6.2 线性变换的矩阵

之前在介绍抽象的线性空间时, 在确定一组基的前提下, 我们通过向量的坐标将对抽象向量的讨论, 可以转化为对相应的列向量、矩阵等的讨论. 这儿, 对于线性变换的处理也是类似的.

设 V 是数域 F 上的有限维线性空间, $\dim(V) = n$, 而 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n$ 是 V 的一组基. 设 $\mathscr{A}$ 是 V 上的一个线性变换.

线性变换在一组基下的矩阵　对于 $1 \le i \le n$, 向量 $\mathscr{A}(\boldsymbol{\alpha}_i) \in V$ 可以由 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n$ 唯一地线性表示出来:

$$\mathscr{A}(\boldsymbol{\alpha}_i) = \sum_{j=1}^{n} a_{ji}\boldsymbol{\alpha}_j, \qquad \text{其中 } a_{ji} \in F. \tag{6.1}$$

形式上, 我们可以采用矩阵的写法, 将其写作

$$\begin{pmatrix} \mathscr{A}(\boldsymbol{\alpha}_1) & \cdots & \mathscr{A}(\boldsymbol{\alpha}_n) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_n \end{pmatrix} \underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}}_{\boldsymbol{A}}.$$

之前我们已经提到, 我们以将抽象向量视为列向量的方式, 即式 (6.1) 中所表达的方式, 来理解上面的矩阵乘法. 其中的矩阵 $\boldsymbol{A}$ 称为**线性变换 $\mathscr{A}$ 在基 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n$ 下的矩阵**. 更进一步地, 我们将 $\begin{pmatrix} \mathscr{A}(\boldsymbol{\alpha}_1) & \cdots & \mathscr{A}(\boldsymbol{\alpha}_n) \end{pmatrix}$ 简记为 $\mathscr{A}(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n)$, 从而有

$$\mathscr{A}(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n) = \begin{pmatrix} \boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_n \end{pmatrix} \boldsymbol{A}.$$

例 6.2.1. *在之前的讨论中, 我们已经知道 $1, x, x^2, \dots, x^n$ 是 $\mathbb{R}_n[x]$ 的一组基, 并且求导运算 $\mathscr{D}$ 是 $\mathbb{R}_n[x]$ 的一个线性变换. 下面, 我们来求 $\mathscr{D}$ 在这组基下的矩阵.*

证明. 对于 $i = 0, 1, \dots, n$, 我们有 $\mathscr{D}(x^i) = (x^i)' = ix^{i-1}$, 即,

$$\begin{aligned} \mathscr{D}(1, x, x^2, \dots, x^n) &= (0, 1, 2x, 3x^2, \dots, nx^{n-1}) \\ &= \begin{pmatrix} 1 & x & x^2 & \cdots & x^n \end{pmatrix} \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & 2 & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & & & \ddots & n \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}_{(n+1)\times(n+1)}}_{\text{所求的矩阵}}. \end{aligned}$$

□

注 6.2.2. (1) 若 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 是线性空间 V 的一组基, $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_n$ 是 V 中任意的一组向量, 则存在唯一一个 V 上的线性变换 $\mathscr{A}$, 使得对任意的 $1\le i\le n$ 有 $\mathscr{A}(\boldsymbol{\alpha}_i)=\boldsymbol{\beta}_i$. 其证明比较简单. 教材习题 #6 只要求证明了它的存在性. 而唯一性可以借用接下来的定理说明.

(2) 设 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 是线性空间 V 的一组基, 对于给定的 V 上的线性变换 $\mathscr{A}$, 它在这组基下的矩阵 $\boldsymbol{A}$ 的列向量是由 $\mathscr{A}(\boldsymbol{\alpha}_1),\mathscr{A}(\boldsymbol{\alpha}_2),\dots,\mathscr{A}(\boldsymbol{\alpha}_n)$ 在这组基下的坐标唯一确定; 反之, 若给定一个 n 阶方阵 $\boldsymbol{A}$ 作为线性变换 $\mathscr{A}$ 在这组基下的矩阵, 也就等价于给出了这组基在线性变换 $\mathscr{A}$ 下的像 $\mathscr{A}(\boldsymbol{\alpha}_1),\mathscr{A}(\boldsymbol{\alpha}_2),\dots,\mathscr{A}(\boldsymbol{\alpha}_n)$, 从而也就确定了线性变换 $\mathscr{A}$. 这说明在基给定的条件下, 线性变换 $\mathscr{A}$ 与矩阵 $\boldsymbol{A}$ 有一个一一对应的关系.

(3) 若已知线性空间 V 上的线性变换 $\mathscr{A}$, 以及 $\mathscr{A}$ 在 V 的某组基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ (未知) 下的矩阵为 $\boldsymbol{A}$ (已知), 我们一般很难把这组基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 再还原出来. 例如, V 上的恒等变换在任何基下的矩阵都是单位阵 $\boldsymbol{I}$, 零变换在任何基下的矩阵都是零矩阵 $\boldsymbol{O}$.

定理 6.2.3. 设 $\mathscr{A}$ 是线性空间 V 上的线性变换, 在基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 下的矩阵为 $\boldsymbol{A}$. 设 $\boldsymbol{x}\in V$, $\boldsymbol{y}=\mathscr{A}(\boldsymbol{x})$, 而 $\boldsymbol{x},\boldsymbol{y}$ 在基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 下的坐标分别为 $\boldsymbol{X},\boldsymbol{Y}$, 则 $\boldsymbol{Y}=\boldsymbol{A}\boldsymbol{X}$. 教材定理 6.2.1

证明. **(形式的证明)** 不妨设 $\boldsymbol{X}=(x_1,\dots,x_n)^{\mathsf{T}}$, 则 $\boldsymbol{x}=\sum_i x_i\boldsymbol{\alpha}_i$. 此时,

$$\begin{aligned}\boldsymbol{y}=&\mathscr{A}(\boldsymbol{x})=\mathscr{A}\left(\sum_i x_i\boldsymbol{\alpha}_i\right)\xlongequal{\mathscr{A}\text{ 的线性性质}}\sum_i x_i\mathscr{A}(\boldsymbol{\alpha}_i)=\begin{pmatrix}\mathscr{A}(\boldsymbol{\alpha}_1)&\cdots&\mathscr{A}(\boldsymbol{\alpha}_n)\end{pmatrix}\boldsymbol{X}\\ &=\left(\begin{pmatrix}\boldsymbol{\alpha}_1&\cdots&\boldsymbol{\alpha}_n\end{pmatrix}\boldsymbol{A}\right)\boldsymbol{X}=\begin{pmatrix}\boldsymbol{\alpha}_1&\cdots&\boldsymbol{\alpha}_n\end{pmatrix}(\boldsymbol{A}\boldsymbol{X}).\end{aligned}$$

但是, $\boldsymbol{y}=\begin{pmatrix}\boldsymbol{\alpha}_1&\cdots&\boldsymbol{\alpha}_n\end{pmatrix}\boldsymbol{Y}$. 由坐标表示的唯一性, 我们可以推出 $\boldsymbol{Y}=\boldsymbol{A}\boldsymbol{X}$.

(分量方式的直接证明) 设 $\boldsymbol{A}=(a_{ij})$. 由定义, $\mathscr{A}(\boldsymbol{\alpha}_j)=\sum_i a_{ij}\boldsymbol{\alpha}_i$. 于是, $y=\mathscr{A}(\boldsymbol{x})=\mathscr{A}(\sum_j x_j\boldsymbol{\alpha}_j)=\sum_j x_j\mathscr{A}(\boldsymbol{\alpha}_j)=\sum_j x_j\sum_i a_{ij}\boldsymbol{\alpha}_i=\sum_i(\sum_j a_{ij}x_j)\boldsymbol{\alpha}_i$. 于是, 若 $\boldsymbol{Y}=(y_1,\dots,y_n)^{\mathsf{T}}$, 则 $y_i=\sum_j a_{ij}x_j$. 换言之, $\boldsymbol{Y}=\boldsymbol{A}\boldsymbol{X}$. □

注 6.2.4 (比较两组基的过渡矩阵与线性变换在基下的矩阵). (1) 给定两组基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 和 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_n$. 则基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 到 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_n$ 的过渡矩阵 $\boldsymbol{T}$ 反映的是 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_n$ 用 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 线性表示的系数 (坐标).

(2) 给定基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 和线性变换 $\mathscr{A}$, 则 $\mathscr{A}$ 在基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 下的矩阵 $\boldsymbol{A}$ 反映的是 $\mathscr{A}(\boldsymbol{\alpha}_1),\dots,\mathscr{A}(\boldsymbol{\alpha}_n)$ 用 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 线性表示的系数 (坐标).

注意: (1) 中过渡矩阵的计算中, 我们没有用到 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_n$ 是线性无关的这条性质. 类似地, (2) 中, 我们一般也没有 $\mathscr{A}(\boldsymbol{\alpha}_1),\dots,\mathscr{A}(\boldsymbol{\alpha}_n)$ 线性无关. 关键仅仅为 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 是基. 所以上面 (1) 与 (2) 中的计算实质是一致的, 虽然反映的信息不尽相同.

例 6.2.5. 设 $V=F^n$, $\boldsymbol{A}\in F^{n\times n}$, 而 $\mathscr{A}:V\to V$ 为线性变换. 证明以下几条:

(1) 若 $\mathscr{A}$ 由法则 $\mathscr{A}(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{x}$ 给出, 则 $\mathscr{A}$ 在标准基 $\boldsymbol{e}_1,\dots,\boldsymbol{e}_n$ 下的矩阵恰为 $\boldsymbol{A}$;

(2) 若 $\mathscr{A}$ 在标准基 $\boldsymbol{e}_1,\dots,\boldsymbol{e}_n$ 下的矩阵为 $\boldsymbol{A}$, 则 $\mathscr{A}$ 由法则 $\mathscr{A}(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{x}$ 给出.

证明. (1) 我们记 $\boldsymbol{A}=(a_{ij})_{n\times n}$, 则

$$\mathscr{A}(\boldsymbol{e}_i)=\boldsymbol{A}\boldsymbol{e}_i=\begin{pmatrix}a_{1i}\\ \vdots\\ a_{ni}\end{pmatrix}=\sum_{j=1}^n a_{ji}\boldsymbol{e}_j=\begin{pmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{pmatrix}\begin{pmatrix}a_{1i}\\ \vdots\\ a_{ni}\end{pmatrix}.$$

从而由定义有

$$\mathscr{A}(\boldsymbol{e}_1,\dots,\boldsymbol{e}_n)=\begin{pmatrix}\mathscr{A}(\boldsymbol{e}_1) & \cdots & \mathscr{A}(\boldsymbol{e}_n)\end{pmatrix}=\begin{pmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{pmatrix}\underbrace{\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}}_{\text{恰为 }\boldsymbol{A}}.$$

故 $\mathscr{A}$ 在 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 下的矩阵恰为 $\boldsymbol{A}$.

(2) **(思路一)** 设 $\boldsymbol{a}_1,\dots,\boldsymbol{a}_n$ 依次为 $\boldsymbol{A}$ 的 n 个列向量. 由定义, 对于每个 i, 有

$$\mathscr{A}(\boldsymbol{e}_i)=(\boldsymbol{e}_1\ \cdots\ \boldsymbol{e}_n)\boldsymbol{a}_i=\boldsymbol{a}_i.$$

任何一个数组向量在标准基下的坐标列向量就是自身

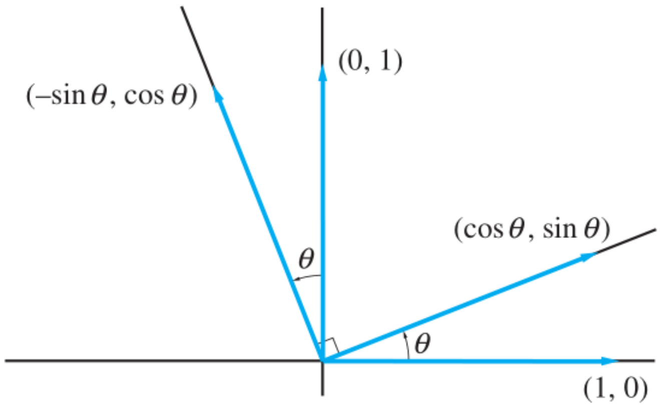
因此, 对任意的 $\boldsymbol{x}=(x_1,\dots,x_n)^{\mathsf{T}}\in F^n$, 我们有

$$\begin{aligned}\mathscr{A}(\boldsymbol{x})&=\mathscr{A}\left(\sum_{i=1}^n x_i\boldsymbol{e}_i\right)=\sum_{i=1}^n x_i\mathscr{A}(\boldsymbol{e}_i)=\sum_{i=1}^n x_i\boldsymbol{a}_i\\&=(\boldsymbol{a}_1\ \cdots\ \boldsymbol{a}_n)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\boldsymbol{A}\boldsymbol{x}.\end{aligned}$$

(思路二) 由注 6.2.2 知在 F^n 上存在唯一的线性变换 $\mathscr{A}$ 使得其在标准基下的矩阵为给定的 $\boldsymbol{A}$. 另一方面, 由法则 $\mathscr{A}':\boldsymbol{x}\mapsto\boldsymbol{A}\boldsymbol{x}$ 给出的 F^n 上的映射也是符合条件的线性变换. 由唯一性可知 $\mathscr{A}=\mathscr{A}'$. □

例 6.2.6. 考虑平面 $\mathbb{R}^2$ 上的逆时针旋转角度 θ 的线性变换 $\mathscr{A}$. 用几何的方法容易看出, 在 $\mathscr{A}$ 的作用下,

$$\boldsymbol{e}_1=(1,0)^{\mathsf{T}}\mapsto(\cos(\theta),\sin(\theta))^{\mathsf{T}},\qquad \boldsymbol{e}_2=(0,1)^{\mathsf{T}}\mapsto(-\sin(\theta),\cos(\theta))^{\mathsf{T}}.$$



这说明 $\mathscr{A}$ 在 $\mathbb{R}^2$ 的标准基 $\boldsymbol{e}_1,\boldsymbol{e}_2$ 下的矩阵为

$$\boldsymbol{A}=\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}.$$

并且, 对任意的向量 $\boldsymbol{x}=(a,b)^{\mathsf{T}}\in\mathbb{R}^2$,

$$\mathscr{A}(\boldsymbol{x})=\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix}.$$

线性变换在不同基下的矩阵 接下来的讨论希望理解: 例 6.2.5 中 $\mathscr{A}$ 在其它基下的矩阵是否仍然为 $\boldsymbol{A}$. 若否, 则新的矩阵与原来的 $\boldsymbol{A}$ 有什么关系?

定理 6.2.7. 设线性变换 $\mathscr{A}:V\to V$ 在基 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n$ 与 $\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n$ 下的矩阵分别为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$. 设基 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n$ 到 $\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n$ 的过渡矩阵为 $\boldsymbol{T}$, 即, 教材定理 6.2.2

$$\begin{pmatrix}\boldsymbol{\beta}_1 & \cdots & \boldsymbol{\beta}_n\end{pmatrix}=\begin{pmatrix}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_n\end{pmatrix}\boldsymbol{T},$$

则 $\boldsymbol{B}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$.

证明. 若 $\boldsymbol{t}_1,\ldots,\boldsymbol{t}_n$ 依次为 $\boldsymbol{T}$ 的列向量, 则其中 $\boldsymbol{t}_i$ 为 $\boldsymbol{\beta}_i$ 在基 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n$ 下的坐标. 由定理 6.2.3 可知, $\boldsymbol{A}\boldsymbol{t}_i$ 为 $\mathscr{A}(\boldsymbol{\beta}_i)$ 在基 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n$ 下的坐标, 再由坐标变换公式可知, $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{t}_i$ 为 $\mathscr{A}(\boldsymbol{\beta}_i)$ 在基 $\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n$ 下的坐标. 从而由定义可知,

$$\boldsymbol{B}=(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{t}_1,\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{t}_2,\ldots,\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{t}_n)=\boldsymbol{T}^{-1}\boldsymbol{A}(\boldsymbol{t}_1,\ldots,\boldsymbol{t}_n)=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}.\qquad\square$$

例 6.2.8. 设线性变换 $\mathscr{A}$ 在基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 下的矩阵为 $\boldsymbol{A}=\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}$, 求 $\mathscr{A}$ 在基 $\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1$ 下的矩阵 $\boldsymbol{B}$.

解. 不难看出, 基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 到基 $\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1$ 的过渡矩阵为 $\boldsymbol{T}=\begin{pmatrix}&&1\\&1&\\1&&\end{pmatrix}$, 并且 $\boldsymbol{T}^{-1}=\boldsymbol{T}$. 从而,

$$\begin{aligned}\boldsymbol{B}=&\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}\\&\xlongequal{r_1\leftrightarrow r_3}\begin{pmatrix}7&8&9\\4&5&6\\1&2&3\end{pmatrix}\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}\xlongequal{c_1\leftrightarrow c_3}\begin{pmatrix}9&8&7\\6&5&4\\3&2&1\end{pmatrix}.\end{aligned}$$

□

例 6.2.9. 设 $\mathscr{A}: F^3\to F^3$ 为线性变换, 满足

教材例题 6.2.2

$$\begin{aligned}\boldsymbol{\alpha}_1=&(2,3,5)^{\mathsf{T}} &\mapsto\quad& \boldsymbol{\beta}_1=(1,2,0)^{\mathsf{T}},\\\boldsymbol{\alpha}_2=&(0,1,2)^{\mathsf{T}} &\mapsto\quad& \boldsymbol{\beta}_2=(2,4,-1)^{\mathsf{T}},\\\boldsymbol{\alpha}_3=&(1,0,0)^{\mathsf{T}} &\mapsto\quad& \boldsymbol{\beta}_3=(3,0,5)^{\mathsf{T}}.\end{aligned}$$

求

(1) $\mathscr{A}$ 在 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 下的矩阵 $\boldsymbol{A}$;

(2) $\mathscr{A}$ 在标准基下的矩阵 $\boldsymbol{B}$.

解. 注意到 $\det(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=1\neq 0$, 故 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 是 F^3 的一组基.

(1) 由定义,

$$\begin{pmatrix}\boldsymbol{\alpha}_1&\boldsymbol{\alpha}_2&\boldsymbol{\alpha}_3\end{pmatrix}\boldsymbol{A}=\mathscr{A}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=\begin{pmatrix}\boldsymbol{\beta}_1&\boldsymbol{\beta}_2&\boldsymbol{\beta}_3\end{pmatrix},$$

从而

$$\begin{aligned}\boldsymbol{A}&=\begin{pmatrix}\boldsymbol{\alpha}_1&\boldsymbol{\alpha}_2&\boldsymbol{\alpha}_3\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{\beta}_1&\boldsymbol{\beta}_2&\boldsymbol{\beta}_3\end{pmatrix}\\&=\begin{pmatrix}2&0&1\\3&1&0\\5&2&0\end{pmatrix}^{-1}\begin{pmatrix}1&2&3\\2&4&0\\0&-1&5\end{pmatrix}\xlongequal{\text{留作热身题}}\begin{pmatrix}4&9&-5\\-10&-23&15\\-7&-16&13\end{pmatrix}.\end{aligned}$$

具体计算如下:

$$\left(\begin{array}{ccc:ccc}2&0&1&1&2&3\\3&1&0&2&4&0\\5&2&0&0&-1&5\end{array}\right)\xrightarrow{-2r_1\to r_2}\left(\begin{array}{ccc:ccc}2&0&1&1&2&3\\1&1&-1&1&2&-3\\5&2&0&0&-1&5\end{array}\right)\xrightarrow{r_1\leftrightarrow r_2}$$

$$\left(\begin{array}{ccc:ccc}1&1&-1&1&2&-3\\2&0&1&1&2&3\\5&2&0&0&-1&5\end{array}\right)\xrightarrow[-5r_1\to r_3]{-2r_1\to r_2}\left(\begin{array}{ccc:ccc}1&1&-1&1&2&-3\\0&-2&3&-1&-2&9\\0&-3&5&-5&-11&20\end{array}\right)\xrightarrow{-r_3\to r_2}$$

$$\left(\begin{array}{ccc:ccc}1&1&-1&1&2&-3\\0&1&-2&4&9&-11\\0&-3&5&-5&-11&20\end{array}\right)\xrightarrow[3r_2\to r_3]{-r_2\to r_1}\left(\begin{array}{ccc:ccc}1&0&1&-3&-7&8\\0&1&-2&4&9&-11\\0&0&-1&7&16&-13\end{array}\right)\xrightarrow[-2r_3\to r_2]{r_3\to r_1}$$

$$\left(\begin{array}{ccc:ccc}1&0&0&4&9&-5\\0&1&0&-10&-23&15\\0&0&-1&7&16&-13\end{array}\right)\xrightarrow{-r_3}\left(\begin{array}{ccc:ccc}1&0&0&4&9&-5\\0&1&0&-10&-23&15\\0&0&1&-7&-16&13\end{array}\right).$$

(2) **(思路一)** $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 在标准基下的坐标就为向量本身, 故由定理 6.2.3 可知, 对于 $i=1,2,3$, 有 $\boldsymbol{\beta}_i=\boldsymbol{B}\boldsymbol{\alpha}_i$, 从而

$$\begin{pmatrix}\boldsymbol{\beta}_1&\boldsymbol{\beta}_2&\boldsymbol{\beta}_3\end{pmatrix}=\boldsymbol{B}\begin{pmatrix}\boldsymbol{\alpha}_1&\boldsymbol{\alpha}_2&\boldsymbol{\alpha}_3\end{pmatrix},$$

故

$$\begin{aligned}\boldsymbol{B}&=\begin{pmatrix}\boldsymbol{\beta}_1&\boldsymbol{\beta}_2&\boldsymbol{\beta}_3\end{pmatrix}\begin{pmatrix}\boldsymbol{\alpha}_1&\boldsymbol{\alpha}_2&\boldsymbol{\alpha}_3\end{pmatrix}^{-1}\\&=\begin{pmatrix}1&2&3\\2&4&0\\0&-1&5\end{pmatrix}\begin{pmatrix}2&0&1\\3&1&0\\5&2&0\end{pmatrix}^{-1}\xlongequal{\text{留作热身题}}\begin{pmatrix}3&-20&11\\0&-16&10\\5&-15&7\end{pmatrix}.\end{aligned}$$

具体计算如下:

$$\left(\begin{array}{ccc}1&2&3\\2&4&0\\0&-1&5\\\hdashline 2&0&1\\3&1&0\\5&2&0\end{array}\right)\xrightarrow{c_1\leftrightarrow c_3}\left(\begin{array}{ccc}3&2&1\\0&4&2\\5&-1&0\\\hdashline 1&0&2\\0&1&3\\0&2&5\end{array}\right)\xrightarrow{-2c_1\to c_3}\left(\begin{array}{ccc}3&2&-5\\0&4&2\\5&-1&-10\\\hdashline 1&0&0\\0&1&3\\0&2&5\end{array}\right)\xrightarrow{-3c_2\to c_3}$$

$$\left(\begin{array}{ccc}3 & 2 & -11\\ 0 & 4 & -10\\ 5 & -1 & -7\\ \hdashline 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 2 & -1\end{array}\right) \xrightarrow{-c_3} \left(\begin{array}{ccc}3 & 2 & 11\\ 0 & 4 & 10\\ 5 & -1 & 7\\ \hdashline 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 2 & 1\end{array}\right) \xrightarrow{-2c_3\to c_2} \left(\begin{array}{ccc}3 & -20 & 11\\ 0 & -16 & 10\\ 5 & -15 & 7\\ \hdashline 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{array}\right).$$

(思路二) 从标准基到基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 的过渡矩阵显然就是方阵 $\begin{pmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \boldsymbol{\alpha}_3\end{pmatrix}$, 于是, 从基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 到标准基的过渡矩阵是方阵 $\begin{pmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \boldsymbol{\alpha}_3\end{pmatrix}^{-1}$. 由此可知, $\boldsymbol{B} = \begin{pmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \boldsymbol{\alpha}_3\end{pmatrix} \boldsymbol{A} \begin{pmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \boldsymbol{\alpha}_3\end{pmatrix}^{-1}$. 由于在第一小问的计算中知道了 $\boldsymbol{A} = \begin{pmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \boldsymbol{\alpha}_3\end{pmatrix}^{-1} \begin{pmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \boldsymbol{\beta}_3\end{pmatrix}$, 因此我们接下来的计算可以化为前面一个思路里的讨论. □

习题 6.2.10. 设 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3$ 构成向量空间 V 的一组基, 而 V 上的线性变换 $\mathscr{A}$ 在这组基下的矩阵为 $\begin{pmatrix}0 & -6 & 1\\ 0 & 5 & -1\\ 1 & -2 & 7\end{pmatrix}$. 计算 $\mathscr{A}(3\boldsymbol{\beta}_1 - 4\boldsymbol{\beta}_2 + 5\boldsymbol{\beta}_3)$.

矩阵的相似

回忆. 在定理 *6.2.7* 中, 设线性变换 $\mathscr{A}: V \to V$ 在基 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n$ 与 $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_n$ 下的矩阵分别为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$. 设基 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n$ 到 $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_n$ 的过渡矩阵为 $\boldsymbol{T}$, 即,

$$\begin{pmatrix}\boldsymbol{\beta}_1 & \dots & \boldsymbol{\beta}_n\end{pmatrix} = \begin{pmatrix}\boldsymbol{\alpha}_1 & \dots & \boldsymbol{\alpha}_n\end{pmatrix} \boldsymbol{T},$$

则 $\boldsymbol{B} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$.

一般地, 考察 $\boldsymbol{B} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$, 其中 $\boldsymbol{T}$ 可逆, 则矩阵 $\boldsymbol{B}$ 和 $\boldsymbol{A}$ 本质上反映了同一个东西 (线性变换 $\mathscr{A}$).

定义 6.2.11. 设 $\boldsymbol{A}, \boldsymbol{B}$ 为数域 F 上的两个 n 阶方阵. 若存在 F 上的同阶可逆方阵 $\boldsymbol{T}$ 使得 $\boldsymbol{B} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$, 则称 $\boldsymbol{B}$ 是由 $\boldsymbol{A}$ 做**相似变换**得到的, 称 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ **(在 F 上) 相似** (*similar*), 记作 $A \sim B$. 若 $F = \mathbb{R}$, 则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **实相似**; 若 $F = \mathbb{C}$, 则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **复相似**.

命题 6.2.12. 同阶方阵的相似关系为等价关系, 即 教材命题 6.2.1

(反身性) $\boldsymbol{A}$ 与自身相似;

(对称性) 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 则反过来 $\boldsymbol{B}$ 也与 $\boldsymbol{A}$ 相似;

(传递性) 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 相似, 则 $\boldsymbol{A}$ 也与 $\boldsymbol{C}$ 相似.

证明. 用定义可以直接验证, 略. □

因此, 当 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是相似的时候, 我们也称它们为**相似等价**的.

推论 6.2.13. 有限维线性空间 V 上的一个线性变换在 V 的不同基下的矩阵是相似的.

命题 6.2.14 (相似矩阵的基本性质)**.** 设 $\boldsymbol{A}, \boldsymbol{A}_i, \boldsymbol{B}, \boldsymbol{B}_i$ 都是 F 上的方阵.

(1) 若 $\boldsymbol{A} \sim \boldsymbol{B}$, 则 $\boldsymbol{A}^m \sim \boldsymbol{B}^m$, 其中 m 为正整数.

(2) 若 $\boldsymbol{A} \sim \boldsymbol{B}$, 设 $f(x)$ 是一个一元多项式, 则矩阵多项式 $f(\boldsymbol{A}) \sim f(\boldsymbol{B})$.

(3) 若 $\boldsymbol{A}_i \sim \boldsymbol{B}_i$, $i = 1, 2, \ldots, s$, 则 $\operatorname{diag}(\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_s) \sim \operatorname{diag}(\boldsymbol{B}_1, \boldsymbol{B}_2, \ldots, \boldsymbol{B}_s)$.

(4) 对于任意的排列 $(i_1, i_2, \ldots, i_s) \in \mathcal{S}_s$, 我们有 $\operatorname{diag}(\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_s) \sim \operatorname{diag}(\boldsymbol{A}_{i_1}, \boldsymbol{A}_{i_2}, \ldots, \boldsymbol{A}_{i_s})$

(5) 若 $\boldsymbol{A} \sim \boldsymbol{B}$, 且 $\boldsymbol{A}$ 可逆, 则 $\boldsymbol{B}$ 也可逆, 且 $\boldsymbol{A}^{-1} \sim \boldsymbol{B}^{-1}$.

(6) 相似的矩阵有相同的行列式、迹和秩.

证明. (1) 若 $\boldsymbol{A} = \boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{T}$, 则对任意的正整数 m, 有

$$\boldsymbol{A}^m = \underbrace{(\boldsymbol{T}^{-1}\boldsymbol{B}\underbrace{\boldsymbol{T})(\boldsymbol{T}^{-1}}_{\text{相乘抵消}}\boldsymbol{B}\boldsymbol{T})\cdots(\boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{T})}_{m\text{组}} = \boldsymbol{T}^{-1}\boldsymbol{B}^m\boldsymbol{T},$$

(2) 若 $\boldsymbol{A} = \boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{T}$, 而 $f(x) = \sum_{i=0}^m a_i x^i$, 那么

$$\begin{aligned} f(\boldsymbol{A}) &= \sum_{i=0}^m a_i \boldsymbol{A}^i = \sum_{i=0}^m a_i (\boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{T})^m = \sum_{i=0}^m a_i \boldsymbol{T}^{-1}\boldsymbol{B}^m\boldsymbol{T} \\ &= \boldsymbol{T}^{-1}\left(\sum_{i=0}^m a_i \boldsymbol{B}^i\right)\boldsymbol{T} = \boldsymbol{T}^{-1} f(\boldsymbol{B})\boldsymbol{T}. \end{aligned}$$

(3) 若 $\boldsymbol{A}_i = \boldsymbol{T}_i^{-1}\boldsymbol{B}_i\boldsymbol{T}_i$, $i = 1, 2, \ldots, s$, 那么

$$\begin{aligned} \boldsymbol{A} &= \operatorname{diag}(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_s) = \operatorname{diag}(\boldsymbol{T}_1^{-1}\boldsymbol{B}_1\boldsymbol{T}_1, \ldots, \boldsymbol{T}_s^{-1}\boldsymbol{B}_s\boldsymbol{T}_s) \\ &= \operatorname{diag}(\boldsymbol{T}_1^{-1}, \ldots, \boldsymbol{T}_s^{-1}) \operatorname{diag}(\boldsymbol{B}_1, \ldots, \boldsymbol{B}_s) \operatorname{diag}(\boldsymbol{T}_1, \ldots, \boldsymbol{T}_s) \\ &= \underbrace{\operatorname{diag}(\boldsymbol{T}_1, \ldots, \boldsymbol{T}_s)^{-1}}_{\boldsymbol{T}^{-1}} \operatorname{diag}(\boldsymbol{B}_1, \ldots, \boldsymbol{B}_s) \underbrace{\operatorname{diag}(\boldsymbol{T}_1, \ldots, \boldsymbol{T}_s)}_{\text{记作 } \boldsymbol{T}}. \end{aligned}$$

(4) 由于任何一个排列都可以通过交换相邻位置, 化为顺序排列, 通过化归的方法, 这儿我们只需考虑 $s = 2$ 且 $(i_1, i_2) = (2, 1)$ 的情形. 此时, 设 $\boldsymbol{A}_1$ 与 $\boldsymbol{A}_2$ 分别为 n_1 阶方阵和 n_2 阶方阵, 并令 $\boldsymbol{T} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{I}_{n_1} \\ \boldsymbol{I}_{n_2} & \boldsymbol{O} \end{pmatrix}$. 不难验证, $\boldsymbol{T}$ 可逆, 满足 $\boldsymbol{T}^{-1} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{I}_{n_2} \\ \boldsymbol{I}_{n_1} & \boldsymbol{O} \end{pmatrix}$, 并且有 $\boldsymbol{T}\,\mathrm{diag}(\boldsymbol{A}_1, \boldsymbol{A}_2)\boldsymbol{T}^{-1} = \mathrm{diag}(\boldsymbol{A}_2, \boldsymbol{A}_1)$.

这儿的证明与思路请用分块矩阵的初等变换的方法来理解