$F$ 中的元素也可以自然地被视为一个 $1\times 1$ 矩阵.

**相等** 两个**矩阵相等**是指它们的维数以及它们的对应元素相等. 因此, 作为矩阵, $\begin{pmatrix}0 & 0\end{pmatrix}$ 与 $\begin{pmatrix}0\\0\end{pmatrix}$ 是不相等的.

**向量** 仅有一行的矩阵称为一个**行向量** (row vector), 仅有一列的矩阵称为一个**列向量** (column vector). 一个 $n$ 元数组既可以表示成行向量的的形式, 也可以表示成列向量的形式, 因此在处理相关问题时, 需要留心数组与矩阵区别. 一般而言在解方程时, 一般习惯于将解写成列向量的形式. 另外, 在我们的课堂里, 一般习惯上用 (黑斜体) 小写字母 (主要为拉丁字母和希腊字母) 表示向量.

对于 $m\times n$ 矩阵 $\boldsymbol{A}=(a_{ij})$, 它的第 $i$ 个行向量是 $(a_{i1},a_{i2},\dots,a_{in})$ (暂记为 $\boldsymbol{\alpha}_i$), 是 $\boldsymbol{A}$ 的第 $i$ 行元素构成的行向量. $\boldsymbol{A}$ 的第 $j$ 个列向量是

$$\begin{pmatrix}a_{1j}\\a_{2j}\\\vdots\\a_{mj}\end{pmatrix}$$

(暂记为 $\boldsymbol{\beta}_j$), 是 $\boldsymbol{A}$ 的第 $j$ 列元素构成的的列向量. 此时,

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1\\\boldsymbol{\alpha}_2\\\vdots\\\boldsymbol{\alpha}_m\end{pmatrix}=\begin{pmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_n\end{pmatrix}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n),$$

从而 $\boldsymbol{A}$ 即可以表示成其行向量按列依次排列, 也可以表示成其列向量按行依次排列.

本节更多的定义, 需要学生课后自行阅读.

## 4.2 矩阵的运算

**标量乘法** 矩阵的**标量乘法** (数乘) 是向量的标量乘法的推广. 设 $\boldsymbol{A}=(a_{ij})$ 是一个矩阵, $\lambda\in F$ 是一个标量 (常数), 则 $\lambda\boldsymbol{A}$ 是将 $\boldsymbol{A}$ 的每个元素都乘以 $\lambda$ 后得到矩阵: $\lambda\boldsymbol{A}\coloneqq(\lambda a_{ij})$.

**例 4.2.1.** 设 $\boldsymbol{A}=\begin{pmatrix}4&8&2\\6&8&10\end{pmatrix}$, 则 $\frac{1}{2}\boldsymbol{A}=\begin{pmatrix}2&4&1\\3&4&5\end{pmatrix}$, 而 $3\boldsymbol{A}=\begin{pmatrix}12&24&6\\18&24&30\end{pmatrix}$.

**矩阵加法**　矩阵的加法是向量的加法的推广: 两个相同维数的矩阵的加法通过对应元素的加法得到. 具体来说, 若 $\boldsymbol{A}=(a_{ij})$ 和 $\boldsymbol{B}=(b_{ij})$ 都是 $m\times n$ 矩阵, 则 $\boldsymbol{A}+\boldsymbol{B}$ 也是一个 $m\times n$ 矩阵, 其第 $(i,j)$ 元素为 $a_{ij}+b_{ij}$, 即 $\boldsymbol{A}+\boldsymbol{B}\coloneqq(a_{ij}+b_{ij})$.

**例 4.2.2.**

$$\begin{pmatrix}3&2&1\\4&5&6\end{pmatrix}+\begin{pmatrix}2&2&2\\1&2&3\end{pmatrix}=\begin{pmatrix}5&4&3\\5&7&9\end{pmatrix},$$
$$\begin{pmatrix}2\\1\\8\end{pmatrix}+\begin{pmatrix}-8\\3\\2\end{pmatrix}=\begin{pmatrix}-6\\4\\10\end{pmatrix}.$$

相同维数的矩阵的减法 $\boldsymbol{A}-\boldsymbol{B}$ 定义为 $\boldsymbol{A}+(-1)\boldsymbol{B}$, 从而新得到的矩阵的元素为对应元素的减法, 即 $\boldsymbol{A}-\boldsymbol{B}\coloneqq(a_{ij}-b_{ij})$.

**例 4.2.3.**

$$\begin{pmatrix}2&4\\3&1\end{pmatrix}-\begin{pmatrix}4&5\\2&3\end{pmatrix}=\begin{pmatrix}-2&-1\\1&-2\end{pmatrix}.$$

我们习惯上会将 $m\times n$ 的所有元素为 0 的矩阵记作 $\boldsymbol{O}_{m\times n}$ 或 $\boldsymbol{O}$, 称之为 ($m\times n$ 的) **零矩阵**. 不难验证,

$$\boldsymbol{A}+\boldsymbol{O}=\boldsymbol{O}+\boldsymbol{A}=\boldsymbol{A}.$$

**矩阵的乘法及线性方程组**　以上定义的矩阵的数乘与加法明显脱胎于数组向量的相应运算. 接下来要介绍的矩阵乘法则要微妙一些. 当然，我们可以把矩阵乘法定义为像加法一样的坐标操作, 事实上, 矩阵的哈达玛德乘法 (Hadamard multiplication) 就是这样的. 但我们的动机并不仅仅是为了下定义, 而是为了下一个有用的定义用于解决问题. 我们这儿采用的矩阵的乘法的定义, 某种意义上来说, 来源于线性方程组的简化表示. 需要提及的是, 在数学里引入新的概念或者工具时, 我们经常会问: 它与之前已知的数学对象之间有怎样的联系, 以及引入这样的新的数学对象有什么意义, 或者说我们为什么要引入这一的新的对象?

**行向量 $\times$ 列向量**　设

$$\boldsymbol{\alpha}=\begin{pmatrix}a_1&a_2&\cdots&a_n\end{pmatrix}\qquad\text{与}\qquad\boldsymbol{\beta}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$$

分别是一个 $1\times n$ 的行向量和一个 $n\times 1$ 的列向量, 即分别为一个 $n$ 维行向量和一个 $n$ 维列向量. 定义矩阵乘法

$$\boldsymbol{\alpha}\boldsymbol{\beta} := a_1b_1 + a_2b_2 + \cdots + a_nb_n.$$

这是一个标量, 即 $1\times 1$ 矩阵. 它可以视作之前提到的数组向量的点乘的一个变形.

**例 4.2.4.**

$$\begin{pmatrix} 2 & 1 & -3 & 4 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \\ 1 \\ -2 \end{pmatrix} = 2\cdot 3 + 1\cdot 2 + (-3)\cdot 1 + 4\cdot(-2) = -3.$$

**例 4.2.5.** 线性方程 $3x_1 + 2x_2 + 5x_3 = 4$ 可以表示成

$$\begin{pmatrix} 3 & 2 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 4$$

的形式.

**矩阵 $\times$ 列向量** 给定一个 $m\times n$ 矩阵 $\boldsymbol{A}=(a_{ij})$ 和一个 $n$ 维列向量 $\boldsymbol{b}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$, 并设 $\boldsymbol{A}=\begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \vdots \\ \boldsymbol{\alpha}_m \end{pmatrix}$ 是由其行向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m$ 按列有序排列得到的. 注意到对于 "行向量 $\times$ 列向量" 的乘法, 我们已有定义, 因此, 对于矩阵乘积 $\boldsymbol{A}\boldsymbol{b}$, 我们可以将其定义为一个 $m$ 维列向量, 由这些 $\boldsymbol{\alpha}_i\boldsymbol{b}$ 按列有序排列得到:

$$\boldsymbol{A}\boldsymbol{b} = \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \vdots \\ \boldsymbol{\alpha}_m \end{pmatrix}\boldsymbol{b} := \begin{pmatrix} \boldsymbol{\alpha}_1\boldsymbol{b} \\ \boldsymbol{\alpha}_2\boldsymbol{b} \\ \vdots \\ \boldsymbol{\alpha}_m\boldsymbol{b} \end{pmatrix}.$$

具体来说, $m$ 维列向量 $\boldsymbol{Ab}$ 的第 $i$ 个元素为 $a_{i1}b_1+a_{i2}b_2+\cdots+a_{in}b_n$. 由于一个 $n$ 维行向量可以视作一个 $1\times n$ 的矩阵, 这儿定义的乘法可以视作前述的 "行向量 $\times$ 列向量" 的乘法的推广. 需要注意的是, 这儿的乘法要求矩阵 $\boldsymbol{A}$ 的列数等于列向量 $\boldsymbol{b}$ 的行数.

**例 4.2.6.**

$$\begin{pmatrix} -3 & 1 \\ 2 & 5 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} -3\cdot 2+1\cdot 4 \\ 2\cdot 2+5\cdot 4 \\ 4\cdot 2+2\cdot 4 \end{pmatrix} = \begin{pmatrix} -2 \\ 24 \\ 16 \end{pmatrix}.$$

**例 4.2.7.** 线性方程组

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

可以表示成

$$\begin{pmatrix} 3 & 2 & 1 \\ 1 & -2 & 5 \\ 2 & 1 & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix}$$

的形式.

**注 4.2.8.** 在接下来的讲义中, 我们希望能反映科学家和工程师实际应用线性代数的方式, 而这意味着在定义和证明中处理的更多是矩阵的列, 而不是矩阵的元素. 其中的一个核心课题就是用另外一个观点来看待矩阵与向量的乘积 $\boldsymbol{Ab}$, 将其作为关于 $\boldsymbol{A}$ 的列的一个线性组合.

若令 $\boldsymbol{A}$ 的列向量依次为 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n$, 则 $b_1\boldsymbol{\beta}_1+b_2\boldsymbol{\beta}_2+\cdots+b_n\boldsymbol{\beta}_n$ 是 $n$ 个 $m$ 维列向量的线性组合, 从而仍然是一个 $m$ 维列向量. 研究其第 $i$ 个分量, 不难看出相应的元素为 $b_1a_{i1}+b_2a_{i2}+\cdots+b_na_{in}$, 即 $m$ 维列向量 $\boldsymbol{Ab}$ 的第 $i$ 个分量. 这说明

$$\boldsymbol{Ab}=b_1\boldsymbol{\beta}_1+b_2\boldsymbol{\beta}_2+\cdots+b_n\boldsymbol{\beta}_n.$$

特别地, 这意味着 $\boldsymbol{Ab}$ 是 $\boldsymbol{A}$ 的列向量的一个线性组合.

这种现代方法简化了许多论述, 且将向量空间思想和线性方程组的研究联系在一起.

**例 4.2.9.** 对于 $m \times n$ 的线性方程组

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

若令

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

则利用上面引入的矩阵的乘法, 原线性方程组可以写成 $\boldsymbol{Ax} = \boldsymbol{b}$ 的形式.

若是结合注 *4.2.8* 中的观察, 可以看出, 线性方程组 $\boldsymbol{Ax} = \boldsymbol{b}$ 有解的充要条件是常数项列向量 $\boldsymbol{b}$ 是系数矩阵 $\boldsymbol{A}$ 的列向量的一个线性组合.

**矩阵 $\times$ 矩阵** 更一般地, 给定两个矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 并假定 $\boldsymbol{B} = \begin{pmatrix} \boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_p \end{pmatrix}$ 由其列向量 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_p$ 按行依次排列得到. 这些列向量的行数即为 $\boldsymbol{B}$ 的行数. 若这些行数与 $\boldsymbol{A}$ 的列数相等, 则乘法 $\boldsymbol{A\beta}_j$ 是有定义的, 而且得到的是一个列向量. 于是, 我们将定义矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的乘法, 规定其为一个新的矩阵, 记作 $\boldsymbol{AB}$, 由这些列向量 $\boldsymbol{A\beta}_1, \boldsymbol{A\beta}_2, \ldots, \boldsymbol{A\beta}_p$ 按行依次排列得到:

$$\boldsymbol{AB} \coloneqq \begin{pmatrix} \boldsymbol{A\beta}_1 & \boldsymbol{A\beta}_2 & \cdots & \boldsymbol{A\beta}_p \end{pmatrix}.$$

因此, 若 $\boldsymbol{A} = (a_{ij}) \in F^{m \times n}$, $\boldsymbol{B} = (b_{ij}) \in F^{n \times p}$, 则 $\boldsymbol{AB} = (c_{ij}) \in F^{m \times p}$, 其中 $\boldsymbol{AB}$ 的第 $(i, j)$ 元素为

$$\begin{aligned} c_{ij} &= \boldsymbol{A\beta}_j \text{ 的第 } i \text{ 个分量} = \boldsymbol{A} \text{ 的第 } i \text{ 个行向量乘以 } \boldsymbol{B} \text{ 的第 } j \text{ 个列向量} \\ &= \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}. \end{aligned} \tag{4.1}$$

由于一个 $p$ 维列向量可以视作一个 $p \times 1$ 的矩阵, 这儿定义的乘法可以视作前述的 “矩阵 $\times$ 列向量” 的乘法的推广.

若结合注 4.2.8 来看, 易知 $\boldsymbol{AB}$ 的每个列向量都是 $\boldsymbol{A}$ 的列向量组的一个线性组合.

**习题 4.2.10.** 验证: 矩阵乘法 $\boldsymbol{AB}$ 中的每个行向量都是 $\boldsymbol{B}$ 的行向量组的一个线性组合.

**例 4.2.11.**

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

$$= \begin{pmatrix} -14 & 1 & -3 \\ 12 & 6 & 30 \\ -14 & -2 & -15 \end{pmatrix}.$$

**注 4.2.12.** (1) 矩阵的乘法一定要遵守“第一个矩阵的列数 = 第二个矩阵的行数”这一要求.

(2) 矩阵的乘法一般不满足交换性. <u>学生直接验证</u>:

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}.$$

因此, 在我们谈两个矩阵的乘法时需要注意其顺序. 我们称 $\boldsymbol{AB}$ 为 $\boldsymbol{A}$ 左乘 $\boldsymbol{B}$, 也称之为 $\boldsymbol{B}$ 右乘 $\boldsymbol{A}$.

(3) 消去律对矩阵乘法不成立, 即若 $\boldsymbol{AB} = \boldsymbol{AC}$, 即便 $\boldsymbol{A}$ 不是零矩阵, 仍然有可能 $\boldsymbol{B} \neq \boldsymbol{C}$. 例如, 对于 $\boldsymbol{A} = \begin{pmatrix} 2 & -3 \\ -4 & 6 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 8 & 4 \\ 5 & 5 \end{pmatrix}$, 以及 $\boldsymbol{C} = \begin{pmatrix} 5 & -2 \\ 3 & 1 \end{pmatrix}$, 我们有 $\boldsymbol{AB} = \begin{pmatrix} 1 & -7 \\ -2 & 14 \end{pmatrix} = \boldsymbol{AC}$, 但是 $\boldsymbol{B} \neq \boldsymbol{C}$. 事实上, 为了让消去律成立, 我们一般会需要这儿的 $\boldsymbol{A}$ 在作矩阵乘法时有**逆元**存在. 这一概念将在稍后介绍.

(4) 等价地, 若乘积 $\boldsymbol{AB}$ 为零矩阵, 我们一般也得不到 $\boldsymbol{A}$ 或 $\boldsymbol{B}$ 为零矩阵. 例如, 可直接验证, 在上面一条的例子里, $\boldsymbol{A}(\boldsymbol{B}-\boldsymbol{C}) = \boldsymbol{O}$, 但是 $\boldsymbol{A} \neq \boldsymbol{O} \neq \boldsymbol{B}-\boldsymbol{C}$.

(5) 若矩阵 $\boldsymbol{A}$ 是 $n \times n$ 的, 则称 $\boldsymbol{A}$ 是 $n$ **阶方阵** (*square matrix of order n*). 对于 $n$ 阶方阵 $\boldsymbol{A} = (a_{ij})$, 称 $a_{11}, a_{22}, \ldots, a_{nn}$ 为 $\boldsymbol{A}$ 的**对角元**, 称这些元素在 $\boldsymbol{A}$ 的 (主) **对角线上** (*diagonal*). 若方阵 $\boldsymbol{A}$ 的所有非对角元都是 0, 则称 $\boldsymbol{A}$ 是一个**对角矩阵**; 此时, 可以将该对角矩阵记作

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix} =: \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn}).$$

若 $n$ 阶对角矩阵的所有对角元都是 1, 则称之为 ($n$ 阶) **单位矩阵** (*identity matrix*), 记作 $\boldsymbol{I}_n$ 或 $\boldsymbol{I}$. 不难看出, $\boldsymbol{I} = (\delta_{ij})$, 其中

$$\delta_{ij} \coloneqq \begin{cases} 1 & \text{如果 } i = j, \\ 0 & \text{如果 } i \neq j \end{cases}$$

是 ***Kronecker* 符号** (*Kronecker symbol*). 更一般地, 若 ($n$ 阶) 对角矩阵的所有对角元都是 $\lambda \in F$, 则称之为一个**数量矩阵**, 记作 $\lambda\boldsymbol{I}$; 显然, 它是单位矩阵 $\boldsymbol{I}$ 关于标量 $\lambda$ 所作的标量乘法.

(6) 若 $\boldsymbol{A} \in F^{m\times n}$, 则可直接验证, $\boldsymbol{I}_m\boldsymbol{A} = \boldsymbol{A} = \boldsymbol{A}\boldsymbol{I}_n$. 另外, 在矩阵相乘允许的条件下, 零矩阵与任何矩阵左乘或右乘都得到零矩阵 (维数可能会改变).

(7) 有关矩阵的加法、乘法的运算的基本性质, 由学生自行阅读教材的定理 *4.2.1* 和定理 *4.2.2*. 仅仅强调一下矩阵乘法的结合律: $(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$. 从而一长串矩阵的乘法, 在乘法运算允许 (*conformable for multiplication*) 的条件下, 与计算乘法时 (打括号) 的顺序无关. 刚刚的那个式子可以简记为 $\boldsymbol{ABC}$.

为了验证该结合律, 由乘法的定义, 不妨假定 $\boldsymbol{A} \in F^{m\times n}$, $\boldsymbol{B} \in F^{n\times p}$, $\boldsymbol{C} \in F^{p\times q}$. 为了说明的方便, 在这儿, 对于一个矩阵右下角的双重下标 $ij$, 我们是在表示该矩阵的第 $(i, j)$ 元素. 反复利用公式 (4.1), 我们有

$$\begin{aligned} ((\boldsymbol{AB})\boldsymbol{C})_{ij} &= \sum_{k=1}^{p} (\boldsymbol{AB})_{ik}\boldsymbol{C}_{kj} = \sum_{k=1}^{p} \left( \sum_{l=1}^{n} \boldsymbol{A}_{il}\boldsymbol{B}_{lk} \right) \boldsymbol{C}_{kj}, \\ (\boldsymbol{A}(\boldsymbol{BC}))_{ij} &= \sum_{l=1}^{n} \boldsymbol{A}_{il}(\boldsymbol{BC})_{lj} = \sum_{l=1}^{n} \boldsymbol{A}_{il} \left( \sum_{k=1}^{p} \boldsymbol{B}_{lk}\boldsymbol{C}_{kj} \right). \end{aligned}$$

利用交换求和顺序的技巧, 可以看出上面的两式相等. 由于 $(\boldsymbol{AB})\boldsymbol{C}, \boldsymbol{A}(\boldsymbol{BC}) \in F^{m\times q}$, 而这两个矩阵的元素对应相等, 故有 $(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$.