Chapter 2

矩阵

2.1 向量

我们在第一章 §1.2 已经介绍过了矩阵的行和列，现在我们称 $\mathbf{a} = (a_1, \cdots, a_m)$，$a_i \in \mathbb{R}$ 为 m 维行向量，$\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ 为 n 维列向量。

2.1.1 向量空间 (坐标空间)

我们记 $\mathbb{R}^{1\times n} = \{(a_1, \cdots, a_n) | a_i \in \mathbb{R}\}$，$\mathbb{R}^{n\times 1} = \left\{ \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} | b_i \in \mathbb{R} \right\}$，并将后者简记成 $\mathbb{R}^n$。为了节省空间，我们把 $\mathbb{R}^n$ 中的列向量简记为 $(b_1, \cdots b_n)^t$，即列向量写成行向量的转置。我们在 $\mathbb{R}^n$ 上定义如下的加法和数乘运算：

(1) 设 $\mathbf{x} = (x_1, \cdots, x_n)^t$，$\mathbf{y} = (y_1, \cdots, y_n)^t$，定义加法 $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \cdots, x_n + y_n)^t$；

(2) 设 $\mathbf{x} = (x_1, \cdots, x_n)^t$，$\alpha \in \mathbb{R}$，定义数乘 $\alpha\mathbf{x} = (\alpha x_1, \cdots, \alpha x_n)^t$。

另外我们记 $\mathbf{0} = (0, \cdots, 0)^t$，显然 $\forall \mathbf{x} \in \mathbb{R}^n$，$\mathbf{x} + \mathbf{0} = \mathbf{x}$。以上定义对行向量同理。

向量的加法和数乘满足以下运算律：

(1) 加法交换律：$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$；

(2) 加法结合律：$(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$；

(3) 加法单位元：$\forall \mathbf{x} \in \mathbb{R}^n$，$\mathbf{x} + \mathbf{0} = \mathbf{x}$；

(3) 加法逆元：$\forall \mathbf{x} \in \mathbb{R}^n$，$\exists \mathbf{y} \in \mathbb{R}$，$\mathbf{x} + \mathbf{y} = \mathbf{0}$；

(5) 数乘结合律：$\alpha, \beta \in \mathbb{R}$，则 $(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x})$；

(6) 数乘单位元：$1 \cdot \mathbf{x} = \mathbf{x}$；

(7) 数乘对向量加法的两个分配律：$\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$; $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$。

例 2.1.1 设

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

记 $A = (a_{ij})_{m\times n}$，$\mathbf{b} = (b_1, \cdots, b_m)^t$，则 (L) 可以表示为

$$x_1\mathbf{A}^{(1)} + \cdots + x_n\mathbf{A}^{(n)} = \mathbf{b}.$$

其中 $\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(n)}, \mathbf{b} \in \mathbb{R}^m$。即线性方程组可以视作列向量的线性组合。

2.1.2 线性相关性

设 $\mathbf{v}_1, \cdots, \mathbf{v}_k \in \mathbb{R}^n$，$\alpha_1, \cdots, \alpha_k \in \mathbb{R}$，则 $\alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k$ 称为 $\mathbf{v}_1, \cdots, \mathbf{v}_k$ 的一个线性组合，$\alpha_1, \cdots, \alpha_k \in \mathbb{R}$ 称为该线性组合的系数。

例 2.1.2 在例 2.1.1 中，(L) 相容 $\Longleftrightarrow$ $\mathbf{b}$ 是 $\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(n)}$ 的线性组合。

例 2.1.3 设 $\mathbf{u} = (1, 1, 1)^t$，$\mathbf{v}_1 = (1, 2, 3)^t$，$\mathbf{v}_2 = (1, 0, 0)^t$，再设 $\mathbf{u} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2$，则

$$\begin{pmatrix} \alpha_1 + \alpha_2 \\ 2\alpha_1 \\ 3\alpha_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

即

$$\begin{cases} \alpha_1 + \alpha_2 = 1 \\ 2\alpha_1 = 1 \\ 3\alpha_1 = 1 \end{cases}$$

不相容。所以 $\mathbf{u}$ 不是 $\mathbf{v}_1$，$\mathbf{v}_2$ 的线性组合。

定义 2.1.1 设 $\mathbf{v}_1, \cdots, \mathbf{v}_k \in \mathbb{R}^n$，如果存在不全为 0 的 $\alpha_1, \cdots, \alpha_k \in \mathbb{R}$，使得 $\alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k = \mathbf{0}$ 成立，则称 $\mathbf{v}_1, \cdots, \mathbf{v}_k$ 线性相关；反之，若 $\alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k = \mathbf{0} \Longrightarrow \alpha_1 = \cdots = \alpha_k = 0$，则称 $\mathbf{v}_1, \cdots, \mathbf{v}_k$ 线性无关。

例 2.1.4 对齐次方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases} \tag{H}$$

记 $A = (a_{ij})$。则 (H) 有非平凡解 $\Longleftrightarrow$ $\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(n)}$ 线性相关。

例 2.1.5 判断 $\mathbf{v}_1 = (1, 0, 1)^t$，$\mathbf{v}_2 = (0, 1, 1)^t$，$\mathbf{v}_3 = (1, 1, 1)^t$ 是否线性相关。

解: 设 $\alpha_1\mathbf{v}_1+\alpha_2\mathbf{v}_2+\alpha_3\mathbf{v}_3=\mathbf{0}$，则

$$\begin{cases}\alpha_1+\alpha_3=0\\ \alpha_2+\alpha_3=0\\ \alpha_1+\alpha_2+\alpha_3=0\end{cases}$$

解得 $\alpha_1=\alpha_2=\alpha_3=0$。即 $\mathbf{v},\mathbf{v}_2,\mathbf{v}_3$ 线性无关。

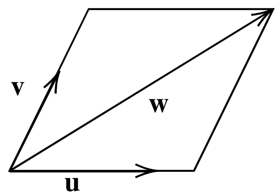
线性相关的几何意义:

(1)$\mathbf{u}\in\mathbb{R}^n$ 线性相关即 $\mathbf{u}=\mathbf{0}$。

(2)$\mathbf{u},\mathbf{v}\in\mathbb{R}^n$ 线性相关，则 $\mathbf{u},\mathbf{v}$ “同向”或“反向”共线 (平行)。

(3)$\mathbf{u},\mathbf{v},\mathbf{w}\in\mathbb{R}^3$ 线性相关，则 $\mathbf{u},\mathbf{v},\mathbf{w}$ 共面。

(2) 共线



(3) 共面

命题 2.1.1 设 $\mathbf{v}_1,\cdots,\mathbf{v}_k\in\mathbb{R}^n$，$1\leqslant i\leqslant k$，则

(i) 如果 $\mathbf{v}_1,\cdots,\mathbf{v}_i$ 线性相关，则 $\mathbf{v}_1,\cdots,\mathbf{v}_k\in\mathbb{R}^n$ 线性相关；

(ii) 如果 $\mathbf{v}_1,\cdots,\mathbf{v}_i,\cdots,\mathbf{v}_k$ 线性无关，则 $\mathbf{v}_1,\cdots,\mathbf{v}_i$ 线性无关；

(iii) $\mathbf{v}_1,\cdots,\mathbf{v}_k$ 线性相关 $\Longleftrightarrow$ $\mathbf{v}_1,\cdots,\mathbf{v}_k$ 中某个向量是其它向量的线性组合；

(iv) 设 $\mathbf{v}\in\mathbb{R}^n$ 且 $\mathbf{v}_1,\cdots,\mathbf{v}_k$ 线性无关，则 $\mathbf{v},\mathbf{v}_1,\cdots,\mathbf{v}_k$ 线性相关 $\Longleftrightarrow$ $\exists!$一组$\alpha_1,\cdots,\alpha_k\in\mathbb{R}$ 使得 $\mathbf{v}=\alpha_1\mathbf{v}_1+\cdots+\alpha_k\mathbf{v}_k$。

证明: 只证 (iv)，(i)(ii)(iii) 留作练习。
($\Longleftarrow$) 方向是显然的，下面证明 ($\Longrightarrow$) 方向。
由于 $\mathbf{v},\mathbf{v}_1,\cdots,\mathbf{v}_k$ 线性相关，故存在不全为 0 的 $\alpha_0,\alpha_1,\cdots,\alpha_k\in\mathbb{R}$ 使得

$$\alpha_0\mathbf{v}+\alpha_1\mathbf{v}_1+\cdots+\alpha_k\mathbf{v}_k=\mathbf{0}. \tag{2.1.1}$$

若 $\alpha_0=0$，则 $\alpha_1,\cdots,\alpha_k$ 不全为 0 且 $\alpha_1\mathbf{v}_1+\cdots+\alpha_k\mathbf{v}_k=\mathbf{0}$，这与 $\mathbf{v}_1,\cdots,\mathbf{v}_k$ 线性无关矛盾！故 $\alpha_0\neq 0$，则移项并在 (2.1.1) 式两边同时除以 α_0 得

$$\mathbf{v}=\left(-\frac{\alpha_1}{\alpha_0}\right)\mathbf{v}_1+\cdots+\left(-\frac{\alpha_k}{\alpha_0}\right)\mathbf{v}_k. \tag{2.1.2}$$

将 $(-\frac{\alpha_i}{\alpha_0})$ 视作新的 α_i 即可。
下证唯一性。若另有 $\mathbf{v}=\lambda_1\mathbf{v}_1+\cdots+\lambda_k\mathbf{v}_k$，则将它与 (2.1.2) 式相减得 $(-\frac{\alpha_1}{\alpha_0}-\lambda_1)\mathbf{v}_1+\cdots+(-\frac{\alpha_k}{\alpha_0}-\lambda_k)\mathbf{v}_k=\mathbf{0}$，则由 $\mathbf{v}_1,\cdots,\mathbf{v}_k$ 线性无关可得 $-\frac{\alpha_i}{\alpha_0}-\lambda_i=0$，$i=1,\cdots,k$，即唯一性得证。 □

例 2.1.6 记 $\mathbf{e}^{(j)}=(0,\cdots,1,\cdots,0)^t\in\mathbb{R}^n$(第 j 个位置是 1，其他为 0)，则

(1)$\mathbf{e}^{(1)},\cdots,\mathbf{e}^{(n)}$ 线性无关；

(2)$\forall\mathbf{x}\in\mathbb{R}^n$，$\exists!$一组$\alpha_1,\cdots,\alpha_n\in\mathbb{R}$ 使得 $\mathbf{x}=\alpha_1\mathbf{e}^{(1)}+\cdots+\alpha_n\mathbf{e}^{(n)}$，即 $\mathbf{x}=(\alpha_1,\cdots,\alpha_n)^t$。

例 2.1.7 求证：$\mathbb{R}^n$ 中任意 $n+1$ 个向量一定线性相关。

证明: 设 $\mathbf{v}_j = (v_{1j}, \cdots, v_{nj})^t$，其中 $j = 1, 2, \cdots, n+1$，若有 $\alpha_1, \cdots, \alpha_{n+1} \in \mathbb{R}$，使得 $\alpha_1\mathbf{v}_1 + \cdots + \alpha_{n+1}\mathbf{v}_{n+1} = \mathbf{0}$ ，则

$$\begin{cases} v_{11}\alpha_1 + \cdots + v_{1,n+1}\alpha_{n+1} = 0 \\ v_{21}\alpha_1 + \cdots + v_{2,n+1}\alpha_{n+1} = 0 \\ \qquad\qquad\vdots \\ v_{n1}\alpha_1 + \cdots + v_{n,n+1}\alpha_{n+1} = 0 \end{cases}$$

即 $\alpha_1, \cdots, \alpha_{n+1}$ 是该齐次方程组的解。由定理 1.2.3 可得该方程组有非平凡解，即 $\mathbf{v}_1, \cdots, \mathbf{v}_{n+1}$ 线性相关。 □

引理 2.1.1 (**线性组合引理**) 设 $\mathbf{u}_1, \cdots, \mathbf{u}_k, \mathbf{v}_1, \cdots, \mathbf{v}_l \in \mathbb{R}^n$ 且每个 $\mathbf{u}_i$ 都是 $\mathbf{v}_1, \cdots, \mathbf{v}_l$ 的线性组合，如果 $k > l$，则 $\mathbf{u}_1, \cdots, \mathbf{u}_k$ 线性相关。

证明: 设

$$\begin{aligned} \mathbf{u}_1 &= a_{11}\mathbf{v}_1 + \cdots + a_{1l}\mathbf{v}_l \\ \mathbf{u}_2 &= a_{21}\mathbf{v}_1 + \cdots + a_{2l}\mathbf{v}_l \\ &\vdots \\ \mathbf{u}_k &= a_{k1}\mathbf{v}_1 + \cdots + a_{kl}\mathbf{v}_l \end{aligned}$$

其中 $a_{ij} \in \mathbb{R}$ 都是已知的。我们想要证明 $\mathbf{u}_1, \cdots, \mathbf{u}_k$ 线性相关，即存在不全为 0 的 $\alpha_1, \cdots, \alpha_k \in \mathbb{R}$ 使得 $\sum_{i=1}^k \alpha_i\mathbf{u}_i = \mathbf{0}$。由于

$$\begin{aligned} &\sum_{i=1}^k \alpha_i\mathbf{u}_i \\ =& \sum_{i=1}^k \alpha_i(\sum_{j=1}^l a_{ij}\mathbf{v}_j) \\ =& \sum_{j=1}^l (\sum_{i=1}^k a_{ij}\alpha_i)\mathbf{v}_j \end{aligned}$$

只要对每个 $j = 1, \cdots, l$, 都有 $\sum_{i=1}^k a_{ij}\alpha_i = 0$, 则 $\sum_{i=1}^k \alpha_i\mathbf{u}_i = \mathbf{0}$ 成立。于是目标变为寻找以 $\alpha_1, \cdots, \alpha_k$ 为变元的方程组

$$\sum_{i=1}^k a_{ij}\alpha_i = 0, \quad j = 1, \cdots, l.$$

的非零解。由于方程组有 k 个变元，l 个方程，$k > l$，故由定理 1.2.3，该方程组有非零解。于是我们证明了原命题。 □

例 2.1.8 设 $\mathbf{u}_1 = \lambda_1\mathbf{v}$，$\mathbf{u}_2 = \lambda_2\mathbf{v}$，则 $\mathbf{u}_1, \mathbf{u}_2$ 线性相关。

2.1.3 极大线性无关组

我们也把向量的集合称为向量组。这一节我们将阐明向量集的一个基本性质：$\mathbb{R}^n$ 的一个向量的集合中一定存在“极大”的线性无关的子集，并且这种子集的元素个数是一个不变量。

定义 2.1.2 设 $T \subset \mathbb{R}^n$ 非空，若 T 中每个非空有限子集都是线性无关的，则称 T 是线性无关集。由例 2.1.7 知 $|T| < n+1$。

定义 2.1.3 设 $S \subset \mathbb{R}^n$ 非空，$T \subset S$ 是线性无关 (子) 集，如果 $\forall \sigma \in S \setminus T$, $T \cup \{\sigma\}$ 是线性相关的，则称 T 是 S 中的极大线性无关集。

引理 2.1.2 (**扩充引理**) 设 $S \subset \mathbb{R}^n$ 且 $T \subset S$ 是一个线性无关集，则存在 S 的一个极大线性无关组 $\widetilde{T}$ 满足 $T \subset \widetilde{T}$。

证明: 若 T 本身是极大的，则证明结束。
否则，存在 $\mathbf{v}_1 \in S \setminus T$ 使得 $T_1 = T \cup \{\mathbf{v}_1\}$ 是 S 中的线性无关集。若 T_1 是极大的，则证明结束。
否则，存在 $\mathbf{v}_2 \in S \setminus T$ 使得 $T_2 = T_1 \cup \{\mathbf{v}_2\}$ 是 S 中的线性无关集。
$\cdots\cdots$ 一直这样重复下去，例 2.1.7 保证了这个过程一定在有限步内终止，于是整个证明完成。□

以后我们会知道，扩充引理对抽象向量空间中的向量集也成立，但终止性由 Zorn 引理保证。

命题 2.1.2 设 $S \subset \mathbb{R}^n$ 且 $S \neq \emptyset$, $S \neq \{\mathbf{0}\}$，则

(i) S 中有极大线性无关组；

(ii) 设 T_1, T_2 是两个极大线性无关组，则 $|T_1| = |T_2|$。

证明: (i) 由条件可得存在向量 $\mathbf{v} \in S$, $\mathbf{v} \neq \mathbf{0}$，对 $\{\mathbf{v}\}$ 应用扩充引理即可。
(ii) 设 $T_1 = \{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$, $T_2 = \{\mathbf{v}_1, \cdots, \mathbf{v}_l\}$，则由极大线性无关组的定义，$\forall \mathbf{u}_i \in T_1$, $i = 1, \cdots, k$，有 $\mathbf{u}_i$ 是 $\mathbf{v}_1, \cdots, \mathbf{v}_l$ 的线性组合。于是由线性组合引理，$k \leqslant l$。同理 $l \leqslant k$，因此 $k = l$。□

推论 2.1.1 设 $S \subset \mathbb{R}^n, T = \{\mathbf{v}_1, \cdots, \mathbf{v}_m\} \subset S$ 是 S 的极大线性无关组，则对 $\forall \mathbf{v} \in S$, $\exists!$一组$\alpha_1, \cdots, \alpha_m \in \mathbb{R}$，使得 $\mathbf{v} = \sum_{i=1}^m \alpha_i \mathbf{v}_i$。

证明: 由于 T 是极大线性无关组，故 $\mathbf{v}, \mathbf{v}, \cdots \mathbf{v}_m$ 线性相关，于是由命题 2.1.1(iv) 知结论成立。□

推论 2.1.2 设 $S \subset \mathbb{R}^n$, $T = \{\mathbf{v}_1, \cdots, \mathbf{v}_m\} \subset S$，如果对 $\forall \mathbf{v} \in S$, $\exists!$一组$\alpha_1, \cdots, \alpha_m \in \mathbb{R}$，使得 $\mathbf{v} = \sum_{i=1}^m \alpha_i \mathbf{v}_i$，则 T 是 S 的极大线性无关组。

证明: 只需证 $\{\mathbf{v}_1, \cdots, \mathbf{v}_m\}$ 线性无关。用反证法。若 $\{\mathbf{v}_1, \cdots, \mathbf{v}_m\}$ 线性相关，由命题 2.1.1(iii)，不妨设 $\mathbf{v}_1 = \alpha_2 \mathbf{v}_2 + \cdots + \alpha_m \mathbf{v}_m$，则

$$\mathbf{v}_1 = 0 \cdot \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_m \mathbf{v}_m = 1 \cdot \mathbf{v}_1$$

即 $\mathbf{v}_1$ 有两种表出方式，这与存在唯一的表出方式矛盾！□

推论 2.1.3 设 $S \subset \mathbb{R}^n$，r 是 S 中极大线性无关组中元素的个数，T 是 S 中的线性无关子集，如果 $|T| = r$，则 T 是极大线性无关组。

证明: 由扩充引理，存在极大线性无关组 $\widetilde{T}$ 满足 $T \subset \widetilde{T}$，又由 $|T| = |\widetilde{T}| = r$ 知 $T = \widetilde{T}$。□

2.1.4 子空间

定义 2.1.4 设 $V \subset \mathbb{R}^n$ 非空，如果对 $\forall \mathbf{x}, \mathbf{y} \in V$ 及 $\alpha \in \mathbb{R}$，有 $\mathbf{x}+\mathbf{y} \in V$, $\alpha\mathbf{x} \in V$ (即 V 对加法和数乘封闭)，则称 V 是 $\mathbb{R}^n$ 的子空间。

注 2.1.1 (1)V 是子空间，则必有 $\mathbf{0} \in V$;
(2)V 是子空间 $\Longleftrightarrow \forall \mathbf{x}, \mathbf{y} \in V$, $\alpha, \beta \in \mathbb{R}$，有 $\alpha\mathbf{x}+\beta\mathbf{y} \in V$。

例 2.1.9 齐次方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases} \tag{H}$$

的所有解的集合 V 是 $\mathbb{R}^n$ 的子空间。

证明: 由 $\mathbf{0} \in V$，故 $V \neq \emptyset$。v 另一方面，若 $\mathbf{u} = (\alpha_1, \cdots, \alpha_n)^t$, $\mathbf{v} = (\beta_1, \cdots, \beta_n)^t \in V$，则对 $\forall \alpha, \beta \in \mathbb{R}$，有

$$\sum_{j=1}^{n} a_{ij}(\alpha\alpha_j + \beta\beta_j) = \alpha(\sum_{j=1}^{n} a_{ij}\alpha_j) + \beta(\sum_{j=1}^{n} a_{ij}\beta_j) = 0.$$

即 $\alpha\mathbf{u}+\beta\mathbf{v}$ 也是 (H) 的解，即封闭性成立。 □

命题 2.1.3 任意多个子空间的交仍然是子空间。

直接验证封闭性即可证明。

但两个子空间的并一般不是子空间 (除非一个包含另一个，试证明之)。因此我们需要考虑同时包含两个子空间的最小的子空间，这就是和空间。

定义 2.1.5 设 $S, T \subset \mathbb{R}^n$ 非空，我们定义集合 $U = \{\mathbf{x}+\mathbf{y} | \mathbf{x} \in S,\ \mathbf{y} \in T\}$，则称 U 是 S, T 的和，记为 $U = S + T$。

命题 2.1.4 设 $U, V \subset \mathbb{R}^n$ 是子空间，则 $U + V$ 也是子空间。

证明: 设 $\mathbf{x}+\mathbf{y} \in U+V$，则 $\exists \mathbf{u}_1, \mathbf{u}_2 \in U$, $\mathbf{v}_1, \mathbf{v}_2 \in V$，使得 $\mathbf{x} = \mathbf{u}_1 + \mathbf{v}_1$, $\mathbf{y} = \mathbf{u}_2 + \mathbf{v}_2$。下面验证封闭性。对 $\forall \alpha, \beta \in \mathbb{R}$，有

$$\alpha\mathbf{x}+\beta\mathbf{y} = \alpha(\mathbf{u}_1+\mathbf{v}_1) + \beta(\mathbf{u}_2+\mathbf{v}_2) = \underbrace{(\alpha\mathbf{u}_1+\beta\mathbf{u}_2)}_{\in U} + \underbrace{(\alpha\mathbf{v}_1+\beta\mathbf{v}_2)}_{\in V}$$

即 $\alpha\mathbf{x}+\beta\mathbf{y} \in U+V$。由注 2.1.1(2) 即得结论。 □

命题 2.1.5 设 $\mathbf{v}_1, \cdots, \mathbf{v}_k \in \mathbb{R}^n$，定义它们的线性包络 (linear span):

$$<\mathbf{v}_1, \cdots, \mathbf{v}_k> = \{\alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k | \alpha_i \in \mathbb{R},\ i = 1, \cdots, k\}.$$

则容易验证 $<\mathbf{v}_1, \cdots, \mathbf{v}_k>$ 是子空间 (也称为由 $\mathbf{v}_1, \cdots, \mathbf{v}_k$ 生成的子空间)。

证明: 设 $\mathbf{u}, \mathbf{v} \in <\mathbf{v}_1, \cdots, \mathbf{v}_k>$，则存在 $\alpha_1, \cdots, \alpha_k, \beta_1, \cdots, \beta_k$ 使得

$$\mathbf{u} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k$$
$$\mathbf{v} = \beta_1\mathbf{v}_1 + \cdots + \beta_k\mathbf{v}_k$$

则对 $\forall\lambda,\mu\in\mathbb{R}$，有

$$\lambda\mathbf{u}+\mu\mathbf{v}=\lambda(\alpha_1\mathbf{v}_1+\cdots+\alpha_k\mathbf{v}_k)+\mu(\beta_1\mathbf{v}_1+\cdots+\beta_k\mathbf{v}_k)=\sum_{i=1}^{k}(\lambda\alpha_i+\mu\beta_i)\mathbf{v}_i\in<\mathbf{v}_1,\cdots,\mathbf{v}_k>.$$

即结论成立。 □

推论 2.1.4 设 $\mathbf{v}_1,\cdots,\mathbf{v}_k\in\mathbb{R}^n$，$V$ 是含有 $\mathbf{v}_1,\cdots,\mathbf{v}_k$ 的子空间，则 $<\mathbf{v}_1,\cdots,\mathbf{v}_k>\subset V$。

证明留作练习。这个推论表明 $<\mathbf{v}_1,\cdots,\mathbf{v}_k>$ 是含有 $\mathbf{v}_1,\cdots,\mathbf{v}_k$ 的最小子空间。

定义 2.1.6 设 U,V 是 $\mathbb{R}^n$ 的子空间，如果 $U\cap V=\{\mathbf{0}\}$，则称 $U+V$ 是直和，记为 $U\oplus V$。

命题 2.1.6 设 $U,V\subset\mathbb{R}^n$ 是子空间，则 $U+V$ 是直和 $\Longleftrightarrow$ 对 $\forall\mathbf{x}\in U+V$, $\exists!\mathbf{u}\in U$和$\mathbf{v}\in V$ 使得 $\mathbf{x}=\mathbf{u}+\mathbf{v}$。

证明: ($\Longrightarrow$) 若 $\mathbf{x}$ 有两种表出方式：$\mathbf{x}=\mathbf{u}+\mathbf{v}=\mathbf{u}_1+\mathbf{v}_1$，则 $\mathbf{u}-\mathbf{u}_1=\mathbf{v}_1-\mathbf{v}\in U\cap V=\{\mathbf{0}\}$，即 $\mathbf{u}=\mathbf{u}_1$，$\mathbf{v}=\mathbf{v}_1$。
($\Longleftarrow$) 设 $\mathbf{w}\in U\cap V$，则 $\mathbf{w}$ 有两种表出方式：$\mathbf{w}=\mathbf{w}+\mathbf{0}=\mathbf{0}+\mathbf{w}$，由唯一性即得 $\mathbf{w}=\mathbf{0}$。 □

注 2.1.2 设 U_1,U_2 是子空间，则 $U_1\cap U_2\subset U_1\subset U_1+U_2$，但 $(U_1+U_2)\cap U_3\neq U_1\cap U_3+U_2\cap U_3$，即分配律一般不成立。例如，设在 $\mathbb{R}^2$ 中 $U_1=<\mathbf{e}_1>$，$U_2=<\mathbf{e}_2>$，则 $U_1\cap U_2=<\mathbf{0}>$，$U_1+U_2=<\mathbf{e}_1,\mathbf{e}_2>$。令 $U_3=<\mathbf{e}_1+\mathbf{e}_2>$，则 $(U_1+U_2)\cap U_3=U_3$，而 $U_1\cap U_3+U_2\cap U_3=<\mathbf{0}>$。

2.1.5 基底与维数

这一小节的内容和下册抽象向量空间的内容是一脉相承的。

定义 2.1.7 设 $V\subset\mathbb{R}^n$ 是子空间，且 $V\neq\{\mathbf{0}\}$，再设 $\mathbf{b}_1,\cdots,\mathbf{b}_d$ 线性无关，并满足 $V=<\mathbf{b}_1,\cdots,\mathbf{b}_d>$，则称 $\mathbf{b}_1,\cdots,\mathbf{b}_d$ 是 V 的一组基 (basis)。

命题 2.1.7 设 $V\subset\mathbb{R}^n$ 是子空间，且 $V\neq\{\mathbf{0}\}$，$\mathbf{b}_1,\cdots,\mathbf{b}_d\in V$，则 $\mathbf{b}_1,\cdots,\mathbf{b}_d$ 是 V 的一组基 $\Longleftrightarrow$ $\mathbf{b}_1,\cdots,\mathbf{b}_d$ 是 V 的一个极大线性无关组。

证明: ($\Longrightarrow$) 由于 $\mathbf{b}_1,\cdots,\mathbf{b}_d$ 是 V 的一组基，即 $\mathbf{b}_1,\cdots,\mathbf{b}_d$ 线性无关，假设它们不是极大线性无关的，则存在 $\mathbf{v}\in V$ 使得 $\mathbf{b}_1,\cdots,\mathbf{b}_d,\mathbf{v}$ 线性无关。另一方面，由基的定义有 $\mathbf{v}\in<\mathbf{b}_1,\cdots,\mathbf{b}_d>$，即存在 $\alpha_1,\cdots,\alpha_d\in\mathbb{R}$，使得 $\mathbf{v}=\alpha_1\mathbf{b}_1+\cdots+\alpha_d\mathbf{b}_d$，移项即可得到 $\mathbf{b}_1,\cdots,\mathbf{b}_d,\mathbf{v}$ 线性相关，矛盾！
($\Longleftarrow$) 设 $\mathbf{b}_1,\cdots,\mathbf{b}_d$ 是 V 中的极大线性无关组，即 $\forall\mathbf{v}\in V$, $\mathbf{b}_1,\cdots,\mathbf{b}_d,\mathbf{v}$ 线性相关，又因为 $\mathbf{b}_1,\cdots,\mathbf{b}_d$ 线性无关，由命题 2.1.1(iv) 及线性包络的定义 (命题 2.1.5) 知 $\mathbf{v}\in<\mathbf{b}_1,\cdots,\mathbf{b}_d>$，由 $\mathbf{v}$ 的任意性即有 $V\subset<\mathbf{b}_1,\cdots,\mathbf{b}_d>$，而另一个方向的包含关系是显然的，因此 $V=<\mathbf{b}_1,\cdots,\mathbf{b}_d>$。 □

注 2.1.3 若 V 有一组基 $\mathbf{b}_1,\cdots,\mathbf{b}_d$，则 $\forall\mathbf{v}\in V$, $\exists!$一组$\alpha_1,\cdots,\alpha_d\in\mathbb{R}$ 使得 $\mathbf{v}=\sum_{i=1}^{d}\alpha_i\mathbf{b}_i$。此时称 $(\alpha_1,\cdots,\alpha_d)^t$ 是 V 在 $\mathbf{b}_1,\cdots,\mathbf{b}_d$ 下的**坐标**。另外，由命题 2.1.2，V 中两组基含有的元素个数相同。

定义 2.1.8 设 $V\subset\mathbb{R}^n$ 是子空间，且 $V\neq\{\mathbf{0}\}$，$\mathbf{b}_1,\cdots,\mathbf{b}_d\in V$ 是 V 的一组基，则称 d 为 V 的维数 (dimension)，简记为 $\dim V=d$。特别地，$\{\mathbf{0}\}$ 的维数为 0。

例 2.1.10 $\mathbb{R}^n$ 有自然的基底 $\mathbf{e}^{(1)}=(1,0,\cdots,0)^t$，$\mathbf{e}^{(2)}=(0,1,\cdots,0)^t,\cdots\cdots,\mathbf{e}^{(n)}=(0,\cdots,0,1)^t$ 。于是 $\dim V=n$。

定理 2.1.1 (**基扩充定理**) 设 $V \subset \mathbb{R}^n$ 是子空间，$\dim V = d > 0$，$\mathbf{v}_1, \cdots, \mathbf{v}_k \in V$ 线性无关，则存在 $\mathbf{v}_{k+1}, \cdots, \mathbf{v}_d \in V$ 使得 $\mathbf{v}_1, \cdots, \mathbf{v}_k, \mathbf{v}_{k+1}, \cdots, \mathbf{v}_d$ 是一组基。

证明: 由扩充引理 (引理 2.1.2)，存在 $\mathbf{v}_{k+1}, \cdots, \mathbf{v}_d \in V$ 使得 $\mathbf{v}_1, \cdots, \mathbf{v}_d$ 是 V 的极大线性无关组。再由命题 2.1.7 可知 $\mathbf{v}_1, \cdots, \mathbf{v}_d$ 是 V 的一组基。 □

定理 2.1.2 (**包含定理**) 设 U, V 是 $\mathbb{R}^n$ 的子空间且 $U \subset V$，则 $U \subsetneq V \Longleftrightarrow \dim U < \dim V$。

证明: $(\Longleftarrow)$ 显然，下面证明 $(\Longrightarrow)$ 方向。
若 $V = \{\mathbf{0}\}$，则结论显然；否则设 $\mathbf{u}_1, \cdots, \mathbf{u}_d$ 是 U 的基，则存在 $\mathbf{v} \in V \setminus U$ 使得 $\mathbf{v} \notin <\mathbf{u}_1, \cdots, \mathbf{u}_d>$。
由命题 2.1.1(iv) 可知 $\mathbf{u}_1, \cdots, \mathbf{u}_d, \mathbf{v}$ 线性无关。再用定理 2.1.1 即可得到 $\dim V \geqslant d + 1 > \dim U$。□

注 2.1.4 该结论对非线性的情形不对。例如，设 S 是方程 $x(x^2 + y^2 - 1) = 0$ 的解集，则 y轴 $\subsetneq S$，但 $\dim$(y 轴)$= 1 = \dim S$。

定理 2.1.3 (**维数公式**) 设 $U_1, U_2 \subset \mathbb{R}^n$ 是子空间，则

$$\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = \dim U_1 + \dim U_2.$$

证明: 若 $U_1 = \{0\}$ 或 $U_2 = \{0\}$，则结论显然成立。
下设 $U_1 \cap U_2 \neq 0$，$\mathbf{w}_1, \cdots, \mathbf{w}_m$ 是 $U_1 \cap U_2$ 的一组基。则由基扩充定理，$\exists \mathbf{u}_1, \cdots, \mathbf{u}_s \in U_1$及$\mathbf{v}_1, \cdots, \mathbf{v}_t \in U_2$ 使得:
(A)$\mathbf{w}_1, \cdots, \mathbf{w}_m, \mathbf{u}_1, \cdots, \mathbf{u}_s$ 是 U_1 的一组基;
(B)$\mathbf{w}_1, \cdots, \mathbf{w}_m, \mathbf{v}_1, \cdots, \mathbf{v}_t$ 是 U_2 的一组基。
下面证明 $\mathbf{w}_1, \cdots, \mathbf{w}_m, \mathbf{u}_1, \cdots, \mathbf{u}_s, \mathbf{v}_1, \cdots, \mathbf{v}_t$ 是线性无关的。
如果存在 $\alpha_1, \cdots, \alpha_m, \lambda_1, \cdots, \lambda_s, \mu_1, \cdots, \mu_t \in \mathbb{R}$ 满足

$$\sum_{i=1}^{m} \alpha_i \mathbf{w}_i + \sum_{i=1}^{s} \lambda_i \mathbf{u}_i + \sum_{i=1}^{t} \mu_i \mathbf{v}_i = \mathbf{0}, \tag{$*$}$$

则令 $\mathbf{w} = \sum_{i=1}^m \alpha_i \mathbf{w}_i$，$\mathbf{u} = \sum_{i=1}^s \lambda_i \mathbf{u}_i$，$\mathbf{v} = \sum_{i=1}^t \mu_i \mathbf{v}_i$。显然 $\mathbf{w} \in U_1 \cap U_2$，$\mathbf{u} \in U_1$，$\mathbf{v} \in U_2$。又因为 $\mathbf{w} + \mathbf{u} = -\mathbf{v}$，故也有 $\mathbf{v} \in U_1$，于是 $\mathbf{v} \in U_1 \cap U_2$。则存在 $\beta_1, \cdots, \beta_m \in \mathbb{R}$ 使得 $\mathbf{v} = \beta_1 \mathbf{w}_1 + \cdots + \beta_m \mathbf{w}_m$。即

$$(\alpha_1 + \beta_1)\mathbf{w}_1 + \cdots + (\alpha_m + \beta_m)\mathbf{w}_m + \lambda_1 \mathbf{u}_1 + \cdots + \lambda_s \mathbf{u}_s = \mathbf{0}.$$

由 (A) 知 $\lambda_1 = \cdots = \lambda_s = 0$，代回 $(*)$ 式得到

$$\sum_{i=1}^{m} \alpha_i \mathbf{w}_i + \sum_{i=1}^{t} \mu_i \mathbf{v}_i = \mathbf{0}$$

再由 (B) 可知 $\alpha_1 = \cdots = \alpha_m = \mu_1 = \cdots = \mu_t = 0$。
即 $\mathbf{w}_1, \cdots, \mathbf{w}_m, \mathbf{u}_1, \cdots, \mathbf{u}_s, \mathbf{v}_1, \cdots, \mathbf{v}_t$ 是线性无关的。
再证 $\mathbf{w}_1, \cdots, \mathbf{w}_m, \mathbf{u}_1, \cdots, \mathbf{u}_s, \mathbf{v}_1, \cdots, \mathbf{v}_t$ 的线性包络是 $U_1 + U_2$。
设 $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 \in U_1 + U_2$，其中 $\mathbf{x}_1 \in U_1$，$\mathbf{x}_2 \in U_2$，则有:

$$\mathbf{x}_1 = a_1 \mathbf{w}_1 + \cdots + a_m \mathbf{w}_m + c_1 \mathbf{u}_1 + \cdots + c_s \mathbf{u}_s;$$
$$\mathbf{x}_2 = b_1 \mathbf{w}_1 + \cdots + b_m \mathbf{w}_m + d_1 \mathbf{v}_1 + \cdots + d_t \mathbf{v}_t.$$

于是

$$\mathbf{x} = (a_1+b_1)\mathbf{w}_1+\cdots+(a_m+b_m)\mathbf{w}_m+c_1\mathbf{u}_1+\cdots+c_s\mathbf{u}_s+d_1\mathbf{v}_1+\cdots+d_t\mathbf{v}_t \in< \mathbf{w}_1,\cdots,\mathbf{w}_m,\mathbf{u}_1,\cdots,\mathbf{u}_s,\mathbf{v}_1,\cdots,\mathbf{v}_t >$$

即 $\mathbf{w}_1,\cdots,\mathbf{w}_m,\mathbf{u}_1,\cdots,\mathbf{u}_s,\mathbf{v}_1,\cdots,\mathbf{v}_t$ 是 U_1+U_2 的基。
因此，$\dim(U_1+U_2) = m+s+t = (m+s)+(m+t)-m = \dim U_1 + \dim U_2 - dim(U_1\cap U_2)$。特别的。当 $U_1\cap U_2 = \{\mathbf{0}\}$ 时，令 $m=0$ 即可得到结论。这样我们就完成了证明。 □

推论 2.1.5 设 $U_1, U_2 \subset \mathbb{R}^n$ 是子空间，则 U_1+U_2 是直和 $\Longleftrightarrow \dim(U_1+U_2) = \dim U_1 + \dim U_2$。

证明留作练习。

注 2.1.5 设 $A\in\mathbb{R}^{m\times n}$，我们有方程组:

$$x_1\mathbf{A}^{(1)}+\cdots+x_n\mathbf{A}^{(n)} = \mathbf{0} \tag{H}$$

$$x_1\mathbf{A}^{(1)}+\cdots+x_n\mathbf{A}^{(n)} = \mathbf{b} \tag{L}$$

则 (H) 有非平凡解 $\Longleftrightarrow \mathbf{A}^{(1)},\cdots,\mathbf{A}^n$ 线性相关；(L) 有解 (相容)$\Longleftrightarrow \mathbf{b}\in< \mathbf{A}^{(1)},\cdots,\mathbf{A}^n >$。

命题 2.1.8 设 $V\subset\mathbb{R}^n$ 是非零子空间，$\mathbf{v}_1,\cdots,\mathbf{v}_k$ 是 V 的一组生成元，则 $\mathbf{v}_1,\cdots,\mathbf{v}_k$ 中的任何极大线性无关组是 V 的一组基。

例 2.1.11 求 $V =< (1,2,1)^t, (1,0,1)^t, (2,2,2)^t >$ 的一组基。

解: 将三个向量按顺序命名为 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$。设 $\alpha_1(1,2,1)^t+\alpha_2(1,0,1)^t = \mathbf{0}$，解得 $\alpha_1=\alpha_2=0$。于是 $\mathbf{v}_1,\mathbf{v}_2$ 线性无关。又 $\mathbf{v}_3 = \mathbf{v}_1+\mathbf{v}_2$，故 $\{\mathbf{v}_1,\mathbf{v}_2\}$ 是 $< (1,2,1)^t, (1,0,1)^t, (2,2,2)^t >$ 的极大线性无关组，因此是 V 的一组基。

2.2 矩阵的秩

这一节我们讨论矩阵的第一个重要的不变量：矩阵的秩。

2.2.1 秩定理

设 $A \in \mathbb{R}^{m\times n}$，我们称矩阵 A 的行向量生成的子空间 $< \mathbf{A}_1, \cdots, \mathbf{A}_m >$ 为 A 的行空间，记作 $V_r(A)$；称矩阵 A 的列向量生成的子空间 $< \mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(n)} >$ 为 A 的列空间，记作 $V_c(A)$。

定义 2.2.1 称行空间的维数 $\dim V_r(A)$ 为矩阵 A 的行秩；列空间的维数 $\dim V_c(A)$ 为 A 的列秩。

本小节的主要结果是下面的秩定理。

定理 2.2.1 设 $A \in \mathbb{R}^{m\times n}$，则 $\dim V_r(A) = \dim V_c(A)$。即矩阵的行秩等于列秩。

下面我们用初等行列变换的工具证明这一结论。

引理 2.2.1 设 B 是 A 通过有限次 (I)、(II) 类初等行变换得到的矩阵，则 $V_r(B) = V_r(A)$。

证明: 设 B 是 A 通过 (I) 类初等行变换得到的矩阵，则

$$\begin{aligned} V_r(A) &=< \mathbf{A}_1, \cdots, \mathbf{A}_i, \cdots, \mathbf{A}_j, \cdots, \mathbf{A}_m > \\ &=< \mathbf{A}_1, \cdots, \mathbf{A}_j, \cdots, \mathbf{A}_i, \cdots, \mathbf{A}_m > \\ &= V_r(B) \end{aligned}$$

设 B 是 A 通过 (II) 类初等变换得到的矩阵，则

$$V_r(B) =< \mathbf{A}_1, \cdots, \mathbf{A}_i, \cdots, \mathbf{A}_j + \lambda \mathbf{A}_i, \cdots, \mathbf{A}_m >$$

因为 $i \neq j$，故 $\mathbf{A}_j = (\mathbf{A}_j + \lambda\mathbf{A}_i) - \lambda\mathbf{A}_i$ ，即 $\mathbf{A}_j \in V_r(B)$，所以 $V_r(A) \subset V_r(B)$。
反之，$\mathbf{A}_j + \lambda\mathbf{A}_i \in V_r(A)$，即 $V_r(B) \subset V_r(A)$。
由于一步变换下行空间不变，故经过有限步后行空间仍不变。综上，$V_r(B) = V_r(A)$。 □

我们将初等行变换中对行的操作改成相应的对列的操作，则可以定义三类初等列变换。与上面的证明过程类似，我们可以得到：

引理 2.2.2 设 B 是 A 通过有限次 (I)、(II) 类初等列变换得到的矩阵，则 $V_c(B) = V_c(A)$。

引理 2.2.3 设 B 由 A 经过有限步 (I)、(II) 类初等行变换得到，则 $\dim V_c(B) = \dim V_c(A)$。

证明: 由引理 2.2.2，不妨设 $\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(d)}$ 是 $\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(d)}, \cdots, \mathbf{A}^{(n)}$ 的一个极大线性无关组，下面我们证明 $\mathbf{B}^{(1)}, \cdots, \mathbf{B}^{(d)}$ 一定是 $\mathbf{B}^{(1)}, \cdots, \mathbf{B}^{(d)}, \cdots, \mathbf{B}^{(n)}$ 的一个极大线性无关组。
设 $(H_A), (H_B)$ 分别是以 A, B 为系数矩阵的齐次线性方程组，由引理 1.2.1 知 $(H_A), (H_B)$ 等价。设有 $\alpha_1, \cdots, \alpha_d \in \mathbb{R}$ 使得

$$\alpha_1\mathbf{B}^{(1)} + \cdots + \alpha_d\mathbf{B}^{(d)} = \mathbf{0},$$

则

$$\alpha_1\mathbf{B}^{(1)} + \cdots + \alpha_d\mathbf{B}^{(d)} + 0 \cdot \mathbf{B}^{(d+1)} + \cdots + 0 \cdot \mathbf{B}^{(n)} = \mathbf{0}$$

即 $(\alpha_1,\cdots,\alpha_d,0,\cdots,0)^t$ 是 (H_B) 的解，从而是 (H_A) 的解。于是

$$\alpha_1\mathbf{A}^{(1)}+\cdots+\alpha_d\mathbf{A}^{(d)}=\mathbf{0},$$

再由 $\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(d)}$ 线性无关知 $\alpha_1=\cdots=\alpha_d=0$，即 $\mathbf{B}^{(1)},\cdots,\mathbf{B}^{(d)}$ 线性无关。
下证极大性。设 $k\in\{d+1,\cdots,n\}$，由于 $\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(d)}$ 是极大线性无关组，故 $\exists\beta_1,\cdots,\beta_d\in\mathbb{R}$ 使得

$$\beta_1\mathbf{A}^{(1)}+\cdots+\beta_d\mathbf{A}^{(d)}-\mathbf{A}^{(k)}=\mathbf{0}.$$

即 $(\beta_1,\cdots,\beta_d,\cdots,0,\cdots,0,-1,0,\cdots,0)^t$ 是 (H_A) 的解，所以它也是 (H_B) 的解，即

$$\beta_1\mathbf{B}^{(1)}+\cdots+\beta_d\mathbf{B}^{(d)}-\mathbf{B}^{(k)}=\mathbf{0}.$$

所以 $\mathbf{B}^{(k)}\in<\mathbf{B}^{(1)},\cdots,\mathbf{B}^{(d)}>$。
故由命题 2.1.8 可得 $\dim V_c(B)=\dim V_c(A)=d$。 □

有了以上的准备工作，我们现在可以来证明定理 2.2.1 了。

证明: 对 A 作以下初等行列变换:

$$A\xrightarrow{(I),(II)\text{类初等行变换}}B=\begin{pmatrix}0&\cdots&0&\square_1&\cdots&*&\cdots&*&\cdots&*\\0&\cdots&0&0&\cdots&\square_2&\cdots&*&\cdots&*\\\vdots&&\vdots&\vdots&&\vdots&\ddots&\vdots&&\vdots\\0&\cdots&0&0&\cdots&0&\cdots&\square_k&\cdots&*\\0&\cdots&0&0&\cdots&0&\cdots&0&\cdots&0\\\vdots&&\vdots&\vdots&&\vdots&&\vdots&&\vdots\\0&\cdots&0&0&\cdots&0&\cdots&0&\cdots&0\end{pmatrix}\begin{matrix}\left.\begin{matrix}\\\\\\\\\end{matrix}\right\}k\text{行}\\\\\\\\\end{matrix}$$

$$\xrightarrow{(I)\text{类初等列变换}}C=\begin{pmatrix}\square_1&*&*&\cdots&\cdots&*\\&\square_2&*&\cdots&\cdots&*\\&&\ddots&&&\vdots\\&&&\square_k&\cdots&*\\0&0&\cdots&0&\cdots&0\\&&\cdots&&&\\0&0&\cdots&0&\cdots&0\end{pmatrix}\begin{matrix}\left.\begin{matrix}\\\\\\\\\end{matrix}\right\}k\text{行}\\\\\\\\\end{matrix}$$
$\underbrace{\qquad\qquad}_{k\text{列}}$

$$\xrightarrow{(II)\text{类初等列变换}}D=\begin{pmatrix}\square_1&0&0&\cdots&\cdots&0\\&\square_2&0&\cdots&\cdots&0\\&&\ddots&&&\vdots\\&&&\square_k&\cdots&0\\0&0&\cdots&0&\cdots&0\\&&\cdots&&&\\0&0&\cdots&0&\cdots&0\end{pmatrix}\begin{matrix}\left.\begin{matrix}\\\\\\\\\end{matrix}\right\}k\text{行}\\\\\\\\\end{matrix}$$
$\underbrace{\qquad\qquad}_{k\text{列}}$

显然 $k=\dim V_c(D)$，于是由引理 2.2.2，$k=\dim V_c(C)=\dim V_c(B)$，再由引理 2.2.3，$k=\dim V_c(A)$。
另一方面，由于 B 是行阶梯型矩阵，容易证明 $\dim V_r(B)=k$，而由引理 2.2.1 即有 $\dim V_r(A)=$

$\dim V_r(B) = k$。
综上，$\dim V_r(A) = \dim V_c(A)$。 □

于是我们就可以定义矩阵的秩如下：

定义 2.2.2 (**矩阵的秩**) 设 $A \in \mathbb{R}^{m\times n}$，则称 $\dim V_r(A)$ 为 A 的秩，记作 $\mathrm{rank}(A)$。

以上的证明过程也告诉了我们一种求矩阵的秩的方法：用初等变换将矩阵化成阶梯型。我们看下面的例子。

例 2.2.1 设 $A = \begin{pmatrix} 1 & 0 & 4 & 5 \\ 2 & 1 & -1 & 3 \\ 4 & 1 & 7 & 13 \end{pmatrix}$，求 $\mathrm{rank}(A)$。

解:

$$A \xrightarrow{r_2-2r_1} \begin{pmatrix} 1 & 0 & 4 & 5 \\ 0 & 1 & -9 & -7 \\ 4 & 1 & 7 & 13 \end{pmatrix} \xrightarrow{r_3-4r_1} \begin{pmatrix} 1 & 0 & 4 & 5 \\ 0 & 1 & -9 & -7 \\ 0 & 1 & -9 & -7 \end{pmatrix} \xrightarrow{r_3-r_2} \begin{pmatrix} 1 & 0 & 4 & 5 \\ 0 & 1 & -9 & -7 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

即 $\mathrm{rank}(A) = 2$。

例 2.2.2 设 $A \in \mathbb{R}^{m\times n}$，求证 $\mathrm{rank}(A) \leqslant \min(m, n)$。

证明: 由于 $V_r(A) \subset \mathbb{R}^{1\times n}$，故 $\dim V_r(A) \leqslant n$。同理 $V_c(A) \subset \mathbb{R}^{m\times 1} \Longrightarrow \dim V_c(A) \leqslant m$。
于是 $\mathrm{rank}(A) \leqslant \min(m, n)$。 □

最后，设 $A \in \mathbb{R}^{m\times n}$，我们称 $\mathrm{rank}(A) = m$ 的矩阵 A 为行满秩的；称 $\mathrm{rank}(A) = n$ 的矩阵 A 为列满秩的。若 $A \in \mathbb{R}^{n\times n}$ 且 $\mathrm{rank}(A) = n$，则称 A 是满秩的。

2.2.2 秩的应用

定理 2.2.2 设 $A \in \mathbb{R}^{m\times n}$，$(H_A)$ 是以 A 为系数矩阵的齐次线性方程组。则

$$(H_A)\ \text{有非平凡解} \Longleftrightarrow \mathrm{rank}(A) < n.$$

证明: 设 (H_A) 为 $x_1\mathbf{A}^{(1)} + \cdots + x_n\mathbf{A}^{(n)} = \mathbf{0}$。
$(\Longrightarrow)$ 设 $(x_1, \cdots, x_n)^t = (\alpha_1, \cdots, \alpha_n)^t$ 是 (H_A) 的非平凡解，则 $\alpha_1\mathbf{A}^{(1)}+\cdots+\alpha_n\mathbf{A}^{(n)} = \mathbf{0}$，即 $\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(n)}$ 线性相关，于是 $\dim V_c(A) < n$，即 $\mathrm{rank}(A) = \dim V_c(A) < n$。
$(\Longleftarrow)$ 由 $\dim V_c(A) < n$ 知 $\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(n)}$ 线性相关，即存在不全为 0 的 $\beta_1, \cdots, \beta_n$ 使得 $\beta_1\mathbf{A}^{(1)} + \cdots + \beta_n\mathbf{A}^{(n)} = \mathbf{0}$，则 $(\beta_1, \cdots, \beta_n)$ 是 (H_A) 的非平凡解。 □

下面的定理是定理 1.2.2 的向量表述。

定理 2.2.3 设 $A \in \mathbb{R}^{m\times n}$，$\mathbf{b} \in \mathbb{R}^m$，(L) 是以 $(A|\mathbf{b})$ 为增广矩阵的线性方程组，则 (L) 相容 $\Longleftrightarrow$ $\mathrm{rank}(A) = \mathrm{rank}(A|\mathbf{b})$。

证明留作练习。

例 2.2.3 判断方程组

$$\begin{cases} x+y=1 \\ 2x-y=2 \\ 5x+2y=5 \end{cases} \tag{L}$$

是否相容。

解: 对 (L) 的增广矩阵作如下初等行变换:

$$B=\left(\begin{array}{cc:c} 1 & 1 & 1 \\ 2 & -1 & 2 \\ 5 & 2 & 5 \end{array}\right) \xrightarrow{r_2-2r_1} \left(\begin{array}{cc:c} 1 & 1 & 1 \\ 0 & -3 & 0 \\ 5 & 2 & 5 \end{array}\right) \xrightarrow{r_3-5r_1} \left(\begin{array}{cc:c} 1 & 1 & 1 \\ 0 & -3 & 0 \\ 0 & -3 & 0 \end{array}\right) \xrightarrow{r_3-r_2} \left(\begin{array}{cc:c} 1 & 1 & 1 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

(其中前两列构成 A)

即 $\operatorname{rank}(A)=\operatorname{rank}(B)=2$，于是 (L) 相容。

在本节的最后，我们考虑齐次线性方程组的解集。例 2.1.9 已经告诉我们该解集是 $\mathbb{R}^n$ 的子空间。那么，这个子空间的维数与方程组的系数矩阵有什么关系呢？这就是下面的对偶定理。

定义 2.2.3 设 $A\in\mathbb{R}^{m\times n}$，以 A 为系数矩阵的齐次线性方程组的解空间记为 V_A。

定理 2.2.4 (**对偶定理** (**方程版**)) 设 $A\in\mathbb{R}^{m\times n}$，则 $\operatorname{rank}(A)+\dim(V_A)=n$。

证明: 不妨设 $r=\operatorname{rank}(A)$, $\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(r)}$ 是 $V_c(A)$ 的一组基, 则 $\forall j\in\{r+1,\cdots,n\}$, $\exists\alpha_{1j},\cdots,\alpha_{rj}\in\mathbb{R}$ 使得

$$\alpha_{1j}\mathbf{A}^{(1)}+\cdots+\alpha_{rj}\mathbf{A}^{(r)}-\mathbf{A}^{(j)}=\mathbf{0}.$$

即 $\mathbf{v}_j=(\alpha_{1j},\cdots,\alpha_{rj},0,\cdots,0,-1,0,\cdots,0)^t$ (第 j 个位置是-1) 是 A 对应的齐次线性方程组的一个解，又注意到 $\mathbf{v}_{r+1},\cdots,\mathbf{v}_n\in V_A$ 是线性无关的 (注意“-1”的位置)，所以 $\dim V_A\geqslant n-r$。
下面任取 $\mathbf{u}=(\beta_1,\cdots,\beta_n)^t\in V_A$，我们的目标是证明 $\mathbf{u}$ 是 $\mathbf{v}_{r+1},\cdots,\mathbf{v}_n$ 的线性组合，从而说明 $\mathbf{v}_{r+1},\cdots,\mathbf{v}_n$ 是 V_A 的一组基。 由于 V_A 是子空间，故 $\mathbf{u}+\beta_{r+1}\mathbf{v}_{r+1}+\cdots+\beta_n\mathbf{v}_n\in V_A$，即形如 $(\lambda_1,\cdots,\lambda_r,0\cdots,0)^t$ (后 $n-r$ 个位置全为 0) 的向量是以 A 为系数矩阵的齐次线性方程组的解，具体写出来就是:

$$\lambda_1\mathbf{A}^{(1)}+\cdots+\lambda_r\mathbf{A}^{(r)}=\mathbf{0}.$$

再由 $\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(r)}$ 线性无关可知 $\lambda_1=\cdots=\lambda_r=0$，即

$$\mathbf{u}=-(\beta_{r+1}\mathbf{v}_{r+1}+\cdots+\beta_n\mathbf{v}_n).$$

于是我们证明了 $\mathbf{v}_{r+1},\cdots,\mathbf{v}_n$ 是 V_A 的一组基，即 $\dim V_A=n-r$。 □

例 2.2.4 设 $A=\begin{pmatrix} 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & 1 \\ 5 & -1 & -2 & 3 \end{pmatrix}$，求 V_A 的一组基。

解: 作如下初等行变换:

$$A\xrightarrow[r_3-5r_1]{r_2-r_1}\begin{pmatrix} 1 & 1 & -1 & 0 \\ 0 & -2 & 1 & 1 \\ 0 & -6 & -3 & 3 \end{pmatrix}\xrightarrow{r_3-3r_2}\begin{pmatrix} 1 & 1 & -1 & 0 \\ 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

于是 $\mathrm{rank}(A)=2$。由对偶定理，$\dim V_A = 4-\mathrm{rank}(A)=4-2=2$。下面求解此齐次方程组。由于方程组等价于

$$\begin{cases} \quad x_1+x_2-x_3=0 \\ -2x_2+x_3+x_4=0 \end{cases}$$

即

$$\begin{cases} x_1=-x_2+x_3 \\ x_4=2x_2-x_3 \end{cases}$$

分别取 $x_2=1$，$x_3=0$ 和 $x_2=0$，$x_3=1$ 得 $V_A =< (-1,1,0,2)^t$，$(1,0,1,-1)^t >$。

2.3 线性映射

用线性映射的观点来看待矩阵是十分重要和基本的。

2.3.1 定义和例子

这一小节我们的主要任务是定义线性映射并讨论它的一些基本的性质。

定义 2.3.1 设 $\varphi:\mathbb{R}^n\to\mathbb{R}^m$ 是映射，如果对 $\forall\mathbf{x},\mathbf{y}\in\mathbb{R}^n$ 及 $\alpha\in\mathbb{R}$，有 $\varphi(\mathbf{x}+\mathbf{y})=\varphi(\mathbf{x})+\varphi(\mathbf{y})$ 及 $\varphi(\alpha\mathbf{x})=\alpha\varphi(\mathbf{x})$ 成立，则称 φ 是线性映射。

注 2.3.1 (1) 对于线性映射 φ，一定有 $\varphi(\mathbf{0})=\mathbf{0}$。这是因为 $\varphi(\mathbf{0}+\mathbf{0})=2\varphi(\mathbf{0})=\varphi(\mathbf{0})$。
(2)$\varphi:\mathbb{R}^n\to\mathbb{R}^m$ 是线性映射 $\Longleftrightarrow\forall\mathbf{x},\mathbf{y}\in\mathbb{R}^n$ 及 $\alpha,\beta\in\mathbb{R}$，有 $\varphi(\alpha\mathbf{x}+\beta\mathbf{y})=\alpha\varphi(\mathbf{x})+\beta\varphi(\mathbf{y})$。
其中，($\Longrightarrow$) 方向利用定义即可证明，($\Longleftarrow$) 方向分别取 $\alpha=\beta=1$ 和 $\alpha=1,\ \beta=0$ 即可验证。

下面来看几个例子。

例 2.3.1 (1)$\varphi:\mathbb{R}^m\to\mathbb{R}^m$，$\mathbf{x}\mapsto\mathbf{x}$ 恒同映射是线性的。
(2)$\varphi:\mathbb{R}^n\to\mathbb{R}^m$，$\mathbf{x}\mapsto\mathbf{0}$ 零映射是线性的。
(3)$\varphi:\mathbb{R}^n\to\mathbb{R}^n$，$\mathbf{x}\mapsto\mathbf{x}+\mathbf{v}$，其中 $\mathbf{v}\in\mathbb{R}^n\setminus\{\mathbf{0}\}$，非平凡的平移映射**不是**线性的。(不满足 $\varphi(\mathbf{0})=\mathbf{0}$.)

命题 2.3.1 设 $\varphi:\mathbb{R}^n\to\mathbb{R}^m$ 是线性映射，$\mathbf{v}_1,\cdots,\mathbf{v}_k\in\mathbb{R}^n$，$\alpha_1,\cdots,\alpha_k\in\mathbb{R}$，则

(i) $\varphi(\alpha_1\mathbf{v}_1+\cdots+\alpha_k\mathbf{v}_k)=\alpha_1\varphi(\mathbf{v}_1)+\cdots+\alpha_k\varphi(\mathbf{v}_k)$;

(ii) 如果 $\mathbf{v}_1,\cdots,\mathbf{v}_k$ 线性相关，则 $\varphi(\mathbf{v}_1),\cdots,\varphi(\mathbf{v}_k)$ 也线性相关。

(iii) 如果 $\varphi(\mathbf{v}_1),\cdots,\varphi(\mathbf{v}_k)$ 线性无关，则 $\mathbf{v}_1,\cdots,\mathbf{v}_k$ 也线性无关。

证明: (i) 对 k 作归纳。$k=1$ 时即线性映射的定义。
设 $k-1$ 时结论成立，则

$$\begin{aligned}&\varphi(\alpha_1\mathbf{v}_1+\cdots+\alpha_{k-1}\mathbf{v}_{k-1}+\alpha_k\mathbf{v}_k)\\=&\varphi(\alpha_1\mathbf{v}_1+\cdots+\alpha_{k-1}\mathbf{v}_{k-1})+\alpha_k\varphi(\mathbf{v}_k)\\=&\alpha_1\varphi(\mathbf{v}_1)+\cdots+\alpha_{k-1}\varphi(\mathbf{v}_{k-1})+\alpha_k\varphi(\mathbf{v}_k)\ \text{(归纳假设)}\end{aligned}$$

(ii) 设 $\beta_1\mathbf{v}_1+\cdots+\beta_k\mathbf{v}_k=\mathbf{0}$，其中 β_i，$i=1,\cdots,k$ 不全为 0，则

$$\mathbf{0}=\varphi(\mathbf{0})=\varphi(\beta_1\mathbf{v}_1+\cdots+\beta_k\mathbf{v}_k)=\beta_1\varphi(\mathbf{v}_1)+\cdots+\beta_k\varphi(\mathbf{v}_k)$$

即 $\varphi(\mathbf{v}_1),\cdots,\varphi(\mathbf{v}_k)$ 线性相关。

(iii) 即 (ii) 的逆否命题。 □

例 2.3.2 设有映射 (函数)$f:\mathbb{R}^n\to\mathbb{R}$，则 f 是线性映射 $\Longleftrightarrow\exists\alpha_1,\cdots,\alpha_n\in\mathbb{R}$，使得对 $\forall\mathbf{x}=(x_1,\cdots,x_n)^t\in\mathbb{R}^n$，有 $f(\mathbf{x})=\alpha_1x_1+\cdots+\alpha_nx_n$。

证明: ($\Longrightarrow$) 设 $\alpha_j=f(\mathbf{e}^{(j)})$，$j=1,\cdots,n$。则由 $\mathbf{x}=x_1\mathbf{e}^{(1)}+\cdots+x_n\mathbf{e}^{(n)}$ 可得

$$f(\mathbf{x})=x_1f(\mathbf{e}^{(1)})+\cdots+x_nf(\mathbf{e}^{(n)})=\alpha_1x_1+\cdots+\alpha_nx_n$$

即得结论。

($\Longleftarrow$) 用定义验证。设 $\mathbf{y} = (y_1, \cdots, y_n)^t$，$\lambda \in \mathbb{R}$，则

$$f(\mathbf{x}+\mathbf{y}) = \alpha_1(x_1+y_1) + \cdots + \alpha_n(x_n+y_n) = f(\mathbf{x}) + f(\mathbf{y})$$
$$f(\lambda\mathbf{x}) = \alpha_1(\lambda x_1) + \alpha_n(\lambda x_n) = \lambda(\alpha_1 x_1 + \cdots + \alpha_n x_n) = \lambda f(\mathbf{x})$$

即得到 f 是线性映射。 □

例 2.3.3 $f: \mathbb{R} \to \mathbb{R},\ x \mapsto x^2$ 不是线性映射。

定理 2.3.1 (**线性映射基本定理**) 设 $\mathbf{b}_1, \cdots, \mathbf{b}_n$ 是 $\mathbb{R}^n$ 的一组基，$\mathbf{v}_1, \cdots, \mathbf{v}_n$ 是 $\mathbb{R}^m$ 中任意 n 个向量，则存在唯一的线性映射 $\varphi: \mathbb{R}^n \to \mathbb{R}^m$ 使得 $\varphi(\mathbf{b}_i) = \mathbf{v}_i$，$i = 1, 2, \cdots, n$。

证明: (分三步，一是构造映射 φ 并说明其良定义，二是证明 φ 的线性性，三是证明唯一性。)

(1) 由基的定义，对 $\forall \mathbf{x} \in \mathbb{R}^n$，存在唯一一组 $\alpha_1, \cdots, \alpha_n \in \mathbb{R}$ 使得 $\mathbf{x} = \alpha_1\mathbf{b}_1 + \cdots + \alpha_n\mathbf{b}_n$。于是我们可以定义

$$\begin{aligned}\varphi: &\mathbb{R}^n \longrightarrow \mathbb{R}^m \\ &\mathbf{x} \longmapsto \sum_{i=1}^{n} \alpha_i \mathbf{v}_i.\end{aligned}$$

注意 $\alpha_1, \cdots, \alpha_n$ 的存在唯一性保证了 φ 是良定义的，且 $\varphi(\mathbf{b}_i) = \mathbf{v}_i$，$i = 1, 2, \cdots, n$ 成立。

(2) 设另有 $\mathbf{y} = \sum_{i=1}^n \beta_i \mathbf{b}_i \in \mathbb{R}^n$ 及 $\lambda, \mu \in \mathbb{R}$，则

$$\begin{aligned}\varphi(\lambda\mathbf{x} + \mu\mathbf{y}) &= \varphi(\lambda \sum_{i=1}^{n} \alpha_i \mathbf{b}_i + \mu \sum_{i=1}^{n} \beta_i \mathbf{b}_i) \\ &= \varphi(\sum_{i=1}^{n} (\lambda\alpha_i + \mu\beta_i)\mathbf{b}_i) \\ &= \sum_{i=1}^{n} (\lambda\alpha_i + \mu\beta_i)\mathbf{v}_i \\ &= \lambda \sum_{i=1}^{n} \alpha_i \mathbf{v}_i + \mu \sum_{i=1}^{n} \beta_i \mathbf{v}_i \\ &= \lambda\varphi(\mathbf{x}) + \mu\varphi(\mathbf{y})\end{aligned}$$

即 φ 的线性性得证。

(3) 设另有 $\psi: \mathbb{R}^n \to \mathbb{R}^m$ 是线性映射并且也满足 $\psi(\mathbf{b}_i) = \mathbf{v}_i$，则对 $\forall \mathbf{x} \in \mathbb{R}^n$，有

$$\psi(\mathbf{x}) = \psi(\sum_{i=1}^{n} \alpha_i \mathbf{b}_i) = \sum_{i=1}^{n} \alpha_i \psi(\mathbf{b}_i) = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \varphi(\mathbf{x}).$$

即 $\psi = \varphi$。于是唯一性成立。 □

例 2.3.4 下面我们考虑两种特殊的线性映射：嵌入与投影。设 $\mathbb{R}^n$ 的标准基为 $\mathbf{e}^{(1)}, \cdots, \mathbf{e}^{(n)}$，$\mathbb{R}^{(m)}$ 的标准基为 $\mathcal{E}^{(1)}, \cdots, \mathcal{E}^{(m)}$。则有：

情形 1：$n \leqslant m$(嵌入)，作线性映射 $\varphi: \mathbb{R}^n \to \mathbb{R}^m$ 且满足：$\varphi(\mathbf{e}^{(j)}) = \mathcal{E}^{(j)}$，$j = 1, \cdots, n$。于是有：

$$\begin{aligned}\varphi((x_1, \cdots, x_n)^t) &= \varphi(x_1\mathbf{e}^{(1)} + \cdots + x_n\mathbf{e}^{(n)}) \\ &= x_1\varphi(\mathbf{e}^{(1)}) + \cdots + x_n\varphi(\mathbf{e}^{(n)}) \\ &= x_1\mathcal{E}^{(1)} + \cdots + \mathcal{E}^{(n)} \\ &= (x_1, \cdots, x_n, \underbrace{0, \cdots, 0}_{n-m\text{个}})^t.\end{aligned}$$

情形 2：$n \geqslant m$(投影)。类似上面的推导可以得到：

$$\varphi((x_1, \cdots, x_n)^t) = ((x_1, \cdots, x_m)^t).$$

例 2.3.5 设线性映射 $\varphi: \mathbb{R}^2 \to \mathbb{R}^2$ 满足

$$\varphi(\mathbf{e}^{(1)}) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \quad \varphi(\mathbf{e}^{(2)}) = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}.$$

那么，对 $\forall \mathbf{x} \in \mathbb{R}^2$，$\mathbf{x} = (x_1, x_2)^t$，有 $\varphi(\mathbf{x}) = (x_1\cos\theta - x_2\sin\theta,\ x_1\sin\theta + x_2\cos\theta)^t$。
此即 $\mathbb{R}^2$ 上的旋转变换。

2.3.2 线性映射下的子空间

这一小节我们讨论子空间在线性映射下的“变与不变”。

命题 2.3.2 设 $\varphi: \mathbb{R}^n \to \mathbb{R}^m$ 是线性映射，则
(i) 如果 $U \subset \mathbb{R}^n$ 是子空间，则 $\varphi(U) \subset \mathbb{R}^m$ 也是子空间，并且 $\dim(U) \geqslant \dim(\varphi(U))$；
(ii) 如果 $V \subset \mathbb{R}^m$ 是子空间，则 $\varphi^{-1}(V)$ 是 $\mathbb{R}^n$ 中的子空间。

证明: (i) 设 $\mathbf{x}, \mathbf{y} \in \varphi(U)$，则存在 $\mathbf{u}, \mathbf{v} \in U$ 使得 $\varphi(\mathbf{u}) = \mathbf{x}$，$\varphi(\mathbf{v}) = \mathbf{y}$。于是对 $\forall \alpha, \beta \in \mathbb{R}$，有

$$\alpha\mathbf{x} + \beta\mathbf{y} = \alpha\varphi(\mathbf{u}) + \beta\varphi(\mathbf{v}) = \varphi(\alpha\mathbf{u} + \beta\mathbf{v}).$$

即 $\varphi(U)$ 是子空间。另设 $\mathbf{x}_1, \cdots, \mathbf{x}_d$ 是 $\varphi(U)$ 的一组基，则存在 $\mathbf{u}_1, \cdots, \mathbf{u}_d \in U$ 使得 $\varphi(\mathbf{u}_i) = \mathbf{x}_i$。于是由命题 2.3.1(iii) 可得 $\mathbf{u}_1, \cdots, \mathbf{u}_d$ 线性无关，所以 $\dim U \geqslant d = \dim(\varphi(U))$。
(ii) 对 $\forall \mathbf{x}, \mathbf{y} \in \varphi^{-1}(V)$，即 $\varphi(\mathbf{x}), \varphi(\mathbf{y}) \in V$，则任取 $\alpha, \beta \in \mathbb{R}$，有 $\alpha\varphi(\mathbf{x}) + \beta\varphi(\mathbf{y}) \in V$，即

$$\varphi(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\varphi(\mathbf{x}) + \beta\varphi(\mathbf{y}) \in V.$$

即 $\alpha\mathbf{x} + \beta\mathbf{y} \in \varphi^{-1}(V)$。即 $\varphi^{-1}(V)$ 是 $\mathbb{R}^n$ 中的子空间。 □

定义 2.3.2 设 $\varphi: \mathbb{R}^n \to \mathbb{R}^m$ 是线性映射，称 $\varphi^{-1}(\mathbf{0})$ 为 φ 的核空间，记作 $\ker(\varphi)$；称 $\varphi(\mathbb{R}^n)$ 为 φ 的像空间，记作 $\mathrm{im}(\varphi)$。由上面的命题，$\ker(\varphi) \subset \mathbb{R}^n$，$\mathrm{im}(\varphi) \subset \mathbb{R}^m$ 都是子空间。

下面的命题十分常用。

命题 2.3.3 设 $\varphi: \mathbb{R}^n \to \mathbb{R}^m$ 是线性映射，则 φ 是单射 $\Longleftrightarrow \ker(\varphi) = \{\mathbf{0}\}$。

证明: $(\Longrightarrow)$ 由于 $\varphi(\mathbf{0}) = \mathbf{0}$，而 φ 是单射，故 $\ker(\varphi) = \{\mathbf{0}\}$。
$(\Longleftarrow)$ 设 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ 且 $\varphi(\mathbf{x}) = \varphi(\mathbf{y})$，则 $\varphi(\mathbf{x} - \mathbf{y}) = \mathbf{0}$，即 $\mathbf{x} - \mathbf{y} \in \ker(\varphi)$，故由 $\ker(\varphi) = \{\mathbf{0}\}$ 立刻得到 $\mathbf{x} - \mathbf{y} = \mathbf{0}$，即 $\mathbf{x} = \mathbf{y}$。

下面我们用线性映射的观点重写对偶定理。

定理 2.3.2 (**对偶定理** (**映射版**)) 设 $\varphi: \mathbb{R}^n \to \mathbb{R}^m$ 是线性映射，则 $\dim(\ker(\varphi)) + \dim(\mathrm{im}(\varphi)) = n$。

证明: 设 $\mathbf{u}_1,\cdots,\mathbf{u}_d$ 是 $\ker(\varphi)$ 的一组基，则由基扩充定理，$\mathbb{R}^n$ 有一组基为:

$$\mathbf{u}_1,\cdots,\mathbf{u}_d,\mathbf{u}_{d+1},\cdots,\mathbf{u}_n.$$

下面我们证明如下的断言：$\varphi(\mathbf{u}_{d+1}),\cdots,\varphi(\mathbf{u}_n)$ 是 $\mathrm{im}(\varphi)$ 的一组基。
(1)$\varphi(\mathbf{u}_{d+1}),\cdots,\varphi(\mathbf{u}_n)$ 线性无关。下面是验证过程。
设 $\beta_{d+1}\varphi(\mathbf{u}_{d+1})+\cdots+\beta_n\varphi(\mathbf{u}_n)=\mathbf{0}$，即 $\varphi(\beta_{d+1}\mathbf{u}_{d+1}+\cdots+\beta_n\mathbf{u}_n)=\mathbf{0}$。于是 $\beta_{d+1}\mathbf{u}_{d+1}+\cdots+\beta_n\mathbf{u}_n\in\ker(\varphi)$。那么，由 $\ker(\varphi)$ 的基的定义，存在 $\beta_1,\cdots,\beta_d\in\mathbb{R}$ 使得 $\beta_{d+1}\mathbf{u}_{d+1}+\cdots+\beta_n\mathbf{u}_n=\beta_1\mathbf{u}_1+\cdots+\beta_d\mathbf{u}_d$，于是由 $\mathbf{u}_1,\cdots,\mathbf{u}_d,\mathbf{u}_{d+1},\cdots,\mathbf{u}_n$ 线性无关知 $\beta_1=\cdots=\beta_n=0$。特别地，$\beta_{d+1}=\cdots=\beta_n=0$，即 $\varphi(\mathbf{u}_{d+1}),\cdots,\varphi(\mathbf{u}_n)$ 线性无关。
(2) 对 $\forall\mathbf{x}\in\mathrm{im}(\varphi)$，$\exists\mathbf{u}\in\mathbb{R}^n$ 使得 $\varphi(\mathbf{u})=\mathbf{x}$。由 $\mathbb{R}^n$ 的基可得 $\exists!$一组$\alpha_1,\cdots,\alpha_n$ 使得 $\mathbf{u}=\sum_{i=1}^n\alpha_i\mathbf{u}_i$。用 φ 作用在上面的等式两边，得到

$$\varphi(\mathbf{u})=\sum_{i=d+1}^{n}\alpha_i\varphi(\mathbf{u}_i).$$

即 $\mathrm{im}(\varphi)\subset<\varphi(\mathbf{u}_{d+1}),\cdots,\varphi(\mathbf{u}_n)>$。由 (1), (2) 可知断言成立，即 $\dim(\mathrm{im}(\varphi))=n-d$。 □

推论 2.3.1 设 $\varphi:\mathbb{R}^n\to\mathbb{R}^m$ 是线性映射，则
(i)φ 是单射 $\Longleftrightarrow\dim(\mathrm{im}(\varphi))=n$;
(ii) 当 $m=n$ 时 φ 是单射 $\Longleftrightarrow$ φ 是满射。

证明: (i)φ 是单射 $\Longleftrightarrow\dim(\ker(\varphi))=0\Longleftrightarrow\dim(\mathrm{im}(\varphi))=n$;
(ii) 由 (i) 知 φ 是单射 $\Longleftrightarrow$ $\mathrm{im}(\varphi))=\mathbb{R}^n\Longleftrightarrow$ φ 是满射。 □

2.3.3 线性映射在标准基下的矩阵表示

在本小节中，我们设 $\mathbb{R}^n$ 的标准基为 $\mathbf{e}^{(1)},\cdots,\mathbf{e}^{(n)}$，$\mathbb{R}^m$ 的标准基为 $\boldsymbol{\mathcal{E}}^{(1)},\cdots,\boldsymbol{\mathcal{E}}^{(m)}$。

定义 2.3.3 设 $\varphi:\mathbb{R}^n\to\mathbb{R}^m$ 是线性映射，$\varphi(\mathbf{e}^{(j)})=(a_{ij},\cdots,a_{mj})^t$，$j=1,\cdots,n$。我们称矩阵

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\ \vdots&\vdots&&\vdots\\ a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix}_{m\times n}=(\varphi(\mathbf{e}^{(1)}),\cdots,\varphi(\mathbf{e}^{(n)}))$$

为 φ 在标准基下的矩阵。注意，由定理 2.3.1，A 是由 φ 唯一确定的，记为 A_φ。

例 2.3.6 (1)$\varphi:\mathbb{R}^n\to\mathbb{R}^m$，$\mathbf{u}\mapsto\mathbf{0}$ 零映射对应的矩阵是零矩阵；
(2)$\varphi:\mathbb{R}^n\to\mathbb{R}^n$，$\mathbf{u}\mapsto\mathbf{u}$ 恒同映射对应的矩阵是单位矩阵

$$A_\varphi=\begin{pmatrix}1&0&\cdots&0\\0&1&\cdots&0\\ \vdots&\vdots&\ddots&\vdots\\0&0&\cdots&1\end{pmatrix}_{n\times n}=E_n.$$

命题 2.3.4 设 $\varphi:\mathbb{R}^n\to\mathbb{R}^m$ 是线性映射，φ 在标准基下的矩阵为 $A=(a_{ij})_{m\times n}$，则对 $\forall\mathbf{x}=(x_1,\cdots,x_n)\in\mathbb{R}^n$，有

$$\varphi(\mathbf{x})=x_1\mathbf{A}^{(1)}+\cdots+x_n\mathbf{A}^{(n)}=\begin{pmatrix}a_{11}x_1+\cdots+a_{1n}x_n\\ \vdots\\ a_{m1}x_1+\cdots+a_{mn}x_n\end{pmatrix}.$$

证明: 由 $\varphi(\mathbf{e}^{(j)}) = \mathbf{A}^{(j)}$, $j = 1, 2, \cdots, n$ 得

$$\begin{aligned}
\varphi(\mathbf{x}) &= \varphi(x_1\mathbf{e}^{(1)} + \cdots + x_n\mathbf{e}^{(n)}) \\
&= x_1\varphi(\mathbf{e}^{(1)}) + \cdots + x_n\varphi(\mathbf{e}^{(n)}) \\
&= x_1\mathbf{A}^{(1)} + \cdots + x_n\mathbf{A}^{(n)} \\
&= \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}.
\end{aligned}$$

□

注 2.3.2 由上述命题有 $\ker(\varphi) = V_A$ 且 $\mathrm{im}(\varphi) = V_c(A)$。

例 2.3.7 设

$$\begin{aligned}
\varphi : &\mathbb{R}^5 \longrightarrow \mathbb{R}^3 \\
&\begin{pmatrix} x_1 \\ \vdots \\ x_5 \end{pmatrix} \longmapsto \begin{pmatrix} x_1 + x_2 + x_3 + x_4 + x_5 \\ x_1 - x_2 - x_3 - x_4 - x_5 \\ 4x_1 + 2x_2 + 2x_3 + 2x_4 + 2x_5 \end{pmatrix}
\end{aligned}$$

. 则 φ 在标准基下的矩阵为

$$A_\varphi = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 & -1 \\ 4 & 2 & 2 & 2 & 2 \end{pmatrix} \xrightarrow{I,II\text{类初等行变换}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & -2 & -2 & -2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

于是 $\mathrm{rank}(A_\varphi) = 2$, $\dim(\mathrm{im}(\varphi)) = 2$。由对偶定理, $\dim(\ker(\varphi)) = 3$。下面解出 $\ker(\varphi)$。

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 0 \\ \quad\;\; x_2 + x_3 + x_4 + x_5 = 0 \end{cases} \Longrightarrow \begin{cases} x_1 = 0 \\ x_2 + x_3 + x_4 + x_5 = 0 \end{cases}$$

于是 $\ker(\varphi) =< (0, -1, 1, 0, 0)^t, (0, -1, 0, 1, 0)^t, (0, -1, 0, 0, 1)^t >$, $\mathrm{im}(\varphi) =< (1, 1, 4)^t, (1, -1, 2)^t >$。

反过来，我们也可以从矩阵出发定义线性映射。

定义 2.3.4 设 $A \in \mathbb{R}^{m\times n}$, 则我们定义 $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ 是满足 $\varphi(\mathbf{e}^{(j)}) = \mathbf{A}^{(j)}$ 的线性映射，称为矩阵 A 对应的线性映射。由定理 2.3.1，这样的 φ 存在且唯一，我们将其记为 φ_A。显然 φ_A 在标准基下的矩阵是 A。

例 2.3.8 证明对偶定理的方程版与映射版是一致的。

证明: 设 $A \in \mathbb{R}^{m\times n}$, 则 $\ker(\varphi_A) = V_A$, $\mathrm{im}(\varphi_A) = V_c(A)$。于是

$$\dim(\ker(\varphi_A)) + \dim(\mathrm{im}(\varphi_A)) = n \Longleftrightarrow \dim(V_A) + \dim(V_c(A)) = n \Longleftrightarrow \dim(V_A) + \mathrm{rank}(A) = n.$$

□

2.4 矩阵的运算

这一节我们主要的任务是从线性映射的角度定义矩阵的运算，并考查其性质。

由 2.3.3 小节的内容，我们已经有了矩阵和线性映射之间的对应关系。

设 $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ 是线性映射，矩阵

$$A_\varphi = (\varphi(\mathbf{e}^{(1)}), \cdots, \varphi(\mathbf{e}^{(n)}))$$

为 φ 在标准基 $(\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(n)})$ 和 $(\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(m)})$ 下的矩阵.

给定矩阵 A，$\varphi_A : \mathbb{R}^n \to \mathbb{R}^m$ 是满足

$$\varphi(\mathbf{e}^{(j)}) = \mathbf{A}^{(j)}$$

的线性映射。

我们记所有 $\mathbb{R}^n \to \mathbb{R}^m$ 的线性映射组成的集合为 $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$。

矩阵的加法和数乘

定理 2.4.1 设 $\Phi : \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R}^{m\times n}$，$\varphi \mapsto A_\varphi$，则 Φ 是双射，并且 Φ^{-1} 是

$$\Psi =: \mathbb{R}^{m\times n} \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m),\ A \mapsto \varphi_A.$$

证明: 由于

$$\begin{aligned}
\Phi \circ \Psi(A) &= \Phi(\varphi_A) && (\Psi\text{的定义}) \\
&= A_{\varphi_A} && (\Phi\text{的定义}) \\
&= (\varphi_A(\mathbf{e}^{(1)}), \cdots, \varphi_A(\mathbf{e}^{(n)})) && (A_{\varphi_A}\text{的定义}) \\
&= (\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(n)}) && (\varphi_A\text{的定义}).
\end{aligned}$$

于是 $\Phi \circ \Psi = \mathrm{id}_{\mathbb{R}^{m\times n}}$。

反之有

$$\begin{aligned}
\Psi \circ \Phi(\varphi) &= \Psi(A_\varphi) && (\Phi\text{的定义}) \\
&= \varphi_{A_\varphi} && (\Psi\text{的定义}).
\end{aligned}$$

对任意的 $j \in \{1, \cdots, n\}$，注意到 $\varphi_{A_\varphi}(\mathbf{e}^{(j)}) = \mathbf{A}_\varphi^{(j)} = \varphi(\mathbf{e}^{(j)})$(由 φ_{A_φ} 的定义)，于是由线性映射基本定理 (唯一性)，有 $\varphi_{A_\varphi} = \varphi$，即对 $\forall \varphi \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$，$\Psi \circ \Phi(\varphi) = \varphi$，即 $\Psi \circ \Phi = \mathrm{id}_{\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)}$。 □

例 2.4.1 (1) 设 $O_{m\times n}$ 是零矩阵，则 $\varphi_{O_{m\times n}}(\mathbf{e}^{(j)}) = \mathbf{O}_{m\times n}^{(j)} = \mathbf{O}_m$，$j = 1, \cdots, n$. 这对应零映射，即 $\forall x_1, \cdots, x_n \in \mathbb{R}$，$\varphi_{O_{m\times n}}(x_1\mathbf{e}^{(1)} + \cdots + x_n\mathbf{e}^{(n)}) = \mathbf{O}_m$。

(2) 设 E_n 是 n 阶单位方阵，则 $\varphi_{E_n}(\mathbf{e}^{(j)}) = \mathbf{E}_n^{(j)} = \mathbf{e}^{(j)}$，$j = 1, \cdots, n$. 则对 $\forall x_1, \cdots, x_n \in \mathbb{R}$，有

$$\begin{aligned}
\varphi_{E_n}(x_1\mathbf{e}^{(1)} + \cdots + x_n\mathbf{e}^{(n)}) &= x_1\varphi_{E_n}(\mathbf{e}^{(1)}) + \cdots + x_n\varphi_{E_n}(\mathbf{e}^{(n)}) \\
&= x_1\mathbf{e}^{(1)} + \cdots + x_n\mathbf{e}^{(n)}.
\end{aligned}$$

即 φ_{E_n} 是恒同映射。

命题 2.4.1 设 $\varphi,\psi\in \mathrm{Hom}(\mathbb{R}^n,\mathbb{R}^m)$，$\alpha\in\mathbb{R}$，则我们有如下定义：

$$\varphi+\psi:\mathbb{R}^n\to\mathbb{R}^m,\ \mathbf{x}\mapsto\varphi(\mathbf{x})+\psi(\mathbf{x});$$
$$\alpha\varphi:\mathbb{R}^n\to\mathbb{R}^m,\ \mathbf{x}\mapsto\alpha\varphi(\mathbf{x}).$$

则 $\varphi+\psi$ 和 $\alpha\varphi$ 都是线性映射 (留作练习)。

设 φ 和 ψ 在标准基下的矩阵分别是 A,B，那么 $\varphi+\psi$ 在标准基下的矩阵

$$A_{\varphi+\psi}=(\mathbf{A}^{(1)}+\mathbf{B}^{(1)},\cdots,\mathbf{A}^{(n)}+\mathbf{B}^{(n)});$$
$$A_{\alpha\varphi}=(\alpha\mathbf{A}^{(1)},\cdots,\alpha\mathbf{A}^{(n)}).$$

证明留作练习。

定义 2.4.1 设 A,B 是 $\mathbb{R}$ 上的 $m\times n$ 矩阵，$\alpha\in\mathbb{R}$。则定义矩阵 $A+B=(\mathbf{A}^{(1)}+\mathbf{B}^{(1)},\cdots,\mathbf{A}^{(n)}+\mathbf{B}^{(n)})$ 及 $\alpha A=(\alpha\mathbf{A}^{(1)},\cdots,\alpha\mathbf{A}^{(n)})$，即矩阵的加法和数乘。若 $A=(a_{ij})$，$B=(b_{ij})$，则 $A+B=(a_{ij}+b_{ij})$，$\alpha A=(\alpha a_{ij})$。我们还可以看到，$A=(\mathbf{A}_1+\mathbf{B}_1,\cdots,\mathbf{A}_m+\mathbf{B}_m)^t$，$\alpha A=(\alpha\mathbf{A}_1,\cdots,\alpha\mathbf{A}_m)^t$。此外，自然地可以定义 $A-B=A+(-1)B$。

例 2.4.2 设 $A=\begin{pmatrix}1&2\\3&4\end{pmatrix},B=\begin{pmatrix}1&0\\2&1\end{pmatrix}$，则 $A-B=\begin{pmatrix}0&2\\1&3\end{pmatrix}$。

有了线性映射和矩阵的加法和数乘的定义以后，我们很容易得到以下推论：

推论 2.4.1 定理 2.4.1 中的 Φ 是线性双射。

证明: 记号与前面相同。双射已经在定理 2.4.1 中证明，下面证明线性性。注意到

$$\begin{aligned}\Phi(\alpha\varphi+\beta\psi)&=A_{\alpha\varphi+\beta\psi}\\&=A_{\alpha\varphi}+A_{\beta\psi}&&\text{(矩阵加法定义)}\\&=\alpha A_\varphi+\beta A_\psi&&\text{(矩阵数乘定义)}\\&=\alpha\Phi(\varphi)+\beta\Phi(\psi)&&(\Phi\text{的定义}).\end{aligned}$$

即得结论。 □

例 2.4.3 设 $A,B\in\mathbb{R}^{m\times n}$，则 $\mathrm{rank}(A+B)\leqslant\mathrm{rank}(A)+\mathrm{rank}(B)$。

证明: 由矩阵加法的定义，显然有：

$$V_c(A+B)=<\mathbf{A}^{(1)}+\mathbf{B}^{(1)},\cdots,\mathbf{A}^{(n)}+\mathbf{B}^{(n)}>.$$

于是 $\forall j\in\{1,2,\cdots,n\}$，$\mathbf{A}^{(j)}+\mathbf{B}^{(j)}\in V_c(A)+V_c(B)$，即 $V_c(A+B)\subset V_c(A)+V_c(B)$。那么，由秩的定义即得：

$$\begin{aligned}\mathrm{rank}(A+B)&=\dim V_c(A+B)\\&\leqslant\dim(V_c(A)+V_c(B))\\&\leqslant\dim(V_c(A))+\dim(V_c(B))\\&\leqslant\mathrm{rank}(A)+\mathrm{rank}(B).\end{aligned}$$

即得结论。 □

矩阵的转置

定义 2.4.2 设 $A \in \mathbb{R}^{m\times n}$, $A = (a_{ij})_{m\times n}$，则 A 的转置 (transpose) 是 $n \times m$ 阶矩阵，记为 A^t。即

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

以下两个性质是显然的。

命题 2.4.2 设 $A \in \mathbb{R}^{m\times n}$，则 $(A^t)^t = A$, $\operatorname{rank}(A) = \operatorname{rank}(A^t)$。

证明: 由定义即有 $(A^t)^t = A$。因为 $\dim(V_r(A^t)) = \dim(V_c(A))$，所以 $\operatorname{rank}(A) = \operatorname{rank}(A^t)$。 □

命题 2.4.3 设 $A, B \in \mathbb{R}^{m\times n}$, $\alpha \in \mathbb{R}$，则 $(A + B)^t = A^t + B^t$, $(\alpha A)^t = \alpha A^t$。证明留作练习。

矩阵的乘法

首先我们考虑线性映射的复合。

命题 2.4.4 设 $\psi \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^s)$, $\varphi \in \operatorname{Hom}(\mathbb{R}^s, \mathbb{R}^m)$，则 $\varphi \circ \psi \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$。

证明: 即下面的交换图。

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "\psi"] \arrow[dr, "\varphi\circ\psi"'] & \mathbb{R}^s \arrow[d, "\varphi"] \\ & \mathbb{R}^m \end{tikzcd}$$

下面验证 $\varphi \circ \psi$ 的线性性。

设 $\alpha, \beta \in \mathbb{R}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$，则

$$\varphi \circ \psi(\alpha\mathbf{x} + \beta\mathbf{y}) = \varphi(\alpha\psi(\mathbf{x}) + \beta\psi(\mathbf{y})) = \alpha\varphi \circ \psi(\mathbf{x}) + \beta\varphi \circ \psi(\mathbf{y}).$$

即得结论。 □

下面我们把线性映射的复合用矩阵表示出来。设 ψ, φ, $\varphi\circ\psi$ 在标准基下的矩阵分别是 B, A, C，则

$$C = (\varphi \circ \psi(\mathbf{e}^{(1)}), \cdots, \varphi \circ \psi(\mathbf{e}^{(n)})) = (\varphi(\mathbf{B}^{(1)}), \cdots, \varphi(\mathbf{B}^{(n)})).$$

设 $B = (b_{kj})_{s\times n}$，即

$$\mathbf{B}^{(j)} = (b_{ij}, \cdots, b_{sj})^t = b_{ij}\boldsymbol{\delta}_1 + \cdots + b_{sj}\boldsymbol{\delta}_s$$

其中 $\boldsymbol{\delta}_1, \cdots, \boldsymbol{\delta}_s$ 是 $\mathbb{R}^s$ 的标准基。则

$$\varphi(\mathbf{B}^{(j)}) = b_{1j}\varphi(\boldsymbol{\delta}_1) + \cdots + b_{sj}\varphi(\boldsymbol{\delta}_s) = b_{1j}\mathbf{A}^{(1)} + \cdots + b_{sj}\mathbf{A}^{(s)}.$$

于是 $\mathbf{C}^{(j)} = b_{1j}\mathbf{A}^{(1)} + \cdots + b_{sj}\mathbf{A}^{(s)}$。令 $C = (c_{ij})_{m\times n}$, $A = (a_{ik})_{m\times s}$，则

$$\begin{aligned} c_{ij} &= b_{1j}a_{i1} + \cdots + b_{sj}a_{is} \\ &= a_{i1}b_{1j} + \cdots + a_{is}b_{sj} \\ &= \sum_{k=1}^{s} a_{ik}b_{kj}. \end{aligned}$$

此即矩阵的乘法。

定义 2.4.3 设 $A \in \mathbb{R}^{m\times s}$，$B \in \mathbb{R}^{s\times n}$，则定义 A 与 B 的乘积是 $m\times n$ 阶矩阵 $C = (c_{ij})_{m\times n}$，其中 $c_{ij} = \sum_{k=1}^{s} a_{ik}b_{kj}$。记作 $C = AB$。

由上面的讨论可知，$\varphi \circ \psi$ 在标准基下的矩阵就是 $C = AB$，即 $\varphi_A \circ \varphi_B = \varphi_{AB}$。

例 2.4.4 设 $(\alpha_1,\ldots,\alpha_s) \in \mathbb{R}^{1\times s}$，$(\beta_1,\ldots,\beta_s)^t \in \mathbb{R}^{s\times 1}$，则

$$(\alpha_1,\ldots,\alpha_s)\begin{pmatrix}\beta_1\\ \vdots\\ \beta_s\end{pmatrix} = \alpha_1\beta_1 + \cdots + \alpha_s\beta_s.$$

于是我们看到，对于 $C = AB = (c_{ij})$，由于

$$c_{ij} = \sum_{k=1}^{s} a_{ik}b_{kj} = (a_{i1},\ldots,a_{is})\begin{pmatrix}b_{ij}\\ \vdots\\ b_{sj}\end{pmatrix} = \mathbf{A}_i\mathbf{B}^{(j)}.$$

因此

$$AB = \begin{pmatrix}\mathbf{A}_1\mathbf{B}^{(1)} & \cdots & \mathbf{A}_1\mathbf{B}^{(n)}\\ \vdots & & \vdots\\ \mathbf{A}_m\mathbf{B}^{(1)} & \cdots & \mathbf{A}_m\mathbf{B}^{(n)}\end{pmatrix}$$

例 2.4.5 (1) 设 $A = \begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$，$B = \begin{pmatrix}2 & 1 & 0\\ 3 & 1 & 3\end{pmatrix}$，则 $AB = \begin{pmatrix}8 & 3 & 6\\ 18 & 7 & 12\end{pmatrix}$，而 BA 没有定义。
(2) 即使 AB 和 BA 都有定义（此时 A, B 必须都是方阵），一般地也有 $AB \neq BA$。例如，若 $A = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$，$B = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$，则 $AB = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$，$BA = \begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$。

例 2.4.6 设 $A \in \mathbb{R}^{m\times n}$, $\operatorname{rank}(A) = 1$, 则 $\exists \alpha_1,\ldots,\alpha_m$ 及 $\beta_1,\ldots,\beta_n \in \mathbb{R}$ 使得 $A = (\alpha_1,\ldots,\alpha_m)^t(\beta_1,\cdots,\beta_n)$。

证明: 由于 $\operatorname{rank}(A) = 1$，即 $\dim V_c(A) = 1$，设 $\boldsymbol{\alpha} = (\alpha_1,\cdots,\alpha_m)^t$ 是 $V_c(A)$ 的一组基，则对 $\forall j = 1,\cdots,n$, $\exists \beta_j \in \mathbb{R}$ 使得 $\mathbf{A}^{(j)} = \beta_j\boldsymbol{\alpha}$，即 $A = (\alpha_1,\cdots,\alpha_m)^t(\beta_1,\cdots,\beta_n)$。 □

矩阵加法、数乘和乘法的运算律

我们首先陈述矩阵关于加法和数乘的运算律，它们的证明都十分容易，留作练习。

设 $A, B, C \in \mathbb{R}^{m\times n}$，$\alpha, \beta \in \mathbb{R}$，则有：

$A + B = B + A$	加法交换
$(A + B) + C = A + (B + C)$	加法结合
$A + O_{m\times n} = A$	加法单位
$A + (-A) = O_{m\times n}$	加法逆
$\alpha(\beta A) = (\alpha\beta)A$	数乘结合
$1 \cdot A = A$	数乘单位
$\alpha(A + B) = \alpha A + \alpha B$	分配律
$(\alpha + \beta)A = \alpha A + \beta A$	分配律

以后我们会看到，$\mathbb{R}^{m\times n}$ 在加法和数乘下构成了一个**向量空间**。

接下来我们考虑矩阵的乘法。首先，矩阵的乘法满足结合律，这可以由映射复合的结合律以及矩阵与线性映射之间的一一对应 (确切地说，是线性同构) 证得。设 $A \in \mathbb{R}^{m\times s}$, $B \in \mathbb{R}^{s\times k}$, $C \in \mathbb{R}^{k\times n}$，我们欲证明 $(AB)C = A(BC)$。证明可以用下面两个图表示，其中用到了定理 1.3.2 和定理 2.4.1。

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "\varphi_C"] \arrow[dr, dashed, "\varphi_B\circ\varphi_C=\varphi_{BC}"'] & \mathbb{R}^k \arrow[d, "\varphi_B"] \arrow[dr, dashed, "\varphi_A\circ\varphi_B=\varphi_{AB}"] & \\
 & \mathbb{R}^s \arrow[r, "\varphi_A"'] & \mathbb{R}^m
\end{tikzcd}$$

$$\begin{array}{ccc}
\varphi_A \circ (\varphi_B \circ \varphi_C) & = & (\varphi_A \circ \varphi_B) \circ \varphi_C \\
\| & & \| \\
\varphi_A \circ \varphi_{BC} & & \varphi_{AB} \circ \varphi_C \\
\| & & \| \\
\varphi_{A(BC)} & = & \varphi_{(AB)C}
\end{array}$$

矩阵乘法与加法之间的左右分配律如下：设 $A \in \mathbb{R}^{m\times s}$，$B, C \in \mathbb{R}^{s\times n}$，则 $A(B+C) = AB + AC$；设 $A \in \mathbb{R}^{n\times k}$，$B, C \in \mathbb{R}^{s\times n}$，则 $(B+C)A = BA + CA$。

矩阵乘法和数乘之间也有结合律：设 $A \in \mathbb{R}^{m\times s}$，$B \in \mathbb{R}^{s\times n}$，$\alpha \in \mathbb{R}$，则 $(\alpha A)B = A(\alpha B) = \alpha(AB)$。

下面考虑矩阵乘法和转置的关系。

命题 2.4.5 设 $A \in \mathbb{R}^{m\times s}$，$B \in \mathbb{R}^{s\times n}$，则 $(AB)^t = B^t A^t$。

证明: 设 $A = (a_{ik})_{m\times s}$，$B = (b_{kj})_{s\times n}$，则 $A^t = (a'_{kj})_{s\times m}$，$B^t = (b'_{jk})_{n\times s}$，其中 $a_{ik} = a'_{ki}$，$b_{kj} = b'_{jk}$。令 $C = AB = (c_{ij})_{m\times n}$，$D = B^t A^t = (d_{ji})_{n\times m}$，则

$$d_{ji} = \sum_{k=1}^{s} b'_{jk} a'_{ki} = \sum_{k=1}^{s} a_{ik} b_{kj} = c_{ij}.$$

即 $D^t = C$。 □

对角矩阵

下面我们考虑一类特殊的矩阵：对角矩阵。由于对角矩阵具有比较简单的形式和运算性质，因此线性代数课程的主线之一就是如何将矩阵化成对角矩阵。

定义 2.4.4 形如 $\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}_{n\times n}$ 的方阵称为对角矩阵。

对角矩阵与其它矩阵的乘法是简单的。

命题 2.4.6 设 $A \in \mathbb{R}^{m\times n}$，则

(i) $\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_m \end{pmatrix} A = \begin{pmatrix} \lambda_1 \mathbf{A}_1 \\ \vdots \\ \lambda_m \mathbf{A}_m \end{pmatrix}$；

(ii) $A\begin{pmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_n\end{pmatrix} = \left(\lambda_1\mathbf{A}^{(1)}, \cdots, \lambda_n\mathbf{A}^{(n)}\right)$.

推论 2.4.2 设 $A \in \mathbb{R}^{m\times n}$，则 $(\lambda E_m)A = A(\lambda E_n) = \lambda A$。特别地，$n$ 阶单位矩阵的倍数与任何 n 阶方阵都可交换。

实际上，与任何 n 阶方阵都可交换的矩阵只能是，n 阶单位矩阵的倍数，证明见下节。

例 2.4.7 $A = \begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$，$B = \begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}$，则 $AB = \begin{pmatrix}1 & 4\\ 3 & 8\end{pmatrix}$，$BA = \begin{pmatrix}1 & 2\\ 6 & 8\end{pmatrix}$，仍然有 $AB \neq BA$。

秩不等式

最后我们从线性映射的角度证明一些关于矩阵的秩的不等式。更多有关秩不等式的内容可以参考习题课讲义。

定理 2.4.2 设 $A \in \mathbb{R}^{m\times s}$，$B \in \mathbb{R}^{s\times n}$，则 $\mathrm{rank}(AB) \leqslant \min(\mathrm{rank}(A), \mathrm{rank}(B))$。

证明：只需证 $\mathrm{rank}(AB) \leqslant \mathrm{rank}(A)$，$\mathrm{rank}(AB) \leqslant \mathrm{rank}(B)$ 即可。对于 $\mathrm{rank}(AB) \leqslant \mathrm{rank}(A)$，只需注意到下面的交换图：

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "\varphi_B"] \arrow[dr, "\varphi_{AB}"'] & \mathbb{R}^s \arrow[d, "\varphi_A"] \\ & \mathbb{R}^m \end{tikzcd} \quad \text{于是} \quad \begin{array}{ccc} \mathrm{im}(\varphi_B) & \subset & \mathbb{R}^s \\ & \Downarrow & \\ \varphi_A(\mathrm{im}(\varphi_B)) & \subset & \varphi_A(\mathbb{R}^s) \\ \| & & \| \\ \mathrm{im}(\varphi_{AB}) & \subset & \mathrm{im}(\varphi_A) \end{array} \text{，因此} \begin{array}{ccc} \dim(\mathrm{im}(\varphi_{AB})) & \leqslant & \dim(\mathrm{im}(\varphi_A)) \\ \| & & \| \\ \mathrm{rank}(AB) & \leqslant & \mathrm{rank}(A) \end{array}.$$

下面证明 $\mathrm{rank}(AB) \leqslant \mathrm{rank}(B)$。设 $\mathbf{x} \in \ker(\varphi_B)$，即 $\varphi_B(\mathbf{x}) = \mathbf{0}$，则

$$\varphi_{AB}(\mathbf{x}) = \varphi_A \circ \varphi_B(\mathbf{x}) = \varphi_A(\varphi_B(\mathbf{x})) = \varphi_A(\mathbf{0}) = \mathbf{0},$$

即 $\mathbf{x} \in \ker(\varphi_{AB})$。由 $\mathbf{x}$ 的任意性知 $\ker(\varphi_B) \subset \ker(\varphi_{AB})$，因此 $\dim(\ker(\varphi_B)) \leqslant \dim(\ker(\varphi_{AB}))$。再利用对偶定理可得 $n - \dim(\mathrm{im}(\varphi_B)) \leqslant n - \dim(\mathrm{im}(\varphi_{AB}))$，所以 $\dim(\mathrm{im}(\varphi_{AB})) \leqslant \dim(\mathrm{im}(\varphi_B))$，即 $\mathrm{rank}(AB) \leqslant \mathrm{rank}(B)$。 □

定理 2.4.3 (Sylvester **不等式**) 设 $A \in \mathbb{R}^{m\times s}$，$B \in \mathbb{R}^{s\times n}$，则 $\mathrm{rank}(AB) \geqslant \mathrm{rank}(A) + \mathrm{rank}(B) - s$。

证明：设 $\mathbf{u}_1, \cdots, \mathbf{u}_d$ 是 $\ker(\varphi_A) \cap \mathrm{im}(\varphi_B)$ 的一组基，则由定理 2.1.1(基扩充定理)，存在 $\mathbf{u}_{d+1}, \cdots, \mathbf{u}_k \in \mathrm{im}(\varphi_B)$ 使得 $\mathbf{u}_1, \cdots, \mathbf{u}_d, \mathbf{u}_{d+1}, \cdots, \mathbf{u}_k$ 是 $\mathrm{im}(\varphi_B)$ 的一组基。下面证明 $\varphi_A(\mathbf{u}_{d+1}), \cdots, \varphi_A(\mathbf{u}_k)$ 线性无关。假设 $\exists \alpha_{d+1}, \cdots, \alpha_k \in \mathbb{R}$ 使得

$$\alpha_{d+1}\varphi_A(\mathbf{u}_{d+1}) + \cdots + \alpha_k\varphi_A(\mathbf{u}_k) = \mathbf{0}. \tag{$*$}$$

我们的证明目标是 $\alpha_{d+1} = \cdots = \alpha_k = 0$。利用 $(*)$ 式及 φ_A 的线性性知 $\varphi_A(\sum_{i=d+1}^k \alpha_i\mathbf{u}_i) = \mathbf{0}$，则 $\sum_{i=d+1}^k \alpha_i\mathbf{u}_i \in \ker(\varphi_A)$。于是 $\sum_{i=d+1}^k \alpha_i\mathbf{u}_i \in \ker(\varphi_A) \cap \mathrm{im}(\varphi_B)$。那么，由 $\mathbf{u}_1, \cdots, \mathbf{u}_d$ 是 $\ker(\varphi_A) \cap \mathrm{im}(\varphi_B)$ 的一组基可知：$\exists \alpha_1, \cdots, \alpha_d \in \mathbb{R}$ 使得 $\alpha_1\mathbf{u}_1 + \cdots + \alpha_d\mathbf{u}_d - \sum_{i=d+1}^k \alpha_i\mathbf{u}_i = \mathbf{0}$。再利用 $\mathbf{u}_1, \cdots, \mathbf{u}_k$ 是 $\mathrm{im}(\varphi_B)$ 的一组基就有：$\alpha_1 = \cdots = \alpha_k = 0$，特别地，$\alpha_{d+1} = \cdots = \alpha_k = 0$，即 $\varphi_A(\mathbf{u}_{d+1}), \cdots, \varphi_A(\mathbf{u}_k)$ 线

性无关成立。
之后，我们只需注意到 $k=\dim(\mathrm{im}(\varphi_B))=\mathrm{rank}(B)$，于是

$$\mathrm{rank}(AB)=\dim(\mathrm{im}(\varphi_{AB}))\geqslant k-d=\mathrm{rank}(B)-d.^{1}$$

又因为

$$\begin{aligned} d&=\dim(\ker(\varphi_A)\cap\mathrm{im}(\varphi_B))\\ &\leqslant\dim(\ker(\varphi_A))=s-\dim(\mathrm{im}(\varphi_A))=s-\mathrm{rank}(A).\end{aligned}$$

结合上面两个不等式即有：$\mathrm{rank}(AB)\geqslant\mathrm{rank}(B)-d\geqslant\mathrm{rank}(B)+\mathrm{rank}(A)-s$。 □

这个证明的方法我们在下册还会再次用到。

推论 2.4.3 设 $P\in\mathbb{R}^{m\times m}$，$Q\in\mathbb{R}^{n\times n}$，$A\in\mathbb{R}^{m\times n}$，则
(i) 如果 $\mathrm{rank}(P)=m$，则 $\mathrm{rank}(PA)=\mathrm{rank}(A)$；
(ii) 如果 $\mathrm{rank}(Q)=n$，则 $\mathrm{rank}(AQ)=\mathrm{rank}(A)$。

证明: 只证明 (i), (ii) 同理。利用定理 2.4.1 和 2.4.2 立刻可得: $\mathrm{rank}(P)+\mathrm{rank}(A)-m\leqslant\mathrm{rank}(PA)\leqslant\min(\mathrm{rank}(P),\mathrm{rank}(A))$，即 $m+\mathrm{rank}(A)-m\leqslant\mathrm{rank}(PA)\leqslant\mathrm{rank}(A)$，即 $\mathrm{rank}(PA)=\mathrm{rank}(A)$。 □

[1]其中 ”$\geqslant$” 来源于：维数 $=$ 极大线性无关组的元素个数 $\geqslant$ 线性无关组的元素个数。

2.5 方阵

这一节我们专门讨论行数和列数相等的矩阵，即方阵。我们将所有方阵的集合 $\mathbb{R}^{n\times n}$ 也记为 $M_n(\mathbb{R})$。$M_n(\mathbb{R})$ 上的加法和数乘显然满足上一节关于一般矩阵的运算律，并且对于矩阵乘法而言，满足乘法的封闭性、结合律、单位元 E_n(无歧义时也记作 E 或 I)，并且乘法对加法满足左右分配律，即满足环的公理，$M_n(\mathbb{R})$ 是环。又因为纯量乘法满足

$$\lambda(AB) = (\lambda A)B = A(\lambda B)$$

集合 $M_n(\mathbb{R})$ 也称为 $\mathbb{R}$ 上的 n 阶矩阵代数。本节中的矩阵如无特殊说明的都是 n 阶方阵。

定义 2.5.1 我们定义方阵的幂如下：$A^k = \underbrace{A\cdots A}_{k个}$。特别地，定义 $A^0 = E$。显然，对 $\forall k, l \in \mathbb{N}$ 都有 $A^{k+l} = A^k A^l$。

注 2.5.1 对 $A, B \in M_n(\mathbb{R})$，我们已经通过反例说明了 $AB \neq BA$；并且，例 2.4.5(2) 还表明，$AB = O$ 不能推出 $A = O$ 或 $B = O$。

例 2.5.1 (1) 设 $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$，则 $D^m = \begin{pmatrix} \lambda_1^m & & \\ & \ddots & \\ & & \lambda_n^m \end{pmatrix}$；

(2) 设 $A = \begin{pmatrix} a & c \\ 0 & b \end{pmatrix}$ $(a \neq b)$，则 $A^m = \begin{pmatrix} a^m & c\frac{a^m-b^m}{a-b} \\ 0 & b^m \end{pmatrix}$。

一般的，我们把第 i 行第 j 列 (以后简称 (i, j) 位置) 的元素为 1，其余位置的元素都为 0 的矩阵记作 E_{ij}，则显然有

$$AE_{ij} = (\mathbf{0}, \cdots, \mathbf{0}, \mathbf{A}^{(i)}, \mathbf{0}, \cdots, \mathbf{0}) = \begin{pmatrix} 0 & \cdots & a_{1i} & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & a_{ni} & \cdots & 0 \end{pmatrix}(\text{第 j 列是 } \mathbf{A}^{(i)}); \tag{2.5.1}$$

$$E_{ij}A = (\mathbf{0}, \cdots, \mathbf{0}, \mathbf{A}_j^t, \mathbf{0}, \cdots, \mathbf{0})^t = \begin{pmatrix} 0 & \cdots & 0 \\ & \cdots & \\ a_{j1} & \cdots & a_{jn} \\ & \cdots & \\ 0 & \cdots & 0 \end{pmatrix}(\text{第 i 行是 } \mathbf{A}_j). \tag{2.5.2}$$

定义 2.5.2 设 $A \in M_n(\mathbb{R})$，如果对 $\forall B \in M_n(\mathbb{R})$，有 $AB = BA$，则称 A 是 $M_n(\mathbb{R})$ 中的中心元。

显然对 $\forall \lambda \in \mathbb{R}$, λE 是中心元。

定理 2.5.1 $M_n(\mathbb{R})$ 中的中心元都是数乘矩阵 λE, $\lambda \in \mathbb{R}$。

证明: 设 $A = (a_{ij})_{n\times n}$ 是中心元，我们解出 A 的形式即可。首先，对 $\forall B = (b_{ij})_{n\times n}$，有 $B = \sum_{i,j=1}^n b_{ij}E_{ij}$。于是，要使 $AB = BA$ 对任意 B 成立，只需 $\forall i, j = 1, \cdots, n$，有 $AE_{ij} = E_{ij}A$ 即可。于是由 (2.5.1) 及 (2.5.2) 式，有 $\forall i \neq j$, $a_{ij} = 0$；而对所有 $i = j$ 则有 $a_{ii} = a_{jj}$。此即 A 是数乘矩阵。 □

例 2.5.2 设 $A \in M_n(\mathbb{R})$，求 $(A + \lambda E)^k$。

解: 利用 $A(\lambda E) = (\lambda E)A = \lambda A$ 即可得到：

$$\begin{aligned}(A+\lambda E)^k &= A^k + \binom{k}{1}A^{k-1}(\lambda E) + \cdots + \binom{k}{k-1}A(\lambda E)^{k-1} + \lambda^k E \\ &= A^k + \binom{k}{1}\lambda A^{k-1} + \cdots + \binom{k}{k-1}\lambda^{k-1}A + \lambda^k E\end{aligned}$$

接下来我们引入一类重要的方阵：可逆矩阵。

定义 2.5.3 设 $A \in M_n(\mathbb{R})$，如果 $\exists B \in M_n(\mathbb{R})$ 使得 $AB = BA = E$，则称 A 是可逆矩阵，且称 B 是 A 的逆矩阵，记为 $B = A^{-1}$。

下面的命题表明逆矩阵如果存在则必唯一。

命题 2.5.1 设 $A, B, C \in M_n(\mathbb{R})$, $AB = BA = E$，若 $CA = E$ 或 $AC = E$ 有一个成立，则必有 $B = C$。

证明: 只证明 $CA = E$ 的情形，另一个同理。由 $B = EB = (CA)B = C(AB) = CE = C$ 即得结论。□

定理 2.5.2 设 $A \in M_n(\mathbb{R})$，则 A 可逆 $\Longleftrightarrow$ A 满秩。

证明: ($\Longrightarrow$) 设 $B \in M_n(\mathbb{R})$ 且 $AB = E$，则 $n = \mathrm{rank}(AB) \leqslant \min(\mathrm{rank}(A), \mathrm{rank}(B)) \leqslant \mathrm{rank}(A) \leqslant n$，即 $\mathrm{rank}(A) = n$ 成立。
($\Longleftarrow$) 设 $\varphi_A : \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{x} \mapsto A\mathbf{x}$ 是 A 对应的线性映射，由于 $\mathrm{rank}(A) = \dim(\mathrm{im}(\varphi_A)) = n$，由推论 2.3.1 知 φ_A 是双射，于是存在 φ_A 的逆映射 φ_A^{-1}(也是双射)。我们验证 φ_A^{-1} 是线性映射。设 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$，由 φ_A 是满射可得 $\exists \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ 使得 $\varphi_A(\mathbf{u}) = \mathbf{x}$，$\varphi_A(\mathbf{v}) = \mathbf{y}$。于是由 φ_A 的线性性可得 $\forall \alpha, \beta \in \mathbb{R}$, $\varphi_A(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha\mathbf{x} + \beta\mathbf{y}$，那么，按照 φ_A^{-1} 的定义就分别有：

$$\alpha\mathbf{u} + \beta\mathbf{v} = \varphi_A^{-1}(\alpha\mathbf{x} + \beta\mathbf{y}),\ \alpha\mathbf{u} + \beta\mathbf{v} = \alpha\varphi_A^{-1}(\mathbf{x}) + \beta\varphi_A^{-1}(\mathbf{y})$$

此即 φ_A^{-1} 是线性映射。于是，可设 φ_A^{-1} 对应的矩阵是 B，即 $\varphi_A^{-1} = \varphi_B$。那么就有

$$\begin{aligned}\varphi_{AB} &= \varphi_A \circ \varphi_B = \varphi_A \circ \varphi_A^{-1} = \mathrm{id}_{\mathbb{R}^n} \Longrightarrow AB = E; \\ \varphi_{BA} &= \varphi_B \circ \varphi_A = \varphi_A^{-1} \circ \varphi_A = \mathrm{id}_{\mathbb{R}^n} \Longrightarrow BA = E.\end{aligned}$$

即 A 可逆且逆矩阵是 B。□

推论 2.5.1 设 $A, B, C \in M_n(\mathbb{R})$，如果 $AB = E$ 或 $CA = E$，则 $B = A^{-1}$ 或 $C = A^{-1}$。

证明: 只证明 $AB = E$ 的情形，另一个同理。由 $AB = E$ 易证 $\mathrm{rank}(A) = n$(见上面的 "$()\Longrightarrow)$")，于是 A 可逆，记 A 的逆为 A^{-1}，则 $AB = E \Longrightarrow A^{'}AB = A^{-1}$，即 $B = A^{-1}$。□

推论 2.5.2 设 $A, B \in M_n(\mathbb{R})$，如果 A, B 都可逆，则 AB 可逆且 $(AB)^{-1} = B^{-1}A^{-1}$。

证明: 由 $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = E$ 即得结论。□

推论 2.5.3 设 $A \in M_n(\mathbb{R})$ 是可逆矩阵，则 A^t 也可逆并且 $(A^t)^{-1} = (A^{-1})^t$。

证明: 注意到 $A^t(A^{-1})^t = (A^{-1}A)^t = E^t = E$ 即可。□

下面我们定义一些其它类型的方阵，它们的性质我们会在后面慢慢学习。

定义 2.5.4 设 $A \in M_n(\mathbb{R})$。如果 $A^t = A$，则称 A 是对称的；如果 $A^t = -A$，则称 A 是斜对称 (反对称) 的；如果存在 $k \in \mathbb{Z}^+$ 使得 $A^k = O$，则称 A 是幂零的；如果 $A^2 = E$，则称 A 是对合的；如果 $A^2 = A$，则称 A 是幂等的。

实际上，直接计算 AB，A^m 或 A^{-1} 在矩阵规模很大时不是一件很容易的事情。为此，我们往往要基于矩阵的特殊性质或者利用更高级的工具来简化计算。在本节的最后，我们给出下面的例子作为一个引子。

例 2.5.3 设 $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$，计算 A^m，$m \geqslant 0$。

解: 简单的计算可以猜测:$A^m = \begin{pmatrix} f_{m-1} & f_m \\ f_m & f_{m+1} \end{pmatrix}$,其中 $f_0 = f_1 = 1$，$f_{m+1} = f_m + f_{m-1}$，即 $\{f_m\}, m = 0, 1, \cdots$ 是斐波那契数列。我们可以用数学归纳法证明这个结论，但我们下面给出一个更直接的计算方法。

引入 $B = \begin{pmatrix} -\frac{\lambda_2}{5} & \frac{1}{5} \\ -\sqrt{5}\lambda_1 & \sqrt{5} \end{pmatrix}$，其中 $\lambda_1 = \dfrac{1+\sqrt{5}}{2}$，$\lambda_2 = \dfrac{1-\sqrt{5}}{2}$。不难计算出

$$B^{-1} = \begin{pmatrix} \sqrt{5} & -\frac{1}{5} \\ \sqrt{5}\lambda_1 & -\frac{\lambda_2}{5} \end{pmatrix}, \ A = B^{-1}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}B.$$

则 $A^m = B^{-1}\begin{pmatrix} \lambda_1^m & 0 \\ 0 & \lambda_2^m \end{pmatrix}B$，代入数值后即可得到 A^m 的通项公式。此外，通过对比元素也可以得到 f_m 的通项公式：

$$f_m = \frac{\sqrt{5}}{5}(\lambda_1^m - \lambda_2^m).$$

另外，由这个通项公式可知，$\lim_{m\to\infty} \dfrac{f_m}{\lambda_1^m} = \dfrac{\sqrt{5}}{5}$，即当 m 充分大时斐波那契数列与等比数列的增长速度相同。

2.6 矩阵的等价

这一节我们来更深入地讨论矩阵的初等变换。

定义 2.6.1 设 $A,B\in\mathbb{R}^{m\times n}$，如果存在 $P\in M_m(\mathbb{R})$，$Q\in M_n(\mathbb{R})$ 且 P,Q 都可逆，使得 $A=PBQ$，则称 A 和 B 初等等价，记为 $A\sim_e B$。

下面我们先验证 $\sim_e$ 是等价关系。
(1)$\forall A\in\mathbb{R}^{m\times n}$，显然 $A=E_mAE_n$，即 $A\sim_e A$；
(2) 若 $A\sim_e B$，即存在可逆矩阵 P,Q 使得 $A=PBQ$，那么 $B=P^{-1}AQ^{-1}$，即 $B\sim_e A$；
(3) 若 $A\sim_e B$，$B\sim_e C$，即存在可逆矩阵 P,Q,S,T 使得 $A=PBQ$，$B=STC$，则 $A=(PS)C(TQ)$，并且由推论 2.5.3 可知 PS,TQ 都可逆，所以 $A\sim_e C$。

命题 2.6.1 设 $A,B\in\mathbb{R}^{m\times n}$，如果 $A\sim_e B$，则 $\operatorname{rank}(A)=\operatorname{rank}(B)$。

证明: $A\sim_e B\Longrightarrow$ 存在可逆矩阵 P,Q 使得 $A=PBQ$。由推论 2.4.3 即得结论。 □

下面我们定义三类初等矩阵，它们都是可逆矩阵，并且分别对应于三类初等变换。

定义 2.6.2 将 E_n 中第 $i,j(i\neq j)$ 两行交换后得到的矩阵称为 (I) 型初等矩阵，记为 $F_{i,j}^{(n)}$(在不引起歧义时可以省略上角标 (n)，即 $F_{i,j}$，以下相同)。

显然 $F_{i,j}=E-E_{ii}-E_{jj}+E_{ij}+E_{ji}$ 并且 $F_{i,j}^2=E$，即 $F_{i,j}$ 可逆。此外，设 $A\in\mathbb{R}^{m\times n}$，则有

$$F_{i,j}^{m}A=\begin{pmatrix}\mathbf{A}_1\\ \vdots\\ \mathbf{A}_j\\ \vdots\\ \mathbf{A}_i\\ \vdots\\ \mathbf{A}_m\end{pmatrix}\qquad\text{(A 的第 i,j 行互换);}$$

$$AF_{i,j}^{n}=(\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(j)},\cdots,\mathbf{A}^{(i)},\cdots,\mathbf{A}^{(n)})\qquad\text{(第 i,j 列互换).}$$

定义 2.6.3 设 $\lambda\in\mathbb{R}$，$i,j\in\{1,\cdots,n\}$，$i\neq j$，将 E_n 中第 j 行乘以 λ 后加到第 i 行得到的矩阵称为 (II) 型初等矩阵，记为 $F_{i,j}^{(n)}(\lambda)$。

显然 $F_{i,j}(\lambda)=E+\lambda E_{ij}$，并且 $[F_{i,j}(\lambda)]^{-1}=F_{i,j}(-\lambda)$(验证之)。设 $A\in\mathbb{R}^{m\times n}$，则有

$$F_{i,j}^{(m)}(\lambda)A=\begin{pmatrix}\mathbf{A}_1\\ \vdots\\ \mathbf{A}_{i-1}\\ \mathbf{A}_i+\lambda\mathbf{A}_j\\ \mathbf{A}_{i+1}\\ \vdots\\ \mathbf{A}_m\end{pmatrix}\qquad\text{(A 的第 i 行加上第 j 行的 }\lambda\text{ 倍)}$$

$$AF_{i,j}^{(n)}(\lambda)=(\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(j)}+\lambda\mathbf{A}^{(i)},\cdots,\mathbf{A}^{(n)})\qquad\text{(A 的第 j 列加上第 i 列的 }\lambda\text{ 倍)}$$

定义 2.6.4 设 $\lambda \in \mathbb{R}, \lambda \neq 0$, $i \in \{1, \cdots, n\}$, 将 E_n 中第 i 行乘以 λ 得到的矩阵称为 (III) 型初等矩阵，记为 $F_i^{(n)}(\lambda)$。

显然 $F_i^{(n)}(\lambda) = E + (\lambda - 1)E_{ii}$ 并且 $[F_i^{(n)}(\lambda)]^{-1} = F_i^{(n)}(\frac{1}{\lambda})$。设 $A \in \mathbb{R}^{m\times n}$，则有

$$F_i^{(m)}(\lambda)A = \begin{pmatrix} \mathbf{A}_1 \\ \vdots \\ \lambda\mathbf{A}_i \\ \vdots \\ \mathbf{A}_m \end{pmatrix} \qquad \text{(A 的第 i 行乘以 } \lambda \text{ 倍)}$$

$$AF_{i,j}^{(n)}(\lambda) = (\mathbf{A}^{(1)}, \cdots, \lambda\mathbf{A}^{(i)}, \cdots, \mathbf{A}^{(n)}) \qquad \text{(A 的第 i 列乘以 } \lambda \text{ 倍)}$$

由以上讨论可以看到，矩阵的初等行变换就是左乘初等矩阵，矩阵的初等列变换就是右乘初等矩阵。

2.7 矩阵的求逆与秩标准型

之前我们定义了方阵的逆，但我们并没有给出具体计算逆矩阵的方法。这一节我们将会给出矩阵求逆的方法，并且定义矩阵的第一种标准型：秩标准型 (在下册我们还会学习矩阵的另外两种标准型：若尔当标准型和有理标准型)。

下面我们给出矩阵求逆的算法，并验证证明其正确性。

设 $A \in M_n(\mathbb{R})$，作 $B = (A \mid E_n)$，对 B 作初等行变换，若 $\operatorname{rank}(A) < n$，则 A 不可逆，算法结束；否则经过有限步后可将 B 中的子矩阵 A 化成 E，则此时 E 的部分就化成了 A^{-1}。这是因为对 B 作初等变换的过程等价于依次对 $(A \mid E)$ 左乘初等矩阵 $P_1, \cdots, P_k$。由于 $P_k \cdots P_1 A = E$，故 $A^{-1} = P_k \cdots P_1$，而 $P_k \cdots P_1 B = (P_k \cdots P_1 A \mid P_k \cdots P_1) = (E \mid A^{-1})$。

下面我们计算一个简单的例子来加深对这个算法的理解。

例 2.7.1 设 $A = \begin{pmatrix} 0 & 2 & 0 \\ 1 & 1 & -1 \\ 2 & 1 & -1 \end{pmatrix}$，求 A^{-1}。

解: 令 $B = \left(\begin{array}{ccc:ccc} 0 & 2 & 0 & 1 & 0 & 0 \\ 1 & 1 & -1 & 0 & 1 & 0 \\ 2 & 1 & -1 & 0 & 0 & 1 \end{array}\right)$，对 B 作以下初等行变换：

$$B = \left(\begin{array}{ccc:ccc} 0 & 2 & 0 & 1 & 0 & 0 \\ 1 & 1 & -1 & 0 & 1 & 0 \\ 2 & 1 & -1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{F_{1,2}} \left(\begin{array}{ccc:ccc} 1 & 1 & -1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 & 0 & 0 \\ 2 & 1 & -1 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{F_{3,1}(-2)} \left(\begin{array}{ccc:ccc} 1 & 1 & -1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & -2 & 1 \end{array}\right) \xrightarrow{F_{1,3}(1)} \left(\begin{array}{ccc:ccc} 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 2 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & -2 & 1 \end{array}\right)$$

$$\xrightarrow{F_2(\frac{1}{2})} \left(\begin{array}{ccc:ccc} 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & -1 & 1 & 0 & -2 & 1 \end{array}\right) \xrightarrow{F_{3,2}(1)} \left(\begin{array}{ccc:ccc} 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & \frac{1}{2} & 0 & 0 \\ \underbrace{0 \quad 0 \quad 1}_{E} & & & \underbrace{\frac{1}{2} \quad -2 \quad 1}_{A^{-1}} & & \end{array}\right)$$

这样我们就计算出了 A^{-1} 并且知道 $A^{-1} = F_{3,2}(1)F_2(\frac{1}{2})F_{1,3}(1)F_{3,1}(-2)F_{1,2}$。

定理 2.7.1 设 $A \in \mathbb{R}^{m\times n}$，$\operatorname{rank}(A) = r$，则存在 $P \in M_m(\mathbb{R})$，$Q \in M_n(\mathbb{R})$ 使得
(i)P, Q 都是有限个初等矩阵的乘积；
(ii)$PAQ = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{m\times n}$。我们把 A 对应的 $\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{m\times n}$ 称为 A 的秩标准型。

证明: 利用定理 2.2.1 的证明过程，我们知道存在若干个 (I)、(II) 型初等矩阵 $P_1, \cdots, P_s$ 和 $Q_1, \cdots, Q_t$ 使得 $P_s \cdots P_1 A Q_1 \cdots Q_s = \begin{pmatrix} \Lambda & O \\ O & O \end{pmatrix}$，其中 Λ 是 $r \times r$ 阶的对角矩阵。于是存在 r 个 (III) 型初等矩阵 $P_{s+1}, \cdots, P_{s+r}$ 使得 $P_{s+r} \cdots P_1 A Q_1 \cdots Q_s = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$。 □

现在我们可以讲清楚定义 2.6.1 中等价关系的等价类是什么了。

推论 2.7.1 $\mathbb{R}^{m\times n}/\sim_e=\left\{\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}\,\middle|\, r=0,\cdots,\min(m,n)\right\}$。

定理 2.7.1 也告诉我们，可逆矩阵一定可以写成若干个初等矩阵的积。

2.8 矩阵的分块

引理 2.8.1 设 $A\in\mathbb{R}^{m\times s}$，$B\in\mathbb{R}^{s\times n}$，则

(i) 若 $A=\begin{pmatrix}A'\\A''\end{pmatrix}$，则 $AB=\begin{pmatrix}A'B\\A''B\end{pmatrix}$;

(ii) 若 $B=(B',B'')$，则 $AB=(AB',AB'')$。

利用矩阵乘法的定义验证即可。

引理 2.8.2 设 $A\in\mathbb{R}^{m\times s}$，$B\in\mathbb{R}^{s\times n}$，并且有如下分块：

$$A=\begin{pmatrix}A_1\\\vdots\\A_p\end{pmatrix},\quad B=(B_1,\dots,B_q).$$

则

1. $AB=\begin{pmatrix}A_1B\\\vdots\\A_pB\end{pmatrix}=(AB_1,\dots,AB_q)$;
2. $AB=\begin{pmatrix}A_1B_1&\cdots&A_1B_q\\\vdots&\vdots&\vdots\\A_pB_1&\cdots&A_pB_q\end{pmatrix}$ 。

这是引理 2.8.1 的自然推广，用矩阵乘法的定义即可验证它。

引理 2.8.3 设 $A\in\mathbb{R}^{m\times s}$，$B\in\mathbb{R}^{s\times n}$，令 $A=(A_1,\ A_2)$，$A_1\in\mathbb{R}^{m\times s_1}$，$A_2\in\mathbb{R}^{m\times s_2}$，$B=\begin{pmatrix}B_1\\B_2\end{pmatrix}$，$B_1\in\mathbb{R}^{s_1\times n}$，$B_2\in\mathbb{R}^{s_2\times n}$，则 $AB=(A_1,\ A_2)\begin{pmatrix}B_1\\B_2\end{pmatrix}=A_1B_1+A_2B_2$。

同样用矩阵乘法的定义即可验证。

引理 2.8.4 设 $A\in\mathbb{R}^{m\times s}$，$B\in\mathbb{R}^{s\times n}$，令 $A=(A_1,\cdots,A_p)$，$A_l\in\mathbb{R}^{m\times s_l}$，$B=\begin{pmatrix}B_1\\\vdots\\B_p\end{pmatrix}$，$B_l\in\mathbb{R}^{s_l\times n}$，则

$$AB=(A_1,\cdots,A_p)\begin{pmatrix}B_1\\\vdots\\B_p\end{pmatrix}=A_1B_1+\cdots+A_pB_p。$$

有了以上引理，我们很容易证明矩阵的分块 (只要是合适的，即分出的块可以做运算) 是保持运算的。

定理 2.8.1 设 $A\in\mathbb{R}^{m\times s}$，$B\in\mathbb{R}^{s\times n}$，令 $A=(A_{ik})$，$A_{ik}\in\mathbb{R}^{m_i\times s_k}$，$B=(B_{kj})$，$B_{kj}\in\mathbb{R}^{s_k\times n_j}$，其中 $i=1,\cdots,p$，$k=1,\cdots l$，$j=1,\cdots,q$。则 $AB=(C_{ij})$，其中 $C_{ij}=A_{i1}B_{1j}+\cdots+A_{il}B_{lj}$。

例 2.8.1 设 $D=\begin{pmatrix}D_1&&\\&\ddots&\\&&D_k\end{pmatrix}$，其中 $D_i\in M_{n_i}(\mathbb{R})$，则 $D^m=\begin{pmatrix}D_1^m&&\\&\ddots&\\&&D_k^m\end{pmatrix}$。

例 2.8.2 (**矩阵的乘法分解**) 设 $A \in \mathbb{R}^{m\times n}$ 且 $\operatorname{rank}(A) = r > 0$，则存在 $B \in \mathbb{R}^{m\times r}$，$C \in \mathbb{R}^{r\times n}$ 使得 $A = BC$ 并且 $\operatorname{rank}(B) = \operatorname{rank}(C) = r$。

证明: 由定理 2.7.1，存在可逆矩阵 $P \in M_m(\mathbb{R})$，$Q \in M_n(\mathbb{R})$ 使得

$$A = P\begin{pmatrix} E_r & O \\ O & O \end{pmatrix} Q = P\begin{pmatrix} E_r \\ O \end{pmatrix}(E_r,\ O)Q.$$

取 $B = P\begin{pmatrix} E_r \\ O \end{pmatrix}$，$C = (E_r,\ O)Q$，则 $A = BC$ 且 $\operatorname{rank}(B) = \operatorname{rank}(C) = r$，即可。 □

下面的引理常用来证明秩不等式。

引理 2.8.5 设 $A \in \mathbb{R}^{m\times n}$，$B \in \mathbb{R}^{k\times l}$，$C \in \mathbb{R}^{k\times n}$，则 $\operatorname{rank}\begin{pmatrix} A & O \\ C & B \end{pmatrix} \geqslant \operatorname{rank}(A) + \operatorname{rank}(B)$，当 $C = O$ 时等号成立。

证明: 不妨设 $\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(p)}$ 是 $V_c(A)$ 的基，$\mathbf{B}^{(1)}, \cdots, \mathbf{B}^{(q)}$ 是 $V_c(B)$ 的基。则对 $M = \begin{pmatrix} A & O \\ C & B \end{pmatrix}$ 的相对应的 $p+q$ 列而言，若 $\sum_{i=1}^{p} \alpha_i \mathbf{M}^{(i)} + \sum_{j=1}^{q} \beta_j \mathbf{M}^{(j)} = \mathbf{0}$，即

$$\alpha_1\begin{pmatrix} \mathbf{A}^{(1)} \\ \mathbf{C}^{(1)} \end{pmatrix} + \cdots + \alpha_p\begin{pmatrix} \mathbf{A}^{(p)} \\ \mathbf{C}^{(p)} \end{pmatrix} + \beta_1\begin{pmatrix} \mathbf{0} \\ \mathbf{B}^{(1)} \end{pmatrix} + \cdots + \beta_q\begin{pmatrix} \mathbf{0} \\ \mathbf{B}^{(q)} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}. \tag{2.8.1}$$

所以 $\sum_{i=1}^{p} \alpha_i \mathbf{M}^{(i)} = \mathbf{0}$，由 $\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(p)}$ 是基即得 $\alpha_1 = \cdots = \alpha_p = 0$。再代回 (2.8.1) 式即得 $\sum_{j=1}^{q} \beta_j \mathbf{M}^{(j)} = \mathbf{0}$，于是 $\beta_1 = \cdots = \beta_q = 0$。所以 $\mathbf{M}^{(1)}, \cdots, \mathbf{M}^{(p)}, \mathbf{M}^{(n+1)}, \cdots, \mathbf{M}^{(n+q)}$ 线性无关。那么就有 $\operatorname{rank}(M) \geqslant p + q = \operatorname{rank}(A) + \operatorname{rank}(B)$。
特别地，当 $C = O$ 时，$\forall j \in \{1, 2, \cdots, n+l\}$，由于

$$\begin{aligned} \mathbf{M}^{(j)} &\in < \mathbf{M}^{(1)}, \cdots, \mathbf{M}^{(p)} > + < \mathbf{M}^{(n+1)}, \cdots, \mathbf{M}^{(n+q)} > \\ &= < \mathbf{M}^{(1)}, \cdots, \mathbf{M}^{(p)}, \mathbf{M}^{(n+1)}, \cdots, \mathbf{M}^{(n+q)} > . \end{aligned}$$

即 $\operatorname{rank}(M) \leqslant p + q$。于是此时只能有 $\operatorname{rank}(M) = \operatorname{rank}(A) + \operatorname{rank}(B)$。 □

例 2.8.3 用上面的引理证明定理 2.4.3(Sylvester 不等式)。

证明: 令 $M = \begin{pmatrix} AB & O \\ O & E_s \end{pmatrix}$，$R = \begin{pmatrix} O & A \\ -B & E_s \end{pmatrix}$，$P = \begin{pmatrix} E_m & A \\ O & E_s \end{pmatrix}$，$Q = \begin{pmatrix} E_n & O \\ -B & E_s \end{pmatrix}$，则 $R = PMQ$ 且 P, Q 满秩，注意到 $\operatorname{rank}(M) = \operatorname{rank}(AB) + s$，而由引理 2.8.4 有 $\operatorname{rank}(R) \geqslant \operatorname{rank}(A) + \operatorname{rank}(B)$，因此

$$\operatorname{rank}(AB) = \operatorname{rank}(M) - s = \operatorname{rank}(R) - s \geqslant \operatorname{rank}(A) + \operatorname{rank}(B) - s.$$

即得结论。 □

2.9 线性流形 (线性方程组解的结构)

在本章的最后，我们从线性映射的角度重新认识线性方程组。

引理 2.9.1 设 $V \subset \mathbb{R}^n$ 是 d 维子空间，则 V 是某个 n 元齐次方程组的解空间。

证明: 若 $V = \{\mathbf{0}\}$，则 $V = V_E$(记号出自定义 2.3.3)。否则，设 $\mathbf{v}_1, \cdots, \mathbf{v}_d$ 是 V 的一组基，则由基扩充定理，它们可以扩充成 $\mathbb{R}^n$ 的一组基 $\mathbf{v}_1, \cdots, \mathbf{v}_d, \mathbf{v}_{d+1}, \cdots, \mathbf{v}_n$。那么，由线性映射基本定理，存在 $\varphi: \mathbb{R}^n \to \mathbb{R}^n$ 满足

$$\varphi(\mathbf{v}_i) = \mathbf{0},\ i = 1, \cdots, d; \qquad \varphi(\mathbf{v}_j) = \mathbf{v}_j,\ j = d+1, \cdots, n.$$

于是 $\dim(\mathrm{im}(\varphi)) = n - d$ ($\mathrm{im}(\varphi) =< \mathbf{v}_{d+1}, \cdots, \mathbf{v}_n >$)，那么由对偶定理，$\dim(\ker(\varphi)) = d$，所以 $V = \ker(\varphi)$(由 φ 的构造显然有 $V \subset \ker(\varphi)$)。因此，设 A 是 φ 在 $\mathbb{R}^n$ 的标准基下的矩阵表示，则有 $V = V_A$。 □

定义 2.9.1 设 $\mathbf{v} \in \mathbb{R}^n$, $V \subset \mathbb{R}^n$ 是子空间，则我们称 $\mathbf{v} + V = \{\mathbf{v} + \mathbf{x} \mid \mathbf{x} \in V\}$ 是一个线性流形。

定理 2.9.1 $S \subset \mathbb{R}^n$ 是线性流形 $\Longleftrightarrow S$ 是 n 元线性方程组的解集。

证明: ($\Longleftarrow$) 只需注意到齐次线性方程组的解集是子空间，而由命题 1.2.1，非齐次线性方程组的解是特解加上齐次方程组的解。于是线性方程组的解集符合线性流形的定义。
($\Longrightarrow$) 设 $S = \mathbf{v} + V$，则由上面的引理，存在列数为 n 的矩阵使得 $V = V_A$。记 $\mathbf{w} = A\mathbf{v}$，则对 $\forall \mathbf{x} = \mathbf{v} + \mathbf{y} \in S$，有 $A\mathbf{y} = 0$，从而 $A\mathbf{x} = A\mathbf{v} = \mathbf{w}$，即 S 中的元素都是方程组 $A\mathbf{x} = \mathbf{w}$ 的解。 □

例 2.9.1 求解方程组 $\begin{cases} x_1 + x_2 - x_3 = 1 \\ x_1 - x_2 + x_3 = 2 \\ 2x_1 = 3 \end{cases}$.

解: 方程组的一个特解为 $\begin{cases} x_1 = \dfrac{3}{2} \\ x_2 = 1 \\ x_3 = \dfrac{3}{2} \end{cases}$，对应齐次方程组的解空间为 $< \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} >$。于是解流形为 $\left\{(\frac{3}{2}, 1, \frac{3}{2})^t + \lambda(0, 1, 1)^t \mid \lambda \in \mathbb{R}\right\}$。

定义 2.9.2 设 $\mathbf{v} + V$ 是 $\mathbb{R}^n$ 中的线性流形，如果 $\dim V = n - 1$，则称 $\mathbf{v} + V$ 是超平面 (hyperplane)。

推论 2.9.1 $P \subset \mathbb{R}^n$ 是超平面 $\Longleftrightarrow \exists \alpha_1, \cdots, \alpha_n \in \mathbb{R}$ 不全为 0 以及 $\beta \in \mathbb{R}$ 使得 P 是 $\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta$ 在 $\mathbb{R}^n$ 中的解集。

证明: ($\Longleftarrow$) 因为 $\alpha_1, \cdots, \alpha_n$ 不全为 0，故 $\mathrm{rank}(\alpha_1, \ldots, \alpha_n) = 1$，并且 $\mathrm{rank}(\alpha_1, \ldots, \alpha_n, \beta) = 1$，于是方程 $\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta$ 的解为 $\mathbf{v} + V$，其中 $\dim V = n - 1$。
($\Longrightarrow$) 设 $P = \mathbf{v} + V$, $\dim V = n - 1$，则由引理 2.9.1，存在列数为 n 的矩阵 A 使得 $V = V_A$，$\dim V_A = n - 1$，则由对偶定理，$\mathrm{rank}(A) = 1$。那么可设 $(\alpha_1, \ldots, \alpha_n)$ 是 A 的非零行，即 $V = V_{(\alpha_1, \cdots, \alpha_n)}$。则由定理 2.9.1 可得 P 是 $\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta$ 的解，其中 $\beta = (\alpha_1, \cdots, \alpha_n)\mathbf{v}$。 □

例 2.9.2 求过点 $(1, 0, 0)^t$, $(0, 1, 0)^t$, $(0, 0, 1)^t$ 的平面方程。

解: 设方程为 $\alpha_1 x_1 + \alpha_2 x_2 + \alpha_n x_n = \beta$，代入点得到 $\alpha_1 = \alpha_2 = \alpha_3 = \beta$，取 $\beta = 1$ 即得到该平面的一个方程为 $x_1 + x_2 + x_3 = 1$。