高等工程數學 (11)

厦門大學
XIAMEN UNIVERSITY
信息学院
（国家示范性软件学院）
School of Informatics
黄 炜
博士，副教授
Dr. Wei Huang

线性空间和线性变换

矩阵论（11）

厦門大學
XIAMEN UNIVERSITY
信息学院
（国家示范性软件学院）
School of Informatics
黄 炜
博士·副教授
Dr. Wei Huang

6 线性空间和线性变换

线性空间和线性变换是重要的数学概念.

线性空间是由具体的几何平面和几何空间的特征经抽象后提炼出来的数学概念.它是前面所说的向量空间概念的拓广.粗略地说,在一个非空集合上定义了线性运算(加法和数乘),并且这种运算满足一定的规则,那么这个非空集就成为一个线性空间.因此,线性空间必有线性运算所确定的代数结构,从而可用数学方法对它进行研究.

线性空间之间一种最简单同时也是最重要的联系是线性变换,也叫做线性映射或线性算子,特别是线性空间到自身的线性变换.通过线性变换的研究可以进一步刻划线性空间的特征和结构,而有限维线性空间之间的线性变换又是借助于矩阵来表示并描述它们的各种性质的.这也是矩阵论之所以在工程技术和社会学科中愈来愈重要的原因.

本章简要地论述线性空间和线性变换的一些基本概念及其重要的性质,讨论矩阵与线性变换之间的关系.

6.1 线性空间

6.1.1 线性空间的定义及例子

由于线性空间的定义涉及数域，所以先叙述数域的含义. 设 F 是包含 0 和 1 的一个数集，如果 F 中任意两个数(它们可以相同)的和、差、积、商(除数不为零)仍是 F 中的数，那么称 F 为**数域**. 例如，全体实数集 $\mathbf{R}$，全体复数集 $\mathbf{C}$，全体有理数集 $\mathbf{Q}$ 等都是数域；但全体正实数集 $\mathbf{R}_+$，全体整数集 $\mathbf{Z}$ 等都不是数域.

定义 6.1-1 设 V 是一非空集，F 是数域. 对于 V 中任意两个元素 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$，定义一个叫做加法的运算，记为“+”，V 中有一个元素 $\boldsymbol{\alpha}+\boldsymbol{\beta}$ 与之对应，称做 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的和，且满足下列规则：

(1) 加法交换律 $\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{\beta}+\boldsymbol{\alpha}$；

(2) 加法结合律 $(\boldsymbol{\alpha}+\boldsymbol{\beta})+\boldsymbol{\gamma}=\boldsymbol{\alpha}+(\boldsymbol{\beta}+\boldsymbol{\gamma})$；

(3) 存在 $\mathbf{0}\in V$，使得对任意 $\boldsymbol{\alpha}\in V$，有

$$\boldsymbol{\alpha}+\mathbf{0}=\boldsymbol{\alpha},$$

这个元素 $\mathbf{0}$ 称为 V 的**零元素**；

(4) 对任意 $\boldsymbol{\alpha}\in V$，存在 $-\boldsymbol{\alpha}\in V$，使得

$$\boldsymbol{\alpha}+(-\boldsymbol{\alpha})=\mathbf{0},$$

称 $-\boldsymbol{\alpha}$ 为 $\boldsymbol{\alpha}$ 的**负元素**.

又在 F 与 V 的元素之间定义一个叫做数乘的运算，对于 F 中任一数 k 与 V 中任一元素 $\boldsymbol{\alpha}$，V 中都有一个元素 $k\boldsymbol{\alpha}$ 与之对应，称它为 k 与 $\boldsymbol{\alpha}$ 的数乘，且满足下列规则：

6.1 线性空间

(5) 对任意 $k\in F$ 和任意 $\boldsymbol{\alpha},\boldsymbol{\beta}\in V$,有

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

(6) 对任意 $\boldsymbol{\alpha}\in V$ 和任意的 $k,l\in F$,有

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

(7) 对任意 $\boldsymbol{\alpha}\in V$ 和任意的 $k,l\in F$,有

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

(8) F 中的数 1,使得对任意 $\boldsymbol{\alpha}\in V$,有

$$1\boldsymbol{\alpha}=\boldsymbol{\alpha}.$$

那么称 V 为数域 F 上的**线性空间**(也称为**向量空间**),记为 $V(F)$. V 中的元素也称为**向量**.

不管 V 中的元素具体是什么,当 F 为实数域 $\mathbf{R}$ 时,称它为**实线性空间**;而当 F 为复数域 $\mathbf{C}$ 时,称它为**复线性空间**.

今后,在不需强调数域时,就以 V 表示线性空间.

例如,前面所讲的向量空间 $\mathbf{R}^n$、$\mathbf{C}^n$ 分别是实线性空间和复线性空间. 它们实际上是由通常的向量加法和数乘向量运算所构成的线性空间.

为了掌握线性空间这一重要概念,再举一些例子说明.

例 1 由 F 中的数形成的 $m\times n$ 矩阵的全体,对于通常定义的矩阵加法和数乘矩阵,构成 F 上的线性空间,记之为 $F^{m\times n}$. ∎

6.1 线性空间

例 2 区间$[a,b]$上的全体连续函数，对于通常定义的函数加法和数乘函数，即

$$(f+g)(x)=f(x)+g(x),\quad (kf)(x)=kf(x),\quad \forall x\in[a,b],$$

构成一个线性空间，记为$C[a,b]$. 而$C^1[a,b]$表示由区间$[a,b]$上连续可微函数全体在通常的函数加法和数乘函数运算下所构成的线性空间. ∎

回顾一下多项式的概念. 设$a_i(i=0,1,2,\cdots,m)$是F中的数，λ为一变量，则形如

$$g(\lambda)=a_0\lambda^m+a_1\lambda^{m-1}+\cdots+a_{m-1}\lambda+a_m$$

式子称为数域F上的一个多项式. 当$a_0\neq 0$时，$g(\lambda)$称为m次多项式，$a_0\lambda^m$称为多项式$g(\lambda)$的首项. 特别地，当$a_0=1$时，称$g(\lambda)$为m次首一多项式. 系数$a_0,a_1,a_2,\cdots,a_m$全为零的多项式，称为零多项式，在不会产生混淆的情况下，以O表示零多项式；零多项式是唯一不定义次数的多项式，它与零次多项式有本质的区别.

例 3 实数域$\mathbf{R}$上的多项式全体，按通常的多项式加法及数与多项式乘法，构成实线性空间，记为$P[\lambda]$. 如果只考虑次数不大于n的多项式全体，则要添加零多项式才在通常的多项式加法及数与多项式乘法下构成一个实线性空间，记它为$P_n[\lambda]$，这是因为零多项式是这种线性空间的零元素. ∎

例 4 数域F按其自身的加法和乘法也构成F上的线性空间.

下面举一个不是常规加法和数乘的例子. ∎

6.1 线性空间

例 5 在正实数集 $\mathbf{R}_+$ 中，定义加法和数乘运算为

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

其中 $a,b \in \mathbf{R}_+$，$k \in \mathbf{R}$. 这里，为了区别于通常的加法和数乘，用"$\oplus$"表示加法（实际是通常的乘法），"$\circ$"表示数乘（实际是乘方），那么 $\mathbf{R}_+$ 是实线性空间.

事实上，不难验证 $\mathbf{R}_+$ 对这两种运算是封闭的，即 $a \oplus b \in \mathbf{R}_+$，$k \circ a \in \mathbf{R}_+$，并且满足规则(1)、(2)及(5)～(8). 又因为 $1 \in \mathbf{R}_+$，且对任意 $a \in \mathbf{R}_+$，有 $a \oplus 1 = a$，同时又有 $a \oplus \frac{1}{a} = a \cdot \frac{1}{a} = 1$，且 $\frac{1}{a} \in \mathbf{R}_+$，所以规则(3)和(4)也是满足的，并且 $\mathbf{R}_+$ 中的零元素是 1，a 的负元素是 $\frac{1}{a}$. ∎

容易验证，线性空间 V 有下列简单性质：

1) 零元素是唯一的.

2) 对任意 $\boldsymbol{\alpha} \in V$，它的负元素唯一，从而可定义 V 中两个元素 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 的减法（记为"－"）为

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

3) 对任意 $\boldsymbol{\alpha} \in V$，有 $0\boldsymbol{\alpha} = \mathbf{0}$， $(-1)\boldsymbol{\alpha} = -\boldsymbol{\alpha}$，

而对任意 $k \in F$，有 $k\mathbf{0} = \mathbf{0}$.

由线性空间的定义可知，若 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m \in V(F)$，$k_1, \cdots, k_m \in F$，则 $k_1\boldsymbol{\alpha}_1 + \cdots + k_m\boldsymbol{\alpha}_m$ 是 $V(F)$ 中的元素，记它为 $\boldsymbol{\beta}$，即

$$\boldsymbol{\beta} = k_1\boldsymbol{\alpha}_1 + \cdots + k_m\boldsymbol{\alpha}_m = \sum_{i=1}^{m} k_i\boldsymbol{\alpha}_i,$$

称 $\boldsymbol{\beta}$ 为 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 的一个线性组合，或称 $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性表出.

6.1 线性空间

6.1.2 基与维数

定义 6.1-2 设 $V(F)$ 是线性空间，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 是 $V(F)$ 中一组向量. 如果在 F 中存在一组不全为零的数 $k_1,k_2,\cdots,k_m$，使

$$\sum_{i=1}^{m} k_i\boldsymbol{\alpha}_i = \mathbf{0}, \tag{6.1-1}$$

则称该向量组线性相关. 否则称为线性无关.

根据这个定义，容易推出下列结论：

1) 当 $m\geqslant 2$ 时，向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关的充要条件是，其中至少有一个向量可由该向量组中其余向量线性表出.

2) 若某个向量组线性无关，则它的任一子向量组必线性无关；若某个向量组中有一个子向量组线性相关，则该向量组必线性相关.

3) 单个零向量组成的向量组是线性相关的，但单个非零向量组成的向量组是线性无关的.

定义 6.1-3 如果在线性空间 V 中可以找到无穷多个线性无关的向量，则称 V 为无限维线性空间. 若在 V 中只有有限多个线性无关的向量，则称 V 为有限维线性空间，并且把最大线性无关向量的个数称为 V 的维数，记为 $\dim V$. 维数为 n 的线性空间 V 称为 n 维线性空间，记为 V^n.

按此定义，例 1 所述的线性空间 $F^{m\times n}$ 是 mn 维的，因为 $F^{m\times n}$ 中任一矩阵 $\boldsymbol{A}=[a_{ij}]$ 可表示为

6.1 线性空间

$$A=[a_{ij}]=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}E_{ij},$$

其中 E_{ij} 表示第 i 行第 j 列处的元素为 1，其余元素均为 0 的 $m\times n$ 矩阵，并且 $\{E_{ij}\mid i=1,2,\cdots,m,j=1,2,\cdots,n\}$ 显然线性无关. 例 3 中的线性空间 $P[\lambda]$ 是无限维的，因为对于任意正整数 N，都有 N 个线性无关的元素 $1,\lambda,\cdots,\lambda^{N-1}$；而线性空间 $P_n[\lambda]$ 是 $n+1$ 维的，因为任一次数不大于 n 的多项式及零多项式，都可由 $n+1$ 个线性无关的多项式 $1,\lambda,\lambda^2,\cdots,\lambda^n$ 线性表出.

矩阵论主要研究有限维线性空间的问题，所以今后就 n 维线性空间 V^n 进行讨论.

定义 6.1-4 线性空间 V^n 中给定顺序的 n 个线性无关向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 组成的向量组称为 V^n 的一个**基**，记为 $\mathscr{B}=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$. $\mathscr{B}$ 中的向量 $\boldsymbol{\alpha}_i(i=1,2,\cdots,n)$ 称为第 i 个**基向量**.

定理 6.1-1 设 $\mathscr{B}$ 是线性空间 V^n 的一个基，则 V 中任一向量 ξ 都可由 $\mathscr{B}$ 唯一地线性表出.

证 由于 V^n 中 $n+1$ 个向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n,\xi$ 必线性相关，故存在不全为零的 $n+1$ 个数 $k_1,k_2,\cdots,k_n,k_{n+1}$，使得

$$\sum_{i=1}^{n}k_i\boldsymbol{\alpha}_i+k_{n+1}\xi=\mathbf{0}.$$

6.1 线性空间

如果 $k_{n+1}=0$，则上式成为 $\sum_{i=1}^{n}k_i\boldsymbol{\alpha}_i=\boldsymbol{0}$. 但$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 是基，故有 $k_i=0(i=1,2,\cdots,n)$. 这与 $k_1,k_2,\cdots,k_n,k_{n+1}$ 不全为零矛盾. 因此 $k_{n+1}\neq 0$，从而有

$$\boldsymbol{\xi}=-\frac{1}{k_{n+1}}\sum_{i=1}^{n}k_i\boldsymbol{\alpha}_i=\sum_{i=1}^{n}\left(-\frac{k_i}{k_{n+1}}\right)\boldsymbol{\alpha}_i,$$

即 $\boldsymbol{\xi}$ 可由 $\mathscr{B}$ 线性表出.

再证唯一性. 设有 $\quad \boldsymbol{\xi}=\sum_{i=1}^{n}x_i\boldsymbol{\alpha}_i \quad$ 和 $\quad \boldsymbol{\xi}=\sum_{i=1}^{n}y_i\boldsymbol{\alpha}_i,$

则

$$\boldsymbol{0}=\sum_{i=1}^{n}x_i\boldsymbol{\alpha}_i-\sum_{i=1}^{n}y_i\boldsymbol{\alpha}_i=\sum_{i=1}^{n}(x_i-y_i)\boldsymbol{\alpha}_i.$$

由$\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n\}$是基，知 $\quad x_i-y_i=0$，即 $x_i=y_i,\quad i=1,\cdots,n.$ ∎

这个定理表明，在 V^n 中取定一个基 $\mathscr{B}$，那么对 V^n 中任一向量 $\boldsymbol{\xi}$，存在唯一的 n 个数 $x_1,\cdots,x_n$，使

$$\boldsymbol{\xi}=\sum_{i=1}^{n}x_i\boldsymbol{\alpha}_i.$$

将上式记为

$$\boldsymbol{\xi}=\{\boldsymbol{\alpha}_1\ \cdots\ \boldsymbol{\alpha}_n\}[x_1,x_2,\cdots,x_n]^{\mathrm{T}}=\mathscr{B}\boldsymbol{x}, \tag{6.1-2}$$

其中 $\boldsymbol{x}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}$，$\boldsymbol{x}$ 称为 $\boldsymbol{\xi}$ 在基 $\mathscr{B}$ 下的**坐标或坐标向量**.

6.1 线性空间

例 6 在 $P_2[\lambda]$中取基 $\mathscr{B}=\{1\ \ \lambda\ \ \lambda^2\}$,则多项式 $g(\lambda)=2\lambda^2-\lambda+1$ 在 $\mathscr{B}$ 下的坐标向量 $\boldsymbol{x}=[1,-1,2]^{\mathrm{T}}$,因为

$$2\lambda^2-\lambda+1=1\cdot 1+(-1)\cdot\lambda+2\cdot\lambda^2=\{1\ \lambda\ \lambda^2\}[1,-1,2]^{\mathrm{T}}.$$

所给出的多项式 $g(\lambda)=2\lambda^2-\lambda+1$ 在另一个基 $\mathscr{B}_1=\{\lambda+1\ \ \lambda+2\ \ \lambda^2\}$下的坐标向量则是 $\boldsymbol{y}=[-3,2,2]^{\mathrm{T}}$. ∎

例 7 将 C 看作复线性空间,则它是 1 维的. 但若将 C 看作实线性空间,则$\{1,\mathrm{i}\}$是它的一个基,其中 i 是虚单位,从而是 2 维的. ∎

这个例子表明,线性空间的维数与所考虑的数域有关.

从上述讨论看出,在线性空间 V^n 中取定一个基 $\mathscr{B}$,那么 V^n 中的元素与 F^n 中的向量之间是一一对应的,并且若 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 在 $\mathscr{B}$ 下的坐标向量分别为 $\boldsymbol{x}$、$\boldsymbol{y}$,则 $\boldsymbol{\alpha}+\boldsymbol{\beta}$ 在 $\mathscr{B}$ 下的坐标向量为 $\boldsymbol{x}+\boldsymbol{y}$,$k\boldsymbol{\alpha}$ 在 $\mathscr{B}$ 下的坐标向量为 $k\boldsymbol{x}$. 因此,V^n 与 F^n 有相同的代数结构. 但向量空间 F^n 比线性空间 V^n 具体,便于应用矩阵运算,所以一般总是把 V^n 上的问题通过取定一个基转化为 F^n 上的问题来讨论.

6.1 线性空间

6.1.3 基变换与坐标变换

从例 6 看出，V^n 中一个元素在不同基下的坐标向量一般是不同的，那么它们之间有何联系呢？为此要讨论 V^n 的两个基之间的变换关系.

定义 6.1-5 设 $\mathscr{B}_\alpha=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$，$\mathscr{B}_\beta=\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$ 是 V^n 的两个基，则每个 $\boldsymbol{\beta}_j(j=1,2,\cdots,n)$ 都可由 $\mathscr{B}_\alpha$ 唯一地线性表出：

$$\boldsymbol{\beta}_j=\sum_{i=1}^{n}p_{ij}\boldsymbol{\alpha}_i=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}[p_{1j},p_{2j},\cdots,p_{nj}]^{\mathrm{T}}.$$

将 $\boldsymbol{\beta}_j(j=1,2,\cdots,n)$ 按顺序排列，并使用矩阵记号，则得

$$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}\begin{bmatrix}p_{11}&\cdots&p_{1n}\\ \vdots&&\vdots\\ p_{n1}&\cdots&p_{nn}\end{bmatrix},\tag{6.1-3}$$

简记为

$$\mathscr{B}_\beta=\mathscr{B}_\alpha\boldsymbol{P},\tag{6.1-3'}$$

其中 n 阶方阵 $\boldsymbol{P}=[p_{ij}]$ 称为基 $\mathscr{B}_\alpha$ 到 $\mathscr{B}_\beta$ 的**变换矩阵**（也称为**过渡矩阵**）.

显然，基变换矩阵 $\boldsymbol{P}$ 中的第 j 个列向量 $\boldsymbol{P}_j=[p_{1j},p_{2j},\cdots,p_{nj}]^{\mathrm{T}}$ 是基 $\mathscr{B}_\beta$ 中第 j 个基向量 $\boldsymbol{\beta}_j$ 在基 $\mathscr{B}_\alpha$ 下的坐标向量.

基变换矩阵 $\boldsymbol{P}$ 是可逆矩阵. 事实上，若 $\boldsymbol{P}$ 不可逆，则存在非零向量 $\boldsymbol{b}$ 使 $\boldsymbol{Pb}=\boldsymbol{0}$，则由 (6.1-3′) 式得

$$\mathscr{B}_\beta\boldsymbol{b}=\mathscr{B}_\alpha\boldsymbol{Pb}=\boldsymbol{0},$$

6.1 线性空间

从而由零元素在基下的坐标向量为零向量得 $\boldsymbol{b}=\boldsymbol{0}$，这与 $\boldsymbol{b}\neq\boldsymbol{0}$ 矛盾.

由于 $\boldsymbol{P}$ 可逆，故由(6.1-3′)式得

$$\mathscr{B}_\alpha = \mathscr{B}_\beta \boldsymbol{P}^{-1},$$

也就是说，$\boldsymbol{P}^{-1}$ 是基 $\mathscr{B}_\beta$ 到 $\mathscr{B}_\alpha$ 的变换矩阵.

例 8 已知 $\boldsymbol{P}_2[\lambda]$的两个基是

$$\mathscr{B}_1 = \{1, \lambda-1, (\lambda-1)^2\}, \quad \mathscr{B}_2 = \{1+\lambda, 2-\lambda, \lambda^2-\lambda+1\},$$

求基 $\mathscr{B}_1$ 到 $\mathscr{B}_2$ 的变换矩阵 $\boldsymbol{P}$.

解 由于

$$1+\lambda = \{1 \quad \lambda-1 \quad (\lambda-1)^2\}[2,1,0]^{\mathrm{T}},$$

$$2-\lambda = \{1 \quad \lambda-1 \quad (\lambda-1)^2\}[1,-1,0]^{\mathrm{T}},$$

$$\lambda^2-\lambda+1 = \{1 \quad \lambda-1 \quad (\lambda-1)^2\}[1,1,1]^{\mathrm{T}},$$

所以
$$\{1+\lambda \quad 2-\lambda \quad \lambda^2-\lambda+1\} = \{1 \quad \lambda-1 \quad (\lambda-1)^2\}\begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

故
$$\boldsymbol{P} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

∎

6.1 线性空间

例 9 已知 $\mathbf{R}^3$ 的两个基

$$\mathscr{B}_1=\left\{\begin{bmatrix}-1\\1\\1\end{bmatrix},\begin{bmatrix}1\\0\\-1\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix}\right\},\quad \mathscr{B}_2=\left\{\begin{bmatrix}1\\0\\1\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix},\begin{bmatrix}1\\1\\1\end{bmatrix}\right\},$$

求 $\mathscr{B}_1$ 到 $\mathscr{B}_2$ 的变换矩阵 $\boldsymbol{P}$.

解 由(6.1-3)式给出

$$\begin{bmatrix}1&0&1\\0&0&1\\1&1&1\end{bmatrix}=\begin{bmatrix}-1&1&0\\1&0&1\\1&-1&1\end{bmatrix}\boldsymbol{P},$$

故

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

现设 V^n 中的元素 ξ 在基 $\mathscr{B}_\alpha$、$\mathscr{B}_\beta$ 下的坐标向量分别为 $\boldsymbol{x}$、$\boldsymbol{y}$，即有

$$\xi=\mathscr{B}_\alpha\boldsymbol{x},\quad \xi=\mathscr{B}_\beta\boldsymbol{y}.$$

将(6.1-3′)式代入上述第二式，并与第一式比较，得

$$\xi=\mathscr{B}_\alpha\boldsymbol{x}=\mathscr{B}_\alpha\boldsymbol{P}\boldsymbol{y},\quad 即\ \mathscr{B}_\alpha(\boldsymbol{x}-\boldsymbol{P}\boldsymbol{y})=\mathbf{0},$$

从而 $\boldsymbol{x}-\boldsymbol{P}\boldsymbol{y}=\mathbf{0}$，即

$$\boldsymbol{y}=\boldsymbol{P}^{-1}\boldsymbol{x}. \tag{6.1-4}$$

(6.1-4)式表示同一个 ξ 在不同基下坐标向量之间的关系，称为**坐标变换公式**.

6.1 线性空间

例 10 已知 $\mathbf{R}^3$ 的两个基

$$\mathscr{B}_1=\{[1,0,0]^{\mathrm{T}},[0,1,0]^{\mathrm{T}},[0,0,1]^{\mathrm{T}}\},$$

$$\mathscr{B}_2=\{[-3,-7,1]^{\mathrm{T}},[3,6,-1]^{\mathrm{T}},[-2,-3,2]^{\mathrm{T}}\},$$

求在这两个基下有相同坐标的所有向量.

解 设 ξ 是所要求的向量,即 ξ 在基 $\mathscr{B}_1$ 下的坐标向量 $\boldsymbol{x}=[x_1,x_2,x_3]^{\mathrm{T}}$ 也是它在 $\mathscr{B}_2$ 下的坐标向量,于是有

$$\xi=\mathscr{B}_1\boldsymbol{x}=\mathscr{B}_2\boldsymbol{x}, \tag{6.1-5}$$

故 $\boldsymbol{x}$ 应满足

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}-3&3&-2\\-7&6&-3\\1&-1&2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix},$$

即

$$\begin{bmatrix}4&-3&2\\7&-5&3\\-1&1&-1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

解之得 $\boldsymbol{x}=k[1,2,1]^{\mathrm{T}}$,$k$ 为任意实数.将 $\boldsymbol{x}$ 代入(6.1-5)式,得到所要求的向量

$$\xi=\mathscr{B}_1\boldsymbol{x}=k\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1\\2\\1\end{bmatrix}=k\begin{bmatrix}1\\2\\1\end{bmatrix},$$

其中 k 为任意实数. ∎

6.1 线性空间

6.1.4 子空间和维数定理

定义 6.1-6 设 W 是线性空间 V 的一个非空子集，如果这个集合 W 关于 V 中所定义的加法和数乘运算也构成线性空间，则称 W 为 V 的**子空间**.

验证 W 是否为 V 的子空间，实际上只需要考察 W 对 V 中所定义的加法和数乘是否封闭就行了，因为线性空间定义中的规则(1)～(8)在 W 对加法和数乘是封闭的情况下必是满足的.

由于子空间不可能比整个空间有更多数目的线性无关向量，所以子空间的维数不会大于整个空间的维数.

按定义，线性空间 V 自身及由 V 的零元素所构成的集合都是 V 的子空间，称它们为 V 的**平凡子空间**，并称单个零元素构成的空间为**零空间**，记为 $\{\mathbf{0}\}$，且由维数定义知，$\dim\{\mathbf{0}\}=0$.

人们感兴趣的当然是非平凡的子空间.

例如，设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是线性空间 $V(F)$ 的 r 个元素，由它们所有可能的线性组合所组成的集合

$$\{\boldsymbol{\alpha}\in V \mid \boldsymbol{\alpha}=\sum_{i=1}^{r}k_i\boldsymbol{\alpha}_i,\quad k_i\in F,\quad (i=1,2,\cdots,r)\}$$

是 $V(F)$ 的一个子空间，称为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ **张成的子空间**，记为 $\mathrm{Span}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r)$.

6.1 线性空间

容易看出，若 W 是 n 维线性空间 V^n 的一个 r 维子空间，且 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\}$ 是 W 的一个基，那么在 V^n 中必可找到 $n-r$ 个元素 $\boldsymbol{\alpha}_{r+1},\cdots,\boldsymbol{\alpha}_n$，使得向量组 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1},\cdots,\boldsymbol{\alpha}_n\}$ 是 V^n 的一个基.

定义 6.1-7 设 W_1、W_2 是线性空间 V 的两个子空间，则

$$W_1 \cap W_2 = \{\xi \in V \mid \xi \in W_1 \text{ 且 } \xi \in W_2\},$$

$$W_1 + W_2 = \{\xi \in V \mid \xi = \xi_1 + \xi_2, \xi_1 \in W_1, \xi_2 \in W_2\},$$

分别称为 W_1 与 W_2 的**交**与**和**.

读者可以验证，$W_1 \cap W_2$，W_1+W_2 都是线性空间 V 的子空间，分别称为 W_1 与 W_2 的**交空间**与**和空间**.

定理 6.1-2 设 W_1、W_2 是线性空间 V^n 的两个子空间，则

$$\dim(W_1+W_2)+\dim(W_1\cap W_2)=\dim W_1+\dim W_2. \tag{6.1-6}$$

证 设 $\dim W_1=n_1$，$\dim W_2=n_2$，$\dim(W_1\cap W_2)=r$，且 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r\}$ 是 $W_1\cap W_2$ 的一个基. 由于 $W_1\cap W_2$ 是 $W_i(i=1,2)$ 的子空间，所以 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r\}$ 可扩充为 W_1 的基：

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_{n_1-r}\},$$

又可扩充为 W_2 的基

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_{n_2-r}\}.$$

下证向量组

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_{n_1-r},\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_{n_2-r}\} \tag{6.1-7}$$

是 W_1+W_2 的一个基，从而(6.1-6)式成立.

6.1 线性空间

显然，W_1+W_2 中的任一元素都可用向量组(6.1-7)线性表出．现设有等式

$$\sum_{i=1}^{r} k_i\boldsymbol{\alpha}_i+\sum_{j=1}^{n_1-r} p_j\boldsymbol{\beta}_j+\sum_{l=1}^{n_2-r} q_l\boldsymbol{\gamma}_l=\mathbf{0}, \tag{6.1-8}$$

记

$$\boldsymbol{\xi}=\sum_{i=1}^{r} k_i\boldsymbol{\alpha}_i+\sum_{j=1}^{n_1-r} p_j\boldsymbol{\beta}_j, \tag{6.1-9}$$

则由 (6.1-8) 式知 $\boldsymbol{\xi}=-\sum_{l=1}^{n_2-r} q_l\boldsymbol{\gamma}_l$，从而 $\boldsymbol{\xi}\in W_2$．而由(6.1-9)式知 $\boldsymbol{\xi}\in W_1$，因此 $\boldsymbol{\xi}\in W_1\cap W_2$，故 $\boldsymbol{\xi}$ 可由$\{\alpha_1,\alpha_2,\cdots,\alpha_r\}$ 线性表出．设 $\boldsymbol{\xi}=\sum_{i=1}^{r} b_i\boldsymbol{\alpha}_i$，

则由 $\boldsymbol{\xi}=-\sum_{l=1}^{n_2-r} q_l\boldsymbol{\gamma}_l=\sum_{i=1}^{r} b_i\boldsymbol{\alpha}_i$ 得 $\sum_{i=1}^{r} b_i\boldsymbol{\alpha}_i+\sum_{l=1}^{n_2-r} q_l\boldsymbol{\gamma}_l=\mathbf{0}$.

但 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_{n_2-r}$ 线性无关，故

$$b_i=0,\quad i=1,2,\cdots,r,\quad q_l=0,\quad l=1,2,\cdots,n_2-r.$$

于是 $\boldsymbol{\xi}=\mathbf{0}$，从而由(6.1-9)式得 $\sum_{i=1}^{r} k_i\boldsymbol{\alpha}_i+\sum_{j=1}^{n_1-r} p_j\boldsymbol{\beta}_j=\mathbf{0}$.

又因 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{n_1-r}$ 线性无关，故

$$k_i=0, i=1,2,\cdots,r, p_j=0, j=1,2,\cdots,n_1-r.$$

这就证明了向量组(6.1-7)线性无关．综合起来便证明了(6.1-7)是和空间 W_1+W_2 的一个基． ■

6.1 线性空间

定理 6.1-2 称为**维数定理**.

和空间 W_1+W_2 的定义仅表明，其任一向量 ξ 可表示为 $\xi_1+\xi_2$，且 $\xi_1\in W_1,\xi_2\in W_2$. 但这种表示式未必唯一. 针对这一情况，提出子空间直和的概念，使分解式唯一.

定义 6.1-8 若 W_1+W_2 中的任一向量只能唯一地分解为 W_1 的一个向量与 W_2 的一个向量之和，则称 W_1+W_2 为 W_1 与 W_2 的**直和**，记作 $W_1\oplus W_2$.

定理 6.1-3 设 W_1 与 W_2 是线性空间 V^n 的两个子空间，则 $W_1+W_2=W_1\oplus W_2$ 的充要条件是下列条件之一成立：

(1) $W_1\cap W_2=\{\mathbf{0}\}$;

(2) 若 $\xi_1+\xi_2=\mathbf{0}$，且 $\xi_i\in W_i(i=1,2)$，则 $\xi_1=\xi_2=\mathbf{0}$;

(3) $\dim(W_1+W_2)=\dim W_1+\dim W_2$.

证 只证(1)，(2)和(3)可由(1)推出.

必要性：若 $W_1\cap W_2\neq\{\mathbf{0}\}$，则存在非零向量 $\boldsymbol{\alpha}\in W_1\cap W_2$，从而 $-\boldsymbol{\alpha}\in W_1\cap W_2$. 因此 W_1+W_2 的零向量既有分解式 $\mathbf{0}=\mathbf{0}+\mathbf{0}$，又有分解式 $\mathbf{0}=\boldsymbol{\alpha}+(-\boldsymbol{\alpha})$，且 $\boldsymbol{\alpha}\in W_1,-\boldsymbol{\alpha}\in W_2$. 这与 $W_1+W_2=W_1\oplus W_2$ 矛盾.

充分性：若 W_1+W_2 不是直和，则 $\xi\in W_1+W_2$ 的分解式不唯一，即存在 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\in W_1$ 和 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\in W_2$，且 $\boldsymbol{\alpha}_1\neq\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1\neq\boldsymbol{\beta}_2$，使 $\xi=\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1=\boldsymbol{\alpha}_2+\boldsymbol{\beta}_2$.
因此 $\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2=\boldsymbol{\beta}_2-\boldsymbol{\beta}_1$. 但 $\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2\in W_1,\boldsymbol{\beta}_2-\boldsymbol{\beta}_1\in W_2$，若记 $\boldsymbol{\delta}=\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2=\boldsymbol{\beta}_2-\boldsymbol{\beta}_1$，则 $\boldsymbol{\delta}\neq\mathbf{0}$，且 $\boldsymbol{\delta}\in W_1\cap W_2$. 这与 $W_1\cap W_2=\{\mathbf{0}\}$ 矛盾. ∎

6.1 线性空间

定理 6.1-4 若 W_1 是线性空间 V^n 的一个真子空间(即 $0<\dim W_1<n$)，则必存在 V^n 的一个真子空间 W_2，使 $W_1 \oplus W_2 = V^n$.

请读者自证. ∎

子空间的交、和及直和的概念都可以推广到多个子空间的情形. 例如，s 个子空间 $W_1, W_2, \cdots, W_s$ 的和 $\sum_{i=1}^{s} W_i$ 是

$$\{\xi \in V \mid \xi = \sum_{i=1}^{s} \xi_i, \text{且 } \xi_i \in W_i, i = 1,2,\cdots,s\}.$$

如果和空间 $\sum_{i=1}^{s} W_i$ 中任一向量 ξ 的分解式

$$\xi = \sum_{i=1}^{s} \xi_i, \quad \xi_i \in W_i (i = 1,2,\cdots,s)$$

唯一，则称它为**直和**，记为

$$W_1 \oplus \cdots + W_s.$$

6.2 线性变换

本节应用矩阵来研究线性变换，说明线性变换所具有的一切性质均可用矩阵等价地描述.

6.2.1 线性变换的定义及矩阵表示

定义 6.2-1 线性空间 V^n 到 V^m 的变换 $\mathscr{T}$ 称为**线性变换**，如果对任意的数 k 及 V^n 中任意向量 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$，恒有

$$\mathscr{T}(\boldsymbol{\alpha}+\boldsymbol{\beta})=\mathscr{T}(\boldsymbol{\alpha})+\mathscr{T}(\boldsymbol{\beta}),\quad \mathscr{T}(k\boldsymbol{\alpha})=k\mathscr{T}(\boldsymbol{\alpha}).$$

记 $\xi=\mathscr{T}(\boldsymbol{\alpha})\in V^m$，则称 ξ 为 $\boldsymbol{\alpha}$ 在 $\mathscr{T}$ 下的**像**，$\boldsymbol{\alpha}$ 称为 ξ 的**原像**.

特别地，当 $\mathscr{T}$ 是 V^n 到自身的一个线性变换，则称 $\mathscr{T}$ 为 V^n **的线性变换**.

例 1 给定 $\boldsymbol{P}\in F^{m\times m}$，$\boldsymbol{Q}\in F^{n\times n}$，不难验证变换

$$\mathscr{T}: \boldsymbol{X}\in F^{m\times n}\rightarrow \boldsymbol{PXQ}\in F^{m\times n}$$

是 $F^{m\times n}$ 的一个线性变换. $\blacksquare$

例 2 对 $P_n[\lambda]$ 中的多项式求导 $\dfrac{\mathrm{d}}{\mathrm{d}\lambda}[\cdot]$，容易验证它是 $P_n[\lambda]$ 的线性变换，记为 $\mathscr{D}$，即

$$\mathscr{D}[p(\lambda)]=\frac{\mathrm{d}}{\mathrm{d}\lambda}p(\lambda),\quad \forall p(\lambda)\in P_n[\lambda]. \quad \blacksquare$$

6.2 线性变换

例 3 V 的恒等变换 $\mathscr{I}$ 定义为

$$\mathscr{I}(\boldsymbol{\alpha}) = \boldsymbol{\alpha}, \quad \forall \boldsymbol{\alpha} \in V.$$

而零变换 $\mathscr{O}$ 定义为 $\mathscr{O}(\boldsymbol{\alpha})=\mathbf{0}, \quad \forall \boldsymbol{\alpha}\in V.$

不难验证，它们都是 V 的线性变换. ■

例 4 $C[a,b]$ 中的函数积分 $\int_a^b \cdot \, dx$，容易验证它是 $C[a,b]$ 的线性变换，记为 $\mathscr{S}$，即

$$\mathscr{S}[f(x)] = \int_a^b f(x)dx, \quad \forall f(x) \in C[a,b].$$ ■

线性变换具有下列性质：

1) $\mathscr{T}(\mathbf{0})=\mathbf{0}$；$\mathscr{T}(-\boldsymbol{\alpha})=-\mathscr{T}(\boldsymbol{\alpha})$，$\forall \boldsymbol{\alpha}\in V$；

2) $\mathscr{T}(\sum_{i=1}^{r} k_i\boldsymbol{\alpha}_i) = \sum_{i=1}^{r} k_i\mathscr{T}(\boldsymbol{\alpha}_i)$，即任意一组向量的线性组合取像，等于它们取像再作同样的线性组合；

3) 一组线性相关的向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$，它们在 $\mathscr{T}$ 下的像 $\mathscr{T}(\boldsymbol{\alpha}_1),\mathscr{T}(\boldsymbol{\alpha}_2),\cdots,\mathscr{T}(\boldsymbol{\alpha}_r)$ 也线性相关.

线性无关的向量在 $\mathscr{T}$ 下的像可能线性相关. 例如，零变换把线性无关的向量都变为零向量，零向量是线性相关的.

6.2　线性变换

对于 V^n 到 V^m 的线性变换 $\mathscr{T}$，如何表示它呢？如果用对 V^n 的每个元素都给出其像的办法来表示 $\mathscr{T}$，显然是不方便，也不能用数学方法研究 $\mathscr{T}$ 的性质. 考虑到 $\mathscr{T}$ 是线性的，且 V^n 中每个元素都可用 V^n 的基唯一地线性表出，而线性变换保持线性组合关系，所以只要给出这个基的每个基向量的像，则 V^n 中每个元素的像也就完全确定了. 这就是有限维线性空间之间的线性变换可以用矩阵表示的基础.

设 $\mathscr{T}$ 是 V^n 到 V^m 的线性变换，在 V^n、V^m 中分别取基 $\mathscr{B}_\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$，$\mathscr{B}_\beta=\{\beta_1,\beta_2,\cdots,\beta_m\}$，则每个 $\boldsymbol{\alpha}_j$ 的像 $\mathscr{T}(\boldsymbol{\alpha}_j)(j=1,2,\cdots,n)$ 可用基 $\mathscr{B}_\beta$ 唯一地线性表出：

$$\mathscr{T}(\boldsymbol{\alpha}_j)=\sum_{i=1}^{m}a_{ij}\boldsymbol{\beta}_i=\{\boldsymbol{\beta}_1\ \cdots\ \boldsymbol{\beta}_m\}[a_{1j},\cdots,a_{mj}]^{\mathrm{T}}.$$

现把 $\mathscr{T}(\boldsymbol{\alpha}_j)(j=1,\cdots,n)$ 按顺序排列，并使用矩阵记号，则有

$$\{\mathscr{T}(\boldsymbol{\alpha}_1)\ \mathscr{T}(\boldsymbol{\alpha}_2)\ \cdots\ \mathscr{T}(\boldsymbol{\alpha}_n)\}=\{\boldsymbol{\beta}_1\boldsymbol{\beta}_2\ \cdots\ \boldsymbol{\beta}_m\}\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m1}&\cdots&a_{mn}\end{bmatrix}. \tag{6.2-1}$$

为了简化记法和便于运算，记

$$\mathscr{T}(\mathscr{B}_\alpha)=\{\mathscr{T}(\boldsymbol{\alpha}_1)\ \cdots\ \mathscr{T}(\boldsymbol{\alpha}_n)\},$$

那么(6.2-1)式可写为

$$\mathscr{T}(\mathscr{B}_\alpha)=\mathscr{B}_\beta\mathbf{A}, \tag{6.2-2}$$

6.2 线性变换

其中 $m\times n$ 矩阵 $\boldsymbol{A}$ 为

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

(6.2-2)式叫做线性变换 $\mathscr{T}$ 的矩阵表示，称 $\boldsymbol{A}$ 为 $\mathscr{T}$ 在基偶$\{\mathscr{B}_\alpha,\mathscr{B}_\beta\}$下的矩阵.

特别地，若 $\mathscr{T}$ 是 V^n 到自身的线性变换，这时 $V^m=V^n$，并规定 $\mathscr{B}_\beta$ 与 $\mathscr{B}_\alpha$ 一致，则(6.2-2)式为

$$\mathscr{T}(\mathscr{B}_\alpha)=\mathscr{B}_\alpha\boldsymbol{A}, \tag{6.2-2$'$}$$

其中 $\boldsymbol{A}$ 为 n 阶方阵. 我们称 $\boldsymbol{A}$ 为 $\mathscr{T}$ 在基 $\mathscr{B}_\alpha$ 下的矩阵.

例 5 求 $P_n[\lambda]$的线性变换 $\mathscr{D}=\dfrac{\mathrm{d}}{\mathrm{d}\lambda}$在基 $\mathscr{B}=\{1,\lambda,\lambda^2,\cdots,\lambda^n\}$下的矩阵.

解 $$\{\mathscr{D}(1)\ \ \mathscr{D}(\lambda)\ \ \cdots\ \ \mathscr{D}(\lambda^n)\}=\{0\ \ 1\ \ \cdots\ \ n\lambda^{n-1}\}=\{1\ \lambda\ \cdots\ \lambda^n\}\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

因此，$\mathscr{D}$ 在 $\mathscr{B}$ 下的矩阵是 $n+1$ 阶方阵

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

∎

6.2 线性变换

例 6 设 $\mathbf{R}^3$ 的线性变换 $\mathscr{T}$ 定义为

$$\mathscr{T}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}5x_1+4x_2+2x_3\\4x_1+5x_2+2x_3\\2x_1+2x_2+2x_3\end{bmatrix}\left(\text{即 } \mathscr{T}(\boldsymbol{x})=\begin{bmatrix}5&4&2\\4&5&2\\2&2&2\end{bmatrix}\boldsymbol{x}\right),$$

求 $\mathscr{T}$ 在基 $\mathscr{B}=\{[2,2,1]^{\mathrm{T}},[-1,1,0]^{\mathrm{T}},[-1,0,2]^{\mathrm{T}}\}$ 下的矩阵 $\boldsymbol{A}$.

解

$$\mathscr{T}\begin{bmatrix}2\\2\\1\end{bmatrix}=\begin{bmatrix}20\\20\\10\end{bmatrix}=\begin{bmatrix}2&-1&-1\\2&1&0\\1&0&2\end{bmatrix}\begin{bmatrix}10\\0\\0\end{bmatrix}=\mathscr{B}\begin{bmatrix}10\\0\\0\end{bmatrix},$$

$$\mathscr{T}\begin{bmatrix}-1\\1\\0\end{bmatrix}=\begin{bmatrix}-1\\1\\0\end{bmatrix}=\begin{bmatrix}2&-1&-1\\2&1&0\\1&0&2\end{bmatrix}\begin{bmatrix}0\\1\\0\end{bmatrix}=\mathscr{B}\begin{bmatrix}0\\1\\0\end{bmatrix},$$

$$\mathscr{T}\begin{bmatrix}-1\\0\\2\end{bmatrix}=\begin{bmatrix}-1\\0\\2\end{bmatrix}=\begin{bmatrix}2&-1&-1\\2&1&0\\1&0&2\end{bmatrix}\begin{bmatrix}0\\0\\1\end{bmatrix}=\mathscr{B}\begin{bmatrix}0\\0\\1\end{bmatrix},$$

因而 $\mathscr{T}$ 在基 $\mathscr{B}$ 下的矩阵 $\boldsymbol{A}$ 为
$$\boldsymbol{A}=\begin{bmatrix}10&0&0\\0&1&0\\0&0&1\end{bmatrix}.$$

∎

6.2　线性变换

有了线性变换 $\mathscr{T}$ 的矩阵表示，那么 V^n 中任一向量 ξ 在 $\mathscr{T}$ 作用下的像 $\mathscr{T}(\xi)$ 就可以用这个方阵表示了. 事实上，$\boldsymbol{\xi}\in V^n$ 可用 $\mathscr{B}_\alpha$ 线性表出：$\boldsymbol{\xi}=\sum_{i=1}^{n}x_i\boldsymbol{\alpha}_i=\mathscr{B}_\alpha\boldsymbol{x}$，从而 $\mathscr{T}(\xi)=\mathscr{T}(\sum_{i=1}^{n}x_i\boldsymbol{\alpha}_i)=\sum_{i=1}^{n}x_i\mathscr{T}(\boldsymbol{\alpha}_i)$，则由(6.2-1)或(6.2-2)式得

$$\mathscr{T}(\boldsymbol{\xi})=\mathscr{T}(\mathscr{B}_\alpha\boldsymbol{x})=\mathscr{T}(\mathscr{B}_\alpha)\boldsymbol{x}=\mathscr{B}_\beta\boldsymbol{A}\boldsymbol{x}. \qquad (6.2\text{-}3)$$

这就是说，若 ξ 在基 $\mathscr{B}_\alpha$ 下的坐标向量是 $\boldsymbol{x}$，则其像 $\mathscr{T}(\boldsymbol{\xi})$ 在基 $\mathscr{B}_\beta$ 下的坐标向量是 $\boldsymbol{A}\boldsymbol{x}$.

从上述讨论可知，取定基偶 $\{\mathscr{B}_\alpha,\mathscr{B}_\beta\}$ 后，线性变换 $\mathscr{T}$ 有唯一的矩阵表示，即 $\mathscr{T}$ 在此基偶下的矩阵 $\boldsymbol{A}$. 反过来，给定 $m\times n$ 矩阵 $\boldsymbol{A}=[a_{ij}]$，是否存在唯一的线性变换 $\mathscr{T}$，它在基偶 $\{\mathscr{B}_\alpha,\mathscr{B}_\beta\}$ 下的矩阵是 $\boldsymbol{A}$ 呢？回答是肯定的.

定理 6.2-1　设 $\mathscr{B}_\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$，$\mathscr{B}_\beta=\{\beta_1,\beta_2,\cdots,\beta_m\}$ 分别是 V^n 与 V^m 的基. 给定 $m\times n$ 矩阵 $\boldsymbol{A}$，存在 V^n 到 V^m 的唯一的线性变换 $\mathscr{T}$，它在基偶 $\{\mathscr{B}_\alpha,\mathscr{B}_\beta\}$ 下的矩阵是 $\boldsymbol{A}$.

证　只证存在性，请读者自证唯一性.

任取 $\boldsymbol{\xi}\in V^n$，设它在基 $\mathscr{B}_\alpha$ 下的坐标向量为 $\boldsymbol{x}$，即 $\boldsymbol{\xi}=\mathscr{B}_\alpha\boldsymbol{x}$，令 $\boldsymbol{\eta}=\mathscr{B}_\beta\boldsymbol{A}\boldsymbol{x}$，则 $\boldsymbol{\eta}\in V^m$，定义 V^n 到 V^m 的变换 $\mathscr{T}$ 为

$$\mathscr{T}(\boldsymbol{\xi})=\boldsymbol{\eta},$$

下证 $\mathscr{T}$ 就是定理中所要求的线性变换.

6.2 线性变换

首先证 $\mathscr{T}$ 是线性的. 设 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2\in V^n$，它们在 $\mathscr{B}_\alpha$ 下的坐标向量分别为 $\boldsymbol{x}_1$、$\boldsymbol{x}_2$，则

$$\boldsymbol{\xi}_1+\boldsymbol{\xi}_2=\mathscr{B}_\alpha\boldsymbol{x}_1+\mathscr{B}_\alpha\boldsymbol{x}_2=\mathscr{B}_\alpha(\boldsymbol{x}_1+\boldsymbol{x}_2).$$

按 $\mathscr{T}$ 的定义有 $\qquad \mathscr{T}(\boldsymbol{\xi}_1)=\mathscr{B}_\beta\boldsymbol{A}\boldsymbol{x}_1,\quad \mathscr{T}(\boldsymbol{\xi}_2)=\mathscr{B}_\beta\boldsymbol{A}\boldsymbol{x}_2,$

从而 $\qquad \mathscr{T}(\boldsymbol{\xi}_1+\boldsymbol{\xi}_2)=\mathscr{B}_\beta\boldsymbol{A}(\boldsymbol{x}_1+\boldsymbol{x}_2)=\mathscr{B}_\beta\boldsymbol{A}\boldsymbol{x}_1+\mathscr{B}_\beta\boldsymbol{A}\boldsymbol{x}_2=\mathscr{T}(\boldsymbol{\xi}_1)+\mathscr{T}(\boldsymbol{\xi}_2).$

类似可证，$\mathscr{T}(k\boldsymbol{\xi})=k\mathscr{T}(\boldsymbol{\xi})$. 因此 $\mathscr{T}$ 是线性变换.

再证 $\mathscr{T}(\mathscr{B}_\alpha)=\mathscr{B}_\beta\boldsymbol{A}$. 显然，$\boldsymbol{\alpha}_i(i=1,2,\cdots,n)$ 在 $\mathscr{B}_\alpha$ 下的坐标向量是 $\boldsymbol{e}_i=[0,\cdots,0,\overset{i}{1},0,\cdots,0]^{\mathrm{T}}$，即第 i 个分量为 1，其余分量为 0 的 n 维列向量. 从而按 $\mathscr{T}$ 的定义有 $\mathscr{T}(\boldsymbol{\alpha}_i)=\mathscr{B}_\beta\boldsymbol{A}\boldsymbol{e}_i$. 因此，

$$\mathscr{T}(\mathscr{B}_\alpha)=\{\mathscr{T}(\boldsymbol{\alpha}_1)\ \cdots\ \mathscr{T}(\boldsymbol{\alpha}_n)\}=\{\mathscr{B}_\beta\boldsymbol{A}\boldsymbol{e}_1\ \cdots\ \mathscr{B}_\beta\boldsymbol{A}\boldsymbol{e}_n\}$$
$$=\mathscr{B}_\beta\boldsymbol{A}[\boldsymbol{e}_1\ \cdots\ \boldsymbol{e}_n]=\mathscr{B}_\beta\boldsymbol{A}\boldsymbol{I}=\mathscr{B}_\beta\boldsymbol{A}. \qquad \blacksquare$$

这个定理表明，取定基偶 $\{\mathscr{B}_\alpha,\mathscr{B}_\beta\}$ 后，V^n 到 V^m 的线性变换 $\mathscr{T}$ 与 $m\times n$ 矩阵 A 是一一对应的. 如果把 V^n 到 V^m 的线性变换全体记作 $\mathscr{L}(V^n,V^m)$，即

$$\mathscr{L}(V^n,V^m)=\{\mathscr{T}\mid \mathscr{T}\text{ 是 }V^n\text{ 到 }V^m\text{ 的线性变换}\},$$

并对任意 $\mathscr{T},\mathscr{S}\in\mathscr{L}(V^n,V^m)$ 定义

$$(\mathscr{T}+\mathscr{S})(\boldsymbol{\alpha})=\mathscr{T}(\boldsymbol{\alpha})+\mathscr{L}(\boldsymbol{\alpha}),\ \forall\boldsymbol{\alpha}\in V^n,$$
$$(k\mathscr{T})(\boldsymbol{\alpha})=k\mathscr{T}(\boldsymbol{\alpha}),\ \forall\boldsymbol{\alpha}\in V^n,\forall k\in F,$$

则容易验证，这种加法及数乘运算满足线性空间定义中的八条规则，故 $\mathscr{L}(V^n,V^m)$ 是数域 F 上的线性空间. 于是 $\mathscr{L}(V^n,V^m)$ 与 $F^{m\times n}$ 有相同的代数结构.

6.2 线性变换

6.2.2 线性变换的零空间和值空间

定义 6.2-2 设 $\mathscr{T}$ 是 V^n 到 V^m 的线性变换,集合

$$\mathscr{N}(\mathscr{T}) = \{\boldsymbol{\xi} \in V^n \mid \mathscr{T}(\boldsymbol{\xi}) = \mathbf{0}\}, \tag{6.2-4}$$

$$\mathscr{R}(\mathscr{T}) = \{\boldsymbol{\beta} \in V^m \mid \boldsymbol{\beta} = \mathscr{T}(\boldsymbol{\alpha}), \boldsymbol{\alpha} \in V^n\} \tag{6.2-5}$$

分别称为 $\mathscr{T}$ 的**核**和 $\mathscr{T}$ 的**值域**.

不难验证,$\mathscr{N}(\mathscr{T})$是 V^n 的子空间,也称它为 $\mathscr{T}$ 的零空间,其维数叫做 $\mathscr{T}$ 的**零度**,记为 $n(\mathscr{T})$;$\mathscr{R}(\mathscr{T})$是 V^m 的子空间,也称为**值空间**,其维数叫做 $\mathscr{T}$ 的秩,记为 $r(\mathscr{T})$,并且,若 $\mathscr{B}_\alpha=\{\alpha_1,\cdots,\alpha_n\}$,$\mathscr{B}_\beta=\{\beta_1,\cdots,\beta_m\}$分别是 V^n、V^m 的基,且 $\mathscr{T}$ 在基偶$\{\mathscr{B}_\alpha,\mathscr{B}_\beta\}$下的矩阵是 $\boldsymbol{A}$,则有

$$\mathscr{R}(\mathscr{T}) = \mathrm{Span}(\mathscr{T}(\boldsymbol{\alpha}_1),\cdots,\mathscr{T}(\boldsymbol{\alpha}_n)),$$

$$r(\mathscr{T}) = r(\boldsymbol{A}).$$

定理 6.2-2 设 $\mathscr{T}$ 是 V^n 到 V^m 的线性变换,则

$$n(\mathscr{T}) + r(\mathscr{T}) = n. \tag{6.2-6}$$

证 设 $n(\mathscr{T})=r$,且$\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r\}$是 $\mathscr{N}(\mathscr{T})$的一个基,则由于 $\mathscr{N}(\mathscr{T})$是 V^n 的子空间,故存在 V^n 中的 $n-r$ 个向量 $\boldsymbol{\alpha}_{r+1},\cdots,\boldsymbol{\alpha}_n$ 使$\{\alpha_1,\alpha_2,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n\}$为 V^n 的基.

下证$\{\mathscr{T}(\alpha_{r+1}),\cdots,\mathscr{T}(\alpha_n)\}$是 $\mathscr{R}(\mathscr{T})$的基,从而(6.2-6)式成立.

任取 $\boldsymbol{\xi} \in V^n$,则 ξ 可表示为 $\xi = \sum\limits_{i=1}^{n} k_i\alpha_i$. 由于 $\boldsymbol{\alpha}_i \in \mathscr{N}(\mathscr{T})(i = 1,2,\cdots,r)$,故 $\mathscr{T}(\boldsymbol{\alpha}_i) =$

6.2 线性变换

$\mathbf{0}, i=1,2,\cdots,r$，从而有

$$\mathscr{T}(\xi)=\sum_{i=1}^{n} k_i\mathscr{T}(\boldsymbol{\alpha}_i)=\sum_{i=r+1}^{n} k_i\mathscr{T}(\boldsymbol{\alpha}_i).$$

也就是说 $\mathscr{R}(\mathscr{T})$ 中的向量 $\mathscr{T}(\xi)$ 可由 $\{\mathscr{T}(\boldsymbol{\alpha}_{r+1}),\cdots,\mathscr{T}(\boldsymbol{\alpha}_n)\}$ 线性表出. 再则，设有等式

$$\sum_{i=r+1}^{n} k_i\mathscr{T}(\boldsymbol{\alpha}_i)=\mathbf{0},$$

则 $\mathscr{T}\left(\sum_{i=r+1}^{n} k_i\boldsymbol{\alpha}_i\right)=\mathbf{0}$，故 $\sum_{i=r+1}^{n} k_i\boldsymbol{\alpha}_i\in\mathscr{N}(T)$.

于是，存在数 $c_j(j=1,2,\cdots,r)$ 使

$$\sum_{j=1}^{r} c_j\boldsymbol{\alpha}_j=\sum_{i=r+1}^{n} k_i\boldsymbol{\alpha}_i,$$

即

$$\sum_{j=1}^{r}(-c_j)\boldsymbol{\alpha}_j+\sum_{i=r+1}^{n} k_i\boldsymbol{\alpha}_i=\mathbf{0}.$$

但 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关，故有

$$c_1=\cdots=c_r=0,\quad k_{r+1}=\cdots=k_n=0.$$

也就是说 $\mathscr{T}(\boldsymbol{\alpha}_{r+1}),\cdots,\mathscr{T}(\boldsymbol{\alpha}_n)$ 线性无关.

综合上述两个方面，即知 $\{\mathscr{T}(\boldsymbol{\alpha}_{r+1}),\cdots,\mathscr{T}(\boldsymbol{\alpha}_n)\}$ 是 $\mathscr{R}(\mathscr{T})$ 的基. ∎

6.2 线性变换

特别地，若 $\mathscr{T}$ 是 V^n 的线性变换，则 $\mathscr{N}(\mathscr{T})$、$\mathscr{R}(\mathscr{T})$ 都是 V^n 的子空间，且它们的维数之和等于 n. 但 $\mathscr{N}(\mathscr{T})+\mathscr{R}(\mathscr{T})$ 却不一定是 V^n. 例如，对于 $P_{n-1}[\lambda]$ 的线性变换 $\mathscr{D}$（即求导）来说，$\mathscr{N}(\mathscr{D})=P_0[\lambda]$，$\mathscr{R}(\mathscr{D})=P_{n-2}[\lambda]$，故

$$n(\mathscr{D})+r(\mathscr{D})=1+(n-1)=n,$$

但
$$\mathscr{N}(\mathscr{D})+\mathscr{R}(\mathscr{D})=P_0[\lambda]+P_{n-2}[\lambda]=P_{n-2}[\lambda]\neq P_{n-1}[\lambda].$$

例 7 设 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}$ 是 V^3 的一个基，且 V^3 的线性变换 $\mathscr{T}$ 在这个基下的矩阵是

$$\boldsymbol{A}=\begin{bmatrix}1&0&2\\-1&2&1\\1&2&5\end{bmatrix},$$

求 $\mathscr{T}$ 的零空间和值空间.

解 设 $\boldsymbol{\xi}=\{\boldsymbol{\alpha}_1\ \boldsymbol{\alpha}_2\ \boldsymbol{\alpha}_3\}[x_1,x_2,x_3]^{\mathrm{T}}\in\mathscr{N}(\mathscr{T})$，则有

$$\mathscr{T}(\boldsymbol{\xi})=\{\mathscr{T}(\boldsymbol{\alpha}_1)\ \mathscr{T}(\boldsymbol{\alpha}_2)\ \mathscr{T}(\boldsymbol{\alpha}_3)\}[x_1,x_2,x_3]^{\mathrm{T}}=\{\boldsymbol{\alpha}_1\ \boldsymbol{\alpha}_2\ \boldsymbol{\alpha}_3\}\begin{bmatrix}1&0&2\\-1&2&1\\1&2&5\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\boldsymbol{0}.$$

由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关，故

$$\begin{bmatrix}1&0&2\\-1&2&1\\1&2&5\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\boldsymbol{0}.$$

6.2　线性变换

解之得$[x_1, x_2, x_3]^{\mathrm{T}} = k[-2, -\frac{3}{2}, 1]^{\mathrm{T}}$，$k$ 为任意实数.

因此，
$$\boldsymbol{\xi} = k(-2\boldsymbol{\alpha}_1 - \frac{3}{2}\boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3),$$

即
$$\mathscr{N}(\mathscr{T}) = \mathrm{Span}(-2\boldsymbol{\alpha}_1 - \frac{3}{2}\boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3).$$

由于 $r(\boldsymbol{A}) = 2$，且 $\boldsymbol{A}$ 的前二列线性无关，故
$$\mathscr{R}(\mathscr{T}) = \mathrm{Span}(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3, 2\boldsymbol{\alpha}_2 + 2\boldsymbol{\alpha}_3). \quad \blacksquare$$

6.2 线性变换

6.2.3 线性变换的最简矩阵表示及不变子空间

由于线性空间的基不是唯一的，同一个线性变换在不同基下的矩阵一般是不相同的，因此就提出下述两个问题：

(1) 如果有另一个基偶$\{\mathscr{B}_{\alpha'},\mathscr{B}_{\beta'}\}$，那么线性变换$\mathscr{T}$在此基偶下的矩阵$\boldsymbol{B}$与$\mathscr{T}$在$\{\mathscr{B}_\alpha,\mathscr{B}_\beta\}$下的矩阵$\boldsymbol{A}$之间有何关系？

(2) 怎样选取基偶，才能使$\mathscr{T}$的矩阵最简？

我们在下面逐一回答这两个问题.

设V^n中有两个基$\mathscr{B}_\alpha$、$\mathscr{B}_{\alpha'}$，$\mathscr{B}_\alpha$到$\mathscr{B}_{\alpha'}$的基变换矩阵是n阶方阵$\boldsymbol{P}$；在V^m中有两个基$\mathscr{B}_\beta$、$\mathscr{B}_{\beta'}$，$\mathscr{B}_\beta$到$\mathscr{B}_{\beta'}$的基变换矩阵是m阶方阵$\boldsymbol{Q}$；而V^n到V^m的线性变换$\mathscr{T}$在基偶$\{\mathscr{B}_\alpha,\mathscr{B}_\beta\}$、$\{\mathscr{B}_{\alpha'},\mathscr{B}_{\beta'}\}$下的矩阵分别为$m\times n$矩阵$\boldsymbol{A}$、$\boldsymbol{B}$，那么由关系式

$$\mathscr{B}_{\alpha'}=\mathscr{B}_\alpha\boldsymbol{P},\quad \mathscr{B}_{\beta'}=\mathscr{B}_\beta\boldsymbol{Q},$$

$$\mathscr{T}(\mathscr{B}_\alpha)=\mathscr{B}_\beta\boldsymbol{A},\quad \mathscr{T}(\mathscr{B}_{\alpha'})=\mathscr{B}_{\beta'}\boldsymbol{B}$$

可以推出
$$\mathscr{T}(\mathscr{B}_{\alpha'})=\mathscr{B}_{\beta'}B=\mathscr{B}_\beta\boldsymbol{QB},$$

$$\mathscr{T}(\mathscr{B}_{\alpha'})=\mathscr{T}(\mathscr{B}_\alpha\boldsymbol{P})=\mathscr{T}(\mathscr{B}_\alpha)\boldsymbol{P}=\mathscr{B}_\beta\boldsymbol{AP}.$$

从而有
$$\mathscr{B}_\beta\boldsymbol{AP}=\mathscr{B}_\beta\boldsymbol{QB},$$

即
$$\mathscr{B}_\beta(\boldsymbol{AP}-\boldsymbol{QB})=\boldsymbol{O}. \tag{6.2-7}$$

由于$\mathscr{B}_\beta$是基，所以有
$$\boldsymbol{AP}=\boldsymbol{QB},\quad \boldsymbol{A}=\boldsymbol{QBP}^{-1},\quad \boldsymbol{B}=\boldsymbol{Q}^{-1}\boldsymbol{AP}. \tag{6.2-8}$$

这就是说，$\boldsymbol{A}$与$\boldsymbol{B}$是等价的.

6.2 线性变换

特别地，如果 $\mathscr{T}$ 是 V^n 到自身的线性变换，即 $\mathscr{T}$ 是 V^n 的线性变换，则在上述推导过程中，令 $V^m=V^n$，$\mathscr{B}_\beta=\mathscr{B}_\alpha$，$\mathscr{B}_{\beta'}=\mathscr{B}_{\alpha'}$，便得

$$\boldsymbol{AP}=\boldsymbol{PB},\quad \boldsymbol{A}=\boldsymbol{PBP}^{-1},\quad \boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{AP}. \tag{6.2-9}$$

(6.2-9)式表明，V^n 的线性变换 $\mathscr{T}$ 在基 $\mathscr{B}_\alpha$ 下的方阵 $\boldsymbol{A}$ 与在基 $\mathscr{B}_{\alpha'}$ 下的方阵 $\boldsymbol{B}$ 相似，且 $\mathscr{B}_\alpha$ 到 $\mathscr{B}_{\alpha'}$ 的基变换矩阵 $\boldsymbol{P}$ 即是相似变换矩阵.

于是，V^n 到 V^m 的线性变换 $\mathscr{T}$ 在不同基偶下的矩阵是等价关系，而 V^n 的线性变换 $\mathscr{T}$ 在不同基下的方阵是相似关系. 从而第二个问题等价于：与 $\boldsymbol{A}$ 等价的所有矩阵中最简矩阵是什么？与方阵 $\boldsymbol{A}$ 相似的所有方阵中最简方阵是什么？

前面所讨论过的矩阵的等价标准形和方阵的 Jordan 标准形（对角矩阵作为 Jordan 矩阵的一种特殊情形）已经解决了这第二个问题.

现在从几何的观点进一步讨论与此有关的问题.

定义 6.2-3 设 $\mathscr{T}$ 是线性空间 V 的一个线性变换，W 是 V 的子空间. 如果对任意 $\xi\in W$ 均有

$$\mathscr{T}(\xi)\in W,$$

则称 W 为 $\mathscr{T}$ 的**不变子空间**.

例如，$\mathscr{T}$ 的零空间 $\mathscr{N}(\mathscr{T})$ 和值空间 $\mathscr{R}(\mathscr{T})$ 都是 $\mathscr{T}$ 的不变子空间，并且 $\mathscr{T}$ 的不变子空间的交空间及和空间也都是 $\mathscr{T}$ 的不变子空间. 请读者自证.

6.2 线性变换

例 8 设 V 是线性空间，$\boldsymbol{\alpha}_i\in V(i=1,2,\cdots,r)$，证明 $W=\mathrm{Span}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$ 是 $\mathscr{T}$ 的不变子空间的充要条件是 $\mathscr{T}(\alpha_i)\in W(i=1,2,\cdots,r)$.

证 必要性显然. 证明充分性如下.

设 ξ 是 W 中任一向量，则 $\xi=\sum\limits_{i=1}^{r}k_i\boldsymbol{\alpha}_i$，从而

$$\mathscr{T}(\boldsymbol{\xi})=\mathscr{T}\left(\sum_{i=1}^{n}k_i\boldsymbol{\alpha}_i\right)=\sum_{i=1}^{n}k_i\mathscr{T}(\boldsymbol{\alpha}_i)\in W,$$

即 W 是 $\mathscr{T}$ 的不变子空间. ∎

利用 $\mathscr{T}$ 的不变子空间可以简化 $\mathscr{T}$ 的矩阵表示. 事实上，若 W 是 V^n 的子空间，且为 $\mathscr{T}$ 的不变子空间，$\dim W=k(1\leqslant k\leqslant n-1)$，则在 W 中取一个基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k\}$，并把它扩充为 V^n 的基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k,\boldsymbol{\alpha}_{k+1},\cdots,\boldsymbol{\alpha}_n\}$，由于 $\mathscr{T}(\boldsymbol{\alpha}_i)\in W(i=1,2,\cdots,k)$，故 $\mathscr{T}(\boldsymbol{\alpha}_i)$ 可用 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$ 线性表出，从而 $\mathscr{T}(\boldsymbol{\alpha}_i)(i=1,2,\cdots,k)$ 在基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 下的坐标向量的后 $n-k$ 个分量必为零. 因此 $\mathscr{T}$ 在基 $\mathscr{B}=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 下的矩阵具有下述形式：

$$\left[\begin{array}{c:c} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \hdashline \underbrace{\boldsymbol{O}}_{k} & \underbrace{\boldsymbol{A}_{22}}_{n-k}\end{array}\right]\begin{array}{l}\}k\\ \}n-k\end{array}.$$

反之，如果 $\mathscr{T}$ 在基 $\mathscr{B}=\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 下的矩阵具有上述形式，则 $W=\mathrm{Span}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k)$ 是 $\mathscr{T}$ 的不变子空间.

6.2 线性变换

特别地，当 $W_1=\mathrm{Span}(\boldsymbol{\alpha}_{k+1},\boldsymbol{\alpha}_{k+2},\cdots,\boldsymbol{\alpha}_n)$ 也是 $\mathscr{T}$ 的不变子空间，则 $\mathscr{T}$ 在基 $\mathscr{B}$ 下的矩阵是分块对角矩阵

$$\begin{array}{cc} \left[\begin{array}{c|c} \boldsymbol{A}_{11} & \boldsymbol{O} \\ \hline \underbrace{\boldsymbol{O}}_{k} & \underbrace{\boldsymbol{A}_{22}}_{n-k} \end{array}\right] & \begin{array}{l} \}k \\ \}n-k \end{array} \end{array},$$

其中 $\boldsymbol{A}_{11}$ 是 k 阶方阵，$\boldsymbol{A}_{22}$ 是 $n-k$ 阶方阵.

更一般地，如果 $W_i(i=1,2,\cdots,s)$ 都是 $\mathscr{T}$ 的不变子空间，且

$$W_1\oplus W_2\oplus\cdots\oplus W_s=V^n,$$

那么在每个 W_i 取一个基 $\{\alpha_{i1},\alpha_{i2},\cdots,\alpha_{in_i}\}(i=1,2,\cdots,s)$，这里 $n_i=\dim W_i$，并把它们按顺序排列成为 V^n 的基

$$\mathscr{B}=\{\alpha_{11},\alpha_{12},\cdots,\alpha_{1n_1},\alpha_{21},\cdots,\alpha_{2n_2},\cdots,\alpha_{s1},\cdots,\alpha_{sn_s}\},$$

则 $\mathscr{T}$ 在 $\mathscr{B}$ 下的矩阵是分块对角矩阵

$$\mathrm{diag}(\boldsymbol{A}_{11},\boldsymbol{A}_{22},\cdots,\boldsymbol{A}_{ss}),$$

其中 $\boldsymbol{A}_{ii}(i=1,2,\cdots,s)$ 是 n_i 阶方阵. 反之，若 $\mathscr{T}$ 在 $\mathscr{B}$ 下的矩阵是上述的分块对角矩阵，则

$$W_i=\mathrm{Span}(\boldsymbol{\alpha}_{i1},\boldsymbol{\alpha}_{i2},\cdots,\boldsymbol{\alpha}_{in_i})\quad(i=1,2,\cdots,s)$$

都是 $\mathscr{T}$ 的不变子空间，且 V^n 是这些子空间的直和.

定义 6.2-4 设 $\mathscr{T}$ 是 $V^n(F)$ 的一个线性变换，如果存在数 $\lambda_0\in F$，$\boldsymbol{\xi}\in V^n(F)$ 且 $\boldsymbol{\xi}\neq 0$，使

6.2 线性变换

$$\mathscr{T}(\xi) = \lambda_0\xi, \tag{6.2-10}$$

则称 λ_0 是 $\mathscr{T}$ 的特征值，称 ξ 为 $\mathscr{T}$ 的属于特征值 λ_0 的特征向量. $V^n(F)$ 的子空间

$$V_{\lambda_0} = \{\xi \in V^n(F) \mid \mathscr{T}(\xi) = \lambda_0\xi\} \tag{6.2-11}$$

称为 $\mathscr{T}$ 关于 λ_0 的**特征子空间**，$\dim V_{\lambda_0}$ 称为 λ_0 的**几何重数**.

显然，V_{λ_0} 是 $\mathscr{T}$ 的不变子空间，且 λ_0 的几何重数就是 $\mathscr{T}$ 的属于 λ_0 的线性无关特征向量的个数.

为了求出 $\mathscr{T}$ 的特征值和特征向量，在 V^n 中取一个基 $\mathscr{B}=\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$，且设 $\mathscr{T}$ 在 $\mathscr{B}$ 下的矩阵是 $\boldsymbol{A}$. 若 ξ 是 $\mathscr{T}$ 的一个特征向量，λ_0 是对应的特征值，即 $\mathscr{T}(\xi)=\lambda_0\xi$，$\xi\neq\boldsymbol{0}$，则 ξ 可用 $\mathscr{B}$ 线性表出：$\xi = \sum_{i=1}^{n} x_i\boldsymbol{\alpha}_i = \mathscr{B}\boldsymbol{x}$，$\boldsymbol{x} = [x_1,x_2,\cdots,x_n]^{\mathrm{T}}$，因此

$$\mathscr{T}(\xi) = \mathscr{T}(\mathscr{B}\boldsymbol{x}) = \mathscr{T}(\mathscr{B})\boldsymbol{x} = \mathscr{B}\boldsymbol{A}\boldsymbol{x}, \lambda_0\xi = \mathscr{B}(\lambda_0\boldsymbol{x}).$$

将此代入(6.2-10)式得 $\mathscr{B}(\boldsymbol{A}\boldsymbol{x})=\mathscr{B}(\lambda_0\boldsymbol{x})$，即 $\mathscr{B}(\boldsymbol{A}\boldsymbol{x}-\lambda_0\boldsymbol{x})=\boldsymbol{0}$. 由此即得 λ_0 和 ξ 在 $\mathscr{B}$ 下的坐标向量 $\boldsymbol{x}$ 所满足的方程为

$$\boldsymbol{A}\boldsymbol{x} = \lambda_0\boldsymbol{x}, \quad \boldsymbol{x} \neq \boldsymbol{0}.$$

也就是说，如果 λ_0 是 $\mathscr{T}$ 的特征值，则 λ_0 必是方阵 $\boldsymbol{A}$ 的特征值. 反之亦然.

于是，线性变换 $\mathscr{T}$ 的特征值问题与方阵 $\boldsymbol{A}$ 的特征值问题一一对应.

6.2 线性变换

例 9 $P_2[t]$的线性变换$\mathscr{T}$定义为

$$\mathscr{T}[p(t)]=p(t)+(t+1)\frac{\mathrm{d}}{\mathrm{d}t}p(t),\quad p(t)\in P_2[t],$$

求$\mathscr{T}$的特征值及对应的特征向量.

解 取$P_2[t]$的一个基$\mathscr{B}=\{1,t,t^2\}$,则$\mathscr{T}$在$\mathscr{B}$下的矩阵是

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

$\boldsymbol{A}$的特征值是$\lambda_1=1,\lambda_2=2,\lambda_3=3$,对应的特征向量分别是

$$k_1[1,0,0]^{\mathrm{T}},\quad k_2[1,1,0]^{\mathrm{T}},\quad k_3[1,2,1]^{\mathrm{T}},$$

其中k_1、k_2、k_3均为任意非零实数.因此,$\mathscr{T}$的特征值为$\lambda_1=1,\lambda_2=2,\lambda_3=3$,而对应的特征向量分别是多项式

$$k_1,\quad k_2(1+t),\quad k_3(1+2t+t^2),$$

其中k_1、k_2、k_3均为任意非零实数. ∎

6.2　线性变换

设 $\mathscr{T}$ 是 V^n 的线性变换，由于 $\mathscr{T}$ 在不同基下的方阵是相似的，而相似方阵有相同的特征多项式，所以 $\mathscr{T}$ 在任一个基下的方阵的特征多项式可以称为 $\mathscr{T}$ 的特征多项式. 因此，如果 $\lambda_1,\cdots,\lambda_s$ 是 $\mathscr{T}$ 的所有不同的特征值，则集合 $Sp(\mathscr{T})=\{\lambda_1,\cdots,\lambda_s\}$ 称为 $\mathscr{T}$ 的**谱**，并且 $\mathscr{T}$ 的特征多项式 $f(\lambda)$ 可以表示为

$$f(\lambda)=(\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_n)^{n_s}\text{，且 } n_1+\cdots+n_s=n,$$

其中 $n_i(i=1,2,\cdots,s)$ 称为特征值 λ_i 的**代数重数**.

与矩阵一样，可以证明，$\mathscr{T}$ 的属于不同特征值的特征向量线性无关，且对 $\mathscr{T}$ 的任一特征值 λ_i，有

$$\dim V_{\lambda_i}\leqslant n_i, \tag{6.2-12}$$

即 $\mathscr{T}$ 的任何特征值的几何重数不大于其代数重数.

由此可知，若 $\mathscr{T}$ 的谱为 $Sp(\mathscr{T})=\{\lambda_1,\cdots,\lambda_s\}$，则

$$V_{\lambda_1}+\cdots+V_{\lambda_s}=V_{\lambda_1}\oplus\cdots\oplus V_{\lambda_s},$$

$$\dim V_{\lambda_1}+\cdots+\dim V_{\lambda_s}\leqslant n_1+\cdots+n_s=n.$$

定义 6.2-5　设 $\mathscr{T}$ 是 V^n 的线性变换，如果存在 V^n 的基 $\mathscr{B}$，使 $\mathscr{T}$ 在 $\mathscr{B}$ 下的方阵是对角矩阵，则称 $\mathscr{T}$ 可对角化.

6.2 线性变换

定理 6.2-3 设 $\mathscr{T}$ 是 V^n 的线性变换，则 $\mathscr{T}$ 可对角化的充要条件是，下列等价条件之一成立：

1) $\mathscr{T}$ 有 n 个线性无关的特征向量；

2) $V_{\lambda_1}\oplus\cdots\oplus V_{\lambda_s}=V^n$，这里 $Sp(\mathscr{T})=\{\lambda_1,\cdots,\lambda_s\}$；

3) $\dim V_{\lambda_i}=n_i(i=1,2,\cdots,s)$，$n_i$ 是 λ_i 的代数重数.

请读者自证. ∎

例 10 证明 $P_2[t]$ 的线性变换 $\mathscr{D}$ 不可对角化.

证 取 $P_2[t]$ 的一个基为 $\mathscr{B}=\{1,t,t^2\}$，$\mathscr{D}$ 在 $\mathscr{B}$ 下的方阵是

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

A 的特征多项式是 $\det(\boldsymbol{A}-\lambda\boldsymbol{I})=(-\lambda)^3$，

$\lambda=0$ 是它的三重根. 由于 $r(\boldsymbol{A}-0\cdot\boldsymbol{I})=r(\boldsymbol{A})=2<3$，所以 $\boldsymbol{A}$ 不可对角化，从而 $\mathscr{D}$ 不可对角化. ∎

例 11 设 $\mathscr{B}=\{\alpha_1,\alpha_2,\alpha_3\}$ 是 V^3 的一个基，且 V^3 的线性变换 $\mathscr{T}$ 在 $\mathscr{B}$ 下的矩阵是

$$\boldsymbol{A}=\begin{bmatrix}4&0&1\\2&3&2\\1&0&4\end{bmatrix},$$

问 $\mathscr{T}$ 可否对角化？若可对角化，求 V^3 的基 $\mathscr{B}_d$，使 $\mathscr{T}$ 在 $\mathscr{B}_d$ 下的矩阵是对角矩阵.

6.2 线性变换

解 $\boldsymbol{A}$ 的特征多项式为

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=(5-\lambda)(3-\lambda)^2,$$

故 $\boldsymbol{A}$ 有两个不同的特征值：$\lambda_1=5$ 和 $\lambda_{2,3}=3$.

对 $\lambda_1=5$，对应的特征向量可取为 $\boldsymbol{x}_1=[1,2,1]^{\mathrm{T}}$.

对 $\lambda_{2,3}=3$，考虑线性方程组 $(\boldsymbol{A}-3\boldsymbol{I})\boldsymbol{x}=\boldsymbol{y}$，即

$$\begin{bmatrix}1&0&1\\2&0&2\\1&0&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}.$$

由于其系数矩阵的秩为 1，故当 $\boldsymbol{y}=\boldsymbol{0}$ 时有两个线性无关的解，例如取为

$$\boldsymbol{x}_2=[1,0,-1]^{\mathrm{T}},\quad \boldsymbol{x}_3=[0,1,0]^{\mathrm{T}}.$$

因此 $\boldsymbol{A}$ 可对角化，故 $\mathscr{T}$ 可对角化.

6.3 内积空间及两类特殊的线性变换

线性空间中元素的基本运算是加法(通过负元素概念也有了减法)和数乘.但几何空间中所讨论过的一些概念,如元素的长度、元素之间的正交性、标准正交基等在线性空间的理论中还没有反映出来.本节引入内积的概念,从而在抽象线性空间中有这些几何方面的概念,并得出相应的结果.

定义 6.3-1 设 V 是线性空间,如果对任意 $\boldsymbol{\alpha},\boldsymbol{\beta}\in V$,都有一个数 $\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle$ 与之对应,并具有下列性质:

1) $\langle\boldsymbol{\alpha},\boldsymbol{\alpha}\rangle\geqslant 0$,当且仅当 $\boldsymbol{\alpha}=\mathbf{0}$ 时才有 $\langle\boldsymbol{\alpha},\boldsymbol{\alpha}\rangle=0$;

2) $\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=\overline{\langle\boldsymbol{\beta},\boldsymbol{\alpha}\rangle}$;

3) $\langle\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\gamma}\rangle=\langle\boldsymbol{\alpha},\boldsymbol{\gamma}\rangle+\langle\boldsymbol{\beta},\boldsymbol{\gamma}\rangle$,$\boldsymbol{\gamma}\in V$;

4) $\langle k\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=k\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle$,$k$ 为任意数;

则称 $\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle$ 为 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的**内积**.定义了内积的线性空间称为**内积空间**.特别地,定义了内积的实线性空间称为**欧氏空间**;定义了内积的复线性空间称为**酉空间**.

例 1 在 $P_n[t]$ 中定义

$$\langle p(t),q(t)\rangle=\int_a^b t^2p(t)q(t)\mathrm{d}t,$$

其中 a、b 是任意给定的两个实数,$a<b$,则容易验证它是内积.在定义了这个内积后,$P_n[t]$ 就成为一个欧氏空间. ∎

6.3 内积空间及两类特殊的线性变换

例 2 在 $C[a,b]$上，对任意两个实函数 $f(x)$、$g(x)$，定义

$$\langle f(x),g(x)\rangle=\int_a^b \rho(x)f(x)g(x)\mathrm{d}x, \tag{6.3-1}$$

其中 $\rho(x)\geqslant 0$ 是给定的权函数，满足下述条件：

(1) 对一切非负实数 r，$\int_a^b |x|^r\rho(x)\mathrm{d}x$ 存在；

(2) 对非负连续实函数 $h(x)$，如果

$$\int_a^b \rho(x)h(x)\mathrm{d}x=0,$$

则在(a,b)内，$h(x)\equiv 0$.

容易验证，(6.3-1)式定义的$\langle\cdot,\cdot\rangle$是内积，从而 $C[a,b]$成为欧氏空间，仍记为 $C[a,b]$. ∎

常用的权函数有

1) $\rho(x)=1, x\in[a,b]$;

2) $\rho(x)=\mathrm{e}^{-x}, x\in[0,+\infty)$;

3) $\rho(x)=\mathrm{e}^{-x^2}, x\in(-\infty,+\infty)$;

4) $\rho(x)=\dfrac{1}{\sqrt{1-x^2}}, x\in(-1,1)$.

在内积空间中可以定义元素的长度，元素之间的正交性，标准正交基等. 由于与第一章论述向量空间时所述的概念类同，所以这里就不一一论述了.

6.3 内积空间及两类特殊的线性变换

下面着重讨论酉空间的两类特殊线性变换. 对于欧氏空间而言，只需做适当修改便有类似的结论.

定义 6.3-2 设 $\mathscr{T}$ 是酉空间 U 的线性变换，如果它保持 U 中任何两个元素的内积不变，即 $\forall \boldsymbol{\alpha},\boldsymbol{\beta}\in U$，恒有

$$\langle \mathscr{T}(\boldsymbol{\alpha}),\mathscr{T}(\boldsymbol{\beta})\rangle = \langle \boldsymbol{\alpha},\boldsymbol{\beta}\rangle, \tag{6.3-2}$$

则称 $\mathscr{T}$ 为**酉变换**(在欧氏空间中，满足(6.3-2)式的线性变换 $\mathscr{T}$ 称为正交变换).

定理 6.3-1 设 $\mathscr{T}$ 是 n 维酉空间 U^n 的线性变换，则 $\mathscr{T}$ 为酉变换当且仅当下列等价条件之一成立：

1) $\mathscr{T}$ 保持元素的长度不变；

2) $\mathscr{T}$ 把一个标准正交基变为一个标准正交基；

3) $\mathscr{T}$ 在任一标准正交基下的矩阵 $\boldsymbol{Q}$ 是酉矩阵，即

$$\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{Q}^{\mathrm{H}} = \boldsymbol{I}.$$

证 1)的证明比较简单，请读者自证.

2) 设 $\{\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n\}$ 是 U^n 的一个标准正交基.

若 $\mathscr{T}$ 是酉变换，则有

$$\langle \mathscr{T}(\boldsymbol{\varepsilon}_i),\mathscr{T}(\boldsymbol{\varepsilon}_j)\rangle = \langle \boldsymbol{\varepsilon}_i,\boldsymbol{\varepsilon}_j\rangle = \delta_{ij} = \begin{cases} 1, i = j, \\ 0, i \neq j. \end{cases}$$

因此 $\{\mathscr{T}(\boldsymbol{\varepsilon}_1),\mathscr{T}(\boldsymbol{\varepsilon}_2),\cdots,\mathscr{T}(\boldsymbol{\varepsilon}_n)\}$ 是标准正交基.

6.3　内积空间及两类特殊的线性变换

反之，设 $\boldsymbol{\alpha}=\sum_{i=1}^{n}x_i\boldsymbol{\varepsilon}_i,\boldsymbol{\beta}=\sum_{j=1}^{n}y_j\boldsymbol{\varepsilon}_j$ 是 U^n 中任意两个元素，则有

$$\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=\left\langle\sum_{i=1}^{n}x_i\boldsymbol{\varepsilon}_i,\sum_{j=1}^{n}y_j\boldsymbol{\varepsilon}_j\right\rangle=\sum_{i=1}^{n}x_i\left\langle\boldsymbol{\varepsilon}_i,\sum_{j=1}^{n}y_j\boldsymbol{\varepsilon}_j\right\rangle=\sum_{i=1}^{n}\sum_{j=1}^{n}x_i\overline{y_j}\langle\boldsymbol{\varepsilon}_i,\boldsymbol{\varepsilon}_j\rangle,$$

$$\langle\mathscr{T}(\boldsymbol{\alpha}),\mathscr{T}(\boldsymbol{\beta})\rangle=\left\langle\mathscr{T}\left(\sum_{i=1}^{n}x_i\boldsymbol{\varepsilon}_i\right),\mathscr{T}\left(\sum_{j=1}^{n}y_j\boldsymbol{\varepsilon}_j\right)\right\rangle=\left\langle\sum_{i=1}^{n}x_i\mathscr{T}(\boldsymbol{\varepsilon}_i),\sum_{j=1}^{n}y_j\mathscr{T}(\boldsymbol{\varepsilon}_j)\right\rangle$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n}x_i\overline{y_j}\langle\mathscr{T}(\boldsymbol{\varepsilon}_i),\mathscr{T}(\boldsymbol{\varepsilon}_j)\rangle.$$

如果 $\mathscr{T}$ 把标准正交基变为标准正交基，则

$$\langle\mathscr{T}(\boldsymbol{\varepsilon}_i),\mathscr{T}(\boldsymbol{\varepsilon}_j)\rangle=\delta_{ij}=\langle\boldsymbol{\varepsilon}_i,\boldsymbol{\varepsilon}_j\rangle,$$

从而 $\langle\mathscr{T}(\boldsymbol{\alpha}),\mathscr{T}(\boldsymbol{\beta})\rangle=\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle$，即 $\mathscr{T}$ 为酉变换.

3) 设 $\{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\}$ 是标准正交基，$\mathscr{T}$ 在这个基下的矩阵是 $\boldsymbol{Q}=[q_{ij}]$，则

$$\mathscr{T}(\boldsymbol{\varepsilon}_j)=\sum_{k=1}^{n}q_{kj}\boldsymbol{\varepsilon}_k,\quad j=1,2,\cdots,n.$$

若 $\mathscr{T}$ 是酉变换，则 $\delta_{ij}=\langle\mathscr{T}(\boldsymbol{\varepsilon}_i),\mathscr{T}(\boldsymbol{\varepsilon}_j)\rangle=\left\langle\sum_{k=1}^{n}q_{ki}\boldsymbol{\varepsilon}_k,\sum_{l=1}^{n}q_{lj}\boldsymbol{\varepsilon}_l\right\rangle$

$$=\sum_{k=1}^{n}\sum_{l=1}^{n}q_{ki}\overline{q_{lj}}\langle\boldsymbol{\varepsilon}_k,\boldsymbol{\varepsilon}_l\rangle=\sum_{k=1}^{n}q_{ki}\overline{q_{kj}},\qquad(6.3\text{-}3)$$

6.3 内积空间及两类特殊的线性变换

即 $Q^{\mathrm{T}}\overline{Q}$ 的第 i 行第 j 列处的元素为 δ_{ij}，故 $Q^{\mathrm{T}}\overline{Q}=I$. 两边取共轭即得

$$Q^{\mathrm{H}}Q=I.$$

由于 $Q^{\mathrm{H}}=Q^{-1}$，所以 $\qquad QQ^{\mathrm{H}}=QQ^{-1}=I.$

反之，若 Q 是酉矩阵，则由(6.3-3)式可知 $\langle\mathscr{T}(\boldsymbol{\varepsilon}_i),\mathscr{T}(\boldsymbol{\varepsilon}_j)\rangle=\delta_{ij}$，即 $\mathscr{T}$ 把标准正交基变为标准正交基，故 $\mathscr{T}$ 是酉变换. ∎

定义 6.3-3 设 $\mathscr{T}$ 是酉空间 U 的线性变换，如果对 U 中任何两个元素 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$，均有

$$\langle\mathscr{T}(\boldsymbol{\alpha}),\boldsymbol{\beta}\rangle=\langle\boldsymbol{\alpha},\mathscr{T}(\boldsymbol{\beta})\rangle, \tag{6.3-4}$$

则称 $\mathscr{T}$ 为**酉对称变换**(在欧氏空间中，满足(6.3-4)式的线性变换 $\mathscr{T}$ 称为对称变换).

定理 6.3-2 设 $\mathscr{T}$ 是 n 维酉空间 U^n 的线性变换，则 $\mathscr{T}$ 为酉对称变换的充要条件是，$\mathscr{T}$ 在标准正交基下的矩阵是 Hermite 矩阵.

证 设 $\{\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n\}$ 是 U^n 的标准正交基，$\mathscr{T}$ 在这个基下的矩阵是 $\boldsymbol{A}=[a_{ij}]$，则

$$\mathscr{T}(\boldsymbol{\varepsilon}_j)=\sum_{k=1}^{n}a_{kj}\boldsymbol{\varepsilon}_k,\quad j=1,2,\cdots,n.$$

若 $\mathscr{T}$ 是酉对称变换，则由 $\langle\mathscr{T}(\boldsymbol{\varepsilon}_i),\boldsymbol{\varepsilon}_j\rangle=\langle\boldsymbol{\varepsilon}_i,\mathscr{T}(\boldsymbol{\varepsilon}_j)\rangle$ 得

$$\left\langle\sum_{k=1}^{n}a_{ki}\boldsymbol{\varepsilon}_k,\boldsymbol{\varepsilon}_j\right\rangle=\left\langle\boldsymbol{\varepsilon}_i,\sum_{k=1}^{n}a_{kj}\boldsymbol{\varepsilon}_k\right\rangle,$$

而上式的左边为 a_{ji}，右边为 $\overline{a_{ij}}$，因而有 $a_{ji}=\overline{a_{ij}}$，即 $\boldsymbol{A}^{\mathrm{T}}=\overline{\boldsymbol{A}}$，亦即 $\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$.

反之，若 $\boldsymbol{A}$ 为 Hermite 矩阵，则可逆推出

$$\langle\mathscr{T}(\boldsymbol{\varepsilon}_i),\boldsymbol{\varepsilon}_j\rangle=\langle\boldsymbol{\varepsilon}_i,\mathscr{T}(\boldsymbol{\varepsilon}_j)\rangle.$$

6.3 内积空间及两类特殊的线性变换

从而对于 U^n 中任何两个元素 $\boldsymbol{\alpha}=\sum_{i=1}^{n}x_i\boldsymbol{\varepsilon}_i,\boldsymbol{\beta}=\sum_{j=1}^{n}y_j\boldsymbol{\varepsilon}_j$，有

$$\langle\mathscr{T}(\boldsymbol{\alpha}),\boldsymbol{\beta}\rangle=\langle\sum_{i=1}^{n}x_i\mathscr{T}(\boldsymbol{\varepsilon}_i),\sum_{j=1}^{n}y_j\boldsymbol{\varepsilon}_j\rangle=\sum_{i=1}^{n}\sum_{j=1}^{n}x_i\overline{y_j}\langle\mathscr{T}(\boldsymbol{\varepsilon}_i),\boldsymbol{\varepsilon}_j\rangle$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n}x_i\overline{y_j}\langle\boldsymbol{\varepsilon}_i,\mathscr{T}(\boldsymbol{\varepsilon}_j)\rangle=\langle\sum_{i=1}^{n}x_i\boldsymbol{\varepsilon}_i,\sum_{j=1}^{n}y_j\mathscr{T}(\boldsymbol{\varepsilon}_j)\rangle=\langle\boldsymbol{\alpha},\mathscr{T}(\boldsymbol{\beta})\rangle,$$

故 $\mathscr{T}$ 是酉对称变换. ■

定理 6.3-3 设 $\mathscr{T}$ 是酉空间 U^n 的酉对称变换，则

1) $\mathscr{T}$ 的特征值必为实数；

2) $\mathscr{T}$ 的属于不同特征值的特征向量必正交；

3) $\mathscr{T}$ 必可酉对角化（即存在标准正交基使 $\mathscr{T}$ 在这个基下为对角矩阵）.

证 1) 设 λ 是酉对称变换 $\mathscr{T}$ 的任一特征值，$\boldsymbol{\xi}\in U^n$ 是对应于 λ 的特征向量，即 $\mathscr{T}(\boldsymbol{\xi})=\lambda\boldsymbol{\xi}$，且 $\boldsymbol{\xi}\neq\mathbf{0}$，那么

$$\lambda\langle\boldsymbol{\xi},\boldsymbol{\xi}\rangle=\langle\lambda\boldsymbol{\xi},\boldsymbol{\xi}\rangle=\langle\mathscr{T}(\boldsymbol{\xi}),\boldsymbol{\xi}\rangle$$

$$=\langle\boldsymbol{\xi},\mathscr{T}(\boldsymbol{\xi})\rangle=\langle\boldsymbol{\xi},\lambda\boldsymbol{\xi}\rangle=\bar{\lambda}\langle\boldsymbol{\xi},\boldsymbol{\xi}\rangle.$$

由于 $\boldsymbol{\xi}\neq\mathbf{0}$，故 $\langle\boldsymbol{\xi},\boldsymbol{\xi}\rangle>0$，从而 $\lambda=\bar{\lambda}$，即 λ 是实数.

6.3 内积空间及两类特殊的线性变换

2）设 λ、μ 是 $\mathscr{T}$ 的两个不同特征值，对应于 λ、μ 的特征向量分别为 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$，则由(1)知 λ、μ 均为实数，故

$$\lambda\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=\langle\lambda\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=\langle\mathscr{T}(\boldsymbol{\alpha}),\boldsymbol{\beta}\rangle=\langle\boldsymbol{\alpha},\mathscr{T}(\boldsymbol{\beta})\rangle=\langle\boldsymbol{\alpha},\mu\boldsymbol{\beta}\rangle=\overline{\mu}\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=\mu\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle.$$

由于 $\lambda\neq\mu$，所以 $\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=0$，即 $\boldsymbol{\alpha}\perp\boldsymbol{\beta}$.

3）对酉空间的维数用数学归纳法. 当 $n=1$ 时，定理显然成立. 现假设维数为 $n-1$ 时定理已成立，要证维数为 n 时定理也成立.

令 $\boldsymbol{\varepsilon}_1$ 是 $\mathscr{T}$ 的属于特征值 λ_1 的单位特征向量，记 $W_1=\mathrm{Span}(\boldsymbol{\varepsilon}_1)$，$W_1$ 显然是 $\mathscr{T}$ 的不变子空间. 于是

$$U^n=W_1\oplus W_1^{\perp},$$

其中 $W_1^{\perp}$ 表示 W_1 的正交补空间. 对任意 $\boldsymbol{\alpha}\in W_1$ 和任意 $\boldsymbol{\beta}\in W_1^{\perp}$，由于

$$\langle\boldsymbol{\alpha},\mathscr{T}(\boldsymbol{\beta})\rangle=\langle\mathscr{T}(\boldsymbol{\alpha}),\boldsymbol{\beta}\rangle=\langle\lambda_1\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=\lambda_1\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=0,$$

所以 $\mathscr{T}(\boldsymbol{\beta})\perp\boldsymbol{\alpha}$，故 $\mathscr{T}(\beta)\in W_1^{\perp}$. 因此 $W_1^{\perp}$ 是 $\mathscr{T}$ 的不变子空间.

如果把 $\mathscr{T}$ 限制作用于 $W_1^{\perp}$，那么它是 $n-1$ 维酉空间 $W_1^{\perp}$ 的一个酉对称变换，记为 $\mathscr{T}_{W_1^{\perp}}$. 根据归纳法假设，在 $W_1^{\perp}$ 中存在标准正交基 $\{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\}$，使限制在 $W_1^{\perp}$ 的酉对称变换 $\mathscr{T}_{W_1^{\perp}}$ 在这个基下的矩阵是 $n-1$ 阶对角矩阵 $\mathrm{diag}(\lambda_2,\lambda_3,\cdots,\lambda_n)$. 因此，若取 $\{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\}$ 为 U^n 的标准正交基，则 $\mathscr{T}$ 在该基下的矩阵是 n 阶对角矩阵 $\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$. ∎

作业

习题 6.1

6. 已知 $\mathbf{R}^3$ 的两个基

$$\mathscr{B}_1=\left\{\begin{bmatrix}1\\-1\\0\end{bmatrix},\begin{bmatrix}0\\1\\-1\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}\right\},\quad \mathscr{B}_2=\left\{\begin{bmatrix}1\\-1\\1\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix},\begin{bmatrix}1\\0\\1\end{bmatrix}\right\}.$$

（1）求 $\mathscr{B}_1$ 到 $\mathscr{B}_2$ 的基变换矩阵 $\boldsymbol{P}$；

（2）求在 $\mathscr{B}_1$、$\mathscr{B}_2$ 下有相同坐标的所有向量.

习题 6.2

4. 已知 $\mathbf{R}^3$ 的线性变换 $\mathscr{T}$ 在基 $\mathscr{B}_1=\left\{\begin{bmatrix}-1\\1\\1\end{bmatrix},\begin{bmatrix}1\\0\\-1\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix}\right\}$

下的矩阵是

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

求 $\mathscr{T}$ 在基 $\mathscr{B}_2=\left\{\begin{bmatrix}1\\-1\\1\end{bmatrix},\begin{bmatrix}1\\0\\1\end{bmatrix},\begin{bmatrix}1\\1\\2\end{bmatrix}\right\}$ 下的矩阵 $\boldsymbol{B}$.

謝謝觀看！

厦門大學
XIAMEN UNIVERSITY

信息学院
（国家示范性软件学院）
School of Informatics

黄 煒
博士，副教授
Dr. Wei Huang