第 1 章　线性空间与线性变换

本章介绍线性空间、线性子空间、线性变换及其矩阵表示、欧氏空间及酉空间的有关概念和结论。这些内容既是线性代数的延伸,也是学习矩阵理论的基础。

1.1　线性空间

1.1.1　线性空间的概念及实例

线性空间也称向量空间,是一个重要的基本概念。粗略地说,线性空间是定义了加法与数乘的一个集合,集合中的元素(简称元)经过这两种运算所得的结果仍属于这个集合,称此集合在这两种运算下封闭。这两种运算满足规定的运算法则,数乘所用的数是取自于定义了和、差、积、商的数集,该数集称为数域。

定义 1.1.1　设 F 是包含 0 和 1 在内的数集,若 F 对于数的加、减、乘、除都封闭,即 F 中任意两个数的和、差、积、商(除数不为零)仍在 F 中,则称 F 是一个数域。

据此,由全体有理数构成的集合、由全体实数构成的集合、由全体复数构成的集合都是数域,分别称为有理数域、实数域、复数域,依次用 $\mathbf{Q}$、$\mathbf{R}$、$\mathbf{C}$ 表示。

本书只涉及实数域 $\mathbf{R}$ 和复数域 $\mathbf{C}$,一般用数域 F 统称。

定义 1.1.2　设 V 为非空集合,F 为一数域(实数域或复数域)。在 V 中定义了加法,即给定一个法则:V 中的任意两个元素 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 有唯一的一个同在 V 中的元素 $\boldsymbol{\gamma}$ 与它们对应,记为 $\boldsymbol{\gamma}=\boldsymbol{\alpha}+\boldsymbol{\beta}$,称为 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的和。在数域 F 与集合 V 的元素之间定义数乘,即给定另一法则:F 中任一数 k 与 V 中任一元素 $\boldsymbol{\alpha}$ 有唯一的一个在 V 中的元素 $\boldsymbol{\delta}$ 与它们对应,记为 $\boldsymbol{\delta}=k\boldsymbol{\alpha}=\boldsymbol{\alpha}k$,称为 k 与 $\boldsymbol{\alpha}$ 的数乘。如果加法和数乘满足下列规则,则称 V 是数域 F 上的线性空间(或 F 上的向量空间)。

加法满足下列 4 条规则:

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

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

(3) 在 V 中有零元素 $\boldsymbol{0}$,使对 V 中任一元素 $\boldsymbol{\alpha}$,都有 $\boldsymbol{\alpha}+\boldsymbol{0}=\boldsymbol{\alpha}$;

(4) 对 V 中每一元素 $\boldsymbol{\alpha}$,都存在一个在 V 中的负元素 $\boldsymbol{\beta}$,使 $\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{0}$。$\boldsymbol{\beta}$ 称为 $\boldsymbol{\alpha}$ 的负元素,记为 $\boldsymbol{\beta}=-\boldsymbol{\alpha}$。

数乘满足下列 4 条规则:

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

(6) 数乘的结合律　　$k(l\boldsymbol{\alpha})=(kl)\boldsymbol{\alpha}$;

(7) 数因子分配律　　$(k+l)\boldsymbol{\alpha}=k\boldsymbol{\alpha}+l\boldsymbol{\alpha}$;

(8) 分配律　　$k(\boldsymbol{\alpha}+\boldsymbol{\beta})=k\boldsymbol{\alpha}+k\boldsymbol{\beta}$。

其中 k、l 为 F 中的任意数,$\boldsymbol{\alpha}$、$\boldsymbol{\beta}$、$\boldsymbol{\gamma}$ 为 V 中的任意元素。

为强调数域 F 的影响,有时将线性空间记为 $V(F)$,其中元素概称为向量,F 中的元素概称为数量或纯量。当数域 F 取实数域 $\mathbf{R}$ 或复数域 $\mathbf{C}$ 时,分别称 V 为实线性空间或复线性空间。

在本书中的符号,除特别声明外,我们约定用“主要符号说明”表中指定的符号。并约定,不加说明时,讨论问题中向量均为列向量。

【例 1.1.1】　分量取自数域 F 上的 n 元向量全体组成的集合 $F^n=\{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\cdots\}$ 按照通常的向量加法和数与向量的数乘,构成 F 上的线性空间(或向量空间)。当 F 为复数域 $\mathbf{C}$ 时,称为 n 元复向量空

间,记成 C^n;当 F 为实数域 $\mathbf{R}$ 时,称为实向量空间,记为 R^n。

事实上,在 F^n 中任取两个向量 $\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$ 和 $\boldsymbol{\beta}=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}$,$k$ 为 F 中任意数,则有

$$\boldsymbol{\alpha}+\boldsymbol{\beta}=(a_1+b_1,a_2+b_2,\cdots,a_n+b_n)^{\mathrm{T}}\in F^n;$$

$$k\boldsymbol{\alpha}=(ka_1,ka_2,\cdots,ka_n)^{\mathrm{T}}\in F^n$$

且易证满足定义 1.1.2 中对加法和数乘两种运算的 8 条运算规则。

【例 1.1.2】 数域 F 上 $m\times n$ 矩阵的全体组成的集合 $F^{m\times n}=\{\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\cdots\}$ 按照矩阵的加法和 F 中数与矩阵的乘法构成 F 上的向量空间,称为矩阵空间。

若把 $m\times n$ 矩阵看做 $m\times n$ 维向量,由例 1.1.1 很容易说明 $F^{m\times n}$ 构成 F 上的向量空间。

【例 1.1.3】 数域 F 上次数小于 n 的一元多项式全体,包括零多项式所组成的集合 $F[x]_n=\{f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}\mid a_i\in F,i=0,1,\cdots,n-1\}$ 按照多项式的加法和 F 中数与多项式的乘法构成 F 上的线性空间,称为多项式空间。

【例 1.1.4】 实函数的全体组成的集合,按照函数的加法和实数与函数的乘法构成实数域 $\mathbf{R}$ 上的线性空间,称为函数空间。

【例 1.1.5】 设 $\boldsymbol{A}$ 为复数域上的 $m\times n$ 矩阵,根据齐次线性方程组解的性质,易证齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的包括零解的所有解的集合构成 $\mathbf{C}$ 上的线性空间,称为方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解空间,也称矩阵 $\boldsymbol{A}$ 的核空间或零空间,常记为 $N(\boldsymbol{A})$。

【例 1.1.6】 设 $\boldsymbol{A}\in C^{m\times n}$,$\boldsymbol{x}\in C^n$,则由 $\boldsymbol{Ax}$ 确定的所有 m 维向量的全体组成的集合

$$V=\{\boldsymbol{y}\mid \boldsymbol{y}=\boldsymbol{Ax}\}$$

构成复数域 $\mathbf{C}$ 上的线性空间,称为 $\boldsymbol{A}$ 的列空间或值空间或 $\boldsymbol{A}$ 的值域,常记为 $R(\boldsymbol{A})$。

【例 1.1.7】 设 $\mathbf{R}$ 是实数域,R^+ 是正实数全体组成的集合。在 R^+ 中元素加法"$\boxplus$"的定义和 R 中的数与 R^+ 的元素间的数乘"$\circ$"的定义为

$$a\boxplus b=ab\qquad k\circ a=a^k$$

其中 $a,b\in R^+$,$k\in\mathbf{R}$,试证明 R^+ 按照上述定义运算"$\boxplus$"与"$\circ$"构成 $\mathbf{R}$ 上的线性空间。

证 首先,由于任意 $a,b\in R^+$,任意 $k\in R$,有

$$a\boxplus b=ab\in R^+,\quad k\circ a=a^k\in R^+$$

即按照所定义的"$\boxplus$",和"$\circ$",对 R^+ 是封闭的。

其次,对所定义的加法和数乘运算满足 8 条运算规则:

任取 $a,b,c\in R^+$,任取 $k,l\in\mathbf{R}$,则

对于加法"$\boxplus$"有

(1) $a\boxplus b=ab=ba=b\boxplus a$ (交换律)

(2) $(a\boxplus b)\boxplus c=(ab)\boxplus c=(ab)c=a(bc)$
$=a\boxplus(bc)=a\boxplus(b\boxplus c)$ (结合律)

(3) $a\boxplus 1=a1=a$ (1 是 R^+ 的零元)

(4) $a\boxplus\dfrac{1}{a}=1$ (任意元的负元是它的倒数)

对于数乘"$\circ$"有

(5) $1\circ a=a^1=a$, $0\circ a=a^0=1$

(6) $k\circ(l\circ a)=k\circ a^l=a^{kl}=(kl)\circ a$ (结合律)

(7) $(k+l)\circ a=a^{(k+l)}=a^ka^l=a^k\boxplus a^l$
$=k\circ a\boxplus l\circ a$ (左分配律)

(8) $k\circ(a\boxplus b)=k\circ(ab)=(ab)^k=a^kb^k$

$$=a^k \boxplus b^k = k \circ a \boxplus k \circ b \qquad \text{（右分配律）}$$

由此，证明了 R^+ 按照所定义的两种运算："$\boxplus$"与"$\circ$"构成 R 上线性空间。

注意：若集合 V 按照所定义的加法与数乘运算不满足封闭性，或者不满足 8 条运算规则中任何一条，则该集合不能构成线性空间。如下面的两个例子。

【例 1.1.8】　次数为 n 的实系数多项式全体组成的集合，按照通常多项式的加法与数乘，不构成实线性空间。因为，按照多项式加法，不满足封闭性。

【例 1.1.9】　非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$，所有解向量的全体所组成的集合，不构成线性空间，因为该集合对加法与数乘均不封闭。

由定义不难推出线性空间的基本性质：

性质 1　在线性空间 V 中，零元素是唯一的，任何一个元素的负元素是唯一的。

性质 2　在线性空间 V 中，对任意 $\boldsymbol{\alpha}\in V, k\in F$，下列关系式成立：

$$0\cdot\boldsymbol{\alpha}=\boldsymbol{0};\qquad (-1)\boldsymbol{\alpha}=-\boldsymbol{\alpha};\qquad k\cdot\boldsymbol{0}=\boldsymbol{0}$$

性质 3　在线性空间 V 中，若 $k\boldsymbol{\alpha}=\boldsymbol{0}$，则必有 $k=0$ 或 $\boldsymbol{\alpha}=\boldsymbol{0}$。

综合以上性质，我们由 $k\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{0}$ 在 $k\neq 0$ 时，可推出 $\boldsymbol{\alpha}=-\dfrac{1}{k}\boldsymbol{\beta}$。

线性空间的元素也称向量，自然，这里所谓向量比几何中所谓向量的含义要广泛得多。

1.1.2　基、维数与坐标

在线性空间 V 中，可以依照 n 元有序数组组成的向量空间，引入线性组合、线性表出、线性相关性、极大无关组、基、坐标等概念及其结论。

（1）线性组合与线性表出

设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 是线性空间 V 中 m 个向量，且 $k_1,k_2,\cdots,k_m$ 是数域 F 中任意 m 个数，则向量

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

称为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的一个线性组合，若向量 $\boldsymbol{\beta}$ 有

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

则称 $\boldsymbol{\beta}$ 可以由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性表出，也可称 $\boldsymbol{\beta}$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的线性组合。

（2）线性组合的线性组合

线性空间 V 中若干向量的线性组合的线性组合还是这若干向量的线性组合。

（3）线性相关与线性无关

在线性空间 V 中对于向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$，若有不全为零的 m 个数 $k_1,k_2,\cdots,k_m$ 存在，使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m=\boldsymbol{0}$$

成立，则称向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关。若只有全为零的数，即仅当 $k_1=k_2=\cdots=k_m=0$ 才使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m=\boldsymbol{0}$$

成立，则称向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关。

在线性无关的向量组中，没有任何向量是其余向量的线性组合。

（4）线性表出的唯一性

在线性空间 V 中设向量 $\boldsymbol{\beta}$ 可由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性表出，即

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

则线性表出系数唯一确定的充分必要条件是向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关。

（5）极大线性无关组

在线性空间 V 中，如果向量组 $\boldsymbol{A}$ 的一个部分组 $\boldsymbol{B}$ 满足：

① 线性无关;

② 向量组 A 中每一个向量都可由部分组 B 线性表出。

则称部分组 B 是向量组 A 的一个极大线性无关组,简称极大无关组。

(6) 向量组的秩

线性空间 V 中一个向量组的极大无关组不是唯一的,但是任意一个极大无关组所含向量个数是相等的,此时极大无关组所含向量的个数称为向量组的秩。

由此可见,线性代数中许多概念值得回忆,这是因为其中相关的大部分内容,大都可以移植到线性空间中,而其本身也是抽象空间的一种具体化。学习数学常有一种方式,即学会把具体的东西抽象化,且能把抽象的东西具体化。在这里线性代数的内容与线性空间中的概念,正提供了这种互相转化的契机。

定义 1.1.3 若 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 是数域 F 上线性空间 V 的 n 个线性无关的向量,且 V 中任意向量 $\boldsymbol{\alpha}$,都可以由 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 线性表出,即

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

则称 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 为 V 的一个基或一组基底,有序数组 $(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$ 为 $\boldsymbol{\alpha}$ 在基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 下的坐标。这里 $(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$ 是被向量 $\boldsymbol{\alpha}$ 与基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 唯一确定的。此时,基中所含基向量的个数称为 V 的维数,并且称 V 为 n 维线性空间,记为 $\dim V=n$。

我们特别约定,当 $V=\{\boldsymbol{0}\}$ 时,有 $\dim V=0$,此外,如例 1.1.4 中函数空间则是无穷维线性空间。本书主要讨论有限维线性空间。

那么 V 中任意一组基所含向量的个数是否相等呢?结论是肯定的,这可以用类似线性代数中的替换定理得以证明。因此 V 的维数是唯一确定的。

【例 1.1.10】 写出实数域 $\mathbf{R}$ 上矩阵空间 $R^{2\times 3}$ 的一组基,求 $\dim R^{2\times 3}$,并求

$$\boldsymbol{A}=\begin{bmatrix}-2 & 6 & 1\\ 0 & -1 & 3\end{bmatrix}$$

在此组基下的坐标。

解 在 $R^{2\times 3}$ 中,设有 2×3 阶矩阵 $\boldsymbol{E}_{ij}$,$\boldsymbol{E}_{ij}$ 中的位于 (i,j) 的元素为 1,其余元素为 0。例如 $\boldsymbol{E}_{12}=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix}$,易证:$\boldsymbol{E}_{11},\boldsymbol{E}_{12},\boldsymbol{E}_{13},\boldsymbol{E}_{21},\boldsymbol{E}_{22},\boldsymbol{E}_{23}$ 是 $R^{2\times 3}$ 的一组基。因为它们线性无关,且任何矩阵 $\boldsymbol{B}\in R^{2\times 3}$ 均可由它们线性表出。

由于基中含有 6 个向量,所以 $\dim R^{2\times 3}=6$,即 $R^{2\times 3}$ 是 6 维线性空间。

又由于 $\boldsymbol{A}=-2\boldsymbol{E}_{11}+6\boldsymbol{E}_{12}+\boldsymbol{E}_{13}+0\boldsymbol{E}_{21}-\boldsymbol{E}_{22}+3\boldsymbol{E}_{23}$,故 $\boldsymbol{A}$ 在上述基下的坐标为 $(-2,6,1,0,-1,3)^{\mathrm{T}}$。

注意:与 R^n 及 C^n 一样,一般线性空间的基也不是唯一的,且同一向量在不同基下的坐标也是不同的。

【例 1.1.11】 设 $R[x]_n=\{$次数小于 n 的变量 x 的实系数多项式 $f(x)=a_0+a_1x+a_2x^2+\cdots+a_{n-1}x^{n-1}$ 全体加上零多项式组成的空间$\}$。试证:

(1) $1,x,x^2,\cdots,x^{n-1}$ 是 $R[x]_n$ 的一组基;

(2) $1,x-a,(x-a)^2,\cdots,(x-a)^{n-1}$ 也是 $R[x]_n$ 的一组基,并求 $f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}$ 在这组基下的坐标。

证 (1) 易证 $R[x]_n$ 中一组向量

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

是线性无关的。

事实上,设有数 $k_0,k_1,k_2,\cdots,k_{n-1}$,使得

$$k_0 \cdot 1 + k_1 \cdot x + k_2 \cdot x^2 + \cdots + k_{n-1} \cdot x^{n-1} = 0$$

要使上式对变量 x 的任何情况都成立的充要条件是

$$k_0 = k_1 = k_2 = \cdots = k_{n-1} = 0$$

所以,向量组 $1, x, x^2, \cdots, x^{n-1}$ 是线性无关的。

对于 $R[x]_n$ 中任一向量 $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1}$ 均可由 $1, x, x^2, \cdots, x^{n-1}$ 线性表出。由定义 1.1.3,此时已证明了向量组 $1, x, x^2, \cdots, x^{n-1}$ 是 $R[x]_n$ 的一组基。

(2) 类似可证明 $R[x]_n$ 中的向量组

$$1, x-a, (x-a)^2, \cdots, (x-a)^{n-1}$$

也是 $R[x]_n$ 中线性无关向量组,且 $R[x]_n$ 中的任一向量

$$f(x) = a_0 + a_1 x + a_2 x^2 \cdots + a_{n-1} x^{n-1}$$

把 $f(x)$ 在 $x=a$ 处按 Taylor 公式展开后,有

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{n-1}(a)}{(n-1)!}(x-a)^{n-1}$$

所以,$1, x-a, (x-a)^2, \cdots, (x-a)^{n-1}$ 也是 $R[x]_n$ 的一组基。且 $f(x)$ 在该基下的坐标为

$$\left[f(a), f'(a), \frac{f''(a)}{2!}, \cdots, \frac{f^{(n-1)}(a)}{(n-1)!} \right]^{\mathrm{T}}$$

此例正说明了线性空间中同一向量在不同基下的坐标是不同的。

1.1.3　基变换与坐标变换

下面将研究线性空间中同一向量在不同基下坐标之间的关系。

定义 1.1.4　设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 与 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n$ 是 n 维线性空间的两个基。基向量 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n$ 用基向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 表示的表达式为

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

也可简写成

$$\boldsymbol{\beta}_i = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix} \quad (i = 1, 2, \cdots, n)$$

写成矩阵形式为

$$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{P} \tag{1.1.1}$$

其中

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

称为由基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 到基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n$ 的过渡矩阵。式(1.1.1)为由基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 到基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n$ 的基变换公式。

易证过渡矩阵 P 是可逆的。

下面给出 n 维线性空间 V 中同一向量在两个不同基下坐标的变换公式。

定理 1.1.1 设 n 维线性空间 V 中向量 $\boldsymbol{\alpha}$ 在两个基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 下的坐标分别为 $(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 和 $(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$,由基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 到基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 的过渡矩阵为 P,则这两个坐标之间变换式为

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \boldsymbol{P} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \text{或} \quad \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \boldsymbol{P}^{-1} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{1.1.2}$$

式(1.1.2)称为坐标变换公式。

事实上,由

$$\boldsymbol{\alpha}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \boldsymbol{\alpha}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

有

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

用式(1.1.1)代入上式,得到

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{P}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是基,所以有

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

这就是所求证的在基变换公式(1.1.1)下的坐标变换公式。

【例 1.1.12】 在 $F[x]_4$ 中,求由基 $\boldsymbol{f}_0=1,\boldsymbol{f}_1=x,\boldsymbol{f}_2=x^2,\boldsymbol{f}_3=x^3$,到基 $\boldsymbol{h}_0=1,\boldsymbol{h}_1=1-x,\boldsymbol{h}_2=1-x-x^2,\boldsymbol{h}_3=1-x-x^2-x^3$ 的过渡矩阵。若多项式 $f(x)$ 在基 $\boldsymbol{f}_0,\boldsymbol{f}_1,\boldsymbol{f}_2,\boldsymbol{f}_3$ 下的坐标为 $(1,0,-2,3)^{\mathrm{T}}$,求 $f(x)$ 在基 $\boldsymbol{h}_0,\boldsymbol{h}_1,\boldsymbol{h}_2,\boldsymbol{h}_3$ 下的坐标。

解 由

$$\begin{aligned} &\boldsymbol{h}_0=\boldsymbol{f}_0 \\ &\boldsymbol{h}_1=1-x=\boldsymbol{f}_0-\boldsymbol{f}_1 \\ &\boldsymbol{h}_2=1-x-x^2=\boldsymbol{f}_0-\boldsymbol{f}_1-\boldsymbol{f}_2 \\ &\boldsymbol{h}_3=1-x-x^2-x^3=\boldsymbol{f}_0-\boldsymbol{f}_1-\boldsymbol{f}_2-\boldsymbol{f}_3 \end{aligned}$$

可得

$$(\boldsymbol{h}_0,\boldsymbol{h}_1,\boldsymbol{h}_2,\boldsymbol{h}_3)=(\boldsymbol{f}_0,\boldsymbol{f}_1,\boldsymbol{f}_2,\boldsymbol{f}_3)\begin{bmatrix}1&1&1&1\\0&-1&-1&-1\\0&0&-1&-1\\0&0&0&-1\end{bmatrix}$$

所以,由基 f_0,f_1,f_2,f_3 到基 h_0,h_1,h_2,h_3 的过渡矩阵为

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

设 $f(x)$在基 $\boldsymbol{h}_0,\boldsymbol{h}_1,\boldsymbol{h}_2,\boldsymbol{h}_3$ 下的坐标为$(y_1,y_2,y_3,y_4)^{\mathrm{T}}$,则由式(1.1.2)有

$$\begin{bmatrix}y_1\\y_2\\y_3\\y_4\end{bmatrix}=\boldsymbol{P}^{-1}\begin{bmatrix}1\\0\\-2\\3\end{bmatrix}$$

于是

$$\begin{bmatrix}y_1\\y_2\\y_3\\y_4\end{bmatrix}=\boldsymbol{P}^{-1}\begin{bmatrix}1\\0\\-2\\3\end{bmatrix}=\begin{bmatrix}1&1&1&1\\0&-1&-1&-1\\0&0&-1&-1\\0&0&0&-1\end{bmatrix}^{-1}\begin{bmatrix}1\\0\\-2\\3\end{bmatrix}$$

$$=\begin{bmatrix}1&1&0&0\\0&-1&1&0\\0&0&-1&1\\0&0&0&-1\end{bmatrix}\begin{bmatrix}1\\0\\-2\\3\end{bmatrix}=\begin{bmatrix}1\\-2\\5\\-3\end{bmatrix}$$

即 $f(x)$在基 $\boldsymbol{h}_0,\boldsymbol{h}_1,\boldsymbol{h}_2,\boldsymbol{h}_3$ 下坐标为$(1,-2,5,-3)^{\mathrm{T}}$。

1.2　线性子空间

与 n 元有序数组组成的向量空间类似,在一般 n 维线性空间中也可引进子空间的概念,且有类似的讨论过程与结论。

1.2.1　线性子空间的概念及实例

定义 1.2.1　设 U 是数域 F 上的 n 维线性空间 V 的一个非空子集,若 U 中的任意元素(或向量)$\boldsymbol{\alpha},\boldsymbol{\beta}$ 以及 F 中的任意数 k,对 V 中所定义的加法与数乘两种运算满足

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

(2)　$k\boldsymbol{\alpha}\in U$。

则 U 也构成数域 F 上的一个线性空间,称 U 是线性空间 V 的一个线性子空间,简称子空间。由定义显然可看出任何子空间必含有 $\boldsymbol{0}$ 元(或零向量)。

由于子空间也是线性空间,因此子空间也有基、维数、坐标等内含。又由于子空间是整个线性空间的子集,因此它不可能比整个线性空间含有更多的线性无关向量。所以,V 的任何一个子空间 U 的维数总不会超过 V 的维数,即有

$$\dim U\leqslant\dim V$$

易看出,任一个非零线性空间至少有两个子空间,一个是仅由零向量构成的非空子集,称为零子空间,记做$\{\boldsymbol{0}\}$;另一个是它自身,即V。这两个子空间称为V的平凡子空间,而V中其他线性子空间称为V的非平凡子空间。

由于零子空间中不含线性无关的向量,所以它没有基,因此,零子空间$\{\boldsymbol{0}\}$当且仅当$\dim\{\boldsymbol{0}\}=0$。

下面介绍一个常用的子空间,它也给出了构造线性子空间的一个方法。

【例 1.2.1】 设V是数域F上的线性空间,$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$是$V$中$S$个向量,$k_1,k_2,\cdots,k_s$是$F$中任意一组数,这组向量的所有可能的线性组合

$$U=\{k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s\}$$

是非空的。容易证明U是V的一个子空间,称U是由向量$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$所生成的子空间,记做

$$U=L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s) \tag{1.2.1}$$

显然,若$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$线性无关,则

$$\dim U=\dim L=s$$

对生成的子空间,下面的定理成立。

定理 1.2.1 $\dim L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)=\operatorname{rank}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)$,其中$\operatorname{rank}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)$表示向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$的秩,而$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)$的基可以是向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$中任何一个极大线性无关组。

【例 1.2.2】 在例 1.1.3 中介绍的多项式空间$F[x]_n$中,次数小于$t(t\leqslant n)$的多项式全体,包括零多项式构成子空间U,由于$1,x,x^2,\cdots,x^{t-1}$是该子空间的一个基,所以其生成子空间,即是

$$U=L(1,x,x^2,\cdots,x^{t-1})$$

显然有

$$\dim L(1,x,x^2,\cdots,x^{t-1})=\operatorname{rank}(1,x,x^2,\cdots,x^{t-1})=t$$

【例 1.2.3】 在例 1.1.5 中介绍的$N(\boldsymbol{A})$是矩阵$\boldsymbol{A}$的核子空间(或者零空间),即

$$N(\boldsymbol{A})=\{\boldsymbol{x}\mid\boldsymbol{Ax}=\boldsymbol{0}\}=L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r})$$

其中$\boldsymbol{A}\in R^{m\times n}$,$\operatorname{rank}(\boldsymbol{A})=r$,$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$是$\boldsymbol{Ax}=\boldsymbol{0}$的一个基础解系,$\dim N(\boldsymbol{A})=n-r$,$\dim N(\boldsymbol{A})$称为$\boldsymbol{A}$的零度。

【例 1.2.4】 在例 1.1.6 中矩阵$\boldsymbol{A}$的值域$R(\boldsymbol{A})=\{\boldsymbol{y}=\boldsymbol{Ax}\mid\boldsymbol{x}\in R^n\}$可以看成由$\boldsymbol{A}$的列向量生成的子空间。因为,若$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$为$m\times n$矩阵$\boldsymbol{A}$的列向量,则它们的任意一个线性组合

$$\begin{aligned}&x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n\\&\quad=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\\&\quad=\boldsymbol{A}(x_1,x_2,\cdots,x_n)^{\mathrm{T}}=\boldsymbol{Ax}=\boldsymbol{y}\end{aligned}$$

这说明,所有乘积$\boldsymbol{A}x$的集合

$$\{\boldsymbol{y}=\boldsymbol{Ax}\mid\boldsymbol{x}\in R^n\}$$

与$\boldsymbol{A}$的列向量组的线性组合的集合$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$相同,即

$$R(\boldsymbol{A})=\boldsymbol{L}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$$

若$\operatorname{rank}(\boldsymbol{A})=r$,则

$$\dim R(\boldsymbol{A})=\operatorname{rank}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=\operatorname{rank}(\boldsymbol{A})=r$$

从例 1.2.3 与例 1.2.4 可以看出

$$\dim R(\boldsymbol{A})+\dim N(\boldsymbol{A})=n \tag{1.2.2}$$

其中,$\dim R(\boldsymbol{A})$称为$\boldsymbol{A}$的秩,而$\dim N(\boldsymbol{A})$称为$\boldsymbol{A}$的零度。对于任意$\boldsymbol{A}\in C^{m\times n}$,式(1.2.2)都是成立的。

定理 1.2.2(基的扩充定理) 设$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_S)$是线性空间$V$的一个子空间,则子空间$\boldsymbol{L}$的任何一个基可扩成$V$的一个基。

证明略。

1.2.2　子空间的交与和

子空间除了可以由线性空间中的元素生成以外，还可以由子空间经过集合运算而生成。

定义 1.2.2　设 U_1 和 U_2 是数域 F 上线性空间 V 的子空间，则由 U_1 与 U_2 所有公共元素（向量）组成的集合，称为 U_1 与 U_2 的交空间，记为 $U_1\cap U_2$，即

$$U_1\cap U_2=\{\boldsymbol{\alpha}\mid\boldsymbol{\alpha}\in U_1\text{ 且 }\boldsymbol{\alpha}\in U_2\}$$

定理 1.2.3　数域 F 上线性空间 V 的两个子空间 U_1 与 U_2 的交空间 $U_1\cap U_2$，仍是 V 的子空间。

证　U_1 与 U_2 是 V 的子空间，则 V 的零向量 $\boldsymbol{0}$ 同属于 U_1 和 U_2，即 $\boldsymbol{0}\in U_1\cap U_2$，故 $U_1\cap U_2$ 是非空子集。

任取 $\boldsymbol{\alpha},\boldsymbol{\beta}\in U_1\cap U_2$，则 $\boldsymbol{\alpha},\boldsymbol{\beta}\in U_1$ 且 $\boldsymbol{\alpha},\boldsymbol{\beta}\in U_2$。由于 U_1 和 U_2 是子空间，故 $\boldsymbol{\alpha}+\boldsymbol{\beta}\in U_1$ 且 $\boldsymbol{\alpha}+\boldsymbol{\beta}\in U_2$，所以 $\boldsymbol{\alpha}+\boldsymbol{\beta}\in U_1\cap U_2$。

再任取 $\boldsymbol{\alpha}\in U_1\cap U_2$，$k\in F$。由于 $\boldsymbol{\alpha}\in U_1$ 和 $\boldsymbol{\alpha}\in U_2$ 且 U_1 和 U_2 都是子空间，所以有 $k\boldsymbol{\alpha}\in U_1$ 且 $k\boldsymbol{\alpha}\in U_2$，因此 $k\boldsymbol{\alpha}\in U_1\cap U_2$。于是，$U_1\cap U_2$ 是 V 的子空间。

定义 1.2.3　设 U_1、U_2 是数域 F 上线性空间 V 的子空间，且 $\boldsymbol{\alpha}\in U_1$，$\boldsymbol{\beta}\in U_2$，则所有 $\boldsymbol{\alpha}+\boldsymbol{\beta}$ 这样的元素的集合称为 U_1 与 U_2 的和，记为 U_1+U_2。即

$$U_1+U_2=\{\boldsymbol{\gamma}\mid\boldsymbol{\gamma}=\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\alpha}\in U_1,\boldsymbol{\beta}\in U_2\}$$

定理 1.2.4　若 U_1 与 U_2 是数域 F 上线性空间 V 的两个子空间，则它们的和 U_1+U_2 也是 V 的子空间，称为 U_1 与 U_2 的和空间。

证　显然 U_1+U_2 是 V 的非空子集。因为 $\boldsymbol{0}=\boldsymbol{0}+\boldsymbol{0}$，故 $\boldsymbol{0}\in U_1+U_2$。

对任意向量 $\boldsymbol{\alpha},\boldsymbol{\beta}\in U_1+U_2$，则 $\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$，$\boldsymbol{\beta}=\boldsymbol{\beta}_1+\boldsymbol{\beta}_2$，其中 $\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1\in U_1$；$\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2\in U_2$。因为 $\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1\in U_1$，$\boldsymbol{\alpha}_2+\boldsymbol{\beta}_2\in U_2$，所以

$$\boldsymbol{\alpha}+\boldsymbol{\beta}=(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)+(\boldsymbol{\beta}_1+\boldsymbol{\beta}_2)=(\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1)+(\boldsymbol{\alpha}_2+\boldsymbol{\beta}_2)\in U_1+U_2$$

对任意 $\boldsymbol{\alpha}\in U_1+U_2$，$k\in F$，则 $\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$。其中 $\boldsymbol{\alpha}_1\in U_1$，$\boldsymbol{\alpha}_2\in U_2$。又 $k\boldsymbol{\alpha}_1\in U_1$，$k\boldsymbol{\alpha}_2\in U_2$，所以 $k\boldsymbol{\alpha}=k(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)=k\boldsymbol{\alpha}_1+k\boldsymbol{\alpha}_2\in U_1+U_2$，故 U_1+U_2 是 V 的子空间。

要注意的是，V 的两个子空间 U_1 与 U_2 的并 $U_1\cup U_2$ 一般不再是 V 的子空间。

【例 1.2.5】　在立体几何空间 R^3 中 U_1 与 U_2 分别表示过原点不重合的直线 l_1 与 l_2 上所有向量形成的子空间。显然，$U_1\cap U_2$ 为 l_1 与 l_2 交点（原点）形成的零子空间，而 U_1+U_2 是由 l_1 与 l_2 所确定的平面上全体向量形成的子空间。

定理 1.2.5　设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 是线性空间 V 中两组向量，则有

$$\begin{aligned}&L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)+L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t)\\&\quad=L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t)\end{aligned}\tag{1.2.3}$$

要证式(1.2.3)成立，只需证明式(1.2.3)等号两边的子空间互相包含即可。

事实上，任取 $\boldsymbol{\gamma}\in L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)+L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t)$，则 $\boldsymbol{\gamma}=\boldsymbol{\alpha}+\boldsymbol{\beta}$，其中，$\boldsymbol{\alpha}\in L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)$，$\boldsymbol{\beta}\in L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t)$。由于 $\boldsymbol{\alpha}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出，$\boldsymbol{\beta}$ 可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 线性表出，显然 $\boldsymbol{\alpha}+\boldsymbol{\beta}$ 就可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 线性表出。则有

$$\boldsymbol{\gamma}\in L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t)$$

于是可得

$$\begin{aligned}&L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)+L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t)\\&\quad\subseteq L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t)\end{aligned}$$

类似还可证明

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

请读者自证。

关于两个子空间的交与和的维数,有以下重要结论。

定理 1.2.6 设 U_1 与 U_2 是数域 F 上线性空间 V 的两个子空间,那么有以下公式成立:

$$\dim U_1+\dim U_2=\dim(U_1+U_2)+\dim(U_1\cap U_2) \tag{1.2.4}$$

上式称为维数公式。

证 设 $\dim U_1=s$, $\dim U_2=t$, $\dim(U_1\cap U_2)=r$。

因为 $U_1\cap U_2$ 是 U_1 与 U_2 的子空间,取 $U_1\cap U_2$ 的一组基

$$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$$

由定理 1.2.2 得知,它可扩充成 U_1 的一组基

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

同样也可扩充成 U_2 的一组基

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

这样

$$U_1=L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_{s-r})$$
$$U_2=L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_{t-r})$$

由定理 1.2.5 有

$$U_1+U_2=L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_{s-r},\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_{t-r})$$

现研究向量组

$$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_{s-r},\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_{t-r}$$

的线性相关性。

设有如下等式:

$$k_1\boldsymbol{\alpha}_1+\cdots+k_r\boldsymbol{\alpha}_r+p_1\boldsymbol{\beta}_1+\cdots+p_{s-r}\boldsymbol{\beta}_{s-r}+q_1\boldsymbol{\gamma}_1+\cdots+q_{t-r}\boldsymbol{\gamma}_{t-r}=\boldsymbol{0}$$

则令

$$\begin{aligned}\boldsymbol{\alpha}&=k_1\boldsymbol{\alpha}_1+\cdots+k_r\boldsymbol{\alpha}_r+p_1\boldsymbol{\beta}_1+\cdots+p_{s-r}\boldsymbol{\beta}_{s-r}\\&=-q_1\boldsymbol{\gamma}_1-\cdots-q_{t-r}\boldsymbol{\gamma}_{t-r}\end{aligned} \tag{1.2.5}$$

可见 $\boldsymbol{\alpha}\in U_1$ 且 $\boldsymbol{\alpha}\in U_2$,于是 $\boldsymbol{\alpha}\in U_1\cap U_2$。

设

$$\boldsymbol{\alpha}=l_1\boldsymbol{\alpha}_1+\cdots+l_r\boldsymbol{\alpha}_r \tag{1.2.6}$$

并将其代入式(1.2.5),则有

$$l_1\boldsymbol{\alpha}_1+\cdots+l_r\boldsymbol{\alpha}_r+q_1\boldsymbol{\gamma}_1+\cdots+q_{t-r}\boldsymbol{\gamma}_{t-r}=\boldsymbol{0}$$

由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_{t-r}$ 是 U_2 的基,所以它们线性无关,由此得到

$$l_1=l_2=\cdots=l_r=q_1=q_2=\cdots=q_{t-r}=0$$

由式(1.2.6),$\boldsymbol{\alpha}=\boldsymbol{0}$。这样,根据式(1.2.5)有

$$k_1\boldsymbol{\alpha}_1+\cdots+k_r\boldsymbol{\alpha}_r+p_1\boldsymbol{\beta}_1+\cdots+p_{s-r}\boldsymbol{\beta}_{s-r}=\boldsymbol{0}$$

由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_{s-r}$ 是 U_1 的基,它们也线性无关,又得到

$$k_1=k_2=\cdots=k_r=p_1=p_2=\cdots=p_{s-r}=0$$

所以向量组

$$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{s-r},\boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_{t-r}$$

线性无关，说明这组向量是 U_1+U_2 的一组基。故有

$$\dim(U_1+U_2)=s+t-r$$

所以，式(1.2.4)表达的维数公式成立。

推论　U_1 与 U_2 是数域 F 上的 n 维线性空间 V 的两个子空间，若它们维数之和大于 n，则 U_1 与 U_2 必含有公共非零向量。

证　已知 $\dim U_1+\dim U_2>n$，由维数公式，必有

$$\dim(U_1\cap U_2)=\dim U_1+\dim U_2-\dim(U_1+U_2)>n-\dim(U_1+U_2)\geqslant 0$$

后一个不等式成立是因为 U_1+U_2 是 V 的子空间，所以有 $\dim(U_1+U_2)\leqslant\dim V=n$，所以 $U_1\cap U_2$ 为非零子空间，必含有公共非零向量。

【例 1.2.6】　已知 $\boldsymbol{\alpha}_1=(1,2,1,0)^{\mathrm{T}},\boldsymbol{\alpha}_2=(-1,1,1,1)^{\mathrm{T}},\boldsymbol{\beta}_1=(2,-1,0,1)^{\mathrm{T}},\boldsymbol{\beta}_2=(1,-1,3,7)^{\mathrm{T}}$，$U_1=L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)$，$U_2=L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2)$，求：

(1)　U_1+U_2 的基与维数；

(2)　$U_1\cap U_2$ 的基与维数。

解　(1)　由定理 1.2.5 得

$$U_1+U_2=L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2)$$

易得 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1$ 是向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 的极大无关组，它是 U_1+U_2 的基，故 $\dim(U_1+U_2)=3$。

(2)　因为 $\dim(U_1+U_2)=3$，$\dim U_1=\dim U_2=2$，由维数公式得

$$\dim(U_1\cap U_2)=1$$

下面求 $U_1\cap U_2$ 的基。

设 $\boldsymbol{\alpha}\in U_1\cap U_2$，则

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

即

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2-k_3\boldsymbol{\beta}_1-k_4\boldsymbol{\beta}_2=\boldsymbol{0}$$

此齐次线性方程组的系数矩阵的秩为 3，因此基础解系含一个解向量，可求出一个基础解系为

$$(k_1,k_2,k_3,k_4)^{\mathrm{T}}=(-1,4,-3,1)^{\mathrm{T}}$$

于是

$$\boldsymbol{\gamma}=-\boldsymbol{\alpha}_1+4\boldsymbol{\alpha}_2=-3\boldsymbol{\beta}_1+\boldsymbol{\beta}_2=(-5,2,3,4)^{\mathrm{T}}$$

所以 $U_1\cap U_2$ 的一个基为 $\boldsymbol{\gamma}=(-5,2,3,4)^{\mathrm{T}}$，故

$$U_1\cap U_2=L(\boldsymbol{\gamma}),\qquad \dim(U_1\cap U_2)=1$$

1.2.3　子空间的直和与补子空间

下面介绍子空间直和，它是子空间和的特殊情况。

定义 1.2.4　设 U_1 与 U_2 是数域 F 上线性空间 V 的两个子空间，若 $U_1\cap U_2=\{\boldsymbol{0}\}$，则称 U_1+U_2 为直和，记做 $U_1\oplus U_2$ 或 $U_1\dot{+}U_2$。

【例 1.2.7】　U_1 与 U_2 分别是 R^3 中过原点的直线与平面，且 U_1 不在 U_2 上。显然 $U_1+U_2=R^3$，且 $U_1\cap U_2=\{\boldsymbol{0}\}$，所以 U_1+U_2 是直和，即 $U_1\oplus U_2$。

【例 1.2.8】　U_1 与 U_2 分别是实数域上齐次线性方程组

$$x_1+x_2+\cdots+x_n=0 \tag{1.2.7}$$

$$x_1=x_2=\cdots=x_n \tag{1.2.8}$$

的解空间,求证 $R^n=U_1\oplus U_2$。

事实上,齐次线性方程组(1.2.7)系数矩阵的秩为 1,基础解系含有 $n-1$ 个解向量。可求出它的一个基础解系为

$$\boldsymbol{e}_1=(-1,1,0,\cdots,0)$$
$$\boldsymbol{e}_2=(-1,0,1,\cdots,0)$$
$$\cdots\cdots$$
$$\boldsymbol{e}_{n-1}=(-1,0,0,\cdots,1)$$

又齐次线性方程组(1.2.8)系数矩阵秩为 $n-1$,基础解系含一个解向量,可求出它的一个基础解系为

$$\boldsymbol{e}_n=(1,1,\cdots,1)$$

显然有 $U_1=L(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_{n-1})$,$U_2=L(\boldsymbol{e}_n)$。若设 $\boldsymbol{\alpha}\in U_1\cap U_2$,则存在 $k_1,k_2,\cdots,k_{n-1},k_n$,使

$$\boldsymbol{\alpha}=k_1\boldsymbol{e}_1+k_2\boldsymbol{e}_2+\cdots+k_{n-1}\boldsymbol{e}_{n-1}=k_n\boldsymbol{e}_n$$

由此可得

$$k_1\boldsymbol{e}_1+\cdots+k_{n-1}\boldsymbol{e}_{n-1}-k_n\boldsymbol{e}_n=0$$

易看出 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_{n-1},\boldsymbol{e}_n$ 线性无关。则必有

$$k_1=k_2=\cdots=k_{n-1}=k_n=0$$

即 $\boldsymbol{\alpha}=\boldsymbol{0}$,于是 $U_1\cap U_2=\{\boldsymbol{0}\}$。

所以,$R^n=U_1\oplus U_2$。

关于两个子空间的直和有以下结论。

定理 1.2.7 设 U_1 与 U_2 是数域 F 上线性空间 V 的两个子空间,下面的条件是等价的:

(1) U_1+U_2 是直和;

(2) $\dim(U_1+U_2)=\dim U_1+\dim U_2$;

(3) U_1+U_2 中每一向量 $\boldsymbol{\alpha}$ 的分解式唯一,即

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\quad \boldsymbol{\alpha}_1\in U_1,\quad \boldsymbol{\alpha}_2\in U_2$$

(4) 如 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t$ 是 U_1 的一个基,$\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s$ 是 U_2 的一个基,则 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s$ 是 U_1+U_2 的一个基。

证略。

子空间直和的概念可以推广到多个子空间情况,这里不再重复。

定义 1.2.5 若数域 F 上的 n 维线性空间 V 可表成两个子空间 U_1 与 U_2 的直和,即

$$V=U_1\oplus U_2$$

则称 U_1,U_2 为线性空间 V 的一对互补子空间,并称 V 有一个直和分解。

可以证明,线性空间 V 的任何一个子空间 U_1 一定存在补子空间 U_2 使 $V=U_1\oplus U_2$,但是,一般地,子空间 U_1 的补子空间不是唯一的。如:

设 R^3 的子空间 $U_1=L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)$,其中 $\boldsymbol{\alpha}_1=(0,1,0)$,$\boldsymbol{\alpha}_2=(0,0,1)$,则由 $\boldsymbol{\alpha}_3=(1,0,0)$ 或 $\boldsymbol{\alpha}_4=(1,1,0)$ 构成的子空间 $L(\boldsymbol{\alpha}_3)$ 或 $L(\boldsymbol{\alpha}_4)$ 均是 U_1 的补子空间。

1.3 线性变换

本节主要研究线性空间中元素之间的关系,介绍映射、线性变换的概念、线性变换的矩阵表示及其有关子空间。

1.3.1 线性变换的概念及实例

若不特别提出,下面所考虑的都是固定在某一数域 F 上的线性空间。

定义 1.3.1　设 V_1,V_2 是两个线性空间，若对于某个指定的法则 T，使 V_1 中任意一元素 $\boldsymbol{\alpha}$ 都能与 V_2 中的一个确定元素 $\boldsymbol{\beta}$ 相对应，则 T 称为由 V_1 到 V_2 的一个映射，记为

$$\boldsymbol{\beta}=T(\boldsymbol{\alpha})$$

$\boldsymbol{\beta}$ 称为 $\boldsymbol{\alpha}$ 在映射 T 下的像，而 $\boldsymbol{\alpha}$ 称为 $\boldsymbol{\beta}$ 的原像。

若 T 同时还满足：

对任意的 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\in V$ 有

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

及对任意 $k\in F$，$\boldsymbol{\alpha}\in V$，有

$$T(k\boldsymbol{\alpha})=kT(\boldsymbol{\alpha})$$

则称 T 为从 V_1 到 V_2 的一个线性映射。

定义 1.3.2　若 T 为线性空间 V 到自身的一个映射，则 T 称为 V 的一个变换。若变换 T 对于 V 中任意的元素 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 和数域 F 中的任意数 k，还同时满足

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

则称 T 为线性空间 V 的一个线性变换。

我们可以把定义 1.3.2 中两个条件用一个表达式来表示，即 T 是 V 的线性变换的充分必要条件是

$$T(k\boldsymbol{\alpha}+l\boldsymbol{\beta})=kT(\boldsymbol{\alpha})+lT(\boldsymbol{\beta})$$

其中，$\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 是 V 中任意向量，k 与 l 是数域 F 中任意数。

以后一般用白斜体拉丁字母 $A,B,C,\cdots,T,\cdots$ 代表线性空间 V 的线性变换，$T(\boldsymbol{\alpha})$ 或 $T\boldsymbol{\alpha}$ 代表元素 $\boldsymbol{\alpha}$ 在线性变换 T 下的像。

线性变换的丰富内容可从下面的几个例子显示出来。

【例 1.3.1】　设 $R^{4\times 4}$ 是实数域 $\mathbf{R}$ 上全体 4 阶方阵的集合，$\boldsymbol{A}$ 为 $R^{4\times 4}$ 中任意一个 4 阶方阵

$$T(\boldsymbol{A})=|\boldsymbol{A}|$$

表示 T 是 $R^{4\times 4}$ 到 $\mathbf{R}$ 的一个映射。

【例 1.3.2】　在平面几何空间 R^2 中，把任一向量 $\boldsymbol{x}=(x_1,x_2)^{\mathrm{T}}$ 在平面上绕原点按逆时针方向旋转 θ 角后，得到向量 $\boldsymbol{y}=(y_1,y_2)^{\mathrm{T}}$。它们之间的关系用以下公式表示：

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}=\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

即

$$\boldsymbol{y}=T(\boldsymbol{x})$$

根据定义，易说明 T 是 R^2 的线性变换。

【例 1.3.3】　关于 $F[x]_n$ 的求导变换 D。对任意 $f(x)\in F[x]_n$ 有 $D[f(x)]=f'(x)$，则求导变换 D 是 $F[x]_n$ 上的一个线性变换。

【例 1.3.4】　设 $C[a,b]$ 表示定义在闭区间 $[a,b]$ 上全体实连续函数组成的实数域上的线性空间。在 $C[a,b]$ 中，变换

$$\int[f(t)]=\int_a^t f(x)\mathrm{d}x$$

其中，$f(t)$ 是 $C[a,b]$ 中任意实连续函数。

用定义不难验证积分变换 $\int$ 是 $C[a,b]$ 上的一个线性变换。

【例 1.3.5】　在线性空间 V 中，

(1) 把 V 中的任意向量 $\boldsymbol{\alpha}$ 变成 $\boldsymbol{\alpha}$ 的变换,称为 V 的恒等变换或单位变换,记做 I,即

$$I(\boldsymbol{\alpha})=\boldsymbol{\alpha}$$

(2) 把 V 中的任意向量 $\boldsymbol{\alpha}$ 变成零向量的变换,称为 V 的零变换,记做 T_0,即

$$T_0(\boldsymbol{\alpha})=\boldsymbol{0}$$

(3) 设 k 是 F 中的某个数,把 V 中任意向量 $\boldsymbol{\alpha}$ 变成 $k\boldsymbol{\alpha}$ 的变换,称为 V 的数乘变换,记做 K,即

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

当 $k=1$ 时,得到单位变换 I;当 $k=0$ 时,得到零变换 T_0。

易证明单位变换、零变换和数乘变换都是线性变换。

线性变换有如下简单性质:

若 T 是 V 的线性变换,则

(1) $T(\boldsymbol{0})=\boldsymbol{0}$,　　$T(-\boldsymbol{\alpha})=-T(\boldsymbol{\alpha})$。

(2) 若 $\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m$,则有

$$T(\boldsymbol{\beta})=k_1T(\boldsymbol{\alpha}_1)+k_2T(\boldsymbol{\alpha}_2)+\cdots+k_mT(\boldsymbol{\alpha}_m)$$

又若 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m=\boldsymbol{0}$,则有

$$k_1T(\boldsymbol{\alpha}_1)+k_2T(\boldsymbol{\alpha}_2)+\cdots+k_mT(\boldsymbol{\alpha}_m)=\boldsymbol{0}$$

其中,$\boldsymbol{\beta},\boldsymbol{\alpha}_i\in V,k_i\in F\ (i=1,2,\cdots,m)$。上式表明,线性变换保持线性关系式不变。

(3) 若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\in V$,且 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关,则 $T(\boldsymbol{\alpha}_1),T(\boldsymbol{\alpha}_2),\cdots,T(\boldsymbol{\alpha}_m)$ 也线性相关。即线性变换把线性相关的向量组变成线性相关的向量组。

要注意的是,线性变换可能把线性无关的向量组变成线性相关的向量组。

1.3.2 线性变换的运算

定义 1.3.3 设线性变换 T_1、T_2 和 T 都是线性空间 V 的线性变换,$k\in F$,若

(1) 对任意向量 $\boldsymbol{\alpha}\in V$,均有 $T_1(\boldsymbol{\alpha})=T_2(\boldsymbol{\alpha})$,则称 T_1 与 T_2 相等,记为 $T_1=T_2$;

(2) 对任意向量 $\boldsymbol{\alpha}\in V$,均有 $T_1(\boldsymbol{\alpha})+T_2(\boldsymbol{\alpha})=T(\boldsymbol{\alpha})$,则称 T 为 T_1 与 T_2 的和,记为 $T=T_1+T_2$;

(3) 对任意向量 $\boldsymbol{\alpha}\in V$,任意 $k\in F$,均有 $T(\boldsymbol{\alpha})=k(T_1(\boldsymbol{\alpha}))$,则 T 为 T_1 与 k 的乘积,记为 $T=kT_1$;

(4) 若对任意向量 $\boldsymbol{\alpha}\in V$,有 $(-T)(\boldsymbol{\alpha})=-T(\boldsymbol{\alpha})$,则 $(-T)$ 为 T 的负变换;

(5) 对于任意向量 $\boldsymbol{\alpha}\in V$,均有 $T(\boldsymbol{\alpha})=T_1(T_2(\boldsymbol{\alpha}))$,则称 T 为 T_1 与 T_2 的积,记为 $T=T_1T_2$;

(6) 设 T_1 是 V 的线性变换,若在 V 中有线性变换 T_2 存在,使对于任意向量 $\boldsymbol{\alpha}\in V$,有

$$T_1T_2(\boldsymbol{\alpha})=T_2T_1(\boldsymbol{\alpha})=I(\boldsymbol{\alpha})$$

这时,线性变换 T_1 称为可逆的。T_2 称为 T_1 的逆变换,记为 $T_1^{-1}=T_2$,且逆变换是唯一的。

要注意的是,一般情况下线性变换的乘积不满足交换律,即 $T_1T_2\neq T_2T_1$。对线性变换的运算,通过验证易得到如下结论:

线性变换的和、数乘、积、逆变换与负变换都仍是线性变换。

这里要指出的是,定义 1.3.3 所提到的线性变换的加法和数乘具有以下性质:

线性变换的加法满足:

(1) $T_1+T_2=T_2+T_1$;

(2) $(T_1+T_2)+T_3=T_1+(T_2+T_3)$;

(3) $T+T_0=T$;

(4) $T+(-T)=T_0$。

线性变换的数乘满足:

(1) $1\cdot T=T$;

(2) $k(lT)=(kl)T$;

(3) $(k+l)T=kT+lT$。

其中 T_1、T_2、T_3、T_0、T、$(-T)$ 都是数域 F 上线性空间 V 的线性变换，k、l 是 F 中任意的数。

由线性变换的加法与数乘定义及其性质，看出线性空间中所有线性变换所构成的集合，在规定的运算下构成 V 的一个新线性子空间。

1.3.3　线性变换的矩阵表示

由于线性变换的概念比较抽象，为了研究方便，现来建立线性变换与矩阵的联系。

首先说明，当数域 F 上 n 维线性空间 V 取定一组基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 后，V 中的一个线性变换 T 可与 $F^{n\times n}$ 中的矩阵 $\boldsymbol{A}$ 有一个一一对应的关系，可形式地表示为

$$\begin{aligned} T[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n] &= [T(\boldsymbol{\alpha}_1),T(\boldsymbol{\alpha}_2),\cdots,T(\boldsymbol{\alpha}_n)] \\ &= (\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\mathbf{A} \end{aligned}$$

事实上，由线性变换 T 唯一地确定了基的像 $T(\boldsymbol{\alpha}_1),T(\boldsymbol{\alpha}_2),\cdots,T(\boldsymbol{\alpha}_n)$，所以它们在给定基下都有确定的坐标。以 $T(\boldsymbol{\alpha}_j)$ 在给定基下的坐标作为矩阵 $\boldsymbol{A}$ 的第 j 列，这就构造了唯一确定的一个 n 阶方阵 $\boldsymbol{A}$。

反之，取定 $F^{n\times n}$ 中一个矩阵 $\boldsymbol{A}$，则 $\boldsymbol{A}$ 的 n 个列唯一确定 n 个向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 分别作为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 的像，就可唯一确定 V 的一个线性变换 T，其中 $\boldsymbol{\beta}_j$ 关于基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 的坐标是 $\boldsymbol{A}$ 的第 j 列元素。

下面给出线性变换矩阵表示的定义。

定义 1.3.4　设 T 是 n 维线性空间 V 的一个线性变换，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是 V 的一个基，设

$$\begin{aligned} T(\boldsymbol{\alpha}_1) &= a_{11}\boldsymbol{\alpha}_1 + a_{21}\boldsymbol{\alpha}_2 + \cdots + a_{n1}\boldsymbol{\alpha}_n \\ T(\boldsymbol{\alpha}_2) &= a_{12}\boldsymbol{\alpha}_1 + a_{22}\boldsymbol{\alpha}_2 + \cdots + a_{n2}\boldsymbol{\alpha}_n \\ &\cdots\cdots \\ T(\boldsymbol{\alpha}_n) &= a_{1n}\boldsymbol{\alpha}_1 + a_{2n}\boldsymbol{\alpha}_2 + \cdots + a_{nn}\boldsymbol{\alpha}_n \end{aligned}$$

用矩阵形式表示为

$$\begin{aligned} T[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n] &= [T(\boldsymbol{\alpha}_1),T(\boldsymbol{\alpha}_2),\cdots,T(\boldsymbol{\alpha}_n)] \\ &= (\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{A} \end{aligned} \tag{1.3.1}$$

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 $\boldsymbol{A}$ 为线性变换 T 在基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 下的矩阵。

显然，线性空间 V 中的零变换、单位变换、数乘变换在任意一个基下的矩阵分别为零矩阵、单位矩阵和数量矩阵。

【例 1.3.6】　求 $F[x]_n$ 的求导变换 D，在基 $1,x,x^2,\cdots,x^{n-1}$ 下的矩阵。

解　因为

$$D(1)=0,\qquad D(x)=1,\qquad D(x^2)=2x,\quad \cdots,\quad D(x^{n-1})=(n-1)x^{n-2}$$

所以

$$D[1,x,x^2,\cdots,x^{n-1}] = [1,x,x^2,\cdots,x^{n-1}]\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & n-1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

所求矩阵为

$$\boldsymbol{A}=\begin{bmatrix}0&1&0&\cdots&0\\0&0&2&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&n-1\\0&0&0&\cdots&0\end{bmatrix}$$

在取定一组基后，线性变换与 n 阶矩阵是一一对应关系，它可以保持运算。

定理 1.3.1 若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是 n 维线性空间 V 的一个基，则在这组基下，每个线性变换按式(1.3.1)都可对应一个 n 阶方阵，这个对应有以下性质：

(1) 线性变换的和对应矩阵的和；

(2) 线性变换的乘积对应矩阵的乘积；

(3) 线性变换与数的乘积对应矩阵与数的乘积；

(4) 可逆的线性变换与可逆矩阵对应，且逆变换对应于逆矩阵。

证略。

利用线性变换的矩阵可以求出一个向量的像。

定理 1.3.2 设线性变换 T 在基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 下的矩阵是 $\boldsymbol{A}$，向量 $\boldsymbol{\alpha}$ 及其像 $T(\boldsymbol{\alpha})$ 在基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 下的坐标分别是 $(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 与 $(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$，则有

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

证 已知

$$[T(\boldsymbol{\alpha}_1),T(\boldsymbol{\alpha}_2),\cdots,T(\boldsymbol{\alpha}_n)]=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{A} \tag{1.3.2}$$

及

$$\boldsymbol{\alpha}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}$$

于是

$$T(\boldsymbol{\alpha})=[T(\boldsymbol{\alpha}_1),T(\boldsymbol{\alpha}_2),\cdots,T(\boldsymbol{\alpha}_n)]\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}$$

将式(1.3.2)代入，有

$$T(\boldsymbol{\alpha})=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{A}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix},\quad \text{而 } T(\boldsymbol{\alpha})=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}$$

又因为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关，所以

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

下面的定理给出同一线性变换在不同基下的矩阵之间的关系，这对以后的讨论很重要。

定理 1.3.3　设线性空间 V 中，线性变换 T 在基(1)$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 与基(2)$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 下的矩阵分别是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$，由基(1) 到基(2) 的过渡矩阵是 $\boldsymbol{P}$，则有 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$。

证　已知

$$T(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{A} \tag{1.3.3}$$

$$T(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)\boldsymbol{B} \tag{1.3.4}$$

$$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{P} \tag{1.3.5}$$

于是

$$\begin{aligned}T(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)&=T[(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{P}]\\&=T(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{P}\\&=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{A}\boldsymbol{P}\\&=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\end{aligned}$$

与式(1.3.4) 比较，有

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

所以，同一线性变换在不同基下的矩阵是相似关系。

【例 1.3.7】　已知 R^3 中线性变换 T 在基 $\boldsymbol{\beta}_1=(-1,1,1)^{\mathrm{T}},\boldsymbol{\beta}_2=(1,0,-1)^{\mathrm{T}},\boldsymbol{\beta}_3=(0,1,1)^{\mathrm{T}}$ 下的矩阵为

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

求：(1)T 在基 $\boldsymbol{\alpha}_1=(1,0,0)^{\mathrm{T}},\boldsymbol{\alpha}_2=(0,1,0)^{\mathrm{T}},\boldsymbol{\alpha}_3=(0,0,1)^{\mathrm{T}}$ 下的矩阵；

(2) 向量 $\boldsymbol{\eta}=(1,0,1)^{\mathrm{T}}$ 及 $T(\boldsymbol{\eta})$ 在 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 下的坐标。

解(1)　由

$$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)\begin{bmatrix}-1&1&0\\1&0&1\\1&-1&1\end{bmatrix}$$

得到基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 到基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 的过渡矩阵 $\boldsymbol{P}$

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

又已知

$$T(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)\begin{bmatrix}1&0&1\\1&1&0\\-1&2&1\end{bmatrix}$$

即线性变换 T 在基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 下的矩阵为

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

设线性变换 T 在基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 下的矩阵为 $\boldsymbol{A}$，即

$$T(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)\boldsymbol{A}$$

由定理 1.3.3 有 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$，求出

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

所以

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

$\boldsymbol{A}$ 即为所求。

(2) 设

$$\boldsymbol{\eta}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix},\qquad 即\quad \begin{bmatrix}1\\ 0\\ 1\end{bmatrix}=\begin{bmatrix}-1 & 1 & 0\\ 1 & 0 & 1\\ 1 & -1 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$$

解此非齐次线性方程组,得

$$x_1=-2,\qquad x_2=-1,\qquad x_3=2$$

所以,$\boldsymbol{\eta}$ 在基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 下的坐标为$(-2,-1,2)^{\mathrm{T}}$。

设 $T(\boldsymbol{\eta})$ 在基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 下的坐标为$(y_1,y_2,y_3)^{\mathrm{T}}$,可由定理 1.3.2 所给出的公式求出,即

$$\begin{bmatrix}y_1\\ y_2\\ y_3\end{bmatrix}=\boldsymbol{B}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$$

其中,$\boldsymbol{B}$ 为线性变换 T 在基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 下的矩阵。由于 $\boldsymbol{B}$ 已在前面求出,所以

$$\begin{bmatrix}y_1\\ y_2\\ y_3\end{bmatrix}=\begin{bmatrix}1 & 0 & 1\\ 1 & 1 & 0\\ -1 & 2 & 1\end{bmatrix}\begin{bmatrix}-2\\ -1\\ 2\end{bmatrix}=\begin{bmatrix}0\\ -3\\ 2\end{bmatrix}$$

*1.3.4 线性映射的矩阵表示

定义 1.3.5 设 T 是由 n 维线性空间 V_1 到 m 维线性空间 V_2 的一个线性映射,$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 分别是 V_1 与 V_2 的一个基,则

$$T(\boldsymbol{\alpha}_1)=a_{11}\boldsymbol{\beta}_1+a_{21}\boldsymbol{\beta}_2+\cdots+a_{m1}\boldsymbol{\beta}_m$$
$$T(\boldsymbol{\alpha}_2)=a_{12}\boldsymbol{\beta}_1+a_{22}\boldsymbol{\beta}_2+\cdots+a_{m2}\boldsymbol{\beta}_m$$
$$\cdots\cdots$$
$$T(\boldsymbol{\alpha}_n)=a_{1n}\boldsymbol{\beta}_1+a_{2n}\boldsymbol{\beta}_2+\cdots+a_{mn}\boldsymbol{\beta}_m$$

可形式地写成

$$T(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m)\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{bmatrix}$$

设

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

上式可写成

$$T(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m)\boldsymbol{A}$$

我们称矩阵 $\boldsymbol{A}$ 为线性映射 T 在 V_1 基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 与 V_2 基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 下的矩阵。上式称为线性映射在一对基下的矩阵表示。

可以证明在 V_1 与 V_2 的一对基下，线性映射 T 与矩阵 $\boldsymbol{A}$ 之间是一一对应的。

定理 1.3.4　设 T 是由 n 维线性空间 V_1 到 m 维线性空间 V_2 的线性映射，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 分别是 V_1 与 V_2 的一个基，T 在给定的一对基下的矩阵为 $\boldsymbol{A}$。若任意向量 $\boldsymbol{\alpha}\in V_1$，且 $T\boldsymbol{\alpha}\in V_2$，在给定基下，它们的坐标分别为 $(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 和 $(y_1,y_2,\cdots,y_m)^{\mathrm{T}}$，则有

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

此定理的证明与定理 1.3.2 的证明类似，留给读者自证。

定理 1.3.4 说明，有了线性映射在一对基下矩阵表示，就可以求出 V_1 中向量 $\boldsymbol{\alpha}$ 的坐标与它在 V_2 中像的坐标之间的关系。

下面的定理将给出线性映射在不同对基下矩阵之间的关系。

定理 1.3.5　设 T 是由 n 维线性空间 V_1 到 m 维线性空间 V_2 的一个线性映射，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 与 $\boldsymbol{\alpha}'_1,\boldsymbol{\alpha}'_2,\cdots,\boldsymbol{\alpha}'_n$ 是 V_1 的两个基，由 $\boldsymbol{\alpha}_i$ 到 $\boldsymbol{\alpha}'_i\,(i=1,2,\cdots,n)$ 的过渡矩阵为 $\boldsymbol{P}_1$。$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 与 $\boldsymbol{\beta}'_1,\boldsymbol{\beta}'_2,\cdots,\boldsymbol{\beta}'_m$ 是 V_2 的两个基，由 $\boldsymbol{\beta}_j$ 到 $\boldsymbol{\beta}'_j\,(j=1,2,\cdots,m)$ 的过渡矩阵为 $\boldsymbol{P}_2$。线性映射 T 在基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 下的矩阵表示为 $\boldsymbol{A}$，在基 $\boldsymbol{\alpha}'_1,\boldsymbol{\alpha}'_2,\cdots,\boldsymbol{\alpha}'_n$ 与 $\boldsymbol{\beta}'_1,\boldsymbol{\beta}'_2,\cdots,\boldsymbol{\beta}'_m$ 下的矩阵表示为 $\boldsymbol{B}$，则有

$$\boldsymbol{B}=\boldsymbol{P}_2^{-1}\boldsymbol{A}\boldsymbol{P}_1$$

证略。由矩阵等价的定义知，矩阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{B}$ 是等价的。

定理 1.3.5 说明，由 n 维线性空间 V_1 到 m 维线性空间 V_2 的一个线性映射 T 有一系列的 $m\times n$ 矩阵表示：$\boldsymbol{A},\boldsymbol{B},\cdots$，它们之间的关系是等价关系。

1.4　与线性变换有关的子空间

1.4.1　线性变换的值域与核

在例 1.1.6 与例 1.1.5 中，我们介绍了矩阵 $\boldsymbol{A}$ 的值域 $R(\boldsymbol{A})$ 与核 $N(\boldsymbol{A})$：

$$\begin{aligned} R(\boldsymbol{A})&=\{\boldsymbol{y}\mid \boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}\} \\ N(\boldsymbol{A})&=\{\boldsymbol{x}\mid \boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}\}\end{aligned}\quad(\text{其中 }\boldsymbol{A}\in C^{m\times n},x\in C^n)$$

若 V 是数域 F 上的 n 维线性空间，T 是 V 的一个线性变换，在 V 中取定一组基之后，我们就建立了 V 的线性变换 T 与数域 F 上的 n 阶方阵 $\boldsymbol{A}$ 的一个一一对应关系。这时若线性变换 T 的矩阵表示就是 n 阶方阵 $\boldsymbol{A}$，则矩阵 $\boldsymbol{A}$ 的值域与核就是线性变换 T 的值域与核。

线性变换的值域与核可定义如下。

定义 1.4.1　如 T 是 n 维线性空间 V 的一个线性变换，T 的所有的像组成的集合称为 T 的值域，用 $R(T)$ 表示，也称为 T 的像空间，记为 TV，于是

$$R(T)=TV=\{\boldsymbol{\beta}\mid\boldsymbol{\beta}=T(\boldsymbol{\alpha}),\quad\boldsymbol{\alpha}\in V\}$$

所有被 T 变成零向量的向量组成的集合称为 T 的核，记为 $\ker(T)$ 或 $T^{-1}(\boldsymbol{0})$；也称为 T 的零空间，记为 $N(T)$，于是

$$N(T)=\{\boldsymbol{\alpha}\mid T(\boldsymbol{\alpha})=\boldsymbol{0},\quad\boldsymbol{\alpha}\in V\}$$

不难验证,线性变换 T 的像空间 $R(T)$ 和零空间 $N(T)$ 都是 V 的子空间。

事实上,由

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

得知,$\boldsymbol{0}\in R(T)$,$R(T)$ 是非空集合,且对线性运算封闭。因此,$R(T)$ 是 V 的子空间。

又由

$$T(\boldsymbol{\alpha})=\boldsymbol{0},\quad T(\boldsymbol{\beta})=\boldsymbol{0}$$

得到

$$T(\boldsymbol{\alpha}+\boldsymbol{\beta})=\boldsymbol{0},\quad T(k\boldsymbol{\alpha})=\boldsymbol{0}$$

即 $N(T)$ 对于线性运算也是封闭的。又由于 $T(\boldsymbol{0})=\boldsymbol{0}$,即 $N(T)$ 非空,所以 $N(T)$ 也是 V 的子空间。

【例 1.4.1】 设 $R[x]_n$ 是由变量 x 的次数小于 n 的实系数多项式全体加上零多项式构成的线性空间,并设求导线性变换为 D,则 $D(R[x]_n)=R'[x]_n=R[x]_{n-1}$。所以,$D$ 的值域 $R(D)$ 是 $R[x]_{n-1}$,而 D 的核 $N(D)$ 是实数域 $\mathbf{R}$,它们都是 $R[x]_n$ 的子空间。

定理 1.4.1 设 T 是 n 维线性空间 V 上的线性变换,把 $R(T)$ 的维数 $\dim R(T)$ 称为 T 的秩,记为 $r(T)$;把 $N(T)$ 的维数 $\dim N(T)$ 称为 T 的零度,记为 $\mathrm{null}(T)$,则有

$$\dim R(T)+\dim N(T)=n$$

即

$$r(T)+\mathrm{null}(T)=\dim V$$

要注意的是,虽然 $R(T)$ 与 $N(T)$ 的维数之和等于 V 的维数 n,但 $R(T)+N(T)$ 不一定就是 V。如例 1.4.1 中,$R(D)$ 就是 $R[x]_{n-1}$,而 $N(D)$ 就是实数域 $\mathbf{R}$,显然

$$R(D)+N(D)\neq R[x]_n$$

1.4.2 线性变换的不变子空间

在本段内容中,将讨论线性变换与子空间的关系,并利用线性变换的不变子空间进行线性变换矩阵的化简。

定义 1.4.2 设 T 是 n 维线性空间 V 上的线性变换,W 是 V 的子空间,若 W 中任意向量 $\boldsymbol{\alpha}$ 有像 $T(\boldsymbol{\alpha})$ 仍在 W 中,则称 W 为 T 的不变子空间,简记 T-子空间。

【例 1.4.2】 线性变换 T 的值域 $R(T)$ 和核 $N(T)$ 都是 T 的不变子空间。事实上,$R(T)$ 是 V 中的向量在 T 下的像集合,显然 $R(T)$ 包含了 $R(T)$ 中向量的像。所以,$R(T)$ 是 T 的不变子空间;而 $N(T)$ 是被 T 变为零向量的向量的集合,$N(T)$ 中向量的像都是零向量,显然在核中,所以 $N(T)$ 也是不变子空间。

【例 1.4.3】 整个线性空间 V 和零子空间 $\{\boldsymbol{0}\}$,对于任何线性变换 T 来说,都是 T 的不变子空间,特别是单位变换 I 的不变子空间。

【例 1.4.4】 任何一个子空间都是数乘变换的不变子空间。

事实上,设 K 是线性空间 V 的数乘变换,W 是 V 的子空间,对任意 $\boldsymbol{\alpha}\in W$,有 $K(\boldsymbol{\alpha})=k\boldsymbol{\alpha}\in W$,这是由于 W 是子空间,对数量乘法是封闭的。

【例 1.4.5】 线性变换 T 的不变子空间的交与和仍为线性变换 T 的不变子空间。即

设 T 是线性空间 V 的线性变换,V_1 与 V_2 是 T 的不变子空间,则它们的交 $V_1\cap V_2$ 与和 V_1+V_2 仍为 T 的不变子空间。

事实上,由于交 $V_1\cap V_2$ 仍是 V 的子空间,若任意 $\boldsymbol{\alpha}\in V_1\cap V_2$,则有 $\boldsymbol{\alpha}\in V_1$ 和 $\boldsymbol{\alpha}\in V_2$。又因 V_1 与 V_2 是 T 的不变子空间,有 $T\boldsymbol{\alpha}\in V_1$,$T\boldsymbol{\alpha}\in V_2$,从而有 $T\boldsymbol{\alpha}\in V_1\cap V_2$,这就证明了 $V_1\cap V_2$ 是 T 的不变子空间。

类似地，可以证明 V_1+V_2 也是 T 的不变子空间。

【例 1.4.6】　若线性空间 V 的子空间 W 由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 生成，即 $W=L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r)$，$T$ 是 V 的线性变换，则 W 是 T 的不变子空间的充要条件是 $T(\boldsymbol{\alpha}_1),T(\boldsymbol{\alpha}_2),\cdots,T(\boldsymbol{\alpha}_r)$ 全属于 W。

证　由不变子空间的定义，可见必要性是显然的。下面证明其充分性。

已知 $T(\boldsymbol{\alpha}_i)\in W(i=1,2,\cdots,r)$，且 $W=L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r)$，所以对任意 $\boldsymbol{\alpha}\in W$，$\boldsymbol{\alpha}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r$ 有

$$T(\boldsymbol{\alpha})=k_1T(\boldsymbol{\alpha}_1)+k_2T(\boldsymbol{\alpha}_2)+\cdots+k_rT(\boldsymbol{\alpha}_r)=\sum_{i=1}^{r}k_iT(\boldsymbol{\alpha}_i)\in W$$

故子空间 W 是线性变换 T 的不变子空间。

下面的定理是利用不变子空间来简化线性变换的矩阵。

定理 1.4.2　设 T 是 n 维线性空间 V 的线性变换，且 V 可分解为 t 个 T 的不变子空间的直和，即

$$V=V_1\oplus V_2\oplus\cdots\oplus V_t$$

若在每个不变子空间 $V_i\quad(i=1,2,\cdots,t)$ 中取基

$$\boldsymbol{\alpha}_{i1},\boldsymbol{\alpha}_{i2},\cdots,\boldsymbol{\alpha}_{ir_i}\tag{1.4.1}$$

其中，$r_1+r_2+\cdots+r_t=n=\dim V$。把这些基合在一起，即

$$\boldsymbol{\alpha}_{11},\boldsymbol{\alpha}_{12},\cdots,\boldsymbol{\alpha}_{1r_1},\boldsymbol{\alpha}_{21},\boldsymbol{\alpha}_{22},\cdots,\boldsymbol{\alpha}_{2r_2},\cdots,\boldsymbol{\alpha}_{t1},\boldsymbol{\alpha}_{t2},\cdots,\boldsymbol{\alpha}_{tr_t}\tag{1.4.2}$$

就是 V 的一个基，则 T 在这个基下的矩阵为准对角矩阵。

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_1 & & & \\ & \boldsymbol{A}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_t\end{bmatrix}\tag{1.4.3}$$

其中 $\boldsymbol{A}_i(i=1,2,\cdots,t)$ 就是 T 在 V_i 的基(1.4.1)下的 $r_i\times r_i$ 阶矩阵。反之亦然。

证　因为 $V_1,V_2,\cdots,V_t$ 都是 T 的不变子空间，所以当 $\boldsymbol{\alpha}_{ij}\in V_i$ 时有

$$T(\boldsymbol{\alpha}_{ij})\in V_i\quad(i=1,2,\cdots,t;j=1,2,\cdots,r_i)$$

于是有

$$\begin{gathered}T(\boldsymbol{\alpha}_{i1})=a_{11}^{(i)}\boldsymbol{\alpha}_{i1}+a_{21}^{(i)}\boldsymbol{\alpha}_{i2}+\cdots+a_{r_i1}^{(i)}\boldsymbol{\alpha}_{ir_i}\\ T(\boldsymbol{\alpha}_{i2})=a_{12}^{(i)}\boldsymbol{\alpha}_{i1}+a_{22}^{(i)}\boldsymbol{\alpha}_{i2}+\cdots+a_{r_i2}^{(i)}\boldsymbol{\alpha}_{ir_i}\\ \cdots\cdots\\ T(\boldsymbol{\alpha}_{ir_i})=a_{1r_i}^{(i)}\boldsymbol{\alpha}_{i1}+a_{2r_i}^{(i)}\boldsymbol{\alpha}_{i2}+\cdots+a_{r_ir_i}^{(i)}\boldsymbol{\alpha}_{ir_i}\\ (i=1,2,\cdots,t)\end{gathered}$$

由此，T 在 V 的基下的矩阵 $\boldsymbol{A}$ 是式(1.4.3)的形式，其中

$$\boldsymbol{A}_i=\begin{bmatrix}a_{11}^{(i)} & a_{12}^{(i)} & \cdots & a_{1r_i}^{(i)}\\ a_{21}^{(i)} & a_{22}^{(i)} & \cdots & a_{2r_i}^{(i)}\\ \vdots & \vdots & & \vdots\\ a_{r_i1}^{(i)} & a_{r_i2}^{(i)} & \cdots & a_{r_ir_i}^{(i)}\end{bmatrix}\quad(i=1,2,\cdots,t)$$

反之，可证：若线性变换 T 在 V 的基式(1.4.2)下的矩阵表示是准对角矩阵式(1.4.3)，则由基式(1.4.1)生成的子空间

$$V_i=L(\boldsymbol{\alpha}_{i1},\boldsymbol{\alpha}_{i2},\cdots,\boldsymbol{\alpha}_{ir_i})\quad(i=1,2,\cdots,t)$$

是 T 的不变子空间。所以，如果线性变换矩阵是准对角矩阵时，那么表明线性空间可分解为不变子空间的直和，它们是相互对应的。

1.5　欧氏空间与酉空间

在前面所研究的线性空间(向量空间)中,向量的基本运算仅限于向量的加法及数与向量的乘法,没有涉及向量的长度、两个向量的夹角等度量概念,但这些概念在很多实际问题中又有着重要的作用。因此,有必要将它们引入到一般的线性空间中。

由于在解析几何中,向量的长度及夹角可以通过向量的数量积来表示。因此,为定义线性空间中的向量长度与夹角等概念,我们首先引入与数量积相似的内积的概念。

1.5.1　欧氏空间的定义与性质

先给出内积与欧氏空间的定义和性质。

定义 1.5.1　设 V 是实数域上的线性空间,若对于 V 中任意两个向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$,按某个规则,恒有唯一的一个实数与之对应,用记号 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 表示,且满足

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

(2) 齐次性:$(k\boldsymbol{\alpha},\boldsymbol{\beta})=k(\boldsymbol{\alpha},\boldsymbol{\beta})$;

(3) 分配律:$(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\gamma})=(\boldsymbol{\alpha},\boldsymbol{\gamma})+(\boldsymbol{\beta},\boldsymbol{\gamma})$;

(4) 非负性:$(\boldsymbol{\alpha},\boldsymbol{\alpha})\geqslant 0$,当且仅当 $\boldsymbol{\alpha}=\boldsymbol{0}$ 时 $(\boldsymbol{\alpha},\boldsymbol{\alpha})=0$。

这里,$\boldsymbol{\alpha},\boldsymbol{\beta}$ 和 $\boldsymbol{\gamma}$ 是 V 中任意向量,k 是任意实数,则称 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 为向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的内积,定义了内积的实线性空间称为 Euclid 空间,简称欧氏空间。

这里要注意的是,欧氏空间的子空间在所定义的内积下,显然也是一个欧氏空间。同一实线性空间,若定义不同的内积,则构成不同的欧氏空间。

【例 1.5.1】　易验证立体几何空间 R^3 中的数量积满足内积的 4 个条件。因此,数量积是一种内积,几何空间是一个具体的欧氏空间。

【例 1.5.2】　在实 n 维向量空间 R^n 中,若对任意两个向量

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

规定

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=a_1b_1+a_2b_2+\cdots+a_nb_n=\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}} \tag{1.5.1}$$

易验证它满足内积的 4 个条件。因此式(1.5.1)是内积,R^n 是欧氏空间。

【例 1.5.3】　在例 1.3.4 的实线性空间 $C[a,b]$ 中,若对它的任意两个连续函数 $f(x),g(x)$,规定

$$[f(x),g(x)]=\int_a^b f(x)g(x)\mathrm{d}x \tag{1.5.2}$$

则由定积分的性质,易验证它满足内积的 4 个条件,因此式(1.5.2)是内积,$C[a,b]$ 是欧氏空间。

【例 1.5.4】　实数域 **R** 上所有 n 阶方阵,依通常的矩阵加法及数与矩阵的乘法,构成的 n^2 维实线性空间 $R^{n\times n}$。若对任意两个矩阵 $\boldsymbol{A}=(a_{ij}),\boldsymbol{B}=(b_{ij})\in R^{n\times n}$,规定

$$(\boldsymbol{A},\boldsymbol{B})=\sum_{i,j=1}^{n}a_{ij}b_{ij} \tag{1.5.3}$$

易验证由式(1.5.3)规定的内积满足内积的 4 个条件。那么,这个实线性空间 $R^{n\times n}$ 是一个欧氏空间。

事实上,矩阵 $\boldsymbol{A}$ 的迹 $\mathrm{tr}(\boldsymbol{A})$ 是 $\boldsymbol{A}$ 主对角线上元素之和,由这个定义有:

(1) $(\boldsymbol{A},\boldsymbol{B})=\sum\limits_{i,j=1}^{n}a_{ij}b_{ij}=\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}^{\mathrm{T}})=\mathrm{tr}(\boldsymbol{B}\boldsymbol{A}^{\mathrm{T}})=(\boldsymbol{B},\boldsymbol{A})$;

(2) $(k\boldsymbol{A},\boldsymbol{B})=\sum\limits_{i,j=1}^{n}ka_{ij}b_{ij}=\mathrm{tr}(k\boldsymbol{A}\boldsymbol{B}^{\mathrm{T}})=k\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}^{\mathrm{T}})=k(\boldsymbol{A},\boldsymbol{B})$;

(3) $(\boldsymbol{A}+\boldsymbol{B},\boldsymbol{C})=\sum\limits_{i,j=1}^{n}(a_{ij}+b_{ij})c_{ij}=\mathrm{tr}[(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C}^{\mathrm{T}}]=\mathrm{tr}(\boldsymbol{A}\boldsymbol{C}^{\mathrm{T}}+\boldsymbol{B}\boldsymbol{C}^{\mathrm{T}})$

$$= \mathrm{tr}(\boldsymbol{A}\boldsymbol{C}^{\mathrm{T}}) + \mathrm{tr}(\boldsymbol{B}\boldsymbol{C}^{\mathrm{T}}) = (\boldsymbol{A},\boldsymbol{C}) + (\boldsymbol{B},\boldsymbol{C});$$

(4) $(\boldsymbol{A},\boldsymbol{A}) = \sum_{i,j=1}^{n} a_{ij}^2 \geqslant 0$，等号成立，当且仅当 $a_{ij} = 0(\forall i,j)$，即 $\boldsymbol{A} = \boldsymbol{O}$。

所以，实线性空间 $R^{n\times n}$ 是一个欧氏空间。

例 1.5.3 与例 1.5.4 表明欧氏空间的概念比几何空间的概念更广泛、更抽象。

从内积的定义（定义 1.5.1），易得到以下内积的基本性质：

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

性质 2　$(\boldsymbol{\alpha},\boldsymbol{0}) = (\boldsymbol{0},\boldsymbol{\alpha}) = 0$；

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

性质 4　$\left(\sum_{i=1}^{n} k_i\boldsymbol{\alpha}_i, \sum_{j=1}^{n} l_j\boldsymbol{\beta}_j\right) = \sum_{i,j=1}^{n} k_i l_j(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_j)$。

1.5.2　度量矩阵及可度量的量

(1) 度量矩阵

设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是 n 维欧氏空间 V 的一个基，$\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 是 V 中任意两个向量，且

$$\boldsymbol{\alpha} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n$$
$$\boldsymbol{\beta} = y_1\boldsymbol{\alpha}_1 + y_2\boldsymbol{\alpha}_2 + \cdots + y_n\boldsymbol{\alpha}_n$$

则

$$\begin{aligned}(\boldsymbol{\alpha},\boldsymbol{\beta}) &= \left(\sum_{i=1}^{n} x_i\boldsymbol{\alpha}_i, \sum_{j=1}^{n} y_j\boldsymbol{\alpha}_j\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)\end{aligned} \tag{1.5.4}$$

若令 $(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j) = a_{ij}(i,j=1,2,\cdots,n)$，则由内积的性质有 $a_{ij} = a_{ji}(i,j=1,2,\cdots,n)$，利用矩阵乘法，将式(1.5.4) 写成矩阵形式，有

$$(\boldsymbol{\alpha},\boldsymbol{\beta}) = (x_1,x_2,\cdots,x_n)\boldsymbol{A}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{y} \tag{1.5.5}$$

其中

$$\boldsymbol{A} = \begin{bmatrix} (\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_n) \\ (\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_n) \\ \vdots & \vdots & & \vdots \\ (\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_n) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

由式(1.5.5) 可看出，一个基向量组中，所有两两基向量的内积就可确定矩阵 $\boldsymbol{A}$。我们称式(1.5.5) 中的 $\boldsymbol{A}$ 为基向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 的度量矩阵。具体定义如下。

定义 1.5.2　设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是 n 维欧氏空间 V 的一个基，令 $a_{ij} = (\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)$，$\boldsymbol{A} = (a_{ij})_{n\times n}$，则称 $\boldsymbol{A}$ 为基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 的度量矩阵。

由定义 1.5.1 第(1) 条，可知 $(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j) = (\boldsymbol{\alpha}_j,\boldsymbol{\alpha}_i)$，即 $\boldsymbol{A}$ 中元素 $a_{ji} = a_{ji}$，显然度量矩阵 $\boldsymbol{A}$ 是实对称矩阵。

下面说明度量矩阵 $\boldsymbol{A}$ 还是正定矩阵，由于 V 中任一个非零向量 $\boldsymbol{\alpha}$ 在基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 下有

$$\boldsymbol{\alpha}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix},\quad \text{其中 } \boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}\neq\boldsymbol{0}$$

由定义 1.5.1 第(4)条,有$(\boldsymbol{\alpha},\boldsymbol{\alpha})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}>0$。

下面给出的定理将说明欧氏空间中不同基的度量矩阵是合同关系。

定理 1.5.1 若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 是 n 维欧氏空间的两个基,它们的度量矩阵分别为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$,由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 到 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 的过渡矩阵是 $\boldsymbol{C}$,那么 $\boldsymbol{B}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$。(留读者自证,或可参考与本教材配套的学习指导书第 1.6 节典型题剖析例 1.15 题)。

有了度量矩阵,欧氏空间中任意两个向量的内积,可通过坐标按式(1.5.5)计算,度量矩阵越简单,向量内积计算也越简单。若基向量两两正交,且长度为 1,这时度量矩阵是单位矩阵,那么任意两个向量内积表达式将最简单。

下面把度量的概念引入欧氏空间,用内积来定义向量的长度,任意两个向量的夹角与任意两个向量的距离等。

(2) 长度、夹角与距离

在几何空间 R^3 中,向量 $\boldsymbol{\alpha}$ 的长度定义为

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

在一般的欧氏空间 V 中,由于内积$(\boldsymbol{\alpha},\boldsymbol{\alpha})\geqslant 0$,所以,可在欧氏空间中,类似的引入长度的定义。

定义 1.5.3 设 $\boldsymbol{\alpha}$ 是欧氏空间中任意向量,则非负实数 $\sqrt{(\boldsymbol{\alpha},\boldsymbol{\alpha})}$,称为 $\boldsymbol{\alpha}$ 向量的长度(或模,或范数),记为 $\|\boldsymbol{\alpha}\|$ 或 $|\boldsymbol{\alpha}|$,即

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

显然,非零向量的长度是正数。长度等于零的向量,当且仅当它是零向量。

【例 1.5.5】 在 n 维欧氏几何空间 R^n 中当向量的内积按定义 1.5.1 定义时,向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 的长度为

$$\|\boldsymbol{x}\|=\sqrt{(\boldsymbol{x},\boldsymbol{x})}=\sqrt{x_1^2+x_2^2+\cdots+x_n^2}$$

对于欧氏空间 $C[a,b]$ 来说,当向量内积按定义 1.5.1 定义时,向量 $\boldsymbol{f}(t)$ 的长度为

$$\|\boldsymbol{f}(t)\|=\sqrt{\int_a^b \boldsymbol{f}^2(t)\,\mathrm{d}t}$$

长度为 1 的向量称为单位向量。若向量 $\boldsymbol{\alpha}\neq\boldsymbol{0}$,则 $\dfrac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|}$(用向量 $\boldsymbol{\alpha}$ 的长度去除 $\boldsymbol{\alpha}$)是一个与 $\boldsymbol{\alpha}$ 同向的单位向量,此过程称为把向量 $\boldsymbol{\alpha}$ 单位化。

根据上述定义,对于欧氏空间中任意的向量 $\boldsymbol{\alpha},\boldsymbol{\beta}\in V$ 及任意数 $k\in\mathbf{R}$,以下各式成立:

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

(2) 柯西—许瓦兹(Cauchy-Schwarz)不等式

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

当且仅当 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 线性相关时等号成立;

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

(4) $\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|\geqslant\big|\,\|\boldsymbol{\alpha}\|-\|\boldsymbol{\beta}\|\,\big|$。

事实上,因为

$$\|k\boldsymbol{\alpha}\|=\sqrt{(k\boldsymbol{\alpha},k\boldsymbol{\alpha})}=\sqrt{k^2(\boldsymbol{\alpha},\boldsymbol{\alpha})}=|k|\sqrt{(\boldsymbol{\alpha},\boldsymbol{\alpha})}=|k|\,\|\boldsymbol{\alpha}\|$$

所以,(1)式成立。

(2)式是很有用的不等式,证明它可考虑任意两个非零向量 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$,和任意实数 $t\in\mathbf{R}$。设 $\boldsymbol{\gamma}=\boldsymbol{\alpha}+t\boldsymbol{\beta}$,

由定义知道

$$
\begin{aligned}
(\boldsymbol{\gamma},\boldsymbol{\gamma}) &= (\boldsymbol{\alpha}+t\boldsymbol{\beta},\boldsymbol{\alpha}+t\boldsymbol{\beta}) \\
&= (\boldsymbol{\alpha},\boldsymbol{\alpha})+2t(\boldsymbol{\alpha},\boldsymbol{\beta})+t^2(\boldsymbol{\beta},\boldsymbol{\beta}) \geqslant 0
\end{aligned}
\tag{1.5.6}
$$

运用初等代数中的知识：关于实系数二次三项式 $at^2+2bt+c(a>0)$，若对任意实数 t 都取非负值，则系数之间必有判别式 $b^2-ac\leqslant 0$ 成立。对于式(1.5.6)可得

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

即

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

两边开方，即得

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

当且仅当 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 线性相关时，等号成立，因为若 $\boldsymbol{\beta}=\boldsymbol{0}$，显然上式中的等号成立；若 $\boldsymbol{\beta}\neq\boldsymbol{0}$，则有 $\boldsymbol{\alpha}=k\boldsymbol{\beta}$($k$ 为常数)，于是有

$$
\begin{aligned}
|(\boldsymbol{\alpha},\boldsymbol{\beta})| &= |(k\boldsymbol{\beta},\boldsymbol{\beta})| = |k|(\boldsymbol{\beta},\boldsymbol{\beta}) = |k|\,\|\boldsymbol{\beta}\|^2 = \|k\boldsymbol{\beta}\|\,\|\boldsymbol{\beta}\| \\
&= \|\boldsymbol{\alpha}\|\,\|\boldsymbol{\beta}\|
\end{aligned}
$$

上式中的等号同样成立。

反之，若上式中的等号成立，即有

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

这时，若 $\boldsymbol{\beta}=\boldsymbol{0}$，显然 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 线性相关。若 $\boldsymbol{\beta}\neq\boldsymbol{0}$，取 $t=\dfrac{\|\boldsymbol{\alpha}\|}{\|\boldsymbol{\beta}\|}$，现研究向量 $\boldsymbol{\alpha}-t\boldsymbol{\beta}$，有

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

注意：在 $(\boldsymbol{\alpha},\boldsymbol{\beta})^2=(\boldsymbol{\alpha},\boldsymbol{\alpha})(\boldsymbol{\beta},\boldsymbol{\beta})$ 条件下，并将 $t=\dfrac{\|\boldsymbol{\alpha}\|}{\|\boldsymbol{\beta}\|}$ 代入上式，有 $\|\boldsymbol{\alpha}-t\boldsymbol{\beta}\|^2=0$。所以 $\boldsymbol{\alpha}-t\boldsymbol{\beta}=\boldsymbol{0}$，因而 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 线性相关。

证明不等式(3)可以用不等式(2)，即

$$
\begin{aligned}
\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2 &= (\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\alpha}+\boldsymbol{\beta}) = (\boldsymbol{\alpha},\boldsymbol{\alpha})+(\boldsymbol{\alpha},\boldsymbol{\beta})+(\boldsymbol{\beta},\boldsymbol{\alpha})+(\boldsymbol{\beta},\boldsymbol{\beta}) \\
&= \|\boldsymbol{\alpha}\|^2+2(\boldsymbol{\alpha},\boldsymbol{\beta})+\|\boldsymbol{\beta}\|^2 \leqslant \|\boldsymbol{\alpha}\|^2+2|(\boldsymbol{\alpha},\boldsymbol{\beta})|+\|\boldsymbol{\beta}\|^2 \\
&\leqslant \|\boldsymbol{\alpha}\|^2+2\|\boldsymbol{\alpha}\|\,\|\boldsymbol{\beta}\|+\|\boldsymbol{\beta}\|^2 = (\|\boldsymbol{\alpha}\|^2+\|\boldsymbol{\beta}\|^2)^2
\end{aligned}
$$

两边开方，有不等式(3)成立，即

$$\|\boldsymbol{\alpha}+\boldsymbol{\beta}\| \leqslant \|\boldsymbol{\alpha}\|+\|\boldsymbol{\beta}\| \qquad \text{（三角不等式）}$$

不等式(3)称为三角不等式，这是因为在通常平面上，上面的向量关系式表示三角形任意两边之和大于第三边。

(4)式留给读者自证。

Cauchy-Schwarz 不等式有十分重要的应用。在欧氏空间 R^n 中，Cauchy-Schwarz 不等式成为

$$
\begin{aligned}
&|x_1y_1+x_2y_2+\cdots+x_ny_n| \\
&\leqslant \sqrt{x_1^2+x_2^2+\cdots+x_n^2}\cdot\sqrt{y_1^2+y_2^2+\cdots+y_n^2}
\end{aligned}
\tag{1.5.7}
$$

其中，$\boldsymbol{\alpha},\boldsymbol{\beta}\in R^n$，$\boldsymbol{\alpha}=(x_1,x_2,\cdots,x_n)$，$\boldsymbol{\beta}=(y_1,y_2,\cdots,y_n)$。

对于欧氏空间 $C[a,b]$，Cauchy-Schwarz 不等式就成为

$$\left|\int_a^b f(x)g(x)\mathrm{d}x\right| \leqslant \left[\int_a^b f^2(x)\mathrm{d}x\right]^{\frac{1}{2}}\left[\int_a^b g^2(x)\mathrm{d}x\right]^{\frac{1}{2}} \tag{1.5.8}$$

不等式(1.5.7)和式(1.5.8)都是数学史上著名的不等式。

有了内积的概念，两个向量之间距离可用 $\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|$ 来表示，记做 $\rho(\boldsymbol{\alpha},\boldsymbol{\beta})$，即

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

它有如下的性质：

(1) **正定性** $\rho(\boldsymbol{\alpha},\boldsymbol{\alpha})=0$, $\rho(\boldsymbol{\alpha},\boldsymbol{\beta})>0$ $(\boldsymbol{\alpha}\neq\boldsymbol{\beta})$;

(2) **对称性** $\rho(\boldsymbol{\alpha},\boldsymbol{\beta})=\rho(\boldsymbol{\beta},\boldsymbol{\alpha})$;

(3) **三角不等式** $\rho(\boldsymbol{\alpha},\boldsymbol{\beta})+\rho(\boldsymbol{\beta},\boldsymbol{\gamma})\geqslant\rho(\boldsymbol{\alpha},\boldsymbol{\gamma})$。

在解析几何中,向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的夹角 $\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle$ 的余弦可以通过内积表示为

$$\cos\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{\|\boldsymbol{\alpha}\|\ \|\boldsymbol{\beta}\|} \tag{1.5.9}$$

在欧氏空间中,由于我们已证明了 Cauchy-Schwarz 不等式

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

即

$$\left|\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{\|\boldsymbol{\alpha}\|\ \|\boldsymbol{\beta}\|}\right|\leqslant 1$$

于是在一般欧氏空间,也可以利用式(1.5.9)引入夹角的概念。

定义 1.5.4 设 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 为欧氏空间 V 中任意两个非零向量,称

$$\theta=<\boldsymbol{\alpha},\boldsymbol{\beta}>=\arccos\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{\|\boldsymbol{\alpha}\|\ \|\boldsymbol{\beta}\|}\quad(0\leqslant\theta\leqslant\pi)$$

为向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的夹角。

定义 1.5.5 设 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 为欧氏空间 V 中任意两个向量,若 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 的内积为零,即 $(\boldsymbol{\alpha},\boldsymbol{\beta})=0$,则称 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 为正交,或互相垂直,记为 $\boldsymbol{\alpha}\perp\boldsymbol{\beta}$。

由定义 1.5.4 看出,只有零向量才自身正交;而且 V 中的零向量与任意向量都正交。

由定义 1.5.4 还可以推出,在欧氏空间中也有勾股定理,即当 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 正交时,有

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

事实上,

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

因 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 正交,有 $(\boldsymbol{\alpha},\boldsymbol{\beta})=0$,所以

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

1.5.3 标准正交基

类似线性空间基的概念,非零的有限维欧氏空间也有基的概念。

定义 1.5.6 欧氏空间 V 中,若有一组非零向量,它们两两正交,则称其为一个正交向量组。

定理 1.5.2 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是欧氏空间的一个正交向量组,则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关。证略。

定义 1.5.7 在 n 维欧氏空间中,由 n 个向量组成的正交向量组称为正交基,由单位向量组成的正交基称为标准正交基。

由定义 1.5.7,若 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 是一组标准正交基,则有

$$(\boldsymbol{\varepsilon}_i,\boldsymbol{\varepsilon}_j)=\begin{cases}1 & \text{当 } i=j\\ 0 & \text{当 } i\neq j\end{cases}\quad(i,j=1,2,\cdots,n) \tag{1.5.10}$$

由式(1.5.10)看出,一组基是标准正交基的充要条件是它的度量矩阵是单位矩阵。前面已说明度量矩阵是正定的,因此它合同于单位矩阵,这说明 n 维欧氏空间中一定存在一组基,它的度量矩阵是单位矩阵,所以在 n 维欧氏空间中一定存在标准正交基。

事实上,对于 n 维欧氏空间 V 中的任何一组基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 都可以用 Schmidt 正交化的方法,先正交化,而后再单位化,使之成为一组标准正交基 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n$,且有

$$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=L(\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n)$$

具体操作过程步骤如下:

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

设
$$\boldsymbol{\beta}_2=k\boldsymbol{\beta}_1+\boldsymbol{\alpha}_2 \tag{1.5.11}$$
利用正交条件$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2)=0$,即
$$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2)=(\boldsymbol{\beta}_1,k\boldsymbol{\beta}_1+\boldsymbol{\alpha}_2)=k(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)+(\boldsymbol{\beta}_1,\boldsymbol{\alpha}_2)=0$$
得到
$$k=-\frac{(\boldsymbol{\beta}_1,\boldsymbol{\alpha}_2)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}$$
代入式(1.5.11),得
$$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\frac{(\boldsymbol{\beta}_1,\boldsymbol{\alpha}_2)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1$$
这样得到的 $\boldsymbol{\beta}_2$ 与 $\boldsymbol{\beta}_1$ 正交。

再用同样的方法求 $\boldsymbol{\beta}_3$,使其与 $\boldsymbol{\beta}_1$ 和 $\boldsymbol{\beta}_2$ 均正交。可取
$$\boldsymbol{\beta}_3=k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2+\boldsymbol{\alpha}_3 \tag{1.5.12}$$
利用正交条件$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_3)=0$,$(\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=0$,可解得
$$k_1=-\frac{(\boldsymbol{\beta}_1,\boldsymbol{\alpha}_3)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)},\qquad k_2=-\frac{(\boldsymbol{\beta}_2,\boldsymbol{\alpha}_3)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}$$
代入式(1.5.12),得到
$$\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3-\frac{(\boldsymbol{\beta}_1,\boldsymbol{\alpha}_3)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\frac{(\boldsymbol{\beta}_2,\boldsymbol{\alpha}_3)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2$$
以此类推,便可得到一组正交向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$,其中
$$\boldsymbol{\beta}_i=\boldsymbol{\alpha}_i-\frac{(\boldsymbol{\beta}_1,\boldsymbol{\alpha}_i)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\cdots-\frac{(\boldsymbol{\beta}_{i-1},\boldsymbol{\alpha}_i)}{(\boldsymbol{\beta}_{i-1},\boldsymbol{\beta}_{i-1})}\boldsymbol{\beta}_{i-1}\quad(i=1,2,\cdots,n)$$
于是 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 向量组构成了 n 维欧氏空间 V 的一个正交基。

然后,再把 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 单位化。令
$$\boldsymbol{\eta}_1=\frac{\boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|},\boldsymbol{\eta}_2=\frac{\boldsymbol{\beta}_2}{\|\boldsymbol{\beta}_2\|},\cdots,\boldsymbol{\eta}_n=\frac{\boldsymbol{\beta}_n}{\|\boldsymbol{\beta}_n\|}$$
这就得到欧氏空间 V 的一组标准正交基。

欧氏空间中任何一组基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 还可用合同变换的方法变成标准正交基,具体可参考与该教材配套的学习指导书第 1.4 节解题方法指导第 12 个问题(标准正交基准法)。

在标准正交基下,向量的内积的表达式特别简单。

定理 1.5.3　$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n$ 是 n 维欧氏空间 V 的一组标准正交基,$\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 为 V 中任意两个向量,有
$$\boldsymbol{\alpha}=x_1\boldsymbol{\eta}_1+x_2\boldsymbol{\eta}_2+\cdots+x_n\boldsymbol{\eta}_n$$
$$\boldsymbol{\beta}=y_1\boldsymbol{\eta}_1+y_2\boldsymbol{\eta}_2+\cdots+y_n\boldsymbol{\eta}_n$$
则
$$(\boldsymbol{\alpha},\boldsymbol{\beta})=x_1y_1+x_2y_2+\cdots+x_ny_n=\sum_{i=1}^{n}x_iy_i \tag{1.5.13}$$

证　注意到
$$(\boldsymbol{\eta}_i,\boldsymbol{\eta}_j)=\begin{cases}1 & \text{当 } i=j\\ 0 & \text{当 } i\neq j\end{cases}$$
我们有
$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\left(\sum_{i=1}^{n}x_i\boldsymbol{\eta}_i,\ \sum_{j=1}^{n}y_j\boldsymbol{\eta}_j\right)=\sum_{i,j=1}^{n}x_iy_j(\boldsymbol{\eta}_i,\boldsymbol{\eta}_j)=\sum_{i=1}^{n}x_iy_j$$
显然,在欧氏空间中,其标准正交基不是唯一的。

定理 1.5.4　设 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 与 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n$ 是 n 维欧氏空间 V 的两个标准正交基,它们之间的

过渡矩阵是 $\boldsymbol{A}=[a_{ij}]_{n\times n}$,则 $\boldsymbol{A}$ 为正交矩阵。

证 由已知条件,设

$$(\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n)=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

因为 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n$ 是标准正交基,有

$$(\boldsymbol{\eta}_i,\boldsymbol{\eta}_j)=\begin{cases}1 & \text{当 } i=j \\ 0 & \text{当 } i\neq j\end{cases} \tag{1.5.14}$$

矩阵 $\boldsymbol{A}$ 的各列是 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n$ 在标准正交基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 下的坐标,即

$$\begin{aligned}\boldsymbol{\eta}_i &= a_{1i}\boldsymbol{\varepsilon}_1+a_{2i}\boldsymbol{\varepsilon}_2+\cdots+a_{ni}\boldsymbol{\varepsilon}_n \\ \boldsymbol{\eta}_j &= a_{1j}\boldsymbol{\varepsilon}_1+a_{2j}\boldsymbol{\varepsilon}_2+\cdots+a_{nj}\boldsymbol{\varepsilon}_n\end{aligned}\quad (i,j=1,2,\cdots,n)$$

将 $\boldsymbol{\eta}_i$ 与 $\boldsymbol{\eta}_j$ 作内积,由定理 1.5.2 及式(1.5.14) 可得到

$$a_{1i}a_{1j}+a_{2i}a_{2j}+\cdots+a_{ni}a_{nj}=\begin{cases}1, i=j \\ 0, i\neq j\end{cases} \tag{1.5.15}$$

这给出了矩阵 $\boldsymbol{A}$ 的列与列之间的关系。式(1.5.15) 实际上表示出如下的矩阵等式:

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

显然,$\boldsymbol{A}$ 是正交矩阵。

正交矩阵具有以下 5 条性质:

定理 1.5.5

(1) 若 $\boldsymbol{A}$ 是正交矩阵,则 $\boldsymbol{A}$ 的行列式的绝对值为 1,即 $\det\boldsymbol{A}=\pm 1$,即 $\boldsymbol{A}$ 是非奇异矩阵;

(2) 若 $\boldsymbol{A}$ 是正交矩阵,则 $\boldsymbol{A}^{-1}$ 与 $\boldsymbol{A}^{\mathrm{T}}$ 也是正交矩阵;

(3) n 阶实方阵 $\boldsymbol{A}$ 是正交矩阵的充分必要条件是:$\boldsymbol{A}$ 的 n 个列向量及 $\boldsymbol{A}$ 的 n 个行向量均构成 R^n 上的标准正交基;

(4) 若 $\boldsymbol{A}$、$\boldsymbol{B}$ 是 n 阶正交矩阵,则 $\boldsymbol{AB}$ 也是 n 阶正交矩阵;

(5) 若 $\boldsymbol{A}$ 是正交矩阵,则对应的线性变换 $T\boldsymbol{x}=\boldsymbol{A}\boldsymbol{x}$ 是正交变换,它在 R^n 上保持向量内积不变,即对 R^n 中任意两个向量 $\boldsymbol{x}_1$ 与 $\boldsymbol{x}_2$,有 $(\boldsymbol{A}\boldsymbol{x}_1,\boldsymbol{A}\boldsymbol{x}_2)=(\boldsymbol{x}_1,\boldsymbol{x}_2)$,从而保持向量的范数不变,即对 R^n 中任意向量 $\boldsymbol{x}$,有 $\|\boldsymbol{A}\boldsymbol{x}\|=\|\boldsymbol{x}\|$。

以上有些性质在线性代数里提过,应是大家熟知的。正交变换在实际应用中比较广泛,如坐标平面上的旋转变换就是一个正交变换。与性质(5) 有关的还有下面两个结论,请读者自证。

(1) 欧氏空间 V 中线性变换 T 是正交变换的充分必要条件是对 V 中任意两个向量 $\boldsymbol{x}_1$、$\boldsymbol{x}_2$,都有 $(T\boldsymbol{x}_1,T\boldsymbol{x}_2)=(\boldsymbol{x}_1,\boldsymbol{x}_2)$;

(2) 欧氏空间 V 中线性变换 T 是正交变换的充分必要条件是 T 在标准交基下的矩阵是正交矩阵。

*1.5.4 酉空间介绍

欧氏空间是定义了内积后的实线性空间。下面将看到酉空间是定义了内积后的复线性空间。欧氏空间与酉空间统称为内积空间。

定义 1.5.8 设 V 是复数域上的线性空间,若对 V 中任意两个向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$,按某个法则,恒有唯一的一个复数与它们对应,这个复数用记号 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 表示,且满足以下 4 个条件:

(1) 交换律 $(\boldsymbol{\alpha},\boldsymbol{\beta})=\overline{(\boldsymbol{\beta},\boldsymbol{\alpha})}$,其中 $\overline{(\boldsymbol{\beta},\boldsymbol{\alpha})}$ 是 $(\boldsymbol{\beta},\boldsymbol{\alpha})$ 的共轭复数;

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

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

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

这里 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 与 $\boldsymbol{\gamma}$ 是 V 中的任意向量，k 为任意复数，则称 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 为向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的内积，而这样的线性空间称为酉空间（或复内积空间）。

【例 1.5.6】　设向量 $\boldsymbol{\alpha}=(x_1,x_2,\cdots,x_n)$ 与向量 $\boldsymbol{\beta}=(y_1,y_2,\cdots,y_n)$ 是 n 维复线性空间 C^n 中任意两个向量，定义其内积为

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=x_1\overline{y_1}+x_2\overline{y_2}+\cdots+x_n\overline{y_n}\tag{1.5.16}$$

易验证式(1.5.16)满足定义 1.5.8 中的 4 个条件，所以它就是内积，因此 C^n 是一个酉空间。

由式(1.5.16)可得到

$$(\boldsymbol{\alpha},\boldsymbol{\alpha})=x_1\overline{x_1}+x_2\overline{x_2}+\cdots+x_n\overline{x_n}=\sum_{i=1}^{n}|x_i|^2$$

比较酉空间与欧氏空间的定义可看出，除内积 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 为复数或实数的差别外，还有两个定义中条件(1)有差别，其他条件都相同。因此，关于酉空间的讨论与欧氏空间的讨论很相似，有一套平行的理论。下面只列出重要的结论，不做详细论述。

(1) 首先由酉空间内积的定义得到酉空间中内积的性质。

性质 1　设 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 是酉空间 V 中任意两个向量，k 是复数域中任意数，则

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

性质 2　设 $\boldsymbol{\alpha}$ 是酉空间中非零向量，则

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

性质 3　设 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 与 $\boldsymbol{\gamma}$ 是酉空间 V 中任意三个向量，则

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

性质 4　设 V 是酉空间，$\mathbf{C}$ 是复数域，$\boldsymbol{\alpha}_i,\boldsymbol{\beta}_j\in V$，$k_i,l_j\in\mathbf{C}$，则

$$\left(\sum_{i=1}^{n}k_i\boldsymbol{\alpha}_i,\sum_{j=1}^{n}l_j\boldsymbol{\beta}_j\right)=\sum_{i,j=1}^{n}k_i\overline{l_j}(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_j)$$

(2) 和欧氏空间一样，在酉空间中可类似定义向量 $\boldsymbol{\alpha}$ 的长度为 $\sqrt{(\boldsymbol{\alpha},\boldsymbol{\alpha})}$，记做 $\|\boldsymbol{\alpha}\|$（或 $|\boldsymbol{\alpha}|$）。

(3) 柯西 — 许瓦兹(Cauchy-Schwarz)不等式仍成立，即对酉空间中任意两个向量 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 有

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

或

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

当且仅当 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 线性相关时，等号成立。

(4) 应用 Cauchy-Schwarz 不等式可定义酉空间中任意两个非零向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的夹角 θ，有

$$\cos^2\theta=\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})(\boldsymbol{\beta},\boldsymbol{\alpha})}{\|\boldsymbol{\alpha}\|^2\,\|\boldsymbol{\beta}\|^2}$$

当 $(\boldsymbol{\alpha},\boldsymbol{\beta})=0$ 时，称向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 正交或互相垂直。

(5) 在 n 维酉空间中，同样可以定义正交基和标准正交基。若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是 n 维酉空间 V 的两两正交向量组，则它构成 V 的一个正交基，若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 还都是单位向量，则它构成 V 的一个标准正交基。

(6) 任意非零酉空间都存在正交基与标准正交基。

(7) 在酉空间中，任意一组线性无关向量可用施密特正交化方法正交化，再单位化，并可扩充成为 V 的一个标准正交基。

(8) 设 $\boldsymbol{A}$ 为 n 阶复矩阵，用 $\overline{\boldsymbol{A}}$ 表示以 $\boldsymbol{A}$ 的元素的共轭复数作元素的矩阵，若 $\boldsymbol{A}$ 满足 $(\overline{\boldsymbol{A}})^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{A}(\overline{\boldsymbol{A}})^{\mathrm{T}}=\boldsymbol{I}$，记 $(\overline{\boldsymbol{A}})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{H}}$，即若 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{I}$，则 $\boldsymbol{A}$ 称为酉矩阵。

(9) 酉矩阵的逆矩阵是酉矩阵，两个酉矩阵的乘积也是酉矩阵。

(10) 设 T 是酉空间 V 中的线性变换，如果对 V 中任意向量 x 有 $(\boldsymbol{x},\boldsymbol{x})=(T\boldsymbol{x},T\boldsymbol{x})$，则 T 为 V 的酉变换。

(11) 酉空间 V 的线性变换 T 为酉变换的充要条件是，对 V 中任意两个向量 $\boldsymbol{x},\boldsymbol{y}$ 都有 $(\boldsymbol{x},\boldsymbol{y})=$

$(T\boldsymbol{x},T\boldsymbol{y})$。

(12) 酉空间中,酉变换在标准正交基下的矩阵 $\boldsymbol{A}$ 是酉矩阵。

(13) 酉空间中,若矩阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$,则 $\boldsymbol{A}$ 称为厄米特(Hermite)矩阵。若 $\boldsymbol{A}^{\mathrm{H}}=-\boldsymbol{A}$,则称 $\boldsymbol{A}$ 为反 Hermite 矩阵。

习　题　1

1-1　判断以下集合对于所给的线性运算是否构成实数域上的线性空间:

(1) 设 $\boldsymbol{A}$ 是 3 阶实数矩阵,$\boldsymbol{A}$ 的实系数多项式 $f(\boldsymbol{A})$ 的全体,对于矩阵的加法和数与矩阵乘法;

(2) 平面上不平行于某一向量的全体向量所组成的集合,对于通常向量的加法和数与向量的乘法。

1-2　实数域 $\mathbf{R}$ 上的集合

$$V=\{f(x)\}f(x)=a_0+a_1x+\cdots+a_nx^n,\quad a_i\in\mathbf{R},a_n\neq 0\}$$

当次数等于定数 $n(n\geqslant 1)$ 时,对于通常多项式加法和数乘运算,判断 V 是否构成 R 上的线性空间。

1-3　设 V_C 是复数集合 $\mathbf{C}$ 定义在复数域上的线性空间,V_R 是复数集合 $\mathbf{C}$ 定义在实数域上的线性空间,试判断

(1) V_R 中向量组 $2,\mathrm{i}$ 的线性相关性;

(2) V_C 中向量组 $2,\mathrm{i}$ 的线性相关性;

其中 $\mathrm{i}=\sqrt{-1}$。

1-4　设 $V^{3\times 3}$ 为实数域上全体 3 阶实对称矩阵构成的线性空间,求它的基与维数。

1-5　求证向量组 $\boldsymbol{\alpha}_1=(1,1,\cdots,1)^{\mathrm{T}},\boldsymbol{\alpha}_2=(1,\cdots,1,0)^{\mathrm{T}},\cdots,\boldsymbol{\alpha}_n=(1,0,\cdots,0)^{\mathrm{T}}$ 是线性空间 F^n 的一个基,并求向量 $\boldsymbol{\alpha}=(t_1,t_2,\cdots,t_n)^{\mathrm{T}}$ 在该基下的坐标。

1-6　设 $\boldsymbol{\alpha}_1=1,\quad \boldsymbol{\alpha}_2=x-1,\quad \boldsymbol{\alpha}_3=(x-1)^2$;

$\boldsymbol{\beta}_1=2,\quad \boldsymbol{\beta}_2=x-2,\quad \boldsymbol{\beta}_3=(x-2)^2$

是 $F[x]_2$ 中的两个基。求由基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 到基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 的过渡矩阵 $\boldsymbol{P}$。

*1-7　若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 为 n 维线性空间 V 的一个基,试证 $\boldsymbol{\alpha}_1,(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2),\cdots,(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\cdots+\boldsymbol{\alpha}_n)$ 也是 V 的一个基。又若向量 $\boldsymbol{\alpha}$ 在前一个基下的坐标是 $(n,n-1,\cdots,2,1)^{\mathrm{T}}$,求 $\boldsymbol{\alpha}$ 在后一个基下的坐标。(中国科技大学研究生入学试题)

1-8　判断下列各线性空间中,满足一定条件的子集在指定的线性运算下是否构成各线性空间的子空间?

(1) 在 n 维线性空间 F^n 中,满足 $x_1+x_2+\cdots+x_n=0$ 的全体向量 $(x_1,x_2,\cdots,x_n)$ 的集合 U_1 对通常的向量加法与数乘运算能否构成 F^n 的子空间?是几维子空间?

(2) 在 F^n 中,满足 $x_1x_2\cdots x_n=0$ 的全体向量 $(x_1,x_2,\cdots,x_n)$ 的集合 U_2 对通常的向量加法与数乘运算能否构成 F^n 的子空间?

(3) 线性空 R^2 中,子集 $U_3=\{\boldsymbol{\alpha}\mid\boldsymbol{\alpha}=(x,x)\mid,x\in R\}$ 与子集 $U_4=\left\{\boldsymbol{\alpha}\mid\boldsymbol{\alpha}=\left(x,\ \frac{1}{2}x(x+1)\right),x\in\mathbf{R}\right\}$ 对于下面指定的加法与数乘运算能否构成 R^2 的子空间?

其中对于 $\boldsymbol{\alpha}=(x_1,\quad x_2),\boldsymbol{\beta}=(y_1,\quad y_2)$ 及 $k\in\mathbf{R}$,规定:

加法运算 $\boldsymbol{\alpha}\oplus\boldsymbol{\beta}=(x_1+y_1,\ x_2+y_2+x_1y_1)$

数乘运算 $k\odot\boldsymbol{\alpha}=\left\{kx_1,\ kx_2+\frac{1}{2}k(k-1)x_1^2\right\}$

1-9　设线性空间 $R^{2\times 2}$ 的两个子集为:

(1) $U_1=\{\boldsymbol{A}\mid\det\boldsymbol{A}=0,\boldsymbol{A}\in R^{2\times 2}\}$;

(2) $U_2=\{\boldsymbol{B}\mid\boldsymbol{B}^2=\boldsymbol{B},\boldsymbol{B}\in R^{2\times 2}\}$,

证明 U_1 与 U_2 均不构成 $R^{2\times2}$ 的子空间。

* 1-10　U 为实数域上多项式空间 $F[x]_4$ 的子空间

$$U=\{f(x)\mid f(x)=a_0+a_1x+a_2x^2+a_3x^3,\quad a_0+a_1+a_2=0,\quad a_1+a_2+a_3=0\}$$

求它的基与维数。

1-11　在 F^4 中求向量组 $\{\boldsymbol{\alpha}_i\}$ 生成的子空间与由向量组 $\{\boldsymbol{\beta}_i\}$ 生成的子空间的交与和的基及维数

(1) $\begin{cases}\boldsymbol{\alpha}_1=(2,1,3,1)\\ \boldsymbol{\alpha}_2=(1,2,0,1)\end{cases}$，　$\begin{cases}\boldsymbol{\beta}_1=(-1,1,-3,1)\\ \boldsymbol{\beta}_2=(1,1,1,1)\end{cases}$；

(2) $\begin{cases}\boldsymbol{\alpha}_1=(1,2,-1,-2)\\ \boldsymbol{\alpha}_2=(3,1,1,1)\\ \boldsymbol{\alpha}_3=(-1,0,1,-1)\end{cases}$，　$\begin{cases}\boldsymbol{\beta}_1=(2,5,-6,-5)\\ \boldsymbol{\beta}_2=(-1,2,-7,3)\end{cases}$

* 1-12　设线性空间 $R^{m\times n}$ 的两个子空间为

$$U_1=\{\boldsymbol{A}=(a_{ij})\in R^{m\times n}\mid a_{ij}=0,\text{当 } i>j \text{ 时}\},$$
$$U_2=\{\boldsymbol{A}=(a_{ij})\in R^{m\times n}\mid a_{ij}=0,\text{当 } i\leqslant j \text{ 时}\},$$

试证明 $R^{m\times n}=U_1\oplus U_2$。

1-13　判断下面所定义的变换 T，哪些是线性变换，哪些不是？

(1) 数域 F 上的线性空间 V 中，$T(\xi)=\xi+\boldsymbol{\alpha}$，其中 ξ 是 V 中任意一个向量，$\boldsymbol{\alpha}$ 是 V 中一个固定向量；

(2) 在 R^3 中，$T(x_1,x_2,x_3)=(x_1^2,\quad x_2+x_3,\quad x_3^2)$，其中 (x_1,x_2,x_3) 是 R^3 中任意一个向量；

(3) 在 $R^{n\times n}$ 中，$T(\boldsymbol{Z})=\boldsymbol{BZC}$，其中 $\boldsymbol{Z}$ 是 $R^{n\times n}$ 中任意一个 n 阶矩阵，$\boldsymbol{B},\boldsymbol{C}$ 是 $R^{n\times n}$ 中两个固定矩阵。

* 1-14　求证在所有二阶方阵组成的线性空间 $F^{2\times2}$ 中，用给定矩阵 $\begin{bmatrix}a & b\\ c & d\end{bmatrix}$ 左乘和右乘二阶方阵的两个变换 T_1 与 T_2 是 $F^{2\times2}$ 的线性变换，并求这两个线性变换在由以下矩阵

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

组成的基下的矩阵。

1-15　设 $T_1(x_1,x_2)=(x_2,-x_1)$，$T_2=(x_1,x_2)=(x_1,-x_2)$ 是线性空间 F^2 的两个变换，证明 T_1 与 T_2 均是 F^2 的线性变换，并求 T_1+T_2，T_1T_2 及 T_2T_1。

1-16　设 R^3 中的变换 T 定义为

$$T\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}x+y+z\\ 2x-y+z\\ y-z\end{bmatrix}$$

(1) 试证明 T 是 R^3 的一个线性变换；

(2) 求 T 在基 $\boldsymbol{\varepsilon}_1=(1,0,0)^{\mathrm{T}}$，$\boldsymbol{\varepsilon}_2=(0,1,0)^{\mathrm{T}}$，$\boldsymbol{\varepsilon}_3=(0,0,1)^{\mathrm{T}}$ 下的矩阵；

(3) 求 T 在基 $\boldsymbol{\alpha}_1=(1,1,1)^{\mathrm{T}}$，$\boldsymbol{\alpha}_2=(1,-1,2)^{\mathrm{T}}$，$\boldsymbol{\alpha}_3=(0,1,1,)^{\mathrm{T}}$ 下的矩阵。

* 1-17　设三维线性空间 V 的两个基为

(Ⅰ) $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$；　　(Ⅱ) $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$

且 $(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)\boldsymbol{P}$，这里 $\boldsymbol{P}$ 为过渡矩阵，其中

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

V 中线性变换 T 满足以下条件：

$$T(\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2+3\boldsymbol{\alpha}_3)=\boldsymbol{\beta}_1+\boldsymbol{\beta}_2$$
$$T(2\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+2\boldsymbol{\alpha}_3)=\boldsymbol{\beta}_2+\boldsymbol{\beta}_3$$
$$T(\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2+4\boldsymbol{\alpha}_3)=\boldsymbol{\beta}_1+\boldsymbol{\beta}_3$$

试求:(1) T 在基(Ⅱ)下的矩阵。

(2) $T\boldsymbol{\beta}_1$ 在基(Ⅰ)下的坐标。

1-18　已知向量空间 F^3 的一个线性变换 T 为 $T(x_1,x_2,x_3)=(0,x_3,0)$,其中 $x_i\in F(i=1,2,3)$。求 T 的核 $N(T)$ 及 T 的值域 $R(T)$ 的维数及一个基。

*1-19　设 n 维线性空间 V

$$V=\{a_mx^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0\mid m<n,a_i\in F\}$$

T 是 V 的一个线性变换,且对于任意 $f(x)\in V,T(f(x))=xf'(x)-f(x)$。

(1) 求 $N(T)$ 及 $R(T)$;

(2) 证明 $V=N(T)\oplus R(T)$。

*1-20　设 V 是实数域 $\mathbf{R}$ 上全体二阶方阵按通常的加法和数乘构成的线性空间,在 V 中定义一个线性变换 T:$T(\boldsymbol{A})=\boldsymbol{AM}=\boldsymbol{MA}$,其中 $\boldsymbol{A}$ 是 V 中任意一个二阶矩阵,而 $\boldsymbol{M}=\begin{bmatrix}1&2\\0&3\end{bmatrix}$,求 T 的核的维数与一个基。

1-21　设数域 F 上 4 维线性空间 V 中线性变换 T 在基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 下的矩阵为

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

试证明子空间 $W=L(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\ \boldsymbol{\alpha}_1+\boldsymbol{\alpha}_3,\ \boldsymbol{\alpha}_1+\boldsymbol{\alpha}_4)$ 是 T 的不变子空间。

1-22　设 $\boldsymbol{\alpha}=(x_1,x_2),\boldsymbol{\beta}=(y_1,y_2)$ 为二维实线性空间 R^2 中任意两个向量,分别定义实数 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 如下,试判断它们是否为 R^2 中 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的内积,并分别说明 R^2 对所定义的内积是否构成欧氏空间?

(1) $(\boldsymbol{\alpha},\boldsymbol{\beta})=x_1y_2+x_2y_1$;

(2) $(\boldsymbol{\alpha},\boldsymbol{\beta})=\sqrt{x_1^2y_1^2+x_2^2y_2^2}$;

(3) $(\boldsymbol{\alpha},\boldsymbol{\beta})=3x_1y_1+5x_2y_2$。

1-23　设 $\mathbf{R}$ 上线性子空间

$U=\left\{\begin{bmatrix}a&b\\0&a\end{bmatrix}\middle|\,a,b\in\mathbf{R}\right\}$,对 U 中任意的二阶方阵 $\boldsymbol{A}=\begin{bmatrix}a_1&a_2\\0&a_1\end{bmatrix}$,$\boldsymbol{B}=\begin{bmatrix}b_1&b_2\\0&b_1\end{bmatrix}$,定义内积 $(\boldsymbol{A},\boldsymbol{B})=a_1b_1+a_2b_2$,证明 U 对所定义的内积构成欧氏空间。

1-24　设 $C[a,b]$ 是定义在区间 $[a,b]$ 上所有连续实函数所构成的线性空间,对任意的 $f(x)$,$g(x),h(x)\in C[a,b]$,规定 $(f,g)=\int_a^b f(x)g(x)\mathrm{d}x$

(1) 证明线性空间 $C[a,b]$ 对于规定的实数 (f,g) 构成欧氏空间。

(2) 当 $a=0,b=1$,且取 $f(x)=3x,g(x)=x+2l$,问 l 取何值时,$f(x)$ 与 $g(x)$ 正交?

1-25　在 R^4 中

(1) 设 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 按通常定义,求向量 $\boldsymbol{\alpha}=(1,2,2,3)$ 与向量 $\boldsymbol{\beta}=(3,1,5,1)$ 的夹角;

(2) 求一个与向量 $\boldsymbol{\alpha}_1=(1,1,-1,1),\boldsymbol{\alpha}_2=(1,-1,-1,1),\boldsymbol{\alpha}_3=(2,1,1,3)$ 正交的单位向量。

1-26　在三维欧氏空间 R^3 中,对任意向量 $\boldsymbol{\alpha}=(x_1,x_2,x_3),\boldsymbol{\beta}=(y_1,y_2,y_3)\in R^3$,定义内积 $(\boldsymbol{\alpha},\boldsymbol{\beta})=x_1y_1+x_2y_2+x_3y_3$,现给出 R^3 中 3 个向量 $\boldsymbol{\alpha}_1=(-1,0,0),\boldsymbol{\alpha}_2=(-1,-2,0),\boldsymbol{\alpha}_3=(-1,0,-1)$,

(1) 求以 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 为基的度量矩阵;

(2) 用两种方法求 R^3 的标准正交基。

第 2 章　矩阵的相似标准形

n 阶矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 间的关系有：等价（$\boldsymbol{B}=\boldsymbol{PAQ}$），或合同（$\boldsymbol{B}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$），或相似（$\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{AP}$），其中，$\boldsymbol{P}$、$\boldsymbol{Q}$ 为可逆矩阵。我们知道数域 F 上的 n 阶方阵不一定都能相似于一个对角矩阵，对于不能相似于对角矩阵的矩阵，人们也希望能找到它所能相似的矩阵的最简形式，即相似标准形，这样便于了解这个矩阵的性质和进行某些计算。

2.1　相似矩阵

2.1.1　相似矩阵及其性质

定义 2.1.1　设 $\boldsymbol{A}$、$\boldsymbol{B}$ 为 n 阶矩阵，若存在可逆矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{AP}$$

则称矩阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{B}$ 相似，记做 $\boldsymbol{A} \sim \boldsymbol{B}$。

矩阵的相似是一种等价关系，即矩阵的相似关系具有：

自反性　$\boldsymbol{A} \sim \boldsymbol{A}$；

对称性　若 $\boldsymbol{A} \sim \boldsymbol{B}$，则 $\boldsymbol{B} \sim \boldsymbol{A}$；

传递性　若 $\boldsymbol{A} \sim \boldsymbol{B}$，$\boldsymbol{B} \sim \boldsymbol{C}$，则 $\boldsymbol{A} \sim \boldsymbol{C}$。

2.1.2　矩阵与对角矩阵相似的条件

一个矩阵与一个对角矩阵相似的问题，在代数与矩阵论中有重要的意义。为此，我们介绍一个矩阵与一个对角矩阵相似的条件。

定理 2.1.1　n 阶矩阵 $\boldsymbol{A}$ 与一个对角矩阵相似的充分必要条件是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量。

线性代数中曾给过此定理的详细证明，并给出找可逆矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 为对角矩阵的方法，在此，我们不再赘述。

由于 $\boldsymbol{A}$ 属于不同特征值的特征向量必线性无关，所以有如下定理。

定理 2.1.2　若 n 阶矩阵 $\boldsymbol{A}$ 有 n 个不同的特征值，则 $\boldsymbol{A}$ 必可与对角矩阵相似。

此定理是 $\boldsymbol{A}$ 可与对角矩阵相似的充分条件，但不是必要条件。当特征值为单根时，必有一个线性无关的特征向量；当特征值为重根时，可根据以下定理来判断。

定理 2.1.3　n 阶矩阵 $\boldsymbol{A}$ 可与对角矩阵相似的充要条件是对 $\boldsymbol{A}$ 的任意一个 k 重特征值 λ，矩阵 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 的秩 $\mathrm{rank}(\lambda\boldsymbol{I}-\boldsymbol{A}) = n-k$，从而对应于 k 重特征值 λ 恰有 k 个线性无关的特征向量。

【例 2.1.1】　研究下列矩阵是否可与对角矩阵相似。

（1）$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$；

（2）$\boldsymbol{A} = \begin{bmatrix} 5 & 6 & -3 \\ -1 & 0 & 1 \\ 1 & 2 & -1 \end{bmatrix}$；

（3）$\boldsymbol{A} = \begin{bmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{bmatrix}$

解　(1) $|\lambda \boldsymbol{I}-\boldsymbol{A}|=(\lambda+1)^2(\lambda-5)$

属于特征值 $\lambda_1=\lambda_2=-1$ 的与线性无关的特征向量有两个，因为此时

$$\lambda \boldsymbol{I}-\boldsymbol{A}=\begin{bmatrix}-2 & -2 & -2\\ -2 & -2 & -2\\ -2 & -2 & -2\end{bmatrix}$$

$\mathrm{rank}(\lambda \boldsymbol{I}-\boldsymbol{A})=1$，与线性无关的特征向量有 $3-1=2$ 个，从而 $\boldsymbol{A}$ 一定可与对角矩阵 $\begin{bmatrix}-1 & & \\ & -1 & \\ & & 5\end{bmatrix}$ 相似。

(2) $|\lambda \boldsymbol{I}-\boldsymbol{A}|=(\lambda-2)(\lambda^2-2\lambda-2)$

特征值为 $\lambda_1=2,\lambda_2=1+\sqrt{3},\lambda_3=1-\sqrt{3}$，由定理 2.1.2 可知，$\boldsymbol{A}$ 必可与对角矩阵相似。

(3) $|\lambda \boldsymbol{I}-\boldsymbol{A}|=(\lambda-1)^2(\lambda+2)$

特征值为 $\lambda_1=\lambda_2=1,\lambda_3=-2$，对于 $\lambda_1=\lambda_2=1$，有

$$\lambda \boldsymbol{I}-\boldsymbol{A}=\begin{bmatrix}-2 & -1 & 0\\ 4 & 2 & 0\\ -4 & 8 & 3\end{bmatrix}\rightarrow\begin{bmatrix}2 & 1 & 0\\ 0 & 10 & 3\\ 0 & 0 & 0\end{bmatrix}$$

$$\mathrm{rank}(\lambda \boldsymbol{I}-\boldsymbol{A})=2$$

属于 $\lambda_1=\lambda_2=1$ 这个二重特征值的线性无关的特征向量只有 $n-\mathrm{rank}(\lambda \boldsymbol{I}-\boldsymbol{A})=3-2=1$ 个，从而 $\boldsymbol{A}$ 一定不能与对角矩阵相似。

【例 2.1.2】　设

$$\boldsymbol{A}=\begin{bmatrix}4 & 6 & 0\\ -3 & -5 & 0\\ -3 & -6 & 1\end{bmatrix}$$

求 $\boldsymbol{A}^{100}$。

分析：这是一个求矩阵的高次幂的问题，可以将求 $\boldsymbol{A}^{100}$ 的问题转化为求与 $\boldsymbol{A}$ 相似的对角矩阵 $\boldsymbol{B}$ 的 100 次幂的问题，而后者计算要比前者容易得多。

解

$$|\lambda \boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-4 & -6 & 0\\ 3 & \lambda+5 & 0\\ 3 & 6 & \lambda-1\end{vmatrix}=(\lambda+2)(\lambda-1)^2$$

对于 $\lambda_1=-2$，求得属于它的线性无关特征向量

$$\boldsymbol{P}_1=\begin{bmatrix}-1\\ 1\\ 1\end{bmatrix}$$

对于 $\lambda_2=\lambda_3=1$，求得属于它的线性无关特征向量

$$\boldsymbol{P}_2=\begin{bmatrix}-2\\ 1\\ 0\end{bmatrix},\qquad \boldsymbol{P}_3=\begin{bmatrix}0\\ 0\\ 1\end{bmatrix}$$

于是

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

不难求得

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

于是

$$P^{-1}AP=\begin{bmatrix}-2 & & \\ & 1 & \\ & & 1\end{bmatrix},\quad A=P\begin{bmatrix}-2 & & \\ & 1 & \\ & & 1\end{bmatrix}P^{-1},$$

$$A^{100}=P\begin{bmatrix}-2 & & \\ & 1 & \\ & & 1\end{bmatrix}^{100}P^{-1}=\begin{bmatrix}-1 & -2 & 0\\ 1 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}\begin{bmatrix}2^{100} & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 2 & 0\\ -1 & -1 & 0\\ -1 & -2 & 1\end{bmatrix}$$

$$=\begin{bmatrix}-2^{100} & -2 & 0\\ 2^{100} & 1 & 0\\ 2^{100} & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 2 & 0\\ -1 & -1 & 0\\ -1 & -2 & 1\end{bmatrix}=\begin{bmatrix}-2^{100}+2 & -2^{101}+2 & 0\\ 2^{100}-1 & 2^{101}-1 & 0\\ 2^{100}-1 & 2^{101}-2 & 1\end{bmatrix}$$

【例 2.1.3】　证明两个对角矩阵相似的充要条件是对角线上元素相同，仅可能排列位置不同。

证　设 A、B 是两个对角矩阵，且 $A\sim B$，则由相似矩阵的定义可知，必存在可逆矩阵 P，使

$$B=P^{-1}AP$$

从而

$$\lambda I-B=\lambda I-P^{-1}AP=P^{-1}\lambda IP-P^{-1}AP$$
$$=P^{-1}(\lambda I-A)P$$
$$|\lambda I-B|=|P^{-1}||\lambda I-A||P|=|\lambda I-A|$$

当 A、B 为对角矩阵时，上式成立的充要条件是对角线上元素相同，仅可能排列位置不同。

2.1.3　相似不变量

定义 2.1.2　设函数 f 定义在矩阵集合 M 上，若对于任意两个相似的矩阵 A、$B\in M$，有

$$f(A)=f(B)$$

则称 f 为相似不变量。

【例 2.1.4】　证明对任意两个相似矩阵，有以下结论：

(1) 行列式是相似不变量；

(2) 特征多项式是相似不变量；

(3) 特征值是相似不变量；

(4) 矩阵的迹是相似不变量。

证　(1) 若 $A\sim B$，则由定义 2.1.1 可知存在可逆矩阵 P，使得

$$P^{-1}AP=B$$

从而

$$|B|=|P^{-1}AP|=|P^{-1}||A||P|=|A|$$

故行列式是相似不变量。

(2) 由例 2.1.3 可知

$$|\lambda I-B|=|\lambda I-A|$$

故矩阵的特征多项式也是相似不变量。

(3) 由于矩阵的特征值就是它的特征多项式的根，故由(2) 可知矩阵的特征值也是相似不变量。

(4) 若 $A\sim B$，则存在可逆矩阵 P，使得

$$P^{-1}AP=B$$

由于任何 n 阶矩阵 $C=(c_{ij})_{n\times n}$ 和 $D=(d_{ij})_{n\times n}$，矩阵 CD 的迹等于矩阵 DC 的迹，即

$$\mathrm{tr}(\boldsymbol{CD})=\sum_{i=1}^{n}\left(\sum_{j=1}^{n}c_{ij}d_{ji}\right)=\sum_{j=1}^{n}\left(\sum_{i=1}^{n}d_{ji}c_{ij}\right)=\mathrm{tr}(\boldsymbol{DC})$$

所以

$$\mathrm{tr}(\boldsymbol{B})=\mathrm{tr}(\boldsymbol{P}^{-1}\boldsymbol{AP})=\mathrm{tr}[\boldsymbol{P}^{-1}(\boldsymbol{AP})]=\mathrm{tr}[(\boldsymbol{AP})\boldsymbol{P}^{-1}]=\mathrm{tr}[\boldsymbol{A}(\boldsymbol{PP}^{-1})]=\mathrm{tr}(\boldsymbol{A})$$

故矩阵的迹也是相似不变量。

2.2 λ-矩阵及其标准形

2.2.1 λ-矩阵

在矩阵理论中,我们把矩阵定义为数的阵列,即它的元素是数域 F 中的数,统称为数字矩阵。现在,我们把数字矩阵加以推广,设 λ 是数域 F 上的一个未定元,$f(\lambda)$,$g(\lambda)$,…,是 F 上 λ 的多项式,现在我们引进 λ-矩阵。

定义 2.2.1 用未定元 λ 的多项式为元素组成的矩阵

$$(f_{ij}(\lambda))=\begin{bmatrix} f_{11}(\lambda) & f_{12}(\lambda) & \cdots & f_{1n}(\lambda) \\ f_{21}(\lambda) & f_{22}(\lambda) & \cdots & f_{2n}(\lambda) \\ \vdots & \vdots & & \vdots \\ f_{m1}(\lambda) & f_{m2}(\lambda) & \cdots & f_{mn}(\lambda) \end{bmatrix}$$

称为 λ-矩阵。用 $\boldsymbol{A}(\lambda)$,$\boldsymbol{B}(\lambda)$,$\boldsymbol{C}(\lambda)$ 等来表示 λ-矩阵。

例如,方阵 $\boldsymbol{A}$ 的特征矩阵 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 就是一个 λ-矩阵,数字矩阵当然也可看做λ-矩阵的特例。

由于 λ 的多项式的和、差、积仍为 λ 的多项式,因此对于 λ-矩阵自然可同数字矩阵一样定义各种运算。两个 λ-矩阵的和、差、积、相等、数乘完全采用与数字矩阵相同的定义,这里不再赘述。

对于正方 λ-矩阵的行列式、余子式、代数余子式及一般 λ-矩阵的子式也采取与数字矩阵相同的定义。只需对 λ-矩阵的秩、可逆等的相关概念再明确一下。

定义 2.2.2 λ-矩阵 $\boldsymbol{A}(\lambda)$ 中不为零的子式的最高阶数 r 称为 $\boldsymbol{A}(\lambda)$ 的秩,记为 $\mathrm{rank}\boldsymbol{A}(\lambda)=r$,简写为 $r[\boldsymbol{A}(\lambda)]=r$。零矩阵的秩规定为零。

例如

$$\mathrm{rank}\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}=2,\qquad \mathrm{rank}\begin{bmatrix}\lambda & \lambda-1\\ 0 & 0\end{bmatrix}=1$$

当 n 阶 λ-矩阵的秩为 n 时,称该 λ-矩阵为满秩的或非奇异的。

定义 2.2.3 对于 n 阶 λ-矩阵 $\boldsymbol{A}(\lambda)$,若有 n 阶 λ-矩阵 $\boldsymbol{B}(\lambda)$,使

$$\boldsymbol{A}(\lambda)\boldsymbol{B}(\lambda)=\boldsymbol{B}(\lambda)\boldsymbol{A}(\lambda)=\boldsymbol{I}_n$$

则称 $\boldsymbol{A}(\lambda)$ 为可逆的 λ-矩阵,并称 $\boldsymbol{B}(\lambda)$ 为 $\boldsymbol{A}(\lambda)$ 的逆矩阵。

与数字矩阵类似,λ-矩阵 $\boldsymbol{A}(\lambda)$ 若有逆矩阵,一定是唯一的,故可记为 $\boldsymbol{B}(\lambda)=\boldsymbol{A}^{-1}(\lambda)$。

对于数字方阵,满秩即可逆,而对于 λ-矩阵,满秩却不一定可逆。

定理 2.2.1 n 阶 λ-矩阵 $\boldsymbol{A}(\lambda)$ 可逆的充要条件是

$$|\boldsymbol{A}(\lambda)|=d$$

其中,d 为一个非零常数。

证 必要性:若 $\boldsymbol{A}(\lambda)$ 可逆,则

$$\boldsymbol{A}(\lambda)\boldsymbol{A}^{-1}(\lambda)=\boldsymbol{I}_n$$

$$|\boldsymbol{A}(\lambda)||\boldsymbol{A}^{-1}(\lambda)|=|\boldsymbol{I}_n|=1$$

而 $|\boldsymbol{A}(\lambda)|$ 与 $|\boldsymbol{A}^{-1}(\lambda)|$ 都是 λ 的多项式,故只能是零次多项式,即 $|\boldsymbol{A}(\lambda)|=d\neq 0$。

充分性:若 $|\boldsymbol{A}(\lambda)|=d\neq 0$,设 $\boldsymbol{A}^*(\lambda)$ 是 $\boldsymbol{A}(\lambda)$ 的伴随矩阵,它是一个 λ-矩阵,此时

$$
\frac{1}{d}\boldsymbol{A}^{*}(\lambda)=\frac{1}{d}\begin{bmatrix} A_{11}(\lambda) & A_{21}(\lambda) & \cdots & A_{n1}(\lambda) \\ A_{12}(\lambda) & A_{22}(\lambda) & \cdots & A_{n2}(\lambda) \\ \vdots & \vdots & & \vdots \\ A_{1n}(\lambda) & A_{2n}(\lambda) & \cdots & A_{nn}(\lambda) \end{bmatrix}
$$

也是一个 λ- 矩阵，且

$$
\boldsymbol{A}(\lambda)\left[\frac{1}{d}\boldsymbol{A}^{*}(\lambda)\right]=\left[\frac{1}{d}\boldsymbol{A}^{*}(\lambda)\right]\boldsymbol{A}(\lambda)=\frac{1}{d}\begin{bmatrix} d & & & \\ & d & & \\ & & \ddots & \\ & & & d \end{bmatrix}=\boldsymbol{I}_n
$$

故 $\boldsymbol{A}(\lambda)$ 可逆，且逆矩阵为$\frac{1}{d}\boldsymbol{A}^{*}(\lambda)$。

由定理 2.2.1 可看出，一个 n 阶 λ- 矩阵 $\boldsymbol{A}(\lambda)$ 满秩，但不一定可逆，只有当其行列式为非零常数时，$\boldsymbol{A}(\lambda)$ 才可逆。例如：

$$
\boldsymbol{A}(\lambda)=\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix},\qquad \operatorname{rank}[\boldsymbol{A}(\lambda)]=2
$$

$\boldsymbol{A}(\lambda)$ 满秩，但 $|\boldsymbol{A}(\lambda)|=\lambda^2$ 不是非零常数，故 $\boldsymbol{A}(\lambda)$ 不可逆。显然，如果 $\boldsymbol{A}(\lambda)$ 可逆，那么其逆矩阵就不是 λ- 矩阵了。

对 λ- 矩阵也可引入初等变换及两个 λ- 矩阵的等价的概念。

λ- 矩阵的初等变换可以用以下记号表示：

(1) $r_i \leftrightarrow r_j$ 表示互换第 i 行和第 j 行；

(2) kr_i 表示用数 k 乘第 i 行；

(3) $r_i+f(\lambda)r_j$ 表示把第 j 行的 $f(\lambda)$ 倍加到第 i 行。

以上表示对 λ- 矩阵做行初等变换，将 r 换做 c，即表示相应的列初等变换。由单位矩阵经过一次初等变换，便可得初等矩阵。

(1) 互换 i,j 两行(列)，其相应的初等矩阵为

$$
\boldsymbol{E}(i,j)=\begin{array}{c}
\begin{bmatrix}
1 & & & & & & & & & & \\
& \ddots & & & & & & & & & \\
& & 1 & & & & & & & & \\
& & & 0 & & & & 1 & & & \\
& & & & 1 & & & & & & \\
& & & & & \ddots & & & & & \\
& & & & & & 1 & & & & \\
& & & 1 & & & & 0 & & & \\
& & & & & & & & 1 & & \\
& & & & & & & & & \ddots & \\
& & & & & & & & & & 1
\end{bmatrix}
\begin{array}{c} ~\\ ~\\ ~\\ i\\ ~\\ ~\\ ~\\ j\\ ~\\ ~\\ ~ \end{array}\\
\begin{array}{ccccccccccc} ~ & ~ & ~ & i & ~ & ~ & ~ & j & ~ & ~ & ~ \end{array}
\end{array}
$$

(2) 用数 $k\neq 0$ 乘第 i 行(第 i 列)，其相应的初等矩阵为

$$
\boldsymbol{E}[i(k)]=\begin{array}{c}
\begin{bmatrix}
1 & & & & & & \\
& \ddots & & & & & \\
& & 1 & & & & \\
& & & k & & & \\
& & & & 1 & & \\
& & & & & \ddots & \\
& & & & & & 1
\end{bmatrix}
\begin{array}{c} ~\\ ~\\ ~\\ i\\ ~\\ ~\\ ~ \end{array}\\
\begin{array}{ccccccc} ~ & ~ & ~ & i & ~ & ~ & ~ \end{array}
\end{array}
$$

(3) 用任一多项式 $f(\lambda)$ 乘第 j 行(第 i 列) 并加到第 i 行(第 j 列) 上,其相应的初等矩阵为

$$
\boldsymbol{E}[i,jf(\lambda)]=\begin{array}{cc}
\begin{bmatrix}
1&&&&&&\\
&\ddots&&&&&\\
&&1&\cdots&f(\lambda)&&\\
&&&\ddots&\vdots&&\\
&&&&1&&\\
&&&&&\ddots&\\
&&&&&&1
\end{bmatrix}
&
\begin{matrix}\\ \\ (i)\\ \\ (j)\\ \\ \\ \end{matrix}
\\
\begin{matrix}\qquad\qquad(i)\quad(j)\end{matrix} &
\end{array}
$$

这三种初等矩阵都是满秩的,且对 $\boldsymbol{A}(\lambda)$ 施行行初等变换,相当于对 $\boldsymbol{A}(\lambda)$ 左乘一个相应的初等矩阵;对 $\boldsymbol{A}(\lambda)$ 施行列初等变换,相当于对 $\boldsymbol{A}(\lambda)$ 右乘一个相应的初等矩阵,且施行初等行(列) 变换不改变矩阵的秩。

定义 2.2.4 若 λ- 矩阵 $\boldsymbol{A}(\lambda)$ 经过有限次初等变换,可化为 $\boldsymbol{B}(\lambda)$,则称 $\boldsymbol{B}(\lambda)$ 与 $\boldsymbol{A}(\lambda)$ 等价,记做 $\boldsymbol{B}(\lambda)\cong\boldsymbol{A}(\lambda)$。

λ- 矩阵的等价关系与数字矩阵一样,也满足:

自反性 $\boldsymbol{A}(\lambda)\cong\boldsymbol{A}(\lambda)$;

对称性 若 $\boldsymbol{A}(\lambda)\cong\boldsymbol{B}(\lambda)$,则 $\boldsymbol{B}(\lambda)\cong\boldsymbol{A}(\lambda)$;

传递性 若 $\boldsymbol{A}(\lambda)\cong\boldsymbol{B}(\lambda)$,$\boldsymbol{B}(\lambda)\cong\boldsymbol{C}(\lambda)$,则 $\boldsymbol{A}(\lambda)\cong\boldsymbol{C}(\lambda)$。

可以证明:若两个 λ- 矩阵等价,则它们的秩必相等。

2.2.2 λ- 矩阵的标准形

这节将说明任何一个正方 λ- 矩阵都可等价于对角形矩阵。

定义 2.2.5 对角形 λ- 矩阵

$$
\boldsymbol{D}(\lambda)=\begin{bmatrix}
d_1(\lambda)&&&&&&\\
&d_2(\lambda)&&&&&\\
&&\ddots&&&&\\
&&&d_r(\lambda)&&&\\
&&&&0&&\\
&&&&&\ddots&\\
&&&&&&0
\end{bmatrix}
$$

若满足:

(1) $d_i(\lambda)(i=1,2,\cdots,r)$ 是首项系数为 1 的 λ- 多项式;

(2) $d_i(\lambda)\mid d_{i+1}(\lambda)$,$(i=1,2,\cdots,r-1)$,记号 $d_i(\lambda)\mid d_{i+1}(\lambda)$ 表示 $d_{i+1}(\lambda)$ 能被 $d_i(\lambda)$ 整除;

则称 $\boldsymbol{D}(\lambda)$ 为 λ- 矩阵的法式或标准形,史密斯(Smith) 标准形,简称标准形,也称 $\boldsymbol{D}(\lambda)$ 为法 λ- 矩阵。

定理 2.2.2 任一 λ- 方阵 $\boldsymbol{A}(\lambda)$ 都可经若干次初等变换化为标准形,即任一 λ- 方阵都可与一个法 λ- 方阵等价。

证明略。

【例 2.2.1】 化 $\boldsymbol{A}(\lambda)$ 为标准形,其中

$$
\boldsymbol{A}(\lambda)=\begin{bmatrix}
1-\lambda&\lambda^2&\lambda\\
\lambda&\lambda&-\lambda\\
1+\lambda^2&\lambda^2&-\lambda^2
\end{bmatrix}
$$

解 对 $\boldsymbol{A}(\lambda)$ 施行初等变换

$$\boldsymbol{A}(\lambda)=\begin{bmatrix}1-\lambda & \lambda^2 & \lambda\\ \lambda & \lambda & -\lambda\\ 1+\lambda^2 & \lambda^2 & -\lambda^2\end{bmatrix}\xrightarrow{c_1+c_3}\begin{bmatrix}1 & \lambda^2 & \lambda\\ 0 & \lambda & -\lambda\\ 1 & \lambda^2 & -\lambda^2\end{bmatrix}\xrightarrow{r_3-r_1}\begin{bmatrix}1 & \lambda^2 & \lambda\\ 0 & \lambda & -\lambda\\ 0 & 0 & -\lambda^2-\lambda\end{bmatrix}$$

$$\xrightarrow[r_1+r_3]{r_1-\lambda r_2}\begin{bmatrix}1 & 0 & 0\\ 0 & \lambda & -\lambda\\ 0 & 0 & -\lambda^2-\lambda\end{bmatrix}\xrightarrow[(-1)r_3]{c_3+c_2}\begin{bmatrix}1 & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \lambda(\lambda+1)\end{bmatrix}$$

【例 2.2.2】　任一 n 阶可逆 λ- 矩阵 $\boldsymbol{A}(\lambda)$ 的标准形为 $\boldsymbol{I}_n$。

证　**由定理** 2.2.2

$$\boldsymbol{A}(\lambda)\cong\begin{bmatrix}d_1(\lambda) & & & & & & \\ & d_2(\lambda) & & & & & \\ & & \ddots & & & & \\ & & & d_r(\lambda) & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0\end{bmatrix}=\boldsymbol{D}(\lambda)$$

其中 $d_i(\lambda)\mid d_{i+1}(\lambda),(i=1,2,\cdots,r-1)$，**且** $d_i(\lambda)(i=1,2,\cdots,r)$ **为首项系数为** 1 **的** λ- **多项式。**

由初等变换的定义知两个等价的 λ- **矩阵的行列式中只能相差一个不为零的常数倍，且秩相同，故由** $\boldsymbol{A}(\lambda)$ **可逆，可知**

$$\operatorname{rank}[\boldsymbol{D}(\lambda)]=r=n$$

$$|\boldsymbol{D}(\lambda)|=d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda)=d\neq 0$$

又由于 $d_i(\lambda)(i=1,2,\cdots,n)$ **首项系数为** 1，**故** $d_i(\lambda)=1(i=1,2,\cdots,n)$，**即** $\boldsymbol{A}(\lambda)$ **的标准形为** $\boldsymbol{I}_n$。

2.3　不变因子与初等因子

2.3.1　不变因子

定义 2.3.1　**设** λ- **矩阵** $\boldsymbol{A}(\lambda)$ **的秩为** r，**对正整数** k ，$1\leqslant k\leqslant r$，$\boldsymbol{A}(\lambda)$ **中必有非零的** k **阶子式，称** $\boldsymbol{A}(\lambda)$ **中所有** k **阶子式的首项系数为** 1 **的最大公因式为** $\boldsymbol{A}(\lambda)$ **的** k **阶行列式因子，记为** $D_k(\lambda)$。

由定义可知，一个 n **阶** λ- **矩阵** $\boldsymbol{A}(\lambda)$ **的** n **阶行列式因子** $D_n(\lambda)=|\boldsymbol{A}(\lambda)|$，**并且** $D_{k-1}(\lambda)$ **能整除任一个** $k-1$ **阶子式，而由行列式的展开可知一个** k **阶行列式可表示为** k **个** $k-1$ **阶子式的代数和，从而** $D_{k-1}(\lambda)$ **能整除任一个** k **阶子式，因此** $D_{k-1}(\lambda)$ **可以整除** $D_k(\lambda)$，**即** $D_{k-1}(\lambda)\mid D_k(\lambda),(k=1,2,\cdots,n-1)$。

若 $\boldsymbol{A}(\lambda)$ **的秩为** r，**由定义可知它的** r **个行列式因子** $D_1(\lambda),D_2(\lambda),\cdots,D_r(\lambda)$ **不为零，而** $D_{r+1}(\lambda)=0,\cdots,D_n(\lambda)=0$，**于是我们可得到** r **个首项系数为** 1 **的多项式**

$$\begin{aligned}d_1(\lambda)&=D_1(\lambda)\\ d_2(\lambda)&=\frac{D_2(\lambda)}{D_1(\lambda)}\\ d_3(\lambda)&=\frac{D_3(\lambda)}{D_2(\lambda)}\\ &\cdots\cdots\\ d_r(\lambda)&=\frac{D_r(\lambda)}{D_{r-1}(\lambda)}\end{aligned}\tag{2.3.1}$$

定义 2.3.2　**在式**(2.3.1) **中的** r **个多项式** $d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda)$ **称为** $\boldsymbol{A}(\lambda)$ **的不变因子，其中** r **为** $\boldsymbol{A}(\lambda)$ **的秩。**

有了 λ- **矩阵的行列式因子和不变因子，就可以引入两个** λ- **矩阵等价的充要条件。**

定理 2.3.1　设有 λ-矩阵 $\boldsymbol{A}(\lambda)$ 和 $\boldsymbol{B}(\lambda)$，若 $\boldsymbol{A}(\lambda)\cong\boldsymbol{B}(\lambda)$，则它们有相同的行列式因子或不变因子。

证　设 $\boldsymbol{A}(\lambda)\cong\boldsymbol{B}(\lambda)$，即 $\boldsymbol{A}(\lambda)$ 可经若干次初等变换化为 $\boldsymbol{B}(\lambda)$，因此我们只需说明任一种初等变换不改变行列式因子即可。

对于前两种变换，是显然成立的，以下只就第三种变换加以说明。

设将 $\boldsymbol{A}(\lambda)$ 的第 j 行乘以 $f(\lambda)$ 加到第 i 行上，得到 $\boldsymbol{B}(\lambda)$，这时 $\boldsymbol{B}(\lambda)$ 中的 k 阶子式 Δ_k 只可能有以下三种情形。

(1) Δ_k 不含第 i 行。显然此时 Δ_k 为 $\boldsymbol{A}(\lambda)$ 的某一个 k 阶子式。

(2) Δ_k 含第 i 行，不含第 j 行。此时把 Δ_k 分解成两个行列式之和

$$\Delta_k=\begin{vmatrix}\cdots\cdots\\ a_{it_1}(\lambda)+f(\lambda)a_{jt_1}(\lambda)+\cdots+a_{it_k}(\lambda)+f(\lambda)a_{jt_k}(\lambda)\\ \cdots\cdots\end{vmatrix}$$

$$=\begin{vmatrix}\cdots\cdots\\ a_{it_1}(\lambda)\cdots a_{it_k}(\lambda)\\ \cdots\cdots\end{vmatrix}+f(\lambda)\begin{vmatrix}\cdots\cdots\\ a_{jt_1}(\lambda)\cdots a_{jt_k}(\lambda)\\ \cdots\cdots\end{vmatrix}$$

上式中第一个 k 阶子式显然是 $\boldsymbol{A}(\lambda)$ 的一个 k 阶子式，第二个 k 阶子式是 $\boldsymbol{A}(\lambda)$ 的某一个 k 阶子式再冠以一个正负号。

(3) Δ_k 含第 i 行，也含第 j 行。此时把 Δ_k 分解成两个行列式之和后，第一个 k 阶子式是 $\boldsymbol{A}(\lambda)$ 的一个 k 阶子式，第二个 k 阶子式有两行相同，为零。

综合以上三种情况，均有 $\boldsymbol{A}(\lambda)$ 的 k 阶子式的最大公因式 $\boldsymbol{D}_k(\lambda)_{\boldsymbol{A}}$ 必整除 Δ_k，即 $\boldsymbol{D}_k(\lambda)_{\boldsymbol{A}}$ 整除 $\boldsymbol{B}(\lambda)$ 中任意 k 阶子式，从而 $\boldsymbol{D}_k(\lambda)_{\boldsymbol{A}}\mid\boldsymbol{D}_k(\lambda)_{\boldsymbol{B}}$。

反过来也可证明 $\boldsymbol{D}_k(\lambda)_{\boldsymbol{B}}\mid\boldsymbol{D}_k(\lambda)_{\boldsymbol{A}}$，故 $\boldsymbol{D}_k(\lambda)_{\boldsymbol{A}}=\boldsymbol{D}_k(\lambda)_{\boldsymbol{B}}$，从而不变因子也相同。

定理 2.3.2　若

$$\boldsymbol{D}(\lambda)=\begin{bmatrix}d_1(\lambda) &&&&&&\\ & d_2(\lambda) &&&&&\\ && \ddots &&&&\\ &&& d_r(\lambda) &&&\\ &&&& 0 &&\\ &&&&& \ddots &\\ &&&&&& 0\end{bmatrix}$$

为 $\boldsymbol{A}(\lambda)$ 的标准形，则 $d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda)$ 恰为 $\boldsymbol{A}(\lambda)$ 的 r 个不变因子。

证　由定理 2.3.1 可知，$\boldsymbol{A}(\lambda)$ 的不变因子与 $\boldsymbol{D}(\lambda)$ 的不变因子完全相同，而对 $\boldsymbol{D}(\lambda)$ 来说，各阶行列式因子为

$$D_1(\lambda)=d_1(\lambda)$$
$$D_2(\lambda)=d_1(\lambda)d_2(\lambda)$$
$$\cdots\cdots$$
$$D_r(\lambda)=d_1(\lambda)d_2(\lambda)\cdots d_r(\lambda)$$

由此得 $D(\lambda)$ 的不变因子，即 $\boldsymbol{A}(\lambda)$ 的不变因子为

$$d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda)$$

进而可知，两个 λ-矩阵若有相同的行列式因子或不变因子，则它们的标准形相同，从而它们等价。

推论　两个 λ-矩阵等价的充要条件是它们有相同的行列式因子或不变因子。

例 2.2.1 中，$\boldsymbol{A}(\lambda)$ 的标准形为

$$A(\lambda)=\begin{bmatrix}1 & & \\ & \lambda & \\ & & \lambda(\lambda+1)\end{bmatrix}$$

故 $A(\lambda)$ 的不变因子为

$$d_1(\lambda)=1$$
$$d_2(\lambda)=\lambda$$
$$d_3(\lambda)=\lambda(\lambda+1)$$

【例 2.3.1】 求

$$A(\lambda)=\begin{bmatrix}\lambda^3+1 & & \\ & \lambda & \\ & & \lambda^2-1\end{bmatrix}$$

的标准形及不变因子。

解　易求得 $A(\lambda)$ 的行列式因子为

$$D_3(\lambda)=\lambda(\lambda^2-1)(\lambda^3+1)$$
$$D_2(\lambda)=\lambda+1$$
$$D_1(\lambda)=1$$

故其不变因子为

$$d_1(\lambda)=1$$
$$d_2(\lambda)=\lambda+1$$
$$d_3(\lambda)=\lambda(\lambda-1)(\lambda^3+1)$$

从而其标准形为

$$\begin{bmatrix}1 & & \\ & \lambda+1 & \\ & & \lambda(\lambda-1)(\lambda^3+1)\end{bmatrix}$$

2.3.2　初等因子

根据代数基本定理可知，任一复系数多项式在复数域内都可以分解为一次因式的乘积，因此，在复数域上将 λ-矩阵 $A(\lambda)$ 的各不变因子再分解为：

$$\begin{aligned}
d_1(\lambda)&=(\lambda-\lambda_1)^{k_{11}}(\lambda-\lambda_2)^{k_{12}}\cdots(\lambda-\lambda_s)^{k_{1s}}\\
d_2(\lambda)&=(\lambda-\lambda_1)^{k_{21}}(\lambda-\lambda_2)^{k_{22}}\cdots(\lambda-\lambda_s)^{k_{2s}}\\
&\cdots\cdots\\
d_r(\lambda)&=(\lambda-\lambda_1)^{k_{r1}}(\lambda-\lambda_2)^{k_{r2}}\cdots(\lambda-\lambda_s)^{k_{rs}}
\end{aligned}\tag{2.3.2}$$

这里 r 为 $A(\lambda)$ 的秩，$\lambda_1,\lambda_2,\cdots,\lambda_s$ 为 $A(\lambda)$ 互不相等的特征值，且由不变因子的依次整除性，有

$$k_{ij}\geqslant 0\text{，且 } k_{1j}\leqslant k_{2j}\leqslant\cdots\leqslant k_{rj}\quad(j=1,2,\cdots,s)$$

在式(2.3.2)中，所有指数大于零的因子都称为是 $A(\lambda)$ 的初等因子。

定义 2.3.3　在 $A(\lambda)$ 的不变因子的分解式(2.3.2)中，若 $k_{ij}>0$，则称相应的因式 $(\lambda-\lambda_j)^{k_{ij}}$ 为 $A(\lambda)$ 的一个初等因子。初等因子的全体，称为 $A(\lambda)$ 的初等因子组。按定义，在计算初等因子个数时，重复的初等因子应按重数计算。

在例 2.2.1 中，$A(\lambda)$ 的不变因子为 $1,\lambda,\lambda(\lambda+1)$，故其初等因子组为 $\lambda,\lambda,\lambda+1$。

【例 2.3.2】 求 n 阶 λ-矩阵

$$A(\lambda)=\begin{bmatrix}\lambda-a & b_1 & & & \\ & \lambda-a & b_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & b_{n-1} \\ & & & & \lambda-a\end{bmatrix}$$

的初等因子组，其中 $b_1b_2\cdots b_{n-1}\neq 0$。

解 易见 $D_n(\lambda)=(\lambda-a)^n$。删去第 1 列第 n 行后所得的 $n-1$ 阶子式的值为常数 $b_1b_2\cdots b_{n-1}\neq 0$，故 $D_{n-1}(\lambda)=1$。又因为 $D_i(\lambda)\mid D_{i+1}(\lambda)$，故 $D_1(\lambda)=D_2(\lambda)=\cdots=D_{n-2}(\lambda)=1$，故 $A(\lambda)$ 的不变因子为

$$d_1(\lambda)=d_2(\lambda)=\cdots=d_{n-1}(\lambda)=1$$
$$d_n(\lambda)=(\lambda-a)^n$$

从而 $A(\lambda)$ 的初等因子组只有 $(\lambda-a)^n$。

由初等因子的定义可知，若给定 $A(\lambda)$ 的不变因子，则它的初等因子就被唯一地决定了。反之，若给定 $A(\lambda)$ 的秩和初等因子组，不变因子也能被唯一确定。这是因为把 $A(\lambda)$ 的所有初等因子按不同的一次因式分类，并按各因式的降幂排列：

$$\begin{array}{ll}(\lambda-\lambda_1)^{k_{11}}(\lambda-\lambda_1)^{k_{21}}\cdots(\lambda-\lambda_1)^{k_{r1}}, & k_{11}\geqslant k_{21}\geqslant\cdots\geqslant k_{r1}\\ (\lambda-\lambda_2)^{k_{12}}(\lambda-\lambda_2)^{k_{22}}\cdots(\lambda-\lambda_2)^{k_{r2}}, & k_{12}\geqslant k_{22}\geqslant\cdots\geqslant k_{r2}\\ \cdots\cdots & \cdots\cdots\\ (\lambda-\lambda_s)^{k_{1s}}(\lambda-\lambda_s)^{k_{2s}}\cdots(\lambda-\lambda_s)^{k_{rs}}, & k_{1s}\geqslant k_{2s}\geqslant\cdots\geqslant k_{rs}\end{array}$$

在这个表中共有 r 列，每行中因式的个数可能不够 r 个，在不够的地方补上 1，于是每一列的乘积就是一个不变因子，即

$$d_r(\lambda)=(\lambda-\lambda_1)^{k_{11}}(\lambda-\lambda_2)^{k_{12}}\cdots(\lambda-\lambda_s)^{k_{1s}}$$
$$\cdots\cdots$$
$$d_1(\lambda)=(\lambda-\lambda_1)^{k_{r1}}(\lambda-\lambda_2)^{k_{r2}}\cdots(\lambda-\lambda_s)^{k_{rs}}$$

【例 2.3.3】 $A(\lambda)$ 是 6 阶 λ-矩阵，秩为 5，初等因子组为 $\lambda,\lambda^2,\lambda^3,(\lambda+2),(\lambda+2)^3,\lambda-1,\lambda-1$，求其不变因子。

解 将初等因子按降幂排列

$$\begin{array}{l}\lambda^3,\ \lambda^2,\lambda,1,1\\ (\lambda+2)^3,\lambda+2,1,1,1\\ \lambda-1,\lambda-1,1,1,1\end{array}$$

由于 $A(\lambda)$ 的秩为 5，故有 5 个不变因子，且

$$d_5(\lambda)=\lambda^3(\lambda+2)^3(\lambda-1)$$
$$d_4(\lambda)=\lambda^2(\lambda+2)(\lambda-1)$$
$$d_3(\lambda)=\lambda$$
$$d_2(\lambda)=1$$
$$d_1(\lambda)=1$$

前面我们介绍了通过 λ-矩阵 $A(\lambda)$ 的各阶行列式因子，求出 $A(\lambda)$ 的不变因子，进而可求出 $A(\lambda)$ 的初等因子组。实际上，求 $A(\lambda)$ 的初等因子组还有下面较为简洁的方法。

定理 2.3.3 将秩为 r 的对角形 λ-矩阵

$$A(\lambda)=\begin{bmatrix}f_1(\lambda) & & & & & \\ & f_2(\lambda) & & & & \\ & & \ddots & & & \\ & & & f_r(\lambda) & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0\end{bmatrix}$$

中主对角线元素 $f_i(\lambda),(i=1,2,\cdots,r)$ 分解为一次因式的方幂的乘积，则这些一次因式的方幂就是 $\boldsymbol{A}(\lambda)$ 的初等因子组。其中 $f_i(\lambda)$ 为首项系数为 1 的多项式。

定理 2.3.3 中的 $\boldsymbol{A}(\lambda)$ 虽然是对角矩阵，但主对角线上的元素并不要求“依次整除性”，因此，用定理 2.3.3 求初等因子组在实际计算中较为便利。例如，矩阵

$$\begin{bmatrix} \lambda^2+\lambda & & \\ & \lambda & \\ & & (\lambda+1)^2 \end{bmatrix}$$

虽然主对角线上元素并不满足依次整除性。但由定理 2.3.3 可以很方便地求出其初等因子组为

$$\lambda,\lambda,\lambda+1,(\lambda+1)^2$$

定理 2.3.4　设有分块对角形 λ-矩阵

$$\boldsymbol{A}(\lambda)=\boldsymbol{A}_1(\lambda)\oplus\boldsymbol{A}_2(\lambda)\oplus\cdots\oplus\boldsymbol{A}_k(\lambda)$$

则 $\boldsymbol{A}_1(\lambda),\boldsymbol{A}_2(\lambda),\cdots,\boldsymbol{A}_k(\lambda)$ 的初等因子组的全体就是 $\boldsymbol{A}(\lambda)$ 的初等因子组。

定义 2.3.4　设 $\boldsymbol{A}$ 是 n 阶数字矩阵，则称 $\boldsymbol{A}$ 的特征矩阵 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 的行列式因子为 $\boldsymbol{A}$ 的行列式因子，称 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 的不变因子为 $\boldsymbol{A}$ 的不变因子，称 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 的初等因子为 $\boldsymbol{A}$ 的初等因子。

【例 2.3.4】　求

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

的初等因子组。

解

$$\lambda\boldsymbol{I}-\boldsymbol{A}=\begin{bmatrix} \lambda-1 & -2 & 0 \\ 0 & \lambda-2 & 0 \\ 2 & 2 & \lambda+1 \end{bmatrix}$$

$$D_3(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=(\lambda+1)(\lambda-1)(\lambda-2)$$
$$D_2(\lambda)=1$$
$$D_1(\lambda)=1$$

于是，不变因子为

$$d_1(\lambda)=1$$
$$d_2(\lambda)=1$$
$$d_3(\lambda)=(\lambda+1)(\lambda-1)(\lambda-2)$$

初等因子组为：$\lambda+1,\lambda-1,\lambda-2$。

下面说明，行列式因子、不变因子、初等因子都是相似不变量，即有如下定理。

定理 2.3.5　两个相似矩阵的 k 阶行列式因子是相同的。

证　设 n 阶矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似，则存在 n 阶可逆矩阵 $\boldsymbol{P}$，使

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

从而

$$\lambda\boldsymbol{I}-\boldsymbol{B}=\boldsymbol{P}^{-1}(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}$$

如果能够证明 $(\lambda\boldsymbol{I}-\boldsymbol{A})$ 与 $(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}$ 有相同的 k 阶行列式因子，$\boldsymbol{P}^{-1}(\lambda\boldsymbol{I}-\boldsymbol{A})$ 与 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 也有相同的 k 阶行列式因子，从而就可证得 $\boldsymbol{P}^{-1}(\lambda\boldsymbol{I}-\boldsymbol{A})$ 与 $\boldsymbol{P}^{-1}(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}=\lambda\boldsymbol{I}-\boldsymbol{B}$ 的 k 阶行列式因子相同，进而有 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 与 $\lambda\boldsymbol{I}-\boldsymbol{B}$ 的 k 阶行列式因子相同。

为此，用 α_j 表示 $(\lambda\boldsymbol{I}-\boldsymbol{A})$ 的列向量，用 β_j 表示 $(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}$ 的列向量

$$\boldsymbol{\alpha}_j=\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}\qquad \boldsymbol{\beta}_j=\begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix}\qquad (j=1,2,\cdots,n)$$

则由矩阵的运算为

$$(\beta_1,\beta_2,\cdots,\beta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix}$$

即

$$\beta_j=\sum_{k=1}^{n} p_{kj}\alpha_k \qquad (j=1,2,\cdots,n)$$

这个式子表明$(\lambda \boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}$的子行列式为$\lambda \boldsymbol{I}-\boldsymbol{A}$的子行列式的常系数线性组合,故$\lambda \boldsymbol{I}-\boldsymbol{A}$的$k$阶子行列式的所有因式也是$(\lambda \boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}$的同阶子行列式的因式。

由于$[(\lambda \boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}]\boldsymbol{P}^{-1}=\lambda \boldsymbol{I}-\boldsymbol{A}$,所以$(\lambda \boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}$的$k$阶子行列式的所有因式也是$\lambda \boldsymbol{I}-\boldsymbol{A}$的同阶子行列式的因式,从而$\lambda \boldsymbol{I}-\boldsymbol{A}$与$(\lambda \boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}$的$k$阶子行列式的最大公因式相同,即它们的$k$阶行列式因子相同。

类似可以证明$\boldsymbol{P}^{-1}(\lambda \boldsymbol{I}-\boldsymbol{A})$与$\lambda \boldsymbol{I}-\boldsymbol{A}$的$k$阶行列式因子相同,定理得证。

2.4 Jordan 标准形

在 2.1 节中曾指出:一个n阶矩阵不一定能找到与之相似的对角矩阵。但总可以相似于一个形式上比对角矩阵稍复杂的 Jordan 矩阵。Jordan 矩阵在数值计算中经常采用,利用它不仅容易求出矩阵的方幂,还在矩阵函数、矩阵级数、微分方程等方面有着广泛的应用。

2.4.1 矩阵的 Jordan 标准形

定义 2.4.1 形如

$$\boldsymbol{J}=\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & & 1 \\ & & & & \lambda \end{bmatrix}$$

的m阶方阵称为m阶 Jordan 块,其中λ可以是实数,也可以是复数。特别指出,若λ是矩阵$\boldsymbol{A}$的特征值,则称它为$\boldsymbol{A}$的特征值λ的 Jordan 块,例如若矩阵$\boldsymbol{A}$的特征值为 3 和i,则称

$$\begin{bmatrix} 3 & 1 \\ & 3 \end{bmatrix},\begin{bmatrix} i & 1 & \\ & i & 1 \\ & & i \end{bmatrix}$$

为$\boldsymbol{A}$的特征值 3 的 2 阶 Jordan 块和$\boldsymbol{A}$的特征值i的 3 阶 Jordan 块。作为特例,一阶方阵是 1 阶 Jordan 块。

定义 2.4.2 设m_i阶 Jordan 块为

$$\boldsymbol{J}_i=\begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}_{m_i} \qquad (i=1,2,\cdots,s)$$

则这些 Jordan 块的直和

$$\boldsymbol{J}=\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{bmatrix}$$

称为 n 阶 Jordan 矩阵，其中 $n = m_1 + m_2 + \cdots + m_s$。

可见，Jordan 矩阵就是由若干个 Jordan 块组成的分块对角矩阵。例如：

$$\begin{bmatrix} 2 & & & & & \\ & -1 & 1 & & & \\ & & -1 & & & \\ & & & 3 & 1 & \\ & & & & 3 & 1 \\ & & & & & 3 \end{bmatrix}$$

就是一个 6 阶 Jordan 矩阵，它由 3 个 Jordan 块组成。

由于 1 阶 Jordan 块就是 1 阶矩阵，所以对角矩阵是 Jordan 矩阵的特例，例如

$$\begin{bmatrix} 3 & & \\ & 3 & \\ & & -2 \end{bmatrix}$$

可以看做由 3 个 1 阶 Jordan 块组成。

定义 2.4.3　与矩阵 $\boldsymbol{A}$ 相似的任何 Jordan 矩阵，称为矩阵 $\boldsymbol{A}$ 的 Jordan 标准形矩阵，简称 Jordan 标准形。

2.4.2　Jordan 标准形的求法

求 Jordan 标准形，一般常用下面两种方法。

方法 1：　初等因子法

定理 2.4.1　m 阶 Jordan 块

$$\boldsymbol{J}_m = \begin{bmatrix} \lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{bmatrix}_m$$

只有一个初等因子 $(\lambda - \lambda_0)^m$。

定理的证明仿照例 2.3.2 即可。将此结果推广到下面的一般情况。

用 $\boldsymbol{J}_{m_i}(\lambda_i)$ 表示主对角线上元素为 λ_i，阶数为 m_i 的 Jordan 块，则 Jordan 矩阵

$$\boldsymbol{J} = \begin{bmatrix} J_{m_1}(\lambda_1) & & & \\ & J_{m_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{m_s}(\lambda_s) \end{bmatrix}$$

的初等因子组为

$$(\lambda - \lambda_1)^{m_1}, (\lambda - \lambda_2)^{m_2}, \cdots, (\lambda - \lambda_s)^{m_s}$$

定理 2.4.2　n 阶矩阵 $\boldsymbol{A}$ 与 n 阶 Jordan 标准形矩阵 $\boldsymbol{J}$ 相似的充分必要条件是它们的不变因子相同。

证　必要性：若 $\boldsymbol{A}$ 相似于 $\boldsymbol{J}$，则由定理 2.3.5，它们的 k 阶行列式因子相同，从而不变因子也相同。

充分性：设 n 阶矩阵 $\boldsymbol{A}$ 与 n 阶 Jordan 标准形 $\boldsymbol{J}$ 的不变因子 $d_i(\lambda)(i = 1, 2, \cdots, n)$ 相同，则由初等因子的定义，可以求出它们的初等因子组，这样，相应的 n 阶 Jordan 标准型 $\boldsymbol{J}$ 就完全确定了。另外，在 n 维线性空间 V 中，取基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$，把在这组基下以 $\boldsymbol{A}$ 为矩阵的线性变换记做 T，在 V 中再选另一组基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n$，使 T 在这组基下的矩阵是 Jordan 标准形 $\boldsymbol{J}$，由于同一个线性变换 T 在两组不同的基下的矩阵是相似的，所以得到 $\boldsymbol{A}$ 相似于 $\boldsymbol{J}$。

定理 2.4.3 两个 n 阶矩阵相似的充分必要条件是这两个矩阵的初等因子是相同的。

结合定理 2.3.5 和定理 2.4.2,该定理是不难证明的。

定理 2.4.4 设 A 为复数域上任一个 n 阶矩阵,则 A 必可与一个 Jordan 标准形矩阵相似,若不计对角线上各子块的顺序,是唯一的。

证 设 n 阶矩阵 A 的初等因子组为

$$(\lambda-\lambda_1)^{m_1},(\lambda-\lambda_2)^{m_2},\cdots,(\lambda-\lambda_s)^{m_s}$$

其中,$\lambda_1,\lambda_2,\cdots,\lambda_s$ 可能有相同的,指数 $m_1,m_2,\cdots,m_s$ 中也可能有相同的,但总有 $\sum\limits_{i=1}^{s}m_i=n$。

每个初等因子 $(\lambda-\lambda_i)^{m_i}$,对应着一个 Jordan 块 $\boldsymbol{J}_i$,其阶数为 m_i,对角线上元素为 λ_i,即

$$\boldsymbol{J}_i=\begin{bmatrix}\lambda_i & & & \\ & \lambda_i & & \\ & & \ddots & \\ & & & \lambda_i\end{bmatrix}_{m_i}\qquad(i=1,2,\cdots,s)$$

这些 Jordan 块的直和构成一个 Jordan 矩阵 $\boldsymbol{J}$,即

$$\boldsymbol{J}=\begin{bmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s\end{bmatrix}$$

则 $\boldsymbol{J}$ 与 A 的初等因子组相同,故 A 与 $\boldsymbol{J}$ 相似。

若还有另外一个 Jordan 标准形矩阵 $\boldsymbol{J}_1$ 与 A 相似,则 $\boldsymbol{J}_1$ 与 A 有相同的初等因子,从而 $\boldsymbol{J}_1$ 与 $\boldsymbol{J}$ 有相同的初等因子组,从而 $\boldsymbol{J}$ 与 $\boldsymbol{J}_1$ 除了其中对角块的排列次序外是相同的。

推论 复矩阵 A 可对角化的充要条件是 A 的初等因子全为一次因式。

【例 2.4.1】 已知 6 阶矩阵 A 的初等因子组为

$$\lambda-2,(\lambda-1)^2,(\lambda+3)^3$$

求 A 的 Jordan 标准形。

解 初等因子 $\lambda-2$ 对应的 Jordan 块为

$$[2]$$

初等因子 $(\lambda-1)^2$ 对应的 Jordan 块为

$$\begin{bmatrix}1 & 1\\ & 1\end{bmatrix}$$

初等因子 $(\lambda+3)^3$ 对应的 Jordan 块为

$$\begin{bmatrix}-3 & 1 & \\ & -3 & 1\\ & & -3\end{bmatrix}$$

故 A 的 Jordan 标准形为

$$\boldsymbol{J}=\begin{bmatrix}2 & & & & & \\ & 1 & 1 & & & \\ & & 1 & & & \\ & & & -3 & 1 & \\ & & & & -3 & 1\\ & & & & & -3\end{bmatrix}$$

【例 2.4.2】　求矩阵

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

的 Jordan **标准形**。

解　A **的特征矩阵为**

$$\lambda I-A=\begin{bmatrix}\lambda+1&-1&0\\4&\lambda-3&0\\-1&0&\lambda-2\end{bmatrix}$$

行列式因子

$$D_3(\lambda)=|\lambda I-A|=(\lambda-2)(\lambda-1)^2$$

由于有二阶子行列式

$$\begin{vmatrix}\lambda+1&-1\\4&\lambda-3\end{vmatrix}=(\lambda-1)^2$$

$$\begin{vmatrix}4&0\\-1&\lambda-2\end{vmatrix}=4(\lambda-2)$$

故

$$D_2(\lambda)=1,\qquad D_1(\lambda)=1$$

从而 A **的不变因子为**

$$d_1(\lambda)=1$$
$$d_2(\lambda)=1$$
$$d_3(\lambda)=(\lambda-2)(\lambda-1)^2$$

于是,A **的初等因子组为**$(\lambda-2),(\lambda-1)^2$,**故** A **的** Jordan **标准形为**

$$J=\begin{bmatrix}2&&\\&1&1\\&&1\end{bmatrix}$$

【例 2.4.3】　求矩阵

$$A=\begin{bmatrix}-1&-2&6\\-1&0&3\\-1&-1&4\end{bmatrix}$$

的 Jordan **标准形**。

解

$$\lambda I-A=\begin{bmatrix}\lambda+1&2&-6\\1&\lambda&-3\\1&1&\lambda-4\end{bmatrix}\xrightarrow{r_1\leftrightarrow r_3}\begin{bmatrix}1&1&\lambda-4\\1&\lambda&-3\\\lambda+1&2&-6\end{bmatrix}$$

$$\xrightarrow[r_3-(\lambda+1)r_1]{r_2-r_1}\begin{bmatrix}1&1&\lambda-4\\0&\lambda-1&1-\lambda\\0&1-\lambda&-\lambda^2+3\lambda-2\end{bmatrix}$$

$$\xrightarrow[c_3-(\lambda-4)c_1]{c_2-c_1}\begin{bmatrix}1&0&0\\0&\lambda-1&1-\lambda\\0&1-\lambda&-\lambda^2+3\lambda-2\end{bmatrix}$$

$$\xrightarrow{r_3+r_2}\begin{bmatrix}1&0&0\\0&\lambda-1&1-\lambda\\0&0&-(\lambda-1)^2\end{bmatrix}\xrightarrow{c_3+c_2}\begin{bmatrix}1&0&0\\0&\lambda-1&0\\0&0&(\lambda-1)^2\end{bmatrix}$$

则 A 的初等因子组为

$$\lambda-1,\qquad (\lambda-1)^2$$

故 A 的 Jordan 标准形为

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

【例 2.4.4】 求矩阵 A 的 Jordan 标准形,其中 A 为元素皆为 1 的 4 阶矩阵。

解 A 的特征矩阵为

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

则 $\boldsymbol{D}_4(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=\lambda^3(\lambda-4)$ 观察所有 3 阶子行列式,可以看到它无外乎下面两种类型:

$$\begin{vmatrix}\lambda-1&-1&-1\\-1&\lambda-1&-1\\-1&-1&\lambda-1\end{vmatrix}=\lambda^2(\lambda-3)$$

$$\begin{vmatrix}\lambda-1&-1&-1\\-1&\lambda-1&-1\\-1&-1&-1\end{vmatrix}=-\lambda^2$$

故

$$\boldsymbol{D}_3(\lambda)=\lambda^2$$

观察所有 2 阶子行列式,可以看到它无外乎下面三种类型:

$$\begin{vmatrix}\lambda-1&-1\\-1&\lambda-1\end{vmatrix}=\lambda(\lambda-2),\quad \begin{vmatrix}\lambda-1&-1\\-1&-1\end{vmatrix}=-\lambda,\quad \begin{vmatrix}-1&-1\\-1&-1\end{vmatrix}=0$$

故

$$\boldsymbol{D}_2(\lambda)=\lambda$$

显然

$$\boldsymbol{D}_1(\lambda)=1$$

从而 A 的不变因子为

$$d_1(\lambda)=1$$
$$d_2(\lambda)=\lambda$$
$$d_3(\lambda)=\lambda$$
$$d_4(\lambda)=\lambda(\lambda-4)$$

故 A 的初等因子组为:$\lambda,\lambda,\lambda,\lambda-4$。

从而 A 的 Jordan 标准形为

$$\boldsymbol{J}=\begin{bmatrix}0&&&\\&0&&\\&&0&\\&&&4\end{bmatrix}$$

可以看出,当 A 的初等因子全为一次因式时,实际上 A 就相似于一个对角阵 $\boldsymbol{J}$。

方法二:波尔曼法

以下介绍一种求 n 阶矩阵 A 的 Jordan 标准形的较为简便的方法——波尔曼法。其基本步骤如下。

第一步:求出 A 的所有特征值 $\lambda_i(i=1,2,\cdots,n)$。

第二步:对每个不同的特征值 λ_i 和每个 $j(j=1,2,\cdots,n)$ 求矩阵 $(\lambda_i\boldsymbol{I}-\boldsymbol{A})^j$ 的秩,记为

$$r_j(\lambda_i)=r(\lambda_i\boldsymbol{I}-\boldsymbol{A})^j$$

在计算秩时,若对某个 j_0,使

$$r_{j_0}(\lambda_i) = r_{j_0+1}(\lambda_i)$$

则对所有的 $j \geqslant j_0$,都有

$$r_j(\lambda_i) = r_{j_0}(\lambda_i)$$

第三步:对每个 $\lambda_i(i=1,2,\cdots,n)$ 求关于 $\lambda=\lambda_i$ 的约当块的阶数 j 和约当块的个数 $b_j(\lambda_i)$。

$$b_1(\lambda_i) = n - 2r_1(\lambda_i) + r_2(\lambda_i)$$

$$b_j(\lambda_i) = r_{j+1}(\lambda_i) - 2r_j(\lambda_i) + r_{j-1}(\lambda_i), \quad j \geqslant 2$$

这里,需要说明的是,若求出 $b_j(\lambda_i)=s$,则说明有 s 个关于 $\lambda=\lambda_i$ 的 j 阶 Jordan 块。

第四步:写出与 $\boldsymbol{A}$ 相似的 Jordan 标准形,它由 $\boldsymbol{A}$ 的每个特征值 λ_i 的 $b_j(\lambda_i)$ 个关于 $\lambda=\lambda_i$ 的 j 阶 Jordan 块的直和组成。

我们以两个例题来说明波尔曼法。

【例 2.4.5】　用波尔曼法解例 2.4.3 题。

解　第一步:求 $\boldsymbol{A}$ 的特征值

$$|\lambda\boldsymbol{I}-\boldsymbol{A}| = \begin{vmatrix} \lambda+1 & 2 & -6 \\ 1 & \lambda & -3 \\ 1 & 1 & \lambda-4 \end{vmatrix} = (\lambda-1)^3$$

特征值为 $\lambda_1=\lambda_2=\lambda_3=1$。

第二步:求 $(\lambda_i\boldsymbol{I}-\boldsymbol{A})^j$ 的秩

$$r_1(1) = r(\lambda\boldsymbol{I}-\boldsymbol{A}) = r(\boldsymbol{I}-\boldsymbol{A}) = r\begin{bmatrix} 2 & 2 & -6 \\ 1 & 1 & -3 \\ 1 & 1 & -3 \end{bmatrix} = 1$$

$$r_2(1) = r(\boldsymbol{I}-\boldsymbol{A})^2 = r\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 0$$

$$r_3(1) = 0$$

这里 $\lambda=1$ 为一个三重特征值,再无其他特征值,且 $\boldsymbol{A}$ 为 3 阶矩阵,故只求这三个就可以了。

第三步:求 Jordan 块的个数和阶数

$$b_1(1) = n - 2r_1(1) + r_2(1) = 3 - 2\times 1 + 0 = 1$$

$$b_2(1) = r_3(1) - 2r_2(1) + r_1(1) = 0 - 2\times 0 + 1 = 1$$

这说明 $\boldsymbol{A}$ 的约当标准形 $\boldsymbol{J}$ 必有 1 个关于 $\lambda=1$ 的 1 阶 Jordan 块和 1 个关于 $\lambda=1$ 的 2 阶 Jordan 块,它们的直和已是 3 阶,故不必再求 $b_3(1)$ 了。所以 $\boldsymbol{A}$ 的 Jordan 标准形为

$$\boldsymbol{J} = \begin{bmatrix} 1 & & \\ & 1 & 1 \\ & & 1 \end{bmatrix}$$

【例 2.4.6】　用波尔曼法解例 2.4.4 题。

解　$\boldsymbol{A}$ 是元素皆为 1 的 4 阶矩阵,$n=4$。

第一步:求 $\boldsymbol{A}$ 的特征值

$$|\lambda\boldsymbol{I}-\boldsymbol{A}| = \lambda^3(\lambda-4)$$

$$\lambda_1=\lambda_2=\lambda_3=0, \lambda_4=4$$

第二步:求 $(\lambda_i\boldsymbol{I}-\boldsymbol{A})^j$ 的秩

下面分别求 $\lambda=0$ 与 $\lambda=4$ 时,$(\lambda\boldsymbol{I}-\boldsymbol{A})$ 的秩:

$$r_1(0) = r(0\boldsymbol{I}-\boldsymbol{A}) = r(-\boldsymbol{A}) = 1$$

$$r_2(0) = r(0\boldsymbol{I}-\boldsymbol{A})^2 = r(4\boldsymbol{A}) = 1$$

因为$r_1(0)=r_2(0)$,所以对 $j\geqslant 1$,都有 $r_j(0)=r_1(0)$,即有

$$r_3(0)=r_4(0)=1$$

$$r_1(4)=r(4\boldsymbol{I}-\boldsymbol{A})=3$$

$$r_2(4)=r(4\boldsymbol{I}-\boldsymbol{A})^2=3$$

于是

$$r_3(4)=r_4(4)=3$$

第三步:求 Jordan 块的个数和阶数

下面分别求 $\lambda=0$ 与 $\lambda=4$ 时,Jordan 块的个数和阶数

$$b_1(0)=n-2r_1(0)+r_2(0)=4-2+1=3$$

$$b_2(0)=r_3(0)-2r_2(0)+r_1(0)=1-2+1=0$$

$$b_1(4)=n-2r_1(4)+r_2(4)=4-6+3=1$$

$$b_2(4)=r_3(4)-2r_2(4)+r_1(4)=3-6+3=0$$

这说明 $\boldsymbol{A}$ 的约当标准形 $\boldsymbol{J}$ 必有 3 个关于 $\lambda=0$ 的 1 阶 Jordan 块和 1 个关于 $\lambda=4$ 的 1 阶 Jordan 块,故

$$\boldsymbol{J}=\begin{bmatrix}0&&&\\&0&&\\&&0&\\&&&4\end{bmatrix}$$

习 题 2

2-1 下列矩阵能否与对角矩阵相似?若能,求出可逆矩阵 $\boldsymbol{P}$,使 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 为对角矩阵:

(1) $\boldsymbol{A}=\begin{bmatrix}3&4\\5&2\end{bmatrix}$; (2) $\boldsymbol{A}=\begin{bmatrix}5&-3&2\\6&-4&4\\4&-4&5\end{bmatrix}$; (3) $\boldsymbol{A}=\begin{bmatrix}0&1&0\\-4&4&0\\-2&1&2\end{bmatrix}$。

2-2 已知矩阵 $\boldsymbol{A}$ 和向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 分别为

$$\boldsymbol{A}=\begin{bmatrix}3&1&1\\1&2&0\\1&0&2\end{bmatrix},\quad \boldsymbol{\alpha}=\begin{bmatrix}1\\-1\\-1\end{bmatrix},\quad \boldsymbol{\beta}=\begin{bmatrix}1\\2\\2\end{bmatrix}$$

(1) 求 $\boldsymbol{A}$ 的特征值和特征向量;

(2) $\boldsymbol{A}$ 是否可以对角化?若可以,求出 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 为对角矩阵;

(3) 利用 $\boldsymbol{AX}=\lambda\boldsymbol{X}$,计算 $\boldsymbol{A}^{100}\boldsymbol{\alpha}$ 及 $\boldsymbol{A}^{100}\boldsymbol{\beta}$。

*2-3 已知 $\xi=\begin{bmatrix}1\\1\\-1\end{bmatrix}$ 是 $\boldsymbol{A}=\begin{bmatrix}2&-1&2\\5&a&3\\-1&b&-2\end{bmatrix}$ 的特征向量,

(1) 求 a,b 及 ξ 所对应的特征值;

(2) $\boldsymbol{A}$ 是否可以相似于对角矩阵,试述理由。

2-4 化下列 λ-矩阵为标准形:

(1) $\begin{bmatrix}\lambda^2-4&\lambda+2\\\lambda+2&(\lambda+2)^2\end{bmatrix}$; (2) $\begin{bmatrix}\lambda^2+\lambda&0&0\\0&\lambda&0\\0&0&(\lambda+1)^2\end{bmatrix}$;

(3) $\begin{bmatrix}0&0&0&\lambda^2\\0&0&\lambda^2-\lambda&0\\0&(\lambda-1)^2&0&0\\(\lambda^2-\lambda)&0&0&0\end{bmatrix}$; (4) $\begin{bmatrix}\lambda^3-\lambda&2\lambda^2\\\lambda^2+5\lambda&3\lambda\end{bmatrix}$。