^2 - B^2$ 成立吗?

6. 举反例说明下列命题是错误的:

- (1) 若 $A^2 = O$, 则 $A = O$;
 (2) 若 $A^2 = A$, 则 $A = O$ 或 $A = I$;
 (3) 若 $AX = AY$ 且 $A \neq O$, 则 $X = Y$.

7. 设 $A = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix}$, 求 $A^2, A^3, \dots, A^k$.

8. 设 $A = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$, 求 A^n .

9. 设 A 、 B 为 n 阶矩阵, 且 A 为对称阵, 证明 $B^T AB$ 也是对称阵.

10. 设 A 、 B 都是 n 阶对称阵, 证明 AB 是对称阵的充分必要条件是 $AB = BA$.

2.3 逆矩阵

当 $m = n$ 时, 线性变换 (2-2) 变为:

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\ \quad \quad \quad \dots\dots\dots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{cases} \quad (2-5)$$

其表示成矩阵相乘的形式为:

$$y = Ax. \quad (2-6)$$

其中, 系数矩阵 $A = (a_{ij})$ 是一个 n 阶方阵.

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

若 A 和 y 已知, 要从式 (2-6) 解出 x , 我们以 A 的伴随矩阵 A^* 左乘式 (2-6) 两端, 并利用式 (2-4), 可得:

$$A^*y = A^*Ax = |A|Ix = |A|x.$$

故当 $|A| \neq 0$ 时, 可解出:

$$x = \frac{1}{|A|} A^*y,$$

若记 $B = \frac{1}{|A|} A^*$, 则上式记作:

$$x = By. \quad (2-7)$$

式(2-7)用矩阵的乘法表示了从 y 到 x 的线性变换,称为线性变换(2-5)的逆变换.

下面通过式(2-6)和式(2-7)分析它们对应的变换矩阵 A 与 B 之间的关系.将式(2-7)代入式(2-6),可得:

$$y = Ax = A(By) = (AB)y,$$

可见 AB 为恒等变换的变换矩阵,即 $AB = I$.

类似地,将式(2-6)代入式(2-7),可得:

$$x = By = B(Ax) = (BA)x,$$

从而又有 $BA = I$.于是有:

$$AB = BA = I.$$

由此我们引入逆矩阵的定义.

定义 2-10 (逆矩阵) 对于 n 阶矩阵 A ,如果存在一个 n 阶矩阵 B ,使

$$AB = BA = I,$$

则称矩阵 A 是**可逆的**(或 A 是**可逆矩阵**),并将 B 称为 A 的**逆矩阵**,简称**逆阵**.

显然,若矩阵 A 是可逆的,则它的逆矩阵是唯一的.事实上,设 B 和 C 都是 A 的逆矩阵,则有:

$$B = BI = B(AC) = (BA)C = IC = C.$$

由此,我们常把 A 的逆矩阵记作 A^{-1} .即若 $AB = BA = I$,则 $B = A^{-1}$.

定理 2-2 矩阵 A 可逆的充要条件是 $|A| \neq 0$,这时:

$$A^{-1} = \frac{1}{|A|} A^*, \quad (2-8)$$

其中 A^* 为 A 的伴随矩阵.

证 必要性.若 A 可逆,则 $AA^{-1} = I$,即 $|A||A^{-1}| = |AA^{-1}| = |I| = 1$,故 $|A| \neq 0$.

充分性.若 $|A| \neq 0$,则由 $AA^* = A^*A = |A|I$ 及矩阵乘法的运算规律知:

$$A \frac{1}{|A|} A^* = \frac{1}{|A|} A^* A = I,$$

根据逆矩阵的定义,矩阵 A 可逆.

等式 $A \frac{1}{|A|} A^* = \frac{1}{|A|} A^* A = I$ 同时也说明:

$$A^{-1} = \frac{1}{|A|} A^*. \blacksquare$$

当 $|A| = 0$ 时,称 A 为奇异矩阵,否则称 A 为非奇异矩阵.定理 2-2 说明,可逆矩阵就是非奇异矩阵.

推论 若 $AB = I$ (或 $BA = I$),则 $B = A^{-1}$.

证 不妨设 $AB = I$,则有 $|A||B| = |I| = 1$,故 $|A| \neq 0$,因而 A^{-1} 存在,于是:

$$B = IB = (A^{-1}A)B = A^{-1}(AB) = A^{-1}I = A^{-1}. \blacksquare$$

逆矩阵满足下列运算规律:

- ① 若 A 可逆, 则 A^{-1} 也可逆, 且 $(A^{-1})^{-1} = A$;
- ② 若 A 可逆, 则当 $\lambda \neq 0$ 时, 数乘矩阵 λA 也可逆, 且 $(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}$;
- ③ 若 A, B 为同阶矩阵且均可逆, 则 AB 也可逆, 且 $(AB)^{-1} = B^{-1}A^{-1}$;
- ④ 若 A 可逆, 则 A^T 也可逆, 且 $(A^T)^{-1} = (A^{-1})^T$.

根据逆矩阵的定义及矩阵乘法的运算规律, ①和②显然是成立的; 对于③, 由于:

$$AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I,$$

由推论知, $(AB)^{-1} = B^{-1}A^{-1}$;

对于④, 由于:

$$A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I,$$

故 $(A^T)^{-1} = (A^{-1})^T$.

当 A 可逆且 k 为正整数时, 还可定义:

$$A^{-k} = (A^{-1})^k,$$

这样, 当 A 可逆且 λ, μ 为整数时, 有:

$$A^\lambda A^\mu = A^{\lambda+\mu}, \quad (A^\lambda)^\mu = A^{\lambda\mu}.$$

例 2-6 求二阶矩阵

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

的逆矩阵, 其中 $ad - bc \neq 0$.

解 因为 $|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \neq 0$, 所以矩阵 A 可逆, 根据式 (2-8), 有:

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

例 2-7 求方阵

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 3 & 4 & 3 \end{pmatrix}$$

的逆矩阵.

解 因为 $|A| = \begin{vmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 3 & 4 & 3 \end{vmatrix} = 16 \neq 0$, 所以矩阵 A 可逆, 而 $|A|$ 的余子式为:

$$M_{11} = -5, \quad M_{12} = 3, \quad M_{13} = 9,$$

$$M_{21} = -6, \quad M_{22} = -6, \quad M_{23} = -2,$$

$$M_{31} = 1, \quad M_{32} = -7, \quad M_{33} = -5,$$

故:

$$A^* = \begin{pmatrix} M_{11} & -M_{21} & M_{31} \\ -M_{12} & M_{22} & -M_{32} \\ M_{13} & -M_{23} & M_{33} \end{pmatrix} = \begin{pmatrix} -5 & 6 & 1 \\ -3 & -6 & 7 \\ 9 & 2 & -5 \end{pmatrix},$$

所以:

$$A^{-1} = \frac{1}{|A|} A^* = \frac{1}{16} \begin{pmatrix} -5 & 6 & 1 \\ -3 & -6 & 7 \\ 9 & 2 & -5 \end{pmatrix} = \begin{pmatrix} -\frac{5}{16} & \frac{3}{8} & \frac{1}{16} \\ -\frac{3}{16} & -\frac{3}{8} & \frac{7}{16} \\ \frac{9}{16} & \frac{1}{8} & -\frac{5}{16} \end{pmatrix}.$$

例 2-8 设

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 3 & 4 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ 0 & 1 \end{pmatrix},$$

求矩阵 U , 使其满足:

$$AUB = C.$$

解 根据题意, $|A| = 16 \neq 0$, $|B| = 2 \neq 0$, 故 A^{-1} 和 B^{-1} 存在. 要使 $AUB = C$, 只需:

$$U = A^{-1}AUBB^{-1} = A^{-1}CB^{-1}.$$

而利用上例及例 2-6 的结果, 有:

$$A^{-1} = \begin{pmatrix} -\frac{5}{16} & \frac{3}{8} & \frac{1}{16} \\ -\frac{3}{16} & -\frac{3}{8} & \frac{7}{16} \\ \frac{9}{16} & \frac{1}{8} & -\frac{5}{16} \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} \\ -2 & 1 \end{pmatrix},$$

于是:

$$\begin{aligned}
 U = A^{-1}CB^{-1} &= \begin{pmatrix} -\frac{5}{16} & \frac{3}{8} & \frac{1}{16} \\ -\frac{3}{16} & -\frac{3}{8} & \frac{7}{16} \\ \frac{9}{16} & \frac{1}{8} & -\frac{5}{16} \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} \\ -2 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} \frac{7}{16} & -\frac{7}{8} \\ -\frac{15}{16} & -\frac{1}{8} \\ \frac{13}{16} & \frac{11}{8} \end{pmatrix} \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} \frac{77}{32} & -\frac{35}{32} \\ -\frac{37}{32} & \frac{11}{32} \\ -\frac{49}{32} & \frac{31}{32} \end{pmatrix}.
 \end{aligned}$$

例 2-9 设 $P = \begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix}$, $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, $AP = PA$, 求 A^n .

解 由 $|P| = 2$ 知, $P^{-1} = \frac{1}{2} \begin{pmatrix} 4 & -2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$. 又由 $AP = PA$ 知:

$$A = PAP^{-1}, \quad A^2 = PAP^{-1}PA P^{-1} = PA^2P^{-1}, \quad \dots, \quad A^n = PA^nP^{-1},$$

而 $A^n = \begin{pmatrix} 1 & 0 \\ 0 & 2^n \end{pmatrix}$, 从而:

$$\begin{aligned}
 A^n &= \begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2^n \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -2^{n-1} & 2^{n-1} \end{pmatrix} = \begin{pmatrix} 2-2^n & 2^n-1 \\ 2-2^{n+1} & 2^{n+1}-1 \end{pmatrix}.
 \end{aligned}$$

定义 2-11 (矩阵多项式) 设

$$\varphi(x) = a_0 + a_1x + \dots + a_mx^m \quad (a_m \neq 0)$$

为关于 x 的 m 次多项式, A 为 n 阶矩阵, I 为 n 阶单位阵, 记:

$$\varphi(A) = a_0I + a_1A + \dots + a_mA^m,$$

则称 $\varphi(A)$ 为矩阵 A 的 m 次多项式.

因为方阵 A 的各次幂及单位矩阵 I 所涉及的乘法运算都满足交换律, 所以方阵 A 的两个多项式 $\varphi_1(A)$ 和 $\varphi_2(A)$ 的乘积也满足交换律, 即总有:

$$\varphi_1(A)\varphi_2(A) = \varphi_2(A)\varphi_1(A),$$

从而 A 的几个多项式可以像数 x 的多项式一样相乘或分解. 例如:

$$(I - A)^3 = I - 3A + 3A^2 - A^3,$$

$$I - A - 6A^2 = (I - 3A)(I + 2A).$$

我们常用例 2-9 中计算 A^k 的方法来计算 A 的多项式 $\varphi(A)$, 即如果 $A = P\Lambda P^{-1}$, 则 $A^k = P\Lambda^k P^{-1}$, 从而:

$$\begin{aligned}\varphi(A) &= a_0 I + a_1 A + \cdots + a_m A^m \\ &= Pa_0 I P^{-1} + Pa_1 \Lambda P^{-1} + \cdots + Pa_m \Lambda^m P^{-1} \\ &= P(a_0 I + a_1 \Lambda + \cdots + a_m \Lambda^m) P^{-1} \\ &= P\varphi(\Lambda) P^{-1}.\end{aligned}$$

进而, 如果 $\Lambda = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 为对角阵, 则 $\Lambda^k = \text{diag}(\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k)$, 从而:

$$\begin{aligned}\varphi(\Lambda) &= a_0 I + a_1 \Lambda + \cdots + a_m \Lambda^m \\ &= a_0 \begin{pmatrix} 1 & & \\ & 1 & \\ & & \ddots \\ & & & 1 \end{pmatrix} + a_1 \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{pmatrix} + \cdots + a_m \begin{pmatrix} \lambda_1^m & & \\ & \lambda_2^m & \\ & & \ddots \\ & & & \lambda_n^m \end{pmatrix} \\ &= \begin{pmatrix} \varphi(\lambda_1) & & \\ & \varphi(\lambda_2) & \\ & & \ddots \\ & & & \varphi(\lambda_n) \end{pmatrix}.\end{aligned}$$

例 2-10 设 $P = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 2 & -1 \end{pmatrix}$, $A = \begin{pmatrix} 1 & & \\ & 2 & \\ & & -1 \end{pmatrix}$, $AP = PA$, 求 $\varphi(A) =$

$$A^3 + 2A^2 - 3A.$$

解 由 $|P| = 6$ 知 P 可逆, 从而:

$$A = P\Lambda P^{-1}, \quad \varphi(A) = P\varphi(\Lambda)P^{-1}.$$

而由 $\varphi(1) = 0$, $\varphi(2) = 10$, $\varphi(-1) = 4$ 知 $\varphi(\Lambda) = \text{diag}(0, 10, 4)$, 于是:

$$\begin{aligned}\varphi(A) &= P\varphi(\Lambda)P^{-1} \\ &= \begin{pmatrix} -1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} 0 & & \\ & 10 & \\ & & 4 \end{pmatrix} \frac{1}{|P|} P^* \\ &= \frac{2}{6} \begin{pmatrix} 0 & 5 & 2 \\ 0 & 0 & 2 \\ 0 & 10 & -2 \end{pmatrix} \begin{pmatrix} -2 & 3 & 1 \\ 2 & 0 & 2 \\ 2 & 3 & -1 \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} 14 & 6 & 8 \\ 4 & 6 & -2 \\ 16 & -6 & 22 \end{pmatrix} = \frac{2}{3} \begin{pmatrix} 7 & 3 & 4 \\ 2 & 3 & -1 \\ 8 & -3 & 11 \end{pmatrix}.\end{aligned}$$

习题 2.3

1. 求下列矩阵的逆矩阵:

(1) $\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix};$

(2) $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix};$

(3) $\begin{pmatrix} 1 & 2 & -1 \\ 3 & 4 & -2 \\ 5 & -4 & 1 \end{pmatrix};$

(4) $\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix} \quad (a_1 a_2 \cdots a_n \neq 0).$

2. 利用逆矩阵解下列线性方程组:

(1) $\begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ 2x_1 + 2x_2 + 5x_3 = 2 \\ 3x_1 + 5x_2 + x_3 = 3 \end{cases};$

(2) $\begin{cases} x_1 - x_2 - x_3 = 2 \\ 2x_1 - x_2 - 3x_3 = 1 \\ 3x_1 + 2x_2 - 5x_3 = 0 \end{cases}.$

3. 已知线性变换:

$$\begin{cases} x_1 = 2y_1 + 2y_2 + y_3 \\ x_2 = 3y_1 + y_2 + 5y_3 \\ x_3 = 3y_1 + 2y_2 + 3y_3 \end{cases},$$

求从变量 y_1, y_2, y_3 到变量 x_1, x_2, x_3 的线性变换.4. 设 A 为 3 阶矩阵且 $|A| = \frac{1}{2}$, 求 $|(2A)^{-1} - 5A^*|$.5. 设 $A = \begin{pmatrix} 0 & 3 & 3 \\ 1 & 1 & 0 \\ -1 & 2 & 3 \end{pmatrix}$, $AB = A + 2B$, 求 B .6. 设 $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$, $AB + I = A^2 + B$, 求 B .7. 设 $A = \text{diag}(1, -2, 1)$ 且 $A^*BA = 2BA - 8I$, 求 B .8. 已知矩阵 A 的伴随阵 $A^* = \text{diag}(1, 1, 1, 8)$ 且 $ABA^{-1} = BA^{-1} + 3I$, 求 B .9. 设 $P^{-1}AP = A$, 其中 $P = \begin{pmatrix} -1 & -4 \\ 1 & 1 \end{pmatrix}$, $A = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$, 求 A^{11} .

10. 设 $AP = PA$, 其中 $P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix}$, $A = \text{diag}(-1, 1, 5)$, 求 $\varphi(A) =$

$$A^8(5I - 6A + A^2).$$

11. 设 $A^k = O$ (k 为正整数), 证明 $(I - A)^{-1} = I + A + A^2 + \cdots + A^{k-1}$.

12. 设方阵 A 满足 $A^2 - A - 2I = O$, 证明 A 和 $A + 2I$ 都可逆, 并求 A^{-1} 和 $(A + 2I)^{-1}$.

13. 设矩阵 A 可逆, 证明其伴随矩阵 A^* 也可逆, 且 $(A^*)^{-1} = (A^{-1})^*$.

14. 设 n 阶矩阵 A 的伴随矩阵为 A^* , 证明:

(1) 若 $|A| = 0$, 则 $|A^*| = 0$; (2) $|A^*| = |A|^{n-1}$.

15. 满足下列条件的 n 阶矩阵 $A = (a_{ij})$ 称为严格对角占优矩阵: 对任意 $i = 1, 2, \dots, n$, 都有:

$$|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}|,$$

试证明严格对角占优矩阵是可逆的.

2.4 矩阵的分块

对于行数和列数都比较高的矩阵, 常采用分块的方法来简化运算, 我们先给出分块矩阵的定义.

定义 2-12 (分块矩阵) 矩阵 A 用若干条纵虚线和横虚线分成许多个小块, 以每一块为形式上的元素放在相应位置所得到的矩阵称为 A 的分块矩阵.

通常将分块矩阵每一个小矩阵 (也称为原矩阵的子块) 用原矩阵字母加下标符号的矩阵来表示, 其下标用来说明该子块在分块矩阵中相应的行列位置.

例如可将 3×4 矩阵

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}$$

按下述四种方法 (当然还有其他方法) 分块:

$$(i) A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}; \quad (ii) A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix};$$

$$(iii) \quad A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}; \quad (iv) \quad A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}.$$

分法 (i) 可记为:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

其中:

$$A_{11} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad A_{12} = \begin{pmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{pmatrix},$$

$$A_{21} = \begin{pmatrix} a_{31} & a_{32} \end{pmatrix}, \quad A_{22} = \begin{pmatrix} a_{33} & a_{34} \end{pmatrix}.$$

即 $A_{11}, A_{12}, A_{21}, A_{22}$ 为 A 的子块, 而 A 形式上称为以这些子块为元素的分块矩阵.

同理, 分法 (ii) 可记为:

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix},$$

这时

$$A_{11} = (a_{11}), \quad A_{12} = (a_{12} \quad a_{13}), \quad A_{13} = (a_{14}),$$

$$A_{21} = \begin{pmatrix} a_{21} \\ a_{31} \end{pmatrix}, \quad A_{22} = \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}, \quad A_{23} = \begin{pmatrix} a_{24} \\ a_{34} \end{pmatrix}.$$

分法 (iii) 和分法 (iv) 的分块矩阵请读者自己给出.

注意记号 A_{ij} 与 A_{ij} 的区别: 前者 (粗体) 表示矩阵 A 的分块矩阵在第 i 行第 j 列位置的子块, 而后者表示行列式 $|A|$ 的 (i, j) 元 a_{ij} 的代数余子式.

分块矩阵的运算规则与普通矩阵类似, 现分别说明如下:

(1) 设 λ 为常数, A 与 B 是同型矩阵, 并且具有相同的分块方法, 如:

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1s} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rs} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & \cdots & B_{1s} \\ \vdots & & \vdots \\ B_{r1} & \cdots & B_{rs} \end{pmatrix},$$

其中 A_{ij} 与 B_{ij} 为同型矩阵 ($i=1, 2, \dots, r; j=1, 2, \dots, s$), 那么:

$$A+B = \begin{pmatrix} A_{11}+B_{11} & \cdots & A_{1s}+B_{1s} \\ \vdots & & \vdots \\ A_{r1}+B_{r1} & \cdots & A_{rs}+B_{rs} \end{pmatrix},$$

$$\lambda A = \begin{pmatrix} \lambda A_{11} & \cdots & \lambda A_{1s} \\ \vdots & & \vdots \\ \lambda A_{r1} & \cdots & \lambda A_{rs} \end{pmatrix}.$$

(2) 设 A 为 $m \times p$ 矩阵, B 为 $p \times n$ 矩阵, 分块成:

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1t} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rt} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & \cdots & B_{1s} \\ \vdots & & \vdots \\ B_{t1} & \cdots & B_{ts} \end{pmatrix},$$

其中 $A_{11}, A_{12}, \dots, A_{1t}$ 的列数对应等于 $B_{1j}, B_{2j}, \dots, B_{tj}$ 的行数, 那么:

$$AB = \begin{pmatrix} C_{11} & \cdots & C_{1s} \\ \vdots & & \vdots \\ C_{r1} & \cdots & C_{rs} \end{pmatrix},$$

其中:

$$C_{ij} = \sum_{k=1}^t A_{ik} B_{kj} \quad (i=1, 2, \dots, r; j=1, 2, \dots, s).$$

例 2-11 设

$$A = \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 2 & 0 & -1 \\ 1 & 0 & 4 & 2 \\ 0 & 1 & 1 & 0 \end{pmatrix},$$

求 AB .

解 把 A, B 分块如下:

$$A = \left(\begin{array}{cc|cc} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & 1 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right) = \begin{pmatrix} I & A_{12} \\ O & I \end{pmatrix}, \quad B = \left(\begin{array}{cc|cc} 1 & -1 & 1 & -1 \\ 0 & 2 & 0 & -1 \\ \hline 1 & 0 & 4 & 2 \\ 0 & 1 & 1 & 0 \end{array} \right) = \begin{pmatrix} B_{11} & B_{12} \\ I & B_{22} \end{pmatrix},$$

则:

$$AB = \begin{pmatrix} I & A_{12} \\ O & I \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ I & B_{22} \end{pmatrix} = \begin{pmatrix} B_{11} + A_{12} & B_{12} + A_{12}B_{22} \\ I & B_{22} \end{pmatrix},$$

而:

$$B_{11} + A_{12} = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} -1 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 2 & 3 \end{pmatrix},$$

$$B_{12} + A_{12}B_{22} = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} -1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} -3 & -2 \\ 9 & 4 \end{pmatrix} = \begin{pmatrix} -2 & -3 \\ 9 & 3 \end{pmatrix},$$

于是:

$$AB = \left(\begin{array}{cc|cc} 0 & 0 & -2 & -3 \\ 2 & 3 & 9 & 3 \\ \hline 1 & 0 & 4 & 2 \\ 0 & 1 & 1 & 0 \end{array} \right).$$

注意, 为了使分块后的矩阵计算更方便, 应尽可能使分成的小块中 $\mathbf{O}$ 矩阵或单位矩阵多一些.

$$(3) \text{ 设 } A = \begin{pmatrix} A_{11} & \cdots & A_{1s} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rs} \end{pmatrix}, \text{ 则 } A^T = \begin{pmatrix} A_{11}^T & \cdots & A_{r1}^T \\ \vdots & & \vdots \\ A_{1s}^T & \cdots & A_{rs}^T \end{pmatrix}.$$

(4) 设 A 为 n 阶方阵, 若 A 的分块矩阵中的非零子块都是方阵, 且只出现在对角线上, 非对角线上的子块都为零矩阵, 即:

$$A = \begin{pmatrix} A_1 & & \mathbf{O} \\ & A_2 & \\ & & \ddots \\ \mathbf{O} & & & A_p \end{pmatrix},$$

其中 $A_i (i=1, 2, \dots, p)$ 都是方阵, 那么 A 叫做分块对角矩阵.

上述分块对角矩阵的行列式具有性质:

$$|A| = |A_1| |A_2| \cdots |A_p|.$$

由此性质知, 若 $|A_i| \neq 0 (i=1, 2, \dots, p)$, 则 $|A| \neq 0$, A^{-1} 仍然是分块对角矩阵, 且:

$$A^{-1} = \begin{pmatrix} A_1^{-1} & & \mathbf{O} \\ & A_2^{-1} & \\ & & \ddots \\ \mathbf{O} & & & A_p^{-1} \end{pmatrix}.$$

推而广之, 若 A 的分块矩阵的主对角及其上方的子块都是方阵, 而主对角线下方的所有子块都为零矩阵, 即有:

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1p} \\ & A_{22} & \cdots & A_{2p} \\ & & \ddots & \\ \mathbf{O} & & & A_{pp} \end{pmatrix},$$

则 A 叫做分块上三角矩阵.

例 2-12 设 $A = \begin{pmatrix} 1 & 3 & 0 & 0 & 0 \\ 1 & 4 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$, 求 A^{-1} .

$$\text{解 } A = \begin{pmatrix} 1 & 3 & 0 & 0 & 0 \\ 1 & 4 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} A_1 & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & A_2 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & A_3 \end{pmatrix},$$

$$\text{而由: } A_1 = \begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix} \text{ 得 } A_1^{-1} = \begin{pmatrix} 4 & -3 \\ -1 & 1 \end{pmatrix},$$

$$A_2 = (5) \quad \text{得 } A_2^{-1} = \left(\frac{1}{5}\right),$$

$$A_3 = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \text{ 得 } A_3^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

$$\text{所以: } A^{-1} = \begin{pmatrix} A_1^{-1} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & A_2^{-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & A_3^{-1} \end{pmatrix} = \begin{pmatrix} 4 & -3 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{5} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

两种特殊的矩阵分块在本书后续章节中经常用到, 这就是按行分块和按列分块.
对于 $m \times n$ 矩阵

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

记:

$$\alpha_i^T = (a_{i1}, a_{i2}, \cdots, a_{in}) \quad (i=1, 2, \cdots, m), \quad a_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad (j=1, 2, \cdots, n),$$

分别表示 A 的 m 个行向量和 n 个列向量, 则 A 可写成:

$$A = \begin{pmatrix} \alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_m^T \end{pmatrix}$$

或

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

这里的 α_i^T ($i=1,2,\cdots,m$) 其实是列向量 $\alpha_i = \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{pmatrix}$ 的转置向量, 故可以参与转置矩阵相关的运算.

通常, 我们使用小写粗体字母带下标 (有时也可以不带下标) 来表示列向量 (列矩阵), 使用小写粗体字母的转置带下标 (有时也可以不带下标) 来表示行向量 (行矩阵), 这是一种惯用的表示方法, 以后还会用到.

按照上述表示方式, $m \times p$ 矩阵 $A = (a_{ij})$ 与 $p \times n$ 矩阵 $B = (b_{ij})$ 的乘积矩阵 $C = AB = (c_{ij})_{m \times n}$ 可以写成:

$$AB = \begin{pmatrix} \alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_m^T \end{pmatrix} (b_1, b_2, \cdots, b_n) = \begin{pmatrix} \alpha_1^T b_1 & \alpha_1^T b_2 & \cdots & \alpha_1^T b_n \\ \alpha_2^T b_1 & \alpha_2^T b_2 & \cdots & \alpha_2^T b_n \\ \vdots & \vdots & & \vdots \\ \alpha_m^T b_1 & \alpha_m^T b_2 & \cdots & \alpha_m^T b_n \end{pmatrix} = (c_{ij})_{m \times n},$$

其中 $\begin{pmatrix} \alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_m^T \end{pmatrix}$ 是将矩阵 A 按行分成 m 块, $(b_1, b_2, \cdots, b_n)$ 是将矩阵 B 按列分成 n 块.

$$c_{ij} = \alpha_i^T b_j = (a_{i1}, a_{i2}, \cdots, a_{ip}) \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{pmatrix} = \sum_{k=1}^p a_{ik} b_{kj}.$$

对于 $m \times n$ 矩阵 A , 我们还有以下重要定理:

定理 2-3 矩阵 $A = O$ 的充分必要条件是方阵 $A^T A = O$.

证 定理的必要性是显然成立的. 下面我们来证明充分性. 设:

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

将 A 按列分块, 即 $A = (a_1, a_2, \cdots, a_n)$, 其中 $a_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$ ($j=1,2,\cdots,n$), 则由

$$A^T A = \begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{pmatrix} (a_1, a_2, \dots, a_n) = \begin{pmatrix} a_1^T a_1 & a_1^T a_2 & \cdots & a_1^T a_n \\ a_2^T a_1 & a_2^T a_2 & \cdots & a_2^T a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n^T a_1 & a_n^T a_2 & \cdots & a_n^T a_n \end{pmatrix} = O$$

知:

$$a_i^T a_j = 0 \quad (i=1, 2, \dots, n; j=1, 2, \dots, n),$$

特别地, 当 $j=1, 2, \dots, n$ 时, 有:

$$a_j^T a_j = 0,$$

$$\text{即: } a_j^T a_j = (a_{1j}, a_{2j}, \dots, a_{mj}) \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = a_{1j}^2 + a_{2j}^2 + \cdots + a_{mj}^2 = 0.$$

又因为 $a_{1j}, a_{2j}, \dots, a_{mj}$ 为实数, 故:

$$a_{1j} = a_{2j} = \cdots = a_{mj} = 0 \quad (j=1, 2, \dots, n),$$

于是:

$$A = O.$$

证毕.

本定理阐明了矩阵 A 与方阵 $A^T A$ 之间的一种关系. 特别地, 当 A 为一列向量 a 时, 由于 $a^T a$ 是一个 1×1 矩阵, 也就是一个数, 这时定理 2-3 可叙述为: 列向量 $a = 0$ 的充分必要条件是 $a^T a = 0$.

根据矩阵与其转置矩阵的关系, 定理 2-3 中的 $A^T A = O$ 也可改为 $AA^T = O$.

下面我们分析矩阵与对角矩阵相乘的计算方法.

设 $A_m = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m)$ 是 m 阶对角阵, A 是按行分块的 $m \times n$ 矩阵, 即:

$$A = \begin{pmatrix} \alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_m^T \end{pmatrix},$$

则:

$$A_m A = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{pmatrix} \begin{pmatrix} \alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_m^T \end{pmatrix} = \begin{pmatrix} \lambda_1 \alpha_1^T \\ \lambda_2 \alpha_2^T \\ \vdots \\ \lambda_m \alpha_m^T \end{pmatrix},$$

可见以对角阵左乘矩阵 A (如果可以相乘的话) 的结果是 A 的每一行乘以对角阵中与该行同行号的行中的对角元.

同理, 设 $A_n = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ 是 n 阶对角阵, A 是按列分块的 $m \times n$ 矩阵, 即:

$$A = (a_1, a_2, \dots, a_n),$$

则:

$$AA_n = (a_1, a_2, \dots, a_n) \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{pmatrix} = (\lambda_1 a_1, \lambda_2 a_2, \dots, \lambda_n a_n),$$

可见以对角阵右乘矩阵 A (如果可以相乘的话) 的结果是 A 的每一列乘以对角阵中与该列同列号的列中的对角元.

重写线性方程组 (1-29) 如下:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases}, \quad (2-9)$$

记:

$$A = (a_{ij}), \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \quad B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \end{pmatrix},$$

这时, A 叫做系数矩阵, x 叫做未知数向量, b 叫做常数项向量, B 叫做增广矩阵, 按分块矩阵的记法, B 也可写作:

$$B = (A:b) \text{ 或 } B = (A, b) = (a_1, a_2, \dots, a_n, b),$$

利用矩阵的乘法, 方程组 (2-9) 可记作:

$$Ax = b. \quad (2-10)$$

方程组 (2-10) 以向量 x 为未知元, 其解称为方程组 (2-9) 的解向量.

如果把系数矩阵 A 按行分块为 $A = \begin{pmatrix} \alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_m^T \end{pmatrix}$, 则方程组 (2-10) 可记作:

$$\begin{pmatrix} \alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_m^T \end{pmatrix} x = b \text{ 或 } \begin{cases} \alpha_1^T x = b_1 \\ \alpha_2^T x = b_2 \\ \dots\dots\dots \\ \alpha_m^T x = b_m \end{cases}, \quad (2-11)$$

这相当于把方程组 (2-9) 的第 i 个方程

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i$$

记作:

$$\alpha_i^T x = b_i \quad (i = 1, 2, \dots, m).$$

习题 2.4

1. 利用分块矩阵计算
$$\begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 3 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & -3 \end{pmatrix}.$$

2. 设 $A = \begin{pmatrix} 3 & 4 & 0 & 0 \\ 4 & -3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 2 & 2 \end{pmatrix}$, 求 $|A^8|$ 和 A^4 .

3. 设 n 阶矩阵 A 和 m 阶矩阵 B 都可逆, C 是任意一个 m 行 n 列矩阵, 证明:

(1) 分块矩阵 $\begin{pmatrix} O & A \\ B & O \end{pmatrix}$ 可逆, 且 $\begin{pmatrix} O & A \\ B & O \end{pmatrix}^{-1} = \begin{pmatrix} O & B^{-1} \\ A^{-1} & O \end{pmatrix};$

(2) 分块矩阵 $\begin{pmatrix} A & O \\ C & B \end{pmatrix}$ 可逆, 且 $\begin{pmatrix} A & O \\ C & B \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & O \\ -B^{-1}CA^{-1} & B^{-1} \end{pmatrix};$

(3) 分块矩阵 $\begin{pmatrix} B & C \\ O & A \end{pmatrix}$ 可逆, 且 $\begin{pmatrix} B & C \\ O & A \end{pmatrix}^{-1} = \begin{pmatrix} B^{-1} & -B^{-1}CA^{-1} \\ O & A^{-1} \end{pmatrix}.$

4. 求下列矩阵的逆矩阵:

(1) $\begin{pmatrix} 5 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 8 & 3 \\ 0 & 0 & 5 & 2 \end{pmatrix};$

(2) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 4 \end{pmatrix}.$

2.5 矩阵的初等变换与初等矩阵

本节将引进矩阵的初等变换并建立矩阵的秩的概念, 进一步用初等变换讨论矩阵的秩的性质.

2.5.1 矩阵的初等变换

矩阵的初等变换是矩阵的一种十分重要的运算, 它在解线性方程组、求逆矩阵及矩阵理论的进一步探讨中都起着重要的作用. 先来看一个例子及其解法.

引例 求解线性方程组:

$$\begin{aligned}
 & \begin{cases} 2x_1 - x_2 - x_3 + x_4 = 2 & \textcircled{1} \\ x_1 + x_2 - 2x_3 + x_4 = 4 & \textcircled{2} \\ 4x_1 - 6x_2 + 2x_3 - 2x_4 = 4 & \textcircled{3} \\ 3x_1 + 6x_2 - 9x_3 + 7x_4 = 9 & \textcircled{4} \end{cases} \quad (2-13) \\
 \text{解} \quad (2-13) & \xrightarrow[\textcircled{3} \times \frac{1}{2}]{\textcircled{1} \leftrightarrow \textcircled{2}} \begin{cases} x_1 + x_2 - 2x_3 + x_4 = 4 & \textcircled{1} \\ 2x_1 - x_2 - x_3 + x_4 = 2 & \textcircled{2} \\ 2x_1 - 3x_2 + x_3 - x_4 = 2 & \textcircled{3} \\ 3x_1 + 6x_2 - 9x_3 + 7x_4 = 9 & \textcircled{4} \end{cases} \quad (B_1) \\
 (B_1) & \xrightarrow[\textcircled{4} - \textcircled{1} \times 3]{\begin{smallmatrix} \textcircled{2} - \textcircled{3} \\ \textcircled{3} - \textcircled{1} \times 2 \end{smallmatrix}} \begin{cases} x_1 + x_2 - 2x_3 + x_4 = 4 & \textcircled{1} \\ 2x_2 - 2x_3 + 2x_4 = 0 & \textcircled{2} \\ -5x_2 + 5x_3 - 3x_4 = -6 & \textcircled{3} \\ 3x_2 - 3x_3 + 4x_4 = -3 & \textcircled{4} \end{cases} \quad (B_2) \\
 (B_2) & \xrightarrow[\textcircled{4} - \textcircled{2} \times 3]{\begin{smallmatrix} \textcircled{2} \times \frac{1}{2} \\ \textcircled{3} + \textcircled{2} \times 5 \end{smallmatrix}} \begin{cases} x_1 + x_2 - 2x_3 + x_4 = 4 & \textcircled{1} \\ x_2 - x_3 + x_4 = 0 & \textcircled{2} \\ 2x_4 = -6 & \textcircled{3} \\ x_4 = -3 & \textcircled{4} \end{cases} \quad (B_3) \\
 (B_3) & \xrightarrow[\textcircled{4} - \textcircled{3} \times 2]{\textcircled{3} \leftrightarrow \textcircled{4}} \begin{cases} x_1 + x_2 - 2x_3 + x_4 = 4 & \textcircled{1} \\ x_2 - x_3 + x_4 = 0 & \textcircled{2} \\ x_4 = -3 & \textcircled{3} \\ 0 = 0 & \textcircled{4} \end{cases} \quad (B_4)
 \end{aligned}$$

这里 $(2-13) \rightarrow (B_1)$ 是为消 x_1 做准备; $(B_1) \rightarrow (B_2)$ 是保留①中的 x_1 , 消去②、③、④中的 x_1 ; $(B_2) \rightarrow (B_3)$ 保留②中的 x_2 并将其系数变为1, 然后消去③、④中的 x_2 , 与此同时恰好也把③、④中的 x_3 消去了; $(B_3) \rightarrow (B_4)$ 保留③中的 x_4 并将其系数变为1, 然后消去④中的 x_4 , 与此同时恰好也把常数项消去了, 得到恒等式 $0=0$ (如果常数项不能消去, 就将得到矛盾方程 $0=1$, 即说明原方程组无解). 至此消元完毕.

方程组 (B_4) 呈阶梯形, 它是含4个未知数、3个有效方程的方程组, 故应有一个未知数是自由的, 称为自由未知数, 另外三个未知数称为非自由未知数. 我们只要把阶梯形每步台阶的第一个未知数取为非自由未知数 (这里为 x_1 、 x_2 、 x_4), 剩下的 x_3 就是自由未知数.

利用方程组 (B_4) , 使用“回代”的方法可求出原方程组 $(2-13)$ 的解: 由③得 $x_4 = -3$; 将 $x_4 = -3$ 代入②得 $x_2 = x_3 + 3$; 将 $x_4 = -3$, $x_2 = x_3 + 3$ 代入①得 $x_1 = x_3 + 4$. 于是解得:

$$\begin{cases} x_1 = x_3 + 4 \\ x_2 = x_3 + 3, \\ x_4 = -3 \end{cases}$$

其中自由未知数 x_3 可以任意取值. 一般令 $x_3 = c$, 则方程组的解 (向量) 可记作:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} c+4 \\ c+3 \\ c \\ -3 \end{pmatrix}, \text{ 即 } \mathbf{x} = c \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 4 \\ 3 \\ 0 \\ -3 \end{pmatrix}, \quad (2-14)$$

其中 c 为任意常数.

在上述的消元过程中, 始终把方程组看做一个整体, 即不是着眼于某一个方程的变形, 而是着眼于整个方程组变形成另一个方程组. 其中用到三种变换, 分别为: (i) 交换方程的次序 (如 (2-13) $\rightarrow (B_1)$ 中的 ① $\leftrightarrow$ ②、 $(B_3) \rightarrow (B_4)$ 中的 ③ $\leftrightarrow$ ④等); (ii) 以不等于 0 的数乘某个方程 (如 (2-13) $\rightarrow (B_1)$ 中的 ③ $\times \frac{1}{2}$ 、 $(B_2) \rightarrow (B_3)$ 中的 ② $\times \frac{1}{2}$ 等); (iii) 一个方程组加上 (或减去) 另一个方程的 k 倍 (如 $(B_2) \rightarrow (B_3)$ 中的 ③ + ② $\times 5$ 和 ④ - ② $\times 3$ 等). 由于这三种变换都是可逆的, 即:

若 $(A) \xrightarrow{\text{①} \leftrightarrow \text{②}} (B)$, 则 $(B) \xrightarrow{\text{①} \leftrightarrow \text{②}} (A)$;

若 $(A) \xrightarrow{\text{①} \times k} (B)$, 则 $(B) \xrightarrow{\text{①} \times \frac{1}{k}} (A)$;

若 $(A) \xrightarrow{\text{①} + \text{②} \times k} (B)$, 则 $(B) \xrightarrow{\text{①} - \text{②} \times k} (A)$.

因此变换前的方程组与变换后的方程组是同解的, 这三种变换都是方程组的同解变换, 所以最后求得的解 (2-14) 是方程组 (2-13) 的全部解.

上述变换过程实际上只对方程组的各未知数的系数及常数进行运算, 未知数本身并未参与运算. 因此, 若记方程组 (2-13) 的增广矩阵为:

$$\mathbf{B} = (\mathbf{A}, \mathbf{b}) = \begin{pmatrix} 2 & -1 & -1 & 1 & 2 \\ 1 & 1 & -2 & 1 & 4 \\ 4 & -6 & 2 & -2 & 4 \\ 3 & 6 & -9 & 7 & 9 \end{pmatrix},$$

那么上述对方程组的变换完全可以转换为对矩阵 $\mathbf{B}$ 的变换. 把方程组的上述三种同解变换移置到矩阵上, 就得到矩阵的三种初等行变换.

定义 2-13 (初等行变换) 分别将下列三种变换称为矩阵的第一初等行变换、第二初等行变换和第三初等行变换:

(I) 对调两行 (对调 i, j 两行, 记作 $r_i \leftrightarrow r_j$);

(II) 以数 k ($k \neq 0$) 乘某一行中的所有元素 (第 i 行乘 k , 记作 $r_i \times k$);

(III) 把某一行所有元素的 k 倍加到另一行对应的元素上去 (第 j 行的 k 倍加

到第 i 行上, 记作 $r_i + r_j \times k$ 或简写作 $r_i + kr_j$).

显然, 三种初等行变换都是可逆的, 且其逆变换与原变换都属于同一类型:

变换 $r_i \leftrightarrow r_j$ 的逆变换是其自身; 变换 $r_i \times k$ 的逆变换是 $r_i \times \frac{1}{k}$ (或记作 $r_i \div k$); 变换 $i + kr_j$ 的逆变换是 $i + (-k)r_j$ (或记作 $i - kr_j$).

把上述定义中的“行”换成“列”, 即得到矩阵的三种初等列变换的定义(把相应记号中的“ r ”换成“ c ”即可). 与初等行变换一样, 三种初等列变换也都是可逆的, 且其逆变换与原变换属于同一类型.

矩阵的初等行变换和初等列变换统称为矩阵的**初等变换**.

如果矩阵 A 经有限次初等行变换变成矩阵 B , 就说矩阵 A 与矩阵 B 行等价, 记作 $A \sim_r B$; 如果矩阵 A 经有限次初等列变换变成矩阵 B , 就说矩阵 A 与矩阵 B 列等价, 记作 $A \sim_c B$; 如果矩阵 A 经有限次初等变换变成矩阵 B , 就说矩阵 A 与矩阵 B 等价, 记作 $A \sim B$.

矩阵之间的等价关系具有下列性质:

- ① 反身性 $A \sim A$;
- ② 对称性 若 $A \sim B$, 则 $B \sim A$;
- ③ 传递性 若 $A \sim B$, $B \sim C$, 则 $A \sim C$.

上面三条性质对行变换和列变换也均适用.

下面用矩阵的初等行变换来解方程组 (2-13), 请读者将其过程与该节引例的求解过程一一对照.

$$\begin{aligned}
 B &= \begin{pmatrix} 2 & -1 & -1 & 1 & 2 \\ 1 & 1 & -2 & 1 & 4 \\ 4 & -6 & 2 & -2 & 4 \\ 3 & 6 & -9 & 7 & 9 \end{pmatrix} \\
 \begin{matrix} r_1 \leftrightarrow r_2 \\ \sim \\ r_3 \times \frac{1}{2} \end{matrix} & \begin{pmatrix} 1 & 1 & -2 & 1 & 4 \\ 2 & -1 & -1 & 1 & 2 \\ 2 & -3 & 1 & -1 & 2 \\ 3 & 6 & -9 & 7 & 9 \end{pmatrix} = B_1 \\
 \begin{matrix} r_2 - r_1 \\ r_3 - 2r_1 \\ \sim \\ r_4 - 3r_1 \end{matrix} & \begin{pmatrix} 1 & 1 & -2 & 1 & 4 \\ 0 & 2 & -2 & 2 & 0 \\ 0 & -5 & 5 & -3 & -6 \\ 0 & 3 & -3 & 4 & -3 \end{pmatrix} = B_2 \\
 \begin{matrix} r_2 \div 2 \\ r_3 + 5r_2 \\ \sim \\ r_4 - 3r_2 \end{matrix} & \begin{pmatrix} 1 & 1 & -2 & 1 & 4 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 2 & -6 \\ 0 & 0 & 0 & 1 & -3 \end{pmatrix} = B_3
 \end{aligned}$$

$$\begin{matrix} r_3 \leftrightarrow r_4 \\ \sim \\ r_4 - 2r_3 \end{matrix} \begin{pmatrix} 1 & 1 & -2 & 1 & 4 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = B_4$$

由方程组 B_4 得到解 (2-14) 的回代过程, 也可用矩阵的初等行变换来完成, 即:

$$\begin{matrix} r_1 - r_2 \\ \sim \\ r_2 - r_3 \end{matrix} B_4 \begin{pmatrix} 1 & 0 & -1 & 0 & 4 \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = B_5,$$

B_5 对应方程组:

$$\begin{cases} x_1 - x_3 = 4 \\ x_2 - x_3 = 3, \\ x_4 = -3 \end{cases}$$

取 x_3 为自由未知数, 并令 $x_3 = c$, 即得:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} c+4 \\ c+3 \\ c \\ -3 \end{pmatrix} = c \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 4 \\ 3 \\ 0 \\ -3 \end{pmatrix},$$

其中 c 为任意常数. 上式与 (2-14) 完全相同.

矩阵 B_4 和 B_5 都称为行阶梯形矩阵, 特点是: 可画一条向右下方走的阶梯线, 线的左下方元素全为 0; 每步台阶只有一行高 (即阶梯线的竖线的高度为一行), 台阶数就是非零行的行数, 而阶梯线的横线将走到下一非零行的第一个非零元及其左方零元之间 (即保证阶梯线的竖线后面为该行的第一个非零元), 当最后一个非零行走完后, 阶梯线结束.

行阶梯形矩阵 B_5 进一步称为行最简形矩阵, 特点是: 所有非零行的第一个非零元为 1, 且这些非零元所在列的其他元素都为 0.

读者可以类似地给出列阶梯形矩阵和列最简形矩阵的定义.

可以用归纳法证明 (这里从略): 对于任何矩阵 $A_{m \times n}$, 总可以经过有限次初等行 (列) 变换把它变为行 (列) 阶梯形矩阵和行 (列) 最简形矩阵.

利用初等行变换, 把一个矩阵化为行阶梯形矩阵和行最简形矩阵是一种很重要的运算. 由引例的第二种解法可知, 要解线性方程组 (2-13), 只需先把其增广矩阵化为行最简形矩阵 B_5 , 然后通过简单观察就可直接写出线性方程组的解 (2-14); 反之, 由方程组的解 (2-14) 也可以写出矩阵 B_5 . 由此, 我们给出 (证明从略) 结论:

一个矩阵的行最简形矩阵是唯一确定的, 其行阶梯形矩阵中非零行的行数也是唯一确定的.

若对行最简形矩阵再施以初等列变换, 则可得到一种形状更简单的矩阵, 称为标准形. 例如:

$$B_5 = \begin{pmatrix} 1 & 0 & -1 & 0 & 4 \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\substack{c_3 \leftrightarrow c_4 \\ c_4 + c_1 + c_2 \\ \sim \\ c_5 - 4c_1 - 3c_2 + 3c_3}]{\quad} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = F.$$

矩阵 F 称为矩阵 B (引例中方程组的增广矩阵) 的标准形, 特点是: F 的左上角是一个单位矩阵, 其余元素全为 0.

因此, 对任意 $m \times n$ 矩阵 A , 总可以经过初等变换 (初等行变换和初等列变换) 把它化为标准形:

$$F = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}_{m \times n},$$

此标准形由 m 、 n 、 r 三个数完全确定, 其中 r 就是行阶梯形矩阵中非零行的行数. 所有与 A 等价的矩阵组成一个集合, 标准形 F 是这个集合中形状最简单的元素.

2.5.2 初等矩阵

为了进一步讨论初等变换的性质, 我们引入初等矩阵的概念.

定义 2-14 (初等矩阵) 由单位矩阵经过一次初等变换得到的矩阵称为初等矩阵.

三种初等变换对应三种初等矩阵:

(I) 把单位矩阵的第 i 、 j 两行对调 (或 i 、 j 两列对调) 得初等矩阵:

$$I(i, j) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & \cdots & 1 \\ & & & 1 & & \\ & & \vdots & & \ddots & \vdots \\ & & & 1 & \cdots & 0 \\ & & & & & 1 & \ddots \\ & & & & & & 1 \end{pmatrix} \begin{matrix} \leftarrow \text{第 } i \text{ 行} \\ \\ \\ \leftarrow \text{第 } j \text{ 行} \end{matrix},$$

用 m 阶初等矩阵 $I_m(i, j)$ 左乘矩阵 $A = (a_{ij})_{m \times n}$, 得:

$$I_m(i, j)A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{matrix} \leftarrow \text{第 } i \text{ 行} \\ \\ \leftarrow \text{第 } j \text{ 行} \end{matrix},$$

其结果相当于对矩阵 A 施以第一种初等行变换 $r_i \leftrightarrow r_j$; 类似地, 以 n 阶初等矩阵 $I_n(i, j)$ 右乘矩阵 A , 相当于对矩阵 A 施以第一种初等列变换 $c_i \leftrightarrow c_j$.

(II) 以数 k ($k \neq 0$) 乘单位矩阵的第 i 行 (或第 i 列), 得初等矩阵:

$$I(i(k)) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & k & \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} \leftarrow \text{第 } i \text{ 行},$$

可以验证: 以 m 阶初等矩阵 $I_m(i(k))$ 左乘矩阵 A , 其结果相当于以数 k 乘矩阵 A 的第 i 行 ($r_i \times k$); 以 n 阶初等矩阵 $I_n(i(k))$ 右乘矩阵 A , 其结果相当于以数 k 乘矩阵 A 的第 i 列 ($c_i \times k$).

(III) 以数 k 乘以矩阵的第 j 行加到第 i 行上或以 k 乘以矩阵的第 i 列加到第 j 列上, 得初等矩阵:

$$I(ij(k)) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & k \\ & & & \ddots & \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} \begin{matrix} \leftarrow \text{第 } i \text{ 行} \\ \\ \leftarrow \text{第 } j \text{ 行} \end{matrix},$$

可以验证: 以 m 阶初等矩阵 $I_m(ij(k))$ 左乘矩阵 A , 其结果相当于把 A 的第 j 行乘 k 加到第 i 行上 ($r_i + kr_j$); 以 n 阶初等矩阵 $I_n(ij(k))$ 右乘矩阵 A , 其结果相当于把 A 的第 i 列乘 k 加到第 j 列上 ($c_j + kc_i$).

显然以上三种初等矩阵都是可逆的, 其逆矩阵是同一类型的初等矩阵, 即:

$$I(i, j)^{-1} = I(i, j); \quad I(i(k))^{-1} = I(i(\frac{1}{k})); \quad I(ij(k))^{-1} = I(ij(-k)).$$

归纳上面的讨论, 可得:

性质 1 设 A 是一个 $m \times n$ 矩阵, 对 A 施以一次初等行变换, 相当于在 A 的左边乘以相应的 m 阶初等矩阵; 对 A 施以一次初等列变换, 相当于在 A 的右边乘以相应的 n 阶初等矩阵.

性质 2 方阵 A 可逆的充分必要条件是存在有限个初等矩阵 $P_1, P_2, \dots, P_l$ 使 $A = P_1 P_2 \cdots P_l$.

证 充分性. 设 $A = P_1 P_2 \cdots P_l$, 而 $P_1, P_2, \dots, P_l$ 可逆, 即 $|P_i| \neq 0$ ($i=1, 2, \dots, l$), 故 $|A| = |P_1| |P_2| \cdots |P_l| \neq 0$, 即 A 可逆.

必要性. 设 n 阶方阵 A 可逆, 且 A 的标准形矩阵为 F , 由 $F \sim A$ 知 F 经有限次初等变换可化为 A , 即有初等矩阵 $P_1, P_2, \dots, P_l$ 使:

$$A = P_1 \cdots P_s F P_{s+1} \cdots P_l,$$

因为 A 可逆, 所以 F 也可逆, 即:

$$F = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}_{n \times n}$$

中的 $r = n$, 即 $F = I$, 从而:

$$A = P_1 P_2 \cdots P_l. \blacksquare$$

在上面的性质中, 若令 $P^{-1} = P_1 P_2 \cdots P_l$, 则有 $PA = I$, 进而说明方阵 A 可逆的充分必要条件是存在可逆矩阵 P 使 $PA = I$.

利用上述初等矩阵的性质, 可以证明矩阵的初等变换的如下基本性质:

定理 2-4 设 A 与 B 是 $m \times n$ 阶矩阵, 那么:

- (1) $A \sim B$ 的充分必要条件是存在 m 阶可逆矩阵 P 使 $PA = B$;
- (2) $A \sim B$ 的充分必要条件是存在 n 阶可逆矩阵 Q 使 $AQ = B$;
- (3) $A \sim B$ 的充分必要条件是存在 m 阶可逆矩阵 P 和 n 阶可逆矩阵 Q 使 $PAQ = B$.

证 (1) $A \sim B \Leftrightarrow A$ 经有限次初等行变换变为 B ;

$\Leftrightarrow$ 存在有限个 m 阶初等矩阵 $P_1, P_2, \dots, P_s$ 使 $P_1 P_2 \cdots P_s A = B$;

$\Leftrightarrow$ 存在 m 阶可逆矩阵 P 使 $PA = B$.

类似可证明 (2) 和 (3). $\blacksquare$

定理 2-4 把矩阵的初等变换与矩阵的乘法联系了起来, 从而可以依据矩阵乘法的运算规律得到初等变换的运算规律, 也可以利用矩阵初等变换来研究矩阵的乘法. 下面先给出定理 2-4 的一个推论, 然后介绍一种利用初等变换求逆矩阵的方法.

推论 方阵 A 可逆的充分必要条件是 $A \sim I$.

证 A 可逆 $\Leftrightarrow$ 存在可逆矩阵 P 使 $PA = I \Leftrightarrow A \sim I$. $\blacksquare$

该推论也同时说明所求的可逆矩阵 P 就是 A 的逆矩阵, 下面给出求 P 的方法.

由定理 2-4 (1) 可知, 如果 $A \sim B$, 即 A 经一系列初等行变换变为 B , 则有可逆矩阵 P 使 $PA = B$, 结合等式 $PI = P$, 有 $P(A, I) = (B, P)$, 从而:

$$(A, I) \sim (B, P).$$

上式说明, 如果对分块矩阵 (A, I) 作初等行变换, 那么, 当 A 变为 B 时, I 就变为 P . 结合上述推论, 令 $B = I$, 则有 $P = A^{-1}$.

综上, 要求一个方阵 A 的逆矩阵, 只需对矩阵 (A, I) 作初等行变换, 那么, 当 A 变为 I 时, I 就变为 A^{-1} .

例 2-13 设 $A = \begin{pmatrix} -2 & -1 & 4 \\ 1 & -1 & -6 \\ 1 & 2 & 2 \end{pmatrix}$, 求其行最简形矩阵 E , 并求一个可逆矩阵 P

使 $PA = E$.

解 如前所述, 只需对 (A, I) 作初等行变换把 A 化为 E , 同时也就得到了由 I

变成的 $\mathbf{P}$. 运算如下:

$$(\mathbf{A}, \mathbf{I}) = \begin{pmatrix} -2 & -1 & 4 & 1 & 0 & 0 \\ 1 & -1 & -6 & 0 & 1 & 0 \\ 1 & 2 & 2 & 0 & 0 & 1 \end{pmatrix} \begin{matrix} r_1 \leftrightarrow r_3 \\ r_2 - r_1 \\ \sim \\ r_3 + 2r_1 \end{matrix} \begin{pmatrix} 1 & 2 & 2 & 0 & 0 & 1 \\ 0 & -3 & -8 & 0 & 1 & -1 \\ 0 & 3 & 8 & 1 & 0 & 2 \end{pmatrix}$$

$$\begin{matrix} r_3 + r_2 \\ r_1 + \frac{2}{3}r_2 \\ \sim \\ r_2 \times (-\frac{1}{3}) \end{matrix} \begin{pmatrix} 1 & 0 & -\frac{10}{3} & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 1 & \frac{8}{3} & 0 & -\frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} = (\mathbf{E}, \mathbf{P}).$$

故 $\mathbf{E} = \begin{pmatrix} 1 & 0 & -\frac{10}{3} \\ 0 & 1 & \frac{8}{3} \\ 0 & 0 & 0 \end{pmatrix}$ 为 $\mathbf{A}$ 的行最简形矩阵, 而使 $\mathbf{PA} = \mathbf{E}$ 的可逆矩阵为:

$$\mathbf{P} = \begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & -\frac{1}{3} & \frac{1}{3} \\ 1 & 1 & 1 \end{pmatrix}.$$

当然, 我们可以继续对上述的 $(\mathbf{E}, \mathbf{P})$ 作针对第三行乘以倍数的初等行变换如 $r_3 \times k$ 、 $r_1 + kr_3$ 、 $r_2 + kr_3$ 等, 并不会改变行最简形矩阵 $\mathbf{E}$, 但所求的 $\mathbf{P}$ 会发生变化. 由此说明本例中使 $\mathbf{PA} = \mathbf{E}$ 的 $\mathbf{P}$ 不是唯一的.

例 2-14 设 $\mathbf{A} = \begin{pmatrix} 0 & 3 & -2 \\ -2 & 0 & 3 \\ 1 & -2 & 0 \end{pmatrix}$, 证明 $\mathbf{A}$ 可逆, 并求 $\mathbf{A}^{-1}$.

解 只要将 $(\mathbf{A}, \mathbf{I})$ 作初等行变换能把 $\mathbf{A}$ 化为 $\mathbf{I}$, 则 $\mathbf{I}$ 就相应地变成了 $\mathbf{A}^{-1}$, 同时也证明了 $\mathbf{A}$ 可逆. 运算如下:

$$(\mathbf{A}, \mathbf{I}) = \begin{pmatrix} 0 & 3 & -2 & 1 & 0 & 0 \\ -2 & 0 & 3 & 0 & 1 & 0 \\ 1 & -2 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{matrix} r_1 \leftrightarrow r_3 \\ \sim \\ r_2 + 2r_1 \end{matrix} \begin{pmatrix} 1 & -2 & 0 & 0 & 0 & 1 \\ 0 & -4 & 3 & 0 & 1 & 2 \\ 0 & 3 & -2 & 1 & 0 & 0 \end{pmatrix}$$

$$\begin{matrix} r_2 + r_3 \\ r_3 + 3r_2 \\ \sim \\ r_2 \times (-1) \end{matrix} \begin{pmatrix} 1 & -2 & 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & -1 & -1 & -2 \\ 0 & 0 & 1 & 4 & 3 & 6 \end{pmatrix} \begin{matrix} r_2 + r_3 \\ \sim \\ r_1 + 2r_2 \end{matrix} \begin{pmatrix} 1 & 0 & 0 & 6 & 4 & 9 \\ 0 & 1 & 0 & 3 & 2 & 4 \\ 0 & 0 & 1 & 4 & 3 & 6 \end{pmatrix} = (\mathbf{I}, \mathbf{A}^{-1}),$$

故 $\mathbf{A}$ 可逆, 且 $\mathbf{A}^{-1} = \begin{pmatrix} 6 & 4 & 9 \\ 3 & 2 & 4 \\ 4 & 3 & 6 \end{pmatrix}.$

例 2-15 设 $A = \begin{pmatrix} 2 & 1 & -3 \\ 1 & 2 & -2 \\ -3 & 3 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 0 & 3 \end{pmatrix}$, 求解满足方程 $AX = B$ 的矩阵 X .

解 设可逆矩阵 P 使 $PA = \tilde{A}$ 为行最简形矩阵, 则有:

$$P(A, B) = (PA, PB) = (\tilde{A}, PB),$$

上式表示对矩阵 (A, B) 施以初等行变换把 A 化为 $\tilde{A}$ 时就把 B 变成了 PB . 若 $\tilde{A} = I$, 则 A 可逆, 且 $A^{-1} = P$, 根据矩阵的乘法, 这时所给方程有唯一解:

$$X = A^{-1}B = PB.$$

因此, 我们应将 (A, B) 施以初等行变换把 A 化为 I , 则 B 变成的结果即为所求, 运算如下:

$$\begin{aligned} (A, B) &= \left(\begin{array}{ccc|cc} 2 & 1 & -3 & 1 & -1 \\ 1 & 2 & -2 & 2 & 0 \\ -3 & 3 & 2 & 0 & 3 \end{array} \right) \xrightarrow[r_1 - 2r_2]{r_3 + r_1 + r_2} \left(\begin{array}{ccc|cc} 0 & -3 & 1 & -3 & -1 \\ 1 & 2 & -2 & 2 & 0 \\ 0 & 6 & -3 & 3 & 2 \end{array} \right) \\ &\xrightarrow[r_3 + 2r_2]{r_1 \leftrightarrow r_2} \left(\begin{array}{ccc|cc} 1 & 2 & -2 & 2 & 0 \\ 0 & -3 & 1 & -3 & -1 \\ 0 & 0 & -1 & -3 & 0 \end{array} \right) \xrightarrow[r_1 + \frac{2}{3}r_2]{r_2 + r_3, r_3 \times (-1)} \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 4 & -\frac{2}{3} \\ 0 & -3 & 0 & -6 & -1 \\ 0 & 0 & -1 & -3 & 0 \end{array} \right) \\ &\xrightarrow[r_3 \times (-1)]{r_2 \times (-\frac{1}{3})} \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 4 & -\frac{2}{3} \\ 0 & 1 & 0 & 2 & \frac{1}{3} \\ 0 & 0 & 1 & 3 & 0 \end{array} \right) = (I, PB), \end{aligned}$$

从而方程的唯一解为:

$$X = PB = \begin{pmatrix} 4 & -\frac{2}{3} \\ 2 & \frac{1}{3} \\ 3 & 0 \end{pmatrix}.$$

我们一般将形如 $AX = B$ 的以矩阵 X 为未知项的方程称为矩阵方程.

以上两例实际上给出了当 A 为 3 阶甚至更高阶的可逆方阵时, 求解矩阵方程 $AX = B$ 的方法 (求 A^{-1} 实际上是求矩阵方程 $AX = I$ 的解). 该方法就是把方程 $AX = B$ 的增广矩阵 (A, B) 通过初等行变换化为行最简形, 从而求得方程的解, 这与求解线性方程组 $Ax = b$ 时将增广矩阵 (A, b) 化为行最简形矩阵的方法是一样的.

2.5.3 矩阵的秩

在 2.5.1 节中我们知道, 给定一个 $m \times n$ 矩阵 A , 它的标准形

$$F = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}_{m \times n}$$

由数 r 完全确定, 该 r 就是行阶梯形矩阵中非零行的行数, 这个数便是所谓的矩阵 A 的秩. 但由于这个数的唯一性尚未证明, 因此下面用另一种方法给出矩阵秩的定义.

定义 2-15 (k 阶子式) 给定一个 $m \times n$ 矩阵 A , 任取其 k 行与 k 列 ($k \leq \min\{m, n\}$), 位于这些行列交叉处的 k^2 个元素, 不改变它们在 A 中的位置次序而得到的 k 阶行列式称为矩阵 A 的 k 阶子式.

由上述定义可知: 矩阵 A 的所有子式都是一个行列式, 其结果是一个数; $m \times n$ 矩阵 A 的 k 阶子式共有 $C_m^k \cdot C_n^k$ 个.

定义 2-16 (矩阵的秩) 对于矩阵 A , 若存在一个正整数 r , 使得 A 的所有 $r+1$ 阶子式 (如果存在的话) 均为 0, 但至少有一个 A 的 r 阶子式 D 不为 0, 那么这个 D 称为矩阵 A 的最高阶非零子式, 而这个数 r 称为矩阵 A 的秩, 记作 $R(A)$. 并规定零矩阵的秩为 0.

由上述定义及转置行列式的性质, 我们有: 矩阵 A 的秩是由 A 唯一确定的; 但矩阵 A 的最高阶非零子式不一定唯一; 对于 $m \times n$ 矩阵 A , 有 $0 \leq R(A) \leq \min\{m, n\}$; $R(A^T) = R(A)$.

由行列式按行 (或列) 展开的性质可知, 如果矩阵 A 的所有 $r+1$ 阶子式均为 0, 则矩阵 A 的所有比 $r+1$ 阶更高阶的子式 (如果存在的话) 也均为 0, 因此在上述定义中把 r 阶非零子式称为最高阶非零子式, 而 A 的秩 $R(A)$ 就是 A 的非零子式的最高阶数. 于是, 我们有结论: 若矩阵 A 的某个 p 阶子式不为 0, 则 $R(A) \geq p$; 若矩阵 A 的所有 q 阶子式均为 0, 则 $R(A) < q$.

对于 n 阶方阵 A , 其 n 阶子式只有一个 $|A|$, 故当 $|A| \neq 0$ 时 $R(A) = n$, 当 $|A| = 0$ 时 $R(A) < n$. 可见可逆矩阵的秩等于矩阵的阶数, 不可逆矩阵的秩小于矩阵的阶数. 鉴于此, 我们把可逆矩阵 (非奇异矩阵) 又叫做满秩矩阵, 把不可逆矩阵 (奇异矩阵) 又叫做降秩矩阵.

利用矩阵的秩, 我们将在 3.2 节中证明矩阵方程 $AX = B$ 的一个重要的结论:

定理 2-5 矩阵方程 $AX = B$ 有解的充分必要条件是 $R(A) = R(A, B)$.

例 2-16 求矩阵

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 3 & 7 \\ 3 & -5 & 1 \end{pmatrix} \text{ 和 } B = \begin{pmatrix} 1 & -1 & 0 & 1 & -3 \\ 0 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

的秩.

解 对于矩阵 A , 经计算, 其唯一的一个 3 阶子式 $|A| = 0$, 但其 2 阶子式 $\begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} = -1 \neq 0$, 故 $R(A) = 2$; 对于矩阵 B , 它是一个行阶梯形矩阵, 其非零行只有 3 行, 故其所有 4 阶子式均为 0, 而以三个非零行的第一个非零元为主对角元的

3 阶子式 $\begin{vmatrix} 1 & -1 & 1 \\ 0 & 4 & 2 \\ 0 & 0 & 3 \end{vmatrix} = 12 \neq 0$, 因此 $R(B) = 3$.

从本例可以看出, 对于一般的行数与列数较高的矩阵, 按定义 2-16 求秩是很麻烦的. 然而对于行阶梯形矩阵, 它的秩就是非零行的行数, 即可以直接求出. 这就启示我们: 若能将一个一般的矩阵通过初等变换化为行阶梯形矩阵, 又能进一步证明两个等价矩阵的秩相等, 那么就可以通过与矩阵 A 等价的行阶梯形矩阵来求矩阵 A 的秩了. 事实上, 有如下定理:

定理 2-6 若 $A \sim B$, 则 $R(A) = R(B)$.

证 先证明: 若 A 经过一次初等行变换变为 B , 则 $R(A) = R(B)$.

设 $R(A) = r$ 且 A 的某个 r 阶子式 D 不为 0.

当 $A \xrightarrow{r_i \leftrightarrow r_j} B$ 或 $A \xrightarrow{r_i \times k} B$ 时, 在 B 中总能找到与 D 相对应的 r 阶子式 D_1 使得 $D_1 = D$ 或 $D_1 = -D$ 或 $D_1 = kD$, 因此 $D_1 \neq 0$, 从而有 $R(B) \geq r$.

当 $A \xrightarrow{r_i + kr_j} B$ 时, 因为对于作变换 $r_i \leftrightarrow r_j$ 时结论成立, 所以不妨只考虑 $A \xrightarrow{r_i + kr_j} B$ 的情形. 又分两种情况: ① D 中不包括 A 的第 1 行, 这时 D 也是 B 的 r 阶非零子式, 故 $R(B) \geq r$; ② D 中包括 A 的第 1 行, 这时把 B 中与 D 对应的 r 阶子式 D_1 选为 (这里 $p \geq 2$):

$$D_1 = \begin{vmatrix} r_1 + kr_2 \\ r_p \\ \vdots \\ r_q \end{vmatrix} = \begin{vmatrix} r_1 \\ r_p \\ \vdots \\ r_q \end{vmatrix} + k \begin{vmatrix} r_2 \\ r_p \\ \vdots \\ r_q \end{vmatrix} = D + kD_2.$$

若 $p = 2$, 则 $D_1 = D \neq 0$; 若 $p > 2$, 则 D_2 也是 B 的 r 阶子式, 由 $D_1 - kD_2 = D \neq 0$ 知 D_1 和 D_2 不同时为 0. 总之, B 至少存在一个不为 0 的 r 阶子式, 故 $R(B) \geq r$.

综上, 当 A 经过一次初等行变换变为 B 时, $R(A) \leq R(B)$. 由于初等行变换的可逆性, 故这时也可使 B 经过一次初等行变换变为 A , 因而有 $R(B) \leq R(A)$. 因此 $R(A) = R(B)$.

由数学归纳法可知, 经过有限次初等行变换不会改变矩阵的秩.

进一步, 若 A 经过有限次初等列变换变为 B , 则相应地, A^T 必经过有限次初等行变换变为 B^T , 从而有 $R(A^T) = R(B^T)$, 而 $R(A) = R(A^T)$, $R(B) = R(B^T)$, 因此 $R(A) = R(B)$.

总之, 若 A 经过有限次初等变换变为 B (即 $A \sim B$), 则 $R(A) = R(B)$. ■

结合定理 2-4 (3), 可得:

推论 若存在可逆矩阵 P 、 Q 使 $PAQ = B$, 则 $R(A) = R(B)$.

例 2-17 设矩阵

$$A = \begin{pmatrix} 3 & 2 & 0 & 5 & 0 \\ 3 & -2 & 3 & 6 & -1 \\ 2 & 0 & 1 & 5 & -3 \\ 1 & 6 & -4 & -1 & 4 \end{pmatrix},$$

求 A 的秩, 并求 A 的一个最高阶非零子式.

解 先求 A 的秩, 为此对 A 作初等行变换把它变成行阶梯形矩阵:

$$A = \begin{pmatrix} 3 & 2 & 0 & 5 & 0 \\ 3 & -2 & 3 & 6 & -1 \\ 2 & 0 & 1 & 5 & -3 \\ 1 & 6 & -4 & -1 & 4 \end{pmatrix} \xrightarrow[r_4 - r_3]{\substack{r_1 \leftrightarrow r_4 \\ r_2 - r_4}} \begin{pmatrix} 1 & 6 & -4 & -1 & 4 \\ 0 & -4 & 3 & 1 & -1 \\ 2 & 0 & 1 & 5 & -3 \\ 1 & 2 & -1 & 0 & 3 \end{pmatrix} \\ \xrightarrow[r_3 \div 2]{\substack{r_3 - r_1 - r_4 \\ r_4 - r_1}} \begin{pmatrix} 1 & 6 & -4 & -1 & 4 \\ 0 & -4 & 3 & 1 & -1 \\ 0 & -4 & 3 & 3 & -5 \\ 0 & -4 & 3 & 1 & -1 \end{pmatrix} \xrightarrow[r_4 - r_2]{\substack{r_3 - r_2 \\ r_4 - r_2}} \begin{pmatrix} 1 & 6 & -4 & -1 & 4 \\ 0 & -4 & 3 & 1 & -1 \\ 0 & 0 & 0 & 2 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

因行阶梯形矩阵有 3 个非零行, 故 $R(A) = 3$.

由 $R(A) = 3$ 知 A 的最高阶非零子式一定为 3 阶行列式, 而所有的 3 阶行列式共有 $C_4^3 \cdot C_5^3 = 40$ 个, 要从这 40 个行列式中找出一个非零的行列式是很麻烦的. 但通过观察 A 的行阶梯形矩阵, 我们只提取 A 中第 1、2、4 列 (即 A 的行阶梯形矩阵中非零行的第一个非零元所在的列) 构成矩阵 A_0 , 即:

$$A_0 = \begin{pmatrix} 3 & 2 & 5 \\ 3 & -2 & 6 \\ 2 & 0 & 5 \\ 1 & 6 & -1 \end{pmatrix},$$

则经过前述同样的初等行变换可将 A_0 化为行阶梯形矩阵 (A 的行阶梯形矩阵的第 1、2、4 列):

$$\begin{pmatrix} 1 & 6 & -1 \\ 0 & -4 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix},$$

从而 $R(A_0) = 3$, 且 A_0 的 3 阶子式一定是 A 的 3 阶子式, 其总个数为 $C_4^3 \cdot C_3^3 = 4$,

要在这4个子式中找一个非零子式要容易得多. 现计算由 A_0 的第1、3、4行构成的三阶子式:

$$\begin{vmatrix} 3 & 2 & 5 \\ 2 & 0 & 5 \\ 1 & 6 & -1 \end{vmatrix} = \begin{vmatrix} 3 & 2 & 5 \\ 2 & 0 & 5 \\ -8 & 0 & -16 \end{vmatrix} = 16 \begin{vmatrix} 2 & 5 \\ 1 & 2 \end{vmatrix} = -16 \neq 0,$$

因此这个子式便是 A 的一个最高阶非零子式.

例 2-18 设 $A = \begin{pmatrix} 1 & 2 & -2 & 3 \\ -2 & -4 & 4 & -6 \\ 2 & 8 & -2 & 0 \\ -1 & 0 & 3 & -6 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$, 求矩阵 A 和矩阵 $B = (A, b)$ 的秩.

解 对 B 作初等行变换把它变为行阶梯形矩阵 $\tilde{B} = (\tilde{A}, \tilde{b})$, 则 $\tilde{A}$ 就是 A 的行阶梯形矩阵, 进而可从 $\tilde{B} = (\tilde{A}, \tilde{b})$ 中同时看出 $R(A)$ 和 $R(B)$.

$$B = \begin{pmatrix} 1 & 2 & -2 & 3 & 1 \\ -2 & -4 & 4 & -6 & 2 \\ 2 & 8 & -2 & 0 & 3 \\ -1 & 0 & 3 & -6 & 4 \end{pmatrix} \xrightarrow{\substack{r_2+2r_1 \\ r_3-2r_1 \\ \sim \\ r_4+r_1 \\ r_3-2r_4}} \begin{pmatrix} 1 & 2 & -2 & 3 & 1 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & -9 \\ 0 & 2 & 1 & -3 & 5 \end{pmatrix}$$

$$\xrightarrow{\substack{r_2 \leftrightarrow r_4 \\ \sim \\ r_4 + \frac{4}{9}r_3}} \begin{pmatrix} 1 & 2 & -2 & 3 & 1 \\ 0 & 2 & 1 & -3 & 5 \\ 0 & 0 & 0 & 0 & -9 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

因此 $R(A) = 2$, $R(B) = 3$.

从矩阵 B 的行阶梯形矩阵可知, 本例中 A 与 b 所对应的线性方程组 $Ax = b$ 是无解的, 这是因为行阶梯形矩阵的第3行表示矛盾方程 $0 = -9$.

例 2-19 设 $A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 2 & 6 \\ -1 & \lambda & 3 \\ 1 & -1 & \mu \end{pmatrix}$ 且 $R(A) = 2$, 求 λ 与 μ 的值.

解 $A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 2 & 6 \\ -1 & \lambda & 3 \\ 1 & -1 & \mu \end{pmatrix} \xrightarrow{\substack{r_2-2r_1 \\ r_3+r_1 \\ \sim \\ r_4-r_1}} \begin{pmatrix} 1 & 3 & 5 \\ 0 & -4 & -4 \\ 0 & \lambda+3 & 8 \\ 0 & -4 & \mu-5 \end{pmatrix} \xrightarrow{\substack{r_4-r_2 \\ r_2 \div (-4) \\ \sim \\ r_3-(\lambda+3)r_2}} \begin{pmatrix} 1 & 3 & 5 \\ 0 & 1 & 1 \\ 0 & 0 & 5-\lambda \\ 0 & 0 & \mu-1 \end{pmatrix},$

因为 $R(A) = 2$, 所以:

$$\begin{cases} 5-\lambda=0 \\ \mu-1=0 \end{cases}, \text{ 即 } \begin{cases} \lambda=5 \\ \mu=1 \end{cases}.$$

下面先归纳前面已经提到的矩阵的秩的一些性质, 再介绍几个常用的矩阵的

秩的性质.

$$\textcircled{1} \quad 0 \leq R(A_{m \times n}) \leq \min\{m, n\};$$

$$\textcircled{2} \quad R(A^T) = R(A);$$

$$\textcircled{3} \quad \text{若 } A \sim B, \text{ 则 } R(A) = R(B);$$

$$\textcircled{4} \quad \text{若 } P, Q \text{ 可逆, 则 } R(PAQ) = R(A);$$

$$\textcircled{5} \quad \max\{R(A), R(B)\} \leq R(A, B) \leq R(A) + R(B). \text{ 特别地:}$$

$$R(A) \leq R(A, b) \leq R(A) + 1,$$

这里 b 为非零列向量.

证 因为 A 或 B 的最高阶非零子式也是 (A, B) 的非零子式, 所以 $R(A) \leq R(A, B)$, $R(B) \leq R(A, B)$, 故 $\max\{R(A), R(B)\} \leq R(A, B)$.

设 $R(A) = r$, $R(B) = s$, 把 A 和 B 分别作初等列变换化为列阶梯形 $\tilde{A}$ 和 $\tilde{B}$, 则 $\tilde{A}$ 和 $\tilde{B}$ 中分别只含有前 r 个和前 s 个非零列, 故可设:

$$A \xrightarrow{c} \tilde{A} = (\tilde{a}_1, \dots, \tilde{a}_r, 0, \dots, 0), \quad B \xrightarrow{c} \tilde{B} = (\tilde{b}_1, \dots, \tilde{b}_s, 0, \dots, 0),$$

则 $(A, B) \xrightarrow{c} (\tilde{A}, \tilde{B})$, 且 $(\tilde{A}, \tilde{B})$ 中只含有 $r+s$ 个非零列, 因此 $R(\tilde{A}, \tilde{B}) \leq r+s$, 而 $R(A, B) = R(\tilde{A}, \tilde{B})$, 故 $R(A, B) \leq r+s = R(A) + R(B)$.

$$\textcircled{6} \quad R(A+B) \leq R(A) + R(B).$$

证 不妨设 A 和 B 都为 $m \times n$ 矩阵. 对矩阵 $(A+B, B)$ 作初等列变换 $c_i - c_{n+i}$ ($i=1, 2, \dots, n$), 则有:

$$(A+B, B) \xrightarrow{c} (A, B),$$

根据性质⑤:

$$R(A+B) \leq R(A+B, B) = R(A, B) \leq R(A) + R(B).$$

利用定理 2-5 还容易得出矩阵的秩的下列性质:

$$\textcircled{7} \quad R(AB) \leq \min\{R(A), R(B)\}.$$

事实上, 令 $AB = C$, 则矩阵方程 $AX = C$ 有解 $X = B$, 于是根据定理 2-5 有 $R(A) = R(A, C)$, 而由性质⑤有 $R(C) \leq R(A, C)$, 因此 $R(C) \leq R(A)$.

又由 $AB = C$ 知 $B^T A^T = C^T$, 故类似地可得 $R(C^T) \leq R(B^T)$, 再由性质②有 $R(C) \leq R(B)$.

综上, $R(C) \leq \min\{R(A), R(B)\}$, 即 $R(AB) \leq \min\{R(A), R(B)\}$.

本书第 3 章还将进一步讨论定理 2-5 和性质⑦的应用, 并将在 3.5 节中进一步证明矩阵的秩还具有如下性质:

$$\textcircled{8} \quad \text{若 } A_{m \times p} B_{p \times n} = O, \text{ 则 } R(A) + R(B) \leq p.$$

例 2-20 设 A 为 n 阶方阵, 证明 $R(A+I) + R(A-I) \geq n$.

证 因 $(A+I) + (I-A) = 2I$, 由性质⑥:

$$R(A+I)+R(I-A) \geq R(2I) = n,$$

而 $R(A-I) = R(I-A)$, 因此:

$$R(A+I)+R(A-I) \geq n.$$

例 2-21 证明: 若 $A_{m \times p} B_{p \times n} = C$ 且 $R(A) = p$, 则 $R(C) = R(B)$.

证 由 $R(A) = p$ 知 A 的行最简形矩阵为 $\begin{pmatrix} I_p \\ O \end{pmatrix}_{m \times p}$, 并有 m 阶可逆阵 P 使

$$PA = \begin{pmatrix} I_p \\ O \end{pmatrix}_{m \times p}, \text{ 于是:}$$

$$PC = PAB = \begin{pmatrix} I_p \\ O \end{pmatrix}_{m \times p} \cdot B = \begin{pmatrix} B \\ O \end{pmatrix}_{m \times n}.$$

由性质④知 $R(C) = R(PC)$, 而 $R\begin{pmatrix} B \\ O \end{pmatrix} = R(B)$, 故:

$$R(C) = R(B).$$

例 2-21 中矩阵 A 的秩等于它的列数, 我们把这样的矩阵称为列满秩矩阵. 当 A 为方阵时, 列满秩矩阵就成为满秩矩阵, 即可逆矩阵, 这时例 2-21 就是性质④的特例.

结合性质②, 我们有: 一个矩阵左乘一个列满秩矩阵后并不改变它的秩; 一个矩阵右乘一个行满秩矩阵(即矩阵的秩等于它的行数)后也不会改变它的秩.

例 2-21 有一种重要的特殊情形, 即 $C = O$, 这时结论为: 设 $AB = O$ 且 A 为列满秩矩阵, 则 $B = O$; 同样, 设 $AB = O$ 且 B 为行满秩矩阵, 则 $A = O$.

我们把上述结论称为矩阵乘法的消去律.

习题 2.5

1. (1) 设 $A = \begin{pmatrix} 4 & 1 & -2 \\ 2 & 2 & 1 \\ 3 & 1 & -1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -3 \\ 2 & 2 \\ 3 & -1 \end{pmatrix}$, 求 X 使 $AX = B$;
- (2) 设 $A = \begin{pmatrix} 0 & 2 & 1 \\ 2 & -1 & 3 \\ -3 & 3 & -4 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -3 & 1 \end{pmatrix}$, 求 X 使 $XA = B$.
2. 解下列矩阵方程:
 - (1) $\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} X = \begin{pmatrix} 4 & -6 \\ 2 & 1 \end{pmatrix}$;

$$(2) \quad X \begin{pmatrix} 2 & 1 & -1 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 3 \\ 4 & 3 & 2 \end{pmatrix};$$

$$(3) \quad \begin{pmatrix} 1 & 4 \\ -1 & 2 \end{pmatrix} X \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 0 & -1 \end{pmatrix};$$

$$(4) \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} X \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -4 & 3 \\ 2 & 0 & -1 \\ 1 & -2 & 0 \end{pmatrix}.$$

3. 用初等行变换把下列矩阵化为行最简形矩阵:

$$(1) \quad \begin{pmatrix} 1 & 0 & 2 & -1 \\ 2 & 0 & 3 & 1 \\ 3 & 0 & 4 & 3 \end{pmatrix};$$

$$(2) \quad \begin{pmatrix} 0 & 2 & -3 & 1 \\ 0 & 3 & -4 & 3 \\ 0 & 4 & -7 & -1 \end{pmatrix};$$

$$(3) \quad \begin{pmatrix} 1 & -1 & 3 & -4 & 3 \\ 3 & -3 & 5 & -4 & 1 \\ 2 & -2 & 3 & -2 & 0 \\ 3 & -3 & 4 & -2 & -1 \end{pmatrix};$$

$$(4) \quad \begin{pmatrix} 2 & 3 & 1 & -3 & -7 \\ 1 & 2 & 0 & -2 & -4 \\ 3 & -2 & 8 & 3 & 0 \\ 2 & -3 & 7 & 4 & 3 \end{pmatrix}.$$

4. 设 $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 \end{pmatrix}$, 求一个可逆矩阵 P 使 PA 为行最简形.

5. 设 $A = \begin{pmatrix} -5 & 3 & 1 \\ 2 & -1 & 1 \end{pmatrix}$:

(1) 求一个可逆矩阵 P 使 PA 为行最简形;

(2) 求一个可逆矩阵 Q 使 QA^T 为行最简形.

6. 试利用矩阵的初等变换求下列方阵的逆矩阵:

$$(1) \quad \begin{pmatrix} 3 & 2 & 1 \\ 3 & 1 & 5 \\ 3 & 2 & 3 \end{pmatrix};$$

$$(2) \quad \begin{pmatrix} 3 & -2 & 0 & -1 \\ 0 & 2 & 2 & 1 \\ 1 & -2 & -3 & -2 \\ 0 & 1 & 2 & 1 \end{pmatrix}.$$

7. 设 $A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix}$, $AX = 2X + A$, 求 X .

8. 在秩是 r 的矩阵中, 有没有等于 0 的 $r-1$ 阶子式? 有没有等于 0 的 r 阶子式?

9. 从矩阵 A 中划去一行得到矩阵 B , 问 A 与 B 的秩有什么关系?

10. 求一个秩是 4 的矩阵, 它的两个行向量是 $(1, 0, 1, 0, 0)$ 和 $(1, -1, 0, 0, 0)$.

11. 求下列矩阵的秩, 并求一个最高阶非零子式:

$$(1) \begin{pmatrix} 3 & 1 & 0 & 2 \\ 1 & -1 & 2 & -1 \\ 1 & 3 & -4 & 4 \end{pmatrix}; \quad (2) \begin{pmatrix} 3 & 2 & -1 & -3 & -1 \\ 2 & -1 & 3 & 1 & -3 \\ 7 & 0 & 5 & -1 & -8 \end{pmatrix};$$

$$(3) \begin{pmatrix} 2 & 1 & 8 & 3 & 7 \\ 2 & -3 & 0 & 7 & -5 \\ 3 & -2 & 5 & 8 & 0 \\ 1 & 0 & 3 & 2 & 0 \end{pmatrix}.$$

12. 设 $A = \begin{pmatrix} 1 & -2 & 3k \\ -1 & 2k & -3 \\ k & -2 & 3 \end{pmatrix}$, 问 k 为何值可使:

(1) $R(A) = 1$; (2) $R(A) = 2$; (3) $R(A) = 3$.

13. 设 A 、 B 都是 $m \times n$ 矩阵, 证明 $A \sim B$ 的充分必要条件是 $R(A) = R(B)$.

14. 证明 $R(A) = 1$ 的充分必要条件是存在非零列向量 a 及非零行向量 b^T 使 $A = ab^T$.

15. 设 A 为列满秩矩阵, $AB = C$, 证明线性方程 $Bx = 0$ 与 $Cx = 0$ 同解.

16. 设 A 为 $m \times n$ 矩阵, 证明方程 $AX = I_m$ 的充分必要条件是 $R(A) = m$.

2.6 矩阵的应用 (一)

2.6.1 矩阵在生物种群建模中的应用

例 2-22 (昆虫种群模型) 一种种群具有三个生命阶段: 卵、幼虫与成虫. 一群此种群的昆虫在某一特定的环境下 2 天时间间隔观察到: 20% 的卵没有成活, 20% 的卵成活但没有发育成幼虫, 60% 的卵发育成幼虫; 在同一时间间隔内, 10% 的幼虫没有成活, 30% 的幼虫成活但没有发育成成虫, 60% 的幼虫发育成成虫, 成虫 80% 存活; 在同一时间间隔内, 平均每只成虫产下 0.25 只虫卵. 假设初始状态有 10 千只卵、8 千只幼虫与 6 千只成虫, 计算 2 天、20 天、200 天后该昆虫种群的结构数量.

解 取初始时间为第 0 天, 设第 $2k$ 天后该昆虫种群的结构数量为 $x^{(k)} = (a_k, b_k, c_k)^T$, 这里 a_k, b_k, c_k 分别表示 $2k$ 天后该昆虫种群中活着的虫卵、幼虫、成虫的数量 (单位: 千只). 而题设中 $x^{(0)} = (a_0, b_0, c_0)^T = (10, 8, 6)^T$. 考查繁衍矩阵:

$$A = \begin{pmatrix} 0.2 & 0 & 0.25 \\ 0.6 & 0.3 & 0 \\ 0 & 0.6 & 0.8 \end{pmatrix},$$

其第一行的元素从左向右分别表示在 2 天时间间隔内虫卵到虫卵、幼虫到虫卵、成虫产虫卵的百分比;第二行元素从左向右分别表示在 2 天时间间隔内虫卵发育成幼虫、幼虫到幼虫、成虫产幼虫的百分比;第三行元素从左向右分别表示在 2 天时间间隔内虫卵发育成成虫、幼虫发育成成虫、成虫成活的百分比,则有:

$$A\mathbf{x}^{(0)} = \begin{pmatrix} 0.2 & 0 & 0.25 \\ 0.6 & 0.3 & 0 \\ 0 & 0.6 & 0.8 \end{pmatrix} \begin{pmatrix} 10 \\ 8 \\ 6 \end{pmatrix} = \begin{pmatrix} 0.2 \times 10 + 0 \times 8 + 0.25 \times 6 \\ 0.6 \times 10 + 0.3 \times 8 + 0 \times 6 \\ 0 \times 10 + 0.6 \times 8 + 0.8 \times 6 \end{pmatrix} = \begin{pmatrix} 3.5 \\ 8.4 \\ 9.6 \end{pmatrix},$$

最后的向量表示在一个 2 天时间间隔内: 活着的虫卵数为 3.5 千只 (10 千只虫卵 20% 的成活率加上 6 千只成虫 25% 的产卵率); 活着的幼虫数为 8.4 千只 (10 千只虫卵 60% 的发育率加上 8 千只幼虫 30% 的保有率); 活着的成虫数为 9.6 千只 (8 千只幼虫 60% 的发育率加上 6 千只成虫 80% 的成活率), 即 $\mathbf{x}^{(1)} = (3.5, 8.4, 9.6)^T$. 依此类推, 可知:

$$\mathbf{x}^{(k)} = A\mathbf{x}^{(k-1)} = \cdots = A^k \mathbf{x}^{(0)},$$

进而第 20 天和第 200 天后该昆虫种群的结构数量分别为:

$$\mathbf{x}^{(10)} = \begin{pmatrix} 0.2 & 0 & 0.25 \\ 0.6 & 0.3 & 0 \\ 0 & 0.6 & 0.8 \end{pmatrix}^{10} \begin{pmatrix} 10 \\ 8 \\ 6 \end{pmatrix} = \begin{pmatrix} 3.331 \\ 2.969 \\ 10.299 \end{pmatrix},$$

$$\mathbf{x}^{(100)} = \begin{pmatrix} 0.2 & 0 & 0.25 \\ 0.6 & 0.3 & 0 \\ 0 & 0.6 & 0.8 \end{pmatrix}^{100} \begin{pmatrix} 10 \\ 8 \\ 6 \end{pmatrix} = \begin{pmatrix} 0.284 \\ 0.253 \\ 0.877 \end{pmatrix}.$$

因此, 该昆虫种群随着时间的推移呈现数量减少的趋势.

例 2-23 (人口模型) 如果以年为时间单位记录人口数量, 年龄取周岁. 设某地区最大年龄为 m 岁, 在第 t 年为 i 岁的人数为 $x_i(t)$ (这里 $i = 0, 1, 2, \dots, m$; $t = 0, 1, 2, \dots$), 试研究函数 $x_i(t)$ 的变化规律.

解 作如下假设:

(1) 设 $d_i(t)$ 为第 t 年 i 岁人口的平均死亡率, 它是这一年中 i 岁人口的死亡数与基数之比, 即:

$$d_i(t) = \frac{x_i(t) - x_{i+1}(t+1)}{x_i(t)},$$

或写成:

$$x_{i+1}(t+1) = [1 - d_i(t)]x_i(t).$$

(2) 设 $b_i(t)$ 为第 t 年 i 岁女性的生育率, 它是每位育龄女性 (生育区间设为 $[i_1, i_2]$) 平均生育的婴儿数, $k_i(t)$ 为第 t 年 i 岁人口的女性比. 由此可知第 t 年出生的人数为:

$$f(t) = \sum_{i=i_1}^{i_2} b_i(t) k_i(t) x_i(t).$$

(3) 设 $d_{00}(t)$ 为第 t 年婴儿的死亡率, 则 $x_0(t) = [1 - d_{00}(t)]f(t)$.

(4) 令 $h_i(t) = \frac{b_i(t)}{\sum_{i=i_1}^{i_2} b_i(t)} = \frac{b_i(t)}{\beta(t)}$, 它表示 i 岁女性的总生育率, 则 $b_i(t) = \beta(t)h_i(t)$,

如果假设 t 年后女性生育率保持不变, 则:

$$\beta(t) = b_{i_1}(t) + b_{i_1+1}(t) + \cdots + b_{i_2}(t) = b_{i_1}(t) + b_{i_1+1}(t+1) + \cdots + b_{i_2}(t+i_2-i_1).$$

可见, $\beta(t)$ 为每位妇女一生中平均生育的婴儿数, 称为总和婴儿数, 它反映了人口变化的基本因素.

令 $b_i^*(t) = [1 - d_0(t)][1 - d_{00}(t)]h_i(t)k_i(t)$ (这里 $i = i_1, \dots, i_2$), 则根据上面的假设可以建立如下模型:

$$\begin{aligned} x_1(t+1) &= [1 - d_0(t)][1 - d_{00}(t)]f(t) \\ &= [1 - d_0(t)][1 - d_{00}(t)] \sum_{i=i_1}^{i_2} b_i(t)k_i(t)x_i(t) \\ &= [1 - d_0(t)][1 - d_{00}(t)]\beta(t) \sum_{i=i_1}^{i_2} h_i(t)k_i(t)x_i(t) \\ &= \beta(t) \sum_{i=i_1}^{i_2} b_i^*(t)x_i(t), \end{aligned}$$

$$x_2(t+1) = [1 - d_1(t)]x_1(t),$$

.....

$$x_m(t+1) = [1 - d_{m-1}(t)]x_{m-1}(t).$$

为了全面系统地反映一个时期内人口数量的状况, 令:

$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_m(t) \end{pmatrix},$$

$$\mathbf{A}(t) = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1-d_1(t) & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1-d_2(t) & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1-d_{m-1}(t) & 0 \end{pmatrix}_{m \times m},$$

$$\mathbf{B}(t) = \begin{pmatrix} 0 & \cdots & 0 & b_{i_1}^*(t) & \cdots & b_{i_2}^*(t) & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \cdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}_{m \times m},$$

其中 $b_{i_1}^*(t), \dots, b_{i_2}^*(t)$ 分别出现在 $i_1, \dots, i_2$ 列, 则有方程:

$$\mathbf{x}(t+1) = \mathbf{A}(t)\mathbf{x}(t) + \beta(t)\mathbf{B}(t)\mathbf{x}(t),$$

即:

$$\mathbf{x}(t+1) = [\mathbf{A}(t) + \beta(t)\mathbf{B}(t)]\mathbf{x}(t).$$

这里 $\beta(t)$ 表示的每位妇女一生中平均生育的婴儿数是可以控制的, 而 $\mathbf{x}(t)$ 也称为该方程的状态变量.

在稳定的社会环境下, 死亡率、生育模式、女性比例、婴儿存活率是可以假设不变的, 故可假定 $\mathbf{A}(t)$ 、 $\mathbf{B}(t)$ 都与 t 无关, 即可令:

$$\mathbf{A}(t) = \mathbf{A} = \begin{pmatrix} 0 & 0 & 0 \cdots & 0 & 0 \\ 1-d_1 & 0 & 0 \cdots & 0 & 0 \\ 0 & 1-d_2 & 0 \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & 1-d_{m-1} & 0 \end{pmatrix}_{m \times m}, \quad \mathbf{B}(t) = \mathbf{B} = \begin{pmatrix} 0 \cdots 0 b_{i_1}^* \cdots b_{i_2}^* 0 \cdots 0 \\ 0 \cdots 0 0 \cdots 0 0 \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 \cdots 0 0 \cdots 0 0 \cdots 0 \end{pmatrix}_{m \times m},$$

从而:

$$\mathbf{x}(t+1) = [\mathbf{A} + \beta(t)\mathbf{B}]\mathbf{x}(t).$$

上式表示, 只要总和生育率确定下来, 人口的变化规律就可以确定下来.

2.6.2 矩阵在力学中的应用 (一)

定义 2-17 在一连续介质内部, 任取一封闭曲面 Σ , 它把介质分成 Σ 内和 Σ 外两部分. 设 Σ 上点 P 的外单位法向量为 $\mathbf{v}$, 点 P 附近取面元 $\Delta\sigma$, 作用在 $\Delta\sigma$ 的面力向量为 $\Delta\mathbf{f}$. 假定介质在 Σ 外部对 Σ 内部的作用力表现为作用在 Σ 上的面力, 如

果保持法向量 $\mathbf{v}$ 不变而令 $\Delta\sigma \rightarrow 0$ 时, $\frac{\Delta\mathbf{f}}{\Delta\sigma}$ 有确定的极限, 这个极限向量 (与点 $\Delta\sigma$

和法向量 $\mathbf{v}$ 有关) 称为应力 (stress), 可以记为 $\mathbf{T}(P, \mathbf{v})$, 力学中也记为 $\overset{\mathbf{v}}{\mathbf{T}}$, 即:

$$\overset{\mathbf{v}}{\mathbf{T}} = \mathbf{T}(P, \mathbf{v}) = \lim_{\Delta\sigma \rightarrow 0} \frac{\Delta\mathbf{f}}{\Delta\sigma}.$$

有了应力的一般概念, 现在考虑一种特殊的模型, 观察图 2-4 的直角坐标系中的正六面体介质.

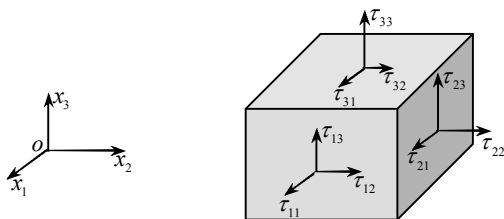


图 2-4

其中坐标轴与六面体的棱平行. 设平面 ΔS_1 是该六面体的一个表面, 其外法线方向指向 x_1 轴的正向. 将作用在 ΔS_1 上的应力矢量记为 $\overset{1}{T}$, 其沿坐标轴 x_1 、 x_2 、 x_3 方向的分量分别为 $\overset{1}{T}_1$ 、 $\overset{1}{T}_2$ 、 $\overset{1}{T}_3$; 同理 ΔS_2 、 ΔS_3 也是该六面体的表面, 其外法线方向分别指向 x_2 、 x_3 轴的正向, 表面上作用的应力沿坐标轴 x_1 、 x_2 、 x_3 方向的分量分别为 $\overset{2}{T}_1$ 、 $\overset{2}{T}_2$ 、 $\overset{2}{T}_3$ 和 $\overset{3}{T}_1$ 、 $\overset{3}{T}_2$ 、 $\overset{3}{T}_3$. 如果记 $\overset{i}{T}_j = \tau_{ij}$ ($i, j = 1, 2, 3$), 并将应力写成矩阵形式, 则有:

$$\begin{pmatrix} \overset{1}{T}_1 & \overset{1}{T}_2 & \overset{1}{T}_3 \\ \overset{2}{T}_1 & \overset{2}{T}_2 & \overset{2}{T}_3 \\ \overset{3}{T}_1 & \overset{3}{T}_2 & \overset{3}{T}_3 \end{pmatrix} = \begin{pmatrix} \tau_{11} & \tau_{12} & \tau_{13} \\ \tau_{21} & \tau_{22} & \tau_{23} \\ \tau_{31} & \tau_{32} & \tau_{33} \end{pmatrix}.$$

力学中可以证明: 只要知道应力分量 τ_{ij} , 就可以写出作用在外法线单位向量 $\boldsymbol{\nu}$ 的任意平面上的应力分量 $\overset{\boldsymbol{\nu}}{T}_1$ 、 $\overset{\boldsymbol{\nu}}{T}_2$ 、 $\overset{\boldsymbol{\nu}}{T}_3$, 即有如下柯西公式:

$$\overset{\boldsymbol{\nu}}{T}_j = \nu_1 \tau_{1j} + \nu_2 \tau_{2j} + \nu_3 \tau_{3j}, \quad (2-15)$$

其中 ν_1 、 ν_2 、 ν_3 为向量 $\boldsymbol{\nu}$ 沿坐标轴 x_1 、 x_2 、 x_3 方向的分量.

例 2-24 设物体内部某点的应力分量 (在 xyz 坐标系内) 为:

$$\begin{pmatrix} \tau_{11} & \tau_{12} & \tau_{13} \\ \tau_{21} & \tau_{22} & \tau_{23} \\ \tau_{31} & \tau_{32} & \tau_{33} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix},$$

试求作用在穿过该处的平面 $x + 3y + z = 1$ 外侧 (远离原点一侧) 上的应力分量.

解 该平面的单位法向量为 $x_1, x_2, x_3 = (1, 3, 1) / \sqrt{11}$, 根据式 (2-15), 有:

$$\begin{aligned} \overset{\boldsymbol{\nu}}{T}_1 &= \nu_1 \tau_{11} + \nu_2 \tau_{21} + \nu_3 \tau_{31} = \frac{1 \times 0 + 3 \times 1 + 1 \times 2}{\sqrt{11}} = \frac{5}{\sqrt{11}}, \\ \overset{\boldsymbol{\nu}}{T}_2 &= \nu_1 \tau_{12} + \nu_2 \tau_{22} + \nu_3 \tau_{32} = \frac{1 \times 1 + 3 \times 2 + 1 \times 0}{\sqrt{11}} = \frac{7}{\sqrt{11}}, \\ \overset{\boldsymbol{\nu}}{T}_3 &= \nu_1 \tau_{13} + \nu_2 \tau_{23} + \nu_3 \tau_{33} = \frac{1 \times 2 + 3 \times 0 + 1 \times 1}{\sqrt{11}} = \frac{3}{\sqrt{11}}. \end{aligned}$$

2.6.3 矩阵在经济分析中的应用——投入产出模型

投入产出模型是研究一个经济系统各部门之间“投入”与“产出”关系的线性模型，它是20世纪30年代由美国经济学家 Wassily Leontief 建立起来的。1949年他用含500个变量的500个线性方程组成的方程组来描述美国经济，这项工作的成果使他于1973年获得诺贝尔经济学奖，他的成就和获奖成为各国科学界用线性代数建立工程和经济模型的强大动力，也推动了线性代数的迅速发展。

一方面，在一个经济系统中，每个部门作为一个生产者，要为该系统内的各个部门（包括本部门）进行生产提供一定产品，又要满足系统外部对该产品的需求，即为“产出”；另一方面，每个部门为了生产其产品，必然又是消耗者，要消耗本部门 and 该系统内部其他部门所生产的产品，如原材料、设备、能源、人力等，即为“投入”，而且需要获得合理的利润。这些物质各方面的消耗和新创造的价值等于它的总产值，这就是“投入”与“产出”总的平衡关系。

1. 模型的构建

表2-1是某个经济系统中 n 个部门的投入产出平衡表。

表2-1 投入产出平衡表

生产部门	消耗部门 部门间流量	1	2	...	n	最终产品				总产品
						消费	积累	...	合计	
1		x_{11}	x_{12}	...	x_{1n}				y_1	x_1
2		x_{21}	x_{22}	...	x_{2n}				y_2	x_2
$\vdots$		$\vdots$	$\vdots$		$\vdots$				$\vdots$	$\vdots$
n		x_{n1}	x_{n2}	...	x_{nn}				y_n	x_n
新创造价值	劳动报酬	v_1	v_2	...	v_n					
	纯收入	m_1	m_2	...	m_n					
	合计	z_1	z_2	...	z_n					
总产品价值		x_1	x_2	...	x_n					

表2-1左上角部分由 n 个部门组成，每个部门既是生产部门，又是消耗部门。量 x_{ij} 表示第 j 部门消耗第 i 部门的产品，称为部门间的流量，它可按实物量计算，也可按价值量计算（这里采用价值量计算方法）。这一部分为投入产出平衡表最基本的部分。

表2-1右上角部分，每一行反映了某一部门从总产品中扣除补偿生产消耗后的余量，即不参加生产周转的最终产品的分配情况。其中 $y_1, y_2, \dots, y_n$ 分别表示第1, 第2, ...

..., 第 n 生产部门的最终产品, 而 $x_1, x_2, \dots, x_n$ 分别表示第1, 第2, ..., 第 n 生产部门的总产品, 也就是对应消耗部门的总产品价值.

表 2-1 左下角部分, 每一列表示该部门创建的价值 (也称净价值), 第 i 个部门新创造的价值为 z_i , 包括劳动报酬 v_i 和纯收入 m_i .

表 2-1 右下角部分反映国民收入的再分配, 比较复杂, 在此我们暂不讨论.

从表 2-1 的每一行来看, 每一生产部门分配给其他各部门的生产消耗的产品, 加上该部门的最终产品的价值, 应等于它的总产品, 即:

$$x_i = \sum_{j=1}^n x_{ij} + y_i \quad (i=1, 2, \dots, n), \quad (1)$$

方程组 (1) 称为产品分配平衡方程组, 其中 $\sum_{j=1}^n x_{ij}$ 为第 i 部门分配给各部门生产消耗的产品总和.

从表 2-1 的每一列来看, 每一消耗部门消耗其他各部门生产消耗的产品, 加上该部门新创造的价值, 应等于它的总产品的价值, 即:

$$x_j = \sum_{i=1}^n x_{ij} + z_j \quad (j=1, 2, \dots, n), \quad (2)$$

方程组 (2) 称为消耗平衡方程组, 其中 $\sum_{i=1}^n x_{ij}$ 为第 j 部门消耗各部门生产消耗的产品总和.

由方程组 (1) 和方程组 (2) 知:

$$\begin{aligned} \sum_{i=1}^n x_i &= \sum_{i=1}^n \left(\sum_{j=1}^n x_{ij} + y_i \right) = \sum_{i=1}^n \sum_{j=1}^n x_{ij} + \sum_{i=1}^n y_i, \\ \sum_{j=1}^n x_j &= \sum_{j=1}^n \left(\sum_{i=1}^n x_{ij} + z_j \right) = \sum_{j=1}^n \sum_{i=1}^n x_{ij} + \sum_{j=1}^n z_j, \end{aligned}$$

从而:

$$\sum_{i=1}^n y_i = \sum_{j=1}^n z_j,$$

即各部门最终产品的总和等于各部门新创造价值总和.

第 j 部门生产单位产品直接消耗第 i 部门的产品量称为第 j 部门对第 i 部门的直接消耗系数, 以 a_{ij} 表示, 即:

$$a_{ij} = \frac{x_{ij}}{x_j} \quad (i, j=1, 2, \dots, n).$$

换言之, a_{ij} 也就是第 j 部门生产单位产品需要第 i 部门直接分配给第 j 部门的产品量.

各部门间的直接消耗系数构成的 n 阶矩阵

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

称为直接消耗系数矩阵.

直接消耗系数 a_{ij} ($i, j = 1, 2, \dots, n$) 具有下列性质:

$$\textcircled{1} \quad 0 \leq a_{ij} < 1 \quad (i, j = 1, 2, \dots, n);$$

$$\textcircled{2} \quad \sum_{i=1}^n a_{ij} < 1 \quad (j = 1, 2, \dots, n).$$

将 $x_{ij} = a_{ij}x_j$ 代入方程组 (1), 得:

$$x_i = \sum_{j=1}^n a_{ij}x_j + y_i \quad (i = 1, 2, \dots, n). \quad (3)$$

若记 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, $\mathbf{y} = (y_1, y_2, \dots, y_n)^T$, 则上式可以写成:

$$\mathbf{x} = A\mathbf{x} + \mathbf{y} \text{ 或 } (I - A)\mathbf{x} = \mathbf{y}. \quad (4)$$

将 $x_{ij} = a_{ij}x_j$ 代入方程组 (1), 得:

$$x_j = \sum_{i=1}^n a_{ij}x_j + z_j \quad (j = 1, 2, \dots, n). \quad (5)$$

若记 $D = \text{diag}(\sum_{i=1}^n a_{i1}, \sum_{i=1}^n a_{i2}, \dots, \sum_{i=1}^n a_{in})$, $\mathbf{z} = (z_1, z_2, \dots, z_n)^T$, 则上式可以写成:

$$\mathbf{x} = D\mathbf{x} + \mathbf{z} \text{ 或 } (I - D)\mathbf{x} = \mathbf{z}. \quad (6)$$

由直接消耗系数的性质②知, $(I - A)^T$ 和 $I - D$ 都是严格对角占优矩阵, 故 $I - A$ 和 $I - D$ 都可逆 (参见习题 2.3 中的题 15).

2. 模型的求解和应用

利用投入产出数学模型进行经济分析时, 先要根据该经济系统报告期的数据求出直接消耗系数矩阵 A , 并假设在未来计划期内直接消耗系数 a_{ij} ($i, j = 1, 2, \dots, n$) 不发生变化, 可由方程组 (4) 和方程组 (6) 求得平衡方程组的解:

$$\mathbf{x} = (I - A)^{-1} \mathbf{y} \text{ 或 } \mathbf{x} = (I - D)^{-1} \mathbf{z}.$$

根据最终产品情况, 通过投入产出模型可求出该系统内各部门的总产值, 从而制定出合理的生产规划.

例 2-25 设有一个经济系统包括三个部门, 在某一个生产周期内各部门的直接消耗系数及最终产品分别为:

$$A = \begin{pmatrix} 0.25 & 0.1 & 0.1 \\ 0.2 & 0.2 & 0.1 \\ 0.1 & 0.1 & 0.2 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 245 \\ 90 \\ 175 \end{pmatrix},$$

求各部分的总产量及部门间的流量.

解 由已知条件得: $I-A=\begin{pmatrix} 0.75 & -0.1 & -0.1 \\ -0.2 & 0.8 & -0.1 \\ -0.1 & -0.1 & 0.8 \end{pmatrix},$

从而: $(I-A)^{-1}=\frac{10}{891}\begin{pmatrix} 126 & 18 & 18 \\ 34 & 118 & 19 \\ 20 & 17 & 116 \end{pmatrix},$

所以: $x=(I-A)^{-1}y=\frac{10}{891}\begin{pmatrix} 126 & 18 & 18 \\ 34 & 118 & 19 \\ 20 & 17 & 116 \end{pmatrix}\begin{pmatrix} 245 \\ 90 \\ 175 \end{pmatrix}=\begin{pmatrix} 400 \\ 250 \\ 300 \end{pmatrix},$

即第 1、2、3 部门的总产量分别为 400、250、300.

将 $x_1=400$ 、 $x_2=250$ 、 $x_3=300$ 代入 $x_{ij}=a_{ij}x_j$ ($i,j=1,2,3$), 可得各部门间的流量表如表 2-2 所示 (表中前 4 列).

表 2-2 各部门间的流量					
生产部门 \ 消耗部门 部门间流量	1	2	3	y	x
1	100	25	30	245	400
2	80	50	30	90	250
3	40	25	60	175	300