第二章　矩阵

1.3.4　子空间的生成

定义 1.31 设 S 是 $\mathbb{R}^n$ 的非空子集. 则由 S 中元素的所有线性组合构成的集合称为由 S 生成的子空间. 记为 $\langle S\rangle$. 集合 S 中的元素称为子空间 $\langle S\rangle$ 的一组生成元.

命题 1.32 设 S 是 $\mathbb{R}^n$ 的非空子集.

(i) $\langle S\rangle$ *是子空间;*

(ii) 设 U 是 $\mathbb{R}^n$ 的子空间且 $S \subset U$. 则 $\langle S\rangle \subset U$.

证明. (i) 设 $\mathbf{x}, \mathbf{y} \in \langle S\rangle$. 则存在 $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_\ell \in S$, $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_\ell \in \mathbb{R}$ 使得

$$\mathbf{x} = \sum_{i=1}^{k} \alpha_i \mathbf{u}_i \quad 和 \quad \mathbf{y} = \sum_{j=1}^{\ell} \beta_j \mathbf{v}_j.$$

则对任意 $\lambda, \mu \in \mathbb{R}$, 我们有

$$\lambda\mathbf{x}+\mu\mathbf{y} = \lambda\left(\sum_{i=1}^{k} \alpha_i \mathbf{u}_i\right)+\mu\left(\sum_{j=1}^{\ell} \beta_j \mathbf{v}_j\right) = \sum_{i=1}^{k}(\lambda\alpha_i)\mathbf{u}_i+\sum_{j=1}^{\ell}(\mu\beta_j)\mathbf{v}_j.$$

故 $\lambda\mathbf{x} + \mu\mathbf{y} \in \langle S\rangle$. 根据命题 1.18 $\langle S\rangle$ 是子空间.

(ii) 因为 $S \subset U$, 所以 $\langle S\rangle \subset U$ (命题 1.18). □

当 S 是有限集 $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ 时, $\langle S\rangle$ 也记作 $\langle \mathbf{v}_1, \ldots, \mathbf{v}_k\rangle$.

例 1.33 设 $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. 则

$$\langle \mathbf{v} \rangle = \{\lambda \mathbf{v} \mid \lambda \in \mathbb{R}\} \quad \text{和} \quad \langle \mathbf{u}, \mathbf{v} \rangle = \{\alpha \mathbf{u} + \beta \mathbf{v} \mid \alpha, \beta \in \mathbb{R}\}.$$

例 1.34 设 $\mathbf{e}_1, \mathbf{e}_2 \in \mathbb{R}^2$,

$$V_1 = \langle \mathbf{e}_1 \rangle,\ V_2 = \langle \mathbf{e}_2 \rangle \text{ 和 } V_3 = \langle \mathbf{e}_1 + \mathbf{e}_2 \rangle.$$

计算 $(V_1 + V_2) \cap V_3$ 和 $V_1 \cap V_3 + V_2 \cap V_3$.
解. 注意到 $V_1 + V_2 = \langle \mathbf{e}_1, \mathbf{e}_2 \rangle = \mathbb{R}^2$. 故 $(V_1 + V_2) \cap V_3 = V_3$. 而 $V_1 \cap V_3 = V_2 \cap V_3 = \{\mathbf{0}\}$. 我们有

$$V_1 \cap V_3 + V_2 \cap V_3 = \{\mathbf{0}\} + \{\mathbf{0}\} = \{\mathbf{0}\}.$$

例 1.35 设 U, W 是 $\mathbb{R}^n$ 中的子空间. 则 $U + W = \langle U \cup W \rangle$.
证明. 显然 $U \cup W \subset U + W$. 根据命题 *1.32 (ii)*,

$$\langle U \cup W \rangle \subset U + W.$$

反之, 设 $\mathbf{x} \in U + W$. 则存在 $\mathbf{u} \in U$ 和 $\mathbf{w} \in W$ 使得 $\mathbf{x} = \mathbf{u} + \mathbf{w}$. 于是, $\mathbf{x} \in \langle U \cup W \rangle$, 即 $U + W \subset \langle U \cup W \rangle$. $\square$

例 1.36 线性组合引理可重述为: 设 $\mathbf{v}_1, \ldots, \mathbf{v}_k; \mathbf{w}_1, \ldots, \mathbf{w}_\ell$ 是 $\mathbb{R}^n$ 中的两组向量. 设对任意 $j \in \{1, 2, \ldots, \ell\}$,

$$\mathbf{w}_j \in \langle \mathbf{v}_1, \ldots, \mathbf{v}_k \rangle.$$

如果 $\ell > k$, 则 $\mathbf{w}_1, \ldots, \mathbf{w}_\ell$ 线性相关.

1.3.5 子空间的直和

定义 1.37 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 1.38 设 U, V 是 $\mathbb{R}^n$ 中子空间. 则 $U + V$ 是直和当且仅当对任意 $\mathbf{x} \in U + V$, 存在唯一的 $\mathbf{u} \in U$, $\mathbf{v} \in V$ 使得

$$\mathbf{x} = \mathbf{u} + \mathbf{v}.$$

证明. 设 $U + V$ 是直和且 $\mathbf{x} \in U + V$. 设

$$\mathbf{x} = \mathbf{u} + \mathbf{v} = \mathbf{u}' + \mathbf{v}',$$

其中 $\mathbf{u}, \mathbf{u}' \in U$, $\mathbf{v}, \mathbf{v}' \in V$. 则

$$\mathbf{u} - \mathbf{u}' = \mathbf{v}' - \mathbf{v} \implies \mathbf{u} - \mathbf{u}' \in U \cap V.$$

因为 $U \cap V = \{\mathbf{0}\}$, 所以 $\mathbf{u} = \mathbf{u}'$. 进而 $\mathbf{v} = \mathbf{v}'$.

反之, 假设 $U + V$ 不是直和. 则存在非零向量 $\mathbf{x} \in U \cap V$. 故 $\mathbf{0} = \mathbf{0} + \mathbf{0} = \mathbf{x} + (-\mathbf{x})$. 与要求的唯一性矛盾. □

2 子空间的基底和维数

2.1 极大线性无关组

例 2.1 设 $\mathbf{u}_1, \ldots, \mathbf{u}_m \in \mathbb{R}^n$ 且 $m > n$. 则 $\mathbf{u}_1, \ldots, \mathbf{u}_m$ 线性相关.

证明. 因为 $\mathbf{u}_1, \ldots, \mathbf{u}_m$ 都是 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 的线性组合, 所以根据线性组合引理和 $m > n$ 可知, $\mathbf{u}_1, \ldots, \mathbf{u}_m$ 线性相关. □

定义 2.2 设 $S \subset \mathbb{R}^n$. 有限非空子集 $T \subset S$ 称为 S 的一个极大线性无关组. 如果

(i) T 中的向量线性无关;

(ii) 对任意 $\mathbf{v} \in S$, $\mathbf{v} \in \langle T \rangle$.

注解 2.3 如果 S 的一个极大线性无关组 $T = \{\mathbf{u}_1, \ldots, \mathbf{u}_d\}$. 我们经常略去集合符号简称 $\mathbf{u}_1, \ldots, \mathbf{u}_d$ 是 S 的一个极大线性无关组.

命题 2.4 设 $S \subset \mathbb{R}^n$.

(i) (可扩充性) 设 $\mathbf{u}_1, \ldots, \mathbf{u}_k$ 是 S 中线性无关的向量. 则存在 S 中极大线性无关组且它包含 $\mathbf{u}_1, \ldots, \mathbf{u}_k$.

(ii) (等势性) 设 $\mathbf{v}_1, \ldots, \mathbf{v}_\ell \in S$ 线性无关, $\mathbf{w}_1, \ldots, \mathbf{w}_m$ 是 S 的极大线性无关组. 则 $\mathbf{v}_1, \ldots, \mathbf{v}_\ell$ 是 S 的极大线性无关组当且仅当 $\ell = m$.

证明. (i) 如果对任意 $\mathbf{u} \in S$, $\mathbf{u}, \mathbf{u}_1, \ldots, \mathbf{u}_k$ 线性相关, 则 $\mathbf{u} \in \langle \mathbf{u}_1, \ldots, \mathbf{u}_k \rangle$ (第五讲命题 1.11 (iv)). 于是, $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ 是 S 中的一个极大线性无关组且它包含 $\mathbf{u}_1, \ldots, \mathbf{u}_k$. 否

则, 存在 $\mathbf{u}_{k+1} \in S$ 使得 $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}$ 线性无关. 对向量 $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}$ 重复上述推理过程, 我们要么得出

$$\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}\}$$

是 S 的极大线性无关组, 要么证明存在向量 $\mathbf{u}_{k+2} \in S$ 使得 $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \mathbf{u}_{k+2}$ 线性无关. 由例 2.1 可知，S 中不可能有 $n+1$ 个线性无关的向量. 故上述推理过程重复有限步后, 我们就会得到一个 S 中的极大线性无关组且它包含 $\mathbf{u}_1, \ldots, \mathbf{u}_k$.

(ii) 设 $\mathbf{v}_1, \ldots, \mathbf{v}_\ell$ 是 S 的极大线性无关组. 因为

$$\mathbf{w}_1, \ldots, \mathbf{w}_m \in \langle \mathbf{v}_1, \ldots, \mathbf{v}_\ell \rangle,$$

所以 $m \leq \ell$ (线性组合引理). 同理可知, $\ell \leq m$.

反之, 设 $\ell = m$. 由 (i) 可知, S 有一个极大线性无关组

$$\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_{m+s}$$

由极大线性无关组的定义可知 $\mathbf{v}_i \in \langle \mathbf{w}_1, \ldots, \mathbf{w}_m \rangle$, $i = 1, 2, \ldots, m+s$. 假设 $s > 0$. 根据线性组合引理, $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_{m+s}$ 线性相关, 矛盾. 故 $s = 0$. 即 $\mathbf{v}_1, \ldots, \mathbf{v}_m$ 是 S 的极大线性无关组. □

注解 2.5 由上述命题可知, 任何含有非零向量的集合必有极大线性无关组.

例 2.6 向量 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 $\mathbb{R}^n$ 的一组极大线性无关组.

例 2.7 证明 $\mathbf{e}_1+\mathbf{e}_2,\mathbf{e}_1-\mathbf{e}_2,\mathbf{e}_3,\ldots,\mathbf{e}_n$ 是 $\mathbb{R}^n$ 的一个极大线性无关组.

证明. 设 $\alpha_1,\alpha_2,\alpha_3,\ldots,\alpha_n\in\mathbb{R}$ 使得

$$\alpha_1(\mathbf{e}_1+\mathbf{e}_2)+\alpha_2(\mathbf{e}_1-\mathbf{e}_2)+\alpha_3\mathbf{e}_3+\cdots+\alpha_n\mathbf{e}_n=\mathbf{0}.$$

则

$$(\alpha_1+\alpha_2)\mathbf{e}_1+(\alpha_1-\alpha_2)\mathbf{e}_2+\alpha_3\mathbf{e}_3+\cdots+\alpha_n\mathbf{e}_n=\mathbf{0}.$$

因为 $\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$ 线性无关, 所以 $\alpha_1+\alpha_2=0$, $\alpha_1-\alpha_2=0$, $\alpha_3=0$, $\ldots$, $\alpha_n=0$. 故 $\alpha_1=\alpha_2=0$. 于是,

$$\mathbf{e}_1+\mathbf{e}_2,\mathbf{e}_1-\mathbf{e}_2,\mathbf{e}_3,\ldots,\mathbf{e}_n$$

线性无关. 由命题 *2.4 (ii)* 可知, $\mathbf{e}_1+\mathbf{e}_2,\mathbf{e}_1-\mathbf{e}_2,\mathbf{e}_3,\ldots,\mathbf{e}_n$ 是 $\mathbb{R}^n$ 的一个极大线性无关组.

2.2 子空间的基底

定义 2.8 设 $U\subset\mathbb{R}^n$ 是子空间且 $U\neq\{\mathbf{0}\}$. 向量集

$$\{\mathbf{u}_1,\ldots\mathbf{u}_d\}\subset U$$

称为 U 的一组基, 如果对任意 $\mathbf{u}\in U$, 存在唯一的

$$\alpha_1,\ldots,\alpha_d\in\mathbb{R},$$

使得 $\mathbf{u}=\alpha_1\mathbf{u}_1+\cdots+\alpha_d\mathbf{u}_d$.

如注解 2.3 所述, 当 $\{\mathbf{u}_1, \dots \mathbf{u}_d\}$ 是子空间 U 的一组基时, 我们也简称 $\mathbf{u}_1, \dots, \mathbf{u}_d$ 是 U 的一组基.

命题 2.9 设 $U \subset \mathbb{R}^n$ 是子空间且 $U \neq \{\mathbf{0}\}$. 向量集

$$\{\mathbf{u}_1, \dots \mathbf{u}_d\}$$

是 U 的一组基当且仅当该向量集是 U 的极大线性无关组.

证明. 设 $\{\mathbf{u}_1, \dots, \mathbf{u}_d\}$ 是 U 的基. 再设 $\alpha_1, \dots, \alpha_d \in \mathbb{R}$ 使得

$$\alpha_1 \mathbf{u}_1 + \cdots + \alpha_d \mathbf{u}_d = \mathbf{0}.$$

由基底定义可知, $\alpha_1 = \cdots = \alpha_d = \mathbf{0}$. 故 $\mathbf{u}_1, \dots, \mathbf{u}_d$ 线性无关. 再由基底定义中条件 (ii) 可知, $\mathbf{u}_1, \dots \mathbf{u}_d$ 是 U 的极大线性无关组.

反之, 设 $\mathbf{u}_1, \dots \mathbf{u}_d$ 是 U 的一个极大线性无关组. 根据第五讲命题 1.11 (iv), $\mathbf{u}_1, \dots \mathbf{u}_d$ 是 U 的一组基. □

由例 2.6 可知, $\mathbf{e}_1, \dots, \mathbf{e}_n$ 是 $\mathbb{R}^n$ 中的一组基. 称之为 $\mathbb{R}^n$ 的标准基.

推论 2.10 设 $S \subset \mathbb{R}^n$ 包含非零向量. 则 S 中的极大线性无关组是 $\langle S \rangle$ 的一组基.

证明. 设 $\mathbf{u}_1, \dots, \mathbf{u}_d$ 是 S 的一个极大线性无关组. 则

$$S \subset \langle \mathbf{u}_1, \dots, \mathbf{u}_d \rangle.$$

由第五讲命题 1.25, $\langle S\rangle \subset \langle \mathbf{u}_1, \ldots, \mathbf{u}_d\rangle$. 而另一个方向的包含关系是显然的. 于是, $\langle S\rangle = \langle \mathbf{u}_1, \ldots, \mathbf{u}_d\rangle$. 即 $\mathbf{u}_1, \ldots, \mathbf{u}_d$ 是 $\langle S\rangle$ 的极大线性无关组. 根据上述命题, $\mathbf{u}_1, \ldots, \mathbf{u}_d$ 是 $\langle S\rangle$ 的一组基. □

例 2.11 设

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \mathbf{z} = \begin{pmatrix} 5 \\ 1 \\ 6 \end{pmatrix}.$$

求 $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}\rangle$ 的一组基.

解. 根据上述推论, 我们只要求出 $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ 中的一个极大线性无关组即可.

$$(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ 3 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 \\ 0 & -3 \\ 0 & -3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 \\ 0 & -3 \\ 0 & 0 \end{pmatrix}.$$

于是, 以矩阵 $(\mathbf{x}, \mathbf{y})$ 为系数矩阵的齐次线性方程组只有平凡解. 故 $\mathbf{x}, \mathbf{y}$ 线性无关. 由第二章第一讲例 *1.9*, $\mathbf{x}, \mathbf{y}, \mathbf{z}$ 线性相关. 故 $\mathbf{x}, \mathbf{y}$ 是 $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}\rangle$ 的一组基. 类似地可知, $\mathbf{x}, \mathbf{z}$ 和 $\mathbf{y}, \mathbf{z}$ 都是 $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}\rangle$ 的基

定理 2.12 *(基扩充定理)* 设 $\mathbf{u}_1, \ldots, \mathbf{u}_k$ 是子空间 U 中线性无关的向量. 则 U 有一组基包含 $\mathbf{u}_1, \ldots, \mathbf{u}_k$.

证明. 根据命题 2.4 (i), U 中有极大线性无关组包含 $\mathbf{u}_1, \ldots, \mathbf{u}_k$. 再根据上述命题, 该极大线性无关组是 U 的一组基. □

2.3 维数

定义 2.13 设 U 是 $\mathbb{R}^n$ 中的子空间. 如果 $\mathbf{u}_1, \ldots, \mathbf{u}_d$ 是 U 的一组基, 则 U 的维数*(dimension)*定义为 d. 当 $U = \{\mathbf{0}\}$ 时, 其维数定义为 0. 该维数记为 $\dim(U)$.

根据命题 2.4 (ii) 和命题 2.9, 子空间维数是良定义的. 因为坐标空间 $\mathbb{R}^n$ 有标准基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$, 所以 $\dim(\mathbb{R}^n) = n$.

命题 2.14 设 U, W 是 $\mathbb{R}^n$ 的两个子空间且 $U \subset W$. 则 $\dim(U) \leq \dim(W)$. 进而, $U = W$ 当且仅当 $\dim(U) = \dim(W)$.

证明. 设 $\mathbf{u}_1, \ldots, \mathbf{u}_d$ 是 U 的一组基. 由假设 $U \subset W$ 和基扩充定理 2.12, W 有一组基包含 $\mathbf{u}_1, \ldots, \mathbf{u}_d$. 故 $\dim(U) \leq \dim(W)$.

再设 $d = \dim(W)$. 则 $\mathbf{u}_1, \ldots, \mathbf{u}_d$ 也是 W 的基 (命题 2.4 (ii)). 故 $W \subset \langle u_1, \ldots, u_d \rangle = U$. □

命题 2.15 设 U, W 是 $\mathbb{R}^n$ 的两个子空间. 则

$$\dim(U + W) + \dim(U \cap W) = \dim(U) + \dim(W).$$

证明. 令 $\mathbf{v}_1, \ldots, \mathbf{v}_k$ 是 $U \cap W$ 的一组基. 由基扩充定理, U 有基底 $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_{k+\ell}$; W 有基底 $\mathbf{v}_1, \ldots, \mathbf{v}_k$, $\mathbf{w}_{k+1}, \ldots, \mathbf{w}_{k+m}$.

断言. $S=\{\mathbf{v}_1,\dots,\mathbf{v}_k,\mathbf{u}_{k+1},\dots,\mathbf{u}_{k+\ell},\mathbf{w}_{k+1},\dots,\mathbf{w}_{k+m}\}$ 是 $U+W$ 的一组基.

断言的证明. 设 $\alpha_1,\dots,\alpha_k,\beta_1,\dots,\beta_\ell,\gamma_1,\dots,\gamma_m\in\mathbb{R}$ 使得

$$\alpha_1\mathbf{v}_1+\cdots+\alpha_k\mathbf{v}_k+\beta_1\mathbf{u}_{k+1}+\cdots+\beta_\ell\mathbf{u}_{k+\ell}+\underbrace{\gamma_1\mathbf{w}_{k+1}+\cdots+\gamma_m\mathbf{w}_{k+m}}_{\mathbf{w}}=\mathbf{0}.$$

则 $\mathbf{w}\in U\cap W$. 于是, 存在 $\lambda_1,\dots,\lambda_k\in\mathbb{R}$ 使得

$$\mathbf{w}=\lambda_1\mathbf{v}_1+\cdots+\lambda_k\mathbf{v}_k.$$

即

$$(-\gamma_1)\mathbf{w}_1+\cdots+(-\gamma_m)\mathbf{w}_m+\lambda_1\mathbf{v}_1+\cdots+\lambda_k\mathbf{v}_k=\mathbf{0}.$$

因为 $\mathbf{w}_1,\dots,\mathbf{w}_m,\mathbf{v}_1,\dots,\mathbf{v}_k$ 线性无关, 所以

$$\gamma_1=\cdots=\gamma_m=0.$$

故

$$\alpha_1\mathbf{v}_1+\cdots+\alpha_k\mathbf{v}_k+\beta_1\mathbf{u}_{k+1}+\cdots+\beta_\ell\mathbf{u}_{k+\ell}=\mathbf{0}.$$

因为 $\mathbf{u}_1,\dots,\mathbf{u}_\ell,\mathbf{v}_1,\dots,\mathbf{v}_k$ 线性无关, 所以 $\alpha_1=\cdots=\alpha_k=\beta_1=\cdots=\beta_\ell=0$. 于是, S 中的向量线性无关.

再设 $\mathbf{x}\in U+W$. 则存在 $\mathbf{y}\in U$ 和 $\mathbf{z}\in W$ 使得 $\mathbf{x}=\mathbf{y}+\mathbf{z}$. 因为 $\mathbf{y}$ 是 $\mathbf{v}_1,\dots,\mathbf{v}_k,\mathbf{u}_{k+1},\dots,\mathbf{u}_{k+\ell}$ 的线性组合, $\mathbf{z}$ 是 $\mathbf{v}_1,\dots,\mathbf{v}_k,\mathbf{w}_{k+1},\dots,\mathbf{w}_{k+m}$ 的线性组合. 故 $\mathbf{x}$ 是 S

中的向量的线性组合. 即 S 是 $U+W$ 的一个极大线性无关组. 根据命题 2.9, 断言成立.

由断言可知 $\dim(U+W)=k+\ell+m$. 因为

$$\dim(U\cap W)=k,\ \dim(U)=k+\ell,\ \dim(W)=k+m,$$

所以

$$\dim(U+W)+\dim(U\cap W)=\dim(U)+\dim(W). \quad \square$$

注解 2.16 在上述命题证明中, 当 $U\cap W=\{\mathbf{0}\}$, 则集合 $\{\mathbf{v}_1,\ldots,\mathbf{v}_k\}$ 是空集, 即 $k=0$.

例 2.17 设 $U,W\subset\mathbb{R}^n$ 是子空间, $\dim(U)=d>0$, $\dim(W)=n-1$. 证明 $\dim(U\cap W)\geq d-1$.

证明. 由维数公式可知:

$$\begin{aligned}\dim(U\cap W)&=\dim(U)+\dim(W)-\dim(U+W)\\&\geq d+n-1-n \quad (\because\ \dim(U+W)\leq n)\\&=d-1.\end{aligned}$$

例 2.18 设 $U,W\subset\mathbb{R}^n$ 是子空间. 证明 $V+W$ 是直和当且仅当 $\dim(V+W)=\dim(V)+\dim(W)$.

证明. 由维数公式可知

$$\begin{aligned}\dim(V+W)=\dim(V)+\dim(W)&\\ \iff \dim(U\cap W)=0&\\ \iff U\cap W=\{\mathbf{0}\}.&\quad \square\end{aligned}$$

3 矩阵的秩

3.1 初等行变换下的不变量

计算过程(算法)中不变的量往往反映被计算对象的基本特征. 本小节研究关于矩阵初等行变换下的不变量.

定义 3.1 设 $A \in \mathbb{R}^{m\times n}$. 由 A 的行向量 $\vec{A}_1, \ldots, \vec{A}_m$ 在行空间 $\mathbb{R}^{1\times n}$ 中生成的子空间称为 A 的行空间. 记为 $V_r(A)$. 由 A 的列向量 $\vec{A}^{(1)}, \ldots, \vec{A}^{(n)}$ 在列空间 $\mathbb{R}^m$ 中生成的子空间称为 A 的列空间. 记为 $V_c(A)$. 矩阵 A 的行秩是 $\dim(V_r(A))$, 列秩是 $\dim(V_c(A))$.

本节的主要结论是一个矩阵的行秩等于列秩.

引理 3.2 设 $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{m\times n}$ 是通过一次初等行变换得到的. 则 $V_r(A) = V_r(B)$. 特别地, A 和 B 的行秩相同.

证明. 当 B 是由 A 通过一次第一类初等行变换得到的, 则 A 和 B 有共同的行向量. 故 $V_r(A) = V_r(B)$.

当 B 是由 A 通过一次第二类初等行变换得到的, 设 $\vec{A}_k = \vec{B}_k$, $k = 1, \ldots, j-1, j+1, \ldots, m$, 而 $\vec{B}_j = \vec{A}_j + \lambda \vec{A}_i$, 其中 $i \neq j$ 和 $\lambda \in \mathbb{R}$. 则 B 的所有行向量在 $V_r(A)$ 中. 故 $V_r(B) \subset V_r(A)$ (第二章第一讲命题 1.25 (ii)). 因为 $i \neq j$, 所以 $\vec{A}_i = \vec{B}_i$. 故 $\vec{A}_j = \vec{B}_j - \lambda \vec{B}_i$. 由此可知, A 的行向量都在 $V_r(B)$ 中. 同样的命题蕴含 $V_r(A) \subset V_r(B)$.

当 B 是由 A 通过一次第三类初等行变换得到的, 设 $\vec{A}_k = \vec{B}_k$, $k = 1, \ldots, j-1, j+1, \ldots, m$, 而 $\vec{B}_j = \lambda \vec{A}_j$, 其中 $\lambda \in \mathbb{R} \setminus \{0\}$. 则 B 的所有行向量在 $V_r(A)$ 中. 故 $V_r(B) \subset V_r(A)$ (第二章第一讲命题 1.25 (ii)). 因为 $\lambda \neq 0$, 所以 $\vec{A}_j = \lambda^{-1} \vec{B}_j$. 故 A 的行向量都在 $V_r(B)$ 中. 同样的命题蕴含 $V_r(A) \subset V_r(B)$. $\square$

例 3.3 *第二类初等行变换得:*

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} =: B.$$

注意到 $V_c(A) \neq V_c(B)$.

一个令人意外得结论是:

引理 3.4 *设 $A \in \mathbb{R}^{m \times n}$. 设 $B \in \mathbb{R}^{m \times n}$ 是通过一次初等行变换得到的. 则 A 和 B 的列秩相同.*

证明. 设以 A 和 B 为系数矩阵的齐次线性方程组分别是 H_A 和 H_B. 则 H_A 和 H_B 等价 (第一章第一讲命题 2.2). 不妨设 $\vec{A}^{(1)}, \ldots, \vec{A}^{(r)}$ 是 A 中列向量的极大线性无关组.

断言. $\vec{B}^{(1)}, \ldots, \vec{B}^{(r)}$ 线性无关.

断言的证明. 设 $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$ 使得

$$\alpha_1 \vec{B}^{(1)} + \cdots + \alpha_r \vec{B}^{(r)} = \mathbf{0}_m.$$

则 $\alpha_1\vec{B}^{(1)}+\cdots+\alpha_r\vec{B}^{(r)}+0\vec{B}^{(r+1)}+\cdots+0\vec{B}^{(n)}=\mathbf{0}_m$. 换言之, 列向量

$$\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_r\\ 0\\ \vdots\\ 0\end{pmatrix}\in\mathbb{R}^n$$

是方程组 H_B 的一个解. 故它也是 H_A 的一个解. 即

$$\alpha_1\vec{A}^{(1)}+\cdots+\alpha_r\vec{A}^{(r)}+0\vec{A}^{(r+1)}+\cdots+0\vec{A}^{(n)}=\mathbf{0}_m.$$

我们得到 $\alpha_1\vec{A}^{(1)}+\cdots+\alpha_r\vec{A}^{(r)}=\mathbf{0}_m$. 因为 $\vec{A}^{(1)}$, $\ldots$, $\vec{A}^{(r)}$ 线性无关, 所以 $\alpha_1=\cdots=\alpha_r=0$. 于是, $\vec{B}^{(1)},\ldots,\vec{B}^{(r)}$ 线性无关. 断言成立.

设 $j\in\{r+1,\ldots,n\}$. 则存在 $\beta_1,\ldots,\beta_r\in\mathbb{R}$ 使得 $\vec{A}^{(j)}=\beta_1\vec{A}^{(1)}+\cdots+\beta_r\vec{A}^{(r)}$. 这是因为 $\vec{A}^{(1)}$, $\ldots$, $\vec{A}^{(r)}$ 是 A 中列向量的极大线性无关组. 我们得到

$$\begin{aligned}&\beta_1\vec{A}^{(1)}+\cdots+\beta_r\vec{A}^{(r)}\\ &+0\vec{A}^{(r+1)}+\cdots+0\vec{A}^{(j-1)}+(-1)\vec{A}^{(j)}\\ &+0\vec{A}^{(j+1)}+\cdots+0\vec{A}^{(n)}=\mathbf{0}_m.\end{aligned}$$

换言之, 列向量

$$j \to \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_r \\ 0 \\ \vdots \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n$$

是 H_A 的解. 于是, 它也是 H_B 的解. 我们有

$$\vec{B}^{(j)} = \beta_1 \vec{B}^{(1)} + \cdots + \beta_r \vec{B}^{(r)}.$$

再根据断言可知 $\vec{B}^{(1)}, \ldots, \vec{B}^{(r)}$ 是 B 中列向量的极大线性无关组. 由推论 2.10 可知,

$$\dim(V_c(A)) = r = \dim(V_c(B)). \quad \square$$

在矩阵中互换两列的位置称为第一类初等列变换, 把一列通乘一个实数加到另一列上称为第二类初等列变换, 把一列通乘一个非零实数称为第三类初等列变换. 它们统称为初等列变换. 类似引理 3.2，我们有

引理 3.5 设 $A\in\mathbb{R}^{m\times n}$. 设 $B\in\mathbb{R}^{m\times n}$ 是通过一次初等列变换得到的. 则 $V_c(A)=V_c(B)$. 特别地, A 和 B 的列秩相同.

3.2 矩阵的秩

定理 3.6 设 $A \in \mathbb{R}^{m\times n}$. 则 A 的行秩等于它的列秩.

证明. 由第一章第一讲命题 2.3 可知, A 可以通过初等行变换化为阶梯型矩阵

$$B = \begin{pmatrix} 0 & \cdots & 0 & \bullet & * & \cdots & * & * & * & * & \cdots & * & * & * & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \bullet & * & * & \cdots & * & * & * & \cdots & * \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & \bullet & * & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

其中 $\bullet$ 代表非零实数, $*$ 代表实数.

设 B 中有 k 行是非零向量. 注意到这 k 行从左至右第一个非零坐标出现的位置两两不同. 故这 k 行线性无关. 由此得出, 这 k 行是行空间 $V_r(B)$ 的一组基. 我们得到 B 的行秩等于 k.

对 B 做第二类列变换得到

$$C=\begin{pmatrix}
0 & \cdots & 0 & \bullet & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0\\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \bullet & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0\\
\vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \vdots\\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & \bullet & 0 & \cdots & 0\\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0\\
\vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \vdots\\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0
\end{pmatrix}.$$

注意到 C 中只有 k 列非零, 且这 k 列从上到下第一个非零坐标出现的位置两两不同. 故这 k 列线性无关. 由此得出, 这 k 列是列空间 $V_c(C)$ 的一组基. 我们有 C 的列秩等于 k.

根据引理 3.5, B 的列秩也等于 k. 故 B 的行秩和列秩相等. 再根据引理 3.2 和 3.4, A 的行秩和列秩相等. □

定义 3.7 矩阵 $A\in\mathbb{R}^{m\times n}$ 行秩称为它的 *秩*. 记为 $\operatorname{rank}(A)$.

由上述定理可知, 矩阵的秩既是它的行秩也是它的列秩.

对矩阵的初等行变换和列变换统称矩阵的初等变换.

推论 3.8 *设 B 是实矩阵 A 通过有限次初等变换得到的. 则 $\operatorname{rank}(A)=\operatorname{rank}(B)$.*

证明. 设 B 是实矩阵 A 通过一次初等行（列）变换得到的. 根据上一讲引理 3.3 (引理 3.5) 可知, A 和 B 的行（列）秩相同. 根据定理 3.6, 它们的秩相同. □

例 3.9 设

$$A = \begin{pmatrix} 1 & 0 & 4 & 5 \\ 2 & 1 & -1 & 3 \\ 4 & 1 & 7 & 13 \end{pmatrix}.$$

计算 $\operatorname{rank}(A)$.

解. 利用初等行变换

$$A \longrightarrow \begin{pmatrix} 1 & 0 & 4 & 5 \\ 0 & 1 & -9 & -7 \\ 0 & 1 & -9 & -7 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 4 & 5 \\ 0 & 1 & -9 & -7 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

于是, $\operatorname{rank}(A) = 2$.

另解. 利用初等列变换

$$A \longrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & -9 & -7 \\ 4 & 1 & -9 & -7 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 4 & 1 & 0 & 0 \end{pmatrix}.$$

于是, $\operatorname{rank}(A) = 2$.

例 3.10 设 $A \in \mathbb{R}^{m\times n}$. 证明*:* $\operatorname{rank}(A) \le \min(m, n)$.

证明. 因为 $V_r(A)$ 是 $\mathbb{R}^{1\times n}$ 的子空间, 所以 $\dim(V_r(A)) \le n$ *(*上一讲命题 *2.14).* 故 $\operatorname{rank}(A) \le n$. 因为 $V_c(A)$ 是 $\mathbb{R}^m$ 的子空间, 所以 $\dim(V_c(A)) \le m$ *(*上一讲命题 *2.14).* 故 $\operatorname{rank}(A) \le m$. □

定义 3.11 设 $A \in \mathbb{R}^{m\times n}$. 如果 $\operatorname{rank}(A) = m$, 则称 A 是行满秩的. 如果 $\operatorname{rank}(A) = n$, 则称 A 是列满秩的. 当 $m = n$ 且 $\operatorname{rank}(A) = n$ 时, 称 A 是满秩的.

例 3.12 设 $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{m\times k}$. 令

$$(A, B) = (\vec{A}^{(1)}, \dots, \vec{A}^{(n)}, \vec{B}^{(1)}, \dots, \vec{B}^{(k)}) \in \mathbb{R}^{m\times(n+k)}.$$

证明: $\operatorname{rank}((A, B)) \le \operatorname{rank}(A) + \operatorname{rank}(B)$.

证明. 设 $C = (A, B)$. 不妨设 $\vec{A}^{(1)}, \dots, \vec{A}^{(n)}$ 的一个极大线性无关组是 $\vec{A}^{(1)}, \dots, \vec{A}^{(s)}$; $\vec{B}^{(1)}, \dots, \vec{B}^{(k)}$ 的一个极大线性无关组是 $\vec{B}^{(1)}, \dots, \vec{B}^{(t)}$. 则矩阵 C 的列向量在

$$V = \langle \vec{A}^{(1)}, \dots, \vec{A}^{(s)}, \vec{B}^{(1)}, \dots, \vec{B}^{(t)} \rangle$$

中. 于是, $V_c(C) \subset V$ (第二章第一讲命题 1.25). 故 $\operatorname{rank}(C) \le \dim(V) \le s + t$. $\square$

另证. 注意到 $V_c(C) = V_c(A) + V_c(B)$. 由子空间的维数公式可知:

$$\begin{aligned}
\dim(V_c(C)) &= \dim(V_c(A) + V_c(B)) \\
&= \dim(V_c(A)) + \dim(V_c(B)) - \dim(V_c(A) \cap V_c(B)) \\
&\le \dim(V_c(A)) + \dim(V_c(B)).
\end{aligned}$$

故 $\operatorname{rank}(C) \le \operatorname{rank}(A) + \operatorname{rank}(B)$. $\square$

定义 3.13 设 $A \in \mathbb{R}^{m\times n}$. 矩阵 A 的**转置**是在 $\mathbb{R}^{n\times m}$ 中的矩阵. 它在第 j 行, 第 i 列处的元素等于 A 在第 i 行, 第 j 列处的元素, 其中 $i = 1, \ldots, m$, $j = 1, \ldots, n$. 矩阵 A 的转置记为 A^t.

例 3.14 设

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix} \Longrightarrow A^t = \begin{pmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{pmatrix}.$$

可直接验证 $(A^t)^t = A$.

例 3.15 设 $A \in \mathbb{R}^{m\times n}$. 则 $\mathrm{rank}(A) = \mathrm{rank}(A^t)$.

证明. 不妨设 $\vec{A}_1, \ldots, \vec{A}_r$ 是 A 的行向量中的一个极大线性无关组. 则 $\mathrm{rank}(A) = r$. 可直接验证 $\vec{A}_1^t, \ldots, \vec{A}_r^t$ 是 A^t 中列向量的一个极大线性无关组. 故 $\mathrm{rank}(A^t) = r$.