5.4 基与维数

回忆. (i) F^n 的非空子集 V 是一个**子空间**当且仅当对任意正整数 m, 对任意数组向量 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m \in V$ 和任意标量 $\lambda_1, \ldots, \lambda_m \in F$, 都有线性组合 $\sum_{i=1}^m \lambda_i \boldsymbol{\alpha}_i \in V$.

(ii) F^n 的子空间 V 是由向量组 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m \in F^n$ 生成的 (张成的) 当且仅当 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m \in V$ 并且任意的数组向量 $\boldsymbol{a} \in V$ 都是 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m$ 的线性组合, 即, 存在标量 $\lambda_1, \ldots, \lambda_m \in F$ 使得 $\boldsymbol{a} = \sum_{i=1}^m \lambda_i \boldsymbol{\alpha}_i$. 此时记作 $V = \langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m \rangle$.

引理 5.4.1. 设 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_i$ 是 F^n 中线性无关的一组向量, 而 $\boldsymbol{\alpha}_{i+1} \in F^n \setminus \langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_i \rangle$, 则 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_i, \boldsymbol{\alpha}_{i+1}$ 也是一组线性无关的向量.

证明. 用反证法和教材课后习题 #13. □

下面的定理说明 F^n 的任何子空间都是某一组向量生成的.

定理 5.4.2. 设非空集合 V 是 F^n 的子空间, 则存在线性无关的向量组 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 使得

$$V = \langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r \rangle .$$

证明. 若 V 是零空间, 上面的 $r = 0$.

若 V 不是零空间, 先任意选取非零向量 $\boldsymbol{\alpha}_1 \in V$. 考虑如下递归的选取方法: 假设已经选出了线性无关的向量 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_i \in V$, 其中 $i \geq 1$. 若 $\langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_i \rangle \subsetneq V$, 则任意选取 $\boldsymbol{\alpha}_{i+1} \in V \setminus \langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_i \rangle$. 由上面的引理 5.4.1 可知, $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_i, \boldsymbol{\alpha}_{i+1}$ 是线性无关的向量组.

另一方面, 由推论 5.2.11(1), F^n 中线性无关的向量组的长度最多为 n. 这说明上述的递归选取的方法无法无限地继续下去. 从而存在非负整数 $r \leq n$, 使得 $V = \langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r \rangle$. □

注 5.4.3. 定理中的线性无关的有序向量组 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 的选取一般并不唯一.

引理 5.4.4. 对于向量组 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \in F^n$, 以下几条等价:

(1) $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ 线性无关;

(2) 对任意的 $\boldsymbol{b} \in \langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \rangle$, $\boldsymbol{b}$ 可以由 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ 唯一地线性表示出来;

(3) 存在 $\boldsymbol{b} \in \langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \rangle$, $\boldsymbol{b}$ 可以由 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ 唯一地线性表示出来.

证明. 对于 (1) $\Rightarrow$ (2), 任取 $\boldsymbol{b} \in \langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \rangle$. 假定我们同时有 $\boldsymbol{b} = \sum_i \lambda_i \boldsymbol{a}_i$ 和 $\boldsymbol{b} = \sum_i \lambda'_i \boldsymbol{a}_i$, 其中 $\lambda_i, \lambda'_i \in F$. 于是, $\mathbf{0} = \boldsymbol{b} - \boldsymbol{b} = \sum_i (\lambda_i - \lambda'_i) \boldsymbol{a}_i$. 由于 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ 线性无关, 这说明对任何的 i 都有 $\lambda_i - \lambda'_i = 0$. 这说明了线性表示的唯一性.

(2) $\Rightarrow$ (3) 是显然的.

对于 (3) $\Rightarrow$ (1), 依条件, 假设 $\boldsymbol{b} = \sum_{i=1}^m \lambda_i \boldsymbol{a}_i \in \langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \rangle$. 用反证法, 反设 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ 线性相关, 则存在不全为零的 $k_1, \ldots, k_n \in F$ 使得 $\sum_{i=1}^m k_i \boldsymbol{a}_i = \mathbf{0}$. 则此时 $\boldsymbol{b} = \sum_{i=1}^m (\lambda_i + k_i) \boldsymbol{a}_i$ 给出了 $\boldsymbol{b}$ 另外一个不同的关于 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ 的线性表示. 这与 (3) 的假设相矛盾. □

定义 5.4.5. 对于 F^n 的子空间 V, 若存在线性无关的向量组 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r \in V$ 满足 $V = \langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r \rangle$, 则称该向量组构成 V 的一组**基**. 由引理 5.4.4 可知, 这等价于说 V 中的任意向量 $\boldsymbol{x}$ 都可以唯一地由 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 线性表示出来:

$$\boldsymbol{x} = \sum_{i=1}^r \lambda_i \boldsymbol{\alpha}_i.$$

我们称数组 $(\lambda_1, \ldots, \lambda_r) \in F^r$ 是 $\boldsymbol{x}$ 在基 $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r\}$ 下的**坐标向量** (或简称为**坐标**).

注 5.4.6. 在上述定义中,

(1) 基 $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r\}$ 是 V (无限多个向量构成的向量组) 的一个极大无关组 (这意味着 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 是线性无关的, 而任取 $\boldsymbol{\alpha}_{r+1} \in V$ 后, $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{r+1}$ 是线性相关的).

(2) 若 $\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_t\}$ 是 V 的另外一组基, 则 $t \leq n$, 且 $\langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r \rangle = \langle \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_t \rangle$, 从而向量组 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 与 $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_t$ 等价. 而它们都是线性无关的向量组, 从而由推论 5.3.15 知 $r = t$. 这说明该线性子空间的基的长度不依赖于基的具体选取. 这个公共的长度 r 称为子空间 V 的**维数**, 记作 $\dim(V) = r$. 依定义, 这是生成 V 的任何线性无关的向量组的长度; 按照前面推广后的定义, 这也是无限长度的向量组 V 的极大无关组的长度. 不难看出, $\dim(V) \leq n$.

(3) F^n 中一组向量 $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r\}$ 的任一极大无关组都构成了子空间 $\langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r \rangle$ 的一组基, 特别地, 该向量组的秩, 即其极大无关组的公共长度, 等于它们生成的子空间 $\langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r \rangle$ 的维数.

(4) 由此可知, 若 $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_r \in F^n$ 也生成了线性空间 V, 其中 $r = \dim(V)$, 则 $\operatorname{rank}(\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_r) = r$. 由此, 由定理 5.3.22(1)可知, $\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_r$ 线性无关, 从而是 V 的一组基.

注 5.4.7. (1) 用一 (有限长度的) 向量组来生成线性子空间 V 时, 从生成集中删除冗余向量的操作, 在余下的集合变成线性无关时必须停止. 如果再多删一个向量, 该向量将不是剩下向量的线性组合, 从而这个较小的集合将不再生成 V. 所以, 基是一个尽可能小的生成集.

(2) 若 S 是 V 的一组基, 在 S 中再添加进一个新的向量, 比如是从 V 中取的一个 $\boldsymbol{w}$, 则新的集合不再是线性无关了, 这是因为 S 生成 V, 因此 $\boldsymbol{w}$ 是 S 中元素的线性组合. 所以, 基还是尽可能大的线性无关集.

例 5.4.8. 回忆: F^n 有一组**基本向量** $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$, 其中 $\boldsymbol{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, 而 1 在 $\boldsymbol{e}_i$ 的第 i 个位置上. 容易验证: $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ 构成了 F^n 的一组基, 称为**自然基**或**标准基**. 对于任何的 $\boldsymbol{b} = (b_1, \ldots, b_n) \in F^n$, 有唯一的线性表示: $\boldsymbol{b} = \sum_{i=1}^n b_i \boldsymbol{e}_i$. 特别地, F^n 作为 F 上的线性空间的维数为 n.

例 5.4.9. (1) 在 $\mathbb{R}^2$ 中, 向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 不共线 $\Leftrightarrow$ $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 线性无关 $\Leftrightarrow$ $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 是 $\mathbb{R}^2$ 的一组基. $\mathbb{R}^2$ 作为自身的平凡的子空间是 2 维的.

(2) 在 $\mathbb{R}^3$ 中, 向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 不共面 $\Leftrightarrow$ $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 线性无关 $\Leftrightarrow$ $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 是 $\mathbb{R}^3$ 的一组基. $\mathbb{R}^3$ 作为自身的平凡的子空间是 3 维的. 例如, 由于

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

向量组 $\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$ 构成了 $\mathbb{R}^3$ 的一组基.

例 5.4.10. (a) 矩阵 $\boldsymbol{A}$ 的行向量组的任何一个极大无关组都构成了行空间 $\mathrm{Row}(\boldsymbol{A})$ 的一组基. 若 $\boldsymbol{A}$ 通过一系列初等行变换后得到矩阵 $\boldsymbol{B}$, 不难看出行空间 $\mathrm{Row}(\boldsymbol{A}) = \mathrm{Row}(\boldsymbol{B})$. 此时, 若 $\boldsymbol{B}$ 是阶梯形矩阵, 则它的非零行的向量也显然构成了这个线性子空间的一组基.

(b) 若 $\boldsymbol{A}$ 通过一系列初等行变换后得到矩阵 $\boldsymbol{B}$, 一般而言, $\mathrm{Col}(\boldsymbol{A}) \neq \mathrm{Col}(\boldsymbol{B})$. 例如, 我们可以通过 $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ 和 $\boldsymbol{B} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ 看到这一点.

很多应用问题可以通过从一个坐标系转化为另一坐标系, 使得问题简化. 在一个向量空间里, 转换坐标系其实和从一组基转换为另一组基在本质上是相同的. 因此, 接下来, 我们来讨论与基的转换相关的问题.

若 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 是 F^n 中子空间 V 的一组基, $\boldsymbol{b}\in V$, 则存在唯一的一组标量 $\lambda_i\in F$, $1\le i\le r$, 使得 $\boldsymbol{b}=\sum_{i=1}^r\lambda_i\boldsymbol{\alpha}_i$. 若将所有的数组向量看成列向量, 则该表达式可以写成

$$\boldsymbol{b}=\begin{pmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r\end{pmatrix}\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_r\end{pmatrix}.$$

显然, $\boldsymbol{\lambda}=(\lambda_1,\lambda_2,\dots,\lambda_r)^{\mathsf{T}}$ 是 $\boldsymbol{b}$ 在 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_r$ 下的坐标向量.

在接下来的讨论中, 除非特别说明, 我们总把数组向量看成列向量.

设 $V\subseteq F^n$ 有两组不同的基: $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 和 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_r$. 接下来讨论相应的坐标的变换. 由于 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_r$ 可以由 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 线性表示, 存在标量 $t_{ji}\in F$, 使得 $\boldsymbol{\beta}_i=\sum_{j=1}^r t_{ij}\boldsymbol{\alpha}_j$, $1\le i\le r$. 注意 t 的下标的顺序! 用矩阵的形式写出, 即有

$$\begin{pmatrix}\boldsymbol{\beta}_1 & \cdots & \boldsymbol{\beta}_r\end{pmatrix}=\begin{pmatrix}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_r\end{pmatrix}\underbrace{\begin{pmatrix}t_{11} & \cdots & t_{r1}\\ \vdots & & \vdots\\ t_{1r} & \cdots & t_{rr}\end{pmatrix}}_{\boldsymbol{T}},$$

其中 $\boldsymbol{T}\in F^{r\times r}$ 称为从 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 到 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_r$ 的**过渡矩阵**或**转移矩阵** (transition matrix).

注 5.4.11. (1) *在有些国外的教材里, 从 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 到 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_r$ 的***转移矩阵***指的是矩阵 $\boldsymbol{T}^{-1}$; 其出发点是下面马上要提到的坐标转换公式. 请在具体阅读时仔细辨别.*

(2) *在上面的讨论中, 设有矩阵 $\boldsymbol{T}'=(t'_{ij})\in F^{r\times r}$ 满足*

$$\begin{pmatrix}\boldsymbol{\beta}_1 & \cdots & \boldsymbol{\beta}_r\end{pmatrix}=\begin{pmatrix}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_r\end{pmatrix}\boldsymbol{T}',$$

则 $\boldsymbol{\beta}_j=\sum_{i=1}^r\boldsymbol{\alpha}_i t'_{ij}$, 即 $\boldsymbol{T}'$ 的第 j 个列向量是向量 $\boldsymbol{\beta}_j$ 在基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 下的坐标向量, 即 $\boldsymbol{T}$ 的第 j 个列向量. 这说明 $\boldsymbol{T}'=\boldsymbol{T}$. 我们可以将这一事实简述为 "过渡矩阵 $\boldsymbol{T}$ 由 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 与 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_r$ 唯一确定".

(3) *容易看出, $\boldsymbol{I}_r$ 是 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 到自身的过渡矩阵.*

(4) 若从 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_r$ 到 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 的过渡矩阵为 $\boldsymbol{S}$, 即

$$\begin{pmatrix}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_r\end{pmatrix} = \begin{pmatrix}\boldsymbol{\beta}_1 & \cdots & \boldsymbol{\beta}_r\end{pmatrix}\boldsymbol{S}.$$

此时, 可以推出

$$\begin{pmatrix}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_r\end{pmatrix} = \begin{pmatrix}\boldsymbol{\beta}_1 & \cdots & \boldsymbol{\beta}_r\end{pmatrix}\boldsymbol{S} = \left(\begin{pmatrix}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_r\end{pmatrix}\boldsymbol{T}\right)\boldsymbol{S} = \begin{pmatrix}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_r\end{pmatrix}(\boldsymbol{TS}).$$

这说明 $\boldsymbol{TS}$ 即为 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 到自身的过渡矩阵 $\boldsymbol{I}_r$. 由此可以看出, $\boldsymbol{S}$ 与 $\boldsymbol{T}$ 为互逆矩阵. 特别地, 它们都是可逆矩阵.

接着上面的的讨论. 任取向量 $\boldsymbol{v}\in V$, 假设它在基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 与 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_r$ 下的坐标分别为 $\boldsymbol{X}=(x_1,\dots,x_r)^\mathsf{T}$ 与 $\boldsymbol{Y}=(y_1,\dots,y_r)^\mathsf{T}$, 即有

$$\begin{aligned}\boldsymbol{v} &= \begin{pmatrix}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_r\end{pmatrix}\boldsymbol{X}\\ &= \begin{pmatrix}\boldsymbol{\beta}_1 & \cdots & \boldsymbol{\beta}_r\end{pmatrix}\boldsymbol{Y} = \begin{pmatrix}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_r\end{pmatrix}\boldsymbol{TY}.\end{aligned}$$

从而 $\boldsymbol{X}$ 与 $\boldsymbol{TY}$ 同为 $\boldsymbol{v}$ 在基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 下的坐标. 由坐标表示的唯一性知 $\boldsymbol{X}=\boldsymbol{TY}$. 从而在从旧的基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 到新的基 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_r$ 的变换过程中, 从旧的坐标 $\boldsymbol{X}$ 到新的坐标 $\boldsymbol{Y}$ 满足**坐标变换公式**:

$$\boldsymbol{Y}=\boldsymbol{T}^{-1}\boldsymbol{X}.$$

例 5.4.12. 已知 $V=\mathbb{R}^3$ 的两组基 $\boldsymbol{\alpha}_1=(1,1,1)^\mathsf{T}$, $\boldsymbol{\alpha}_2=(1,0,-1)^\mathsf{T}$, $\boldsymbol{\alpha}_3=(1,0,1)^\mathsf{T}$, 和 $\boldsymbol{\beta}_1=(1,2,1)^\mathsf{T}$, $\boldsymbol{\beta}_2=(2,3,4)^\mathsf{T}$, $\boldsymbol{\beta}_3=(3,4,3)^\mathsf{T}$.

(1) 求 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 到 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 的过渡矩阵 $\boldsymbol{T}$.

(2) 对于 $\boldsymbol{u}=3\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3$, 求 $\boldsymbol{u}$ 在 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 下的坐标.

解. (1) 由定义有

$$\begin{pmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \boldsymbol{\beta}_3\end{pmatrix} = \begin{pmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \boldsymbol{\alpha}_3\end{pmatrix}\boldsymbol{T}.$$

记 $\boldsymbol{B}=\begin{pmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \boldsymbol{\beta}_3\end{pmatrix}$, $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \boldsymbol{\alpha}_3\end{pmatrix}$, 则有 $\boldsymbol{B}=\boldsymbol{AT}$. 由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 和 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 都是 $\mathbb{R}^3$ 的基, $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是列满秩的方阵, 从而都是可逆矩阵. 此时, 可以推出, $\boldsymbol{T}=\boldsymbol{A}^{-1}\boldsymbol{B}$.

考虑分块矩阵的行运算,

$$\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\end{pmatrix} \xrightarrow{\text{经过一系列初等行变换}} \begin{pmatrix}\boldsymbol{I} & \boldsymbol{X}\end{pmatrix},$$

由于进行一系列初等行变换相当于左乘可逆矩阵, 我们必然有 $\boldsymbol{X} = \boldsymbol{A}^{-1}\boldsymbol{B}$, 从而可以求出 $\boldsymbol{T}$.

用这个办法, 我们看到

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \end{pmatrix} = \left(\begin{array}{ccc:ccc} 1 & 1 & 1 & 1 & 2 & 3 \\ 1 & 0 & 0 & 2 & 3 & 4 \\ 1 & -1 & 1 & 1 & 4 & 3 \end{array}\right) \xrightarrow{\text{(计算留作热身题)}} \left(\begin{array}{ccc:ccc} 1 & 0 & 0 & 2 & 3 & 4 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 & -1 \end{array}\right)$$

从而过渡矩阵为

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

(2) (a) 显然, $\boldsymbol{u}$ 在 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 下的坐标为 $\boldsymbol{X} = (3, 2, -1)^{\mathsf{T}}$. 利用初等行变换, 我们有

$$\begin{pmatrix} \boldsymbol{T} & \boldsymbol{X} \end{pmatrix} = \left(\begin{array}{ccc:c} 2 & 3 & 4 & 3 \\ 0 & -1 & 0 & 2 \\ -1 & 0 & -1 & -1 \end{array}\right) \xrightarrow{\text{(计算留作热身题)}} \left(\begin{array}{ccc:c} 1 & 0 & 0 & -5/2 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 7/2 \end{array}\right)$$

因此, $\boldsymbol{u}$ 在 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3$ 下的坐标为 $\boldsymbol{Y} = \boldsymbol{T}^{-1}\boldsymbol{X} = (-5/2, -2, 7/2)^{\mathsf{T}}$.

(b) 换一个思路来考虑. 可直接算出 $\boldsymbol{u} = (4, 3, 0)^{\mathsf{T}}$. 为了计算 $\boldsymbol{u}$ 在 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3$ 下的坐标 $\boldsymbol{Y}$, 由定义, 我们有

$$\boldsymbol{u} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3)\boldsymbol{Y} = \boldsymbol{B}\boldsymbol{Y}.$$

而 $\boldsymbol{B}$ 是一个可逆方阵. 因此, $\boldsymbol{Y} = \boldsymbol{B}^{-1}\boldsymbol{u}$. 用初等行变换, 我们看到

$$\begin{pmatrix} \boldsymbol{B} & \boldsymbol{u} \end{pmatrix} = \left(\begin{array}{ccc:c} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 3 \\ 1 & 4 & 3 & 0 \end{array}\right) \xrightarrow{\text{(计算留作热身题)}} \left(\begin{array}{ccc:c} 1 & 0 & 0 & -5/2 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 7/2 \end{array}\right).$$

由此看出, $\boldsymbol{u}$ 在 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3$ 下的坐标为 $\boldsymbol{Y} = \boldsymbol{B}^{-1}\boldsymbol{u} = (-5/2, -2, 7/2)^{\mathsf{T}}$. $\square$

设 V 是 F^n 的一个 r 维子空间, 则 V 的基为它的一个极大无关组 (长度必为 r). 若 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_s$ 是 V 中线性无关的向量, 则仿照定理 5.4.2 的证明, 可以陆续添加 $\boldsymbol{\alpha}_{s+1}, \ldots$, 使得 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_s, \boldsymbol{\alpha}_{s+1}, \ldots$ 仍然线性无关, 并且生成了 V, 从而成为 V 的一组基. 显然, 新添加的向量的个数必为 $r - s$. 此时的 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 称为 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_s$ 的一组**扩充基**. 若 $s < r$, 一般而言, $\boldsymbol{\alpha}_{s+1}, \ldots, \boldsymbol{\alpha}_r$ 的选取不唯一.

例 5.4.13. 对于给定的行向量 $\boldsymbol{\alpha}_1=(1,0,1,0)$ 和 $\boldsymbol{\alpha}_2=(1,-1,2,0)$, 求包含 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 的 $\mathbb{R}^4$ 的一组基.

解. 我们只需找到 $\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4\in\mathbb{R}^4$, 使得 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_4$ 线性无关即可. 选法不唯一. 由于 $\boldsymbol{\alpha}_1$ 和 $\boldsymbol{\alpha}_2$ 写成了行向量的形式, $\boldsymbol{\alpha}_3$ 和 $\boldsymbol{\alpha}_4$ 也需要写成行向量.

容易看出, 若取 $\boldsymbol{\alpha}_3=(0,0,1,0)$, $\boldsymbol{\alpha}_4=(0,0,0,1)$, 则矩阵

$$\begin{pmatrix}\boldsymbol{\alpha}_1\\\boldsymbol{\alpha}_2\\\boldsymbol{\alpha}_3\\\boldsymbol{\alpha}_4\end{pmatrix}=\begin{pmatrix}1&0&1&0\\1&-1&2&0\\0&0&1&0\\0&0&0&1\end{pmatrix}$$

的行列式不为 0, 从而这四个向量线性无关. □

定理 5.4.14. *对于数组空间 F^n 有:* 教材定理 5.4.2

(1) *若 V 是 F^n 的 r 维子空间, 则 V 中的任意 $r+1$ 个向量线性相关;*

(2) *若 V 是 F^n 的 r 维子空间, 则 V 中的任意 r 个线性无关的向量构成 V 的一组基;*

(3) *设 $U\subseteq V$ 是 F^n 的两个子空间, 则 $\dim(U)\le\dim(V)$;*

(4) *设 $U\subseteq V$ 是 F^n 的两个子空间, 且 $\dim(U)=\dim(V)$, 则 $U=V$.*

证明. (1) 设 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 是 V 的一组基, $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_{r+1}\in V$. 则向量组 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_{r+1}$ 可以由 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 线性表示, 从而由定理 5.3.22(3) 知

$$\operatorname{rank}(\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_{r+1})\le\operatorname{rank}(\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r)=r,$$

其中的等号是因为 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 线性无关. 这说明向量组 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_{r+1}$ 的秩严格小于它的长度. 再由定理 5.3.22(2) 知 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_{r+1}$ 线性相关.

(2) 设 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 是 V 的一组基, $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_r\in V$ 线性无关. 则向量组 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_r$ 可以由 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 线性表示, 即存在 r 阶矩阵 $\boldsymbol{T}$ 使得

$$\begin{pmatrix}\boldsymbol{\beta}_1&\cdots&\boldsymbol{\beta}_r\end{pmatrix}=\begin{pmatrix}\boldsymbol{\alpha}_1&\cdots&\boldsymbol{\alpha}_r\end{pmatrix}\boldsymbol{T}.$$

由于 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_r$ 线性无关, 利用矩阵的秩等于其列秩, 我们有

$$\operatorname{rank}(\boldsymbol{T})\ge\operatorname{rank}\left(\begin{pmatrix}\boldsymbol{\alpha}_1&\cdots&\boldsymbol{\alpha}_r\end{pmatrix}\boldsymbol{T}\right)=\operatorname{rank}\left(\begin{pmatrix}\boldsymbol{\beta}_1&\cdots&\boldsymbol{\beta}_r\end{pmatrix}\right)=r.$$

由于 $\boldsymbol{T}$ 是 r 阶方阵, 这说明 $\boldsymbol{T}$ 是一个满秩方阵, 即, $\boldsymbol{T}$ 可逆. 从而 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 与 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_r$ 等价, 即 $V=\langle\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r\rangle=\langle\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_1\rangle$. 由于 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_r$ 线性无关, 这说明 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_r$ 是 V 的一组基.

(3) 设 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 是 U 的一组基, $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_s$ 是 V 的一组基. 此时, $U \subseteq V$ 等价于说 $\langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r \rangle \subseteq \langle \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_s \rangle$, 而这等价于说 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 可以由 $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_s$ 线性表示 (定理 5.2.7), 从而由定理 5.3.22(3) 可知,

$$\dim(U) = \operatorname{rank}(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r) \leq \operatorname{rank}(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_s) = \dim(V).$$

(4) 利用上面的记号, 则 $r = s$. 上面的 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 是 r 维子空间 V 中线性无关的 r 个向量, 由 (2) 知 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 也是 V 的一组基, 从而

$$U = \langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r \rangle = V. \qquad \square$$

5.5 线性方程组解集的结构

对于线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{5.4}$$

我们将讨论解的存在性、解的唯一性，以及解的 “形状”.

线性方程组解的存在性和唯一性

定理 5.5.1. 设 $\boldsymbol{A} \in F^{m \times n}$, 而 $\boldsymbol{x} \in F^n$ 和 $\boldsymbol{b} \in F^m$ 为列向量. 记 $\overline{\boldsymbol{A}} = (\boldsymbol{A}\ \boldsymbol{b})$ 为相应的增广矩阵. 则

(1) 方程组 (5.4) 有解 $\Leftrightarrow \operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\overline{\boldsymbol{A}})$;

(2) 方程组 (5.4) 有解且唯一 $\Leftrightarrow \operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\overline{\boldsymbol{A}}) = n$.

证明. 设 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in F^m$ 为 $\boldsymbol{A}$ 的列向量.

(1) 我们有

$$\begin{aligned}
\text{方程组 } \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \text{ 有解} &\Leftrightarrow \text{ 存在 } \boldsymbol{x} = (x_1, \ldots, x_n)^{\mathsf{T}} \in F^n \text{ 使得 } x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{b} \\
&\Leftrightarrow \boldsymbol{b} \text{ 可以由 } \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \text{ 线性表示} \\
&\xLeftrightarrow{\text{定理 5.3.22(5)}} \operatorname{rank}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) = \operatorname{rank}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n, \boldsymbol{b}) \\
&\xLeftrightarrow{\text{列秩等于秩}} \operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\overline{\boldsymbol{A}}).
\end{aligned}$$