《矩阵分析》期末复习

第1章　矩阵理论

1、 设$\mathbb{F}$是一个域，$\mathbb{F}$上所有$m\times n$矩阵集合记为$\mathbb{F}^{m\times n}$或$M_{m\times n}(\mathbb{F})$；所有$n$维列（行）向量集合记为$\mathbb{F}^n$。

2、 设$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\in M_{m\times n}(\mathbb{F})$。称$n\times m$阶矩阵$\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}\in$ $M_{n\times m}(\mathbb{F})$为矩阵$A$的<u>转置</u>，记为$A^T$或$A'$。$A^H=\begin{pmatrix} \bar{a}_{11} & \bar{a}_{21} & \cdots & \bar{a}_{m1} \\ \bar{a}_{12} & \bar{a}_{22} & \cdots & \bar{a}_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{a}_{1n} & \bar{a}_{2n} & \cdots & \bar{a}_{mn} \end{pmatrix}\in M_{n\times m}(\mathbb{F})$，其中$\bar{a}_{ij}$是$a_{ij}$的共轭，$A^H$称为 A 的<u>共轭转置</u>。

3、 设$A=\left(a_{ij}\right)_{n\times n}\in M_n(\mathbb{F})$，定义$A$的迹$\mathrm{tr}(A)$为$A$的对角线上元素之和，即$\mathrm{tr}(A)=\sum_{i=1}^{n}a_{ii}$。

A的迹$\mathrm{tr}(A)$也等于A的特征值（包括重根）之和，即$\mathrm{tr}(A)=\sum_{i=1}^{n}\lambda_i$。

4、 设$A\in M_{m\times n}(\mathbb{F})$，如果$A$中存在非零$r$级子式，同时所有$r+1$级子式（如果有的话）都等于$0$，则称$A$的<u>秩</u>（rank）为$r$。如果$A=0$，则称$A$的秩为$0$。通常用$\mathrm{rank}\,A$或$\mathrm{rank}(A)$表示矩阵$A$的秩。$A$的行向量组的秩称为$A$的<u>行秩</u>，$A$的列向量组的秩称为为$A$的<u>列秩</u>。

5、 矩阵秩的重要结论：

(1) $\mathrm{rank}\,A=\mathrm{rank}\,A^T=\mathrm{rank}(A^TA)=\mathrm{rank}(AA^T)$

(2) $\mathrm{rank}(A_{m\times n})+\mathrm{rank}(B_{n\times s})-n\le\mathrm{rank}(AB)\le\min\{\mathrm{rank}\,A,\mathrm{rank}\,B\}$

(3) $\mathrm{rank}\,A-\mathrm{rank}\,B\le\mathrm{rank}(A\pm B)\le\mathrm{rank}\,A+\mathrm{rank}\,B$

(4) $\max\{\mathrm{rank}\,A,\mathrm{rank}\,B\}\le\mathrm{rank}(A,B)\le\mathrm{rank}\,A+\mathrm{rank}\,B$

(5) 若$A_{m\times n},B_{n\times s}$且$AB=O$，则$\mathrm{rank}\,A+\mathrm{rank}\,B\le n$

(6) 若P,Q可逆，则$\mathrm{rank}(PA)=\mathrm{rank}(AQ)=\mathrm{rank}(PAQ)=\mathrm{rank}\,A$

(7) 若A列满秩（即$\mathrm{rank}(A_{m\times n})=n$）则$\mathrm{rank}(AB)=\mathrm{rank}(B)$且有左消去律$AB=O\Longrightarrow B=O, AB=AC\Longrightarrow B=C$。（同理，$A$行满秩时有右消去律）

(8) A为n阶满秩矩阵$\Leftrightarrow\mathrm{rank}\,A=n\Leftrightarrow|A|\neq0\Leftrightarrow A$为可逆矩阵

6、 A和B<u>等价</u>（$A\cong B$）$\Longleftrightarrow$存在可逆矩阵P和Q，使得$B=PAQ$

充要条件：矩阵的秩相等

7、 A和B<u>合同</u>（$A\simeq B$）$\Longleftrightarrow$存在可逆矩阵C，使得$B=C^TAC$；

充要条件：正负惯性指数相同；正负特征值个数相等（实对称矩阵）

8、 A和B<u>相似</u>（$A\sim B$）$\Longleftrightarrow$存在可逆矩阵P，使得$B=P^{-1}AP$

充分条件：特征值相同，且都可以相似对角化

必要条件：秩相同：行列式相同；特征值相同；迹相同

【例题】

1、 设$n \geq 2$，$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & 3 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \in M_n(F)$，则$A^n =$ ________.

【方法一】

设e_i是$n \times 1$的列向量，且第i个元素为 1

由$Ae_1 = e_2$，$A^2e_1 = e_3$，$A^3e_1 = e_4$，…，$A^ne_1 = 3e_1$，两边同除e_1得$A^n = 3I_n$

【方法二】

当$n = 2$时，$A = \begin{pmatrix} 0 & 3 \\ 1 & 0 \end{pmatrix}$，$A^2 = \begin{pmatrix} 0 & 3 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 3 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} = 3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 3I_2$

当$n = 2$时，$A = \begin{pmatrix} 0 & 0 & 3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$，$A^2 = \begin{pmatrix} 0 & 0 & 3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 3 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{pmatrix}$

$$A^3 = A^2A = \begin{pmatrix} 0 & 3 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = 3\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 3I_3$$

归纳递推，得$A^n = 3I_n$.

2、 已知矩阵$A = \begin{bmatrix} 2 & -1 & 2 \\ 4 & -2 & 4 \\ 2 & -1 & 2 \end{bmatrix}$，求矩阵$A^{100}$。

由于$A = \begin{bmatrix} 2 & -1 & 2 \\ 4 & -2 & 4 \\ 2 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}[2 \quad -1 \quad 2]$，

故$A^{100} = \begin{bmatrix} 2 & -1 & 2 \\ 4 & -2 & 4 \\ 2 & -1 & 2 \end{bmatrix}^{100} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}[2 \quad -1 \quad 2] \ldots \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}[2 \quad -1 \quad 2]}_{100\text{个}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}[2 \quad -1 \quad 2]} =$

$\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}\underbrace{[2 \quad -1 \quad 2] \ldots \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}}_{99\text{个}[2 \quad -1 \quad 2]\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}}[2 \quad -1 \quad 2] = 2^{99}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}[2 \quad -1 \quad 2] = 2^{99}A$

3、 设$A = \begin{bmatrix} \lambda & 2 & 4 \\ 0 & \lambda & 3 \\ 0 & 0 & \lambda \end{bmatrix}$，求矩阵$A^n$。

令$A = \begin{bmatrix} \lambda & 2 & 4 \\ 0 & \lambda & 3 \\ 0 & 0 & \lambda \end{bmatrix} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} + \begin{bmatrix} 0 & 2 & 4 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} = \lambda E + B$

则$B^2 = \begin{bmatrix} 0 & 0 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$，$B^k = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, k \geq 3$

于是$A^n = (\lambda E + B)^n = \sum_{k=0}^{n} C_n^k (\lambda E)^{n-k}B^k = (\lambda E)^n + C_n^1(\lambda E)^{n-1}B + C_n^2(\lambda E)^{n-2}B^2 =$

$$\lambda^n E+n\lambda^{n-1}B+\frac{n(n-1)}{2}\lambda^{n-2}B^2=\begin{bmatrix}\lambda^n & 0 & 0\\ 0 & \lambda^n & 0\\ 0 & 0 & \lambda^n\end{bmatrix}+n\begin{bmatrix}\lambda^{n-1} & 0 & 0\\ 0 & \lambda^{n-1} & 0\\ 0 & 0 & \lambda^{n-1}\end{bmatrix}\begin{bmatrix}0 & 2 & 4\\ 0 & 0 & 3\\ 0 & 0 & 0\end{bmatrix}+$$

$$\frac{n(n-1)}{2}\begin{bmatrix}\lambda^{n-2} & 0 & 0\\ 0 & \lambda^{n-2} & 0\\ 0 & 0 & \lambda^{n-2}\end{bmatrix}\begin{bmatrix}0 & 0 & 6\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}=\begin{bmatrix}\lambda^n & 0 & 0\\ 0 & \lambda^n & 0\\ 0 & 0 & \lambda^n\end{bmatrix}+n\begin{bmatrix}0 & 2\lambda^{n-1} & 4\lambda^{n-1}\\ 0 & 0 & 3\lambda^{n-1}\\ 0 & 0 & 0\end{bmatrix}+$$

$$\frac{n(n-1)}{2}\begin{bmatrix}0 & 0 & 6\lambda^{n-2}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}=\begin{bmatrix}\lambda^n & 2n\lambda^{n-1} & 4n\lambda^{n-1}+3n(n-1)\lambda^{n-2}\\ 0 & \lambda^n & 3n\lambda^{n-1}\\ 0 & 0 & \lambda^n\end{bmatrix}$$

4、 设A是 3 阶非零方阵，$B=\begin{pmatrix}1 & 1 & 3\\ 1 & 2 & 4\\ 0 & 1 & 1\end{pmatrix}$且$AB=0$，则$\mathrm{rank}(A)=$ ________.

因为A是 3 阶非零方阵，所以$\mathrm{rank}(A)\geq 1$

因为$B=\begin{pmatrix}1 & 1 & 3\\ 1 & 2 & 4\\ 0 & 1 & 1\end{pmatrix}\to\begin{pmatrix}1 & 1 & 3\\ 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}$，所以$\mathrm{rank}(B)=2$

因为$AB=0$，所以$\mathrm{rank}(A)+\mathrm{rank}(B)\leq 3$，即$\mathrm{rank}(A)\leq 1$

综上，$\mathrm{rank}(A)=1$.

5、 设$A\in M_{n\times m}(\mathbb{F})$，$B\in M_{m\times n}(\mathbb{F})$.

(1) 若$\mathrm{rank}(AB)=n$，证明$\mathrm{rank}(BA)=n$.

因为$A\in M_{n\times m}(\mathbb{F})$，$B\in M_{m\times n}(\mathbb{F})$，$\mathrm{rank}(AB)=n\leq\min\{\mathrm{rank}(A),\mathrm{rank}(B)\}$

所以$\mathrm{rank}(A)\geq n$，$\mathrm{rank}(B)\geq n$，$n\leq m$

因为$\mathrm{rank}(A)\leq n$，$\mathrm{rank}(B)\leq n$，所以$\mathrm{rank}(A)=n$，$\mathrm{rank}(B)=n$，即B列满秩

所以$\mathrm{rank}(BA)=\mathrm{rank}(A)=n$.

(2) 若$\mathrm{rank}(BA)=n$，问$\mathrm{rank}(AB)=n$是否成立，并说明理由.

不成立。例如，取$A=(1\quad -1)$和$B=\begin{pmatrix}1\\ 1\end{pmatrix}$。此时$BA=\begin{pmatrix}1 & -1\\ 1 & -1\end{pmatrix}$的秩为 1，但$AB=0$的秩为 0。

6、 设$A\in M_n(F)$，证明：$A^3=A$当且仅当$\mathrm{rank}(A)+\mathrm{rank}(I_n-A)+\mathrm{rank}(I_n+A)=2n$.

(1) 必要性证明（$A^3=A\to\mathrm{rank}(A)+\mathrm{rank}(I_n-A)+\mathrm{rank}(I_n+A)=2n$）：

因为$A^3=A$，所以A的特征值只可能是 0，-1，1

A在某个基下可以对角化为分块对角形式：$A\sim\begin{pmatrix}I_r & 0 & 0\\ 0 & -I_s & 0\\ 0 & 0 & 0_t\end{pmatrix}$

其中，I_r是$r\times r$的单位矩阵，$-I_s$是$s\times s$的负单位矩阵，0_t是$t\times t$的零矩阵，并且$r+s+t=n$

因此，$\mathrm{rank}(A)=r+s$，$\mathrm{rank}(I_n-A)=t+s$，$\mathrm{rank}(I_n+A)=t+r$

所以，$\mathrm{rank}(A)+\mathrm{rank}(I_n-A)+\mathrm{rank}(I_n+A)=r+s+t+s+t+r=2(r+s+t)=2n$

(2) 充分性证明（$\mathrm{rank}(A)+\mathrm{rank}(I_n-A)+\mathrm{rank}(I_n+A)=2n\to A^3=A$）：

若A满足$\mathrm{rank}(A)+\mathrm{rank}(I_n-A)+\mathrm{rank}(I_n+A)=2n$

设$\mathrm{rank}(A)=r+s$，$\mathrm{rank}(I_n-A)=t+s$，$\mathrm{rank}(I_n+A)=t+r$

则有：$r+s+t+s+t+r=2(r+s+t)=2n$

这表明矩阵A的特征值只有0，-1，1，所以$A^3 = A$

第2章　行列式的定义及性质

1、 设$\mathbb{F}$是一个域。对$n \in \mathbb{N}^*$，我们定义映射det：$M_n(\mathbb{F}) \to \mathbb{F}$如下：

$$A = (a_{ij}) \to \det(A) = \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} a_{2j_2} \dots a_{nj_n}$$

其中$j_1 j_2 \dots j_n$跑遍$1,2,\dots,n$的所有排列。映射det称为行列式映射；det(A)称为n阶矩阵A的行列式，也称为一个n阶行列式，det(A)也记为$|A|$，或$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n1} & \cdots & a_{nn} \end{vmatrix}$。

2、 设A, B都是n阶方阵，则$|AB| = |A| \cdot |B|$。

3、 设A, B分别为m与n阶行列式，则$\begin{vmatrix} A & C \\ 0 & B \end{vmatrix} = \begin{vmatrix} A & 0 \\ C & B \end{vmatrix} = |A||B|$，$\begin{vmatrix} C & A \\ B & 0 \end{vmatrix} = \begin{vmatrix} 0 & A \\ B & C \end{vmatrix} = (-1)^{mn}|A||B|$。

4、 设$A = (a_{ij})_{n\times n} \in M_n(\mathbb{F})$。对任意$1 \le i$，$j \le n$，在矩阵A中划去第i行、第j列，剩余$n-1$行$n-1$列仍按原来次序构成的$n-1$阶方阵的行列式，记为$M_{ij}$，称为$a_{ij}$的余子式，即

$$M_{ij} = \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix}$$，$A_{ij} = (-1)^{i+j} M_{ij}$称为$a_{ij}$的代数余子式。

5、 对任意$1 \le i \le n$，$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n1} & \cdots & a_{nn} \end{vmatrix} = \sum_{j=1}^{n} a_{ij} A_{ij}$

对任意$1 \le j \le n$，$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n1} & \cdots & a_{nn} \end{vmatrix} = \sum_{i=1}^{n} a_{ij} A_{ij}$

6、 设$A = (a_{ij})_{n\times n}$是n阶方阵，称n阶方阵$A^* = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$为A的伴随矩阵，其中$A_{ij}$是$a_{ij}$的代数余子式。

7、 设$A \in M_n(\mathbb{F})$，则

(1) $AA^* = A^*A = |A|I_n$

(2) A可逆的充分必要条件是$|A| \neq 0$

(3) A可逆的充分必要条件是存在$B \in M_n(\mathbb{F})$使得$AB = I_n$或$BA = I_n$成立，而且此时$A^{-1} = B$。

【例题】

1、 设$A, B \in M_n(F)$，$|A| = 2$，A^*是A的伴随矩阵且$|B| = 3$，则$|5A^*B^{-1}| =$ ________.

$$|5A^*B^{-1}| = 5^n\left||A|A^{-1}B^{-1}\right| = \frac{5^n|A|^n}{|A||B|} = \frac{5^n \cdot 2^{n-1}}{3}$$

第3章　向量组的线性相关性

1、 如果存在一组不全为零的数$k_1, k_2, \dots, k_s \in F$使得$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = 0$，则称向量组$\alpha_1, \alpha_2, \dots, \alpha_s$<u>线性相关</u>，否则，称向量组$\alpha_1, \alpha_2, \dots, \alpha_s$<u>线性无关</u>。

【例题】

1、 设α_1，α_2，α_3线性无关，$\beta_1 = \alpha_1 + 4\alpha_2 + t\alpha_3$，$\beta_2 = 2\alpha_1 + \alpha_2 - \alpha_3$，$\beta_3 = \alpha_1 - \alpha_3$，则$\beta_1$，$\beta_2$，$\beta_3$线性相关的充要条件是$t =$ ________.

$$(\beta_1, \beta_2, \beta_3) = (\alpha_1, \alpha_2, \alpha_3)\begin{pmatrix}1 & 2 & 1\\ 4 & 1 & 0\\ t & -1 & -1\end{pmatrix}$$

记$A = (\alpha_1, \alpha_2, \alpha_3)$，$B = (\beta_1, \beta_2, \beta_3)$，$M = \begin{pmatrix}1 & 2 & 1\\ 4 & 1 & 0\\ t & -1 & -1\end{pmatrix}$，则$B = AM$

因为α_1，α_2，α_3线性无关，所以$\mathrm{rank}(A) = \mathrm{rank}(\alpha_1, \alpha_2, \alpha_3) = 3$
因为β_1，β_2，β_3线性相关，所以$\mathrm{rank}(B) = \mathrm{rank}(\beta_1, \beta_2, \beta_3) < 3$
因为$\mathrm{rank}(B) = \mathrm{rank}(AM) \le \min\{\mathrm{rank}(A), \mathrm{rank}(M)\}$，所以$\mathrm{rank}(M) < 3$

所以$|M| = \begin{vmatrix}1 & 2 & 1\\ 4 & 1 & 0\\ t & -1 & -1\end{vmatrix} = \begin{vmatrix}1 & 2 & 1\\ 4 & 1 & 0\\ t+1 & 1 & 0\end{vmatrix} = (-1)^{1+3}\begin{vmatrix}4 & 1\\ t+1 & 1\end{vmatrix} = 3 - t = 0$，即$t = 3$.

2、 给定向量组$\alpha_1 = (3,1,2,0)$，$\alpha_2 = (0,-1,-2,3)$，$\alpha_3 = (6,2,4,0)$，$\alpha_4 = (-1,1,-2,0)$，$\alpha_5 = (1,0,0,1)$.

(1) 求α_1，α_2，α_3，α_4，α_5的秩.

$$(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) = \begin{pmatrix}3 & 0 & 6 & -1 & 1\\ 1 & -1 & 2 & 1 & 0\\ 2 & -2 & 4 & -2 & 0\\ 0 & 3 & 0 & 0 & 1\end{pmatrix} \to \begin{pmatrix}0 & 3 & 0 & -4 & 1\\ 1 & -1 & 2 & 1 & 0\\ 0 & 0 & 0 & -4 & 0\\ 0 & 3 & 0 & 0 & 1\end{pmatrix} \to \begin{pmatrix}1 & -1 & 2 & 1 & 0\\ 0 & 3 & 0 & 0 & 1\\ 0 & 0 & 0 & -4 & 0\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}$$

故$\mathrm{rank}(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) = 3$.

(2) 求α_1，α_2，α_3，α_4，α_5的一个极大线性无关组，并将α_1，α_2，α_3，α_4，α_5中其余向量表为该极大线性无关组中向量的线性组合.

$$(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) \to \begin{pmatrix}1 & -1 & 2 & 1 & 0\\ 0 & 3 & 0 & 0 & 1\\ 0 & 0 & 0 & -4 & 0\\ 0 & 0 & 0 & 0 & 0\end{pmatrix} \to \begin{pmatrix}1 & 0 & 2 & 0 & \frac{1}{3}\\ 0 & 1 & 0 & 0 & \frac{1}{3}\\ 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}$$

故α_1，α_2，α_3，α_4，α_5的一个极大线性无关组是α_1，α_2，α_4

此时：$\alpha_3 = 2\alpha_1$，$\alpha_5 = \frac{1}{3}\alpha_1 + \frac{1}{3}\alpha_2$.

第4章　线性方程组

1、 线性方程组解的存在性定理

设$A \in M_{m\times n}(\mathbb{F})$，$\beta \in \mathbb{F}^m$，则

(1) $AX=\beta$有解当且仅当 $\mathrm{rank}\, A=\mathrm{rank}(A,\beta)$

$AX=\beta$有唯一解当且仅当$\mathrm{rank}\, A=\mathrm{rank}(A,\beta)=n$

$AX=\beta$有无穷多解当且仅当$\mathrm{rank}\, A=\mathrm{rank}(A,\beta)<n$

(2) 齐次线性方程组$AX=0$总有零解，而且

$AX=0$只有零解当且仅当$\mathrm{rank}\, A=n$

$AX=0$有非零解当且仅当$\mathrm{rank}\, A<n$

特别地，当$A\in M_n(\mathbb{F})$时，$AX=0$只有零解当且仅当$|A|\neq 0$，$AX=0$有非零解当且仅当$|A|=0$。

【例题】

1、 设$A=\begin{pmatrix}-2&1&1\\1&-2&1\\1&1&-2\end{pmatrix}$且$B=\begin{pmatrix}1&3\\2&a\\a&0\end{pmatrix}$，则矩阵方程$AX=B$有解当且仅当$a=$________.

$$(A|B)=\left(\begin{array}{ccc|cc}-2&1&1&1&3\\1&-2&1&2&a\\1&1&-2&a&0\end{array}\right)\to\left(\begin{array}{ccc|cc}1&-\frac{1}{2}&-\frac{1}{2}&-\frac{1}{2}&-\frac{3}{2}\\0&-\frac{3}{2}&\frac{3}{2}&\frac{5}{2}&a+\frac{3}{2}\\0&\frac{3}{2}&-\frac{3}{2}&a+\frac{1}{2}&\frac{3}{2}\end{array}\right)\to\left(\begin{array}{ccc|cc}1&-\frac{1}{2}&-\frac{1}{2}&-\frac{1}{2}&-\frac{3}{2}\\0&\frac{3}{2}&-\frac{3}{2}&a+\frac{1}{2}&\frac{3}{2}\\0&0&0&a+3&a+3\end{array}\right)$$

因为矩阵方程$AX=B$有解，所以$\mathrm{rank}(A)=\mathrm{rank}(A|B)$，即$a+3=0$，即$a=-3$.

2、 讨论$a\in F$为何值时线性方程组

$$\begin{cases}x_1+x_2+ax_3=a^2\\x_1+ax_2+x_3=a\\(a+1)x_1+(a+1)x_2+2x_3=a+1\end{cases}$$

(1) 无解并说明理由.

记线性方程组的增广矩阵为

$$(A,B)=\begin{pmatrix}1&1&a&a^2\\1&a&1&a\\a+1&a+1&2&a+1\end{pmatrix}\to\begin{pmatrix}1&1&a&a^2\\0&a-1&1-a&a(1-a)\\0&0&(a+2)(1-a)&(a+1)^2(1-a)\end{pmatrix}$$

因为当$a=-2$时，$\mathrm{rank}(A)=2<3=\mathrm{rank}(A,B)$，故此时线性方程组无解.

(2) 有唯一解并求其解.

当$a\neq 1$ 且 $a\neq -2$时，$\mathrm{rank}(A)=\mathrm{rank}(A,B)=3$，故此时线性方程组有唯一解.

$$\text{此时：}(A,B)\to\begin{pmatrix}1&1&a&a^2\\0&1&-1&-a\\0&0&1&\frac{(a+1)^2}{(a+2)}\end{pmatrix}\to\begin{pmatrix}1&0&0&\frac{-a-1}{(a+2)}\\0&1&0&\frac{1}{(a+2)}\\0&0&1&\frac{(a+1)^2}{(a+2)}\end{pmatrix}$$

故唯一解为$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\left(\frac{-a-1}{(a+2)},\frac{1}{(a+2)},\frac{(a+1)^2}{(a+2)}\right)^T$.

(3) 有无穷多解并求其通解.

当$a=1$时，$\mathrm{rank}(A)=\mathrm{rank}(A,B)=1$，故此时线性方程组有无穷多解.

此时：$(A,B)\to\begin{pmatrix}1&1&1&1\\0&0&0&0\\0&0&0&0\end{pmatrix}$

故通解为$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=k_1\begin{pmatrix}-1\\1\\0\end{pmatrix}+k_2\begin{pmatrix}-1\\0\\1\end{pmatrix}+\begin{pmatrix}1\\0\\0\end{pmatrix},k_1,k_2\in F.$

3、 设$A\in M_{m\times n}(\mathbb{R})$，$\beta\in\mathbb{R}^m$，证明线性方程组$A'AX=A'\beta$总有解。
因为$\operatorname{rank}A=\operatorname{rank}A'A\le\operatorname{rank}(A'A,A'\beta)=\operatorname{rank}A'(A,\beta)\le\operatorname{rank}A'=\operatorname{rank}A$，
所以$\operatorname{rank}A'A=\operatorname{rank}(A'A,A'\beta)$，故线性方程组$A'AX=A'\beta$总有解。

第5章　线性空间

1、 设V是非空集合，F是数域。如果存在下述两种运算：
- 加法：$V\times V\to V$，$(\alpha,\beta)\mapsto\alpha+\beta$
- 数乘：$F\times V\to V$，$(k,\alpha)\mapsto k\cdot\alpha$

满足以下 8 条公理：
(1) 加法交换律：$\alpha+\beta=\beta+\alpha$
(2) 加法结合律：$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$
(3) 存在零元：存在$0\in V$使得对V中任意元素α都有$0+\alpha=\alpha$（0称为V的零元或零向量）
(4) 存在负元：对任意的$\alpha\in V$都存在$\alpha'\in V$使得$\alpha+\alpha'=0$（α'称为α的负元）
(5) 数乘结合律：$(k_1\cdot k_2)\cdot\alpha=k_1\cdot(k_2\cdot\alpha)$
(6) 数乘单位律：$1\cdot\alpha=\alpha$
(7) 分配律 I：$(k_1+k_2)\cdot\alpha=k_1\cdot\alpha+k_2\cdot\alpha$
(8) 分配律 II：$k\cdot(\alpha+\beta)=k\cdot\alpha+k\cdot\beta$

其中α,β,γ是V中任意元素，k,k_1,k_2是F中任意数，则称$(V,F,+,\cdot)$为线性空间或向量空间，简称V是F上的线性空间或向量空间，V中的元素称为向量。

2、 设V是线性空间
(1) 若存在$n\in\mathbb{N}^*$个线性无关的向量$\alpha_1,\alpha_2,\dots,\alpha_n\in V$使得V中所有向量均可由这n个向量线性表出，则称V是n维线性空间，这n个线性无关的向量$\alpha_1,\alpha_2,\dots,\alpha_n$叫作V的一个基。
(2) 若V仅含有零元，则称V为零空间或0维线性空间。
(3) 若V是n维线性空间，其中$n\in\mathbb{N}$，则称V是有限维线性空间，并称n是V的维数，记为$\dim_F V$，简记为$\dim V$。如果V不是有限维线性空间，那么称V是无限维线性空间。
【例子】

2×2矩阵的维数为 4，其一组基为$\left\{\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\right\}$

2×2上三角矩阵的维数为 3，其一组基为$\left\{\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\right\}$

3、 设$\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n$和$\eta_1,\eta_2,\dots,\eta_n$是n维线性空间V的两个基，则存在$A\in M_n(F)$使得$(\eta_1,\eta_2,\dots,\eta_n)=(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n)A$，我们称A为从基$\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n$到基$\eta_1,\eta_2,\dots,\eta_n$的过渡矩阵。

4、 设V_1,V_2是线性空间V的两个子空间，若V_1+V_2中的每个向量α可唯一分解为$\alpha=\alpha_1+\alpha_2$，其中$\alpha_1\in V_1$，$\alpha_2\in V_2$，则称和V_1+V_2为直和，记为$V_1\oplus V_2$。

【例题】

1、 考虑集合$\mathbb{R}^+=\{a\in\mathbb{R}|a>0\}$，定义两种运算：

$\oplus$：$\mathbb{R}^+\times\mathbb{R}^+\to\mathbb{R}^+$，$(a,b)\mapsto ab$（通常的乘法），

$\circ$：$\mathbb{R}\times\mathbb{R}^+\to\mathbb{R}^+$，$(k,a)\mapsto a^k$（通常的方幂）。

证明：$(\mathbb{R}^+,\mathbb{R},\oplus,\circ)$是线性空间。

根据线性空间的定义，需要验证运算满足(1)~(8)条公理。我们逐条来验证：任取$a,b,c\in\mathbb{R}^+$，$k,l\in\mathbb{R}$，

(1) 因为$a\oplus b=ab=ba=b\oplus a$，所以加法交换律成立。

(2) 因为$(a\oplus b)\oplus c=(ab)c=a(bc)=a\oplus(b\oplus c)$，所以加法结合律成立。

(3) 因为$1\oplus a=1a=a$，所以加法有零元

（注意：1与a进行$\oplus$运算后仍为a，故零元是1）。

(4) 因为$a\oplus\frac{1}{a}=a\frac{1}{a}=1$，所以加法有负元

（注意：a与$\frac{1}{a}$进行$\oplus$运算后为零元1，故负元是$\frac{1}{a}$）。

(5) 因为$l\circ(k\circ a)=l\circ a^k=a^{lk}=(lk)\circ a$，所以数乘有结合律。

(6) 因为$1\circ a=a^1=a$，所以数乘有单位律。

(7) 因为$(l+k)\circ a=a^{l+k}=a^la^k=(l\circ a)\oplus(k\circ a)$，所以分配律 I 成立。

(8) 因为$l\circ(a\oplus b)=(ab)^l=a^lb^l=(l\circ a)\oplus(l\circ b)$，所以分配律 II 成立。

2、 已知三位向量空间的两个基为$\alpha_1=(1,1,1)^T$，$\alpha_2=(1,0,-1)^T$，$\alpha_3=(1,0,1)^T$，$\beta_1=(1,2,1)^T$，$\beta_2=(2,3,4)^T$，$\beta_3=(3,4,3)^T$。

(1) 求由基$\alpha_1,\alpha_2,\alpha_3$到基$\beta_1,\beta_2,\beta_3$的过渡矩阵P

由$(\beta_1,\beta_2,\beta_3)=(\alpha_1,\alpha_2,\alpha_3)P$得$P=(\alpha_1,\alpha_2,\alpha_3)^{-1}(\beta_1,\beta_2,\beta_3)$

由$(\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3)=\begin{pmatrix}1&1&1&1&2&3\\1&0&0&2&3&4\\1&-1&1&1&4&3\end{pmatrix}\to\begin{pmatrix}1&0&0&2&3&4\\0&1&0&0&-1&0\\0&0&1&-1&0&-1\end{pmatrix}$

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

(2) 设ξ在基$\alpha_1,\alpha_2,\alpha_3$下的坐标为$(1,1,-1)$，求$\xi$在基$\beta_1,\beta_2,\beta_3$下的坐标。

由于$\begin{pmatrix}\xi_1\\\xi_2\\\xi_3\end{pmatrix}=P^{-1}\begin{pmatrix}1\\1\\-1\end{pmatrix}=\begin{pmatrix}-\frac{1}{2}&-\frac{3}{2}&-2\\0&-1&0\\\frac{1}{2}&\frac{3}{2}&1\end{pmatrix}\begin{pmatrix}1\\1\\-1\end{pmatrix}=\begin{pmatrix}0\\-1\\1\end{pmatrix}$

故ξ在基β_1,β_2,β_3下的坐标为$(0,-1,1)$。

第6章 欧几里得空间

1、 设V是实线性空间，f是从集合$V\times V$到$\mathbb{R}$的映射，并且满足下面的条件：

(1) $f(\alpha,\beta)=f(\beta,\alpha)$

(2) $f(\alpha+\beta,\gamma)=f(\alpha,\gamma)+f(\beta,\gamma)$

(3) $f(k\alpha,\beta)=kf(\alpha,\beta)$

(4) $f(\alpha,\alpha)\geq0$，等号成立当且仅当$\alpha=0$

其中α,β,γ是V中任意向量，k是任意实数，则称$f(\alpha,\beta)$为α,β的内积，以后将α,β简记为(α,β)。带有内积的有限维实线性空间称为欧几里得空间，简称为欧氏空间。

2、设$V\in\mathbb{R}^n$是n维实线性空间，$\alpha=(x_1,\dots,x_n)',\beta=(y_1,\dots,y_n)'\in V$，通常的内积定义为

$$(\alpha,\beta)=\alpha'\beta=(x_1,x_2,\dots,x_n)\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}=x_1y_1+x_2y_2+\cdots+x_ny_n。$$

3、设V是欧氏空间，$\alpha\in V$，定义$|\alpha|=(\alpha,\alpha)$为α的长度，或范数。长度为1的向量称为单位向量。

4、设α,β为欧氏空间V中非零向量，定义它们的夹角为$\theta=\arccos\frac{(\alpha,\beta)}{|\alpha||\beta|}$。

5、设$\alpha_1,\alpha_2,\dots,\alpha_n$是欧氏空间$V$的一个基。定义$\alpha_1,\alpha_2,\dots,\alpha_n$的度量矩阵或格拉姆矩阵为

$$\begin{pmatrix}(\alpha_1,\alpha_1)&(\alpha_1,\alpha_2)&\cdots&(\alpha_1,\alpha_n)\\(\alpha_2,\alpha_1)&(\alpha_2,\alpha_2)&\cdots&(\alpha_2,\alpha_n)\\\vdots&\vdots&\ddots&\vdots\\(\alpha_n,\alpha_1)&(\alpha_n,\alpha_2)&\cdots&(\alpha_n,\alpha_n)\end{pmatrix}。$$

6、设V是欧氏空间，$\alpha,\beta\in V$。如果$(\alpha,\beta)=0$，则称α,β正交或者说α正交于β，记为$\alpha\perp\beta$。如果有一组非零向量两两正交，则称这组向量是正交向量组。

7、设$\alpha\in V$，W是V的子空间，如果对任意$\beta\in W$都有$(\alpha,\beta)=0$，则称α与W正交，记为$\alpha\perp W$。设V_1,V_2都是V的子空间，如果对任意$\alpha_1\in V_1$以及$\alpha_2\in V_2$，有$(\alpha_1,\alpha_2)=0$，则称V_1与V_2正交，记为$V_1\perp V_2$。

8、设n欧氏空间V的向量$\alpha_1,\alpha_2,\dots,\alpha_k$满足对任意$1\le i,j\le k$，$(\alpha_i,\alpha_j)=\delta_{ij}$（当$i=j$时，$(\alpha_i,\alpha_j)=1$，即向量与自身的内积为1；当$i\neq j$时，$(\alpha_i,\alpha_j)=0$，即不同向量之间的内积为0），则称$\alpha_1,\alpha_2,\dots,\alpha_k$是$V$的标准正交向量组。若$e_1,e_2,\dots,e_n$是$V$的标准正交向量组，则称$e_1,e_2,\dots,e_n$是$V$的标准正交基（规范正交基）。

9、设$\alpha_1,\alpha_2,\dots,\alpha_n$是欧氏空间$V$的一组线性无关的向量，$W=L(\alpha_1,\alpha_2,\dots,\alpha_n)$，则存在$W$的一个标准正交基$e_1,e_2,\dots,e_n$使得对任意$1\le m\le n$都有$L(e_1,e_2,\dots,e_m)=L(\alpha_1,\alpha_2,\dots,\alpha_m)$。

10、求标准正交向量组的过程叫做格拉姆-施密特标准正交化过程，简称施密特标准正交化，即

$$\begin{cases}\beta_1=\alpha_1\\\beta_2=\alpha_2-\dfrac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1\\\beta_3=\alpha_3-\dfrac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\dfrac{(\alpha_2,\beta_2)}{(\beta_2,\beta_2)}\beta_2\\\vdots\\\beta_m=\alpha_m-\dfrac{(\alpha_m,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\cdots-\dfrac{(\alpha_m,\beta_{m-1})}{(\beta_{m-1},\beta_{m-1})}\beta_{m-1}\\\vdots\\\beta_n=\alpha_n-\dfrac{(\alpha_n,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\cdots-\dfrac{(\alpha_n,\beta_{n-1})}{(\beta_{n-1},\beta_{n-1})}\beta_{n-1}\end{cases}$$

【例题】

1、 证明：设α,β是欧氏空间V中的两个向量，则$(\alpha,\beta)^2\le(\alpha,\alpha)(\beta,\beta)$，等号成立当且仅当$\alpha,\beta$线性相关。

当α,β中有一个是零向量时，结论显然成立。

设α,β都是非零向量，令$f(t)=(\alpha+t\beta,\alpha+t\beta)=(\beta,\beta)t^2+2(\alpha,\beta)t+(\alpha,\alpha)$，则$f(t)$是首项系数为正数的一元二次多项式。对任意$t\in R$，由定义知，$f(t)=(\alpha+t\beta,\alpha+t\beta)\ge0$。因此$f(t)$的判别式$\Delta=4(\alpha,\beta)^2-4(\alpha,\alpha)(\beta,\beta)\le0$。所以$(\alpha,\beta)^2\le(\alpha,\alpha)(\beta,\beta)$，等号成立当且仅当存在实数t使得$\alpha+t\beta=0$，即$\alpha,\beta$线性相关。

2、 证明：设$\alpha_1,\alpha_2,\dots,\alpha_n$是欧氏空间V的一个基，则$\alpha_1,\alpha_2,\dots,\alpha_n$的度量矩阵A是正定矩阵。

任取非零向量$X=(x_1,x_2,\dots,x_n)'\in\mathbb{R}^m$。

令$\gamma=\sum_{i=0}^n x_i\alpha_i=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n\in V$，则$\gamma\ne0$，从而$X'AX=$

$$(x_1,x_2,\dots,x_n)\begin{pmatrix}(\alpha_1,\alpha_1)&(\alpha_1,\alpha_2)&\cdots&(\alpha_1,\alpha_n)\\(\alpha_2,\alpha_1)&(\alpha_2,\alpha_2)&\cdots&(\alpha_2,\alpha_n)\\\vdots&\vdots&\ddots&\vdots\\(\alpha_n,\alpha_1)&(\alpha_n,\alpha_2)&\cdots&(\alpha_n,\alpha_n)\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\left(\sum_{i=0}^n x_i\alpha_i,\sum_{i=0}^n x_i\alpha_i\right)=(\gamma,\gamma)>0,$$

故A是正定矩阵。

3、 证明：设V_1与V_2是欧氏空间V的两个正交的子空间，则V_1与V_2的和为直和。

任取$\alpha\in V_1\cap V_2$，因为V_1与V_2是正交的，所以$(\alpha,\alpha)=0$，从而$\alpha=0$。故V_1与V_2的和为直和。

4、 设$\alpha_1=\begin{pmatrix}1\\1\\1\end{pmatrix},\alpha_2=\begin{pmatrix}1\\0\\2\end{pmatrix},\alpha_3=\begin{pmatrix}0\\0\\1\end{pmatrix}\in\mathbb{R}^3$。证明：$\alpha_1,\alpha_2,\alpha_3$是$\mathbb{R}^3$的一个基，并用施密特标准正交化将其化为标准正交基。

由$\begin{vmatrix}1&1&0\\1&0&0\\1&2&1\end{vmatrix}=-1$知，$\alpha_1,\alpha_2,\alpha_3$是$\mathbb{R}^3$的一个基。

另$\beta_1=\alpha_1=\begin{pmatrix}1\\1\\1\end{pmatrix},\ \beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1=\begin{pmatrix}0\\-1\\1\end{pmatrix},\ \beta_3=\alpha_3-\frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_2,\beta_2)}{(\beta_2,\beta_2)}\beta_2=\begin{pmatrix}-\frac13\\\frac16\\\frac16\end{pmatrix},$

则β_1,β_2,β_3是$\mathbb{R}^3$的一个正交基。

再令$e_1=\frac{\beta_1}{|\beta_1|}=\begin{pmatrix}\frac{1}{\sqrt3}\\\frac{1}{\sqrt3}\\\frac{1}{\sqrt3}\end{pmatrix},\ e_2=\frac{\beta_2}{|\beta_2|}=\begin{pmatrix}0\\-\frac{1}{\sqrt2}\\\frac{1}{\sqrt2}\end{pmatrix},\ e_3=\frac{\beta_3}{|\beta_3|}=\begin{pmatrix}-\frac{2}{\sqrt6}\\\frac{1}{\sqrt6}\\\frac{1}{\sqrt6}\end{pmatrix},$

则e_1,e_2,e_3是$\mathbb{R}^3$的一个标准正交基。

第7章　正交变换与正交矩阵

1、 设V是欧氏空间，$\sigma\in End(V)$。如果对任意$\alpha,\beta\in V$都有$(\alpha,\beta)=\big(\sigma(\alpha),\sigma(\beta)\big)$，则称$\sigma$是V上的正交变换。

2、 设$A\in M_n(\mathbb{R})$。如果$A'A=I_n$，则称A是正交矩阵。所有n级正交矩阵组成的集合记为$O(n)$,

称为n级正交群。

第8章　实对称矩阵及其极分解

1、 对称矩阵$A\in M_n(\mathbb{R})$的复特征值都是实数。

2、 任一n级实对称矩阵A都正交相似于对角矩阵。

3、 设A是实对称矩阵，A正定当且仅当A的特征值都是正数，A半正定当且仅当A的特征值都是非负数。

4、 设$A\in M_n(\mathbb{R})$，则一定存在n级正交矩阵R和半正定矩阵P使得$A=RP$（或者$A=PR$），其中P由A唯一确定。如果A是可逆的，那么R也由A唯一确定且P正定。这种分解称为矩阵的极分解。

5、 设A是$m\times n$实矩阵，则m级半正定矩阵AA'或n级半正定矩阵$A'A$的正特征值（λ_i或μ_i）的正平方根（$\sqrt{\lambda_i}$或$\sqrt{\mu_i}$）称为矩阵 A 的正奇异值，简称奇异值。（$\lambda_i=\mu_i>0$）

6、 设 $A\in M_{m\times n}(\mathbb{R})$，则存在 $U\in O(m)$，$V\in O(n)$ 和分块矩阵 $S=\begin{pmatrix}\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\dots,\sqrt{\lambda_r}) & 0\\ 0 & 0\end{pmatrix}$使得$A=USV^T$，其中$\sqrt{\lambda_1},\sqrt{\lambda_2},\dots,\sqrt{\lambda_r}$是A的奇异值，这种分解称为矩阵A的奇异值分解。（U和V为方阵，S和A的形状相同）

【例题】

1、 证明：设A，B都是实对称矩阵，B是正定矩阵，则A正定当且仅当AB的特征值都是正数。
因为B正定，所以存在可逆实矩阵P使得$B=P'P$，从而$AB=AP'P=P^{-1}PAP'P$。
故AB与PAP'相似，从而AB与PAP'具有相同的特征值。
又因为PAP'与A合同，故A正定$\Leftrightarrow PAP'$正定$\Leftrightarrow PAP'$的特征值$>0\Leftrightarrow AB$的特征值>0。
故A正定当且仅当AB的特征值都是正数。

2、 已知矩阵$A=\begin{bmatrix}1&1\\1&1\\0&0\end{bmatrix}$，求矩阵A的奇异值和奇异值分解。

【方法一】

矩阵AA^H为$AA^H=\begin{bmatrix}1&1\\1&1\\0&0\end{bmatrix}\begin{bmatrix}1&1&0\\1&1&0\end{bmatrix}=\begin{bmatrix}2&2&0\\2&2&0\\0&0&0\end{bmatrix}$

根据特征方程$\left|\lambda I-AA^H\right|=0$，有$\begin{vmatrix}\lambda-2&-2&0\\-2&\lambda-2&0\\0&0&\lambda\end{vmatrix}=\lambda\begin{vmatrix}\lambda-2&-2\\-2&\lambda-2\end{vmatrix}=\lambda^2(\lambda-4)=0$

解得特征值为$\lambda_1=4$，$\lambda_2=\lambda_3=0$

故矩阵A的奇异值为 2，奇异值矩阵$\Sigma=[2]$，分块矩阵$S=\begin{pmatrix}\Sigma&0\\0&0\end{pmatrix}=\begin{bmatrix}2&0\\0&0\\0&0\end{bmatrix}$

当$\lambda=4$时，$\lambda I-AA^H=\begin{bmatrix}4-2&-2&0\\-2&4-2&0\\0&0&4\end{bmatrix}\to\begin{bmatrix}2&-2&0\\-2&2&0\\0&0&4\end{bmatrix}\to\begin{bmatrix}1&-1&0\\0&0&1\\0&0&0\end{bmatrix}$

当$\lambda = 0$时，$\lambda I - AA^H = \begin{bmatrix} 0-2 & -2 & 0 \\ -2 & 0-2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} -2 & -2 & 0 \\ -2 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

故$\lambda_1 = 4$对应的特征向量为$u_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}$（单位化），$U_1 = u_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}$（由非零特征值对应的特征向量组成），$\lambda_2 = \lambda_3 = 0$对应的特征向量为$u_2 = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}$和$u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$（施密特正交化并单位化），$U_2 = (u_2, u_3) = \begin{bmatrix} -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & 0 \\ 0 & 1 \end{bmatrix}$（由零特征值对应的特征向量组成）

于是$V_1 = A^H U_1 \Sigma^{-H} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$

矩阵$A^H A$为$A^H A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$

根据特征方程$\left| \mu I - A^H A \right| = 0$，有$\begin{vmatrix} \mu - 2 & -2 \\ -2 & \mu - 2 \end{vmatrix} = \mu(\mu - 4) = 0$

解得特征值为$\mu_1 = 4$，$\mu_2 = 0$

当$\mu = 0$时，$\mu I - A^H A = \begin{bmatrix} 0-2 & -2 \\ -2 & 0-2 \end{bmatrix} \to \begin{bmatrix} -2 & -2 \\ -2 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$

故$\mu_2 = 0$对应的特征向量为$v_2 = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$（单位化），$V_2 = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$（由零特征值对应的特征向量组成）

于是矩阵U表示为$U = (U_1, U_2) = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$，矩阵V表示为$V = (V_1, V_2) = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$

因此，矩阵A的奇异值分解为$A = USV^H = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}^H$

【方法二】

矩阵AA'为$AA' = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

根据特征方程$|\lambda I-AA'|=0$，有$\begin{vmatrix}\lambda-2 & -2 & 0\\ -2 & \lambda-2 & 0\\ 0 & 0 & \lambda\end{vmatrix}=\lambda\begin{vmatrix}\lambda-2 & -2\\ -2 & \lambda-2\end{vmatrix}=\lambda^2(\lambda-4)=0$

解得特征值为$\lambda_1=4$，$\lambda_2=\lambda_3=0$

故矩阵A的奇异值为 2，分块矩阵$S=\begin{pmatrix}\mathrm{diag}\left(\sqrt{\lambda_1}\right) & 0\\ 0 & 0\end{pmatrix}=\begin{bmatrix}2 & 0\\ 0 & 0\\ 0 & 0\end{bmatrix}$

当$\lambda=4$时，$\lambda I-AA'=\begin{bmatrix}4-2 & -2 & 0\\ -2 & 4-2 & 0\\ 0 & 0 & 4\end{bmatrix}\to\begin{bmatrix}2 & -2 & 0\\ -2 & 2 & 0\\ 0 & 0 & 4\end{bmatrix}\to\begin{bmatrix}1 & -1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}$

当$\lambda=0$时，$\lambda I-AA'=\begin{bmatrix}0-2 & -2 & 0\\ -2 & 0-2 & 0\\ 0 & 0 & 0\end{bmatrix}\to\begin{bmatrix}-2 & -2 & 0\\ -2 & -2 & 0\\ 0 & 0 & 0\end{bmatrix}\to\begin{bmatrix}1 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$

故$\lambda_1=4$对应的特征向量为$u_1=\begin{bmatrix}\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\\ 0\end{bmatrix}$（单位化），$\lambda_2=\lambda_3=0$对应的特征向量为$u_2=\begin{bmatrix}-\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\\ 0\end{bmatrix}$和$u_3=\begin{bmatrix}0\\ 0\\ 1\end{bmatrix}$（施密特正交化并单位化），于是矩阵U表示为$U=(u_1,u_2,u_3)=\begin{bmatrix}\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0\\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0\\ 0 & 0 & 1\end{bmatrix}$

矩阵$A'A$为$A'A=\begin{bmatrix}1 & 1 & 0\\ 1 & 1 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\\ 0 & 0\end{bmatrix}=\begin{bmatrix}2 & 2\\ 2 & 2\end{bmatrix}$

根据特征方程$|\mu I-A'A|=0$，有$\begin{vmatrix}\mu-2 & -2\\ -2 & \mu-2\end{vmatrix}=\mu(\mu-4)=0$

解得特征值为$\mu_1=4$，$\mu_2=0$

当$\mu=4$时，$\mu I-A'A=\begin{bmatrix}4-2 & -2\\ -2 & 4-2\end{bmatrix}\to\begin{bmatrix}2 & -2\\ -2 & 2\end{bmatrix}\to\begin{bmatrix}1 & -1\\ 0 & 0\end{bmatrix}$

当$\mu=0$时，$\mu I-A'A=\begin{bmatrix}0-2 & -2\\ -2 & 0-2\end{bmatrix}\to\begin{bmatrix}-2 & -2\\ -2 & -2\end{bmatrix}\to\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}$

故$\mu_1=4$对应的特征向量为$v_1=\begin{bmatrix}\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\end{bmatrix}$（单位化），$\mu_2=0$对应的特征向量为$v_2=\begin{bmatrix}-\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\end{bmatrix}$（单位化），于是矩阵V表示为$V=(v_1,v_2)=\begin{bmatrix}\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\end{bmatrix}$

因此，矩阵A的奇异值分解为$A=USV'=\begin{bmatrix}\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0\\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}2 & 0\\ 0 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\end{bmatrix}'$

第9章　最小二乘问题

1、 设$A\in M_{m\times n}(\mathbb{R}),\beta\in\mathbb{R}^m$。当$AX=\beta$无解时，称$A'AX=A'\beta$的解为$AX=\beta$的最小二乘解，称用最小二乘解作为$AX=\beta$的近似解的方法为最小二乘法。

【例题】

1、 证明：线性方程组$A'AX=A'\beta$总是有解。

由于A是实矩阵，有$\mathrm{rank}(A)=\mathrm{rank}(A^TA)$

则$\mathrm{rank}(A)=\mathrm{rank}(A^TA)\le\mathrm{rank}(A^TA|A^T\beta)=\mathrm{rank}\left(A^T(A|\beta)\right)\le\mathrm{rank}(A^T)=\mathrm{rank}(A)$

于是$\mathrm{rank}(A^TA|A^T\beta)=\mathrm{rank}(A)=\mathrm{rank}(A^TA)$，从而$A'AX=A'\beta$总是有解。

2、 已知某种材料在生产过程中的废品率y与某种化学成分x有关。下表记录了某工厂生产中y与相应的x的几次数值，找出y对x的一个近似公式。

废品率y与某化学成分x的关系表

废品率y/%	1.00	0.9	0.9	0.81	0.60	0.56	0.35
某化学成分占比x/%	3.6	3.7	3.8	3.9	4.0	4.1	4.2

把表中数值画出图来看，发现它的变化趋势近似于一条直线，因此我们决定选取x的一次式$ax+b$来表达，即求解齐次线性方程组$\begin{cases}3.6a+b=1.00\\3.7a+b=0.9\\3.8a+b=0.9\\3.9a+b=0.81\\4.0a+b=0.60\\4.1a+b=0.56\\4.2a+b=0.35\end{cases}$的最小二乘解。

令$A=\begin{bmatrix}3.6&1\\3.7&1\\3.8&1\\3.9&1\\4.0&1\\4.1&1\\4.2&1\end{bmatrix}$，$B=\begin{bmatrix}1.00\\0.9\\0.9\\0.81\\0.60\\0.56\\0.35\end{bmatrix}$，则最小二乘解a, b所满足的方程组为$A^TA\begin{pmatrix}a\\b\end{pmatrix}=A^TB$，即

$\begin{cases}106.75a+27.3b=19.675\\27.3a+7b=5.12\end{cases}$，解得$\begin{cases}a=-1.05\\b=4.81\end{cases}$

故y对x的近似公式为$y=-1.05x+4.81$

3、 已知数据点(1,3)，(3,1)，(5,7)，(4,6)，(7,4)，求最小二乘的拟合直线。

【方法一】

题目即求解齐次线性方程组$\begin{cases}a+b=3\\3a+b=1\\5a+b=7\\4a+b=6\\7a+b=4\end{cases}$的最小二乘解

令$A=\begin{bmatrix}1&1\\3&1\\5&1\\4&1\\7&1\end{bmatrix}$，$B=\begin{bmatrix}3\\1\\7\\6\\4\end{bmatrix}$，则最小二乘解a, b所满足的方程组为$A^TA\begin{pmatrix}a\\b\end{pmatrix}=A^TB$，即

$\begin{cases}100a+20b=93\\20a+5b=21\end{cases}$，解得$\begin{cases}a=0.45\\b=2.4\end{cases}$，故最小二乘的拟合直线为$y=0.45x+2.4$

【方法二】

$\overline{x}=\frac{\sum_{i=1}^{5}x_i}{5}=\frac{1+3+5+4+7}{5}=\frac{20}{5}=4$，$\overline{y}=\frac{\sum_{i=1}^{5}y_i}{5}=\frac{3+1+7+6+4}{5}=\frac{21}{5}=4.2$

最小二乘拟合的代价函数可以表示为

$$D_{LS}^{(1)}=\sum_{i=1}^{5}[m(x_i-\overline{x})+(y_i-\overline{y})]^2$$

$$=(-3m-1.2)^2+(-m-3.2)^2+(m+2.8)^2+1.8^2+(3m-0.2)^2$$

$\frac{D_{LS}^{(1)}}{\partial m}=40m+18=0\Rightarrow m=-0.45$

则最小二乘的拟合直线为$-0.45(x-4)+(y-4.2)=0$，即$y=0.45x+2.4$

第10章 范数、序列、级数

10.1 向量范数

1、 设V是数域F（一般为实数域R或复数域C）上的线性空间，用$\|x\|$表示按照某个法则确定的与向量 x 对应的实数，且满足

(1) 非负性：当$x\neq 0$，$\|x\|>0$；当且仅当$x=0$时，$\|x\|=0$；

(2) 齐次性：$\|kx\|=|k|\cdot\|x\|$，k为任意数；$|k|$表示k的绝对值或模长；

(3) 三角不等式：对于V中任何向量x,y都有$\|x+y\|\leq\|x\|+\|y\|$。

则称实数$\|x\|$是向量x的<u>向量范数</u>。

2、 设向量$x=(x_1,x_2,\dots,x_n)^T$，对任意数$p\geq 1$，称$\|x\|_p=(\sum_{i=1}^{n}|x_i|^p)^{\frac{1}{p}}$为向量x的<u>p-范数</u>。常用的p-范数有下述三种：

(1) 1-范数：$\|x\|_1=\sum_{i=1}^{n}|x_i|$

(2) 2-范数：$\|x\|_2=(\sum_{i=1}^{n}|x_i|^2)^{\frac{1}{2}}=(x^Hx)^{\frac{1}{2}}$也称为欧式范数（$x^H$为x的共轭转置）

(3) ∞-范数：$\|x\|_\infty=\lim\limits_{p\to\infty}\|x\|_p=\max|x_i|,(i=1,2,\dots,n)$

3、 设V是n维线性空间，$\|x\|_\alpha$和$\|x\|_\beta$为任意两种向量范数（不限于 p 范数），则总存在正数c_1,c_2，对V中所有向量$x\in V$恒有$c_1\|x\|_\beta\leq\|x\|_\alpha\leq c_2\|x\|_\beta$。

4、 若实数$\|x\|$是向量x的向量范数，则对于$\forall c>0$，实数$c\|x\|$也是向量x的向量范数。

【例题】

1、 设$\|x\|$是向量x的范数，证明：$\|x+y\|\geq|\|x\|-\|y\||$。

由$\|x\|=\|x-y+y\|\leq\|x-y\|+\|y\|$得$\|x\|-\|y\|\leq\|x-y\|$

由$\|y\|=\|x-x+y\|\leq\|x\|+\|-x+y\|=\|x\|+\|x-y\|$得$\|x\|-\|y\|\geq-\|x-y\|$

故$\|x-y\|\geq|\|x\|-\|y\||$，再由$\|-y\|=\|y\|$得$\|x+y\|\geq|\|x\|-\|y\||$

2、 证明：$\|x\|_p=(\sum_{i=1}^{n}|x_i|^p)^{\frac{1}{p}}$为向量$x=(x_1,x_2,\dots,x_n)^T$的p-范数。

(1) 非负性：显然满足

(2) 齐次性：

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

(3) 三角不等式：若设$y=(y_1,y_2,\dots,y_n)^T$，则

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

3、 设$\|x\|_\infty$是向量x的∞-范数，证明：$\|x\|_\infty=\max|x_i|,(i=1,2,\dots,n)$

令$\alpha=\max|x_i|$，则$\beta_i=\frac{|x_i|}{\alpha}\le 1\ (i=1,2,\dots,n)$，即$|x_i|=\alpha\cdot\beta_i$

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

由于$1\le(\sum_{i=1}^{n}{\beta_i}^p)^{\frac{1}{p}}\le n^{\frac{1}{p}}$，故$\lim\limits_{p\to\infty}(\sum_{i=1}^{n}{\beta_i}^p)^{\frac{1}{p}}=1$

因此$\|x\|_\infty=\lim\limits_{p\to\infty}\|x\|_p=\alpha=\max|x_i|$

10.2 矩阵范数

1、 对于任何一个矩阵$A\in C^{m\times n}$，用$\|A\|$表示按照某个法则确定的与矩阵A对应的实数，且满足：

(1) 非负性：当$A\ne 0$时，$\|A\|>0$；当且仅当$A=0$时，$\|A\|=0$

(2) 齐次性：$\|kA\|=|k|\cdot\|A\|$，k为任意复数

(3) 三角不等式：对于任何两个同类型矩阵A,B都有$\|A+B\|\le\|A\|+\|B\|$

(4) 矩阵乘法相容性：若A与B可乘，有$\|AB\|\le\|A\|\cdot\|B\|$

则称对应于A的这个实数$\|A\|$是矩阵A的矩阵范数。

2、 若$\|A\|_\alpha$与$\|A\|_\beta$是任意两种矩阵范数，则总存在正数c_1,c_2，对于任意矩阵A恒有$c_1\|A\|_\beta\le\|A\|_\alpha\le c_2\|A\|_\beta$。

3、 若$A=(a_{ij})\in C^{m\times n}$，规定$\|A\|_F=\left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{\frac{1}{2}}$为矩阵A的 Frobenius 范数，是向量范数中欧氏范数的形式推广，其满足以下性质：

(1) 若$A=(\alpha_1,\alpha_2,\dots,\alpha_n)$，则$\|A\|_F^2=\sum_{i=1}^{n}\|\alpha_i\|_2^2$

(2) $\|A\|_F^2=tr(A^HA)=\sum_{i=1}^{n}\lambda_i(A^HA)$，其中$\lambda_i(A^HA)$表示n阶方阵$A^HA$的第i个特征值，$tr(A^HA)$是$A^HA$的迹。

(3) 对于任何m阶酉矩阵U与n阶酉矩阵V，都有等式$\|A\|_F=\|UA\|_F=\|A^H\|_F=\|AV\|_F=\|UAV\|_F$

【例题】

1、 证明：对$A=(a_{ij})\in C^{m\times n}$，$\|A\|=\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|$是矩阵范数（向量1-范数的形式推广）。

(1) 非负性：显然满足

(2) 齐次性：

$$\|kA\| = \sum_{i=1}^{m}\sum_{j=1}^{n}\left|k \cdot a_{ij}\right| = |k| \cdot \sum_{i=1}^{m}\sum_{j=1}^{n}\left|a_{ij}\right| = |k| \cdot \|A\|$$

(3) 三角不等式：若设$B = (b_{ij}) \in C^{m\times n}$，则

$$\|A+B\| = \sum_{i=1}^{m}\sum_{j=1}^{n}\left|a_{ij}+b_{ij}\right| \le \sum_{i=1}^{m}\sum_{j=1}^{n}\left(\left|a_{ij}\right|+\left|b_{ij}\right|\right) = \sum_{i=1}^{m}\sum_{j=1}^{n}\left|a_{ij}\right| + \sum_{i=1}^{m}\sum_{j=1}^{n}\left|b_{ij}\right| = \|A\| + \|B\|$$

(4) 矩阵乘法相容性：若$A = (a_{ij}) \in C^{m\times p}$，$B = (b_{ij}) \in C^{p\times n}$，则

$$\|AB\| = \sum_{i=1}^{m}\sum_{j=1}^{n}\left|\sum_{k=1}^{p} a_{ik}b_{kj}\right| \le \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p}\left|a_{ik}\right| \cdot \left|b_{kj}\right| \le \sum_{i=1}^{m}\sum_{j=1}^{n}\left[\left(\sum_{k=1}^{p}\left|a_{ik}\right|\right)\left(\sum_{k=1}^{p}\left|b_{kj}\right|\right)\right]$$

$$= \left(\sum_{i=1}^{m}\sum_{k=1}^{p}\left|a_{ik}\right|\right)\left(\sum_{j=1}^{n}\sum_{k=1}^{p}\left|b_{kj}\right|\right) = \|A\| \cdot \|B\|$$

2、 若实数$\|A\|$是矩阵A的矩阵范数，则对于$\forall c > 0$，实数$c\|A\|$也是矩阵A的矩阵范数吗？
由于矩阵范数需须满足矩阵乘法相容性，即若A与B可乘，有$\|AB\| \le \|A\| \cdot \|B\|$。
由于$c\|AB\| \le c\|A\| \cdot c\|B\|$，不等式两边同时约去c，得$\|AB\| \le c\|A\| \cdot \|B\|$。
因此，须满足$\|AB\| \le \|A\| \cdot \|B\| \le c\|A\| \cdot \|B\|$，
即$c \ge 1$时，实数$c\|A\|$也是矩阵A的矩阵范数。

10.3 诱导范数（算子范数）

1、 设$\|x\|_\alpha$是向量范数，$\|A\|_\beta$是矩阵范数，若对于任何矩阵A与向量x都有$\|Ax\|_\alpha \le \|A\|_\beta \cdot \|x\|_\alpha$，则称$\|A\|_\beta$为与向量范数$\|x\|_\alpha$相容的矩阵范数。

2、 上述所定义的矩阵范数称为由向量范数$\|x\|_\alpha$所诱导的诱导范数，也称算子范数。

3、 单位矩阵 E 的（任何）诱导范数$\|E\|_i = 1$。

证明：将$A = E$代入$\|Ax\|_\alpha \le \|A\|_i \cdot \|x\|_\alpha$，得$\|E\|_i = \max\limits_{x\ne 0}\frac{\|Ex\|_\alpha}{\|x\|_\alpha} = \max\limits_{x\ne 0}\frac{\|x\|_\alpha}{\|x\|_\alpha} = 1$

4、 由向量p-范数$\|x\|_p$所诱导的矩阵范数称为矩阵p-范数，即$\|A\|_p = \max\limits_{x\ne 0}\frac{\|Ax\|_p}{\|x\|_p}$。设$A = (a_{ij})_{m\times n}$，则

(1) $\|A\|_1 = \max\limits_{1\le j\le n}\left(\sum_{i=1}^{m}\left|a_{ij}\right|\right)$，称$\|A\|_1$是列和范数。

(2) $\|A\|_2 = \max\limits_{1\le j\le n}\left(\lambda_j(A^HA)\right)^{\frac{1}{2}}$，$\lambda_j(A^HA)$表示矩阵$A^HA$的第j个特征值。称$\|A\|_2$是谱范数，即$\|A\|_2$是A的最大正奇异值。

(3) $\|A\|_\infty = \max\limits_{1\le i\le m}\left(\sum_{j=1}^{n}\left|a_{ij}\right|\right)$，称$\|A\|_\infty$是行和范数。

5、 若A是正规矩阵（$A^HA = AA^H$），则$\|A\|_2 = \max|\lambda_i(A)|$。

6、 设$\|A\|$是矩阵范数，则存在向量范数$\|x\|$，满足$\|Ax\| \le \|A\| \cdot \|x\|$。（方法：任给非零向量α，定义向量范数$\|A\|_\alpha = \left\|xa^H\right\|$）

7、 设$A \in C^{n\times n}$，A的n个特征值为$\lambda_1, \lambda_2, \dots, \lambda_n$，称$\rho(A) = \max\{|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|\}$是A的谱半径。

8、 $A\in C^{m\times n}$，则$\rho(A)\le\|A\|$，其中$\|A\|$是A的任何一种范数。
证明：设λ是A的任何一个特征值，即$Ax=\lambda x, x\neq 0$。故$|\lambda|\cdot\|x\|=\|\lambda x\|=\|Ax\|\le\|A\|\cdot\|x\|$，于是$|\lambda|\le\|A\|$。由于$\lambda$是A的任一个特征值，因此，$\rho(A)\le\|A\|$。

【例题】

1、 证明：设$\|x\|_\alpha$是向量范数，则$\|A\|_i=\max\limits_{x\neq 0}\frac{\|Ax\|_\alpha}{\|x\|_\alpha}$满足矩阵范数定义，且$\|A\|_i$是与向量范数$\|x\|_\alpha$相容的矩阵范数。

(1) 证明$\|A\|_i=\max\limits_{x\neq 0}\frac{\|Ax\|_\alpha}{\|x\|_\alpha}$满足矩阵范数定义：

非负性、齐次性：显然满足

三角不等式：

$$\|A+B\|_i=\max_{x\neq 0}\frac{\|(A+B)x\|_\alpha}{\|x\|_\alpha}=\max_{x\neq 0}\frac{\|Ax+Bx\|_\alpha}{\|x\|_\alpha}\le\max_{x\neq 0}\frac{\|Ax\|_\alpha+\|Bx\|_\alpha}{\|x\|_\alpha}$$
$$\le\max_{x\neq 0}\frac{\|Ax\|_\alpha}{\|x\|_\alpha}+\max_{x\neq 0}\frac{\|Bx\|_\alpha}{\|x\|_\alpha}=\|A\|_i+\|B\|_i$$

矩阵乘法相容性：

$$\|AB\|_i=\max_{x\neq 0}\frac{\|ABx\|_\alpha}{\|x\|_\alpha}=\max_{x\neq 0}\left(\frac{\|A(Bx)\|_\alpha}{\|Bx\|_\alpha}\frac{\|Bx\|_\alpha}{\|x\|_\alpha}\right)\le\max_{Bx\neq 0}\frac{\|A(Bx)\|_\alpha}{\|Bx\|_\alpha}\cdot\max_{x\neq 0}\frac{\|Bx\|_\alpha}{\|x\|_\alpha}\le\|A\|_i\cdot\|B\|_i$$

(2) 证明$\|A\|_i$是与向量范数$\|x\|_\alpha$相容的矩阵范数：

由$\|A\|_i=\max\limits_{x\neq 0}\frac{\|Ax\|_\alpha}{\|x\|_\alpha}$得$\|A\|_i\ge\frac{\|Ax\|_\alpha}{\|x\|_\alpha}$，即$\|Ax\|_\alpha\le\|A\|_i\cdot\|x\|_\alpha$，这表明$\|A\|_i$与$\|x\|_\alpha$相容。

2、 已知矩阵范数$\|A\|_F=\left(\sum_{i=1}^m\sum_{j=1}^n\left|a_{ij}\right|^2\right)^{\frac{1}{2}}$，求与之相容的一个向量范数。

取$\alpha=(1,0,\dots,0)^T$，设$x=(x_1,x_2,\dots,x_n)^T$，则$\|x\|=\left\|x\alpha^H\right\|=\left(\sum_{i=1}^i|x_i|^2\right)^{\frac{1}{2}}=\|x\|_2$

10.4 矩阵序列与极限

1、 设矩阵序列$\{A_k\}$，其中$A_k=\left(a_{ij}^{(k)}\right)\in C^{m\times n}$，若$m\times n$个数列$\left\{a_{ij}^{(k)}\right\}$ $(i=1,2,\dots,m;j=1,2,\dots,n)$都收敛，便称矩阵序列$\{A_k\}$收敛。若$\lim\limits_{k\to\infty}a_{ij}^{(k)}=a_{ij}$，则$\lim\limits_{k\to\infty}A_k=A=\left(a_{ij}\right)$，称A为矩阵序列$\{A_k\}$的极限。若把向量看成矩阵的特例，向量序列收敛的定义类似可得。

2、 3级矩阵的Jordan（若尔当）标准形：设A的若尔当标准形为J，考虑A的特征多项式$f_A(\lambda)$有无重根，分以下三种情况。

(1) 若$f_A(\lambda)=(\lambda-\lambda_1)^3$，则A相似于
$$\begin{cases}\begin{pmatrix}\lambda_1&&\\&\lambda_1&\\&&\lambda_1\end{pmatrix}\text{当且仅当 }\dim V_{\lambda_1}=3\\ \begin{pmatrix}\lambda_1&&\\&\lambda_1&1\\&&\lambda_1\end{pmatrix}\text{当且仅当 }\dim V_{\lambda_1}=2\\ \begin{pmatrix}\lambda_1&1&\\&\lambda_1&1\\&&\lambda_1\end{pmatrix}\text{当且仅当 }\dim V_{\lambda_1}=1\end{cases}$$

(2) 若$f_A(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)^2$，其中$\lambda_1\neq\lambda_2$，则$\dim V_{\lambda_1}=1$且A相似于

$$\begin{cases}\begin{pmatrix}\lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_2\end{pmatrix} \text{当且仅当 } \dim V_{\lambda_2}=2 \\ \begin{pmatrix}\lambda_1 & & \\ & \lambda_2 & 1 \\ & & \lambda_2\end{pmatrix} \text{当且仅当 } \dim V_{\lambda_2}=1\end{cases}$$

(3) 若$f_A(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3)$，其中$\lambda_1,\lambda_2,\lambda_3$互不相同，则$\dim V_{\lambda_i}=1, i=1,2,3$，且A相似于$\begin{pmatrix}\lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3\end{pmatrix}$

3、 若对矩阵A的某一种范数$\|A\|<1$，则$\lim\limits_{k\to\infty}A^k=0$。

4、 已知矩阵序列：$A,A^2,\dots,A^k,\dots$，则$\lim\limits_{k\to\infty}A^k=0$的充要条件是$\rho(A)<1$。

5、 设 A 的 Jordan 标准形 $J=\operatorname{diag}\big(J_1(\lambda_1),J_2(\lambda_2),\dots,J_r(\lambda_r)\big)$，其中 $J_i(\lambda_i)=\begin{bmatrix}\lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & 0 & & \ddots & 1 \\ & & & & \lambda_i\end{bmatrix}_{d_i\times d_i}$，$(i=1,2,\dots,r)$，于是$A^k=P\operatorname{diag}\left(J_1^k(\lambda_1),J_2^k(\lambda_2),\dots,J_r^k(\lambda_r)\right)P^{-1}$，其中$J_i^k(\lambda_i)=\begin{bmatrix}\lambda_i^k & C_k^1\lambda_i^{k-1} & \cdots & & C_k^{d_i-1}\lambda_i^{k-d_i+1} \\ & \lambda_i^k & C_k^1\lambda_i^{k-1} & & \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & C_k^1\lambda_i^{k-1} \\ & & & & \lambda_i^k\end{bmatrix}$，其中$C_k^l=\begin{cases}\frac{k(k-1)\cdots(k-l+1)}{l!} & \text{当 } l\le k \\ 0 & \text{当 } l>k\end{cases}$

【例题】

1、 求矩阵$A=\begin{bmatrix}-1 & -2 & 6 \\ -1 & 0 & 3 \\ -1 & -1 & 4\end{bmatrix}$的 Jordan 标准形。

令$|\lambda E-A|=\begin{vmatrix}\lambda+1 & 2 & -6 \\ 1 & \lambda & -3 \\ 1 & 1 & \lambda-4\end{vmatrix}=(\lambda-1)^3=0$得A的特征值为$\lambda_1=\lambda_2=\lambda_3=1$

当$\lambda_i=1$时，由$\lambda_iE-A=\begin{bmatrix}2 & 2 & -6 \\ 1 & 1 & -3 \\ 1 & 1 & -3\end{bmatrix}\to\begin{bmatrix}1 & 1 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0\end{bmatrix}$得$\dim V_{\lambda_i}=3-\operatorname{rank}(\lambda_i-A)=2$

故A的 Jordan 标准形为$J=\begin{bmatrix}1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1\end{bmatrix}$

2、 判断矩阵序列A^k的敛散性。

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

由于$A^k=\begin{bmatrix}1 & k\\0 & 1\end{bmatrix}$且$\lim\limits_{k\to\infty}k=\infty$，故$\lim\limits_{k\to\infty}A^k$发散。

(2) $A=\begin{bmatrix}0.9 & 1\\0 & 0.9\end{bmatrix}$

由于A的特征值$\lambda_1=\lambda_2=0.9$，故$\rho(A)<1$，故A^k收敛，且$\lim\limits_{k\to\infty}A^k=0$。

(3) $A=\begin{bmatrix}1 & 0 & 0\\0 & 0.9 & 1\\0 & 0 & 0.9\end{bmatrix}$

由于$A^k=\begin{bmatrix}1 & 0 & 0\\0 & 0.9^k & k0.9^{k-1}\\0 & 0 & 0.9^k\end{bmatrix}$且$\lim\limits_{k\to\infty}0.9^k=0$，$\lim\limits_{k\to\infty}k0.9^{k-1}=0$，故$A^k$收敛，且

$\lim\limits_{k\to\infty}A^k=\begin{bmatrix}1 & 0 & 0\\0 & 0 & 0\\0 & 0 & 0\end{bmatrix}$。

(4) $A=\begin{bmatrix}0.3 & 0.8\\0.6 & 0.1\end{bmatrix}$

由于$\|A\|_1=0.9$，故$\rho(A)<1$，故A^k收敛，且$\lim\limits_{k\to\infty}A^k=0$。

(5) $A=\begin{bmatrix}1 & 1 & 0\\0 & 1 & 0\\0 & -1 & 1\end{bmatrix}$

令$|\lambda E-A|=\begin{vmatrix}\lambda-1 & -1 & 0\\0 & \lambda-1 & 0\\0 & 1 & \lambda-1\end{vmatrix}=(\lambda-1)^3=0$得A的特征值为$\lambda_1=\lambda_2=\lambda_3=1$，故$\rho(A)=1$，若J是A的 Jordan 标准形，由于$A^k=PJ^kP^{-1}$，所以只需判别$J^k$的敛散性即可。当$\lambda_i=1$时，由$\lambda_iE-A=\begin{bmatrix}0 & -1 & 0\\0 & 0 & 0\\0 & 1 & 0\end{bmatrix}\to\begin{bmatrix}0 & 1 & 0\\0 & 0 & 0\\0 & 0 & 0\end{bmatrix}$得$\dim V_{\lambda_i}=3-\operatorname{rank}(\lambda_i-A)=2$，所以$J=\begin{bmatrix}1 & 0 & 0\\0 & 1 & 1\\0 & 0 & 1\end{bmatrix}$，$J^k=\begin{bmatrix}1 & 0 & 0\\0 & 1 & k\\0 & 0 & 1\end{bmatrix}$且$\lim\limits_{k\to\infty}k=\infty$，故$A^k$发散。

(6) $A=\begin{bmatrix}\frac{7}{3} & \frac{1}{3} & -\frac{4}{3}\\0 & 1 & 0\\-\frac{2}{3} & \frac{1}{3} & \frac{5}{3}\end{bmatrix}$

令$|\lambda E-A|=\begin{vmatrix}\lambda-\frac{7}{3} & -\frac{1}{3} & \frac{4}{3}\\0 & \lambda-1 & 0\\\frac{2}{3} & -\frac{1}{3} & \lambda-\frac{5}{3}\end{vmatrix}=(\lambda-1)^2(\lambda-3)=0$得A的特征值为$\lambda_1=\lambda_2=1$，$\lambda_3=3$，故$\rho(A)=3>1$，故$A^k$发散。

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

令 $|\lambda E-A|=\begin{vmatrix}\lambda-\frac{2}{3} & 0 & \frac{1}{3}\\ 0 & \lambda-1 & 0\\ \frac{1}{6} & 0 & \lambda-\frac{5}{6}\end{vmatrix}=(\lambda-1)^2\left(\lambda-\frac{1}{2}\right)=0$ 得 A 的特征值为 $\lambda_1=\lambda_2=1$，$\lambda_3=\frac{1}{2}$，故$\rho(A)=1$，所以需要判别J^k的敛散性。当$\lambda_1=\lambda_2=1$时，由$\lambda_1 E-A=\begin{bmatrix}\frac{1}{3} & 0 & \frac{1}{3}\\ 0 & 0 & 0\\ \frac{1}{6} & 0 & \frac{1}{6}\end{bmatrix}\to\begin{bmatrix}1 & 0 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$得$\dim V_{\lambda_1}=3-\operatorname{rank}(\lambda_1-A)=2$，所以$J=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & \frac{1}{2}\end{bmatrix}$，

$J^k=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & \left(\frac{1}{2}\right)^k\end{bmatrix}$且$\lim\limits_{k\to\infty}\left(\frac{1}{2}\right)^k=0$，故$A^k$收敛。

(8) $A=\begin{bmatrix}3 & 1 & -4\\ 0 & 1 & 0\\ -2 & 1 & 5\end{bmatrix}$

令 $|\lambda E-A|=\begin{vmatrix}\lambda-3 & -1 & 4\\ 0 & \lambda-1 & 0\\ 2 & -1 & \lambda-5\end{vmatrix}=(\lambda-1)^2(\lambda-7)=0$ 得 A 的特征值为 $\lambda_1=\lambda_2=1$，$\lambda_3=7$，故$\rho(A)=7>1$，故A^k发散。

10.5 矩阵幂级数

1、 设$A_k=\left(a_{ij}^{(k)}\right)\in C^{m\times n}$，若$m\times n$个常数项级数

$$\sum_{k=1}^{\infty}a_{ij}^{(k)}=a_{ij}^{(1)}+a_{ij}^{(2)}+\cdots+a_{ij}^{(k)}+\cdots(i=1,2,\cdots,m;\ j=1,2,\cdots,n)$$

都收敛时，称矩阵级数

$$\sum_{k=1}^{\infty}A_k=A_1+A_2+\cdots+A_k+\cdots$$

收敛。若常数项级数的和为a_{ij}，则矩阵级数的和为$A=(a_{ij})$。若mn个常数项级数都绝对收敛，则称矩阵级数绝对收敛。

2、 设$A_k=\left(a_{ij}^{(k)}\right)\in C^{m\times n}$，则矩阵级数$\sum_{k=1}^{\infty}A_k$绝对收敛的充要条件是正项数项级数$\sum_{k=1}^{\infty}\|A_k\|$收敛，其中$\|A\|$为任何一种矩阵范数。

3、 常用级数的敛散性：

(1) p 级数：$\sum_{n=1}^{\infty}\frac{1}{n^p}$

当$p>1$时收敛，当$p\le 1$时发散。

(2) 等比级数：$\sum_{n=1}^{\infty}q^n$

当$q<1$时收敛，当$q\ge 1$时发散，其中q为正数。

(3) 交错级数：$\sum_{n=1}^{\infty}(-1)^{n-1}u_n, u_n>0$

莱布尼茨准测：若(1)$\{u_n\}$单调减少；(2)$\lim\limits_{n\to\infty} u_n = 0$，则$\sum_{n=1}^{\infty}(-1)^{n-1}u_n$收敛。

4、 设$A_k = (a_{ij}) \in C^{m\times n}$，称形如

$$\sum_{k=0}^{\infty} c_k A^k = c_0E + c_1A + c_2A^2 + \cdots + c_kA^k + \cdots$$

的矩阵级数为矩阵幂级数。

5、 若矩阵A的某一种范数$\|A\|$在幂级数$c_0 + c_1x + c_2x^2 + \cdots + c_kx^k + \cdots$的收敛域内，则矩阵幂级数$c_0E + c_1A + c_2A^2 + \cdots + c_kA^k + \cdots$绝对收敛。

补充：给定幂级数$c_0 + c_1x + c_2x^2 + \cdots + c_kx^k + \cdots$，如果$\lim\limits_{n\to\infty}\left|\frac{c_{n+1}}{c_n}\right| = \rho$，则该幂级数的收敛半径为$R = \begin{cases} \frac{1}{\rho}, 0 < \rho < +\infty \\ \infty, \rho = 0 \\ 0, \rho = +\infty \end{cases}$，开区间$(-R, R)$称为它的收敛区间，再考察$x = \pm R$时$c_0 + c_1x + c_2x^2 + \cdots + c_kx^k + \cdots$的收敛性，可得出该级数收敛点的全体，称之为收敛域。

6、 设幂级数$\sum_{k=0}^{\infty} c_kx^k$的收敛半径为R，A为n阶方阵。若$\rho(A) < R$，则矩阵幂级数$\sum_{k=0}^{\infty} c_kA^k$绝对收敛；若$\rho(A) > R$，则矩阵幂级数$\sum_{k=0}^{\infty} c_kA^k$发散。

7、 矩阵幂级数$E + A + A^2 + \cdots + A^k + \cdots$绝对收敛的充要条件是$\rho(A) < 1$，且其和为$(E - A)^{-1}$。

【例题】

1、 设$A = \begin{bmatrix} 0.2 & 0.5 & 0.1 \\ 0.1 & 0.5 & 0.3 \\ 0.2 & 0.4 & 0.2 \end{bmatrix}$，判断矩阵幂级数$E + A + A^2 + \cdots + A^k + \cdots$是否绝对收敛。

因为幂级数$1 + x + x^2 + \cdots + x^k + \cdots$的收敛半径为 1，而$\|A\|_\infty = 0.9 < 1$，故矩阵幂级数$E + A + A^2 + \cdots + A^k + \cdots$绝对收敛。

2、 判断下列幂级数是否绝对收敛。

(1) $1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{k!}x^k + \cdots$

由$\rho = \lim\limits_{n\to\infty}\left|\frac{c_{n+1}}{c_n}\right| = \lim\limits_{n\to\infty}\left|\frac{\frac{1}{n+1!}}{\frac{1}{n!}}\right| = \lim\limits_{n\to\infty}\left|\frac{1}{n+1}\right| = 0$得收敛半径$R = \infty$，故绝对收敛。

(2) $1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots + (-1)^k\frac{1}{(2k)!}x^{2k} + \cdots$

由$\rho = \lim\limits_{n\to\infty}\left|\frac{c_{n+1}}{c_n}\right| = \lim\limits_{n\to\infty}\left|\frac{(-1)^{(n+1)}\frac{1}{(2(n+1))!}}{(-1)^n\frac{1}{(2n)!}}\right| = \lim\limits_{n\to\infty}\left|\frac{1}{(2n+2)(2n+1)}\right| = 0$得收敛半径$R = \infty$，故绝对收敛。

(3) $x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots + (-1)^k\frac{1}{(2k+1)!}x^{2k+1} + \cdots$

由$\rho = \lim\limits_{n\to\infty}\left|\frac{c_{n+1}}{c_n}\right| = \lim\limits_{n\to\infty}\left|\frac{(-1)^{(n+1)}\frac{1}{(2(n+1)+1)!}}{(-1)^n\frac{1}{(2n+1)!}}\right| = \lim\limits_{n\to\infty}\left|\frac{1}{(2n+3)(2n+2)}\right| = 0$得收敛半径$R = \infty$，故绝对收敛。

3、 设$A=\begin{bmatrix}0.2&0.5&0.1\\0.1&0.5&0.3\\0.2&0.4&0.2\end{bmatrix}$，求$E+A+A^2+\cdots+A^k+\cdots$的和。

由$\|A\|_\infty=0.9$得$\rho(A)<1$，故$E+A+A^2+\cdots+A^k+\cdots$的和为$(E-A)^{-1}$。

由$E-A=\begin{bmatrix}0.8&-0.5&-0.1\\-0.1&0.5&-0.3\\-0.2&-0.4&0.8\end{bmatrix}$得$E+A+A^2+\cdots+A^k+\cdots=(E-A)^{-1}=\begin{bmatrix}2&\frac{22}{7}&\frac{10}{7}\\1&\frac{31}{7}&\frac{25}{14}\\1&3&\frac{5}{2}\end{bmatrix}$

4、 已知$A=\begin{bmatrix}-2&1\\-1&0\end{bmatrix}$，分别讨论矩阵幂级数$\sum_{k=1}^{\infty}\frac{1}{k}A^k$与$\sum_{k=1}^{\infty}\frac{1}{k^2}A^k$的敛散性。

令$|\lambda E-A|=\begin{vmatrix}\lambda+2&-1\\1&\lambda\end{vmatrix}=(\lambda+1)^2=0$得A的特征值为$\lambda_1=\lambda_2=-1$，故$\rho(A)=1$。

当$\lambda_i=-1$时，由$\lambda_i E-A=\begin{bmatrix}1&-1\\1&-1\end{bmatrix}\to\begin{bmatrix}1&-1\\0&0\end{bmatrix}$得$\dim V_{\lambda_i}=2-\mathrm{rank}(\lambda_i-A)=1$，所以存在可逆矩阵P使得$P^{-1}AP=J=\begin{bmatrix}-1&1\\0&-1\end{bmatrix}$，$P^{-1}A^kP=J^k=\begin{bmatrix}(-1)^k&k(-1)^{k-1}\\0&(-1)^k\end{bmatrix}$

(1) 矩阵幂级数$\sum_{k=1}^{\infty}\frac{1}{k}A^k$

由$\rho=\lim\limits_{n\to\infty}\left|\frac{c_{n+1}}{c_n}\right|=\lim\limits_{n\to\infty}\left|\frac{\frac{1}{n+1}}{\frac{1}{n}}\right|=\lim\limits_{n\to\infty}\left|\frac{n}{n+1}\right|=1$得收敛半径$R=1$，所以不能用定理来判断矩阵幂级数$\sum_{k=1}^{\infty}\frac{1}{k}A^k$的敛散性，只能用定义来验证其敛散性。

矩阵幂级数$\sum_{k=1}^{\infty}\frac{1}{k}A^k=P\left(\sum_{k=1}^{\infty}\frac{1}{k}J^k\right)P^{-1}=P\begin{bmatrix}\sum_{k=1}^{\infty}\frac{(-1)^k}{k}&\sum_{k=1}^{\infty}(-1)^{k-1}\\0&\sum_{k=1}^{\infty}\frac{(-1)^k}{k}\end{bmatrix}P^{-1}$

由于数项级数$\sum_{k=1}^{\infty}\frac{(-1)^k}{k}$收敛，$\sum_{k=1}^{\infty}(-1)^{k-1}$发散，故矩阵幂级数$\sum_{k=1}^{\infty}\frac{1}{k}A^k$发散。

(2) 矩阵幂级数$\sum_{k=1}^{\infty}\frac{1}{k^2}A^k$

由$\rho=\lim\limits_{n\to\infty}\left|\frac{c_{n+1}}{c_n}\right|=\lim\limits_{n\to\infty}\left|\frac{\frac{1}{(n+1)^2}}{\frac{1}{n^2}}\right|=\lim\limits_{n\to\infty}\left|\frac{n^2}{(n+1)^2}\right|=1$得收敛半径$R=1$，所以不能用定理来判断矩阵幂级数$\sum_{k=1}^{\infty}\frac{1}{k^2}A^k$的敛散性，只能用定义来验证其敛散性。

矩阵幂级数$\sum_{k=1}^{\infty}\frac{1}{k^2}A^k=P\left(\sum_{k=1}^{\infty}\frac{1}{k^2}J^k\right)P^{-1}=P\begin{bmatrix}\sum_{k=1}^{\infty}\frac{(-1)^k}{k^2}&\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\\0&\sum_{k=1}^{\infty}\frac{(-1)^k}{k^2}\end{bmatrix}P^{-1}$

由于数项级数$\sum_{k=1}^{\infty}\frac{(-1)^k}{k^2}$和$\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}$均收敛，故矩阵幂级数$\sum_{k=1}^{\infty}\frac{1}{k^2}A^k$收敛。

5、 判断下列矩阵幂级数的敛散性。

(1) $\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}1&1\\0&1\end{bmatrix}^k$

矩阵幂级数$\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}1&1\\0&1\end{bmatrix}^k=\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}1&k\\0&1\end{bmatrix}=\begin{bmatrix}\sum_{k=1}^{\infty}\frac{1}{k^2}&\sum_{k=1}^{\infty}\frac{1}{k}\\0&\sum_{k=1}^{\infty}\frac{1}{k^2}\end{bmatrix}$

由于数项级数$\sum_{k=1}^{\infty}\frac{1}{k^2}$收敛，$\sum_{k=1}^{\infty}\frac{1}{k}$发散，故此矩阵幂级数发散。

(2) $\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}-1 & 1\\ 0 & -1\end{bmatrix}^k$

矩阵幂级数$\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}-1 & 1\\ 0 & -1\end{bmatrix}^k=\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}(-1)^k & k(-1)^{k-1}\\ 0 & (-1)^k\end{bmatrix}=\begin{bmatrix}\sum_{k=1}^{\infty}\frac{(-1)^k}{k^2} & \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\\ 0 & \sum_{k=1}^{\infty}\frac{(-1)^k}{k^2}\end{bmatrix}$

由于数项级数$\sum_{k=1}^{\infty}\frac{(-1)^k}{k^2}$和$\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}$均收敛，故此矩阵幂级数收敛（不是绝对收敛）。

(3) $\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}-1 & 1 & 0\\ 0 & -1 & 1\\ 0 & 0 & -1\end{bmatrix}^k$

矩 阵 幂 级 数$\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}-1 & 1 & 0\\ 0 & -1 & 1\\ 0 & 0 & -1\end{bmatrix}^k=\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}(-1)^k & k(-1)^{k-1} & \frac{1}{2}k(k-1)(-1)^{k-2}\\ 0 & (-1)^k & k(-1)^{k-1}\\ 0 & 0 & (-1)^k\end{bmatrix}=$

$$\begin{bmatrix}\sum_{k=1}^{\infty}\frac{(-1)^k}{k^2} & \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k} & \frac{1}{2}\sum_{k=1}^{\infty}\frac{k-1}{k}(-1)^{k-2}\\ 0 & \sum_{k=1}^{\infty}\frac{(-1)^k}{k^2} & \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\\ 0 & 0 & \sum_{k=1}^{\infty}\frac{(-1)^k}{k^2}\end{bmatrix}$$

由于数项级数$\sum_{k=1}^{\infty}\frac{k-1}{k}(-1)^{k-2}$发散，故此矩阵幂级数发散。

(4) $\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}1 & 0 & 0\\ 0 & -1 & 1\\ 0 & 0 & -1\end{bmatrix}^k$

矩阵幂级数$\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}1 & 0 & 0\\ 0 & -1 & 1\\ 0 & 0 & -1\end{bmatrix}^k=\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}(-1)^k & 0 & 0\\ 0 & (-1)^k & k(-1)^{k-1}\\ 0 & 0 & (-1)^k\end{bmatrix}=$

$$\begin{bmatrix}\sum_{k=1}^{\infty}\frac{(-1)^k}{k^2} & 0 & 0\\ 0 & \sum_{k=1}^{\infty}\frac{(-1)^k}{k^2} & \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\\ 0 & 0 & \sum_{k=1}^{\infty}\frac{(-1)^k}{k^2}\end{bmatrix}$$

由于数项级数$\sum_{k=1}^{\infty}\frac{(-1)^k}{k^2}$和$\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}$均收敛，故此矩阵幂级数收敛（不是绝对收敛）。