基与维数

回忆

1. 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 $\sum_{i=1}^m \lambda_i \boldsymbol{\alpha}_i \in V$.
2. 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 $V = \langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m \rangle$.

F^n 的任何子空间都是某一组向量生成的.

定理 1

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

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

注

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 1

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.

定义

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

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 $r=t$. 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, $\dim(V)\le n$.

注

在上述定义中,

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

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 2

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 3

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 4

- (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 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 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 $\lambda_i \in F$, $1 \leq i \leq r$，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，$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_r)^{\mathrm{T}}$ 是 $\boldsymbol{b}$ 在 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r$ 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: $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 和 $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_r$. 接下来讨论相应的坐标的变换. 由于 $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_r$ 可以由 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 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 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 到 $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_r$ 的过渡矩阵或转移矩阵 (transition matrix).

关于过渡矩阵 $\boldsymbol{T}$, 我们能说些什么?

注

1. 在有些国外的教材里，从 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_r$ 到 $\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_r$ 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 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_r$ 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 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_r$ 与 $\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_r$ 唯一确定".

注

3. 容易看出, $\boldsymbol{I}_r$ 是 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 到自身的过渡矩阵.
4. 若从 $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_r$ 到 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 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 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 与 $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_r$ 下的坐标分别为 $\boldsymbol{X} = (x_1, \ldots, x_r)^{\mathsf{T}}$ 与 $\boldsymbol{Y} = (y_1, \ldots, y_r)^{\mathsf{T}}$, 即有

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

从而 $\boldsymbol{X}$ 与 $\boldsymbol{T}\boldsymbol{Y}$ 同为 $\boldsymbol{v}$ 在基 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 下的坐标. 由坐标表示的唯一性知 $\boldsymbol{X} = \boldsymbol{T}\boldsymbol{Y}$. 从而在从旧的基 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 到新的基 $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_r$ 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

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.

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 1 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 6

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.

定理 7

对于数组空间 F^n 有:

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

线性方程组解集的结构

对于线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{1}$$

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 8

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 (1) 有解 $\Leftrightarrow \operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\overline{\boldsymbol{A}})$;
2. 方程组 (1) 有解且唯一 $\Leftrightarrow \operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\overline{\boldsymbol{A}}) = n$.

推论 9

关于 n 个未知元的齐次线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 总有零解. 它有非零解的充要条件是 $\operatorname{rank}(\boldsymbol{A}) < n$. 特别地, 若 $\boldsymbol{A}$ 是 n 阶方阵, 则齐次线性方程组有非零解的充要条件是 $\det(\boldsymbol{A}) = 0$.