第三章　线性空间与线性变换

正如前面已经介绍的那样, 在许多理论研究和实际应用中都需要处理有序数组(n元向量)、矩阵、映射等对象. 为了用整体的观点考虑和研究这些对象, 我们有必要引入由这些对象及其线性运算构成的代数系统, 这就是线性空间, 它是研究客观世界中线性问题的重要理论, 即使对于非线性问题, 经过局部化后, 也可以运用线性空间的理论对其展开理论研究. 在本章, 我们将讨论线性空间的结构、介绍线性空间之间保持加法与数乘的映射——线性映射, 并重点研究一个线性空间到其自身的线性映射——线性变换.

§3.1　线性空间

通过前两章的学习, 我们已经知道数域F上所有n元向量组成的集合有加法、数量乘法运算, 它们满足加法交换律、结合律等八条运算法则(性质**??**); 数域F上所有$m\times n$矩阵组成的集合有矩阵的加法、数量乘法运算, 它们也满足加法交换律、结合律等八条运算法则(性质**??**). 类似的情况还有很多. 由此受到启发, 人们从这些结构相似的代数系统中抽象出共同的属性, 而略去具体元素与具体运算的实际含义. 这一抽象的代数系统就是线性空间.

定义 3.1.1 *设V是一个非空集合, F是一个数域. 在V的元素之间定义一种代数运算称为***加法***, 即对任意$\boldsymbol{\alpha},\boldsymbol{\beta}\in V$, 在$V$中都有唯一的一个元素$\boldsymbol{\gamma}$与它们对应, 称之为$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$的和, 记为$\boldsymbol{\gamma}=\boldsymbol{\alpha}+\boldsymbol{\beta}$. 在数域$F$和集合$V$的元素之间还定义了一种运算称为***数量乘法***, 即对任一$k\in F$及任一$\boldsymbol{\alpha}\in V$, 在V中都有唯一的一个元素$\boldsymbol{\delta}$与它们对应, 称之为k与$\boldsymbol{\alpha}$的***数量乘积***, 记为$\boldsymbol{\delta}=k\boldsymbol{\alpha}$. 如果加法与数量乘法满足下面八条运算规则, 就称V是F上的***线性空间***. 这八条运算规则是:*

(1) $\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{\beta}+\boldsymbol{\alpha}$;

(2) $(\boldsymbol{\alpha}+\boldsymbol{\beta})+\boldsymbol{\gamma}=\boldsymbol{\alpha}+(\boldsymbol{\beta}+\boldsymbol{\gamma})$;

(3) *在V中存在一个元素$\boldsymbol{\theta}$, 对任一$\boldsymbol{\alpha}\in V$, 都有$\boldsymbol{\alpha}+\boldsymbol{\theta}=\boldsymbol{\alpha}$, 其中$\boldsymbol{\theta}$称为$V$的***零元素***;*

(4) *对任一$\boldsymbol{\alpha}\in V$, 都存在V中的元素$\boldsymbol{\sigma}$使得$\boldsymbol{\alpha}+\boldsymbol{\sigma}=\boldsymbol{\theta}$, 其中$\boldsymbol{\sigma}$称为$\boldsymbol{\alpha}$的***负元素***, 记为$-\boldsymbol{\alpha}$;*

(5) $1\boldsymbol{\alpha}=\boldsymbol{\alpha}$;

(6) $(kl)\boldsymbol{\alpha}=k(l\boldsymbol{\alpha})$;

(7) $(k+l)\boldsymbol{\alpha}=k\boldsymbol{\alpha}+l\boldsymbol{\alpha}$;

(8) $k(\boldsymbol{\alpha}+\boldsymbol{\beta})=k\boldsymbol{\alpha}+k\boldsymbol{\beta}$.

这里$\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$是$V$中任意元素, k,l是F中的任意数.

线性空间的元素也称为**向量**. 显然, 这里所谓的向量比前面所述n元向量的涵义更广泛.

例 3.1.1 数域F上全体n元向量的集合对向量的加法及数与向量的数量乘法, 构成F上的线性空间(称之为**向量空间**), 记为F^n.

例 3.1.2 数域F上全体$m\times n$矩阵的集合对矩阵的加法及数与矩阵的数量乘法, 构成F上的线性空间(称之为**矩阵空间**), 记为$F^{m\times n}$.

例 3.1.3 设

$$F[x]_n=\{a_0+a_1x+a_2x^2+\cdots+a_{n-1}x^{n-1}|a_i\in F,i=0,1,\ldots,n-1\}$$

表示数域F上所有次数小于n的一元多项式再添加零多项式构成的集合, $F[x]_n$中的两个多项式

$$f(x)=\sum_{i=0}^{n-1}a_ix^i,\quad g(x)=\sum_{i=0}^{n-1}b_ix^i$$

有加法运算:

$$f(x)+g(x)=\sum_{i=0}^{n-1}(a_i+b_i)x^i$$

数域F中的数k与多项式$f(x)$有数量乘法运算:

$$kf(x)=\sum_{i=0}^{n-1}(ka_i)x^i$$

容易验证, 上述的加法和数量乘法运算满足下列八条运算规则:

(1) 交换律: $f(x)+g(x)=g(x)+f(x)$;

(2) 结合律: $[f(x)+g(x)]+h(x)=f(x)+[g(x)+h(x)]$;

(3) 存在零多项式$\boldsymbol{\theta}$: 对任一$f(x)\in F[x]_n$, 都有$f(x)+\boldsymbol{\theta}=f(x)$;

(4) 存在负多项式: 对任一$f(x)=\sum\limits_{i=0}^{n-1}a_ix^i\in F[x]_n$,都有$-f(x)=\sum\limits_{i=0}^{n-1}(-a_i)x^i\in F[x]_n$, 满足$f(x)+(-f(x))=\boldsymbol{\theta}$;

(5) $1\ f(x)=f(x)$;

(6) $(kl)f(x) = k(lf(x))$;

(7) $(k+l)f(x) = kf(x) + lf(x)$;

(8) $k(f(x) + g(x)) = kf(x) + kg(x)$.

这里$f(x), g(x), h(x) \in F[x]_n$, $k, l \in F$. 因此, $F[x]_n$对多项式的加法及数与多项式的乘法也构成F上的线性空间. 同样地, 以数域F中的数为系数的全体1元多项式的集合$F[x]$对多项式的加法及数与多项式的乘法, 也构成F上的线性空间(称之为**多项式空间**).

例 3.1.4 在闭区间$[a, b]$上一切连续实函数的集合$C[a, b]$上定义函数的加法及实数与函数的数量乘法:

$$\begin{aligned}(f+g)(x) &= f(x) + g(x), \quad \forall\ x \in \mathbf{R}\\ (kf)(x) &= k(f(x)), \qquad\quad \forall\ x \in \mathbf{R}\end{aligned}$$

其中$f, g \in C[a, b], k \in \mathbf{R}$, 则$C[a, b]$构成实数域$\mathbf{R}$上的线性空间(称之为**函数空间**).

例 3.1.5 设$\boldsymbol{A} \in F^{m\times n}$, 则齐次线性方程组$\boldsymbol{AX} = \mathbf{0}$的所有解的集合构成数域$F$上的线性空间. 这个空间称为方程组$\boldsymbol{AX} = \mathbf{0}$**的解空间**, 也称为**矩阵$\boldsymbol{A}$的核**(或**零空间**), 常用$\mathrm{N}(\boldsymbol{A})$表示.

例 3.1.6 设$\mathbf{R}^+$是全体正实数的集合, $\mathbf{R}$是实数域. 在$\mathbf{R}^+$中定义元素的加法“$\oplus$”及$\mathbf{R}$中的数与$\mathbf{R}^+$的元素之间的数量乘法“$\cdot$”:

$$a \oplus b = ab, \quad k \cdot a = a^k$$

其中$a, b \in \mathbf{R}^+, k \in \mathbf{R}$, 则$\mathbf{R}^+$关于运算“$\oplus$”和“$\cdot$”构成$\mathbf{R}$上的线性空间.

证 显然, 如上定义的“$\oplus$”和“$\cdot$”对$\mathbf{R}^+$是封闭的, 即它们确实可作为$\mathbf{R}^+$的运算. 下面逐一验证它们也满足线性空间定义中的八条运算规则. 任取$a, b, c \in \mathbf{R}^+, k, l \in \mathbf{R}$, 则

(1) $a \oplus b = ab = ba = b \oplus a$;

(2)$(a \oplus b) \oplus c = (ab) \oplus c = (ab)c = a(bc) = a \oplus (bc) = a \oplus (b \oplus c)$;

(3) $a \oplus 1 = a1 = a$;

(4) $a \oplus \frac{1}{a} = a \times \frac{1}{a} = 1$;

(5) $1 \cdot a = a^1 = a$;

(6) $(kl) \cdot a = a^{kl} = a^{lk} = \left(a^l\right)^k = k \cdot a^l = k \cdot (l \cdot a)$;

(7) $(k+l)\cdot a = a^{k+l} = a^k a^l = a^k \oplus a^l = k\cdot a \oplus l\cdot a$;

(8) $k\cdot(a\oplus b) = k\cdot(ab) = (ab)^k = a^k b^k = a^k \oplus b^k = k\cdot a \oplus k\cdot b$.

所以, $\mathbf{R}^+$关于“$\oplus$”与“$\cdot$”构成$\mathbf{R}$上的线性空间.

这个例子表明, 线性空间的加法与数量乘法完全是借用$\mathbf{R}^n$及$\mathbf{R}^{m\times n}$中运算的叫法. 而与通常的加、乘并无必然联系.

根据线性空间V的八条运算规则, 可得出下述几个简单但很有用的性质.

性质 3.1.1 线性空间V的零向量唯一, 任一向量的负向量唯一.

证明 设$\boldsymbol{\theta}_1, \boldsymbol{\theta}_2$是$V$的两个零向量, 由$\boldsymbol{\theta}_1$是零向量可得

$$\boldsymbol{\theta}_1 + \boldsymbol{\theta}_2 = \boldsymbol{\theta}_2$$

由$\boldsymbol{\theta}_2$是零向量可得

$$\boldsymbol{\theta}_1 + \boldsymbol{\theta}_2 = \boldsymbol{\theta}_1$$

所以$\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$, 即零向量唯一.

任取$\boldsymbol{\alpha} \in V$, 设$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$是$\boldsymbol{\alpha}$的两个负向量, 则

$$\boldsymbol{\alpha} + \boldsymbol{\beta}_1 = \boldsymbol{\theta}, \quad \boldsymbol{\alpha} + \boldsymbol{\beta}_2 = \boldsymbol{\theta}$$

于是

$$\boldsymbol{\beta}_1 = \boldsymbol{\beta}_1 + \boldsymbol{\theta} = \boldsymbol{\beta}_1 + (\boldsymbol{\alpha} + \boldsymbol{\beta}_2) = (\boldsymbol{\beta}_1 + \boldsymbol{\alpha}) + \boldsymbol{\beta}_2 = \boldsymbol{\theta} + \boldsymbol{\beta}_2 = \boldsymbol{\beta}_2$$

即$\boldsymbol{\alpha}$的负向量唯一.

性质 3.1.2 设V是数域F上的线性空间, 则对任意$\boldsymbol{\alpha} \in V, k \in F$, 都有

(1) $0\boldsymbol{\alpha} = \boldsymbol{\theta}, k\boldsymbol{\theta} = \boldsymbol{\theta}$;

(2) $k(-\boldsymbol{\alpha}) = (-k)\boldsymbol{\alpha} = -(k\boldsymbol{\alpha})$;

(3) 若$k\boldsymbol{\alpha} = \boldsymbol{\theta}$, 则$k=0$或$\boldsymbol{\alpha} = \boldsymbol{\theta}$.

证 利用线性空间运算的八条运算规则可得

(1) $k\boldsymbol{\alpha} = (k+0)\boldsymbol{\alpha} = k\boldsymbol{\alpha} + 0\boldsymbol{\alpha}$, 故$0\boldsymbol{\alpha} = \boldsymbol{\theta}$. 类似地, 由$k\boldsymbol{\alpha} + k\boldsymbol{\theta} = k(\boldsymbol{\alpha} + \boldsymbol{\theta}) = k\boldsymbol{\alpha}$可得$k\boldsymbol{\theta} = \boldsymbol{\theta}$.

(2) 由$k(-\boldsymbol{\alpha}) + k\boldsymbol{\alpha} = k[(-\boldsymbol{\alpha}) + \boldsymbol{\alpha}] = k\boldsymbol{\theta} = \boldsymbol{\theta}$可得

$$k(-\boldsymbol{\alpha}) = -(k\boldsymbol{\alpha})$$

同理可证, $(-k)\boldsymbol{\alpha} = -(k\boldsymbol{\alpha})$.

(3) 若$k=0$, 则$k\boldsymbol{\alpha} = 0\boldsymbol{\alpha} = \boldsymbol{\theta}$; 若$k \neq 0$, 则由$k\boldsymbol{\alpha} = \boldsymbol{\theta}$得

$$\boldsymbol{\alpha} = 1\boldsymbol{\alpha} = (\frac{1}{k} \times k)\boldsymbol{\alpha} = \frac{1}{k}(k\boldsymbol{\alpha}) = \frac{1}{k}\boldsymbol{\theta} = \boldsymbol{\theta}$$

§3.2 基、维数与坐标

由于线性空间中的两种线性运算与F^n中的两种线性运算都满足性质**??**, 于是, 可以仿照F^n中的做法, 在线性空间中引入线性组合、线性表出、线性相关与线性无关、向量组的秩以及极大无关组等概念, 并且对F^n中向量组成立的性质、结论等也完全可以推广到线性空间中, 这里不再一一赘述. 下面仅举例说明.

例 3.2.1 在线性空间$\mathbf{R}^{2\times 2}$中, 因为

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = 1\cdot\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 2\cdot\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + 3\cdot\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + 4\cdot\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

所以, 矩阵$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$是矩阵$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$的线性组合.

例 3.2.2 线性空间$\mathbf{R}^{2\times 2}$中的一组向量(矩阵)

$$\boldsymbol{I}_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \boldsymbol{I}_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \boldsymbol{I}_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \boldsymbol{I}_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

是线性无关的.

证 假设

$$k_1\cdot\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + k_2\cdot\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + k_3\cdot\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + k_4\cdot\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

整理得

$$\begin{bmatrix} k_1 & k_2 \\ k_3 & k_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

所以$k_1 = k_2 = k_3 = k_4 = 0$, 于是$\boldsymbol{I}_{11}, \boldsymbol{I}_{12}, \boldsymbol{I}_{21}, \boldsymbol{I}_{22}$ 线性无关.

例 3.2.3 线性空间$F[x]_4$中的一组向量(多项式)

$$x^3 + x, x^2 + 1, x + 1$$

是线性无关的.

证 假设

$$k_1(x^3+x)+k_2(x^2+1)+k_3(x+1)=\boldsymbol{\theta}$$

整理得

$$k_1x^3+k_2x^2+(k_1+k_3)x+(k_2+k_3)=\boldsymbol{\theta}$$

由零多项式及多项式相等的定义得$k_1=k_2=k_1+k_3=k_2+k_3=0$, 从而$k_1=k_2=k_3=0$, 于是$x^3+x, x^2+1, x+1$是多项式空间$F[x]_4$中线性无关的向量组.

例 3.2.4 试求函数空间$C[a,b]$中的向量(函数)组

$$\boldsymbol{\alpha}_1=\mathbf{1},\quad \boldsymbol{\alpha}_2=\cos^2 x,\quad \boldsymbol{\alpha}_3=\cos 2x$$

的一个极大无关组.

证 显然

$$\cos 2x=2\cdot\cos^2 x-\mathbf{1}$$

因此$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性相关. 又因$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$不能相互线性表出, 所以$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$线性无关, 并为向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$的一个极大无关组, 且$\mathrm{r}\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}=2$.

§3.2.1 基与维数

定义 3.2.1 若能从线性空间V中找到有限个向量$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m$使$V$中任一向量均可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m$线性表出, 则称$V$是**有限维线性空间**; 否则, 就称$V$是**无限维线性空间**.

若V是有限维线性空间, 则存在$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m\in V$使$V$中任一向量均可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m$线性表出. 任取$V$中$k$个向量$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_k$, 只要$k>m$, 则$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_k$线性相关. 这就表明$V$中线性无关向量组所含向量最大个数不超过$m$, 即有上界.

例 3.2.5 向量空间F^n、矩阵空间$F^{m\times n}$、多项式空间$F[x]_n$都是有限维的, 而多项式空间$F[x]$与函数空间$C[a,b]$都是无限维的.

证 令$\boldsymbol{\varepsilon}_1=(1,0,\ldots,0),\boldsymbol{\varepsilon}_2=(0,1,\ldots,0),\ldots,\boldsymbol{\varepsilon}_n=(0,\ldots,0,1)$, 则$\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2,\ldots,\boldsymbol{\varepsilon}_n\in F^n$. 对任意$\boldsymbol{\alpha}=(a_1,a_2,\ldots,a_n)\in F^n$, 有

$$\boldsymbol{\alpha}=a_1\boldsymbol{\varepsilon}_1+a_2\boldsymbol{\varepsilon}_2+\cdots+a_n\boldsymbol{\varepsilon}_n$$

所以F^n是有限维的.

令$\boldsymbol{I}_{ij}(i=1,2,\ldots,m;j=1,2,\ldots,n)$表示$(i,j)-$元为1, 其余元素均为零的$m\times n$矩阵, 则这些矩阵均在$F^{m\times n}$中, 对任意$\boldsymbol{A}=[a_{ij}]\in F^{m\times n}$, 有

$$\boldsymbol{A}=\sum_{i,j}a_{ij}\boldsymbol{I}_{ij}$$

故$F^{m\times n}$是有限维的.

令$f_1=1,f_2=x,f_3=x^2,\ldots,f_n=x^{n-1}$, 则对任意$f(x)=a_0+a_1x+a_2x^2+\cdots+a_{n-1}x^{n-1}\in F[x]_n$, 均有

$$f(x)=a_0f_1+a_1f_2+\cdots+a_{n-1}f_n$$

所以$F[x]_n$是有限维的.

由定义3.2.1后的讨论可知, 有限维线性空间中线性无关向量组所含向量的个数有上界, 故只要表明$F[x]$中有任意多个向量组成的线性无关向量组就可说明它是无限维的. 对任一正整数N, 考虑$F[x]$中的N个多项式$1,x,x^2,\ldots,x^{N-1}$, 显然, 对任意N个不全为零的数$a_0,a_1,a_2,\ldots,a_{N-1}$, 多项式

$$a_0+a_1x+a_2x^2+\cdots+a_{N-1}x^{N-1}$$

不是零多项式, 即$a_0+a_1x+\cdots+a_{N-1}x^{N-1}\neq\boldsymbol{\theta}$, 所以$1,x,x^2,\ldots,x^{N-1}$线性无关. 由此得, $F[x]$是无限维的.

同理可证, $C[a,b]$也是无限维的.

本书只讨论有限维线性空间, 但在某些例题中也会使用$F[x]$与$C[a,b]$.

定义 3.2.2 *设V是数域F上的一个有限维线性空间. 若V中m个向量$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m$满足*

(1) $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m$线性无关,

(2) V中任一向量均可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m$线性表出,

*则称$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m$是$V$的一个(组)***基***, 称$m$为$V$的***维数***, 记为$\dim_F V$, 简记为$\dim V$.*

从上述定义不难看出, 线性空间的基实际上就是线性空间这个"向量组"的一个极大无关组, 因此, 基包含的向量个数是唯一确定的. 我们规定只含零向量的线性空间是0维的.

例 3.2.6 *F^n的维数是n, $F^{m\times n}$的维数是$m\times n$, $F[x]_n$的维数是n.*

容易验证, $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\ldots,\boldsymbol{\varepsilon}_n$; $\boldsymbol{I}_{11},\ldots,\boldsymbol{I}_{mn}$与$1,x,x^2,\ldots,x^{n-1}$均线性无关, 再根据例3.2.5, F^n有基

$$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\ldots,\boldsymbol{\varepsilon}_n$$

称之为F^n**的自然基**, $F^{m\times n}$有基

$$\boldsymbol{I}_{11},\boldsymbol{I}_{12},\ldots,\boldsymbol{I}_{1n},\boldsymbol{I}_{21},\boldsymbol{I}_{22},\ldots,\boldsymbol{I}_{2n},\ldots,\boldsymbol{I}_{m1},\boldsymbol{I}_{m2},\ldots,\boldsymbol{I}_{mn}$$

称之为$F^{m\times n}$**的自然基**, $F[x]_n$有基

$$1,x,x^2,\ldots,x^{n-1}$$

称之为$F[x]_n$**的自然基**.

定理 3.2.1 n维线性空间V中任意n个线性无关的向量均构成V的一个基.

证 任取V中n个线性无关的向量$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n$, 只需证$V$中任一向量均可由$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n$线性表出.

任取$\boldsymbol{\alpha}\in V$及V的一个基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$, 则$\boldsymbol{\alpha},\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$线性表出. 根据定理**??**, $\boldsymbol{\alpha},\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m$线性相关, 又$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n$线性无关, 所以$\boldsymbol{\alpha}$可由$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n$线性表出.

§3.2.2 坐标

定义 3.2.3 设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$是数域$F$上的$n$维线性空间$V$的一个基, 则$V$中任一向量$\boldsymbol{\alpha}$可唯一地表示为

$$\boldsymbol{\alpha}=a_1\boldsymbol{\alpha}_1+a_2\boldsymbol{\alpha}_2+\cdots+a_n\boldsymbol{\alpha}_n$$

其中$a_i\in F, i=1,2,\cdots,n$, 称有序数组$a_1,a_2,\ldots,a_n$为$\boldsymbol{\alpha}$**关于基**$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$**的坐标**, 记为

$$(a_1,a_2,\ldots,a_n)^{\mathrm{T}}$$

例 3.2.7 F^n中任一向量$\boldsymbol{\alpha}=(a_1,a_2,\ldots,a_n)$关于基$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\ldots,\boldsymbol{\varepsilon}_n$的坐标为$(a_1,a_2,\ldots,a_n)^{\mathrm{T}}$; $F^{m\times n}$中任一矩阵$\boldsymbol{A}=[a_{ij}]$关于基$\boldsymbol{I}_{11},\ldots,\boldsymbol{I}_{1n}$, $\ldots$, $\boldsymbol{I}_{m1}$, $\ldots$, $\boldsymbol{I}_{mn}$的坐标为$(a_{11},\ldots,a_{1n},\ldots,a_{m1},\ldots,a_{mn})^{\mathrm{T}}$; $F[x]_n$中任一多项式$f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}$关于基$1,x,\ldots,x^{n-1}$的坐标为$(a_0,a_1,\ldots,a_{n-1})^{\mathrm{T}}$.

例 3.2.8 已知$\mathbf{R}^3$中的3个向量:

$$\boldsymbol{\alpha}_1=(1,1,1),\quad \boldsymbol{\alpha}_2=(1,1,0),\quad \boldsymbol{\alpha}_3=(1,0,0)$$

(1) 证明: $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$是$\mathbf{R}^3$的一个基;

(2) 求向量$\boldsymbol{\alpha}=(1,2,3)$关于基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$的坐标.

解 (1) 因$\mathbf{R}^3$是3维向量空间, 故只需证明$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性无关. 以$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$为列构造矩阵$\boldsymbol{A}$, 并利用初等变换将之化为阶梯形:

$$\boldsymbol{A}=\begin{bmatrix}1&1&1\\1&1&0\\1&0&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&1&1\\0&1&1\\0&0&1\end{bmatrix}$$

由此可得$\mathrm{r}\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}=3$, 即$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性无关.

(2) 令

$$\boldsymbol{\alpha}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+x_3\boldsymbol{\alpha}_3$$

则有,

$$\begin{cases}x_1+x_2\ \ +x_3=1\\x_1+x_2\qquad\ \ =2\\x_1\qquad\qquad\ \ =3\end{cases}$$

解得$x_1=3,x_2=-1,x_3=-1$. 故$\boldsymbol{\alpha}$关于基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$的坐标为$(3,-1,-1)^{\mathrm{T}}$.

例 3.2.9 在$\mathbf{R}^{2\times2}$中证明矩阵

$$\boldsymbol{A}_1=\begin{bmatrix}1&1\\1&1\end{bmatrix},\boldsymbol{A}_2=\begin{bmatrix}1&1\\-1&-1\end{bmatrix},\boldsymbol{A}_3=\begin{bmatrix}1&-1\\1&-1\end{bmatrix},\boldsymbol{A}_4=\begin{bmatrix}-1&1\\1&-1\end{bmatrix}$$

构成一个基, 并求矩阵$\boldsymbol{A}=\begin{bmatrix}1&2\\3&4\end{bmatrix}$关于这个基的坐标.

解 已知矩阵空间$\mathbf{R}^{2\times2}$的维数为4, 根据定理3.2.1, 只需验证$\boldsymbol{A}_1,\boldsymbol{A}_2,\boldsymbol{A}_3,\boldsymbol{A}_4$线性无关.

令$k_1\boldsymbol{A}_1+k_2\boldsymbol{A}_2+k_3\boldsymbol{A}_3+k_4\boldsymbol{A}_4=\mathbf{0}$, 则有

$$\begin{cases}k_1+k_2+k_3-k_4=0\\k_1+k_2-k_3+k_4=0\\k_1-k_2+k_3+k_4=0\\k_1-k_2-k_3-k_4=0\end{cases}$$

容易验证上述齐次线性方程组只有零解, 即$k_1 = k_2 = k_3 = k_4 = 0$, 所以$\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3$, $\boldsymbol{A}_4$线性无关.

令$\boldsymbol{A} = x_1\boldsymbol{A}_1 + x_2\boldsymbol{A}_2 + x_3\boldsymbol{A}_3 + x_4\boldsymbol{A}_4$, 则有

$$\begin{cases} x_1 + x_2 + x_3 - x_4 = 1 \\ x_1 + x_2 - x_3 + x_4 = 2 \\ x_1 - x_2 + x_3 + x_4 = 3 \\ x_1 - x_2 - x_3 - x_4 = 4 \end{cases}$$

不难验证上述线性方程组有唯一解: $x_1 = \frac{5}{2}, x_2 = -1, x_3 = -\frac{1}{2}$, $x_4 = 0$. 所以, $\boldsymbol{A} = \frac{5}{2}\boldsymbol{A}_1 - \boldsymbol{A}_2 - \frac{1}{2}\boldsymbol{A}_3 + 0\boldsymbol{A}_4$.

对一般的线性空间V, 当取定一个基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$后, V中向量$\boldsymbol{\alpha}$与其关于$\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, ..., $\boldsymbol{\alpha}_n$的坐标$(a_1, a_2, \ldots, a_n)^{\mathrm{T}}$一一对应, 可表示为

$$\boldsymbol{\alpha} \longleftrightarrow (a_1, a_2, \ldots, a_n)^{\mathrm{T}}$$

在同一个基下, 设

$$\boldsymbol{\beta} \longleftrightarrow (b_1, b_2, \ldots, b_n)^{\mathrm{T}}$$

则

$$\begin{gathered} \boldsymbol{\alpha} + \boldsymbol{\beta} \longleftrightarrow (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)^{\mathrm{T}} \\ k\boldsymbol{\alpha} \longleftrightarrow (ka_1, ka_2, \ldots, ka_n)^{\mathrm{T}} \end{gathered}$$

这样, 线性空间中抽象元素之间的运算可转化为具体的n元向量间的线性运算. 这一点如用第§3.6.4节中线性空间同构的语言来描述, 就是显而易见的了.

§3.2.3 基变换与坐标变换

通过例3.2.8可以看到, $\boldsymbol{\alpha}$关于基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$的坐标, 与关于$\mathbf{R}^3$的自然基的坐标是不一样的. 一般地, 同一个向量关于不同基的坐标是不同的. 本节我们就来讨论同一个向量在不同基下的坐标之间的关系.

定义 3.2.4 设V是数域F上的n维线性空间, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$与$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots,$ $\boldsymbol{\beta}_n$是V的两个基, 且

$$\begin{cases} \boldsymbol{\beta}_1 = a_{11}\boldsymbol{\alpha}_1 + a_{21}\boldsymbol{\alpha}_2 + \cdots + a_{n1}\boldsymbol{\alpha}_n \\ \boldsymbol{\beta}_2 = a_{12}\boldsymbol{\alpha}_1 + a_{22}\boldsymbol{\alpha}_2 + \cdots + a_{n2}\boldsymbol{\alpha}_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \boldsymbol{\beta}_n = a_{1n}\boldsymbol{\alpha}_1 + a_{2n}\boldsymbol{\alpha}_2 + \cdots + a_{nn}\boldsymbol{\alpha}_n \end{cases} \tag{3.1}$$

式(3.1)可形式上写成如下矩阵的形式:

$$[\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n] = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\, \boldsymbol{A} \tag{3.2}$$

其中

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

则称$\boldsymbol{A}$是基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$到基$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$的**过渡矩阵**. 称式(3.1)和式(3.2)为**基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$到基$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$的基变换公式**.

式(3.2)之所以称为是“形式上”的, 是因为这里将向量作为矩阵的元素. 可以证明这个形式写法满足以下规则:

设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$与$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$是$V$中两个向量组, $\boldsymbol{A}, \boldsymbol{B} \in F^{n\times n}$, $k \in F$, 则

$$\begin{aligned} ([\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\, \boldsymbol{A})\, \boldsymbol{B} &= [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\, (\boldsymbol{AB}) \\ [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\, \boldsymbol{A} + [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\, \boldsymbol{B} &= [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\, (\boldsymbol{A} + \boldsymbol{B}) \\ [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\, \boldsymbol{A} + [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n]\, \boldsymbol{A} &= [\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1, \boldsymbol{\alpha}_2 + \boldsymbol{\beta}_2, \ldots, \boldsymbol{\alpha}_n + \boldsymbol{\beta}_n]\, \boldsymbol{A} \\ (k\, [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n])\, \boldsymbol{A} &= [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\, (k\boldsymbol{A}) \\ &= k\, ([\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\, \boldsymbol{A}) \end{aligned}$$

根据定义3.2.4, 过渡矩阵$\boldsymbol{A}$的每列都是向量关于基的坐标, 因此过渡矩阵$\boldsymbol{A}$是被基唯一确定的. 特别地, 一个基到自身的过渡矩阵为$\boldsymbol{I}$.

如果设基$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$到基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$的过渡矩阵为$\boldsymbol{B}$, 即$[\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n] = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n]\, \boldsymbol{B}$, 将之代入式(3.2)有

$$[\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n] = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n]\, (\boldsymbol{BA})$$

于是$\boldsymbol{BA} = \boldsymbol{I}$, 即$\boldsymbol{A}$是可逆的, 且$\boldsymbol{B} = \boldsymbol{A}^{-1}$. 综上所述, 关于过渡矩阵可得:

(1) 过渡矩阵是唯一确定的;

(2) 过渡矩阵是可逆的;

(3) 若$\boldsymbol{A}$是基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$到基$\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, …, $\boldsymbol{\beta}_n$的过渡矩阵, 则$\boldsymbol{A}^{-1}$是$\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, …, $\boldsymbol{\beta}_n$到$\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, …, $\boldsymbol{\alpha}_n$的过渡矩阵;

(4) 基变换公式(3.2)可按普通矩阵那样进行运算.

在定义3.2.4的条件下, 任取$\boldsymbol{\gamma} \in V$, 设$\boldsymbol{\gamma}$关于基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$的坐标为$(x_1, x_2, \ldots, x_n)^{\mathrm{T}}$, 关于基$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$的坐标为$(y_1, y_2, \ldots, y_n)^{\mathrm{T}}$, 即有

$$\boldsymbol{\gamma} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n$$

和

$$\boldsymbol{\gamma} = y_1\boldsymbol{\beta}_1 + y_2\boldsymbol{\beta}_2 + \cdots + y_n\boldsymbol{\beta}_n$$

将上面两式形式地改写为

$$\boldsymbol{\gamma} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{3.3}$$

$$\boldsymbol{\gamma} = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

将式(3.2)代入上式, 得到

$$\boldsymbol{\gamma} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\, \boldsymbol{A} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

将上式与式(3.3)对比, 根据坐标的唯一性可得

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{3.4}$$

于是有如下结果:

定理 3.2.2 设V是数域F上的n维线性空间, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$与$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$是$V$的两个基, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$到$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$的过渡矩阵为$\boldsymbol{A}$. 任取$\boldsymbol{\gamma} \in$

V, 设$\boldsymbol{\gamma}$关于基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$的坐标为$(x_1, x_2, \ldots, x_n)^{\mathrm{T}}$, 关于基$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$的坐标为$(y_1, y_2, \ldots, y_n)^{\mathrm{T}}$, 则

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \text{或} \quad \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \boldsymbol{A}^{-1} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{3.5}$$

称式(3.5)为**基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$到基$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$的坐标变换公式**.

例 3.2.10 已知$\boldsymbol{\beta}_1 = (1, 1), \boldsymbol{\beta}_2 = (-1, 1)$为$\mathbf{R}^2$的一个基, 求$\mathbf{R}^2$的自然基$\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2$到基$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$的过渡矩阵, 并计算$\boldsymbol{\alpha} = (1, 3)$关于基$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$的坐标.

解 因$\boldsymbol{\beta}_1 = \boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2, \boldsymbol{\beta}_2 = -\boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2$, 整理得

$$[\boldsymbol{\beta}_1, \boldsymbol{\beta}_2] = [\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2] \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

即所求过渡矩阵为

$$\boldsymbol{A} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

设$\boldsymbol{\alpha}$关于基$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$的坐标为$(y_1, y_2)^{\mathrm{T}}$, 因$\boldsymbol{\alpha} = \boldsymbol{\varepsilon}_1 + 3\boldsymbol{\varepsilon}_2$, 即关于基$\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2$的坐标为$(1, 3)^{\mathrm{T}}$(见图3.2.1), 由式(3.5)得

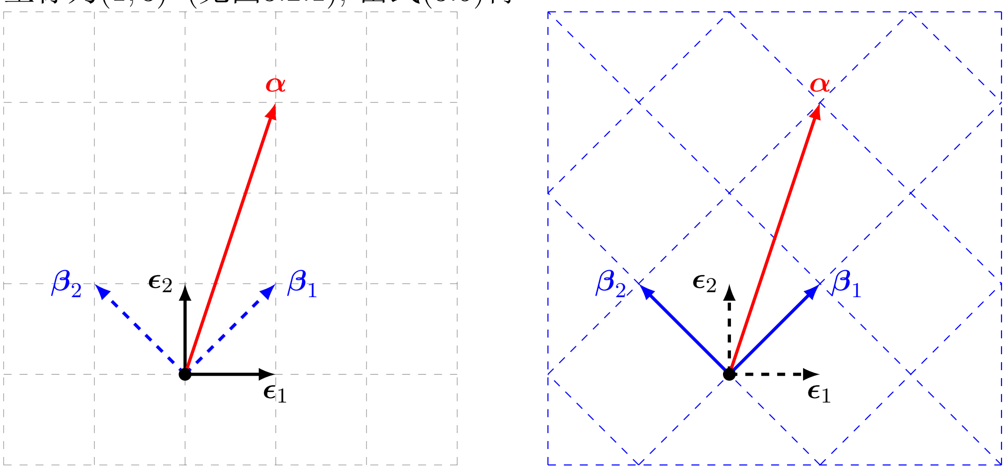

图 3.2.1: 坐标变换

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \boldsymbol{A}^{-1} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

即$\boldsymbol{\alpha}$关于$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$的坐标为$(2, 1)^{\mathrm{T}}$(见图3.2.1).

例 3.2.11 求$F[x]_4$的自然基到基

$$g_0 = 1, g_1 = 1 + x, g_2 = 1 + x + x^2, g_3 = 1 + x + x^2 + x^3$$

的过渡矩阵. 已知多项式$h(x)$关于基g_0, g_1, g_2, g_3的坐标为$(7, 0, 8, -2)^{\mathrm{T}}$, 求$h(x)$关于自然基的坐标.

解 将$F[x]_4$的自然基记为$f_0 = 1, f_1 = x, f_2 = x^2, f_3 = x^3$, 则有

$$\begin{cases} g_0 = f_0 \\ g_1 = f_0 + f_1 \\ g_2 = f_0 + f_1 + f_2 \\ g_3 = f_0 + f_1 + f_2 + f_3 \end{cases}$$

整理得

$$[g_0, g_1, g_2, g_3] = [f_0, f_1, f_2, f_3] \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

故基f_0, f_1, f_2, f_3到基g_0, g_1, g_2, g_3的过渡矩阵为

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

设$h(x)$关于基f_0, f_1, f_2, f_3的坐标为$(x_1, x_2, x_3, x_4)^{\mathrm{T}}$, 则由坐标变换公式(3.5)得$(x_1, x_2, x_3, x_4)^{\mathrm{T}} = \boldsymbol{A}(7, 0, 8, -2)^{\mathrm{T}} = (13, 6, 6, -2)^{\mathrm{T}}$.

性质 3.2.1 设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$是数域$F$上线性空间$V$的一个基, 且

$$[\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n] = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\, \boldsymbol{A}$$

其中$\boldsymbol{A} \in F^{n\times n}$, 则$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$是$V$的一个基当且仅当$\boldsymbol{A}$是可逆矩阵.

证 充分性: 只需证明$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$线性无关. 不妨设$k_1\boldsymbol{\beta}_1 + k_2\boldsymbol{\beta}_2 + \cdots + k_n\boldsymbol{\beta}_n = \mathbf{0}$, 即

$$[\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n] \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\, \boldsymbol{A} \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} = \mathbf{0}$$

由$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$线性无关可得

$$\boldsymbol{A}\begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} = \mathbf{0}$$

由$\boldsymbol{A}$为可逆矩阵可知$k_1 = k_2 = \cdots = k_n = 0$, 因此, $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$线性无关, 进而是$V$的一个基.

必要性: 由定理**??**可知, 只需证明齐次线性方程组$\boldsymbol{AX} = \mathbf{0}$仅有零解. 设$(k_1, k_2, \ldots, k_n)^{\mathrm{T}}$为方程组$\boldsymbol{AX} = \mathbf{0}$的任意一个解, 则

$$\begin{aligned} & k_1\boldsymbol{\beta}_1 + k_2\boldsymbol{\beta}_2 + \cdots + k_n\boldsymbol{\beta}_n \\ = & [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n]\begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\,\boldsymbol{A}\begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} = \mathbf{0} \end{aligned}$$

由$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$是$V$的一个基可知, $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$线性无关, 因此, $k_1 = k_2 = \cdots = k_n = 0$, 即方程组$\boldsymbol{AX} = \mathbf{0}$仅有零解, 从而$\boldsymbol{A}$为可逆矩阵.

§3.3 线性子空间

在研究代数系统时, 有很多途径, 其中之一就是通过研究一部分元素来获取整个系统的某些性质.

§3.3.1 线性子空间的概念

在平面解析几何中, 过原点的共线向量集合保留了平面向量的加法与数量乘法运算, 依然构成$\mathbf{R}$上的线性空间. 数域F上的多项式空间$F[x]$中的子集$F[x]_n$也是F上一个线性空间. 一般地,

定义 3.3.1 设V是数域F上的线性空间, W是V的非空子集. 若W对V的两种线性运算也构成F上的线性空间, 则称W是V的**线性子空间**, 简称**子空间**.

显然, $F[x]_n$是$F[x]$的子空间, 然而, 例3.1.6中的线性空间$\mathbf{R}^+$却不是实数域$\mathbf{R}$上的线性空间$\mathbf{R}$的子空间, 这是因为这两个线性空间的线性运算的定义不相同.

例 3.3.1 设V是数域F上线性空间, 则V一定包含零向量$\boldsymbol{\theta}$. 同时, V本身及$\{\boldsymbol{\theta}\}$都是V的子空间, 称它们为**平凡子空间**. V的其他子空间, 如果还有的话, 均称为**非平凡子空间**.

例 3.3.2 设$\boldsymbol{A} \in F^{m\times n}$, 则矩阵$\boldsymbol{A}$的零空间$\mathrm{N}(\boldsymbol{A})$为$F^n$的子空间.

那么, 一个数域F上线性空间V的非空子集W什么情况下才能构成V的子空间呢?

首先, W应对V的两种线性运算**封闭**, 即对任意$\boldsymbol{\alpha}, \boldsymbol{\beta} \in W, k \in F$均有$\boldsymbol{\alpha}+\boldsymbol{\beta}$, $k\boldsymbol{\alpha} \in W$, 这样$V$的两种运算也可作为$W$的两种运算. 其次, W应对这两种线性运算满足线性空间的八条运算规则. 不难发现, W显然满足定义3.1.1中第(1)、(2)、(5)、(6)、(7)、(8)条性质, 这是因为W中的向量也是V中的向量. 现在唯一需要确定的是, V中的零向量与W中向量的负向量是否也在W中. 实际上, 由W对V的两种运算的封闭性即可保证第(3)、(4)条性质也成立, 这里只需取$k=0$和$k=-1$就会有$0\boldsymbol{\alpha}=\boldsymbol{\theta}, (-1)\boldsymbol{\alpha}=-\boldsymbol{\alpha} \in W$. 于是得到子空间的下述判别定理:

定理 3.3.1 设V是数域F上的线性空间, W是V的非空子集. 若W满足
(1)对$\forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in W$, 均有$\boldsymbol{\alpha}+\boldsymbol{\beta} \in W$,
(2)对$\forall \boldsymbol{\alpha} \in W, \forall k \in F$, 均有$k\boldsymbol{\alpha} \in W$, 则$W$是$V$的子空间.

例 3.3.3 考虑向量空间$\mathbf{R}^3$中的两个非空集合

$$\begin{aligned} V_1 &= \{(x,y,z) \in \mathbf{R}^3 \mid x+y+z=0\} \\ V_2 &= \{(x,y,z) \in \mathbf{R}^3 \mid x+y+z=1\} \end{aligned}$$

问V_1和V_2是否构成$\mathbf{R}^3$的子空间?

解 V_1可以看成齐次线性方程组$x+y+z=0$的解的集合, 因此V_1是$\mathbf{R}^3$的子空间.

考虑$\boldsymbol{\alpha} = (2,-3,2) \in V_2$, $k=4$, 则$k\boldsymbol{\alpha} = 4(2,-3,2) = (8,-12,8)$. 因$8+(-12)+8=4\neq 1$, 故$k\boldsymbol{\alpha} \notin V_2$, 即$V_2$对数乘运算不封闭. 因此, V_2不构成$\mathbf{R}^3$的子空间(也可通过V_2不包含零向量来说明V_2不构成子空间).

显然, V_1和V_2均可以看作空间中的两个平面, 如图3.3.1所示.

对一般的线性空间, 可由其中的任一组向量构造一种特殊且很重要的子空间.

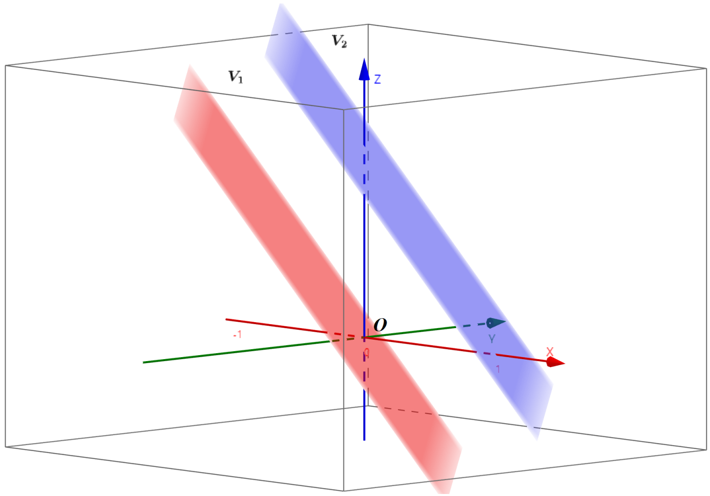

图 3.3.1: V_1与V_2

定理 3.3.2 设V是数域F上的线性空间, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m$是$V$中$m$个向量, 则$V$的子集合

$$\left\{\sum_{i=1}^{m} k_i\boldsymbol{\alpha}_i \mid k_1, k_2, \ldots, k_m \in F\right\}$$

构成V的子空间, 称为**由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m$生成的子空间**, 记为$L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m)$.

证 显然V是一个非空集合.

任取$L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m)$中的两个向量$\boldsymbol{\beta} = \sum\limits_{i=1}^{m} k_i\boldsymbol{\alpha}_i, \boldsymbol{\gamma} = \sum\limits_{i=1}^{m} l_i\boldsymbol{\alpha}_i$, 其中, $k_i, l_i \in F, i = 1, 2, \ldots, m$及$c \in F$, 因有

$$\boldsymbol{\beta} + \boldsymbol{\gamma} = \sum_{i=1}^{m} k_i\boldsymbol{\alpha}_i + \sum_{i=1}^{m} l_i\boldsymbol{\alpha}_i = \sum_{i=1}^{m}(k_i + l_i)\boldsymbol{\alpha}_i$$

$$c\boldsymbol{\beta} = c(\sum_{i=1}^{m} k_i\boldsymbol{\alpha}_i) = \sum_{i=1}^{m}(ck_i)\boldsymbol{\alpha}_i$$

都属于$L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m)$, 故$L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m)$构成$V$的子空间.

显然, $\mathbf{R}^2 = L((1,0),(0,1))$, $F[x]_n = L(1, x, x^2, \ldots, x^{n-1})$, 而 $C[-\pi, \pi]$有无穷多个子空间$C_n = L(\cos nx, \sin nx)(n = 1, 2\cdots)$.

由向量空间$\mathbf{R}^3$中两个向量$\boldsymbol{\alpha} = (1, 4, 1), \boldsymbol{\beta} = (1, -1, 1)$生成的子空间$L(\boldsymbol{\alpha}, \boldsymbol{\beta})$如图3.3.2所示. 扫描右侧二维码, 选择$\mathbf{R}^3$中两个不同的向量, 观察它们生成的子空间.

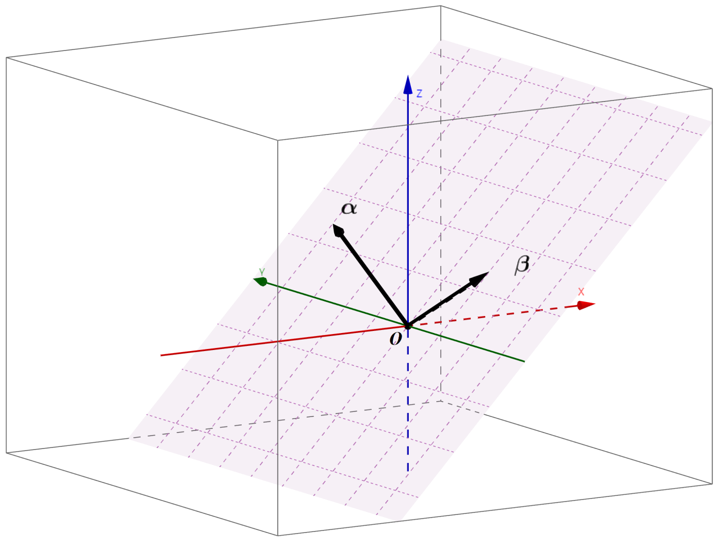

图 3.3.2: 生成子空间

例 3.3.4 设$\boldsymbol{X}_1,\boldsymbol{X}_2,...,\boldsymbol{X}_t$是齐次线性方程组$\boldsymbol{AX}=\boldsymbol{0}$的一个基础解系, 则

$$\mathrm{N}(\boldsymbol{A})=L\left(\boldsymbol{X}_1,\boldsymbol{X}_2,...,\boldsymbol{X}_t\right)$$

虽然基础解系能生成解空间, 但解空间的生成向量组未必都是基础解系. 不过有一点可以肯定, 即解空间的生成向量组一定包含一个基础解系.

例 3.3.5 设$\boldsymbol{A}\in F^{m\times n}$, 把$\boldsymbol{A}$按列分块$\boldsymbol{A}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...,\boldsymbol{\alpha}_n]$, 则$L\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...,\boldsymbol{\alpha}_n\right)$是$F^m$的子空间, 称之为**矩阵$\boldsymbol{A}$的列空间**或**$\boldsymbol{A}$的值域**, 常记为$\mathrm{Col}(\boldsymbol{A})$或者$\mathrm{R}(\boldsymbol{A})$.

显然, 线性方程组$\boldsymbol{AX}=\boldsymbol{b}$有解的充分必要条件是$\boldsymbol{b}\in\mathrm{Col}(\boldsymbol{A})$. 此外, 由$\boldsymbol{A}$的行向量组生成的子空间, 称为**矩阵$\boldsymbol{A}$的行空间**, 记为$\mathrm{Row}(\boldsymbol{A})$.

$\mathrm{N}(\boldsymbol{A})$、$\mathrm{Col}(\boldsymbol{A})$、$\mathrm{Row}\left(\boldsymbol{A}\right)$是与$\boldsymbol{A}$相关的三个重要的向量空间, 在许多理论和实际问题中都有应用.

例 3.3.6 设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...,\boldsymbol{\alpha}_s$与$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,...,\boldsymbol{\beta}_t$是线性空间$V$的两组向量, 则$L\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...,\boldsymbol{\alpha}_s\right)=L\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,...,\boldsymbol{\beta}_t\right)$的充分必要条件是$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...,\boldsymbol{\alpha}_s\}\cong\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,...,\boldsymbol{\beta}_t\}$.

证 充分性: 设$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s\}\cong\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_t\}$. 任取$\boldsymbol{\gamma}\in L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s)$, 则$\boldsymbol{\gamma}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s$线性表出. 又$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s$可由$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_t$线性表出, 故$\boldsymbol{\gamma}$可由$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_t$线性表出. 所以$\boldsymbol{\gamma}\in L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_t)$, 由此得$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s)\subseteq L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_t)$.

同理可证$L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_t)\subseteq L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s)$. 于是$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s) = L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_t)$.

必要性: 设$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s)=L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_t)$. 因为$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s\in L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s)=L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_t)$, 所以$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s$可由$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_t$线性表出. 同理, $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_t$也可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s$线性表出. 所以$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s\}\cong\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_t\}$.

例 3.3.7 向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m$的极大无关组都是生成子空间$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m)$的基, 且有$\dim(L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m))=\mathrm{r}\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m\}$.

证 任取$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m$的一个极大无关组$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\ldots,\boldsymbol{\alpha}_{i_r}$, 则每个$\boldsymbol{\alpha}_j(j=1,2,\ldots,m)$均可由$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\ldots,\boldsymbol{\alpha}_{i_r}$线性表出. 又$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m)$中任一向量$\boldsymbol{\alpha}$均可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m$线性表出, 故$\boldsymbol{\alpha}$也可由$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\ldots,\boldsymbol{\alpha}_{i_r}$线性表出. 从而$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\ldots,\boldsymbol{\alpha}_{i_r}$是$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m)$的基.

根据线性空间维数和向量组秩的定义, 立即可得$\dim(L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m))=\mathrm{r}\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m\}$.

例 3.3.8 设$\boldsymbol{A}\in F^{m\times n}$, $\mathrm{r}(\boldsymbol{A})=r(1\leqslant r<n)$, 任取齐次线性方程组$\boldsymbol{AX}=\boldsymbol{0}$的一个基础解系$\boldsymbol{X}_1,\boldsymbol{X}_2,\ldots,\boldsymbol{X}_{n-r}$, 容易看出它们构成$\mathrm{N}(\boldsymbol{A})$的一个基, 因此$\dim(\mathrm{N}(\boldsymbol{A}))=n-r$.

由上例可得, 矩阵$\boldsymbol{A}$的零空间$\mathrm{N}(\boldsymbol{A})$与行空间$\mathrm{Row}(\boldsymbol{A})$之间的一个重要关系:

定理 3.3.3 设$\boldsymbol{A}\in F^{m\times n}$, 则

$$\dim(\mathrm{N}(\boldsymbol{A}))+\dim(\mathrm{Row}(\boldsymbol{A}))=n \tag{3.6}$$

例 3.3.9 已知$\mathbf{R}^4$中的3个向量

$$\boldsymbol{\alpha}_1=(1,2,0,1),\boldsymbol{\alpha}_2=(-1,1,1,1),\boldsymbol{\alpha}_3=(4,14,2,8)$$

求$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)$的一个基及维数, 并将这个基扩充为$\mathbf{R}^4$的一个基.

解 根据例3.3.8, $L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)$ 的基可取$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$的一个极大无关组. 以$\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$为列向量构造矩阵$\boldsymbol{A}$, 用初等行变换将之化为阶梯形:

$$\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 4 \\ 2 & 1 & 14 \\ 0 & 1 & 2 \\ 1 & 1 & 8 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -1 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

故得$\mathrm{r}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\} = 2$, 且$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$是其一个极大无关组. 所以$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$是$L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)$的一个基, 并且$L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)$的维数为2.

因$\mathbf{R}^4$是4维向量空间, 故只需找$\boldsymbol{\alpha}_4$, $\boldsymbol{\alpha}_5 \in \mathbf{R}^4$使$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_4, \boldsymbol{\alpha}_5$线性无关. 以$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$为前两列、4元基本向量$\boldsymbol{\varepsilon}_3, \boldsymbol{\varepsilon}_4$为后两列构造4阶方阵

$$\boldsymbol{B} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

容易验证$\boldsymbol{B}$是满秩的, 故$\mathrm{r}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\varepsilon}_3, \boldsymbol{\varepsilon}_4\} = 4$. 取$\boldsymbol{\alpha}_4 = \boldsymbol{\varepsilon}_3, \boldsymbol{\alpha}_5 = \boldsymbol{\varepsilon}_4$, 则$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$, $\boldsymbol{\alpha}_4, \boldsymbol{\alpha}_5$线性无关, 即是$\mathbf{R}^4$的一个基.

§3.3.2 子空间的运算

子空间作为子集合, 可进行集合的运算.

定理 3.3.4 设W_1, W_2是数域F上线性空间V的两个子空间, 则$W_1 \cap W_2$也是V的子空间, 称之为W_1与W_2的**交空间**.

证 因W_1, W_2是V的子空间, 故V的零向量$\boldsymbol{\theta}$同时属于W_1, W_2, 即$\boldsymbol{\theta} \in W_1 \cap W_2$. 所以, $W_1 \cap W_2$是V的非空子集.

任取$\boldsymbol{\alpha}, \boldsymbol{\beta} \in W_1 \cap W_2$, 则$\boldsymbol{\alpha}, \boldsymbol{\beta} \in W_1$. 因$W_1$是子空间, 故$\boldsymbol{\alpha} + \boldsymbol{\beta} \in W_1$. 同理可证$\boldsymbol{\alpha} + \boldsymbol{\beta} \in W_2$. 所以$\boldsymbol{\alpha} + \boldsymbol{\beta} \in W_1 \cap W_2$.

任取$\boldsymbol{\alpha} \in W_1 \cap W_2, k \in F$, 则由$\boldsymbol{\alpha} \in W_1$可得$k\boldsymbol{\alpha} \in W_1$. 同理可证$k\boldsymbol{\alpha} \in W_2$. 所以$k\boldsymbol{\alpha} \in W_1 \cap W_2$

由定理3.3.1可知$W_1 \cap W_2$是V的子空间.

定理 3.3.5 设W_1, W_2是数域F上线性空间V的两个子空间, 则下列集合

$$\{\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 \mid \boldsymbol{\alpha}_1 \in W_1, \boldsymbol{\alpha}_2 \in W_2\}$$

也是V的子空间, 称之为W_1与W_2的**和空间**, 记为$W_1 + W_2$.

证 因$\boldsymbol{\theta}=\boldsymbol{\theta}+\boldsymbol{\theta}$, 故$\boldsymbol{\theta}\in W_1+W_2$. 因此$W_1+W_2$是$V$的非空子集.

任取$\boldsymbol{\alpha},\boldsymbol{\beta}\in W_1+W_2$, 则$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\beta}=\boldsymbol{\beta}_1+\boldsymbol{\beta}_2$, 这里$\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1\in W_1,\boldsymbol{\alpha}_2$, $\boldsymbol{\beta}_2\in W_2$. 因$\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1\in W_1,\boldsymbol{\alpha}_2+\boldsymbol{\beta}_2\in W_2$, 故

$$\begin{aligned}\boldsymbol{\alpha}+\boldsymbol{\beta}&=(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)+(\boldsymbol{\beta}_1+\boldsymbol{\beta}_2)\\&=(\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1)+(\boldsymbol{\alpha}_2+\boldsymbol{\beta}_2)\in W_1+W_2\end{aligned}$$

任取$\boldsymbol{\alpha}\in W_1+W_2,k\in F$, 则$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$, 其中$\boldsymbol{\alpha}_1\in W_1,\boldsymbol{\alpha}_2\in W_2$. 又$k\boldsymbol{\alpha}_1\in W_1$, $k\boldsymbol{\alpha}_2\in W_2$, 故

$$k\boldsymbol{\alpha}=k\left(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2\right)=k\boldsymbol{\alpha}_1+k\boldsymbol{\alpha}_2\in W_1+W_2$$

于是W_1+W_2是V的子空间.

必须指出的是, 两个子空间W_1与W_2的并$W_1\cup W_2$一般不再是子空间.

例 3.3.10 已知

$$W_1=\left\{\begin{bmatrix}a_1&b_1\\0&c_1\end{bmatrix}\mid a_1,b_1,c_1\in\mathbf{R}\right\}$$

$$W_2=\left\{\begin{bmatrix}a_2&0\\b_2&c_2\end{bmatrix}\mid a_2,b_2,c_2\in\mathbf{R}\right\},\quad W_3=\left\{\begin{bmatrix}0&0\\a_3&0\end{bmatrix}\mid a_3\in\mathbf{R}\right\}$$

是矩阵空间$\mathbf{R}^{2\times2}$的三个子空间. 试讨论它们的交与和.

解 首先, 易得$W_1\cap W_3=\{\mathbf{0}\},W_2\cap W_3=W_3$,

$$W_1\cap W_2=\left\{\begin{bmatrix}a_4&0\\0&b_4\end{bmatrix}\mid a_4,b_4\in\mathbf{R}\right\}$$

其次, 任取$\boldsymbol{A}=[a_{ij}]\in\mathbf{R}^{2\times2}$, 因

$$\boldsymbol{A}=\begin{bmatrix}a_{11}&a_{12}\\0&a_{22}\end{bmatrix}+\begin{bmatrix}0&0\\a_{21}&0\end{bmatrix}$$

故$\boldsymbol{A}\in W_1+W_2,\boldsymbol{A}\in W_1+W_3$. 由此得

$$W_1+W_2=\mathbf{R}^{2\times2},\quad W_1+W_3=\mathbf{R}^{2\times2}$$

此外, $W_2+W_3=W_2$.

在上例中, 虽然$W_1+W_2=W_1+W_3=\mathbf{R}^{2\times2}$, 但是这两个和却有很大的不同. 在$W_1+W_2=\mathbf{R}^{2\times2}$ 中, 每个矩阵均可分解为W_1与W_2中各一个矩阵的和, 但分解式不唯一. 而在$W_1+W_3=\mathbf{R}^{2\times2}$中, 每个矩阵均可唯一地分解为$W_1$与$W_3$中各一个矩阵的和. 我们称后一个和为**直和**, 记为$W_1\oplus W_3=\mathbf{R}^{2\times2}$.

例 3.3.11 设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s$与$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t$是数域$F$上线性空间$V$中的两组向量, 则

$$L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s) + L(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t) = L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t)$$

证 任取$\boldsymbol{\gamma} \in L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s) + L(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t)$, 则$\boldsymbol{\gamma} = \boldsymbol{\alpha} + \boldsymbol{\beta}$, 其中$\boldsymbol{\alpha} \in L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s), \boldsymbol{\beta} \in L(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t)$. 因为$\boldsymbol{\alpha}$可由$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s$线性表出, $\boldsymbol{\beta}$可由$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t$线性表出, 故$\boldsymbol{\gamma} = \boldsymbol{\alpha} + \boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s$, $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t$线性表出, 即$\boldsymbol{\gamma} \in L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t)$. 由此得

$$L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s) + L(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t) \subseteq L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t)$$

反之, 任取$\boldsymbol{\gamma} \in L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t)$, 则

$$\boldsymbol{\gamma} = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_s\boldsymbol{\alpha}_s + l_1\boldsymbol{\beta}_1 + l_2\boldsymbol{\beta}_2 + \cdots + l_t\boldsymbol{\beta}_t$$

令

$$\boldsymbol{\alpha} = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_s\boldsymbol{\alpha}_s, \quad \boldsymbol{\beta} = l_1\boldsymbol{\beta}_1 + l_2\boldsymbol{\beta}_2 + \cdots + l_t\boldsymbol{\beta}_t$$

则$\boldsymbol{\gamma} = \boldsymbol{\alpha} + \boldsymbol{\beta}$, 且有$\boldsymbol{\alpha} \in L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s)$, $\boldsymbol{\beta} \in L(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t)$. 由此得$\boldsymbol{\gamma} \in L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s) + L(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t)$, 于是

$$L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t) \subseteq L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s) + L(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t).$$

综上所述, 结论得证.

定理 3.3.6 设W_1, W_2是数域F上线性空间V的两个有限维子空间, 则

$$\dim(W_1) + \dim(W_2) = \dim(W_1 + W_2) + \dim(W_1 \cap W_2)$$

称上式为**维数公式**.

证 设$\dim(W_1) = s, \dim(W_2) = t, \dim(W_1 \cap W_2) = r$, 只需证明$W_1 + W_2$的维数为$= s + t - r$.

取$W_1 \cap W_2$的一个基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r$, 因$W_1 \cap W_2$是$W_1$和$W_2$的子空间, 故由$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r$可分别扩充为$W_1$和$W_2$的基, 设之为$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_{r+1}, \ldots, \boldsymbol{\beta}_s$和$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\gamma}_{r+1}, \ldots, \boldsymbol{\gamma}_t$, 则

$$W_1 = L\left(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_{r+1}, \ldots, \boldsymbol{\beta}_s\right), W_2 = L\left(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\gamma}_{r+1}, \ldots, \boldsymbol{\gamma}_t\right)$$

于是

$$W_1+W_2=L\left(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}_{r+1},\ldots,\boldsymbol{\beta}_s,\boldsymbol{\gamma}_{r+1},\ldots,\boldsymbol{\gamma}_t\right)$$

下面只需证明$\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}_{r+1},\ldots,\boldsymbol{\beta}_s,\boldsymbol{\gamma}_{r+1},\ldots,\boldsymbol{\gamma}_t$线性无关.

令

$$k_1\boldsymbol{\alpha}_1+\cdots+k_r\boldsymbol{\alpha}_r+k_{r+1}\boldsymbol{\beta}_{r+1}+\cdots+k_s\boldsymbol{\beta}_s+k_{s+1}\boldsymbol{\gamma}_{r+1}+\cdots+k_{s+t-r}\boldsymbol{\gamma}_t=\boldsymbol{\theta} \tag{3.7}$$

则

$$\begin{aligned}\boldsymbol{\alpha}&=k_1\boldsymbol{\alpha}_1+\cdots+k_r\boldsymbol{\alpha}_r+k_{r+1}\boldsymbol{\beta}_{r+1}+\cdots+k_s\boldsymbol{\beta}_s\\&=-k_{s+1}\boldsymbol{\gamma}_{r+1}-\cdots-k_{s+t-r}\boldsymbol{\gamma}_t\end{aligned}$$

既属于W_1又属于W_2, 即$\boldsymbol{\alpha}\in W_1\cap W_2$, 于是有$\boldsymbol{\alpha}=l_1\boldsymbol{\alpha}_1+\cdots+l_r\boldsymbol{\alpha}_r$. 将之代入上式整理得

$$l_1\boldsymbol{\alpha}_1+\cdots+l_r\boldsymbol{\alpha}_r+k_{s+1}\boldsymbol{\gamma}_{r+1}+\cdots+k_{s+t-r}\boldsymbol{\gamma}_t=\boldsymbol{\theta}$$

因$\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_r,\boldsymbol{\gamma}_{r+1},\ldots,\boldsymbol{\gamma}_t$是$W_2$的基, 故其线性无关, 所以由上式可得

$$k_{s+1}=k_{s+2}=\cdots=k_{s+t-r}=0$$

将之代入式(3.7)得

$$k_1\boldsymbol{\alpha}_1+\cdots+k_r\boldsymbol{\alpha}_r+k_{r+1}\boldsymbol{\beta}_{r+1}+\cdots+k_s\boldsymbol{\beta}_s=\boldsymbol{\theta}$$

因$\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}_{r+1},\ldots,\boldsymbol{\beta}_s$是$W_1$的基, 故其线性无关, 所以由上式得

$$k_1=k_2=\cdots=k_s=0$$

于是有$\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}_{r+1},\ldots,\boldsymbol{\beta}_s,\boldsymbol{\gamma}_{r+1},\ldots,\boldsymbol{\gamma}_t$线性无关.

维数公式是有限维线性空间的一个重要结论, 利用它可得到关于维数和空间运算的更进一步的结果.

例 3.3.12 已知$\mathbf{R}^3$中的两组向量:

$$\boldsymbol{\alpha}_1=(-1,-1,-2),\boldsymbol{\alpha}_2=(2,1,1),\boldsymbol{\alpha}_3=(-1,0,1)$$

$$\boldsymbol{\beta}_1=(1,2,1),\boldsymbol{\beta}_2=(-1,-1,-1)$$

令$W_1=L\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\right),W_2=L\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\right)$, 求$W_1+W_2$与$W_1\cap W_2$.

解 首先, 通过求生成向量组的秩和极大无关组, 不难得到

$$\dim W_1 = 2, \dim W_2 = 2, \dim\left(W_1 + W_2\right) = 3$$

并且

$$W_1 = L\left(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\right), \quad W_1 + W_2 = L\left(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_1\right)$$

由维数公式可得$\dim\left(W_1 \cap W_2\right) = 1$.

设$\boldsymbol{\gamma} \in W_1 \cap W_2$, 则

$$\boldsymbol{\gamma} = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 = k_3\boldsymbol{\beta}_1 + k_4\boldsymbol{\beta}_2$$

即$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 - k_3\boldsymbol{\beta}_1 - k_4\boldsymbol{\beta}_2 = \boldsymbol{\theta}$, 按分量写出即为

$$\begin{cases} -k_1 + 2k_2 - k_3 + k_4 = 0 \\ -k_1 + k_2 - 2k_3 + k_4 = 0 \\ -2k_1 + k_2 - k_3 + k_4 = 0 \end{cases}$$

由于系数矩阵的秩为3, 所以上述齐次线性方程组的基础解系恰含一个解向量. 求出一个基础解系为$(k_1, k_2, k_3, k_4)^{\mathrm{T}} = (1, -1, 1, 4)^{\mathrm{T}}$, 于是

$$\boldsymbol{\gamma} = \boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2 = \boldsymbol{\beta}_1 + 4\boldsymbol{\beta}_2 = (-3, -2, -3)$$

由此即得$W_1 \cap W_2 = L(\boldsymbol{\gamma})$.

扫描右侧二维码, 找到子空间$W_1 + W_2$与$W_1 \cap W_2$, 并验证维数公式.

维数公式

§3.4 欧氏空间(一)

众所周知, 几何中的向量或物理中的矢量都有大小和方向. 如何定义线性空间中向量的大小和方向呢？本节将在实向量空间(即$\mathbf{R}^n$的子空间)上, 通过引入内积, 赋予向量大小和方向的内涵.

§3.4.1 向量的内积

在平面解析几何中, 根据几何直观, 可定义有向线段$\boldsymbol{a}$与有向线段$\boldsymbol{b}$的数性积(也称点乘或内积)为

$$\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}| \cos\langle\boldsymbol{a}, \boldsymbol{b}\rangle \tag{3.8}$$

这里$|\boldsymbol{a}|$表示$\boldsymbol{a}$的长度, $\langle \boldsymbol{a}, \boldsymbol{b} \rangle$表示$\boldsymbol{a}$与$\boldsymbol{b}$的夹角. 利用数性积, 可表示有向线段的度量

$$|\boldsymbol{a}| = \sqrt{\boldsymbol{a} \cdot \boldsymbol{a}}, \quad \cos\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{|\boldsymbol{a}||\boldsymbol{b}|} \tag{3.9}$$

建立平面直角坐标系Oxy, 平移$\boldsymbol{a}, \boldsymbol{b}$使之起点落在原点$O$上, 则终点坐标即为$\boldsymbol{a}, \boldsymbol{b}$对应的2元实向量. 设$\boldsymbol{a}$对应$(a_1, a_2)$, $\boldsymbol{b}$对应(b_1, b_2), 可以证明

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_1 b_1 + a_2 b_2 \tag{3.10}$$

同样地, 对于三维几何空间中对应3元实向量(a_1, a_2, a_3), (b_1, b_2, b_3)的有向线段$\boldsymbol{a}, \boldsymbol{b}$有

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$$

对一般的n元$(n \geqslant 4)$实向量$\boldsymbol{\alpha} = (a_1, a_2, \ldots, a_n)$, $\boldsymbol{\beta} = (b_1, b_2, \ldots, b_n)$, 由于没有直观的几何意义, 故无法像式(3.8)那样引入内积, 因此人们仿照式(3.10)定义$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$的内积.

定义 3.4.1 设$\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbf{R}^n$,

$$\boldsymbol{\alpha} = (a_1, a_2, \ldots, a_n), \quad \boldsymbol{\beta} = (b_1, b_2, \ldots, b_n)$$

则$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$的**内积**$(\boldsymbol{\alpha}, \boldsymbol{\beta})$规定为

$$(\boldsymbol{\alpha}, \boldsymbol{\beta}) = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \tag{3.11}$$

若$\boldsymbol{\alpha}, \boldsymbol{\beta}$均为列向量, 则

$$(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{\alpha}$$

若$\boldsymbol{\alpha}, \boldsymbol{\beta}$均为行向量,则

$$(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \boldsymbol{\alpha} \boldsymbol{\beta}^{\mathrm{T}} = \boldsymbol{\beta} \boldsymbol{\alpha}^{\mathrm{T}}$$

根据定义容易证明, 向量的内积具有下列性质:

性质 3.4.1 对任意$\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbf{R}^n, k \in \mathbf{R}$, 均有

(1) $(\boldsymbol{\alpha}, \boldsymbol{\beta}) = (\boldsymbol{\beta}, \boldsymbol{\alpha})$;

(2) $(\boldsymbol{\alpha} + \boldsymbol{\beta}, \boldsymbol{\gamma}) = (\boldsymbol{\alpha}, \boldsymbol{\gamma}) + (\boldsymbol{\beta}, \boldsymbol{\gamma})$;

(3) $(k\boldsymbol{\alpha}, \boldsymbol{\beta}) = k(\boldsymbol{\alpha}, \boldsymbol{\beta})$;

(4) $(\boldsymbol{\alpha}, \boldsymbol{\alpha}) \geqslant 0$, 当且仅当$\boldsymbol{\alpha} = \boldsymbol{\theta}$时等号成立.

在解析几何中, 有向线段的数性积也具有上述四条性质. 由此可见, 这里定义的向量内积的确是有向线段数性积的推广. 由式(3.11)定义的内积, 称为**标准内积**.

定义 3.4.2 定义了内积的实向量空间称为**Euclid空间**, 简称为**欧氏空间**[1].

§3.4.2 向量的度量

利用内积, 可根据解析几何中的结果引入向量的长度和夹角等度量概念. 根据式(3.9)中的第一式, 首先引入

定义 3.4.3 欧氏空间中的**向量$\boldsymbol{\alpha}$的长度**$|\boldsymbol{\alpha}|$规定为

$$|\boldsymbol{\alpha}| = \sqrt{(\boldsymbol{\alpha}, \boldsymbol{\alpha})}$$

若$\boldsymbol{\alpha} = (a_1, a_2, \ldots, a_n)$, 则

$$|\boldsymbol{\alpha}| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$

当$n = 2$(或3)时, 上式恰为平面(或空间)直角坐标系下, 起点为原点, 终点为点(a_1, a_2) (或(a_1, a_2, a_3))的有向线段的长度公式.

显然, $|\boldsymbol{\alpha}| = 0$当且仅当$\boldsymbol{\alpha} = \boldsymbol{\theta}$. 当$|\boldsymbol{\alpha}| = 1$时, 称$\boldsymbol{\alpha}$为**单位向量**. 当$\boldsymbol{\alpha} \neq \boldsymbol{\theta}$时, $\frac{1}{|\boldsymbol{\alpha}|}\boldsymbol{\alpha}$一定是单位向量, 称之为**对$\boldsymbol{\alpha}$单位化**.

例如, $\boldsymbol{\alpha} = \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$是单位向量, $\boldsymbol{\beta} = (1, 2)$不是单位向量, 但$\frac{1}{|\boldsymbol{\beta}|}\boldsymbol{\beta} = \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$是单位向量.

根据式(3.9)中的第二式, 可引入向量夹角的概念. 首先需要一个预备性结论.

定理 3.4.1 设V是欧氏空间, 则对任意$\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$均有

$$|(\boldsymbol{\alpha}, \boldsymbol{\beta})| \leqslant |\boldsymbol{\alpha}||\boldsymbol{\beta}| \tag{3.12}$$

上式称为**Cauchy-Schwarz不等式**.

[1]本节的欧氏空间特指定义了标准内积的实向量空间, 在第§3.5节, 我们将介绍一般的欧氏空间定义.

证 当$\boldsymbol{\beta}=\boldsymbol{\theta}$时, $(\boldsymbol{\alpha},\boldsymbol{\beta})=0, |\boldsymbol{\beta}|=0$, 此时式(3.12)显然成立.

假设$\boldsymbol{\beta}\neq\boldsymbol{\theta}$, 任取实数$x$, 则$\boldsymbol{\alpha}+x\boldsymbol{\beta}\in V$. 根据性质3.4.1,

$$(\boldsymbol{\alpha}+x\boldsymbol{\beta},\boldsymbol{\alpha}+x\boldsymbol{\beta})\geqslant 0$$

于是有

$$(\boldsymbol{\beta},\boldsymbol{\beta})x^2+2(\boldsymbol{\alpha},\boldsymbol{\beta})x+(\boldsymbol{\alpha},\boldsymbol{\alpha})\geqslant 0$$

因上式对任意实数x均成立, 且$(\boldsymbol{\beta},\boldsymbol{\beta})>0$, 故

$$[2(\boldsymbol{\alpha},\boldsymbol{\beta})]^2-4(\boldsymbol{\beta},\boldsymbol{\beta})(\boldsymbol{\alpha},\boldsymbol{\alpha})\leqslant 0$$

整理后得

$$|(\boldsymbol{\alpha},\boldsymbol{\beta})|^2=(\boldsymbol{\alpha},\boldsymbol{\beta})^2\leqslant(\boldsymbol{\alpha},\boldsymbol{\alpha})(\boldsymbol{\beta},\boldsymbol{\beta})=|\boldsymbol{\alpha}|^2|\boldsymbol{\beta}|^2$$

上式不等号两边同时开平方, 即得

$$|(\boldsymbol{\alpha},\boldsymbol{\beta})|\leqslant|\boldsymbol{\alpha}||\boldsymbol{\beta}|$$

请读者自己证明: 式(3.12)等号成立的充分必要条件是$\boldsymbol{\alpha},\boldsymbol{\beta}$线性相关.

现在我们可以定义向量的夹角了.

定义 3.4.4 *设V是欧氏空间, $\boldsymbol{\alpha},\boldsymbol{\beta}\in V$且$\boldsymbol{\alpha},\boldsymbol{\beta}$均不是零向量, 则$\boldsymbol{\alpha}$***与***$\boldsymbol{\beta}$***的夹角***$\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle$规定为*

$$\cos\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{|\boldsymbol{\alpha}||\boldsymbol{\beta}|}$$

这里$0\leqslant\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle\leqslant\pi$.

若向量$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$的夹角为$\frac{\pi}{2}$, 则$\cos\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=0$, 于是$(\boldsymbol{\alpha},\boldsymbol{\beta})=0$. 反之, 由$(\boldsymbol{\alpha},\boldsymbol{\beta})=0$又可得$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$的夹角为$\frac{\pi}{2}$. 因此, 利用内积可判断两个向量是否垂直(正交).

定义 3.4.5 *若$(\boldsymbol{\alpha},\boldsymbol{\beta})=0$, 则称向量$\boldsymbol{\alpha}$与向量$\boldsymbol{\beta}$***正交***, 记为$\boldsymbol{\alpha}\perp\boldsymbol{\beta}$.*

若$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$均为2元向量或均为3元向量, 则$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$正交对应有向线段的垂直. 此外, 零向量与任一向量均正交.

例 3.4.1 *设$\boldsymbol{A}\in\mathbf{R}^{m\times n}$, 则对任意$\boldsymbol{\alpha}\in\mathrm{Row}\,(\boldsymbol{A})$与任意$\boldsymbol{\beta}\in\mathrm{N}(\boldsymbol{A})$, 均有$\boldsymbol{\alpha}\perp\boldsymbol{\beta}$.*

证 设$\boldsymbol{A}$的行向量组为$\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\ldots,\boldsymbol{\gamma}_m$, 并且

$$\boldsymbol{\gamma}_i=(a_{i1},a_{i2},\ldots,a_{in}),\qquad \text{其中} i=1,2,\ldots,m$$

则$\boldsymbol{\gamma}_i$是齐次线性方程组$\boldsymbol{AX}=\mathbf{0}$中第$i$个方程的系数向量. 因$\boldsymbol{\beta}\in\mathrm{N}(\boldsymbol{A})$, 故$\boldsymbol{\beta}$是$\boldsymbol{AX}=\mathbf{0}$的解向量. 设$\boldsymbol{\beta}=(c_1,c_2,\ldots,c_n)$, 则

$$a_{i1}c_1+a_{i2}c_2+\cdots+a_{in}c_n=0$$

即$(\boldsymbol{\gamma}_i,\boldsymbol{\beta})=0$, 其中$i=1,2,\ldots,m$. 又因为$\boldsymbol{\alpha}\in\mathrm{Row}\,(\boldsymbol{A})$, 所以存在$k_i\in\mathbf{R}$使

$$\boldsymbol{\alpha}=k_1\boldsymbol{\gamma}_1+k_2\boldsymbol{\gamma}_2+\cdots+k_m\boldsymbol{\gamma}_m$$

于是

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=k_1(\boldsymbol{\gamma}_1,\boldsymbol{\beta})+k_2(\boldsymbol{\gamma}_2,\boldsymbol{\beta})+\cdots+k_m(\boldsymbol{\gamma}_m,\boldsymbol{\beta})=0$$

即$\boldsymbol{\alpha}\perp\boldsymbol{\beta}$.

在几何空间中有两个简单而重要的结论: 三角不等式和勾股定理. 下面证明这两个结论在欧氏空间中也成立.

性质 3.4.2 设V是欧氏空间, $\boldsymbol{\alpha}$与$\boldsymbol{\beta}$是V中任意两个向量, 则有

(1) 三角不等式: $|\boldsymbol{\alpha}+\boldsymbol{\beta}|\leqslant|\boldsymbol{\alpha}|+|\boldsymbol{\beta}|$;

(2) 勾股定理: 若$\boldsymbol{\alpha}\perp\boldsymbol{\beta}$, 则$|\boldsymbol{\alpha}+\boldsymbol{\beta}|^2=|\boldsymbol{\alpha}|^2+|\boldsymbol{\beta}|^2$.

证 (1) 因$|\boldsymbol{\alpha}+\boldsymbol{\beta}|^2=(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\alpha}+\boldsymbol{\beta})=(\boldsymbol{\alpha},\boldsymbol{\alpha})+2(\boldsymbol{\alpha},\boldsymbol{\beta})+(\boldsymbol{\beta},\boldsymbol{\beta})$

$$\leqslant|\boldsymbol{\alpha}|^2+2|\boldsymbol{\alpha}||\boldsymbol{\beta}|+|\boldsymbol{\beta}|^2=(|\boldsymbol{\alpha}|+|\boldsymbol{\beta}|)^2$$

故$|\boldsymbol{\alpha}+\boldsymbol{\beta}|\leqslant|\boldsymbol{\alpha}|+|\boldsymbol{\beta}|$.

(2) 因$\boldsymbol{\alpha}\perp\boldsymbol{\beta}$, 故$(\boldsymbol{\alpha},\boldsymbol{\beta})=0$. 于是

$$|\boldsymbol{\alpha}+\boldsymbol{\beta}|^2=(\boldsymbol{\alpha},\boldsymbol{\alpha})+2(\boldsymbol{\alpha},\boldsymbol{\beta})+(\boldsymbol{\beta},\boldsymbol{\beta})=|\boldsymbol{\alpha}|^2+|\boldsymbol{\beta}|^2$$

§3.4.3 标准正交基

有了上述必要的准备工作, 现在就可以在欧氏空间中建立“直角坐标系”了.

从解析几何中可知, 建立“直角坐标系”的关键在于选取两条或三条彼此垂直且长度全为1的有向线段. 以此为线索, 首先讨论两两正交的向量构成的向量组.

定义 3.4.6 设V是一个欧氏空间, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m$是$V$中$m$个非零向量, 若$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m$两两正交, 则称$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m$是**正交向量组**; 由单位向量构成的正交向量组称为**标准正交向量组**, 或**正交单位向量组**.

由于对向量的单位化不改变向量的正交性, 因此正交向量组单位化即得标准正交向量组. 这样, 寻找标准正交向量组的关键就是构造正交向量组. 首先给出正交向量组的一个重要性质:

定理 3.4.2 设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m$是欧氏空间$V$的一个正交向量组, 则$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m$线性无关.

证 令

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m = \boldsymbol{\theta} \tag{3.13}$$

上式两端同时与$\boldsymbol{\alpha}_1$作内积:

$$(k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m, \boldsymbol{\alpha}_1) = (\boldsymbol{\theta}, \boldsymbol{\alpha}_1)$$

展开后得

$$k_1(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1) + k_2(\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1) + \cdots + k_m(\boldsymbol{\alpha}_m, \boldsymbol{\alpha}_1) = 0 \tag{3.14}$$

已知$\boldsymbol{\alpha}_1$与$\boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m$均正交, 故

$$(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_1) = 0, \quad 其中 i = 2, 3, \ldots, m$$

将之代入式(3.14)得

$$k_1(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1) = 0$$

又$\boldsymbol{\alpha}_1 \neq \boldsymbol{\theta}$, 故$(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1) > 0$, 于是$k_1 = 0$.

同理可得$k_2 = k_3 = \cdots = k_m = 0$. 所以, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m$线性无关.

显然, 线性无关的向量组不一定是正交向量组. 能否在线性无关向量组的基础上, 构造出一个正交向量组呢? 下面以两个2元线性无关向量为例, 探讨构造相应的标准正交向量组的方法.

设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$是两个线性无关的2元

实向量. 因2元实向量对应平面上的有向线段(见图3.4.1), 故可用平面上的两条起点相同的有向线段$\overrightarrow{OA_1}$, $\overrightarrow{OA_2}$分别表示$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$. 由假设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$线性无关可得$\overrightarrow{OA_1}, \overrightarrow{OA_2}$不共线. 过$A_2$点向$\overrightarrow{OA_1}$引垂线且交$\overrightarrow{OA_1}$于$A_3$点, 则

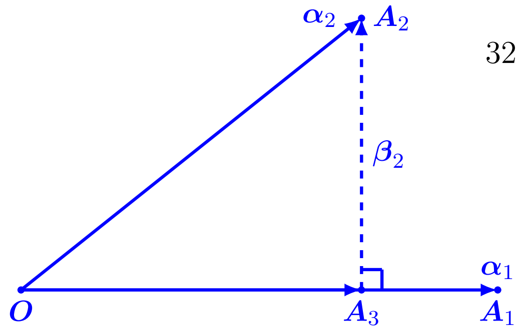

图 3.4.1: 平面向量的正交化

$$\overrightarrow{OA_2} = \overrightarrow{OA_3} + \overrightarrow{A_3A_2}$$

因$\overrightarrow{OA_3}$与$\overrightarrow{OA_1}$共线, 且$\overrightarrow{OA_1} \neq \boldsymbol{\theta}$, 故$\overrightarrow{OA_3}$可表示为$k_1\overrightarrow{OA_1}$, 于是

$$\overrightarrow{A_3A_2} = \overrightarrow{OA_2} - k_1\overrightarrow{OA_1} = \boldsymbol{\alpha}_2 - k_1\boldsymbol{\alpha}_1$$

令$\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1, \boldsymbol{\beta}_2$为有向线段$\overrightarrow{A_3A_2}$对应的2元向量, 即$\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 - k_1\boldsymbol{\beta}_1$, 则$\boldsymbol{\beta}_2 \perp \boldsymbol{\beta}_1$, 且它们均不为零向量, 否则$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$线性相关. 再令$\boldsymbol{\eta}_1 = \frac{1}{|\boldsymbol{\beta}_1|}\boldsymbol{\beta}_1, \boldsymbol{\eta}_2 = \frac{1}{|\boldsymbol{\beta}_2|}\boldsymbol{\beta}_2$, 则$\boldsymbol{\eta}_1, \boldsymbol{\eta}_2$即为所求的标准正交向量组. 这里系数$k_1$是可以确定的. 实际上, 由$\boldsymbol{\beta}_2 \perp \boldsymbol{\beta}_1$可得$(\boldsymbol{\beta}_2, \boldsymbol{\beta}_1) = 0$, 于是

$$0 = (\boldsymbol{\beta}_2, \boldsymbol{\beta}_1) = (\boldsymbol{\alpha}_2 - k_1\boldsymbol{\beta}_1, \boldsymbol{\beta}_1) = (\boldsymbol{\alpha}_2, \boldsymbol{\beta}_1) - k_1(\boldsymbol{\beta}_1, \boldsymbol{\beta}_1)$$

又$\boldsymbol{\beta}_1 \neq \boldsymbol{\theta}$, 故$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_1) > 0$, 于是可得

$$k_1 = \frac{(\boldsymbol{\alpha}_2, \boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1, \boldsymbol{\beta}_1)}$$

并且容易证明$\{\boldsymbol{\alpha}_1\} \cong \{\boldsymbol{\beta}_1\}, \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\} \cong \{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2\}$.

扫描右侧二维码, 了解如何通过$\mathbf{R}^3$中三个向量组成的线性无关的向量组构造相应的正交向量组.

将上述过程一般化, 即得下述结论:

定理 3.4.3 设V是欧氏空间, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m$是$V$中的线性无关向量组, 则$V$中存在标准正交的向量组$\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_m$, 使得

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_i\} \cong \{\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_i\}, \quad 其中 i = 1, 2, \ldots, m$$

证 对i作数学归纳法:

(1) 令$\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1, \boldsymbol{\eta}_1 = \frac{1}{|\boldsymbol{\beta}_1|}\boldsymbol{\beta}_1$, 则$\{\boldsymbol{\alpha}_1\} \cong \{\boldsymbol{\beta}_1\}$, 进而$\{\boldsymbol{\alpha}_1\} \cong \{\boldsymbol{\eta}_1\}$. 这说明结论在$i = 1$时成立.

(2) 设存在正交向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\dots,\boldsymbol{\beta}_{i-1}(i\geqslant 2)$, 满足

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_s\}\cong\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\dots,\boldsymbol{\beta}_s\},\quad \text{其中}s=1,2,\dots,i-1$$

令$\boldsymbol{\eta}_s=\frac{1}{|\boldsymbol{\beta}_s|}\boldsymbol{\beta}_s$, 其中$s=1,2,\dots,i-1$, 则$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\dots,\boldsymbol{\eta}_{i-1}$是标准正交向量组, 并且也满足

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_s\}\cong\{\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\dots,\boldsymbol{\eta}_s\},\quad \text{其中}s=1,2,\dots,i-1$$

(3) 对$2\leqslant i\leqslant m$, 在(2)的假设下, 令

$$\boldsymbol{\beta}_i=\boldsymbol{\alpha}_i-\frac{(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\frac{(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2-\cdots-\frac{(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_{i-1})}{(\boldsymbol{\beta}_{i-1},\boldsymbol{\beta}_{i-1})}\boldsymbol{\beta}_{i-1}\tag{3.15}$$

因$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\dots,\boldsymbol{\beta}_{i-1}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_{i-1}$线性表出, 故$\boldsymbol{\beta}_i$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_i$线性表出且$\boldsymbol{\alpha}_i$的系数为1. 又已知$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_i$线性无关, 所以$\boldsymbol{\beta}_i\neq\boldsymbol{\theta}$. 此外, 容易证明$\boldsymbol{\beta}_i$与$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\dots,\boldsymbol{\beta}_{i-1}$都正交, 故$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\dots,\boldsymbol{\beta}_i$是正交向量组. 根据式(3.15), $\boldsymbol{\alpha}_i$可由$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\dots,\boldsymbol{\beta}_i$线性表出, 所以

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_s\}\cong\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\dots,\boldsymbol{\beta}_s\},\quad \text{其中}s=1,2,\dots,i$$

令

$$\boldsymbol{\eta}_i=\frac{1}{|\boldsymbol{\beta}_i|}\boldsymbol{\beta}_i$$

则$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\dots,\boldsymbol{\eta}_i$是标准正交向量组, 并且也满足

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_s\}\cong\{\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\dots,\boldsymbol{\eta}_s\},\ \text{其中}s=1,2,\dots,i$$

由数学归纳法原理可知, 定理结论正确.

这个定理采用了归纳式的构造性证明方法, 证明过程提供了一种有效的寻找标准正交向量组的方法, 通常称之为**Schmidt正交化方法**.

例 3.4.2 试用Schmidt正交化方法将向量组$\boldsymbol{\alpha}_1=(1,1,0,0),\boldsymbol{\alpha}_2=(1,0,1,0)$, $\boldsymbol{\alpha}_3=(-1,0,0,1)$化为标准正交向量组.

解 先正交化: 令

$$\begin{aligned}
\boldsymbol{\beta}_1&=\boldsymbol{\alpha}_1\\
\boldsymbol{\beta}_2&=\boldsymbol{\alpha}_2-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1=\boldsymbol{\alpha}_2-\frac{1}{2}\boldsymbol{\beta}_1=\left(\frac{1}{2},-\frac{1}{2},1,0\right)\\
\boldsymbol{\beta}_3&=\boldsymbol{\alpha}_3-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2\\
&=\boldsymbol{\alpha}_3+\frac{1}{2}\boldsymbol{\beta}_1+\frac{1}{3}\boldsymbol{\beta}_2=\left(-\frac{1}{3},\frac{1}{3},\frac{1}{3},1\right)
\end{aligned}$$

再单位化: 令

$$\begin{aligned}
\boldsymbol{\eta}_1 &= \frac{1}{|\boldsymbol{\beta}_1|}\boldsymbol{\beta}_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right) \\
\boldsymbol{\eta}_2 &= \frac{1}{|\boldsymbol{\beta}_2|}\boldsymbol{\beta}_2 = \left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0\right) \\
\boldsymbol{\eta}_3 &= \frac{1}{|\boldsymbol{\beta}_3|}\boldsymbol{\beta}_3 = \left(-\frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{3}{2\sqrt{3}}\right)
\end{aligned}$$

则$\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3$为所求标准正交向量组.

定义 3.4.7 设V是欧氏空间, 则V中由正交向量组构成的基称为**正交基**, 而V中由标准正交向量组构成的基则称为**标准正交基**.

例 3.4.3 欧氏空间$\mathbf{R}^n$的自然基$\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_n$是标准正交基.

例 3.4.4 求下列实系数齐次线性方程组

$$\begin{cases}
x_1 + x_2 + x_3 + x_4 = 0 \\
x_1 + 2x_2 + 3x_3 + 4x_4 = 0 \\
2x_1 + 3x_2 + 4x_3 + 5x_4 = 0
\end{cases}$$

的解空间的一个标准正交基.

解 由于解空间的基就是基础解系, 故由它们经过Schmidt正交化方法得到的标准正交向量组就是标准正交基.

容易求得上述方程组的一个基础解系:

$$\boldsymbol{X}_1 = (1, -2, 1, 0)^{\mathrm{T}}, \boldsymbol{X}_2 = (2, -3, 0, 1)^{\mathrm{T}}$$

下面对$\boldsymbol{X}_1, \boldsymbol{X}_2$正交化, 单位化:

令 $\boldsymbol{\beta}_1 = \boldsymbol{X}_1$,

$$\begin{aligned}
\boldsymbol{\beta}_2 &= \boldsymbol{X}_2 - \frac{(\boldsymbol{X}_2, \boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1, \boldsymbol{\beta}_1)}\boldsymbol{\beta}_1 = \boldsymbol{X}_2 - \frac{8}{6}\boldsymbol{\beta}_1 = \left(\frac{2}{3}, -\frac{1}{3}, -\frac{4}{3}, 1\right)^{\mathrm{T}} \\
\boldsymbol{\eta}_1 &= \frac{1}{|\boldsymbol{\beta}_1|}\boldsymbol{\beta}_1 = \left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, 0\right)^{\mathrm{T}} \\
\boldsymbol{\eta}_2 &= \frac{1}{|\boldsymbol{\beta}_2|}\boldsymbol{\beta}_2 = \left(\frac{2}{\sqrt{30}}, -\frac{1}{\sqrt{30}}, -\frac{4}{\sqrt{30}}, \frac{3}{\sqrt{3}}\right)^{\mathrm{T}}
\end{aligned}$$

则$\boldsymbol{\eta}_1, \boldsymbol{\eta}_2$就是解空间的一个标准正交基.

若一个齐次线性方程组的基础解系只含一个解向量, 则解空间的标准正交基只含一个单位解向量. 我们规定, 只含一个向量的标准正交向量组就是一个单位向量.

由定理3.4.3不难发现, 存在基的欧氏空间也一定存在标准正交基. 由于本节所讨论的欧氏空间均为$\mathbf{R}^n$的子空间, 所以除$\{\boldsymbol{\theta}\}$以外, 它们都存在基. 由此可得:

定理 3.4.4 *设V是欧氏空间, $V \subseteq \mathbf{R}^n$且$V \neq \{\boldsymbol{\theta}\}$, 则$V$一定存在标准正交基.*

标准正交基可以形象地理解为几何空间的直角坐标系在欧氏空间的一个推广, 许多问题的讨论或需借助标准正交基来进行, 或在标准正交基下讨论比较方便.

§3.4.4 正交矩阵

在第§3.2节, 我们已经知道, 数域F上线性空间的一个基到另一个基的过渡矩阵为数域F上的可逆矩阵. 在欧氏空间中, 一个标准正交基到另一个标准正交基的过渡矩阵显然是一个可逆的实矩阵. 除此以外, 它还有什么特性呢? 对此我们有如下结果.

定理 3.4.5 *设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$与$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$是欧氏空间$V$的两个标准正交基, $\boldsymbol{A}$为由基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$到基$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$的过渡矩阵, 则$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}$.*

证 根据定义3.2.4, 由已知条件可得$[\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n] = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\,\boldsymbol{A}$, 其中$\boldsymbol{A} = [a_{ij}] \in \mathbf{R}^{n\times n}$, 于是

$$\boldsymbol{\beta}_i = \sum_{k=1}^{n} a_{ki}\boldsymbol{\alpha}_k \quad i = 1, 2, \ldots, n$$

因$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$与$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$是两个标准正交基, 故

$$\left(\boldsymbol{\beta}_i, \boldsymbol{\beta}_j\right) = \left(\sum_{k=1}^{n} a_{ki}\boldsymbol{\alpha}_k, \sum_{l=1}^{n} a_{lj}\boldsymbol{\alpha}_l\right) = \sum_{k=1}^{n} a_{ki}a_{kj} = \begin{cases} 0, i \neq j \\ 1, i = j \end{cases} \tag{3.16}$$

其中$i, j = 1, 2, \ldots, n$. 将式(3.16)中最后一个等式表示为矩阵形式即为

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}$$

定义 3.4.8 设$\boldsymbol{A} \in \mathbf{R}^{n\times n}$, 若$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}$, 则称$\boldsymbol{A}$是**正交矩阵**.

显然, 若$\boldsymbol{A}$是正交矩阵, 则$\boldsymbol{A}$可逆, 且$\boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathrm{T}}$.

设$\boldsymbol{A}$是n阶正交矩阵, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$是$\boldsymbol{A}$的列向量组, 则$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n \in \mathbf{R}^n$. 根据$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}$,

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]^{\mathrm{T}}[\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n] = \begin{bmatrix} \boldsymbol{\alpha}_1^{\mathrm{T}} \\ \boldsymbol{\alpha}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{\alpha}_n^{\mathrm{T}} \end{bmatrix}[\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]$$

$$= \begin{bmatrix} \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_n \\ \boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_n \\ \vdots & \vdots & & \vdots \\ \boldsymbol{\alpha}_n^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_n^{\mathrm{T}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n^{\mathrm{T}}\boldsymbol{\alpha}_n \end{bmatrix} = \boldsymbol{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

由此得

$$(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j) = \boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{\alpha}_j = \begin{cases} 0, i \neq j \\ 1, i = j \end{cases}, \quad i, j = 1, 2, \ldots, n$$

即$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$是标准正交向量组. 上述推导过程是可逆的, 因此有

定理 3.4.6 设$\boldsymbol{A} \in \mathbf{R}^{n\times n}$, 则$\boldsymbol{A}$是正交矩阵的充分必要条件是$\boldsymbol{A}$的列(行)向量组是标准正交的.

关于行向量组的标准正交性请读者自证.

例 3.4.5 在$\mathbf{R}^2$中, 将直角坐标系Oxy绕原点逆时针旋转$\boldsymbol{\theta}$角, 得到新坐标系$Ox'y'$(如图3.4.2所示). 不难得到, 平面空间的这两个标准正交基$\overrightarrow{\boldsymbol{i}}$, $\overrightarrow{\boldsymbol{j}}$和$\overrightarrow{\boldsymbol{i}'}$, $\overrightarrow{\boldsymbol{j}'}$之间具有下列关系:

$$\begin{cases} \overrightarrow{\boldsymbol{i}'} = \overrightarrow{\boldsymbol{i}}\cos\boldsymbol{\theta} + \overrightarrow{\boldsymbol{j}}\sin\boldsymbol{\theta} \\ \overrightarrow{\boldsymbol{j}'} = -\overrightarrow{\boldsymbol{i}}\sin\boldsymbol{\theta} + \overrightarrow{\boldsymbol{j}}\cos\boldsymbol{\theta} \end{cases}$$

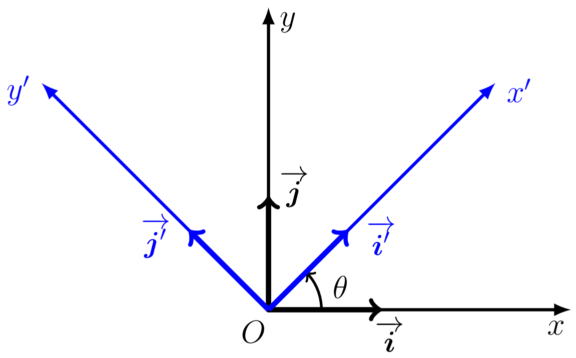

图 3.4.2: 坐标旋转

上式可改写为

$$[\vec{\boldsymbol{i}}', \vec{\boldsymbol{j}}'] = [\vec{\boldsymbol{i}}, \vec{\boldsymbol{j}}]\begin{bmatrix} \cos\boldsymbol{\theta} & -\sin\boldsymbol{\theta} \\ \sin\boldsymbol{\theta} & \cos\boldsymbol{\theta} \end{bmatrix}$$

这就是从基$\vec{\boldsymbol{i}}, \vec{\boldsymbol{j}}$到基$\vec{\boldsymbol{i}}', \vec{\boldsymbol{j}}'$的基变换公式, 而

$$\begin{bmatrix} \cos\boldsymbol{\theta} & -\sin\boldsymbol{\theta} \\ \sin\boldsymbol{\theta} & \cos\boldsymbol{\theta} \end{bmatrix}$$

为从基$\vec{\boldsymbol{i}}, \vec{\boldsymbol{j}}$到基$\vec{\boldsymbol{i}}', \vec{\boldsymbol{j}}'$的过渡矩阵, 是一个正交矩阵.

扫描右侧二维码, 通过拖动滑动条选择不同的角度值$\boldsymbol{\theta}$, 观察两个标准正交基的关系, 以及它们之间的过渡矩阵对应的旋转变换.

例 3.4.6 设$\boldsymbol{\alpha}$是n元实的列向量, $\boldsymbol{I}$是n阶单位矩阵, 令$\boldsymbol{A} = \boldsymbol{I} - 2\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}$, 证明: 若$\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha} = 1$, 则$\boldsymbol{A}$是正交矩阵.

证 由

$$\boldsymbol{A}^{\mathrm{T}} = \left(\boldsymbol{I} - 2\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}\right)^{\mathrm{T}} = \boldsymbol{I}^{\mathrm{T}} - 2\left(\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}\right)^{\mathrm{T}} = \boldsymbol{I} - 2\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}} = \boldsymbol{A}$$

有

$$\begin{aligned} \boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}^2 = \left(\boldsymbol{I} - 2\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}\right)^2 &= \boldsymbol{I}^2 - 4\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}} + \left(2\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}\right)^2 \\ &= \boldsymbol{I} - 4\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}} + 4\left(\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}\right)^2 \\ &= \boldsymbol{I} - 4\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}} + 4\left(\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}\right)\left(\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}\right) \\ &= \boldsymbol{I} - 4\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}} + 4\boldsymbol{\alpha}\left(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha}\right)\boldsymbol{\alpha}^{\mathrm{T}} \\ &= \boldsymbol{I} - 4\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}} + 4\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}} = \boldsymbol{I} \end{aligned}$$

故$\boldsymbol{A}$是正交矩阵.

若$\boldsymbol{\alpha}$为单位向量, $n = 2$(或3), 则例3.4.6中的矩阵$\boldsymbol{A}$表示关于过原点且

与$\boldsymbol{\alpha}$垂直的直线(或平面)的镜面反射. 扫描右侧二维码, 通过选择不同的方向角确定单位向量$\boldsymbol{\alpha}$, 观察$\boldsymbol{A}$对应的镜面反射.

例 3.4.7 设$\boldsymbol{A} \in \mathbf{R}^{n\times n}$, 证明: 若$\boldsymbol{A}$可逆, 则$\boldsymbol{A}$可表示为

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R} \tag{3.17}$$

其中$\boldsymbol{Q}$是n阶正交矩阵, $\boldsymbol{R}$是n阶可逆上三角矩阵. 式(3.17)称为实方阵$\boldsymbol{A}$的**正交分解**.

证 因可逆上三角矩阵的逆矩阵也是上三角矩阵, 故式(3.17)等价于

$$\boldsymbol{A}\boldsymbol{R}' = \boldsymbol{Q}$$

其中$\boldsymbol{R}'$是可逆上三角矩阵, $\boldsymbol{Q}$是正交矩阵.

设$\boldsymbol{A}$的列向量组为$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$, 则由$\boldsymbol{A}$可逆得$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$是欧氏空间$\mathbf{R}^n$的一个基. 将$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$用Schmidt正交化方法正交化、单位化:

令$\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1$

$$\boldsymbol{\beta}_i = \boldsymbol{\alpha}_i - \frac{(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1, \boldsymbol{\beta}_1)}\boldsymbol{\beta}_1 - \frac{(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2, \boldsymbol{\beta}_2)}\boldsymbol{\beta}_2 - \cdots - \frac{(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_{i-1})}{(\boldsymbol{\beta}_{i-1}, \boldsymbol{\beta}_{i-1})}\boldsymbol{\beta}_{i-1} \tag{3.18}$$

其中$i = 2, 3, \ldots, n$. 因$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_{i-1}$可由$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_{i-1}$线性表出, 故式(3.18)又可表示为

$$\boldsymbol{\beta}_i = k_{1i}\boldsymbol{\alpha}_1 + k_{2i}\boldsymbol{\alpha}_2 + \cdots + k_{i-1,i}\boldsymbol{\alpha}_{i-1} + \boldsymbol{\alpha}_i, i = 1, 2, \ldots, n$$

再令

$$\boldsymbol{\eta}_i = \frac{1}{|\boldsymbol{\beta}_i|}\boldsymbol{\beta}_i = r_{1i}\boldsymbol{\alpha}_1 + r_{2i}\boldsymbol{\alpha}_2 + \cdots + r_{ii}\boldsymbol{\alpha}_i, i = 1, 2, \ldots, n \tag{3.19}$$

则$r_{ii} > 0 (i = 1, 2, \ldots, n)$.

把式(3.19)改写成矩阵形式

$$[\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_n] = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix} \tag{3.20}$$

令

$$\boldsymbol{Q}=[\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\ldots,\boldsymbol{\eta}_n],\quad \boldsymbol{R}'=\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix}$$

则式(3.20)成为

$$\boldsymbol{A}\boldsymbol{R}'=\boldsymbol{Q}$$

因$r_{11},r_{22},\ldots,r_{nn}>0$, 故$\mathrm{r}(\boldsymbol{R}')=n$, $\boldsymbol{R}'$是可逆的上三角矩阵. 又列向量组$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\ldots,\boldsymbol{\eta}_n$标准正交, 故$n$阶实方阵$\boldsymbol{Q}$是正交矩阵.

如果一个实系数线性方程组$\boldsymbol{AX}=\boldsymbol{b}$的系数矩阵$\boldsymbol{A}$存在正交分解$\boldsymbol{A}=\boldsymbol{QR}$, 则解方程组$\boldsymbol{AX}=\boldsymbol{b}$可通过解下列两个线性方程组来进行:

$$\boldsymbol{QY}=\boldsymbol{b},\quad \boldsymbol{RX}=\boldsymbol{Y}$$

前一个方程组的解为

$$\boldsymbol{Y}=\boldsymbol{Q}^{-1}\boldsymbol{b}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{b}$$

后一个方程组直接回代即可得解. 这种求解方法与利用系数矩阵的三角分解求解(见第**??**节)的方法思想相同, 但正交分解的适用范围更广一些.

§3.4.5 *最小二乘法

人们在解决实际问题的过程中, 经常需要处理线性方程组, 然而在很多情况下方程组都是无解的, 导致这种情况发生的原因是复杂的, 有可能是环境发生了微小变化, 也有可能是读取数据时出现偏差, 还有可能是处理数据时出现计算误差等等. 因此, 我们不能因为一个方程组无解就轻易断言问题本身无解. 人们通常的做法是, 求一组数使之最大限度地适合原方程组, 进而把它就视为原方程组的解, 实现这一目标的方法之一是最小二乘法.

为了介绍最小二乘法, 首先要引入距离的概念.

平面上的两个向量$\boldsymbol{a}$与$\boldsymbol{b}$的距离应理解为两向量位置的差异程度. 把$\boldsymbol{a}$与$\boldsymbol{b}$的起点放在一起, 则终点之间的距离远近恰能表示$\boldsymbol{a}$与$\boldsymbol{b}$的相似程度, 而终点的距离恰为$\boldsymbol{a}-\boldsymbol{b}$的长度, 即$|\boldsymbol{a}-\boldsymbol{b}|$, 可称为$\boldsymbol{a}$与$\boldsymbol{b}$的距离. 将之推广到欧氏空间, 我们有

定义 3.4.9 设V是欧氏空间, $\boldsymbol{\alpha},\boldsymbol{\beta}\in V$, 则$\boldsymbol{\alpha}$**到**$\boldsymbol{\beta}$**的距离**$d(\boldsymbol{\alpha},\boldsymbol{\beta})$规定为

$$d(\boldsymbol{\alpha},\boldsymbol{\beta})=|\boldsymbol{\alpha}-\boldsymbol{\beta}|$$

容易证明向量的距离具有下述性质:

性质 3.4.3 设$\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$是欧氏空间$V$中任意三个向量, 则

(1) $d(\boldsymbol{\alpha},\boldsymbol{\beta})=d(\boldsymbol{\beta},\boldsymbol{\alpha})$;

(2) $d(\boldsymbol{\alpha},\boldsymbol{\beta})\geqslant 0$,当且仅当$\boldsymbol{\alpha}=\boldsymbol{\beta}$时等号成立;

(3) $d(\boldsymbol{\alpha},\boldsymbol{\beta})\leqslant d(\boldsymbol{\alpha},\boldsymbol{\gamma})+d(\boldsymbol{\gamma},\boldsymbol{\beta})$.

在几何空间中, 一条起点在平面$\boldsymbol{\pi}$上的有向线段$\boldsymbol{a}$到$\boldsymbol{\pi}$的距离应为$\boldsymbol{a}$与$\boldsymbol{\pi}$上所有有向线段的距离中的最小者. 从图3.4.3中几何直观上不难看出, 若$\boldsymbol{\pi}$上的有向线段$\boldsymbol{b}$满足$\boldsymbol{a}-\boldsymbol{b}$垂直于$\boldsymbol{\pi}$, 则$d(\boldsymbol{a},\boldsymbol{b})$是所有$d(\boldsymbol{a},\boldsymbol{c})$中的最小者, 这里$\boldsymbol{c}$是$\boldsymbol{\pi}$上任一有向线段. 这一结果在欧氏空间也成立.

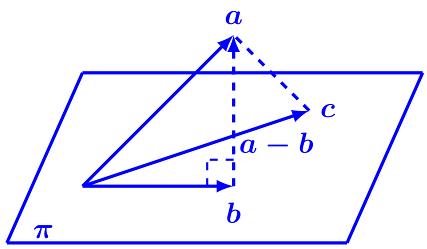

图 3.4.3: 向量到平面的距离

定义 3.4.10 设V是欧氏空间, W是V的子空间. 对$\boldsymbol{\alpha}\in V$, 若$\boldsymbol{\alpha}$与W中每个向量都正交, 则称**向量$\boldsymbol{\alpha}$与子空间$\boldsymbol{W}$正交**, 记为$\boldsymbol{\alpha}\perp W$.

显然, 只有零向量才与其所在子空间正交.

定理 3.4.7 设V是欧氏空间, W是V的子空间. 任取$\boldsymbol{\alpha}\in V$, 若存在$\boldsymbol{\beta}\in W$满足$\boldsymbol{\alpha}-\boldsymbol{\beta}\perp W$, 则对任意$\boldsymbol{\gamma}\in W$均有

$$d(\boldsymbol{\alpha},\boldsymbol{\beta})\leqslant d(\boldsymbol{\alpha},\boldsymbol{\gamma})$$

证 根据定义3.4.9,

$$[d(\boldsymbol{\alpha},\boldsymbol{\gamma})]^2=|\boldsymbol{\alpha}-\boldsymbol{\gamma}|^2=|(\boldsymbol{\alpha}-\boldsymbol{\beta})+(\boldsymbol{\beta}-\boldsymbol{\gamma})|^2$$

因$\boldsymbol{\alpha}-\boldsymbol{\beta}\perp W,\boldsymbol{\beta}-\boldsymbol{\gamma}\in W$, 故$\boldsymbol{\alpha}-\boldsymbol{\beta}\perp\boldsymbol{\beta}-\boldsymbol{\gamma}$. 利用勾股定理,

$$\begin{aligned}[d(\boldsymbol{\alpha},\boldsymbol{\gamma})]^2&=|\boldsymbol{\alpha}-\boldsymbol{\beta}|^2+|\boldsymbol{\beta}-\gamma|^2\\&\geqslant|\boldsymbol{\alpha}-\boldsymbol{\beta}|^2=[d(\boldsymbol{\alpha},\boldsymbol{\beta})]^2\end{aligned}$$

于是

$$d(\boldsymbol{\alpha},\boldsymbol{\beta})\leqslant d(\boldsymbol{\alpha},\boldsymbol{\gamma})$$

在上述定理中, $d(\boldsymbol{\alpha},\boldsymbol{\beta})$可视为$\boldsymbol{\alpha}$到子空间$W$的距离. 从这个角度出发, 就可给出最小二乘法.

已知无解线性方程组

$$\boldsymbol{AX}=\boldsymbol{b} \tag{3.21}$$

其中$\boldsymbol{A}=[a_{ij}]\in\mathbf{R}^{m\times n},\boldsymbol{b}=(b_1,b_2,\ldots,b_m)^{\mathrm{T}}\in\mathbf{R}^m,\boldsymbol{X}=(x_1,x_2,\ldots,x_n)^{\mathrm{T}}$. 若存在$c_1^*,c_2^*,\ldots,c_n^*\in\mathbf{R}$使得

$$\sum_{i=1}^{m}\left(a_{i1}c_1^*+a_{i2}c_2^*+\cdots+a_{in}c_n^*-b_i\right)^2 \tag{3.22}$$

最小, 则称$\boldsymbol{X}^*=(c_1^*,c_2^*,\ldots,c_n^*)^{\mathrm{T}}$为方程组(3.21)的**最小二乘解**.

考虑欧氏空间$\mathbf{R}^m$及其子空间$\mathrm{Col}(\boldsymbol{A})$. 令$\boldsymbol{b}^*=\boldsymbol{AX}^*$, 则$\boldsymbol{b}^*\in\mathrm{Col}(\boldsymbol{A})$且式(3.22) 即为

$$|\boldsymbol{b}^*-\boldsymbol{b}|^2=[d(\boldsymbol{b}^*,\boldsymbol{b})]^2$$

由此可知, 求$\boldsymbol{X}^*$使式(3.22)最小等价于求$\boldsymbol{b}^*\in\mathrm{Col}(\boldsymbol{A})$, 使$\boldsymbol{b}^*$与$\boldsymbol{b}$的距离最小. 根据定理3.4.7, 只需求$\boldsymbol{b}^*\in\mathrm{Col}(\boldsymbol{A})$, 使$(\boldsymbol{b}-\boldsymbol{b}^*)\perp\mathrm{Col}(\boldsymbol{A})$, 即应使$\mathrm{Col}(\boldsymbol{A})$中的任一向量$\boldsymbol{AX}$ ($\boldsymbol{X}\in\mathbf{R}^n$)与$\boldsymbol{b}-\boldsymbol{b}^*$的内积均为零, 即

$$(\boldsymbol{b}-\boldsymbol{b}^*,\boldsymbol{AX})=(\boldsymbol{b}-\boldsymbol{b}^*)^{\mathrm{T}}\boldsymbol{AX}=0$$

由于上式应对任一$\boldsymbol{X}\in\mathbf{R}^n$均成立, 故应有

$$(\boldsymbol{b}-\boldsymbol{b}^*)^{\mathrm{T}}\boldsymbol{A}=\mathbf{0}$$

即

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}^*=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$$

所以, $\boldsymbol{X}^*$应满足线性方程组$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$. 由上述推导过程得到下述结论:

定理 3.4.8 设$\boldsymbol{AX}=\boldsymbol{b}$是无解线性方程组, 则其最小二乘解是下列线性方程组

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} \tag{3.23}$$

的解.

下面的例子表明线性方程组(3.23)总有解.

例 3.4.8 任取$\boldsymbol{A}\in\mathbf{R}^{m\times n},\boldsymbol{b}\in\mathbf{R}^m$, 则线性方程组$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$一定有解.

证 只需证明系数矩阵$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$与增广矩阵$[\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}]$有相同的秩.

显然, $\mathrm{r}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right) \leqslant \mathrm{r}\left(\left[\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}\right]\right)$. 另一方面,

$$\left[\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}\right] = \boldsymbol{A}^{\mathrm{T}}[\boldsymbol{A},\boldsymbol{b}]$$

故$\mathrm{r}\left(\left[\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}\right]\right) \leqslant \mathrm{r}\left(\boldsymbol{A}^{\mathrm{T}}\right)$, 再根据例??有$\mathrm{r}\left(\boldsymbol{A}^{\mathrm{T}}\right) = \mathrm{r}(\boldsymbol{A}) = \mathrm{r}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)$, 故

$$\mathrm{r}\left(\left[\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}\right]\right) \leqslant \mathrm{r}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)$$

综上所述, 得

$$\mathrm{r}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right) = \mathrm{r}\left(\left[\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}\right]\right)$$

定理3.4.8和例3.4.8告诉我们, 实数域上的任一无解线性方程组都有最小二乘解.

例 3.4.9 物理学中的胡克定律指出, 在弹性范围内, 一个匀质弹簧的长度x是作用力y的线性函数, 可设之为

$$y = y_0 + kx \tag{3.24}$$

其中k称为该弹簧的倔强系数. 现有一匀质弹簧, 未受力时的长度为6.1cm. 当弹簧被分别施加2N, 4N以及6N的外力时, 测得其长度分别为7.6cm, 8.7cm以及10.4cm. 求这个弹簧的倔强系数(结果保留一位小数).

解 已知4组对应数据

x_i/cm	6.1	7.6	8.7	10.4
y_i/N	0	2	4	6

将之代入式(3.24)得

$$\begin{cases} y_0 + 6.1k = 0 \\ y_0 + 7.6k = 2 \\ y_0 + 8.7k = 4 \\ y_0 + 10.4k = 6 \end{cases} \tag{3.25}$$

不难验证方程组(3.25)无解.

令

$$\boldsymbol{A}=\begin{bmatrix}1 & 6.1\\ 1 & 7.6\\ 1 & 8.7\\ 1 & 10.4\end{bmatrix},\quad \boldsymbol{X}=\begin{bmatrix}y_0\\ k\end{bmatrix},\quad \boldsymbol{b}=\begin{bmatrix}0\\ 2\\ 4\\ 6\end{bmatrix}$$

构造方程组

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$$

因为矩阵$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$可逆, 故

$$\boldsymbol{X}^{*}=\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}=(-8.6,1.4)^{\mathrm{T}}$$

由此得该弹簧的倔强系数$k=1.4\mathrm{N/cm}$.

在例3.4.9中, 由式(3.24)和式(3.25)可知, 目标线性方程是一条过点$A(6.1,0)$, $B(7.6,2)$, $C(8.7,4)$, $D(10.4,6)$的直线. 然而, 线性方程组(3.25)无解导致不存在这样的直线. 由式(3.22)可知, 直线$y=f(x)=1.4x-8.6$是满足

拟合直线

$$(f(x_A)-y_A)^2+(f(x_B)-y_B)^2+(f(x_C)-y_C)^2+(f(x_D)-y_D)^2 \tag{3.26}$$

最小的一条直线, 其中x_A,x_B,x_C,x_D 分别为A,B,C,D的横坐标, y_A,y_B,y_C,y_D分别为A,B,C,D的纵坐标. 如图3.4.4所示, 式(3.26)即为四个正方形的面积.

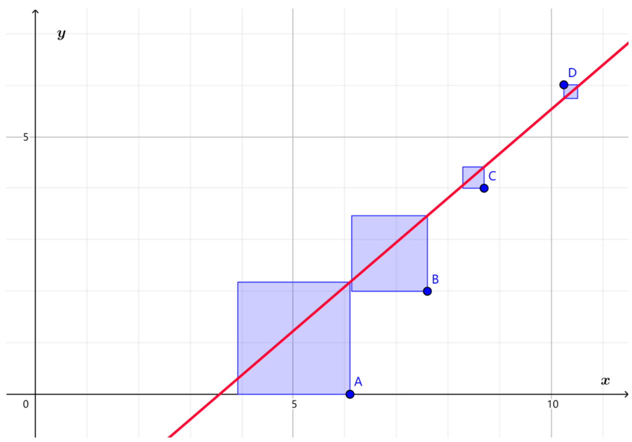

图 3.4.4: 拟合直线

扫描右侧二维码, 选择任意一条直线$y=f(x)$, 观察其对应的式(3.26)的值, 验证直线$y=1.4x-8.6$是满足式(3.26)最小的直线.

§3.5 *欧氏空间(二)

在实向量空间中, 通过引入内积, 赋予了向量大小和方向的内涵, 建立了欧氏空间. 本节将在实数域$\mathbf{R}$上的线性空间中引入内积, 赋予向量大小和方向的内涵, 建立一般的欧氏空间.

§3.5.1 实线性空间的内积

定义 3.5.1 设V是实数域$\mathbf{R}$上的一个线性空间. 若对V中任意两个向量$\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, 有唯一确定的记作$(\boldsymbol{\alpha},\boldsymbol{\beta})$的实数与之对应, 并且满足下列条件:

(1) $(\boldsymbol{\alpha},\boldsymbol{\beta})=(\boldsymbol{\beta},\boldsymbol{\alpha})$,

(2) $(k\boldsymbol{\alpha},\boldsymbol{\beta})=k(\boldsymbol{\alpha},\boldsymbol{\beta})$,

(3) $(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\gamma})=(\boldsymbol{\alpha},\boldsymbol{\gamma})+(\boldsymbol{\beta},\boldsymbol{\gamma})$,

(4) $(\boldsymbol{\alpha},\boldsymbol{\alpha})\geqslant 0$, 当且仅当$\boldsymbol{\alpha}=\boldsymbol{\theta}$时$(\boldsymbol{\alpha},\boldsymbol{\alpha})=0$,

这里$\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$是$V$中的任意向量, k是任意实数, 则称$(\boldsymbol{\alpha},\boldsymbol{\beta})$为向量$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$的**内积**. 这种定义了内积的实线性空间称为**Euclid空间**, 简称**欧氏空间**.

显然, 上节中的内积与欧氏空间是定义3.5.1的特殊情形.

例 3.5.1 设$\boldsymbol{\alpha}=(a_1,a_2,\ldots,a_n)$, $\boldsymbol{\beta}=(b_1,b_2,\ldots,b_n)$是$\mathbf{R}^n$中任意两个向量, 规定

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=a_1b_1+2a_2b_2+\cdots+na_nb_n$$

容易验证它是$\mathbf{R}^n$上的一个内积, 且与$\mathbf{R}^n$上的标准内积(式(3.11))不同.

事实上, 在同一个实线性空间V上可以定义不同的内积, 构成不同的欧氏空间, 其差别在于“尺度”不同.

例 3.5.2 设V是定义在闭区间$[a,b]$上的所有连续实函数构成的函数空间$C[a,b]$. 对于V中任意两个函数$f(x),g(x)$, 规定

$$(f(x),g(x))=\int_a^b f(x)g(x)\mathrm{d}x \tag{3.27}$$

则容易证明它满足定义3.5.1中的四个条件, 所以V对按式(3.27)引入的内积构成一个欧氏空间. 这个内积称为**函数空间$C[a,b]$的标准内积**.

例 3.5.3 在线性空间$\mathbf{R}^{n\times n}$中, 对任意两个n阶矩阵$\boldsymbol{A},\boldsymbol{B}$, 规定

$$(\boldsymbol{A},\boldsymbol{B})=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}b_{ij}=\operatorname{tr}\left(\boldsymbol{A}\boldsymbol{B}^{\mathrm{T}}\right) \tag{3.28}$$

容易验证它满足定义3.5.1中的四个条件, 因此$\mathbf{R}^{n\times n}$对于式(3.28)所定义的内积构成一个欧氏空间.

类似地, 可以证明: 在上节的欧氏空间$\mathbf{R}^{n}$中所叙述的内容对于本节欧氏空间V也都是对的. 为明确起见, 把主要结论概述如下:

(1) 向量$\boldsymbol{\alpha}$的长度规定为$|\boldsymbol{\alpha}|=\sqrt{(\boldsymbol{\alpha},\boldsymbol{\alpha})}$;

(2) Cauchy − Schwarz不等式: $|(\boldsymbol{\alpha},\boldsymbol{\beta})|\leqslant|\boldsymbol{\alpha}||\boldsymbol{\beta}|$;

(3) 两个非零向量$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$的夹角$\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle$由其余弦值确定

$$\cos\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{|\boldsymbol{\alpha}||\boldsymbol{\beta}|}$$

(4) 三角不等式: $|\boldsymbol{\alpha}+\boldsymbol{\beta}|\leqslant|\boldsymbol{\alpha}|+|\boldsymbol{\beta}|$;

(5) 勾股定理: $|\boldsymbol{\alpha}+\boldsymbol{\beta}|^2=|\boldsymbol{\alpha}|^2+|\boldsymbol{\beta}|^2$, 当$\boldsymbol{\alpha}$ 与$\boldsymbol{\beta}$ 正交时;

(6) 两个向量$\boldsymbol{\alpha},\boldsymbol{\beta}$正交的充要条件是$(\boldsymbol{\alpha},\boldsymbol{\beta})=0$.

下面我们对一般欧氏空间的内容再作一些补充.

§3.5.2 度量矩阵

设V是一个n维实线性空间, $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$是$V$的一个基, $\boldsymbol{\alpha}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n,\boldsymbol{\beta}=y_1\boldsymbol{\alpha}_1+y_2\boldsymbol{\alpha}_2+\cdots+y_n\boldsymbol{\alpha}_n$是$V$中的任意两个向量. 下面考察它们的内积:

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\left(\sum_{i=1}^{n}x_i\boldsymbol{\alpha}_i,\sum_{j=1}^{n}y_j\boldsymbol{\alpha}_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}x_iy_j\left(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j\right)$$

利用矩阵的乘法, 上式还可表示为

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Y} \tag{3.29}$$

其中

$$\boldsymbol{X}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix},\quad \boldsymbol{Y}=\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}$$

分别是$\boldsymbol{\alpha},\boldsymbol{\beta}$关于基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$的坐标,

$$\boldsymbol{A}=\begin{bmatrix}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_n)\\ (\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_n)\\ \vdots & \vdots & & \vdots\\ (\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_n)\end{bmatrix}$$

称为**基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$的度量矩阵**. 显然$\boldsymbol{A}$是对称矩阵.

由此可知, 只要知道了一个基的度量矩阵, 亦即知道了任意两个基向量的内积, 那么整个空间中任意两个向量的内积就可立即得到. 这就是说, 度量矩阵完全确定了内积. 显然度量矩阵越简单, 内积的表达式也越简单. 如果度量矩阵是单位矩阵, 即基向量两两正交, 并且长度都是1, 那么内积的表达式将最简单.

§3.5.3 标准正交基

定理 3.5.1 设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$是欧氏空间$V$的一个正交向量组, 则$\boldsymbol{\alpha}_1$,$\boldsymbol{\alpha}_2$, $\ldots$,$\boldsymbol{\alpha}_n$线性无关.

证明同定理3.4.2.

显然, n维欧氏空间的一个基是标准正交基的充分必要条件是: 它的度量矩阵是单位矩阵.

在标准正交基下, 向量的内积表达式是简单的, 即

定理 3.5.2 设$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\ldots,\boldsymbol{\varepsilon}_n$为$n$维欧氏空间$V$的一个标准正交基, $\boldsymbol{\alpha}=x_1\boldsymbol{\varepsilon}_1+x_2\boldsymbol{\varepsilon}_2+\cdots+x_n\boldsymbol{\varepsilon}_n,\boldsymbol{\beta}=y_1\boldsymbol{\varepsilon}_1+y_2\boldsymbol{\varepsilon}_2+\cdots+y_n\boldsymbol{\varepsilon}_n$, 则

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\sum_{i=1}^{n}x_iy_i$$

证 由式(3.29)及标准正交基的度量矩阵是单位矩阵立即可得上式.

欧氏空间$\mathbf{R}^n$中定义的标准内积式(3.11)正是上式在$\mathbf{R}^n$的自然基下的坐标表达式.

定理 3.5.3 设$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\ldots,\boldsymbol{\varepsilon}_n$是欧氏空间$V$的一个标准正交基, 则任一$V$中的向量$\boldsymbol{\alpha}$可以表示成

$$\boldsymbol{\alpha}=(\boldsymbol{\alpha},\boldsymbol{\varepsilon}_1)\,\boldsymbol{\varepsilon}_1+(\boldsymbol{\alpha},\boldsymbol{\varepsilon}_2)\,\boldsymbol{\varepsilon}_2+\cdots+(\boldsymbol{\alpha},\boldsymbol{\varepsilon}_n)\,\boldsymbol{\varepsilon}_n$$

或

$$\boldsymbol{\alpha}=\sum_{i=1}^{n}\left(\boldsymbol{\alpha},\boldsymbol{\varepsilon}_i\right)\boldsymbol{\varepsilon}_i$$

证 设向量$\boldsymbol{\alpha}$在基$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\ldots,\boldsymbol{\varepsilon}_n$下的坐标为$x_1,x_2,\ldots,x_n$, 即

$$\boldsymbol{\alpha}=x_1\boldsymbol{\varepsilon}_1+x_2\boldsymbol{\varepsilon}_2+\cdots+x_n\boldsymbol{\varepsilon}_{\boldsymbol{n}}$$

作向量$\boldsymbol{\alpha}$与向量$\boldsymbol{\varepsilon}_1$的内积, 有

$$\left(\boldsymbol{\alpha},\boldsymbol{\varepsilon}_1\right)=x_1\left(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_1\right)+x_2\left(\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_1\right)+\cdots+\boldsymbol{x}_n\left(\boldsymbol{\varepsilon}_n,\boldsymbol{\varepsilon}_1\right)=x_1$$

同理可得

$$x_2=\left(\boldsymbol{\alpha},\boldsymbol{\varepsilon}_2\right),\ldots,x_n=\left(\boldsymbol{\alpha},\boldsymbol{\varepsilon}_n\right)$$

在一般的欧氏空间里是否存在标准正交基呢? 下面的定理给出了肯定的回答.

定理 3.5.4 *在n维欧氏空间V里必有标准正交基.*

同样地, 可以证明: 定理3.4.4在一般的欧氏空间中仍然成立, 其方法仍称为Schmidt正交化方法. 因为在n维欧氏空间V中必存在一个基, 所以可通过Schmidt正交化方法由此求出一个标准正交基.

本节对实数域$\mathbf{R}$上的线性空间引入了内积, 建立了一般的欧氏空间, 又称实内积空间. 对于复数域$\mathbf{C}$上的线性空间, 应该如何定义内积使之成为复内积空间呢? 感兴趣的读者可参见有关的参考书.

§3.6 *线性映射

在许多数学分支和实际问题中都会遇到线性空间之间的映射. 例如, 解析几何中的坐标变换, 二次型理论中的线性替换等等. 并且这种映射保持加法和数量乘法两种运算, 称之为线性映射. 实际上, 正是对线性映射的研究促进了矩阵理论的发展. 例如, Cayley在研究线性变换的复合时提出了矩阵乘法的定义, 并进一步研究了逆矩阵的问题. 本节研究线性映射的理论.

§3.6.1 映射

定义 3.6.1 *设S,T是两个集合. 如果有一个确定的法则, 使S中每个元素x, 都有T中唯一确定的元素y与之对应, 就称这个法则是S到T的一个***映射***. 一个集合S到其自身的映射称为S上的***变换***.*

常用$\boldsymbol{\sigma}, \boldsymbol{\tau}, \ldots$表示映射. 如果$\boldsymbol{\sigma}$是$S$到$T$的映射, 则记为

$$\boldsymbol{\sigma} : S \to T$$

如果$x \in S$通过$\boldsymbol{\sigma}$对应$y \in T$, 则记为

$$\boldsymbol{\sigma} : x \to y \quad \text{或} \quad \boldsymbol{\sigma}(x) = y$$

此时称y为x在$\boldsymbol{\sigma}$下的**像**, 称x为y在$\boldsymbol{\sigma}$下的**原像**.

例 3.6.1 设$S = T = \mathbf{R}, \boldsymbol{\sigma}(x) = x^2$, 则$\boldsymbol{\sigma}$是$S$到$T$的一个映射.

例 3.6.2 在解析几何中, 设S表示空间中所有点的集合, $T = \mathbf{R}^3$, 则在建立空间直角坐标系后, 空间的点与其坐标的对应关系是S到T的一个映射.

例 3.6.3 设$S = \mathbf{R}[x]_{n+1}, T = \mathbf{R}[x]_n, \mathcal{D}(f(x)) = f'(x), f(x) \in S$, 则$\mathcal{D}$是$S$到$T$的一个映射.

由上面三个例子可知:

(1) S与T可以是相同的集合, 也可以是不同的集合;

(2) 对S中每个元素x, 需要有T中唯一确定的元素与它对应;

(3) 一般来说, T的元素不一定都是S中元素的像.

设$\boldsymbol{\sigma} : S \to T$, 记$\boldsymbol{\sigma}(S) = \{\boldsymbol{\sigma}(x) \mid x \in S\}$, 称之为$S$在映射$\boldsymbol{\sigma}$下的**像集**. 显然, $\boldsymbol{\sigma}(S) \subseteq T$.

定义 3.6.2 设$\boldsymbol{\sigma}$是S到T的映射. 若$\boldsymbol{\sigma}(S) = T$, 则称$\boldsymbol{\sigma}$为**满射**; 若对任意$a$, $b \in S, a \neq b$, 均有$\boldsymbol{\sigma}(a) \neq \boldsymbol{\sigma}(b)$, 则称$\boldsymbol{\sigma}$为**单射**; 若$\boldsymbol{\sigma}$既是满射又是单射, 则称$\boldsymbol{\sigma}$是**双射**, 也称为**一一对应**.

不难看出, 例3.6.3中的映射是满射但不是单射, 例3.6.2中的映射是一一对应, 例3.6.1中的映射既不是满射也不是单射.

定义 3.6.3 设$\boldsymbol{\sigma}, \boldsymbol{\tau}$是$S$到$T$的两个映射. 若对$\forall a \in S$都有$\boldsymbol{\sigma}(a) = \boldsymbol{\tau}(a)$, 则称$\boldsymbol{\sigma}$与$\boldsymbol{\tau}$**相等**, 记为$\boldsymbol{\sigma} = \boldsymbol{\tau}$.

此外, 映射之间有时可以进行运算.

定义 3.6.4 设$\boldsymbol{\sigma}$是集合S到T的映射, $\boldsymbol{\tau}$是集合T到U的映射, $\boldsymbol{\tau}$与$\boldsymbol{\sigma}$的**乘积** $\boldsymbol{\tau\sigma}$定义为$\boldsymbol{\tau\sigma}(a) = \boldsymbol{\tau}(\boldsymbol{\sigma}(a))$, $a \in S$.

例 3.6.4 设$S = \mathbf{R}^{m\times n}, T = \mathbf{R}^{n\times m}$, 则$\boldsymbol{\sigma}(\boldsymbol{A}) = \boldsymbol{A}^{\mathrm{T}}, \boldsymbol{A} \in \mathbf{R}^{m\times n}$是$S$到$T$的一个映射, $\boldsymbol{\tau}(\boldsymbol{B}) = \mathrm{r}(\boldsymbol{B}), \boldsymbol{B} \in \mathbf{R}^{n\times m}$是$T$到自然数集合$\mathbf{N}$的一个映射, $\boldsymbol{\tau\sigma}(\boldsymbol{A}) = \boldsymbol{\tau}(\boldsymbol{\sigma}(\boldsymbol{A})) = \mathrm{r}(\boldsymbol{A}^{\mathrm{T}})$是$S$到$\mathbf{N}$的一个映射.

可以证明, 映射的乘法不满足交换律, 但满足结合律.

§3.6.2 线性映射的概念

在解析几何中, 常需要把空间中的点向某一固定平面作投影, 例如向Oxy面投影. 在线性代数中, 这实际上是实数域$\mathbf{R}$上的3维向量空间$\mathbf{R}^3$到自身的一个投影映射$\boldsymbol{p}$:

$$\boldsymbol{p}(x,y,z)=(x,y,0),\quad (x,y,z)\in\mathbf{R}^3$$

不难发现

$$\begin{gathered}\boldsymbol{p}\left[(x_1,y_1,z_1)+(x_2,y_2,z_2)\right]=\boldsymbol{p}\left(x_1,y_1,z_1\right)+\boldsymbol{p}\left(x_2,y_2,z_2\right)\\ \boldsymbol{p}\left[k\left(x_1,y_1,z_1\right)\right]=k\boldsymbol{p}\left(x_1,y_1,z_1\right)\end{gathered}$$

其中(x_1,y_1,z_1)与(x_2,y_2,z_2)是$\mathbf{R}^3$中任意向量, k是任一实数. 即$\boldsymbol{p}$保持$\mathbf{R}^3$中的线性运算.

扫描右侧二维码, 选择$\mathbf{R}^3$中的向量, 并验证$\boldsymbol{p}$的线性性质.

函数的定积分是$\mathbf{R}$上的线性空间$C[a,b]$到$\mathbf{R}$的一个映射, 它也具有下述保持线性运算的性质:

$$\int_a^b[f(x)+g(x)]\mathrm{d}x=\int_a^b f(x)\mathrm{d}x+\int_a^b g(x)\mathrm{d}x$$

$$\int_a^b kf(x)\mathrm{d}x=k\int_a^b f(x)\mathrm{d}x$$

其中$f(x),g(x)$是$C[a,b]$中任意的连续函数, k是任一实数.

定义 3.6.5 设V_1与V_2是数域F上的两个线性空间. 若V_1到V_2的一个映射$\boldsymbol{\sigma}$保持加法运算和数量乘法运算, 即对任意的$\boldsymbol{\alpha}$, $\boldsymbol{\beta}\in V_1, k\in F$, 均有

$$\boldsymbol{\sigma}(\boldsymbol{\alpha}+\boldsymbol{\beta})=\boldsymbol{\sigma}(\boldsymbol{\alpha})+\boldsymbol{\sigma}(\boldsymbol{\beta}),\quad \boldsymbol{\sigma}(k\boldsymbol{\alpha})=k\boldsymbol{\sigma}(\boldsymbol{\alpha})\tag{3.30}$$

则称$\boldsymbol{\sigma}$是V_1到V_2的一个**线性映射**.

例 3.6.5 设V是数域F上的一个线性空间, 映射$\boldsymbol{\varepsilon}:V\to V$定义为

$$\boldsymbol{\varepsilon}(\boldsymbol{\alpha})=\boldsymbol{\alpha},\quad \forall\ \boldsymbol{\alpha}\in V$$

易证$\boldsymbol{\varepsilon}$是一个线性映射, 称为**恒等映射**(或**单位映射**).

设V_1与V_2是数域F上的两个线性空间, 映射$\boldsymbol{\sigma}:V_1\to V_2$定义为

$$\boldsymbol{\sigma}(\boldsymbol{\alpha})=\boldsymbol{\theta},\quad \forall\ \boldsymbol{\alpha}\in V_1$$

易证$\boldsymbol{\sigma}$是一个线性映射, 称为**零映射**, 记为$\underline{\mathbf{0}}$.

例 3.6.6 设$\boldsymbol{A}$是数域F上的一个$m\times n$矩阵, 映射$\boldsymbol{\sigma_A}:F^n\to F^m$定义为

$$\boldsymbol{\sigma_A}(\boldsymbol{\alpha})=\boldsymbol{A\alpha},\quad \forall\ \boldsymbol{\alpha}\in F^n$$

易证$\boldsymbol{\sigma_A}$是F^n到F^m的一个线性映射.

注: 这里将F^n, F^m中的元素视为列矩阵, 例??中的旋转, 切变, 投影映射都是通过上例中的对应法则定义的线性映射.

例 3.6.7 映射$\mathcal{D}:\mathbf{R}[x]_{n+1}\to\mathbf{R}[x]_n$定义为

$$\mathcal{D}[f(x)]=\frac{\mathrm{d}}{\mathrm{d}x}f(x),\quad \forall\ f(x)\in\mathbf{R}[x]_{n+1}$$

易证$\mathcal{D}$是一个线性映射.

例 3.6.8 映射$\mathcal{S}:\mathbf{R}[x]_n\to\mathbf{R}[x]_{n+1}$定义为

$$\mathcal{S}[f(x)]=\int_0^x f(t)\mathrm{d}t,\quad \forall\ f(x)\in\mathbf{R}[x]_n$$

易证$\mathcal{S}$是一个线性映射.

例 3.6.9 映射$\boldsymbol{\sigma}:\mathbf{R}^2\to\mathbf{R}^3$定义为

$$\boldsymbol{\sigma}\left(a_1,a_2\right)=\left(a_1+2a_2,a_1+a_2,a_2\right)$$

其中$a_1,a_2\in\mathbf{R}$, 证明$\boldsymbol{\sigma}$是一个线性映射.

证明 因

$$\begin{aligned}
&\boldsymbol{\sigma}\left[\left(a_1,a_2\right)+\left(b_1,b_2\right)\right]=\boldsymbol{\sigma}\left(a_1+b_1,a_2+b_2\right)\\
=&\left(a_1+b_1+2a_2+2b_2,a_1+b_1+a_2+b_2,a_2+b_2\right)\\
=&\left(a_1+2a_2,a_1+a_2,a_2\right)+\left(b_1+2b_2,b_1+b_2,b_2\right)\\
=&\boldsymbol{\sigma}\left(a_1,a_2\right)+\boldsymbol{\sigma}\left(b_1,b_2\right)\\
&\boldsymbol{\sigma}\left[k\left(a_1,a_2\right)\right]=\boldsymbol{\sigma}\left(ka_1,ka_2\right)\\
=&\left(ka_1+2ka_2,ka_1+ka_2,ka_2\right)\\
=&k\left(a_1+2a_2,a_1+a_2,a_2\right)\\
=&k\boldsymbol{\sigma}\left(a_1,a_2\right)
\end{aligned}$$

其中$a_1,a_2,b_1,b_2,k\in\mathbf{R}$, 故$\boldsymbol{\sigma}$是一个线性映射.

扫描右侧二维码, 拖动红色向量,了解例3.6.9中的线性映射, 并验证它的

线性性质.

性质 3.6.1 设$\boldsymbol{\sigma}$是线性空间V_1到V_2的线性映射, 则

(1) $\boldsymbol{\sigma}(\boldsymbol{\theta}_1)=\boldsymbol{\theta}_2, \boldsymbol{\theta}_i \in V_i; \boldsymbol{\sigma}(-\boldsymbol{\alpha})=-\boldsymbol{\sigma}(\boldsymbol{\alpha})$;

(2) $\boldsymbol{\sigma}$保持线性组合与线性关系式不变;

(3) $\boldsymbol{\sigma}$把线性相关的向量组变成线性相关的向量组.

证明 (1) $\boldsymbol{\sigma}(\boldsymbol{\theta}_1)=\boldsymbol{\sigma}(0\boldsymbol{\alpha})=0\boldsymbol{\sigma}(\boldsymbol{\alpha})=\boldsymbol{\theta}_2$,

$$\boldsymbol{\sigma}(\boldsymbol{\alpha})+\boldsymbol{\sigma}(-\boldsymbol{\alpha})=\boldsymbol{\sigma}[\boldsymbol{\alpha}+(-\boldsymbol{\alpha})]=\boldsymbol{\sigma}(\boldsymbol{\theta})=\boldsymbol{\theta}$$

故$\boldsymbol{\sigma}(-\boldsymbol{\alpha})=-\boldsymbol{\sigma}(\boldsymbol{\alpha})$.

(2) 设$\boldsymbol{\alpha}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m$, 则

$$\boldsymbol{\sigma}(\boldsymbol{\alpha})=k_1\boldsymbol{\sigma}\left(\boldsymbol{\alpha}_1\right)+k_2\boldsymbol{\sigma}\left(\boldsymbol{\alpha}_2\right)+\cdots+k_m\boldsymbol{\sigma}\left(\boldsymbol{\alpha}_m\right)$$

即线性组合的像等于像的线性组合且组合系数相同.

(3) 由(1)与(2)可证(3).

$\boldsymbol{\sigma}$也可能把线性无关的向量组变成线性相关的向量组, 即(3)的逆不成立.

例 3.6.10 考虑投影映射$\boldsymbol{p}:\mathbf{R}^3\rightarrow\mathbf{R}^3$, 容易验证: $\mathbf{R}^3$中的三个线性无关向量$\boldsymbol{\alpha}_1=(1,1,1), \boldsymbol{\alpha}_2=(1,1,0), \boldsymbol{\alpha}_3=(1,0,0)$的像$\boldsymbol{p}\left(\alpha_1\right)=(1,1,0), \boldsymbol{p}\left(\alpha_2\right)=(1,1,0), \boldsymbol{p}\left(\alpha_3\right)=(1,0,0)$是线性相关的.

§3.6.3 线性映射的矩阵表示

设V_1,V_2是数域F上的线性空间, $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$是$V_1$的一个基, $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m$是$V_2$的一个基, $\boldsymbol{\sigma}$是V_1到V_2的一个线性映射, 则

$$\left\{\begin{array}{l}
\boldsymbol{\sigma}\left(\boldsymbol{\alpha}_1\right)=a_{11}\boldsymbol{\beta}_1+a_{21}\boldsymbol{\beta}_2+\cdots+a_{m1}\boldsymbol{\beta}_m \\
\boldsymbol{\sigma}\left(\boldsymbol{\alpha}_2\right)=a_{12}\boldsymbol{\beta}_1+a_{22}\boldsymbol{\beta}_2+\cdots+a_{m2}\boldsymbol{\beta}_m \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\boldsymbol{\sigma}\left(\boldsymbol{\alpha}_n\right)=a_{1n}\boldsymbol{\beta}_1+a_{2n}\boldsymbol{\beta}_2+\cdots+a_{mn}\boldsymbol{\beta}_m
\end{array}\right. \tag{3.31}$$

上式可形式地写成矩阵形式

$$\left[\boldsymbol{\sigma}\left(\boldsymbol{\alpha}_1\right),\boldsymbol{\sigma}\left(\boldsymbol{\alpha}_2\right),\ldots,\boldsymbol{\sigma}\left(\boldsymbol{\alpha}_n\right)\right]=\left[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m\right]\boldsymbol{A}$$

其中

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

称为**线性映射$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$与基$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m$下的矩阵表示**.

显然, $\boldsymbol{\sigma}$在给定的一对基下的矩阵表示$\boldsymbol{A}$是唯一确定的, 而在不同基下的矩阵表示一般不相同.

有了线性映射$\boldsymbol{\sigma}$在一对基下的矩阵表示$\boldsymbol{A}$之后, 可以得到线性空间V_1中向量$\boldsymbol{\alpha}$与它在V_2中的像$\boldsymbol{\sigma}(\boldsymbol{\alpha})$之间的坐标关系.

任取$\boldsymbol{\alpha}\in V_1$, 不妨设$\boldsymbol{\alpha}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n$, 它的像

$$\boldsymbol{\sigma}(\boldsymbol{\alpha})=y_1\boldsymbol{\beta}_1+y_2\boldsymbol{\beta}_2+\cdots+y_m\boldsymbol{\beta}_m=[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m]\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

又

$$\begin{aligned}\boldsymbol{\sigma}(\boldsymbol{\alpha})&=\boldsymbol{\sigma}\left(\sum_{i=1}^{n}x_i\boldsymbol{\alpha}_i\right)=[\boldsymbol{\sigma}(\boldsymbol{\alpha}_1),\boldsymbol{\sigma}(\boldsymbol{\alpha}_2),\ldots,\boldsymbol{\sigma}(\boldsymbol{\alpha}_n)]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\\ &=([\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m]\,\boldsymbol{A})\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m]\,\boldsymbol{A}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\end{aligned}$$

于是, 由坐标的唯一性可得

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}=\boldsymbol{A}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{3.32}$$

式(3.32)称为线性映射在给定基$\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n$与$\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_m$下**向量坐标变换公式**.

定理 3.6.1 设V_1, V_2是数域F上的线性空间, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$是$V_1$的一个基, $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m$是$V_2$的一个基. 对$F^{m\times n}$中给定的矩阵$\boldsymbol{A} = [a_{ij}]_{m\times n}$, 存在唯一的线性映射$\boldsymbol{\sigma}: V_1 \to V_2$, 它在这两个基下的矩阵表示为$\boldsymbol{A}$.

证 对任意的$\boldsymbol{\alpha} = \sum_{i=1}^n x_i \boldsymbol{\alpha}_i \in V_1$, 取

$$\boldsymbol{\beta} = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m]\, \boldsymbol{A} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in V_2$$

定义映射$\boldsymbol{\sigma}: V_1 \longrightarrow V_2$为$\boldsymbol{\sigma}(\boldsymbol{\alpha}) = \boldsymbol{\beta}$. 容易验证$\boldsymbol{\sigma}$是$V_1$到$V_2$的线性映射. 事实上, 任取$\boldsymbol{\gamma}_j = \sum\limits_{i=1}^n x_{ij}\boldsymbol{\alpha}_i \in V_1, j = 1, 2, k \in F$, 则有

$$\begin{aligned}
\boldsymbol{\sigma}(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2) &= [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m]\, \boldsymbol{A} \begin{bmatrix} x_{11} + x_{12} \\ x_{21} + x_{22} \\ \vdots \\ x_{n1} + x_{n2} \end{bmatrix} \\
&= [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m]\, \boldsymbol{A} \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix} + [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m]\, \boldsymbol{A} \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix} \\
&= \boldsymbol{\sigma}(\boldsymbol{\gamma}_1) + \boldsymbol{\sigma}(\boldsymbol{\gamma}_2) \\
\boldsymbol{\sigma}(k\boldsymbol{\gamma}_1) &= [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m]\, \boldsymbol{A} \begin{bmatrix} kx_{11} \\ kx_{21} \\ \vdots \\ kx_{n1} \end{bmatrix} \\
&= k\,[\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m]\, \boldsymbol{A} \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix} = k\boldsymbol{\sigma}(\boldsymbol{\gamma}_1)
\end{aligned}$$

又因为

$$\boldsymbol{\sigma}(\boldsymbol{\alpha}_i) = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m]\boldsymbol{A}\boldsymbol{\varepsilon}_i, i = 1, 2, ..., n$$

所以

$$[\boldsymbol{\sigma}(\boldsymbol{\alpha}_1), \boldsymbol{\sigma}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\sigma}(\boldsymbol{\alpha}_n)]$$

$$= [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m] \boldsymbol{A} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m] \boldsymbol{A}$$

最后证明唯一性. 若还有线性映射$\boldsymbol{\tau}: V_1 \to V_2$, 满足

$$[\boldsymbol{\tau}(\boldsymbol{\alpha}_1), \boldsymbol{\tau}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\tau}(\boldsymbol{\alpha}_n)] = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m] \boldsymbol{A}$$

则

$$[\boldsymbol{\tau}(\boldsymbol{\alpha}_1), \boldsymbol{\tau}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\tau}(\boldsymbol{\alpha}_n)] = [\boldsymbol{\sigma}(\boldsymbol{\alpha}_1), \boldsymbol{\sigma}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\sigma}(\boldsymbol{\alpha}_n)]$$

即

$$\boldsymbol{\sigma}(\boldsymbol{\alpha}_i) = \boldsymbol{\tau}(\boldsymbol{\alpha}_i), \quad i = 1, 2, \ldots, n$$

于是, 对任意$\boldsymbol{\alpha} = \sum_{i=1}^n x_i \boldsymbol{\alpha}_i \in V_1$,

$$\begin{aligned} \boldsymbol{\sigma}(\boldsymbol{\alpha}) &= [\boldsymbol{\sigma}(\boldsymbol{\alpha}_1), \boldsymbol{\sigma}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\sigma}(\boldsymbol{\alpha}_n)] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= [\boldsymbol{\tau}(\boldsymbol{\alpha}_1), \boldsymbol{\tau}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\tau}(\boldsymbol{\alpha}_n)] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \boldsymbol{\tau}(\boldsymbol{\alpha}) \end{aligned}$$

从而, $\boldsymbol{\sigma} = \boldsymbol{\tau}$.

因此, 在给定基以后, $\boldsymbol{\sigma}$与其矩阵表示$\boldsymbol{A}$是相互唯一确定的. 若记V_1到V_2的所有线性映射组成的集合为V^*, 则在给定的一对基下, 线性映射与其矩阵表示的对应关系是V^*到$F^{m\times n}$的一一对应.

例 3.6.11 零映射在任意一对基下的矩阵表示均为零矩阵, 恒等映射在任意一对相同的基下的矩阵表示均为单位矩阵(见例3.6.5).

例 3.6.12 求$\mathbf{R}[x]_{n+1}$到$\mathbf{R}[x]_n$的线性映射$\mathcal{D}$:

$$\mathcal{D}[f(x)] = \frac{\mathrm{d}}{\mathrm{d}x}f(x), \quad \forall\ f(x) \in \mathbf{R}[x]_{n+1}$$

在基$1, x, x^2, \ldots, x^n$与基$1, x, x^2, \ldots, x^{n-1}$下的矩阵表示.

解 因$\mathcal{D}(1) = 0, \mathcal{D}(x) = 1, \mathcal{D}\left(x^2\right) = 2x, \ldots, \mathcal{D}\left(x^n\right) = nx^{n-1}$, 将之写成矩阵形式

$$[\mathcal{D}(1), \mathcal{D}(x), \ldots, \mathcal{D}\left(x^n\right)] = [1, x^2, \ldots, x^{n-1}]\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & n \end{bmatrix}_{n\times(n+1)}$$

于是所求矩阵表示为

$$\boldsymbol{D} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & n \end{bmatrix}_{n\times(n+1)}$$

注: 对$\mathbf{R}[x]_{n+1}$到$\mathbf{R}[x]_{n+1}$的线性映射$\mathcal{D}'$:

$$\mathcal{D}'[f(x)] = \frac{\mathrm{d}}{\mathrm{d}x}f(x), \quad \forall\ f(x) \in \mathbf{R}[x]_{n+1}$$

其在基$1, x, x^2, \ldots, x^n$与基$1, x, x^2, \ldots, x^n$下的矩阵表示为

$$\boldsymbol{D}' = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}_{(n+1)\times(n+1)}$$

例 3.6.13 求$\mathbf{R}[x]_n$到$\mathbf{R}[x]_{n+1}$的线性映射$\mathcal{S}$:

$$\mathcal{S}[f(x)] = \int_0^x f(t)\mathrm{d}t, \quad \forall\ f(x) \in \mathbf{R}[x]_n$$

在基$1, x, x^2, \ldots, x^{n-1}$与基$1, x, x^2, \ldots, x^n$下的矩阵表示.

解 $\mathcal{S}(1)=\int_0^x 1\ \mathrm{d}t=x, \mathcal{S}(x)=\int_0^x t\ \mathrm{d}t=\frac{1}{2}x^2,\ \mathcal{S}\left(x^2\right)=\int_0^x t^2\ \mathrm{d}t=\frac{1}{3}x^3, \ldots, \mathcal{S}\left(x^{n-1}\right)=\int_0^x t^{n-1}\ \mathrm{d}t=\frac{1}{n}x^n$, 将之写成矩阵形式

$$\left[\mathcal{S}(1), \mathcal{S}(x), \ldots, \mathcal{S}\left(x^{n-1}\right)\right]=\left[1, x, \ldots, x^n\right]\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ 0 & \frac{1}{2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \frac{1}{n} \end{bmatrix}_{(n+1)\times n}$$

于是所求矩阵表示为

$$\boldsymbol{S}=\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ 0 & \frac{1}{2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \frac{1}{n} \end{bmatrix}_{(n+1)\times n}$$

例 3.6.14 求例3.6.9中定义的线性映射$\boldsymbol{\sigma}$在$\mathbf{R}^2$的自然基$\boldsymbol{\alpha}_1=(1,0), \boldsymbol{\alpha}_2=(0,1)$与$\mathbf{R}^3$的自然基$\boldsymbol{\beta}_1=(1,0,0), \boldsymbol{\beta}_2=(0,1,0), \boldsymbol{\beta}_3=(0,0,1)$下的矩阵表示.

解 由定义可知$\boldsymbol{\sigma}\left(\boldsymbol{\alpha}_1\right)=(1,1,0)=\boldsymbol{\beta}_1+\boldsymbol{\beta}_2, \boldsymbol{\sigma}\left(\boldsymbol{\alpha}_2\right)=(2,1,1)=2\boldsymbol{\beta}_1+\boldsymbol{\beta}_2+\boldsymbol{\beta}_3$, 将之写为矩阵的形式

$$\left[\boldsymbol{\sigma}\left(\boldsymbol{\alpha}_1\right), \boldsymbol{\sigma}\left(\boldsymbol{\alpha}_2\right)\right]=\left[\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3\right]\begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

于是所求矩阵表示为

$$\boldsymbol{A}=\begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

实际上, 给定$\mathbf{R}^2, \mathbf{R}^3$的自然基, 例3.6.14中矩阵$\boldsymbol{A}$唯一确定的线性映射即为例3.6.9中所定义映射.

一般地, 线性空间的基是不唯一的, 若选取不同的一对基, 则同一个线性映射的矩阵表示也不同, 它们之间有什么关系呢? 对此有如下结果.

定理 3.6.2 设$\boldsymbol{\sigma}$是V_1到V_2的一个线性映射, $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$与$\boldsymbol{\alpha}'_1,\boldsymbol{\alpha}'_2,\ldots,\boldsymbol{\alpha}'_n$是$V_1$的两个基, 且

$$[\boldsymbol{\alpha}'_1,\boldsymbol{\alpha}'_2,\ldots,\boldsymbol{\alpha}'_n]=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n]\boldsymbol{P} \tag{3.33}$$

$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m$与$\boldsymbol{\beta}'_1,\boldsymbol{\beta}'_2,\ldots,\boldsymbol{\beta}'_m$是$V_2$的两个基, 且

$$[\boldsymbol{\beta}'_1,\boldsymbol{\beta}'_2,\ldots,\boldsymbol{\beta}'_m]=[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m]\boldsymbol{Q} \tag{3.34}$$

若线性映射$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$与$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m$下的矩阵表示为$\boldsymbol{A}$, 在基$\boldsymbol{\alpha}'_1,\boldsymbol{\alpha}'_2,\ldots,\boldsymbol{\alpha}'_n$与$\boldsymbol{\beta}'_1,\boldsymbol{\beta}'_2,\ldots,\boldsymbol{\beta}'_m$下的矩阵表示为$\boldsymbol{B}$, 则

$$\boldsymbol{B}=\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{P}$$

证 根据假设有

$$[\boldsymbol{\sigma}(\boldsymbol{\alpha}_1),\boldsymbol{\sigma}(\boldsymbol{\alpha}_2),\ldots,\boldsymbol{\sigma}(\boldsymbol{\alpha}_n)]=[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m]\,\boldsymbol{A} \tag{3.35}$$

$$[\boldsymbol{\sigma}(\boldsymbol{\alpha}'_1),\boldsymbol{\sigma}(\boldsymbol{\alpha}'_2),\ldots,\boldsymbol{\sigma}(\boldsymbol{\alpha}'_n)]=[\boldsymbol{\beta}'_1,\boldsymbol{\beta}'_2,\ldots,\boldsymbol{\beta}'_m]\,\boldsymbol{B} \tag{3.36}$$

将式(3.33)、(3.34)代入式(3.36)整理可得

$$[\boldsymbol{\sigma}(\boldsymbol{\alpha}_1),\boldsymbol{\sigma}(\boldsymbol{\alpha}_2),\ldots,\boldsymbol{\sigma}(\boldsymbol{\alpha}_n)]\,\boldsymbol{P}=[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m]\,\boldsymbol{Q}\boldsymbol{B} \tag{3.37}$$

将式(3.35)代入式(3.37)得

$$[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m]\,\boldsymbol{A}\boldsymbol{P}=[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m]\,\boldsymbol{Q}\boldsymbol{B}$$

因$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m$线性无关, 故

$$\boldsymbol{A}\boldsymbol{P}=\boldsymbol{Q}\boldsymbol{B}$$

又由于$\boldsymbol{Q}$是可逆方阵, 所以

$$\boldsymbol{B}=\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{P}$$

因此, 一个线性映射$\boldsymbol{\sigma}: V_1\to V_2$有一系列$F^{m\times n}$中的矩阵表示, 它们彼此相抵. 反之, 可以证明, 在$F^{m\times n}$中所有相抵的矩阵代表同一个线性映射. 人们自然会问: 能否找到一对适当的基, 使得$\boldsymbol{\sigma}$在该对基下的矩阵表示最简单呢? 请读者思考.

§3.6.4 线性空间的同构

前面已经提到, 线性空间是代数系统$F^n, F^{m\times n}, F[x]$等的抽象与提高, 并且后者的许多概念和性质完全可推广到线性空间, 那么为什么不同的线性空间会有相同的性质呢? 在此我们做一简单介绍.

定义 3.6.6 设V_1, V_2是数域F上的线性空间, $\boldsymbol{\sigma}$是V_1到V_2的一个线性映射. 若$\boldsymbol{\sigma}$是双射, 则称$\boldsymbol{\sigma}$是V_1到V_2的**同构映射**. 若V_1与V_2之间存在同构映射, 则称V_1与V_2同构, 记为$V_1 \cong V_2$.

当两个线性空间同构时, 它们在线性相关性方面的性质完全相同. 例如, 在F^n与$F[x]_n$之间定义映射

$$\boldsymbol{\sigma}: (a_1, a_2, \ldots, a_n) \to a_1 + a_2x + a_3x^2 + \cdots + a_nx^{n-1}$$

则$\boldsymbol{\sigma}$是F^n与$F[x]_n$之间的同构映射. 因此, $F[x]_n$具有与F^n完全相同的线性相关性结论. 更进一步, 我们有

定理 3.6.3 数域F上的每个n维线性空间都同构于向量空间F^n.

证 设V是数域F上的线性空间, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$是$V$的一个基, 对任一$\boldsymbol{\alpha} \in V$, 设$(a_1, a_2, \ldots, a_n)^{\mathrm{T}}$是$\boldsymbol{\alpha}$关于基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$的坐标, 构造$V$到$F^n$的映射

$$\boldsymbol{\sigma}: \boldsymbol{\alpha} \to (a_1, a_2, \ldots, a_n)$$

则容易证明$\boldsymbol{\sigma}$是双射且满足

$$\boldsymbol{\sigma}(\boldsymbol{\alpha}+\boldsymbol{\beta}) = \boldsymbol{\sigma}(\boldsymbol{\alpha}) + \boldsymbol{\sigma}(\boldsymbol{\beta}), \quad \boldsymbol{\sigma}(k\boldsymbol{\alpha}) = k\boldsymbol{\sigma}(\boldsymbol{\alpha})$$

其中$\boldsymbol{\alpha}, \boldsymbol{\beta} \in V, k \in F$, 所以$\boldsymbol{\sigma}$是$V$到$F^n$的同构映射, 于是$V \cong F^n$.

根据这个定理, 可得F上任意两个n维线性空间都同构, 并且通过F^n可获知F上所有n维线性空间的性质.

§3.7 *线性变换

本节讨论线性空间V到其自身的线性映射, 称这样的线性映射为线性空间V上的**线性变换**. 显然, 上节中有关线性映射的结果对线性变换都成立.

例 3.7.1 设V是数域F上的一个线性空间. 取定$k \in F$, 定义映射$\boldsymbol{\sigma}: V \to V$为

$$\boldsymbol{\sigma}(\boldsymbol{\alpha}) = k\boldsymbol{\alpha}, \quad \forall\ \boldsymbol{\alpha} \in V$$

易证$\boldsymbol{\sigma}$是V上的一个线性变换, 称之为**数乘变换**. 事实上,

$$\begin{aligned}\boldsymbol{\sigma}(a\boldsymbol{\alpha} + b\boldsymbol{\beta}) &= k(a\boldsymbol{\alpha} + b\boldsymbol{\beta}) = k(a\boldsymbol{\alpha}) + k(b\boldsymbol{\beta}) \\ &= a(k\boldsymbol{\alpha}) + b(k\boldsymbol{\beta}) = a\boldsymbol{\sigma}(\boldsymbol{\alpha}) + b\boldsymbol{\sigma}(\boldsymbol{\beta})\end{aligned}$$

其中$\boldsymbol{\alpha}, \boldsymbol{\beta} \in V, a, b \in F$.

特别地, 当$k = 0$ 时, 称此变换为**零变换**, 记为$\underline{\mathbf{0}}$, 即

$$\underline{\mathbf{0}}(\boldsymbol{\alpha}) = \boldsymbol{\theta}, \quad \forall\ \boldsymbol{\alpha} \in V$$

当$k = 1$ 时, 称此变换为**恒等变换**或**单位变换**, 记为$\boldsymbol{\varepsilon}$, 即

$$\boldsymbol{\varepsilon}(\boldsymbol{\alpha}) = \boldsymbol{\alpha}, \quad \forall\ \boldsymbol{\alpha} \in V$$

扫描右侧二维码, 拖动滑动条, 选择不同的$k \in \mathbf{R}$, 观察由其定义的$\mathbf{R}^2$上的数乘变换.

例 3.7.2 在$F[x]$中, 定义变换

$$\boldsymbol{\sigma}[f(x)] = [f(x)]^2, \quad \forall\ f(x) \in F[x]$$

因为对任意的$a, b \in F, f(x), g(x) \in F[x]$, 有

$$\begin{aligned}&\boldsymbol{\sigma}[af(x) + bg(x)] \\ =&[af(x) + bg(x)]^2 \\ =&a^2[f(x)]^2 + b^2[g(x)]^2 + 2abf(x)g(x)\end{aligned}$$

而

$$a\boldsymbol{\sigma}[f(x)] + b\boldsymbol{\sigma}[g(x)] = a[f(x)]^2 + b[g(x)]^2$$

所以

$$\boldsymbol{\sigma}[af(x) + bg(x)] \neq a\boldsymbol{\sigma}[f(x)] + b\boldsymbol{\sigma}[g(x)]$$

由此可知, 该变换不是线性变换.

例 3.7.3 设$\boldsymbol{\sigma}$是V上的线性变换, $\boldsymbol{\varepsilon}$是V上的恒等变换, 则$\boldsymbol{\sigma\varepsilon} = \boldsymbol{\varepsilon\sigma} = \boldsymbol{\sigma}$.

定义 3.7.1 设$\boldsymbol{\sigma}$是线性空间V上的一个变换. 若存在V上的另一个变换$\boldsymbol{\tau}$, 使

$$\boldsymbol{\tau}\boldsymbol{\sigma} = \boldsymbol{\sigma}\boldsymbol{\tau} = \boldsymbol{\varepsilon}$$

则称$\boldsymbol{\sigma}$是**可逆变换**, 称$\boldsymbol{\tau}$是$\boldsymbol{\sigma}$的**逆变换**, 这里$\boldsymbol{\varepsilon}$是V上的恒等变换.

不难证明, 变换$\boldsymbol{\sigma}$可逆当且仅当$\boldsymbol{\sigma}$是双射, 并且当$\boldsymbol{\sigma}$可逆时其逆唯一, 记为$\boldsymbol{\sigma}^{-1}$.

定理 3.7.1 可逆线性变换的逆变换也是线性变换.

证 设$\boldsymbol{\sigma}$是可逆线性变换,$\boldsymbol{\sigma}^{-1}$是它的逆变换. 任取$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in V, k \in F$, 令

$$\boldsymbol{\sigma}^{-1}(\boldsymbol{\alpha}_1) = \boldsymbol{\beta}_1, \quad \boldsymbol{\sigma}^{-1}(\boldsymbol{\alpha}_2) = \boldsymbol{\beta}_2$$

则由可逆变换的定义可得

$$\boldsymbol{\sigma}(\boldsymbol{\beta}_1) = \boldsymbol{\alpha}_1, \quad \boldsymbol{\sigma}(\boldsymbol{\beta}_2) = \boldsymbol{\alpha}_2$$

已知$\boldsymbol{\sigma}$是线性的, 故

$$\boldsymbol{\sigma}(\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2) = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2$$

由此得

$$\boldsymbol{\sigma}^{-1}(\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2) = \boldsymbol{\beta}_1 + \boldsymbol{\beta}_2 = \boldsymbol{\sigma}^{-1}(\boldsymbol{\alpha}_1) + \boldsymbol{\sigma}^{-1}(\boldsymbol{\alpha}_2)$$

同理, 由$k\boldsymbol{\alpha}_1 = k\boldsymbol{\sigma}(\boldsymbol{\beta}_1) = \boldsymbol{\sigma}(k\boldsymbol{\beta}_1)$可得

$$\boldsymbol{\sigma}^{-1}(k\boldsymbol{\alpha}_1) = k\boldsymbol{\beta}_1 = k\boldsymbol{\sigma}^{-1}(\boldsymbol{\alpha}_1)$$

所以, $\boldsymbol{\sigma}^{-1}$也是线性的.

此外, 还可以定义线性变换的加法、乘法、数乘运算, 并且其和、乘积以及数乘积仍然是线性变换.

定义 3.7.2 设$\boldsymbol{\sigma}, \boldsymbol{\tau}$是数域$F$上线性空间$V$上的线性变换, $k \in F$, $\boldsymbol{\sigma}$与$\boldsymbol{\tau}$的**和**, $\boldsymbol{\sigma} + \boldsymbol{\tau}$定义为

$$(\boldsymbol{\sigma} + \boldsymbol{\tau})(\boldsymbol{\alpha}) = \boldsymbol{\sigma}(\boldsymbol{\alpha}) + \boldsymbol{\tau}(\boldsymbol{\alpha}), \quad \forall\ \boldsymbol{\alpha} \in V$$

$\boldsymbol{\sigma}$与$\boldsymbol{\tau}$的**乘积**, $\boldsymbol{\sigma}\boldsymbol{\tau}$定义为

$$(\boldsymbol{\sigma}\boldsymbol{\tau})(\boldsymbol{\alpha}) = \boldsymbol{\sigma}(\boldsymbol{\tau}(\boldsymbol{\alpha})) \quad \forall\ \boldsymbol{\alpha} \in V$$

k与$\boldsymbol{\sigma}$的**数乘**, $k\boldsymbol{\sigma}$定义为

$$(k\boldsymbol{\sigma})(\boldsymbol{\alpha}) = k\boldsymbol{\sigma}(\boldsymbol{\alpha}), \quad \forall\ \boldsymbol{\alpha} \in V$$

性质 3.7.1 设$\boldsymbol{\sigma}, \boldsymbol{\tau}$是数域$F$上线性空间$V$上的线性变换, $k \in F$, 则$\boldsymbol{\sigma}+\boldsymbol{\tau}$、$\boldsymbol{\sigma\tau}$、$k\boldsymbol{\sigma}$都是$V$上的线性变换.

证 对于任意$\boldsymbol{\alpha}, \boldsymbol{\beta} \in V, a \in F$, 有

$$\begin{aligned}(\boldsymbol{\sigma}+\boldsymbol{\tau})(\boldsymbol{\alpha}+\boldsymbol{\beta}) &= \boldsymbol{\sigma}(\boldsymbol{\alpha}+\boldsymbol{\beta})+\boldsymbol{\tau}(\boldsymbol{\alpha}+\boldsymbol{\beta}) \\ &= \boldsymbol{\sigma}(\boldsymbol{\alpha})+\boldsymbol{\sigma}(\boldsymbol{\beta})+\boldsymbol{\tau}(\boldsymbol{\alpha})+\boldsymbol{\tau}(\boldsymbol{\beta}) \\ &= (\boldsymbol{\sigma}+\boldsymbol{\tau})(\boldsymbol{\alpha})+(\boldsymbol{\sigma}+\boldsymbol{\tau})(\boldsymbol{\beta}) \\ (\boldsymbol{\sigma}+\boldsymbol{\tau})(a\boldsymbol{\alpha}) &= \boldsymbol{\sigma}(a\boldsymbol{\alpha})+\boldsymbol{\tau}(a\boldsymbol{\alpha}) = a\boldsymbol{\sigma}(\boldsymbol{\alpha})+a\boldsymbol{\tau}(\boldsymbol{\alpha}) \\ &= a(\boldsymbol{\sigma}(\boldsymbol{\alpha})+\boldsymbol{\tau}(\boldsymbol{\alpha})) = a(\boldsymbol{\sigma}+\boldsymbol{\tau})(\boldsymbol{\alpha})\end{aligned}$$

因此, $\boldsymbol{\sigma}+\boldsymbol{\tau}$是$V$上的线性变换. 同理可证$\boldsymbol{\sigma\tau}$、$k\boldsymbol{\sigma}$也是$V$上的线性变换.

容易验证, V上线性变换的加法与数量乘法满足线性空间定义中的8条运算法则, 因此数域F上线性空间V的所有线性变换组成的集合也是一个线性空间.

§3.7.1 线性变换的矩阵表示

设$\boldsymbol{\sigma}$是数域F上线性空间V上的线性变换, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$为$V$的一个基,

$$\boldsymbol{\sigma}\left(\boldsymbol{\alpha}_j\right) = \sum_{i=1}^{n} a_{ij}\boldsymbol{\alpha}_i, \quad j = 1, 2, \ldots, n$$

上式可形式地写成矩阵形式

$$[\boldsymbol{\sigma}(\boldsymbol{\alpha}_1), \boldsymbol{\sigma}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\sigma}(\boldsymbol{\alpha}_n)] = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\, \boldsymbol{A}$$

其中

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

称为**线性变换$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$下的矩阵表示**.

例 3.7.4 零变换在任一基下的矩阵表示均为零矩阵; 恒等变换在任一基下的矩阵表示均为单位矩阵; 数乘变换在任一基下矩阵表示均为数量矩阵.

例 3.7.5 $\mathbf{R}[x]_{n+1}$上的线性变换$\mathcal{D}$:

$$\mathcal{D}[f(x)] = \frac{\mathrm{d}}{\mathrm{d}x}f(x), \quad \forall\ f(x) \in \mathbf{R}[x]_{n+1}$$

在基$1, x, x^2, \ldots, x^n$下的矩阵表示

$$\boldsymbol{D} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}_{(n+1)\times(n+1)}$$

与第§3.6.3节中的讨论一样, 有了线性变换$\boldsymbol{\sigma}$在一个基下的矩阵表示$\boldsymbol{A}$之后, 可以得到线性空间V中向量$\boldsymbol{\alpha}$与它的像$\boldsymbol{\sigma}(\boldsymbol{\alpha})$之间的坐标关系.

定理 3.7.2 设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$是数域$F$上$n$维线性空间$V$的一个基, $\boldsymbol{\sigma}$是V上的一个线性变换, 其在基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$下的矩阵表示为$\boldsymbol{A}$. 若$\boldsymbol{\alpha} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n, \boldsymbol{\sigma}(\boldsymbol{\alpha}) = y_1\boldsymbol{\alpha}_1 + y_2\boldsymbol{\alpha}_2 + \cdots + y_n\boldsymbol{\alpha}_n$, 则

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

在n维线性空间中取定一组基后, 其上的一个线性变换就与一个n阶方阵对应. 这个对应还保持线性变换的运算.

定理 3.7.3 设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$是数域$F$上$n$维线性空间$\boldsymbol{V}$的一个基, V上的两个线性变换$\boldsymbol{\sigma}$与$\boldsymbol{\tau}$在这个基下矩阵表示分别为$\boldsymbol{A}$与$\boldsymbol{B}$, 则

(1) 线性变换的和$\boldsymbol{\sigma} + \boldsymbol{\tau}$在基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$下的矩阵表示为$\boldsymbol{A} + \boldsymbol{B}$;

(2) 线性变换的数乘积$k\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$下的矩阵表示为$k\boldsymbol{A}$;

(3) 线性变换的乘积$\boldsymbol{\sigma}\boldsymbol{\tau}$ 在基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$下的矩阵表示为$\boldsymbol{A}\boldsymbol{B}$;

(4) 线性变换$\boldsymbol{\sigma}$可逆当且仅当$\boldsymbol{A}$可逆, 并且$\boldsymbol{\sigma}^{-1}$在基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$下的矩阵表示为$\boldsymbol{A}^{-1}$.

证 由已知得

$$[\boldsymbol{\sigma}(\boldsymbol{\alpha}_1), \boldsymbol{\sigma}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\sigma}(\boldsymbol{\alpha}_n)] = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\,\boldsymbol{A}$$

$$[\boldsymbol{\tau}(\boldsymbol{\alpha}_1), \boldsymbol{\tau}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\tau}(\boldsymbol{\alpha}_n)] = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\,\boldsymbol{B}$$

(1)
$$\begin{aligned}
&[(\boldsymbol{\sigma}+\boldsymbol{\tau})(\boldsymbol{\alpha}_1), (\boldsymbol{\sigma}+\boldsymbol{\tau})(\boldsymbol{\alpha}_2), \ldots, (\boldsymbol{\sigma}+\boldsymbol{\tau})(\boldsymbol{\alpha}_n)] \\
&= [\boldsymbol{\sigma}(\boldsymbol{\alpha}_1)+\boldsymbol{\tau}(\boldsymbol{\alpha}_1), \boldsymbol{\sigma}(\boldsymbol{\alpha}_2)+\boldsymbol{\tau}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\sigma}(\boldsymbol{\alpha}_n)+\boldsymbol{\tau}(\boldsymbol{\alpha}_n)] \\
&= [\boldsymbol{\sigma}(\boldsymbol{\alpha}_1), \boldsymbol{\sigma}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\sigma}(\boldsymbol{\alpha}_n)] + [\boldsymbol{\tau}(\boldsymbol{\alpha}_1), \boldsymbol{\tau}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\tau}(\boldsymbol{\alpha}_n)] \\
&= [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\,\boldsymbol{A} + [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\,\boldsymbol{B} \\
&= [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\,(\boldsymbol{A}+\boldsymbol{B})
\end{aligned}$$

(2)
$$\begin{aligned}
&[(k\boldsymbol{\sigma})(\boldsymbol{\alpha}_1), (k\boldsymbol{\sigma})(\boldsymbol{\alpha}_2), \ldots, (k\boldsymbol{\sigma})(\boldsymbol{\alpha}_n)] \\
&= [k[\boldsymbol{\sigma}(\boldsymbol{\alpha}_1)], k[\boldsymbol{\sigma}(\boldsymbol{\alpha}_2)], \ldots, k[\boldsymbol{\sigma}(\boldsymbol{\alpha}_n)]] \\
&= k\,[\boldsymbol{\sigma}(\boldsymbol{\alpha}_1), \boldsymbol{\sigma}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\sigma}(\boldsymbol{\alpha}_n)] \\
&= k\,([\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\,\boldsymbol{A}) \\
&= [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\,(k\boldsymbol{A})
\end{aligned}$$

(3) 设$\boldsymbol{B} = [b_{ij}]$, 则$\boldsymbol{\tau}(\boldsymbol{\alpha}_j) = \sum\limits_{i=1}^{n} b_{ij}\boldsymbol{\alpha}_i$, 从而

$$\begin{aligned}
&[(\boldsymbol{\sigma}\boldsymbol{\tau})(\boldsymbol{\alpha}_1), (\boldsymbol{\sigma}\boldsymbol{\tau})(\boldsymbol{\alpha}_2), \ldots, (\boldsymbol{\sigma}\boldsymbol{\tau})(\boldsymbol{\alpha}_n)] \\
&= [\boldsymbol{\sigma}[\boldsymbol{\tau}(\boldsymbol{\alpha}_1)], \boldsymbol{\sigma}[\boldsymbol{\tau}(\boldsymbol{\alpha}_2)], \ldots, \boldsymbol{\sigma}[\boldsymbol{\tau}(\boldsymbol{\alpha}_n)]] \\
&= \left[\boldsymbol{\sigma}(\sum_{i=1}^{n} b_{i1}\boldsymbol{\alpha}_i), \boldsymbol{\sigma}(\sum_{i=1}^{n} b_{i2}\boldsymbol{\alpha}_i), \ldots, \boldsymbol{\sigma}(\sum_{i=1}^{n} b_{in}\boldsymbol{\alpha}_i)\right] \\
&= \left[\sum_{i=1}^{n} b_{i1}\boldsymbol{\sigma}(\boldsymbol{\alpha}_i), \sum_{i=1}^{n} b_{i2}\boldsymbol{\sigma}(\boldsymbol{\alpha}_i), \ldots, \sum_{i=1}^{n} b_{in}\boldsymbol{\sigma}(\boldsymbol{\alpha}_i)\right] \\
&= [\boldsymbol{\sigma}(\boldsymbol{\alpha}_1), \boldsymbol{\sigma}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\sigma}(\boldsymbol{\alpha}_n)]\,\boldsymbol{B} \\
&= ([\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\,\boldsymbol{A})\,\boldsymbol{B} \\
&= [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\,(\boldsymbol{A}\boldsymbol{B})
\end{aligned}$$

(4) 设$\boldsymbol{\sigma}$可逆, 并且 $\left[\boldsymbol{\sigma}^{-1}(\boldsymbol{\alpha}_1), \boldsymbol{\sigma}^{-1}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\sigma}^{-1}(\boldsymbol{\alpha}_n)\right] = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\,\boldsymbol{A}'$, 由(3)得

$$\begin{aligned}
&\left[(\boldsymbol{\sigma}\boldsymbol{\sigma}^{-1})(\boldsymbol{\alpha}_1), (\boldsymbol{\sigma}\boldsymbol{\sigma}^{-1})(\boldsymbol{\alpha}_2), \ldots, (\boldsymbol{\sigma}\boldsymbol{\sigma}^{-1})(\boldsymbol{\alpha}_n)\right] \\
&= [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\,(\boldsymbol{A}\boldsymbol{A}') \\
&= [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n]\,\boldsymbol{I}
\end{aligned}$$

于是$\boldsymbol{A}\boldsymbol{A}'=\boldsymbol{I}$, $\boldsymbol{A}$可逆, 且$\boldsymbol{A}'=\boldsymbol{A}^{-1}$.

反之亦然.

例 3.7.6 在$\mathbf{R}^2$中, 规定

$$\boldsymbol{\sigma}(a_1,a_2)=(a_1+a_2,a_2)$$

其中$a_1,a_2\in\mathbf{R}$. 证明$\boldsymbol{\sigma}$是可逆的线性变换, 并求$\boldsymbol{\sigma}^{-1}$.

解 因

$$\begin{aligned}\boldsymbol{\sigma}[k(a_1,a_2)+l(b_1,b_2)]&=\boldsymbol{\sigma}(ka_1+lb_1,ka_2+lb_2)\\&=(ka_1+lb_1+ka_2+lb_2,ka_2+lb_2)\\&=(ka_1+ka_2,ka_2)+(lb_1+lb_2,lb_2)\\&=k\boldsymbol{\sigma}(a_1,a_2)+l\boldsymbol{\sigma}(b_1,b_2)\end{aligned}$$

其中$k,l,a_i,b_i\in\mathbf{R},i=1,2$, 故$\boldsymbol{\sigma}$是线性变换.

取$\boldsymbol{\varepsilon}_1=(1,0),\boldsymbol{\varepsilon}_2=(0,1)$, 则$\boldsymbol{\sigma}(\boldsymbol{\varepsilon}_1)=\boldsymbol{\sigma}(1,0)=(1,0)=\boldsymbol{\varepsilon}_1,\boldsymbol{\sigma}(\boldsymbol{\varepsilon}_2)=\boldsymbol{\sigma}(0,1)=(1,1)=\boldsymbol{\varepsilon}_1+\boldsymbol{\varepsilon}_2$, 将之写为矩阵形式

$$[\boldsymbol{\sigma}(\boldsymbol{\varepsilon}_1),\boldsymbol{\sigma}(\boldsymbol{\varepsilon}_2)]=[\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2]\begin{bmatrix}1&1\\0&1\end{bmatrix}$$

因$\boldsymbol{\sigma}$在基$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2$下的矩阵表示$\boldsymbol{A}=\begin{bmatrix}1&1\\0&1\end{bmatrix}$可逆, 故$\boldsymbol{\sigma}$可逆, 且$\boldsymbol{\sigma}^{-1}$在基$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2$下的矩阵表示为$\boldsymbol{A}^{-1}=\begin{bmatrix}1&-1\\0&1\end{bmatrix}$. 由此得

$$\begin{aligned}\boldsymbol{\sigma}^{-1}(a_1,a_2)&=\boldsymbol{\sigma}^{-1}(a_1\boldsymbol{\varepsilon}_1+a_2\boldsymbol{\varepsilon}_2)=a_1\boldsymbol{\sigma}^{-1}(\boldsymbol{\varepsilon}_1)+a_2\boldsymbol{\sigma}^{-1}(\boldsymbol{\varepsilon}_2)\\&=\left[\boldsymbol{\sigma}^{-1}(\boldsymbol{\varepsilon}_1),\boldsymbol{\sigma}^{-1}(\boldsymbol{\varepsilon}_2)\right]\begin{bmatrix}a_1\\a_2\end{bmatrix}=[\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2]\,\boldsymbol{A}^{-1}\begin{bmatrix}a_1\\a_2\end{bmatrix}\\&=[\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2]\begin{bmatrix}1&-1\\0&1\end{bmatrix}\begin{bmatrix}a_1\\a_2\end{bmatrix}=[\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2]\begin{bmatrix}a_1-a_2\\a_2\end{bmatrix}\\&=(a_1-a_2)\boldsymbol{\varepsilon}_1+a_2\boldsymbol{\varepsilon}_2=(a_1-a_2,a_2)\end{aligned}$$

一般地, 线性变换的矩阵表示与线性空间基的选取有关, 在不同的基下同一个线性变换的矩阵表示一般不相同. 它们之间有什么关系呢? 由定理3.6.2可得

定理 3.7.4 设$\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$与$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n$是数域$F$上$n$维线性空间$V$的两个基, $\boldsymbol{\sigma}$是V上的一个线性变换. 若$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$与$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n$下的矩阵表示分别为$\boldsymbol{A}$与$\boldsymbol{B}$, 由$\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\ldots,\boldsymbol{\alpha}_n$到$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n$的过渡矩阵为$\boldsymbol{P}$, 则$\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$.

例 3.7.7 考虑$\mathbf{R}^2$上的线性变换$\boldsymbol{\sigma}$:

$$\boldsymbol{\sigma}\left(a_1,a_2\right)=\left(2a_1+a_2,a_1+2a_2\right)$$

其中$a_1,a_2\in\mathbf{R}$. 比较$\boldsymbol{\sigma}$在自然基$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2$下的矩阵表示$\boldsymbol{A}$与在基$\boldsymbol{\alpha}_1=(1,1)$, $\boldsymbol{\alpha}_2=(-1,1)$下的矩阵表示$\boldsymbol{B}$.

解 由$\boldsymbol{\sigma}\left(\boldsymbol{\varepsilon}_1\right)=(2,1)=2\boldsymbol{\varepsilon}_1+\boldsymbol{\varepsilon}_2,\boldsymbol{\sigma}\left(\boldsymbol{\varepsilon}_2\right)=(1,2)=\boldsymbol{\varepsilon}_1+2\boldsymbol{\varepsilon}_2$, 可得

$$[\boldsymbol{\sigma}(\boldsymbol{\varepsilon}_1),\boldsymbol{\sigma}(\boldsymbol{\varepsilon}_2)]=[\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2]\begin{bmatrix}2&1\\1&2\end{bmatrix}$$

于是$\boldsymbol{A}=\begin{bmatrix}2&1\\1&2\end{bmatrix}$.

由$\boldsymbol{\sigma}\left(\boldsymbol{\alpha}_1\right)=(3,3)=3\boldsymbol{\alpha}_1+0\boldsymbol{\alpha}_2,\boldsymbol{\sigma}\left(\boldsymbol{\alpha}_2\right)=(-1,1)=0\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$, 可得

$$[\boldsymbol{\sigma}(\boldsymbol{\alpha}_1),\boldsymbol{\sigma}(\boldsymbol{\alpha}_2)]=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2]\begin{bmatrix}3&0\\0&1\end{bmatrix}$$

于是$\boldsymbol{B}=\begin{bmatrix}3&0\\0&1\end{bmatrix}$.

由$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2$到$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$的过渡矩阵为$\boldsymbol{P}=\begin{bmatrix}1&-1\\1&1\end{bmatrix}$, 且有

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix}1&-1\\1&1\end{bmatrix}^{-1}\begin{bmatrix}2&1\\1&2\end{bmatrix}\begin{bmatrix}1&-1\\1&1\end{bmatrix}=\begin{bmatrix}3&0\\0&1\end{bmatrix}=\boldsymbol{B}.$$

扫描右侧二维码, 选择例3.7.7中两个不同的基, 比较同一个线性变换$\boldsymbol{\sigma}$在这两个不同基下的矩阵表示. 选出你认为更加简单的矩阵表示.

§3.7.2 线性变换的特征值与特征向量

线性变换的矩阵表示对研究线性变换必不可少, 并且一般地同一个线性变换在不同的基下的矩阵表示也不同. 因此人们当然希望能找到一个适当的基, 使线性变换在这个基下的矩阵表示尽可能简单, 如对角矩阵(见例3.7.7). 在什么条件下线性变换的矩阵表示为对角矩阵呢?

设线性变换$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$下的矩阵表示为对角矩阵

$$\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

则

$$[\boldsymbol{\sigma}(\boldsymbol{\alpha}_1), \boldsymbol{\sigma}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\sigma}(\boldsymbol{\alpha}_n)] = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

于是

$$[\boldsymbol{\sigma}(\boldsymbol{\alpha}_1), \boldsymbol{\sigma}(\boldsymbol{\alpha}_2), \ldots, \boldsymbol{\sigma}(\boldsymbol{\alpha}_n)] = [\lambda_1\boldsymbol{\alpha}_1, \lambda_2\boldsymbol{\alpha}_2, \ldots, \lambda_n\boldsymbol{\alpha}_n]$$

由此得

$$\boldsymbol{\sigma}(\boldsymbol{\alpha}_i) = \lambda_i\boldsymbol{\alpha}_i, \quad i = 1, 2, \ldots, n \tag{3.38}$$

定义 3.7.3 设$\boldsymbol{\sigma}$为数域F上n维线性空间V上的一个线性变换. 若对F中的一个数λ, 存在V中一个非零向量$\boldsymbol{\alpha}$, 使得

$$\boldsymbol{\sigma}(\boldsymbol{\alpha}) = \lambda\boldsymbol{\alpha}$$

则称λ为$\boldsymbol{\sigma}$的**特征值**, $\boldsymbol{\alpha}$为$\boldsymbol{\sigma}$**属于**λ**的一个特征向量**.

在例3.7.7中, 当选择自然基$\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2$来表示$\mathbf{R}^2$时$\boldsymbol{\sigma}(\boldsymbol{\alpha}_1) = (3,3) = 3\boldsymbol{\alpha}_1, \boldsymbol{\sigma}(\boldsymbol{\alpha}_2) = (-1,1) = \boldsymbol{\alpha}_2$, 因此, $\boldsymbol{\alpha}_1 = (1,1), \boldsymbol{\alpha}_2 = (-1,1)$分别为$\boldsymbol{\sigma}$属于特征值$3, 1$的特征向量.

根据定义3.7.3, 由式(3.38)可知: 若线性变换$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$下的矩阵表示是对角阵$\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, 则$\lambda_1, \lambda_2, \ldots, \lambda_n$是$\boldsymbol{\sigma}$的特征值, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$是$\boldsymbol{\sigma}$分别属于$\lambda_1, \lambda_2, \ldots, \lambda_n$的特征向量, 并且线性无关. 反之, 若$\boldsymbol{\sigma}$有$n$个线性无关的特征向量$\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\ldots$, $\boldsymbol{\alpha}_n$, 设它们对应的特征值分别为$\lambda_1, \lambda_2, \ldots, \lambda_n$, 则$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$下的矩阵表示为$\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. 于是有

定理 3.7.5 设$\boldsymbol{\sigma}$是数域F上n维线性空间V上的一个线性变换, 则$\boldsymbol{\sigma}$的矩阵表示可以在某一个基下为对角矩阵的充分必要条件是$\boldsymbol{\sigma}$有n个线性无关的特征向量.

例如, 例3.7.7中2维向量空间$\mathbf{R}^2$上的线性变换$\boldsymbol{\sigma}$在由特征向量$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$构成的基下的矩阵表示为对角矩阵. 需要注意的是, 线性变换的特征值与特征向量的计算依赖于矩阵的特征值与特征向量, 还要借助行列式的有关知识. 这些问题将分别在第四章和第五章解决.

习题三

1. 有没有只含一个向量的线性空间? 有没有只含两个向量的线性空间? 有没有只含m个向量的线性空间?

2. 次数等于$n(n \geqslant 1)$的实系数多项式全体, 对于多项式的加法和数量乘法, 是否构成实数域上的线性空间?

3. 在几何空间中, 按通常的向量加法和数与向量的数量乘法, 下列集合是否是实数域$\mathbf{R}$上的线性空间?

(1) 过原点平面$\boldsymbol{\pi}$上所有向量集合;

(2) 位于第一卦限, 以原点为始点的向量集合;

(3) 位于第一、三卦限, 以原点为始点的向量集合;

(4) x轴上的向量集合;

(5) 平面$\boldsymbol{\pi}$上, 不平行于某向量的向量集合.

4. 全体实对称(反称, 上三角形)矩阵, 对于矩阵的加法和数量乘法, 是否构成实数域上的线性空间?

5. 平面上全体向量, 对于通常的加法和如下定义的数量乘法

$$k \cdot \boldsymbol{\alpha} = \boldsymbol{\theta}$$

是否构成实数域上的线性空间?

6. 设$\mathbf{R}^+$是正实数集合, $\mathbf{R}$是实数域, 规定

$$a \oplus b = ab, \quad k \cdot a = a$$

其中$a, b \in \mathbf{R}^+, k \in \mathbf{R}$, 问: $\mathbf{R}^+$关于运算"$\oplus$"和"$\cdot$"是否构成$\mathbf{R}$上的线性空间?

7. 在函数空间中,下列函数组是否线性相关?

(1) $1, \cos^2 x, \cos 2x$;

(2) $1, \sin^2 x, \cos^2 x$.

8. 在$\mathbf{R}^3$中, 求向量$\boldsymbol{\alpha}=(3,7,1)$关于基

$$\boldsymbol{\alpha}_1=(1,3,5),\quad \boldsymbol{\alpha}_2=(6,3,2),\quad \boldsymbol{\alpha}_3=(3,1,0)$$

的坐标.

9. 设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$是3维向量空间$V$的一个基, 令

$$\begin{cases}\boldsymbol{\beta}_1=\ \ \boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2+3\boldsymbol{\alpha}_3\\ \boldsymbol{\beta}_2=2\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2+4\boldsymbol{\alpha}_3\\ \boldsymbol{\beta}_3=3\boldsymbol{\alpha}_1+\ \ \boldsymbol{\alpha}_2+3\boldsymbol{\alpha}_3\end{cases}$$

(1) 证明: $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$也是$V$的一个基;

(2) 求V中的向量$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2+2\boldsymbol{\alpha}_3$关于基$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$的坐标.

10. 已知$\mathbf{R}^3$的两个基:

$$\begin{cases}\boldsymbol{\alpha}_1=(1,1,1)\\ \boldsymbol{\alpha}_2=(1,1,0)\\ \boldsymbol{\alpha}_3=(1,0,0)\end{cases}\qquad \begin{cases}\boldsymbol{\beta}_1=(0,0,1)\\ \boldsymbol{\beta}_2=(0,1,1)\\ \boldsymbol{\beta}_3=(1,1,1)\end{cases}$$

求所有关于这两个基有相同坐标的向量.

11. 求数域F上n阶对称矩阵所构成的线性空间的维数和一个基.

12. 求数域F上n阶反称矩阵所构成的线性空间的维数和一个基.

13. 证明: 若$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$能将线性空间$V$中每个向量唯一地线性表出, 则$\boldsymbol{\alpha}_1,\ \boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$是$V$的一个基.

14. 证明: 若$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$能将线性空间$V$中的每个向量线性表出, 又对$V$中某一向量$\boldsymbol{\alpha}$表示方法唯一, 则$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$是$V$的一个基.

15. 试求线性空间$F^{2\times 2}$的两个不同的基, 并求向量

$$\boldsymbol{A}=\begin{bmatrix}-1 & 3\\ 0 & 2\end{bmatrix}$$

在所求基下的坐标.

16. 证明: 实数域上全体3阶方阵的集合是$\mathbf{R}$上的9维线性空间, 并且当

$$\boldsymbol{A}=\begin{bmatrix}0 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 1\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{bmatrix}$$

时, $\boldsymbol{I}, \boldsymbol{A}, \boldsymbol{A}^2, \boldsymbol{B}, \boldsymbol{AB}, \boldsymbol{A}^2\boldsymbol{B}, \boldsymbol{B}^2, \boldsymbol{AB^2}, \boldsymbol{A}^2\boldsymbol{B}^2$为此空间的一个基.

17. 已知$\mathbf{R}^3$的两个基:

$$\begin{cases} \boldsymbol{\alpha}_1 = (1,2,1) \\ \boldsymbol{\alpha}_2 = (2,3,3) \\ \boldsymbol{\alpha}_3 = (3,7,1) \end{cases} \quad \begin{cases} \boldsymbol{\beta}_1 = (3,1,4) \\ \boldsymbol{\beta}_2 = (5,2,1) \\ \boldsymbol{\beta}_3 = (1,1,-6) \end{cases}$$

(1) 求基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$到基$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3$的过渡矩阵;

(2) 求向量$\boldsymbol{\gamma} = (1,0,0)$关于基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$的坐标;

(3) 已知向量$\boldsymbol{\xi}$关于基$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3$的坐标为$(-1,0,2)^{\mathrm{T}}$, 求$\boldsymbol{\xi}$关于基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\ \boldsymbol{\alpha}_3$的坐标.

18. 在$F[x]_4$中, 求基$1, x, x^2, x^3$到基$1, 1+x, 1+x+x^2, 1+x+x^2+x^3$的过渡矩阵. 已知在后一个基下$g(x)$的坐标为$(1,0,-2,5)^{\mathrm{T}}$, $f(x)$的坐标为$(7,0,8,-2)^{\mathrm{T}}$, 试求$f(x)+g(x)$在这两个基下的坐标.

19. 在$\mathbf{R}^{2\times 2}$中,证明

$$\boldsymbol{\alpha}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \boldsymbol{\alpha}_2 = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$$
$$\boldsymbol{\alpha}_3 = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \quad \boldsymbol{\alpha}_4 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

构成一个基, 并求矩阵

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

关于基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4$的坐标.

20. 在$\mathbf{R}^{2\times 2}$中, 证明下列两组矩阵是两个基:

$$\boldsymbol{\alpha}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \boldsymbol{\alpha}_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \boldsymbol{\alpha}_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \boldsymbol{\alpha}_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\boldsymbol{\beta}_1 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad \boldsymbol{\beta}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad \boldsymbol{\beta}_3 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{\beta}_4 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

求从基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4$到基$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{\beta}_4$的过渡矩阵, 并求矩阵$\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ 2 & -3 \end{bmatrix}$关于这两个基的坐标.

21. 设有$\mathbf{R}^3$的两个子集合

$$W_1 = \{(x_1, x_2, x_3) \mid x_1 - 2x_2 + 2x_3 = 0\}$$
$$W_2 = \{(x_1, x_2, x_3) \mid x_1 - 2x_2 + 2x_3 = 1\}$$

证明: W_1是子空间, W_2不是子空间, 并对这一结果做出几何解释.

22. 设W_1, W_2是$\mathbf{R}^n$的两个子空间, 证明: $W_1 \cap W_2$是$\mathbf{R}^n$的子空间.

23. 证明: $L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s) = L(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t)$的充分必要条件是$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s\} \cong \{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_t\}$.

24. 证明: 齐次线性方程组的解空间的每个基都是该方程组的一个基础解系.

25. 设V是n维向量空间, W是V的k维子空间$(k < n)$, 证明: W的任一个基都能扩充为V的一个基.

26. 设W是$\mathbf{R}^4$的2维子空间, W的一个基为$\boldsymbol{\alpha}_1 = (1, 2, 0, 1), \boldsymbol{\alpha}_2 = (-1, 1,$ 1,1), 试将W的这个基扩充为$\mathbf{R}^4$的基.

27. 设W是$\mathbf{R}^4$的向量$(1, -2, 5, -3), (2, 3, 1, -4), (3, 8, -3, -5)$生成的子空间,

(1) 求W的一个基和维数;

(2) 把W的这个基扩充为$\mathbf{R}^4$的一个基.

28. 设

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & -1 & 0 & 1 \\ 2 & 3 & 1 & -1 & 0 \\ 0 & 1 & 3 & -1 & -2 \\ 4 & 1 & -13 & 3 & 10 \end{bmatrix}$$

(1) 求$\boldsymbol{A}$的列空间$\mathrm{Col}(\boldsymbol{A})$和行空间$\mathrm{Row}(\boldsymbol{A})$的基与维数;

(2) 求$\boldsymbol{A}$的零空间$\mathrm{N}(\boldsymbol{A})$的基和维数.

29. 设$\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$是线性空间$V$中的向量, 试说明它们的线性组合的全体构成$V$的子空间.

30. 验证形如

$$\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}, \quad a \in F$$

的2阶矩阵的全体构成线性空间$F^{2\times 2}$的子空间.

31. 设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$是$n$维向量空间$V$的一个基, 又$V$中向量$\boldsymbol{\alpha}_{n+1}$关于这个基的坐标全不为零, 证明: $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n, \boldsymbol{\alpha}_{n+1}$中任意$n$个向量都构成$V$的一个基.

32. 设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s$与$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_t$是数域$F$上的两组$n$元向量, $\mathrm{r}\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s\}=r,\mathrm{r}\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_t\}=p$, 证明:

$$\dim\left(L\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_t\right)\right)\leqslant r+p$$

33. 设W_1,W_2是向量空间V的两个非平凡子空间, 证明: 在V中存在向量$\boldsymbol{\alpha}$, 使$\boldsymbol{\alpha}\notin W_1$且$\boldsymbol{\alpha}\notin W_2$. 试在$\mathbf{R}^3$中举例说明这个结论.

34. 已知欧氏空间$\mathbf{R}^4$的三个向量:

$$\boldsymbol{\alpha}=(1,2,-1,1),\quad \boldsymbol{\beta}=(2,3,1,-1),\quad \boldsymbol{\gamma}=(-1,-1,-2,2)$$

求$|\boldsymbol{\alpha}|,|\boldsymbol{\beta}|,|\boldsymbol{\gamma}|$及$\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle,\langle\boldsymbol{\alpha},\boldsymbol{\gamma}\rangle$.

35. 设$\boldsymbol{\alpha},\boldsymbol{\beta}$是欧氏空间$V$的两个向量, 且$\boldsymbol{\alpha}=k\boldsymbol{\beta}$, 证明:

(1) $k>0$的充分必要条件是$\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=0$;

(2) $k<0$的充分必要条件是$\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=\pi$.

36. 设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m$是欧氏空间的一组向量, 令

$$\mathbf{G}\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m\right)=\begin{bmatrix}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1)&(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)&\cdots&(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_m)\\(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1)&(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2)&\cdots&(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_m)\\\vdots&\vdots&&\vdots\\(\boldsymbol{\alpha}_m,\boldsymbol{\alpha}_1)&(\boldsymbol{\alpha}_m,\boldsymbol{\alpha}_2)&\cdots&(\boldsymbol{\alpha}_m,\boldsymbol{\alpha}_m)\end{bmatrix}$$

证明: $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m$线性无关的充分必要条件是$\mathbf{G}\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m\right)$可逆.
称$\mathbf{G}\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m\right)$为向量$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m$的**Gram矩阵**.

37. 在欧氏空间$\mathbf{R}^4$中求与$(1,1,-1,1),(1,-1,-1,1),(2,1,1,3)$都正交的单位向量.

38. 设$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}_3$是一组标准正交向量, 令

$$\boldsymbol{\xi}_1=\frac{1}{3}\left(2\boldsymbol{\eta}_1+2\boldsymbol{\eta}_2-\boldsymbol{\eta}_3\right),\boldsymbol{\xi}_2=\frac{1}{3}\left(2\boldsymbol{\eta}_1-\boldsymbol{\eta}_2+2\boldsymbol{\eta}_3\right),\boldsymbol{\xi}_3=\frac{1}{3}\left(\boldsymbol{\eta}_1-2\boldsymbol{\eta}_2-2\boldsymbol{\eta}_3\right)$$

证明: $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$也是一组标准正交向量.

39. 用Schmidt正交化方法, 由下列向量组分别构造一组标准正交向量:

(1) $\boldsymbol{\alpha}_1=(1,2,3),\boldsymbol{\alpha}_2=(0,2,3),\boldsymbol{\alpha}_3=(0,0,3)$;

(2) $\boldsymbol{\alpha}_1=(-1,0,1),\boldsymbol{\alpha}_2=(0,1,0),\boldsymbol{\alpha}_3=(1,0,1)$;

(3) $\boldsymbol{\alpha}_1=(2,1,3,-1),\boldsymbol{\alpha}_2=(7,4,3,-3),\boldsymbol{\alpha}_3=(5,7,7,8)$.

40. 已知欧氏空间$\mathbf{R}^3$的一个基:

$$\boldsymbol{\alpha}_1=(1,-1,1),\quad \boldsymbol{\alpha}_2=(-1,1,1),\quad \boldsymbol{\alpha}_3=(1,1,-1)$$

由此构造$\mathbf{R}^3$的一个标准正交基.

41. 求齐次线性方程组

$$\begin{cases} 2x_1 + x_2 - x_3 + x_4 - 3x_5 = 0 \\ x_1 + x_2 - x_3 + x_5 = 0 \\ 3x_1 + 2x_2 - 2x_3 + x_4 - 2x_5 = 0 \end{cases}$$

的解空间的一个标准正交基.

42. 把41题中解空间的标准正交基扩充为$\mathbf{R}^5$的标准正交基.

43. 已知

$$\boldsymbol{Q} = \begin{bmatrix} \frac{2}{7} & a & -\frac{3}{7} \\ d & b & c \\ e & -\frac{3}{7} & \frac{2}{7} \end{bmatrix}$$

为正交矩阵, 试确定a, b, c, d, e的值.

44. 已知$\boldsymbol{P} = \left[\frac{2}{3}\ \frac{1}{3}\ \frac{2}{3}\right]^{\mathrm{T}}$, 且矩阵$\boldsymbol{X}$满足$\boldsymbol{X} = 2\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{X} + \boldsymbol{E}$, 这里$\boldsymbol{E}$为元素全为1的3阶方阵. 求$\boldsymbol{X}$.

45. 设$\boldsymbol{A}, \boldsymbol{B}$是$n$阶正交矩阵, 证明: $\boldsymbol{A}^{\mathrm{T}}, \boldsymbol{AB}$也是正交矩阵.

46. 设$\boldsymbol{Q}$为n阶正交矩阵, $\boldsymbol{I} + \boldsymbol{Q}$可逆,这里$\boldsymbol{I}$是$n$阶单位矩阵, 证明:

(1) $(\boldsymbol{I} - \boldsymbol{Q})(\boldsymbol{I} + \boldsymbol{Q})^{-1} = (\boldsymbol{I} + \boldsymbol{Q})^{-1}(\boldsymbol{I} - \boldsymbol{Q})$;

(2) $(\boldsymbol{I} - \boldsymbol{Q})(\boldsymbol{I} + \boldsymbol{Q})^{-1}$是反称矩阵.

47. 设$\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$是欧氏空间$V$中的任意三个向量, 证明:

$$d(\boldsymbol{\alpha}, \boldsymbol{\beta}) \leqslant d(\boldsymbol{\alpha}, \boldsymbol{\gamma}) + d(\boldsymbol{\gamma}, \boldsymbol{\beta})$$

48. 求拟合下列4个点$(0, 1), (2, 0), (3, 1), (3, 2)$的最小二乘直线.

49. 某公司销售部在6月份进行中期统计, 得到当年前5个月的销售额分别为4.0万元, 4.4万元, 5.2万元, 6.4万元以及8.0万元. 统计员发现这些数据接近一个二次多项式函数. 试用最小二乘法求其表达式, 并预估12月份的销售额.

50. 判断下列映射哪些是线性映射:

(1) 映射$\boldsymbol{\sigma} : \mathbf{R}^2 \to \mathbf{R}$, $\boldsymbol{\sigma}(x_1, x_2) = x_1$;

(2) 映射$\boldsymbol{\sigma} : \mathbf{R}^2 \to \mathbf{R}$, $\boldsymbol{\sigma}(x_1, x_2) = x_1^2$;

(3) 映射$\boldsymbol{\sigma} : \mathbf{R}^2 \to \mathbf{R}^3$, $\boldsymbol{\sigma}(x_1, x_2) = (x_1 + x_2, x_1 - x_2, 2x_1 + 3x_2)$.

51. 对于下列线性映射$\boldsymbol{\sigma} : \mathbf{R}^3 \to \mathbf{R}^4$, 求$\boldsymbol{\sigma}(x_1, x_2, x_3)$:

(1) $\boldsymbol{\sigma}(1, 0, 0) = (1, 0, 0, 0), \boldsymbol{\sigma}(0, 1, 0) = (0, 0, 1, 0), \boldsymbol{\sigma}(0, 0, 1) = (0, 0, 0, 1)$;

(2) $\boldsymbol{\sigma}(1,0,0)=(1,1,1,1),\boldsymbol{\sigma}(0,1,0)=(4,2,1,1),\boldsymbol{\sigma}(0,0,1)=(3,1,0,0)$;

(3) $\boldsymbol{\sigma}(1,1,1)=(1,2,3,1),\boldsymbol{\sigma}(1,2,1)=(0,1,2,1),\boldsymbol{\sigma}(1,1,0)=(1,1,1,0)$.

52. 将实数域$\mathbf{R}$看作自身上的线性空间(加法和数乘为实数加法和乘法). 试写出$\mathbf{R}$上线性空间$C[a,b]$到$\mathbf{R}$的一个线性映射.

53. 已知F^4到F^3的一个映射$\boldsymbol{\sigma}(x_1,x_2,x_3,x_4)=(-x_1+x_2+2x_3+x_4,-2x_2+x_3,-x_1-x_2+3x_3+x_4)$, 其中$x_i\in F,i=1,2,3,4$.

(1) 证明$\boldsymbol{\sigma}$是一个线性映射;

(2) 取F^4的一个基: $\boldsymbol{\alpha}_1=(1,0,1,1),\boldsymbol{\alpha}_2=(0,1,0,1),\boldsymbol{\alpha}_3=(0,0,1,0),\boldsymbol{\alpha}_4=(0,0,2,1)$, 以及$F^3$的一个基: $\boldsymbol{\beta}_1=(1,1,1),\boldsymbol{\beta}_2=(1,0,-1),\boldsymbol{\beta}_3=(0,1,0)$, 求$\boldsymbol{\sigma}$在取定基下的矩阵表示.

54. 已知$F[x]_n$到$F[x]_{n+1}$的一个映射$\boldsymbol{\sigma}\left(a_0+a_1x+\cdots+a_{n-1}x^{n-1}\right)=a_0x+\frac{1}{2}a_1x^2+\frac{1}{3}a_2x^3+\cdots+\frac{1}{n}a_{n-1}x^n$, 其中$a_i\in F,i=1,2,\ldots,n-1$.

(1) 证明$\boldsymbol{\sigma}$是一个线性映射;

(2) 取$F[x]_n$的一个基$1,x,\ldots,x^{n-1}$, 以及$F[x]_{n+1}$的一个基$1,x,\ldots,x^n$, 求$\boldsymbol{\sigma}$在取定基下的矩阵表示.

55. 判断下列变换哪些是线性变换:

(1) 在线性空间V中, $\boldsymbol{\sigma}(\boldsymbol{\alpha})=\boldsymbol{\alpha}+\boldsymbol{\beta}$, 其中$\boldsymbol{\beta}\in V$是一个固定的向量;

(2) 在线性空间V中, $\boldsymbol{\sigma}(\boldsymbol{\alpha})=\boldsymbol{\beta}$, 其中$\boldsymbol{\beta}\in V$是一个固定的向量;

(3) 在线性空间$F[x]$中$\boldsymbol{\sigma}(f(x))=f'(x)$;

(4) 在线性空间$F^{n\times n}$中$\boldsymbol{\sigma}(\boldsymbol{X})=\boldsymbol{BXC}$, 其中$\boldsymbol{B},\boldsymbol{C}\in F^{n\times n}$是两个固定的方阵;

(5) 在由全体复数构成的复线性空间中, $\boldsymbol{\sigma}(\boldsymbol{\alpha})=\overline{\boldsymbol{\alpha}}$;

(6) 在$\mathbf{R}^3$中, $\boldsymbol{\sigma}\left(x_1,x_2,x_3\right)=\left(x_1^2,x_2+x_3,x_3^2\right)$;

(7) 在$\mathbf{R}^3$中, $\boldsymbol{\sigma}\left(x_1,x_2,x_3\right)=\left(2x_1-x_2,x_2+x_3,x_1\right)$.

56. 线性变换是否把零向量变为零向量? 能否把非零向量也变为零向量? 一个非零向量的所有原像能否构成一个子空间?

57. 证明: 若某向量组在线性变换下的像是线性无关的, 则该向量组是线性无关的.

58. 在多项式空间$F[x]_n$中, 令$\boldsymbol{\sigma}(f(x))=f'(x)$, 证明: $\boldsymbol{\sigma}\left(F[x]_n\right)=F[x]_{n-1},\boldsymbol{\sigma}^{-1}(0)=\{f(x)\mid f(x)\in F[x]_n\text{且}\boldsymbol{\sigma}(f(x))=0\}=F$.

59. 在$\mathbf{R}^3$中取自然基$\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$, 试求将向量投影到$Oxy$平面上的线性变换$\boldsymbol{\sigma}$的矩阵表示.

60. 求上述线性变换$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1=\boldsymbol{i},\boldsymbol{\alpha}_2=\boldsymbol{j},\boldsymbol{\alpha}_3=\boldsymbol{i}+\boldsymbol{j}+\boldsymbol{k}$下的矩阵表示.

61. 在$\mathbf{R}^3$中, $\boldsymbol{\sigma}$定义如下:

$$\begin{aligned} \boldsymbol{\sigma}(\boldsymbol{\alpha}_1) &= (-5,0,3), & \boldsymbol{\alpha}_1 &= (-1,0,2) \\ \boldsymbol{\sigma}(\boldsymbol{\alpha}_2) &= (0,-1,6), & \boldsymbol{\alpha}_2 &= (0,1,1) \\ \boldsymbol{\sigma}(\boldsymbol{\alpha}_3) &= (-5,-1,9), & \boldsymbol{\alpha}_3 &= (3,-1,0) \end{aligned}$$

求$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$下的矩阵表示.

62. 证明: 把n维线性空间的每一个向量都变为零向量的变换是线性变换, 而且它在任何基下的矩阵表示是零矩阵.

63. 设3维实线性空间的线性变换$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$下的矩阵表示是

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 2 \\ 2 & -1 & 4 \end{bmatrix}$$

求在基$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1,\boldsymbol{\beta}_2=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\beta}_3=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$下$\boldsymbol{\sigma}$的矩阵表示.

64. 设4维实线性空间的线性变换$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$下的矩阵表示是

$$\begin{bmatrix} 1 & 0 & 2 & -3 \\ 2 & -1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ -2 & 3 & 1 & 4 \end{bmatrix}$$

求在基$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4,\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4,\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4,\boldsymbol{\beta}_4=\boldsymbol{\alpha}_4$下$\boldsymbol{\sigma}$的矩阵表示.

65. 给定$\mathbf{R}^3$的两个基:

$$\begin{cases} \boldsymbol{\alpha}_1=(1,0,1) \\ \boldsymbol{\alpha}_2=(2,1,0) \\ \boldsymbol{\alpha}_3=(1,1,1) \end{cases} \quad \begin{cases} \boldsymbol{\beta}_1=(1,2,-1) \\ \boldsymbol{\beta}_2=(2,2,-1) \\ \boldsymbol{\beta}_3=(2,-1,-1) \end{cases}$$

定义线性变换: $\boldsymbol{\sigma}(\boldsymbol{\alpha}_i)=\boldsymbol{\beta}_i\ (i=1,2,3)$.

(1) 写出由基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$到基$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$的过渡矩阵;

(2) 写出$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$下的矩阵表示;

(3) 写出$\boldsymbol{\sigma}$在基$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$下的矩阵表示.

66. 在$F[x]_4$中, $\boldsymbol{\sigma}[f(x)]=f'(x)$, 求$\boldsymbol{\sigma}$在基$1,1+x,1+x+x^2,1+x+x^2+x^3$下的矩阵表示.

67. 在$F[x]_n$中, $\boldsymbol{\sigma}[f(x)]=f'(x)$, 求$\boldsymbol{\sigma}$在基$1,x,x^2,\ldots,x^{n-1}$下的矩阵表示.

68. 在$F[x]_n$中, $\boldsymbol{\sigma}[f(x)] = f'(x)$, 求$\boldsymbol{\sigma}$在基$1, x-c, (x-c)^2/2!, \ldots,$ $(x-c)^{n-1}/(n-1)!$下的矩阵表示.

69. 设V是函数

$$\begin{aligned}&\boldsymbol{\alpha}_1 = \mathrm{e}^{ax}\cos bx,\ \ \boldsymbol{\alpha}_2 = \mathrm{e}^{ax}\sin bx,\ \ \boldsymbol{\alpha}_3 = x\mathrm{e}^{ax}\cos bx\\ &\boldsymbol{\alpha}_4 = x\mathrm{e}^{ax}\sin bx, \boldsymbol{\alpha}_5 = (x^2/2)\,\mathrm{e}^{ax}\cos bx, \boldsymbol{\alpha}_6 = (x^2/2)\,\mathrm{e}^{ax}\sin bx\end{aligned}$$

的所有实系数线性组合构成的实数域上的一个6维线性空间. 求微分变换$\mathcal{D}$在基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_6$下的矩阵表示.

70. 已知数域F上3维线性空间V上的线性变换$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$下的矩阵表示为

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

(1) 求$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_3, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1$下的矩阵表示;

(2) 求$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1, k\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$下的矩阵表示, 其中$k \in F$且$k \neq 0$;

(3) 求$\boldsymbol{\sigma}$在基$\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$下的矩阵表示.