第一章　线性空间与线性变换

§1.1　线 性 空 间

线性空间是线性代数的基本概念之一，在大学《工程数学》中已经学过，这里对其加以复习和提高.

1. 线性空间的定义

定义 1　设 V 是非空集合，P 为数域，在 V 中定义了一种代数运算，叫做加法，就是说，给定了一个法则，对于 V 中任意两个元素 α 与 β，在 V 中都有唯一的一个元素 γ 与它们对应，称为 α 与 β 的**和**，记成 $\gamma=\alpha+\beta$. 在数域 P 与集合 V 的元素之间还定义了一种运算，叫做数乘，就是说，对于 P 中任一数 k 与 V 中任一元素 α，在 V 中都有唯一的元素 δ 与它们对应，称为 k 与 α 的**数乘**，记成 $\delta=k\alpha$. 如果加法与数乘满足下述规则，则称 V 为数域 P 上的**线性空间**（有时也称为在 P 上的**向量空间**）.

加法满足下列四条规则：

(1) $\alpha+\beta=\beta+\alpha$;

(2) $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$;

(3) 在 V 中有一个元素 0，使对 V 中任一元素 α 都有

$$\alpha+0=\alpha,$$

具有这个性质的元素 0 称为 V 的**零元素**；

(4) 对于 V 中每一个元素 α，都有 V 中元素 β，使得 $\alpha+\beta=0$，β 称为 α 的**负元素**，记为 $-\alpha$.

数乘满足下列四条规则：

(5) $1\alpha=\alpha$;

(6) $k(l\alpha)=(kl)\alpha$;

(7) $(k+l)\alpha=k\alpha+l\alpha$;

(8) $k(\alpha+\beta)=k\alpha+k\beta$,

其中,k,l 为 P 中任何数,α,β,γ 为 V 中任意元素.

由定义知,几何空间全部向量组成的集合是一个实数域上的线性空间;分量属于数域 P 的全体 n 元数组 $(x_1,x_2,\cdots,x_n)^T$ 构成数域 P 上的一个线性空间,这个线性空间我们常用 P^n 来表示.

设 $x=(x_1,x_2,\cdots,x_n)\in P^n$,$y=(y_1,y_2,\cdots,y_n)\in P^n$,$k\in P$,则有

$$x+y\triangleq(x_1+y_1,x_2+y_2,\cdots,x_n+y_n),$$

$$kx\triangleq(kx_1,kx_2,\cdots,kx_n).$$

当 P 为复数域 C 时,上述线性空间称为 n 元**复向量空间**,记成 C^n;当 P 为实数域 R 时,上述线性空间称为 n 元**实向量空间**,记成 R^n.

例 1 复数域 C 上次数不超过 n 的一元多项式全体 $C_n[x]$,按通常多项式加法和数与多项式乘法,构成一个复数域 C 上的线性空间.

例 2 元素属于复数域 C 的 $m\times n$ 矩阵,按矩阵的加法和矩阵与数的数乘,构成复数域 C 上的线性空间,用 $C^{m\times n}$ 表示.

例 3 全体实函数,按函数的加法和数与函数的数量乘法,构成一个实数域 R 上的线性空间.

例 4 给定 $A\in C^{m\times n}$,记

$$R(A)=\{y\,|\,y=Ax,x\in C^n\},$$

$$N(A)=\{x\,|\,Ax=0,x\in C^n\},$$

按 C^n 中的加法和数乘运算,则 $R(A)$ 和 $N(A)$ 都是复数域 C 上的线性空间.

证 设 $y_1,y_2\in R(A)$,则存在 $x_1,x_2\in C^n$,使得 $y_1=Ax_1$,$y_2=Ax_2$. 又 C^n 为线性空间,故 $x_1+x_2\in C^n$,因此 $A(x_1+x_2)\in R(A)$.

又 $A(x_1+x_2)=Ax_1+Ax_2=y_1+y_2$，故 $y_1+y_2\in R(A)$. 同理，当 $k\in C$ 时有 $ky_1\in R(A)$. 由于 C^n 为线性空间，容易验证 $R(A)$中的加法和数乘满足 8 条规则（这 8 条规则，有时称为线性空间 8 条公理），故 $R(A)$为 C 上的线性的空间.

对于 $N(A)$，也可类似证明. 设 $x_1,x_2\in N(A)$，即 $Ax_1=0,Ax_2=0$，因此 $A(x_1+x_2)=Ax_1+Ax_2=0$，故 $x_1+x_2\in N(A)$；设 $k\in C$，$A(kx_1)=kAx_1=0$，故 $kx_1\in N(A)$. 同样，可验证 $N(A)$中的加法和数乘满足 8 条规则，故 $N(A)$为 C 上的线性空间.

例 5 仅由 C 上线性空间 V 中的零元素 0 构成的单元素集合 $0=\{0\mid 0\in V\}$，按 V 中的运算定义运算，则 0 是 C 上的一个线性空间，称为**零空间**.

事实上，对于 0 中的元素 0，以及 C 中的 k，显然有 $0+0=0$，$k0=0\in 0$，并容易验证它满足 8 条公理. 因此，它是 C 上的线性空间.

注 当 $b\neq 0$ 时，相容的线性非齐次方程组 $Ax=b$ 的解的全体 $S=\{x\mid Ax=b,x\in C^n\}$，按 C^n 中的运算，就不是线性空间. 用反证法，若 S 是线性空间，那么由 $x_1,x_2\in S$ 应有 $x_1+x_2\in S$，但 $A(x_1+x_2)=Ax_1+Ax_2=2b\neq b$，故 $x_1+x_2\overline{\in} S$，所以 S 不是线性空间.

例 6 n 阶线性齐次微分方程

$$L[x]=\frac{\mathrm{d}^n x}{\mathrm{d}t^n}+a_1\frac{\mathrm{d}^{n-1}x}{\mathrm{d}t^{n-1}}+\cdots+a_n x=0$$

的解的全体 $S=\{x(t)\mid L[x]=0\}$，以普通函数的加法、数乘为运算，构成 C 上的线性空间.

2. 基、维与坐标

定义 2 设线性空间 V 中，有 n 个元素 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 满足

(1) $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关；

(2) V 中任一元素 α 总可由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表示；

则 $\alpha_1,\alpha_2,\cdots,\alpha_n$称为线性空间 V 的一个**基**，n 称为线性空间 V 的**维**

数，记为 $\dim(V)=n$.

维数为 n 的线性空间称为 **n 维线性空间**，记为 V_n.

若 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为 V_n 的一个基，则对任意元素 $\alpha\in V_n$，都存在一组有序数 $x_1,x_2,\cdots,x_n$，使得

$$\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n.$$

容易证明，这组数是唯一的. 事实上，若有另一组数 $y_1,y_2,\cdots,y_n$，使得

$$\alpha=y_1\alpha_1+y_2\alpha_2+\cdots+y_n\alpha_n,$$

则有

$$(x_1-y_1)\alpha_1+(x_2-y_2)\alpha_2+\cdots+(x_n-y_n)\alpha_n=0,$$

由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的线性无关性知

$$x_1=y_1,\ x_2=y_2,\ \cdots,\ x_n=y_n.$$

反之，任给一组有序数组 $x_1,x_2,\cdots,x_n$，总有唯一的元素 α 可由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表示为

$$\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n\in V_n.$$

从而可知，若 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为 V_n 的一个基，则 V_n 中元素的全体可表示为

$$V_n=\{\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n\,|\,x_1,x_2,\cdots,x_n\in R\}.$$

这样，V 中的元素 α 与有序数组 $x_1,x_2,\cdots,x_n$ 之间构成一一对应关系. 因此，可用这组有序数表示 α.

定义 3 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为线性空间 V_n 的一个基，对于任一元素 $\alpha\in V_n$，有且仅有一组有序数 $x_1,x_2,\cdots,x_n$，使得

$$\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n,$$

则称 $x_1,x_2,\cdots,x_n$ 为元素 α 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的**坐标**，记作

$$\alpha=(x_1,x_2,\cdots,x_n).$$

例 7 求线性空间 $C_n[x]$的基、维数及向量 p 的坐标.

解 在线性空间 $C_n[x]$中，它的一个基为

$$p_1=1,\ p_2=x,\ p_3=x^2,\ \cdots,\ p_{n+1}=x^n,$$

其维数为 $n+1$. 任何次数不超过 n 的多项式

$$p=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

可表示为

$$p=a_0p_1+a_1p_2+\cdots+a_np_{n+1},$$

因此，p 在这个基下的坐标为

$$p=(a_0,a_1,\cdots,a_{n+1}).$$

如果在 $C_n[x]$中取另一个基

$$p_1'=1,\ p_2'=x-a,\ \cdots,\ p_{n+1}'=(x-a)^n,$$

则把 p 在 $x=a$ 处的按 Taylor 公式展开后，有

$$p=p(a)+p'(a)(x-a)+\cdots+\frac{p^{(n)}(a)}{n!}(x-a)^n,$$

即得 p 在基 $p_1',p_2',\cdots,p_{n-1}'$下的坐标为

$$p=\left(p(a),p'(a),\cdots,\frac{p^{(n)}(a)}{n!}\right).$$

证明留给读者.

例 8 在 n 维线性空间 R^n 中，它的一个基为

$$\varepsilon_1=(1,0,\cdots,0),$$
$$\varepsilon_2=(0,1,\cdots,0),$$
$$\cdots\cdots$$
$$\varepsilon_n=(0,0,\cdots,1).$$

对于任一向量 $\alpha=(a_1,a_2,\cdots,a_n)\in R^n$，有

$$\alpha=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n,$$

所以，$(a_1,a_2,\cdots,a_n)$为向量 α 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标.

可以证明

$$\varepsilon_1'=(1,1,\cdots,1),$$
$$\varepsilon_2'=(0,1,\cdots,1),$$
$$\cdots\cdots$$
$$\varepsilon_n'=(0,0,\cdots,1)$$

也是 R^n 中的一个基. 在基 $\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n'$下，对任一向量 $\alpha=(a_1,a_2,\cdots,a_n)\in R^n$，有

$$\alpha = a_1\varepsilon_1' + (a_2 - a_1)\varepsilon_2' + \cdots + (a_n - a_{n-1})\varepsilon_n'.$$

所以，α 在基 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$ 下的坐标为

$$\alpha = (a_1, a_2 - a_1, \cdots, a_n - a_{n-1}).$$

例 9 线性空间 $C^{m\times n}$ 的一个基为

$$E_{ij} = \begin{array}{c} \begin{pmatrix} & & 0 & & \\ & & \vdots & & \\ 0 & \cdots & 1 & \cdots & 0 \\ & & \vdots & & \\ & & 0 & & \end{pmatrix} \begin{matrix} \\ \\ \}\, i \\ \\ \\ \end{matrix} \\ \underbrace{\qquad\qquad}_{j} \end{array} \qquad (i = 1, 2, \cdots, m; j = 1, 2\cdots, n).$$

在 E_{ij} 中，除第 i 行第 j 列的元素是 1 外，其余元素都是 0.

对于任一 $m\times n$ 矩阵

$$M = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

都可表示为

$$\begin{aligned} M &= a_{11}E_{11} + a_{12}E_{12} + \cdots + a_{1n}E_{1n} + a_{21}E_{21} + \cdots + a_{mn}E_{mn} \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}E_{ij}, \end{aligned}$$

因此，$\dim(C^{m\times n}) = m\times n$，矩阵 $M = (a_{ij})_{m\times n}$ 的坐标为

$$M = (a_{11}, \cdots, a_{1n}, a_{21}, \cdots, a_{2n}, \cdots, a_{m1}, \cdots, a_{mn}).$$

引进了线性空间 V_n 的基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 以后，不仅把 V_n 中抽象的向量 α 与具体的有序数组向量

$$(x_1, x_2, \cdots, x_n)$$

联系起来了，而且还把 V_n 中抽象的线性运算与有序数组向量的线性运算联系起来了.

设 $\alpha, \beta \in V_n$，$\lambda \in R$，记

$$\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n,$$
$$\beta=y_1\alpha_1+y_2\alpha_2+\cdots+y_n\alpha_n.$$

于是有

$$\alpha+\beta=(x_1+y_1)\alpha_1+(x_2+y_2)\alpha_2+\cdots+(x_n+y_n)\alpha_n,$$
$$\lambda\alpha=(\lambda x_1)\alpha_1+(\lambda x_2)\alpha_2+\cdots+(\lambda x_n)\alpha_n,$$

则 $\alpha+\beta$ 的坐标为

$$(x_1+y_1,\cdots,x_n+y_n)=(x_1,\cdots,x_n)+(y_1,\cdots,y_n),$$

$\lambda\alpha$ 的坐标为

$$(\lambda x_1,\lambda x_2,\cdots,\lambda x_n)=\lambda(x_1,x_2,\cdots,x_n).$$

这样，线性空间 V_n 与其对应的坐标空间 R^n 从代数结构上看，就没有本质上的区别了.

定义 4 设 V 与 V^* 同为域 P 上的两个线性空间，若 V 与 V^* 的元素之间可建立一一对应关系

$$x\longleftrightarrow x' \qquad (x\in V, x'\in V^*),$$

且当

$$x\longleftrightarrow x', \quad y\longleftrightarrow y'$$

时，必有

$$x+y\longleftrightarrow x'+y',$$
$$kx\longleftrightarrow kx' \qquad (k\in P),$$

则称在域 P 上的线性空间 V 与 V^* 是**同构的**，且称一一对应为 V 与 V^* 之间的**同构对应**.

显然，任何域 P 上的 n 维线性空间都与 P^n 同构. 因此，域 P 上维数相等的线性空间也都彼此同构.

3. 基变换与坐标变换

众所周知，在线性空间 V_n 中含有 n 个向量的线性无关组不是唯一的，因此它的基是可供选择的. 同一个元素在不同的基下应有不同的坐标，那么不同坐标之间有怎样的关系呢？下面来讨论这个问题.

设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 和 $\beta_1,\beta_2,\cdots,\beta_n$ 是线性空间中的两个不同的基，令

$$
\begin{aligned}
\beta_1 &= p_{11}\alpha_1 + p_{21}\alpha_2 + \cdots + p_{n1}\alpha_n,\\
\beta_2 &= p_{12}\alpha_1 + p_{22}\alpha_2 + \cdots + p_{n2}\alpha_n,\\
&\cdots\cdots\\
\beta_n &= p_{1n}\alpha_1 + p_{2n}\alpha_2 + \cdots + p_{nn}\alpha_n.
\end{aligned}
\tag{1.1}
$$

把 n 个有序元素 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 记作 $(\alpha_1,\alpha_2,\cdots,\alpha_n)$，利用向量和矩阵的形式，将(1.1)式表示为

$$
\begin{pmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_n\end{pmatrix}=\begin{pmatrix}p_{11} & p_{21} & \cdots & p_{n1}\\ p_{12} & p_{22} & \cdots & p_{n2}\\ \cdots & & & \\ p_{1n} & p_{2n} & \cdots & p_{nn}\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix}=P^T\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix}
$$

或

$$
(\beta_1,\beta_2,\cdots,\beta_n) = (\alpha_1,\alpha_2,\cdots,\alpha_n)P, \tag{1.2}
$$

其中

$$
P=\begin{pmatrix}p_{11} & p_{12} & \cdots & p_{1n}\\ p_{21} & p_{22} & \cdots & p_{2n}\\ \cdots & & & \\ p_{n1} & p_{n2} & \cdots & p_{nn}\end{pmatrix}.
$$

(1.1)式或(1.2)式称为**基变换公式**. 矩阵 P 称为基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 到基 $\beta_1,\beta_2,\cdots,\beta_n$ 的**过渡矩阵**.

由于 $\beta_1,\beta_2,\cdots,\beta_n$ 线性无关，故过渡矩阵 P 为可逆矩阵.

定理 1 设元素 $\alpha\in V_n$，它在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的坐标为 $(x_1,x_2,\cdots,x_n)$，在基 $\beta_1,\beta_2,\cdots,\beta_n$ 下的坐标为 $(x_1',x_2',\cdots,x_n')$，若两个基满足关系式(1.2)，则有坐标变换公式

$$
\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}=P\begin{pmatrix}x_1'\\ x_2'\\ \vdots\\ x_n'\end{pmatrix} \quad 或 \quad \begin{pmatrix}x_1'\\ x_2'\\ \vdots\\ x_n'\end{pmatrix}=P^{-1}\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}. \tag{1.3}
$$

证 因为

$$(\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\alpha=(\beta_1,\beta_2,\cdots,\beta_n)\begin{pmatrix}x_1'\\x_2'\\\vdots\\x_n'\end{pmatrix}$$

$$=(\alpha_1,\alpha_2,\cdots,\alpha_n)P\begin{pmatrix}x_1'\\x_2'\\\vdots\\x_n'\end{pmatrix},$$

由于 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关，故(1.3)式成立.

定理1的逆命题也成立. 若线性空间 V_n 中任一元素的两种坐标满足变换公式(1.3)，则两个基满足基变换公式(1.2).

例 10 设 R^4 空间中的向量 α 在基 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 下的表达式为

$$\alpha=\alpha_1-2\alpha_2+3\alpha_3+\alpha_4,$$

求 α 在基 $\beta_1,\beta_2,\beta_3,\beta_4$ 下的坐标，这里

$$\beta_1=\alpha_1+3\alpha_2-5\alpha_3+7\alpha_4$$

$$\beta_2=\alpha_2+2\alpha_3-3\alpha_4,$$

$$\beta_3=\alpha_3+2\alpha_4,$$

$$\beta_4=\alpha_4.$$

解 从基 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 到基 $\beta_1,\beta_2,\beta_3,\beta_4$ 的过渡矩阵为

$$P=\begin{pmatrix}1&0&0&0\\3&1&0&0\\-5&2&1&0\\7&-3&2&1\end{pmatrix},$$

由此可得

$$P^{-1}=\begin{pmatrix}1&0&0&0\\-3&1&0&0\\11&-2&1&0\\-38&7&-2&1\end{pmatrix}.$$

于是，由(1.3)式可得 α 在基 $\beta_1,\beta_2,\beta_3,\beta_4$ 下的坐标为

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \end{pmatrix} = P^{-1}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 11 & -2 & 1 & 0 \\ -38 & 7 & -2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -5 \\ 18 \\ -57 \end{pmatrix},$$

因而，α 在基 $\beta_1,\beta_2,\beta_3,\beta_4$ 下的表达式为

$$\alpha = \beta_1 - 5\beta_2 + 18\beta_3 - 57\beta_4.$$

§1.2 线性子空间

1. 子空间

定义 1 设 V 是线性空间，S 为 V 中的一个非空子集. 如果 S 对于 V 中所定义的加法和数乘两种运算也构成一个线性空间，则称 S 为 V 的一个**线性子空间**，简称**子空间**.

定理 1 线性空间 V 的非空子集 S 构成子空间的充分必要条件是，S 对 V 中的线性运算具有封闭性.

子空间既然也是一个线性空间，上面讲到的基、维数、坐标等概念也可用到子空间中. 由于子空间中不可能比整个线性空间 V 中有更多数目的线性无关的向量，所以 V 的任意一个子空间 S 的维数不可能超过 V 的维数，即

$$\dim(S) \leqslant \dim(V).$$

显然，线性空间 V 是它自身的一个子空间.

在线性空间中，由单个零向量构成的非空子集是一个线性子空间，称为**零子空间**，记作 0，其维数规定为 0，即 $\dim(0)=0$.

零子空间和线性空间 V 自身这两个子空间叫做 V 的**平凡子空间**，而 V 中其它线性子空间叫做**非平凡子空间**.

例 1 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是线性空间 V 中的一组向量，$k_1,k_2,\cdots,k_s$ 是任意一组数，它们的线性组合

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s$$

构成的集合是非空集，而且对线性运算是封闭的. 因此，它是 V 的一个子空间 S，并称 S 是由向量 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 所**生成的子空间**，记作

$$S=\mathrm{Span}\{\alpha_1,\alpha_2,\cdots,\alpha_s\}.$$

显然，若 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关，则 $\dim(S)=s$.

例 1 不仅给出了构造线性子空间 S 的一个方法，而且在有限维空间中，任一子空间都可以由这个方法得到.

例 2 在 $C_n[x]$ 中次数不高于 $r-1$ $(r\leqslant n)$ 的多项式的全体构成一个 r 维子空间. 显然

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

是该子空间的一个基，子空间可表示为

$$S=\mathrm{Span}\{1,x,x^2,\cdots,x^{r-1}\}.$$

例 3 在 R^n 中前 k 个分量为 0 的一切 n 元数组

$$(0,\cdots,0,x_{k+1},\cdots,x_n)\qquad(k\leqslant n)$$

的全体构成一个 $n-k$ 维子空间. 记

$$\begin{aligned}\varepsilon_1&=(1,0,0,\cdots,0,0),\\ \varepsilon_2&=(0,1,0,\cdots,0,0),\\ &\cdots\cdots\\ \varepsilon_n&=(0,0,0,\cdots,0,1),\end{aligned}$$

则上述子空间可表示为

$$S=\mathrm{Span}\{\varepsilon_{k+1},\cdots,\varepsilon_n\}.$$

例 4 齐次线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0,\\ \cdots\cdots\\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0\end{cases}$$

的全体解向量构成一个子空间. 这个子空间叫做这个齐次线性方

程的**解空间**.解空间的基就是方程组的**基础解系**,解空间的维数为 $n-r$,其中 r 是方程组系数矩阵的秩.

2. 子空间的交与和

定理 2 设 V_1,V_2都是线性空间 V 的子空间,用 $V_1\cap V_2$ 表示 V_1 与 V_2 中公共元素的集合,则 $V_1\cap V_2$ 也是 V 的子空间,称为子空间 V_1 与 V_2 的**交**.

证 由于 $0\in V_1,0\in V_2$,所以 $0\in V_1\cap V_2$,从而 $V_1\cap V_2$ 是非空集.

若 $\alpha,\beta\in V_1\cap V_2$,则 $\alpha,\beta\in V_1$ 且 $\alpha,\beta\in V_2$.由于 V_1,V_2 是子空间,故

$$\alpha+\beta\in V_1,\quad k\alpha\in V_1,$$
$$\alpha+\beta\in V_2,\quad k\alpha\in V_2,$$

因此

$$\alpha+\beta\in V_1\cap V_2;$$
$$k\alpha\in V_1\cap V_2.$$

所以,$V_1\cap V_2$ 是 V 的子空间.

例 5 V_1,V_2 分别表示齐次线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0,\\ \cdots\cdots\\ a_{s1}x_1+a_{s2}x_2+\cdots+a_{sn}x_n=0\end{cases}$$

和

$$\begin{cases}b_{11}x_1+b_{12}x_2+\cdots+b_{1n}x_n=0,\\ b_{21}x_1+b_{22}x_2+\cdots+b_{2n}x_n=0,\\ \cdots\cdots\\ b_{t1}x_1+b_{t2}x_2+\cdots+b_{tn}x_n=0\end{cases}$$

的解空间,则 $V_1\cap V_2$ 就是齐次方程组

$$
\begin{cases}
a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0, \\
a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0, \\
\cdots\cdots \\
a_{s1}x_1+a_{s2}x_2+\cdots+a_{sn}x_n=0, \\
b_{11}x_1+b_{12}x_2+\cdots+b_{1n}x_n=0, \\
b_{21}x_1+b_{22}x_2+\cdots+b_{2n}x_n=0, \\
\cdots\cdots \\
b_{t1}x_1+b_{t2}x_2+\cdots+b_{tn}x_n=0
\end{cases}
$$

的解空间.

定理 3　设 V_1,V_2 是线性空间 V 的子空间,用 V_1+V_2 表示形如 $\alpha_1+\alpha_2$(其中 $\alpha_1\in V_1$,$\alpha_2\in V_2$)的向量组成的集合. 则 V_1+V_2 也是 V 的子空间,称为 V_1 与 V_2 的**和**.

证　显然,V_1+V_2 是非空子集. 若 $\alpha,\beta\in V_1+V_2$,则 α,β 可表示为

$$
\alpha=\alpha_1+\alpha_2 \qquad (\alpha_1\in V_1,\alpha_2\in V_2),
$$
$$
\beta=\beta_1+\beta_2 \qquad (\beta_1\in V_1,\beta_2\in V_2),
$$

于是

$$
\alpha+\beta=(\alpha_1+\beta_1)+(\alpha_2+\beta_2),
$$
$$
k\alpha=k\alpha_1+k\alpha_2.
$$

由于 V_1,V_2 都是子空间,所以

$$
\alpha_1+\beta_1\in V_1, \quad \alpha_2+\beta_2\in V_2,
$$
$$
k\alpha_1\in V_1, \quad k\alpha_2\in V_2,
$$

按照 V_1+V_2 的定义可知

$$
\alpha+\beta\in V_1+V_2, \quad k\alpha\in V_1+V_2,
$$

因此,V_1+V_2 是子空间.

在线性空间 V 中,V_1 和 V_2 分别是向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 及 β_1,$\beta_2,\cdots,\beta_t$ 生成的子空间,它们的和记作

$$
V_1+V_2=\mathrm{Span}\{\alpha_1,\alpha_2,\cdots,\alpha_s\}+\mathrm{Span}\{\beta_1,\beta_2,\cdots,\beta_t\}
$$

$$=\mathrm{Span}\{\alpha_1,\alpha_2,\cdots,\alpha_s,\beta_1,\beta_2,\cdots,\beta_t\},$$

而它们的交记为

$$\begin{aligned}V_1\cap V_2&=\mathrm{Span}\{\alpha_1,\alpha_2,\cdots,\alpha_s\}\cap\mathrm{Span}\{\beta_1,\beta_2,\cdots,\beta_t\}\\&=\mathrm{Span}\{\gamma_1,\gamma_2,\cdots,\gamma_r\}\qquad(r\leqslant s,t),\end{aligned}$$

其中 $\gamma_1,\gamma_2,\cdots,\gamma_r$ 是 V_1,V_2 的公共元素.

例 6 设 $\alpha_1=(1,2,1,0)$,$\alpha_2=(-1,1,1,1)$,$\beta_1=(2,-1,0,1)$,$\beta_2=(1,-1,3,7)$,求 $V_1=\mathrm{Span}\{\alpha_1,\alpha_2\}$ 与 $V_2=\mathrm{Span}\{\beta_1,\beta_2\}$的和及交的维数和它们的基.

解 因为和

$$\begin{aligned}V_1+V_2&=\mathrm{Span}\{\alpha_1,\alpha_2\}+\mathrm{Span}\{\beta_1,\beta_2\}\\&=\mathrm{Span}\{\alpha_1,\alpha_2,\beta_1,\beta_2\},\end{aligned}$$

向量组 $\alpha_1,\alpha_2,\beta_1,\beta_2$ 的秩为 3,且 $\alpha_1,\alpha_2,\beta_1$ 是它的一个极大线性无关组,所以

$$\dim(V_1+V_2)=3,$$

而 $\alpha_1,\alpha_2,\beta_1$ 是 V_1+V_2 的一个基,即

$$V_1+V_2=\mathrm{Span}\{\alpha_1,\alpha_2,\beta_1\}.$$

下面求交 $V_1\cap V_2$ 的基,设向量 $\alpha\in V_1\cap V_2$,则有 k_1,k_2,l_1,l_2,使得

$$\alpha=k_1\alpha_1+k_2\alpha_2=l_1\beta_1+l_2\beta_2,$$

即

$$k_1\alpha_1+k_2\alpha_2-l_1\beta_1-l_2\beta_2=0,$$

由此得到关于 k_1,k_2,l_1,l_2 的齐次线性方程组

$$\begin{cases}k_1-k_2-2l_1-l_2=0,\\2k_1+k_2+l_1+l_2=0,\\k_1+k_2-3l_2=0,\\k_2-l_1-7l_2=0,\end{cases}$$

其基础解系为$(1,-4,3,-1)^T$,即

$$k_1=1,\quad k_2=-4,\quad l_1=3,\quad l_2=-1.$$

又因为

$$\alpha=\alpha_1-4\alpha_2=3\beta_1-\beta_2=(5,-2,-3,-4),$$

故

$$\dim(V_1\cap V_2)=1,$$

而 $\alpha=(5,-2,-3,-4)$是 $V_1\cap V_2$ 的一个基,即

$$V_1\cap V_2=\mathrm{Span}\{\alpha\}.$$

关于子空间的下列定理是经常用到的:

定理 4 两个向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 及 $\beta_1,\beta_2,\cdots,\beta_t$ 张成相同子空间的充要条件是这两个向量组等价,即这两个向量组可相互线性表示.

证 必要性. 若 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 及 $\beta_1,\beta_2,\cdots,\beta_t$ 张成相同的子空间,即

$$\mathrm{Span}\{\alpha_1,\alpha_2,\cdots,\alpha_s\}=\mathrm{Span}\{\beta_1,\beta_2,\cdots,\beta_t\},$$

则每个 $\alpha_i(i=1,2,\cdots,s)$作为 $\mathrm{Span}\{\beta_1,\beta_2,\cdots,\beta_t\}$中的向量,都可以由向量组 $\beta_1,\beta_2,\cdots,\beta_t$ 线性表示;同样,每个 $\beta_j(j=1,2,\cdots,t)$作为 $\mathrm{Span}\{\alpha_1,\alpha_2,\cdots,\alpha_s\}$中的向量,都可由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表示,所以这两个向量组等价.

充分性. 若两个向量组等价,则 $\mathrm{Span}\{\alpha_1,\alpha_2,\cdots,\alpha_s\}$中的向量都是 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的线性组合,它们都可由 $\beta_1,\beta_2,\cdots,\beta_t$ 线性表示;同理,$\mathrm{Span}\{\beta_1,\beta_2,\cdots,\beta_t\}$中的每个向量也都可由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表示. 因此,它们张成的子空间相同,即

$$\mathrm{Span}\{\alpha_1,\alpha_2,\cdots,\alpha_s\}=\mathrm{Span}\{\beta_1,\beta_2,\cdots,\beta_t\}.$$

定理 5 $V_1=\mathrm{Span}\{\alpha_1,\alpha_2,\cdots,\alpha_s\}$的维数 $\dim(V_1)$等于向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的秩.

证 设向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的秩为 r,并设 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 为它的一个极大线性无关组,则 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 与 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 等价. 因此 $\mathrm{Span}\{\alpha_1,\alpha_2,\cdots,\alpha_s\}$中的每个向量都可由 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 线性表示,根据定义,$\mathrm{Span}\{\alpha_1,\alpha_2,\cdots,\alpha_s\}$的维数为 r.

定理 6 设 V_r 是 n 维线性空间 V_n 的一个 r 维子空间,$\alpha_1,\alpha_2,$

$\cdots,\alpha_r$ 是 V_r 的一个基，则这组向量可以扩充为整个空间的基，也就是说，在 V_n 中可以找到 $n-r$ 个向量，使得

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r,\ \alpha_{r+1},\ \cdots,\ \alpha_n$$

为 V_n 的一个基.

证明留作练习.

关于两个子空间交与和的维数，有下面的定理：

定理 7(维数公式) 若 V_1,V_2 为线性空间 V 的子空间，则

$$\dim(V_1+V_2)=\dim(V_1)+\dim(V_2)-\dim(V_1\cap V_2).$$

证 设 $V_1,V_2,V_1\cap V_2$ 的维数依次为 n_1,n_2,r. 取 $V_1\cap V_2$ 的一个基为

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r,$$

由定理 6 知，它们可以扩充为 V_1 的一个基

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r,\ \beta_1,\ \beta_2,\ \cdots,\ \beta_{n_1-r},$$

也可以扩充成 V_2 的一个基

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r,\ \gamma_1,\ \gamma_2,\ \cdots,\ \gamma_{n_2-r}.$$

下面证明向量组

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r,\ \beta_1,\ \beta_2,\ \cdots,\ \beta_{n_1-r},\ \gamma_1,\ \gamma_2,\ \cdots,\ \gamma_{n_2-r}$$

是 V_1+V_2 的一个基. 由于

$$V_1=\mathrm{Span}\{\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_1,\ \beta_2,\ \cdots,\ \beta_{n_1-r}\},$$

$$V_2=\mathrm{Span}\{\alpha_1,\alpha_2,\cdots,\alpha_r,\gamma_1,\ \gamma_2,\ \cdots,\ \gamma_{n_2-r}\},$$

所以

$$V_1+V_2=\mathrm{Span}\{\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_1,\beta_2,\cdots,\beta_{n_1-r},\gamma_1,\gamma_2,\cdots,\gamma_{n_2-r}\}.$$

现在只须证明

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r,\ \beta_1,\ \beta_2,\ \cdots,\ \beta_{n_1-r},\ \gamma_1,\ \gamma_2,\ \cdots,\ \gamma_{n_2-r}$$

线性无关即可. 设

$$h_1\alpha_1+h_2\alpha_2+\cdots+h_r\alpha_r+k_1\beta_1+k_2\beta_2+\cdots+k_{n_1-r}\beta_{n_1-r}$$
$$+l_1\gamma_1+l_2\gamma_2+\cdots+l_{n_2-r}\gamma_{n_2-r}=0,$$

则

$$\alpha = h_1\alpha_1 + h_2\alpha_2 + \cdots + h_r\alpha_r + k_1\beta_1 + k_2\beta_2 + \cdots + k_{n_1-r}\beta_{n_1-r}$$
$$= -l_1\gamma_1 - l_2\gamma_2 - \cdots - l_{n_2-r}\gamma_{n_2-r}.$$

由上式可知 $\alpha\in V_1$，$\alpha\in V_2$，因此 $\alpha\in V_1\cap V_2$，故 α 可由 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表示. 设

$$\alpha = p_1\alpha_1 + p_2\alpha_2 + \cdots + p_r\alpha_r,$$

则

$$\alpha = p_1\alpha_1 + p_2\alpha_2 + \cdots + p_r\alpha_r = -l_1\gamma_1 - l_2\gamma_2 - \cdots - l_{n_2-r}\gamma_{n_2-r},$$

即

$$p_1\alpha_1 - p_2\alpha_2 + \cdots + p_r\alpha_r + l_1\gamma_1 + l_2\gamma_2 + \cdots + l_{n_2-r}\gamma_{n_2-r} = 0,$$

由于 $\alpha_1,\alpha_2,\cdots,\alpha_r,\gamma_1,\gamma_2,\cdots,\gamma_{n_2-r}$ 线性无关，故

$$p_1 = p_2 = \cdots = p_r = l_1 = l_2 = \cdots = l_{n_2-r} = 0,$$

从而 $\alpha=0$. 于是有

$$h_1\alpha_1 + h_2\alpha_2 + \cdots + h_r\alpha_r + k_1\beta_1 + k_2\beta_2 + \cdots + k_{n_1-r}\beta_{n_1-r} = 0.$$

再由 $\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_1,\beta_2,\cdots,\beta_{n_1-r}$ 的线性无关性，得

$$h_1 = h_2 = \cdots = h_r = k_1 = k_2 = \cdots = k_{n_1-r} = 0.$$

所以，$\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_1,\beta_2,\cdots,\beta_{n_1-r},\gamma_1,\gamma_2,\cdots,\gamma_{n_2-r}$ 线性无关，故它们是 V_1+V_2 的一个基. 因此

$$\begin{aligned}\dim(V_1+V_2) &= n_1+n_2-r \\ &= \dim(V_1)+\dim(V_2)-\dim(V_1\cap V_2).\end{aligned}$$

例 6 就是定理 7 的具体例子.

推论 如果 n 维线性空间 V_n 的两个子空间 V_1、V_2 的维数之和大于 n，则 V_1、V_2 必含有非零的公共向量.

证 由假设有

$$\begin{aligned}\dim(V_1\cap V_2) &= \dim(V_1)+\dim(V_2)-\dim(V_1+V_2) \\ &> n-\dim(V_1+V_2)\geqslant 0,\end{aligned}$$

最后一个不等式成立是由于 V_1+V_2 是 V_n 的子空间，因此有

$$\dim(V_1+V_2)\leqslant \dim(V_n) = n.$$

所以，$V_1\cap V_2$ 为非零子空间，必含有非零向量.

下面介绍子空间的特殊情况——子空间的直和.

定义 2 设 V_1,V_2 是线性空间 V 的两个子空间，若其和 V_1+V_2 中每个向量 α 的分解式

$$\alpha=\alpha_1+\alpha_2 \qquad (\alpha_1\in V_1,\alpha_2\in V_2)$$

是唯一的，则和 V_1+V_2 称为**直和**，记为 $V_1\oplus V_2$.

定理 8 V_1+V_2 是 $V_1\oplus V_2$ 的充要条件是

$$V_1\cap V_2=\{0\}.$$

证 必要性. 任取 $\alpha\in V_1\cap V_2$，则零向量 0 可表示为

$$0=\alpha+(-\alpha) \qquad (\alpha\in V_1,\alpha\in V_2).$$

因为 V_1+V_2 是直和，所以 $\alpha=-\alpha=0$. 于是有

$$V_1\cap V_2=\{0\}.$$

充分性. 如果 V_1+V_2 中某个向量 α 有两种表示方法，即

$$\alpha=\alpha_1+\alpha_2=\beta_1+\beta_2 \qquad (\alpha_1,\beta_1\in V_1;\alpha_2,\beta_2\in V_2),$$

则必有

$$(\alpha_1-\beta_1)+(\alpha_2-\beta_2)=0.$$

如果 $\alpha_2\neq\beta_2$，则

$$\alpha_1-\beta_1=\beta_2-\alpha_2\neq 0,$$

这说明 V_1 和 V_2 的公共元素 $\alpha_1-\beta_1$ 与 $\beta_2-\alpha_2$ 不为 0，这与 $V_1\cap V_2=\{0\}$矛盾. 故必有

$$\alpha_1=\beta_1, \quad \alpha_2=\beta_2,$$

这说明了 α 的分解式是唯一的，从而 V_1+V_2 是 $V_1\oplus V_2$.

定理 9 V_1+V_2 是 $V_1\oplus V_2$ 的充要条件是

$$\dim(V_1+V_2)=\dim(V_1)+\dim(V_2).$$

证 由维数公式及定理 8 即得.

因此，若 V_1+V_2 是 $V_1\oplus V_2$，在 V_1 及 V_2 中分别取一个基

$$\alpha_1, \alpha_2, \cdots, \alpha_r$$

及

$$\beta_1,\ \beta_2,\ \cdots,\ \beta_{r_2}$$

组成的向量组

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r,\ \beta_1,\ \beta_2,\ \cdots,\ \beta_{r_2}$$

就是 $V_1 \oplus V_2$ 的一个基.

定理 10 设 V_1 是 V_n 的一个子空间，则一定存在 V_n 的另一个子空间 V_2，使得 $V_n = V_1 \oplus V_2$.

证 取 V_1 的一个基 $\alpha_1, \alpha_2, \cdots, \alpha_r$，把它扩充为 V_n 的一个基

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r,\ \alpha_{r+1},\ \cdots,\ \alpha_n.$$

令 $V_2 = \mathrm{Span}\{\alpha_{r+1}, \cdots, \alpha_n\}$，则 V_2 即满足要求.

若线性空间 V_n 表示成直和

$$V_n = V_1 \oplus V_2,$$

则称 V_1, V_2 **互补**，即 V_1 为 V_2 的补空间，V_2 为 V_1 的补空间；并称 $V_n = V_1 \oplus V_2$ 为**直和分解**，即所谓**空间分解**.

例 7 在 R^3 中，设 $V_1 = \mathrm{Span}\{e_1, e_2\}$，$V_2 = \mathrm{Span}\{e_1, e_3\}$，则 $V_1 \cap V_2 = \mathrm{Span}\{e_1\} \neq \{0\}$，因而 $V_1 + V_2$ 不是 $V_1 \oplus V_2$.

例 8 在 R^4 中，设 $V_1 = \mathrm{Span}\{2e_1 - e_2, e_1\}$，$V_2 = \mathrm{Span}\{e_3 - e_4, e_4 + e_1\}$. 若设 $\alpha \in V_1 \cap V_2$，则存在 $k_1, k_2, l_1, l_2 \in R$，使得

$$k_1(2e_1 - e_2) + k_2 e_1 = l_1(e_3 - e_4) + l_2(e_4 + e_1),$$

由此可得

$$(2k_1 + k_2 - l_2)e_1 + k_1 e_2 - l_1 e_3 - (l_1 - l_2)e_4 = 0.$$

注意到 e_1, e_2, e_3, e_4 线性无关，则有

$$k_1 = k_2 = l_1 = l_2 = 0,$$

即 $\alpha = 0$，于是

$$V_1 \cap V_2 = \{0\}.$$

因此，$R^4 = V_1 \oplus V_2$.

例 9 在 R^n 中取

$$V_k = \{(x_1, \cdots, x_k, 0, \cdots, 0)\},$$

$$V_k' = \{(0, \cdots, 0, x_{k+1}, \cdots, x_n)\},$$

则

$$R^n = V_k \oplus V_k^{\perp}.$$

直和分解还可推广到多个子空间，例如

$$R^n = \mathrm{Span}\{e_1\} \oplus \mathrm{Span}\{e_2\} \oplus \cdots \oplus \mathrm{Span}\{e_n\},$$

其中 $e_i (i=1,2,\cdots,n)$ 为单位坐标向量.

综上所述，若 V_1, V_2 是 V 的子空间，则以下均是 V_1+V_2 为直和的等价条件：

(1) $V_1 \oplus V_2 = \{0\}$；

(2) V_1+V_2 中任意一个向量 α 的分解式唯一；

(3) 零向量的分解式唯一，即若

$$\alpha_1 + \alpha_2 = 0 \qquad (\alpha_1 \in V_2, \alpha_2 \in V_2),$$

则必有

$$\alpha_1 = \alpha_2 = 0;$$

(4) V_1 与 V_2 的基合在一起构成 V_n 的基；

(5) $\dim(V_1)+\dim(V_2)=\dim(V_1+V_2)$.

§1.3 内 积 空 间

在线性空间中，元素之间仅限于加法及数乘两种线性运算. 但在三维欧氏空间中，也就是在向量代数中，向量的数量积是一个重要的概念，它是引入向量正交、长度和两向量夹角等概念的基础. 为了使这些应用较广的概念能在抽象的线性空间中得到反映，我们有必要将这些概念加以拓广，建立线性空间的内积概念，由此形成内积空间.

1. 酉空间与欧氏空间

定义 1 设 V 是复数域 C 上的向量空间，$\alpha,\beta,\gamma \in V$；$a,b \in C$. 在 V 中定义了一个复值函数 (α,β)，它满足下列条件：

(1) 对称性：$(\alpha,\beta)=\overline{(\beta,\alpha)}$；

(2) 线性性：$(a\alpha+b\beta,\gamma)=a(\alpha,\gamma)+b(\beta,\gamma)$；

(3) 正定性：$(\alpha,\alpha)\geqslant 0$，当且仅当 $\alpha=0$ 时 $(\alpha,\alpha)=0$，

这时称函数 (α,β) 为向量 α 与 β 的**内积**，$\sqrt{(\alpha,\alpha)}$ 为向量 α 的**范数**或**长度**，记作 $\|\alpha\|=\sqrt{(\alpha,\alpha)}$.

我们称上述定义了向量内积及范数的线性空间为**酉空间**(Unitary Space).

若 V 为实数域 R 上的线性空间，$\alpha,\beta,\gamma\in V$；$a,b\in R$，且满足下列条件：

(1) 对称性：$(\alpha,\beta)=(\beta,\alpha)$；

(2) 线性性：$(a\alpha+b\beta,\gamma)=a(\alpha,\gamma)+b(\beta,\gamma)$；

(3) 正定性：$(\alpha,\alpha)\geqslant 0$，当且仅当 $\alpha=0$ 时 $(\alpha,\alpha)=0$，

则称 α,β 的实值函数 (α,β) 为 α 与 β 的**内积**，α 的**范数**定义为 $\|\alpha\|=\sqrt{(\alpha,\alpha)}$.

定义了上述内积及向量范数的实线性空间称为**欧几里得空间**(Euclide Space)，简称**欧氏空间**. 今后用 C^n 表示 n 维酉空间，R^n 表示 n 维欧氏空间.

几何空间中矢量的数量积显然具备内积定义中所列的性质，因此，几何空间是一个具体的欧氏空间.

例 1 在向量空间 R^n 中，对向量

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

定义内积为

$$(\alpha,\beta)=a_1b_1+a_2b_2+\cdots+a_nb_n,$$

显然它满足定义 1 中的条件，所以它是一个具体的欧氏空间，今后仍用 R^n 表示这个空间.

例 2 在闭区间 $[a,b]$ 上所有连续函数构成的线性空间 $C[a,b]$ 中，对函数 $f(x)$，$g(x)$ 定义内积为

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

则 $C[a,b]$ 为欧氏空间. 因为对任意连续函数 $f(x)$，$g(x)\in$

$C[a,b]$，有 $f(x)g(x)\in C[a,b]$，所以 (f,g) 是唯一确定的实数，同时满足

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

(2) $$\begin{aligned}(f+g,h)&=\int_a^b [f(x)+g(x)]h(x)\mathrm{d}x\\&=\int_a^b f(x)h(x)\mathrm{d}x+\int_a^b g(x)h(x)\mathrm{d}x\\&=(f,h)+(g,h);\end{aligned}$$

(3) $$\begin{aligned}(kf,g)&=\int_a^b [kf(x)]g(x)\mathrm{d}x\\&=k\int_a^b f(x)g(x)\mathrm{d}x\\&=k(f,g)\qquad(k\in R);\end{aligned}$$

(4) $(f,f)=\int_a^b f^2(x)\mathrm{d}x\geqslant 0$，当且仅当 $f=0$ 时 $(f,f)=0$.

例 1、例 2 表明，欧氏空间的概念比普遍几何空间的概念更广泛、更抽象. 至于酉空间，只不过是复数形式的欧氏空间，它是将线性空间 V 定义在复数域 C 上，除了对称性 $(\alpha,\beta)=\overline{(\beta,\alpha)}$ 外，其它性质与欧氏空间相同.

2. 正交性和施密特(Schmidt)正交化方法

有了内积，即可引进正交性概念. 设 $\alpha,\beta\in C^n$（或 R^n），有下面的正交性定义：

定义 2 若 $(\alpha,\beta)=0$，则称 α 与 β **正交**，记为 $\alpha\perp\beta$.

显然，正交性是相互的. 因为若 $(\alpha,\beta)=0$，则 $(\beta,\alpha)=\overline{(\alpha,\beta)}=\overline{0}=0$，从而由 $\alpha\perp\beta$ 必有 $\beta\perp\alpha$.

由正交的定义易知，零向量和任何向量正交.

定理 1 若 $\alpha_1,\alpha_2,\cdots,\alpha_k$ 是空间 C^n（或 R^n）中的一个非零正交系，即 $\alpha_i\neq 0$，$(\alpha_i,\alpha_j)=0\,(i\neq j;i,j=1,2,\cdots,k)$，则 $\alpha_1,\alpha_2,\cdots,\alpha_k$ 必线性无关. 简言之，非零正交系必线性无关.

证 假设存在线性关系

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_k\alpha_k=0,$$

则对任何 α_j，必有

$$\left(\sum_{i=1}^{k}k_i\alpha_i,\ \alpha_j\right)=\sum_{i=1}^{k}k_i(\alpha_i,\alpha_j)=k_j(\alpha_j,\alpha_j)=0,$$

因为 $\alpha_j\neq0$，所以 $(\alpha_j,\alpha_j)>0$，于是有

$$k_j=0\qquad(j=1,2,\cdots,k).$$

这就证明了正交系 $\alpha_1,\alpha_2,\cdots,\alpha_k$ 是线性无关的.

定义 3 设 $\alpha_1,\alpha_2,\cdots,\alpha_k$ 为 C^n 中的一个正交系，若 $\|\alpha_i\|=\sqrt{(\alpha_i,\alpha_i)}=1,i=1,2,\cdots,k$，则称 $\alpha_1,\alpha_2,\cdots,\alpha_k$ 为 C^n 中的一个**标准正交系.**

下面讨论空间 C^n 中标准正交系的存在性，并介绍 Schmidt 正交化方法，从而为在空间中建立标准正交基打下基础.

定理 2 在空间 C^n 中必存在标准正交系 $\alpha_1,\alpha_2,\cdots,\alpha_n$，即

$$(\alpha_i,\alpha_j)=\delta_{ij}=\begin{cases}1, & i=j\\0, & i\neq j\end{cases}\qquad(i,j=1,2,\cdots,n),$$

其中 δ_{ij} 称为克罗内克尔-得尔塔(Kronecker delta).

证 我们以一种构造性的方法来证明标准正交系的存在. 因为 $\dim(C^n)=n$，所以必有基

$$\beta_1,\beta_2,\cdots,\beta_n,$$

现在用这组基向量的线性组合来建立标准正交系 $\{\alpha_j\}$.

首先令 $\alpha_1=\dfrac{\beta_1}{\|\beta_1\|}$，于是

$$\|\alpha_1\|=1.$$

假定已构造好标准正交系

$$\alpha_1,\alpha_2,\cdots,\alpha_r,$$

其中，α_j 是 $\beta_1,\beta_2,\cdots,\beta_j(j=1,2,\cdots,r)$ 的线性组合. 现确定 α_{r+1} 如下：令

$$u=\beta_{r+1}-(k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r),$$

为确定系数 $k_1,k_2,\cdots,k_r$，令

$$(u,\alpha_j) = (\beta_{r+1} - \sum_{i=1}^{r} k_i\alpha_i, \alpha_j) = 0 \qquad (j = 1,2,\cdots,r).$$

由于 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 两两正交，所以这些方程可简化为

$$(\beta_{r+1},\alpha_j)-k_j=0 \qquad (j=1,2,\cdots,r),$$

即

$$k_j=(\beta_{r+1},\alpha_j) \qquad (j=1,2,\cdots,r).$$

向量 u 显然不是零向量，否则 $\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_{r+1}$ 线性相关. 从而 $\beta_1,\beta_2,\cdots,\beta_{r+1}$ 也线性相关，这与 $\beta_1,\beta_2,\cdots,\beta_n$ 为一个基相矛盾.

令 $\alpha_{r+1}=\dfrac{u}{\|u\|}$，于是求得 α_{r+1}. 继续此法，最后即得到标准正交系 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 它可以作为 C^n 的标准正交基.

定理 2 的证明中所用的方法称为 **Schmidt 正交化方法**.

为了便于记忆，下面我们介绍从 C^n 的一个基 $\beta_1,\beta_2,\cdots,\beta_n$ 出发，求出 C^n 的一个标准正交基的规范方法——Schmidt 正交化公式. 令

$$u_1=\beta_1,$$

$$u_2=\beta_2-\frac{(\beta_2,u_1)}{(u_1,u_1)}u_1,$$

$$u_3=\beta_3-\frac{(\beta_3,u_1)}{(u_1,u_1)}u_1-\frac{(\beta_3,u_2)}{(u_2,u_2)}u_2,$$

$$\cdots\cdots$$

$$u_n=\beta_n-\frac{(\beta_n,u_1)}{(u_1,u_1)}u_1-\frac{(\beta_n,u_2)}{(u_2,u_2)}u_2-\cdots-\frac{(\beta_n,u_{n-1})}{(u_{n-1},u_{n-1})}u_{n-1},$$

则 $u_1,u_2,\cdots,u_n$ 两两正交. 再将 u_i 单位化，即令

$$\alpha_i=\frac{u_i}{\|u_i\|} \qquad (i=1,2,\cdots,n),$$

从而，向量系 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 就是 C^n 的一个标准正交基.

定理 3 设 $e_1,e_2,\cdots,e_n$ 为酉空间 C^n 中的一个标准正交基，任意向量 α 在这个基下的坐标为 $(x_1,x_2,\cdots,x_n)$，则

$$x_i=(\alpha,e_i) \qquad (i=1,2,\cdots,n).$$

B

证 因为 $\alpha=\sum_{j=1}^{n}x_je_j$，所以

$$(\alpha,e_i)=(\sum_{j=1}^{n}x_je_j,\ e_i)=\sum_{j=1}^{n}x_j(e_j,e_i)$$
$$=x_i(e_i,e_i)=x_i\qquad(i=1,2,\cdots,n).$$

定理 4 设 $e_1,e_2,\cdots,e_n;e_1',e_2',\cdots,e_n'$分别为酉空间 C^n 中的两组标准正交基.则 $e_1,e_2,\cdots,e_n$ 到 $e_1',e_2',\cdots,e_n'$的过渡矩阵为酉矩阵.

证 设 $e_i'=\sum_{j=1}^{n}c_{ji}e_j$，则

$$\delta_{ij}=(e_i',e_j')=(\sum_{l=1}^{n}c_{li}e_l,\sum_{k=1}^{n}c_{kj}e_k)$$
$$=\sum_{l=1}^{n}c_{li}(e_l,\sum_{k=1}^{n}c_{kj}e_k)=\sum_{l=1}^{n}c_{li}\,\overline{(\sum_{k=1}^{n}c_{kj}e_k,e_l)}$$
$$=\sum_{l=1}^{n}c_{li}\sum_{k=1}^{n}\overline{c_{kj}}\,\overline{(e_k,e_l)}=\sum_{l=1}^{n}c_{li}\sum_{k=1}^{n}\overline{c_{kj}}(e_l,e_k)$$
$$=\sum_{l=1}^{n}\sum_{k=1}^{n}c_{li}\,\overline{c_{kj}}\delta_{lk}=\sum_{l=1}^{n}c_{li}\,\overline{c_{lj}},$$

从而可见

$$U^T\overline{U}=I,$$

或

$$U^*U=I\qquad(U^*=\overline{U^T}),$$

即 U 为酉矩阵.

推论 若 $e_1,e_2,\cdots,e_n$ 及 $e_1',e_2',\cdots,e_n'$分别为 n 维欧氏空间的两组标准正交基，则从 $e_1,e_2,\cdots,e_n$ 到 $e_1',e_2',\cdots,e_n'$的过渡矩阵 Q 为正交阵，即

$$Q^TQ=I.$$

3. 内积的表示 Cauchy-Schwarz 不等式

我们知道，引入了基的概念之后，线性空间 V 中的向量就可

以用坐标来表示. 因此,向量的内积也可以用它们的坐标来表示.

设 C^n 中的一个基为

$$\eta_1,\ \eta_2,\ \cdots,\ \eta_n,$$

它们不一定正交. 在这个基下,向量 $\alpha,\beta\in C^n$,分别用坐标表示为

$$\alpha=x_1\eta_1+x_2\eta_2+\cdots+x_n\eta_n,$$

$$\beta=y_1\eta_1+y_2\eta_2+\cdots+y_n\eta_n.$$

根据内积的定义,有关系式

$$\begin{aligned}(\alpha,\beta) &= \left(\sum_{i=1}^{n}x_i\eta_i,\ \sum_{j=1}^{n}y_j\eta_j\right)\\ &= \sum_{i,j=1}^{n}x_i\,\overline{y_j}(\eta_i,\eta_j)\\ &= \sum_{i,j=1}^{n}h_{ij}x_i\,\overline{y_j},\end{aligned}$$

其中 $h_{ij}=(\eta_i,\eta_j),i,j=1,2,\cdots,n.$

由于 $(\eta_i,\eta_j)=\overline{(\eta_j,\eta_i)}$,因此

$$h_{ij}=\overline{h_{ji}}\qquad (i,j=1,2,\cdots,n),$$

而且 h_{ij} 只与基有关,与向量 α,β 无关,即

$$(\alpha,\beta)=y^{*}Hx,$$

其中 $H=(h_{ij})_{n\times n},x=(x_1,x_2,\cdots,x_n)^T,y=(y_1,y_2,\cdots,y_n)^T$,且 $H=H^{*}$,即 H 是一个 Hermite 矩阵.

由于 $(\alpha,\alpha)\geqslant 0$,当且仅当 $\alpha=0$ 时,才有 $(\alpha,\alpha)=0$,因此,由 $(\alpha,\alpha)=x^{*}Hx$ 知,H 为正定矩阵.

特别地,若取空间的基为标准正交基

$$e_1,\ e_2,\ \cdots,\ e_n,$$

由于 $(e_i,e_j)=\delta_{ij},\ i,j=1,2,\cdots,n$,这时内积坐标表达式变成了如下的简单形式:

$$(\alpha,\beta)=\sum_{i=1}^{n}x_i\,\overline{y_i}.$$

同理可知,欧氏空间中的内积可表示为

$$(\alpha,\beta)=y^T H x$$

的形式，其中 H 为正定矩阵.

在给定的基是标准正交基时，上式可化为

$$(\alpha,\beta)=\sum_{i=1}^{n} x_i y_i.$$

当基为标准正交基时，有

$$\|\alpha\|^2=\sum_{i=1}^{n}|x_i|^2 \qquad (\alpha\in C^n),$$

或

$$\|\alpha\|^2=\sum_{i=1}^{n}x_i^2 \qquad (\alpha\in R^n).$$

定理 5（Cauchy-Schwarz 不等式） 设有酉空间 C^n，则对于任意的 $\alpha,\beta\in C^n$，恒有不等式

$$|(\alpha,\beta)|\leqslant\|\alpha\|\ \|\beta\|.$$

证 若 $\beta\neq0$，则对任意实数 λ，有

$$(\alpha-\lambda\beta,\alpha-\lambda\beta)=(\alpha,\alpha)-\bar{\lambda}(\alpha,\beta)-\lambda(\beta,\alpha)+\lambda\bar{\lambda}(\beta,\beta)\geqslant0.$$

取 $\lambda=\dfrac{(\alpha,\beta)}{(\beta,\beta)}$，以正数 (β,β) 乘上式两端，并整理得

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

两端开方取算术根，得

$$|(\alpha,\beta)|\leqslant\|\alpha\|\ \|\beta\|.$$

若 α,β 线性相关，设 $\alpha=k\beta$，则

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

若 α,β 线性无关，必有 $\alpha-\lambda\beta\neq0$，则

$$(\alpha-\lambda\beta,\alpha-\lambda\beta)>0,$$

于是有

$$|(\alpha,\beta)|<\|\alpha\|\ \|\beta\|.$$

由 Cauchy-Schwarz 不等式可得以下两个三角不等式：

(1) $\|\alpha+\beta\|\leqslant\|\alpha\|+\|\beta\|$；

(2) $\|\alpha-\beta\| \geqslant |\|\alpha\| - \|\beta\||$.

今后，用 $\|\alpha-\beta\|$ 表示向量 α,β 之间的距离，记作 $\rho(\alpha,\beta)$，即

$$\rho(\alpha,\beta)=\|\alpha-\beta\|.$$

它有如下性质：

(1) 正定性：$\rho(\alpha,\alpha)=0$，$\rho(\alpha,\beta)>0$；

(2) 对称性：$\rho(\alpha,\beta)=\rho(\beta,\alpha)$；

(3) 三角不等式：$\rho(\alpha,\beta)+\rho(\beta,\gamma)\geqslant\rho(\alpha,\gamma)$.

4. 空间的正交分解

线性空间的直和分解在前面已介绍过，现在有了内积，就可以考虑一种特殊的直和分解，即空间的正交分解.

设 W 为酉空间 C^n 的一个 k 维子空间，$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_k$ 是它的一个正交基，可以添补 $n-k$ 个正交向量 $\varepsilon_{k+1},\cdots,\varepsilon_n$，使得 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 成为 C^n 的一个正交基. 记 $\varepsilon_{k-1},\cdots,\varepsilon_n$ 所张成的子空间为

$$\mathrm{Span}\{\varepsilon_{k+1},\cdots,\varepsilon_n\},$$

其维数为 $n-k$，且其中任一向量 α 和子空间 W 中的任何向量都正交. 因此，称这个 $n-k$ 维子空间为 W 的**正交补空间**，记作 $W^\perp$，即

$$W^\perp=\mathrm{Span}\{\varepsilon_{k+1},\cdots,\varepsilon_n\}.$$

这样，便得到空间 C^n 的一种直和分解

$$C^n=W\oplus W^\perp.$$

空间 C^n 还可以进一步分解成若干个互相正交的子空间的直和. 特别地，对于给定的正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，空间 C^n 可以分解成直和

$$C^n=\mathrm{Span}\{\varepsilon_1\}\oplus\mathrm{Span}\{\varepsilon_2\}\oplus\cdots\oplus\mathrm{Span}\{\varepsilon_n\}.$$

这就是说，C^n 可以分解成 n 个一维正交子空间的直和.

欧氏空间 R^n 也可以作同样的正交分解.

例 3 设欧氏空间 R^4 中标准正交基为 e_1,e_2,e_3,e_4，它的子空间 $W=\mathrm{Span}\{\alpha_1,\alpha_2\}$，其中 $\alpha_1=e_1+e_2$，$\alpha_2=e_1+e_2-e_3$，求 $W^\perp$.

解 设 $W^\perp$ 中的向量

$$\beta=x_1e_1+x_2e_2+x_3e_3+x_4e_4,$$

由$(\alpha_1,\beta)=0,(\alpha_2,\beta)=0$ 得

$$\begin{cases}x_1+x_2=0,\\x_1+x_2-x_3=0,\end{cases}$$

其基础解系为

$$\beta_1=(1,-1,0,0)^T,$$
$$\beta_2=(0,0,0,1)^T,$$

于是有

$$W^{\perp}=\mathrm{Span}\{\beta_1,\beta_2\},$$

因此

$$R^4=W\oplus W^{\perp}.$$

§1.4 线性变换

1. 映射(变换)

定义 1 设有两个非空集 A 与 B,所谓集 A 到集 B 的一个**映射**,就是指定一个法则,它使 A 中的任一元素 α 都能与 B 中一个确定的元素 β 相对应.如果映射 T 使 $\alpha\in A$ 与 $\beta\in B$ 对应,则记为

$$\beta=T(\alpha),$$

β 称为 α 在映射 T 下的**象**,而 α 称为 β 在映射 T 下的**原象**. A 称为映射 T 的**原集**,象的全体称为**象集**,记作 $T(A)$,即

$$T(A)=\{\beta=T(\alpha)\mid\alpha\in A\}.$$

显然,$T(A)\subset B$.

例 1 设 A 是全体整数的集,B 是全体偶数的集,定义

$$T(n)=2n\qquad(n\in A),$$

则 T 是 A 到 B 的一个映射.

例 2 设 M 是数域 R 上全体 n 阶方阵的集合,定义

$$T(A)=|A|,$$

则 T 是 $R^{n\times n}$ 到 R 的一个映射.

定义 2 设 V_n, V_m 分别为数域 P 上的 n 维和 m 维线性空间，T 是一个从 V_n 到 V_m 的映射，如果映射 T 满足

(1) $T(\alpha+\beta)=T(\alpha)+T(\beta)$, $\quad \alpha,\beta\in V_n$;

(2) $T(k\alpha)=kT(\alpha)$, $\quad \alpha\in V_n, k\in R$,

则称 T 为 V_n 到 V_m 的**线性映射**或**线性算子**.

简言之，线性映射就是使线性运算保持对应的映射.

特别地，若 $V_n=V_m$，则称 T 为线性空间 V_n 中的**线性变换**.

下面讨论线性空间 V_n 中的线性变换的几个性质：

(1) $T(0)=0$, $T(-\alpha)=-T(\alpha)$, $\quad 0,-\alpha\in V_n$;

(2) 若

$$\beta=k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r,$$

则

$$T(\beta)=k_1T(\alpha_1)+k_2T(\alpha_2)+\cdots+k_rT(\alpha_r).$$

若对于任意的 $\alpha\in V_n$，恒有 $T(\alpha)=0$，则称 T 为**零变换**，记为 0，即 $0(\alpha)=0, \forall\ \alpha\in V_n$；若对于任意的 $\alpha\in V_n$，恒有 $T(\alpha)=\alpha$，则称 T 为**恒等变换**，记作 I，即 $I(\alpha)=\alpha, \forall\ \alpha\in V_n$;

(3) 若 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性相关，则 $T(\alpha_1),T(\alpha_2),\cdots,T(\alpha_n)$ 也线性相关.

但其逆不真. 例如，零变换将任一组线性无关的向量都变为零向量.

例 3 在 R^2 中定义线性变换 T 为

$$T[(x,y)]=(2x+y,x),$$

则 T 是 R^2 中的一个线性变换.

证 显然，$T(x,y)\in R^2$，又因为

$$\begin{aligned}T[(x_1,y_1)+(x_2,y_2)]&=T[(x_1+x_2,y_1+y_2)]\\&=(2(x_1+x_2)+(y_1+y_2),x_1+x_2)\\&=(2x_1+y_1,x_1)+(2x_2+y_2,x_2)\\&=T[(x_1,y_1)]+T[(x_2,y_2)],\end{aligned}$$

$$T[k(x,y)]=T[(kx,ky)]=(2kx+ky,kx)$$
$$=k(2x+y,x)=kT[(x,y)],$$
所以，T 是 R^2 中的一个线性变换.

2. 线性变换的运算

(1) 线性变换的相等

设 T_1, T_2 为两个线性变换，若 $T_1(\alpha)=T_2(\alpha)$，$\forall\ \alpha\in V_n$，则称 T_1 与 T_2 **相等**，记作 $T_1=T_2$.

设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为 V_n 的一个基，则 $T_1=T_2$ 的充要条件是
$$T_1(\alpha_i)=T_2(\alpha_i)\qquad(i=1,2,\cdots,n).$$

(2) 线性变换的和

设 T_1, T_2 为 V_n 中的两个线性变换，对于任一个元素 $\alpha\in V_n$，与 $T_1(\alpha)+T_2(\alpha)$ 相对应的变换称为 T_1 与 T_2 的**和**，记作 T_1+T_2，即
$$(T_1+T_2)\alpha=T_1(\alpha)+T_2(\alpha).$$

(3) 线性变换的数乘

设 $\alpha\in V_n$，$k\in R$，则 $kT(\alpha)$ 对应的变换称为 k 与线性变换 T 的**数乘**，记为 kT，即
$$(kT)\alpha=kT(\alpha).$$

(4) 线性变换的乘积

设 T_1, T_2 为 V_n 中的两个线性变换，$\alpha\in V_n$，与 $T_1(T_2(\alpha))$ 对应的变换称为 T_1 与 T_2 的**积**，记作 T_1T_2，即，
$$(T_1T_2)\alpha=T_1(T_2(\alpha)).$$

定理 1 设 T_1, T_2 是线性空间 V_n 中的两个线性变换，则 T_1+T_2，kT，T_1T_2 都是 V_n 的线性变换.

证明留给读者.

注意，线性变换的乘积一般不满足交换律，即一般有
$$T_1T_2\neq T_2T_1.$$

下面的例子说明了这一点.

例 4 在 R^2 中定义线性变换

$$T_1[(x,y)]=(y,x),$$
$$T_2[(x,y)]=(x,0).$$

因为

$$T_1T_2[(x,y)]=T_1[(x,0)]=(0,x),$$
$$T_1T_2[(x,y)]=T_2[(y,x)]=(y,0),$$

所以，在这里有

$$T_1T_2\neq T_2T_1.$$

例 5 设 n 阶方阵 $(a_{ij})=(\alpha_1,\alpha_2,\cdots,\alpha_n)$，其中 $\alpha_j=(a_{1j},a_{2j},\cdots,a_{nj})^T$，定义 R^n 中的变换 $Y=T(X)$ 为

$$T(X)=AX \qquad (X\in R^n),$$

则 T 为线性变换.

证 设 $\alpha,\beta\in R^n$，则

$$T(\alpha+\beta)=A(\alpha+\beta)=A\alpha+A\beta=T(\alpha)+T(\beta),$$
$$T(k\alpha)=A(k\alpha)=k(A\alpha)=kT(\alpha),$$

所以，T 为线性变换.

最后指出，如果两个线性变换满足

$$T_1T_2=T_2T_1=I,$$

这里 I 是恒等变换，则称 T_1 与 T_2 **互为逆变换**，记作

$$T_1^{-1}=T_2,\ T_2^{-1}=T_1.$$

3. 线性算子的矩阵表示

设 T 是线性空间 V_n 到 V_m 的一个线性算子，$\alpha_1,\alpha_2,\cdots,\alpha_n$ 为 V_n 中的一个基，$\beta_1,\beta_2,\cdots,\beta_m$ 为 V_m 中的一个基. 而 V_n 中各个基向量的象 $T(\alpha_j)$ 可由 V_m 中给定的基线性表示：

$$T(\alpha_j)=\sum_{i=1}^{m}a_{ij}\beta_i \qquad (j=1,2,\cdots,n), \tag{1.4}$$

令

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}. \tag{1.5}$$

若记

$$(T(\alpha_1),T(\alpha_2),\cdots,T(\alpha_n))=T(\alpha_1,\alpha_2,\cdots,\alpha_n),$$

则有

$$T(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\beta_1,\beta_2,\cdots,\beta_m)A. \tag{1.6}$$

这时，称 A 为线性算子 T 在 V_n 和 V_m **给定基下的矩阵**. 也就是说，在给定空间的基下，线性算子 T 可以唯一地由 $m\times n$ 阶矩阵 A 确定.

反之，当线性空间 V_n 和 V_m 分别给定基$\{\alpha_i\}$和$\{\beta_i\}$后，任意的 $m\times n$ 阶矩阵 A 通过对应关系(1.4)式确定了一个唯一的从 V_n 到 V_m 的线性算子 T. 这样，在空间的基给定后，从 V_n 到 V_m 的线性算子 T 和 $m\times n$ 阶矩阵 A 构成一一对应.

定理 2 设 T 为 V_n 中的线性变换，它在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵为 A，如果 α 与 $T(\alpha)$在此基下的坐标分别为$(x_1,x_2,\cdots,x_n)$与$(y_1,y_2,\cdots,y_n)$，则

$$(y_1,y_2,\cdots,y_n)^T=A(x_1,x_2,\cdots,x_n)^T. \tag{1.7}$$

证 已知

$$\alpha=\sum_{i=1}^{n}x_i\alpha_i,\quad T(\alpha)=\sum_{i=1}^{n}y_i\alpha_i,$$

由于

$$\begin{aligned} T(\alpha)&=T(\sum_{i=1}^{n}x_i\alpha_i)\\ &=\sum_{i=1}^{n}x_iT(\alpha_i)\\ &=\sum_{i=1}^{n}x_i\sum_{j=1}^{n}a_{ji}\alpha_j \end{aligned}$$

$$= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ji} x_i \right) \alpha_j ,$$

所以

$$y_j = \sum_{i=1}^{n} a_{ji} x_i \qquad (j = 1,2,\cdots,n),$$

因此

$$(y_1, y_2, \cdots, y_n)^T = A(x_1, x_2, \cdots, x_n)^T.$$

公式(1.7)称为线性变换 T 的**解析表示**.

下面的定理给出了同一线性变换在不同的基下,所对应的矩阵之间的关系.

定理 3 设在线性空间 V_n 中,从基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 到 $\beta_1, \beta_2, \cdots, \beta_n$ 的过渡矩阵为 P,V_n 中线性变换 T 在这两组基下的矩阵依次为 A 及 B,则 $B = P^{-1}AP$.

证 由假设知

$$(\beta_1, \beta_2, \cdots, \beta_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)P,$$
$$T(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)A,$$
$$T(\beta_1, \beta_2, \cdots, \beta_n) = (\beta_1, \beta_2, \cdots, \beta_n)B,$$

并由 $\beta_1, \beta_2, \cdots, \beta_n$ 的线性无关性知,P 为可逆矩阵,故

$$\begin{aligned}
(\beta_1, \beta_2, \cdots, \beta_n)B &= T(\beta_1, \beta_2, \cdots, \beta_n) \\
&= T[(\alpha_1, \alpha_2, \cdots, \alpha_n)P] \\
&= [T(\alpha_1, \alpha_2, \cdots, \alpha_n)]P \\
&= (\alpha_1, \alpha_2, \cdots, \alpha_n)AP \\
&= (\beta_1, \beta_2, \cdots, \beta_n)P^{-1}AP,
\end{aligned}$$

再由 $\beta_1, \beta_2, \cdots, \beta_n$ 的线性无关性知

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

即 A 与 B 为相似矩阵,记作 $A \sim B$.

相似矩阵有如下运算性质:

若 $B_1 = P^{-1}A_1P$,$B_2 = P^{-1}A_2P$,则

$$B_1+B_2=P^{-1}(A_1+A_2)P,$$
$$B_1B_2=P^{-1}(A_1A_2)P.$$

由此可知，若 $B=P^{-1}AP$，$f(x)$为某数域上多项式，则

$$f(B)=P^{-1}f(A)P.$$

定理 4 设 A 与 B 分别为线性变换 T_1 与 T_2 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵，则在这个基下有：

(1) T_1+T_2 的矩阵为 $A+B$；

(2) kT_1 的矩阵为 $kA,k\in R$；

(3) T_1T_2 的矩阵为 AB；

(4) 若 T_1 可逆，则 T_1^{-1} 的矩阵为 A^{-1}.

例 6 在由次数不超过 n 的一元多项式集合构成的线性空间 $P_n[x]$中，取基为

$$\alpha_1=1,\ \alpha_2=x,\ \alpha_3=x^2,\ \cdots,\ \alpha_{n+1}=x^n.$$

设 D 为微分算子，则

$$D\alpha_1=0\alpha_1+0\alpha_2+\cdots+0\alpha_{n+1},$$
$$D\alpha_2=1\alpha_1+0\alpha_2+\cdots+0\alpha_{n+1},$$
$$D\alpha_3=0\alpha_1+2\alpha_2+\cdots+0\alpha_{n+1},$$
$$\cdots\cdots$$
$$D\alpha_{n+1}=0\alpha_1+0\alpha_2+\cdots+n\alpha_n+0\alpha_{n+1},$$

故 D 在基 $\alpha_1,\alpha_2,\cdots,\alpha_{n+1}$下的矩阵为

$$A=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

若取基为

$$\beta_1=1,\ \beta_2=x,\ \beta_3=\frac{1}{2!}x^2,\ \cdots,\ \beta_{n+1}=\frac{1}{n!}x^n,$$

则由基 $\alpha_1,\alpha_2,\cdots,\alpha_{n+1}$到基 $\beta_1,\beta_2,\cdots,\beta_{n-1}$的过渡矩阵为

$$P=\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \frac{1}{2!} & \cdots & 0 & 0 \\ \cdots \\ 0 & 0 & 0 & \cdots & \frac{1}{(n-1)!} & 0 \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{n!} \end{bmatrix},$$

且有

$$P^{-1}=\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2! & \cdots & 0 & 0 \\ \cdots \\ 0 & 0 & 0 & \cdots & (n-1)! & 0 \\ 0 & 0 & 0 & \cdots & 0 & n! \end{bmatrix},$$

故 D 在基 $\beta_1,\beta_2,\cdots,\beta_{n+1}$ 下的矩阵为

$$B=P^{-1}AP=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

4. 线性变换的值域与核

定义 3 设 T 是线性空间 V_n 到 V_m 的一个线性算子，T 的全体象组成的集合称为 T 的**值域**，用 $R(T)$表示，也称为 T 的**象空间**，记为 TV_n. 于是

$$R(T)=TV_n=\{T\alpha \mid \alpha\in V_n\}\subset V_m. \tag{1.8}$$

所有被 T 变成零向量的向量构成的集合称为 T 的**核**，记为 $\mathrm{Ker}(T)$或 $T^{-1}(0)$，有时也称 $\mathrm{Ker}(T)$为 T 的**零空间**，记为 $N(T)$，

即

$$N(T)=\mathrm{Ker}(T)=\{\alpha \mid T\alpha=0,\ \alpha\in V_n\}\subset V_n. \quad (1.9)$$

不难证明，线性算子 T 的象空间 $R(T)$ 是 V_m 的子空间，零空间 $N(T)$ 是 V_n 的子空间.

事实上，由

$$T(\alpha)+T(\beta)=T(\alpha+\beta),\quad kT(\alpha)=T(k\alpha)$$

可知，$R(T)$ 是非空集且对线性运算封闭，因此，$R(T)$ 为 V_m 的子空间.

另一方面，由

$$T(\alpha)=0,\quad T(\beta)=0,$$

有

$$T(\alpha+\beta)=0,\quad T(k\alpha)=0.$$

这就是说，$N(T)$ 对于线性运算也是封闭的. 又因为 $T(0)=0$，即 $N(T)$ 非空，所以 $N(T)$ 是 V_n 的子空间.

称 $R(T)$ 的维数 $\dim R(T)$ 为 **T 的秩**，记为 $r(T)$；称 $N(T)$ 的维数 $\dim N(T)$ 为 **T 的零度**，记为 $\mathrm{null}(T)$.

例如，在线性空间 $P_n[x]$ 中，令线性算子 T 为微分算子 D，则 $DP=P'$. 显然 D 的值域 $R(D)$ 就是 $P_{n-1}[x]$，D 的核 $N(D)$ 就是数域 R(子空间).

定理 5 设 T 为线性空间 V_n 上的线性变换，$\alpha_1,\alpha_2,\cdots,\alpha_n$ 为 V_n 的一个基，在该基下的矩阵为 A，则

(1) $R(T)=\mathrm{Span}\{T(\alpha_1),T(\alpha_2),\cdots,T(\alpha_n)\}$, (1.10)

即 T 的值域 $R(T)$ 是由基的象生成的子空间；

(2) $r(T)=r(A)$.

证 (1) 设 $\alpha\in V_n$，α 可由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表示为

$$\alpha=a_1\alpha_1+a_2\alpha_2+\cdots+a_n\alpha_n,$$

于是

$$T(\alpha)=a_1T(\alpha_1)+a_2T(\alpha_2)+\cdots+a_nT(\alpha_n), \quad (1.11)$$

所以

$$T(\alpha)\in \mathrm{Span}\{T(\alpha_1),T(\alpha_2),\cdots,T(\alpha_n)\},$$

即

$$R(T)\subset \mathrm{Span}\{T(\alpha_1),T(\alpha_2),\cdots,T(\alpha_n)\}.$$

而(1.11)式又表明，基象的线性组合仍为一个象，故

$$\mathrm{Span}\{T(\alpha_1),T(\alpha_2),\cdots,T(\alpha_n)\}\subset R(T).$$

于是

$$R(T)=\mathrm{Span}\{T(\alpha_1),T(\alpha_2),\cdots,T(\alpha_n)\}.$$

(2) 由定理 5 的(1)知，T 的秩等于基象组 $T(\alpha_1),T(\alpha_2),\cdots,T(\alpha_n)$的秩. 另一方面，$A$ 是由基象组坐标按列排列而成的. 在 V_n 中，当基取定后，向量与它的坐标相对应，此为 V_n 与 P^n 的同构对应. 同构对应保持向量组的一切线性关系，因此，基象组的秩与它们的坐标组(即 A 的列向量组)的秩相等.

定理 6 设 T 是线性空间 V_n 上的线性变换，则

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

即

$$\dim R(T)+\dim N(T)=n. \tag{1.12}$$

证 设 $\dim N(T)=r$，在 $N(T)$中取定一个基 $\alpha_1,\alpha_2,\cdots,\alpha_r$，再把它扩充成 V_n 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n$. 根据定理 5 可知

$$R(T)=\mathrm{Span}\{T(\alpha_1),T(\alpha_2),\cdots,T(\alpha_n)\},$$

但 $T(\alpha_i)=0$, $i=1,2,\cdots,r$，所以

$$R(T)=\mathrm{Span}\{T(\alpha_{r+1}),\cdots,T(\alpha_n)\}.$$

下面证明 $T(\alpha_{r+1}),\cdots,T(\alpha_n)$线性无关，因而它为 $R(T)$的一个基. 事实上，设

$$k_{r+1}T(\alpha_{r-1})+\cdots+k_nT(\alpha_n)=0,$$

则

$$T(k_{r+1}\alpha_{r-1}+\cdots+k_n\alpha_n)=0,$$

这说明

$$k_{r-1}\alpha_{r-1}+\cdots+k_n\alpha_n\in N(T),$$

因此，可被 $N(T)$ 的基 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表示：

$$k_{r+1}\alpha_{r+1}+\cdots+k_n\alpha_n=k_1\alpha_1+\cdots+k_r\alpha_r.$$

由于 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关，所以

$$k_1=k_2=\cdots=k_n=0,$$

从而 $T(\alpha_{r+1}),\cdots,T(\alpha_n)$ 线性无关，故 T 的秩为 $n-r$. 这样即证明了(1.12)式.

必须指出，虽然子空间 $R(T)$ 与 $N(T)$ 的维数之和为 V_n 的维数 n，但 $R(T)+N(T)$ 并不一定就是 V_n. 例如在线性空间 $P_n[x]$ 中，令

$$Dp=p',$$

则 $R(D)$ 就是 $P_{n-1}[x]$，而 $N(D)$ 就是子空间 R，显然

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

定义 4 设 T 是 V_n 上线性变换，W 是 V_n 的子空间，如果对于任意的 $\alpha\in W$，有 $T(\alpha)\in W$，则称 W 为 T 的**不变子空间**，简称 **T-子空间**.

例 7 $R(T)$、$N(T)$ 都是 T-子空间.

按定义，$R(T)$ 显然包含了 $R(T)$ 中向量的象，所以，$R(T)$ 是 T 的不变子空间；而 $N(T)$ 是被 T 变为零的向量的集合，$N(T)$ 中向量的象都是零，自然在 $N(T)$ 中. 因此，$N(T)$ 也是不变子空间.

例 8 若线性变换 T_1 与 T_2 可交换，即 $T_1T_2=T_2T_1$，则 T_2 的核与值域都是 T_1-子空间.

任取 $\alpha\in N(T_2)$，则

$$T_2(T_1(\alpha))=(T_1T_2)(\alpha)=T_1(T_2(\alpha))=T_1(0)=0,$$

所以，$T_1(\alpha)$ 在 T_2 下的象是零，即 $T_1(\alpha)\in N(T_2)$，这就证明了 $N(T_2)$ 是 T_1-子空间. 在 $R(T_2)$ 中任取向量 $T_2(\beta)$，则

$$T_1[T_2(\beta)]=T_2[T_1(\beta)]\in R(T_2).$$

因此，$R(T_2)$ 也是 T_1-子空间.

例 9 如果线性空间 V_n 的子空间 W 由 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 所张成，即

$$W=\mathrm{Span}\{\alpha_1,\alpha_2,\cdots,\alpha_r\},$$

则 W 是 T-子空间的充分必要条件是 $T(\alpha_1),T(\alpha_2),\cdots,T(\alpha_r)$ 全都属于 W.

证 必要性是显然的，下面证明充分性.

如果 $T(\alpha_i)\in W$，$i=1,2,\cdots,r$，由于

$$\alpha=\sum_{i=1}^{r}k_i\alpha_i \qquad (\forall\ \alpha\in W),$$

所以

$$T(\alpha)=\sum_{i=1}^{r}k_iT(\alpha_i)\in W.$$

§1.5 特征值与特征向量

众所周知，在有限维线性空间 V 中，基被取定以后，线性变换 T 就可以用矩阵 A 来表示. 为了利用矩阵研究线性变换，对于每个给定的线性变换，总希望求得一个基，使得它的矩阵是最简形式，为此先介绍特征值与特征向量.

1. 线性变换的特征值与特征向量

定义 1 设 T 是数域 P 上线性空间 V 中的线性变换. 如果对于数域 P 中某一数 λ，存在非零向量 α，使得

$$T(\alpha)=\lambda\alpha, \tag{1.13}$$

则称 λ 为 T 的一个**特征值**，而 α 称为 T 的对应于特征值 λ 的一个**特征向量**.

从几何上看，特征向量的方向经线性变换后，仍保持在原方向的同一直线上. 当 $\lambda>0$ 时，其方向不变；当 $\lambda<0$ 时，其方向变成反向. 因此，T 的特征向量经变换后，要么“伸长或缩短”，要么“反向伸长或反向缩短”. 特别地，当 $\lambda=0$ 时，特征向量经线性变换后变成零向量.

若 α 是对应于特征值 λ 的特征向量，则 $k\alpha$ 也是对应特征值 λ 的特征向量，其中 $k\in P$，即

$$T(k\alpha)=\lambda(k\alpha).$$

这说明特征向量不是被特征值所唯一决定的. 相反，特征值却被特征向量所唯一决定. 因此，一个特征向量只能属于一个特征值.

下面研究确定特征值与特征向量的方法.

设数域 P 上的线性空间 V_n 的一组基为 $\alpha_1,\alpha_2,\cdots,\alpha_n$，线性变换 T 在这个基下的矩阵为 A，又设与特征值 λ 对应的特征向量 α 在这个基下的坐标为 $(x_1,x_2,\cdots,x_n)$，则与 $T(\alpha)=\lambda\alpha$ 对应的矩阵等式为

$$A\alpha=\lambda\alpha, \tag{1.14}$$

式中 α 为其坐标表达式，即 $\alpha=(x_1,x_2,\cdots,x_n)^T$. 于是得

$$(\lambda I-A)\alpha=0,$$

即

$$\begin{cases}(\lambda-a_{11})x_1-a_{12}x_2-\cdots-a_{1n}x_n=0,\\ -a_{21}x_1+(\lambda-a_{22})x_2-\cdots-a_{2n}x_n=0,\\ \cdots\cdots\\ -a_{n1}x_1-a_{n2}x_2-\cdots+(\lambda-a_{nn})x_n=0.\end{cases} \tag{1.15}$$

由于 $\alpha\neq0$，故(1.15)式有非零解. 我们知道，(1.15)式有非零解的充要条件是它的系数行列式为零，即

$$\det(\lambda I-A)=\det\begin{bmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \cdots & & & \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{bmatrix}=0. \tag{1.16}$$

定义 2 设 A 为 n 阶方阵，矩阵 $\lambda I-A$ 称为 A 的**特征矩阵**，其行列式 $\det(\lambda I-A)$ 称为 A 的**特征多项式**，$\det(\lambda I-A)=0$ 为 A 的**特征方程**，其根称为 A 的**特征值**(或**特征根**).

$\det(\lambda I - A)$的展开式是 λ 的 n 次多项式，记为

$$f(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n, \qquad (1.17)$$

这里的系数 $a_1, a_2, \cdots, a_n$ 与 A 中的元素 $a_{ij}(i, j=1, 2\cdots, n)$存在一定的关系.

为了揭示 $a_1, a_2, \cdots, a_n$ 与 $a_{ij}(i, j=1, 2, \cdots, n)$之间的关系，我们拓广矩阵迹的概念，引入所谓方阵 A 的阶迹的概念.

定义 3 在 n 阶方阵 $A=(a_{ij})$中任取行次和列次相同的 k 行 k 列$(1\leqslant k\leqslant n)$，位于这 k 行 k 列交叉处的 k^2 个元素按原位置组成的 k 阶行列式之和，称为方阵 A 的 **k 阶迹**，记为 $tr^{[k]}(A)$，即

$$tr^{[k]}(A) = \sum_{1\leqslant i_1 < i_2 < \cdots < i_k \leqslant n} \begin{vmatrix} a_{i_1 i_1} & \cdots & a_{i_1 i_k} \\ \cdots & & \\ a_{i_k i_1} & \cdots & a_{i_k i_k} \end{vmatrix} \quad (k = 1, 2, \cdots, n), \qquad (1.18)$$

其中 $a_{i_t i_s}(1\leqslant i_t, i_s\leqslant n; t, s=1, 2, \cdots, k; i_1 < i_2 < \cdots < i_k)$是位于 A 中第 i_t 行第 i_s 列的交叉处的元素.

n 阶方阵 A 的 k 阶迹 $tr^{[k]}(A)$为 C_n^k 个 k 阶行列式之和$(k=1, 2, \cdots, n)$.

例如，A 的 1 阶迹

$$tr^{[1]}(A) = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}, \qquad (1.19)$$

就是 A 的迹；

A 的 2 阶迹为

$$tr^{[2]}(A) = \sum_{1\leqslant i<j\leqslant n} \begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \cdots + \begin{vmatrix} a_{11} & a_{1n} \\ a_{n1} & a_{nn} \end{vmatrix}$$

$$+\cdots+\begin{vmatrix} a_{n-1,n-1} & a_{n-1,n} \\ a_{n,n-1} & a_{n,n} \end{vmatrix};$$

A 的 n 阶迹为

$$tr^{[n]}(A)=\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & & \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}=|A|.$$

定理 1 n 阶方阵 A 的特征多项式 $f(\lambda)=|\lambda I-A|$ 可写成如下形式

$$f(\lambda)=\sum_{k=1}^{n}(-1)^k tr^{[k]}(A)\lambda^{n-k}, \tag{1.20}$$

其中

$$tr^{[0]}(A)=1,$$

$$tr^{[k]}(A)=\sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n}\begin{vmatrix} a_{i_1i_1} & \cdots & a_{i_1i_k} \\ \cdots & & \\ a_{i_ki_1} & \cdots & a_{i_ki_k} \end{vmatrix}$$

$$(k=1,2,\cdots,n).$$

公式(1.20)称为 A 的特征多项式的**阶迹表达式**(证略).

用(1.20)式计算方阵特征多项式的系数

$$a_k=(-1)^k tr^{[k]}(A) \qquad (k=1,2,\cdots,n),$$

只用到方阵 $A=(a_{ij})$ 的元素 a_{ij} 组成的顺序主子式的和,计算起来比较规范,这反映了 $f(\lambda)$ 的系数与方阵 A 的元素 a_{ij} 之间的独特关系.

例 1 设线性变换 T 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵为

$$A=\begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix},$$

求线性变换 T 的特征值与特征向量.

解 按特征多项式的阶迹表达式,有

$$
\begin{aligned}
f(\lambda) &= \det(\lambda I - A) \\
&= \lambda^3 - tr^{[1]}(A)\lambda^2 - tr^{[2]}(A)\lambda - tr^{[3]}(A) \\
&= \lambda^3 - (-1+3+2)\lambda^2 \\
&\quad + \left(\begin{vmatrix} -1 & 1 \\ -4 & 3 \end{vmatrix} + \begin{vmatrix} -1 & 0 \\ 1 & 2 \end{vmatrix} + \begin{vmatrix} 3 & 0 \\ 0 & 2 \end{vmatrix} \right)\lambda - \begin{vmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{vmatrix} \\
&= \lambda^3 - 4\lambda^2 + 5\lambda - 2 \\
&= (\lambda-2)(\lambda-1)^2,
\end{aligned}
$$

令 $f(\lambda)=0$，得 T（或 A）的特征值为

$$\lambda_1=2, \ \lambda_2=\lambda_3=1.$$

对于 $\lambda_1=2$，齐次线性方程组

$$\begin{cases} 3x_1 - x_2 = 0, \\ 4x_1 - x_2 = 0, \\ -x_1 = 0 \end{cases}$$

的基础解系为 $(0,0,1)^T$，与 $\lambda_1=2$ 相应的全部特征向量为 $k(0,0,1)^T$，其中 $k\neq 0$.

对于 $\lambda_2=\lambda_3=1$，齐次线性方程组

$$\begin{cases} 2x_1 - x_2 = 0, \\ 4x_1 - 2x_2 = 0, \\ -x_1 - x_3 = 0 \end{cases}$$

的基础解系为 $(1,2,-1)^T$，与二重根 $\lambda_2=\lambda_3=1$ 相应的全部特征向量为 $k(1,2,-1)^T$，其中 $k\neq 0$.

例 2 在线性空间 $P_n[x]$ 中，设线性变换 D 为

$$Df(x)=f'(x),$$

则线性变换 D 在基

$$\beta_1=1, \ \beta_2=x, \ \beta_3=\frac{1}{2!}x^2, \ \cdots, \ \beta_{n+1}=\frac{1}{n!}x^n$$

下的矩阵为

$$B=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

其特征多项式为

$$\det(\lambda I-B)=\begin{vmatrix} \lambda & -1 & 0 & \cdots & 0 \\ 0 & \lambda & -1 & \cdots & 0 \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & -1 \\ 0 & 0 & 0 & \cdots & \lambda \end{vmatrix}.$$

因此，B 的特征值只有 0. 通过解相应的齐次线性方程组可知，属于特征值 0 的线性无关的特征向量组只是任一非零常数(即零次多项式). 这表明导数为零的多项式只能是零或非零常数.

例 3 在控制论中常见的所谓**相伴矩阵**(或称**友矩阵**)为

$$A=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{pmatrix},$$

其中 $a_1,a_2,\cdots,a_n$ 为常数，求 A 的特征值与特征向量.

解 A 的特征多项式为

$$\det(\lambda I-A)=\det\begin{pmatrix} \lambda & -1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & -1 & \cdots & 0 & 0 \\ \cdots & & & & & \\ 0 & 0 & 0 & \cdots & \lambda & -1 \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_2 & \lambda+a_1 \end{pmatrix},$$

将行列式的第 $2,3,\cdots,n$ 列分别乘以 $\lambda,\lambda^2,\cdots,\lambda^{n-1}$后都加到第一列上去，得

$$\det(\lambda I - A) = \det\begin{bmatrix} 0 & -1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & -1 & \cdots & 0 & 0 \\ \cdots \\ 0 & 0 & 0 & \cdots & \lambda & -1 \\ f(\lambda) & a_{n-1} & a_{n-2} & \cdots & a_2 & \lambda + a_1 \end{bmatrix}, \tag{1.21}$$

其中，$f(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$.

设 $\lambda_i(i=1,2,\cdots,n)$ 为 $f(\lambda)$ 的根，即 A 的特征值. 欲求属于特征值 λ_i 的特征向量 $\alpha_i=(x_1,x_2,\cdots,x_n)^T$，必须先解齐次线性方程组 $(\lambda_i I-A)\alpha_i=0$，即

$$\begin{cases} \lambda_i x_1 - x_2 = 0, \\ \lambda_i x_2 - x_3 = 0, \\ \cdots\cdots \\ \lambda_i x_{n-1} - x_n = 0, \\ a_n x_1 + a_{n-1}x_2 + \cdots + a_2 x_{n-1} + (\lambda_i + a_1)x_n = 0, \end{cases} \tag{1.22}$$

亦即

$$\begin{cases} x_2=\lambda_i x_1, \\ x_3=\lambda_i x_2, \\ \cdots\cdots \\ x_n=\lambda_i x_{n-1}, \\ \lambda_i x_n+a_1x_n+a_2x_{n-1}+\cdots+a_{n-1}x_2+a_nx_1=0. \end{cases}$$

从第一个等式开始，依次将上一个等式代入下一个等式，得

$$\begin{cases} x_2=\lambda_i x_1, \\ x_3=\lambda_i^2 x_1, \\ \cdots\cdots \\ x_n=\lambda_i^{n-1}x_1, \\ (\lambda_i^n+a_1\lambda_i^{n-1}+\cdots+a_{n-1}\lambda_i+a_n)x_1=0, \end{cases}$$

注意到

$$f(\lambda_i)=\lambda_i^n+a_1\lambda_i^{n-1}+\cdots+a_{n-1}\lambda_i+a_n=0,$$

从而求得解向量为

$$\alpha_i=(x_1,x_2,\cdots,x_n)^T=(1,\lambda_i,\lambda_i^2,\cdots,\lambda_i^{n-1})^T x_1,$$

其中 x_1 为任意实数.

令 $x_1=1$，则(1.22)式的基础解系为$(1,\lambda_i,\lambda_i^2,\cdots,\lambda_i^{n-1})^T,i=1,2,\cdots,n$，即属于特征值 $\lambda=\lambda_i$ 的全体特征向量为

$$k(1,\lambda_i,\lambda_i^2,\cdots,\lambda_i^{n-1})^T \qquad (k\neq 0,i=1,2,\cdots,n).$$

假定 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为相异的特征值，则以与之相应的线性无关的特征向量为列向量所构成的可逆矩阵

$$P=\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \cdots & & & \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix}$$

满足

$$P^{-1}AP=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n).$$

在线性变换中，方阵的特征多项式是很重要的. 假定 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为方阵 A 的 n 个相异的特征值，则它的特征多项式 $f(\lambda)$可以分解成一次因式的乘积：

$$f(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n). \qquad (1.23)$$

将(1.23)式的右端展开，再根据根与系数的关系，有

$$f(\lambda)=\lambda^n+(-1)(\sum_{i=1}^{n}\lambda_i)\lambda^{n-1}+(-1)^2(\sum_{i<j}\lambda_i\lambda_j)\lambda^{n-2}+\cdots$$
$$+(-1)^k(\sum_{i<j<\cdots<h}\lambda_i\lambda_j\cdots\lambda_h)\lambda^{n-k}+\cdots+(-1)^n\prod_{i=1}^{n}\lambda_i. \qquad (1.24)$$

比较(1.20)式与(1.24)式，有下面的结果：

$$tr^{[1]}(A)=\sum_{i=1}^{n}a_{ii}=\sum_{i=1}^{n}\lambda_i, \qquad (1.25)$$

$$tr^{[n]}(A) = |A| = \prod_{i=1}^{n} \lambda_i. \qquad (1.26)$$

(1.25)式和(1.26)式表明，矩阵 A 的全体特征值的和为 A 的迹 $a_{11}+a_{22}+\cdots+a_{nn}$，而矩阵 A 的全体特征值的积为 $|A|$. 这反映了 $f(\lambda)$ 的系数 $a_i(i=1,2,\cdots,n)$ 与它的根 $\lambda_i(i=1,2,\cdots,n)$ 及 A 的元素 $a_{ij}(i,j=1,2\cdots,n)$ 之间存在着独特、微妙的关系，这个关系可从下式得到全面反映：

$$f(\lambda) = |\lambda I - A| = \sum_{k=0}^{n}(-1)^k tr^{[k]}(A)\lambda^{n-k} = \prod_{i=1}^{n}(\lambda - \lambda_i).$$

2. 矩阵的特征值与特征向量之间的关系

定义 4 设 T 是线性空间 V 的一个线性变换，对于 T 的任一特征值 λ_0，适合条件

$$T(\alpha)=\lambda_0\alpha$$

的向量组成的集，即属于 λ_0 的全体特征向量再加上零向量组成的集，称为 T 的一个**特征子空间**，记为 V_{λ_0}，用集合的记法可写成

$$V_{\lambda_0}=\{\alpha \mid T(\alpha)=\lambda_0\alpha, \alpha\in V\}.$$

显然，V_{λ_0} 的维数就是属于特征值 λ_0 的线性无关的特征向量的个数，称 $\dim(V_{\lambda_0})$ 为特征值 λ_0 的**几何重数**；而称 λ_0 在特征多项式 $f(\lambda)$ 中根的重数为特征值 λ_0 的**代数重数**.

定理 2 线性变换 T 的任一特征值的几何重数不大于它的代数重数.

证 首先取定 V_{λ_0} 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_k(k=\dim(V_{\lambda_0})>0)$，然后把它扩充为 V 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_k,\alpha_{k+1},\cdots,\alpha_n$. 于是有

$$T(\alpha_1)=\lambda_0\alpha_1,$$

$$T(\alpha_2)=\lambda_0\alpha_2,$$

$$\cdots\cdots$$

$$T(\alpha_k)=\lambda_0\alpha_k,$$

$$T(\alpha_{k+1})=a_{1,k+1}\alpha_1+\cdots+a_{k,k+1}\alpha_k+\cdots+a_{n,k+1}\alpha_n,$$

……

$$T(\alpha_n)=a_{1n}\alpha_1+\cdots+a_{kn}\alpha_k+\cdots+a_{nn}\alpha_n,$$

即在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的线性变换 T 所对应的矩阵为

$$A=\begin{pmatrix} \lambda_0 & & & a_{1,k-1} & \cdots & a_{1n} \\ & \ddots & & \cdots & & \\ & & \lambda_0 & a_{k,k-1} & \cdots & a_{kn} \\ & & & a_{k-1,k+1} & \cdots & a_{k+1,n} \\ & 0 & & \cdots & & \\ & & & a_{n,k+1} & \cdots & a_{nn} \end{pmatrix},$$

从而，线性变换 T 的特征多项式为

$$f(\lambda)=|\lambda I-A|=(\lambda-\lambda_0)^k|\lambda I-A_0|,$$

其中

$$A_0=\begin{pmatrix} a_{k+1,k+1} & \cdots & a_{k+1,n} \\ \cdots & & \\ a_{n,k+1} & \cdots & a_{nn} \end{pmatrix}.$$

由于 $|\lambda I-A_0|$ 为 $n-k$ 次多项式，记为 $\varphi(\lambda)$，因而有

$$f(\lambda)=(\lambda-\lambda_0)^k\varphi(\lambda),$$

故 λ_0 在 $f(\lambda)$ 中的重数不小于 k，即线性变换 T 的特征值的几何重数不大于其代数重数。

关于特征值与特征向量，有下面几个命题：

（1）方阵 A 与 A^T 有相同的特征多项式，因而有相同的特征值.

（2）若 λ 为 A 的一个特征值，α 为对应的特征向量，则 $k\lambda$，λ^m 分别为 kA，A^m 的特征值，α 仍为对应的特征向量，其中 k 为常数，m 为正整数.

（3）设 λ 为可逆矩阵 A 的某个特征值，α 为对应的特征向量，则 $\frac{1}{\lambda}$ $(\lambda\neq0)$ 为 A^{-1} 的特征值，对应的特征向量仍为 α.

（4）相似矩阵有相同的特征多项式.

以上各命题的证明留作练习.

我们知道,在有限维线性空间中,基被选定之后,特征值就是线性变换在这组基下矩阵的特征多项式的根.随着基的不同,线性变换的矩阵一般也不同,但这些矩阵是相似的.上面的命题(4)指出了相似矩阵有相同的特征多项式,这说明线性变换的特征多项式与基的选择无关,它直接由线性变换所决定.因此,以后就可以说线性变换的特征多项式了.

下面讨论分属于不同特征值所对应的特征向量之间的关系.

定理 3 矩阵 A 的相异特征值所对应的特征向量必线性无关.

证 设 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 A 的 n 个相异的特征值,$\alpha_1,\alpha_2,\cdots,\alpha_n$ 是与它们相应的特征向量.下面用归纳法证明,当 $n=1$ 时,定理显然成立,因为任一非零向量线性无关.假设对 $n-1$ 个相异特征值的情况定理成立,下面证明对 n 个情况也成立.设

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_{n-1}\alpha_{n-1}+k_n\alpha_n=0,$$

上式左乘矩阵 A,并由 $A\alpha_i=\lambda_i\alpha_i(i=1,2\cdots,n)$,得

$$k_1\lambda_1\alpha_1+k_2\lambda_2\alpha_2+\cdots+k_{n-1}\lambda_{n-1}\alpha_{n-1}+k_n\lambda_n\alpha_n=0,$$

从上面两式中消去 α_n,得

$$k_1(\lambda_1-\lambda_n)\alpha_1+k_2(\lambda_2-\lambda_n)\alpha_2+\cdots+k_{n-1}(\lambda_{n-1}-\lambda_n)\alpha_{n-1}=0.$$

由于假设 $\alpha_1,\alpha_2,\cdots,\alpha_{n-1}$ 线性无关,又因为 $\lambda_i-\lambda_n\neq0$ $(i=1,2\cdots,n-1)$,因此,$k_1=k_2=\cdots=k_{n-1}=0$,从而 $k_n=0$.所以 n 个相异特征值对应的特征向量线性无关.

定理 3 告诉我们,若方阵 A 有 n 个相异特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$,则对应的线性变换 T 就有 n 个线性无关的特征向量 $\alpha_1,\alpha_2,\cdots,\alpha_n$,以这一组特征向量为基,线性变换在此基下的矩阵为对角形矩阵 $\mathrm{diag}\ (\lambda_1,\lambda_2,\cdots,\lambda_n)$.

但是,一般来说,方阵 A 未必都有 n 个相异的特征值,下面讨论这种情况.

定义 5 设线性变换 T 有 t 个不同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_t(t\leqslant$

$n)$，$V_{\lambda_1},V_{\lambda_2},\cdots,V_{\lambda_t}$分别为相应的特征子空间，其维数依次为$r_1,r_2,\cdots,r_t$，相应的基为$\alpha_{11},\alpha_{12},\cdots,\alpha_{1r_1};\alpha_{21},\alpha_{22},\cdots,\alpha_{2r_2};\cdots\cdots;\alpha_{t1},\alpha_{t2},\cdots,\alpha_{tr_t}$，称上面的向量组为线性变换$T$的一个**特征向量系**.

由于特征子空间$V_{\lambda_i}(i=1,2,\cdots,t)$中的基$\alpha_{i1},\cdots,\alpha_{ir_i}$是线性无关的，而相异特征值所对应的特征向量也是线性无关的，故特征向量系是线性无关的向量组.

一般来说，线性空间V_n中的线性变换T的特征向量系所含的向量个数不大于n，这是因为线性变换T的任一特征值的几何重数不大于它的代数重数. 当特征向量系所含向量个数等于n时，便说它是**完全的特征向量系**.

对于一个线性变换T，如果它具有完全特征向量系，以此完全特征向量系为基，则线性变换T在此基下的矩阵为对角形矩阵. 此时称线性变换T的矩阵A为**单纯矩阵**，矩阵A的所有特征值的几何重数与代数重数相等. 如果线性变换T的特征向量系中的向量个数小于空间维数n，则此线性变换在任一组基下的矩阵不能为对角形，此时线性变换的矩阵A称为**非单纯矩阵**(或**亏损矩阵**). 因此，T的矩阵A在某一个基下成对角形的充分必要条件是T有完全的特征向量系，即T的特征子空间$V_{\lambda_1},V_{\lambda_2},\cdots,V_{\lambda_t}$的维数之和等于空间的维数$n$.

关于非单纯矩阵(亏损矩阵)化为标准形的问题，将在下一章讨论.

将方阵作相似变换化为对角形是一个比较复杂的问题，但是将方阵作相似变换化为三角阵则较为容易. 为此，先介绍下面的引理及许尔(Schur)定理.

引理　若n元复向量$\alpha_1=(c_1,c_2,\cdots,c_n)^T$的范数

$$\|\alpha_1\|=\sqrt{(\alpha_1,\alpha_1)}=1,$$

则有酉矩阵U以α_1为它的第1列向量.

证　先考虑n个未知量$x_1,x_2,\cdots,x_n$的线性齐次方程

$$\overline{\alpha}_1^T\alpha=\overline{c}_1x_1+\overline{c}_2x_2+\cdots+\overline{c}_nx_n=0, \qquad (1.27)$$

式中 $\alpha=(x_1,x_2,\cdots,x_n)^T$.

取(1.27)式的任一非零解 β_2，令

$$\alpha_2=\frac{\beta_2}{\sqrt{(\beta_2,\beta_2)}},$$

则 α_2 为与 α_1 正交的单位向量.

设已求得 $r-1$ 个两两正交的单位向量 $\alpha_1,\alpha_2,\cdots,\alpha_{r-1}$，若 β_r 是线性齐次方程组

$$\overline{\alpha}_1^T\alpha=0,\ \overline{\alpha}_2^T\alpha=0,\ \cdots,\ \overline{\alpha}_{r-1}^T\alpha=0$$

的任一非零解，令

$$\alpha_r=\frac{\beta_r}{\sqrt{(\beta_r,\beta_r)}},$$

则 α_r 为与 $\alpha_1,\alpha_2,\cdots,\alpha_{r-1}$ 都正交的单位向量. 这样可一直做下去，直到求出 α_n. 由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为列向量的矩阵 U 即为所求的矩阵.

定理 4(**Schur 定理**) 设 A 为 n 阶方阵，$\lambda_1,\lambda_2,\cdots,\lambda_n$ 为其特征值，不论它们为实数还是复数，总存在相似酉矩阵 U，将 A 化为三角阵，即 $U^{-1}AU=B$，B 为三角阵，且 B 的对角线元素为 $\lambda_1,\lambda_2,\cdots,\lambda_n$.

证 用归纳法证明.

设 λ_1 为 A 的一个特征值，相应的特征向量为 α_1，不妨设它为单位向量. 由引理知，有一个以 α_1 为第 1 列向量的酉矩阵 $U_1=(\alpha_1,\alpha_2,\cdots,\alpha_n)$. 因为

$$(U_1^{-1}A)\alpha_1=U_1^{-1}(A\alpha_1)=\lambda_1U^{-1}\alpha_1=\lambda_1\overline{U}^T\alpha_1=\lambda_1(1,0,\cdots,0)^T,$$

故

$$U_1^{-1}AU_1=\begin{bmatrix}\lambda_1 & \cdots\\ 0 & A_1\end{bmatrix},$$

这里 A_1 是 $n-1$ 阶方阵. 显然当 $n=2$ 时，$U_1^{-1}AU_1$ 是三角阵.

现假定对一般的 $n-1$，定理 4 的结论成立，即对任一$(n-1)$

阶方阵，总有酉矩阵 V，使得

$$V^{-1}A_1V=B_1,$$

其中 B_1 为三角阵，则对于 n，取

$$U_2=\begin{pmatrix}1 & 0\\ 0 & V\end{pmatrix},$$

则 U_2 也是酉矩阵. 令 $U=U_1U_2$，则 U 也是酉矩阵. 故

$$\begin{aligned}U^{-1}AU &=U_2^{-1}(U_1^{-1}AU_1)U_2\\ &=U_2^{-1}\begin{pmatrix}\lambda_1 & \cdots\\ 0 & A_1\end{pmatrix}U_2\\ &=\begin{pmatrix}1 & 0\\ 0 & V^{-1}\end{pmatrix}\begin{pmatrix}\lambda_1 & \cdots\\ 0 & A_1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & V\end{pmatrix}\\ &=\begin{pmatrix}\lambda_1 & \cdots\\ 0 & B_1\end{pmatrix}\\ &=B.\end{aligned}$$

推论 1 如果 A 为 Hermite 矩阵，则存在酉矩阵 U，使得

$$U^*AU=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n),$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 A 的特征值.

推论 2 如果 A 为实对称矩阵，则存在正交矩阵 Q，使得

$$Q^TAQ=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n).$$

习 题 一

1. 检验以下集合对于所指的线性运算是否构成实数域上的线性空间：

(1) 设 A 是 n 阶实数矩阵. A 的实系数多项式 $f(A)$ 的全体，对于矩阵的加法和数乘；

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

(3) 全体实数的二元数列，对于如下定义的加法⊕和数乘。

运算：

$$(a_1,b_1)\oplus(a_2,b_2)=(a_1+a_2,b_1+b_2+a_1a_2),$$

$$k\circ(a_1,b_1)=(ka_1,kb_1+\frac{k(k-1)}{2}a_1^2);$$

(4) 设 R^+ 是一切正实数集合，定义如下加法和数乘运算：

$$a\oplus b=ab,$$

$$k\circ a=a^k,$$

其中 $a,b\in R^+,k\in R$；

(5) 二阶常系数非齐次线性微分方程的解的集合，对于通常函数的加法和数乘；

(6) 设 $V=\{x\,|\,x=c_1\sin t+c_2\sin 2t+\cdots+c_k\sin kt,c_i\in R,0\leqslant t\leqslant 2\pi\}$，$V$ 中元素对于通常的加法与数乘，并证明：$\{\sin t,\sin 2t,\cdots,\sin kt\}$是 V 的一个基，试指出确定 c_i 的方法.

2. 求下列线性空间的维数与一个基：

(1) $R^{n\times n}$中全体对称(反对称、上三角)矩阵构成的实数域 R 上的空间；

(2) 第 1 题(4)中的空间；

(3) 实数域 R 上由矩阵 A 的全体实系数多项式组成的空间，其中

$$A=\begin{pmatrix}1&0&0\\0&\omega&0\\0&0&\omega^2\end{pmatrix},\ \omega=\frac{-1+\sqrt{3}\mathrm{i}}{2},\ \omega^2=\bar{\omega},\ \omega^3=1.$$

3. 设 $P[x]$表示实数系数多项式全体构成的线性空间，问下列向量集合是否构成 $P[x]$的子空间：

(1) $\{p(x)\,|\,p(1)=0\}$；

(2) $\{p(x)\,|\,p(x)$的常数项为零$\}$；

(3) $\{p(x)\,|\,p(x)=p(-x)\}$；

(4) $\{p(x)\,|\,p(x)=-p(-x)\}$.

4. 证明下列向量集合组成线性子空间，并求基和维数：

(1) 第偶数个坐标为零的所有 n 维向量；

(2) 形如$(\alpha,\beta,\alpha,\beta,\alpha,\beta,\cdots)^T$ 的所有的 n 维向量，其中 α,β 为任意数.

5. 在 R^4 中，求由基 ξ_1,ξ_2,ξ_3,ξ_4 到 $\eta_1,\eta_2,\eta_3,\eta_4$ 的过渡矩阵，并求向量 ξ 在指定基下的坐标，设

$$(1)\begin{cases}\xi_1=(1,2,-1,0)^T,\\ \xi_2=(1,-1,1,1)^T,\\ \xi_3=(-1,2,1,1)^T,\\ \xi_4=(-1,-1,0,1)^T,\end{cases}\qquad \begin{cases}\eta_1=(2,1,0,1)^T,\\ \eta_2=(0,1,2,2)^T,\\ \eta_3=(-2,1,1,2)^T,\\ \eta_4=(1,3,1,2)^T,\end{cases}$$

$\xi=(1,0,0,0)$在 ξ_1,ξ_2,ξ_3,ξ_4 下的坐标；

$$(2)\begin{cases}\xi_1=(1,1,1,1)^T,\\ \xi_2=(1,1,-1,-1)^T,\\ \xi_3=(1,-1,1,-1)^T,\\ \xi_4=(1,-1,-1,1)^T,\end{cases}\qquad \begin{cases}\eta_1=(1,1,0,1)^T,\\ \eta_2=(2,1,3,1)^T,\\ \eta_3=(1,1,0,0)^T,\\ \eta_4=(0,1,-1,-1)^T,\end{cases}$$

$\xi=(1,0,0,-1)$在 $\eta_1,\eta_2,\eta_3,\eta_4$ 下的坐标.

6. 在 R^4 中给定两个基

$$\begin{cases}\xi_1=(1,0,0,0)^T,\\ \xi_2=(0,1,0,0)^T,\\ \xi_3=(0,0,1,0)^T,\\ \xi_4=(0,0,0,1)^T,\end{cases}\qquad \begin{cases}\eta_1=(2,1,-1,1)^T,\\ \eta_2=(0,3,1,0)^T,\\ \eta_3=(5,3,2,1)^T,\\ \eta_4=(6,6,1,3)^T,\end{cases}$$

求一非零向量，使它在两个基下有相同的坐标.

7. 设

$$A=\begin{pmatrix}1&0&0\\0&1&0\\3&1&2\end{pmatrix},$$

求 $R^{3\times3}$中全体与 A 可交换的矩阵所生成子空间的维数和一个基.

8. 设 K,S 是实系数多项式空间 $P[x]$中的两个子集，其定义为

$$K=\{p(x)\mid p(-x)=-p(x),\forall\ x\in R\},$$
$$S=\{p(x)\mid p(-x)=p(x),\forall\ x\in R\},$$

证明：

$$P[x]=K\oplus S.$$

9. 设 V_1 与 V_2 分别是齐次线性方程组

$$x_1+x_2+\cdots+x_n=0$$

和

$$x_1=x_2=\cdots=x_n$$

的解空间，证明：

$$R^n=V_1\oplus V_2.$$

10. 设

$$V_1=\{A=(a_{ij})\in R^{m\times n}\mid a_{ij}=0,\text{当 } i>j \text{ 时}\},$$
$$V_2=\{A=(a_{ij})\in R^{m\times n}\mid a_{ij}=0,\text{当 } i\leqslant j \text{ 时}\},$$

证明：

$$R^{m\times n}=V_1\oplus V_2.$$

11. 证明：和空间 $\sum\limits_{i=1}^{k}V_i$ 为直和的充要条件是

$$V_i\cap\sum_{j\neq i}V_j=\{0\}\qquad(i=1,2,\cdots,k).$$

12. 求下列由向量 $\{\alpha_i\}$ 生成的子空间与由向量 $\{\beta_i\}$ 生成的子空间的交与和的维数和基：

(1) $\begin{cases}\alpha_1=(1,2,1,0)^T,\\ \alpha_2=(-1,1,1,1)^T,\end{cases}$ $\quad\begin{cases}\beta_1=(2,-1,0,1)^T,\\ \beta_2=(1,-1,3,7)^T;\end{cases}$

(2) $\begin{cases}\alpha_1=(1,2,-1,-2)^T,\\ \alpha_2=(3,1,1,1)^T,\\ \alpha_3=(-1,0,1,-1)^T,\end{cases}$ $\quad\begin{cases}\beta_1=(2,5,-6,-5)^T,\\ \beta_2=(-1,2,-7,3)^T.\end{cases}$

13. 设 $A=(a_{ij})$ 是一个 n 阶正定矩阵，而 $\alpha=(x_1,x_2,\cdots,x_n)^T$，$\beta=(y_1,y_2,\cdots,y_n)^T$，在 R^n 中定义内积为 $(\alpha,\beta)=\alpha^TA\beta$，试证明：在这个定义下，$R^n$ 为欧氏空间.

14. 设

$$V=\left\{\begin{pmatrix} a & a+b \\ c & c \end{pmatrix}\middle| a,b,c\in R\right\},$$

作出 V 到 R^3 的同构对应.

15. 设 X 是微分方程 $x''+x=0$ 的解的全体

$$X=\{x\,|\,x=a\cos t+b\sin t, a,b\in R, 0\leqslant t\leqslant 2\pi\},$$

证明:X 与 R^2 同构.

16. 在 R^4 中求一单位向量与 $(1,1,-1,1)^T,(1,-1,-1,1)^T,(2,1,1,3)^T$ 正交.

17. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4,\varepsilon_5$ 是 R^5 中一组标准正交基,$V=\mathrm{Span}\{\alpha_1,\alpha_2,\alpha_3\}$,其中 $\alpha_1=\varepsilon_1+\varepsilon_5,\alpha_2=\varepsilon_1-\varepsilon_2+\varepsilon_4,\alpha_3=2\varepsilon_1+\varepsilon_2+\varepsilon_3$,求 V 的一组标准正交基.

18. 设 $V=C^3$,并有§1.3 节例 1 所定义的内积,若 $\alpha=(2,1+\mathrm{i},\mathrm{i})^T,\beta=(2-\mathrm{i},2,1+2\mathrm{i})^T$,计算 (α,β),$\|\alpha\|$,$\|\beta\|$ 及距离 $\rho(\alpha,\beta)$,并验证 Cauchy-Schwarz 不等式.

19. 用 Schmidt 正交化方法,将内积空间 V 的给定子集 S 正交化,再找出 V 的标准正交基,并求出给定向量在标准正交基下的坐标表达式:

(1) $V=R^4,S=\{(1,2,2,-1)^T,(1,1,-5,3)^T,(3,2,8,-7)^T\},\alpha=(3,1,1,-3)^T$;

(2) $V=R^4,S=\{(2,1,3,-1)^T,(1,4,3,-3)^T,(1,1,-6,0)^T,(5,7,7,8)^T\},\alpha=(2,1,3,-1)^T$;

(3) $V=P_3[x]$,定义内积为 $(f,g)=\int_0^1 f(t)g(t)\mathrm{d}t$,$S=\{1,x,x^2\}$,$f(x)=1+x$.

20. 用向量 $\alpha_1=(1,0,2,1)^T,\alpha_2=(2,1,2,3)^T,\alpha_3=(0,1,-2,1)^T$ 生成子空间 V,求 V 的正交补 $V^{\perp}$ 的基底及正交补空间 $V^{\perp}$.

21. 判别下面所定义的变换,哪些是线性变换,哪些不是?

(1) 在线性空间 V 中,$T(\xi)=\xi+\alpha$,其中 $\alpha\in V$ 是一固定

向量；

(2) 把复数域看作复数域上的线性空间，$T(\xi)=\bar{\xi}$；

(3) 在 R^3 中，$T(x_1,x_2,x_3)=(x_1^2,x_2+x_3,x_3^2)$；

(4) 在 R^3 中，$T(x_1,x_2,x_3)=(2x_1-x_2,x_2+x_3,x_1)$；

(5) 在 $R^{n\times n}$ 中，$T(Z)=BZC$，其中 $B,C\in C^{n\times n}$ 是两个固定的矩阵；

(6) 在 $P[x]$ 中，$T[p(x)]=p(x+1)$；

(7) 在 $P[x]$ 中，$T[p(x)]=p(x_0)$，其中 $x_0\in R$ 是一个固定的数.

22. 求下列线性变换在指定基下的矩阵：

(1) 第 21 题(4)中变换 T 在基 $\varepsilon_1=(1,0,0),\varepsilon_2=(0,1,0),\varepsilon_3=(0,0,1)$ 下的矩阵；

(2) 六个函数

$$\xi_1=e^{\alpha x}\cos\beta x,\qquad \xi_2=e^{\alpha x}\sin\beta x,$$
$$\xi_3=xe^{\alpha x}\cos\beta x,\qquad \xi_4=xe^{\alpha x}\sin\beta x,$$
$$\xi_5=\frac{1}{2}x^2e^{\alpha x}\cos\beta x,\quad \xi_6=\frac{1}{2}x^2e^{\alpha x}\sin\beta x$$

的所有实系数线性组合构成实数域上一个六维空间，求微分变换 D 在基 $\{\xi_i, i=1,2,\cdots,6\}$ 下的矩阵；

(3) 已知 R^3 中线性变换 T 在基 $\eta=(-1,1,1)^T,\eta=(1,0,-1)^T,\eta=(0,1,1)^T$ 下的矩阵为

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix},$$

求在基 $\varepsilon_1=(1,0,0)^T,\varepsilon_2=(0,1,0)^T,\varepsilon_3=(0,0,1)^T$ 下的矩阵；

(4) 在 R^3 中，线性变换 T 定义如下：

$$\begin{cases} T(\eta_1)=(-5,0,3)^T, \\ T(\eta_2)=(0,-1,6)^T, \\ T(\eta_3)=(-5,-1,9)^T, \end{cases}$$

其中

$$\begin{cases}\eta_1=(-1,0,2)^T,\\ \eta_2=(0,1,1)^T,\\ \eta_3=(3,-1,0)^T,\end{cases}$$

求 T 在基 $\varepsilon_1=(1,0,0)^T,\varepsilon_2=(0,1,0)^T,\varepsilon_3=(0,0,1)^T$ 下的矩阵.

23. 证明:用给定矩阵$\begin{pmatrix}a & b\\ c & d\end{pmatrix}$左乘和右乘二阶方阵,是所有二阶方阵组成的空间中的两个线性变换.并求这两个线性变换在由以下矩阵

$$E_{11}=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix},\quad E_{21}=\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix},$$

$$E_{12}=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix},\quad E_{22}=\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}$$

组成的基下的矩阵.

24. 设 ξ_1,ξ_2,ξ_3,ξ_4 是四维线性空间 V 的一组基,已知线性变换 T 在这组基下的矩阵为

$$\begin{pmatrix}1 & 0 & 2 & 1\\ -1 & 2 & 1 & 3\\ 1 & 2 & 5 & 5\\ 2 & -2 & 1 & -2\end{pmatrix},$$

(1) 求 T 在基 $\eta_1=\xi_1-2\xi_2+\xi_4,\eta_2=3\xi_2-\xi_3-\xi_4,\eta_3=\xi_3+\xi_4,\eta_4=2\xi_4$ 下的矩阵;

(2) 求 T 的核与值域;

(3) 在 T 的核中选一组基,把它扩充成 V 的一组基,并求 T 在这组基下的矩阵;

(4) 在 T 的值域中选一组基,把它扩充成 V 的一组基,并求 T 在这组基下的矩阵.

25. T 是数域 P 上的 n 维线性空间 V 的一个线性变换.证明:如果 T 在任意一组基下的矩阵都相同,则 T 是数乘变换.

26. 设 $A^2=A, B^2=B$，证明：

(1) A 与 B 有相同的值域$\Longleftrightarrow AB=B, BA=A$；

(2) A 与 B 有相同的核$\Longleftrightarrow AB=A, BA=B$.

27. 求复数域上线性空间 V 的线性变换 T 的特征值与特征向量. 已知 T 在一组基下的矩阵为

(1) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$；

(2) $\begin{bmatrix} 0 & 2 & 1 \\ -2 & 0 & 3 \\ -1 & -3 & 0 \end{bmatrix}$；

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

(4) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{bmatrix}$.

28. 在第 27 题中，哪些是单纯矩阵？对于单纯矩阵，求出矩阵 P，使得 $P^{-1}AP$ 为对角阵 Λ.

29. 设 T 是线性空间 V 上的线性变换. 证明：T 可逆的充要条件是 T 没有等于零的特征值.

30. 若方阵 $A\neq 0$，但 $A^k=0$（k 为某一正整数），则 A 不可能相似于对角阵.

31. 设 $A^2=I$，试证：A 的特征值只能为 ± 1；若特征值都等于 1，则 $A=I$.

32. 证明：幂等矩阵 A（即 $A^2=A$）与对角阵相似，并且 $A\sim \mathrm{diag}(I_r, 0)$，其中 $\mathrm{rank}\ A=r$.

33. 设 z 是 A 的特征向量. 证明：z 也是 $f(A)$ 的特征向量，这

里，$f(x)$为x的任一多项式．对应于z的A的特征值和$f(A)$的特征值什么关系？

34．证明：不论A、B为怎样的两个n阶方阵，AB与BA有相同的特征多项式，因此有相同的特征值和迹．

35．设$AB=BA$，证明：A与B有公共的特征向量．

第二章　λ-矩阵与 Jordan 标准形

在第一章§1.5 节中，我们研究了单纯矩阵化为对角阵的问题，而非单纯矩阵则不能与对角形相似. 本章将引进 λ-矩阵这个工具，将矩阵化为 Jordan 标准形.

§2.1　λ-矩　　阵

在矩阵理论中，我们把矩阵定义为数的阵列，即它的元素是数域中的数，统称数字矩阵. 现在我们把数字矩阵加以推广，引进 λ-矩阵. 所谓 **λ-矩阵**，就是以 λ 的复系数多项式为元素的矩阵，故有时也称多项式矩阵. 以下以 $A(\lambda)$，$B(\lambda)$，…来表示 λ-矩阵.

例如，方阵 A 的特征矩阵 $\lambda I - A$ 就是一个 λ-矩阵，数字矩阵当然也可看作 λ-矩阵的特例.

λ-矩阵作为矩阵，它们的运算显然与数字矩阵有相同的规律. 同时可定义正方 λ-矩阵的行列式，记为 $\det A(\lambda)$. 显然有

$$\det(A(\lambda)B(\lambda)) = \det A(\lambda) \cdot \det B(\lambda).$$

有了正方 λ-矩阵的行列式的概念，也就有 λ-矩阵的子式、秩等概念.

定义 1　λ-矩阵 $A(\lambda)$ 中不恒等于零的子式的最高阶数 r 称为 λ-矩阵 $A(\lambda)$ 的**秩**，记为 $\operatorname{rank} A(\lambda)$，即 $\operatorname{rank} A(\lambda) = r$. 若 $\det A(\lambda) \not\equiv 0$ 或 $\det A(\lambda) \equiv 0$，则分别称 $A(\lambda)$ 为**满秩或降秩**.

有了 λ-矩阵的秩，自然联想到它的初等变换及两个 λ-矩阵等价的概念.

定义 2　λ-矩阵的**初等变换**是指下面的三种变换：

(1) 任两行(列)互换；

(2) 用数 $k\neq 0$ 乘某行(列);

(3) 用 λ 的多项式 $\varphi(\lambda)$ 乘某行(列),并加到另一行(列)上去.

上述三种初等变换有相应的三种初等矩阵:

$$P(i,j),\quad P[i(k)],\quad P[j(\varphi),i],$$

并且,在施行行变换时左乘初等矩阵,在施行列变换时右乘初等矩阵.

上面定义的三种初等矩阵是满秩的.满秩阵左乘或右乘一个矩阵并不改变它的秩.

定义 3 若 λ-矩阵 $A(\lambda)$ 经过有限次初等变换化为 λ-矩阵 $B(\lambda)$,则称 $B(\lambda)$ 与 $A(\lambda)$ **等价**,记为

$$A(\lambda)\sim B(\lambda).$$

λ-矩阵的等价关系与数字矩阵一样,满足下面的等价律:

(1) 反身律:$A(\lambda)\sim A(\lambda)$;

(2) 对称律:若 $A(\lambda)\sim B(\lambda)$,则 $B(\lambda)\sim A(\lambda)$;

(3) 传递律:若 $A(\lambda)\sim B(\lambda)$,$B(\lambda)\sim C(\lambda)$,则 $A(\lambda)\sim C(\lambda)$.

若两个 λ-矩阵等价,则它们的秩必相等.但此命题的逆命题却不成立,这与数字矩阵不同.例如

$$A(\lambda)=\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix},\quad B(\lambda)=\begin{pmatrix}1 & -\lambda\\ 1 & \lambda\end{pmatrix},$$

它们的行列式分别为 λ^2 和 2λ,显然秩都是 2.但是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因子,从而 $A(\lambda)$ 与 $B(\lambda)$ 不等价.因此,秩相等只是 λ-矩阵等价的必要条件,而并非充分条件.那么两个 λ-矩阵等价的充分必要条件是什么呢?这个问题将在§2.2节中回答.

下面研究如何将 λ-矩阵化为标准形.

引理 设 λ-矩阵 $A(\lambda)$ 的元素 $a_{11}(\lambda)\neq 0$,且 $A(\lambda)$ 至少有一个元素不能被它整除,则可得到一个 λ-矩阵 $B(\lambda)$ 与 $A(\lambda)$ 等价;$B(\lambda)$ 中的首行首列的元素 $b_{11}(\lambda)\neq 0$,且次数比 $a_{11}(\lambda)$ 的次数低.

证 先设 $A(\lambda)$ 的第 1 行上有元素 $a_{1j}(\lambda)$ 不能被 $a_{11}(\lambda)$ 整除,

此时必有

$$a_{1j}(\lambda)=a_{11}(\lambda)\varphi(\lambda)+b_{11}(\lambda),$$

这里 $b_{11}(\lambda)\neq 0$，且次数比 $a_{11}(\lambda)$ 的次数低. 把第一列的元素乘以 $-\varphi(\lambda)$ 加到第 j 列上，即得到与 $A(\lambda)$ 等价的矩阵，它的第 1 行第 j 列上的元素为 $b_{11}(\lambda)$. 把第 1 列与第 j 列对调，将 $b_{11}(\lambda)$ 移到第 1 行第 1 列的位置，这样得到的 λ-矩阵 $B(\lambda)$ 显然合乎引理的要求.

同理，若 $A(\lambda)$ 的第 1 列上有元素 $a_{i1}(\lambda)$ 不能被 $a_{11}(\lambda)$ 整除，也可仿照上面的做法，达到引理的要求.

现设第 1 行和第 1 列的所有元素都能被 $a_{11}(\lambda)$ 整除，而 $a_{ij}(\lambda)$ 不能被 $a_{11}(\lambda)$ 整除. 由于 $a_{i1}(\lambda)$ 能被 $a_{11}(\lambda)$ 整除，设 $a_{i1}(\lambda)=a_{11}(\lambda)\varphi(\lambda)$. 将第 1 行乘以 $-\varphi(\lambda)$ 加到第 i 行上，这样得到的矩阵，其元素

$$a_{i1}'(\lambda)=0,a_{ij}'=a_{ij}(\lambda)-a_{1j}(\lambda)\varphi(\lambda).$$

显然，$a_{11}(\lambda)$ 也不能整除 $a_{ij}'(\lambda)$. 再把第 i 行加到第 1 行，这时 $a_{11}(\lambda)$ 没有改变（注意 $a_{i1}'(\lambda)=0$），而第 1 行第 j 列的元素变为 $[1-\varphi(\lambda)]a_{1j}(\lambda)+a_{ij}(\lambda)$，它不能被 $a_{11}(\lambda)$ 整除. 此时化为第一种情况，于是引理得证.

定理 1 若 $A(\lambda)$ 为 m 行 n 列的 λ-矩阵，rank $A(\lambda)=r$，则

$$A(\lambda)\sim\begin{pmatrix} d_1(\lambda) & 0 & \cdots & 0 & \cdots & 0 \\ 0 & d_2(\lambda) & \cdots & 0 & \cdots & 0 \\ \cdots & & & & & \\ 0 & 0 & \cdots & d_r(\lambda) & \cdots & 0 \\ \cdots & & & & & \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}. \tag{2.1}$$

这里，$d_i(\lambda)$ 为首一多项式（首项系数为 1 的多项式），并且 $d_i(\lambda)\mid d_{i+1}(\lambda)$，$i=1,2,\cdots,r-1$. 这种形式的 λ-矩阵称为 $A(\lambda)$ 的 **Smith 标准形**.

证 若 rank $A(\lambda)=0$，则 $A(\lambda)$ 为零矩阵，没有讨论的必要. 现设 rank $A(\lambda)=r>0$，且 $A(\lambda)$ 中 $a_{11}(\lambda)\neq 0$，若不然，可通过行、列的对调做到这一点.

如果 $A(\lambda)$的元素不是全部被 $a_{11}(\lambda)$整除，则由引理 1 可找到一个等价矩阵，使其左上角的元素次数比 $a_{11}(\lambda)$低. 若这时左上角元素还不能整除矩阵的全体元素，则可用同样的方法，逐步降低左上角元素的次数，直到得到一个等价矩阵 $B(\lambda)$，其左上角元素 $b_{11}(\lambda)$可整除其它所有元素，且 $b_{11}(\lambda)$为首一多项式. 由于 $b_{11}(\lambda)$可整除其它所有元素，可将适当的多项式乘以第 1 行后加到其它各行，使第 1 列中除 $b_{11}(\lambda)$外的其余元素全为 0；用同样的方法可将第 1 行中除 $b_{11}(\lambda)$外的其余元素全化为 0. 这样得到与 $A(\lambda)$等价的形如下式的矩阵

$$\begin{bmatrix} d_1(\lambda) & 0 & 0 & \cdots & 0 \\ 0 & c_{22}(\lambda) & c_{23}(\lambda) & \cdots & c_{2n}(\lambda) \\ 0 & c_{32}(\lambda) & c_{33}(\lambda) & \cdots & c_{3n}(\lambda) \\ \cdots & & & & \\ 0 & c_{m2}(\lambda) & c_{m3}(\lambda) & \cdots & c_{mn}(\lambda) \end{bmatrix}, \tag{2.2}$$

且 $d_1(\lambda)$(即 $b_{11}(\lambda)$)可整除所有的 $c_{ij}(\lambda)$.

现在对上面的$(m-1)\times(n-1)$子阵$[c_{ij}(\lambda)]$进行同样的变换. 注意，这样的初等变换也是对矩阵(2.2)式的初等变换. 这样把(2.2)式化为 $A(\lambda)$的等价矩阵

$$\begin{bmatrix} d_1(\lambda) & 0 & 0 & \cdots & 0 \\ 0 & d_2(\lambda) & 0 & \cdots & 0 \\ 0 & 0 & f_{33}(\lambda) & \cdots & f_{3n}(\lambda) \\ \cdots & & & & \\ 0 & 0 & f_{m3}(\lambda) & \cdots & f_{mn}(\lambda) \end{bmatrix}, \tag{2.3}$$

这里 $d_1(\lambda)\,|\,d_2(\lambda)\,|\,f_{ij}(\lambda)$.

这样一直做下去，即可得到(2.1)式右端的形式，即主对角线上元素顺次为 $d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda)$，其它元素全为 0，每个 $d_i(\lambda)$为首一多项式，且满足 $d_i(\lambda)\,|\,d_{i+1}(\lambda)$，$i=1,2,\cdots,r-1$.

例 1 将下面的 λ-矩阵化为 Smith 标准形：

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

解 对 λ-矩阵 $A(\lambda)$施行初等变换,可得

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

$$\xrightarrow{c_1+c_3}\begin{pmatrix}1 & \lambda^2 & \lambda\\ 0 & \lambda & -\lambda\\ 1 & \lambda^2 & -\lambda^2\end{pmatrix}$$

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

$$\xrightarrow{r_3\times(-1)}\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \lambda(\lambda+1)\end{pmatrix}.$$

§2.2 不变因子及初等因子

在§2.1节中已经指出,秩相等是两个 λ-矩阵等价的必要条件,此外还需具备什么性质才能使两个 λ-矩阵等价呢? 下面来讨论这个问题.

设 rank $A(\lambda)=r$,当 $k<r$ 时、$A(\lambda)$的 k **阶行列式因子**定义为 $A(\lambda)$的所有 k 阶子式的最高公因式(取首一多项式);当 $k>r$ 时规定为零. $A(\lambda)$的 k 阶行式因子记作 $D_k(\lambda)$,并规定 $D_0(\lambda)=1$.

定理 1 λ-矩阵 $A(\lambda)$的各阶行列式因子 $D_k(\lambda)(k=1,2,\cdots,n)$在初等变换下不变,即等价的 λ-矩阵有相同的行列式因子 $D_k(\lambda)(k=1,2,\cdots,n)$. 特别地,等价的 λ-矩阵有相同的秩.

证 我们就初等行变换来证明,且证明过程中将 $k\leqslant r$ 或 $k>r$ 的情况都包括进去.

首先,对于前两种初等行变换而言,所得的两个等价的 λ-矩阵显然有相同的 k 阶行列式因子 $D_k(\lambda)$, $k=1,2,\cdots,n$.

其次,考虑第三种初等变换,即用初等矩阵 $P[j(\varphi),i]$右乘 λ-矩阵 $A(\lambda)$,得 $P[j(\varphi),i]A(\lambda)$. 它们的各个对应的 k 阶子式中如果不包含第 i 行或同时包含第 i 行和第 j 行,这样对应的子式是相等的. 如果只含第 i 行而不含第 j 行,则 $P[j(\varphi),i]A(\lambda)$的子式可表示为 $G_k(\lambda)+\varphi(\lambda)H_k(\lambda)$,式中 $G_k(\lambda)$和 $H_k(\lambda)$为 $A(\lambda)$的两个 k 阶子式. 因此,$P[j(\varphi),i]A(\lambda)$的 k 阶行列式因子 $D_k'(\lambda)$可被 $A(\lambda)$的 k 阶行列式因子 $D_k(\lambda)$整除. 但第三种初等变换的逆变换仍为第三种初变换,因此 $D_k(\lambda)$又可被 $D_k'(\lambda)$整除. 又因为它们都是首一多项式,所以 $D_k'(\lambda)=D_k(\lambda)$, $k=1,2,\cdots,r$.

由于 $D_k(\lambda)$在初等变换下不变,因此当 $k>r$ 时,$A(\lambda)$的 $D_k(\lambda)$都为零,就是说有相同的秩.

定理 2 λ-矩阵 $A(\lambda)$的 Smith 标准形

$$\operatorname{diag}[d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda),0,\cdots,0]$$

是唯一的,且 $d_k(\lambda)=\dfrac{D_k(\lambda)}{D_{k-1}(\lambda)}$, $k=1,2,\cdots,r$. 而当 $k>r$ 时,$d_k(\lambda)=0$.

证 因为

$$A(\lambda)\sim\operatorname{diag}[d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda),0,\cdots,0],$$

所以它们有相同的行列式因子. 但 $A(\lambda)$的 Smith 标准形为对角阵,且 $d_i(\lambda)|d_{i+1}(\lambda)$, $i=1,2,\cdots,r-1$;$d_i=0,i>r$. 故

$$D_1(\lambda)=d_1(\lambda);$$

$$D_2(\lambda)=d_1(\lambda)d_2(\lambda);$$

$$\cdots\cdots$$

$$D_r(\lambda)=d_1(\lambda)d_2(\lambda)\cdots d_r(\lambda);$$

$$D_k(\lambda)=0\qquad(k>r).$$

因此

$$d_k(\lambda)=\frac{D_k(\lambda)}{D_{k-1}(\lambda)}\qquad(k=1,2,\cdots,r).$$

由于行列式因子在初等变换下不变，所以 $d_k(\lambda)$ 也就唯一地被确定.

今后称 $d_k(\lambda)$ 为 $A(\lambda)$ 的**第 k 个不变因子**.

有了 λ-矩阵的行列式因子或不变因子，就可解决两个 λ-矩阵等价的充要条件的问题.

定理 3 设有 λ-矩阵 $A(\lambda)$ 与 $B(\lambda)$，则 $A(\lambda)\sim B(\lambda)$ 的充要条件是它们有相同的行列式因子或有相同的不变因子.

证 如果 $A(\lambda)\sim B(\lambda)$，则它们有相同的行列式因子，因而有相同的不变因子. 反之，若它们有相同的行列式因子或相同的不变因子，它们就有相同的 Smith 标准形，因而也等价.

考虑到，任一复系数多项式都可以分解为一次因式之积. 假设 $A(\lambda)$ 的各个不变因子的分解式为

$$\left.\begin{aligned}d_1(\lambda)&=(\lambda-\lambda_1)^{e_{11}}(\lambda-\lambda_2)^{e_{12}}\cdots(\lambda-\lambda_s)^{e_{1s}},\\ d_2(\lambda)&=(\lambda-\lambda_1)^{e_{21}}(\lambda-\lambda_2)^{e_{22}}\cdots(\lambda-\lambda_s)^{e_{2s}},\\ &\cdots\cdots\\ d_r(\lambda)&=(\lambda-\lambda_1)^{e_{r1}}(\lambda-\lambda_2)^{e_{r2}}\cdots(\lambda-\lambda_s)^{e_{rs}},\end{aligned}\right\}\qquad(2.4)$$

这里 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 互不相等. 因为 $d_k(\lambda)\mid d_{k+1}(\lambda)$，所以

$$e_{1j}\leqslant e_{2j}\leqslant\cdots\leqslant e_{rj}\qquad(j=1,2,\cdots,s).$$

由于 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 相异，$e_{r1},e_{r2},\cdots,r_{rs}$ 全不为零. 但 $e_{1j},e_{2j},\cdots,e_{r-1,j}$，$j=1,2,\cdots,s$ 可能有为零的，如果某个 $e_{ij}=0$ $(j=1,2,\cdots,s;i=1,2,\cdots,r-1)$，必有

$$e_{1j}=e_{2j}=\cdots=e_{i-1,j}=0.$$

在(2.4)式中，所有指数大于零的因子

$$(\lambda-\lambda_j)^{e_{ij}} \qquad (j=1,2,\cdots,s;i=1,2,\cdots,r,e_{ij}>0)$$

都称为 $A(\lambda)$的**初等因子**.

由初等因子的定义可知,如果给定 $A(\lambda)$的不变因子,那么它的初等因子就唯一地决定了.反之,如果给定 $A(\lambda)$的所有初等因子及 $A(\lambda)$的秩,则它的不变因子也被唯一决定.这是因为由初等因子的定义易知,如果把 $A(\lambda)$的所有初等因子按不同的一次因子分类,并按各因子的幂的大小排列出来:

$$(\lambda-\lambda_1)^{e_{11}},\ (\lambda-\lambda_1)^{e_{21}},\ \cdots,\ (\lambda-\lambda_1)^{e_{r1}} \qquad (e_{11}\leqslant e_{21}\leqslant\cdots\leqslant e_{r1}),$$

$$(\lambda-\lambda_2)^{e_{12}},\ (\lambda-\lambda_2)^{e_{22}},\ \cdots,\ (\lambda-\lambda_2)^{e_{r2}} \qquad (e_{12}\leqslant e_{22}\leqslant\cdots\leqslant e_{r2}),$$

$$\cdots\cdots$$

$$(\lambda-\lambda_s)^{e_{1s}},\ (\lambda-\lambda_s)^{e_{2s}},\ \cdots,\ (\lambda-\lambda_s)^{e_{rs}} \qquad (e_{1s}\leqslant e_{2s}\leqslant\cdots\leqslant e_{rs}).$$

在这个表里共有 r 列,每列里的因子个数可能不同,在空白处都用 1 补上,于是第 i 列上各式的积

$$(\lambda-\lambda_1)^{e_{i1}}(\lambda-\lambda_2)^{e_{i2}}\cdots(\lambda-\lambda_s)^{e_{is}}$$

便是第 i 个不变因子.因此,当已知一个 λ-矩阵 $A(\lambda)$的秩 r 后,求不变因子的问题等价于求初等因子的问题.

定理 4 设 λ-矩阵

$$A(\lambda)=\begin{pmatrix} B(\lambda) & 0 \\ 0 & C(\lambda) \end{pmatrix}$$

为分块对角形,则 $B(\lambda)$,$C(\lambda)$的各个初等因子的全体构成 $A(\lambda)$的全体初等因子.

证 先将 $B(\lambda)$与 $C(\lambda)$分别化为 Smith 标准形

$$B(\lambda)\sim\begin{pmatrix} b_1(\lambda) & & & & & \\ & \ddots & & & & \\ & & b_{r_1}(\lambda) & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix},$$

$$
C(\lambda)\sim\begin{pmatrix}
c_1(\lambda) & & & & & \\
 & \ddots & & & & \\
 & & c_{r_2}(\lambda) & & & \\
 & & & 0 & & \\
 & & & & \ddots & \\
 & & & & & 0
\end{pmatrix},
$$

这里，$A(\lambda)$的秩 $r=r_1+r_2$.

将 $b_k(\lambda)$、$c_l(\lambda)$分解为不同的一次因子的幂积：

$$b_k(\lambda)=(\lambda-\lambda_1)^{e_{k1}}(\lambda-\lambda_2)^{e_{k2}}\cdots(\lambda-\lambda_t)^{e_{kt}}\qquad(k=1,2,\cdots,r_1),$$

$$c_l(\lambda)=(\lambda-\lambda_1)^{i_{l1}}(\lambda-\lambda_2)^{i_{l2}}\cdots(\lambda-\lambda_t)^{i_{lt}}\qquad(l=1,2,\cdots,r_2),$$

因此，$B(\lambda)$与 $C(\lambda)$的初等因子分别是

$$(\lambda-\lambda_1)^{e_{11}},(\lambda-\lambda_1)^{e_{21}},$$

$$\cdots\cdots,$$

$$(\lambda-\lambda_1)^{e_{r_1 1}},\cdots,(\lambda-\lambda_t)^{e_{1t}},\cdots,(\lambda-\lambda_t)^{e_{r_1 t}}$$

及

$$(\lambda-\lambda_1)^{i_{11}},(\lambda-\lambda_1)^{i_{21}},$$

$$\cdots\cdots,$$

$$(\lambda-\lambda_1)^{i_{r_2 1}},\cdots,(\lambda-\lambda_t)^{i_{1t}},\cdots,(\lambda-\lambda_t)^{i_{r_2 t}}$$

中不为常数的多项式.

现在来证明 $B(\lambda)$、$C(\lambda)$的初等因子就是 $A(\lambda)$的全部初等因子. 不失一般性，先考虑 $B(\lambda)$、$C(\lambda)$中只含$(\lambda-\lambda_1)$的那些初等因子，将

$$e_{11},\ e_{21},\ \cdots,\ e_{r_1 1},\ i_{11},\ i_{21},\ \cdots,\ i_{r_2 1}$$

按大小重新排列，记为

$$j_1,j_2,\cdots,j_r,$$

满足

$$0\leqslant j_1\leqslant j_2\leqslant\cdots\leqslant j_r.$$

由 $A(\lambda)$的分块对角形易见，在矩阵 $B(\lambda)$和 $C(\lambda)$上施行初等

变换，都可看作在 $A(\lambda)$ 上施行初等变换，从而

$$A \sim \begin{bmatrix} b_1(\lambda) & & & & & & & & \\ & \ddots & & & & & & & \\ & & b_r(\lambda) & & & & & & \\ & & & c_1(\lambda) & & & & & \\ & & & & \ddots & & & & \\ & & & & & c_s(\lambda) & & & \\ & & & & & & 0 & \cdots & 0 \\ & & & & & & \vdots & & \vdots \\ & & & & & & 0 & \cdots & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} (\lambda-\lambda_1)^{j_1}\varphi_1(\lambda) & & & & & & \\ & (\lambda-\lambda_1)^{j_2}\varphi_2(\lambda) & & & & & \\ & & \ddots & & & & \\ & & & (\lambda-\lambda_1)^{j_r}\varphi_r(\lambda) & & & \\ & & & & 0 & \cdots & 0 \\ & & & & \vdots & & \vdots \\ & & & & 0 & \cdots & 0 \end{bmatrix},$$

其中多项式 $\varphi_1(\lambda),\cdots,\varphi_r(\lambda)$ 中都不含 $(\lambda-\lambda_1)$. 因此，在行列式因子 $D_r(\lambda),D_{r-1}(\lambda),\cdots,D_1(\lambda)$ 中，$(\lambda-\lambda_1)$ 的幂指数分别为

$$\sum_{k=1}^{r} j_k,\ \sum_{k=1}^{r-1} j_k,\ \cdots,\ \sum_{k=1}^{2} j_k,\ j_1.$$

再由不变因子 $d_k(\lambda)$ 与行列式因子 $D_k(\lambda)$ 的关系

$$d_k(\lambda)=\frac{D_k(\lambda)}{D_{k-1}(\lambda)} \qquad (k=1,2,\cdots,r)$$

知，$d_r(\lambda),d_{r-1}(\lambda),\cdots,d_1(\lambda)$ 中 $(\lambda-\lambda_1)$ 的幂指数分别为 $j_r,j_{r-1},\cdots,j_1$. 这就是说，$A(\lambda)$ 中含有 $(\lambda-\lambda_1)$ 的初等因子为

$$(\lambda-\lambda_1)^{j_k} \qquad (j_k\neq 0,k=1,2,\cdots,r).$$

对 $\lambda-\lambda_2,\lambda-\lambda_3,\cdots,\lambda-\lambda_t$，可类似地进行讨论. 这样就证明了

$B(\lambda)$,$C(\lambda)$的初等因子就是 $A(\lambda)$的初等因子.

下面还须证明,除 $B(\lambda)$,$C(\lambda)$的初等因子外,$A(\lambda)$再没有其它初等因子了.

因为 $A(\lambda)$的秩为 r,所以 $D_r(\lambda)$为 $A(\lambda)$的所有初等因子的乘积. 考虑到

$$D_r(\lambda)=b_1(\lambda)\cdots b_{r_1}(\lambda)c_1(\lambda)\cdots c_{r_2}(\lambda),$$

因此,若$(\lambda-a)^k$ 为 $A(\lambda)$ 初等因子,则它必包含在某个 $b_i(\lambda)$, $i=1,2,\cdots,r_1$ 或 $c_j(\lambda)$, $j=1,2,\cdots,r_2$ 中. 即 $A(\lambda)$的初等因子包含在 $B(\lambda)$和 $C(\lambda)$的初等因子中. 至此定理 4 全部证毕.

由归纳法可知,对于一般的分块对角形矩阵

$$A(\lambda)=\begin{pmatrix} B_1(\lambda) & & & & \\ & B_2(\lambda) & & & \\ & & \ddots & & \\ & & & B_t(\lambda) & \\ & & & & 0 \end{pmatrix},$$

各子块 $B_i(\lambda)$, $i=1,2,\cdots,t$ 的初等因子的全体构成 $A(\lambda)$的全部初等因子.

例 1 求下面矩阵 $A(\lambda)$的初等因子、不变因子和 Smith 标准形:

$$A(\lambda)=\begin{pmatrix} 3\lambda-5 & (\lambda+2)^2 & 4\lambda+5 & (\lambda-1)^2 \\ \lambda+7 & (\lambda+2)^2 & \lambda+7 & 0 \\ \lambda-1 & 0 & 2\lambda-1 & (\lambda-1)^2 \\ 0 & 0 & (\lambda-2)(\lambda-5) & 0 \end{pmatrix}.$$

解 对 $A(\lambda)$施行初等变换,可得

$$A(\lambda)\xrightarrow{r_1-r_3}\begin{pmatrix} 2\lambda+6 & (\lambda+2)^2 & 2\lambda+6 & 0 \\ \lambda+7 & (\lambda+2)^2 & \lambda+7 & 0 \\ \lambda-1 & 0 & 2\lambda-1 & (\lambda-1)^2 \\ 0 & 0 & (\lambda-2)(\lambda-5) & 0 \end{pmatrix}$$

$$\xrightarrow{c_3-c_1}\begin{bmatrix} 2\lambda+6 & (\lambda+2)^2 & 0 & 0 \\ \lambda+7 & (\lambda+2)^2 & 0 & 0 \\ \lambda-1 & 0 & \lambda & (\lambda-1)^2 \\ 0 & 0 & (\lambda-2)(\lambda-5) & 0 \end{bmatrix}$$

$$\xrightarrow{r_1-r_2}\begin{bmatrix} \lambda-1 & 0 & 0 & 0 \\ \lambda+7 & (\lambda+2)^2 & 0 & 0 \\ \lambda-1 & 0 & \lambda & (\lambda-1)^2 \\ 0 & 0 & (\lambda-2)(\lambda-5) & 0 \end{bmatrix}$$

$$\xrightarrow{r_3-r_1}\begin{bmatrix} \lambda-1 & 0 & 0 & 0 \\ \lambda+7 & (\lambda+2)^2 & 0 & 0 \\ 0 & 0 & \lambda & (\lambda-1)^2 \\ 0 & 0 & (\lambda-2)(\lambda-5) & 0 \end{bmatrix}.$$

至此，$A(\lambda)$已化为分块对角阵

$$\begin{bmatrix} A_1(\lambda) & 0 \\ 0 & A_2(\lambda) \end{bmatrix},$$

其中

$$A_1(\lambda)=\begin{pmatrix} \lambda-1 & 0 \\ \lambda+7 & (\lambda+2)^2 \end{pmatrix},$$

$$A_2(\lambda)=\begin{pmatrix} \lambda & (\lambda-1)^2 \\ (\lambda-2)(\lambda-5) & 0 \end{pmatrix}.$$

对于 $A_1(\lambda)$，有

$$D_1(\lambda)=1,\quad D_2(\lambda)=(\lambda-1)(\lambda+2)^2,$$

所以

$$d_1(\lambda)=1,\quad d_2(\lambda)=(\lambda-1)(\lambda+2)^2,$$

其初等因子为$(\lambda-1)$和$(\lambda+2)^2$；而对于 $A_2(\lambda)$，有

$$D_1(\lambda)=1,\quad D_2(\lambda)=(\lambda-2)(\lambda-5)(\lambda-1)^2,$$

所以

$$d_1(\lambda)=1,\quad d_2(\lambda)=(\lambda-2)(\lambda-5)(\lambda-1)^2,$$

其初等因子为$(\lambda-2)$,$(\lambda-5)$和$(\lambda-1)^2$.

由定理 4,$A(\lambda)$的初等因子为

$$\lambda-1,\ (\lambda-1)^2,\ \lambda-2,\ \lambda-5,\ (\lambda+2)^2,$$

又 rank $A(\lambda)=4$,故 $A(\lambda)$的不变因子为

$$d_4(\lambda)=(\lambda-2)(\lambda-5)(\lambda-1)^2(\lambda+2)^2,$$

$$d_3(\lambda)=\lambda-1,$$

$$d_2(\lambda)=d_1(\lambda)=1,$$

因此,$A(\lambda)$的 Smith 标准形为

$$\begin{pmatrix}1 & & & \\ & 1 & & \\ & & \lambda-1 & \\ & & & (\lambda-2)(\lambda-5)(\lambda-1)^2(\lambda+2)^2\end{pmatrix}.$$

例 2 求 λ-矩阵

$$A(\lambda)=\begin{pmatrix}\lambda-a & c_1 & & & \\ & \lambda-a & c_2 & & \\ & & \ddots & \ddots & \\ & & & \lambda-a & c_{n-1} \\ & & & & \lambda-a\end{pmatrix}$$

的不变因子和初等因子,其中 $c_1,c_2,\cdots,c_{n-1}$为非零常数.

解 因为 $A(\lambda)$为 n 阶方阵,所以 $D_n(\lambda)=(\lambda-a)^n$,去掉第 1 列和第 n 行后余下的$(n-1)$阶子式为非零常数 $c_1c_2\cdots c_{n-1}$,故 $D_{n-1}(\lambda)=1$,因此

$$D_1(\lambda)=D_2(\lambda)=\cdots=D_{n-1}(\lambda)=1,$$

于是 $A(\lambda)$的不变因子为

$$d_1(\lambda)=d_2(\lambda)=\cdots=d_{n-1}(\lambda)=1,$$

$$d_n(\lambda)=(\lambda-a)^n;$$

$A(\lambda)$的初等因子为$(\lambda-a)^n$，其 Smith 标准形为

$$A(\lambda)\sim\begin{bmatrix}1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & (\lambda-a)^n\end{bmatrix}.$$

例 3 在控制论中，常见的矩阵是相伴矩阵(友矩阵)：

$$A=\begin{bmatrix}0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & 1\\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1\end{bmatrix},$$

求它的特征矩阵的不变因子，并将特征矩阵化为 Smith 标准形.

解 设 A 的特征矩阵为 $A(\lambda)$，即

$$A(\lambda)=\lambda I-A=\begin{bmatrix}\lambda & -1 & 0 & \cdots & 0 & 0\\ 0 & \lambda & -1 & \cdots & 0 & 0\\ \cdots & & & & & \\ 0 & 0 & 0 & \cdots & \lambda & -1\\ a_n & a_{n-1} & a_{n-2} & \cdots & a_2 & \lambda+a_1\end{bmatrix}.$$

将 $A(\lambda)$的第 $2,3,\cdots,n$ 列依次乘以 $\lambda,\lambda^2,\cdots,\lambda^{n-1}$后都加到第 1 列上去，得

$$A(\lambda)\sim\begin{bmatrix}0 & -1 & 0 & \cdots & 0 & 0\\ 0 & \lambda & -1 & \cdots & 0 & 0\\ \cdots & & & & & \\ 0 & 0 & 0 & \cdots & \lambda & -1\\ f(\lambda) & a_{n-1} & a_{n-2} & \cdots & a_2 & \lambda+a_1\end{bmatrix}.$$

把上式右端的矩阵记为 $B(\lambda)$，其中

$$f(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-2}\lambda^2+a_{n-1}\lambda+a_n,$$

且

$$\det A(\lambda)=\det B(\lambda)=(-1)^{n+1}f(\lambda)\cdot(-1)^{n-1}=f(\lambda),$$

即 $D_n(\lambda)=f(\lambda)$. 由于将 $B(\lambda)$ 的第 1 列第 n 行去掉后剩下的 $n-1$ 阶子式为 $(-1)^{n-1}$，故

$$D_1(\lambda)=D_2(\lambda)=\cdots=D_{n-1}(\lambda)=1,$$

从而

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

所以 $A(\lambda)$ 的 Smith 标准形为

$$\begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & f(\lambda) \end{bmatrix}.$$

注 如果矩阵 A 的特征矩阵的不变因子为

$$1,\ 1,\ \cdots,\ 1,\ f(\lambda),$$

则称 A 为属于 $f(\lambda)$ 的**相伴矩阵**(或**友矩阵**).

§2.3 Jordan 标准形

虽然非单纯矩阵不能相似于对角阵，但它能相似于一个形式上比对角阵稍复杂的约当(Jordan)标准形 J. 由于 Jordan 标准形的独特结构揭示了两个矩阵相似的本质关系，故在数值计算和理论推导中经常采用. 利用它不仅容易求出 A 的乘幂，还可以讨论矩阵函数和矩阵级数，求解矩阵微分方程. 因此，Jordan 标准形的理论在数学、力学和计算方法中得到广泛的应用.

定义 1 形如

$$J_i=\begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{bmatrix}_{m_i\times m_i} \tag{2.5}$$

的方阵称为 **m_i 阶 Jordan 块**，其中 λ 可以是实数，也可以是复数.

例如

$$\begin{pmatrix} 3 & 1 \\ & 3 \end{pmatrix},\quad \begin{pmatrix} \mathrm{i} & 1 & \\ & \mathrm{i} & 1 \\ & & \mathrm{i} \end{pmatrix},\quad \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix},$$

都是 Jordan 块. 特别地，一阶方阵是一阶 Jordan 块.

定义 2 由若干个 Jordan 块组成的分块对角阵

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_t \end{pmatrix}, \tag{2.6}$$

其中 $J_i(i=1,2,\cdots,t)$ 为 m_i 阶 Jordan 块，且当 $\sum\limits_{i=1}^{t} m_i = n$ 时，称为 **n 阶 Jordan 标准形**，记为 J.

例如

$$J = \begin{pmatrix} 3 & 1 & & & & & & & \\ 0 & 3 & & & & & & & \\ & & \mathrm{i} & 1 & 0 & & & & \\ & & 0 & \mathrm{i} & 1 & & & & \\ & & 0 & 0 & \mathrm{i} & & & & \\ & & & & & 0 & 1 & 0 & 0 \\ & & & & & 0 & 0 & 1 & 0 \\ & & & & & 0 & 0 & 0 & 1 \\ & & & & & 0 & 0 & 0 & 0 \end{pmatrix}$$

是 9 阶 Jordan 标准形.

下面讨论任何一个矩阵 A 与 Jordan 标准形 J 相似的条件，以及如何将矩阵 A 化为 Jordan 标准形 J. 显然，对角阵为 Jordan 矩阵的特例，其中每个 Jordan 块都是一阶的.

由§2.2节的例2知，(2.5)式的 m_i 阶 Jordan 块的特征矩阵 $\lambda I-J_i$ 的初等因子为 $(\lambda-\lambda_i)^{m_i}$，因此，Jordan 标准形(2.6)式的特征矩阵 $\lambda I-J$ 的初等因子是同各个 Jordan 块的初等因子合在一起构成的，即 $\lambda I-J$ 的全部初等因子是

$$(\lambda-\lambda_1)^{m_1},\ (\lambda-\lambda_2)^{m_2},\ \cdots,\ (\lambda-\lambda_t)^{m_t},$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_t$ 中可能有相同的值，指数 $m_1,m_2,\cdots,m_t$ 中也可能有相同的值，但总有

$$\sum_{i=1}^{t} m_i = n.$$

在(2.6)式中，Jordan 标准形除了各个 Jordan 块的排列次序外，完全由三个数所决定，即由所含 Jordan 块的个数 t、每个 Jordan 块的阶数 m_i 以及主对角线上的元素 λ_i 所决定. 这三个数都在 J 的全部初等因子中得到全面反映. 所以，如果不计较 J 中各 Jordan 块的排列顺序，则 Jordan 标准形可由它的全部初等因子所唯一决定.

定理 1 $A\sim J \Longleftrightarrow (\lambda I-A)\sim(\lambda I-J)$.

证 定理的充分性证明涉及到较多的预备知识，故从略. 下面证明其必要性.

若 $A\sim J$，则必存在可逆矩阵 P，使得

$$P^{-1}AP=J,$$

故

$$P^{-1}(\lambda I-A)P=\lambda I-J,$$

因此

$$(\lambda I-A)\sim(\lambda I-J).$$

定理 2 每个 n 阶复矩阵 A 都与一个 Jordan 标准形 J 相似. 这个 Jordan 标准形在不计其中 Joradn 块的排列次序时，完全由矩阵 A 唯一决定. 即每个矩阵都有 Jordan 标准形.

证 设 n 阶矩阵 A 的特征矩阵 $\lambda I-A$ 的初等因子(也称为矩阵 A 的初等因子)为

$$(\lambda-\lambda_1)^{m_1},\ (\lambda-\lambda_2)^{m_2},\ \cdots,\ (\lambda-\lambda_t)^{m_t}, \tag{2.7}$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_t$ 及 $m_1,m_2,\cdots,m_t$ 中可能有相同的，且 $\sum\limits_{i=1}^{t} m_i=n$. 每个初等因子 $(\lambda-\lambda_i)^{m_i}$ 对应于一个 Jordan 块 J_i，其阶数为 m_i，对角线元素为 λ_i. 这些 Jordan 块构成一个 Jordan 标准形 J，经过计算可知，J 的全部初等因子就是(2.7)式. 因为 J 与 A 有相同的初等因子，亦即 $\lambda I-A$ 与 $\lambda I-J$ 有相同的初等因子，因此 $(\lambda I-1)\sim(\lambda I-J)$. 由定理 1 知，$A\sim J$.

如果还有约当标准形 J'，使得 $J'\sim A$，则 J' 与 A 也有相同的初等因子. 因此，J' 与 J 在不计其 Jordan 块的排列次序时是相同的. 这样就证明了唯一性.

推论 复矩阵 A 可对角化的充要条件是 A 的特征矩阵的初等因子全为一次.

例 1 求矩阵

$$A=\begin{pmatrix}-1 & -2 & 6\\ -1 & 0 & 3\\ -1 & -1 & 4\end{pmatrix}$$

的 Jordan 标准形.

解 先求 $\lambda I-A$ 的初等因子.

$$\lambda I-A=\begin{pmatrix}\lambda+1 & 2 & -6\\ 1 & \lambda & -3\\ 1 & 1 & \lambda-4\end{pmatrix}\sim\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda-1 & 0\\ 0 & 0 & (\lambda-1)^2\end{pmatrix},$$

因此，$\lambda I-A$ 的初等因子为 $\lambda-1$，$(\lambda-1)^2$，矩阵 A 的 Jordan 标准形为

$$J=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}.$$

例 2 求矩阵

$$A=\begin{pmatrix}2 & 0 & 2 & 1\\ 6 & 1 & 4 & 4\\ 10 & 0 & 0 & 4\\ 7 & 0 & -7 & 2\end{pmatrix}$$

的 Jordan 标准形.

解 矩阵 A 的特征多项式为

$$\begin{aligned}D_4(\lambda)&=\det(\lambda I-A)\\ &=\det\begin{pmatrix}\lambda-2 & 0 & -2 & -1\\ -6 & \lambda-1 & -4 & -4\\ -10 & 0 & \lambda & -4\\ -7 & 0 & 7 & \lambda-2\end{pmatrix}\\ &=(\lambda-2)(\lambda-1)^3,\end{aligned}$$

去掉 $\lambda I-A$ 的第 4 行第 4 列后的余子式 M_{44} 为

$$\begin{aligned}M_{44}&=\det\begin{pmatrix}\lambda-2 & 0 & -2\\ -6 & \lambda-1 & -4\\ -10 & 0 & \lambda\end{pmatrix}\\ &=(\lambda-1)(\lambda^2-2\lambda-20),\end{aligned}$$

同理可得

$$\begin{aligned}M_{12}&=\det\begin{pmatrix}-6 & -4 & -4\\ -10 & \lambda & -4\\ -7 & 7 & \lambda-2\end{pmatrix}\\ &=-2(3\lambda^2+28\lambda-40),\end{aligned}$$

因为 M_{44} 与 M_{12} 互素,所以 $D_3(\lambda)=1$,进而有

$$D_2(\lambda)=D_1(\lambda)=D_0(\lambda)=1.$$

因此,$\lambda I-A$ 的不变因子为

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

$$d_4(\lambda)=(\lambda-2)(\lambda-1)^3,$$

初等因子为 $(\lambda-2)$ 和 $(\lambda-1)^3$,所以 A 的 Jordan 标准形为

$$J=\begin{pmatrix}2&0&0&0\\0&1&1&0\\0&0&1&1\\0&0&0&1\end{pmatrix}.$$

例 3 求矩阵

$$A=\begin{pmatrix}3&1&&\\-4&-1&&\\&&2&1\\&&-1&0\end{pmatrix}$$

的 Jordan 标准形.

解 将 A 写成分块形式

$$A=\begin{pmatrix}A_1&\\&A_2\end{pmatrix},$$

其中

$$A_1=\begin{pmatrix}3&1\\-4&-1\end{pmatrix},\quad A_2=\begin{pmatrix}2&1\\-1&0\end{pmatrix}.$$

先分别求出子矩阵 A_1,A_2 的 Jordan 标准形. 由

$$\lambda I-A_1=\begin{pmatrix}\lambda-3&-1\\4&\lambda+1\end{pmatrix}$$

可知,A_1 的不变因子为 $d_1(\lambda)=1,d_2(\lambda)=(\lambda-1)^2$,故初等因子为 $(\lambda-1)^2$. 因此

$$A_1\sim J_1=\begin{pmatrix}1&1\\0&1\end{pmatrix}.$$

再由

$$\lambda I-A_2=\begin{pmatrix}\lambda-2&-1\\1&\lambda\end{pmatrix}$$

可知,A_2 的不变因子为 $d_1(\lambda)=1,d_2(\lambda)=(\lambda-1)^2$,故初等因子为 $(\lambda-1)^2$,因此

$$A_2 \sim J_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

于是得到 A 的 Jordan 标准形为

$$J = \begin{pmatrix} J_1 & \\ & J_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & 0 & 1 \end{pmatrix}.$$

上面介绍了将方阵 A 化为 Jordan 标准形的方法. 我们知道，若 $A \sim J$，则必存在满秩方阵 P，使得 $P^{-1}AP = J$. 在通常情况下，如果只需找到 Jrodan 标准形，则不必求 P. 但若利用 Jordan 标准形求解微分方程组，就少不了求 P. 下面介绍求 P 的一般方法.

设 P 为 n 阶方阵，J 的某个 Jordan 块 J_i 为 m_i 阶，其主对角线元素为 λ_i，它们在 J 中的行数与列数分别为 $m_1+m_2+\cdots+m_{i-1}+1$，$m_1+m_2+\cdots+m_{i-1}+2$，$\cdots$，$m_1+m_2+\cdots+m_{i-1}+m_i$. 记 P 的各个列向量分别为 $\eta_1, \eta_2, \cdots, \eta_n$，则

$$P = (\eta_1, \cdots, \eta_{m_1+\cdots+m_{i-1}+1}, \cdots, \eta_{m_1+\cdots+m_{i-1}+m_i}, \cdots, \eta_n).$$

为了方便起见，将编号 $m_1+\cdots+m_{i-1}+1$，$m_1+\cdots+m_{i-1}+2$，$\cdots$，$m_1+\cdots+m_{i-1}+m_i$ 记为 $1, 2, \cdots, m_i$，将 λ_i 记为 λ，并由 $P^{-1}AP = J$ 得 $AP = PJ$，从而

$$A(\cdots, \eta_1, \cdots, \eta_{m_i}, \cdots) = (\cdots, \eta_1, \cdots, \eta_{m_i}, \cdots)J.$$

由此得

$$\begin{aligned} &A\eta_1 = \lambda\eta_1, \\ &A\eta_2 = \eta_1 + \lambda\eta_2, \\ &\cdots\cdots \\ &A\eta_{m_i} = \eta_{m_i-1} + \lambda\eta_{m_i}, \end{aligned}$$

即

$$\begin{aligned} &(\lambda I - A)\eta_1 = 0, \\ &(\lambda I - A)\eta_2 = -\eta_2, \end{aligned}$$

$$\cdots\cdots$$

$$(\lambda I - A)\eta_{m_i} = -\eta_{m_i-1}.$$

若记 $H=\lambda I-A$，以上各方程即为

$$H\eta_1 = 0,$$

$$H\eta_2 = -\eta_1,$$

$$\cdots\cdots$$

$$H\eta_{m_i} = -\eta_{m_i-1}.$$

以 $S(H)$ 表示由 H 的列向量张成的线性空间. 因为 $\eta_1 \in S(H)$ 是 $H\eta_2=-\eta_1$ 有解的充要条件，而 $\eta_1\in S(H)$ 的充要条件又是 η_1 与 $S(H)^{\perp}$ 中的一切向量正交，这又等价于与 $S(H)^{\perp}$ 的一个基正交. $S(H)^{\perp}$ 的一个基即为方程 $H^TX=0$ 的一个基础解系. 因此，为了使 $H\eta_2=-\eta_1$ 对 η_2 有解，η_1 应这样求出：

(1) 求出 $H^TX=0$ 的一个基础解系 $p_1,p_2,\cdots,p_s$. 以 $p_1^T,p_2^T,\cdots,p_s^T$ 为行向量构成矩阵 B.

(2) 以 $\begin{pmatrix} H \\ B \end{pmatrix} X=0$ 的一个非零向量作为 η_1.

(3) 同理可得，对 $k=2,3,\cdots,m_i-1$，向量 η_k 为

$$\begin{pmatrix} H \\ B \end{pmatrix} X = \begin{pmatrix} -\eta_{k-1} \\ 0 \end{pmatrix}$$

的一个非零解.

(4) η_{m_i} 为方程 $HX=-\eta_{m_i-1}$ 的非零解.

按上述步骤，可依次求出 $\eta_1,\eta_2,\cdots,\eta_{m_i}$.

例 4 设

$$A=\begin{pmatrix} -4 & 2 & 10 \\ -4 & 3 & 7 \\ -3 & 1 & 7 \end{pmatrix},$$

求 A 的 Jordan 标准形 J，并求 P，使得 $P^{-1}AP=J$.

解 由初等变换可得

$$\lambda I - A = \begin{bmatrix} \lambda+4 & -2 & -10 \\ 4 & \lambda-3 & -7 \\ 3 & -1 & \lambda-7 \end{bmatrix}$$

$$\xrightarrow{r_2 - r_3} \begin{bmatrix} \lambda+4 & -2 & -10 \\ 1 & \lambda-2 & -\lambda \\ 3 & -1 & \lambda-7 \end{bmatrix}$$

$$\xrightarrow[c_3 - 5c_2]{c_1 + 3c_2} \begin{bmatrix} \lambda-2 & -2 & 0 \\ 3\lambda-5 & \lambda-2 & -6\lambda+10 \\ 0 & -1 & \lambda-2 \end{bmatrix}$$

$$\xrightarrow{c_3 + 2c_1} \begin{bmatrix} \lambda-2 & -2 & 2(\lambda-2) \\ 3\lambda-5 & \lambda-2 & 0 \\ 0 & -1 & \lambda-2 \end{bmatrix}$$

$$\xrightarrow{r_1 - 2r_3} \begin{bmatrix} \lambda-2 & 0 & 0 \\ 3\lambda-5 & \lambda-2 & 0 \\ 0 & -1 & \lambda-2 \end{bmatrix}$$

$$\xrightarrow{r_2 + (\lambda-2)r_3} \begin{bmatrix} \lambda-2 & 0 & 0 \\ 3\lambda-5 & 0 & (\lambda-2)^2 \\ 0 & -1 & \lambda-2 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} 1 & \lambda-2 & 0 \\ 0 & (\lambda-2)^2 & 3\lambda-5 \\ 0 & 0 & \lambda-2 \end{bmatrix}$$

$$\xrightarrow{c_2 - (\lambda-2)c_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & (\lambda-2)^2 & 3\lambda-5 \\ 0 & 0 & \lambda-2 \end{bmatrix}$$

$$\xrightarrow{r_2-3r_3}\begin{pmatrix}1 & 0 & 0\\ 0 & (\lambda-2)^2 & 1\\ 0 & 0 & \lambda-2\end{pmatrix}$$

$$\xrightarrow{r_3-(\lambda-2)r_2}\begin{pmatrix}1 & 0 & 0\\ 0 & (\lambda-2)^2 & 0\\ 0 & -(\lambda-2)^3 & 0\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & (\lambda-2)^3\end{pmatrix},$$

所以

$$A\sim J=\begin{pmatrix}2 & 1 & 0\\ 0 & 2 & 1\\ 0 & 0 & 2\end{pmatrix}.$$

令

$$H=2I-A=\begin{pmatrix}6 & -2 & -10\\ 4 & -1 & -7\\ 3 & -1 & -5\end{pmatrix},$$

求解方程组 $H^TX=0$,即

$$\begin{pmatrix}6 & 4 & 3\\ -2 & -1 & -1\\ -10 & -7 & -5\end{pmatrix}X=0,$$

解得 $X=k\begin{pmatrix}1\\ 0\\ -2\end{pmatrix}$,故可取 $B=(1,0,-2)$.

由方程组

$$\begin{pmatrix}H\\ B\end{pmatrix}X=0,$$

即

$$\begin{pmatrix} 6 & -2 & -10 \\ 4 & -1 & -7 \\ 3 & -1 & -5 \\ 1 & 0 & -2 \end{pmatrix} X = 0.$$

解得 $X = k\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$，取 $\eta_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$.

由

$$\begin{pmatrix} H \\ B \end{pmatrix} X = \begin{pmatrix} -\eta_1 \\ 0 \end{pmatrix},$$

即

$$\begin{pmatrix} 6 & -2 & -10 \\ 4 & -1 & -7 \\ 3 & -1 & -5 \\ 1 & 0 & -2 \end{pmatrix} X = \begin{pmatrix} -2 \\ -1 \\ -1 \\ 0 \end{pmatrix},$$

解得 $X = k\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$，取 $\eta_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

最后由

$$HX = -\eta_2,$$

即

$$\begin{pmatrix} 6 & -2 & -10 \\ 4 & -1 & -7 \\ 3 & -1 & -5 \end{pmatrix} X = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix},$$

解得 $X = k\begin{pmatrix} -1 \\ -3 \\ 0 \end{pmatrix}$，取 $\eta_3 = \begin{pmatrix} -1 \\ -3 \\ 0 \end{pmatrix}$.

因此

$$P=(\eta_1,\eta_2,\eta_3)=\begin{pmatrix}2&0&-1\\1&1&-3\\1&0&0\end{pmatrix}.$$

检验得

$$AP=PJ=\begin{pmatrix}4&2&-2\\2&3&-5\\2&1&0\end{pmatrix},$$

即

$$P^{-1}AP=J.$$

故所得的 P 确为所求.

例 5 求解微分方程组

$$\begin{cases}\dfrac{\mathrm{d}x_1}{\mathrm{d}t}=-x_1+x_2,\\ \dfrac{\mathrm{d}x_2}{\mathrm{d}t}=-4x_1+3x_2,\\ \dfrac{\mathrm{d}x_3}{\mathrm{d}t}=x_1+2x_3.\end{cases}$$

解 将方程组写成矩阵形式

$$\frac{\mathrm{d}X}{\mathrm{d}t}=AX,$$

其中

$$X=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix},\quad A=\begin{pmatrix}-1&1&0\\-4&3&0\\1&0&2\end{pmatrix}.$$

$\lambda I-A$ 的初等因子为$(\lambda-2),(\lambda-1)^2$,所以

$$A\sim J=\begin{pmatrix}2&0&0\\0&1&1\\0&0&1\end{pmatrix}.$$

设 $P^{-1}AP=J$,即 $AP=PJ$. 令

$$P=(\eta_1,\eta_2,\eta_3),$$

则

$$\begin{cases} A\eta_1=2\eta_1, \\ A\eta_2=\eta_2, \\ A\eta_3=\eta_2+\eta_3. \end{cases}$$

由 $A\eta_1=2\eta_1$ 即 $(2I-A)\eta_1=0$，得 $\eta_1=(0,0,1)^T$.

对于特征值 1，记

$$H=I-A=\begin{bmatrix} 2 & -1 & 0 \\ 4 & -2 & 0 \\ -1 & 0 & -1 \end{bmatrix},$$

则 $H^TX=0$ 的基础解系为 $(2,-1,0)^T$，取 $B=(2,-1,0)$. 由

$$\begin{pmatrix} H \\ B \end{pmatrix} X=0,$$

得基础解系为 $(-1,-2,1)^T$，取 $\eta_2=(-1,-2,1)^T$. 由

$$HX=-\eta_2,$$

得一个非零解 $(1,1,0)^T$，取 $\eta_3=(1,1,0)^T$. 因此

$$P=(\eta_1,\eta_2,\eta_3)=\begin{bmatrix} 0 & -1 & 1 \\ 0 & -2 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

作线性变换

$$X=PY,$$

其中 $Y=(y_1,y_2,y_3)^T$，代入 $\dfrac{dX}{dt}=AX$，得

$$P\frac{dY}{dt}=APY,$$

即

$$\frac{dY}{dt}=P^{-1}APY=JY,$$

因此

$$
\begin{cases}
\dfrac{\mathrm{d}y_1}{\mathrm{d}t}=2y_1, \\
\dfrac{\mathrm{d}y_2}{\mathrm{d}t}=y_2+y_3, \\
\dfrac{\mathrm{d}y_3}{\mathrm{d}t}=y_3.
\end{cases}
$$

由第一、三个方程解得

$$y_1=k_1\mathrm{e}^{2t},\quad y_3=k_3\mathrm{e}^{t}.$$

代入第二个方程得

$$\frac{\mathrm{d}y_2}{\mathrm{d}t}=y_2+k_3\mathrm{e}^{t},$$

解得

$$y_2=\mathrm{e}^{t}(k_2+k_3t).$$

又由 $X=PY$，得

$$
\begin{cases}
x_1=-y_2+y_3=\mathrm{e}^{t}(k_3-k_2-k_3t), \\
x_2=-2y_2+y_3=\mathrm{e}^{t}(k_3-2k_2-2k_3t), \\
x_3=y_1+y_2=k_1\mathrm{e}^{2t}+\mathrm{e}^{t}(k_2+k_3t).
\end{cases}
$$

其中 k_1,k_2,k_3 为任意常数.

§2.4 Cayley-Hamilton 定理　最小多项式

1. Cayley-Hamilton 定理

先给出 λ-矩阵的多项式写法. 例如，二阶三次 λ-矩阵

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

可写为

$$A(\lambda)=\begin{pmatrix} \lambda^3+0\lambda^2+\lambda+1 & 0\lambda^3+\lambda^2-\lambda+1 \\ 0\lambda^3+0\lambda^2+\lambda-1 & \lambda^3+\lambda^2+0\lambda+2 \end{pmatrix}$$

$$=\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\lambda^3+\begin{pmatrix}0 & 1\\0 & 1\end{pmatrix}\lambda^2+\begin{pmatrix}1 & -1\\0 & 1\end{pmatrix}\lambda+\begin{pmatrix}1 & 1\\-1 & 2\end{pmatrix}.$$

一般地，一个次数不超过 m 的 n 阶 λ-矩阵 $B(\lambda)$ 总可写为

$$B(\lambda)=B_0\lambda^m+B_1\lambda^{m-1}+\cdots+B_{m-1}\lambda+B_m,\qquad(2.8)$$

其中 $B_i(i=0,1,\cdots,m)$ 为 n 阶常数矩阵，且称(2.8)式为 λ-矩阵 $B(\lambda)$ 的**多项式写法**.

同普通的关于某个变量的多项式一样，形如

$$a_0A^m+a_1A^{m-1}+\cdots+a_{m-1}A+a_mI$$

的式子称为**方阵 A 的多项式**，其中 a_i 为常数，m 为正整数.

例如，设 λ 的多项式 $\varphi(\lambda)$ 为

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

A 为 n 阶方阵，则

$$\varphi(A)=a_0A^m+a_1A^{m-1}+\cdots+a_{m-1}A+a_mI$$

也是 n 阶方阵.

下面的定理表示矩阵 A 与其特征多项式 $f(\lambda)=\det(\lambda I-A)$ 之间的一个重要关系.

定理 1（Cayley-Hamilton 定理） 设 $f(\lambda)$ 为矩阵 A 的特征多项式，则

$$f(A)=0.\qquad(2.9)$$

证 设 B 为 A 的特征矩阵 $(\lambda I-A)$ 的伴随矩阵，由伴随矩阵的定义知，伴随矩阵 B 的元素都是 $(\lambda I-A)$ 的各元素的代数余子式，即为 λ 的次数不超过 $n-1$ 的多项式. 所以 B 是 $n-1$ 次 λ-矩阵，记为

$$B=B_0\lambda^{n-1}+B_1\lambda^{n-2}+\cdots+B_{n-2}\lambda+B_{n-1},$$

并且有

$$B(\lambda I-A)=(\lambda I-A)B=f(\lambda)I,$$

即

$$(B_0\lambda^{n-1}+B_1\lambda^{n-2}+\cdots+B_{n-2}\lambda+B_{n-1})(\lambda I-A)$$
$$=(\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n)I,$$

比较两边系数，有

$$
\begin{aligned}
&-B_{n-1}A=a_nI,\\
&-B_{n-2}A+B_{n-1}=a_{n-1}I,\\
&-B_{n-3}A+B_{n-2}=a_{n-2}I,\\
&\cdots\cdots\\
&-B_0A+B_1=a_1I,\\
&B_0=I.
\end{aligned}
$$

依次用 $I,A,A^2,\cdots,A^n$ 右乘以上各式，再两边相加得

$$0=a_nI+a_{n-1}A+\cdots+a_1A^{n-1}+A^n,$$

即

$$f(A)=0.$$

例 1 设

$$A=\begin{pmatrix}1 & 0 & -1\\ 0 & \omega & \sqrt{2}\mathrm{i}\\ 0 & 0 & \omega^2\end{pmatrix},$$

其中 $\omega=\dfrac{-1+\sqrt{3}\mathrm{i}}{2}$，计算 A^{100}.

解 易见矩阵 A 的特征多项式为

$$(\lambda-1)(\lambda-\omega)(\lambda-\omega^2)=\lambda^3-1,$$

由定理 1，有 $A^3-I=0$，即 $A^3=I$，所以

$$A^{100}=(A^3)^{33}\cdot A=A.$$

例 2 设

$$A=\begin{pmatrix}1 & 0 & 0\\ 1 & 0 & 1\\ 0 & 1 & 0\end{pmatrix},$$

求证：

$$A^n=A^{n-2}+A^2-I \qquad (n\geqslant 3).$$

证 易知，矩阵 A 的特征多项式为

$$f(\lambda)=(\lambda^2-1)(\lambda-1),$$

令

$$\begin{aligned}
g(\lambda)&=\lambda^n-\lambda^{n-2}-\lambda^2+1\\
&=\lambda^{n-2}(\lambda^2-1)-(\lambda^2-1)\\
&=(\lambda^2-1)(\lambda^{n-2}-1)\\
&=(\lambda^2-1)(\lambda-1)(\lambda^{n-3}+\lambda^{n-4}+\cdots+\lambda+1)\qquad(\lambda\geqslant 3),
\end{aligned}$$

由定理 1 可知

$$g(A)=0,$$

因此

$$A^n=A^{n-2}+A^2-I\qquad(n\geqslant 3).$$

2. 最小多项式

Cayley-Hamilton 定理表明，对于任何方阵 A，总可以找到变量 λ 的多项式 $\varphi(\lambda)$，使得 $\varphi(\lambda)=0$.

定义 1 凡使 $\varphi(\lambda)=0$ 的 λ 的多项式 $\varphi(\lambda)$ 称为矩阵 A 的**零化多项式**(一般取系数为 1).

由定义 1 可知，Cayley-Hamilton 定理表明，A 的特征多项式是方阵 A 的一个零化多项式. 显然，A 的零化多项式不是唯一的. 因为零化多项式乘以任意一个多项式后仍为零化多项式. 而且，A 的特征多项式也不一定是 A 的次数最低的零化多项式.

例如，设

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

它的特征多项式为 $f(\lambda)=(\lambda-2)^2(\lambda-3)$，故 $f(A)=0$. 设 $m(\lambda)=(\lambda-2)(\lambda-3)$，容易验证 $m(A)=0$.

定义 2 矩阵 A 的次数最低的首一零化多项式称为 A 的**最小多项式**，记为 $m(\lambda)$.

定理 2 多项式 $\varphi(\lambda)$ 是矩阵 A 的零化多项式的充要条件是

$m(\lambda)\mid\varphi(\lambda)$. 特别地，有 $m(\lambda)\mid f(\lambda)$，即 A 的最小的多项式为其特征多项式的因式.

证 设 $\varphi(\lambda)$是 A 的零化多项式，其次数自然不比 $m(\lambda)$的次数低，以 $m(\lambda)$除 $\varphi(\lambda)$，得

$$\varphi(\lambda)=q(\lambda)m(\lambda)+r(\lambda),$$

这里，$r(\lambda)=0$ 或 $r(\lambda)$的次数低于 $m(\lambda)$的次数. 于是

$$\varphi(A)=q(A)m(A)+r(A).$$

由 $\varphi(A)=0$ 和 $m(A)=0$ 得 $r(A)=0$. 若 $r(A)\neq 0$，则 $r(\lambda)$为 A 的次数低于 $m(\lambda)$的零化多项式，这与 $m(\lambda)$为最小多项式矛盾. 因此，$r(\lambda)\equiv 0$，即 $m(\lambda)\mid\varphi(\lambda)$.

反之，若 $m(\lambda)\mid\varphi(\lambda)$，则 $\varphi(\lambda)=q(\lambda)m(\lambda)$，于是

$$\varphi(A)=q(A)m(A)=0,$$

即 $\varphi(\lambda)$为 A 的零化多项式.

定理 3 相似矩阵有相同的最小的多项式.

证 设 $A\sim B$，则存在可逆矩阵 P，使得

$$A=P^{-1}BP,$$

若 $f(B)=0$，则有

$$f(A)=f(P^{-1}BP)=P^{-1}f(B)P=0.$$

注 定理 3 的逆定理不成立，即最小多项式相同的两矩阵不一定相似. 例如，设

$$A=\begin{bmatrix}2&&\\&3&\\&&3\end{bmatrix},\quad B=\begin{bmatrix}2&&\\&2&\\&&3\end{bmatrix},$$

它们的特征多项式分别为$(\lambda-2)(\lambda-3)^2$ 和$(\lambda-2)^2(\lambda-3)$，由它们的特征多项式不同知它们不相似. 但是它们有相同的最小多项式$(\lambda-2)(\lambda-3)$.

由定理 2 可知，$m(\lambda)\mid f(\lambda)$，那么商$\dfrac{f(\lambda)}{m(\lambda)}$等于什么呢？下面的

4 定理回答了这个问题.

定理 4 n 阶矩阵 A 的最小多项式等于它的特征矩阵 $(\lambda I-A)$ 中的第 n 个不变因子 $d_n(\lambda)$，因而

$$\frac{f(\lambda)}{m(\lambda)}=\frac{\det(\lambda I-A)}{m(\lambda)}=\frac{\det(\lambda I-A)}{d_n(\lambda)}=D_{n-1}(\lambda).$$

证明参阅张远达《线性代数原理》第 215 页.

推论 1 方阵 A 的特征多项式的根必是其最小多项式的根.

证 将 A 的特征矩阵 $\lambda I-A$ 化为 Smith 标准形

$$\lambda I-A\sim\begin{pmatrix} d_1(\lambda) & & & \\ & d_2(\lambda) & & \\ & & \ddots & \\ & & & d_n(\lambda)\end{pmatrix}.$$

于是

$$f(\lambda)=\det(\lambda I-A)=d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda).$$

若 λ_i 为 $f(\lambda)$ 的一个根，则由 $f(\lambda_i)=0$ 可知，λ_i 必为某个 $d_j(\lambda)$ 的根. 又由 $d_j(\lambda)\mid d_n(\lambda)$ 可知，λ_i 为 $d_n(\lambda)$ 的根，即 λ_i 为 $m(\lambda)$ 的根.

由推论 1 知，若 $f(\lambda)$ 分解为不同的一次因式的幂积，如

$$f(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_t)^{m_t},$$

其中 $\sum\limits_{i=1}^{t}m_i=n$，且每个 $m_i>0$，当 $i\neq j$ 时，$\lambda_i\neq\lambda_j$，因此

$$m(\lambda)=(\lambda-\lambda_1)^{l_1}(\lambda-\lambda_2)^{l_2}\cdots(\lambda-\lambda_t)^{l_t},$$

其中 $1\leqslant l_i\leqslant m_i$，$i=1,2,\cdots,t$. 特别地，当每个 $m_i=1$ 时，有 $t=n$，且每个 $l_i=1$，因而这时 $m(\lambda)=f(\lambda)$. 这就是下面的推论：

推论 2 若矩阵 A 的特征值互异，则它的最小多项式就是特征多项式.

例 3 求下列方阵的特征多项式与最小多项式：

(1) $\begin{pmatrix} 3 & 1 & \\ & 3 & \\ & & 3\end{pmatrix}$;

(2) $\begin{pmatrix} a & c_1 & & & \\ & a & c_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & c_{n-1} \\ & & & & a \end{pmatrix}$;

(3) $\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} \\ 0 & 1 & \cdots & 0 & -a_{n-2} \\ \cdots & & & & \\ 0 & 0 & \cdots & 0 & -a_2 \\ 0 & 0 & \cdots & 1 & -a_1 \end{pmatrix}$.

解 (1) 特征多项式为

$$f(\lambda)=\det\begin{pmatrix} \lambda-3 & -1 & \\ & \lambda-3 & \\ & & \lambda-3 \end{pmatrix}=(\lambda-3)^3,$$

最小多项式可能是

$$(\lambda-3),\quad (\lambda-3)^2,\quad (\lambda-3)^3$$

中的一个. 通过计算可知

$$A-3I=\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\neq 0,$$

$$(A-3I)^2=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}=0,$$

所以,最小多项式为

$$m(\lambda)=(\lambda-3)^2.$$

或者用初等变换将 $\lambda I-A$ 化为 Smith 标准形:

$$\lambda I-A \longrightarrow \begin{pmatrix} \lambda-3 & -1 & 0 \\ 0 & \lambda-3 & 0 \\ 0 & 0 & \lambda-3 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} -1 & \lambda-3 & 0 \\ \lambda-3 & 0 & 0 \\ 0 & 0 & \lambda-3 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} -1 & \lambda-3 & 0 \\ 0 & (\lambda-3)^2 & 0 \\ 0 & 0 & \lambda-3 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} -1 & \lambda-3 & 0 \\ 0 & (\lambda-3)^2 & 0 \\ 0 & 0 & \lambda-3 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda-3 & 0 \\ 0 & 0 & (\lambda-3)^2 \end{pmatrix},$$

故最小多项式为

$$m(\lambda)=d_3(\lambda)=(\lambda-3)^2.$$

(2) 由§2.2节的例2知,特征矩阵 $\lambda I-A$ 的第 n 个不变因子为 $d_n(\lambda)=(\lambda-a)^n$,故

$$m(\lambda)=(\lambda-a)^n.$$

(3) 由§2.2节的例3知,$d_n(\lambda)=f(\lambda)$,故 $m(\lambda)=f(\lambda)$,即友矩阵的最小多项式就是它的特征多项式.

习 题 二

1. 化下列矩阵为Smith标准形:

(1) $\begin{pmatrix} 1-\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ 1+\lambda^2 & \lambda^2 & -\lambda^2 \end{pmatrix}$;

(2) $\begin{pmatrix} 0 & 0 & 0 & \lambda^2 \\ 0 & 0 & \lambda^2-\lambda & 0 \\ 0 & (\lambda-1)^2 & 0 & 0 \\ \lambda^2-\lambda & 0 & 0 & 0 \end{pmatrix}$;

(3) $\begin{pmatrix} 3\lambda^2+2\lambda-3 & 2\lambda-1 & \lambda^2+2\lambda-3 \\ 4\lambda^2+3\lambda-5 & 3\lambda-2 & \lambda^2+3\lambda-4 \\ \lambda^2+\lambda-4 & \lambda-2 & \lambda-1 \end{pmatrix}$;

(4) $\begin{pmatrix} 2\lambda & 3 & 0 & 1 & \lambda \\ 4\lambda & 3\lambda-6 & 0 & \lambda+2 & 2\lambda \\ 0 & 6\lambda & \lambda & 2\lambda & 0 \\ \lambda-1 & 0 & \lambda-1 & 0 & 0 \\ 3\lambda-3 & 1-\lambda & 2\lambda-2 & 0 & 0 \end{pmatrix}$.

2. 求下列 λ 矩阵的不变因子：

(1) $\begin{pmatrix} \lambda-2 & -1 & 0 \\ 0 & \lambda-2 & -1 \\ 0 & 0 & \lambda-2 \end{pmatrix}$;

(2) $\begin{pmatrix} \lambda+\alpha & \beta & 1 & 0 \\ -\beta & \lambda+\alpha & 0 & 1 \\ 0 & 0 & \lambda+\alpha & \beta \\ 0 & 0 & -\beta & \lambda+\alpha \end{pmatrix}$;

(3) $\begin{pmatrix} \lambda & -1 & 0 & 0 \\ 0 & \lambda & -1 & 0 \\ 0 & 0 & \lambda & -1 \\ 5 & 4 & 3 & \lambda+2 \end{pmatrix}$;

(4) $\begin{pmatrix} 0 & 0 & 1 & \lambda+2 \\ 0 & 1 & \lambda+2 & 0 \\ 1 & \lambda+2 & 0 & 0 \\ \lambda+2 & 0 & 0 & 0 \end{pmatrix}$.

3. 求下列λ-矩阵的初等因子：

(1) $\begin{pmatrix} \lambda^3+2 & \lambda^3+1 \\ 2\lambda^3-\lambda^2-\lambda+3 & 2\lambda^3-\lambda^2-\lambda+2 \end{pmatrix}$;

(2) $\begin{pmatrix} \lambda^3-2\lambda^2+2\lambda-1 & \lambda^2-2\lambda-1 \\ 2\lambda^3-2\lambda^2+\lambda-1 & 2\lambda^2-2\lambda \end{pmatrix}$.

4. 求下列矩阵的 Jordan 标准形：

(1) $\begin{pmatrix} 13 & 16 & 16 \\ -5 & -7 & -6 \\ -6 & -8 & -7 \end{pmatrix}$;

(2) $\begin{pmatrix} 4 & 5 & -2 \\ -2 & -2 & 1 \\ -1 & -1 & 1 \end{pmatrix}$;

(3) $\begin{pmatrix} 3 & 7 & -3 \\ -2 & -5 & 2 \\ -4 & -10 & 3 \end{pmatrix}$;

(4) $\begin{pmatrix} 1 & 1 & -1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{pmatrix}$;

(5) $\begin{pmatrix} 0 & 3 & 3 \\ -1 & 8 & 6 \\ 2 & -14 & -10 \end{pmatrix}$;

(6) $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

5. 设矩阵

$$A=\begin{pmatrix} 1 & 4 & 2 \\ 0 & -3 & 4 \\ 0 & 4 & 3 \end{pmatrix},$$

求 A^5.

6. 设矩阵

$$A=\begin{pmatrix} 2 & -1 & -1 \\ 2 & -1 & -2 \\ -1 & 1 & 2 \end{pmatrix},$$

求 A 的 Jordan 标准形 J,并求相似变换矩阵 P,使得 $P^{-1}AP=J$.

7. 设矩阵

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

试计算 $2A^8-3A^5+A^4+A^2-4I$.

8. 证明:任意可逆矩阵 A 的逆矩阵 A^{-1}可以表示为 A 的多项式.

9. 设矩阵

$$A=\begin{pmatrix}2 & -1\\ 1 & 3\end{pmatrix},$$

试计算$(A^4-5A^3+6A^2+6A-8I)^{-1}$.

10. 已知3阶矩阵A的三个特征值为1,-1,2,试将A^{2n}表示为A的二次式.

11. 求下列矩阵的最小多项式:

(1) $\begin{bmatrix}3 & 1 & -1\\ 0 & 2 & 0\\ 1 & 1 & 1\end{bmatrix}$;

(2) $\begin{bmatrix}4 & -2 & 2\\ -5 & 7 & -5\\ -6 & 7 & -4\end{bmatrix}$;

(3) n阶单位阵I_n;

(4) n阶方阵A,其元素均为1;

(5) $B=\begin{bmatrix}a_0 & a_1 & a_2 & a_3\\ -a_1 & a_0 & -a_3 & a_2\\ -a_2 & a_3 & a_0 & -a_1\\ -a_3 & -a_2 & a_1 & a_0\end{bmatrix}$.

第三章 矩阵分析及矩阵函数

前两章主要研究了矩阵的代数运算性质，本章将研究矩阵的极限、微分、积分等分析性质及理论. 这些理论已广泛应用于各个领域，而且已成为应用数学中最基础的内容之一.

§3.1 基本概念

1. 矩阵序列的极限

按照一定的顺序，将可数个 n 阶方阵排成一列：

$$A_1,\ A_2,\ \cdots,\ A_m,\ \cdots, \tag{3.1}$$

称这列有次序的矩阵为**矩阵序列**，称 A_m 为**矩阵序列的一般项**.

定义 1 设 $A_m=(a_{ij}^{(m)})_{n\times n}$，$m=1,2,\cdots$，如果对任意的 i,j，有 $\lim\limits_{m\to\infty}a_{ij}^{(m)}=a_{ij}$，$i,j=1,2,\cdots,n$，则称矩阵序列 $\{A_m\}$**收敛**于矩阵 $A=(a_{ij})_{n\times n}$，记为 $\lim\limits_{m\to\infty}A_m=A$；否则，称矩阵序列 $\{A_m\}$ 为**发散的**.

性质 1 设 $\{A_m\}$ 与 $\{B_m\}$ 为两个 n 阶矩阵序列，如果

$$\lim_{m\to\infty}A_m=A,\quad \lim_{m\to\infty}B_m=B,$$

则

$$\lim_{m\to\infty}(kA_m+lB_m)=kA+lB,$$

$$\lim_{m\to\infty}(A_mB_m)=AB,$$

其中，k,l 为任意的常数.

性质 2 如果 $\lim\limits_{m\to\infty}A_m=A$，且 A^{-1}，A_m^{-1} 都存在，$m=1,2,\cdots$，则 $\lim\limits_{m\to\infty}A_m^{-1}$ 存在，且

$$\lim_{m\to\infty} A_m^{-1} = A^{-1}.$$

2. 矩阵级数

定义 2 如果给定一个 n 阶矩阵序列

$$A_1,\ A_2,\ \cdots,\ A_m,\ \cdots,$$

则由这 n 阶矩阵序列构成的表达式

$$A_1 + A_2 + \cdots + A_m + \cdots \tag{3.2}$$

叫做 **n 阶矩阵级数**，记为 $\sum_{m=1}^{\infty} A_m$，即

$$\sum_{m=1}^{\infty} A_m = A_1 + A_2 + \cdots + A_m + \cdots,$$

其中第 m 项 A_m 叫做**级数的一般项**.

作级数(3.2)式的前 m 项的和

$$S_m = A_1 + A_2 + \cdots + A_m,$$

称 S_m 为级数(3.2)式的**部分和**；并称矩阵序列

$$S_1,\ S_2,\ \cdots,\ S_m,\ \cdots$$

为级数(3.2)式的**部分和序列**.

定义 3 如果级数(3.2)式的部分和序列的极限存在，设为 S，则称 S 为级数(3.2)式的**和**，并称级数是**收敛的**，记为 $S = \sum_{m=1}^{\infty} A_m$；否则，称级数(3.2)式是**发散的**.

显然，$S = \sum_{m=1}^{\infty} A_m$ 收敛的充要条件为对任意的 $i,j=1,2,\cdots,n$，$\sum_{m=1}^{\infty} a_{ij}^{(m)}$ 都收敛.

3. 函数矩阵及其极限

定义 4 如果矩阵 A 的每一个元素 a_{ij} 都是变量 t 的函数，则称矩阵

$$A(t)=\begin{pmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \\ \cdots & & & \\ a_{m1}(t) & a_{m2}(t) & \cdots & a_{mn}(t) \end{pmatrix}$$

为函数矩阵.

定义 5 如果对任意的 $i,j,1\leqslant i\leqslant m,1\leqslant j\leqslant n$,都有 $\lim\limits_{t\to t_0} a_{ij}(t)=a_{ij}$,则称矩阵 $A(t)=(a_{ij}(t))_{m\times n}$ 在 $t\to t_0$ 时**极限为** $A=(a_{ij})_{m\times n}$.

性质 3 设 $\lim\limits_{t\to t_0}A(t)=A,\lim\limits_{t\to t_0}B(t)=B$.

(1) 如果 $A(t),B(t)$ 都是 $m\times n$ 阶矩阵,则

$$\lim_{t\to t_0}[A(t)+B(t)]=A+B=\lim_{t\to t_0}A(t)+\lim_{t\to t_0}B(t);$$

(2) 如果 $A(t)$ 与 $B(t)$ 分别为 $m\times n$ 阶及 $n\times r$ 阶矩阵,则

$$\lim_{t\to t_0}[A(t)B(t)]=AB=\lim_{t\to t_0}A(t)\lim_{t\to t_0}B(t);$$

(3) 设 k 为常数,则

$$\lim_{t\to t_0}[kA(t)]=kA=k\lim_{t\to t_0}A(t).$$

定义 6 设函数矩阵 $A(t)$ 中所有元素 $a_{ij}(t)$ 在 t_0 处连续,则称 $A(t)$ 在 $\mathbf{t_0}$ **处连续**. 如果所有元素 $a_{ij}(t)$ 在 (a,b) 内每一点都连续,则称 $A(t)$ 在 $(\mathbf{a},\mathbf{b})$ **内连续**. 如果 $A(t)$ 在 (a,b) 内连续,并且所有的 $a_{ij}(t)$ 在 a 点右连续,在 b 点左连续,则称 $A(t)$ 在 $[\boldsymbol{a},\boldsymbol{b}]$ **上连续**.

对多个自变量的情况,可类似地定义极限、连续等概念.

§ 3.2 函数矩阵的微分和积分

1. 函数矩阵的微分和积分

定义 1 设函数矩阵 $A(t)$ 中所有元素 $a_{ij}(t)$ 都在 t_0 点或在某区间内可微,则称函数矩阵 $A(t)$ 在 t_0 点或某区间内是**可微的**. 若 $A(t)$ 可微,其导数定义如下:

$$A'(t)=(a_{ij}'(t)),$$

同样,$A(t)$的高阶导数定义为

$$A''(t)=(A'(t))',\ \cdots,\ A^{(n)}(t)=[A^{(n-1)}(t)]'.$$

性质 1 设函数矩阵 $A(t)$,$B(t)$都可微.

(1) 若 k 为常数,则

$$[kA(t)]'=kA'(t);$$

(2) 若 $A(t)$与 $B(t)$可以相加,则

$$[A(t)+B(t)]'=A'(t)+B'(t);$$

(3) 若 $A(t)$与 $B(t)$可以相乘,则

$$[A(t)B(t)]'=A'(t)B(t)+A(t)B'(t);$$

特别地,若 $A(t)$或 $B(t)$为常数矩阵 A 或 B,则有

$$[AB(t)]'=AB'(t),\quad [A(t)B]'=A'(t)B.$$

注 由于矩阵乘法不满足交换律,故不能颠倒性质 1 的(3)中乘积因子之间的顺序. 例如:

$$[A^2(t)]'=A'(t)A(t)+A(t)A'(t)\neq 2A(t)A'(t).$$

性质 2 设 $A(u)$是可微的函数矩阵,$u=f(t)$是一元可微函数,则

$$[A(f(t))]'_t=A_u'(u)f'(t).$$

性质 3 若 $A(t)$及其逆矩阵 $A^{-1}(t)$都是可微的,则

$$[A^{-1}(t)]'=-A^{-1}(t)A'(t)A^{-1}(t).$$

证 因为 $A(t)A^{-1}(t)=I$,两边求导数得

$$A(t)[A^{-1}(t)]'+A'(t)A^{-1}(t)=0,$$

故

$$[A^{-1}(t)]'=-A^{-1}(t)A'(t)A^{-1}(t).$$

对多元函数,也有类似于上述公式的偏导数公式.

例 1 设向量 $X=(x_1(t),x_2(t),\cdots,x_n(t))^T$ 及对称矩阵 $A(t)=(a_{ij}(t))_{n\times n}$都是可微的,求二次型 X^TAX 的导数.

解 $[X^TAX]'=(X^T)'(AX)+X^T(AX)'$

$$=(X^T)'AX+X^TA'X+X^TAX'.$$

由于$(X^T)'AX$是一阶矩阵，因此其转置矩阵等于自身，即

$$(X^T)'AX=[(X^T)'AX]^T=X^TA^TX'=X^TAX',$$

因此

$$(X^TAX)'=X^TA'X+2X^TAX'.$$

特别地，当A为常值对称矩阵时，由于$A'=0$，故

$$(X^TAX)'=2X^TAX'.$$

例 2 设函数矩阵

$$A(t)=\begin{pmatrix}1 & t^2\\ t & 0\end{pmatrix},$$

试计算：

(1) $A'(t)$，$A''(t)$，$A'''(t)$；

(2) $\dfrac{\mathrm{d}}{\mathrm{d}t}|A(t)|$；

(3) $[A^{-1}(t)]'$.

解 (1) $A'(t)=\begin{pmatrix}0 & 2t\\ 1 & 0\end{pmatrix}$,

$$A''(t)=\begin{pmatrix}0 & 2\\ 0 & 0\end{pmatrix},$$

$$A'''(t)=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}.$$

(2) 因为$|A(t)|=-t^3$，故$\dfrac{\mathrm{d}}{\mathrm{d}t}|A(t)|=-3t^2$.

(3) 先求$A(t)$的逆矩阵.

$$\begin{aligned}A^{-1}(t)&=\frac{1}{|A(t)|}A^*(t)\\&=-\frac{1}{t^3}\begin{pmatrix}0 & -t^2\\ -t & 1\end{pmatrix}\\&=\begin{pmatrix}0 & \dfrac{1}{t}\\ \dfrac{1}{t^2} & -\dfrac{1}{t^3}\end{pmatrix},\end{aligned}$$

再求$[A^{-1}(t)]'$.

$$[A^{-1}(t)]'=\begin{pmatrix} 0 & -\dfrac{1}{t^2} \\ -\dfrac{2}{t^3} & \dfrac{3}{t^4} \end{pmatrix}.$$

注 若按公式$[A^{-1}(t)]'=-A^{-1}(t)A'(t)A^{-1}(t)$来直接求解例2中的(3),则计算较为复杂.

定义2 设函数矩阵$A(t)=(a_{ij}(t))_{m\times n}$中的每个元素$a_{ij}(t)$在$[a,b]$上是可积的,则称$A(t)$在$[a,b]$上**可积分**,且定义

$$\int_a^b A(t)\mathrm{d}t=\left(\int_a^b a_{ij}(t)\mathrm{d}t\right)_{m\times n}.$$

例3 设函数矩阵

$$A(t)=\begin{pmatrix} \sin t & -\cos t \\ \cos t & \sin t \end{pmatrix},$$

求$\int_0^t A(t)\mathrm{d}t$和$\left(\int_0^{t^2} A(t)\mathrm{d}t\right)'$.

解

$$\int_0^t A(t)\mathrm{d}t=\begin{pmatrix} \int_0^t \sin t\mathrm{d}t & \int_0^t(-\cos t)\mathrm{d}t \\ \int_0^t \cos t\mathrm{d}t & \int_0^t \sin t\mathrm{d}t \end{pmatrix}$$

$$=\begin{pmatrix} 1-\cos t & -\sin t \\ \sin t & 1-\cos t \end{pmatrix},$$

$$\left(\int_0^{t^2} A(t)\mathrm{d}t\right)'=A(t^2)\cdot 2t$$

$$=2t\begin{pmatrix} \sin t^2 & -\cos t^2 \\ \cos t^2 & \sin t^2 \end{pmatrix}.$$

2. 数量函数关于矩阵的微分

在场论中,我们对数量函数$u(x,y,z)$定义梯度如下:

$$\operatorname{grad} u=\nabla u=\left(\frac{\partial u}{\partial x},\frac{\partial u}{\partial y},\frac{\partial u}{\partial z}\right),$$

这可以理解为数量函数 $u(x,y,z)$ 对向量 (x,y,z) 的导数. 下面我们将这一概念推广为数量函数对矩阵的导数.

定义 3 设 $y=f(X)=f(x_1,x_2,\cdots,x_n)$ 对 $x_1,x_2,\cdots,x_n$ 有偏导数，定义 $y=f(X)$ 对向量 $X=(x_1,x_2\cdots,x_n)^T$ 的**导数**为

$$\frac{\mathrm{d}f}{\mathrm{d}X}=\left(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\cdots,\frac{\partial f}{\partial x_n}\right)^T=\operatorname{grad}f,$$

而数量函数 $y=f(X)$ 对向量 $X^T=(x_1,x_2,\cdots,x_n)$ 的导数定义为

$$\frac{\mathrm{d}f}{\mathrm{d}X^T}=\left(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\cdots,\frac{\partial f}{\partial x_n}\right).$$

显然

$$\left(\frac{\mathrm{d}f}{\mathrm{d}X^T}\right)^T=\frac{\mathrm{d}f}{\mathrm{d}X}.$$

一般地，若 $y=f(X)=f(x_{11},x_{12},\cdots,x_{1n},x_{21},x_{22},\cdots,x_{2n},\cdots,x_{m1},x_{m2},\cdots,x_{mn})$ 对每个 x_{ij} 有偏导数，则定义数量函数 $y=f(X)$ 对矩阵 $X=(x_{ij})_{m\times n}$ 的导数为

$$\frac{\mathrm{d}f}{\mathrm{d}X}=\begin{pmatrix}\frac{\partial f}{\partial x_{11}} & \frac{\partial f}{\partial x_{12}} & \cdots & \frac{\partial f}{\partial x_{1n}}\\ \frac{\partial f}{\partial x_{21}} & \frac{\partial f}{\partial x_{22}} & \cdots & \frac{\partial f}{\partial x_{2n}}\\ \cdots & & & \\ \frac{\partial f}{\partial x_{m1}} & \frac{\partial f}{\partial x_{m2}} & \cdots & \frac{\partial f}{\partial x_{mn}}\end{pmatrix}.$$

按定义容易证明下面性质：

性质 4 设 $f(X),g(X)$ 为矩阵 X 的数量函数，如果 $\frac{\mathrm{d}f}{\mathrm{d}X}$ 及 $\frac{\mathrm{d}g}{\mathrm{d}X}$ 存在，则

(1) $\frac{\mathrm{d}}{\mathrm{d}X}(f(X)+g(X))=\frac{\mathrm{d}f(X)}{\mathrm{d}X}+\frac{\mathrm{d}g(X)}{\mathrm{d}X}$;

(2) $\frac{\mathrm{d}}{\mathrm{d}X}(f(X)g(X))=f(X)\frac{\mathrm{d}g(X)}{\mathrm{d}X}+g(X)\frac{\mathrm{d}f(X)}{\mathrm{d}X}$.

例 4 设矩阵及其数量函数分别为

$$X=\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdots & & & \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix},$$

$$\begin{aligned} y = f(X) &= x_{11}^2 + x_{12}^2 + \cdots + x_{1n}^2 + x_{21}^2 + x_{22}^2 + \cdots \\ &\quad + x_{2n}^2 + \cdots + x_{m1}^2 + x_{m2}^2 + \cdots + x_{mn}^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}^2, \end{aligned}$$

求$\dfrac{\mathrm{d}y}{\mathrm{d}X}$.

解 由于

$$\frac{\partial f}{\partial x_{ij}}=2x_{ij} \qquad (i=1,2,\cdots,m;j=1,2,\cdots n),$$

则

$$\frac{\mathrm{d}y}{\mathrm{d}X}=\left(\frac{\partial f}{\partial x_{ij}}\right)_{m\times n}=(2x_{ij})_{m\times n}=2X.$$

设 X 同例 4，由于

$$tr(XX^T) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}^2 = f(X),$$

由例 4 可得到关于迹的导数公式.

性质 5 设 $X=(x_{ij})_{m\times n}$，$B=(b_{ij})_{n\times m}$，$A=(a_{ij})_{m\times m}$，则

(1) $\dfrac{\mathrm{d}}{\mathrm{d}X}tr(XX^T)=2X$；

(2) $\dfrac{\mathrm{d}}{\mathrm{d}X}tr(BX)=\dfrac{\mathrm{d}}{\mathrm{d}X}tr(X^TB^T)=B^T$；

(3) $\dfrac{\mathrm{d}}{\mathrm{d}X}tr(X^TAX)=(A+A^T)X$.

例 5 设二次型

$$y = f(X) = X^TAX = \sum_{i,j=1}^{n} a_{ij}x_ix_j,$$

其中，$X=(x_1,x_2,\cdots,x_n)^T$，$A=(a_{ij})_{n\times n}$，且 $A^T=A$，求$\dfrac{\mathrm{d}y}{\mathrm{d}X}$.

解 因为

$$\frac{\mathrm{d}y}{\mathrm{d}X}=\left(\frac{\partial y}{\partial x_1},\frac{\partial y}{\partial x_2},\cdots,\frac{\partial y}{\partial x_n}\right)^T,$$

故只要求出$\dfrac{\partial y}{\partial x_k}$，$k=1,2,\cdots,n$.

由于

$$\begin{aligned}y&=\sum_{i=1}^n\sum_{j=1}^n a_{ij}x_ix_j\\&=\sum_{i\neq k}\sum_{j\neq k}a_{ij}x_ix_j+2\sum_{i\neq k}a_{ik}x_ix_k+a_{kk}x_k^2,\end{aligned}$$

故

$$\begin{aligned}\frac{\partial y}{\partial x_k}&=2\sum_{i\neq k}a_{ik}x_i+2a_{kk}x_k\\&=2\sum_{i=1}^n a_{ik}x_i\\&=2\sum_{j=1}^n a_{kj}x_j,\end{aligned}$$

从而

$$\begin{aligned}\frac{\mathrm{d}y}{\mathrm{d}X}&=\left(2\sum_{j=1}^n a_{1j}x_j,2\sum_{j=1}^n a_{2j}x_j,\cdots,2\sum_{j=1}^n a_{nj}x_j\right)^T\\&=2AX.\end{aligned}$$

3. 向量函数对向量的微分

设

$$X=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\quad a(X)=\begin{pmatrix}a_1(X)\\a_2(X)\\\vdots\\a_n(X)\end{pmatrix},$$

则

$$X^T=(x_1,x_2,\cdots,x_n),$$

$$a^T(X)=(a_1(X),a_2(X),\cdots,a_m(X)).$$

定义 4 设 $a_1(X),a_2(X),\cdots,a_m(X)$ 对 $x_i(i=1,2,\cdots,n)$ 的偏导数都存在，定义向量函数 $a^T(X)$ 对向量 X 的**导数**为一个 $n\times m$ 阶矩阵，它的第 i 列向量就是 $a_i(X)$ 对向量 X 的导数，即

$$\frac{\mathrm{d}a^T(X)}{\mathrm{d}X}=\begin{pmatrix} \frac{\partial a_1(X)}{\partial x_1} & \frac{\partial a_2(X)}{\partial x_1} & \cdots & \frac{\partial a_m(X)}{\partial x_1} \\ \frac{\partial a_1(X)}{\partial x_2} & \frac{\partial a_2(X)}{\partial x_2} & \cdots & \frac{\partial a_m(X)}{\partial x_2} \\ \cdots & & & \\ \frac{\partial a_1(X)}{\partial x_n} & \frac{\partial a_2(X)}{\partial x_n} & \cdots & \frac{\partial a_m(X)}{\partial x_n} \end{pmatrix}_{n\times m}.$$

同理，$a(X)$ 对 X^T 的导数定义为

$$\frac{\mathrm{d}a(X)}{\mathrm{d}X^T}=\begin{pmatrix} \frac{\partial a_1(X)}{\partial x_1} & \frac{\partial a_1(X)}{\partial x_2} & \cdots & \frac{\partial a_1(X)}{\partial x_n} \\ \frac{\partial a_2(X)}{\partial x_1} & \frac{\partial a_2(X)}{\partial x_2} & \cdots & \frac{\partial a_2(X)}{\partial x_n} \\ \cdots & & & \\ \frac{\partial a_m(X)}{\partial x_1} & \frac{\partial a_m(X)}{\partial x_2} & \cdots & \frac{\partial a_m(X)}{\partial x_n} \end{pmatrix}_{m\times n}.$$

易知

$$\frac{\mathrm{d}a^T(X)}{\mathrm{d}X}=\left(\frac{\mathrm{d}a(X)}{\mathrm{d}X^T}\right)^T,$$

$$\frac{\mathrm{d}X}{\mathrm{d}X^T}=I=\frac{\mathrm{d}X^T}{\mathrm{d}X}.$$

性质 6 设 A 为 $s\times m$ 阶常值矩阵，$f(X)$ 为向量 X 的数量函数，$a(X),b(X)$ 为 X 的 m 维列向量函数，则

(1) $\frac{\mathrm{d}}{\mathrm{d}X}(a^T(X)+b^T(X))=\frac{\mathrm{d}a^T(X)}{\mathrm{d}X}+\frac{\mathrm{d}b^T(X)}{\mathrm{d}X}$；

(2) $\frac{\mathrm{d}}{\mathrm{d}X}(f(X)a^T(X))=\frac{\mathrm{d}f(X)}{\mathrm{d}X}a^T(X)+f(X)\frac{\mathrm{d}a^T(X)}{\mathrm{d}X}$；

(3) $\dfrac{\mathrm{d}}{\mathrm{d}X^T}(Aa(X))=A\dfrac{\mathrm{d}a(X)}{\mathrm{d}X^T}$；

(4) $\dfrac{\mathrm{d}}{\mathrm{d}X}(a^T(X)b(X))=\dfrac{\mathrm{d}a^T(X)}{\mathrm{d}X}b(X)+\dfrac{\mathrm{d}b^T(X)}{\mathrm{d}X}a(X)$；

(5) $\dfrac{\mathrm{d}}{\mathrm{d}X^T}(a^T(X)b(X))=b^T(X)\dfrac{\mathrm{d}a(X)}{\mathrm{d}X^T}+a^T(X)\dfrac{\mathrm{d}b(X)}{\mathrm{d}X^T}$.

证 (1)、(2)的证明是显然的；(3)作为练习；(4)、(5)的证明类似，仅以(4)为例证明之. 由于

$$a^T(X)b(X)=\sum_{i=1}^{m}a_i(X)b_i(X),$$

因此，$a^T(X)b(X)$是向量 X 的数量函数，按数量函数对向量的导数的定义，可得

$$\frac{\mathrm{d}(a^T(X)b(X))}{\mathrm{d}X}=\begin{pmatrix}\sum\limits_{i=1}^{m}\left(\dfrac{\partial a_i(X)}{\partial x_1}b_i(X)+a_i(X)\dfrac{\partial b_i(X)}{\partial x_1}\right)\\ \sum\limits_{i=1}^{m}\left(\dfrac{\partial a_i(X)}{\partial x_2}b_i(X)+a_i(X)\dfrac{\partial b_i(X)}{\partial x_2}\right)\\ \vdots\\ \sum\limits_{i=1}^{m}\left(\dfrac{\partial a_i(X)}{\partial x_n}b_i(X)+a_i(X)\dfrac{\partial b_i(X)}{\partial x_n}\right)\end{pmatrix}$$

$$=\begin{pmatrix}\sum\limits_{i=1}^{m}\dfrac{\partial a_i(X)}{\partial x_1}b_i(X)\\ \sum\limits_{i=1}^{m}\dfrac{\partial a_i(X)}{\partial x_2}b_i(X)\\ \vdots\\ \sum\limits_{i=1}^{m}\dfrac{\partial a_i(X)}{\partial x_n}b_i(X)\end{pmatrix}+\begin{pmatrix}\sum\limits_{i=1}^{m}a_i(X)\dfrac{\partial b_i(X)}{\partial x_1}\\ \sum\limits_{i=1}^{m}a_i(X)\dfrac{\partial b_i(X)}{\partial x_2}\\ \vdots\\ \sum\limits_{i=1}^{m}a_i(X)\dfrac{\partial b_i(X)}{\partial x_n}\end{pmatrix}$$

$$=\frac{\mathrm{d}a^T(X)}{\mathrm{d}X}b(X)+\frac{\mathrm{d}b^T(X)}{\mathrm{d}X}a(X).$$

例 6 设 $A=(a_{ij})_{m\times n}$为数量矩阵，$X=(x_1,x_2,\cdots,x_n)^T$，求线性向量函数 $y=AX$ 关于向量 X^T 的导数.

解 由性质 6 的(3)可得

$$\frac{\mathrm{d}y}{\mathrm{d}X^T}=\frac{\mathrm{d}}{\mathrm{d}X^T}(AX)=A\frac{\mathrm{d}X}{\mathrm{d}X^T}=AI=A.$$

例 6 也可按下列方法求解：由于

$$y=\begin{bmatrix}\sum_{j=1}^{n}a_{1j}x_j\\ \sum_{j=1}^{n}a_{2j}x_j\\ \vdots\\ \sum_{j=1}^{n}a_{mj}x_j\end{bmatrix},$$

按向量函数的导数的定义，有

$$\frac{\mathrm{d}y}{\mathrm{d}X^T}=\begin{bmatrix}\frac{\partial\sum_{j=1}^{n}a_{1j}x_j}{\partial x_1} & \frac{\partial\sum_{j=1}^{n}a_{1j}x_j}{\partial x_2} & \cdots & \frac{\partial\sum_{j=1}^{n}a_{1j}x_j}{\partial x_n}\\ \cdots & & & \\ \frac{\partial\sum_{j=1}^{n}a_{mj}x_j}{\partial x_1} & \frac{\partial\sum_{j=1}^{n}a_{mj}x_j}{\partial x_2} & \cdots & \frac{\partial\sum_{j=1}^{n}a_{mj}x_j}{\partial x_n}\end{bmatrix}$$

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

$$=A.$$

§3.3 向量和矩阵的范数

从计算数学的角度看，在研究计算方法的收敛性和稳定性问题时，范数起到了十分重要的作用。下面对向量和矩阵的范数作进一步的讨论。

1. 向量的范数

定义 1 设 V 是数域 P(实数域或复数域)上的性线空间. 如果对任意 $x\in V$,都对应一个数 $\|x\|\in P$,且满足下列三条性质:

(1) 正定性:对$\forall\ x\in V$,且 $x\neq 0$,有 $\|x\|>0$;

(2) 齐次性:对$\forall\ k\in P$,$x\in V$,有

$$\|kx\|=|k|\,\|x\|;$$

(3) 三角不等式:对$\forall\ x,y\in V$,有

$$\|x+y\|\leqslant\|x\|+\|y\|,$$

则称 V 为**赋范线性空间**,称 $\|x\|$ 是 V 中**向量 x 的范数**.

由定义 1 可知,向量 x 的范数是按照一定的规律与 x 对应的非负函数,这个函数是什么并没有说明. 但只要满足定义中的三条性质,这个函数就是 x 的一种范数. 因此,我们有时称定义 1 中的三条性质为**范数公理**. 因此可以说,凡满足范数公理的实值函数都可以定义为向量的范数.

性质 1 向量范数具有下列性质:

(1) $\|x\|=0$ 当且仅当 $x=0$;

(2) 当 $\|x\|\neq 0$ 时,$\left\|\dfrac{1}{\|x\|}\cdot x\right\|=1$;

(3) 对$\forall\ x\in V$,有 $\|-x\|=\|x\|$;

(4) 对$\forall\ x,y\in V$,有 $\|x\|-\|y\|\leqslant\|x-y\|$.

证 (1) 若 $x=0$,则

$$\|x\|=\|0\|=\|0\cdot x\|=|0|\,\|x\|=0,$$

若 $x\neq 0$,则

$$\|x\|>0,$$

故当 $\|x\|=0$ 时,$x=0$;

(2) $\left\|\dfrac{1}{\|x\|}\cdot x\right\|=\dfrac{1}{\|x\|}\|x\|=1$;

(3) $\|-x\|=|-1|\,\|x\|=\|x\|$;

(4) $\|x\|=\|(x-y)+y\|\leqslant\|x-y\|+\|y\|$,故

$$\|x\| - \|y\| \leqslant \|x-y\|.$$

例 1 设 $x=(x_1,x_2,\cdots,x_n)^T$ 为 C^n 中的任一向量,规定

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|,$$

证明:$\|x\|_1$ 是 C^n 的一种向量范数,称为**1-范数**.

证 (1) 对$\forall\ x=(x_1,x_2,\cdots,x_n)^T\in C^n$,若 $x\neq 0$,则 $x_1,x_2,\cdots,x_n$ 不全为零,故

$$\|x\|_1 = |x_1|+|x_2|+\cdots+|x_n|>0;$$

(2) 对$\forall\ k\in C$,$x=(x_1,x_2,\cdots,x_n)^T\in C^n$,则

$$\begin{aligned}\|kx\|_1 &= \|(kx_1,kx_2,\cdots,kx_n)^T\|_1 \\ &= \sum_{i=1}^{n}|kx_i| \\ &= |k|\sum_{i=1}^{n}|x_i| \\ &= |k|\,\|x\|_1;\end{aligned}$$

(3) 对$\forall\ x=(x_1,x_2,\cdots,x_n)^T\in C^n$,$\forall\ y=(y_1,y_2,\cdots,y_n)^T\in C^n$,则

$$\begin{aligned}\|x+y\|_1 &= \sum_{i=1}^{n}|x_i+y_i| \\ &\leqslant \sum_{i=1}^{n}|x_i| + \sum_{i=1}^{n}|y_i| \\ &= \|x\|_1 + \|y\|_1;\end{aligned}$$

由定义 1 知,$\|x\|_1 = \sum_{i=1}^{n}|x_i|$ 是 C^n 中的向量范数.

例 2 对任意的 $x=(x_1,x_2,\cdots,x_n)^T\in C^n$,规定

$$\|x\|_\infty = \max_{1\leqslant i\leqslant n}|x_i|,$$

则 $\|x\|_\infty$ 是 C^n 中的一种向量范数,称为**∞-范数**.

证 (1) 当 $x=(x_1,x_2,\cdots,x_n)^T\neq 0$ 时,x 必有一个分量不为零,故

$$\|x\|_\infty = \max_{1\leqslant i\leqslant n}|x_i| > 0;$$

(2) 对$\forall\ k\in C$,$\forall\ x=(x_1,x_2,\cdots,x_n)^T\in C^n$,则

$$\begin{aligned}\|kx\|_\infty &= \max_{1\leqslant i\leqslant n}|kx_i| \\ &= |k|\max_{1\leqslant i\leqslant n}|x_i| \\ &= |k|\,\|x\|_\infty;\end{aligned}$$

(3) 对$\forall\ x=(x_1,x_2,\cdots,x_n)^T\in C^n$,$\forall\ y=(y_1,y_2,\cdots,y_n)^T\in C^n$,则

$$\begin{aligned}\|x+y\|_\infty &= \max_{1\leqslant i\leqslant n}|x_i+y_i| \\ &\leqslant \max_{1\leqslant i\leqslant n}(|x_i|+|y_i|) \\ &\leqslant \max_{1\leqslant i\leqslant n}|x_i| + \max_{1\leqslant i\leqslant n}|y_i| \\ &= \|x\|_\infty + \|y\|_\infty,\end{aligned}$$

故 $\|x\|_\infty = \max\limits_{1\leqslant i\leqslant n}|x_i|$ 是C^n中的向量范数.

例 3 对$\forall\ x=(x_1,x_2,\cdots,x_n)^T\in C^n$,规定

$$\|x\|_p = \left(\sum_{i=1}^n|x_i|^p\right)^{\frac{1}{p}}\quad(1\leqslant p<+\infty),$$

则$\|x\|_p$是C^n中的一种向量范数,称为***p*-范数**.

证 (1) 当$x\neq0$时,x至少有一个分量不为零,故

$$\|x\|_p = \left(\sum_{i=1}^n|x_i|^p\right)^{\frac{1}{p}} > 0;$$

(2) 对$\forall\ k\in C$,$\forall\ x=(x_1,x_2,\cdots,x_n)^T\in C^n$,则

$$\begin{aligned}\|kx\|_p &= \|(kx_1,kx_2,\cdots,kx_n)\|_p \\ &= \left(\sum_{i=1}^n|kx_i|^p\right)^{\frac{1}{p}} \\ &= |k|\left(\sum_{i=1}^n|x_i|^p\right)^{\frac{1}{p}} \\ &= |k|\,\|x\|_p;\end{aligned}$$

(3) 对$\forall\ x=(x_1,x_2,\cdots,x_n)^T\in C^n$,$\forall\ y=(y_1,y_2,\cdots,y_n)^T\in C^n$,则

$$\| x+y \|_p = \left(\sum_{i=1}^{n} |x_i + y_i|^p \right)^{\frac{1}{p}},$$

由于

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p \right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i| \right)^{\frac{1}{p}},$$

其中 $1 \leqslant p < +\infty$，从而得到

$$\| x+y \|_p \leqslant \| x \|_p + \| y \|_p,$$

故 $\| x \|_p$ 是 C^n 的向量范数.

注 （1）当 $p=1$ 时，$\| x \|_p = \| x \|_1$；

（2）当 $p=2$ 时，$\| x \|_2 = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{\frac{1}{2}}$，称它为 **2-范数**，2-范数是酉空间范数；当 x_i 为实数时，$\| x \|_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}$ 为欧氏空间范数；

（3）$\lim\limits_{p \to +\infty} \| x \|_p = \| x \|_\infty$.

事实上，设

$$|x_{i_0}| = \max_{1 \leqslant i \leqslant n} |x_i| = \| x \|_\infty,$$

因为

$$|x_{i_0}|^p \leqslant \sum_{i=1}^{n} |x_i|^p \leqslant n |x_{i_0}|^p,$$

所以

$$|x_{i_0}| \leqslant \| x \|_p \leqslant n^{\frac{1}{p}} \cdot |x_{i_0}|.$$

由于

$$\lim_{p \to +\infty} n^{\frac{1}{p}} = 1,$$

故

$$\lim_{p \to +\infty} \| x \|_p = |x_{i_0}| = \| x \|_\infty.$$

1-范数、2-范数、∞-范数是 C^n 中三种常用的向量范数. 但 C^n 中可以定义无穷多种向量范数，下面给出由已知的某种范数构造

出新的向量范数的一种方法.

例 4 设 $\|y\|_\alpha$ 是 C^m 上的一种向量范数,给定矩阵 $A\in C^{m\times n}$,且矩阵 A 的 n 个列向量线性无关,对任意 $x=(x_1,x_2,\cdots,x_n)^T\in C^n$,规定

$$\|x\|_\beta=\|Ax\|_\alpha,$$

则 $\|x\|_\beta$ 是 C^n 中的向量范数.

证 (1) 设 $A_1,A_2,\cdots,A_n$ 是矩阵 A 的 n 个线性无关的列向量,从而对任何 $x=(x_1,x_2,\cdots,x_n)^T\neq 0$,有

$$\begin{aligned}Ax&=(A_1,A_2,\cdots,A_n)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\\&=x_1A_1+x_2A_2+\cdots+x_nA_n\\&\neq 0.\end{aligned}$$

由于 $\|y\|_\alpha$ 是 C^m 上的向量范数,故 $\|Ax\|_\alpha>0$,即

$$\|x\|_\beta=\|Ax\|_\alpha>0;$$

(2) 设 $\forall\ k\in C$,$\forall\ x\in C^n$,则

$$\begin{aligned}\|kx\|_\beta&=\|A(kx)\|_\alpha=\|kAx\|_\alpha\\&=|k|\,\|Ax\|_\alpha\\&=|k|\,\|x\|_\beta;\end{aligned}$$

(3) 对 $\forall\ x,y\in C^n$,有

$$\begin{aligned}\|x+y\|_\beta&=\|A(x+y)\|_\alpha\\&=\|Ax+A\beta\|_\alpha\\&\leqslant\|Ax\|_\alpha+\|Ay\|_\alpha\\&=\|x\|_\beta+\|y\|_\beta.\end{aligned}$$

故 $\|x\|_\beta$ 是 C^n 中的向量范数.

这个例子说明,对于 C^m 中给定的一种向量 α-范数,对于任意给定的 $m\times n$ 矩阵 A 或任意给定的一个从 C^n 到 C^m 的线性变换

A,都可以定义 C^n 中向量的一种范数. 因而 C^n 中可以定义无穷多种向量范数.

例 5 设 V 是 n 维(复的或实的)线性空间,$e_1,e_2,\cdots,e_n$ 是 V 的一组基,则对 $\forall x\in V$,x 有唯一表示式

$$x = x_1e_1 + x_2e_2 + \cdots + x_ne_n,$$

规定

$$\| x \|_E = \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}},$$

证明:$\| x \|_E$ 是 V 中元素的一种范数.

证 (1) 若 $x=x_1e_1+x_2e_2+\cdots+x_ne_n\neq 0$,且 $e_1,e_2,\cdots,e_n$ 线性无关,故至少有一个坐标 $x_{i_0}\neq 0$,因此,$\| x \|_E>0$;

(2) 对 $\forall k\in P$,$\forall x=x_1e_1+x_2e_2+\cdots+x_ne_n\in V$,则

$$\begin{aligned}\| kx \|_E &= \left\| \sum_{i=1}^{n} (kx_i)e_i \right\|_E \\ &= \left(\sum_{i=1}^{n} |kx_i|^2\right)^{\frac{1}{2}} \\ &= |k|\left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}} \\ &= |k| \, \| x \|_E;\end{aligned}$$

(3) 对 $\forall x = \sum_{i=1}^{n} x_ie_i$,$\forall y = \sum_{i=1}^{n} y_ie_i \in V$,有

$$\begin{aligned}\| x + y \|_E &= \left\| \sum_{i=1}^{n} (x_i + y_i)e_i \right\|_E \\ &= \left(\sum_{i=1}^{n} |x_i + y_i|^2\right)^{\frac{1}{2}},\end{aligned}$$

由 C^n 中 2-范数的三角不等式得

$$\begin{aligned}\| x + y \|_E &\leqslant \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}} + \left(\sum_{i=1}^{n} |y_i|^2\right)^{\frac{1}{2}} \\ &= \| x \|_E + \| y \|_E.\end{aligned}$$

故 $\| x \|_E$ 是 V 中的一种向量范数.

由上述例题可知，线性空间 V 中可以有无穷多种范数. 而这些范数之间是否存在某种关系呢？对有限维线性空间，下面给出两种范数相互等价的概念.

定义 2 设 V 是有限维线性空间，$\|x\|_\alpha$ 与 $\|x\|_\beta$ 是 V 中任意两种范数，如果存在正数 M、m，使对一切 $x\in V$，都有

$$m\|x\|_\beta \leqslant \|x\|_\alpha \leqslant M\|x\|_\beta,$$

称 $\|x\|_\alpha$ 与 $\|x\|_\beta$ 是**等价的**.

定理 1 有限维线性空间中的任何两种向量范数都是等价的.

证 设 V 是 n 维线性空间，$e_1, e_2, \cdots, e_n$ 是 V 的一组基，则对 $\forall\ x\in V$，x 有唯一的表达式

$$x = x_1e_1 + x_2e_2 + \cdots + x_ne_n.$$

下面先证明 V 中任何一种范数 $\|x\|$ 都是坐标 $x_1, x_2, \cdots, x_n$ 的连续函数. 令

$$\varphi(x_1, x_2, \cdots, x_n) = \|x\|,$$

则对 $\forall\ y = y_1e_1 + y_2e_2 + \cdots + y_ne_n \in V$，由于

$$\begin{aligned}
|\varphi(x_1, x_2, \cdots, x_n) - \varphi(y_1, y_2, \cdots, y_n)|
&= |\,\|x\| - \|y\|\,| \\
&\leqslant \|x - y\| \\
&= \left\|\sum_{i=1}^{n}(x_i - y_i)e_i\right\| \\
&\leqslant \sum_{i=1}^{n}(|x_i - y_i|\,\|e_i\|) \\
&\leqslant \sqrt{\sum_{i=1}^{n}\|e_i\|^2}\sqrt{\sum_{i=1}^{n}|x_i - y_i|^2} \\
&= k\sqrt{\sum_{i=1}^{n}|x_i - y_i|^2},
\end{aligned}$$

其中 $k=\sqrt{\sum_{i=1}^{n}\|e_i\|^2}$ 是确定的常数.

因此,当 $x=(x_1,x_2,\cdots,x_n)$ 与 $y=(y_1,y_1,\cdots,y_n)$ 无限接近时,$|\varphi(x_1,x_2,\cdots,x_n)-\varphi(y_1,y_2,\cdots,y_n)|$ 无限趋于 0,即 $\varphi(x_1,x_2,\cdots,x_n)=\|x\|$ 是坐标 $x_1,x_2,\cdots,x_n$ 的连续函数.

现在来证明定理的结论.设 $\|x\|_\alpha$ 与 $\|x\|_\beta$ 是 V 中任意两种范数,即要证明存在正数 M,m,使对$\forall\ x\in V$,都有

$$m\|x\|_\beta\leqslant\|x\|_\alpha\leqslant M\|x_\beta\|.$$

当 $x=0$ 时上式显然成立.

当 $x\neq0$ 时,则 $\|x\|_\beta\neq0$.由于 $\|x\|_\alpha$,$\|x\|_\beta$ 都是 $x_1,x_2,\cdots,x_n$ 的连续函数,因此

$$f(x_1,x_2,\cdots,x_n)=\frac{\|x\|_\alpha}{\|x\|_\beta}$$

是 $x_1,x_2,\cdots,x_n$ 的连续函数.考虑单位球面

$$S=\{(\xi_1,\xi_2,\cdots,\xi_n)\mid|\xi_1|^2+|\xi_2|^2+\cdots+|\xi_n|^2=1\},$$

由于 S 是有限闭集,且 S 上的点均不为零,因此 $f(x_1,x_2,\cdots,x_n)$ 在 S 上连续.根据多元连续函数的性质,$f(x_1,x_2,\cdots,x_n)$ 在 S 上取得最大值 M 与最小值 m,即存在 $x_0,y_0\in S$,使得

$$M=\max_{x\in S}\frac{\|x\|_\alpha}{\|x\|_\beta}=\frac{\|x_0\|_\alpha}{\|x_0\|_\beta}>0,$$

$$m=\max_{x\in S}\frac{\|x\|_\alpha}{\|x\|_\beta}=\frac{\|y_0\|_\alpha}{\|y_0\|_\beta}>0.$$

又由于对$\forall\ x=x_1e_1+x_2e_2+\cdots+x_ne_n\in V$,且 $x\neq0$,都有

$$\xi=\frac{x}{\sqrt{|x_1|^2+|x_2|^2+\cdots+|x_n|^2}}\in S,$$

从而

$$m\leqslant\frac{\|\xi\|_\alpha}{\|\xi\|_\beta}=\frac{\|x\|_\alpha}{\|x\|_\beta}\leqslant M,$$

即

$$m\|x\|_\beta \leqslant \|x\|_\alpha \leqslant M\|x\|_\beta.$$

下面在赋范线性空间中引入序列的极限.

定义 3 设

$$x_1,\ x_2,\ \cdots,\ x_n,\ \cdots$$

是线性空间 V 中的元素序列,如果存在 $x\in V$,使得

$$\lim_{m\to+\infty}\|x_m - x\|_\alpha = 0,$$

则称序列$\{x_m\}$**按 α-范数收敛于 x**.

由于线性空间 V 中可以定义多种范数,因此,V 中序列就有多种收敛. 例如,在 C^n 中有 1-范数收敛、2-范数收敛,等等. 这些收敛之间有什么样的关系呢? 它们与向量序列按坐标收敛(§3.1 节中定义 1)又有怎样的关系呢? 对有限维线性空间 V 来说,有下面的结论.

定理 2 (1) 在有限维线性空间中,若序列$\{x_m\}$按某种范数收敛于 x,则$\{x_m\}$按任何范数都收敛于 x,即在有限维线性空间中按范数收敛是等价的;

(2) 在有限维线性空间中,按范数收敛于 x 等价于按坐标收敛于 x.

证 (1) 设 $\|x\|_\alpha$,$\|x\|_\beta$ 是有限维线性空间 V 中的任意两种范数,若序列$\{x_m\}$按 α-范数收敛于 x,即

$$\lim_{m\to+\infty}\|x_m - x\|_\alpha = 0,$$

由定理 1 知,存在常数 $m_0>0$,使对一切 $x\in V$,有

$$\|x\|_\beta \leqslant \frac{1}{m_0}\|x\|_\alpha.$$

由于 $x_m, x\in V$,故 $x_m - x\in V$. 因此

$$0\leqslant \|x_m - x\|_\beta \leqslant \frac{1}{m_0}\|x_m - x\|_\alpha \to 0 \qquad (m\to+\infty),$$

故

$$\lim_{m\to+\infty}\|x_m - x\|_\beta = 0,$$

即序列$\{x_m\}$按 β-范数收敛于 x.

(2) 由(1)可知，$\{x_m\}$按任何范数收敛与按例 5 中 $\|x\|_E$ 范数收敛等价，故

$$\lim_{m\to\infty}\|x_m-x\|=0$$

$$\Longleftrightarrow \lim_{m\to\infty}\|x_m-x\|_E=\lim_{m\to\infty}\left(\sum_{i=1}^{n}|\xi_i^{(m)}-\xi_i|^2\right)^{\frac{1}{2}}=0$$

$$\Longleftrightarrow \lim_{m\to\infty}\xi_i^{(m)}=\xi_i \qquad (i=1,2,\cdots,n),$$

其中，$x_m=(\xi_1^{(m)},\xi_2^{(m)},\cdots,\xi_n^{(m)})$，$x=(\xi_1,\xi_2,\cdots,\xi_n)$，$m=1,2,\cdots$. 从而序列按范数收敛等价于按坐标收敛.

定理 2 的(2)说明：向量序列$\{x_m\}$收敛于 x(指按坐标收敛)与$\{x_m\}$按范数收敛于 x 是一致. 由于 $m\times n$ 阶矩阵可以看成 $m\times n$ 维线性空间中的向量，因此，矩阵范数以及矩阵序列按范数收敛也有同样的定义，而且$\{A_m\}$收敛于 A(按 § 3.1 节中定义)与$\{A_m\}$按矩阵范数收敛于 A 也是一致的.

2. 矩阵的相容范数

上面已经说过，$m\times n$ 阶矩阵可以看成 $m\times n$ 维线性空间中的向量，因此，矩阵范数已有定义了. 但矩阵在应用中具有双重性：一方面 $m\times n$ 阶矩阵可以看成 $m\times n$ 维线性空间中的向量，另一方面任意矩阵 $A_{m\times n}$都可看成是 $C^n\to C^m$ 的映射 A. 因此，作为矩阵的真正有意义的范数，应同时反映这二重特征，所以有必要引入矩阵范数的相容性概念.

定义 4 设 $\|\cdot\|_{m\times n}$，$\|\cdot\|_{n\times p}$和 $\|\cdot\|_{m\times p}$分别是线性空间 $C^{m\times n}$，$C^{n\times p}$和 $C^{m\times p}$中的给定范数，如果对$\forall A\in C^{m\times n}$，$\forall B\in C^{n\times p}$，恒有不等式

$$\|AB\|_{m\times p}\leqslant\|A\|_{m\times n}\|B\|_{n\times p},$$

则称这三种范数是**相容的**.

特别地，当 $p=1$ 时，B 与 AB 分别是 n 维与 m 维列向量，这时定义 4 中的不等式可简单地表示为

$$\| Ax \|_\beta \leqslant \| A \|_{m\times n} \| x \|_\alpha.$$

特别地，当 $A\in C^{n\times n}$，且 α-范数与 β-范数相同时，定义 4 中的不等式为

$$\| Ax \|_\alpha \leqslant \| A \| \ \| x \|_\alpha,$$

这时，称方阵 A 的范数 $\| A \|_{n\times n}$ 与向量的 α-范数 $\| x \|_\alpha$ 是相容的.

例 6 设 $A=(a_{ij})_{n\times n}\in C^{n\times n}$，令

$$\| A \|_F = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{\frac{1}{2}},$$

证明：$\| A \|_F$ 是一种与向量的 2-范数 $\| x \|_2$ 相容的方阵范数. 称它为方阵 A 的 **Frobenius 范数，简称 F-范数.**

证 因为方阵 A 可以看成是线性空间 C^{n^2} 中的向量，因此 $\| A \|_F$ 实质上是向量 A 的 2-范数，故 $\| A \|_F$ 是范数. 下面证明 $\| A \|_F$ 与向量 x 的 2-范数 $\| x \|_2$ 相容，即证明

$$\| Ax \|_2 \leqslant \| A \|_F \| x \|_2.$$

设 $A_i=(a_{i1},a_{i2},\cdots,a_{in})$，$i=1,2,\cdots,n$；$x=(x_1,x_2,\cdots,x_n)^T$，则

$$Ax = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix} x = \begin{pmatrix} A_1x \\ A_2x \\ \vdots \\ A_nx \end{pmatrix},$$

由 Cauchy 不等式得

$$\begin{aligned} |A_ix|^2 &= |a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n|^2 \\ &\leqslant \sum_{j=1}^{n} |a_{ij}|^2 \sum_{j=1}^{n} |x_j|^2 \\ &= \| A_i \|_2^2 \| x \|_2^2. \end{aligned}$$

所以

$$\begin{aligned} \| Ax \|_2^2 &= \sum_{i=1}^{n} |A_ix|^2 \\ &\leqslant \left(\sum_{i=1}^{n} \| A_i \|_2^2 \right) \| x \|_2^2 \end{aligned}$$

$$= \left(\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \right) \|x\|_2^2$$
$$= \|A\|_F^2 \|x\|_2^2,$$

从而

$$\|Ax\|_2 \leqslant \|A\|_F \|x\|_2.$$

方阵 A 的 F-范数除了与向量 x 的 2-范数相容外，还有下面的重要性质.

性质 2 对$\forall A, B \in C^{n\times n}$，恒有

$$\|AB\|_F \leqslant \|A\|_F \|B\|_F.$$

证 设 $B_1, B_2, \cdots, B_n$ 为 B 的 n 个列向量，则

$$\|B\|_F^2 = \|B_1\|_2^2 + \|B_2\|_2^2 + \cdots + \|B_n\|_2^2,$$

故

$$\begin{aligned}
\|AB\|_F^2 &= \|(AB_1, AB_2, \cdots, AB_n)\|_F \\
&= \|AB_1\|_2^2 + \|AB_2\|_2^2 + \cdots + \|AB_n\|_2^2 \\
&\leqslant \|A\|_F^2 \|B_1\|_2^2 + \|A\|_F^2 \|B_2\|_2^2 + \cdots \\
&\quad + \|A\|_F^2 \|B_n\|_2^2 \\
&= \|A\|_F^2 (\|B_1\|_2^2 + \|B_2\|_2^2 + \cdots + \|B_n\|_2^2) \\
&= \|A\|_F^2 \|B\|_F^2,
\end{aligned}$$

因此

$$\|AB\|_F \leqslant \|AB\|_F \|B\|_F.$$

方阵 A 的 F-范数与向量 x 的 2-范数相容的性质反映了这样一个事实：A 的 F-范数 $\|A\|_F$ 是象 Ax 的 2-范数 $\|Ax\|_2$ 与原象 x 的 2-范数 $\|x\|_2$ 之比的一个上界，即

$$\|Ax\|_2 \leqslant \|A\|_F \|x\|_2.$$

因此可以用 $\|A\|_F$ 来评价变换 A 的结果. 但是，这种估计比较粗糙，因为恒等变换 E_n 的 F-范数为 $\sqrt{n}$. 现在的问题是象 Ax 的范数 $\|Ax\|_\beta$ 与原象 x 的范数 $\|x\|_\alpha$ 之比的最小上界(即上确界)是否是 A 的范数呢？如果是 A 的范数，用它来评价变换 A 的

结果应该是最精确的. 这问题就是下面所讨论的 A 的算子范数问题.

3. 算子范数

例 7 设 $A\in C^{m\times n}$, $x\in C^n$, $\|x\|_\alpha$ 与 $\|Ax\|_\beta$ 分别是 C^n 与 C^m 中的范数，令

$$\|A\| = \sup_{x\neq 0}\frac{\|Ax\|_\beta}{\|x\|_\alpha} = \sup_{\|x\|_\alpha=1}\|Ax\|_\beta,$$

证明：$\|A\|$ 为 A 的范数，并称它为 A 的**算子范数**.

证 (1) 正定性. 对$\forall$ $A_{m\times n}\neq 0$，则必存在非零向量 $x_0\in C^n$，使得 $Ax_0\neq 0$，因此

$$\|Ax_0\|_\beta > 0, \quad \|x_0\|_\alpha > 0,$$

从而

$$\|A\| = \sup_{x\neq 0}\frac{\|Ax\|_\beta}{\|x\|_\alpha} \geqslant \frac{\|Ax_0\|_\beta}{\|x_0\|_\alpha} > 0;$$

(2) 齐次性. $\forall$ $k\in C$，有

$$\begin{aligned}\|kA\| &= \sup_{x\neq 0}\frac{\|(kA)x\|_\beta}{\|x\|_\alpha} \\ &= \sup_{x\neq 0}\frac{|k|\,\|Ax\|_\beta}{\|x\|_\alpha} \\ &= |k|\sup_{x\neq 0}\frac{\|Ax\|_\beta}{\|x\|_\alpha} \\ &= |k|\,\|Ax\|,\end{aligned}$$

(3) 三角不等式. 对$\forall$ A、$B\in C^{m\times n}$，有

$$\begin{aligned}\|A+B\| &= \sup_{x\neq 0}\frac{\|(A+B)x\|_\beta}{\|x\|_\alpha} \\ &\leqslant \sup_{x\neq 0}\frac{\|Ax\|_\beta + \|Bx\|_\beta}{\|x\|_\alpha} \\ &\leqslant \sup_{x\neq 0}\frac{\|Ax\|_\beta}{\|x\|_\alpha} + \sup_{x\neq 0}\frac{\|Bx\|_\beta}{\|x\|_\alpha}\end{aligned}$$

$$= \| A \| + \| B \|,$$

从而 $\| A \|$ 为 A 的范数.

定理 3 设 $A \in C^{m \times n}$, $x \in C^n$, $\| A \|$ 为 A 的算子范数,则

(1) 对$\forall\ x \in C^n$,恒有

$$\| Ax \|_\beta \leqslant \| A \| \, \| x \|_\alpha,$$

即 A 的算子范数与向量的 β-范数及 α-范数是相容的;

(2) 若存在正数 M,使得对$\forall\ x \in C^n$,有

$$\| Ax \|_\beta \leqslant M \| x \|_\alpha,$$

则 $\| A \| \leqslant M$,即 $\| A \|$ 是使得上述不等式成立的最小常数;

(3) 对$\forall\ A \in C^{m \times n}$, $B \in C^{n \times p}$,恒有

$$\| AB \| \leqslant \| A \| \, \| B \|.$$

特别地,若 $A \in C^{n \times n}$,则

$$\| A^2 \| \leqslant \| A \|^2,\ \| A^3 \| \leqslant \| A \|^3,\ \cdots,\ \| A^n \| \leqslant \| A \|^n.$$

显然,A 的算子范数与 $\| Ax \|_\beta$ 和 $\| x \|_\alpha$ 的定义有关.当 A 为方阵时,若 α-范数与 β-范数为同一范数,则 $\| A \|$ 仅与向量 α-范数有关,为了方便,我们称此矩阵数 $\| A \|$ 为从属于向量的 α-范数的算子范数.

例 8 证明:方阵 $A = (a_{ij})_{n \times n} \in C^{n \times n}$ 从属于向量的 1-范数的算子范数 $\| A \|_1$ 为

$$\| A \|_1 = \max_{1 \leqslant j \leqslant n} \sum_{i=1}^{n} |a_{ij}|.$$

证 对$\forall\ x = (x_1, x_2, \cdots, x_n)^T \in C^n$,则

$$Ax = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum\limits_{j=1}^{n} a_{1j} x_j \\ \sum\limits_{j=1}^{n} a_{2j} x_j \\ \vdots \\ \sum\limits_{j=1}^{n} a_{nj} x_j \end{pmatrix},$$

故向量 Ax 的 1-范数为

$$
\begin{aligned}
\| Ax \|_1 &= \left|\sum_{j=1}^{n} a_{1j}x_j\right| + \left|\sum_{j=1}^{n} a_{2j}x_j\right| + \cdots + \left|\sum_{j=1}^{n} a_{nj}x_j\right| \\
&= \sum_{i=1}^{n} \left|\sum_{j=1}^{n} a_{ij}x_j\right| \\
&\leqslant \sum_{i=1}^{n}\sum_{j=1}^{n} (|a_{ij}| \cdot |x_j|) \\
&= \sum_{j=1}^{n}\left(\sum_{i=1}^{n} |a_{ij}| \cdot |x_j|\right) \\
&\leqslant \max_{1\leqslant j\leqslant n}\left(\sum_{i=1}^{n} |a_{ij}|\right) \sum_{j=1}^{n} |x_j| \\
&= \max_{1\leqslant j\leqslant n} \sum_{i=1}^{n} |a_{ij}| \, \| x \|_1.
\end{aligned}
$$

由定理 3 的(2)可知

$$\| A \|_1 \leqslant \max_{1\leqslant j\leqslant n} \sum_{i=1}^{n} |a_{ij}|.$$

下面再证相反的不等式. 由于 j 只取有限个值,故有 j_0,使得

$$\max_{1\leqslant j\leqslant n} \sum_{i=1}^{n} |a_{ij}| = \sum_{i=1}^{n} |a_{ij_0}|.$$

设 x_0 是第 j_0 个分量为 1 而其它分量为 0 的列向量,即

$$x_0 = (0,\cdots,0,1,0,\cdots,0)^T,$$

则 $\| x_0 \|_1 = 1$,而且

$$Ax_0 = (a_{1j_0}, a_{2j_0}, \cdots, a_{nj_0})^T,$$

故

$$\| Ax_0 \|_1 = \sum_{i=1}^{n} |a_{ij_0}| = \max_{1\leqslant j\leqslant n} \sum_{i=1}^{n} |a_{ij}|.$$

由于 $\| A \|_1$ 是算子范数,所以

$$\| Ax_0 \|_1 \leqslant \| A \|_1 \| x_0 \|_1 = \| A \|_1,$$

因此

$$\max_{1\leqslant j\leqslant n}\sum_{i=1}^{n}|a_{ij}| = \| Ax_0 \|_1 \leqslant \| A \|_1.$$

从而

$$\| A \|_1 = \max_{1\leqslant j\leqslant n}\sum_{i=1}^{n}|a_{ij}|.$$

由于方阵 A 的 1-范数等于 A 的列向量的 1-范数的最大值，故又称 $\| A \|_1$ 为方阵 A 的**列范数**.

例 9 证明：$A=(a_{ij})_{n\times n}\in C^{n\times n}$ 从属于向量的∞-范数 $\| x \|_\infty = \max\limits_{1\leqslant j\leqslant n}|x_j|$ 的算子范数 $\| A \|_\infty$ 为

$$\| A \|_\infty = \max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}|a_{ij}|.$$

证 对∀ $x=(x_1,x_2,\cdots,x_n)^T\in C^n$，则

$$\begin{aligned}\| Ax \|_\infty &= \max_{1\leqslant i\leqslant n}\left|\sum_{j=1}^{n}a_{ij}x_j\right| \\ &\leqslant \max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}|a_{ij}| \max_{1\leqslant i\leqslant n}|x_j| \\ &\leqslant \max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}|a_{ij}| \; \| x \|_\infty.\end{aligned}$$

由于 $\| A \|_\infty$ 为算子范数，由定理 3 的(2)知

$$\| A \|_\infty \leqslant \max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}|a_{ij}|.$$

下面再证相反的不等式. 由于 i 只取 $1,2,\cdots,n$ 这有限个值，故存在 i_0 使得

$$\max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}|a_{ij}| = \sum_{j=1}^{n}|a_{i_0j}|,$$

取 $x_0=(\xi_1,\xi_2,\cdots,\xi_n)^T$，其中 ξ_j 如下：

$$\xi_j = \begin{cases}\dfrac{|a_{i_0j}|}{a_{i_0j}}, & \text{当 } a_{i_0j}\neq 0 \text{ 时} \\ 0, & \text{当 } a_{i_0j}=0 \text{ 时}\end{cases} \qquad (j=1,2,\cdots,n),$$

则

$$\| x_0 \|_\infty = \max_{1\leqslant j\leqslant n} |\xi_j| = 1,$$

且 $|\xi_j|\leqslant 1$. 由于对 $\forall\, i\in\{1,2,\cdots,n\}$,有

$$\begin{aligned}\left|\sum_{j=1}^{n} a_{ij}\xi_j\right| &\leqslant \sum_{j=1}^{n} |a_{ij}|\cdot|\xi_j| \leqslant \sum_{j=1}^{n}|a_{ij}| \\ &\leqslant \sum_{j=1}^{n}|a_{i_0 j}| = \sum_{j=1}^{n}|a_{i_0 j}\xi_j|.\end{aligned}$$

从而

$$\begin{aligned}\| Ax_0 \|_\infty &= \left\| \left(\sum_{j=1}^{n} a_{1j}\xi_j, \sum_{j=1}^{n} a_{2j}\xi_j, \cdots, \sum_{j=1}^{n} a_{nj}\xi_j \right)^T \right\|_\infty \\ &= \max_{1\leqslant i\leqslant n} \left| \sum_{j=1}^{n} a_{ij}\xi_j \right| \\ &= \sum_{j=1}^{n} |a_{i_0 j}\xi_j| \\ &= \sum_{j=1}^{n} |a_{i_0 j}| \\ &= \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}|.\end{aligned}$$

故

$$\begin{aligned}\| A \|_\infty &= \| A \|_\infty \| x_0 \|_\infty \\ &\geqslant \| Ax_0 \|_\infty \\ &= \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}|.\end{aligned}$$

综上所证即得

$$\| A \|_\infty = \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}|.$$

由于方阵的∞-范数 $\| A \|_\infty$ 等于行向量的 1-范数的最大值,因此又称方阵 A 的∞-范数 $\| A \|_\infty$ 为**行范数**.

例 10　证明:方阵 $A=(a_{ij})_{n\times n}\in C^{n\times n}$ 从属于向量的 2-范数

$$\|x\|_2=\left(\sum_{i=1}^{n}|x_i|^2\right)^{\frac{1}{2}}$$

的算子范数 $\|A\|_2$ 为

$$\|A\|_2=\sqrt{\lambda_1},$$

其中,λ_1 是 $\overline{A}^TA$ 的最大特征值,故又称 $\|A\|_2$ 为 A 的 λ_1-范数.

证 因为 $\overline{(\overline{A}^TA)}=\overline{A}^TA$,故 $\overline{A}^TA$ 是 n 阶厄米特(Hermite)矩阵,它对应的厄米特二次型为

$$f(x)=\overline{x}^T(\overline{A}^TA)x=(\overline{Ax})^TAx\geqslant 0,$$

因此,$f(x)$ 为正定或半正定的. 故 $\overline{A}^TA$ 的 n 个特征值都大于等于零. 不妨设这 n 个特征值为

$$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n\geqslant 0,$$

设 $\xi_1,\xi_2,\cdots,\xi_n$ 分别是属于 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 的单位正交特征向量.

对$\forall\ x\in C^n$,且 $\|x\|_2=1$,则 x 可表示为

$$x=a_1\xi_1+a_2\xi_2+\cdots+a_n\xi_n,$$

其中

$$\|x\|_2^2=\overline{x}^Tx=|a_1|^2+|a_2|^2+\cdots+|a_n|^2=1.$$

则

$$\begin{aligned}(\overline{A}^TA)x&=a_1(\overline{A}^TA)\xi_1+a_2(\overline{A}^TA)\xi_2+\cdots+a_n(\overline{A}^TA)\xi_n\\&=a_1\lambda_1\xi_1+a_2\lambda_2\xi_2+\cdots+a_n\lambda_n\xi_n,\end{aligned}$$

故

$$\begin{aligned}\|Ax\|_2^2&=(\overline{Ax})^T(Ax)\\&=\overline{x}^T(\overline{A}^TAx)\\&=(\overline{a}_1\overline{\xi}_1^T+\overline{a}_2\overline{\xi}_2^T+\cdots+\overline{a}_n\overline{\xi}_n^T)\\&\quad\cdot(a_1\lambda_1\xi_1+a_2\lambda_2\xi_2+\cdots+a_n\lambda_n\xi_n)\\&=\lambda_1|a_1|^2+\lambda_2|a_2|^2+\cdots+\lambda_n|a_n|^2\\&\leqslant\lambda_1(|a_1|^2+|a_2|^2+\cdots+|a_n|^2)\\&=\lambda_1\|x\|_2^2=\lambda_1.\end{aligned}$$

故

$$\| Ax \|_2 \leqslant \sqrt{\lambda_1} = \sqrt{\lambda_1} \| x \|_2,$$

则由定理 3 的(2)得

$$\| A \|_2 \leqslant \sqrt{\lambda_1}.$$

又因为 $\xi_1 = 1 \cdot \xi_1 + 0 \cdot \xi_2 + \cdots + 0 \cdot \xi_n$,故

$$\| \xi_1 \|_2 = 1,$$

从而

$$\| A\xi_1 \|_2^2 = \overline{\xi}_1^T \overline{A}^T A\xi_1 = \overline{\xi}_1^T \lambda_1 \xi_1 = \lambda_1,$$

即

$$\| A\xi_1 \|_2 = \sqrt{\lambda_1},$$

因此

$$\| A \|_2 = \| A \|_2 \| \xi_1 \|_2 \geqslant \| A\xi_1 \|_2 = \sqrt{\lambda_1}.$$

故

$$\| A \|_2 = \sqrt{\lambda_1}.$$

与 $\| A \|_1$ 和 $\| A \|_\infty$ 比较,$\| A \|_2$ 在计算上没有前两种方便,但 $\| A \|_2$ 具有一些特殊性质,使它在理论上有广泛的用途.

§3.4 矩 阵 函 数

我们已经研究了以实数为自变量且取值为实数的函数,以及以复数为自变量且取值为复数的复变函数,现在开始研究以矩阵为变量且取值为矩阵的函数,并称这类函数为矩阵函数. 在本节中,我们先通过收敛的矩阵级数给出矩阵函数的定义,然后再讨论矩阵函数的计算等问题.

1. 矩阵级数收敛性的判别法

由于在有限维线性空间中,矩阵序列按范数收敛与按坐标收

敛是等价的，因此可立即得到定理 1.

定理 1（Cauchy 收敛准则） $\sum_{m=1}^{\infty} A_m$ 收敛，当且仅当矩阵序列 $S_m = A_1 + A_2 + \cdots + A_m$ 收敛. 即，当且仅当仅对 $\forall\ \varepsilon > 0$ 存在 $N > 0$，对任何正整数 p, q，只要 $q > p > N$，都有 $\left\| \sum_{m=p}^{q} A_m \right\| < \varepsilon$.

定理 2 若数项级数 $\sum_{m=1}^{\infty} \|A_m\|$ 收敛，则矩阵级数 $\sum_{m=1}^{\infty} A_m$ 收敛.

证 由于数项级数 $\sum_{m=1}^{\infty} \|A_m\|$ 收敛，由数项级数的 Cauchy 收敛准则知，对 $\forall\ \varepsilon > 0$，存在 $N > 0$，当 $q > p > N$ 时，都有

$$\sum_{m=p}^{q} \|A_m\| < \varepsilon,$$

从而

$$\left\| \sum_{m=p}^{q} A_m \right\| \leqslant \sum_{m=p}^{q} \|A_m\| < \varepsilon.$$

由定理 1 知，矩阵级数 $\sum_{m=1}^{\infty} A_m$ 收敛.

特别地，对于方阵 A，若 $\sum_{k=1}^{\infty} \|A\|^k$ 收敛，则矩阵幂级数 $\sum_{k=1}^{\infty} A^k$ 收敛.

定理 3 设幂级数 $\sum_{k=0}^{\infty} a_k \lambda^k$ 的收敛半径为 R，则当方阵 A 的范数 $\|A\| < R$ 时，矩阵幂级数 $\sum_{k=0}^{\infty} a_k A^k$ 收敛.

证 由于 $\|A\| < R$，故数项级数 $\sum_{k=0}^{\infty} |a_k| \|A\|^k$ 收敛. 又

$$\|a_k A_k\| = |a_k| \|A^k\| \leqslant |a_k| \|A\|^k,$$

由数项级数的比较判别法知，数项级数 $\sum\limits_{k=0}^{\infty}\|a_kA^k\|$ 收敛. 又由定理 2 知，矩阵幂级数 $\sum\limits_{k=0}^{\infty}a_kA^k$ 收敛.

2. 矩阵函数的定义及性质

定义 1 设幂级数 $f(\lambda)=\sum\limits_{k=0}^{\infty}a_k\lambda^k$ 的收敛半径为 R，当 $\|A\|<R$ 时，定义方阵 A 的级数

$$a_0I + a_1A + a_2A^{-}2\cdots + a_kA^k + \cdots$$

的和为 $f(A)$，即

$$f(A)=a_0I+a_1A+a_2A^2+\cdots+a_kA^k+\cdots$$
$$=\sum_{k=0}^{\infty}a_kA^k.$$

特别地，当 $f(\lambda)=a_0+a_1\lambda+a_2\lambda^2+\cdots+a_k\lambda^k$ 时，矩阵 A 的多项式$f(A)$为

$$f(A)=a_0I+a_1A+a_2A^2+\cdots+a_kA^k.$$

由于

$$e^{\lambda}=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}$$
$$=1+\lambda+\frac{1}{2!}\lambda^2+\cdots+\frac{1}{k!}\lambda^k+\cdots,$$

$$\sin\lambda=\sum_{k=0}^{\infty}(-1)^k\frac{1}{(2k+1)!}\lambda^{2k+1}$$
$$=\lambda-\frac{1}{3!}\lambda^3+\frac{1}{5!}\lambda^5+\cdots$$
$$+(-1)^k\frac{1}{(2k+1)!}\lambda^{2k-1}+\cdots,$$

$$\cos\lambda=\sum_{k=0}^{\infty}(-1)^k\frac{1}{(2k)!}\lambda^{2k}$$

$$
= 1 - \frac{1}{2!}\lambda^2 + \frac{1}{4!}\lambda^4 + \cdots
$$

$$
+ (-1)^k \frac{1}{(2k)!}\lambda^{2k} + \cdots,
$$

且它们的收敛半径都为$+\infty$. 因此,对任何方阵 A,都有

$$
\mathrm{e}^A = \sum_{k=0}^{\infty} \frac{1}{k!}A^k
$$

$$
= I + A + \frac{1}{2!}A^2 + \cdots + \frac{1}{k!}A^k + \cdots,
$$

$$
\sin A = \sum_{k=0}^{\infty}(-1)^k \frac{1}{(2k+1)!}A^{2k+1}
$$

$$
= A - \frac{1}{3!}A^3 + \frac{1}{5!}A^5 + \cdots
$$

$$
+ (-1)^k \frac{1}{(2k+1)!}A^{2k+1} + \cdots,
$$

$$
\cos A = \sum_{k=0}^{\infty}(-1)^k \frac{1}{(2k)!}A^{2k}
$$

$$
= I - \frac{1}{2!}A^2 + \frac{1}{4!}A^4 + \cdots
$$

$$
+ (-1)^k \frac{1}{(2k)!}A^{2k} + \cdots.
$$

对方阵 A 的函数 e^A,$\sin A$,$\cos A$,容易验证下列性质:

性质 1 对$\forall$ $A,B\in C^{n\times n}$,$k,l\in C$,有

(1) $\mathrm{e}^{kA}\mathrm{e}^{lA}=\mathrm{e}^{(k+l)A}$;

(2) $(\mathrm{e}^A)^{-1}=\mathrm{e}^{-A}$;

(3) 当 $AB=BA$ 时,$\mathrm{e}^A\mathrm{e}^B=\mathrm{e}^B\mathrm{e}^A=\mathrm{e}^{A+B}$;

(4) $\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{At})=A\mathrm{e}^{At}=\mathrm{e}^{At}A$;

(5) $\frac{\mathrm{d}}{\mathrm{d}t}(\sin At)=A\cos At=\cos At\cdot A$;

(6) $\frac{\mathrm{d}}{\mathrm{d}t}(\cos At)=-A\sin At=-\sin At\cdot A$.

3. 矩阵函数的计算方法

矩阵函数的计算问题，是矩阵在应用中的关键问题. 矩阵函数的计算是相当复杂的，例如，简单的矩阵函数 A^{101}就要计算 100 次矩阵 A 的乘积；若 A 为 5 阶方阵，则要进行 22500 次加法和乘法运算. 因此，研究如何方便地计算矩阵函数是非常有意义的. 本节将讨论四种计算法.

I. 递推公式计算法

设 $f(\lambda)=|\lambda I-A|$，根据 Cayley-Hamilton 定理知，$f(A)=0$，由此可得 A 的递推关系式，从而计算给定的矩阵 A 的函数.

例 1 设 4 阶方阵 A 的特征值为 $\pi,-\pi,0,0$，求 $\sin A,\cos A$.

解 因为 A 的特征值为 $\pi,-\pi,0,0$，故 A 的特征多项式 $f(\lambda)$ 为

$$f(\lambda)=\lambda^2(\lambda^2-\pi^2)=\lambda^4-\lambda^2\pi^2.$$

由 $f(A)=0$ 得

$$A^4-\pi^2A^2=0,$$

即

$$A^4=\pi^2A^2.$$

因此

$$\begin{aligned}
&A^5=A^4A=\pi^2A^3=\pi^{5-3}A^3,\\
&A^7=A^5A^2=\pi^2A^5=\pi^4A^3=\pi^{7-3}A^3,\\
&A^9=A^7A^2=\pi^4A^5=\pi^6A^3=\pi^{9-3}A^3,\\
&\cdots\cdots\\
&A^{2k+1}=\pi^{(2k+1)-3}A^3=\pi^{2k-2}A^3,\\
&\cdots\cdots
\end{aligned}$$

从而

$$\begin{aligned}
\sin A=A-\frac{1}{3!}A^3+\frac{1}{5!}A^5-\frac{1}{7!}A^7+\cdots\\
+(-1)^k\frac{1}{(2k+1)!}A^{2k+1}+\cdots
\end{aligned}$$

$$
\begin{aligned}
&= A - \frac{1}{3!}A^3 + \frac{1}{5!}\pi^2 A^3 - \frac{1}{7!}\pi^4 A^3 + \cdots \\
&\quad + (-1)^k \frac{1}{(2k+1)!}\pi^{2k-2}A^3 + \cdots \\
&= A + A^3\Big(-\frac{1}{3!} + \frac{1}{5!}\pi^2 - \frac{1}{7!}\pi^4 + \cdots \\
&\quad + (-1)^k \frac{1}{(2k+1)!}\pi^{2k-2} + \cdots \Big) \\
&= A + \frac{1}{\pi^3}A^3\Big[-\pi + \Big(\pi - \frac{1}{3!}\pi^3 + \frac{1}{5!}\pi^5 - \frac{1}{7!}\pi^7 + \cdots \\
&\quad + (-1)^k \frac{1}{(2k+1)!}\pi^{2k+1} + \cdots \Big) \Big] \\
&= A + \frac{1}{\pi^3}A^3(-\pi + \sin\pi) \\
&= A - \frac{1}{\pi^2}A^3.
\end{aligned}
$$

同理可得

$$
\begin{aligned}
\cos A &= I - \frac{1}{2!}A^2 + \frac{1}{4!}A^4 + \cdots + (-1)^k \frac{1}{(2k)!}A^{2k} + \cdots \\
&= I - \frac{2}{\pi^2}A^2.
\end{aligned}
$$

Ⅱ. 利用 Jordan 标准形的计算法

由递推公式计算法知，若 A 是有限阶方阵，则由矩阵幂级数定义的矩阵函数 $f(A)$ 与矩阵 A 的某一多项式相等. 因此，对给定的有限阶方阵 A，计算 $f(A)$ 的问题，就是计算矩阵多项式的问题，因而关键是计算 A^n 的问题. 下面就 A 为各种不同矩阵情况下的计算问题进行讨论.

(1) A 为对角矩阵

设

$$
A = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix},
$$

则

$$A^m = \begin{pmatrix} a_1^m & & & \\ & a_2^m & & \\ & & \ddots & \\ & & & a_n^m \end{pmatrix}.$$

(2) A 为对角形分块矩阵

设

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix},$$

其中 $A_1, A_2, \cdots, A_k$ 为 A 的子方阵. 由于分块矩阵的乘积与矩阵乘积类似，故对上述分块矩阵 A，有

$$A^m = \begin{pmatrix} A_1^m & & & \\ & A_2^m & & \\ & & \ddots & \\ & & & A_k^m \end{pmatrix}.$$

(3) A 为一般矩阵

由于对任何方阵 A，总有 A 的 Jordan 标准形 J 及满秩方阵 P，使得 $A=PJP^{-1}$，因此

$$A^m = PJ^mP^{-1}.$$

若

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix},$$

其中

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}_{l_i \times l_i},$$

$\lambda_i(i=1,2,\cdots,k)$为 A 的 l_i 重特征根，且 $l_1+l_2+\cdots+l_k=n$，则

$$A^m = PJ^mP^{-1} = P\begin{pmatrix} J_1^m & & & \\ & J_2^m & & \\ & & \ddots & \\ & & & J_k^m \end{pmatrix}P^{-1}.$$

由上述讨论知，对一般的 n 阶方阵 A，要计算 A^m，实质上是计算 A 的 Jordan 块 J_i 的函数 J_i^m. 并且通过上述(1)、(2)、(3)的讨论可知，A 的多项式及 A 的幂级数的计算问题亦可化为计算 A 的 Jordan 块的函数.

(4) 计算 Jordan 块 J_i 的函数 $f(J_i)$

设

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}_{k \times k},$$

令

$$H_i = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}_{k \times k},$$

则

$$J_i = \lambda_i I + H_i,$$

即

$$H_i = J_i - \lambda_i I.$$

又

$$H_i^2 = \begin{pmatrix} 0 & 0 & 1 & & \\ & 0 & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & \ddots & 0 \\ & & & & 0 \end{pmatrix}_{k\times k},$$

$$H_i^3 = \begin{pmatrix} 0 & 0 & 0 & 1 & \cdots & 0 \\ & \ddots & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \ddots & 1 \\ & & & \ddots & \ddots & 0 \\ & & & & \ddots & 0 \\ & & & & & 0 \end{pmatrix}_{k\times k},$$

……

$$H_i^p = 0 \qquad (p \geqslant k).$$

设函数 $f(\lambda)$在 λ_i 处的 Taylor 展开式为

$$f(\lambda) = \sum_{m=0}^{\infty} \frac{f^{(m)}(\lambda_i)}{m!}(\lambda - \lambda_i)^m,$$

则

$$\begin{aligned} f(J_i) &= f(\lambda_i)I + \frac{f'(\lambda_i)}{1!}(J_i - \lambda_i I) + \frac{f''(\lambda_i)}{2!}(J_i - \lambda_i I)^2 \\ &\quad + \cdots + \frac{f^{(m)}(\lambda_i)}{m!}(J_i - \lambda_i I)^m + \cdots \\ &= f(\lambda_i)I + \frac{f'(\lambda_i)}{1!}H_i + \frac{f''(\lambda_i)}{2!}H_i^2 + \cdots \\ &\quad + \frac{f^{(k-1)}(\lambda_i)}{(k-1)!}H_i^{k-1} + 0 + 0 + \cdots \end{aligned}$$

$$= \begin{pmatrix} f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \frac{f''(\lambda_i)}{2!} & \cdots & \frac{f^{(k-1)}(\lambda_i)}{(k-1)!} \\ 0 & f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \cdots & \frac{f^{(k-2)}(\lambda_i)}{(k-2)!} \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & \frac{f'(\lambda_i)}{1!} \\ 0 & 0 & 0 & \cdots & f(\lambda_i) \end{pmatrix}.$$

由上述讨论可知，对于给定的一般矩阵 A 及函数 $f(\lambda)$，计算 $f(A)$ 的步骤如下：

第一步，经过相似变换将 A 化成 A 的 Jordan 标准形 J，并求相似变换矩阵 P，使得 $A=PJP^{-1}$. 其中：

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix},$$

$$J_i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}_{t_i \times t_i} \qquad (i = 1,2,\cdots,k);$$

第二步，计算 $f(J)$.

$$f(J) = \begin{pmatrix} f(J_1) & & & \\ & f(J_2) & & \\ & & \ddots & \\ & & & f(J_k) \end{pmatrix},$$

其中

$$f(J_i)=\begin{pmatrix} f(\lambda_i) & f'(\lambda_i) & \dfrac{f''(\lambda_i)}{2!} & \cdots & \dfrac{f^{(l_i-1)}(\lambda_i)}{(l_i-1)!} \\ 0 & f(\lambda_i) & f'(\lambda_i) & \cdots & \dfrac{f^{(l_i-2)}(\lambda_i)}{(l_i-2)!} \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & f(\lambda_i) \end{pmatrix};$$

第三步，计算 $f(A)$.

$$f(A)=Pf(J)P^{-1}.$$

例 2 设 $A\in C^{n\times n}$，它的 Jordan 标准形为

$$J=\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix},$$

其中

$$J_i=\begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}_{l_i\times l_i} \quad (i=1,2,\cdots,k),$$

且 $A=PJP^{-1}$，试写出 e^{At}，$\sin At$，A^m.

解 (1) 计算 e^{At}. 此时

$$f(\lambda)=e^{\lambda t},$$
$$f'(\lambda)=te^{\lambda t},$$
$$f''(\lambda)=t^2e^{\lambda t},$$
$$\cdots\cdots$$
$$f^{(l_i-1)}(\lambda)=t^{l_i-1}e^{\lambda t},$$

故

$$e^{At} = Pe^{Jt}P^{-1} = P\begin{pmatrix} e^{J_1t} & & & \\ & e^{J_2t} & & \\ & & \ddots & \\ & & & e^{J_kt} \end{pmatrix}P^{-1},$$

其中

$$e^{J_it} = \begin{pmatrix} e^{\lambda_it} & te^{\lambda_it} & \frac{1}{2!}t^2e^{\lambda_it} & \cdots & \frac{1}{(l_i-1)!}t^{l_i-1}e^{\lambda_it} \\ 0 & e^{\lambda_it} & te^{\lambda_it} & \cdots & \frac{1}{(l_i-2)!}t^{l_i-2}e^{\lambda_it} \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & e^{\lambda_it} \end{pmatrix}_{l_i\times l_i},$$

$(i = 1,2,\cdots k)$;

(2) 计算 $\sin At$. 此时

$$f(\lambda) = \sin\lambda t,$$

$$f'(\lambda) = t\cos\lambda t = t\sin(\lambda t + \frac{\pi}{2}),$$

$$f''(\lambda) = t^2\sin(\lambda t + \frac{2\pi}{2}),$$

$$\cdots\cdots$$

$$f^{(l_i-1)}(\lambda) = t^{l_i-1}\sin(\lambda t - \frac{l_i-1}{2}\pi),$$

故

$$\sin At = P\sin Jt \cdot P^{-1}$$

$$= P\begin{pmatrix} \sin J_1t & & & \\ & \sin J_2t & & \\ & & \ddots & \\ & & & \sin J_kt \end{pmatrix}P^{-1},$$

其中

$\sin J_it =$

$$\begin{bmatrix} \sin\lambda_i t & t\sin(\lambda_i t+\frac{\pi}{2}) & \cdots & \frac{1}{(l_i-1)!}t^{l_i-1}\sin(\lambda_i t+(l_i-1)\frac{2}{\pi}) \\ 0 & \sin\lambda_i t & \cdots & \frac{1}{(l_i-2)!}t^{l_i-2}\sin(\lambda_i t+(l_i-2)\frac{2}{\pi}) \\ \cdots & & & \\ 0 & 0 & \cdots & \sin\lambda_i t \end{bmatrix}_{l_i\times l_i}$$

$(i=1,2,\cdots,k)$.

(3) 计算 A^m. 此时

$$f(\lambda)=\lambda^m,$$

$$f'(\lambda)=m\lambda^{m-1},$$

$$f''(\lambda)=m(m-1)\lambda^{m-2},$$

$$\cdots\cdots$$

$$f^{(l_i-1)}(\lambda)=m(m-1)\cdots(m-l_i)\lambda^{m-l_i+1}$$

(假设 $m\geqslant l_i-1$),

则

$$A^m=PJ^mP^{-1}=P\begin{bmatrix} J_1^m & & & \\ & J_2^m & & \\ & & \ddots & \\ & & & J_k^m \end{bmatrix}P^{-1},$$

其中

$$J_i^m=\begin{bmatrix} \lambda_i^m & C_m^1\lambda_i^{m-1} & C_m^2\lambda_i^{m-2} & \cdots & C_m^{l_i-1}\lambda_i^{m-l_i+1} \\ 0 & \lambda_i^m & C_m^1\lambda_i^{m-1} & \cdots & C_m^{l_i-2}\lambda_i^{m-l_i+2} \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & \lambda_i^m \end{bmatrix}_{l_i\times l_i}$$

$(i=1,2,\cdots,k)$.

例 3 设

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

求 $e^A, e^{At}, \sin A$.

解 先求 A 的 Jordan 标准形. 由于

$$\lambda I - A = \begin{bmatrix} \lambda-2 & 0 & 0 \\ -1 & \lambda-1 & -1 \\ -1 & 1 & \lambda-3 \end{bmatrix}$$

$$\xrightarrow{c_1 \leftrightarrow c_2} \begin{bmatrix} 0 & \lambda-2 & 0 \\ \lambda-1 & -1 & -1 \\ 1 & -1 & \lambda-3 \end{bmatrix}$$

$$\xrightarrow{r_1 \leftrightarrow r_3} \begin{bmatrix} 1 & -1 & \lambda-3 \\ \lambda-1 & -1 & -1 \\ 0 & \lambda-2 & 0 \end{bmatrix}$$

$$\xrightarrow[c_3-(\lambda-3)c_1]{c_2+c_1} \begin{bmatrix} 1 & 0 & 0 \\ \lambda-1 & \lambda-2 & -\lambda^2+4\lambda-4 \\ 0 & \lambda-2 & 0 \end{bmatrix}$$

$$\xrightarrow[r_3-r_2]{r_2-(\lambda-1)r_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda-2 & -\lambda^2+4\lambda-4 \\ 0 & 0 & \lambda^2-4\lambda+4 \end{bmatrix}$$

$$\xrightarrow{r_2+r_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda-2 & 0 \\ 0 & 0 & (\lambda-2)^2 \end{bmatrix},$$

故 $\lambda I - A$ 的不变因子为

$$d_1(\lambda) = 1, \quad d_2(\lambda) = \lambda - 2, \quad d_3(\lambda) = (\lambda-2)^2.$$

从而, A 的初等因子为

$$(\lambda-2)^2, \quad \lambda-2,$$

因此, A 的 Jordan 标准形为

$$J = \begin{bmatrix} J_1 & \\ & J_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

再求相似变换矩阵 P. 令

$$P = (\eta_1, \eta_2, \eta_3),$$

由

$$AP = PJ,$$

可得

$$\eta_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \eta_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \eta_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

故

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

而且

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

(1) 计算 e^A. 此时 $f_1(\lambda)=e^\lambda$，$f_1'(\lambda)=e^\lambda$，故

$$f_1(J_1) = \begin{bmatrix} e^2 & e^2 \\ 0 & e^2 \end{bmatrix}, \quad f_1(J_2) = (e^2),$$

$$e^J = f_1(J) = \begin{bmatrix} f_1(J_1) & \\ & f_1(J_2) \end{bmatrix} = \begin{bmatrix} e^2 & e^2 & 0 \\ 0 & e^2 & 0 \\ 0 & 0 & e^2 \end{bmatrix}.$$

从而

$$\begin{aligned} e^A &= Pe^J P^{-1} \\ &= \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} e^2 & e^2 & 0 \\ 0 & e^2 & 0 \\ 0 & 0 & e^2 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & e^2 & e^2 \\ e^2 & e^2 & 0 \\ e^2 & e^2 & e^2 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \end{aligned}$$

$$=\begin{pmatrix} e^2 & 0 & 0 \\ e^2 & 2e^2 & -e^2 \\ e^2 & e^2 & 0 \end{pmatrix};$$

(2) 计算 e^{At}. 此时，$f_2(\lambda)=e^{\lambda t}$，$f_2'(\lambda)=te^{\lambda t}$，故

$$e^{Jt}=\begin{pmatrix} f_2(J_1) & \\ & f_2(J_2) \end{pmatrix}=\begin{pmatrix} e^{2t} & te^{2t} & 0 \\ 0 & e^{2t} & 0 \\ 0 & 0 & e^{2t} \end{pmatrix},$$

从而

$$e^{At}=Pe^{Jt}P^{-1}=\begin{pmatrix} e^{2t} & 0 & 0 \\ te^{2t} & (1+t)e^{2t} & -te^{2t} \\ te^{2t} & te^{2t} & (1-t)e^{2t} \end{pmatrix};$$

(3) 计算 $\sin A$. 此时，$f_3(\lambda)=\sin\lambda$，$f_3'(\lambda)=\cos\lambda$，故

$$\sin J=f_3(J)=\begin{pmatrix} \sin 2 & \cos 2 & 0 \\ 0 & \sin 2 & 0 \\ 0 & 0 & \sin 2 \end{pmatrix},$$

$$\sin A=P\sin J\cdot P^{-1}=\begin{pmatrix} \sin 2 & 0 & 0 \\ \cos 2 & \sin 2+\cos 2 & -\cos 2 \\ \cos 2 & \cos 2 & \sin 2-\cos 2 \end{pmatrix}.$$

在一般情况下，将 A 化成 Jordan 标准形时，求相似变换矩阵 P 是很麻烦的. 另外，A 的特征多项式也不一定是其最小多项式. 但对友矩阵 A，有下列性质：

性质 2 设友矩阵 A 为

$$A=\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \cdots & & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & -a_{n-3} & \cdots & -a_2 & -a_1 \end{pmatrix}_{n\times n},$$

则：

(1) 当 A 有 n 个不同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 时，Jordan 标准形

J 及相似变换矩阵 P 分别为

$$J=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

$$P=\begin{pmatrix}1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\ \cdots & & & \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1}\end{pmatrix};$$

(2) A 的特征多项式等于 A 的最小多项式，且

$$|\lambda I - A| = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_{n-1}\lambda + a_n.$$

例 4 设

$$A=\begin{pmatrix}0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6\end{pmatrix},$$

求 e^{At}.

解 由于 A 为友矩阵，故

$$\begin{aligned}|\lambda I - A| &= \lambda^3 + 6\lambda^2 + 11\lambda + 6 \\ &= (\lambda+1)(\lambda+2)(\lambda+3),\end{aligned}$$

从而

$$\lambda_1 = -1,\ \lambda_2 = -2,\ \lambda_3 = -3.$$

故

$$P=\begin{pmatrix}1 & 1 & 1 \\ -1 & -2 & -3 \\ 1 & 4 & 9\end{pmatrix},$$

因此

$$P^{-1}=\frac{1}{2}\begin{pmatrix}6 & 5 & 1\\ -6 & -8 & -2\\ 2 & 3 & 1\end{pmatrix}.$$

又

$$J=\begin{pmatrix}-1 & & \\ & -2 & \\ & & -3\end{pmatrix},$$

于是

$$e^{Jt}=\begin{pmatrix}e^{-t} & & \\ & e^{-2t} & \\ & & e^{-3t}\end{pmatrix},$$

$$e^{At}=Pe^{Jt}P^{-1}$$

$$=\frac{1}{2}\begin{pmatrix}6e^{-t}-6e^{-2t}+2e^{-3t} & 5e^{-t}-8e^{-2t}+3e^{-3t} & e^{-t}-2e^{-2t}+e^{-3t}\\ -6e^{-t}+12e^{-2t}-6e^{-3t} & -5e^{-t}+16e^{-2t}-9e^{-3t} & -e^{-t}+4e^{-2t}-3e^{-3t}\\ 6e^{-t}-24e^{-2t}+18e^{-3t} & 5e^{-t}-32e^{-2t}+27e^{-3t} & e^{-t}-8e^{-2t}+9e^{-3t}\end{pmatrix}.$$

Ⅲ. 拉格朗日插值法

定义 2 设矩阵 $A\in C^{n\times n}$的最小多项式为

$$m(\lambda)=(\lambda-\lambda_1)^{l_1}(\lambda-\lambda_2)^{l_2}\cdots(\lambda-\lambda_t)^{l_t},$$

则称集合

$$\{(\lambda_i,l_i)\mid i=1,2,\cdots,t\}$$

为 A 的**谱**,记为 σ_A.

定义 3 设 $A\in C^{n\times n}$,称函数 $f(\lambda)$在 A 的谱 σ_A **上给定**,是指给定了

$$f(\lambda_i),\ f'(\lambda_i),\ f''(\lambda_i),\ \cdots,\ f^{(l_i-1)}(\lambda_i)\qquad (i=1,2,\cdots,t),$$

其中 $\sum_{i=1}^{t}l_i$ 为 A 的最小多项式 $m(\lambda)$的次数. 若 $f(\lambda)$在 σ_A 上给定,则记为 $f(\sigma_A)$.

定理 4 设 $A\in C^{n\times n}$,$g_1(\lambda)$,$g_2(\lambda)$为复系数多项式,则

$$g_1(A)=g_2(A)\Longleftrightarrow g_1(\sigma_A)=g_2(\sigma_A).$$

任何由矩阵级数定义的矩阵函数 $f(A)$，对每一个固定的矩阵 A，由 Cayley-Hamilton 定理易知，$f(A)$与矩阵 A 的某一多项式$g(A)$相等；定理 4 又告诉我们，只要多项式 $g_1(\lambda)$与 $g(\lambda)$在 A 的谱上的值相等，则 $g_1(A)=g(A)$，从而 $f(A)=g(A)=g_1(A)$. 这说明对任何函数 $f(\lambda)$及多项式 $g(\lambda)$，只要 $f(\sigma_A)=g(\sigma_A)$，则 $f(A)=g(A)$. 因此，可以推广定义 1 如下：

定义 4 对任何函数 $f(\lambda)$，若存在复系数多项式 $g(\lambda)$，使得

$$f(\sigma_A) = g(\sigma_A),$$

则定义矩阵函数 $f(A)$为 $g(A)$，即

$$f(A) = g(A).$$

这个定义不仅给出了矩阵函数的较广的定义，同时也指出了计算 $f(A)$的一种新方法，即只须找出与 $f(\lambda)$在 σ_A 上取值相同的多项式 $g(\lambda)$，则 $f(A)=g(A)$. 由此可见，为了方便计算 $f(A)$，就应找出次数最低的多项式 $g(\lambda)$，使得 $g(\sigma_A)=f(\sigma_A)$，如何找呢？下面介绍拉格朗日插值法.

给定函数 $f(\lambda)$以及 n 阶方阵 A，设 A 的最小多项式为

$$m(\lambda) = (\lambda - \lambda_1)^{l_1}(\lambda - \lambda_2)^{l_2}\cdots(\lambda - \lambda_t)^{l_t},$$

其中，$\sum_{i=1}^{t} l_i = m \leqslant n$.

若多项式 $g(\lambda)$使得 $g(\sigma_A)=f(\sigma_A)$，即

$$f(\lambda_i) = g(\lambda_i),\ f'(\lambda_i) = g'(\lambda_i),\ \cdots,\ f^{(l_i-1)}(\lambda_i) = g^{(l_i-1)}(\lambda_i)$$
$$(i = 1,2,\cdots,t),$$

则 $f(A)=g(A)$. 由于在 σ_A 上已知多项式 $g(\lambda)$的上述 m 个值，因此可确定 $g(\lambda)$的 m 个系数，由此可见，$g(\lambda)$最多是 $m-1$ 次多项式. 若用 A 的特征多项式代替它的最小多项式，则 $g(\lambda)$最多也只不过是 $n-1$ 次多项式. 下面介绍确定 $g(\lambda)$的方法.

(1) 设 A 的特征多项式 $\varphi(\lambda)=|\lambda I-A|$ 没有重根，即 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为互不相同的特征根. 此时，A 的最小多项式 $m(\lambda)=\varphi(\lambda)$都是 n 次的，因此 $g(\lambda)$是 $n-1$ 次的，且 $g(\lambda)$在 n 个不同的点 λ_1,λ_2,

$\cdots,\lambda_n$ 上取值是已知的，即

$$f(\lambda_i)=g(\lambda_i)\qquad(i=1,2,\cdots,n).$$

由拉格朗日插值公式得

$$g(\lambda)=\sum_{i=1}^{n}f(\lambda_i)l_i(\lambda),$$

其中

$$l_i(\lambda)=\frac{(\lambda-\lambda_1)\cdots(\lambda-\lambda_{i-1})(\lambda-\lambda_{i+1})\cdots(\lambda-\lambda_n)}{(\lambda_i-\lambda_1)\cdots(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})\cdots(\lambda_i-\lambda_n)}.$$

于是

$$f(A)=g(A)=\sum_{i=1}^{n}f(\lambda_i)l_i(A),$$

其中

$$l_i(A)=\frac{(A-\lambda_1I)\cdots(A-\lambda_{i-1}I)(A-\lambda_{i-1}I)\cdots(A-\lambda_nI)}{(\lambda_i-\lambda_1)\cdots(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})\cdots(\lambda_i-\lambda_n)},$$

$$(i=1,2,\cdots,n).$$

(2) 设 A 的特征多项式 $\varphi(\lambda)=|\lambda I-A|$ 有重根，但最小多项式 $m(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_m)$ 没有重根，则

$$f(A)=\sum_{i=1}^{m}f(\lambda_i)l_i(A),$$

其中

$$l_i(A)=\frac{(A-\lambda_1I)\cdots(A-\lambda_{i-1}I)(A-\lambda_{i+1}I)\cdots(A-\lambda_mI)}{(\lambda_i-\lambda_1)\cdots(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})\cdots(\lambda_i-\lambda_m)},$$

$$(i=1,2,\cdots,m).$$

(3) A 的最小多项式 $m(\lambda)$ 也有重根，设

$$m(\lambda)=(\lambda-\lambda_1)^{l_1}(\lambda-\lambda_2)^{l_2}\cdots(\lambda-\lambda_t)^{l_t},$$

其中，$\sum_{i=1}^{t}l_i=m\leqslant n$. 下面用类似于确定极点处留数的办法来确定 $g(\lambda)$ 的系数.

设 $g(\lambda)$ 是所要求的 $m-1$ 次多项式，因此 $\dfrac{g(\lambda)}{m(\lambda)}$ 为真分式，将

$\frac{g(\lambda)}{m(\lambda)}$分解成下式：

$$\frac{g(\lambda)}{m(\lambda)} = \sum_{i=1}^{t}\left(\frac{a_{i1}}{(\lambda-\lambda_i)^{l_i}} + \frac{a_{i2}}{(\lambda-\lambda_i)^{l_i-1}} + \cdots + \frac{a_{il_i}}{\lambda-\lambda_i}\right).$$

如果定出 a_{ij}，则

$$g(\lambda) = m(\lambda)\sum_{i=1}^{t}\left(\frac{a_{i1}}{(\lambda-\lambda_i)^{l_i}} + \frac{a_{i2}}{(\lambda-\lambda_i)^{l_i-1}} + \cdots + \frac{a_{il_i}}{\lambda-\lambda_i}\right).$$

令

$$\begin{aligned}\varphi_i(\lambda) &= \frac{m(\lambda)}{(\lambda-\lambda_i)^{l_i}} \\ &= (\lambda-\lambda_1)^{l_1}\cdots(\lambda-\lambda_{i-1})^{l_{i-1}}(\lambda-\lambda_{i+1})^{l_{i+1}}\cdots(\lambda-\lambda_t)^{l_t},\end{aligned}$$

则

$$g(\lambda) = \sum_{i=1}^{t}\varphi_i(\lambda)[a_{i1} + a_{i2}(\lambda-\lambda_i) + \cdots + a_{il_i}(\lambda-\lambda_i)^{l_i-1}].$$

现在的问题是如何确定 a_{ij}，$i=1,2,\cdots,t$；$j=1,2,\cdots,l_i$.

先求 a_{i1}. 由于

$$\begin{aligned}(\lambda-\lambda_i)^{l_i}\frac{g(\lambda)}{m(\lambda)} &= [a_{i1} + a_{i2}(\lambda-\lambda_i) + \cdots + a_{il_i}(\lambda-\lambda_i)^{l_i-1}] \\ &\quad + (\lambda-\lambda_i)^{l_i}\sum_{k\neq i}\frac{1}{(\lambda-\lambda_k)^{l_k}}[a_{k1} + a_{k2}(\lambda-\lambda_k) \\ &\quad + \cdots + a_{kl_k}(\lambda-\lambda_k)^{l_k-1}]\end{aligned}$$

则

$$\begin{aligned}a_{i1} &= \lim_{\lambda\to\lambda_i}(\lambda-\lambda_i)^{l_i}\frac{g(\lambda)}{m(\lambda)} \\ &= \lim_{\lambda\to\lambda_i}\frac{g(\lambda)}{(\lambda-\lambda_1)^{l_1}\cdots(\lambda-\lambda_{i-1})^{l_{i-1}}(\lambda-\lambda_{i+1})^{l_{i+1}}\cdots(\lambda-\lambda_t)^{l_t}} \\ &= \frac{g(\lambda_i)}{(\lambda_i-\lambda_1)^{l_1}\cdots(\lambda_i-\lambda_{i-1})^{l_{i-1}}(\lambda_i-\lambda_{i+1})^{l_{i+1}}\cdots(\lambda_i-\lambda_t)^{l_t}} \\ &= \frac{f(\lambda_i)}{(\lambda_i-\lambda_1)^{l_1}\cdots(\lambda_i-\lambda_{i-1})^{l_{i-1}}(\lambda_i-\lambda_{i+1})^{l_{i+1}}\cdots(\lambda_i-\lambda_t)^{l_t}}.\end{aligned}$$

再求 a_{i2}. 此时 $l_i\geqslant 2$，对 $(\lambda-\lambda_i)^{l_i}\frac{g(\lambda)}{m(\lambda)}$ 的表达式两边对 λ 求导，

并令 $\lambda \to \lambda_i$，得

$$a_{i2} = \lim_{\lambda \to \lambda_i} \frac{\mathrm{d}}{\mathrm{d}\lambda}\left[(\lambda - \lambda_i)^{l_i} \frac{g(\lambda)}{m(\lambda)}\right]$$

$$= \frac{\mathrm{d}}{\mathrm{d}\lambda}\left[(\lambda - \lambda_i)^{l_i} \frac{g(\lambda)}{m(\lambda)}\right]\Bigg|_{\lambda=\lambda_i}.$$

由于

$$f^{(k)}(\lambda_i) = g^{(k)}(\lambda_i) \qquad (k = 0,1,2,\cdots,l_i - 1),$$

故

$$a_{i2} = \frac{\mathrm{d}}{\mathrm{d}\lambda}\left[(\lambda - \lambda_i)^{l_i} \frac{f(\lambda)}{m(\lambda)}\right]\Bigg|_{\lambda=\lambda_i}.$$

类似地可求得

$$a_{ij} = \frac{1}{(j-1)!} \frac{\mathrm{d}^{j-1}}{\mathrm{d}\lambda^{j-1}}\left[(\lambda - \lambda_i)^{l_i} \frac{f(\lambda)}{m(\lambda)}\right]\Bigg|_{\lambda=\lambda_i},$$

其中，$i=1,2,\cdots,t$；$j=1,2,\cdots,l_i$.

综上所述，求 $f(A)$的步骤如下：

第一步，求 A 的最小多项式，并分解为如下形式：

$$m(\lambda) = (\lambda - \lambda_1)^{l_1}(\lambda - \lambda_2)^{l_2}\cdots(\lambda - \lambda_t)^{l_t};$$

第二步，求 a_{ij}，$i=1,2,\cdots,t$；$j=1,2,\cdots,l_i$.

$$a_{ij} = \frac{1}{(j-1)!} \frac{\mathrm{d}^{j-1}}{\mathrm{d}\lambda^{j-1}}\left[(\lambda - \lambda_i)^{l_i} \frac{f(\lambda)}{m(\lambda)}\right]\Bigg|_{\lambda=\lambda_i};$$

第三步，计算 $f(A)$.

$$f(A) = g(A) = \sum_{i=1}^{t} \varphi_i(A)[a_{i1}I + a_{i2}(A - \lambda_i I) + \cdots + a_{il_i}(A - \lambda_i I)^{l_i - 1}],$$

其中

$$\varphi_i(A) = (A - \lambda_1 I)^{l_1}(A - \lambda_2 I)^{l_2}\cdots(A - \lambda_{i-1} I)^{l_{i-1}} \cdot (A - \lambda_{i+1} I)^{l_{i+1}}\cdots(A - \lambda_t I)^{l_t} \quad (i = 1,2,\cdots,t).$$

注 (1) $\varphi_i(A)$只与 A 有关，与函数 $f(\lambda)$无关；a_{ij}与 $f(\lambda)$及 A 都有关.

(2) 若用 A 的特征多项式代替它的最小多项式,用上述方法计算 $f(A)$仍有效.

例5 设

$$A=\begin{pmatrix}1 & 3\\ 2 & 2\end{pmatrix},$$

求 e^{At}.

解 设 $f(\lambda)=e^{\lambda t}$. 因为

$$|\lambda I-A|=\begin{vmatrix}\lambda-1 & -3\\ -2 & \lambda-2\end{vmatrix}=(\lambda+1)(\lambda-4),$$

从而 A 的特征根为

$$\lambda_1=-1,\quad \lambda_2=4.$$

按情况(1)的公式可得

$$f(A)=f(\lambda_1)l_1(A)+f(\lambda_2)l_2(A),$$

其中

$$l_1(A)=\frac{A-\lambda_2 I}{\lambda_1-\lambda_2}=\frac{A-4I}{-1-4}=-\frac{1}{5}(A-4I),$$

$$l_2(A)=\frac{A-\lambda_1 I}{\lambda_2-\lambda_1}=\frac{A+I}{4+1}=\frac{1}{5}(A+I),$$

$$f(\lambda_1)=f(-1)=e^{-t},$$

$$f(\lambda_2)=f(4)=e^{4t}.$$

故

$$\begin{aligned}e^{At}&=e^{-t}\cdot\frac{-1}{5}(A-4I)+e^{4t}\cdot\frac{1}{5}(A+I)\\&=-\frac{1}{5}e^{-t}\begin{pmatrix}-3 & 3\\ 2 & -2\end{pmatrix}+\frac{1}{5}e^{4t}\begin{pmatrix}2 & 3\\ 2 & 3\end{pmatrix}\\&=\frac{1}{5}\begin{pmatrix}3e^{-t}+2e^{4t} & -3e^{-t}+3e^{4t}\\ -2e^{-t}+2e^{4t} & 2e^{-t}+3e^{4t}\end{pmatrix}.\end{aligned}$$

例6 设

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix},$$

求 e^{At}.

解 设 $f(\lambda)=e^{\lambda t}$. 易知

$$\varphi(\lambda) = |\lambda I - A| = (\lambda - 2)^3(\lambda + 2),$$
$$m(\lambda) = (\lambda - 2)(\lambda + 2).$$

显然,$\varphi(\lambda)$有重根,而 $m(\lambda)$无重根,按情况(2)可知

$$e^{At} = f(A) = f(\lambda_1)l_1(A) + f(\lambda_2)l_2(A),$$

其中,$\lambda_1=2$,$\lambda_2=-2$. 又

$$l_1(A) = \frac{A - \lambda_2 I}{\lambda_1 - \lambda_2} = \frac{1}{4}(A + 2I),$$

$$l_2(A) = \frac{A - \lambda_1 I}{\lambda_2 - \lambda_1} = \frac{-1}{4}(A - 2I),$$

从而

$$\begin{aligned} e^{At} &= f(A) \\ &= \frac{1}{4}e^{2t}(A + 2I) - \frac{1}{4}e^{-2t}(A - 2I) \\ &= \frac{1}{2}\begin{pmatrix} \mathrm{sh}2t + 2\mathrm{ch}2t & \mathrm{sh}2t & \mathrm{sh}2t & \mathrm{sh}2t \\ \mathrm{sh}2t & \mathrm{sh}2t + 2\mathrm{ch}2t & -\mathrm{sh}2t & -\mathrm{sh}2t \\ \mathrm{sh}2t & -\mathrm{sh}2t & \mathrm{sh}2t + 2\mathrm{ch}2t & -\mathrm{sh}2t \\ \mathrm{sh}2t & -\mathrm{sh}2t & -\mathrm{sh}2t & \mathrm{sh}2t + 2\mathrm{ch}2t \end{pmatrix}. \end{aligned}$$

若用特征多项式 $\varphi(\lambda)$代替最小多项式 $m(\lambda)$求插值多项式,由于 $\varphi(\lambda)$次数较高,因而计算较繁. 事实上,$\lambda_1=2$和 $\lambda_2=-2$ 分别是 $\varphi(\lambda)$的三重根和单根,按情况(3)可解例 6 如下:

$$a_{11} = \left.\frac{(\lambda - \lambda_1)f(\lambda)}{\varphi(\lambda)}\right|_{\lambda=\lambda_1} = \left.\frac{e^{\lambda t}}{\lambda + 2}\right|_{\lambda=2} = \frac{e^{2t}}{4},$$

$$a_{12} = \frac{1}{(2-1)!}\frac{d^{2-1}}{d\lambda^{2-1}}\left[\frac{(\lambda - \lambda_1)^3 f(\lambda)}{\varphi(\lambda)}\right]\Bigg|_{\lambda=\lambda_1}$$

$$= \frac{d}{d\lambda}\left(\frac{e^{\lambda t}}{\lambda+2}\right)\bigg|_{\lambda=2} = \frac{1}{16}(4t-1)e^{2t},$$

$$a_{13} = \frac{1}{2!}\frac{d^2}{d\lambda^2}\left((\lambda-\lambda_1)^3\frac{f(\lambda)}{\varphi(\lambda)}\right)\bigg|_{\lambda=\lambda_1}$$

$$= \frac{1}{2}\frac{d^2}{d\lambda^2}\left(\frac{e^{\lambda t}}{\lambda+2}\right)\bigg|_{\lambda=2}$$

$$= \frac{1}{64}(8t^2-4t-1)e^{2t},$$

$$a_{21} = \frac{(\lambda-\lambda_2)f(\lambda)}{\varphi(\lambda)}\bigg|_{\lambda=\lambda_2} = \frac{e^{\lambda t}}{(\lambda-2)^3}\bigg|_{\lambda=-2}$$

$$= \frac{1}{64}e^{-2t},$$

$$\varphi_1(\lambda) = \frac{\varphi(\lambda)}{(\lambda-\lambda_1)^3} = \frac{(\lambda-2)^3(\lambda+2)}{(\lambda-2)^3} = \lambda+2,$$

$$\varphi_2(\lambda) = \frac{\varphi(\lambda)}{\lambda-\lambda_2} = \frac{(\lambda-2)^3(\lambda+2)}{\lambda+2} = (\lambda-2)^3,$$

故

$$\varphi_1(A) = A+2I,\quad \varphi_2(A) = (A-2I)^3,$$

从而

$$f(A) = e^{At}$$
$$= \varphi_1(A)[a_{11}I + a_{12}(A-2I) + a_{13}(A-2I)^2]$$
$$+ a_{21}\varphi_2(A).$$

将 $a_{11}, a_{21}, a_{12}, a_{13}, \varphi_1(A), \varphi_2(A)$ 代入上式，并利用

$$m(A) = (A-2I)(A+2I) = 0$$

化简，即得

$$e^{At} = \frac{1}{4}e^{2t}(A+2I) - \frac{1}{4}e^{-2t}(A-2I).$$

Ⅳ. 待定系数法

按矩阵函数的定义，只须求出多项式 $g(\lambda)$，使得 $f(\sigma_A) = g(\sigma_A)$. 由于 $f(\lambda)$ 在 σ_A 上给定，从而确定了 m 个条件，因此，可用这 m 个条件确定 $g(\lambda)$ 的系数. 即令

$$g(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_{m-1}\lambda^{m-1},$$

其中，m 为 A 的最小多项式的次数. 由条件 $f(\sigma_A)=g(\sigma_A)$列出方程组，解出 $a_0,a_1,a_2,\cdots,a_{m-1}$，从而求出 $g(\lambda)$，进而计算 $f(A)=g(A)$.

例 7 设

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

求 e^{At}.

解 容易算得

$$\varphi(\lambda)=|\lambda I-A|=(\lambda-1)^2(\lambda-2),$$
$$m(\lambda)=(\lambda-1)(\lambda-2).$$

由于 $m(\lambda)$是 2 次多项式，且 $\lambda_1=1,\lambda_2=2$，故 $g(\lambda)$是 1 次多项式. 设

$$g(\lambda)=a_0+a_1\lambda,$$

由于 $f(\lambda)=e^{\lambda t}$，且 $f(\lambda_1)=g(\lambda_1)$，$f(\lambda_2)=g(\lambda_2)$，故

$$\begin{cases}e^t=a_0+a_1,\\ e^{2t}=a_0+2a_1,\end{cases}$$

于是解得

$$\begin{cases}a_0=2e^t-e^{2t},\\ a_1=e^{2t}-e^t.\end{cases}$$

从而

$$\begin{aligned}f(A)=e^{At}=g(A)&=a_0I+a_1A\\ &=(2e^t-e^{2t})I+(e^{2t}-e^t)A\\ &=\begin{bmatrix}2e^t-e^{2t} & 0 & 2e^t-2e^{2t}\\ 0 & e^t & 0\\ e^{2t}-e^t & 0 & 2e^{2t}-e^t\end{bmatrix}.\end{aligned}$$

例 8 设

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

求 e^{At}.

解 由于

$$\varphi(\lambda)=|\lambda I-A|=(\lambda-2)(\lambda+1)^2,$$

且$(\lambda-2)(\lambda+1)$不是 A 的零化多项式，故 A 的最小多项式 $m(\lambda)$ 为

$$m(\lambda)=\varphi(\lambda)=(\lambda-2)(\lambda+1)^2.$$

从而 $g(\lambda)$为 2 次多项式，设

$$g(\lambda)=a_0+a_1\lambda+a_2\lambda^2,$$

由于 $f(\lambda)=e^{\lambda t}$，又 $m(\lambda)$的根 $\lambda_1=2$ 是单根，$\lambda_2=-1$ 是二重根，由 $f(\sigma_A)=g(\sigma_A)$得

$$\begin{cases}f(\lambda_1)=g(\lambda_1),\\f(\lambda_2)=g(\lambda_2),\\f'(\lambda_2)=g'(\lambda_2),\end{cases}$$

即

$$\begin{cases}e^{2t}=a_0+2a_1+4a_2,\\e^{-t}=a_0-a_1+a_2,\\te^{-t}=a_1-2a_2,\end{cases}$$

解得

$$\begin{cases}a_0=\dfrac{1}{9}(e^{2t}+8e^{-t}+6te^{-t}),\\a_1=\dfrac{1}{9}(2e^{2t}-2e^{-t}+3te^{-t}),\\a_2=\dfrac{1}{9}(e^{2t}-e^{-t}-3te^{-t}).\end{cases}$$

从而

$$e^{At}=f(A)=g(A)$$

$$= a_0 I + a_1 A + a_2 A^2$$

$$= \frac{1}{9}\begin{bmatrix} e^{2t} + (8+6t)e^{-t} & 2e^{2t} - (2-3t)e^{-t} & e^{2t} - (1+3t)e^{-t} \\ 2e^{2t} - (2+6t)e^{-t} & 4e^{2t} + (5-3t)e^{-t} & 2e^{2t} - (2-3t)e^{-t} \\ 4e^{2t} + (4+6t)e^{-t} & 8e^{2t} + (3t-8)e^{-t} & 4e^{2t} + (5-3t)e^{-t} \end{bmatrix}.$$

习　题　三

1. 证明下列问题：

(1) 若矩阵序列$\{A_m\}$收敛于 A，则$\{A_m^T\}$收敛于 A^T，$\{\overline{A}_m\}$收敛于 $\overline{A}$；

(2) 若方阵级数 $\sum\limits_{m=0}^{\infty} c_m A^m$ 收敛，则

$$\left(\sum_{m=0}^{\infty} c_m A^m\right)^T = \sum_{m=0}^{\infty} c_m (A^T)^m.$$

2. 已知方阵序列$\{A_m\}$收敛于 A，且 A_m^{-1} 及 A^{-1}都存在，证明：

(1) $\lim\limits_{m\to\infty} |A_m| = |A|$；

(2) $\lim\limits_{m\to\infty} A_m^{-1} = A^{-1}$.

3. 设函数矩阵

$$A(t) = \begin{bmatrix} \sin t & \cos t & t \\ \dfrac{\sin t}{t} & e^t & t^2 \\ 1 & 0 & t^3 \end{bmatrix},$$

其中 $t \neq 0$，计算 $\lim\limits_{t\to 0} A(t)$、$\dfrac{d}{dt}A(t)$、$\dfrac{d^2}{dt^2}A(t)$、$\dfrac{d}{dt}|A(t)|$，$\left|\dfrac{d}{dt}A(t)\right|$.

4. 设函数矩阵

$$A(x) = \begin{bmatrix} e^{2x} & xe^{x} & x^2 \\ e^{-x} & 2e^{2x} & 0 \\ 3x & 0 & 0 \end{bmatrix},$$

计算 $\int_0^1 A(x)dx$ 和 $\dfrac{d}{dx}\left(\int_0^{x^2} A(t)dt\right)$.

5. 设 $y=(y_1(t),y_2(t),\cdots,y_n(t))^T$，$A$ 为 n 阶常数对称矩阵，$f(y)=y^TAy$. 证明：

(1) $\dfrac{\mathrm{d}f}{\mathrm{d}y}=2y^TA\dfrac{\mathrm{d}y}{\mathrm{d}t}$；

(2) $\dfrac{\mathrm{d}}{\mathrm{d}t}\|y\|_2^2=2y^T\dfrac{\mathrm{d}y}{\mathrm{d}t}$.

6. 证明关于迹的下列公式：

(1) $\dfrac{\mathrm{d}}{\mathrm{d}X}tr(XX^T)=\dfrac{\mathrm{d}}{\mathrm{d}X}tr(X^TX)=2X$；

(2) $\dfrac{\mathrm{d}}{\mathrm{d}X}tr(BX)=\dfrac{\mathrm{d}}{\mathrm{d}X}(X^TB^T)=B^T$；

(3) $\dfrac{\mathrm{d}}{\mathrm{d}X}tr(X^TAX)=(A+A^T)X$，

其中

$$X=(x_{ij})_{m\times n},\quad B=(b_{ij})_{n\times m},\quad A=(a_{ij})_{m\times m}.$$

7. 证明：

$$\frac{\mathrm{d}}{\mathrm{d}X^T}(a^Tb)=b^T\frac{\mathrm{d}a}{\mathrm{d}X^T}+a^T\frac{\mathrm{d}b}{\mathrm{d}X^T},$$

其中，$a(X)$，$b(X)$ 为向量函数.

8. 在 R^2 中将向量 $(x_1,x_2)^T$ 表示成平面上直角坐标系 x_1，x_2 中的点 $(x_1,x_2)^T$，分别画出下列不等式决定的向量 $x=(x_1,x_2)^T$ 全体所对应的几何图形：

(1) $\|x\|_1\leqslant 1$；

(2) $\|x\|_2\leqslant 1$；

(3) $\|x\|_\infty\leqslant 1$.

9. 证明：对任何 $x,y\in C^n$，总有

$$\bar{x}^Ty+\bar{y}^Tx=\frac{1}{2}(\|x-y\|_2^2-\|x-y\|_2^2).$$

10. 证明：对 $x\in C^n$，有

$$\|x\|_\infty\leqslant\|x\|_2\leqslant\|x\|_1.$$

11. 设 $a_1,a_2,\cdots,a_n$ 是正实数，证明：对任意 $X=(x_1,x_2,\cdots,$

$x_n)^T \in C^n$，

$$\| X \| = \left(\sum_{i=1}^{n} a_i |x_i|^2 \right)^{\frac{1}{2}}$$

是 C^n 中的向量范数.

12. 证明：

$$\| A \| = n \max_{1 \leqslant i,j \leqslant n} |a_{ij}|$$

是矩阵 $A=(a_{ij})_{n\times n}$的范数，并且与向量的 1-范数相容.

13. 设 $A \in C^{n\times n}$，且 A 可逆，证明：

$$\| A^{-1} \| \geqslant \| A \|^{-1}.$$

14. 设 $A \in C^{n\times n}$，且 $\| A \| < 1$，证明：$I-A$ 可逆，而且有

(1) $\| (I-A)^{-1} \| \leqslant \dfrac{1}{1-\| A \|}$；

(2) $\| (I-A)^{-1}-I \| \leqslant \dfrac{\| A \|}{1-\| A \|}$.

15. 设 $A,B \in C^{n\times n}$，$k,l \in C$，证明：

(1) $e^{kA}e^{lA}=e^{(k+l)A}$，特别地，$(e^A)^{-1}=e^{-A}$；

(2) 当 $AB=BA$ 时，$e^A e^B=e^B e^A=e^{A+B}$；

(3) $\dfrac{d}{dt}e^{At}=Ae^{At}=e^{At}A$；

(4) $\sin(A\pm B)=\sin A\cos B\pm\cos A\sin B$.

16. 求下列三类矩阵的矩阵函数 $\cos A$，$\sin A$，e^{A^2}：

(1) 当 A 为幂等矩阵($A^2=A$)时；

(2) 当 A 为对合矩阵($A^2=I$)时；

(3) 当 A 为幂零矩阵($A^2=0$)时.

17. 若矩阵 A 的特征值的实部全为负，则

$$\lim_{t\to+\infty} e^{At} = 0.$$

18. 计算 e^{At} 和 $\sin At$，其中：

(1) $A=\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$；

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

(3) $A=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ -6 & -11 & -6\end{bmatrix}$.

19. 计算下列矩阵函数：

(1) $A=\begin{bmatrix}2 & 2 & 1\\ 1 & 3 & 1\\ 1 & 2 & 2\end{bmatrix}$，求 A^{1000}；

(2) $A=\begin{bmatrix}4 & 2 & -5\\ 6 & 4 & -9\\ 5 & 3 & -7\end{bmatrix}$，求 e^A；

(3) $A=\begin{pmatrix}0 & -1\\ 4 & 4\end{pmatrix}$，求 $\arcsin\dfrac{A}{4}$；

(4) $A=\begin{pmatrix}16 & 8\\ 8 & 4\end{pmatrix}$，求 $(I+A)^{-1}$ 及 $A^{\frac{1}{2}}$.

20. 证明：

$$\sin^2 A+\cos^2 A=I,$$
$$e^{A+2\pi i I}=e^A,$$

其中 A 为任意方阵.

21. 若 A 是反实对称(反 Hermite)矩阵，则 e^A 为实正交(酉)矩阵.

22. 若 A 是 Hermite 矩阵，则 e^{Ai} 是酉矩阵，并说明当 $n=1$ 时，此结论的意义.

23. 将下列矩阵函数表示成矩阵幂级数，并说明对 A 的限制：

(1) $\mathrm{sh}A$；

(2) $\ln(I+A)$；

(3) $\mathrm{arctg}A$.

24. 设 $A\in C^{m\times n}$，证明：

(1) $|e^{A}|=e^{tr(A)}$;

(2) $\| e^{A} \| \leqslant e^{\| A \|}$.

25. 设 $A\in C^{n\times n}$，且 A 可逆. 若 λ 是 A 的任一特征值，则

$$\frac{1}{\| A^{-1} \|_2} \leqslant |\lambda| \leqslant \| A \|_2.$$

第四章 矩阵微分方程

利用矩阵表示线性微分方程的定解问题，形式比较简单，而矩阵函数又使线性微分方程的求解问题得到简化. 不仅如此，矩阵微分方程还是系统工程和控制理论的重要数学基础. 本章先讨论线性定常系统状态方程的求解，再讨论线性时变系统的状态方程及状态转移矩阵，并给出其级数解法.

§4.1 线性定常系统的状态方程

线性定常系统问题常通过一阶常系数线性方程组的定解问题来描述.

1. 线性常系数齐次微分方程组的解

设有一阶线性常系数齐次微分方程组

$$\begin{cases} \dfrac{\mathrm{d}x_1}{\mathrm{d}t} = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\ \dfrac{\mathrm{d}x_2}{\mathrm{d}t} = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\ \cdots\cdots \\ \dfrac{\mathrm{d}x_n}{\mathrm{d}t} = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n, \end{cases}$$

其中，$x_i = x_i(t)$是自变量 t 的函数，$a_{ij} \in C$，$i, j = 1, 2, \cdots, n$. 设

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix},$$

则上述方程组写成矩阵形式为

$$\frac{\mathrm{d}x}{\mathrm{d}t} = Ax.$$

若未知函数 $x(t)$ 不是列向量，而是 $n \times m$ 矩阵

$$x(t) = \begin{bmatrix} x_{11}(t) & x_{12}(t) & \cdots & x_{1m}(t) \\ x_{21}(t) & x_{22}(t) & \cdots & x_{2m}(t) \\ \cdots & & & \\ x_{n1}(t) & x_{n2}(t) & \cdots & x_{nm}(t) \end{bmatrix},$$

则方程

$$\frac{\mathrm{d}x}{\mathrm{d}t} = Ax$$

就是 $n \times m$ 个未知函数的线性微分方程组. 下面将讨论该方程组满足初始条件

$$x(t)\Big|_{t=t_0} = x(t_0) = \begin{bmatrix} x_{11}(t_0) & x_{12}(t_0) & \cdots & x_{1m}(t_0) \\ x_{21}(t_0) & x_{22}(t_0) & \cdots & x_{2m}(t_0) \\ \cdots & & & \\ x_{n1}(t_0) & x_{n2}(t_0) & \cdots & x_{nm}(t_0) \end{bmatrix}$$

的定解问题.

定理 1 设定解问题为：

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = Ax, \\ x(t)|_{t=t_0} = x(t_0), \end{cases} \tag{4.1}$$

其中，$x(t)$ 是 t 的可微函数的 $n \times m$ 矩阵，$x(t_0)$ 是 $n \times m$ 阶常数矩阵，A 是给定的 n 阶常数方阵，则

(1) 定解问题(4.1)的解为

$$x(t) = \mathrm{e}^{A(t-t_0)}x(t_0),$$

并且这个解是唯一的；

(2) 解 $x(t)$ 的秩与 t 的取值无关.

证 (1) 因为

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{e}^{A(t-t_0)}x(t_0)]$$
$$= A\mathrm{e}^{A(t-t_0)}x(t_0)$$
$$= Ax(t),$$

且

$$x(t)|_{t=t_0} = [\mathrm{e}^{A(t-t_0)}x(t_0)]|_{t=t_0}$$
$$= Ix(t_0) = x(t_0),$$

因此 $x(t) = \mathrm{e}^{A(t-t_0)}x(t_0)$ 是(4.1)式的解.

若 $v(t)$也是(4.1)式的解,令

$$y(t) = \mathrm{e}^{-At}v(t),$$

则

$$\frac{\mathrm{d}y(t)}{\mathrm{d}t} = \mathrm{e}^{-At}(-A)v(t) + \mathrm{e}^{-At}\frac{\mathrm{d}v(t)}{\mathrm{d}t}$$
$$= -\mathrm{e}^{-At}Av(t) + \mathrm{e}^{-At}Av(t)$$
$$= 0,$$

故对$\forall\, t \in C$,有

$$y(t) = k,$$

从而

$$y(t_0) = k,$$

由初始条件知

$$y(t_0) = \mathrm{e}^{-At_0}v(t_0) = \mathrm{e}^{-At_0}x(t_0),$$

故

$$\mathrm{e}^{-At_0}x(t_0) = k.$$

从而

$$y(t) = \mathrm{e}^{-At}v(t) = k = \mathrm{e}^{-At_0}x(t_0),$$

因此

$$v(t) = \mathrm{e}^{A(t-t_0)}x(t_0),$$

即(4.1)的解是唯一的.

(2) 由于$\mathrm{e}^{A(t-t_0)}$对于任何的t都是可逆矩阵,因而$x(t)$的秩与

$x(t_0)$的秩相同，故 $x(t)$的秩与 t 的取值无关.

这个定理表明，线性齐次方程组的定解问题(4.1)与普通齐次微分方程的定解问题具有相同形式的解. 但矩阵乘法不满足交换律，因此要注意初始矩阵 $x(t_0)$是右乘.

2. 线性常系数非齐次微分方程组的解

设 $A=(a_{ij})_{n\times n}$ 与 $B=(b_{ij})_{n\times m}$ 是常数矩阵，而

$$x(t)=\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix},\quad u(t)=\begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_m(t) \end{bmatrix}$$

都是函数向量，其中 $u_1(t),u_2(t),\cdots,u_m(t)$是已知函数，则称

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}=Ax(t)+Bu(t)$$

为**线性常系数非齐次微分方程组**. 下面考虑该方程组的定解问题：

$$\begin{cases} \dfrac{\mathrm{d}x(t)}{\mathrm{d}t}=Ax(t)+Bu(t), \\ x(t)|_{t=t_0}=x(t_0). \end{cases} \tag{4.2}$$

设非齐次方程组(4.2)对应的齐次方程组为

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}=Ax(t),$$

则它的解为

$$x(t)=\mathrm{e}^{At}c,$$

其中 c 为任意常数列向量. 与普通微分方程情况一样，下面用常数变易法来求非齐次方程组的解. 设

$$x(t)=\mathrm{e}^{At}c(t)$$

为(4.2)式的解，则

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}=A\mathrm{e}^{At}c(t)+\mathrm{e}^{At}c'(t),$$

代入(4.2)式得

$$c'(t)=\mathrm{e}^{-At}Bu(t),$$

故

$$c(t)=\int_{t_0}^{t}e^{-At}Bu(t)\mathrm{d}t+c,$$

从而(4.2)式的解为

$$x(t)=e^{At}\int_{t_0}^{t}e^{-At}Bu(t)\mathrm{d}t+e^{At}c,$$

其中 c 为任意常数列向量.为了满足初始条件,应有

$$x(t_0)=0+e^{At_0}c,$$

因此

$$c=e^{-At_0}x(t_0),$$

从而定解问题(4.2)的定解为

$$x(t)=e^{A(t-t_0)}x(t_0)+\int_{t_0}^{t}e^{A(t-v)}Bu(v)\mathrm{d}v.$$

3. n 阶常系数微分方程的解

设 $a_1,a_2,\cdots,a_n$ 为常数,$u(t)$为已知函数,称

$$y^{(n)}+a_1y^{(n-1)}+a_2y^{(n-2)}+\cdots+a_ny=u(t)$$

为 n 阶常系数微分方程.当 $u(t)\neq 0$ 时,称为非齐次的;否则,称为齐次的.

由于常系数线性微分方程组的矩阵形式解已经得到,因此,人们自然想把高阶微分方程化为方程组来求解.下面先考虑 n 阶常系数线性齐次方程的定解问题:

$$\begin{cases}y^{(n)}+a_1y^{(n-1)}+a_2y^{(n-2)}+\cdots+a_ny=0,\\ y^{(i)}(t)\big|_{t=0}=y_0^{(i)},\qquad i=0,1,\cdots,n-1.\end{cases}\tag{4.3}$$

令

$$\begin{cases}x_1=y,\\ x_2=y'=x_1',\\ \cdots\cdots\\ x_n=y^{(n-1)}=x_{n-1}',\end{cases}$$

从而

$$\begin{cases} x_1' = x_2, \\ x_2' = x_3, \\ \cdots\cdots \\ x_{n-1}' = x_n, \\ x_n' = -a_n x_1 - a_{n-1} x_2 - \cdots - a_1 x_n. \end{cases}$$

令

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix},$$

$$x(0) = \begin{pmatrix} x_1(0) \\ x_2(0) \\ \vdots \\ x_n(0) \end{pmatrix} = \begin{pmatrix} y_0 \\ y_0' \\ \vdots \\ y_0^{(n-1)} \end{pmatrix},$$

则定解问题(4.3)可写成

$$\begin{cases} \dfrac{\mathrm{d}x(t)}{\mathrm{d}t} = Ax(t), \\ x(t)\big|_{t=0} = x(0), \end{cases} \tag{4.4}$$

其中

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{pmatrix},$$

显然,A 为友矩阵.

由定理 1 得,定解问题(4.4)的解为

$$x(t) = e^{At}x(0).$$

由于定解问题(4.3)的解是(4.4)式的解的第一个分量,从而

定解问题(4.3)的解为

$$
\begin{aligned}
y &= (1,0,0,\cdots,0)x(t) \\
&= (1,0,0,\cdots,0)\mathrm{e}^{At}x(0) \\
&= (1,0,0,\cdots,0)\mathrm{e}^{At}\begin{pmatrix} y_0 \\ y_0' \\ \vdots \\ y_0^{(n-1)} \end{pmatrix}.
\end{aligned}
$$

对于 n 阶常系数线性非齐次方程的定解问题

$$
\begin{cases} y^{(n)} + a_1 y^{(n-1)} + a_2 y^{(n-2)} + \cdots + a_n y = u(t), \\ y^{(i)}(t)\big|_{t=0} = y_0^{(i)}, \qquad i = 0,1,\cdots,n-1, \end{cases} \tag{4.5}
$$

可作类似讨论,进而得到定解问题(4.5)的解是方程组

$$
\begin{cases} \dfrac{\mathrm{d}x(t)}{\mathrm{d}t} = Ax(t) + Bu(t), \\ x(t)\big|_{t=0} = x(0) \end{cases}
$$

的解的第一个分量,其中

$$
x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} = \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix},
$$

$$
x(0) = \begin{pmatrix} x_1(0) \\ x_2(0) \\ \vdots \\ x_n(0) \end{pmatrix} = \begin{pmatrix} y_0 \\ y_0' \\ \vdots \\ y_0^{(n-1)} \end{pmatrix},
$$

$$
A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.
$$

由定解问题(4.2)知

$$x(t) = e^{At}x(0) + \int_0^t e^{A(t-v)}Bu(v)dv,$$

从而定解问题(4.5)的解为

$$y(t) = (1,0,\cdots,0)\left(e^{At}x(0) + \int_0^t e^{A(t-v)}Bu(v)dv\right).$$

由上述讨论可知，不论是一阶常系数线性微方程组，还是 n 阶常系数线性微分方程、求定解问题的解的关键在于计算矩阵函数 e^{At}，而计算矩阵函数 e^{At} 是第三章的基本内容，从而求常系数一阶线性微分方程组及 n 阶线性微分方程的定解问题的解的问题已经解决. 作为例子，下面考虑一种特殊情况，即矩阵 A 为友矩阵且有 n 个不同的特征值的情况.

例 1 设矩阵 A 为友矩阵，且有 n 个不同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$，求定解问题

$$\begin{cases} \dfrac{dx(t)}{dt} = Ax(t) + Bu(t), & (4.6) \\ x(t)|_{t=0} = x(0) & (4.7) \end{cases}$$

的解.

解 方法 1 由友矩阵的性质知，矩阵 A 的 Jordan 标准形 J 与相似变换矩阵 P 分别为

$$J = P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

$$P = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\ \cdots & & & \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix}.$$

令

$$x = Pz,$$

则

$$\frac{\mathrm{d}x}{\mathrm{d}t} = P\frac{\mathrm{d}z}{\mathrm{d}t},$$

从而方程组(4.6)变为

$$P\frac{\mathrm{d}z}{\mathrm{d}t} = APz + Bu(t),$$

因此

$$\frac{\mathrm{d}z}{\mathrm{d}t} = P^{-1}APz + P^{-1}Bu(t),$$

即

$$\begin{pmatrix} \frac{\mathrm{d}z_1}{\mathrm{d}t} \\ \frac{\mathrm{d}z_2}{\mathrm{d}t} \\ \vdots \\ \frac{\mathrm{d}z_n}{\mathrm{d}t} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} + P^{-1} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} u(t).$$

设

$$P^{-1}\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} u(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_n(t) \end{pmatrix},$$

则上述方程组可改写为

$$\begin{cases} \frac{\mathrm{d}z_1}{\mathrm{d}t} = \lambda_1 z_1 + u_1(t), \\ \frac{\mathrm{d}z_2}{\mathrm{d}t} = \lambda_2 z_2 + u_2(t), \\ \cdots\cdots \\ \frac{\mathrm{d}z_n}{\mathrm{d}t} = \lambda_n z_n + u_n(t). \end{cases}$$

显然，上面 n 个方程都是一阶线性微分方程，因此. 可以求出 z_1，z_2，…，z_n，从而求出向量 z，并由 $x=Pz$ 得到方程组(4.6)的解，再由初始条件(4.7)确定定解问题的解.

方法 2　由于 A 的特征值互不相同，利用拉格朗日插值公式求解也是比较简单的. 事实上，定解问题的解为

$$x(t)=e^{At}x(0)+\int_0^t e^{A(t-v)}Bu(v)dv,$$

因此，问题归结为求 e^{At}，按拉格朗日公式(也可按 Jordan 标准形等方法)求 e^{At}，有

$$e^{At}=\sum_{i=1}^{n}e^{\lambda_i t}l_i(A),$$

$$e^{A(t-v)}=\sum_{i=1}^{n}e^{(t-v)\lambda_i}l_i(A),$$

其中

$$l_i(A)=\frac{(A-\lambda_1 I)\cdots(A-\lambda_{i-1}I)(A-\lambda_{i+1}I)\cdots(A-\lambda_n I)}{(\lambda_i-\lambda_1)\cdots(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})\cdots(\lambda_i-\lambda_n)},$$

$$(i=1,2,\cdots,n).$$

故定解问题的解为

$$x(t)=\sum_{i=1}^{n}e^{\lambda_i t}l_i(A)x(0)+\sum_{i=1}^{n}[e^{\lambda_i t}\int_0^t e^{-\lambda_i v}l_i(A)Bu(v)dv].$$

例 2　求常系数线性齐次微分方程组

$$\begin{cases}\dfrac{dy_1(t)}{dt}=2y_1+2y_2-y_3,\\[2mm]\dfrac{dy_2(t)}{dt}=-y_1-y_2+y_3,\\[2mm]\dfrac{dy_3(t)}{dt}=-y_1-2y_2+2y_3\end{cases}$$

在初始条件

$$y(0)=\begin{pmatrix}y_1(0)\\y_2(0)\\y_3(0)\end{pmatrix}=\begin{pmatrix}1\\1\\3\end{pmatrix}$$

下的解.

解 系数矩阵为

$$A=\begin{pmatrix}2 & 2 & -1\\ -1 & -1 & 1\\ -1 & -2 & 2\end{pmatrix},$$

从而定解问题的解为

$$y(t)=\mathrm{e}^{At}y(0),$$

下面求 e^{At}.

方法 1 易求出 A 的 Jordan 标准形及相似变换矩阵分别为

$$J=\begin{pmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix},\quad P=\begin{pmatrix}1 & 1 & 1\\ -1 & 0 & 0\\ -1 & 0 & 1\end{pmatrix},$$

从而

$$\mathrm{e}^{At}=P\mathrm{e}^{Jt}P^{-1}=P\begin{pmatrix}\mathrm{e}^t & t\mathrm{e}^t & 0\\ 0 & \mathrm{e}^t & 0\\ 0 & 0 & \mathrm{e}^t\end{pmatrix}P^{-1}.$$

故定解问题的解为

$$y(t)=\mathrm{e}^{At}y(0)=\begin{pmatrix}\mathrm{e}^t\\ \mathrm{e}^t\\ 3\mathrm{e}^t\end{pmatrix}.$$

方法 2 用拉格朗日插值公式求 e^{At}. 由于 A 的最小多项式 $m(\lambda)=(\lambda-1)^2$, 因此可设多项式 $g(\lambda)$ 为

$$g(\lambda)=a_0+a_1\lambda,$$

由

$$\begin{cases}f(1)=g(1),\\ f'(1)=g'(1),\end{cases}$$

得

$$\begin{cases}f(1)=\mathrm{e}^{\lambda t}|_{\lambda=1}=\mathrm{e}^t=a_0+a_1=g(1),\\ f'(1)=t\mathrm{e}^{\lambda t}|_{\lambda=1}=t\mathrm{e}^t=a_1=g'(1),\end{cases}$$

解得

$$\begin{cases} a_0 = e^t - te^t, \\ a_1 = te^t, \end{cases}$$

从而

$$g(\lambda) = (1 - t)e^t + te^t\lambda,$$

故

$$\begin{aligned} f(A) = g(A) &= (1 - t)e^tI + te^tA \\ &= e^t\begin{bmatrix} 1 + t & 2t & -t \\ -t & -2t + 1 & t \\ -t & -2t & t + 1 \end{bmatrix}, \end{aligned}$$

因此定解问题的解为

$$y(t) = e^{At}y(0) = \begin{bmatrix} e^t \\ e^t \\ 3e^t \end{bmatrix}.$$

例 3 求常系数线性非齐次方程组

$$\begin{cases} \dfrac{dy_1(t)}{dt} = 2y_1 - y_2 + y_3 + e^{2t}, \\ \dfrac{dy_2(t)}{dt} = 3y_2 - y_3, \\ \dfrac{dy_3(t)}{dt} = 2y_1 + y_2 + 3y_3 + te^{2t}, \end{cases}$$

在初始条件

$$y(0) = \begin{bmatrix} y_1(0) \\ y_2(0) \\ y_3(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

下的解.

解 将方程组写成向量方程

$$\frac{dy(t)}{dt} = Ay(t) + u(t),$$

其中

$$A=\begin{pmatrix}2 & -1 & 1\\ 0 & 3 & -1\\ 2 & 1 & 3\end{pmatrix},$$

$$y(t)=\begin{pmatrix}y_1(t)\\ y_2(t)\\ y_3(t)\end{pmatrix},\quad u(t)=\begin{pmatrix}e^{2t}\\ 0\\ te^{2t}\end{pmatrix}.$$

由非齐次方程组的定解问题公式得

$$y(t)=e^{At}y(0)+\int_0^t e^{A(t-\tau)}u(\tau)d\tau.$$

由于 A 的 Jordan 标准形 J 及相似变换矩阵 P 分别为

$$J=\begin{pmatrix}2 & 1 & 0\\ 0 & 2 & 0\\ 0 & 0 & 4\end{pmatrix},$$

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

而且

$$P^{-1}=\begin{pmatrix}-\frac{1}{2} & 0 & \frac{1}{2}\\ 1 & 1 & 0\\ \frac{1}{2} & 0 & \frac{1}{2}\end{pmatrix},$$

于是

$$\begin{aligned}e^{At}y(0)&=Pe^{Jt}P^{-1}y(0)\\ &=\begin{pmatrix}-1 & 0 & 1\\ 1 & 1 & -1\\ 1 & 0 & 1\end{pmatrix}\begin{pmatrix}e^{2t} & te^{2t} & 0\\ 0 & e^{2t} & 0\\ 0 & 0 & e^{4t}\end{pmatrix}\begin{pmatrix}-\frac{1}{2} & 0 & \frac{1}{2}\\ 1 & 1 & 0\\ \frac{1}{2} & 0 & \frac{1}{2}\end{pmatrix}\begin{pmatrix}1\\ 1\\ 1\end{pmatrix}\\ &=\begin{pmatrix}-2te^{2t}+e^{4t}\\ 2(1+t)e^{2t}-e^{4t}\\ 2te^{2t}+e^{4t}\end{pmatrix},\end{aligned}$$

$$e^{A(t-v)}u(v) = Pe^{Jt}P^{-1}u(v)$$

$$= \begin{pmatrix} -1 & 0 & 1 \\ 1 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^{2(t-v)} & (t-v)e^{2(t-v)} & 0 \\ 0 & e^{2(t-v)} & 0 \\ 0 & 0 & e^{4(t-v)} \end{pmatrix}$$

$$\times \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} e^{2v} \\ 0 \\ ve^{2v} \end{pmatrix}$$

$$= \frac{e^{2t}}{2} \begin{pmatrix} 1 - 2t + e^{2(t-v)} + v + ve^{2(t-v)} \\ 1 + 2t - e^{2(t-v)} - v - ve^{2(t-v)} \\ -1 + 2t + e^{2(t-v)} - v + ve^{2(t-v)} \end{pmatrix},$$

代入公式，定解问题的解为

$$y(t) = e^{At}y(0) + \int_0^t e^{A(t-v)}u(v)\mathrm{d}v$$

$$= e^{2t} \begin{pmatrix} \frac{11}{8}e^{2t} - \frac{3}{8} - \frac{7}{8}t - \frac{3}{4}t^2 \\ -\frac{11}{8}e^{2t} + \frac{19}{8} + \frac{11}{4}t + \frac{3}{4}t^2 \\ \frac{11}{8}e^{2t} - \frac{3}{8} + \frac{5}{4}t + \frac{3}{4}t^2 \end{pmatrix}.$$

例 4 设某一动态微分方程为

$$y''' + 7y'' + 14y' + 8y = 6u(t),$$

其中，y 为系统的输出函数，$u(t)$ 为系统的输入函数，求 $y(t)$.

解 令

$$\begin{cases} x_1 = y, \\ x_2 = y', \\ x^3 = y'', \end{cases}$$

则

$$\begin{cases} x_1' = x_2, \\ x_2' = x_3, \\ x_3' = y''' = -8x_1 - 14x_2 - 7x_3 + 6u(t), \end{cases}$$

写成向量方程组为

$$x' = Ax + Bu(t),$$

其中

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -8 & -14 & -7 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix}.$$

由于 A 为友矩阵，故特征多项式为

$$\varphi(\lambda) = \lambda^3 + 7\lambda^2 + 14\lambda + 8 = (\lambda+1)(\lambda+2)(\lambda+4),$$

故 A 的特征值分别为

$$\lambda_1 = -1, \quad \lambda_2 = -2, \quad \lambda_3 = -4.$$

从而 A 的 Jordan 标准形 J 及相似变换矩阵 P 分别为

$$J = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -4 \end{bmatrix},$$

$$P = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & -4 \\ 1 & 4 & 16 \end{bmatrix},$$

而且

$$P^{-1} = \frac{1}{6}\begin{bmatrix} 16 & 12 & 2 \\ -12 & -15 & -3 \\ 2 & 3 & 1 \end{bmatrix}.$$

方法 1　计算 e^{At} 及 $e^{A(t-v)}$，并代入非齐次线性方程组的定解公式

$$x(t) = e^{At}x(0) + \int_0^t e^{A(t-v)}Bu(v)\mathrm{d}v,$$

其中

$$x(0)=\begin{pmatrix}k_1\\k_2\\k_3\end{pmatrix}$$

为任意给定的初始条件.因而原方程组的解为

$$\begin{aligned}y(t)&=(1,0,\cdots,0)x(t)\\&=(1,0,\cdots,0)\left(e^{At}x(0)+\int_0^t e^{A(t-v)}Bu(v)\mathrm{d}v\right).\end{aligned}$$

此种方法在计算 e^{At} 及 $e^{A(t-v)}$ 的过程中显得繁杂,下面介绍另一种方法.

方法 2 令

$$x=Pz,$$

其中 $z=(z_1,z_2,z_3)^T$,则

$$x'=Pz',$$

因而

$$Pz'=APz+Bu(t).$$

两边左乘 P^{-1},得

$$\begin{aligned}z'&=P^{-1}APz+P^{-1}Bu(t)\\&=\begin{pmatrix}-1&0&0\\0&-2&0\\0&0&-4\end{pmatrix}\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix}+\begin{pmatrix}2\\-3\\1\end{pmatrix}u(t).\end{aligned}$$

由 $y=(1,0,\cdots,0)x$, $x=Pz$ 得

$$\begin{aligned}y&=(1,0,\cdots,0)\begin{pmatrix}1&1&1\\-1&-2&-4\\1&4&16\end{pmatrix}\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix}\\&=z_1+z_2+z_3,\end{aligned}$$

因此,只须求出 z_1,z_2,z_3 即可.解方程组

$$\begin{cases}z_1'=-z_1+2u(t),\\z_2'=-2z_2-3u(t),\\z_3'=-4z_3+u(t),\end{cases}$$

得

$$\begin{cases} z_1 = e^{-t}\left(k_1 + \int 2u(t)e^{t}dt\right), \\ z_2 = e^{-2t}\left(k_2 + \int -3u(t)e^{2t}dt\right), \\ z_3 = e^{-4t}\left(k_3 + \int u(t)e^{4t}dt\right), \end{cases}$$

从而原方程组的解为

$$y = k_1e^{-t} + k_2e^{-2t} + k_3e^{-4t} + 2e^{-t}\int u(t)e^{t}dt$$
$$- 3e^{-2t}\int u(t)e^{2t}dt + e^{-4t}\int u(t)e^{4t}dt,$$

其中,k_1,k_2,k_3 为三个独立的任意常数.

§4.2 线性时变系统的状态方程

1. 线性时变系统的转移矩阵

线性时变系统是指变系数的线性微分方程组.

定义 1 设 n 阶方阵 $A(t)$ 在 $[t_0,t_1]$ 上连续,$x(t)$ 是 $n\times m$ 阶未知矩阵,则称

$$\frac{dx(t)}{dt} = A(t)x(t) \tag{4.8}$$

为**变系数的齐次微分方程组**.

定义 2 设

$$\Phi_j(t,t_0) = \begin{bmatrix} x_{1j}(t,t_0) \\ x_{2j}(t,t_0) \\ \vdots \\ x_{nj}(t,t_0) \end{bmatrix} \qquad (j = 1,2,\cdots,n)$$

满足条件

$$\frac{d\Phi_j(t,t_0)}{dt} = A(t)\Phi_j(t,t_0),$$

且

$$\Phi_j(t,t_0)\big|_{t=t_0}=\begin{pmatrix}x_{1j}(t_0,t_0)\\x_{2j}(t_0,t_0)\\\vdots\\x_{nj}(t_0,t_0)\end{pmatrix}=\begin{pmatrix}0\\\vdots\\0\\1\\0\\\vdots\\0\end{pmatrix}\begin{matrix}\left.\vphantom{\begin{matrix}0\\\vdots\\0\end{matrix}}\right\}j-1\text{个}\\\\\\\\\end{matrix},$$

则称 n 阶方阵

$$\begin{aligned}\Phi(t,t_0)&=(\Phi_1(t,t_0),\Phi_2(t,t_0),\cdots,\Phi_n(t,t_0))\\&=\begin{pmatrix}x_{11}(t,t_0)&x_{12}(t,t_0)&\cdots&x_{1n}(t,t_0)\\x_{21}(t,t_0)&x_{22}(t,t_0)&\cdots&x_{2n}(t,t_0)\\\cdots&&&\\x_{n1}(t,t_0)&x_{n2}(t,t_0)&\cdots&x_{nn}(t,t_0)\end{pmatrix}\end{aligned}$$

为方程组(4.8)的**转移矩阵**,有时又称它为**基本矩阵**.

显然,$\Phi(t_0,t_0)=I_n$.但有时,只要 n 阶方阵的每个列向量都是方程

$$\frac{\mathrm{d}x}{\mathrm{d}t}=A(t)x$$

的解,则称它为转移矩阵.

例 1 证明:矩阵

$$\Phi(t,t_0)=\begin{pmatrix}\mathrm{e}^{2(t-t_0)}\cos(t-t_0)&-\mathrm{e}^{2t-t_0}\sin(t-t_0)\\\mathrm{e}^{t-2t_0}\sin(t-t_0)&\mathrm{e}^{t-t_0}\cos(t-t_0)\end{pmatrix}$$

是方程组

$$\begin{pmatrix}x_1'(t)\\x_2'(t)\end{pmatrix}=\begin{pmatrix}2&-\mathrm{e}^t\\\mathrm{e}^{-t}&1\end{pmatrix}\begin{pmatrix}x_1(t)\\x_2(t)\end{pmatrix}$$

的转移矩阵.

证 容易验证

$$\Phi_1(t,t_0) = \begin{pmatrix} e^{2(t-t_0)}\cos(t-t_0) \\ e^{t-2t_0}\sin(t-t_0) \end{pmatrix},$$

$$\Phi_2(t,t_0) = \begin{pmatrix} -e^{2t-t_0}\sin(t-t_0) \\ e^{t-t_0}\cos(t-t_0) \end{pmatrix}$$

是方程组的解，又因为

$$\Phi_1(t_0,t_0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \Phi_2(t_0,t_0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

故

$$\Phi(t,t_0) = (\Phi_1(t,t_0),\Phi_2(t,t_0))$$

为方程组的转移矩阵.

性质 1 n 阶方阵 $\Phi(t,t_0)$是方程组(4.8)的转移矩阵的充要条件是 $\Phi(t,t_0)$是定解问题

$$\begin{cases} \dfrac{\mathrm{d}x(t)}{\mathrm{d}t} = A(t)x(t), \\ x(t)\big|_{t=t_0} = I_n \end{cases} \tag{4.9}$$

的解.

证 必要性. 设 $\Phi(t,t_0)$为方程组(4.8)式的转移矩阵. 显然，$\Phi(t_0,t_0)=I_n$. 剩下的问题是只需证明 $\Phi(t,t_0)$是方程组(4.9)的解. 按转移矩阵的定义得

$$\frac{\mathrm{d}\Phi_j(t,t_0)}{\mathrm{d}t} = A(t)\Phi_j(t,t_0) \qquad (j=1,2,\cdots,n),$$

因此

$$\begin{aligned} \frac{\mathrm{d}\Phi(t,t_0)}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}(\Phi_1(t,t_0),\Phi_2(t,t_0),\cdots,\Phi_n(t,t_0)) \\ &= \left(\frac{\mathrm{d}\Phi_1(t,t_0)}{\mathrm{d}t},\frac{\mathrm{d}\Phi_2(t,t_0)}{\mathrm{d}t},\cdots,\frac{\mathrm{d}\Phi_n(t,t_0)}{\mathrm{d}t}\right) \\ &= (A(t)\Phi_1(t,t_0),A(t)\Phi_2(t,t_0),\cdots,A(t)\Phi_n(t,t_0)) \\ &= A(t)(\Phi_1(t,t_0),\Phi_2(t,t_0),\cdots,\Phi_n(t,t_0)) \\ &= A(t)\Phi(t,t_0). \end{aligned}$$

充分性. 设

$$\boldsymbol{\Phi}(t,t_0)=(\boldsymbol{\Phi}_1(t,t_0),\boldsymbol{\Phi}_2(t,t_0),\cdots,\boldsymbol{\Phi}_n(t,t_0))$$

是定解问题的解,于是

$$\frac{\mathrm{d}\boldsymbol{\Phi}(t,t_0)}{\mathrm{d}t}=\boldsymbol{A}(t)\boldsymbol{\Phi}(t,t_0),$$

即

$$\frac{\mathrm{d}\boldsymbol{\Phi}_j(t,t_0)}{\mathrm{d}t}=\boldsymbol{A}(t)\boldsymbol{\Phi}_j(t,t_0)\qquad(j=1,2,\cdots,n).$$

又由

$$\boldsymbol{\Phi}(t_0,t_0)=\boldsymbol{I}_n$$

知

$$\boldsymbol{\Phi}_j(t_0,t_0)=\begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix}\begin{array}{l}\left.\begin{array}{l}\\ \\ \\ \end{array}\right\}j-1\text{个}\\ \\ \\ \\ \\ \end{array},$$

因此,$\boldsymbol{\Phi}(t,t_0)$为方程组(4.8)的转移矩阵.

性质 2 设 $\boldsymbol{\Phi}(t,t_0)$是方程组(4.8)的转移矩阵,则定解问题

$$\begin{cases}\dfrac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t}=\boldsymbol{A}(t)\boldsymbol{x}(t),\\ \boldsymbol{x}(t)\big|_{t=t_0}=\boldsymbol{x}(t_0)\end{cases}\tag{4.10}$$

的解为

$$\boldsymbol{x}(t)=\boldsymbol{\Phi}(t,t_0)\boldsymbol{x}(t_0).$$

证 因为

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t}=\frac{\mathrm{d}\boldsymbol{\Phi}(t,t_0)}{\mathrm{d}t}\boldsymbol{x}(t_0),$$

由性质 1 知

$$\frac{\mathrm{d}\Phi(t,t_0)}{\mathrm{d}t} = A(t)\Phi(t,t_0),$$

从而

$$\begin{aligned}\frac{\mathrm{d}x(t)}{\mathrm{d}t} &= \frac{\mathrm{d}\Phi(t,t_0)}{\mathrm{d}t}x(t_0)\\ &= A(t)\Phi(t,t_0)x(t_0)\\ &= A(t)x(t),\end{aligned}$$

而且

$$x(t)\big|_{t=t_0} = \Phi(t_0,t_0)x(t_0) = I_n x(t_0) = x(t_0),$$

故 $x(t)=\Phi(t,t_0)x(t_0)$是定解问题的解.

性质 2 表明,若 $x(t_0)$表示系统的初始状态,$x(t)$表示系统在 t 时刻的状态,则对初始状态作用 $\Phi(t,t_0)$就得(转移到)t 时刻的状态.这既说明了把 $\Phi(t,t_0)$称为转移矩阵的原因,也说明了求一个系统的转移矩阵的重要性.

性质 3 设 $\Phi(t,t_0)$为方程组(4.8)的转移矩阵,则

(1) $\Phi(t,t_0) = \Phi(t,t_1)\Phi(t_1,t_0)$;

(2) $\Phi^{-1}(t,t_0) = \Phi(t_0,t)$.

证 (1) 对任意的 $x(t_0)$,有

$$x(t) = \Phi(t,t_0)x(t_0)$$

是定解问题

$$\begin{cases}\dfrac{\mathrm{d}x(t)}{\mathrm{d}t} = A(t)x(t),\\ x(t)\big|_{t=t_0} = x(t_0)\end{cases}$$

的解,故

$$x(t_1) = \Phi(t_1,t_0)x(t_0),$$

即 $x(t) = \Phi(t,t_0)x(t_0)$ 也是定解问题

$$\begin{cases}\dfrac{\mathrm{d}x(t)}{\mathrm{d}t} = A(t)x(t),\\ x(t)\big|_{t=t_1} = x(t_1)\end{cases}$$

的解. 由性质 2 可知,该定解问题的解为

$$x(t) = \Phi(t,t_1)x(t_1).$$

由定解问题解的唯一性得

$$\begin{aligned} x(t) &= \Phi(t,t_0)x(t_0) \\ &= \Phi(t,t_1)x(t_1) \\ &= \Phi(t,t_1)\Phi(t_1,t_0)x(t_0), \end{aligned}$$

由 $x(t_0)$的任意性得

$$\Phi(t,t_0) = \Phi(t,t_1)\Phi(t_1,t_0).$$

(2) 在(1)中令 $t=t_0$,得

$$I_n = \Phi(t_0,t_0) = \Phi(t_0,t_1)\Phi(t_1,t_0),$$

即

$$\Phi^{-1}(t_1,t_0) = \Phi(t_0,t_1),$$

由 t_1 的任意性得

$$\Phi^{-1}(t,t_0) = \Phi(t_0,t).$$

2. 状态转移矩阵的求法

根据性质 2 ,求定解问题(4.10)的解的关键在于求方程(4.8)的转移矩阵,而根据性质 1,求方程组(4.8)的转移矩阵的关键在于求定解问题(4.9)的解,因此干脆直接找定解问题(4.10)的解.

若 $A(t)$为一阶矩阵,即 $A(t)=a(t)$为数值函数,则定解问题为

$$\begin{cases} x'(t) = a(t)x(t), \\ x(t)\big|_{t=t_0} = x(t_0). \end{cases}$$

它是普通微分方程的定解问题,由《高等数学》中的微分方程知识知,它的解为

$$x(t) = e^{\int_{t_0}^{t} a(v)\mathrm{d}v}x(t_0).$$

因此,可以假设定解问题(4.10)也有类似形式的解:

$$x(t) = e^{\int_{t_0}^{t} A(v)\mathrm{d}v}x(t_0). \qquad (4.11)$$

应该指出的是，(4.11)式是否真的是定解问题(4.10)的解，还必须验证. 由矩阵的幂级数定义，有

$$x(t)=\left[I+\int_{t_0}^{t}A(v)\mathrm{d}v+\frac{1}{2!}\left(\int_{t_0}^{t}A(v)\mathrm{d}v\right)^2\right.$$

$$\left.+\cdots+\frac{1}{k!}\left(\int_{t_0}^{t}A(v)\mathrm{d}v\right)^k+\cdots\right]x(t_0),$$

则 $x(t)$ 的导数为

$$x'(t)=\left[A(t)+\frac{1}{2!}\left(\int_{t_0}^{t}A(v)\mathrm{d}vA(t)+A(t)\int_{t_0}^{t}A(v)\mathrm{d}v\right)\right.$$

$$\left.+\cdots\right]x(t_0),$$

而

$$A(t)x(t)=\left[A(t)+A(t)\int_{t_0}^{t}A(v)\mathrm{d}v+\frac{1}{2!}A(t)\left(\int_{t_0}^{t}A(v)\mathrm{d}v\right)^2\right.$$

$$\left.+\cdots\right]x(t_0).$$

在一般情况下，矩阵乘法不满足交换律，因此不一定有

$$x'(t)=A(t)x(t),$$

但若有

$$\int_{t_0}^{t}A(v)\mathrm{d}v\cdot A(t)=A(t)\int_{t_0}^{t}A(v)\mathrm{d}v, \qquad (4.12)$$

则

$$x'(t)=A(t)x(t).$$

事实上，若(4.12)式成立，则

$$A(t)\left(\int_{t_0}^{t}A(v)\mathrm{d}v\right)^2=A(t)\int_{t_0}^{t}A(v)\mathrm{d}v\int_{t_0}^{t}A(v)\mathrm{d}v$$

$$=\int_{t_0}^{t}A(v)\mathrm{d}vA(t)\int_{t_0}^{t}A(v)\mathrm{d}v$$

$$=\int_{t_0}^{t}A(v)\mathrm{d}v\cdot\int_{t_0}^{t}A(v)\mathrm{d}v\cdot A(t)$$

$$= \left(\int_{t_0}^{t} A(v)\mathrm{d}v \right)^2 \cdot A(t).$$

同理可得

$$A(t)\left(\int_{t_0}^{t} A(v)\mathrm{d}v \right)^k = \left(\int_{t_0}^{t} A(v)\mathrm{d}v \right)^k A(t) \qquad (k = 3,4,\cdots),$$

因此

$$x'(t) = A(t)x(t),$$

而且

$$x(t)\big|_{t=t_0} = I_n x(t_0) = x(t_0),$$

故(4.11)式是定解问题(4.10)的解. 由于条件(4.12)使用时不方便,下面将它变为对 $A(t)$的条件,进而得到求解定解问题(4.10)的一个充分条件.

定理 1 若对任意的 t_1、t_2,有

$$A(t_1)A(t_2) = A(t_2)A(t_1),$$

则定解问题(4.10)的解为

$$\begin{aligned} x(t) &= \mathrm{e}^{\int_{t_0}^{t} A(v)\mathrm{d}v} x(t_0) \\ &= \left[I + \int_{t_0}^{t} A(v)\mathrm{d}v + \frac{1}{2!}\left(\int_{t_0}^{t} A(v)\mathrm{d}v \right)^2 + \cdots \right. \\ &\quad \left. + \frac{1}{k!}\left(\int_{t_0}^{t} A(v)\mathrm{d}v \right)^k + \cdots \right] x(t_0), \end{aligned}$$

而且,方程组(4.8)的转移矩阵为

$$\begin{aligned} \Phi(t,t_0) &= \mathrm{e}^{\int_{t_0}^{t} A(v)\mathrm{d}v} \\ &= I + \int_{t_0}^{t} A(v)\mathrm{d}v + \frac{1}{2!}\left(\int_{t_0}^{t} A(v)\mathrm{d}v \right)^2 + \cdots \\ &\quad + \frac{1}{k!}\left(\int_{t_0}^{t} A(v)\mathrm{d}v \right)^k + \cdots. \end{aligned}$$

证 由于对任意的 t_1,t_2,有

$$A(t_1)A(t_2) = A(t_2)A(t_1),$$

故

$$\int_{t_0}^{t} A(v)\mathrm{d}v \cdot A(t) = \int_{t_0}^{t} A(v)A(t)\mathrm{d}v$$
$$= \int_{t_0}^{t} A(t)A(v)\mathrm{d}v$$
$$= A(t)\int_{t_0}^{t} A(v)\mathrm{d}v,$$

即公式(4.12)成立，从而(4.11)是定解问题(4.10)的解. 特别地，取$x(t_0)=I_n$时，由性质 1 知

$$\Phi(t,t_0) = \mathrm{e}^{\int_{t_0}^{t} A(v)\mathrm{d}v}$$

是方程组(4.8)的转移矩阵.

例 2 设

$$A(t) = \begin{pmatrix} 0 & \dfrac{1}{(1+t)^2} \\ 0 & 0 \end{pmatrix},$$

求方程组

$$x'(t) = A(t)x(t)$$

的转移矩阵.

解 对任意的 t_1、t_2、有

$$A(t_1)A(t_2) = \begin{pmatrix} 0 & \dfrac{1}{(1+t_1^2)} \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & \dfrac{1}{(1+t_2)^2} \\ 0 & 0 \end{pmatrix}$$
$$= 0$$
$$= A(t_2)A(t_1),$$

故方程组的转移矩阵为

$$\Phi(t,t_0) = \mathrm{e}^{\int_{t_0}^{t} A(v)\mathrm{d}v}$$
$$= I + \int_{t_0}^{t} A(v)\mathrm{d}v + \frac{1}{2!}\left(\int_{t_0}^{t} A(v)\mathrm{d}v\right)^2 + \cdots.$$

由于

$$\int_{t_0}^{t} A(v)\mathrm{d}v = \begin{pmatrix} 0 & \displaystyle\int_{t_0}^{t} \frac{1}{(1+v)^2}\mathrm{d}v \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \dfrac{t-t_0}{(t+1)(t_0+1)} \\ 0 & 0 \end{pmatrix},$$

$$\left(\int_{t_0}^{t} A(v)\mathrm{d}v\right)^2$$

$$= \begin{pmatrix} 0 & \dfrac{t-t_0}{(t+1)(t_0+1)} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & \dfrac{t-t_0}{(t+1)(t_0+1)} \\ 0 & 0 \end{pmatrix}$$

$$= 0,$$

从而

$$\left(\int_{t_0}^{t} A(v)\mathrm{d}v\right)^k = 0 \qquad (k = 3,4,\cdots),$$

故方程组的转移矩阵为

$$\Phi(t,t_0) = I + \int_{t_0}^{t} A(v)\mathrm{d}v$$

$$= \begin{pmatrix} 1 & \dfrac{t-t_0}{(t+1)(t_0+1)} \\ 0 & 1 \end{pmatrix}.$$

显然，定理 1 中的条件是定解问题(4.10)求解的一个充分条件，但并不一定是必要条件．下面说明，不管 $A(t)$ 是否满足该条件，但只要它能使级数

$$I + \int_{t_0}^{t} A(v)\mathrm{d}v + \int_{t_0}^{t} A(v_1)\mathrm{d}v_1\int_{t_0}^{v_1} A(v_2)\mathrm{d}v_2$$
$$+ \int_{t_0}^{t} A(v_1)\mathrm{d}v_1\int_{t_0}^{v_1} A(v_2)\mathrm{d}v_2\int_{t_0}^{v_2} A(v_3)\mathrm{d}v_3 + \cdots$$

一致收敛，则矩阵

$$\Phi(t,t_0) = I + \int_{t_0}^{t} A(v)\mathrm{d}v + \int_{t_0}^{t} A(v_1)\mathrm{d}v_1\int_{t_0}^{v_1} A(v_2)\mathrm{d}v_2$$
$$+ \int_{t_0}^{t} A(v_1)\mathrm{d}v_1\int_{t_0}^{v_1} A(v_2)\mathrm{d}v_2\int_{t_0}^{v_2} A(v_3)\mathrm{d}v_3 + \cdots$$

必为方程组(4.8)的转移矩阵，上述级数称为 **Peano-Baker 级数**.

事实上，若 Peano-Baker 级数一致收敛，则可逐项求导：

$$
\begin{aligned}
\Phi'(t,t_0) &= A(t) + A(t)\int_{t_0}^{t} A(v_2)\mathrm{d}v_2 \\
&\qquad + A(t)\int_{t_0}^{t} A(v_2)\mathrm{d}v_2\int_{t_0}^{v_2} A(v_3)\mathrm{d}v_3 + \cdots \\
&= A(t)\Big[I + \int_{t_0}^{t} A(v_2)\mathrm{d}v_2 \\
&\qquad - \int_{t_0}^{t} A(v_2)\mathrm{d}v_2\int_{t_0}^{v_2} A(v_3)\mathrm{d}v_3 + \cdots\Big] \\
&= A(t)\Phi(t,t_0),
\end{aligned}
$$

而且

$$\Phi(t_0,t_0) = I.$$

由性质 1 知，$\Phi(t,t_0)$为方程组(4.8)的转移矩阵.

当 $A(t)$满足什么条件时，Peano-Baker 级数才一致收敛呢？下面给出一个充分条件.

定理 2 若 $A(t)$的各元素在$[t_0,t_1]$上有界，则 Peano-Baker 级数为方程组(4.8)的矩阵.

证 为了证明 Peano-Baker 级数为方程组(4.8)的转移矩阵，只需证明 Peano-Baker 级数在$[t_0,t_1]$上一致收敛即可. 由于矩阵级数一致收敛等价于矩阵的各元素级数一致收敛. 设

$$A(t) = (a_{ij}(t))_{n\times n},$$

并设 M 是各元素的公共上界，即

$$|a_{ij}(t)| \leqslant M \qquad (i,j = 1,2,\cdots,n;t_0 \leqslant t \leqslant t_1),$$

则对任何 i,j 及 $t\in[t_0,t_1]$，有

$$\left|\int_{t_0}^{t} a_{ij}(t)\mathrm{d}t\right| \leqslant M(t_1 - t_0),$$

$$
\begin{aligned}
&\left|\int_{t_0}^{t} a_{ij}(v_1)\mathrm{d}v_1\int_{t_0}^{v_1} a_{ij}(v_2)\mathrm{d}v_2\right| \\
&\qquad = \left|\int_{t_0}^{t}\left[a_{ij}(v_1)\int_{t_0}^{v_1} a_{ij}(v_2)\mathrm{d}v_2\right]\mathrm{d}v_1\right|
\end{aligned}
$$

$$\leqslant \int_{t_0}^{t}\left[|a_{ij}(v_1)|\cdot\left|\int_{t_0}^{v_1}a_{ij}(v_2)\mathrm{d}v_2\right|\right]\mathrm{d}v_1$$

$$\leqslant \int_{t_0}^{t}\left[|a_{ij}(v_1)|\cdot\int_{t_0}^{v_1}|a_{ij}(v_2)|\mathrm{d}v_2\right]\mathrm{d}v_1$$

$$\leqslant M^2\int_{t_0}^{t}(v_1-t_0)\mathrm{d}v_1$$

$$=\frac{1}{2!}M^2(t-t_0)^2$$

$$\leqslant \frac{1}{2!}M^2(t_1-t_0)^2.$$

类似地，有

$$\left|\int_{t_0}^{t}a_{ij}(v_1)\mathrm{d}v_1\int_{t_0}^{v_1}a_{ij}(v_2)\mathrm{d}v_2\int_{t_0}^{v_2}a_{ij}(v_3)\mathrm{d}v_3\right|$$

$$\leqslant \frac{1}{2}M^3\int_{t_0}^{t}(t-t_0)^2\mathrm{d}t$$

$$=\frac{1}{3!}M^3(t-t_0)^3$$

$$\leqslant \frac{1}{3!}M^3(t_1-t_0)^3,$$

……

易知，级数

$$M(t_1-t_0)+\frac{1}{2!}M^2(t_1-t_0)^2+\frac{1}{3!}(t_1-t_0)^3+\cdots$$

是收敛的，从而

$$\int_{t_0}^{t}a_{ij}(v)\mathrm{d}v+\int_{t_0}^{t}a_{ij}(v_1)\mathrm{d}v_1\int_{t_0}^{v_1}a_{ij}(v_2)\mathrm{d}v_2$$
$$+\int_{t_0}^{t}a_{ij}(v_1)\mathrm{d}v_1\int_{t_0}^{v_1}a_{ij}(v_2)\mathrm{d}v_2\int_{t_0}^{v_2}a_{ij}(v_3)\mathrm{d}v_3+\cdots$$

一致收敛．故 Peano-Baker 级数一致收敛．

例 3 求时变系统

$$\begin{pmatrix}x_1'\\x_2'\end{pmatrix}=\begin{pmatrix}0&1\\0&t\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$$

的转移矩阵．

解 由于

$$A(t)=\begin{pmatrix}0 & 1\\ 0 & t\end{pmatrix},$$

显然，$A(t)$各元素在有限区间上总是有界的，故 $A(t)$的 Peano-Baker 级数即为时变系统的转移矩阵. 下面计算该级数的各项：

$$\int_{t_0}^{t}A(v)\mathrm{d}v=\int_{t_0}^{t}\begin{pmatrix}0 & 1\\ 0 & v\end{pmatrix}\mathrm{d}v=\begin{bmatrix}0 & t-t_0\\ 0 & \frac{1}{2}(t^2-t_0^2)\end{bmatrix},$$

$$\int_{t_0}^{t}A(v_1)\mathrm{d}v_1\int_{t_0}^{v_1}A(v_2)\mathrm{d}v_2=\int_{t_0}^{t}\begin{pmatrix}0 & 1\\ 0 & v_1\end{pmatrix}\begin{bmatrix}0 & v_1-t_0\\ 0 & \frac{1}{2}(v_1^2-t_0^2)\end{bmatrix}\mathrm{d}v_1$$

$$=\int_{t_0}^{t}\begin{bmatrix}0 & \frac{1}{2}(v_1^2-t_0^2)\\ 0 & \frac{1}{2}v_1(v_1^2-t_0^2)\end{bmatrix}\mathrm{d}v_1$$

$$=\begin{bmatrix}0 & \frac{1}{6}(t-t_0)^2(t+2t_0)\\ 0 & \frac{1}{8}(t^2-t_0^2)\end{bmatrix},$$

……

因此时变系统的转移矩阵为

$$\Phi(t,t_0)=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}+\begin{bmatrix}0 & t-t_0\\ 0 & \frac{1}{2}(t^2-t_0^2)\end{bmatrix}$$

$$+\begin{bmatrix}0 & \frac{1}{6}(t-t_0)^2(t+2t_0)\\ 0 & \frac{1}{8}(t^2-t_0^2)\end{bmatrix}+\cdots$$

$$=\begin{bmatrix}1 & (t-t_0)+\frac{1}{6}(t-t_0)^2(t+2t_0)+\cdots\\ 0 & 1+\frac{1}{2}(t^2-t_0^2)+\frac{1}{8}(t^2-t_0^2)+\cdots\end{bmatrix}.$$

3. 非齐次时变系统的定解问题

设 $A(t)$ 为 n 阶方阵，$B(t)$ 为 $n\times m$ 阶矩阵，且 $A(t)$ 与 $B(t)$ 在所讨论的区间上连续，$x(t)$ 与 $u(t)$ 分别为 n 维及 m 维列向量. 现在求解非齐次时变系统的定解问题

$$\begin{cases} x'(t) = A(t)x(t) + B(t)u(t), \\ x(t)\big|_{t=t_0} = x(t_0). \end{cases} \tag{4.13}$$

由于非齐次时变系统所对应的齐次时变系统为

$$x'(t) = A(t)x(t),$$

若 $\Phi(t,t_0)$ 为它的转移矩阵，则它的通解为

$$x(t) = \Phi(t,t_0)c,$$

其中 c 为任意的 n 维列向量. 下面用常数变易法求非齐次时变系统的解. 设

$$x(t) = \Phi(t,t_0)c(t)$$

为非齐次时变系统的解，则

$$\begin{aligned} x'(t) &= \Phi'(t,t_0)c(t) + \Phi(t,t_0)c'(t) \\ &= A(t)\Phi(t,t_0)c(t) + \Phi(t,t_0)c'(t), \end{aligned}$$

代入非齐次时变系统得

$$\begin{aligned} &A(t)\Phi(t,t_0)c(t) + \Phi(t,t_0)c'(t) \\ &\quad = A(t)\Phi(t,t_0)c(t) + B(t)u(t), \end{aligned}$$

则

$$c'(t) = \Phi^{-1}(t,t_0)B(t)u(t) = \Phi(t_0,t)B(t)u(t),$$

故

$$c(t) = c + \int_{t_0}^{t}\Phi(t_0,t)B(t)u(t)\mathrm{d}t.$$

从而，非齐次时变系统的解为

$$\begin{aligned} x(t) &= \Phi(t,t_0)c(t) \\ &= \Phi(t,t_0)\left[c + \int_{t_0}^{t}\Phi(t_0,t)B(t)u(t)\mathrm{d}t\right] \end{aligned}$$

$$= \Phi(t,t_0)c + \int_{t_0}^{t} \Phi(t,v)B(v)u(v)\mathrm{d}v.$$

为了满足初始条件

$$x(t)\big|_{t=t_0} = x(t_0),$$

应有

$$x(t_0) = \Phi(t_0,t_0)c + 0 = Ic = c,$$

即

$$c = x(t_0).$$

故定解问题(4.13)的解为

$$x(t) = \Phi(t,t_0)x(t_0) + \int_{t_0}^{t} \Phi(t,v)B(v)u(v)\mathrm{d}v,$$

其中,$\Phi(t,t_0)$为相应的齐次时变系统的转移矩阵.

在定常系统中,A 为常数矩阵,且

$$\Phi(t,t_0) = \mathrm{e}^{A(t-t_0)},$$

故

$$x(t) = \mathrm{e}^{A(t-t_0)}x(t_0) + \int_{t_0}^{t} \mathrm{e}^{A(t-v)}B(v)u(v)\mathrm{d}v.$$

由此可见,定常系统是时变系统的特例.

例 4 求系统

$$\begin{cases} \begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = \begin{pmatrix} 0 & \dfrac{1}{(t+1)^2} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \\ x(t)\big|_{t=t_0} = x(t_0) \end{cases}$$

的解.

解 由于

$$A(t) = \begin{pmatrix} 0 & \dfrac{1}{(1+t)^2} \\ 0 & 0 \end{pmatrix}, \quad Bu(t) = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

由例 2 知,系统所对应的齐次系统

$$x'(t) = A(t)x(t)$$

的转移矩阵为

$$\Phi(t,t_0) = \begin{bmatrix} 1 & \dfrac{t-t_0}{(t+1)(t_0+1)} \\ 0 & 1 \end{bmatrix}.$$

故系统的解为

$$\begin{aligned} x(t) &= \Phi(t,t_0)x(t_0) + \int_{t_0}^{t} \Phi(t,v)Bu(v)\mathrm{d}v \\ &= \begin{bmatrix} 1 & \dfrac{t-t_0}{(t+1)(t_0+1)} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t_0) \\ x_2(t_0) \end{bmatrix} \\ &\quad + \int_{t_0}^{t} \begin{bmatrix} 1 & \dfrac{t-v}{(t+1)(v+1)} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mathrm{d}v \\ &= \begin{bmatrix} x_1(t_0) + \dfrac{t-t_0}{(t+1)(t_0+1)} x_2(t_0) \\ x_2(t_0) \end{bmatrix} \\ &\quad + \int_{t_0}^{t} \begin{bmatrix} 1 + \dfrac{t-v}{(t+1)(v+1)} \\ 1 \end{bmatrix} \mathrm{d}v \\ &= \begin{bmatrix} x_1(t_0) + \dfrac{t-t_0}{t+1}\left(\dfrac{x_2(t_0)}{t_0+1} + t\right) + \ln\dfrac{t+1}{t_0+1} \\ x_2(t_0) + t - t_0 \end{bmatrix}. \end{aligned}$$

习 题 四

1．求下列微分方程组的通解：

(1) $\begin{cases} \dfrac{\mathrm{d}x_1}{\mathrm{d}t} = x_1 + 2x_2, \\ \dfrac{\mathrm{d}x_2}{\mathrm{d}t} = 4x_1 + 3x_2; \end{cases}$

(2) $\begin{cases} \dfrac{\mathrm{d}x_1}{\mathrm{d}t} = x_2 + x_3, \\ \dfrac{\mathrm{d}x_2}{\mathrm{d}t} = x_1 + x_2 - x_3, \\ \dfrac{\mathrm{d}x_3}{\mathrm{d}t} = x_2 + x_3. \end{cases}$

2. 求微分方程组$\dfrac{\mathrm{d}x}{\mathrm{d}t}=Ax$满足初始条件$x(0)=\xi$的解：

(1) $A=\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$, $\xi=\begin{pmatrix} 3 \\ 3 \end{pmatrix}$；

(2) $A=\begin{bmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \\ 2 & 0 & 1 \end{bmatrix}$, $\xi=\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

3. 求$\dfrac{\mathrm{d}x}{\mathrm{d}t}=Ax+Bu(t)$满足条件$x(0)=\xi$的解：

(1) $A=\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$, $B=\begin{pmatrix} -1 \\ 4 \end{pmatrix}$, $u(t)=\mathrm{e}^{-t}$, $\xi=\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$；

(2) $A=\begin{bmatrix} -6 & 1 & 0 \\ -11 & 0 & 1 \\ -6 & 0 & 0 \end{bmatrix}$, $B=\begin{bmatrix} 2 \\ 6 \\ 2 \end{bmatrix}$, $u(t)=1$, $\xi=\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$.

4. 求方程

$$y''' + 6y'' + 11y' + 6y = \mathrm{e}^{-t}$$

满足 $y(0) = y'(0) = y''(0) = 0$ 的解.

5. 试证明：若 A 为 2 阶方阵，其特征值为 λ_1 和 λ_2，特征向量为 P_1 和 P_2，则方程

$$\frac{\mathrm{d}x}{\mathrm{d}t} = Ax$$

的解一定能表示成

$$x = c_1\mathrm{e}^{\lambda_1 t}P_1 + c_2\mathrm{e}^{\lambda_2 t}P_2,$$

其中，常数 c_1 与 c_2 由下式确定：

$$x(0) = c_1P_1 + c_2P_2.$$

然后利用这一结论求解定解问题：

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=\begin{pmatrix}0 & 1\\ -5 & -6\end{pmatrix},\\ x(0)=\begin{pmatrix}1\\1\end{pmatrix}\end{cases}$$

的解. 并将这一结论推广到 n 阶方阵情形.

6. 已知 $\Phi(t,t_0)$ 是方程组

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}=A(t)x(t)$$

的转移矩阵，试证：

$$\frac{\mathrm{d}}{\mathrm{d}t_0}\Phi(t,t_0)=-\Phi(t,t_0)A(t_0).$$

7. 求时变系统

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=A(t)x(t),\\ x(t)\big|_{t=t_0}=x_0\end{cases}$$

的解，其中，$A(t)$，x_0 分别如下：

(1) $A(t)=\begin{pmatrix}1 & \mathrm{e}^{-t}\\ 0 & 1\end{pmatrix}$，　$x_0=\begin{pmatrix}1\\1\end{pmatrix}$，　$t_0=0$；

(2) $A(t)=\begin{pmatrix}0 & \dfrac{1}{t^2} & 1\\ 0 & 0 & \dfrac{1}{(1+t)^2}\\ 0 & 0 & 0\end{pmatrix}$，　$x_0=\begin{pmatrix}1\\1\\1\end{pmatrix}$；

(3) $A(t)=\begin{pmatrix}t & 1\\ 2 & t\end{pmatrix}$，　$x_0=\begin{pmatrix}1\\-1\end{pmatrix}$，　$t_0=0$.

8. 求下列定解问题的解：

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=A(t)x(t)+B(t)u(t),\\ x(t_0)=x_0,\end{cases}$$

其中：

(1) $A(t)=\begin{pmatrix} 1 & e^{t} \\ 0 & 1 \end{pmatrix}$, $B(t)u(t)=\begin{pmatrix} t \\ 1 \end{pmatrix}$, $x_0=\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $t_0=0$;

(2) $A(t)=\begin{pmatrix} 0 & \frac{1}{t^2} & 0 \\ 0 & 0 & \frac{1}{(1+t)^2} \\ 0 & 0 & 0 \end{pmatrix}$, $B(t)u(t)=\begin{pmatrix} t \\ 1 \\ 0 \end{pmatrix}$, $x_0=\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

第五章　广义逆矩阵

在线性代数中已对线性方程组 $AX=b$ 进行了较完整的讨论，它可以无解，或恰有一组解，或有无穷多组解.初看起来，似乎无解的矛盾方程组(或称不相容方程组)最为乏味并且没有实际意义，但事实却相反.在某些实际问题中，如数据处理、多元分析、最优化理论、现代控制理论、网络理论等学科中，所遇到的方程组往往是不相容方程组.此时，我们不能求得 $AX=b$ 的解，而只能将要求合理地改为：寻求 $X\in C^n$，使 $\|AX-b\|$ 为最小.当 $\|\cdot\|$ 为欧氏范数时，这样的解称为线性方程组的最小二乘解，或最小剩余解.如果 $AX=b$ 是相容的且有无穷多组解，在这无穷多组解中，往往要求的是范数最小的解.问题是这样的解是否能统一地表示成紧凑形式 $X=Gb$，其中 G 是某个矩阵.这个问题的回答是肯定的.这就是引出逆矩阵的实际背景.

E. H. Moore 于 1920 年在美国数学会上提出了他的广义逆矩阵的一个论文摘要，论文发表在他死后的 1935 年.本世纪 30 年代，我国的曾远荣先生把它推广到了 Hilbert 空间线性算子中，他还是把不加可分条件的完备内积空间叫做 Hilbert 空间的创始人之一.由于不知其用途，逆矩阵一直未被重视，直到 50 年代，由于数学的发展，需要广义逆矩阵概念的要求日益增多.1955 年，R. Penrose 发表了与 Moore 等价的广义逆矩阵的理论，现在就称为 Moore-Penrose 广义逆矩阵，常记成 A^+.同年，Rao 提出了一个更一般的广义逆矩阵概念，现在叫做 g 逆，常记成 A^-.此后，广义逆矩阵的理论才逐步发展起来，并开始广泛地应用于许多学科中.

§5.1 和相容方程组求解问题相应的广义逆矩阵 A^-

1. 广义逆矩阵 A^- 的定义及性质

设线性方程组

$$AX = b \tag{5.1}$$

是相容的，其中 $A \in C^{m\times n}$，$X \in C^n$，$b \in C^m$.

注 $AX=b$ 相容 $\Longleftrightarrow b \in R(A)$（矩阵 A 的象空间）.

定理 1 对任何 $b \in R(A)$，有矩阵 G，使得 $Gb=X$ 是相容方程组(5.1)的解的充要条件是，G 满足

$$AGA = A. \tag{5.2}$$

证 必要性. 若有矩阵 G，对任何 $b \in R(A)$，使得 $X=Gb$ 是(5.1)的解，要证明 $AGA=A$. 事实上，对任意 $Z \in C^n$，令 $b=AZ \in R(A)$. 由假设，$X=Gb$ 是 $AX=b$ 的解，因此

$$AGb = b,$$

代入 $b=AZ$，有 $AGAZ=AZ$. 由 $Z \in C^n$ 的任意性即得 $AGA=A$.

充分性. 若有矩阵 G 满足 $AGA=A$，证明对任意的 $b \in R(A)$，$X=Gb$ 是 $AX=b$ 的解. 事实上，因为 $b \in R(A)$，所以存在 $Z \in C^n$，使 $AZ=b$. 对于这个 Z，由于 $AGA=A$，故有 $AGAZ=AZ$，即 $AGb=b$. 这说明 $X=Gb$ 是(5.1)的解.

定义 1 设 $A \in C^{m\times n}$，若存在 $G \in C^{n\times m}$，使得

$$AGA = A, \tag{①}$$

则称 G 为 A 的 **g 逆**，记作 A^-.

因此有 $AA^-A=A$.

注 (1)若 A 存在逆矩阵 A^{-1}，则 A^{-1}是 A 的 g 逆 A^-；但反之不真.

(2)$X=A^-b$ 是相容方程 $AX=b$ 的解.

2. 矩阵 A 的左逆 A_L^{-1} 与右逆 A_R^{-1}

在没有证明 g 逆的存在性之前，请注意如下事实：若 $G\in C^{n\times m}$，使 $AG=I$，或 $GA=I$，则 G 必为 A 的 g 逆. 但这样的 G 并不一定是 A 的逆，因此可引入下面的定义：

定义 2 设 $A\in C^{m\times n}$，若有 $G\in C^{n\times m}$，使得

$$AG = I(\text{或 } GA = I),$$

则称 G 为 A 的**右逆**（或**左逆**），记为 A_R^{-1}（或 A_L^{-1}），即

$$AA_R^{-1} = I(\text{或 } A_L^{-1}A = I).$$

在一般情况下，$A_R^{-1}\neq A_L^{-1}$. 若 $A_R^{-1}=A_L^{-1}$，则 A^{-1}存在，且 $A^{-1}=A_R^{-1}=A_L^{-1}$.

定理 2 若 A 是行（或列）满秩，则必存在 A 的右逆（或左逆），且

$$A_R^{-1} = A^*(AA^*)^{-1}(\text{或 } A_L^{-1} = (A^*A)^{-1}A^*). \tag{5.3}$$

证 因为 A 是行（或列）满秩，所以 AA^*（或 A^*A）为满秩方阵，因此有

$$AA^*(AA^*)^{-1} = I_{m\times m} = (AA^*)^{-1}AA^*$$

$$(\text{或}(A^*A)^{-1}A^*A = I_{n\times n} = A^*A(A^*A)^{-1}),$$

所以有

$$A_R^{-1} = A^*(AA^*)^{-1}(\text{或 } A_L^{-1} = (A^*A)^{-1}A^*).$$

定理 2 所定义的右逆（或左逆）的**性质**为：

$$\left.\begin{array}{l}
(1)\ AA_R^{-1}A=A,\ AA_L^{-1}A=A; \\
(2)\ A_R^{-1}AA_R^{-1}=A_R^{-1},\ A_L^{-1}AA_L^{-1}=A_L^{-1}; \\
(3)\ (AA_R^{-1})^*=AA_R^{-1},\ (AA_L^{-1})^*=AA_L^{-1}; \\
(4)\ (A_R^{-1}A)^*=A_R^{-1}A,\ (A_L^{-1}A)^*=A_L^{-1}A.
\end{array}\right\} \quad (RL)$$

证 以(4)为例证之，其余请读者自证.

$$\begin{aligned}(A_R^{-1}A)^* &= [A^*(AA^*)^{-1}A]^* \\ &= A^*[(AA^*)^{-1}]^*(A^*)^*\end{aligned}$$

$$\begin{aligned}&= A^*[(AA^*)^*]^{-1}A \\ &= A^*(AA^*)^{-1}A \\ &= A_R^{-1}A.\end{aligned}$$

3. g 逆 A^- 的存在性

定理 3 非零矩阵 $A\in C^{m\times n}$ 总有 g 逆 A^-.

证 下面给出定理 3 的构造性证明，即不仅证明它的存在性，而且给出构造 g 逆的方法. 分两种情况讨论：

(1) A 是行(或列)满秩，即 $\operatorname{rank} A=m\leqslant n$(或 $\operatorname{rank} A=n\leqslant m$)，这时 A 有右(或左)逆，而左、右逆都是 g 逆，因此知 A^- 存在，且

$$A^-=A_R^{-1}=A^*(AA^*)^{-1}(\text{或 } A^-=A_L^{-1}=(A^*A)^{-1}A^*).$$

(2) A 既不是行满秩也不是列满秩，即

$$\operatorname{rank} A=r<\min\{m,n\} \qquad (r>0).$$

对这类矩阵进行满秩分解，即有列满秩矩阵 $C\in C^{m\times r}$ 与行满秩矩阵 $D\in C^{r\times n}$ 使得 $A=CD$，因此存在 C_L^{-1} 与 D_R^{-1}. 因为

$$A(D_R^{-1}C_L^{-1})A=CD(D_R^{-1}C_L^{-1})CD=CD=A,$$

所以

$$A^-=D_R^{-1}C_L^{-1}. \tag{5.4}$$

关于满秩分解的问题，我们可以这样来进行：对 A 进行一系列行、列初等变换，使 A 变成

$$PAQ=\begin{pmatrix}A_r & 0\\ 0 & 0\end{pmatrix},$$

其中 A_r 为 r 阶满秩方阵. 于是

$$A=P^{-1}\begin{pmatrix}A_r & 0\\ 0 & 0\end{pmatrix}Q^{-1}=P^{-1}\begin{pmatrix}A_r\\ 0\end{pmatrix}(I_r\quad 0)Q^{-1},$$

取 $C=P^{-1}\begin{pmatrix}A_r\\ 0\end{pmatrix}\in C^{m\times r}$，$D=(I_r\quad 0)Q^{-1}\in C^{r\times n}$ 即可.

请读者证明，这样构造出来的 A^- 满足(RL)式的四个性质. 例

如性质(3)：$(AA^-)^*=AA^-$. 事实上

$$(AA^-)^* = (CDD_R^{-1}C_L^{-1})^* = (CC_L^{-1})^* = CC_L^{-1}$$
$$= CDD_R^{-1}C_L^{-1} = AA^-.$$

例 1 设

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

求 A^-.

解 因为 rank $A=2$，所以 A 为行满秩，故

$$A^-=A_R^{-1}=A^*(AA^*)^{-1}.$$

于是有

$$A^-=\begin{bmatrix}1 & 0\\ 2 & -1\\ -1 & 2\end{bmatrix}\left[\begin{pmatrix}1 & 2 & -1\\ 0 & -1 & 2\end{pmatrix}\begin{bmatrix}1 & 0\\ 2 & -1\\ -1 & 2\end{bmatrix}\right]^{-1}$$
$$=\begin{bmatrix}1 & 0\\ 2 & -1\\ -1 & 2\end{bmatrix}\begin{pmatrix}6 & -4\\ -4 & 5\end{pmatrix}^{-1}$$
$$=\begin{bmatrix}1 & 0\\ 2 & -1\\ -1 & 2\end{bmatrix}\cdot\frac{1}{14}\begin{pmatrix}5 & 4\\ 4 & 6\end{pmatrix}$$
$$=\frac{1}{14}\begin{bmatrix}5 & 4\\ 6 & 2\\ 3 & 8\end{bmatrix}.$$

例 2 设

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

求 A^-.

解 因为 rank $A=2$，所以 A 为列满秩. 故

$$A^-=A_L^{-1}=(A^*A)^{-1}A^*$$

$$= \left[\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 1 \end{pmatrix} \right]^{-1} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix}$$

$$= \frac{1}{11} \begin{pmatrix} -4 & 7 & 1 \\ 7 & -4 & 1 \end{pmatrix}.$$

4. g 逆的一般表达式

定理 4(g 逆的一般表达式) 设 $A \in C^{m\times n}$，A^- 为 A 的一个 g 逆，则对任意 $V, W \in C^{n\times m}$，

$$G = A^- + V(I_m - AA^-) + (I_n - A^- A)W \tag{5.5}$$

也是 A 的 g 逆. 并且对任意一个 A 的 g 逆 G_1，必存在 $V_1, W_1 \in C^{n\times m}$，使 G_1 表成(5.5)式的形式. 因此，(5.5)式是 A 的 g 逆的一般表达式.

证 对任何 $V, W \in C^{n\times m}$，考虑

$$\begin{aligned} AGA &= AA^- A + AV(I_m - AA^-)A + A(I_n - A^- A)WA \\ &= AA^- A + AV(A - AA^- A) + (A - AA^- A)WA \\ &= A + AV(A - A) - (A - A)WA \\ &= A, \end{aligned}$$

故 G 为 A 的 g 逆.

可见 A 的 g 逆并不唯一，其全体记成 $A\{1\}$.

任取 A 的一个 g 逆 G_1，令 $V_1 = G_1 - A^-$，$W_1 = G_1AA^-$，代入(5.5)式的右端，有

$$\begin{aligned} & A^- + V_1(I_m - AA^-) + (I_n - A^- A)W_1 \\ &\quad = A^- + (G_1 - A^-)(I_m - AA^-) \\ &\qquad + (I_n - A^- A)G_1AA^- \\ &\quad = A^- + G_1(I_m - AA^-) - A^-(I_m - AA^-) \\ &\qquad + G_1AA^- - A^- AG_1AA^- \\ &\quad = G_1 + A^- AA^- - A^- AA^- \\ &\quad = G_1. \end{aligned}$$

即对任给的 G_1，存在 $V_1, W_1 \in C^{n\times m}$ 使(5.5)式成立.

5. 广义逆 A^- 的性质

性质 1 $(A^-)^* = (A^*)^-$.

证 因为

$$AA^-A = A,$$

所以

$$A^* = (AA^-A)^* = A^*(A^-)^*A^*,$$

由 g 逆定义，有

$$(A^*)^- = (A^-)^*.$$

性质 2 $A(A^*A)^-(A^*A)=A$（或$(A^*A)(A^*A)^-A^*=A^*$）.

证 首先注意这样一个事实：对任何矩阵 A，若 $A^*A=0$，则 $A=0$. 因此，要证明性质 2，只需证明

$$[A(A^*A)^-(A^*A)-A]^*[A(A^*A)^-(A^*A)-A]=0$$

即可. 实事上，

$$\begin{aligned}
&[A(A^*A)^-(A^*A)-A]^*[A(A^*A)^-(A^*A)-A] \\
&\quad=[A^*A(A^*A)^-A^*-A^*][A(A^*A)^-(A^*A)-A] \\
&\quad=[A^*A(A^*A)^--I]A^*[A(A^*A)^-(A^*A)-A] \\
&\quad=[A^*A(A^*A)^--I][(A^*A)(A^*A)^-(A^*A)-A^*A] \\
&\quad=[A^*A(A^*A)^--I][A^*A-A^*A] \\
&\quad=0.
\end{aligned}$$

性质 3 $AGA=A \Longleftrightarrow A^*AGA=A^*A$.

证 若 $AGA=A$，则 $A^*AGA=A^*A$. 反之，若 $A^*AGA=A^*A$，则

$$\begin{aligned}
&(AGA-A)^*(AGA-A) \\
&\quad=(A^*G^*A^*-A^*)(AGA-A) \\
&\quad=(A^*G^*-I)(A^*AGA-A^*A) \\
&\quad=0,
\end{aligned}$$

所以有

$$AGA = A.$$

性质 4 $\operatorname{rank} A^- \geqslant \operatorname{rank} A$.

证 由不等式

$$\operatorname{rank}(AB) \leqslant \min\{\operatorname{rank} A, \operatorname{rank} B\}$$

及等式

$$A = AA^- A$$

即得.

6. 反射 g 逆

对于满秩矩阵 A 有

$$(A^{-1})^{-1} = A,$$

对 g 逆此式一般不成立，但对某些特殊的 g 逆，这种反射性质却是成立的.

定义 3 设 $G \in C^{n\times m}$ 满足

$$AGA = A \qquad ①$$

和

$$GAG = G \qquad ②$$

则称 G 为 A 的一个反射 g 逆，记成 A_r^-，其全体记成 $A\{1,2\}$.

显然，反射 g 逆必为 g 逆，但反之不真，即 $A\{1,2\} \subset A\{1\}$.

对于反射 g 逆来说，有

$$(A_r^-)_r^- = A.$$

事实上，由 $AGA=A$ 有 $G=A^-$，又由 $GAG=G$ 有 $A^-AA^-=A^-$，所以 $(A^-)^-=A$.

按定义易知：

(1) 若 A 是行满秩，则 A 的右逆 A_R^{-1} 是 A 的反射 g 逆，即 $A_R^{-1}=A_r^-$；

若 A 是列满秩，则 A 的左逆 A_L^{-1} 是 A 的反射 g 逆，即 $A_L^{-1} = A_r^-$；

若 A 满秩分解成 $A=CD$，则 $A^- = D_R^{-1}C_L^{-1}$ 是 A 的反射 g 逆.

(2) 若已知 A 的一个 g 逆 A^-，则 $G=A^-AA^-$ 是 A 的反射 g 逆；

若已知 A 的两个 g 逆 G_1,G_2，则 $G=G_1AG_2$ 是 A 的反逆 g 逆.

证 (1) 由右、左逆性质即得.

(2) 只证后一种情况. 设 G_1,G_2 为 A 的两个 g 逆，则

$$AGA = AG_1AG_2A = AG_2A = A;$$

$$GAG = G_1AG_2AG_1AG_2 = G_1AG_1AG_2 = G_1AG_2 = G.$$

7. 广义逆 A^- 的计算

设 $A\in C^{m\times n}$，分两种情况讨论：

(1)若 A 是行(或列)满秩，则

$A^- = A_R^{-1} = A^*(AA^*)^{-1}$(或 $A^- = A_L^{-1} = (A^*A)^{-1}A^*$)；

(2)若 A 既非行满秩又非列满秩，即 rank $A=r<\min\{m,n\}$，这时有三种方法计算 A 的 g 逆 A^-.

方法 1　将 A 满秩分解成 $A=CD$，其中 C 为列满秩，D 为行满秩，则

$$A^- = D_R^{-1}C_L^{-1},$$

其中，$D_R^{-1} = D^*(DD^*)^{-1}$，$C_L^{-1} = (C^*C)^{-1}C^*$.

方法 2　对 A 仅施行一系列行初等变换，即左乘一个可逆矩阵 P，使得

$$PA = \begin{pmatrix} C \\ 0 \end{pmatrix},$$

其中 C 为 $r\times n$ 阶矩阵，rank $C=r$. 把 $\begin{pmatrix} C \\ 0 \end{pmatrix}$ 记为 A_1，则 A 的 g 逆为

$$A^- = A_1^- P,$$

其中，$A_1^- = (C_R^{-1}\quad 0)_{n\times m}$，$C_R^{-1} = C^*(CC^*)^{-1}$.

证　先证 $A_1^- = (C_R^{-1}\quad 0)$.

$$A_1A_1^-A_1 = \begin{pmatrix} C \\ 0 \end{pmatrix}(C_R^{-1}\quad 0)\begin{pmatrix} C \\ 0 \end{pmatrix} = \begin{pmatrix} C \\ 0 \end{pmatrix}(C_R^{-1}C)$$

$$= \begin{pmatrix} CC_R^{-1}C \\ 0 \end{pmatrix} = \begin{pmatrix} C \\ 0 \end{pmatrix} = A_1.$$

再证 $A^- = A_1^- P$（注意 $A = P^{-1}A_1$）.

$$AA^- A = P^{-1}A_1A_1^- PP^{-1}A_1 = P^{-1}A_1A^- A_1 = P^{-1}A_1 = A.$$

方法 3　对矩阵 A 仅施行列初等变换，即右乘一个可逆矩阵 Q，使得

$$AQ = A_1 = (C \quad 0),$$

其中 C 是 $m\times r$ 阶矩阵，rank $C=r$，则 A 的 g 逆为

$$A^- = QA_1^-,$$

其中，$A_1^- = \begin{pmatrix} C_L^{-1} \\ 0 \end{pmatrix}_{n\times m}$，$C_L^{-1} = (C^*C)^{-1}C^*$.

证　先证 $A_1^- = \begin{pmatrix} C_L^{-1} \\ 0 \end{pmatrix}$.

$$A_1A_1^- A_1 = (C \quad 0)\begin{pmatrix} C_L^{-1} \\ 0 \end{pmatrix}(C \quad 0)$$

$$= (CC_L^{-1}C \quad 0) = (C \quad 0) = A;$$

再证 $A^- = QA_1^-$（注意 $A = A_1Q^{-1}$）.

$$AA^- A = A_1Q^{-1}QA_1^- AQ^{-1} = A_1A_1^- A_1Q^{-1}$$

$$= A_1Q^{-1} = A.$$

例 3　设

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 2 \\ 2 & 4 & 0 \end{pmatrix},$$

求 A 的 g 逆 A^-.

解　rank $A=2<3$，所以 A 既非行满秩又非列满秩，属于第二种情况. 现用三种方法来解：

解法 1　用第三种方法，对 A 仅施行列初等变换. 求 A_1 和 Q 可以如下一并进行：

$$\begin{pmatrix} A \\ I \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 2 \\ 2 & 4 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 2 & 0 & 0 \\ 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \\ 1 & 0 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} A_1 \\ Q \end{pmatrix},$$

求得

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{pmatrix} = (C \quad 0) = AQ,$$

$$Q = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 0 \end{pmatrix},$$

$$\begin{aligned} C_L^{-1} &= (C^* C)^{-1} C^* \\ &= \left[\begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 0 \end{pmatrix} \right]^{-1} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 5 & 0 \\ 0 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \end{pmatrix} \\ &= \frac{1}{20} \begin{pmatrix} 4 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \end{pmatrix} \\ &= \frac{1}{20} \begin{pmatrix} 4 & 0 & 8 \\ 0 & 10 & 0 \end{pmatrix} \\ &= \frac{1}{10} \begin{pmatrix} 2 & 0 & 4 \\ 0 & 5 & 0 \end{pmatrix}, \end{aligned}$$

于是

$$A_1^- = \begin{pmatrix} C_L^{-1} \\ 0 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 2 & 0 & 4 \\ 0 & 5 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

故

$$\begin{aligned} A^- &= QA_1^- \\ &= \frac{1}{10}\begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 2 & 0 & 4 \\ 0 & 5 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \frac{1}{10}\begin{pmatrix} 2 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix}. \end{aligned}$$

解法 2 用第二种方法，对 A 仅施行行初等变换．求 A_1 和 P 可如下一并进行：

$$\begin{aligned} (A \quad I) &= \begin{pmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 \\ 2 & 4 & 0 & 0 & 0 & 1 \end{pmatrix} \\ &\to \begin{pmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 & 0 & 1 \end{pmatrix} \\ &= (A_1, \quad P), \end{aligned}$$

求得

$$A_1 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} C \\ 0 \end{pmatrix} = PA,$$

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

$$C_R^{-1} = C^*(CC^*)^{-1}$$

$$= \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 2 \end{bmatrix} \left[\begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 2 \end{bmatrix} \right]^{-1}$$

$$= \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 4 \end{pmatrix}^{-1}$$

$$= \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 2 \end{bmatrix} \cdot \frac{1}{20} \begin{pmatrix} 4 & 0 \\ 0 & 5 \end{pmatrix}$$

$$= \frac{1}{10} \begin{bmatrix} 2 & 0 \\ 4 & 0 \\ 0 & 5 \end{bmatrix},$$

于是

$$A_1^- = (C_K^{-1} \quad 0) = \frac{1}{10} \begin{bmatrix} 2 & 0 & 0 \\ 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix},$$

故

$$\begin{aligned} A^- &= A_1^- P \\ &= \frac{1}{10} \begin{bmatrix} 2 & 0 & 0 \\ 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \\ &= \frac{1}{10} \begin{bmatrix} 2 & 0 & 0 \\ 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix}. \end{aligned}$$

解法 3 由方法 1，将 A 满秩分解成 $A=CD$，则 $A^- = D_R^{-1}C_L^{-1}$. 为此对 A 施行一系列初等变换.

先作行初等变换：

$$(A \quad I) = \begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 \\ 2 & 4 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\to \begin{pmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 & 0 & 1 \end{pmatrix}$$

$$= (A_1 \quad P),$$

其中

$$A_1 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} = PA,$$

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix};$$

再作列初等变换：

$$\begin{pmatrix} A_1 \\ I \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \\ 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} A_2 \\ Q \end{pmatrix},$$

其中

$$A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} A_3 & 0 \\ 0 & 0 \end{pmatrix} = A_1Q = PAQ,$$

$$Q = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

进一步计算得

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

$$Q^{-1} = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$A = P^{-1}A_2Q^{-1} = P^{-1}\begin{pmatrix} A_3 & 0 \\ 0 & 0 \end{pmatrix}Q^{-1}$$

$$= P^{-1}\begin{pmatrix} A_3 \\ 0 \end{pmatrix}(I \quad 0)Q^{-1}.$$

取

$$C = P^{-1}\begin{pmatrix} A_3 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 0 \end{pmatrix},$$

$$D = (I \quad 0)Q^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

于是

$$D_R^{-1} = D^*(DD^*)^{-1}$$

$$= \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{pmatrix}\left[\begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{pmatrix}\right]^{-1}$$

$$= \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{pmatrix} \cdot \frac{1}{5}\begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$$

$$=\frac{1}{5}\begin{pmatrix}1&0\\2&0\\0&5\end{pmatrix},$$

$$\begin{aligned}C_L^{-1}&=(C^*C)^{-1}C^*\\&=\left[\begin{pmatrix}1&0&2\\0&2&0\end{pmatrix}\begin{pmatrix}1&0\\0&2\\2&0\end{pmatrix}\right]^{-1}\begin{pmatrix}1&0&2\\0&2&0\end{pmatrix}\\&=\begin{pmatrix}5&0\\0&4\end{pmatrix}^{-1}\begin{pmatrix}1&0&2\\0&2&0\end{pmatrix}\\&=\frac{1}{20}\begin{pmatrix}4&0\\0&5\end{pmatrix}\begin{pmatrix}1&0&2\\0&2&0\end{pmatrix}\\&=\frac{1}{10}\begin{pmatrix}2&0&4\\0&5&0\end{pmatrix},\end{aligned}$$

所以

$$\begin{aligned}A^-&=D_R^{-1}C_L^{-1}\\&=\frac{1}{5}\begin{pmatrix}1&0\\2&0\\0&5\end{pmatrix}\cdot\frac{1}{10}\begin{pmatrix}2&0&4\\0&5&0\end{pmatrix}\\&=\frac{1}{50}\begin{pmatrix}2&0&4\\4&0&8\\0&25&0\end{pmatrix}.\end{aligned}$$

由此可见，用方法 1 的计算程序是：对 A 施行初等变换，把它变为 $A_1=\begin{pmatrix}A_2&0\\0&0\end{pmatrix}$，即求可逆矩阵 P,Q，使得 $PAQ=A_1$，即 $A=P^{-1}A_1Q^{-1}$，然后令

$$C=P^{-1}\begin{pmatrix}A_2\\0\end{pmatrix},\quad D=(I\quad 0)Q^{-1},$$

并计算

$$C_L^{-1}=(C^*C)^{-1}C^*,\quad D_R^{-1}=D^*(DD^*)^{-1},$$

于是

$$A^- = D_R^{-1}C_L^{-1}.$$

用方法 1 计算时，要求出 P,Q,P^{-1},Q^{-1}，计算量虽然较大，但具有的特征较多；而方法 2 和方法 3 的计算量较小，具有的特征也少.

下例用方法 2 和方法 3 来求 g 逆.

例 4 设矩阵

$$A = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix},$$

求 A^-.

解 rank $A=2$，属于第二种情况.

解法 1 对 A 仅施行列初等变换，计算次序为：

$$A = (C \quad 0) = A_1,$$

$$A_1^- = \begin{pmatrix} C_L^{-1} \\ 0 \end{pmatrix},$$

$$A^- = QA_1^-.$$

下面先求出 A_1 及 Q.

$$\begin{pmatrix} A \\ I \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

故

$$A_1 = \begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} = (C \quad 0),$$

$$Q = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

其中

$$C = \begin{pmatrix} 0 & 2 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

于是

$$\begin{aligned} C_L^{-1} &= (C^* C)^{-1} C^* \\ &= \left(\begin{pmatrix} 0 & 1 & 0 & 1 \\ 2 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \right)^{-1} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 2 & 0 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 \\ 1 & 6 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 2 & 0 & 1 & 1 \end{pmatrix} \\ &= \frac{1}{11} \begin{pmatrix} 6 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 2 & 0 & 1 & 1 \end{pmatrix} \\ &= \frac{1}{11} \begin{pmatrix} -2 & 6 & -1 & 5 \\ 4 & -1 & 2 & 1 \end{pmatrix}, \end{aligned}$$

$$A_1^- = \frac{1}{11} \begin{pmatrix} -2 & 6 & -1 & 5 \\ 4 & -1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\begin{aligned} A^- &= Q A_1^- \\ &= \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \cdot \frac{1}{11} \begin{pmatrix} -2 & 6 & -1 & 5 \\ 4 & -1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

$$= \frac{1}{11}\begin{pmatrix} -2 & 6 & -1 & 5 \\ 0 & 0 & 0 & 0 \\ 4 & -1 & 2 & 1 \end{pmatrix}.$$

解法 2　对 A 仅施行行初等变换，计算次序为

$$PA = \begin{pmatrix} C \\ 0 \end{pmatrix} = A_1,$$

$$A_1^{-1} = (C_R^{-1} \quad 0),$$

$$A^- = A_1^- P.$$

下面先求出 A_1 和 P.

$$(A \quad I) = \begin{pmatrix} 0 & 0 & 2 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\to \begin{pmatrix} 0 & 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\to \begin{pmatrix} 0 & 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\to \begin{pmatrix} 0 & 0 & 0 & 1 & 2 & 0 & -2 \\ 0 & 0 & -1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 & 2 & 0 & -2 \end{pmatrix},$$

故

$$A_1 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} C \\ 0 \end{pmatrix},$$

$$P = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & 1 & -1 \\ 1 & 2 & 0 & -2 \end{pmatrix}.$$

其中

$$C = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

于是

$$\begin{aligned} C_R^{-1} &= C^*(CC^*)^{-1} \\ &= \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & -1 \end{pmatrix} \left(\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & -1 \end{pmatrix} \right)^{-1} \\ &= \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & -1 \end{pmatrix} \cdot \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & -2 \end{pmatrix}, \end{aligned}$$

$$A_1^- = (C_R^{-1} \quad 0) = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & -2 & 0 & 0 \end{pmatrix},$$

$$A^- = A_1^- P$$

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

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

8. 用 A^- 表示相容方程的通解

我们知道，相容的线性方组必有解，问题是它的所有解如何表示. 下面给出用系数矩阵的 g 逆表示的通解.

定理 5（相容方程组的通解） 相容方程组

$$
AX = b
$$

的通解为

$$
X = A^- b + (I_n - A^- A)Y, \tag{5.6}
$$

其中 Y 为 C^n 中的任意元素.

证 首先证明(5.6)式中的 X 确为方程组的解. 事实上，由于 $AX=b$ 为相容方程组，故必有 X_0，使得 $AX_0=b$，因此对(5.6)式中的 X，有

$$
\begin{aligned}
AX &= A[A^- b + (I_n - A^- A)Y] \\
&= AA^- AX_0 - (A - AA^- A)Y \\
&= AX_0 = b.
\end{aligned}
$$

其次证明，对 $AX=b$ 的任何一个解 X_0，都有 $Y\in C^n$，使 X_0 表示成(5.6)的形式. 事实上，取 $Y=X_0$，有

$$
\begin{aligned}
A^- b + (I_n - A^- A)X_0 &= A^- b + X_0 - A^- AX_0 \\
&= A^- b + X_0 - A^- b \\
&= X_0.
\end{aligned}
$$

特别地，当 $b=0$ 时，$AX=b=0$ 即为齐次线性方程组，而齐次线性方程组总是有解的，因此，有如下的结果：

推论 齐次线性方程组

$$AX = 0$$

的通解为

$$X = (I_n - A^- A)Y, \tag{5.6'}$$

其中 Y 为 C^n 中任意元素.

注 由于 A^-b 是相容方程组的解,由(5.6)和(5.6)′可得相容方程组的结构为:

线性非齐次方程组的通解为它的一个特解加上对应的齐次方程组的通解.

例 5 求解线性方程组

$$\begin{cases} x_1 + 2x_2 - x_3 = 1, \\ -x_2 + 2x_3 = 2. \end{cases}$$

解 系数矩阵和常数项矩阵为:

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

$$\operatorname{rank} A = \operatorname{rank}(A \quad b) = 2,$$

故方程组相容,其通解为

$$X = A^- b + (I - A^- A)Y \qquad (Y \in C^3).$$

因为 A 为行满秩,所以

$$\begin{aligned}
A^- = A_R^{-1} &= A^*(AA^*)^{-1} \\
&= \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ -1 & 2 \end{pmatrix} \left[\begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ -1 & 2 \end{pmatrix} \right]^{-1} \\
&= \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 6 & -4 \\ -4 & 5 \end{pmatrix}^{-1} \\
&= \begin{pmatrix} 1 & 2 \\ 2 & -1 \\ -1 & 2 \end{pmatrix} \cdot \frac{1}{14} \begin{pmatrix} 5 & 4 \\ 4 & 6 \end{pmatrix}
\end{aligned}$$

$$
=\frac{1}{14}\begin{pmatrix}5 & 4\\ 6 & 2\\ 3 & 8\end{pmatrix},
$$

从而

$$
\begin{aligned}
A^- A &= \frac{1}{14}\begin{pmatrix}5 & 4\\ 6 & 2\\ 3 & 8\end{pmatrix}\begin{pmatrix}1 & 2 & -1\\ 0 & -1 & 2\end{pmatrix}\\
&= \frac{1}{14}\begin{pmatrix}5 & 6 & 3\\ 6 & 10 & -2\\ 3 & -2 & 13\end{pmatrix}.
\end{aligned}
$$

因此,通解为

$$
\begin{aligned}
X &= A^- b + (I - A^- A)Y\\
&= \frac{1}{14}\begin{pmatrix}5 & 4\\ 6 & 2\\ 3 & 8\end{pmatrix}\begin{pmatrix}1\\ 2\end{pmatrix} + \left(I - \frac{1}{14}\begin{pmatrix}5 & 6 & 3\\ 6 & 10 & -2\\ 3 & -2 & 13\end{pmatrix}\right)\begin{pmatrix}c_1\\ c_2\\ c_3\end{pmatrix}\\
&= \frac{1}{14}\begin{pmatrix}13\\ 10\\ 19\end{pmatrix} + \frac{1}{14}\begin{pmatrix}9 & -6 & -3\\ -6 & 4 & 2\\ -3 & 2 & 1\end{pmatrix}\begin{pmatrix}c_1\\ c_2\\ c_3\end{pmatrix},
\end{aligned}
$$

其中,$Y=(c_1,c_2,c_3)^T$,或

$$
\begin{cases}
x_1 = \dfrac{1}{14}(13 + 9c_1 - 6c_2 - 3c_3),\\[2ex]
x_2 = \dfrac{1}{14}(10 - 6c_1 + 4c_2 - 2c_3),\\[2ex]
x_3 = \dfrac{1}{14}(19 - 3c_1 + 2c_2 + c_3),
\end{cases}
$$

其中,c_1,c_2,c_3 为任意常数.

例 6 求解线性方程组

$$
\begin{cases}
2x_3 = 2\\
x_1 + x_2 = 0\\
x_3 = 1,\\
x_1 + x_2 + x_3 = 1.
\end{cases}
$$

解 由所给方程组可知

$$A=\begin{pmatrix}0&0&2\\1&1&0\\0&0&1\\1&1&1\end{pmatrix},\quad b=\begin{pmatrix}2\\0\\1\\1\end{pmatrix},$$

$$\operatorname{rank} A=\operatorname{rank}(A\quad b)=2,$$

故方程组相容. 利用例 4 解法 2 的结果, 有

$$A^{-}=\begin{pmatrix}0&\frac{1}{2}&0&0\\0&\frac{1}{2}&0&0\\0&-1&0&1\end{pmatrix},$$

代入通解的公式

$$X=A^{-}b+(I-A^{-}A)Y,$$

即得

$$X=\begin{pmatrix}0\\0\\1\end{pmatrix}+\begin{pmatrix}\frac{1}{2}&-\frac{1}{2}&0\\-\frac{1}{2}&\frac{1}{2}&0\\0&0&0\end{pmatrix}\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix},$$

即

$$\begin{cases}x_1=\frac{1}{2}c_1-\frac{1}{2}c_2,\\x_2=-\frac{1}{2}c_1+\frac{1}{2}c_2,\\x_3=1,\end{cases}$$

其中, c_1, c_2 为任意常数.

例 7 求齐次线性方程组

$$\begin{cases}x_1+2x_2-x_3=0,\\-x_2+2x_3=0\end{cases}$$

的通解.

解 由例 5 的结果及公式(5.6)′,得

$$
\begin{aligned}
X &= (I - A^{-}A)Y \\
&= \left[I - \frac{1}{14}\begin{pmatrix} 5 & 4 \\ 6 & 2 \\ 3 & 8 \end{pmatrix}\begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \right]\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \\
&= \frac{1}{14}\begin{pmatrix} 9c_1 - 6c_2 - 3c_3 \\ -6c_1 + 4c_2 + 2c_3 \\ -3c_1 + 2c_2 + c_3 \end{pmatrix},
\end{aligned}
$$

其中,$Y=(c_1,c_2,c_3)^T$ 为 C^3 中任意元素.

§5.2 相容方程组的极小范数解和广义逆 A_m^-

1. 相容方程组的极小范数解与广义逆 A_m^- 的引出

前面讲过,对于相容方程组

$$AX = b, \tag{5.1}$$

其中 $A \in C^{m\times n}$,若能找到 A 的一个 g 逆 A^-,则(5.1)式的通解可表示成

$$X = A^- b + (I - A^- A)Y \qquad (Y \in C^n).$$

现在要在这些解中找出一个解 X,使其范数最小.即要找出(5.1)式的解 X_0,对(5.1)式的任何解 X,都有

$$\|X_0\| \leqslant \|X\|.$$

这样的解 X_0 称为(5.1)式的**极小范数解**.显然,极小范数解 X_0 一般与 b 有关.对于任意 $b\in R(A)$,可以证明极小范数解存在且唯一.现在的问题是如何根据 b 求出 X_0,即是否有这样的矩阵 G,使得对$\forall\ b\in R(A)$,$X_0=Gb$ 为(5.1)式的极小范数解?这样的 G 存在且可能不止一个(注意,G 的不唯一并不与最小范数解的唯一性矛盾,因为对不同的矩阵 A 和 B,对某些向量 b,仍可能有 $Ab=Bb$).

为此，让我们先来考察，如果 G 存在，则 G 应该有哪些特征？

定理 1 设 $G\in C^{n\times m}$，对任意 $b\in R(A)$，使 Gb 都是方程(5.1)的极小范数解的充要条件是，G 满足

$$AGA = A \qquad ①$$

和

$$(GA^{*}) = GA. \qquad ①$$

引理 1 设 $G\in A\{1\}$，对任意 $b\in R(A)$，$X=Gb$ 是方程 $AX=b$ 的极小范数解的充要条件是，对任意的 $Y,Z\in C^{n}$，G 应使

$$Z^{*}(GA)^{*}(I-GA)Y=0. \qquad (5.7)$$

证 必要性. 对任意 $b\in R(A)$，或对任意 $Z\in C^{n}$，使 $b=AZ$. 设 $X_0=Gb$ 是 $AX=b$ 的极小范数解，因为 $G\in A\{1\}$，故方程的通解为

$$\begin{aligned} X &= Gb+(I-GA)Y \\ &= GAZ+(I-GA)Y \qquad (Y\in C^{n}, Z\in C^{n}), \end{aligned}$$

通解 X 的范数

$$\begin{aligned} \|X\|^2 &= (X,X) \\ &= (Gb+(I-GA)Y, Gb+(I-GA)Y) \\ &= \|Gb\|^2+\|(I-GA)Y\|^2 \\ &\quad +2\mathrm{Re}[Gb,(I-GA)Y] \\ &= \|GAZ\|^2+\|(I-GA)Y\|^2 \\ &\quad +2\mathrm{Re}[Z^{*}(GA)^{*}(I-GA)Y], \end{aligned} \qquad (5.8)$$

为了使 Gb 是极小范数解，必须且只需

$$\|(I-GA)Y\|^2+2\mathrm{Re}[Z^{*}(GA)^{*}(I-GA)Y]\geqslant 0. \qquad (5.9)$$

上式中前一项只与 Y 有关，而后一项不仅与 Y 有关，而且与 Z 也有关. 为了对任意的 Y 和 Z 使(5.9)式都成立，则矩阵 G 应使(5.7)式成立. 否则，必有 Y_0,Z_0 使(5.7)式不成立，即

$$Z_0^{*}(GA)^{*}(I-GA)Y_0\neq 0,$$

不妨假设

$$\mathrm{Re}[Z_0^*(GA)^*(I-GA)Y_0]<0,$$

(不然，若 $\mathrm{Re}[Z_0^*(GA)^*(I-GA)Y_0]=0$，则 $\mathrm{Im}[Z_0^*(GA)^*(I-GA)Y_0]\neq 0$，取 $Y_1=\mathrm{i}Y_0$ 或 $-\mathrm{i}Y_0$；若 $\mathrm{Re}[Z_0^*(GA)^*(I-GA)Y_0]>0$，取 $Y_2=-Y_0$，总可使假设成立). 由于 Z_0、Y_0 是固定的，可取 $k>0$ 足够大，使得

$$\|(I-GA)Y_0\|^2+k\cdot 2\mathrm{Re}[Z_1^*(GA)^*I-GA)Y_0]<0,$$

此时令 $Z_1=kZ_0$，$b_1=AZ_1$，则有

$$\|(I-GA)Y_0\|_2+2\mathrm{Re}[Z_0^*(GA)^*(I-GA)Y_0]<0. \tag{5.10}$$

对于这里的 $b_1=AZ_1$，Gb_1 就不是相容方程 $AX=b_1$ 的极小范数解，因为对这个方程的另一个解

$$X_1=Gb_1+(I-GA)Y_0$$

的范数有

$$\|X_1\|^2=\|Gb_1\|^2+\|(I-GA)Y_0\|^2 + 2\mathrm{Re}[Z_1^*(GA)^*(I-GA)Y_0],$$

由(5.10)式知

$$\|X_1\|<\|Gb_1\|.$$

这说明，如果对 $\forall\ Y,Z\in C^n$，(5.7)式不成立，那么 Gb 就可能不是 $AX=b$ 的极小范数解. 这样就证明了必要性.

充分性. 假设对任意 $Y,Z\in C^n$，(5.7)式都成立，那么由(5.8)式知，对 $AX=b$ 的任何解

$$X=Gb+(I-GA)Y$$

的范数 $\|X\|$，都有

$$\|X\|^2=\|Gb\|^2+\|(I-GA)Y\|^2\geqslant\|Gb\|^2,$$

即

$$\|X\|\geqslant\|Gb\|,$$

这说明 Gb 是 $AX=b$ 的极小范数解.

定理 1 的证明 必要性. 设 G 对任意 $b\in R(A)$，都使 $X=Gb$ 是 $AX=b$ 的极小范数解，现要证明 G 具有性质①和④. 因为 $X=Gb$ 对任意 $b\in R(A)$ 都是 $AX=b$ 的解，由定理 1，G 必须具有性质①，即 G 必为 A 的 g 逆. 又由引理 1，对任意 $Y,Z\in C^n$，(5.7)式都成立，因此

$$(GA)^*(I-GA)=0,$$

或

$$(GA)^*=(GA)^*GA,$$

由此有

$$\begin{aligned}GA&=[(GA)^*]^*=[(GA)^*GA]^*\\&=(GA)^*GA=(GA)^*,\end{aligned}$$

即 G 具有性质④.

充分性. 设 $G\in C^{n\times m}$ 具有性质①和④，现要证明对任何 $b\in R(A)$，Gb 都是 $AX=b$ 的极小范数解. 由性质①知，G 是 A 的 g 逆，因此，Gb 是相容方程组 $AX=b$ 的解. 又由于 G 具有性质①和④故有

$$\begin{aligned}(GA)^*(I-GA)&=GA(I-GA)\\&=GA-GAGA\\&=GA-GA=0,\end{aligned}$$

因此，对任意的 $Y,Z\in C^n$，有

$$Z^*(GA)^*(I-GA)Y=0,$$

由引理可知，$X=Gb$ 是 $AX=b$ 的极小范数解.

推论 极小范数解是唯一的.

证 设 X_1 也是方程组 $AX=b$ 的一个极小范数解，即 $\|X_1\|=\|Gb\|$，由(5.6)式所给解的通式，对某个 $Y\in C^n$ 有

$$X_1=Gb+(I-GA)Y, \tag{5.11}$$

由引理及(5.8)式有

$$\|X_1\|^2=\|Gb\|^2+\|(I-GA)Y\|^2.$$

因为 $\|X_1\|=\|Gb\|$，故

$$\|(I-GA)Y\|=0,$$

这样

$$(I-GA)Y=0,$$

因此，由(5.11)式得

$$X_1=Gb+(I-GA)Y=Gb,$$

这就证明了极小范数解 Gb 的唯一性.

定义 1 对任何 $b\in C^m$，使 $X=Gb$ 是相容方程 $AX=b$ 的极小范数解的矩阵 G，称为 A 的**极小范数 g 逆**，记为 A_m^-.

定理 2 矩阵 G 是矩阵 A 的极小范数 g 逆的充要条件是 G 满足：

(1) $AGA=A$；

(2) $(GA)^*=GA$.

注 (1)A 的极小范数 g 逆的全体记为 $A\{1,4\}$；

(2)按定义，$X=A_m^-b$ 是相容方程 $AX=b$ 的极小范数解，虽然 A_m^- 不是唯一的，但极小范数解 A_m^-b 是唯一的.

2. 广义逆 A_m^- 的计算

(1) 当 A 为行(或列)满秩时，则

$$A_m^-=A_R^{-1}=A^*(AA^*)^{-1}$$

$$(\text{或 } A_m^-=A_L^{-1}=(A^*A)^{-1}A^*); \qquad (5.12)$$

(2)当 rank $A=r<\min\{m,n\}$时，将 A 满秩分解为 $A=CD$，其中 C 为列满秩，D 为行满秩，则

$$A_m^-=D_R^{-1}C_L^{-1}. \qquad (5.13)$$

在一般情况下，用满秩分解来求 A_m^- 是很麻烦的，我们可以这样做：

(3)对于 $A\in C^{m\times n}$，有

$$A_m^-=A^*(AA^*)^-. \qquad (5.14)$$

证 设 $G=A^*(AA^*)^-$，要证明 $G=A_m^-$，即

$$AGA=A,\quad (GA)^*=GA.$$

$$(GA)^* = [A^*(AA^*)^- A]^* = A^*[(AA^*)^-]^* A$$
$$= A^*[(AA^*)^*]^- A = A^*(AA^*)^- A = GA,$$
$$(AGA - A)(AGA - A)^*$$
$$= [AA^*(AA^*)^- A - A][AA^*(AA^*)^- A - A]^*$$
$$= [AA^*(AA^*)^- A - A][A^*(AA^*)^- AA^* - A^*]$$
$$= [AA^*(AA^*)^- AA^* - AA^*][(AA^*)^- AA^* - I]$$
$$= (AA^* - AA^*)[(AA^*)^- AA^* - I]$$
$$= 0,$$

即

$$AGA = A.$$

例 1 求方程组 $AX=b$ 的最小范数解，其中

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

解 因为 rank A=rank $(A \quad b)=2$，所以方程组相容. 由于 A 为行满秩，故 $A_m^- = A_R^{-1} = A^*(AA^*)^{-1}$，由 § 5.1 节例 1 知

$$A_m^- = A_R^{-1} = \frac{1}{14}\begin{pmatrix} 5 & 4 \\ 6 & 2 \\ 3 & 8 \end{pmatrix},$$

因此，最小范数解为

$$X = A_m^- b = \frac{1}{14}\begin{pmatrix} 5 & 4 \\ 6 & 2 \\ 3 & 8 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{1}{14}\begin{pmatrix} 13 \\ 10 \\ 19 \end{pmatrix}.$$

例 2 求方程组

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1, \\ x_1 + x_3 = 0, \\ 2x_1 + 2x_3 = 0, \\ 2x_1 + 4x_2 + 6x_3 = 2 \end{cases}$$

的最小范数解与通解.

解 由所给方程组可知

$$A=\begin{pmatrix}1&2&3\\1&0&1\\2&0&2\\2&4&6\end{pmatrix},\quad b=\begin{pmatrix}1\\0\\0\\2\end{pmatrix},$$

由于

$$\operatorname{rank} A=\operatorname{rank}(A\quad b)=2,$$

所以，方程组相容. 因为 $r=2<\min\{3,4\}$，用满秩分解法来求 $A_m^-=D_R^{-1}C_L^{-1}$. 易求得

$$P=\begin{pmatrix}1&0&0&0\\-1&1&0&0\\0&-2&1&0\\-2&0&0&1\end{pmatrix},\quad Q=\begin{pmatrix}1&0&-1\\0&1&-1\\0&0&1\end{pmatrix},$$

使得

$$PAQ=A_1=\begin{pmatrix}1&2&0\\0&-2&0\\0&0&0\\0&0&0\end{pmatrix}=\begin{pmatrix}A_2&0\\0&0\end{pmatrix},$$

其中

$$A_2=\begin{pmatrix}1&2\\0&-2\end{pmatrix}.$$

再求得

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

$$Q^{-1}=\begin{pmatrix}1&0&1\\0&1&1\\0&0&1\end{pmatrix},$$

$$C = P^{-1}\begin{pmatrix} A_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & -2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 0 \\ 2 & 0 \\ 2 & 4 \end{pmatrix},$$

$$D = (I \quad 0)Q^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

$$D_R^{-1} = D^*(DD^*)^{-1} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \\ 1 & 1 \end{pmatrix},$$

$$C_L^{-1} = (C^*C)^{-1}C^* = \frac{1}{10}\begin{pmatrix} 0 & 2 & 4 & 0 \\ 1 & -1 & -2 & 2 \end{pmatrix},$$

$$A_m^- = D_R^{-1}C_L^{-1} = \frac{1}{30}\begin{pmatrix} -1 & 5 & 10 & -2 \\ 2 & -4 & -8 & 4 \\ 1 & 1 & 2 & 2 \end{pmatrix},$$

故得方程组的最小范数解为

$$X = A_m^- b = \frac{1}{30}\begin{pmatrix} -1 & 5 & 10 & -2 \\ 2 & -4 & -8 & 4 \\ 1 & 1 & 2 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{6} \\ \frac{1}{3} \\ \frac{1}{6} \end{pmatrix}.$$

而方程组的通解为

$$X = Gb + (I - GA)Y$$

$$= \begin{pmatrix} -\frac{1}{6} \\ \frac{1}{3} \\ \frac{1}{6} \end{pmatrix} + \left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right.$$

$$-\frac{1}{30}\begin{pmatrix}-1 & 5 & 10 & -2\\ 2 & -4 & -8 & 4\\ 1 & 1 & 2 & 2\end{pmatrix}\begin{pmatrix}\begin{pmatrix}1 & 2 & 3\\ 1 & 0 & 1\\ 2 & 0 & 2\\ 2 & 4 & 6\end{pmatrix}\end{pmatrix}\begin{pmatrix}c_1\\ c_2\\ c_3\end{pmatrix}$$

$$=\frac{1}{3}\begin{pmatrix}-\frac{1}{2}-c_1+c_2-c_3\\ 1+c_1+c_2-c_3\\ \frac{1}{2}-c_1-c_2+c_3\end{pmatrix}.$$

3. 广义逆的通式

对于 $A\in C^{m\times n}$，A_m^- 不是唯一的，现给出它的通式.

引理 2 设 $A\in C^{m\times n}$，则

$$G\in A\{1,4\}\Longleftrightarrow GA=A_m^-A,\quad G\in C^{n\times m}.$$

证 充分性. 由于

$$AGA=AA_m^-A=A,$$

故 $G\in A\{1\}$，再由

$$(GA)^*=(A_m^-A)^*=A_m^-A=GA,$$

故 $G\in A\{1,4\}$.

必要性. 若 $G\in A\{1,4\}$，即

$$AGA=A,\quad (GA)^*=GA,$$

于是

$$\begin{aligned}A_m^-A&=A_m^-AGA=(A_m^-A)(GA)\\ &=(A_m^-A)^*(GA)^*=A^*(A_m^-)^*A^*G^*\\ &=(AA_m^-A)^*G^*=A^*G^*\\ &=(GA)^*=GA.\end{aligned}$$

定理 3 设 $A\in C^{m\times n}$，则 A 的任何极小范数 g 逆 G 必可表示成

$$G=A_m^-+Z(I-AA_m^-)\qquad (Z\in C^{n\times m}),\qquad (5.15)$$

其中，A_m^- 是 A 的某一个极小范数 g 逆.

证 首先证明对$\forall$ $Z\in C^{n\times m}$，(5.15)式所确定的 $G\in A\{1,4\}$. 事实上，

$$\begin{aligned} GA &= [A_m^- + Z(I - AA_m^-)]A \\ &= A_m^- A + Z(A - AA_m^- A) \\ &= A_m^- A + Z(A - A) \\ &= A_m^- A, \end{aligned}$$

由引理 2 知，$G\in A\{1,4\}$.

再证 $A\{1,4\}$中任何元素都能表成(5.15)式的形式. 由于 $GA = A_m^- A$，于是

$$0 = (G - A_m^-)A = (G - A_m^-)AA_m^-,$$

因而

$$\begin{aligned} G &= A_m^- + G - A_m^- \\ &= A_m^- + (G - A_m^-) - (G - A_m^-)AA_m^- \\ &= A_m^- + (G - A_m^-)(I - AA_m^-). \end{aligned}$$

取 $Z=G-A_m^-$ 即可.

§5.3 矛盾方程组的最小二乘解和广义逆 A_l^-

1. 矛盾方程组的最小二乘解与广义逆矩阵 A_l^- 的引出

设

$$AX = b \qquad (A \in C^{m\times n}, X \in C^n, b \in C^m) \tag{5.16}$$

是不相容方程组(矛盾方程组)，它在一般意义下无解，现在要求这样的解，使它的剩余(误差)范数(C^m 中的 2-范数)为最小：

$$\| AX - b \|_2 = \min,$$

这样的解称为不相容方程组(5.16)的**最小二乘解**.

注 最小二乘解并不是方程组 $AX=b$ 的解.

现在的问题是:是否有这样的矩阵G,对于任意的向量b,都使$X=Gb$为方程组$AX=b$的最小二乘解?下面给出这样的矩阵G的特征.

定理 1 对任何$b\in C^m$,$X=Gb$都是不相容方程$AX=b$的最小二乘解的充要条件是,矩阵$G\in C^{n\times m}$满足

$$AGA=A \tag{①}$$

和

$$(AG)^*=AG. \tag{③}$$

证 必要性.若Gb为不相容方程组$AX=b$的最小二乘解,对任何$X\in C^n$,考虑剩余范数

$$\begin{aligned}\|AX-b\|^2 &= \|AGb-b+A(X-Gb)\|^2 \\ &= \|AGb-b\|^2+\|A(X-Gb)\|^2 \\ &\quad +2\mathrm{Re}[(AGb-b)^*A(X-Gb)],\end{aligned}$$

则

$$\begin{aligned}&\|AX-b\|^2-\|AGb-b\|^2 \\ &=\|A(X-Gb)\|^2 \\ &\quad +2\mathrm{Re}[b^*(AG-I)^*A(X-Gb)].\end{aligned} \tag{5.17}$$

与§5.2节引理1的证法类似,可得:为了对任意b和X,使得

$$\|AX-b\|^2-\|AGb-b\|^2\geqslant 0,$$

当且仅当对任何$b\in C^m$,$X\in C^n$,有

$$b^*(AG-I)^*A(X-Gb)=0.$$

由X,b的任意性,有

$$(AG-I)^*A=0,$$

即

$$(AG)^*A=A. \tag{5.18}$$

右乘G得

$$(AG)^*AG=AG,$$

两边取转置共轭得

$$(AG)^* = [(AG)^* AG]^* = (AG)^* AG = AG,$$

这就是性质③,将它代入(5.18)式即有

$$AGA = A,$$

这就是是性质①.

充分性. 若 G 满足

$$AGA = A, \quad (AG)^* = AG,$$

则有

$$(AG)^* A = AGA = A,$$

于是

$$(AG)^* A - A = [(AG)^* - I]A = (AG - I)^* A = 0,$$

因此,对任何 $X \in C^n$,由(5.17)式有

$$\| AX - b \|^2 - \| AGb - b \|^2 = \| A(X - Gb) \|^2 \geqslant 0,$$

故 Gb 是 $AX=b$ 的最小二乘解.

定义 1 对于 $A \in C^{m \times n}$,若有 $G \in C^{n \times m}$满足

$$AGA = A \qquad ①$$

和

$$(AG)^* = AG, \qquad ③$$

则称为 A 的**最小二乘 g 逆**,记成 A_l^-,其全体记成 $A\{1,3\}$.

2. 广义逆矩阵 A_l^- 的计算

(1) 当 A 行(或列)满秩时,有

$$A_l^- = A_R^{-1} = A^* (AA^*)^{-1}$$

$$(或\ A_l^- = A_L^{-1} = (A^* A)^{-1} A^*). \qquad (5.19)$$

(2) 当 $\operatorname{rank} A = r < \min\{m, n\}$时,将 A 满秩分解成 $A = CD$,其中 C 为列满秩,D 为行满秩,则

$$A_l^- = A^- = D_R^{-1} C_L^{-1}.$$

在一般情况下,用满秩分解法求 A_l^- 比较麻烦,此时可按下法来求:

(3) 对于 $A \in C^{m \times n}$,有

$$A_l^- = (A^* A)^- A^*. \tag{5.20}$$

证 令 $G=(A^*A)^-A^*$，只须验证 G 具有性质①和③即可. 由§5.1节中广义逆 A^- 的性质(3)：

$$AGA = A \Longleftrightarrow A^*AGA = A^*A$$

及

$$A^*AGA = A^*A(A^*A)^- A^*A = A^*A$$

知

$$AGA = A,$$

此即性质①. 又

$$AG = A(A^*A)^- A^*,$$

两边取共轭转置，有

$$\begin{aligned}(AG)^* &= [A(A^*A)^- A^*]^* = A[(A^*A)^-]^* A^* \\ &= A[(A^*A)^*]^- A^* = A(A^*A)^- A^* \\ &= AG,\end{aligned}$$

此即性质③.

例 1 求矛盾方程组

$$\begin{cases} x_1 + 2x_2 = 1, \\ 2x_1 + x_2 = 0, \\ x_1 + x_2 = 0 \end{cases}$$

的最小二乘解.

解 由所给方程组，有

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

方法 1 用求极值的方法求最小二乘解. 设 X 为最小二乘解，由它所产生的误差平方和为

$$\begin{aligned}\varepsilon^2 &= \| AX - b \|^2 \\ &= (AX - b, AX - b) \\ &= (AX - b)^*(AX - b)\end{aligned}$$

$$= \left(\begin{pmatrix}1 & 2\\2 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} - \begin{pmatrix}1\\0\\0\end{pmatrix}\right)^* \left(\begin{pmatrix}1 & 2\\2 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} - \begin{pmatrix}1\\0\\0\end{pmatrix}\right)$$

$$= \begin{pmatrix}x_1 + 2x_2 - 1\\2x_1 + x_2\\x_1 + x_2\end{pmatrix}^* \begin{pmatrix}x_1 + 2x_2 - 1\\2x_1 + x_2\\x_1 + x_2\end{pmatrix}$$

$$= (x_1 + 2x_2 - 1)^2 + (2x_1 + x_2)^2 + (x_1 + x_2)^2,$$

于是

$$\frac{\partial \varepsilon^2}{\partial x_1} = 2(x_1 + 2x_2 - 1) + 2(2x_1 + x_2)\cdot 2 + 2(x_1 + x_2)$$

$$= 12x_1 + 10x_2 - 2$$

$$= 0,$$

$$\frac{\partial \varepsilon^2}{\partial x_2} = 2(x_1 + 2x_2 - 1)\cdot 2 + 2(2x_1 + x_2) + 2(x_1 + x_2)$$

$$= 10x_1 + 12x_2 - 4$$

$$= 0.$$

由此解得

$$x_1 = -\frac{4}{11},\quad x_2 = \frac{7}{11},$$

此即矛盾方程组的最小二乘解.

方法 2　A 为列满秩,故

$$A_l^- = A_L^{-1} = (A^* A)^{-1} A^*$$

$$= \left(\begin{pmatrix}1 & 2 & 1\\2 & 1 & 1\end{pmatrix}\begin{pmatrix}1 & 2\\2 & 1\\1 & 1\end{pmatrix}\right)^{-1}\begin{pmatrix}1 & 2 & 1\\2 & 1 & 1\end{pmatrix}$$

$$= \begin{pmatrix}6 & 5\\5 & 6\end{pmatrix}^{-1}\begin{pmatrix}1 & 2 & 1\\2 & 1 & 1\end{pmatrix}$$

$$= \frac{1}{11}\begin{pmatrix}6 & -5\\-5 & 6\end{pmatrix}\begin{pmatrix}1 & 2 & 1\\2 & 1 & 1\end{pmatrix}$$

$$= \frac{1}{11}\begin{pmatrix} -4 & 7 & 1 \\ 7 & -4 & 1 \end{pmatrix},$$

由此得最小二乘解为

$$X = A_l^- b = \frac{1}{11}\begin{pmatrix} -4 & 7 & 1 \\ 7 & -4 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{11}\begin{pmatrix} -4 \\ 7 \end{pmatrix}.$$

这里两种解法所得结果一致.

将 X 代入误差平方的公式得

$$\| AX - b \|^2 = \frac{1}{11}.$$

注 一般来说，最小二乘解不是唯一的，因此两种解法的结果未必一致. 但在 A 的列向量线性无关时，解是唯一的. 证明见后面定理 3 的推论.

例 2 求不相容方程组

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1, \\ x_1 + x_3 = 0, \\ 2x_1 + 2x_3 = 1, \\ 2x_1 + 4x_2 + 6x_3 = 3 \end{cases}$$

的最小二乘解.

解 由所给方程组可知

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 0 & 2 \\ 2 & 4 & 6 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 3 \end{pmatrix},$$

于是 rank $A=2$, rank$(A \quad b)=3$，故此方程组为不相容方程组.

§5.2 节例 2 中的矩阵 A 即为此处的矩阵 A，又由满秩分解所得的 A^- 也是 A_l^-，故由前面的结果知

$$A_l^- = \frac{1}{30}\begin{pmatrix} -1 & 5 & 10 & -2 \\ 2 & -4 & -8 & 4 \\ 1 & 1 & 2 & 2 \end{pmatrix},$$

于是最小二乘解

$$X = A_l^- b = \frac{1}{30}\begin{pmatrix} -1 & 5 & 10 & -2 \\ 2 & -4 & -8 & 4 \\ 1 & 1 & 2 & 2 \end{pmatrix}\cdot\begin{pmatrix} 1 \\ 0 \\ 1 \\ 3 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix},$$

经计算，其最小误差平方和为

$$\| AA_l^- b - b \|^2 = \frac{10}{25}.$$

3. 广义逆矩阵 A_l^- 的通式

如前所述，因为不相容方程组的最小二乘解不是唯一的，因此 A 的广义逆 A_l^- 也不是唯一的. 现在要找出 $A\{1,3\}$ 中元素的通式.

引理 设 $A_l^- \in A\{1,3\}$，则 $G\in A\{1,3\}$ 的充要条件是

$$AG = AA_l^-. \tag{5.21}$$

证 充分性. 设 G 满足(5.21)式，要证明 $G\in A\{1,3\}$. 事实上，在(5.21)式两边右乘 A，由于 $A_l^- \in A\{1\}$，于是有

$$AGA = AA_l^- A = A,$$

故 $G\in A\{1\}$. 又因为 $A_l^- \in A\{3\}$，于是有

$$(AG)^* = (AA_l^-)^* = AA_l^- = AG,$$

故 $G\in A\{3\}$.

必要性. 设 $G\in A\{1,3\}$，要证明 $AG=AA_l^-$. 事实上，

$$\begin{aligned} AA_l^- &= AGAA_l^- = (AG)^*(AA_l^-)^* \\ &= G^* A^* (A_l^-)^* A^* = G^*(AA_l^- A)^* \\ &= G^* A^* = (AG)^* = AG. \end{aligned}$$

定理 2 设 $A\in C^{m\times n}$，$A_l^-\in A\{1,3\}$，则 $A\{1,3\}$中任何元素都可表示成

$$G=A_l^-+(I_n-A_l^-A)Z \qquad (Z\in C^{n\times m}). \qquad (5.22)$$

证 首先证明，对于任何 $Z\in C^{n\times m}$，(5.22)式所确定的 $G\in A\{1,3\}$. 事实上，

$$\begin{aligned} AG &= AA_l^-+A(I_n-A_l^-A)Z \\ &= AA_l^-+(A-AA_l^-A)Z \\ &= AA_l^-, \end{aligned}$$

由引理知，$G\in A\{1,3\}$.

再证明对$\forall\ G\in A\{1,3\}$，存在 $Z\in C^{n\times m}$，使 G 具有(5.22)式的形式. 事实上，取 $Z=G-A_l^-$ 即可. 因为由引理有 $AG=AA_l^-$，所以

$$\begin{aligned} &A_l^-+(I_n-A_l^-A)(G-A_l^-) \\ &\quad = A_l^-+G-A_l^--A_l^-AG+A_l^-AA_l^- \\ &\quad = A_l^-+G-A_l^--A_l^-AA_l^-+A_l^-AA_l^- \\ &\quad = G. \end{aligned}$$

定理 3 不相容方程组 $AX=b$ 的最小二乘解的通式为

$$X=A_l^-b+(I_n-A_l^-A)Y \qquad (Y\in C^n). \qquad (5.23)$$

证 先证(5.23)式中的 X 确为最小二乘解. 因为 A_l^-b 是 $AX=b$ 的最小二乘解，所以 $\|AA_l^-b-b\|$ 取最小值，而

$$\begin{aligned} A[A_l^-b+(I_n-A_l^-A)Y] &= AA_l^-b+(A-AA_l^-A)Y \\ &= AA_l^-b, \end{aligned}$$

所以，$X=A_l^-b+(I-A_l^-A)Y$ 也为最小二乘解.

再证 $AX=b$ 的任一个最小二乘解 X_0 必可表成(5.23)式的形式. 事实上，由于 X_0，A_l^-b 都是方程的最小二乘解，故有

$$\|AX_0-b\|=\|AA_l^-b-b\|=\min.$$

由于 A_l^- 具有性质①和③，故有

$$(AA_l^-)^*=AA_l^-, \quad AA_l^-A=A,$$

因此

$$(AA_l^- - I)^* A = 0,$$

代入(5.17)式，有

$$\begin{aligned}
&\| AX_0 - b \|^2 - \| AA_l^- b - b \|^2 \\
&= \| A(X_0 - A_l^- b) \|^2 \\
&\quad + 2\mathrm{Re}[b^*(AA_l^- - I)^* A(X - A_l^- b)] \\
&= \| A(X_0 - A_l^- b \|^2.
\end{aligned}$$

又由 $\| AX_0 - b \| = \| AA_l^- b - b \|$ 知

$$\| A(X_0 - A_l^- b) \|^2 = 0,$$

于是

$$A(X_0 - A_l^- b) = 0.$$

这说明 $X_0 - A_l^- b$ 为 $AX=0$ 的一个解. 由齐次方程组的通解公式(5.6)知，存在 $Y_0 \in C^n$，使得

$$X_0 - A_l^- b = (I - A_l^- A)Y_0,$$

即

$$X_0 = A_l^- b + (I - A_l^- A)Y_0.$$

推论 当 A 为列满秩时，最小二乘解是唯一的.

证 当 A 为列满秩时，其左逆 $G=(A^*A)^{-1}A^*$ 具有性质①和②. 先证明它是唯一的满足性质③的左逆.

A 的左逆的通式可写为：

$$A_L^{-1} = G + B = (A^*A)^{-1}A^* + B,$$

其中 $B \in C^{n\times m}$，且 $BA=0$.

A_L^{-1} 具有性质③的充要条件是 $(AB)^* = AB$，即

$$\begin{aligned}
(AA_L^{-1})^* = AA_L^{-1} &\Longleftrightarrow [A(G+B)]^* = A(G+B) \\
&\Longleftrightarrow (AB)^* = AB.
\end{aligned}$$

若 B 为非零矩阵，由于 A 为列满秩，则 AB 必为非零矩阵. 于是

$$\begin{aligned}
&[(AB)^* - AB]^*[(AB)^* - AB] \\
&= (AB)(AB)^* + (AB)^*(AB) \\
&\quad - ABAB - (AB)^*(AB)^*
\end{aligned}$$

$$= (AB)(AB)^{*} + (AB)^{*}(AB),$$

上式中用到了 $BA=0$ 这一条件.

由于 AB 为非零矩阵，$(AB)(AB)^{*}$ 为 Hermite 矩阵，其对角线元素为 AB 各行元素的平方和，因此不全为零，且非零者必为正数，$(AB)^{*}(AB)$ 也如此. 这样，$(AB)(AB)^{*}+(AB)^{*}(AB)$ 非零，因此

$$(AB)^{*} - AB \neq 0.$$

所以 $A_l^{-1}=G+B=(A^{*}A)^{-1}A^{*}+B$ 当 $B\neq 0$ 时不满足性质③. 这就证明了 $G=(A^{*}A)^{-1}A^{*}$ 是唯一满足性质③的左逆.

由于此 G 具有性质①和③，由引理知，对任何最小二乘 g 逆 A_l^{-}，有

$$AG = AA_l^{-}.$$

所以

$$\begin{aligned} G - A_l^{-} &= I(G - A_l^{-}) = GA(G - A_l^{-}) \\ &= G(AG - AA_l^{-}) = 0, \end{aligned}$$

这证明了左逆 G 是唯一的最小二乘 g 逆，故(5.23)式中的解为

$$\begin{aligned} X &= A_l^{-}b + (I - A_l^{-}A)Y \\ &= Gb + (I - GA)Y \\ &= Gb + (I - I)Y \\ &= Gb = (A^{*}A)^{-1}A^{*}b, \end{aligned}$$

这就证明了当 A 为列满秩时，其最小二乘解是唯一的，而且证明了其最小二乘 g 逆也是唯一的，就是具有如下形式的左逆：

$$G = (A^{*}A)^{-1}A^{*}.$$

值得注意的是，左逆并不唯一.

在最小二乘法、曲线拟合和多元线性回归分析中常常要计算矛盾方程组的最小二乘解，广义逆矩阵的理论使求矛盾方程组的最小二乘解的方法标准化、规格化了，整个求解过程归结为求 A 的广义逆 A_l^{-}，用不着求误差平方和的极值等一套繁琐的步骤.

§ 5.4 线性方程组的极小最小二乘解和广义逆 A^+

1. 线性方程组的极小最小二乘解及广义逆 A^+ 的引出

在§5.3节中已经知道，对任何 $b\in C^m$，不相容方程组

$$AX = b \tag{5.24}$$

的最小二乘解可以表成

$$X = A_l^- b + (I_n - A_l^- A)Y \qquad (Y \in C^n),$$

而且

$$AX = A[A_l^- b + (I_n - A_l^- A)Y] = AA_l^- b.$$

这说明，对任何 $b\in C^m$，不相容方程组(5.24)的所有最小二乘解都是方程

$$AX = AA_l^- b \tag{5.25}$$

的解. 因为(5.25)式是相容方程组，现在想从(5.24)式的所有最小二乘解中找出一个范数最小的，称之为不相容方程(5.24)的**极小最小二乘解**. 对于这样的解，能否找出一个矩阵 $G\in C^{n\times m}$，使得 $X=Gb$ 就是(5.24)的极小最小二乘解呢？回答是肯定的. 在§5.2节中已知相容方程组 $AX=b$ 的极小范数解为 $X=A_m^- b$，因此，方程组(5.25)的极小范数解为

$$X = A_m^- AA_l^- b. \tag{5.26}$$

如果 $X=A_m^- AA_l^- b$ 还是(5.24)式的最小二乘解，则它就是不相容方程组(5.24)的极小最小二乘解. 事实上正是如此，这可由矩阵 $G=A_m^- AA_l^-$ 具有性质①和③推出. 而且，G 具有更多的性质.

定理 1 矩阵 $G=A_m^- AA_l^-$ 具有下列性质：

(1) $AGA=A$;

(2) $GAG=G$;

(3) $(AG)^*=AG$;

(4) $(GA)^*=GA$.

证 由 A_m^-, A_l^- 的性质,容易知道

$$AGA = AA_m^- AA_l^- A = AA_l^- A = A;$$

$$\begin{aligned} GAG &= A_m^- AA_l^- AA_m^- AA_l^- \\ &= A_m^- AA_m^- AA_l^- = A_m^- AA_l^- = G; \end{aligned}$$

$$AG = AA_m^- AA_l^- = AA_l^-,$$

$$(AG)^* = (AA_l^-)^* = AA_l^- = AG;$$

$$GA = A_m^- AA_l^- A = A_m^- A,$$

$$(GA)^* = (A_m^- A)^* = A_m^- A = GA.$$

同时也说明了 $X=A_m^- AA_l^- b$ 是 $AX=b$ 的极小最小二乘解.

定义 1 设 $A\in C^{m\times n}$,若 $G\in C^{n\times m}$ 满足如下四个性质:

(1) $AGA=A$;

(2) $GAG=G$;

(3) $(AG)^*=AG$;

(4) $(GA)^*=GA$,

则称矩阵 G 为矩阵 A 的 **Moore-Penrose 广义逆或极小最小二乘 g 逆**,记成 A^+.

显然,$A_m^- AA_l^-$ 就是 A 的 Moore-Penrose 广义逆.

定理 2 对于任意 $A\in C^{m\times n}$,其 Moore-Penrose 广义逆 A^+ 存在且唯一.

证 存在性. 显然,$A_m^- AA_l^-$ 就是 A 的 Moore-Penrose 广义逆. 下面证明唯一性. 设 G_1,G_2 都是 A 的 Moore-Penrose 广义逆,则

$$\begin{aligned} G_1 &= G_1AG_1 = G_1AG_2AG_1 \\ &= (G_1A)(G_2A)G_1 = (G_1A)^*(G_2A)^*G_1 \\ &= A^*G_1^*A^*G_2^*G_1 = (AG_1A)^*G_2^*G_1 \\ &= A^*G_2^*G_1 = (G_2A)^*G_1 \\ &= G_2AG_1 = G_2AG_2AG_1 \\ &= G_2(AG_2)^*(AG_1)^* = G_2G_2^*A^*G_1^*A^* \end{aligned}$$

$$
\begin{aligned}
&= G_2 G_2^* (AG_1A)^* = G_2 G_2^* A^* \\
&= G_2 (AG_2)^* = G_2 A G_2 \\
&= G_2.
\end{aligned}
$$

推论 1 A^+的表达式为

$$A^+ = A_m^- A A_l^-.$$

推论 2 不相容方程组 $AX=b$ 的极小最小二乘解为

$$X = A^+ b.$$

关于 Moore-Penrose 广义逆 A^+，必须注意到：A^+具有性质①，即 A^+也是一个 g 逆. 因此，$X=A^+b$ 是相容方程组 $AX=b$ 的一个特解，$X=A^+b+(I-A^+A)Y$，$Y\in C^n$，是其一般解，$G=A^+ + V(I-AA^+)+(I-A^+A)W$，$V,W\in C^{n\times m}$，是 A 的 g 逆的一般表达式. 由于 A^+是 g 逆，所以它具有 g 逆的一切性质，特别地，$(A^+)^*=(A^*)^-$. 由于 A^+又具有性质②，所以 A^+是一个反射 g 逆，即$(A^+)^-=A$. 由 A^+具有性质①和④知，A^+是极小范数 g 逆，$X=A^+b$ 是相容方程组 $AX=b$ 的极小范数解. 又由 A^+具有性质①和③知，A^+是最小二乘 g 逆，$X=A^+b$ 是不相容方程组 $AX=b$ 的最小二乘解，而且是具有极小范数的最小二乘解，即极小最小二乘解.

定理 3 $\operatorname{rank} A=\operatorname{rank} A^+=\operatorname{rank} A^+A=\operatorname{rank} AA^+$.

证 由 $A=AA^+A$，$A^+=A^+AA^+$ 知

$$
\begin{aligned}
\operatorname{rank} A &= \operatorname{rank} AA^+A \leqslant \operatorname{rank} A^+A \leqslant \operatorname{rank} A^+ \\
&= \operatorname{rank} A^+AA^+ \leqslant \operatorname{rank} AA^+ \leqslant \operatorname{rank} A.
\end{aligned}
$$

注 上面我们虽然证明了广义逆 A^+的存在性与唯一性，并且 $\operatorname{rank} A=\operatorname{rank} A^+$，$(A^+)^+=A$，似乎 A^+与 A^{-1}的性质完全一致，其实不然. 例如，若 $A=(0)_{m\times n}$，则 $A^+=(0)_{n\times m}$. 又若 A，B 皆可逆，则$(AB)^{-1}=B^{-1}A^{-1}$，但对于 A^+此性质不成立，即使$(A^2)^+$也未必等于$(A^+)^2$. 例如，设

$$A=\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix},$$

不难验证

$$A^{+}=\frac{1}{2}\begin{pmatrix}1 & 0\\ -1 & 0\end{pmatrix},$$

显然 $A^2=A$，故

$$(A^2)^{+}=A^{+}=\frac{1}{2}\begin{pmatrix}1 & 0\\ -1 & 0\end{pmatrix},$$

但

$$(A^{+})^2=\frac{1}{4}\begin{pmatrix}1 & 0\\ -1 & 0\end{pmatrix}\neq (A^2)^{+}.$$

2. 广义逆矩阵 A^+的常用性质

定理 4 设 $A\in C^{m\times n}$，则

(1) $(A^{+})^{+}=A$；

(2) $(A^{+})^{*}=(A^{*})^{+}$；

(3) $(\lambda A)^{+}=\frac{1}{\lambda}A^{+}$，$\lambda\in C$，$\lambda\neq 0$；

(4) $A^{*}=A^{*}AA^{+}=A^{+}AA^{*}$；

(5) $A=AA^{*}(A^{*})^{+}=(A^{*})^{+}A^{*}A$；

(6) $(A^{*}A)^{+}=A^{+}(A^{*})^{+}$

(7) $A^{+}=(A^{*}A)^{+}A^{*}=A^{*}(AA^{*})^{+}$；

(8) $(UAV)^{+}=V^{*}A^{+}U^{*}$，其中 $U^{*}U=I_m$，$V^{*}V=I_n$；

(9) $A^{+}AB=A^{+}AC\Longleftrightarrow AB=AC$；

(10) $m-\text{rank}\,(I_m-AA^{+})=n-\text{rank}(I_n-A^{+}A)$
$=\text{rank}\,A=\text{rank}\,A^{+}$.

3. 广义逆 A^+的计算方法

下面介绍 A^+的各种算法，有些算法在前面虽已讲过，但为了完整起见，仍分述如下：

(1) 如果 A 为满秩方阵，则 $A^{+}=A^{-1}$；

(2)如果 $A=\mathrm{diag}(d_1,d_2,\cdots,d_n)$，$d_i\in C$，$i=1,2,\cdots,n$，则

$$A^+=\mathrm{diag}(d_1^+,d_2^+,\cdots,d_n^+),$$

其中

$$d_i^+=\begin{cases}0,当\ d_i=0\ 时,\\ \dfrac{1}{d_i},当\ d_i\neq 0\ 时;\end{cases}$$

(3)如果 A 为行满秩矩阵，则

$$A^+=A_R^{-1}=A^*(AA^*)^{-1};$$

(4)如果 A 为列满秩矩阵，则

$$A^+=A_L^{-1}=(A^*A)^{-1}A^*;$$

(5)如果 A 为降秩的 $m\times n$ 矩阵，可用满秩分解求 A^+，即将 A 满秩分解成 $A=CD$，其中 $C\in C^{m\times r}$，$D\in C^{r\times n}$，且

$$\mathrm{rank}\ C=\mathrm{rank}\ D=r=\mathrm{rank}\ A<\min\{m,n\},$$

$$C_L^{-1}=(C^*C)^{-1}C^*,\quad D_R^{-1}=D^*(DD^*)^{-1},$$

则

$$A^+=D_R^{-1}C_L^{-1}.$$

例 1 求方程组

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

的极小最小二乘解.

解 由方程组知

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

$$\mathrm{rank}\ A=2,\quad \mathrm{rank}(A\quad b)=3,$$

故此方程组不相容.按前述，它的极小最小二乘解为 $X=A^+b$，为此先求 A^+.由 §5.1 节例 3 的解法可知

$$A^{+}=\frac{1}{25}\begin{bmatrix}1 & 0 & 2\\ 2 & 0 & 4\\ 0 & 25 & 0\end{bmatrix},$$

因此，方程组的极小最小二乘解为

$$X=A^{+}b=\frac{1}{25}\begin{bmatrix}1 & 0 & 2\\ 2 & 0 & 4\\ 2 & 25 & 0\end{bmatrix}\begin{bmatrix}1\\ 1\\ 3\end{bmatrix}=\frac{1}{25}\begin{bmatrix}7\\ 14\\ 25\end{bmatrix}.$$

例 2 求不相容方程组

$$\begin{cases}x_1+2x_2+3x_3=1,\\ x_1+x_3=0,\\ 2x_1+2x_3=1,\\ 2x_1+4x_2+6x_3=3\end{cases}$$

的极小最小二乘解.

解 由§5.3节例2知

$$A^{+}=\frac{1}{30}\begin{bmatrix}-1 & 5 & 10 & -2\\ 2 & -4 & -8 & 4\\ 1 & 1 & 2 & 2\end{bmatrix},$$

故方程组的极小最小二乘解为

$$X=A^{+}b=\frac{1}{30}\begin{bmatrix}-1 & 5 & 10 & -2\\ 2 & -4 & -8 & 4\\ 1 & 1 & 2 & 2\end{bmatrix}\begin{bmatrix}1\\ 0\\ 1\\ 3\end{bmatrix}=\frac{1}{10}\begin{bmatrix}1\\ 2\\ 3\end{bmatrix}.$$

注意，这时 $\|AX-b\|=\dfrac{\sqrt{10}}{5}$，$\|X\|=\dfrac{\sqrt{14}}{10}$ 均为最小.

请读者思考一下，若给定线性方程组 $AX=b$，在各种情况下，$X=A^{+}b$ 代表什么？这时你对 Moore-Penrose 广义逆 A^{+} 能说些什么？

习 题 五

1. 证明 AA^- 与 A^-A 均为幂等矩阵，即

$$(A^-A)^2 = A^-A, \quad (AA^-)^2 = AA^-.$$

2. 设 $A \in C^{m\times n}$，试证：

$$(A^-)^T \in A^T\{1\}.$$

3. 设 $A \in C^{m\times n}$，则

(1) $A^-A = I_n \Longleftrightarrow \text{rank } A = n$；

(2) $AA^- = I_m \Longleftrightarrow \text{rank } A = m$.

4. 证明：

(1) $\text{rank }(A^-A) = \text{rank } A$；

(2) $\text{rank }(A_r^-) = \text{rank } A$.

5. 验证：

(1) $A_m^- = A^*(AA^*)^-$；

(2) $A_l^- = (A^*A)^-A^*$.

6. 求下列矩阵的 g 逆 A^-：

(1) $A_1 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \\ 1 & 0 & 2 \end{pmatrix}$；

(2) $A_2 = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}$.

并分别求方程 $A_1X = b_1$ 和 $A_2X = b_2$ 的通解，其中：

$$b_1 = (1,0,1,1)^T, \quad b_2 = (2,1,1)^T.$$

7. 求下列矩阵的极小范数 g 逆 A_m^-：

(1) $A_1 = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 4 \end{pmatrix}$；

(2) $A_2=\begin{pmatrix}1 & 0 & 3\\ 2 & 3 & 0\\ 1 & 1 & 1\end{pmatrix}$.

并分别求方程 $A_1X=b_1$ 和 $A_2X=b_2$ 的最小范数解，其中：

$$b_1=(1,-1)^T,\quad b_2=(3,0,1)^T.$$

8. 设矩阵 $G\in A\{1,3\}$，则 $X\in C^n$ 是不相容线性方程组 $AX=b$ 的最小二乘解的充要条件是，对任意 $b\in C^m$，X 是方程组

$$AX=AGb$$

的解.

9. 求下列矩阵的最小二乘 g 逆 A_l：

(1) $A_1=\begin{pmatrix}1 & 2\\ 2 & 1\\ 1 & 1\end{pmatrix}$；

(2) $A_2=\begin{pmatrix}1 & 0 & 1 & 1\\ 2 & 1 & 2 & 1\\ 2 & 0 & 2 & 2\\ 4 & 2 & 4 & 2\end{pmatrix}$.

并分别求不相容方程 $A_1X=b_1$ 和 $A_2X=b_2$ 的最小二乘解，其中：

$$b_1=(1,0,0)^T,\quad b_2=(0,1,0,1)^T.$$

10. 用满秩分解法求下列矩阵的极小最小二乘 g 逆 A^+：

(1) $A_1=\begin{pmatrix}0 & 0 & 1\\ 0 & 0 & 2\\ 1 & 1 & 0\\ 1 & 1 & 1\end{pmatrix}$；

(2) $A_2=\begin{pmatrix}1 & 1 & 2\\ 0 & 2 & 2\\ 1 & 0 & 1\\ 1 & 0 & 1\end{pmatrix}$.

并分别求不相容方程 $A_1X=b_1$ 及 $A_2X=b_2$ 的极小最小二乘解，其中：

$b_1(1,2,1,2)^T$, $\quad b_2=(0,1,0,0)^T$.

11. 证明：线性方程组 $A\xi=\eta$ 有解的充要条件是 $AA^+\eta=\eta$ 和 $\operatorname{rank} A=\operatorname{rank}(A \quad \eta)$.

12. 设 $A\in C^{m\times n}$，$B\in C^{p\times q}$，$C\in C^{m\times q}$，那么矩阵方程 $AXB=C$，相容的充要条件是，对某个 A^-，B^-，有

$$AA^-CB^-B=C$$

成立，且方程的通解为

$$X=A^-CB^-+(Z-A^-AZBB^-) \qquad (\forall\, Z\in C^{m\times p}).$$

13. 设 $A\in R^{m\times n}$，$\operatorname{rank} A=r$，U，V 是正交阵，若

$$A=U\begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}V^T,$$

其中 Δ 为 r 阶下三角阵，证明：

$$A^+=V\begin{pmatrix}\Delta^{-1} & 0\\ 0 & 0\end{pmatrix}U^T.$$

附录一　矩阵乘积的秩

引理　在秩为 r 的 $m\times n$ 阶矩阵 A 中，任取 s 行，得到 $s\times n$ 阶矩阵 B，则 B 的秩不小于 $r+s-m$；同理，任取 A 的 s 列，得到 $m\times s$ 阶矩阵 C，则 C 的秩不小于 $r+s-n$.

证　因 A 中必有 r 个行是线性无关的，而剩下的 $m-r$ 行在 $s>m-r$ 时即使全部包含于 B 内，这时 B 也要包含上述 r 行中的 $r+s-m$ 行，故 B 中至少有 $r+s-m$ 行是线性无关的，即 $\operatorname{rank} B\geqslant r+s-m$，这说明当 $s>m-r$ 时结论成立，当 $s\leqslant m-r$ 时，$r+s-m\leqslant 0$，结论自然成立.

定理 1　设 A，B 为同一数域上的 $m\times n$ 与 $n\times q$ 阶矩阵，则

$$\begin{aligned}\operatorname{rank} A+\operatorname{rank} B-n&\leqslant \operatorname{rank}(AB)\\&\leqslant \min\{\operatorname{rank} A,\operatorname{rank} B\}.\end{aligned}\qquad(1.1)$$

证　显然，方程组 $BX=0$ 的解向量 X 也满足方程组 $(AB)X=0$，记

$$U=\{X\mid BX=0\},V=\{X\mid (AB)X=0\},$$

则 $U\subset V$，于是

$$\dim U=n-\operatorname{rank} B\leqslant n-\operatorname{rank}(AB)=\dim V,$$

即

$$\operatorname{rank}(AB)\leqslant \operatorname{rank} B.$$

又由于

$$\begin{aligned}\operatorname{rank}(AB)=\operatorname{rank}(AB)^T&=\operatorname{rank}(B^TA^T)\\&\leqslant \operatorname{rank} A^T=\operatorname{rank} A,\end{aligned}$$

因此

$$\operatorname{rank}(AB)\leqslant \min\{\operatorname{rank} A,\operatorname{rank} B\}.$$

设 $\operatorname{rank} A=r_1$，$\operatorname{rank} B=r_2$，$\operatorname{rank}(AB)=r$. 注意到，矩阵 A 可

通过初等行变换和初等列变换化为标准形，即存在满秩矩阵 $P\in C^{m\times m}$，$Q\in C^{n\times n}$，使得

$$PAQ=\begin{pmatrix}I_{r_1}&0\\0&0\end{pmatrix}\in C^{m\times n}.$$

设 $Q^{-1}B=\begin{pmatrix}B_1\\B_2\end{pmatrix}$，其中 $B_1\in C^{r_1\times q}$，$B_2\in C^{(n-r_1)\times q}$，且

$$\operatorname{rank}(PAQQ^{-1}B)=\operatorname{rank}(PAB)=\operatorname{rank}(AB)=r,$$

而

$$PAQQ^{-1}B=\begin{pmatrix}I_{r_1}&0\\0&0\end{pmatrix}\begin{pmatrix}B_1\\B_2\end{pmatrix}=\begin{pmatrix}B_1\\0\end{pmatrix},$$

比较上式两端矩阵的秩知道

$$\operatorname{rank}B_1=r;$$

而且 $\operatorname{rank}(Q^{-1}B)=\operatorname{rank}B=r_2$，因为 B_1 为 $Q^{-1}B$ 的前 r_1 行，故由引理知，

$$\operatorname{rank}B_1\geqslant r_1+r_2-n.$$

由上面结果即有

$$\operatorname{rank}(AB)\geqslant\operatorname{rank}A+\operatorname{rank}B-n.$$

定理 2 设 $A\in C^{m\times n}$，则

$$\operatorname{rank}A=\operatorname{rank}AA^*=\operatorname{rank}A^*A.\tag{1.2}$$

证 由定理 1 可知，$\operatorname{rank}(A^*A)\leqslant\operatorname{rank}A$，现在只需证明 $\operatorname{rank}(A^*A)\geqslant\operatorname{rank}A$ 即可．

考虑线性方程组 $A^*AX=0$，设 $X=(x_1,x_2,\cdots,x_n)^T$ 是方程组的一组解．将 $A^*AX=0$ 两边左乘 X^*，得

$$X^*A^*AX=0,$$

即

$$(AX)^*AX=0,$$

所以

$$AX=0.$$

这说明 $A^*AX=0$ 的解都是 $AX=0$ 的解，即

$$\{X\mid A^*AX=0\}\subset\{X\mid AX=0\}.$$

于是

$$n-\operatorname{rank}(A^*A)\leqslant n-\operatorname{rank}A,$$

即

$$\operatorname{rank}(A^*A)=\operatorname{rank}A,$$

并且有

$$\operatorname{rank}AA^*=\operatorname{rank}(A^*)^*A^*=\operatorname{rank}A^*=\operatorname{rank}A.$$

推论 1 若 $A\in R^{m\times n}$，则

$$\operatorname{rank}A=\operatorname{rank}A^TA=\operatorname{rank}AA^T.$$

注意：对复矩阵 A，上式不一定成立. 例如，设 $A=\begin{pmatrix}1 & \mathrm{i}\\ -\mathrm{i} & 1\end{pmatrix}$，则 $\operatorname{rank}A=1$. 由于

$$A^TA=\begin{pmatrix}1 & -\mathrm{i}\\ \mathrm{i} & 1\end{pmatrix}\begin{pmatrix}1 & \mathrm{i}\\ -\mathrm{i} & 1\end{pmatrix}=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix},$$

故 $\operatorname{rank}(A^TA)=0$.

推论 2 设 $A\in C^{m\times n}$（或 $R^{m\times n}$），则

(1) $\operatorname{rank}A=n\Longleftrightarrow\det(A^*A)\neq0$（或 $\det(A^TA)\neq0$），　(1.3)

(2) $\operatorname{rank}A=m\Longleftrightarrow\det(AA^*)\neq0$（或 $\det(AA^T)\neq0$）.　(1.4)

附录二　分块矩阵的逆

设 A 为 n 阶方阵,将 A 写为分块矩阵

$$A=\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix},$$

其中,A_{11}、A_{22}分别为 n_1 阶和 n_2 阶方阵,A_{12}和 A_{21}分别为 $n_1\times n_2$ 和 $n_2\times n_1$ 阶矩阵,且 $n_1+n_2=n$.

为了求 $\det A$,我们先讨论如何将 A 表示成两个三角形矩阵的乘积.

设 A_1 为 n_1 阶非异矩阵,这时可以对 A 进行初等行变换,将 A_{21}化为零矩阵,即对 A 左乘满秩阵

$$\begin{pmatrix}I_{n_1} & 0\\ -A_{21}A_{11}^{-1} & I_{n_2}\end{pmatrix},$$

于是有

$$\begin{pmatrix}I_{n_1} & 0\\ -A_{21}A_{11}^{-1} & I_{n_2}\end{pmatrix}\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}=\begin{pmatrix}A_{11} & A_{12}\\ 0 & A_{22}-A_{21}A_{11}^{-1}A_{12}\end{pmatrix}, \tag{2.1}$$

因此

$$\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}=\begin{pmatrix}I_{n_1} & 0\\ A_{21}A_{11}^{-1} & I_{n_2}\end{pmatrix}\begin{pmatrix}A_{11} & A_{22}\\ 0 & A_{22}-A_{21}A_{11}^{-1}A_{12}\end{pmatrix}, \tag{2.2}$$

即 A 表示成两个分块三角形矩阵的乘积.

(2.2)式右边第二个矩阵还可以再分解成分块对角矩阵和三角形矩阵之积.若 $\det A_{11}\neq 0$ 时,可对(2.2)式右端第二个矩阵进行初等列变换以消去 A_{12},即右乘矩阵

$$\begin{pmatrix} I_{n_1} & -A_{11}^{-1}A_{12} \\ 0 & I_{n_2} \end{pmatrix},$$

于是可得

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

$$= \begin{pmatrix} I_{n_1} & 0 \\ A_{21}A_{11}^{-1} & I_{n_2} \end{pmatrix} \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{pmatrix} \begin{pmatrix} I_{n_1} & A_{11}^{-1}A_{12} \\ 0 & I_{n_2} \end{pmatrix}. \quad (2.3)$$

在(2.3)式两边取行列式，当 $\det A_{11} \neq 0$ 时，有

$$\det A \neq 0 \Longleftrightarrow \det(A_{22} - A_{21}A_{11}^{-1}A_{12}) \neq 0.$$

类似地，当 $\det A_{22} \neq 0$ 时，有

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

$$= \begin{pmatrix} I_{n_1} & A_{12}A_{22}^{-1} \\ 0 & I_{n_2} \end{pmatrix} \begin{pmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & 0 \\ 0 & A_{22} \end{pmatrix} \begin{pmatrix} I_{n_1} & 0 \\ A_{22}^{-1}A_{21} & I_{n_2} \end{pmatrix}. \quad (2.4)$$

因此，当 $\det A_{22} \neq 0$ 时，有

$$\det A \neq 0 \Longleftrightarrow \det(A_{11} - A_{12}A_{22}^{-1}A_{21}) \neq 0.$$

定理 1 将 n 阶矩阵 A 按上述方法分割成

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

则当 $\det A_{11} \neq 0$ 时，

$$\det A = \det A_{11} \cdot \det(A_{22} - A_{21}A_{11}^{-1}A_{12}); \quad (2.5)$$

当 $\det A_{22} \neq 0$ 时，

$$\det A = \det A_{22} \cdot \det(A_{11} - A_{12}A_{22}^{-1}A_{21}). \quad (2.6)$$

推论 1 设 A 为 $m \times n$ 阶矩阵，B 为 $n \times m$ 阶矩阵，则

$$\det(I_m + AB) = \det(I_n + BA).$$

证 由定理 1，考虑：

$$\det\begin{bmatrix} I_n & B \\ -A & I_m \end{bmatrix} = \det(I_m + AB) = \det(I_n + BA).$$

推论 2 若 $A_{11}A_{21}=A_{21}A_{11}$，则

$$\det\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \det(A_{11}A_{22} - A_{21}A_{12}); \tag{2.7}$$

若 $A_{22}A_{12}=A_{12}A_{22}$，则

$$\det\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \det(A_{22}A_{11} - A_{12}A_{21}). \tag{2.8}$$

证 在 $A_{11}A_{21}=A_{21}A_{11}$的情况下，若 $\det A_{11}\neq 0$，由(2.5)式有

$$\begin{aligned}\det\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} &= \det(A_{11}A_{22} - A_{11}A_{21}A_{11}^{-1}A_{12}) \\ &= \det(A_{11}A_{22} - A_{21}A_{11}A_{11}^{-1}A_{12}) \\ &= \det(A_{11}A_{22} - A_{21}A_{12}).\end{aligned}$$

若 $\det A_{11}=0$，考虑 $A_{11}+xI$. 因为必有常数 k，当 $x>k$ 时，恒有 $\det(A_{11}+xI)\neq 0$，且

$$(A_{11} + xI)A_{21} = A_{21}(A_{11} + xI),$$

于是

$$\det\begin{bmatrix} A_{11} + xI & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \det(A_{11}A_{22} + xA_{22} - A_{21}A_{12}),$$

由于这是关于 x 的恒等式，故当 $x=0$ 时仍成立，即有

$$\det\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \det(A_{11}A_{22} - A_{21}A_{12}).$$

同理可得(2.8)式.

下面讨论 A 分块后逆矩阵的求法. 我们知道，A 分块后，可分解成(2.3)式或(2.4)式所示的三个矩阵的乘积. 因此，为了求 A^{-1}，可求出三个矩阵各自的逆矩阵. 由(2.3)式可得

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1}$$

$$
=\begin{bmatrix} I_{n_1} & -A_{11}^{-1}A_{12} \\ 0 & I_{n_2} \end{bmatrix}\begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & (A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix}\begin{bmatrix} I_{n_1} & 0 \\ -A_{21}A_{11}^{-1} & I_{n_2} \end{bmatrix},
\tag{2.9}
$$

由(2.4)式可得

$$
\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1}
$$

$$
=\begin{bmatrix} I_{n_1} & 0 \\ -A_{22}^{-1}A_{21} & I_{n_2} \end{bmatrix}\begin{bmatrix} (A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1} & 0 \\ 0 & A_{22}^{-1} \end{bmatrix}\begin{bmatrix} I_{n_1} & -A_{12}A_{22}^{-1} \\ 0 & I_{n_2} \end{bmatrix}.
\tag{2.10}
$$

定理 2 设

$$
A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},
$$

(1)若 A_{11} 为非奇异矩阵，则 A 是非奇异矩阵的充要条件是 $(A_{22}-A_{21}A_{11}^{-1}A_{12})$ 为非奇异矩阵(由(2.5)式可证).

(2)若 A_{11}，$(A_{22}-A_{21}A_{11}^{-1}A_{12})$，皆为非奇异矩阵，则 A^{-1} 存在，且由(2.9)式有

$$
A^{-1}=\begin{bmatrix} A_{11}^{-1}+A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \\ -(A_{22}+A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & (A_{22}+A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix}.
$$

(3)若 A_{22} 是非奇异矩阵，则 A 为非奇异矩阵的充分必要条件是 $A_{11}-A_{21}A_{11}^{-1}A_{12}$ 为非奇异矩阵.(由(2.6)可证).

(4)若 A_{22}，$A_{11}-A_{12}A_{22}^{-1}A_{21}$ 皆为非奇异矩阵，则 A^{-1} 存在，且由(2.10)式有

$$
A^{-1}=\begin{bmatrix} (A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1} & -(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{21}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1} & A_{22}^{-1}+A_{22}^{-1}A_{21}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1} \end{bmatrix}.
$$

(5)若 A_{11}、A_{22} 均为非奇异矩阵，当矩阵 $A_{22}-A_{21}A_{11}^{-1}A_{12}$ 和 $A_{11}-A_{12}A_{22}^{-1}A_{21}$ 中有一个为非奇异矩阵时，则另一个及 A 也是非奇异矩阵.

推论 若 $A_{21}=0$,则

(1)A_{11}和 A_{22}均为非奇异性矩阵的充要条件是 A 为非奇异矩阵;

(2) $\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}^{-1}=\begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{pmatrix}$.

习题答案与提示

习 题 一

1. (1)是；(2)非，(3)是；(4)是；(5)非；(6)是，$c_i = \frac{1}{\pi}\int_0^{2\pi} x\sin it\mathrm{d}t$.

2. (1)对 $k=1,2,\cdots,n; l=k,k+1,\cdots,n$，令 $F_{kl}=(a_{ij})_{n\times n}$，其中 $a_{kl}=a_{lk}=1$，其余的 $a_{ij}=0$，则$\{F_{kl}\}$为对称矩阵空间的一组基，维数为$\frac{1}{2}n(n+1)$.

对 $k=1,2,\cdots,n-1; l=k+1,k+2,\cdots,n$，令 $G_{kl}=(a_{ij})_{n\times n}$，其中 $a_{kl}=-a_{lk}=1$，其余的 $a_{ij}=0$，则$\{G_{kl}\}$为反对称矩阵空间的一组基，维数为$\frac{1}{2}n(n-1)$.

对 $k=1,2,\cdots,n; l=k,k+1,\cdots,n$，令 $H_{kl}=(a_{ij})_{n\times n}$，其中 $a_{kl}=1$，其余的 $a_{ij}=0$，则$\{H_{kl}\}$为上三角形矩阵空间的一组基，维数为$\frac{1}{2}n(n+1)$.

(2)R^+中任意非零元素都可作R^+的基，$\dim R^+=1$.

(3)I,A,A^2 为线性空间的一组基，维数是 3.

3. (1)是；(2)是；(3)是；(4)是.

4. (1) 当 $n=2k$ 时，基为

$$e_1 = (1,0,0,\cdots,0)^T,$$

$$e_2 = (0,0,1,0,\cdots,0)^T,$$

$$\cdots\cdots$$

$$e_k = (0,0,\cdots,0,1,0)^T,$$

维数为 k；

当 $n=2k+1$ 时，基为

$$e_1 = (1,0,0,\cdots,0)^T,$$
$$e_2 = (0,0,1,0,\cdots,0)^T,$$
$$\cdots\cdots$$
$$e_{k+1} = (0,0,\cdots,0,1)^T,$$

维数为 $k+1$.

(2) 基为 $(1,0,1,0,1,0,\cdots)^T$ 和 $(0,1,0,1,0,1,\cdots)^T$，维数为 2.

5. (1)过渡矩阵 $C=\begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$，$\xi=(\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4)\begin{pmatrix} \frac{3}{13} \\ \frac{5}{13} \\ -\frac{2}{13} \\ -\frac{3}{13} \end{pmatrix}$；

(2)过渡矩阵 $C=\begin{pmatrix} \frac{3}{4} & \frac{7}{4} & \frac{1}{2} & -\frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} & \frac{1}{2} & \frac{3}{4} \\ -\frac{1}{4} & \frac{3}{4} & 0 & -\frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} & 0 & -\frac{1}{4} \end{pmatrix}$，

$$\xi=(\eta_1,\eta_2,\eta_3,\eta_4)\begin{pmatrix} -2 \\ -\frac{1}{2} \\ 4 \\ -\frac{3}{2} \end{pmatrix}.$$

6. $\xi=(c,c,c,-c)^T$，c 为任意非零常数.

7. 维数是 5，一组基为

$$\begin{pmatrix}3&1&0\\0&0&0\\0&0&3\end{pmatrix},\quad \begin{pmatrix}-\frac{1}{3}&0&0\\1&0&0\\0&0&0\end{pmatrix},\quad \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},$$

$$\begin{pmatrix}-\frac{1}{3}&0&0\\0&0&0\\1&0&0\end{pmatrix},\quad \begin{pmatrix}1&0&0\\0&0&0\\0&1&1\end{pmatrix}.$$

8. 提示：证明 $K\cap S=\{0\}$.

9. 提示：同上.

10. 提示：同上.

12. (1) $\dim \mathrm{Span}\{\alpha_1,\alpha_2,\beta_1,\beta_2\}=3$,

$\dim \mathrm{Span}\{\alpha_1,\alpha_2\}=2$,

$\dim \mathrm{Span}\{\beta_1,\beta_2\}=2$,

故交的维数为 $2+2-3=1$，交的一个基为 $(-5,2,3,4)^T$；和的维数为 3，$\{\alpha_1,\alpha_2,\beta_1\}$ 为一组基；

(2) $\dim \mathrm{Span}\{\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2\}=4$,

$\dim \mathrm{Span}\{\alpha_1,\alpha_2,\alpha_3\}=3$,

$\dim \mathrm{Span}\{\beta_1,\beta_2\}=2$,

故交的维数为 1，基为 β_1；和的维数为 4，$\{\alpha_1,\alpha_2,\alpha_3,\beta_2\}$ 为一组基.

13. 提示：证明如此定义的内积满足欧氏空间定义的三个条件.

14. 令 $V\sim(a,b,c)\in R^3$ 即可.

15. 令 $X\sim(a,b)\in R^2$ 即可.

16. $\left(\frac{4}{\sqrt{26}},0,\frac{1}{\sqrt{26}},-\frac{3}{\sqrt{26}}\right)^T$.

17. $\eta_1=\frac{\sqrt{2}}{2}\varepsilon_1-\frac{\sqrt{2}}{2}\varepsilon_2$.

$$\eta_2=\frac{\sqrt{10}}{10}\varepsilon_1-\frac{2\sqrt{10}}{10}\varepsilon_2+\frac{2\sqrt{10}}{10}\varepsilon_4-\frac{\sqrt{10}}{10}\varepsilon_5,$$

$$\eta_3=\frac{1}{2}\varepsilon_1+\frac{1}{2}\varepsilon_2+\frac{1}{2}\varepsilon_3-\frac{1}{2}\varepsilon_5.$$

18. $(\alpha,\beta)=8+5\mathrm{i}$，$\|\alpha\|=\sqrt{7}$，$\|\beta\|=\sqrt{14}$，$\rho(\alpha,\beta)=\sqrt{5}$.

19. (1) $\eta_1=\left(\frac{1}{\sqrt{10}},\frac{2}{\sqrt{10}},\frac{2}{\sqrt{10}},-\frac{1}{\sqrt{10}}\right)^T$,

$\eta_2=\left(\frac{2}{\sqrt{26}},\frac{3}{\sqrt{26}},\frac{-3}{\sqrt{26}},\frac{2}{\sqrt{26}}\right)^T$,

$\eta_3=\left(\frac{2}{\sqrt{10}},-\frac{1}{\sqrt{10}},-\frac{1}{\sqrt{10}},-\frac{2}{\sqrt{10}}\right)^T$,

$\alpha=\sqrt{10}(\eta_1+\eta_3)$;

(2) $\eta_1=\left(\frac{2}{\sqrt{15}},\frac{1}{\sqrt{15}},\frac{3}{\sqrt{15}},\frac{-1}{\sqrt{15}}\right)^T$

$\eta_2=\left(\frac{3}{\sqrt{23}},\frac{2}{\sqrt{23}},\frac{-3}{\sqrt{23}},\frac{-1}{\sqrt{23}}\right)^T$,

$\eta_3=\left(\frac{1}{\sqrt{127}},\frac{5}{\sqrt{127}},\frac{1}{\sqrt{127}},\frac{10}{\sqrt{127}}\right)^T$,

$\alpha=\sqrt{15}\eta_1$;

(3) $\eta_1=1$,

$\eta_2=2\sqrt{3}x-\sqrt{3}$,

$\eta_3=6\sqrt{5}x^2-6\sqrt{5}x+\sqrt{5}$,

$f(x)=\frac{3}{2}\eta_1+\frac{1}{2\sqrt{3}}\eta_2$.

20. $V^{\perp}$的基为$\beta_1=(-2,2,1,0)^T$,$\beta_2=(-1,-1,0,1)^T$,$V^{\perp}=\mathrm{Span}\{\beta_1,\beta_2\}$.

21. (1) $a=0$,是;$a\neq0$,非;(2)非;(3)非;(4)是;(5)是;(6)是;(7)是.

22. (1) $A=\begin{pmatrix}2 & -1 & 0\\0 & 1 & 1\\1 & 0 & 0\end{pmatrix}$;

(2) $D=\begin{pmatrix}\alpha & \beta & 1 & 0 & 0 & 0\\-\beta & \alpha & 0 & 1 & 0 & 0\\0 & 0 & \alpha & \beta & 1 & 0\\0 & 0 & -\beta & \alpha & 0 & 1\\0 & 0 & 0 & 0 & \alpha & \beta\\0 & 0 & 0 & 0 & -\beta & \alpha\end{pmatrix}$;

(3) $\beta=\begin{pmatrix}-1 & 1 & -2\\2 & 2 & 0\\3 & 0 & 2\end{pmatrix}$;

(4) $\begin{pmatrix}-\frac{5}{7} & \frac{20}{7} & -\frac{20}{7}\\-\frac{4}{7} & -\frac{5}{7} & -\frac{2}{7}\\\frac{27}{7} & \frac{18}{7} & \frac{24}{7}\end{pmatrix}$.

23. 左乘时为$\begin{pmatrix}a & b & 0 & 0\\c & d & 0 & 0\\0 & 0 & a & b\\0 & 0 & c & d\end{pmatrix}$;右乘时为$\begin{pmatrix}a & 0 & c & 0\\0 & a & 0 & c\\b & 0 & d & 0\\0 & b & 0 & d\end{pmatrix}$.

24. (1) $\begin{pmatrix}2 & -3 & 3 & 2\\\frac{2}{3} & -\frac{4}{3} & \frac{10}{3} & \frac{10}{3}\\\frac{10}{3} & -\frac{16}{3} & \frac{40}{3} & \frac{40}{3}\\0 & 1 & -7 & -8\end{pmatrix}$;

(2) $T^{-1}(0)=\mathrm{Span}\{\alpha_1,\alpha_2\}$,其中 $\alpha_1=-2\varepsilon_1-\frac{3}{2}\varepsilon_2+\varepsilon_3$,$\alpha_2=-\varepsilon_1-2\varepsilon_2+\varepsilon_4$;$TV=\mathrm{Span}\{T\varepsilon_1,T\varepsilon_2\}$,其中,$T\varepsilon_1=\varepsilon_1-\varepsilon_2+\varepsilon_3+2\varepsilon_4$,$T\varepsilon_2=2\varepsilon_2+2\varepsilon_3-2\varepsilon_4$;

(3) 将 $T^{-1}(0)$ 的基扩充为 V 的一组基 $\varepsilon_1,\varepsilon_2,\alpha_1,\alpha_2$. 在此基下, T 的矩阵为

$$A=\begin{pmatrix}5 & 2 & 0 & 0\\ \frac{9}{2} & 1 & 0 & 0\\ 1 & 2 & 0 & 0\\ 2 & -2 & 0 & 0\end{pmatrix};$$

(4) TV 的基扩充为 V 的基 $T\varepsilon_1,T\varepsilon_2,\varepsilon_3,\varepsilon_4$. T 在这组基下的矩阵为

$$A=\begin{pmatrix}5 & 2 & 2 & 1\\ \frac{9}{2} & 1 & \frac{3}{2} & 2\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}.$$

25. 提示:设 T 在给定基下的矩阵为 $A=(a_{ij})$,并设 Z 为从旧基到新基的过渡矩阵,由于 T 在任一组基下的矩阵相同,则有 $A=Z^{-1}AZ$,即 $AZ=ZA$,由 A 与一切满秩矩阵可交换即可定出 A 为数量矩阵.

26. 提示:(1)必要性. 若 $AV=BV$,对$\forall\ \alpha\in V$,则 $B\alpha\in BV=AV$,故存在 $\beta\in V$,使 $B\alpha=A\beta$. $AB\alpha=A^2\beta=A\beta=B\alpha$,由 α 的任意性有 $AB=B$. 同理可证 $BA=A$.

充分性. 若 $AB=B$,$BA=A$,对$\forall\ A\alpha\in AV\subset V$. $A\alpha=BA\alpha=B(A\alpha)\in BV$,故 $AV\subset BV$;同理可证 $BV\subset AV$.

(2)必要性. 若 $A^{-1}(0)=B^{-1}(0)$,对$\forall\ \beta\in V$,作为 $\beta-A\beta$,因为 $A(\beta-A\beta)=A\beta-A^2\beta=A\beta-A\beta=0$,所以 $\beta-A\beta\in A^{-1}(0)=B^{-1}(0)$,则 $B(\beta-A\beta)=0$,故 $B\beta=BA\beta$,由 β 的任意性有 $B=BA$;同理,通过作 $\beta-B\beta$ 可得 $A=AB$.

充分性. 若 $A=AB$,$B=BA$,对$\forall\ \alpha\in A^{-1}(0)$,由 $B\alpha=BA\alpha=B(A\alpha)=B(0)=0$,故 $A^{-1}(0)\subset B^{-1}(0)$;同理,由$\forall\ \beta\in B^{-1}(0)$可

得 $B^{-1}(0)\subset A^{-1}(0)$.

27. (1) $\lambda_1=\lambda_2=1,\alpha_1=(1,0,1)^T,\alpha_2=(0,1,0)^T,\lambda_3=-1,\alpha_3=(1,0,-1)^T$;

(2) $\lambda_1=0,\alpha_1=(3,-1,2)^T,\lambda_2=\sqrt{14}\mathrm{i},\alpha_2=(6+\sqrt{14}\mathrm{i},-2+3\sqrt{14}\mathrm{i},-10)^T,\lambda_3=-\sqrt{14}\mathrm{i},\alpha_3=(6-\sqrt{14}\mathrm{i},-2-3\sqrt{14}\mathrm{i},-10)^T$;

(3) $\lambda_1=\lambda_2=1,\alpha_1=(3,-6,20)^T,\lambda_3=-2,\alpha_2=(0,0,1)^T$;

(4) $\lambda_1=\lambda_2=\lambda_3=2,\alpha_1=(1,1,0,0)^T,\alpha_2=(1,0,1,0)^T,\alpha_3=(1,0,0,1)^T,\lambda_4=-2,\alpha_4=(1,-1,-1,-1)^T$.

28. (1),(2),(4)为单纯矩阵,其变换矩阵分别为:

(1) $\begin{pmatrix}1&0&1\\0&1&0\\1&0&-1\end{pmatrix}$;

(2) $\begin{pmatrix}3&6+\sqrt{14}\mathrm{i}&6-\sqrt{14}\mathrm{i}\\-1&-2+3\sqrt{14}\mathrm{i}&-2-3\sqrt{14}\mathrm{i}\\2&-10&-10\end{pmatrix}$;

(4) $\begin{pmatrix}1&1&1&1\\1&0&0&-1\\0&1&0&-1\\0&0&1&-1\end{pmatrix}$.

29. 提示:设 T 在给定基下的矩阵为 A,则由 $\prod_{i=1}^{n}\lambda_i=\det A$ 可得.

30. 提示:用反证法.

31. 提示:A 的特征值为 λ_i,则 A^2 的特征值为 λ_i^2,由 $\lambda_i^2=1$ 有 $\lambda_i=\pm1$.若所有 $\lambda_i=1$,则 $A+I$ 为满秩矩阵,故由$(A+I)(A-I)=A^2-I^2=0$ 有 $A=I$.

32. 设 A 为 n 阶方阵,其秩为 r,且 $A^2=A$,由 $A^2=A$ 知 A 的列向量都是 A 的对应于特征值为 1 的特征向量,由于 $R(A)=r$,

故特征值 1 的几何重数为 r，其代数重数至少为 r. 又 $AX=0$ 的基础解系中的向量个数为 $n-r$，即 A 的特征值 0 的几何重数为 $n-r$，其代数重数不小于 $n-r$. 由于一个 n 阶矩阵的特征值的代数重数之和恰为 n，故特征值 1 和 0 的代数重数分别为 r 和 $n-r$. 可见 A 除了 1 和 0 外无其它特征值，而 1 和 0 的几何重数之和为 n，故 A 为单纯矩阵，所以 $A\sim\mathrm{diag}(I_r,0)$.

33. 提示：由 $Az=\lambda z$ 有 $A(kz)=k\lambda z$，$A^kz=\lambda^kz$，从而有 $f(A)z=f(\lambda)z$.

34. 提示：设 $R(A)=r$，则存在满秩矩阵 P 与 Q，使得 $PAQ=\mathrm{diag}(I_r,0)$，故 $PABP^{-1}=PAQQ^{-1}BP^{-1}=\mathrm{diag}(I_r,0)C$，其中 $C=Q^{-1}BP^{-1}=(c_{ij})$，且 $Q^{-1}BAQ=Q^{-1}BP^{-1}\cdot PAQ=C\mathrm{diag}(I_r,0)$，可验证 $\det(\lambda I-\mathrm{diag}(I_r,0)C)=\det(\lambda I-C\mathrm{diag}(I_r,0))$.

35. 提示：设 A 的任一特征值为 λ，A 的对应于 λ 的特征子空间记为 V_λ. 对 V_λ 中任意向量 Z，$ABZ=BAZ=B\lambda Z=\lambda BZ$，故 $BZ\in V_\lambda$. 因此 V_λ 为线性变换 $T(Z)=BZ$ 的不变子空间，即 $T(Z)=BZ$ 为 V_λ 中的线性变换. 此线性变换的特征向量即为 B 的特征向量，但它又属于 V_λ，由 V_λ 的定义知它又是 A 的特征向量，即 A,B 有公共的特征向量.

习 题 二

1. (1) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda(\lambda+1) \end{pmatrix}$；

(2) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda(\lambda-1) & 0 & 0 \\ 0 & 0 & \lambda(\lambda-1) & 0 \\ 0 & 0 & 0 & \lambda^2(\lambda-1)^2 \end{pmatrix}$；

(3) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda-1 & 0 \\ 0 & 0 & \lambda^3-\lambda^2-\lambda+1 \end{bmatrix}$;

(4) $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \lambda(\lambda-1) & 0 \\ 0 & 0 & 0 & 0 & \lambda^2(\lambda-1) \end{bmatrix}$.

2. (1) $d_1=d_2=1, d_3=(\lambda-2)^3$;

(2) 当 $\beta\neq 0$ 时, $d_1=d_2=d_3=1, d_4=[(\lambda+2)^2+\beta^2)]^2$; 当 $\beta=0$ 时, $d_1=d_2=1, d_3=(\lambda+\alpha)^2, d_4=(\lambda+\alpha)^2$;

(3) $d_1=d_2=d_3=1, d_4=\lambda^4+2\lambda^3+3\lambda^2+4\lambda+5$;

(4) $d_1=d_2=d_3=1, d_4=(\lambda+2)^4$.

3. (1) $\lambda+1, (\lambda+1)^2$;

(2) $\lambda+1, \lambda-1, \lambda-1$.

4. (1) $\begin{bmatrix} -3 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$;

(2) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$;

(3) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & i \end{bmatrix}$;

(4) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$;

(5) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$;

(6) $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

5. $P=\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -2 \\ 0 & 2 & 1 \end{pmatrix}$, $P^{-1}=\frac{1}{5}\begin{pmatrix} 5 & 0 & -5 \\ 0 & 1 & 2 \\ 0 & -2 & 1 \end{pmatrix}$,

$\Lambda=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$, $A=P\Lambda P^{-1}$,

$$A^5=P\Lambda^5(P^{-1})^5=\begin{pmatrix} 1 & 4\times 5^4 & 3\times 5^4-1 \\ 0 & -3\times 5^4 & 4\times 5^4 \\ 0 & 4\times 5^4 & 3\times 5^4 \end{pmatrix}.$$

6. $P^{-1}AP=J=\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, $P=\begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & -1 & -1 \end{pmatrix}$,

(注:P 并不唯一.)

7. $\begin{pmatrix} -3 & 48 & -26 \\ 0 & 95 & -61 \\ 0 & -61 & 34 \end{pmatrix}$.

8. 提示:利用 A 的特征方程及 Cayley-Hamilton 定理.

9. $\begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{pmatrix}$.

10. 提示:$f(\lambda)=(\lambda+1)(\lambda-1)(\lambda-2)$,令 $\lambda^{2n}=f(\lambda)q(\lambda)+a\lambda^2+b\lambda+c$,应有(分别取 $\lambda=1,-1,2$):

$$\begin{cases} a+b+c=1, \\ a-b+c=1, \\ 4a+2b+c=2^{2n}, \end{cases}$$

解出 a、b、c，再利用 Cayley-Hamilton 定理得

$$A^{2n} = \frac{1}{3}[(2^{2n}-1)A^2 - (2^{2n}-4)I].$$

11. (1) $m(\lambda)=(\lambda-2)^2$；

(2) $m(\lambda)=(\lambda-2)(\lambda^2-5\lambda+11)$；

(3) $m(\lambda)=\lambda-1$；

(4) $m(\lambda)=\lambda^2-n\lambda$；

(5) $m(\lambda)=(\lambda-a_0)^2+a_1^2+a_2^2+a_3^2$.

习　题　三

3. $\lim\limits_{t\to 0}A(t)=\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$；

$$\frac{\mathrm{d}}{\mathrm{d}t}A(t)=\begin{pmatrix} \cos t & -\sin t & 1 \\ \dfrac{t\cos t-\sin t}{t^2} & \mathrm{e}^t & 2t \\ 0 & 0 & 3t^2 \end{pmatrix};$$

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}A(t)=\begin{pmatrix} -\sin t & -\cos t & 0 \\ \dfrac{(2-t^2)\sin t-2t\cos t}{t^3} & \mathrm{e}^t & 2 \\ 0 & 0 & 6t \end{pmatrix};$$

$$\frac{\mathrm{d}}{\mathrm{d}t}|A(t)|=\mathrm{e}^t[3t^2\sin t+t^3(\sin t+\cos t)-t-1]$$
$$+t(2\cos t-t\sin t)-t(\sin 2t+\cos 2t);$$

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right|=3t^2\mathrm{e}^t\cos t+3\sin t(t\cos t-\sin t).$$

4. $\displaystyle\int_0^1 A(x)\mathrm{d}x = \begin{pmatrix} \frac{1}{2}(\mathrm{e}^2-1) & 1 & \frac{1}{3} \\ 1-\mathrm{e}^{-1} & \mathrm{e}^2-1 & 0 \\ \frac{3}{2} & 0 & 0 \end{pmatrix}$；

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_0^{x^2} A(t)\mathrm{d}t\right) = 2x\begin{bmatrix} \mathrm{e}^{2x^2} & x^2\mathrm{e}^{x^2} & x^4 \\ \mathrm{e}^{-x^2} & 2\mathrm{e}^{2x^2} & 0 \\ 3x^2 & 0 & 0 \end{bmatrix}.$$

13. 提示：由于 $A^{-1}A=I$，则 $\|A^{-1}A\|=1$，故

$$1 = \|A^{-1}A\| \leqslant \|A^{-1}\|\ \|A\|.$$

14. 提示：$\forall\ x\in C^n$，则

$$\begin{aligned} \|(I-A)x\| &\geqslant \|x\| - \|Ax\| \\ &\geqslant (1-\|A\|)\|x\| \\ &\geqslant 0, \end{aligned}$$

且仅当 $x=0$ 时才有 $(I-A)x=0$，从而推出 $I-A$ 可逆．由

$$(I-A)(I-A)^{-1} = I,$$

可得

$$(I-A)^{-1} = I + A(I-A)^{-1},$$

由此即可证明(1)和(2)．

16. (1) $\cos A=I+(\cos 1-1)A$，　$\sin A=(\sin 1)A$，

$\mathrm{e}^{A^2}=I+(\mathrm{e}-1)A$；

(2) $\cos A=(\cos 1)I$，　$\sin A=(\sin 1)I$，　$\mathrm{e}^{A^2}=\mathrm{e}I$；

(3) $\cos A=I$，　$\sin A=A$，　$\mathrm{e}^{A^2}=I$．

18. (1) $\mathrm{e}^{At}=\begin{bmatrix} \mathrm{e}^{2t} & 0 & 0 \\ 0 & \mathrm{e}^t & 0 \\ 0 & t\mathrm{e}^t & \mathrm{e}^t \end{bmatrix}$，

$$\sin At=\begin{bmatrix} \sin 2t & 0 & 0 \\ 0 & \sin t & 0 \\ 0 & t\cos t & \sin t \end{bmatrix};$$

(2) $\mathrm{e}^{At}=\begin{bmatrix} 1-\frac{t^2}{2} & -t & \frac{-t^2}{2} \\ t & 1 & t \\ \frac{t^2}{2} & t & 1+\frac{t^2}{2} \end{bmatrix}$，

$$\sin At=\begin{bmatrix}0 & -t & 0\\ t & 0 & t\\ 0 & t & 0\end{bmatrix};$$

$$(3)\ e^{At}=\begin{bmatrix}3e^{-t}-3e^{-2t}+e^{-3t} & \frac{5}{2}e^{-t}-4e^{-2t}+\frac{3}{2}e^{-3t} & \frac{1}{2}e^{-t}-e^{-2t}+\frac{1}{2}e^{-3t}\\ -3e^{-t}+6e^{-2t}-3e^{-3t} & \frac{5}{2}e^{-t}+8e^{-2t}-\frac{9}{2}e^{-3t} & -\frac{1}{2}e^{-t}+2e^{-2t}-\frac{3}{2}e^{-3t}\\ 3e^{-t}-12e^{-2t}+9e^{-3t} & \frac{5}{2}e^{-t}-16e^{-2t}+\frac{27}{2}e^{-3t} & \frac{1}{2}e^{-t}-4e^{-2t}+\frac{9}{2}e^{-3t}\end{bmatrix},$$

$$\sin At=\frac{1}{2}\begin{bmatrix}-5\sin t+6\sin 2t-2\sin 3t & -5\sin t+8\sin 2t-3\sin 3t & -\sin t+2\sin 2t-\sin 3t\\ 6\sin t-12\sin 2t+6\sin 3t & 5\sin t-16\sin 2t+9\sin 3t & \sin t-4\sin 2t+3\sin 3t\\ -5\sin t+24\sin 2t-18\sin 3t & -5\sin t+32\sin 2t-27\sin 3t & -\sin t+8\sin 2t-9\sin 3t\end{bmatrix}.$$

19. (1) $A^{1000}=\dfrac{5I-A}{4}+\dfrac{A-I}{4}5^{1000}$;

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

(3) $\arcsin\dfrac{A}{4}=\begin{bmatrix}\dfrac{\pi}{6}-\dfrac{\sqrt{3}}{6} & -\dfrac{\sqrt{3}}{6}\\ \dfrac{2\sqrt{3}}{3} & \dfrac{\pi}{6}+\dfrac{\sqrt{3}}{3}\end{bmatrix}$;

(4) $(I+A)^{-1}=\dfrac{1}{21}\begin{bmatrix}5 & -8\\ -8 & 17\end{bmatrix}$, $A^{\frac{1}{2}}=\dfrac{\sqrt{5}}{10}A$.

23. (1) $\operatorname{sh}A=\sum\limits_{n=0}^{\infty}\dfrac{1}{(2n+1)!}A^{2n+1}, \forall A\in C^{n\times n}$;

(2) $\ln(I+A)=4\sum\limits_{n=1}^{\infty}(-1)^{n-1}\dfrac{1}{n}A^{n}$, $\|A\|<1$;

(3) $\operatorname{arctg}A=\sum\limits_{k=1}^{+\infty}(-1)^{k}\dfrac{1}{2k+1}A^{2k+1}$, $\|A\|<1$.

习 题 四

1. (1) $\begin{pmatrix}x_1\\x_2\end{pmatrix}=\frac{1}{3}\begin{pmatrix}2e^{-t}+e^{5t} & e^{5t}-e^{-t}\\2e^{5t}-2e^{-t} & 2e^{5t}+e^{-t}\end{pmatrix}\begin{pmatrix}c_1\\c_2\end{pmatrix}$;

(2) $\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}2-e^t+te^t & te^t & -2+2e^t+te^t\\-1+e^t & e^t & 1+e^t\\-1-e^t+te^t & te^t & -1+2e^t+te^t\end{pmatrix}\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix}$.

2. (1) $x=\begin{pmatrix}2e^{5t}+e^{-t}\\4e^{5t}-e^{-t}\end{pmatrix}$;

(2) $x=\begin{pmatrix}\frac{1}{2}e^{3t}+\frac{1}{2}e^{-t}\\\frac{1}{4}e^{3t}-\frac{1}{4}e^{-t}\\\frac{1}{2}e^{3t}-\frac{1}{2}e^{-t}\end{pmatrix}$.

3. (1) $x=\begin{pmatrix}\frac{1}{3}[(c_1+c_2+\frac{1}{2})e^{5t}+(2c_1-c_2-\frac{1}{2})e^{-t}]-2te^{-t}\\\frac{1}{3}[(2c_1+2c_2+1)e^{5t}-(2c_1-c_2+1)e^{-t}]+2te^{-t}\end{pmatrix}$;

(2) $x=\begin{pmatrix}e^{-t}-4e^{-2t}+\frac{11}{3}e^{-3t}+\frac{1}{3}\\5e^{-t}-16e^{-2t}+11e^{-3t}\\6e^{-t}-12e^{-2t}+\frac{22}{3}e^{-3t}-\frac{7}{3}\end{pmatrix}$.

4. $y(t)=-\frac{3}{4}e^{-t}+\frac{1}{2}te^{-t}+e^{-2t}+\frac{1}{4}e^{-3t}$.

5. 提示：

$$x=(P_1\quad P_2)\begin{pmatrix}e^{\lambda_1 t} & 0\\0 & e^{\lambda_2 t}\end{pmatrix}(P_1\quad P_2)^{-1}x(0),$$

令

$$(P_1 \quad P_2)^{-1}x(0) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

则

$$x(0) = c_1 P_1 + c_2 P_2,$$

于是

$$x = c_1 e^{\lambda_1 t} P_1 + c_2 e^{\lambda_2 t} P_2,$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2}e^{-5t} + \frac{1}{2}e^{-t} \\ \frac{5}{2}e^{-5t} - \frac{3}{2}e^{-t} \end{pmatrix}.$$

6. 提示：对恒等式 $\Phi(t,t_0)\Phi(t_0,t)=I$ 求导.

7. (1) 提示：

$$A(t_1)A(t_2) = A(t_2)A(t_1),$$

$$\Phi(t,0) = \begin{pmatrix} e^t & e^t - 1 \\ 0 & e^t \end{pmatrix},$$

$$x(t) = \Phi(t,0)\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2e^t - 1 \\ e^t \end{pmatrix};$$

(2) 提示：

$$A(t_1)A(t_2) = A(t_2)A(t_1),$$

$$\Phi(t,t_0) = \begin{pmatrix} 1 & \dfrac{t_0 - t}{tt_0} & \dfrac{-(t-t_0)^2}{(1+t)(1+t_0)tt_0} \\ 0 & 1 & \dfrac{t-t_0}{(1+t)(1+t_0)} \\ 0 & 0 & 1 \end{pmatrix},$$

$$x(t) = \Phi(t,t_0)x_0;$$

(3) 提示：

$$\Phi(t,0) = \begin{pmatrix} 1 + \frac{3}{2}t^2 + \frac{1}{8}t^4 + \cdots & t + \frac{1}{2}t^3 + \cdots \\ 2t + t^3 + \cdots & 1 + \frac{3}{2}t^3 + \frac{1}{8}t^4 + \cdots \end{pmatrix},$$

$$x(t)=\begin{bmatrix}1-t+\frac{3}{2}t^2-\frac{1}{2}t^3+\frac{1}{8}t^4+\cdots\\-1+2t-\frac{3}{2}t^2+t^3-\frac{1}{8}t^4+\cdots\end{bmatrix}.$$

8. (1) $x(t)=\begin{bmatrix}e^t-1\\0\end{bmatrix}$;

(2) $$x(t)=\begin{bmatrix}1+\dfrac{t-t_0}{tt_0}-\dfrac{(t-t_0)^2}{(1+t)(1+t_0)tt_0}+\dfrac{t^2-t_0^2}{2}+\dfrac{t-t_0}{t}-\ln\dfrac{t}{t_0}\\1+\dfrac{t-t_0}{(1+t)(1+t_0)}+(t-t_0)\end{bmatrix}.$$

习 题 五

3. (1) 必要性显然. 充分性. 由第 1 题 A^-A 为幂等矩阵，且

$$\operatorname{rank}(A^-A)\leqslant\operatorname{rank}(A)=\operatorname{rank}(AA^-A)\leqslant\operatorname{rank}(A^-A),$$

故 $\operatorname{rank}(A^-A)=n$，$(A^-A)^{-1}$ 存在. 在 $(A^-A)^2=A^-A$ 两边同乘 $(A^-A)^{-1}$ 即得.

(2) 类似(1)可得.

6. (1) $A_1^-=\begin{bmatrix}\frac{1}{5}&0&0&0\\0&1&0&0\\\frac{2}{5}&0&0&0\end{bmatrix}$,

$A_1X=B_1$ 的通解为

$$X=\begin{bmatrix}\frac{1}{5}+\frac{4}{5}c_1-\frac{2}{5}c_2\\\frac{2}{5}-\frac{2}{5}c_1+\frac{1}{5}c_2\end{bmatrix};$$

(2) $A_2^- = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & -\frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{2}{3} & 0 \\ 0 & \frac{1}{3} & 0 \end{pmatrix}$,

$A_2X=B_2$ 的通解为

$$X = \begin{pmatrix} \frac{2}{3} + \frac{1}{3}c_1 - \frac{1}{3}c_2 - \frac{1}{3}c_4 \\ \frac{1}{3} - \frac{1}{3}c_1 + \frac{2}{3}c_2 + \frac{1}{3}c_3 \\ \frac{1}{3}c_2 + \frac{1}{3}c_3 - \frac{1}{3}c_4 \\ \frac{1}{3} - \frac{1}{3}c_2 - \frac{1}{3}c_3 + \frac{2}{3}c_4 \end{pmatrix}.$$

7. (1) $A_m^- = \frac{1}{5}\begin{pmatrix} 1 & 0 \\ -10 & 5 \\ 2 & 0 \end{pmatrix}$,

$A_1X=b_1$ 的最小范数解为

$$X = \frac{1}{5}\begin{pmatrix} 1 \\ -15 \\ 2 \end{pmatrix};$$

(2) $A_m^- = \frac{1}{154}\begin{pmatrix} 8 & 19 & 9 \\ -10 & 34 & 8 \\ 44 & -11 & 11 \end{pmatrix}$,

$A_2X=b_2$ 的最小范数解为

$$X = \frac{1}{14}\begin{pmatrix} 3 \\ -2 \\ 13 \end{pmatrix}.$$

8. 设 $X_0=Gb$，则 $X=Gb$ 为 $AX=b$ 最小二乘解.

充分性. 若 $AX=AGb$，则

$$\| AX-b \| = \| AGb-b \| = \| AX_0-b \|,$$

故 X 为最小二乘解.

必要性. 若 X 为最小二乘解，即 $\| AX-b \| = \| AX_0-b \|$，则

$$\begin{aligned}\| AX-b \|^2 &= \| AX-AX_0+AX_0-b \|^2 \\ &= \| AX-AX_0 \|^2 + \| AX_0-b \|^2 \\ &\quad + (AX-AX_0)^*(AX_0-b) \\ &\quad + (AX_0-b)^*(AX-AX_0),\end{aligned}$$

由 $\| AX-b \| = \| AX_0-b \|$ 及

$$\begin{aligned}(AX-AX_0)^*(AX_0-b) &= (AX-AGb)^*(AGb-b) \\ &= (X^*-b^*G^*)A^*(AGb-b) \\ &= (X^*-b^*G^*)[(AGA)^*-A^*]b \\ &= 0,\end{aligned}$$

$$\begin{aligned}(AX_0-b)^*(AX-AX_0) &= [(AX-AX_0)^*(AX_0-b)]^* \\ &= 0,\end{aligned}$$

得 $\| AX-AX_0 \|^2=0$，即

$$AX = AX_0 = AGb.$$

9. (1) $A_l^- = \dfrac{1}{11}\begin{pmatrix} -4 & 7 & 1 \\ 7 & -4 & 1 \end{pmatrix}$，

不相容方程 $A_1X=b_1$ 的最小二乘解为

$$X=\frac{1}{11}\begin{pmatrix} -4 \\ 7 \end{pmatrix};$$

(2) $A_l^- = \dfrac{1}{25}\begin{pmatrix} 0 & 1 & 0 & 2 \\ -5 & 3 & -10 & 6 \\ 0 & 1 & 0 & 2 \\ 5 & -2 & 10 & -4 \end{pmatrix}$，

不相容方程 $A_2X=b_2$ 的最小二乘解为

$$X=\frac{1}{25}\begin{pmatrix}3\\9\\3\\-6\end{pmatrix}.$$

10. (1) $A^{+}=\dfrac{1}{22}\begin{pmatrix}-1 & -2 & 6 & 5\\ -1 & -2 & 6 & 5\\ 4 & 8 & -2 & 2\end{pmatrix}$, $X=\begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\\ 1\end{pmatrix}$;

(2) $A^{+}=\dfrac{1}{42}\begin{pmatrix}6 & -10 & 11 & 11\\ 0 & 14 & -7 & -7\\ 6 & 4 & 4 & 4\end{pmatrix}$, $X=\dfrac{1}{21}\begin{pmatrix}-5\\ 7\\ -2\end{pmatrix}$.

11. 提示:充分性. 若 $AA^{-}\eta=\eta$ 及 $\mathrm{rank}\ A=\mathrm{rank}(A\quad \eta)$,则后一条件表明 $A\xi=\eta$ 相容,显然 $\xi=A^{-}\eta$ 为方程组的解.

必要性. 设方程组有解 ξ_0,即 $A\xi_0=\eta$,显然 $\mathrm{rank}\ A=\mathrm{rank}(A\quad \eta)$,且 $AA^{+}\eta=AA^{-}A\xi_0=A\xi_0=\eta$.

12. 提示:必要性.

$$\begin{aligned}AXB = C \Longrightarrow AA^{-}CB^{-}B &= AA^{-}(AXB)B^{-}B\\ &= AXB = C.\end{aligned}$$

充分性. 设对某个 A^{-} 和 B^{-} 有 $AA^{-}CB^{-}B=C$,则 $A^{-}CB^{-}$ 显然是解,且对任何 $Z\in C^{n\times p}$,$X=A^{-}CB^{-}+(Z-A^{-}AZBB^{-})$ 都是解. 设 Y 为任一解,令 $Y=Z$ 就有 $Y=A^{-}CB^{-}+Y-A^{-}(AYB)B^{-}$,即任一解 Y 也满足 $A^{-}CB^{-}+(Z-A^{-}AZBB^{-})$ 的形式.

13. 提示:验证 $A^{+}=V\begin{pmatrix}\Delta^{-1} & 0\\ 0 & 0\end{pmatrix}U^{-}$ 满足 Moore-Penrose 广义逆的四个条件即可.

参 考 文 献

[1] 北京大学几何与代数教研室编. 高等代数. 人民教育出版社, 19[illegible]

[2] 南京大学计算数学专业编. 线性代数. 科学出版社, 1978

[3] 南京大学计算数学专业编. 常微分方程数值解法. 科学出版社, 1979

[4] 程云鹏等编著. 矩阵论. 西北工业大学出版社, 1989

[5] 王朝瑞等编著. 矩阵分析. 北京理工大学出版社, 1990

[6] 吴文达. 广义逆矩阵. 计算机应用与应用数学, **1**, 1974

[7] P. Lancaster, M. Tismenetsky. The Theory of Matrices, 2nd Edition. Academic Press, 1985

[8] S. Fenyö. Modern Mathematical Methods in Techology. Morth-Holland and Publ, 1975

[9] C. R. Rao, S. K. Mitra. Generalized Inverse of Matrices and its Applications. John wiley and Sons, 1971