第1章

线性空间与线性变换

主要内容

1.线性空间的概念
2 基、坐标与维数
3 线性子空间
4 线性变换
5 线性变换的矩阵
6 线性空间的同构

1.1　线性空间的基本概念

1.1.1数域

设P是一个包含数1与0的数集，如果P对于数的加、减、乘、除（除数不为零）四则运算都封闭，则称P是一个数域。

显然：复数集C，实数集R，有理数集Q都是数域。

例1.1 数集 $F=\{a+b\sqrt{2}\,|\,a,b\in Q\}$ 是一个数域。

注：最小的数域是有理数域?

1.1.2线性空间的定义与性质

设V是一非空集合，P是一个数域，在V中定义加法：$v=\alpha+\beta$; 在V与P之间定义数量乘法 $\delta=k\alpha$. 如果V对加法与数量乘法运算封闭，且加法与数量乘法运算满足

1) $\alpha+\beta=\beta+\alpha$
2) $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$
3) $\exists\ 0\in V, \forall\alpha\in V,$ 有$\alpha+0=\alpha$
4) $\forall\alpha\in V, \exists\ \beta\in V, s.t\ \ \alpha+\beta=0$
5) $1\alpha=\alpha$
6) $k(l\alpha)=(kl)\alpha$
7) $(k+l)\alpha=k\alpha+l\alpha$
8) $k(\alpha+\beta)=k\alpha+l\beta$

则称V是数域P上的线性空间或向量空间.

例1.2 判断下列集合是否构成线性空间

1）空间中不平行于一已知向量ξ的全体向量所的集合，是否构成线性空间？

不

2）数域P上次数等于定数$n(n\geq1)$的多项式全体所构成的集合，是否构成复数域上的线性空间？

不

例1.3

1. n维向量空间R^n按照向量的加法以及向量与实数的数乘都构成实线性空间.
2. 全体 $m\times n$实矩阵，在矩阵的加法及数乘两种运算下构成一个实线性空间，记为$R^{m\times n}$.
3. 区间$[a,b]$上的全体连续实函数，按照函数的加法及数与函数的乘法构成一个实线性空间，记为$C[a,b]$.
4. 全体次数小于 n的多项式连同零多项式，按照多项式的加法与数乘构成一个实线性空间，记为 $P_n[x]$.
5. 齐次线性方程组 $AX=0$的全体解向量，在向量的加法及数乘运算下构成一个线性空间，即通常所说的解空间.

注：非齐次线性方程组$AX=b$的全体解向量，在上述两种运算下不构成一个线性空间.

例1.3

1. n维向量空间R^n按照向量的加法以及向量与实数的数乘都构成实线性空间.
2. 全体 $m\times n$实矩阵，在矩阵的加法及数乘两种运算下构成一个实线性空间，记为$R^{m\times n}$.
3. 区间$[a,b]$上的全体连续实函数，按照函数的加法及数与函数的乘法构成一个实线性空间，记为$C[a,b]$.

例1.3

4. 全体次数小于 n的多项式连同零多项式，按照多项式的加法与数乘构成一个实线性空间，记为 $P_n[x]$.
5. 齐次线性方程组 $AX=0$的全体解向量，在向量的加法及数乘运算下构成一个线性空间，即通常所说的解空间.

注：非齐次线性方程组$AX=b$的全体解向量，在上述两种运算下不构成一个线性空间.

6. 仅含有单独一个零向量的集合V也构成一个向量空间，称为零空间，记为$\{0\}$.

7. 设R_+表示所有正实数的集合，在下述的加法与数乘之下，R_+构成实数域上的向量空间。

$$x\oplus y=xy,\quad k\circ x=x^k\quad (x,y\in R_+,k\in R)$$

注：以上例子表明，线性空间是广泛的，因而向量也是广泛的概念，不仅限于几何向量。可以是几何向量，也可以是数、函数、矩阵、多项式，还可以是变换等。

线性空间的性质

性质1 线性空间V中只有一个零向量. (零向量的唯一性)

性质2 V中每个向量只有一个负向量. (负向量的唯一性)

性质3 $0\alpha=0$, $k0=0$, $(-1)\alpha=-\alpha$.

性质4 若$k\alpha=0$, 则$k=0$或者$\alpha=0$.

1.2 基、坐标与维数

1.2.1 向量组的线性相关性

1．有关概念

定义1.3 设V是数域P上的线性空间，对V中的向量(元素) $\beta,\alpha_1,\alpha_2,\cdots,\alpha_m$ 如果存在一组实数 $k_1,k_2,\cdots,k_m\in P$ 使得

$$\beta=k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m$$

则称β是向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 的一个线性组合，也称β可以由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性表示，$k_1,k_2,\ldots,k_m$ 称为线性表示的系数。

定义1.4

设V是数域P上的线性空间，向量组$\alpha_1,\alpha_2,\cdots,\alpha_m\in V$，如果存在一组不全为零的数 $k_1,k_2,\cdots,k_m\in P$ 使得

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m=0$$

则称 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性相关，否则称它们线性无关.

常用的结论

1. $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性无关当且仅当由

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m=0\Rightarrow k_1=k_2=\cdots k_m=0$$

2. 单个向量α线性相关当且仅当 $\alpha=0$.
3. R^n中两个向量线性相关当且仅当它们成比例.
4. 向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性相关当且仅当其中至少有一个向量是其余向量的线性组合.

5. 向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性无关，而$\beta,\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性相关，则β可以由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 唯一线性表示.
6. 线性无关组不含零向量，含零向量的向量组必定线性相关.
7. 如果向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$线性无关，并且可由向量组 $\beta_1,\beta_2,\cdots,\beta_t$ 线性表示，则 $s\le t$.
8. 等价的线性无关向量组必定含有相同个数的向量.

1.2.2 基、坐标与维数

定义1.5 设V是数域P上的线性空间，如果存在n个向量 $\alpha_1,\alpha_2,\cdots,\alpha_n\in V$，满足

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

(2) V中任一向量都可由$\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表示;

则称 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是线性空间V的一组基，n称为V的维数，记为$\dim V$，并称V是n维线性空间.

由定义不难证明：

（1）零空间{0}是零维线性空间，没有基;

（2）n维线性空间V恰有n个线性无关的向量构成的基;

（3）n维线性空间V任意n个线性无关的向量都构成V的一组基。

例1.4 求下列各线性空间的一组基

1) 数域P上全体n阶方阵构成的空间$P^{n\times n}$,

2) $P^{n\times n}$中全体对称矩阵构成数域P上的空间.

解：1) $P^{n\times n}$基为 $E_{ij}\quad i,j=1,2,\cdots,n$

$\dim(P^{n\times n})=n^2$

其中E_{ij}表示第i行j列元素是1，其余元素是0的n阶矩阵

2) 令 $F_{ij}=\begin{cases}E_{ij}+E_{ji}\\ E_{ii}\end{cases}\quad 1\le i\le j\le n$

维数为$\dfrac{n(n+1)}{2}$.

向量的坐标

设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\in V$ 是 n 维线性空间 V 的一组基，对 V 中任意向量 α 均可由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 唯一线性表示为 $\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n$.

记 $X=(x_1,x_2,\ldots,x_n)^T$, 称为向量 α 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标.

也可以记为 $\alpha=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X$.

例1.5 在实系数多项式所构成的实线性空间 $P_3[x]$ 中，

令 $f_1=1,f_2=x,f_3=x^2$

则 f_1,f_2,f_3 线性无关，且 $P_3[x]$ 中任一多项式都可以由 f_1,f_2,f_3 线性表示， $1,x,x^2$ 是 $P_3[x]$ 的一组基，且 $f=a+bx+cx^2$ 在该基下的坐标是 $(a,b,c)^T$.

例1.6 在全体二阶实方阵在矩阵的加法及数乘两种运算下构成的实线性空间 $R^{2\times2}$ 中，令

$$E_{11}=\begin{pmatrix}1&0\\0&0\end{pmatrix}\quad E_{12}=\begin{pmatrix}0&1\\0&0\end{pmatrix}\quad E_{21}=\begin{pmatrix}0&0\\1&0\end{pmatrix}\quad E_{22}=\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

设 $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$ 是 数域 P 上 n 维线性空间 V 的一组基, 设向量 α 在基下的坐标为 $X=(x_1,x_2,\ldots,x_n)^T$, 则有：

$$\alpha=(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)X,$$

于是 V 与数域 P 上 n 维向量空间建立了一一对应关系,

$$\alpha\xleftrightarrow{(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)}(x_1,x_2,\ldots,x_n)^T$$

例1.7. 求 $R^{2\times2}$ 中，向量组

$$A_1=\begin{pmatrix}1&-1\\0&1\end{pmatrix},A_2=\begin{pmatrix}2&-2\\0&2\end{pmatrix},A_3=\begin{pmatrix}1&1\\1&0\end{pmatrix},A_4=\begin{pmatrix}2&0\\1&1\end{pmatrix}$$

的秩及一个极大无关组。

1.2.3 基变换与坐标变换

*n*维线性空间*V*中任意*n*个线性无关的向量都可以作为*V*的一组基. 显然，同一个向量在两组基下的坐标是不同的，下面主要研究同一个向量在不同基下的坐标之间的联系 。

定义1.7 设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$和$\eta_1,\eta_2,\cdots,\eta_n$是*n*维线性空间*V*的两组基，显然它们可以互相线性表示，

$$\begin{cases}\eta_1=c_{11}\varepsilon_1+c_{21}\varepsilon_2+\cdots+c_{n1}\varepsilon_n\,,\\ \eta_2=c_{12}\varepsilon_1+c_{22}\varepsilon_2+\cdots+c_{n2}\varepsilon_n\,,\\ \quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\\ \eta_n=c_{1n}\varepsilon_1+c_{2n}\varepsilon_2+\cdots+c_{nn}\varepsilon_n\,,\end{cases}$$

若将上式用矩阵形式表示，则

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}c_{11}&c_{12}&\cdots&c_{1n}\\ c_{21}&c_{22}&\cdots&c_{2n}\\ \vdots&\vdots&\ddots&\vdots\\ c_{n1}&c_{n2}&\cdots&c_{nn}\end{pmatrix}$$

记

$$C=\begin{pmatrix}c_{11}&c_{12}&\cdots&c_{1n}\\ c_{21}&c_{22}&\cdots&c_{2n}\\ \vdots&\vdots&\ddots&\vdots\\ c_{n1}&c_{n2}&\cdots&c_{nn}\end{pmatrix}$$

矩阵*C*称为由基$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)$到基$(\eta_1,\eta_2,\cdots,\eta_n)$的过渡矩阵

定理1.1 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 和 $\eta_1,\eta_2,\cdots,\eta_n$ 是*n*维线性空间*V*的两组基，由基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 到基 $\eta_1,\eta_2,\cdots,\eta_n$ 的过渡矩阵为*C*, 则*C*是可逆的；且如果向量在这两组基下的坐标分别是*X*与*Y*，则

$$X=CY \quad 或 \quad Y=C^{-1}X$$

上式就是维线性空间*V*中向量在两组基下的新旧坐标之间的坐标变换公式 .

例1.8 在$P_4[x]$中取两组基

$$\begin{cases}\alpha_1=x^3+2x^2-x\\ \alpha_2=x^3-x^2+x+1\\ \alpha_3=-x^3+2x^2+x+1\\ \alpha_4=-x^3-x^2+1\end{cases}\qquad \begin{cases}\beta_1=2x^3+x^2+1\\ \beta_2=x^2+2x+2\\ \beta_3=-2x^3+x^2+x+2\\ \beta_4=x^3+3x^2+x+2\end{cases}$$

求由基 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 到基 $\beta_1,\beta_2,\beta_3,\beta_4$ 的过渡矩阵.

$$(\alpha_1,\alpha_2,\alpha_3,\alpha_4)=(x^3,x^2,x,1)\begin{pmatrix}1&1&-1&-1\\ 2&-1&2&-1\\ -1&1&1&0\\ 0&1&1&1\end{pmatrix}$$

$$(\beta_1,\beta_2,\beta_3,\beta_4)=(x^3,x^2,x,1)\begin{pmatrix}2&0&-2&1\\ 1&1&1&3\\ 0&2&1&1\\ 1&2&2&2\end{pmatrix}$$

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

则有

$$(\beta_1,\beta_2,\beta_3,\beta_4)=(\alpha_1,\alpha_2,\alpha_3,\alpha_4)A^{-1}B$$

例1.9 设R^3中两组基

I $\varepsilon_1=\begin{pmatrix}1\\0\\1\end{pmatrix},\varepsilon_2=\begin{pmatrix}1\\1\\-1\end{pmatrix},\varepsilon_3=\begin{pmatrix}0\\1\\0\end{pmatrix}$

II $\eta_1=\begin{pmatrix}1\\-2\\1\end{pmatrix},\eta_2=\begin{pmatrix}1\\2\\-1\end{pmatrix},\eta_3=\begin{pmatrix}0\\1\\-2\end{pmatrix}$

求

(1) 基I到基II的过渡矩阵；

(2) 向量$\alpha=3\eta_1+2\eta_3$在基I下的坐标以及在自然基e_1,e_2,e_3下的坐标；

(3) 向量$\beta=(4,1,-2)^T$ 在基 I 下的坐标.

解 （1）设基 **I** 到基 **II** 的过渡矩阵为C，记矩阵 $A=(\varepsilon_1,\varepsilon_2,\varepsilon_3)$，$B=(\eta_1,\eta_2,\eta_3)$，则由基变换公式知有 $B=AC$ ，于是

$$C=A^{-1}B=\begin{pmatrix}1&1&0\\0&1&1\\1&-1&0\end{pmatrix}^{-1}\begin{pmatrix}1&1&0\\-2&2&1\\1&-1&-2\end{pmatrix}=\frac{1}{2}\begin{pmatrix}1&0&1\\1&0&-1\\-1&2&1\end{pmatrix}\begin{pmatrix}1&1&0\\-2&2&1\\1&-1&-2\end{pmatrix}=\begin{pmatrix}1&0&-1\\0&1&1\\-2&1&0\end{pmatrix}$$

（2） $\alpha=3\eta_1+2\eta_3$ 在基 **I** 下的坐标为

$$X=CX'=\begin{pmatrix}1&0&-1\\0&1&1\\-2&1&0\end{pmatrix}\begin{pmatrix}3\\0\\2\end{pmatrix}=\begin{pmatrix}1\\2\\-6\end{pmatrix}$$

$\alpha=3\eta_1+2\eta_3$ 在自然基 e_1,e_2,e_3 下的坐标可直接算出

$$\alpha=3\eta_1+2\eta_3=3\begin{pmatrix}1\\-2\\1\end{pmatrix}+2\begin{pmatrix}0\\1\\-2\end{pmatrix}=\begin{pmatrix}3\\-4\\-1\end{pmatrix}$$

（3）自然基 e_1,e_2,e_3 到基 **I** 的过渡矩阵为 $A=(\varepsilon_1,\varepsilon_2,\varepsilon_3)$，所以 $\beta=(4,1,-2)^T$ 在基 **I** 下的坐标为

$$Y'=A^{-1}Y=\frac{1}{2}\begin{pmatrix}1&0&1\\1&0&-1\\-1&2&1\end{pmatrix}\begin{pmatrix}4\\1\\-2\end{pmatrix}=\begin{pmatrix}1\\3\\-2\end{pmatrix}$$

1.3 线性子空间

1.3.1 子空间的概念

定义1.8 设V为数域P上的线性空间，W是线性空间V的非空子集，若W关于V中的线性运算也构成数域P上的线性空间，则称W是V的线性子空间，简称子空间.

对任何线性空间V，显然由中单个零向量构成的子集是的子空间，称为的零子空间,记为{0}; V本身也是V的子空间. 这两个子空间称为V的平凡子空间. V的其它子空间称为V的非平凡子空间.

若W V，且W V，称 W是V的真子空间。

注：线性子空间也是线性空间，所以前面有关基、维数与坐标等概念，对线性子空间也成立.

定理1.2 设 W是线性空间 V的非空子集，则 W是 V的子空间的充要条件是: W对 V中的线性运算封闭.

证明：必要性是显然的，只证充分性.

因W对V中的线性运算封闭，则只需验证满足定义1.2中的八条运算规则.

因为取k=0，$k\alpha=0\in W$；又取$k=-1$，$k\alpha=-\alpha\in W$，满足规则(3), (4)，即W中存在零元素负元素.又因为W V，所以V中加法与数乘关于定义1.2中的其余六条运算规则对W中的元素运算时必须满足.故由定义1.2知：W是线性空间，从而是V的子空间.

例1.10 设$A\in P^{n\times n}$,证明全体与A可交换的n阶矩阵作成$P^{n\times n}$的一个子空间，记为$C(A)$.

思考：设W_1，W_2都是V的真子空间，则存在$\alpha\in V$，

而$\alpha\notin W_1$且$\alpha\notin W_2$同时成立。

例1.11 n阶上三角实矩阵的集合、下三角实矩阵的集合、实对角矩阵的集合都是线性空间$P^{n\times n}$的子空间.

例1.12 函数集合$C_1=\{f(x)\in C[a,b], f(a)=0\}$是线性空间$C$[a,b]的子空间.

例1.13 函数集合$C_2=\{f(x)\in C[a,b], f(a)=1\}$ 不是线性空间C[a,b]的子空间.

例1.14 取线性空间$P_4[x]$的子集

$$W=\{p(x)=a_3x^3+a_2x^2+a_1x+a_0 \mid a_2+a_1+a_0=0, a_i\in R\}$$

证明W是$P_4[x]$的子空间，并求W的维数.

例1.15 设V为数域P上的线性空间，$\alpha_1,\alpha_2,\ldots,\alpha_m$是$V$中的一组向量，则

$$span(\alpha_1,\alpha_2,\cdots,\alpha_m)=\{k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m \mid k_1,k_2,\cdots,k_m\in P\}$$

是V的子空间，称为由向量$\alpha_1,\alpha_2,\ldots,\alpha_m$的生成子空间，$\alpha_1,\alpha_2,\ldots,\alpha_m$称为该子空间的生成元.

生成子空间的重要意义在于：n维线性空间V是由它的基$\alpha_1,\alpha_2,\ldots,\alpha_m$生成的子空间，即$V=span\{\alpha_1,\alpha_2,\ldots,\alpha_m\}$.

易证对一般的线性空间，成立：

(1) $\alpha_1,\alpha_2,\ldots,\alpha_m$的任一极大无关组就是$W=span\{\alpha_1,\alpha_2,\ldots,\alpha_m\}$的一组基，且$\dim W=R\{\alpha_1,\alpha_2,\ldots,\alpha_m\}$

(2) 两个向量组$\alpha_1,\alpha_2,\ldots,\alpha_m$与$\beta_1,\beta_2,\ldots,\beta_s$等价 $\Leftrightarrow$ $span\{\alpha_1,\alpha_2,\ldots,\alpha_m\}=span\{\beta_1,\beta_2,\ldots,\beta_s\}$.

由于线性空间V的子空间W是V的一个子集，因此W中线性无关的向量个数不可能比V中的更多，所以$\dim W\le\dim V$.

进一步有下列定理：

定理1.3 n线性空间V的任何一个子空间W的基都可以扩充成V的一组基.

1.3.2 子空间的交与和

定义1.9 设W_1与W_2是线性空间V的两个子空间，称

$$W_1\cap W_2=\{\alpha \mid \alpha\in W_1\text{且}\alpha\in W_2\}$$

为W_1与W_2的交. 而称

$$W_1+W_2=\{\alpha_1+\alpha_2 \mid \alpha_1\in W_1,\ \alpha_2\in W_2\}$$

为W_1与W_2的和.

定理1.4 设V是数域P上的线性空间，W_1与W_2是V的两个子空间，则$W_1\cap W_2$与W_1+W_2都是V的子空间.

例1.16 设W_1与W_2分别是齐次线性方程组$AX=0$和$BX=0$的解空间，则$W_1\cap W_2$为线性方程组

$$\begin{cases}AX=\mathbf{0}\\BX=\mathbf{0}\end{cases}$$

的解空间.

例1.17

已知$C^{2\times2}$的子空间

$$W_1=\left\{\begin{pmatrix}x & x\\ y & y\end{pmatrix} \middle| x,y\in C\right\},\quad W_2=\left\{\begin{pmatrix}x & -y\\ -x & y\end{pmatrix} \middle| x,y\in C\right\}$$

分别求$W_1\cap W_2$, W_1+W_2 的一组基及维数。

定理1.5 设W_1和W_2是线性空间V的子空间，则

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

定理1.5 设W_1和W_2是线性空间V的子空间，则

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

注 一般地

$$\dim(W_1+W_2)\le\dim(W_1)+\dim(W_2)$$

1.3.3 子空间的直和

定义1.10：设V_1和V_2都是线性空间V的子空间，若对任意的

$$\alpha\in V_1+V_2 \text{ 都有 } \alpha=\alpha_1+\alpha_2,\ (\alpha_i\in V_i)$$

且是唯一的，这个和V_1+V_2就称为直和，记为

$$V_1\oplus V_2$$

定理1.6 设W_1,W_2是线性空间V的子空间，则下列命题等价

(1) W_1+W_2是直和;

(2) W_1+W_2中的零元素分解唯一，即由

$$0=\alpha_1+\alpha_2,\ \alpha_i\in W_i$$

则$\alpha_1=\alpha_2=0$;

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

(4) 对于W_1,W_2分别取一组线性无关的向量组，它们合起来仍然线性无关;

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

例 设α,β 线性无关，则$L(\alpha)+L(\beta)$是直和，而$L(\alpha,\alpha+\beta)+L(\beta)$不是直和.

设$V_1,V_2,\dots,V_s$是线性空间V的子空间，如果和$V_1+V_2+\dots+V_s$中的每个向量α表法唯一，

称和$V_1+V_2+\dots+V_s$是直和，记为 $V_1\oplus V_2\oplus\cdots\oplus V_s$

定理 3：设 $V_1,V_2,\cdots,V_s$ 是线性空间V的子空间，则下列命题相互等价：

（1） $W=V_1+V_2+\cdots+V_s$ 是直和；

（2） 零向量表示法唯一；

（3） $V_i\cap(\sum_{j\ne i}V_j)=\{0\}$.

（4） $\dim(W)=\sum\dim(V_i)$.

定理1.7 设W_1是n维线性空间V的一个子空间，则必存在V的一个子空间W_2，使

$$V=W_1\oplus W_2.$$

例1.18
设V是数域P上的2维空间，V的一组基为α_1, α_2, V的两个子空间为

$$W_1=\{k_0(\alpha_1+\alpha_2)|k_0\in P\},$$
$$W_2=\{k_1\alpha_1+k_2\alpha_2|k_1,k_2\in P, k_1+k_2=0\}$$

证明个$V=W_1\oplus W_2$.

注意：先证明$V=W_1+W_2$，再证明W_1+W_2为直和.

1.4 线性变换

1.4.1 线性变换的概念

定义 1.10 线性映射（变换）的要点：

(i) T是线性空间V到线性空间U的映射：

$$T: V\to U$$

(ii) T保持线性运算：

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

当$V=U$时，T 称为线性空间V上的线性变换。

例1.17 V中的数乘变换T_λ：

c是P中的数，$\forall\alpha\in V$，$T_c(\alpha)=c\alpha$.

特例：c=1，T_λ是恒等变换, 记为I

c=0，T_c是零变换, 记为 O

注：可以在任何线性空间中定义数乘变换.

例1.18 $P^{n\times n}$中的变换 T_A: 设$A\in P^{n\times n}$是一个给定的矩阵，$\forall X\in P^{n\times n}$，$T_A(X)=AX$. 则$T_A$是$P^{n\times n}$的一个线性变换.

例1.19 $P_n[X]$中的微分变换：$T: f(X)\to f'(x)$

注：以上三个例题中的变换都是线性变换.

例1.20 在$P_n[x]$中，定义$T(p)$=1, $\forall p\in P_n[X]$
则T 不是$P_n[x]$上的线性变换.

1.4.2 线性变换T的性质：

1. $T(0)=0$; $T(-\alpha)=-T(\alpha)$.

2. $T\left(\sum_{i=1}^{r}k_i\alpha_i\right)=\sum_{i=1}^{r}k_iT(\alpha_i)$.

注：线性变换保持线性相关性不变。

3. 若$\alpha_1,\alpha_2,\ldots,\alpha_r$线性相关,则$T(\alpha_1),T(\alpha_2),\ldots,T(\alpha_r)$也线性相关。

此命题的逆命题不真，即线性变换可能把线性无关的向量组映射成线性相关的向量组。

4. 若线性变换T是单射,则T把线性无关的向量组变成线性无关的向量组.

1.4.3 线性变换的运算

定义1.11 设V是数域P上的线性空间, T_1, T_2都是V上的线性变换, 则定义如下运算:

(1) 变换的加法: T_1+T_2: $(T_1+T_2)(\alpha)=T_1(\alpha)+T_2(\alpha)$

(2) 变换的数乘: kT: $(kT)(\alpha)=kT(\alpha)$

(3) 变换的乘法: T_1T_2: $(T_1T_2)(\alpha)=T_1(T_2(\alpha))$

(4) 可逆变换: 对变换T_1, 若存在变换T_2, 使得

$T_1T_2=T_2T_1=I$（恒等变换），

则称T_1为可逆变换, T_2是T_1的逆变换, 记为

$T_2=T_1^{-1}$.

关于线性变换的运算，有以下几点值得一提：

(1) 上述线性变换运算的结果仍是线性空间V上的线性变换；

(2) 线性变换T可逆当且仅当T为一一对应的；

(3) 若线性变换T可逆，则其逆变换是唯一的；

(4) 线性变换的乘法一般不满足交换律，即 $T_1T_2 \neq T_2T_1$；

(5) 对线性变换T，当n个T相乘时，常用T的n次幂来表示，即

$$T^n = \overbrace{T \cdot T \cdots T}.$$

1.5 线性变换的矩阵

1.5.1 线性变换在给定基下的矩阵

定义1.12 设T是n维线性空间V上的一个线性变换，$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$是$V$的一组基，如果这组基在线性变换$T$下的象$T(\varepsilon_1), T(\varepsilon_2), \dots, T(\varepsilon_n)$由这组基线性表示为

$$\begin{cases} T(\varepsilon_1) = a_{11}\varepsilon_1 + a_{21}\varepsilon_2 + \cdots + a_{n1}\varepsilon_n, \\ T(\varepsilon_2) = a_{12}\varepsilon_1 + a_{22}\varepsilon_2 + \cdots + a_{n2}\varepsilon_n, \\ \cdots\cdots\cdots\cdots \\ T(\varepsilon_n) = a_{1n}\varepsilon_1 + a_{2n}\varepsilon_2 + \cdots + a_{nn}\varepsilon_n, \end{cases}$$

即

$$T(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)A.$$

$$T(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)A$$

其中

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

则n阶矩阵A就称为线性变换T在基$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$下的矩阵，

其中A的第j列就是$T(\varepsilon_j)$在基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$下的坐标.

显然: 矩阵A由基在T下的像$T(\varepsilon_1), T(\varepsilon_2), \dots, T(\varepsilon_n)$唯一确定.

例1.21 在R^3中，定义$f(a,b,c)=(a,b,0)$为投影变换.

(1) 求f在自然基$e_1=(1,0,0), e_2=(0,1,0), e_3=(0,0,1)$下的矩阵；

(2) 求f在基$\varepsilon_1=(1,0,0), \varepsilon_2=(-1,1,0), \varepsilon_3=(1,1,1)$下的矩阵.

例1.22 取

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

为线性空间$R^{2\times2}$的一组基，设 $A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$. 在$R^{2\times2}$上定义线性变换$T$为 $T(X)=AX, \forall X \in R^{2\times2}$. 求$T$在这一组基$E_{11}, E_{12}, E_{21}, E_{22}$下的矩阵.

n维线性空间V的线性变换与n阶矩阵之间存在一一对应关系，即有

定理1.8 设$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$是线性空间$V$的一组基，$T$ 和S 是V 的两个线性变换，且它们在基$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$下的矩阵分别是$A$和$B$，则

(1) $T+S$在基$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$下 的矩阵是$A+B$；

(2) kT在基$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$下的矩阵是$kA$；

(3) TS在基$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$下的矩阵是$AB$；

(4) T 可逆当且仅当矩阵A可逆，且T^{-1}在基$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$下的矩阵是$A^{-1}$.

(5) 设线性空间V中任一个向量α与其象$T(\alpha)$在基$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$

下的坐标分别为$X=(x_1, x_2, \dots, x_n)^T$与$Y=(y_1, y_2, \dots, y_n)^T$,

则 $Y=AX$.

注：这就是线性变换的坐标变换公式即

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

3 不同基下的变换矩阵关系

两组基: $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$, $\{\beta_1, \beta_2, \dots, \beta_n\}$,

$$(\beta_1, \beta_2, \dots, \beta_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)P$$

$$T(\alpha_1, \alpha_2, \dots, \alpha_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)A$$

$$T(\beta_1, \beta_2, \dots, \beta_n) = (\beta_1, \beta_2, \dots, \beta_n)B$$

则$B=P^{-1}AP$.

定理1.9 线性空间V的一个线性变换T在不同基下的矩阵是相似的. (P29)

例1.23 设T是$R^{2\times2}$上的线性变换，$\forall A\in R^{2\times2}$，有

$$T(A)=\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}A$$

求T在基

$$M_{11}=\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, M_{12}=\begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}, M_{21}=\begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix} M_{22}=\begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$$

下的矩阵.

1.6值域、核与不变子空间

1.6.1值域与核的定义

定义 1.14 (P.29) 线性变换的象空间和零空间

设线性映射T: $V\to U$,

值域 $R(T)=\{\beta\in U \mid \exists\alpha\in V, \beta=T(\alpha)\}$ $\quad U$

核空间 $N(T)=\{\alpha \mid \alpha\in V, T(\alpha)=0\}$, 也可记为Ker$T$.

定理1.10 $N(T)$, $R(T)$分别是V, U的子空间.

基于以上原因，所以T值域又称为T的像空间, T 的核子空间又称为T的零子空间.

1.6.2 值域与核的性质

定义1.14 设T是线性空间V上的线性变换，$R(T)$的维数称为T的秩, 记为rank$T$; 而$N(T)$的维数称为T的零度或亏度, 记为null$T$.

T 的秩=dim$R(T)$; T 的零度=dim$N(T)$

定理1.11 设T是n维线性空间V上的线性变换，且T在V的一组基$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$下的矩阵是$A$，则

(1) T的值域$R(T)$是$T(\varepsilon_1), T(\varepsilon_2), \dots, T(\varepsilon_n)$生成的子空间，即

$$R(T)=\mathrm{span}\{T(\varepsilon_1), T(\varepsilon_2), \dots, T(\varepsilon_n)\}.$$

(2) T的秩 $=r(A)$.

例1.24 由例1.31知R^3上的投影变换$f:(a,b,c)\to(a,b,0)$, 在自然基$e_1=(1,0,0)$, $e_2=(0,1,0)$, $e_3=(0,0,1)$下的矩阵为

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

由定理1.11知的T秩 =2. 事实上, 由例1.34知: R^3上的投影变换f的值域就是xoy平面.

定理1.12 设V, U分别是数域P上的n维和m维线性空间，

$T: V\to U$的线性映射，则

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

设A 是$m\times n$矩阵，称 $R(A)=\{Ax \mid x\in C^n\}$

为矩阵A的值域; $N(A)=\{x \mid x\in C^n, Ax=0\}$

为A的核。

$\dim R(A)$ $\dim N(A)$ 称为A的秩和零度.

推论

（1） $\dim R(T)+\dim N(T)=\dim V$

（2） $\dim R(A)=rank(A)$

（3） $\dim R(A)+\dim N(A)=n$，n为A的列数。

例1.26 设 $A=\begin{pmatrix}1 & 1\\ -1 & -1\end{pmatrix}$ 在$R^{2\times 2}$上的线性变换定义为

$$TX=AX.$$

求T的值域$R(T)$及核子空间$N(T)$基与维数，并问

$R(T)+N(T)$是否是直和？

注：书上P34页：不变子空间

1.7 向量空间的同构

1.7.1 定义

设V, U都是数域P上的线性空间，如果映射

$\sigma: V\to U$，具有如下性质：

i) σ为双射

ii) $\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta), \forall \alpha, \beta\in V$

iii) $\sigma(k\alpha)=k\sigma(\alpha), \forall k\in P, \forall \alpha\in V$

则称σ是V到U的一个同构映射，并称线性空间V和U同构，记作 $V\cong U$.

1.7.2 同构的有关结论

1、数域P上任一n维线性空间都与P^n同构.

2、设V, U是数域P上的线性空间，σ是V到U的同构映射，则有

1) $\sigma(0)=0$，$\sigma(-\alpha)=-\sigma(\alpha)$.

2) $\sigma(k_1\alpha_1+k_2\alpha_2+\ldots+k_r\alpha_r)$

$=k_1\sigma(\alpha_1)+k_2\sigma(\alpha_2)+\ldots+k_r\sigma(\alpha_r)$，

其中 $k_i\in P, \alpha_i\in V, i=1,2,\ldots,r$.

3) V中向量组$\alpha_1, \alpha_2, \ldots, \alpha_r$线性相关（线性无关）的充要条件是它们的象$\sigma(\alpha_1), \sigma(\alpha_2), \ldots, \sigma(\alpha_r)$ 线性相关（线性无关）.

4) $\dim V=\dim U$.

5) $\sigma: V\to U$的逆映射σ^{-1}为U到V的同构映射.

6) 若W是V的子空间，则W在σ下的象集

$$\sigma(W)=\{\sigma(\alpha) \mid \alpha\in W\}$$

是的U子空间，且$\dim W=\dim\sigma(W)$.

注: 由2可知，同构映射保持零元、负元、线性组合及线性相关性，并且同构映射把子空间映成子空间.

3、两个同构映射的乘积还是同构映射.

证: 设σ: $V\to U$, τ: $U\to W$为线性空间的同构映射，则乘积 $\tau\circ\sigma$是 $V\to W$ 的1−1 映射. 任取$\alpha,\beta\in V$, $k\in P$, 有

$$\tau\circ\sigma(\alpha+\beta)=\tau(\sigma(\alpha)+\sigma(\beta))$$
$$=\tau(\sigma(\alpha))+\tau(\sigma(\beta))=\tau\circ\sigma(\alpha)+\tau\circ\sigma(\beta)$$
$$\tau\circ\sigma(k\alpha)=\tau(\sigma(k\alpha))=\tau(k\sigma(\alpha))$$
$$=k\tau(\sigma(\alpha))=k\tau\circ\sigma(\alpha)$$

所以，乘积$\tau\circ\sigma$是 V到W的同构映射.

NJUPT

4、数域P上的两个有限维线性空间V_1, V_2同构 $\Leftrightarrow$ $\dim V_1=\dim V_2$.

证：($\Rightarrow$) 若$V_1\cong V_2$, 由性质2之4) 即得

$$\dim V_1=\dim V_2.$$

($\Leftarrow$) （法一）若$\dim V_1=\dim V_2$,

由性质1, 有$V_1\cong P^n$, $V_2\cong P^n$,

$\therefore$ $V_1\cong V_2$.

NJUPT

"$\Leftarrow$"(法二：构造同构映射)

设$\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n$和$e_1,e_2,\dots,e_n$分别为$V_1$和$V_2$的一组基.

定义σ: $V_1\to V_2$ 使

$$\forall\alpha=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n\in V_1,$$
$$\sigma(\alpha)=a_1e_1+a_2e_2+\cdots+a_ne_n.$$

则σ就是V_1到V_2的一个映射.

又任取$\alpha,\beta\in V$, 设 $\alpha=\sum_{i=1}^{n}a_i\varepsilon_i$, $\beta=\sum_{i=1}^{n}b_i\varepsilon_i$,

若$\sigma(\alpha)=\sigma(\beta)$, 即 $\sum_{i=1}^{n}a_ie_i=\sum_{i=1}^{n}b_ie_i$, 则 $a_i=b_i$,

$i=1,2,\cdots,n$, 从而, $\alpha=\beta$, 所以σ是单射.

NJUPT

任取 $\alpha'\in V_2$, 设 $\alpha'=\sum_{i=1}^{n}a_ie_i$,

则有 $\alpha=\sum_{i=1}^{n}a_i\varepsilon_i\in V_1$, 使 $\sigma(\alpha)=\alpha'$.

所以σ是满射.

再由σ的定义，有 $\sigma(\varepsilon_i)=e_i$, $i=1,2,\cdots,n$

易证，对 $\forall\alpha,\beta\in V_1$, $\forall k\in P$ 有

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

所以σ是V_1到V_2的一个同构映射，故$V_1\cong V_2$.

NJUPT

例1.25 把复数域看成实数域R上的线性空间，证明：$C\cong R^2$.

证法一：证维数相等

首先，$\forall x\in C$, x 可表成 $x=a1+bi$, $a,b\in R$

其次，若 $a1+bi=0$, 则 $a=b=0$.

所以，1，i 为C的一组基， $\dim C=2$.

又$\dim R^2=2$,

所以 $\dim C=\dim R^2$, 故 $C\cong R^2$.

NJUPT

证法二：构造同构映射

作对应 $\sigma: C\to R^2$, $\sigma(a+bi)=(a,b)$.

则σ为C到R^2的一个同构映射.

例1.25 全体正实数R^+关于加法$\oplus$与数量乘法 $\circ$:

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

作成实数域R上的线性空间. 把实数域R看成是自身上的线性空间.

证明：$R^+\cong R$, 并写出一个同构映射.

NJUPT

证：作对应 $\sigma: R^+ \to R$, $\sigma(a)=\ln a$, $\forall a\in R^+$

易证σ为R^+到R的1－1对应.

且对 $\forall a,b\in R^+$, $\forall k\in R$,有

$$\sigma(a\oplus b)=\sigma(ab)=\ln ab=\ln a+\ln b=\sigma(a)+\sigma(b)$$

$$\sigma(k\circ a)=\sigma(a^k)=\ln a^k=k\ln a=k\sigma(a)$$

所以，σ为R^+到R的同构映射．故$R^+\cong R$.

方法二：作对应 $\tau: R\to R^+, \tau(x)=e^x$, $\forall x\in R$

易证：τ为R到R^+的1－1对应，而且也为同构映射.

事实上，τ为σ的逆同构映射.

练习

设集合 $W=\left\{\begin{pmatrix} a & b\\ -b & a\end{pmatrix}\middle| a,b\in R\right\}$

1) 证明：W为$R^{2\times 2}$的子空间，并求出W的维数与一组基.

2) 证明：复数域C看成R上的线性空间与W同构，并写出一个同构映射.

定理1.13 设V,U是有限维线性空间,线性变换

$T:V\to U$.

则T是单射当且仅当$N(T)=\{0\}$;

T是满射当且仅当$R(T)=U$.

定理1.14 设V是n维线性空间,线性变换$T: V\to V$

则以下条件等价：

（1） T是单射；

（2） T是满射；

（3） T是双射.

例1.27 平面上全体向量，对如下定义的加法和数乘

$$\alpha\oplus\beta=\alpha-\beta \qquad k\circ\alpha=-k\alpha$$

则R^2按照上述定义不构成R上的线性空间。

例1.28 设 $A=\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}$, 记

$$L(A)=\left\{B\middle| B\in R^{2\times 2}, AB=BA\right\}$$

求证$L(A)$为$R^{2\times 2}$的线性子空间，并求$\dim L(A)$.

例1.29 设有R^3的两个子空间：

$$W_1=\{(x_1,x_2,x_3)\mid 2x_1+x_2-x_3=0\}$$

$$W_2=\{(x_1,x_2,x_3)\mid x_1+x_2=0,\ 3x_1+2x_2-x_3=0\}$$

分别求子空间W_1+W_2, $W_1\cap W_2$的基与维数.

例1.30 设W_1, W_2分别是齐次线性方程组

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

与

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

的解空间，试证明$R^n=W_1\oplus W_2$.

第2章 内积空间与等距变换

主要内容

2.1 内积空间的基本概念

2.2 标准正交基与Schmidt正交化

2.3 正交子空间

2.4 等距变换

2.1 内积空间的基本概念

设V为数域P上的线性空间，如果按照某种对应法则，使得V中任意两个向量α, β都可以确定一个数(α, β)，且这个对应法则满足：对$\forall \alpha, \beta, \gamma \in V, k \in P$，有

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

（2）齐次性：$(k\alpha, \beta)=k(\alpha, \beta)$

（3）可加性：$(\alpha+\gamma, \beta)=(\alpha, \beta)+(\gamma, \beta)$

（4）正定性：$(\alpha, \alpha)\geq 0$，当且仅当$\alpha=0$时$(\alpha,\alpha)=0$.

则称该对应法则为V上的一个内积，复数(α,β)称为α与β的内积. 定义了内积的线性空间称为内积空间.

当$P=R$时，定义了内积的实线性空间V称为欧几里德空间（简称欧氏空间），也称实内积空间.

当$P=C$时，定义了内积的复线性空间V称为酉空间，也称复内积空间.

例2.1 在实线性空间R^n (C^n)中，对任意两个向量

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

定义内积为

$$(\alpha, \beta)=a_1b_1+a_2b_2+\cdots+a_nb_n=\sum_{i=1}^{n}a_ib_i=\alpha\beta^T \qquad (2.1)$$

$$(\alpha, \beta)=a_1\overline{b_1}+a_2\overline{b_2}+\cdots+a_n\overline{b_n}=\sum_{i=1}^{n}a_i\overline{b_i}=\alpha\beta^H \qquad (2.1')$$

易验证这样定义的(α,β)满足内积的4个条件，所以式(2.1)是R^n的一种内积 (式(2.1')是C^n的一种内积)，此内积称为R^n (C^n)的标准内积，记为$\alpha\beta^T$ ($\alpha\beta^H$).

其中β^H表示的β共轭转置向量，即 $\beta^H=\overline{\beta}^T$.

例2.3 在连续实函数组成的实线性空间中$C[a,b]$，对任意两个连续实函数$f(x), g(x)$，定义内积为

$$(f(x), g(x))=\int_a^b f(x)g(x)dx$$

利用定积分的性质可以证明这样定义的

$$(f(x), g(x))$$

是$C[a,b]$的内积.

向量的长度与夹角

定义2.2 在欧氏空间V中，对$\alpha\in V$，称非负实数$\sqrt{(\alpha, \alpha)}$为向量α的长度，记为记为$\|\alpha\|$. 当 $\|\alpha\|=1$ 时，称α为单位向量.

容易验证，向量的长度具有下列性质：

非负性：$\forall \alpha\in V, \|\alpha\|\geq 0$. 且$\|\alpha\|=0 \Leftrightarrow \alpha=0$.

齐次性：$\forall \alpha\in V,\ k\in R, \|k\alpha\|=|k|\|\alpha\|$；

三角不等式：$\|\alpha+\beta\|\leq\|\alpha\|+\|\beta\|$.

对任意非零向量α，向量$\dfrac{\alpha}{\|\alpha\|}$是与$\alpha$同方向长度的单位向量，由$\alpha$求$\dfrac{\alpha}{\|\alpha\|}$的过程称为把向量$\alpha$单位化.

Cauchy—Schwarz不等式

设V为内积空间，对$\forall \alpha, \beta \in V$, 有

$$\left|(\alpha,\beta)\right|^2 \le (\alpha,\alpha)(\beta,\beta) = \|\alpha\|^2\,\|\beta\|^2 \qquad \textbf{(2.2)}$$

其中等号当且仅当α与β线性相关时成立.

Cauchy—Schwarz不等式

设V为内积空间，对$\forall \alpha, \beta \in V$, 有

$$\left|(\alpha,\beta)\right|^2 \le (\alpha,\alpha)(\beta,\beta) = \|\alpha\|^2\,\|\beta\|^2 \qquad \textbf{(2.2)}$$

其中等号当且仅当α与β线性相关时成立.

$$\left|(\alpha,\beta)\right| \le \sqrt{(\alpha,\alpha)(\beta,\beta)} = \|\alpha\|\,\|\beta\|$$

在R^n中不等式

$$\left|a_1b_1 + a_2b_2 + \cdots + a_nb_n\right| \le \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}\,\sqrt{b_1^2 + b_2^2 + \cdots + b_n^2}$$

在$C[a,b]$中不等式

$$\left|\int_a^b f(x)g(x)dx\right| \le \sqrt{\int_a^b f^2(x)dx}\cdot\sqrt{\int_a^b g^2(x)dx}$$

定义2.3 对欧氏空间V中任意非零向量α, β, 定义

$$\theta = \arccos\frac{(\alpha,\beta)}{\|\alpha\|\,\|\beta\|}$$

为非零向量α与β的夹角. 若$(\alpha,\beta)=0$, 则称向量α与β正交，记为 $\alpha\perp\beta$.

由定义2.3可知，与几何向量一样有

(1) $\forall \alpha\in V$, 有$0\perp\alpha$

(2) $\forall \alpha\in V$, 有$\alpha\perp\alpha \Leftrightarrow \alpha=0$

(3) 若$\forall \alpha,\beta\in V$均是非零元素, 则$\alpha\perp\beta \Leftrightarrow$ α与β的夹角为$\pi/2$.

2.2　标准正交基与Schmidt正交化

一、 标准正交基

定义2.4 在内积空间V中，一组两两正交的非零向量称为V中的正交向量组.

定理2.2 若$\alpha_1,\alpha_2, \ldots, \alpha_r$ 是正交向量组，则 $\alpha_1,\alpha_2, \ldots, \alpha_r$ 线性无关.

定义2.5 设 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$是内积空间$V$的一组基, 且它们两两正交, 则称 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$为$V$ 的一组**正交基**.

当正交基的每一个向量都是**单位向量**时，则称这组正交基为V的**标准正交基**.

显然，由定义2.5知, $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$是$n$维内积空间$V$的标准正交基的充要条件是:

$$(\varepsilon_i,\varepsilon_j) = \begin{cases} 0, & i \ne j; \\ 1, & i = j. \end{cases}$$

标准正交基的特性:

定理2.3 设 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$是内积空间$V$的一组标准正交基，有下列结论成立:

(1) 对$\forall \alpha\in V$, 设向量α在基 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$下的坐标为$X=(x_1, x_2, \ldots, x_n)^T$, 则$x_i=(\alpha, \varepsilon_i)$, $i=1,2, \ldots, n$;

标准正交基的特性:

定理2.3 设 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$是内积空间$V$的一组标准正交基，有下列结论成立:

(1) 对$\forall \alpha \in V$, 设向量α在基 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$下的坐标为$X=(x_1, x_2, \ldots, x_n)^T$, 则$x_i=(\alpha, \varepsilon_i)$, $i=1,2, \ldots, n$;

(2) 若α, β在该基下的坐标分别为 X 和 Y，则 $(\alpha,\beta)=(X,Y)$;

标准正交基的特性:

定理2.3 设 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$是内积空间$V$的一组标准正交基，有下列结论成立:

(1) 对$\forall \alpha \in V$, 设向量α在基 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$下的坐标为$X=(x_1, x_2, \ldots, x_n)^T$, 则$x_i=(\alpha, \varepsilon_i)$, $i=1,2, \ldots, n$;

(2) 若α, β在该基下的坐标分别为 X 和 Y，则 $(\alpha,\beta)=(X,Y)$;

(3) 对$\forall \alpha \in V$, 设向量α在基$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$下的坐标为$X=(x_1, x_2, \ldots, x_n)^T$，则

$$\|\alpha\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

2.2.2 Schmidt正交化方法

给定内积空间V的一组基，能否由此构造出内积空间V的一组标准正交基，如何构造？

设$\alpha_1,\alpha_2, \ldots, \alpha_r$是内积空间$V$的一组线性无关的向量组, 要求的一组两两正交的单位向量$e_1,e_2, \ldots, e_r$，使$e_1, e_2, \ldots, e_r$与$\alpha_1,\alpha_2, \ldots, \alpha_r$等价，这一过程称为把向量组$\alpha_1,\alpha_2, \ldots, \alpha_r$ 标准正交化.

显然，e_i是$\alpha_1,\alpha_2, \ldots, \alpha_r$的线性组合.

施密特正交化过程

设$\alpha_1, \alpha_2, \ldots ,\alpha_r$ 是内积空间V的一个基, 要求V的一个标准正交基, 就是找一组两两正交的单位向量组$e_1, e_2, \ldots, e_r$, 使$\alpha_1, \alpha_2, \ldots,\alpha_r$与$e_1,e_2, \ldots, e_r$等价.

若$\alpha_1, \alpha_2, \ldots, \alpha_r$ 是线性空间V的一组基，

则 $e_1, e_2, \ldots, e_r$ 就是V的一组标准正交基.

$\alpha_1, \alpha_2, \ldots,\alpha_r$ 是内积空间V的一个基, 求V的一组两两正交的单位向量组$e_1, e_2, \ldots, e_r$ 且与$\alpha_1, \alpha_2, \ldots,\alpha_r$等价.

(1) 正交化，取 $\beta_1=\alpha_1$,

$$\beta_2=\alpha_2-\frac{(\beta_1,\alpha_2)}{(\beta_1,\beta_1)}\beta_1,$$

$$\beta_3=\alpha_3-\frac{(\beta_1,\alpha_3)}{(\beta_1,\beta_1)}\beta_1-\frac{(\beta_2,\alpha_3)}{(\beta_2,\beta_2)}\beta_2,$$

$$\cdots\cdots$$

$$\beta_r=\alpha_r-\frac{(\beta_1,\alpha_r)}{(\beta_1,\beta_1)}\beta_1-\frac{(\beta_2,\alpha_r)}{(\beta_2,\beta_2)}\beta_2-\cdots-\frac{(\beta_{r-1},\alpha_r)}{(\beta_{r-1},\beta_{r-1})}\beta_{r-1},$$

那么$\beta_1, \beta_2, \ldots, \beta_r$两两正交，且 $\beta_1, \beta_2, \ldots, \beta_r$与$\alpha_1, \alpha_2, \ldots,\alpha_r$等价.

(2) 单位化，取

$$e_1=\frac{\beta_1}{|\beta_1|},\quad e_2=\frac{\beta_2}{|\beta_2|},\ \cdots, e_r=\frac{\beta_r}{|\beta_r|},$$

向量组$e_1, e_2, \ldots, e_r$即为一组标准正交基.

上述有线性无关向量组通过正交化过程构造标准正交组的过程称为施密特正交化过程.

例2.5 设$P_3[x]$是全体次数小于3的实系数多项式构成一个实线性空间，定义内积为

$$(f(x),g(x))=\int_{-1}^{1}f(x)g(x)dx,$$
$$\forall f(x),g(x)\in P_3[x]$$

不难验证这样定义的$(f(x),g(x))$是$P_3[x]$的内积，

试求$P_3[x]$的一组标准正交基.

在$P_3[x]$ 内积为 $(f(x),g(x))=\int_{-1}^{1}f(x)g(x)dx$,
求$P_3[x]$的一组标准正交基.

2.3 正交子空间

定义2.6 设W_1与W_2是内积空间V的非空子集，若对于任意$\alpha\in W_1$，$\beta\in W_2$，都有$(\alpha,\beta)=0$，则称W_1与W_2互相正交，记为$W_1\perp W_2$；若$\alpha\in V$，对任意$\beta\in W_1$都有$(\alpha,\beta)=0$，则称α与W_1正交，记为$\alpha\perp W_1$.

定义2.7 设V是一个内积空间，集合 $W^{\perp}=\{\alpha \mid \alpha\in V, \alpha\perp W\}$ 称为W的正交补.

注：由定义2.6知：若W_1与W_2是内积空间V两个互相正交的子空间，即$W_1\perp W_2$，则必有$W_1\cap W_2=\{0\}$，所以两个互相正交的子空间之和必为直和.

由定义2.7知：无论W是否为V的子空间，$W^{\perp}$必是V的子空间.

定理2.4 设V是一个n维内积空间, $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_r$, $\varepsilon_{r+1}, \varepsilon_{r+2}, \ldots, \varepsilon_n$ 是V的一组标准正交基 $(1\le r\le n)$，
记 $W=\mathrm{span}\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_r\}$,
$S=\mathrm{span}\{\varepsilon_{r+1}, \varepsilon_{r+2}, \ldots, \varepsilon_n\}$,
则 $S=W^{\perp}$, $W=S^{\perp}$.

下面定理讨论了内积空间分解为互相正交的子空间的直和问题.

定理2.5（内积空间正交直和分解）设V是一个n维内积空间, W是V的子空间，则$V=W\oplus W^{\perp}$.

注：由定理2.5知：对每个$\alpha\in V$, 有唯一的表示：

$$\alpha=\alpha_1+\alpha_2,\quad \alpha_1\in W,\quad \alpha_2\in W^{\perp},$$

称α_1为α沿着空间$W^{\perp}$向W的正交投影，α_2为α沿着空间W向$W^{\perp}$的正交投影

例2.6 设欧氏空间$P_3[x]$中的多项式$f(x)$与$g(x)$的内积定义为

$$(f(x),g(x))=\int_{-1}^{1}f(x)g(x)dx$$

选取$f_1(x)=x$，构造子空间$W=\mathrm{span}(x)$.

(1) 求$W^{\perp}$的一组正交基;

(2) 将$W^{\perp}$分解为两个正交的非零子空间的和.

解 (1) 设 $g(x)=k_0+k_1x+k_2x^2\in W^{\perp}$，则有

$$(f_1(x),g(x))=\int_{-1}^{1}f_1(x)g(x)dx$$

$$=\int_{-1}^{1}x(k_0+k_1x+k_2x^2)dx=0$$

得$k_1=0$，于是

$$W^{\perp}=\{g(x)|g(x)=k_0+k_2x^2,k_0,k_2\in R\}$$

取$W^{\perp}$的一组基$1,x^2$，并进行正交化可得

$$g_1(x)=1,\quad g_2(x)=x^2-1/3,$$

则 $g_1(x),g_2(x)$是$W^{\perp}$的一组正交基.

(2) 令

$$V_1=\mathrm{span}(g_1(x)),\quad V_2=\mathrm{span}(g_2(x))$$

则V_1与V_2正交，且$V=V_1\oplus V_2$.

2.4 等距变换

定义2.8 设T是内积空间V上的一个线性变换，若对$\forall\alpha,\beta\in V$成立$(T(\alpha),T(\beta))=(\alpha,\beta)$，则称$T$是等距变换. 特别地，当$V$是酉空间时，则称$T$是酉变换；当$V$是欧氏空间时，则称$T$是正交变换.

例2.7 在酉空间C^n中，对任意$X\in C^n$，作变换T: $TX=AX$，其中n阶方阵A为酉矩阵，则

$$(TX,TY)=Y^HA^HAX=Y^HX=(X,Y)$$

所以此变换是一个酉变换. 若在欧氏空间R^n中且方阵A为正交阵，则$TX=AX$是正交变换.

例2.8 设$H=I_n-2uu^H\in C^{n\times n}$，且$u\in C^n$，$u^Hu=1$，定义变换

$$H(\alpha)=\alpha-2uu^H\alpha$$

则H是C^n上的酉变换，称为Householder镜象变换，它将α映射为关于与u正交的$n-1$维空间的镜象.

事实上：

$$(H(\alpha),H(\beta))=(\alpha-2uu^H\alpha,\beta-2uu^H\beta)$$

$$=(\beta^H-2\beta^Huu^H)(\alpha-2uu^H\alpha)$$

$$=\beta^H\alpha=(\alpha,\beta).$$

定理2.6 设T是内积空间V上的一个线性变换，则下列命题等价：

(1) $(T(\alpha),T(\beta))=(\alpha,\beta),\forall\alpha,\beta\in V$;

(2) $\|T(\alpha)\|=\|\alpha\|,\forall\alpha\in V$;

定理2.6 设T是内积空间V上的一个线性变换，则下列命题等价：

(1) $(T(\alpha),T(\beta))=(\alpha,\beta), \forall \alpha, \beta \in V$;

(2) $\|T(\alpha)\| = \|\alpha\|, \forall \alpha \in V$;

当V是有限维时, 以上命题进一步与以下命题等价.

(3) 若$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$是$V$的一组标准正交基，则 $T(\varepsilon_1), T(\varepsilon_2), \ldots, T(\varepsilon_n)$也是$V$的一组标准正交基;

(4) T在V的任一组标准正交基$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$下的矩阵是酉矩阵。

第3章 矩阵的若当标准形

3.1 特征值与特征向量

任何线性空间，给定基后，我们对元素进行线性变换或线性运算时，只需用元素的坐标向量以及线性变换的矩阵即可。因此，在后面的内容中着重研究矩阵和向量。对角矩阵的形式比较简单，处理起来较方便，比如求解矩阵方程$Ax=b$时，将矩阵A对角化后很容易得到方程的解。以前我们学习过相似变化对角化。

那么，一个方阵是否总可以通过相似变化将其对角化呢？或者对角化需要什么样的条件呢？如果不能对角化，我们还可以做哪些处理使问题变得简单呢？

3.1.1特征值与特征向量

定义3.1 设T是数域P上的线性空间V上的线性变换，如果存在$\lambda\in P$，及$\xi\in V$，$\xi\neq 0$使得

$$T(\xi)=\lambda\xi \qquad (3.1)$$

则称λ是T的一个特征值，而ξ称为T属于特征值λ的一个特征向量。

特征值与特征向量的求法

设线性变换T在n维线性空间V上基$\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$下矩阵$A=(a_{ij})_{n\times n}$，$\lambda$为$T$的一个特征值，设$\xi$是$\lambda$对应的特征向量，$\xi$在基$\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$下的坐标为$X$，

即 $T(\xi)=\lambda\xi$

$$T(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_1,\alpha_2,\ldots,\alpha_n)A$$

$$\xi=(\alpha_1,\alpha_2,\ldots,\alpha_n)X$$

$$T\xi=T(\alpha_1,\alpha_2,\ldots,\alpha_n)X=(\alpha_1,\alpha_2,\ldots,\alpha_n)AX$$

$$\lambda\xi=\lambda(\alpha_1,\alpha_2,\ldots,\alpha_n)X=(\alpha_1,\alpha_2,\ldots,\alpha_n)\lambda X$$

$T\xi=\lambda\xi$ 当且仅当 $AX=\lambda X$

特征值与特征向量的求法

设线性变换T在n维线性空间V上基$\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$下矩阵$A=(a_{ij})_{n\times n}$，$\lambda$为$T$的一个特征值，设$\xi$是$\lambda$对应的特征向量，$\xi$在基$\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$下的坐标为$X$，

$T(\xi)=\lambda\xi$ 相当于$AX=\lambda X$, 等价于$(\lambda I-A)X=0$, $X\neq 0$. 于是求T的特征值λ对应的特征向量ξ, 可以先求出T在基$\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$下矩阵$A$的特征值$\lambda$, 对应的特征向量$X$, 则$T$的特征值$\lambda$对应的特征向量为

$$(\alpha_1,\alpha_2,\ldots,\alpha_n)X.$$

特征值和特征向量的计算

设λ是n阶方阵A的特征值，x是A属于特征值λ的特征向量，则$Ax=\lambda x$或者

$$(\lambda I-A)x=0$$

则上述齐次线性方程组有非零解的充要条件是系数行列式等于零。即

$$|\lambda I-A|=0$$

记

$$f(\lambda)=|\lambda I-A|=\begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}$$

$$=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$$

称$f(\lambda)=|\lambda I-A|$为矩阵A的特征多项式。

$f(\lambda)=|\lambda I-A|=0$为矩阵$A$的特征方程。

$f(\lambda)$是最高项系数为1的n次多项式。它在复数域中有n个复根$\lambda_1, \lambda_2, \ldots,\lambda_n$是$A$的全部特征值. 如果将相同的写在一起，则$f(\lambda)$可以表示为

$$f(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s}$$

$\lambda_1,\lambda_2,\cdots,\lambda_s$ 两两互异，且$1\le m_i\le n$

$$m_1+m_2+\cdots+m_k=n$$

m_i称为特征值λ_i的代数重数.

求线性变换特征值与特征向量的步骤:

1.求线性变换T在任意基$\alpha_1,\alpha_2,\ldots,\alpha_n$下矩阵$A$的特征值,即求$|\lambda I-A|=0$的根$\lambda_1, \lambda_2, \ldots, \lambda_n$.

2. 对于A的每一个特征值λ_i,求出齐次线性方程组

$(\lambda_i I-A)X=0$的一个基础解系 $\xi_1,\xi_2,\cdots,\xi_s$,则$x=\sum_{i=1}^{s}k_i\xi_i$

就是A属于该特征值的全部特征向量，其中

$k_1,k_2,\cdots,k_s$不全为零.

3. $\beta_j=(\alpha_1,\alpha_2,\cdots,\alpha_n)\xi_j \quad j=1,2,\cdots,s$ 就是线性变换 T 属于这些特征值的线性无关的特征向量.

例3.1 设线性空间V的线性变换T在基$\alpha_1, \alpha_2, \alpha_3$下矩阵是$A$,

其中 $A=\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix}$，求其特征值和特征向量.

解

$$\det(\lambda I-A)=\begin{vmatrix}\lambda-1 & -2 & -2\\ -2 & \lambda-1 & -2\\ -2 & -2 & \lambda-1\end{vmatrix}=0$$

特征值：$\lambda_1=\lambda_2=-1 \qquad \lambda_3=5$

$$-I-A=\begin{pmatrix}-2&-2&-2\\-2&-2&-2\\-2&-2&-2\end{pmatrix}\to\begin{pmatrix}1&1&1\\0&0&0\\0&0&0\end{pmatrix}\quad \xi_1=\begin{pmatrix}1\\0\\-1\end{pmatrix},\xi_2=\begin{pmatrix}0\\1\\-1\end{pmatrix},$$

解齐次线性方程组$(-1I-A)x=0$得属于特征值$\lambda_1=\lambda_2=-1$线性无关的特征向量 为$\xi_1=(1,0,-1)^T$, $\xi_2=(0,1,-1)^T$,故A属于特征值-1的全部特征向量为$k_1\xi_1+k_2\xi_2$, 其中k_1,k_2不全为零。

解齐次线性方程组（$5I-A)x=0$得属于特征值$\lambda_3=5$线性无关的特征向量 为$\xi_3=(1,1,1)^T$, A属于特征值5的全部特征向量为$k_3\xi_3$, $k_3\neq0$.

3.1.2. 矩阵的迹与行列式

矩阵$A=(a_{ij})_{n\times n}$的所有对角元素之和称为矩阵A

的迹，记为 tr$A$. 即

$$\mathrm{tr}A=\sum_{i=1}^{n}a_{ii}$$

因 $f(\lambda)$是n次多项式，它在复数域内有n个根，它们就是矩阵A的全部特征值，即n阶矩阵在复数域内有n个特征值。

定理3.1 设 n阶矩阵$A=(a_{ij})_{n\times n}$ 的n个特征值为$\lambda_1, \lambda_2, \ldots, \lambda_n$，则

(1) $\lambda_1\lambda_2\ldots\lambda_n=|A|$

(2) $\lambda_1+\lambda_2+\ldots+\lambda_n=a_{11}+a_{22}+\ldots+a_{nn}=\mathrm{tr}(A)$.

$$\det A=\prod_{i=1}^{n}\lambda_i \qquad \mathrm{tr}A=\sum_{i=1}^{n}\lambda_i$$

推论1 设A, B分别 $m\times n$ 和 $n\times m$ 矩阵，则

(1) $\mathrm{tr}(AB)=\mathrm{tr}(BA)$

(2) Sylvester定理：

$$\det(\lambda I_m-AB)=\lambda^{m-n}\det(\lambda I_n-BA)$$

推论2 设$A,B\in P^{n\times n}$,则tr(AB)=tr(BA)

证明 设$A=(a_{ij})_{n\times n}$,$B=(b_{ij})_{n\times n}$, 则AB的对角线元素为

$$\sum_{k=1}^{n}a_{ik}b_{ki} \qquad i=1,\mathrm{L}\ n$$

而BA的对角线元素为 $\sum_{k=1}^{n}b_{ik}a_{ki} \qquad i=1,\mathrm{L}\ ,n$

于是 $\mathrm{tr}(AB)=\sum_{i=1}^{n}(\sum_{k=1}^{n}a_{ik}b_{ki})$

改变求和顺序

$$=\sum_{k=1}^{n}(\sum_{i=1}^{n}a_{ik}b_{ki})=\sum_{k=1}^{n}(\sum_{i=1}^{n}b_{ki}a_{ik})=\mathrm{tr}(BA)$$

3.1.3 特征子空间

设V是数域P上的n维线性空间，T是V的线性变换，对T的任一特征值λ_0，T属于λ_0的全部特征向量以及零向量一起所组成的集合构成V的一个子空间，称为T的属于特征值λ_0的特征子空间。记为V_{λ_0},即$V_{\lambda_0}=\{\alpha|T(\alpha)=\lambda_0\alpha\}$。

容易证明：V_{λ_0}是V在T下不变子空间.

定理3.2. T的特征子空间V_{λ_0}的维数是T属于特征值λ_0线性无关的特征向量的最大个数。

注. 特别地，方阵A的特征子空间V_{λ_0}的维数是A属于特征值λ_0线性无关的特征向量的最大个数, 也称为特征值λ_0的几何重数. 一个特征值λ_0的几何重数不会超过它的代数重数.

3.2 矩阵通过相似变换对角化

定义：设T是线性空间V一个线性变换，若存在一组基 $\alpha_1, \alpha_2, \ldots, \alpha_n$使得$T$在该基下的矩阵是对角阵$\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$，就称$T$可以对角化。

设T在基 $\beta_1, \beta_2, \ldots, \beta_n$下的矩阵是$A$，且由基 $\alpha_1, \alpha_2, \ldots, \alpha_n$到基$\beta_1, \beta_2, \ldots, \beta_n$的过渡矩阵是$P$，则 $P^{-1}AP=\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$.

用矩阵语言叙述就是

设$A\in C^{n\times n}$,如果存在可逆矩阵P使得$P^{-1}AP$是对角阵，就称A是可以对角化的。设T可以对角化，则

$$T(\alpha_1\ \alpha_2\ \ldots\alpha_n)=(\alpha_1\ \alpha_2\ \ldots\alpha_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$$

从而$T(\alpha_i)=\lambda_i\alpha_i$, $i=1,2,\ldots,n$, 表明T可以对角化的充要条件是T有n 个线性无关的特征向量。所以A可以对角化的充要条件是A有n个线性无关的特征向量。线性变换T在不同基下的矩阵相似, 而矩阵的相似关系是一个等价关系, 即满足 反身性，对称性，传递性.

且有如下性质：

相似矩阵的性质

设n阶矩阵A与B相似，则

1、$\det A=\det B$;

2、$\det(\lambda I-A)=\det(\lambda I-B)$;

3、A, B有相同的特征值;

4、A, B有相同的迹.

注：以上性质的逆命题均不成立.

例如:
$$A=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}, B=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

A, B均满足以上四条，但是A与B不相似, 事实上，B只能与B自己相似。

所以研究线性变换T的对角化问题，只要研究T在任意基下矩阵的对角化问题即可。

定理3.3设$A\in C^{n\times n}$, $\lambda_1,\lambda_2,\ldots,\lambda_s$是$A$的所有互异的特征值，则以下条件等价：

1）A可以对角化；

2）A有n个线性无关的特征向量；

3）A是单纯矩阵；

4）C^n有直和分解

$C^n=N(\lambda_1 I-A)\oplus N(\lambda_2 I-A)\oplus\ldots\oplus N(\lambda_s I-A)$.

注：1.单纯矩阵：矩阵A的任意k重特征值恰有k个线性无关的特征向量。

2. $N(\lambda_i I-A)=\{X\in C^n \mid AX=\lambda_i X\}$

证明：1）A可以对角化 $\Rightarrow$ 2）A有n个线性无关的特征向量

A可以对角化，则存在可逆矩阵P使得：

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \text{O} & \\ & & & \lambda_n\end{pmatrix}$$

令$P=(x_1, x_2, \ldots, x_n)$，则$AP=P\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \text{O} & \\ & & & \lambda_n\end{pmatrix}$

即$A x_i=\lambda_i x_i$, 由P可逆知：$x_1, x_2, \ldots, x_n$ 线性无关，故A有n个线性无关的特征向量。

2）A有n个线性无关的特征向量 $\Rightarrow$ 3）A是单纯矩阵

设$x_1, x_2,\ldots, x_n$是A的n个线性无关的特征向量，将它们分组为:

$x_{11}, x_{12}, \ldots, x_{1r_1}\in N(\lambda_1 I-A)$;

$x_{21}, x_{22}, \ldots, x_{2r_2}\in N(\lambda_2 I-A)$;

……

$x_{k1}, x_{k2}, \ldots, x_{kr_s}\in N(\lambda_s I-A)$.

则$r_1+r_2+\ldots+r_s=n$, 而$r_i=\dim N(\lambda_i I-A)\le m_i$ 代数重数

又因为 $\sum_{i=1}^{s}(m_i-r_i)=\sum_{i=1}^{s}m_i-\sum_{i=1}^{s}r_i=n-n=0$

所以，$m_i=r_i$，$i=1,2,\ldots,s$.

3）A是单纯矩阵

⇒ 4）C^n有直和分解

C^n=N($\lambda_1 I-A$)⊕N($\lambda_2 I-A$)⊕…⊕N($\lambda_s I-A$)

设 r_i = dimN($\lambda_i I-A$) = m_i，i=1,2, … , s.

因为 N($\lambda_1 I-A$)+N($\lambda_2 I-A$)+···+N($\lambda_k I-A$) C^n.

$$\dim\sum_{i=1}^{s} N(\lambda_i I-A)=\sum_{i=1}^{s}\dim N(\lambda_i I-A)=\sum_{i=1}^{s} m_i=n=\dim C^n.$$

所以N($\lambda_1 I-A$)⊕N($\lambda_2 I-A$)⊕…⊕N($\lambda_s I-A$) = C^n.

4）C^n有直和分解

C^n=N($\lambda_1 I-A$)⊕N($\lambda_2 I-A$)⊕…⊕N($\lambda_s I-A$)

⇒ 1）A可以对角化

取N($\lambda_i I-A$)的基$x_{i1}, x_{i2}, \ldots, x_{ir_i}$, i=1,2,···, s. 则

$$A(x_{11},x_{12},\cdots,x_{1r_1},x_{21},x_{22},\cdots,x_{2r_2},\cdots,x_{k1},x_{k2},\cdots,x_{kr_s})=$$

$$(x_{11},x_{12},\cdots,x_{1r_1},x_{21},x_{22},\cdots,x_{2r_2},\cdots,x_{k1},x_{k2},\cdots,x_{kr_s})\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{pmatrix}$$

令 $P=(x_{11},x_{12},\cdots,x_{1r_1},x_{21},x_{22},\cdots,x_{2r_2},\cdots,x_{k1},x_{k2},\cdots,x_{kr_s})$

则P可逆，且$P^{-1}AP=\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{pmatrix}$.

推论1. n阶方阵A有n个互异的特征值, 则必可对角化 (充分条件).

推论2. n阶矩阵A可与对角矩阵相似的充要条件是对A的任意一个k重特征值 ，均有r($I-A$)=$n-k$.从而对应于k重特征值恰有k个线性无关的特征向量.

证明:

由r($I-A$)=$n-k$, 得到齐次线性方程组($I-A$)X=0的基础解系含有k个向量, 所以A的任一k重特征值恰有k个线性无关的特征向量, 故A有n个线性无关的特征向量, 由定理3.3知A与对角矩阵相似。

例3.3 设 $A=\begin{pmatrix}0&-2&-3\\1&3&3\\0&0&1\end{pmatrix}$ 判断A是否可以对角化，

若能，求出可逆矩阵P使得$P^{-1}AP$是对角阵。

解：det(λI−A)=(λ−1)²(λ−2), A的特征值λ_1= $_2$=1,λ_3=2

(I−A)X=0的基础解系是

$$\alpha_1=\begin{pmatrix}-2\\1\\0\end{pmatrix},\ \alpha_2=\begin{pmatrix}-3\\0\\1\end{pmatrix}$$

(2I−A)X=0的基础解系是 $\alpha_3=\begin{pmatrix}1\\-1\\0\end{pmatrix}$

A有3个线性无关的特征向量α_1,α_2,α_3，所以A可以对角化，令

$$P=(\alpha_1,\alpha_2,\alpha_3)$$

则 $$P^{-1}AP=\begin{pmatrix}1&0&0\\0&1&0\\0&0&2\end{pmatrix}$$

补充例3.1 设 $A=\begin{pmatrix}3&1&0\\0&3&0\\0&0&4\end{pmatrix}$

判断A是否能够对角化？

解：A不能对角化。

因为A的特征值3是2重的，而r(3I−A)=2≠3−2,
所以A不能对角化。

补充例3.2
已知{α_1，α_2，α_3}是线性空间V的基，T是线性空间V如下的线性变换，

$T(\alpha_1) = \alpha_1$
$T(\alpha_2) = 2\alpha_2$
$T(\alpha_3) = \alpha_1 + t\,\alpha_2 + 2\,\alpha_3$

讨论：t为何值，T可以对角化.

解：T在基α_1，α_2，α_3下的矩阵为

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

T可以对角化当且仅当A可以对角化，当且仅当

r(2I−A)=3-2=1,当且仅当t=0.

补充例3.3 证明幂等变换(A^2=A)有对角矩阵表示.

证明：设λ是A的任一特征值，x是对应的特征向量，

则Ax=λx, A^2 x=λ^2x= λx,所以λ＝0，或者1.

设r(A)=r,则AX=0的基础解系含有n-r个线性无关的解向量$\alpha_1, \alpha_2,\cdots,\alpha_{n-r}$,也即A恰有n-r个属于特征值0的线性无关的特征向量。又A^2=A，所以r(I-A)=n-r,故(I-A)X=0基础解系恰含r个解向量$\alpha_{n-r+1},\cdots,\alpha_n$,则$\alpha_1, \alpha_2,\cdots,\alpha_{n-r},\cdots,\alpha_n$线性无关，令

P=($\alpha_1, \alpha_2,\cdots,\alpha_{n-r},\cdots,\alpha_n$),则$P^{-1}AP=\begin{pmatrix}I_r&0\\0&0\end{pmatrix}$

例3.4 设 $A=\begin{pmatrix}4&6&0\\-3&-5&0\\-3&-6&1\end{pmatrix}$

求A的相似对角形以及A^{100}.

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

A特征值 $\lambda_1=-2,\lambda_2=\lambda_3=1$，求得属于特征值-2

的线性无关的特征向量是 $\alpha_1=(-1,1,1)^T$

属于1的线性无关的特征向量 $\alpha_2=(-2,1,0)^T,\alpha_3=(0,0,1)^T$

故取 $P=\begin{pmatrix}-1&-2&0\\1&1&0\\1&0&1\end{pmatrix}$ ，则 $P^{-1}=\begin{pmatrix}1&2&0\\-1&-1&0\\-1&-2&1\end{pmatrix}$

$$P^{-1}AP=\begin{pmatrix}-2&0&0\\0&1&0\\0&0&1\end{pmatrix}\qquad A=P\begin{pmatrix}-2&0&0\\0&1&0\\0&0&1\end{pmatrix}P^{-1}$$

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

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

3.2.3 正规矩阵

定义3.6 n阶复矩阵A如果满足等式$AA^H=A^HA$，则称A是正规矩阵或规范矩阵.

注：如果$AA^H=I$, 称A为酉矩阵.

例如，所有的实对称矩阵，实反对称矩阵，正交矩阵都是实正规矩阵，所有的n阶Hermite矩阵，反Hermite矩阵，酉矩阵都是复正规矩阵.

Schur引理

设数λ_1，λ_2，…，λ_n 是n阶方阵A的特征值，则存在n阶酉矩阵U，使

$$U^H AU = U^{-1}AU = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

定理3.4 n阶方阵A酉相似于对角阵的充要条件是A为正规阵（实或复）.

例3.5 设A,B均为n阶正规矩阵，证明A与B相似当且仅当A与B酉相似.

证明：必要性 由于A, B均为n阶正规矩阵且A与B相似，故有相同的特征值，由定理3.4分别存在n阶酉矩阵U_1,U_2，使得

$$U_1^H AU_1 = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \quad U_2^H BU_2 = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

所以 $B = U_2(U_1^H AU_1)U_2^H = (U_2U_1^H)A(U_1U_2^H) = U^H AU$

其中 $U = U_1U_2^H$ 是酉矩阵，即A与B酉相似.

注（1）不能酉对角化的矩阵仍有可能采用其它可逆变换将其对角化，例如

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \quad A^T = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} \quad AA^T \neq A^T A$$

A不是正规矩阵.

但1,3是A的两个互异特征值，从而有两个线性无关的特征向量，所以A可以相似变换对角化。可见，A可以对角化，但不能酉对角化。

（2）实正规矩阵一般不能通过正交相似变换对角化。

（若特征值全为实数，则可正交相似对角化）

如 $A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$，特征值为1+2i，$AA^T = A^T A = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$ 是正规阵，但不可能正交对角化。

3.3矩阵的Jordan标准形

Jordan Canonical Form

问题：

对线性空间中的线性变换T，求一组基$\{\alpha_1, \alpha_2, ..., \alpha_n\}$，使$T$在基$\alpha_1, \alpha_2, ..., \alpha_n$下矩阵$J$的形式尽可能简单.

J称为若当矩阵.

重点：

如何求矩阵的Jordan标准形.

3.3.1 Jordan矩阵

1 —矩阵

定义3.7 元素均为λ多项式的矩阵称为λ-矩阵，记为$A(\lambda)$.

λ-矩阵的运算与数字矩阵相同. λ-矩阵$A(\lambda)$的秩也定义为$A(\lambda)$中非零子式的最高阶数. 如果n阶λ-矩阵$A(\lambda)$，存在n阶λ-矩阵$B(\lambda)$，使

$$A(\lambda)\,B(\lambda)=B(\lambda)\,A(\lambda)=I_n,$$

则称$A(\lambda)$可逆，称$B(\lambda)$为$A(\lambda)$的逆矩阵，记为$B(\lambda)=A^{-1}(\lambda)$. 可以证明$n$阶$\lambda$-矩阵$A(\lambda)$可逆当且仅当$|A(\lambda)|=d\neq 0$，其中$d$是常数。

λ-矩阵的初等变换

λ-矩阵的初等变换有以下三种：

(1)互换λ-矩阵的两行（列）；

(2)以非零常数乘以λ-矩阵某行（列）（这里不能乘以λ的多项式或零，这样有可能改变原来矩阵的秩和属性）；

(3)将λ-矩阵某行（列）乘以λ的多项式$\varphi(\lambda)$加到另一行（列）上。

λ-矩阵的标准形式：

采用初等变换可将λ-矩阵化为如下形式：

$$A(\lambda)\to\begin{pmatrix} d_1(\lambda) & & & & & & 0\\ & d_2(\lambda) & & & & & \\ & & \ddots & & & & \\ & & & d_r(\lambda) & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ 0 & & & & & & 0 \end{pmatrix}$$

其中，多项式$d_i(\lambda)$是首一多项式（首项系数为1，即最高幂次项的系数为1），且

$$d_i(\lambda)\,|\,d_{i+1}(\lambda),\ i=1,2,\cdots,r-1.$$

(1) λ-矩阵的标准形式不随所采用的初等变换而变，故称$d_i(\lambda)$为$A(\lambda)$的不变因子。

(2) 设$D_i(\lambda)$为$A(\lambda)$的所有i阶子行列式的首一最大公因式，称为$A(\lambda)$的i阶行列式因子。可以证明初等变换不改变$A(\lambda)$的各阶行列式因子，若$A(\lambda)$的秩为r，则有

$$d_i(\lambda)=\frac{D_i(\lambda)}{D_{i-1}(\lambda)},\quad i=1,2,\cdots,r.$$

从而

$$D_i(\lambda)=d_1(\lambda)d_2(\lambda)\cdots d_i(\lambda),\quad i=1,2,\cdots,r.$$

(3) 将$A(\lambda)$每个不变因子化为不可约因式的方幂，这些不可约因式的方幂均称为$A(\lambda)$的初等因子，全体初等因子称为初等因子组。

例3.6 求λ-矩阵

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

的标准形.

$$A(\lambda)\to\begin{pmatrix} 1 & & \\ & \lambda & \\ & & \lambda(\lambda+1) \end{pmatrix}$$

思考：12阶矩阵的不变因子为:

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

初等因子有几个?

定义3.8 对于n阶矩阵A，定义λ-矩阵$\lambda I-A$的不变因子、各阶行列式因子以及初等因子分别为A的不变因子、各阶行列式因子以及初等因子.

2. Jordan矩阵

定义3.9 形如

$$J_i(\lambda_i)=\begin{pmatrix}\lambda_i & 1 & & 0\\ & \lambda_i & \ddots & \\ & & \ddots & 1\\ 0 & & & \lambda_i\end{pmatrix}_{n_i\times n_i}$$

的矩阵，称为若当块，由若干个若当块构成的分块对角阵，称为若当矩阵，记为*J*. 即

$$J=\begin{pmatrix}J_1(\lambda_1) & & & \\ & J_2(\lambda_2) & & \\ & & \ddots & \\ & & & J_s(\lambda_s)\end{pmatrix}$$

是若当矩阵.

例如 $\begin{pmatrix}2&1&0\\0&2&1\\0&0&2\end{pmatrix}$ 是一个3阶若当块

$$J=\begin{pmatrix}2&1&0&0\\0&2&0&0\\0&0&-3&1\\0&0&0&-3\end{pmatrix}$$ 是一个若当矩阵，它由两个若当块

$J_1(2)=\begin{pmatrix}2&1\\0&2\end{pmatrix}$ 以及 $J_2(-3)=\begin{pmatrix}-3&1\\0&-3\end{pmatrix}$ 组成，即

$$J=\begin{pmatrix}J_1(2)&0\\0&J_2(3)\end{pmatrix}$$

注：单独一个若当块也是一个若当矩阵，对角矩阵也是特殊的若当矩阵.

3.3.2 Jordan标准形的存在定理

定理3.5 在复数域上，任何*n*阶方阵*A*均相似于如下的若当矩阵*J*，称为矩阵*A*的Jordan标准形, 即存在阶可逆矩阵*P*使得:

$$P^{-1}AP=J=\begin{pmatrix}J_{n_1}(\lambda_1) & & & \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{n_s}(\lambda_s)\end{pmatrix}\qquad(3.7)$$

其中$J_i(\lambda_i)$ 为n_i阶Jordan块，$\sum_{i=1}^{s} n_i=n$. 即存在*n*阶可逆矩阵*P*, 使$P^{-1}AP=J$.

若不计Jordan块排列次序, 则*A*的Jordan标准形是唯一的.

3.3.3 Jordan标准形的求法

方法1：初等因子法:

先计算若当块的初等因子, 设有若当块

$$J_{n_i}(\lambda_i)=\begin{pmatrix}\lambda_i & 1 & & 0\\ & \lambda_i & \ddots & \\ & & \ddots & 1\\ 0 & & & \lambda_i\end{pmatrix}_{n_i\times n_i}\qquad \lambda I-J_{n_i}(\lambda_i)=\begin{pmatrix}\lambda-\lambda_i & -1 & & 0\\ & \lambda-\lambda_i & \ddots & \\ & & \ddots & -1\\ 0 & & & \lambda-\lambda_i\end{pmatrix}$$

有一个n_i-1阶子式 $\begin{vmatrix}-1 & 0 & \cdots & 0\\ \lambda-\lambda_i & -1 & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & 0 & \lambda-\lambda_i & -1\end{vmatrix}=(-1)^{n_i-1}$

所以不变因子为 $d_1(\lambda)=d_2(\lambda)=\cdots=d_{n_i-1}(\lambda)=1,\ d_{n_i}(\lambda)=(\lambda-\lambda_i)^{n_i}$

于是J_{n_i}( )有唯一的初等因子$(\lambda-\lambda_i)^{n_i}$ 。

反之给定*A*一个初等因子$(\lambda-\lambda_i)^{n_i}$，有唯一的若当块$J_{n_i}(\lambda)$与之对应。于是得到求矩阵*A*若当标准形的方法。

(1)求出特征矩阵 *I*–*A*的初等因子，设为

$$(\lambda-\lambda_1)^{n_1},(\lambda-\lambda_2)^{n_2},\cdots,(\lambda-\lambda_s)^{n_s}$$

(2)对于*A*的每个初等因子 $(\lambda-\lambda_i)^{n_i}$ 写出对应的n_i阶Jordan块

$$J_{n_i}(\lambda_i)=\begin{pmatrix}\lambda_i & 1 & & 0\\ & \lambda_i & \ddots & \\ & & \ddots & 1\\ 0 & & & \lambda_i\end{pmatrix}_{n_i\times n_i}$$

(3) 合成Jordan矩阵：$J=\begin{pmatrix}J_{n_1}(\lambda_1) & & & 0\\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & J_{n_s}(\lambda_s)\end{pmatrix}$

例3.7 求矩阵*A*的Jordan标准形, 其中 $A=\begin{pmatrix}-3&3&-2\\-7&6&-3\\1&-1&2\end{pmatrix}$

解 写出特征矩阵,并进行初等变换得到

$$\lambda I-A=\begin{pmatrix}\lambda+3 & -3 & 2\\ 7 & \lambda-6 & 3\\ -1 & 1 & \lambda-2\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&0\\0&1&0\\0&0&(\lambda-1)(\lambda-2)^2\end{pmatrix}$$

故的初等因子为(−1),(−2)^2,于是得到*A*的两个若当块(1),

及 $\begin{pmatrix}2&1\\0&2\end{pmatrix}$，从而*A*的若当标准形为 $J=\begin{pmatrix}1&0&0\\0&2&1\\0&0&2\end{pmatrix}$

或者是 $J=\begin{pmatrix}2&1&0\\0&2&0\\0&0&1\end{pmatrix}$

方法2. Jordon标准形变换矩阵的求法.

设A是n阶复矩阵,由定理3.5存在阶可逆矩阵P, 使$P^{-1}AP=J$.

以下给出求P和J的方法.

(1) 将P按J的结构写成列块的形式 $P=\begin{pmatrix}P_1 & P_2 & \cdots & P_s\end{pmatrix}$

故 $$A\begin{pmatrix}P_1 & P_2 & \cdots & P_s\end{pmatrix}=\begin{pmatrix}P_1 & P_2 & \cdots & P_s\end{pmatrix}\begin{pmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s\end{pmatrix}$$

从而 $AP_i=P_iJ_i \quad (i=1,2,\cdots,s)$

(2) 求解s个矩阵方程 $AP_i=P_iJ_i \quad (i=1,2,\cdots,s)$

(3) 将s个P_i合成变换矩阵 $P=\begin{pmatrix}P_1 & P_2 & \cdots & P_s\end{pmatrix}$.

例3.7 设 $A=\begin{pmatrix}-4 & 2 & 10\\ -4 & 3 & 7\\ -3 & 1 & 7\end{pmatrix}$

求A的若当标准形J，并求可逆矩阵P，使$P^{-1}AP$=J。

解
$$\lambda I-A=\begin{pmatrix}\lambda+4 & -2 & -10\\ 4 & \lambda-3 & -7\\ 3 & -1 & \lambda-7\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & (\lambda-2)^3\end{pmatrix}$$

有唯一的初等因子$(\lambda-2)^3$，从而的若当标准形为

$$J=\begin{pmatrix}2 & 1 & 0\\ 0 & 2 & 1\\ 0 & 0 & 2\end{pmatrix}$$

设 $P=(x_1,x_2,x_3)$ 使 $P^{-1}AP=J$，得到

$$\begin{aligned}Ax_1&=2x_1\\ Ax_2&=x_1+2x_2\\ Ax_3&=x_2+2x_3\end{aligned}\qquad \text{即} \qquad \begin{aligned}(2I-A)x_1&=0\\ (2I-A)x_2&=-x_1\\ (2I-A)x_3&=-x_2\end{aligned}$$

由$(2I-A)X=0$解得基础解系为$\alpha_1=(2,1,1)^T$,

由$(2I-A)X=-\alpha_1$，得解$\alpha_2=(0,-1,0)^T$

再由 $(2I-A)X=-\alpha_2$,得解$\alpha_3=(-1,-3,0)^T$

于是令$P=(\alpha_1,\alpha_2,\alpha_3)=\begin{pmatrix}2 & 0 & -1\\ 1 & -1 & -3\\ 1 & 0 & 0\end{pmatrix}$ 则P可逆且$P^{-1}AP=\begin{pmatrix}2 & 1 & 0\\ 0 & 2 & 1\\ 0 & 0 & 2\end{pmatrix}$

例3.8 求矩阵A的Jordan 标准形。

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

3.3.4 矩阵Jordan标准形的应用

定理3.6 设A是n阶复方阵，则A可以对角化当且仅当的初等因子都是一次的.

例3.9 设A是n阶复方阵，$A^2=A$，证明A可以对角化.

证 设 $(\lambda-\lambda_1)^{n_1},(\lambda-\lambda_2)^{n_2},\cdots,(\lambda-\lambda_n)^{n_s}$

是A的全部初等因子，则的若当标准形为

$$J=\begin{pmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s\end{pmatrix} \quad \text{其中} \quad J_i=\begin{pmatrix}\lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1\\ & & & \lambda_i\end{pmatrix}_{n_i\times n_i}$$

从而存在可逆矩阵P，使得$P^{-1}AP=J$.

于是得到 $A^2=A \Leftrightarrow J_i^2=J_i$, $i=1,2,\dots,s$.

若某个$n_i>1$，则由于

$$J_i^2=\begin{pmatrix} \lambda_i^2 & 2\lambda_i & 1 & & \\ & \lambda_i^2 & 2\lambda_i & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & \lambda_i^2 & 2\lambda_i \\ & & & & \lambda_i^2 \end{pmatrix} \qquad J_i=\begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}_{n_i\times n_i}$$

显然$J_i^2\neq J_i$，从而，$A^2\neq A$与假设矛盾，故A的初等因子均为一次的，由定理3.6知道A可以对角化.

3.4 *Hamilton –Cayley*定理及矩阵的最小多项式

3.4.1 *Hamilton—Cayley*定理

设A是任意n阶方阵，其特征多项式为

$$f(\lambda)=|\lambda I-A|=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$$

则$f(A)=0$。即有如下定理

定理3.7（*Hamilton—Cayley*）设是$f(\lambda)=|\lambda I-A|$的特征多项式，则$f(A)=0$.

证明 设 $f(\lambda)=|\lambda I-A|=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$

A的若当标准形为: 其中

$$J=\begin{pmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_s \end{pmatrix} \qquad J_i=\begin{pmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix}_{n_i\times n_i}$$

且$\sum_{i=1}^{s} n_i=n$，$(\lambda_i I_{n_i}-J_i)^{n_i}=0$，

从而存在可逆矩阵P, 使得$P^{-1}AP=J$. 于是,

$$P^{-1}f(A)P=f(P^{-1}AP)=(J-\lambda_1 I_n)^{n_1}(J-\lambda_2 I_n)^{n_2}\cdots(J-\lambda_s I_n)^{n_s}$$

$$J-\lambda_i I_n=\begin{pmatrix} J_1-\lambda_i I_{n_1} & & & & & 0 \\ & \ddots & & & & \\ & & J_{i-1}-\lambda_i I_{n_{i-1}} & & & \\ & & & U_i & & \\ & & & & \ddots & \\ 0 & & & & & J_s-\lambda_i I_{n_s} \end{pmatrix}$$

其中

$$U_i=\begin{pmatrix} 0 & 1 & & 0 \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}_{n_i\times n_i}$$

则

$$U_i^{n_i}=0, i=1,2,\cdots,s$$

$$(J-\lambda_i I_n)^{n_i}=\begin{pmatrix} (J_1-\lambda_i I_{n_1})^{n_i} & & & & & \\ & \ddots & & & & \\ & & (J_{i-1}-\lambda_i I_{n_{i-1}})^{n_i} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & (J_s-\lambda_i I_{n_s})^{n_i} \end{pmatrix}$$

$$i=1,2,\cdots,s$$

$$P^{-1}f(A)P=f(P^{-1}AP)=(J-\lambda_1 I_n)^{n_1}(J-\lambda_2 I_n)^{n_2}\cdots(J-\lambda_s I_n)^{n_s}$$

从而 $f(A)=0$.

例3.9 设 $A=\begin{pmatrix} 1 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$，计算矩阵多项式

$$g(A)=2A^8-3A^5+A^4+A^2-4I$$

解：A的特征多项式为

$$f(\lambda)=|\lambda I-A|=(\lambda-1)^2(\lambda+1)=\lambda^3-\lambda^2-\lambda+1$$

取 $g(\lambda)=2\lambda^8-3\lambda^5+\lambda^4+\lambda^2-4$

$$g(\lambda)=f(\lambda)(2\lambda^5+2\lambda^4+4\lambda^3+\lambda^2+4\lambda+1)+5\lambda^2-3\lambda-5$$

由*Hamlton-Cayley*定理得，故

$$g(A)=5A^2-3A-5I=\begin{pmatrix} -3 & 10 & 4 \\ 0 & 0 & -3 \\ 0 & -3 & 0 \end{pmatrix}$$

例3.10 设 $A=\begin{pmatrix}1&0&0\\1&0&1\\0&1&0\end{pmatrix}$

证明 $A^n=A^{n-2}+A^2-I \quad (n\ge 3)$.

证明 A的特征多项式为 $f(\lambda)=|\lambda I-A|=(\lambda^2-1)(\lambda-1)$

令
$$g(\lambda)=\lambda^n-\lambda^{n-2}-\lambda^2+1=(\lambda^2-1)(\lambda^{n-2}-1)$$
$$=(\lambda^2-1)(\lambda-1)(\lambda^{n-3}+\lambda^{n-4}+\cdots+\lambda+1)$$

由*Hamlton-Cayley*定理得 $f(A)=0$, 故
$$g(A)=A^n-A^{n-2}-A^2+I=0$$

移项得 $A^n=A^{n-2}+A^2-I$.

3.4.2 最小多项式

定义3.10 设A是n阶矩阵，使$\varphi(A)=0$的多项式$\varphi(\lambda)$称为矩阵A的化零多项式。

由上知道，一个矩阵的化零多项式不是唯一的。

定义3.11 矩阵A的化零多项式中，次数最低且首一的多项式称为A的最小多项式，记为$m_A(\lambda)$,或者$m(\lambda)$。

定理3.8 多项式$\varphi(\lambda)$是矩阵A的化零多项式当且仅当$m(\lambda)|\ \varphi(\lambda)$.

特别地，有$m(\lambda)|f(\lambda)$，其中$f(\lambda)$是A的特征多项式。

推论1 矩阵A的最小多项式是唯一的.

定理3.9 矩阵A的特征值一定是A的最小多项式的根.

结合定理3.8与3.9知道：一个n阶方阵A的特征多项式与最小多项式有相同的根，可能重数不一定相同。

推论：如果矩阵A的特征值都是单根，则特征多项式与最小多项式相同。

定理3.10 设A是n阶矩阵，$D_{n-1}(\lambda)$是A的n-1阶行列式因子，则A的最小多项式为
$$m(\lambda)=\frac{|\lambda I-A|}{D_{n-1}(\lambda)}=\frac{D_n(\lambda)}{D_{n-1}(\lambda)}=d_n(\lambda)$$
这里$d_n(\lambda)$是A的第n个不变因子。

例3.11 求下列矩阵A的特征多项式与最小多项式，其中
$$A=\begin{pmatrix}-2&-1&0\\0&-2&0\\0&0&-2\end{pmatrix}$$

解：A的特征多项式
$$f(\lambda)=|\lambda I-A|=(\lambda+2)^3$$

方法1 利用定义与定理3.9

A的最小多项式可能是 $\lambda+2,(\lambda+2)^2,(\lambda+2)^3$中的某一个，通过计算
$$A+2I=\begin{pmatrix}0&-1&0\\0&0&0\\0&0&0\end{pmatrix}\ne 0 \qquad (A+2I)^2=\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}=0$$
所以A的最小多项式是$(\lambda-2)^2$.

方法2 利用定理3.10
$$\lambda I-A=\begin{pmatrix}\lambda+2&1&0\\0&\lambda+2&0\\0&0&\lambda+2\end{pmatrix}$$
$$\longrightarrow\begin{pmatrix}1&0&0\\0&\lambda+2&0\\0&0&(\lambda+2)^2\end{pmatrix}$$

由于$d_3(\lambda)=(\lambda-2)^2$，所以 A的最小多项式是 $(\lambda-2)^2$.

例3.12 设矩阵A的特征多项式$|\lambda I-A|=(\lambda-2)^2(\lambda-3)^2$

(1）给出A所有可能的最小多项式；

(2) 给出A所有可能的若当标准形。

定理3.11 n阶矩阵A可以对角化的充要条件是A的最小多项式无重根。

证明 由定理3.9，A的最小多项式是A的最后一个不变因子，

因此A的最小多项式无重根当且仅当A的各个不变因子均没有重根，

当且仅当A的初等因子都是一次的，

当且仅当A可以对角化（根据定理3.6）。

例 3.13 设矩阵

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

求A的最小多项式，并判断A能否对角化。

解

$$\lambda I-A=\begin{pmatrix}\lambda&0&0&-4\\-1&\lambda&0&4\\0&-1&\lambda&3\\0&0&-1&\lambda-4\end{pmatrix}\xrightarrow{\text{初等变换}}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&(\lambda-1)(\lambda+1)(\lambda-2)^2\end{pmatrix}$$

于是A的最小多项式为$(\lambda-1)(\lambda+1)(\lambda-2)^2$，

由于A的最小多项式有重根，所以A不能对角化.

3.5 矩阵特征值估计

本节介绍利用矩阵的元素更准确地估计其特征值在复平面上的分布区域.

3.5.1矩阵特征值的圆盘定理：

定义 3.12 设 $A=(a_{ij})\in C^{n\times n}$，记

$$G_i=\{z\,\|z-a_{ii}|\le R_i, z\in C\}$$ ，

其中 $R_i=R_i(A)=\sum_{\substack{j=1\\j\ne i}}^{n}|a_{ij}|$ （$i=1,2,\cdots,n$），

称 G_i 为矩阵 A 在复平面上的第 i 个盖尔(Gerschgorin)圆，称 R_i 为 G_i 的半径（$i=1,2,\cdots,n$）.

定理 3.12（盖尔定理） 设 $A=(a_{ij})\in C^{n\times n}$，则 A 的一切特征值都在它的 n 个盖尔圆的并集之内，即 A 的任一特征值 λ 满足

$$\lambda\in G=\bigcup_{i=1}^{n}G_i.$$

证明 设 λ 为 A 的特征值，其对应的特征向量为 $x\,(x\ne 0)$，即 $Ax=\lambda x$，写成分量形式为

$$\sum_{j=1}^{n}a_{ij}x_j=\lambda x_i,\quad (i=1,2,\cdots,n)$$

或

$$(\lambda-a_{ii})x_i=\sum_{\substack{j=1\\j\ne i}}^{n}a_{ij}x_j.\quad (i=1,2,\cdots,n) \qquad (3.9)$$

设 x_t 为 x 的各分量中模最大的一个，则 $x_t\ne 0$，在(3.9)式中当 $i=t$ 时有

$$(\lambda-a_{tt})x_t=\sum_{\substack{j=1\\j\ne t}}^{n}a_{tj}x_j, \qquad (3.10)$$

(3.10)式的两边除以 x_t 并取模得

$$|\lambda-a_{tt}|\le\sum_{\substack{j=1\\j\ne t}}^{n}|a_{tj}|\cdot\left|\frac{x_j}{x_t}\right|\le\sum_{\substack{j=1\\j\ne t}}^{n}|a_{tj}|=R_t,$$

所以 $\lambda\in G_t$，即 $\lambda\in G=\bigcup G_i$.

例 3.14 估计矩阵 $A=\begin{pmatrix}1 & 0.2 & 0.1 & 0.3\\ 0.1 & 0 & 0.3 & 0.4\\ 0.3 & 0.2 & 3 & 0.5\\ 0.2 & -0.3 & 0.1 & 2i\end{pmatrix}$ 的特征值的分布范围.

解 A 的四个盖尔圆为

G_1 $|z-1|\le 0.6$， G_2 $|z|\le 0.8$，

G_3 $|z-3|\le 1$， G_4 $|z-2i|\le 0.6$.

所以 A 的特征值都落在这四个盖尔圆的并集内.

在例 3.14 中，圆盘 G_1 与 G_2 相交，$G_1\bigcup G_2$ 构成一个连通区域，而 G_3 与 G_4 是孤立的.

一般地，由矩阵的 k 个相交的盖尔圆的并集构成的连通区域称为一个连通部分，并说它是由 k 个盖尔圆组成. 一个孤立的盖尔圆组成一个连通部分.

圆盘定理 3.12 只说明矩阵的特征值均在其全部盖尔圆的并集内，并没有明确指出哪个盖尔圆中有多少个特征值，圆盘定理 3.13 更准确地说明特征值的分布情况.

定理 3.13（圆盘定理 2） 设矩阵 A 的 n 个盖尔圆中有 k 个互相连通且与其余 $n-k$ 个不相交，则这个连通区域中恰有 A 的 k 个特征值(当 A 的主对角线上有相同元素时，则按重复次数计算，有特征值相同时也按重复次数计算).

证明 记 $D=diag(a_{11},a_{22},\cdots,a_{nn})$， $B=A-D$，定义矩阵族

$$A(t)=D+tB,\quad (0\le t\le 1)$$

则

$$A(0)=D,\quad A(1)=A.$$

不失一般性，假设 A 的前 k 个圆盘组成连通区域 G，且与其它 $n-k$ 个圆盘的并集 $\overline{G}$ 不相交. 记

$$G(t)=\bigcup_{i=1}^{k}\{z\,|\,|z-a_{ii}|\le tR_i(A),z\in C\},$$

$$\overline{G}(t)=\bigcup_{i=k+1}^{n}\{z\,|\,|z-a_{ii}|\le tR_i(A),z\in C\}.$$

则

$$G(t)\subset G(1)=G,\quad \overline{G}(t)\subset\overline{G}(1)=\overline{G}.\qquad(0\le t\le 1)$$

因而对一切 $t\in[0,1]$，$G(t)$ 与 $\overline{G}(t)$ 不相交. 特别地，$G(0)$ 恰含有 $A(0)=D$ 的 k 个特征值 $a_{11},a_{22},\cdots,a_{kk}$.

由圆盘定理 1，对一切 $t\in[0,1]$，$A(t)$ 的特征值全含在 $G(t)\bigcup\overline{G}(t)$ 中，但由于 $G(t)$ 与 $\overline{G}(t)$ 不相交，当 t 由 0 增加到 1 时，由特征值是矩阵元素的连续函数知，$A(t)$ 的特征值不能从 $G(t)$ 跳到 $\overline{G}(t)$，或从 $\overline{G}(t)$ 跳到 $G(t)$. 因此，$G(0)$ 恰含 $A(0)=D$ 的 k 个特征值保证了对一切 $t\in[0,1]$，$G(t)$ 必须恰含有 $A(t)$ 的 k 个特征值，因而 $G=G(1)$ 必须恰含有 A 的 k 个特征值.

从定理 3.13 可知，由一个盖尔圆组成的连通部分有且仅有一个特征值，由两个盖尔圆组成的连通部分有且仅有两个特征值，但可能这两个特征值都落在一个圆盘中，而另一个圆盘中没有特征值.

例如 矩阵 $A=\begin{pmatrix}1 & 0.8\\ -0.5 & 0\end{pmatrix}$ 的特征方程为 $\lambda^2-\lambda+0.4=0$，所以 A 的特征值为

$$\lambda_1=\frac{1+\sqrt{0.6}\,i}{2},\quad \lambda_2=\frac{1-\sqrt{0.6}\,i}{2}.$$

A 的两个盖尔圆为

$$|z-1|\le 0.8,\quad |z|\le 0.5.$$

由于

$$|\lambda_1|=|\lambda_2|=\sqrt{0.4}\approx 0.63>0.5.$$

所以这两个特征值都不落在圆盘 $|z|\le 0.5$ 内.

推论 3.6 设 n 阶矩阵 A 的 n 个盖尔圆两两互不相交(都是孤立的)，则 A 相似于对角矩阵.

推论 3.7 设 n 阶实矩阵 A 的 n 个盖尔圆两两互不相交，则 A 的特征值全为实数.

证明 因为 A 为实矩阵，所以 A 的 n 个盖尔圆都关于实轴对称. 又由这 n 个盖尔圆两两互不相交知，A 的 n 个特征值互不相等，且每个盖尔圆内恰含有一个特征值. 因为，如果实矩阵有复特征值，则一定成对出现，且在复平面上关于实轴对称，所以若有一个复特征值在某个盖尔圆内，则与其成共轭的特征值也一定在该盖尔圆内，这与定理 3.13 的结论相矛盾，所以 A 的特征值都是实数.

隔离矩阵的特征值除了上述方法外，还有其它方法.

设 $A=(a_{ij})\in C^{n\times n}$，D 为 n 阶可逆矩阵，则 $B=D^{-1}AD$ 与 A 有相同的特征值，对 B 应用圆盘定理有时可能得到更为准确的特征值的包含区域. 为便于运算，一个方便的选择是取 D 为正对角矩阵，即

$$D=diag(a_1,a_2,\ldots,a_n).\qquad a_i>0\ (i=1,2,\cdots,n)$$

这时，$B=D^{-1}AD$ 与 A 不仅有相同的特征值，而且具有相同的对角线元素.

例 3.17 隔离矩阵 $A=\begin{pmatrix}20 & 5 & 0.8\\ 4 & 10 & 1\\ 1 & 2 & 10i\end{pmatrix}$ 的特征值.

解 A 的三个盖尔圆为

G_1：$|z-20|\le 5.8$，G_2：$|z-10|\le 5$， G_3：$|z-10i|\le 3$.

G_1 与 G_2 相交，G_3 孤立，选取 $P=diag(1,1,0.5)$ ，则

$$B=P^{-1}AP=\begin{pmatrix}20 & 5 & 0.4\\ 4 & 10 & 0.5\\ 2 & 4 & 10i\end{pmatrix}.$$

的三个盖尔圆为

G_1'：$|z-20|\le 5.4$，G_2'：$|z-10|\le 4.5$，G_3'：$|z-10i|\le 6$.

G_1',G_2',G_3' 互不相交，每个圆盘中恰有 A 的一个特征值．注意到 G_3' 中的特征值就是 G_3 中的特征值，所以 A 的三个特征值分别在 G_1',G_2',G_3 中.

矩阵的谱半径

我们已经知道，若 $A=(a_{ij})\in C^{n\times n}$， A 的特征值为 λ_1，λ_2，$\cdots,\lambda_n$，则

$$\rho(A)=\max_{1\le i\le n}|\lambda_i|$$

称为 A 的谱半径.

矩阵的谱半径及其估计在线性方程组的迭代法收敛性问题上，以及在差分方程组的稳定性问题上都有重要应用．利用圆盘定理，可以得到关于谱半径 $\rho(A)$ 的一些实用性估计.

命题 3.1 设 $A=(a_{ij})\in C^{n\times n}$，则 A 的谱半径 $\rho(A)$ 满足：

（1）$\rho(A)\le \max\limits_{1\le i\le n}\sum\limits_{j=1}^{n}|a_{ij}|=\|A\|_\infty$.

（2）$\rho(A)\le \max\limits_{1\le j\le n}\sum\limits_{i=1}^{n}|a_{ij}|=\|A\|_1$.

证明 设 A 的特征值为 λ_1，λ_2，$\cdots,\lambda_n$，由圆盘定理 3.12 得

$$\lambda\in G=\bigcup_{i=1}^{n}G_i,$$

其中 G_i 为 A 的第 i 个盖尔圆．所以 A 的最大模特征值也在 G 中．在 G_i 中距离原点最远的点有模

$$|a_{ii}|+R_i(A)=|a_{ii}|+\sum_{\substack{j=1\\ j\ne i}}^{n}|a_{ij}|=\sum_{j=1}^{n}|a_{ij}|,$$

则 $$|\lambda_i|\le \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|,\qquad (i=1,2,\cdots,n)$$

所以

$$\rho(A)\le \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|=\|A\|_\infty.$$

类似地对 A 的列盖尔圆讨论可得（2）.

推论 3.1 设 $A=(a_{ij})\in C^{n\times n}$，则 A 的谱半径 $\rho(A)$ 满足：

$$\rho(A)\le \min\{\max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|,\ \max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|\},$$

定理3.13 设$A=(a_{jk})\in C^{n\times n}$,则$A$是Hermite矩阵的充分必要条件是对任意$x\in C^n$, x^HAx是实数.

定理3.14 设 A为n 阶Hermite矩阵，则

(1) A的所有特征值全是实数;

(2) A的不同特征值所对应的特征向量是互相正交的。

定理3.15 设 $A\in C^{n\times n}$，则 A是Hermite矩阵的充分必要条件是存在酉矩阵U使得

$$U^HAU=\Lambda=diag(\lambda_1,\lambda_2,\cdots,\lambda_n)\qquad(3.10)$$

其中$\lambda_1,\lambda_2,\cdots,\lambda_n$均为实数。

定理3.16 设 $A\in R^{n\times n}$，则 A是实对称矩阵的充分必要条件是存在正交矩阵Q使得

$$Q^TAQ=\Lambda=diag(\lambda_1,\lambda_2,\cdots,\lambda_n)\qquad(3.11)$$

其中$\lambda_1,\lambda_2,\cdots,\lambda_n$均为实数。

定理3.17 设A是n阶Hermite矩阵，

$$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$$

是A的特征值，则$\forall x \in C^n$,有

$$\lambda_1 \ge \frac{x^H Ax}{x^H x} \ge \lambda_n$$

第4章　矩阵分解

Matrix Factorization and Decomposition

矩阵分解的概述

- 矩阵的分解：
  $A=A_1+A_2+\ldots+A_k$　　矩阵的和
  $A=A_1A_2\ldots A_m$　　矩阵的乘积
- 矩阵分解的原则：
  – 实际应用的需要
  理论上的需要
  计算上的需要
  - 显示原矩阵的某些特性
  - 矩阵化简的方法之一
- 主要技巧：
  - 各种标准形的理论和计算方法
  - 矩阵的分块

本章首先由**Gauss**消去法推导出矩阵的三角分解（或**LU**分解），然后介绍**QR**分解。满秩分解等。这些分解在计算数学中都扮演着十分重要的角色，尤其是**QR**分解所建立的**QR**方法，它对数值代数理论的发展起着关键的作用。我们还介绍在广义逆矩阵理论中所遇到的矩阵的满秩分解，奇异值分解和谱分解，它们与**QR**分解都是求解各类最小二乘法问题和最优化问题的重要数学工具。

4.1 矩阵的三角分解

4.1.1 Gauss消去法的矩阵表述

设有 n 元线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n \end{cases}$$

其矩阵形式为 $Ax=b$

其中 $A=(a_{ij})_{n\times n}, x=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, b=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$

如果矩阵A非奇异，则线性方程组AX=b的解为

$X=A^{-1}b$,但当n充分大，用公式计算的元素是非常困难的。因此寻求（4.1）的直接解法是很有价值的，这就是解线性方程组的Gauss消去法产生的原因，而Gauss消去法又与矩阵分解有着密切关系。为建立矩阵的三角分解理论，我们用矩阵语言来描述Gauss消去法的程序。

设 $A^{(0)}=A$，其元素 $a_{ij}^{(0)}=a_{ij}$ 记A的k阶顺序主子式为A_k

k=1,2, …,n.如果　$\Delta_1=a_{11}^{(0)}\neq 0$

，令　$c_{i1}=\dfrac{a_{i1}^{(0)}}{a_{11}^{(0)}}\quad (i=2,3,\cdots,n)$

。对应于Gauss消元程序，构造消元矩阵　$L_1=\begin{pmatrix} 1 & & & \\ -c_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ -c_{n1} & -c_{n2} & \cdots & 1 \end{pmatrix}$

$$L_1A^{(0)}=\begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix}=A^{(1)} \qquad L_1^{-1}=\begin{pmatrix} 1 & & & \\ c_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ c_{n1} & c_{n2} & \cdots & 1 \end{pmatrix}$$

由于 $|A|=|A^{(0)}|=|L_1A^{(0)}|$

，所以由A 得 的二阶顺序主子式为 $A_2=a_{11}^{(0)}a_{22}^{(1)}$，如果

，$\Delta_2\neq 0$则 $a_{22}^{(1)}\neq 0$令 $c_{i2}=\dfrac{a_{i2}^{(1)}}{a_{22}^{(1)}}$ $(i=3,4,\cdots,n)$，并构造消元矩阵

$$L_2=\begin{pmatrix}1 & & & & \\ 0 & 1 & & & \\ 0 & -c_{32} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & -c_{n2} & -c_{n3} & \cdots & 1\end{pmatrix}$$

则

$$L_2A^{(1)}=\begin{pmatrix}a_{11}^{(0)} & a_{12}^{(0)} & a_{13}^{(0)} & \cdots & a_{1n}^{(0)} \\ & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} \\ & & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} \\ & & \vdots & & \vdots \\ & & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)}\end{pmatrix}=A^{(2)}L_2^{-1}=\begin{pmatrix}1 & & & & \\ 0 & 1 & & & \\ 0 & c_{32} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & c_{n2} & c_{n3} & \cdots & 1\end{pmatrix}$$

同理由A(2)可得A的3阶顺序主子式$\Delta_3=a_{11}^{(0)}a_{22}^{(1)}a_{33}^{(2)}$ 如果Δ3 ≠0，则

$a_{33}^{(2)}\neq 0$ 继续下去，直到第r步，这时Δr ≠0，则 $a_{rr}^{(r-1)}\neq 0$

$$L_rA^{(r-1)}=\begin{pmatrix}a_{11}^{(0)} & & a_{1r}^{(0)} & a_{1r+1}^{(0)} & & a_{1n}^{(0)} \\ & & & & & \\ & & a_{rr}^{(r-1)} & a_{rr+1}^{(r-1)} & & a_{rn}^{(r-1)} \\ & & & a_{r+1r+1}^{(r)} & & a_{r+1n}^{(r)} \\ & & & \vdots & \vdots & \vdots \\ & & & a_{nr+1}^{(r)} & & a_{nn}^{(r)}\end{pmatrix}=A^{(r)}$$

其中

$$L_r=\begin{pmatrix}1 & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -c_{r+1,r} & 1 & & \\ & & \vdots & \vdots & \ddots & \\ & & -c_{nr} & -c_{n,r+1} & \cdots & 1\end{pmatrix}\quad L_r^{-1}=\begin{pmatrix}1 & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & c_{r+1,r} & 1 & & \\ & & \vdots & \vdots & \ddots & \\ & & c_{nr} & c_{n,r+1} & \cdots & 1\end{pmatrix}$$

如果这时Δr=0,即 $a_{r+1,r+1}^{(r)}=0$

则Gauss消元法过程中断；否则可以一直进行下去，

则在第n-1步，$\Delta_{n-1}\neq 0$,就有

$$L_{n-1}A^{(n-2)}=\begin{pmatrix}a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{13}^{(0)} & a_{1n}^{(0)} \\ & a_{22}^{(1)} & \cdots & a_{23}^{(1)} & a_{2n}^{(1)} \\ & & \ddots & \vdots & \vdots \\ & & & a_{n-1,n-1}^{(n-2)} & a_{n-1,n}^{(n-2)} \\ & & & & a_{nn}^{(n-1)}\end{pmatrix}=A^{(n-1)}$$

$$L_{n-1}=\begin{pmatrix}1 & & \\ & \ddots & & \\ & & 1 & \\ & & -c_{n,n-1} & 1\end{pmatrix}\quad L_{n-1}^{-1}=\begin{pmatrix}1 & & \\ & \ddots & & \\ & & 1 & \\ & & c_{n,n-1} & 1\end{pmatrix}\quad c_{n,n-1}=\frac{a_{n,n-1}^{(n-2)}}{a_{n-1,n-1}^{(n-2)}}$$

由上可知，对矩阵A的Gauss程序能够进行到最后一步

的充要条件是 $a_{11}^{(0)},a_{22}^{(1)},\cdots,a_{n-1,n-1}^{(n-2)}$均不为零，也即A的前

n-1个顺序主子式$\Delta_k\neq 0$(k=1,2, …,n-1)　　　　（4.2）

因为Gauss消元法上述过程用到行、列交换，所以附加条件（4.2）是合理的。我们看到当条件（4.2）成立时，有 $L_{n-1}\ldots L_2L_1A=A^{(n-1)}$,也即$A=A^{(0)}=L_1^{-1}L_2^{-1}\ldots L_{n-1}^{-1}A^{(n-1)}$,令 $L=L_1^{-1}L_2^{-1}\ldots L_{n-1}^{-1}$,则

$$L=\begin{pmatrix}1 & & & & \\ c_{21} & 1 & & & \\ \vdots & \vdots & \ddots & & \\ c_{n-1,1} & c_{n-1,2} & \cdots & 1 & \\ c_{n,1} & c_{n,2} & \cdots & c_{n,n-1} & 1\end{pmatrix}$$

L是一个单位下三角矩阵，我们记$U=A^{(n-1)}$,则U是一个上三角矩阵，且A=LU.

从解线性方程组的观点看，由Gauss消去法我们得到一个简单的非奇异矩阵。即单位下三角矩阵 **L**,使**$L^{-1}A=U$**是一个上三角矩阵，令**$y=L^{-1}b$**,则**Ax=b**可化为

Ux=y　　　　(4.3)

它的第**n**个方程只含x_n,第**n-1**个方程只含x_n和x_{n-1}, …,因而可以依次求出$x_n,x_{n-1}, \ldots,x_1$,进而解出（**4.3**）式。

4.1.2矩阵的三角分解

定义4.1 设矩阵A是n阶复矩阵，如果A可以分解成一个下三角矩阵K与一个上三角矩阵U的乘积，即A=KU，称A可以三角分解；
如果A可以分解为一个单位下三角矩阵L与一个上三角矩阵U的乘积，即A=LU,称A可以LU分解；
如果A=LDU，其中L是单位下三角矩阵，U是单位上三角矩阵，D是对角矩阵，称A可以LDU分解。

注：一个矩阵的三角分解一般不唯一。

矩阵能够三角分解的条件：

定理4.1 设矩阵A是n阶矩阵，则A能够唯一LDU分解的充要条件是A的前n-1个顺序主子式$\Delta_k \neq 0$(k=1,2, …,n-1)
其中L是单位下三角矩阵，U是单位上三角矩阵，
$D=\mathrm{diag}(d_1,d_2,\ldots,d_n)$是对角矩阵，而且$d_1=\Delta_1, d_k=\Delta_k/\Delta_{k-1}$,
(k=2,3, …,n).

推论4.1 设A是n阶矩阵，则A能够唯一分解为A=LU的充要条件是A的前n-1个顺序主子式均不为零，即$\Delta_k \neq 0$(k=1,2, …， n-1)

推论4.2 n阶非奇异矩阵A有三角分解A=KU的充要条件是A的前n-1个顺序主子式均不为零，即$\Delta_k \neq 0$(k=1,2, …， n-1)

求矩阵A的LU分解以及LDU分解。其中

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

解：因为$\Delta_1=2$，$\Delta_2=5$，所以A有唯一的LU分解，令

$$L_1=\begin{pmatrix}1 & 0 & 0\\ -\frac{1}{2} & 1 & 0\\ -1 & 0 & 1\end{pmatrix} \quad L_1^{-1}=\begin{pmatrix}1 & 0 & 0\\ \frac{1}{2} & 1 & 0\\ 1 & 0 & 1\end{pmatrix} \quad \text{则} \quad L_1A^{(0)}=\begin{pmatrix}2 & -1 & 3\\ 0 & \frac{5}{2} & -\frac{1}{2}\\ 0 & 5 & -1\end{pmatrix}=A^{(1)}$$

$$L_2=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -2 & 1\end{pmatrix} \quad L_2^{-1}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 2 & 1\end{pmatrix} \quad \text{则} \quad L_2A^{(1)}=\begin{pmatrix}2 & -1 & 3\\ 0 & \frac{5}{2} & -\frac{1}{2}\\ 0 & 0 & 0\end{pmatrix}=A^{(2)}$$

因为

$$L^{-1}=L_2L_1=\begin{pmatrix}1 & & \\ -\frac{1}{2} & 1 & \\ -1 & -2 & 1\end{pmatrix}, \therefore L=L_1^{-1}L_2^{-1}=\begin{pmatrix}1 & & \\ \frac{1}{2} & 1 & \\ 1 & 2 & 1\end{pmatrix}$$

于是 $A=A^{(0)}=L_1^{-1}A^{(1)}=L_1^{-1}L_2^{-1}A^{(2)}=LU$

从而得到矩阵A的LU分解为 $A=LA^{(2)}=\begin{pmatrix}1 & 0 & 0\\ \frac{1}{2} & 1 & 0\\ 1 & 2 & 1\end{pmatrix}\begin{pmatrix}2 & -1 & 3\\ 0 & \frac{5}{2} & -\frac{1}{2}\\ 0 & 0 & 0\end{pmatrix}$

$$\begin{pmatrix}2 & -1 & 3\\ 0 & \frac{5}{2} & -\frac{1}{2}\\ 0 & 0 & 0\end{pmatrix}=\begin{pmatrix}2 & 0 & 0\\ 0 & \frac{5}{2} & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}1 & -\frac{1}{2} & \frac{3}{2}\\ 0 & 1 & -\frac{1}{5}\\ 0 & 0 & 1\end{pmatrix}=DU$$

故A的LDU分解为

$$A=\begin{pmatrix}1 & 0 & 0\\ \frac{1}{2} & 1 & 0\\ 1 & 2 & 1\end{pmatrix}\begin{pmatrix}2 & 0 & 0\\ 0 & \frac{5}{2} & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}1 & -\frac{1}{2} & \frac{3}{2}\\ 0 & 1 & -\frac{1}{5}\\ 0 & 0 & 1\end{pmatrix}$$

由定理4. 1及推论1知，上述的两个分解都是唯一的。

- 三角分解的存在性和惟一性

  定理4.2 （*P.96*）：
  - 矩阵的k 阶主子式：取矩阵的前k行、前k列得到的行列式，k=1，2， … ， n。
  - 定理4.1: $A\in P^{n\times n}$有惟一LDU分解的充要条件是A的顺序主子式A_k非零，k=1，2，…， n-1。

  证明过程给出了LDV分解的一种算法

4.2 矩阵的满秩分解

设A是复数域C上，秩为r的$m\times n$矩阵，记为$A\in C_r^{m\times n}$

4.2.1 矩阵的满秩分解

设A是复数域C上，秩为r的$m\times n$矩阵，记为$A\in C_r^{m\times n}$(r>0),
如果存在秩为r的$m\times r$矩阵F和秩为r的$r\times n$矩阵G，使得
$A=FG$，称其为A的一个满秩分解。

注1. F称为列满秩矩阵，G称为行满秩矩阵；
2. 矩阵的满秩分解一般不唯一。

满秩分解的存在性定理：

定理4.3 任何非零矩阵均存在满秩分解.

证明： 采用构造性证明方法。设A是秩为r的$m\times n$矩阵，故存在m阶初等矩阵$E_1, E_2, \ldots, E_k$, 使得

$$B = E_k \cdots E_2 E_1 A = \begin{pmatrix} G \\ \cdots \\ O \end{pmatrix} \begin{array}{l} r\text{行} \\ \\ m-r\text{行} \end{array}$$

其中G是$r\times n$行满秩矩阵, 记$P= E_kE_{k-1}\ldots E_1$, 则P可逆，并且$A=P^{-1}B$, 把P^{-1}分块为$P^{-1}=(F, S)$, 其中F为$m\times r$矩阵，则$A=FG$, F是列满秩矩阵，G是行满秩矩阵，这就是矩阵A的一个满秩分解。

上述分解方法虽然能够直接求出G，但还需要由P求出P^{-1}，才能够求出F, 而求P^{-1}还是比较麻烦的，下面我们给出初等变换直接求满秩分解的方法。

设A是秩为r的$m\times n$矩阵，则A有r个列向量线性无关，不妨设前r列线性无关，则A的后$n-r$个列向量可以表示为前r列的线性组合，用分块矩阵表示就是

$$A=(F, A_2)=(F, FQ)$$

其中F是A的前r个列向量构成的$m\times r$列满秩矩阵，Q是一个$n-r$阶方阵，于是 $A=F(I_r, Q)=FG$.
$G=(I_r, Q)$是$r\times(n-r)$行满秩矩阵。

由此得到求矩阵A的满秩分解的初等变换方法：

对A初等行变换化为行最简形式 $\begin{pmatrix} G \\ \cdots \\ O \end{pmatrix}$

再去掉全为零的$m-r$个行, 即得矩阵G, 然后再根据G中单位矩阵I_r对应的列, 找出矩阵A中对应的列向量$\alpha_{j1},\alpha_{j2}, \ldots,\alpha_{jr}$,令

$$F=(\alpha_{j1},\alpha_{j2}, \ldots,\alpha_{jr})$$

$A=FG$就是A的一个满秩分解。

例4.3 求矩阵A的一个满秩分解，其中

$$A=\begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & 3 & 3 & -2 \\ 3 & 5 & 2 & -1 \\ -1 & -3 & 6 & -5 \end{pmatrix}$$

解 为了对A进行初等行变换后化为行阶梯形矩阵B,

$$A=\begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & 3 & 3 & -2 \\ 3 & 5 & 2 & -1 \\ -1 & -3 & 6 & -5 \end{pmatrix} \xrightarrow[r_4+r_1]{\substack{r_2-2r_1 \\ r_3-3r_1}} \begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & -1 & 5 & -4 \\ 0 & -1 & 5 & -4 \\ 0 & -1 & 5 & -4 \end{pmatrix}$$

$$\xrightarrow[r_1+2r_2]{\substack{r_3-r_2 \\ r_4-r_2}} \begin{pmatrix} 1 & 0 & 9 & -7 \\ 0 & -1 & 5 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{r_2\times(-1)} \begin{pmatrix} 1 & 0 & 9 & -7 \\ 0 & 1 & -5 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad G=\begin{pmatrix} 1 & 0 & 9 & -7 \\ 0 & 1 & -5 & 4 \end{pmatrix}$$

G的前两列构成单位矩阵，所以A的前两列构成矩阵F，即

$$F=\begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 5 \\ -1 & -3 \end{pmatrix}$$

所以A的一个满秩分解是：

$$A=FG=\begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 5 \\ -1 & -3 \end{pmatrix}\begin{pmatrix} 1 & 0 & 9 & -7 \\ 0 & 1 & -5 & 4 \end{pmatrix}$$

例4.4 设A是$m\times n$阶矩阵，$r(A)=r$, $A=\begin{pmatrix} X & Y \\ Z & W \end{pmatrix}$

其中X是r阶可逆矩阵，$W=ZX^{-1}Y$.
证明A有如下的满秩分解：

$$A=\begin{pmatrix} X \\ Z \end{pmatrix}(I_r, X^{-1}Y) \text{ 或 } A=\begin{pmatrix} I_r \\ ZX^{-1} \end{pmatrix}(X, Y)$$

提示：

$$\begin{pmatrix} X^{-1} & O \\ ZX^{-1} & I_{m-r} \end{pmatrix}\begin{pmatrix} X & Y \\ Z & W \end{pmatrix}=\begin{pmatrix} I_r & X^{-1}Y \\ O & W-ZX^{-1}Y \end{pmatrix}=\begin{pmatrix} I_r & X^{-1}Y \\ O & O \end{pmatrix}$$

由初等变换求满秩分解的方法知

$$A=\begin{pmatrix} X \\ Z \end{pmatrix}(I_r, X^{-1}Y)$$

4.3矩阵的QR分解

利用酉变换得到矩阵A的QR分解，在数值代数中起着重要作用，它为计算特征值的数值方法提供了理论依据，而且是求解线性方程组的一个重要工具.

定义4.5 如果n阶矩阵A可以分解成一个酉（正交）矩阵Q与一个复（实）上三角矩阵R的乘积，即$A=QR$，此式称为矩阵A的一个QR分解.

矩阵A能够QR分解的一个条件：

定理4.6 如果n阶方阵$A=(a_{ij})_{n\times n}$是可逆复（实）矩阵，则存在n阶酉（正交）矩阵Q和复（实）的正线上三角矩阵R，使$A=QR$.

证明 把矩阵A按照列分块$A=(\alpha_1, \alpha_2, \ldots, \alpha_n)$，则向量组$\alpha_1, \alpha_2, \ldots, \alpha_n$线性无关，由Schmidt正交化方法可得到

$$\beta_1=\alpha_1,$$
$$\beta_2=\alpha_2-k_{21}\beta_1$$
$$\cdots\cdots\cdots\cdots$$
$$\beta_n=\alpha_n-k_{n,n-1}\beta_{n-1}-\ldots-k_{n1}\beta_1.$$

其中$k_{ij}=(\alpha_i, \beta_j)/(\beta_j, \beta_j)$, $(j<i)$，即有

$$\alpha_1=\beta_1,$$
$$\alpha_2=k_{21}\beta_1+\beta_2$$
$$\cdots\cdots\cdots\cdots$$
$$\alpha_n=k_{n1}\beta_1+k_{n2}\beta_2+\ldots+k_{n,n-1}\beta_{n-1}+\beta_n$$

于是

$$(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\beta_1,\beta_2,\cdots,\beta_n)\begin{pmatrix}1 & k_{21} & \cdots & k_{n1}\\ 0 & 1 & \ddots & \vdots\\ \vdots & \vdots & \ddots & k_{n\,n-1}\\ 0 & 0 & \cdots & 1\end{pmatrix}$$

再将$\beta_1,\beta_2,\cdots,\beta_n$单位化得$\gamma_i=\dfrac{1}{\|\beta_i\|}\beta_i \quad (i=1,2,\cdots,n)$

$$A=(\gamma_1,\gamma_2,\cdots,\gamma_n)\begin{pmatrix}\|\beta_1\| & & & \\ & \|\beta_2\| & & \\ & & \ddots & \\ & & & \|\beta_n\|\end{pmatrix}\begin{pmatrix}1 & k_{21} & \cdots & k_{n1}\\ & 1 & \ddots & \vdots\\ & & \ddots & k_{n,n-1}\\ & & & 1\end{pmatrix}$$

令 $Q=(\gamma_1,\gamma_2,\cdots,\gamma_n)$，则$Q$是酉（正交）矩阵，而令

$$R=\begin{pmatrix}\|\beta_1\| & & & \\ & \|\beta_2\| & & \\ & & \ddots & \\ & & & \|\beta_n\|\end{pmatrix}\begin{pmatrix}1 & k_{21} & \cdots & k_{n1}\\ & 1 & \ddots & \vdots\\ & & \ddots & k_{n,n-1}\\ & & & 1\end{pmatrix}=\begin{pmatrix}\|\beta_1\| & \dfrac{(\alpha_2,\beta_1)}{\|\beta_1\|} & \cdots & \dfrac{(\alpha_2,\beta_1)}{\|\beta_1\|}\\ & \|\beta_2\| & \ddots & \vdots\\ & & \ddots & \dfrac{(\alpha_n,\beta_{n-1})}{\|\beta_{n-1}\|}\\ & & & \|\beta_n\|\end{pmatrix}$$

因为$\|\beta_i\|>0(i=1,2,\cdots,n)$是正实数，所以$R$是正线上三角矩阵，因此$A$有$QR$分解。

推论1 设$n\times r$矩阵A的秩是r，则存在n阶酉矩阵Q及r阶复正线上三角矩阵R，使得

$$A=Q\begin{pmatrix}R\\ \cdots\\ O\end{pmatrix}$$

推论2 设A是n阶实对称正定矩阵，则存在正线上三角矩阵R，使得 $A=R^TR$.

例4.5 用Schmidt正交化方法求A的QR分解。其中

$$A=\begin{pmatrix}3 & 14 & 9\\ 6 & 43 & 3\\ 6 & 22 & 15\end{pmatrix}$$

解 A的列向量

$$\alpha_1=\begin{pmatrix}3\\6\\6\end{pmatrix},\alpha_2=\begin{pmatrix}14\\43\\22\end{pmatrix},\alpha_3=\begin{pmatrix}9\\3\\15\end{pmatrix}$$

将其正交化得

$$\beta_1=\begin{pmatrix}3\\6\\6\end{pmatrix}\qquad \beta_2=\alpha_2-\frac{16}{3}\beta_1=\begin{pmatrix}-2\\11\\-10\end{pmatrix}$$

$$\beta_3=\alpha_3+\frac{3}{5}\beta_2-\frac{5}{3}\beta_1=\frac{1}{5}\begin{pmatrix}14\\-2\\-5\end{pmatrix}$$

再单位化得

$$\gamma_1=\frac{1}{3}\begin{pmatrix}1\\2\\2\end{pmatrix},\gamma_2=\frac{1}{15}\begin{pmatrix}-2\\11\\-10\end{pmatrix},\gamma_3=\frac{1}{15}\begin{pmatrix}14\\-2\\-5\end{pmatrix}$$

故得正交矩阵$Q=(\gamma_1, \gamma_2, \gamma_3)$ $\quad Q=\dfrac{1}{15}\begin{pmatrix}5 & -2 & 14\\ 10 & 11 & -2\\ 10 & -10 & -5\end{pmatrix}$

以及上三角矩阵

$$R=\begin{pmatrix}9 & 0 & 0\\ 0 & 15 & 0\\ 0 & 0 & 3\end{pmatrix}\begin{pmatrix}1 & 16/3 & 5/3\\ 0 & 1 & -3/5\\ 0 & 0 & 1\end{pmatrix}=\begin{pmatrix}9 & 48 & 15\\ 0 & 15 & -9\\ 0 & 0 & 3\end{pmatrix}$$

所以

$$A=\frac{1}{15}\begin{pmatrix}5 & -2 & 14\\ 10 & 11 & -2\\ 10 & -10 & -5\end{pmatrix}\begin{pmatrix}9 & 48 & 15\\ 0 & 15 & -9\\ 0 & 0 & 3\end{pmatrix}$$

4.4 矩阵的奇异值分解

利用酉变换得到矩阵A的奇异值分解，在最小二乘法和矩阵的广义逆中起着关键作用。

定义4.6 如果$m\times n$矩阵A的秩为r，则n阶Hermite矩阵A^HA半正定，故特征值均非负实数，特征值可以表示成 $\lambda_1\ge\lambda_2\ge\ldots\ge\lambda_r>\lambda_{r+1}=\ldots=\lambda_n=0$，则称

$$\sigma_i=\sqrt{\lambda_i}\quad(i=1,2,\cdots,n)$$

为矩阵A的奇异值。

定理 4.9 设$A\in C_r^{m\times n}$，$\sigma_1,\sigma_2,\ldots,\sigma_r$是$A$的$r$个正奇异值，则存在酉矩阵$U\in C^{m\times m}$ 和$V\in C^{n\times n}$，使得

$$A=U\begin{pmatrix}\Sigma & 0\\ 0 & 0\end{pmatrix}V^H$$

其中，$\Sigma=diag(\sigma_1,\sigma_2,\cdots,\sigma_r)$，

$$\sigma_1\ge\sigma_2\ge\cdots\ge\sigma_r>0.$$

证 由A^HA是正规矩阵，且半正定，

$$V^HA^HAV=\begin{pmatrix}\Sigma^H\Sigma & 0\\ 0 & 0\end{pmatrix}=\begin{pmatrix}\sigma_1^2 &&&&&&\\ &\sigma_2^2&&&&&\\ &&\ddots&&&&\\ &&&\sigma_r^2&&&\\ &&&&0&&\\ &&&&&\ddots&\\ &&&&&&0\end{pmatrix}$$

$V=\begin{pmatrix}V_1 & V_2\end{pmatrix}$，$V_1\in C_r^{n\times r}$，$V_2\in C_{n-r}^{n\times(n-r)}$

$$V^HA^HAV=\begin{pmatrix}V_1^HA^HAV_1 & V_1^HA^HAV_2\\ V_2^HA^HAV_1 & V_2^HA^HAV_2\end{pmatrix}=\begin{pmatrix}\Sigma^2 & 0\\ 0 & 0\end{pmatrix}$$

$$V_1^HA^HAV_1=\Sigma^H\Sigma,\qquad V_2^HA^HAV_2=0$$

$$(AV_2)^HAV_2=0\qquad AV_2=0$$

令$U_1=AV_1\Sigma^{-1}$，则$U_1^HU_1=I_r$，所以U_1的列向量是标准正交的单位向量。设$U_1=(\gamma_1,\gamma_2,\ldots,\gamma_r)$，将其扩充成$C^n$的标准正交基$\gamma_1,\gamma_2,\ldots,\gamma_r,\gamma_{r+1},\ldots,\gamma_m$.

令$U_2=(\gamma_{r+1},\ldots,\gamma_m)$，则$U=(U_1,U_2)$，$U_1^HU_1=I_r$，$U_2^HU_1=0$，

$\Rightarrow$ $U_1^HAV_1=\Sigma\qquad(AV_2)^HAV_2=0$

$AV_2=0$ $\Rightarrow$ $U_1^HAV_2=0\qquad U_2^HAV_2=0$

$$\Rightarrow\quad U^HAV=\begin{pmatrix}U_1^H\\ U_2^H\end{pmatrix}A\begin{pmatrix}V_1 & V_2\end{pmatrix}$$

$$=\begin{pmatrix}U_1^HAV_1 & U_1^HAV_2\\ U_2^HAV_1 & U_2^HAV_2\end{pmatrix}=\begin{pmatrix}\Sigma & 0\\ 0 & 0\end{pmatrix}$$

定理 4.10 设A是n阶非奇异矩阵，则存在n阶酉矩阵U,V使得

$$A=U\begin{pmatrix}\sigma_1 &&& 0\\ &\sigma_2&&\\ &&\ddots&\\ 0&&&\sigma_n\end{pmatrix}V^H$$

$$\sigma_1\ge\sigma_2\ge\cdots\ge\sigma_n>0.$$

若将U,V分别写成$U=(u_1,u_2,\ldots,u_n)$，$V=(v_1,v_2,\ldots,v_n)$，则

$$A=\sum_{i=1}^n\sigma_iu_iv_i^H\ .$$

例4.8 求矩阵 $A=\begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix}$ 的奇异值分解

一、求$A^H A$的特征值及特征向量

$$A^H A=\begin{pmatrix}1&0&1\\0&1&1\\1&1&2\end{pmatrix}$$

$$|\lambda I-A^H A|=(\lambda-3)(\lambda-1)\lambda$$

$\Rightarrow$ $\lambda_1=3,\lambda_2=1,\lambda_3=0;\sigma_1=\sqrt{3},\sigma_2=1$

$(\lambda_i I-A^H A)x=0$ $\Rightarrow$

$$\alpha_1=\begin{pmatrix}1\\1\\2\end{pmatrix}\quad \alpha_2=\begin{pmatrix}1\\-1\\0\end{pmatrix}\quad \alpha_3=\begin{pmatrix}1\\1\\-1\end{pmatrix}$$

它们两两正交，将其单位化得到

$$\gamma_1=\begin{pmatrix}\frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{6}}\\ \frac{2}{\sqrt{6}}\end{pmatrix}\quad \gamma_2=\begin{pmatrix}\frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}}\\ 0\end{pmatrix}\quad \gamma_3=\begin{pmatrix}\frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{3}}\\ -\frac{1}{\sqrt{3}}\end{pmatrix}$$

构造酉矩阵V

$$\boldsymbol{V}=\begin{pmatrix}\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{6}}&-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{3}}\\ \frac{2}{\sqrt{6}}&0&-\frac{1}{\sqrt{3}}\end{pmatrix}$$

$$U_1=AV_1\Sigma^{-1}=\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\\ 0&0\end{pmatrix}$$

取U_1的列向量生成子空间的正交补的基 $\begin{pmatrix}0\\0\\1\end{pmatrix}$

$$U_2=\begin{pmatrix}0\\0\\1\end{pmatrix}$$

$$U=(U_1,U_2)=\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}&0\\ \frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}&0\\ 0&0&1\end{pmatrix}$$

$$A=U\begin{pmatrix}\sqrt{3}&0&0\\0&1&0\\0&0&0\end{pmatrix}V^H$$

$$=\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}&0\\ \frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}&0\\ 0&0&1\end{pmatrix}\begin{pmatrix}\sqrt{3}&0&0\\0&1&0\\0&0&0\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{6}}&\frac{2}{\sqrt{6}}\\ \frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}&0\\ \frac{1}{\sqrt{3}}&\frac{1}{\sqrt{3}}&-\frac{1}{\sqrt{3}}\end{pmatrix}$$

例4.10 设矩阵A的奇异值分解为 $A=U\begin{pmatrix}\Sigma&0\\0&0\end{pmatrix}V^H$
其中

$$\boldsymbol{\Sigma=diag(\sigma_1,\sigma_2,\text{L}\ ,\sigma_r)}.$$

则U的列向量是矩阵AA^H的特征向量，而V的列向量是A^HA的特征向量。

第 5 章

矩阵分析

同数学分析一样，矩阵分析理论的建立，也是以极限理论为重要基础的，其内容丰富，是研究数值分析方法和其它数学分支以及许多工程问题的重要工具。本章首先讨论n维线性空间C^n中的向量范数与矩阵空间$C^{m\times n}$中矩阵范数的理论与性质，接着讨论矩阵函数的相关概念、性质、求法；函数矩阵的微分与积分，最后介绍矩阵函数在微分方程组中的应用.

5.1 向量范数及其性质

5.1.1 向量范数

定义5.1 设V是数域P（实数域或复数域）上的线性空间，如果对任意向量x，按照某个对应法则，对应于一个实数 $\|x\|$，且满足下列三个条件：

(1) 正定性 $\|x\|\geq 0$，且$\|x\|=0$当且仅当$x=0$；

(2) 齐次性 $\|kx\|=|k|\,\|x\|$；

(3) 三角不等式 $\|x+y\|\leq \|x\|+\|y\|$；

则称$\|x\|$为V上向量x的范数，简称为向量范数. 在线性空间V中定义了范数，就称V是线性赋范空间.

向量范数有以下性质

(1) 当 $\|x\|\neq 0, \left\|\frac{1}{\|x\|}\cdot x\right\|=1$　　(2) $\forall x\in V, \|-x\|=\|x\|$

(3) $\big|\|x\|-\|y\|\big|\leq\|x-y\|$　　(4) $\big|\|x\|-\|y\|\big|\leq\|x+y\|$

仅证明(3)

$\|x\|=\|x-y+y\|\leq\|x-y\|+\|y\|\quad\Rightarrow\|x\|-\|y\|\leq\|x-y\|$

类似的, $\|y\|-\|x\|\leq\|y-x\|=\|-(x-y)\|=\|x-y\|$

$\Rightarrow\|x\|-\|y\|\geq-\|x-y\|$

因此： $\big|\|x\|-\|y\|\big|\leq\|x-y\|$

例5.1 设向量 $x=(x_1,x_2,\cdots,x_n)\in C^n$，规定

$$\|x\|=\sum_{i=1}^n|x_i|$$

则$\|x\|$是C^n上的一个范数. 此范数称为向量x的1-范数，记为$\|x\|_1$，即

$$\|x\|_1=\sum_{i=1}^n|x_i|$$

例5.2 设向量 $x=(x_1,x_2,\cdots,x_n)\in C^n$，规定

$$\|x\|=\max_{1\leq i\leq n}|x_i|$$

则$\|x\|$是C^n上的一个范数. 此范数称为向量x的∞-范数，记为$\|x\|_\infty$，即

$$\|x\|_\infty=\max_{1\leq i\leq n}|x_i|$$

例5.3 设向量 $x=(x_1,x_2,\cdots,x_n)\in C^n$，规定

$$\|x\|=\sqrt{\sum_{i=1}^{n}|x_i|^2}$$

则$\|x\|$是C^n上的一个范数. 此范数称为向量x的2-范数，记为$\|x\|_2$，即

$$\|x\|_2=\sqrt{\sum_{i=1}^{n}|x_i|^2}=\sqrt{x^H x}$$

例5.4 设向量 $x=(x_1,x_2,\cdots,x_n)\in C^n$，规定

$$\|x\|=(\sum_{i=1}^{n}|x_i|^p)^{\frac{1}{p}},\quad p\geq 1$$

则$\|x\|$是C^n上的一个范数. 此范数称为向量x的p-范数，记为$\|x\|_p$，即

$$\|x\|_p=(\sum_{i=1}^{n}|x_i|^p)^{\frac{1}{p}}$$

$$\|x\|_p=\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}$$

证明：当p=1时为向量的1-范数, p>1时, 对$x\neq 0$ 有

(1) $\|x\|_p=\left(|x_1|^p+|x_2|^p+\cdots+|x_n|^p\right)^{\frac{1}{p}}>0,$

且$\|x\|$=0当且仅当x=0.

(2) $$\|\alpha x\|_p=\|(\alpha x_1,\alpha x_2,\cdots,\alpha x_n)\|_p$$
$$=\left(|\alpha x_1|^p+|\alpha x_2|^p+\cdots+|\alpha x_n|^p\right)^{\frac{1}{p}}$$
$$=|\alpha|\left(|x_1|^p+|x_2|^p+\cdots+|x_n|^p\right)^{\frac{1}{p}}=|\alpha|\|x\|_p$$

$$\|x\|_p=\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}$$

证明：(3) 对$\forall x=(x_1,x_2,\ldots,x_n),\ y=(y_1,y_2,\ldots,y_n)\in C^n$

$$\|x\|_p=\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}},\qquad \|y\|_p=\left(\sum_{i=1}^{n}|y_i|^p\right)^{\frac{1}{p}},$$

$$\|x+y\|_p=\left(\sum_{i=1}^{n}|x_i+y_i|^p\right)^{\frac{1}{p}},$$

$$\left(\|x+y\|_p\right)^p=\sum_{i=1}^{n}|x_i+y_i|^p=\sum_{i=1}^{n}|x_i+y_i|^{p-1}|x_i+y_i|$$
$$\leq\sum_{i=1}^{n}|x_i+y_i|^{p-1}|x_i|+\sum_{i=1}^{n}|x_i+y_i|^{p-1}|y_i|$$

应用Hölder不等式

$$\sum_{i=1}^{n}|a_i||b_i|\leq\left[\sum_{i=1}^{n}|a_i|^p\right]^{\frac{1}{p}}\left[\sum_{i=1}^{n}|b_i|^q\right]^{\frac{1}{q}},\ p>1,q>1,\frac{1}{p}+\frac{1}{q}=1.$$

$$\sum_{i=1}^{n}|x_i+y_i|^{p-1}|x_i|\leq\left[\sum_{i=1}^{n}|x_i+y_i|^{(p-1)q}\right]^{\frac{1}{q}}\left[\sum_{i=1}^{n}|x_i|^p\right]^{\frac{1}{p}}$$

$$\sum_{i=1}^{n}|x_i+y_i|^{p-1}|y_i|\leq\left[\sum_{i=1}^{n}|x_i+y_i|^{(p-1)q}\right]^{\frac{1}{q}}\left[\sum_{i=1}^{n}|y_i|^p\right]^{\frac{1}{p}}$$

$$\sum_{i=1}^{n}|x_i+y_i|^p\leq\sum_{i=1}^{n}|x_i+y_i|^{p-1}|x_i|+\sum_{i=1}^{n}|x_i+y_i|^{p-1}|y_i|$$

$$\leq\left[\sum_{i=1}^{n}|x_i+y_i|^p\right]^{1-\frac{1}{p}}\left[\sum_{i=1}^{n}|x_i|^p\right]^{\frac{1}{p}}+\left[\sum_{i=1}^{n}|x_i+y_i|^p\right]^{1-\frac{1}{p}}\left[\sum_{i=1}^{n}|y_i|^p\right]^{\frac{1}{p}}$$

$$\sum_{i=1}^{n}|x_i+y_i|^{p-1}|x_i|\leq\left[\sum_{i=1}^{n}|x_i+y_i|^{(p-1)q}\right]^{\frac{1}{q}}\left[\sum_{i=1}^{n}|x_i|^p\right]^{\frac{1}{p}}$$

$$\sum_{i=1}^{n}|x_i+y_i|^{p-1}|y_i|\leq\left[\sum_{i=1}^{n}|x_i+y_i|^{(p-1)q}\right]^{\frac{1}{q}}\left[\sum_{i=1}^{n}|y_i|^p\right]^{\frac{1}{p}}$$

$$\frac{1}{p}+\frac{1}{q}=1\quad\Rightarrow\frac{1}{q}=1-\frac{1}{p},\ (p-1)q=p$$

$$\sum_{i=1}^{n}|x_i+y_i|^p \le \sum_{i=1}^{n}|x_i+y_i|^{p-1}|x_i| + \sum_{i=1}^{n}|x_i+y_i|^{p-1}|y_i|$$

$$\le \left[\sum_{i=1}^{n}|x_i+y_i|^p\right]^{1-\frac{1}{p}}\left[\sum_{i=1}^{n}|x_i|^p\right]^{\frac{1}{p}} + \left[\sum_{i=1}^{n}|x_i+y_i|^p\right]^{1-\frac{1}{p}}\left[\sum_{i=1}^{n}|y_i|^p\right]^{\frac{1}{p}}$$

$$\Rightarrow \left(\sum_{i=1}^{n}|x_i+y_i|^p\right)^{\frac{1}{p}} \le \left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n}|y_i|^p\right)^{\frac{1}{p}}$$

即 $\|x+y\|_p \le \|x\|_p + \|y\|_p$.

因此，$\|x\|_p$是C^n上的一个范数.

在C^n中常用的p-范数有三类

$p=1$时，x的1-范数：$\|x\|_1 = \sum_{i=1}^{n}|x_i|$

$p=2$时，x的2-范数，也称为即欧氏范数

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n}|x_i|^2}$$

$p\to\infty$时，得到x的∞-范数

$$\|x\|_\infty = \max_{1\le i\le n}|x_i|$$

定理5.1 若记 $\lim\limits_{p\to\infty}\|x\|_p = \|x\|_\infty$，则$\|x\|_\infty = \max\limits_{1\le i\le n}|x_i|$.

证明： 令 $\omega = \max\limits_{1\le i\le n}|x_i|$，则

$$\|x\|_p = \left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}} = \left(\sum_{i=1}^{n}\left|\frac{x_i}{\omega}\right|^p\cdot\omega^p\right)^{\frac{1}{p}} = \omega\underline{\left(\sum_{i=1}^{n}\left|\frac{x_i}{\omega}\right|^p\right)^{\frac{1}{p}}}$$

由于$\left|\frac{x_i}{\omega}\right|^p \le 1$且$\sum_{i=1}^{n}\left|\frac{x_i}{\omega}\right|^p \ge \left|\frac{\omega}{\omega}\right|^p = 1$，因此

$$1 \le \sum_{i=1}^{n}\left|\frac{x_i}{\omega}\right|^p \le n \quad \Rightarrow 1 \le \left(\sum_{i=1}^{n}\left|\frac{x_i}{\omega}\right|^p\right)^{\frac{1}{p}} \le n^{\frac{1}{p}} \xrightarrow{p\to\infty} 1$$

例5.5 设A是任意n阶实对称正定矩阵，n维列向量x，则函数

$$\|x\|_A = (x^T A x)^{1/2}$$

是R^n上的一种向量范数，称为加权范数或椭圆范数.

例5.6 设向量 $x=(3i, 2, -5)^T$，求

$$\|x\|_1, \|x\|_2, \|x\|_\infty$$

5.1.2 向量范数的连续性与等价性

定理5.2 $\|x\|_\alpha$ 与 $\|x\|_\beta$是n维线性空间中任意两种向量范数，则一定存在两个与向量无关的正常数c_1, c_2，使得对所有$x\in V$，有不等式

$$c_1\|x\|_\beta \le \|x\|_\alpha \le c_2\|x\|_\beta \qquad (5.1)$$

证明：设$\alpha_1,\alpha_2,\ldots,\alpha_n$是线性空间$V$的一组基, 于是对$\forall x\in V$有 $x=\sum_{i=1}^{n}x_i\alpha_i$.

于是, $\|x\|_\alpha=\left\|\sum_{i=1}^{n}x_i\alpha_i\right\|_\alpha \overset{记为}{=} \varphi(x_1,x_2,\cdots,x_n)$

同理, $\|y\|_\alpha=\left\|\sum_{i=1}^{n}y_i\alpha_i\right\|_\alpha \overset{记为}{=} \varphi(y_1,y_2,\cdots,y_n)$

$$|\varphi(y_1,y_2,\cdots,y_n)-\varphi(x_1,x_2,\cdots,x_n)|=\left|\|y\|_\alpha-\|x\|_\alpha\right|$$
$$\le\|y-x\|_\alpha=\left\|\sum_{i=1}^{n}(y_i-x_i)\alpha_i\right\|_\alpha$$
$$\le|y_1-x_1|\|\alpha_1\|_\alpha+|y_2-x_2|\|\alpha_2\|_\alpha+\cdots+|y_n-x_n|\|\alpha_n\|_\alpha \xrightarrow{y\to x} 0.$$

于是, $\varphi(x_1,x_2,\cdots,x_n)$是坐标$(x_1,x_2,\cdots,x_n)$的连续函数.

下面利用连续函数的性质证明:

$$c_1\|x\|_\beta\le\|x\|_\alpha\le c_2\|x\|_\beta$$

$x=0$ 结论显然. 当 $x\neq0$ 时, $\|x\|_\alpha\neq0$, $\|x\|_\beta\neq0$, 且

$f(x)\overset{记}{=}\dfrac{\|x\|_\alpha}{\|x\|_\beta}$ 也是 $x_1,x_2,\ldots,x_n$的连续函数.

考虑单位超球面: $S=\left\{(x_1,x_2,\cdots,x_n)\,\middle|\,\sum_{i=1}^{n}|x_i|^2=1\right\}$

由于S为有界闭集, 且$f(x)$在S上的点均不为零且连续, 因此f在S上有最大值M和最小值m, 即

$$0<m\le\frac{\|x\|_\alpha}{\|x\|_\beta}\le M$$

定义5.2 $\|x\|_\alpha$ 与 $\|x\|_\beta$ 是n维线性空间中任意两种向量范数, 若存在两个与向量无关的正常数c_1, c_2, 使得对所有$x\in V$, 有不等式

$$c_1\|x\|_\beta\le\|x\|_\alpha\le c_2\|x\|_\beta$$

则称向量范数$\|\cdot\|_\alpha$ 与 $\|\cdot\|_\beta$ 是等价的.

5.2 矩阵范数及其性质

5.2.1 矩阵范数的定义与性质

定义5.3 设$A\in C^{m\times n}$，按照某个对应法则，对应于一个实数$\|A\|$，且满足下列四个条件：

（1）正定性 $\|A\|\ge0$，且$\|A\|=0$当且仅当A=0;

（2）齐次性 $\|kA\|=|k|\,\|A\|$

（3）三角不等式 $\|A+B\|\le\|A\|+\|B\|$

（4）相容性 $\|AB\|\le\|A\|\|B\|$，其中A,B可乘。

则称$\|A\|$为$C^{m\times n}$上矩阵A的范数。

注：类似于定理5.2，所有满足定义5.3的矩阵范数都是等价的.

例5.6 设矩阵$A=(a_{ij})_{m\times n}\in C^{m\times n}$，定义

$$\|A\|_{m_1}=\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|$$

$$\|A\|_{m_\infty}=\max(m,n)\max_{i,j}|a_{ij}|$$

则$\|A\|_{m_1}$, $\|A\|_{m_\infty}$都是$C^{m\times n}$上的矩阵范数. 此范数分别称为矩阵A的m_1-范数及m_∞范数。

定义5.4 对于$C^{m\times n}$上的矩阵范数$\|\cdot\|_M$, C^m与C^n上的同类向量范数$\|x\|_V$，如果

$$\|Ax\|_V\le\|A\|_M\|x\|_V$$

则称矩阵范数$\|\cdot\|_M$与向量范数$\|\cdot\|_V$相容。

例5.7 设矩阵 $A=(a_{ij})_{m\times n}\in C^{m\times n}$，证明函数

$$\|A\|_F=\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}$$

则$\|A\|_F$是$C^{m\times n}$上的一个矩阵范数. 且与向量范数$\|\cdot\|_2$相容。

此范数称为**Frobenius范数**，简称为**F-范数**或者**m_2-范数**。

定理5.3 设矩阵$A=(a_{ij})\in C^{m\times n}$，$P\in C^{m\times m}$，$Q\in C^{n\times n}$，都是酉矩阵，则

$$\|PA\|_F=\|A\|_F=\|AQ\|_F$$

推论5.1 两个酉相似矩阵的F-范数相同。即若$B=Q^HAQ$,其中Q是酉矩阵，则

$$\|B\|_F=\|A\|_F$$

例5.6 设 $\|\cdot\|_M$ 是$C^{n\times n}$矩阵范数，任取C^n中的非零列向量y，则函数

$$\|x\|_V=\|xy^H\|_M\ (\forall x\in C^n)$$

是C^n中的向量范数，且矩阵范数 $\|\cdot\|_M$ 与向量范数 $\|\cdot\|_V$ 相容。

5.2.2 几种常用的矩阵范数

定理5.4 已知C^m和C^n上的向量范数$\|\cdot\|$，设$A\in C^{m\times n}$,则函数

$$\|A\|=\max_{\|x\|=1}\|Ax\|$$

是$C^{m\times n}$上的矩阵范数，且与已知的向量范数相容。

称为由向量范数导出的矩阵范数，简称为向量的**从属范数**或者**算子范数**。

定理5.5 设 $A=(a_{ij})_{m\times n}\in C^{m\times n}$，$x=(x_1,x_2,\cdots,x_n)^T\in C^n$，则从属于向量的三种范数 $\|x\|_1,\|x\|_2,\|x\|_\infty$ 的矩阵范数依次为：

(1) $\|A\|_1=\max_j\sum_{i=1}^{n}|a_{ij}|$ （5.9）

(2) $\|A\|_2=\sqrt{\lambda_1}$ （5.10）

其中λ_1是A^HA的最大特征值。

(3) $\|A\|_\infty=\max_i\sum_{j=1}^{n}|a_{ij}|$ （5.11）

例5.8 设矩阵

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

求 $\|A\|_{m_1},\|A\|_{m_2},\|A\|_{m_\infty}\ \|A\|_1,\|A\|_2,\|A\|_\infty$

定理5.6 $\|x\|_\alpha$是C^n上的向量范数，$A\in C^{n\times n}$

则一定存在与向量范数相容的矩阵范数$\|A\|_M$，

可定义为

$$\|A\|_M = \max_{x\neq 0}\frac{\|Ax\|_\alpha}{\|x\|_\alpha} \qquad (5.12)$$

则称方阵范数$\|A\|_M$是从属于向量范数$\|x\|_\alpha$的导出范数，或称$\|A\|_M$是由向量范数$\|x\|_\alpha$诱导出的矩阵范数. 也称$\|A\|_M$是与$\|x\|_\alpha$相应的算子范数.

5.2.3 范数的应用—矩阵非异性条件

设$C^{n\times n}$，可以根据范数的大小来判断是否为非奇异矩阵.

定理5.7 设 $C^{n\times n}$，且对设上的某种矩阵范数，有

$\|A\|<1$，则$I-A$非奇异，且

$$\|(I-A)^{-1}\| \le \frac{\|I\|}{1-\|A\|}$$

定理5.8 设$A\in C^{n\times n}$，且设对$C^{n\times n}$上的某种矩阵

范数满足$\|A\|<1$，则

$$\|I-(I-A)^{-1}\| \le \frac{\|A\|}{1-\|A\|}.$$

定义3.13 设A是n阶矩阵，它的特征值的全体称为A的谱，记为$\lambda(A)$，并且称

$$\max_{\lambda\in\lambda(A)} |\lambda|$$

为矩阵A的谱半径，记为$\rho(A)$.

定理5.9 对于任意$A\in C^{n\times n}$，总有

$$\rho(A)\le\|A\|.$$

定理5.10 设$A\in C^{n\times n}$是正规矩阵，则$\rho(A)=\|A\|_2$.

定理5.11 设$A\in C^{n\times n}$是n阶非奇异矩阵，则A的谱范数为

$$\|A\|_2 = \sqrt{\rho(A^HA)} = \sqrt{\rho(AA^H)}.$$

5.3 矩阵序列与矩阵级数

5.3.1 向量序列与矩阵序列

5.3.2 矩阵级数

5.3.1.1 向量序列

定义5.6 设 $x^{(k)}=(x_1^{(k)},x_2^{(k)},\cdots,x_n^{(k)})^T$, $k=1,2,\cdots$ 是空间C^n的一个向量序列, 记为$\{x^{(k)}\}$. 如果当$k\to+\infty$时，它的n个分量数列都收敛，即

$$\lim_{k\to\infty} x_i^{(k)}=a_i, \quad i=1,2,\cdots,n$$

则称向量序列 $\{x^{(k)}\}$是按分量收敛的，向量

$$\alpha=(a_1, a_2, \ldots, a_n)^T$$

称为它的极限向量，记为

$$\lim_{k\to\infty} x^{(k)}=\alpha \quad \text{或} \quad x^{(k)}\xrightarrow[k\to\infty]{}\alpha .$$

如果至少有一个分量数列是发散的，则称该向量序列是发散的.

例如 向量序列 $x^{(k)}=\begin{pmatrix}2^{-k}\\ \dfrac{1}{k+1}\\ 1-\dfrac{\sin k}{k+1}\end{pmatrix}, k=0,1,2,\cdots$

是一个收敛的向量序列，且 $\lim_{k\to\infty} x^{(k)}=\begin{pmatrix}0\\0\\1\end{pmatrix}$

而向量序列 $x^{(k)}=\begin{pmatrix}2^{-k}\\ \dfrac{1}{k+1}\\ \sin k\end{pmatrix}, k=0,1,2,\cdots$

是发散的，因为 $\lim_{k\to\infty}\sin k$不存在.

定义5.7 设 $\{x^{(k)}\}$ 是C^n中的向量序列，$\|x\|$是C^n的任意一个向量范数. 如果存在向量$\alpha\in C^n$，使当$k\to+\infty$ 时

$$\|x^{(k)}-\alpha\|\to 0$$

则称向量序列按向量范数收敛于α.

定理5.11 设 $\{x^{(k)}\}$ 是C^n中的向量序列，它按分量收敛当且仅当它按C^n中的任一向量范数收敛.

5.3.1.2 矩阵序列

设$m\times n$矩阵序列$\{A^{(k)}\}$, 其中

$$A^{(k)}=\begin{pmatrix} a_{11}^{(k)} & a_{12}^{(k)} & \cdots & a_{1n}^{(k)}\\ a_{21}^{(k)} & a_{22}^{(k)} & \cdots & a_{2n}^{(k)}\\ \cdots & \cdots & \cdots & \cdots\\ a_{m1}^{(k)} & a_{m2}^{(k)} & \cdots & a_{mn}^{(k)}\end{pmatrix}, k=1,2,\cdots$$

定义 $\lim_{k\to\infty} a_{ij}^{(k)}=a_{ij} \Leftrightarrow \lim_{k\to\infty} A^{(k)}=A=(a_{ij})_{m\times n}$

不收敛的矩阵序列则称为发散的.

类似于定理定理5.11有

定理5.12 设$\{A^{(k)}\}$, $A\in C^{m\times n}$, $k=1, 2, \ldots$

则 $\lim_{k\to\infty} A^{(k)}=A$ 当且仅当 $\lim_{k\to+\infty}\|A^{(k)}-A\|=0$,

其中$\|\cdot\|$是$C^{m\times n}$上的任一矩阵范数.

命题5.1 设 $\lim_{k\to+\infty} A^{(k)}=A$, $\lim_{k\to+\infty} B^{(k)}=B. \alpha,\beta\in C$,则

(1) $\lim_{k\to+\infty}(\alpha A^{(k)}+\beta B^{(k)})=\alpha A+\beta B;$

(2) $\lim_{k\to+\infty} A^{(k)}B^{(k)}=AB;$

(3) 当$A^{(k)}$与A都可逆时，$\lim_{k\to+\infty}(A^{(k)})^{-1}=A^{-1}$.

(4) $PA^{(k)}Q\xrightarrow[k\to\infty]{} PAQ$

定义5.9 设A为n阶方阵, 且当$k\to+\infty$时, 则$A^k\to 0$, 称A为收敛矩阵，否则就称A为发散矩阵.

例5.11 设

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

则

$$A^k=\begin{pmatrix}(\frac{1}{2})^k & 0 & 0\\ 0 & (\frac{1}{3})^k & k(\frac{1}{3})^{k-1}\\ 0 & 0 & (\frac{1}{3})^k\end{pmatrix}$$

可见 $\lim\limits_{k\to\infty} A^k=0$，所以$A$为收敛矩阵.

定理5.13 n阶方阵A为收敛矩阵的充要条件是A的谱半径 $\rho(A)<1$.

定理5.14 设$A\in C^{n\times n}$，若存在$C^{n\times n}$上的一种矩阵范数$\|\cdot\|$，使得$\|A\|<1$，则 $\lim\limits_{k\to+\infty} A^k=0$.

定理5.15 设$\|\cdot\|$是$C^{n\times n}$上的一种矩阵范数，则对任意矩阵 $A\in C^{n\times n}$，有

$$\rho(A)=\lim_{k\to+\infty}\|A^k\|^{\frac{1}{k}}.$$

5.3.2 矩阵级数

定义5.10 设$\{A^{(k)}\}$是$C^{m\times n}$的矩阵序列，称

$$A^{(1)}+A^{(2)}+\cdots+A^{(k)}+\cdots$$

为矩阵级数，记为$\sum\limits_{k=1}^{\infty}A^{(k)}$. 称 $S_n=\sum\limits_{k=1}^{n}A^{(k)}$ 为 $\{A^{(k)}\}$部分和. 如果 $\lim\limits_{n\to\infty}S_n=S$ 存在，则称矩阵级数$\sum\limits_{k=1}^{\infty}A^{(k)}$ 收敛，且极限为S，记为 $\sum\limits_{k=1}^{\infty}A^{(k)}=S$.

不收敛的矩阵级数称为发散的.

如果mn个数项级数

$$\sum_{k=1}^{\infty}a_{ij}^{(k)},\quad i=1,2,\cdots,m;\quad j=1,2,\cdots,n$$

都绝对收敛,则称矩阵级数$\sum\limits_{k=1}^{\infty}A^{(k)}$绝对收敛.

命题5.2 在$C^{n\times n}$中，$\sum\limits_{k=1}^{\infty}A^{(k)}$ 绝对收敛的充要条件是正项级数$\sum\limits_{k=1}^{\infty}\|A^{(k)}\|$收敛.

必要性：

证明：$\sum\limits_{k=1}^{\infty}A^{(k)}$ 绝对收敛 $\Rightarrow$ $\sum\limits_{k=1}^{\infty}|a_{ij}^{(k)}|\le M$ $\Rightarrow$

$$\sum_{k=1}^{\infty}\|A^{(k)}\|_{m_1}=\sum_{k=1}^{\infty}\left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}^{(k)}|\right)\le mnM$$

$\Rightarrow$ $\sum\limits_{k=1}^{\infty}\|A^{(k)}\|_{m_1}$收敛 $\Rightarrow$ $\sum\limits_{k=1}^{\infty}\|A^{(k)}\|$收敛

充分性：$\sum\limits_{k=1}^{\infty}\|A^{(k)}\|$收敛 $\Rightarrow$ $\sum\limits_{k=1}^{\infty}\|A^{(k)}\|_{m_1}$收敛

$\xrightarrow{|a_{ij}^{(k)}|\le\|A^{(k)}\|_{m_1}}$ $\sum\limits_{k=1}^{\infty}a_{ij}^{(k)}$绝对收敛

定理5.16 (Neumann定理) 方阵A的Neumann级数

$$\sum_{k=0}^{\infty}A^k=I+A+A^2+\cdots+A^k+\cdots$$

收敛的充要条件是$\rho(A)<1$，且收敛时,其和为$(I-A)^{-1}$.

证：充分性： $\rho(A)<1$ $\Rightarrow$ $I-A$可逆

$\Rightarrow$ $(I+A+A^2+\ldots+A^k)(I-A)=I-A^{k+1}$

$\Rightarrow$ $I+A+A^2+\ldots+A^k=(I-A)^{-1}-A^{k+1}(I-A)^{-1}$

$\xrightarrow{\rho(A)<1}$ $I+A+A^2+\ldots+A^k\to(I-A)^{-1}$

必要性：$\sum_{k=0}^{\infty} A^k$ 收敛

➡ $\delta_{ij}+(A)_{ij}+(A^2)_{ij}+\ldots+(A^k)_{ij}+\ldots$ 收敛

➡ $(A^k)_{ij}\to 0$

➡ $A^k=((A^k)_{ij})\to 0$ （收敛矩阵）

➡ $\rho(A)<1$

定理5.17 若n阶方阵A的特征值全部落在幂级数

$\varphi(z)=\sum_{k=0}^{\infty} c_k z^k$ 的收敛域内，则矩阵A的幂级数

$\varphi(A)=\sum_{k=0}^{\infty} c_k A^k$ 是绝对收敛的；

反之，若A存在落在$\varphi(z)$的收敛域之外的特征值，

则$\varphi(A)$是发散的.

推论5.3 若幂级数$\varphi(z)=\sum_{k=1}^{\infty} c_k z^k$在整个复平面上收敛, 则对任何的$n$阶方阵$A$, $\varphi(A)$均收敛.

推论5.4 设幂级数 $\varphi(z)=\sum_{k=1}^{\infty} c_k z^k$的收敛半径为$r$, $A\in C^{n\times n}$. 若存在$C^{n\times n}$上的某一矩阵范数$\|\cdot\|$使得 $\|A\|<r$, 则矩阵幂级数 $\varphi(A)=\sum_{k=1}^{\infty} c_k A^k$ 绝对收敛.

例5.12 判断下列矩阵幂级数的敛散性.

$$\sum_{k=0}^{\infty}\frac{k}{6^k}\begin{pmatrix}1 & -8\\ -2 & 1\end{pmatrix}^k$$

解 令 $A=\frac{1}{6}\begin{pmatrix}1 & -8\\ -2 & 1\end{pmatrix}$，则 $\rho(A)=\frac{5}{6}<1$，

而 $\varphi(z)=\sum_{k=1}^{\infty} kz^k$ 的收敛半径为 $r=1$，

所以 $\varphi(A)=\sum_{k=1}^{\infty} c_k A^k=\sum_{k=0}^{\infty}\frac{k}{6^k}\begin{pmatrix}1 & -8\\ -2 & 1\end{pmatrix}^k$ 绝对收敛.

5.4 矩阵函数

矩阵函数的概念与通常的函数概念类似, 所不同的是这里的自变量和因变量都是阶方阵. 本节介绍矩阵函数的定义及计算方法，并给出矩阵函数的应用以及矩阵的微分与积分.

5.4.1、矩阵函数的定义

定义5.11 设幂级数 $\sum_{k=0}^{\infty} c_k z^k$ 收敛半径为R, 且当 $|z|<R$时, 幂级数收敛于$f(z)$, 即

$$f(z)=\sum_{k=0}^{\infty} c_k z^k\ ,\quad |z|<R$$

如果$A\in C^{n\times n}$满足 $\rho(A)<R$, 则称收敛的矩阵幂级数 $\sum_{k=0}^{\infty} c_k A^k$ 的和为矩阵函数, 记为$f(A)$, 即

$$f(A)=\sum_{k=0}^{\infty} c_k A^k.$$

$$f(A)=\sum_{k=0}^{\infty}c_k A^k,$$

把$f(A)$的方阵A换为At, t为参数，则得到

$$f(At)=\sum_{k=0}^{\infty}c_k\left(At\right)^k,$$

常用的矩阵函数：

(1) $e^A=\sum_{k=0}^{\infty}\frac{1}{k!}A^k,\quad A\in C^{n\times n}$

(2) $\sin A=\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}A^{2k+1},\quad A\in C^{n\times n}$

(3) $\cos A=\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}A^{2k},\quad A\in C^{n\times n}$

(4) $(I-A)^{-1}=\sum_{k=0}^{\infty}A^k,\quad \rho(A)<1$

(5) $\ln(I+A)=\sum_{k=0}^{\infty}\frac{(-1)^k}{k+1}A^{k+1},\quad \rho(A)<1$

定理5.18 若n阶矩阵A, B 满足$AB=BA$

则 $e^Ae^B=e^Be^A=e^{A+B}$.

推论1 $e^Ae^{-A}=e^{-A}e^A=e^0=I$

$(e^A)^{-1}=e^{-A},\quad (e^A)^m=e^{mA}$

其中m是整数。

矩阵函数的性质

若$AB=BA$,则

（1） $\cos(A+B)=\cos A\cos B-\sin A\sin B$

（2） $\sin(A+B)=\sin A\cos B+\cos A\sin B$

5.4.2矩阵函数的计算

1. 利用Hamilton-Cayley定理计算矩阵指数函数、正弦函数、余弦函数

例5.13 已知四阶矩阵A的特征值是π、−π、 0、 0, 求$\sin A$, $\cos A$, e^A.

解 A的特征多项式为

$$f(\lambda)=|\lambda I-A|=\lambda^2(\lambda-\pi)(\lambda+\pi)=\lambda^4-\pi^2\lambda^2$$

由Hamilton-Cayley定理

$$f(A)=A^4-\pi^2A^2=0$$

$A^4=\pi^2A^2, A^5=\pi^2A^3, A^6=\pi^2A^4=\pi^4A^2,\ldots, A^{2k}=\pi^{2(k-1)}A^2$

于是

$$\sin A=A+\sum_{k=1}^{\infty}\frac{(-1)^k}{(2k+1)!}A^{2k+1}=A+\sum_{k=1}^{\infty}\frac{(-1)^k}{(2k+1)!}\pi^{2(k-1)}A^3$$

$$=A+\frac{1}{\pi^3}A^3\sum_{k=1}^{\infty}\frac{(-1)^k}{(2k+1)!}\pi^{2k+1}=A+\frac{1}{\pi^3}(\sin\pi-\pi)A^3=A-\frac{1}{\pi^2}A^3$$

$$\cos A=I+\sum_{k=1}^{\infty}\frac{(-1)^k}{(2k)!}A^{2k}=I+A^2\sum_{k=1}^{\infty}\frac{(-1)^k}{(2k)!}\pi^{2(k-1)}$$

$$=I+\frac{1}{\pi^2}(\cos\pi-1)A^2=I-\frac{2}{\pi^2}A^2$$

$$e^A=\sum_{k=0}^{\infty}\frac{1}{k!}A^k=I+A+\sum_{k=1}^{\infty}\frac{1}{(2k)!}A^{2k}+\sum_{k=1}^{\infty}\frac{1}{(2k+1)!}A^{2k+1}$$

$$=I+A+\sum_{k=1}^{\infty}\frac{1}{(2k)!}\pi^{2(k-1)}A^2+\sum_{k=1}^{\infty}\frac{1}{(2k+1)!}\pi^{2(k-1)}A^3$$

$$=I+A+\frac{\cosh\pi-\pi}{\pi^2}A^2+\frac{\sinh\pi-\pi}{\pi^2}A^3$$

例5.11 已知 $A=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$，求 e^{At}.

解：A的特征多项式多项式为 $|\lambda I-A|=\lambda^2+1$

由Hamilton-Cayley定理知：$A^2+I=0$.

从而：$A^2=-I,\ A^3=-A,\ A^4=I,\ A^5=A,\cdots$

即：$A^{2k}=(-1)^kI,\quad A^{2k+1}=(-1)^kA,\ k=1,2,\cdots$

$$e^{At}=\sum_{k=0}^{\infty}\frac{1}{k!}A^kt^k=\left(1-\frac{t^2}{2!}+\frac{t^4}{4!}-\cdots\right)I+\left(t-\frac{t^3}{3!}+\frac{t^5}{5!}-\cdots\right)A$$

$$=(\cos t)I+(\sin t)A=\begin{pmatrix}\cos t&\sin t\\-\sin t&\cos t\end{pmatrix}$$

2. 待定系数法

利用Jordan标准型求解矩阵函数的方法比较复杂，它需要求若当标准形和变换矩阵.现给出根据最小多项式求解矩阵函数的一种方法.

（1）设A 的最小多项式为

$$m(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s},\quad m=\sum_{i=1}^{s}m_i$$

（2）形式上写出待定多项式

$$g(\lambda)=\sum_{i=0}^{m-1}c_i\lambda^i=c_0+c_1\lambda+\cdots+c_{m-1}\lambda^{m-1}$$

（3）求解关于 $c_0,c_1,\cdots,c_{m-1}$ 的线性方程组

$$g^{(k)}(\lambda_i)=f^{(k)}(\lambda_i)(k=0,1,2,\cdots,i=1,2,\cdots,s)$$

（4）求出$g(\lambda)$即可得$g(A)=f(A)$

例5.12 计算矩阵A的函数$\sqrt{A}$

其中
$$A=\begin{pmatrix}1&2&3&4\\0&1&2&3\\0&0&1&2\\0&0&0&1\end{pmatrix}.$$

解：令 $f(\lambda)=\sqrt{\lambda}$，$A$的特征多项式为 $\varphi(\lambda)=(\lambda-1)^4$

即：$g(\lambda)=c_0+c_1\lambda+c_2\lambda^2+c_3\lambda^3$

解得：
$$g(A)=\frac{5}{16}I+\frac{15}{16}A-\frac{5}{16}A^2+\frac{1}{16}A^3$$

得：
$$f(A)=g(A)=\begin{pmatrix}1&1&1&1\\0&1&1&1\\0&0&1&1\\0&0&0&1\end{pmatrix}$$

例5.13 计算矩阵A的函数，求 e^{At}, $\sin A$

其中
$$A=\begin{pmatrix}3&1&-1\\-2&0&2\\-1&-1&3\end{pmatrix}.$$

解：

$$\lambda I-A=\begin{pmatrix}\lambda-3&-1&1\\2&\lambda&-2\\1&1&\lambda-3\end{pmatrix}\to\begin{pmatrix}1&0&0\\0&\lambda-2&0\\0&0&(\lambda-2)^2\end{pmatrix}.$$

故A的最小多项式为 $m(\lambda)=(\lambda-2)^2$

3.利用相似对角化

设A是可对角化的矩阵，则存在n阶可逆矩阵P, 使得

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

$$f(A)=\sum_{k=0}^{\infty}a_kA^k=\sum_{k=0}^{\infty}a_k(P\Lambda P^{-1})^k=P(\sum_{k=0}^{\infty}a_k\Lambda^k)P^{-1}$$

$$=Pdiag(\sum_{k=0}^{\infty}a_k\lambda_1^k,\sum_{k=0}^{\infty}a_k\lambda_2^k,\cdots,\sum_{k=0}^{\infty}a_k\lambda_n^k)P^{-1}$$

$$=Pdiag(f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n))P^{-1}$$

从而

$$f(At)=Pdiag(f(\lambda_1t),f(\lambda_2t),\cdots,f(\lambda_nt))P^{-1}$$

例5.14 已知

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

求 $e^{At}, \cos A$

解 $\det(\lambda I-A)=(\lambda+2)(\lambda-1)^2$ 所以A的特征值为，

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

对应于$\lambda_1=-2$的特征向量为$(-1,1,1)^T$，对应于$\lambda_2=1$线性无关的特征向量$(-2,1,0)^T, (0,0,1)^T$.

故取 $P=\begin{pmatrix}-1 & -2 & 0\\ 1 & 1 & 0\\ 1 & 0 & 1\end{pmatrix}$ 则 $P^{-1}AP=\begin{pmatrix}-2 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$

$$e^{At}=P\begin{pmatrix}e^{-2t} & 0 & 0\\ 0 & e^t & 0\\ 0 & 0 & e^t\end{pmatrix}P^{-1}=\begin{pmatrix}2e^t-e^{-2t} & 2e^t-2e^{-2t} & 0\\ e^{-2t}-e^t & 2e^{-2t}-e^t & 0\\ e^{-2t}-e^t & 2e^{-2t}-2e^t & e^t\end{pmatrix}$$

$$\cos A=P\begin{pmatrix}\cos(-2) & 0 & 0\\ 0 & \cos 1 & 0\\ 0 & 0 & \cos 1\end{pmatrix}P^{-1}$$

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

例5.15 $A=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}$, 求 $\sin A$.

解 1) 化为Jordan标准形

$A \rightarrow J_1=1, J_2=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$

2) 计算 $\sin J_i$

$$\sin J_1=\sin 1, \sin J_2=\begin{pmatrix}\sin 1 & \frac{1}{1!}\cos 1\\ 0 & \sin 1\end{pmatrix}$$

$$\sin A=\begin{pmatrix}\sin 1 & 0 & 0\\ 0 & \sin 1 & \cos 1\\ 0 & 0 & \sin 1\end{pmatrix}$$

5.4.3 函数矩阵的微分与积分

设元素是实变量t的实函数矩阵

$$A(t)=\begin{pmatrix}a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t)\\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t)\\ \cdots & \cdots & \cdots & \cdots\\ a_{m1}(t) & a_{m2}(t) & \cdots & a_{mn}(t)\end{pmatrix}$$

$A(t)$的所有元素$a_{ij}(t)$定义在同一区间$[a,b]$上.

如果$A(t)$的所有元素$a_{ij}(t)$均在区间$[a,b]$上有界、连续、可微、可积, 就分别称$A(t)$在区间$[a,b]$上有界、连续、可微、可积.

1.矩阵函数的极限与连续

定义5.12 若函数矩阵$A(t)$的所有元素$a_{ij}(t)$, 当$t\to t_0$时，有极限a_{ij}，其中a_{ij}是常数，则称函数矩阵$A(t)$当$t\to t_0$时有极限为A，即 $\lim\limits_{t\to t_0}A(t)=A$，其中$A=(a_{ij})$.

如果$A(t)$的所有元素$a_{ij}(t)$，在$t=t_0$处连续,即

$$\lim_{t\to t_0}a_{ij}(t)=a_{ij}(t_0)$$

则称矩阵$A(t)$函数在$t=t_0$处连续，记为

$$\lim_{t\to t_0}A(t)=A(t_0)$$

性质1 设函数矩阵

$$A(t)=(a_{ij}(t))_{n\times n}, B(t)=(b_{ij}(t))_{n\times n}$$

且 $\lim\limits_{t\to t_0} A(t)=A=(a_{ij})_{n\times n}, \lim\limits_{t\to t_0} B(t)=B=(b_{ij})_{n\times n}$

则

(1) $\lim\limits_{t\to t_0}[A(t)\pm B(t)]=A\pm B;$

(2) $\lim\limits_{t\to t_0}[kA(t)]=kA;$

(3) $\lim\limits_{t\to t_0}[A(t)B(t)]=AB.$

例5.16 设函数矩阵

$$A(t)=\begin{pmatrix}\sin t & e^{-2t}\\ \cos t & t\end{pmatrix}, B(t)=\begin{pmatrix}\cos 2t & \sin 3t\\ 3\sin 2t & \sin 3t-\cos 4t\end{pmatrix}$$

求 $\lim\limits_{t\to\frac{\pi}{2}} A(t)B(t)$

解

$$\lim_{t\to\frac{\pi}{2}} A(t)B(t)=\lim_{t\to\frac{\pi}{2}}\begin{pmatrix}\sin t & e^{-2t}\\ \cos t & t\end{pmatrix}\cdot\lim_{t\to\frac{\pi}{2}}\begin{pmatrix}\cos 2t & \sin 3t\\ 3\sin 2t & \sin 3t-\cos 4t\end{pmatrix}$$

$$=\begin{pmatrix}1 & e^{-\pi}\\ 0 & \frac{\pi}{2}\end{pmatrix}\begin{pmatrix}-1 & -1\\ 0 & -1\end{pmatrix}=\begin{pmatrix}-1 & -1-e^{-\pi}\\ 0 & -1\end{pmatrix}$$

2.函数矩阵的微分及其性质

定义5.13 若函数矩阵$A(t)$的所有元素$a_{ij}(t)$均在t_0(或区间(a,b)上)可微，则称此函数矩阵$A(t)$在t_0(或区间(a,b)上)可导，且把各元素$a_{ij}(t)$在t_0处的导数为元素的矩阵 $\left(\frac{da_{ij}(t)}{dt}\right)_{m\times n}$，称为函数矩阵在$t_0$处的导数，记作

$$A'(t_0)=\left(\frac{da_{ij}(t)}{dt}\right)_{m\times n}\Bigg|_{t=t_0}=\begin{pmatrix}a_{11}'(t_0) & a_{12}'(t_0) & \cdots & a_{1n}'(t_0)\\ a_{21}'(t_0) & a_{22}'(t_0) & \cdots & a_{2n}'(t_0)\\ \cdots & \cdots & \cdots & \cdots\\ a_{m1}'(t_0) & a_{m2}'(t_0) & \cdots & a_{mn}'(t_0)\end{pmatrix}$$

并记以元素$a_{ij}(t)$在t_0处的微分为元素的矩阵称为函数矩阵$A(t)$在t_0处的微分，记作

$$\mathrm{d}A(t_0)=(\mathrm{d}a_{ij}(t_0))_{m\times n}.$$

例5.17 求函数矩阵

$$A(t)=\begin{pmatrix}\frac{\sin 2t}{t} & e^{-2t} & t^2+2t\\ \cos t & \ln(1+t) & 3t-1\\ 1 & 0 & 2\sin t\end{pmatrix}$$

$(t>0)$的导数.

可微函数矩阵的性质

(1) $\frac{d}{dt}(A(t)+B(t))=\frac{d}{dt}A(t)+\frac{d}{dt}B(t);$

(2) $\frac{d}{dt}[cA(t)]=c\frac{d}{dt}A(t)$

(3) $\frac{d}{dt}[A(t)B(t)]=[\frac{d}{dt}A(t)]\cdot B(t)+A(t)\cdot[\frac{d}{dt}B(t)]$

(4) $A(t), t=f(u), \frac{d}{du}A(t)=f'(u)\frac{d}{dt}A(t)$

(5) $\frac{d}{dt}A^{-1}(t)=-A^{-1}(t)[\frac{d}{dt}A(t)]A^{-1}(t)$

例5.18 设函数矩阵 $A(t)=\begin{pmatrix}0 & -t\\ t & 1\end{pmatrix}$

求（1）$A'(t)$

（2）$[A^{-1}(t)]'$

3. 函数矩阵的积分

定义5.14 若函数矩阵$A(t)$的所有元素$a_{ij}(t)$均在$[a,b]$上可积，则称此函数矩阵$A(t)$在$[a,b]$上可积，且把各元素在$[a,b]$上的积分为元素的矩阵，称为函数矩阵$A(t)$在$[a,b]$上的积分，记作

$$\int_a^b A(t)dt=\begin{pmatrix}\int_a^b a_{11}(t)dt & \int_a^b a_{12}(t)dt & \cdots & \int_a^b a_{1n}(t)dt\\ \int_a^b a_{21}(t)dt & \int_a^b a_{22}(t)dt & \cdots & \int_a^b a_{2n}(t)dt\\ \cdots & \cdots & \cdots & \cdots\\ \int_a^b a_{m1}(t)dt & \int_a^b a_{m2}(t)dt & \cdots & \int_a^b a_{mn}(t)dt\end{pmatrix}$$

$$=(\int_a^b a_{ij}(t)dt)_{m\times n}$$

性质3 函数矩阵$A(t)$，$B(t)$在区间$[a,b]$上均可积分，则

(1) $\int_a^b [A(t)+B(t)]dt=\int_a^b A(t)dt+\int_a^b B(t)dt$

(2) $\int_a^b cA(t)dt=c\int_a^b A(t)dt$ c是实常数；

(3) $\int_a^b A(t)Bdt=[\int_a^b A(t)dt]\cdot B$

(4) $\int_a^b A\cdot B(t)dt=A\cdot[\int_a^b B(t)dt]$

其中A, B都是与t无关的常数矩阵。

(5) $\int_a^b A(t)B'(t)dt=A(t)B(t)-\int_a^b [A'(t)B(t)]dt$

$$\frac{d}{dt}\int_a^b A(\tau)d\tau=A(t)$$

例5.19 已知函数矩阵

$$A(t)=\begin{pmatrix}\sin\pi t & 3t^2\\ -2t & e^t\end{pmatrix}$$

求 $\int_0^1 A(t)dt.$

例5.20 设函数矩阵

$$A(t)=\begin{pmatrix}\sin 2t & t^2\\ -2\cos t & 0\end{pmatrix}$$

求 $\frac{d}{dt}[\int_0^{t^3} A(\tau)d\tau].$

5.5 常系数线性微分方程

微分方程 的初值问题：

$$\frac{dx(t)}{dt}=ax(t)$$
$$x(o)=b$$

有唯一解： $x(t)=be^{at}.$

常系数线性微分方程组

设a_{ij}均是常数，考虑关于未定函数

$x_1(t),x_2(t),\cdots,x_n(t)$的微分方程组：

$$\begin{cases}\frac{d}{dt}x_1(t)=a_{11}x_1(t)+a_{12}x_2(t)+\cdots+a_{1n}x_n(t)\\ \frac{d}{dt}x_2(t)=a_{21}x_1(t)+a_{22}x_2(t)+\cdots+a_{2n}x_n(t)\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \frac{d}{dt}x_n(t)=a_{n1}x_1(t)+a_{n2}x_2(t)+\cdots+a_{nn}x_n(t)\end{cases}$$

如果记：

$$A=(a_{ij})_{n\times n},X(t)=\begin{pmatrix}x_1(t)\\ x_2(t)\\ \vdots\\ x_n(t)\end{pmatrix},$$

则，这个方程组可以写成矩阵方程的形式：

$$\frac{d}{dt}X(t)=AX(t)$$

定理5.8

假设$A, X(t)$如前，X_0是已知的n维列向量，则初值问题

$$\begin{cases} \dfrac{d}{dt}X(t) = AX(t) \\ X(0) = X_0 \end{cases}$$

有唯一解

$$X(t) = e^{At}X_0$$

定理5.9

假设$A, X(t)$如前，X_0是已知的n维列向量，则初值问题

$$\begin{cases} \dfrac{d}{dt}X(t) = AX(t) + f(t) \\ X(t_0) = X_0 \end{cases}$$

则通解为

$$X = e^{A(t-t_0)}X(t_0) + \int_{t_0}^{t} e^{A(t-\tau)} f(\tau)d\tau$$

例5.21 设,求下列微分方程组 $\dfrac{dX(t)}{dt} = AX(t)$

满足初始条件 $X(0) = (1,1,1)^T$ 的解，其中

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

解 由定理5.8得到满足初始条件的解为$X = e^{At}X(0)$，由例5.14得到

$$e^{At} = \begin{pmatrix} 2e^t - e^{-2t} & 2e^t - 2e^{-2t} & 0 \\ e^{-2t} - e^t & 2e^{-2t} - e^t & 0 \\ e^{-2t} - e^t & 2e^{-2t} - 2e^t & e^t \end{pmatrix}$$

故所求的解为

$$X(t) = \begin{pmatrix} 2e^t - e^{-2t} & 2e^t - 2e^{-2t} & 0 \\ e^{-2t} - e^t & 2e^{-2t} - e^t & 0 \\ e^{-2t} - e^t & 2e^{-2t} - 2e^t & e^t \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3e^{-2t} + 4e^t \\ 3e^{-2t} - 2e^t \\ 3e^{-2t} - 2e^t \end{pmatrix}$$

例5.22设 $A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 3 \end{pmatrix}$ $f(t) = (e^{2t}, e^{2t}, 0)^T$

求微分方程组 $\dfrac{dX}{dt} = AX + f(t)$ 满足初始条件

$X(0) = (-1,1,0)^T$

的解.

解 利用最小多项式可求得

$$e^{At} = e^{2t}\begin{pmatrix} 1 & 0 & 0 \\ t & 1-t & -t \\ t & -t & 1+t \end{pmatrix}$$

$$e^{-A\tau}f(\tau) = e^{-2\tau}\begin{pmatrix} 1 & 0 & 0 \\ -\tau & 1-\tau & -\tau \\ -\tau & \tau & 1-\tau \end{pmatrix}\begin{pmatrix} e^{2\tau} \\ e^{2\tau} \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

故该方程组的解是

$$X(t) = e^{At}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + e^{At}\begin{pmatrix} t \\ t \\ 0 \end{pmatrix} = e^{2t}\begin{pmatrix} 1 & 0 & 0 \\ t & 1-t & -t \\ t & -t & 1+t \end{pmatrix}\begin{pmatrix} t-1 \\ t+1 \\ 0 \end{pmatrix} = \begin{pmatrix} (t-1)e^{2t} \\ (1-t)e^{2t} \\ -2te^{2t} \end{pmatrix}$$