矩阵理论

王磊

自动化科学与电气工程学院

第四章 矩阵分析

- 向量范数
- 矩阵范数
- 相容范数
- 矩阵扰动分析
- 特征值估计
- 矩阵级数
- 矩阵函数
- 函数矩阵
- 应用：主元分析法

第四章 矩阵分析

矩阵扰动分析（自学）

例4.4.1 考察线性方程组$A\boldsymbol{x}=\boldsymbol{b}$, 其中$x=[x_1,x_2]^T$,

$$A=\begin{bmatrix}1 & 0.99\\ 0.99 & 0.98\end{bmatrix},\qquad b=\begin{bmatrix}1\\ 1\end{bmatrix}$$

例4.4.1 考察线性方程组$A\boldsymbol{x}=\boldsymbol{b}$, 其中$x=[x_1,x_2]^T$,

$$A=\begin{bmatrix}1 & 0.99\\ 0.99 & 0.98\end{bmatrix},\qquad b=\begin{bmatrix}1\\1\end{bmatrix}$$

通过计算得, 该方程的精确解为

$$x_1=100,\qquad x_2=-100$$

例4.4.1 考察线性方程组$A\boldsymbol{x}=\boldsymbol{b}$, 其中$x=[x_1,x_2]^T$,

$$A=\begin{bmatrix}1 & 0.99\\0.99 & 0.98\end{bmatrix},\qquad b=\begin{bmatrix}1\\1\end{bmatrix}$$

通过计算得, 该方程的精确解为

$$x_1=100,\qquad x_2=-100$$

若系数矩阵A有一扰动$\begin{bmatrix}0 & 0\\0 & 0.01\end{bmatrix}$, 并且右端$\boldsymbol{b}$也有一扰动$\begin{bmatrix}0\\0.001\end{bmatrix}$, 则扰动后的线性方程组为

$$\begin{bmatrix} 1 & 0.99 \\ 0.99 & 0.99 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1.001 \end{bmatrix}$$

通过计算得, 该方程组的精确解为$x_1 = -\frac{1}{10}, x_2 = \frac{10}{9}$.

对比两组解向量知，系数矩阵和右端项的微小扰动会引起解的强烈变化．同时，注意例4.4.1本身并没有截断误差和舍入误差，因此原始数据的扰动对问题解的影响程度取决于问题本身的固有性质.

定义4.4.1（病态问题）病态问题是指输出结果相对于输入非常敏感的问题, 输入数据中哪怕是极少（或者极微妙）的扰动也会导致输出的较大改变. 相反的, 对于输入不敏感的问题, 我们称为良态问题.

矩阵扰动分析就是研究矩阵元素的变化对矩阵问题解的影响, 它对矩阵理论和矩阵计算都具有重要意义. 这里仅简要介绍矩阵A的逆矩阵、以A为系数矩阵的线性方程组的扰动分析.

为讨论方便，假定$\|\cdot\|$为$\mathbb{C}^{n\times n}$上满足$\|I\|=1$的相容矩阵范数. 我们需要解决:

（1）什么条件下$A+E$非奇异？

（2）当$A+E$非奇异时，$(A+E)^{-1}$与A^{-1}的近似程度.

定理4.4.1 设A和$E\in\mathbb{C}^{n\times n}$，$\mathrm{B}=A+E$. 若$A$与$B$均非奇异，则

$$\frac{\|\mathrm{B}^{-1}-A^{-1}\|}{\|A^{-1}\|}\le\|A\|\|\mathrm{B}^{-1}\|\frac{\|E\|}{\|A\|}$$

定理4.4.2 设$A \in \mathbb{C}^{n\times n}$是非奇异矩阵，$E \in \mathbb{C}^{n\times n}$，且满足条件$\|A^{-1}E\| < 1$，则$A+E$非奇异，并且有

$$\|(A+E)^{-1}\| \le \frac{\|A^{-1}\|}{1-\|A^{-1}E\|}$$

$$\frac{\|(A+E)^{-1}-A^{-1}\|}{\|A^{-1}\|} \le \frac{\|A^{-1}E\|}{1-\|A^{-1}E\|}$$

推论4.4.1设$A \in \mathbb{C}^{n\times n}$是非奇异矩阵, $E \in \mathbb{C}^{n\times n}$满足条件$\|A^{-1}E\| < 1$, 则$A+E$非奇异, 并且有

$$\frac{\|(A+E)^{-1} - A^{-1}\|}{\|A^{-1}\|} \le \frac{\kappa(A)\dfrac{\|E\|}{\|A\|}}{1-\kappa(A)\dfrac{\|E\|}{\|A\|}}$$

其中,$\kappa(A) = \|A\|\|A^{-1}\|$.

推论表明, $\kappa(A)$反映了A^{-1}对于A的扰动的敏感性$\kappa(A)$愈大, $(A+E)^{-1}$与A^{-1}的相对误差就愈大.

定义4.4.1（条件数） 设n阶矩阵A非奇异，称$\kappa(A)=\|A\|\|A^{-1}\|$ 为A关于求逆的条件数，也称为求解线性方程组$A\boldsymbol{x}=\boldsymbol{b}$的条件数.

注1: 条件数就是用来衡量输出相对于输入敏感度的指标. 良态问题和病态问题就是靠这个指标进行区分的: 如果$\kappa(A)$很大，则矩阵A关于求逆是病态的.

现考查线性方程组

$$Ax = \boldsymbol{b}$$

如果系数矩阵A和右端项$\boldsymbol{b}$分别有扰动项E和$\delta\boldsymbol{b}$, 则扰动后方程组为

$$(A+E)(\boldsymbol{x}+\delta\boldsymbol{x}) = \boldsymbol{b}+\delta\boldsymbol{b}$$

且有如下定理:

定理4.4.3 设$A \in \mathbb{C}^{n\times n}$是非奇异矩阵, $\boldsymbol{b}$和$\delta\boldsymbol{b} \in \mathbb{C}^{n}$, $\boldsymbol{x}$是原线性方程组的解. 如果$E \in \mathbb{C}^{n\times n}$满足条件$\|A^{-1}E\| < 1$, 则上述扰动后方程组有唯一解$\boldsymbol{x}+\delta\boldsymbol{x}$,

且有

$$\frac{\|\delta \boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \frac{\kappa(A)}{1-\kappa(A)\frac{\|E\|}{\|A\|}}\left(\frac{\|E\|}{\|A\|}+\frac{\|\delta \boldsymbol{b}\|}{\|b\|}\right)$$

该式表明, $\kappa(A)$反映了$A\boldsymbol{x}=\boldsymbol{b}$的解$\boldsymbol{x}$的相对误差对于$A$和$\boldsymbol{b}$的相对误差的依赖程度. 在求矩阵的逆或求解线性方程组时, 可以通过变换降低矩阵的条件数, 即所谓的预处理或预条件.

例4.4.2 考查线性方程组

$$\begin{bmatrix} 1 & 10^5 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 10^5 \\ 2 \end{bmatrix}$$

的求解问题.

注2: 条件数$\kappa(A)$大的两个常见原因: 矩阵的列之间的相关性过大（此时矩阵为奇异或接近奇异的, 见例4.4.1）和矩阵的特征值差异大（见例4.4.2）.

第四章 矩阵分析

特 征 值 估 计

定义4.5.1（谱和谱半径） 给定复方阵$A \in \mathbb{C}^{n\times n}$，记$S_p(A) = \{\lambda|\lambda$是$A$的特征值$\}$, 则称$S_p(A)$是矩阵$A$的谱, 称$A$的特征值模的最大值为$A$的谱半径, 记为$\rho(A)$.

例4.5.1 求$A = \begin{bmatrix} 1 & 0 \\ 2 & 2i \end{bmatrix}$的谱半径$\rho(A)$.

思考:矩阵的谱半径可定义矩阵范数吗?

例4.5.2 求$A=\begin{bmatrix}1 & 2\\ 2 & 2\end{bmatrix}$的谱半径$\rho(A)$, $\|A\|_1$, $\|A\|_\infty$, $\|A\|_2$和$\|A\|_F$.

定理4.5.1 复方阵的谱半径不大于它的任一矩阵范数.

定理4.5.2 设$A \in \mathbb{C}^{n\times n}$, 任取正常数$\epsilon$, 则必存在某个矩阵范数$\|\cdot\|$使得$\|A\| \le \rho(A) + \epsilon$.

定理4.5.1 复方阵的谱半径不大于它的任一矩阵范数.

定理4.5.2 设$A \in \mathbb{C}^{n\times n}$, 任取正常数$\epsilon$, 则必存在某个矩阵范数$\|\cdot\|$使得$\|A\| \le \rho(A) + \epsilon$.

注1: 定理4.5.2中构造的矩阵范数与给定的矩阵A有关. 因此, 当用矩阵A构造的矩阵范数来计算矩阵B（$B \neq A$）矩阵范数时, 不等式$\|B\| \le \rho(B) + \epsilon$不一定成立.

定义4.5.2（盖尔圆盘） 设$A=(a_{ij})\in\mathbb{C}^{n\times n}$, 令

$$\delta_i=\sum_{j=1,j\neq i}^{n}\left|a_{ij}\right|,i=1,\cdots,n$$

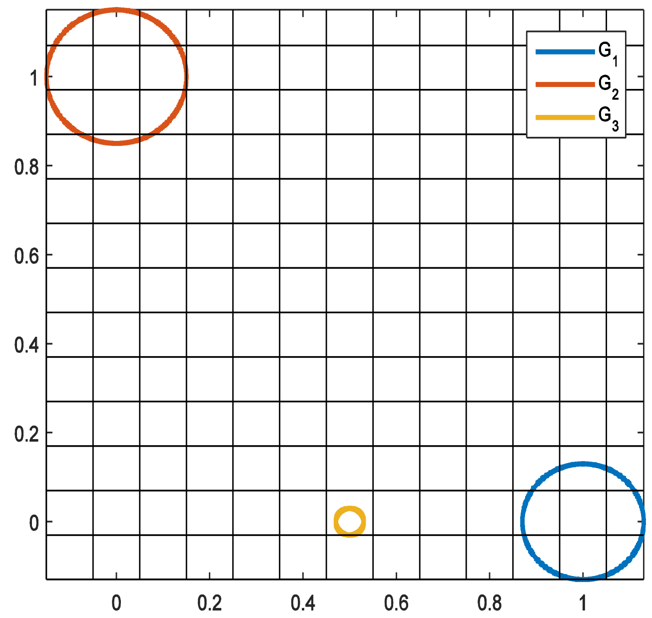
并定义$G_i=\{z\in C||z-a_{ii}|\leq\delta_i\}$, $i=1,\cdots,n$

即G_i是复平面上以a_{ii}为圆心, δ_i为半径的闭圆盘,

称之为矩阵A的一个盖尔圆盘.

例4.5.3 计算矩阵$A=\begin{bmatrix} 1 & 0.02 & 0.11 \\ 0.01 & \mathrm{i} & 0.14 \\ 0.02 & 0.01 & 0.5 \end{bmatrix}$的盖尔圆盘.

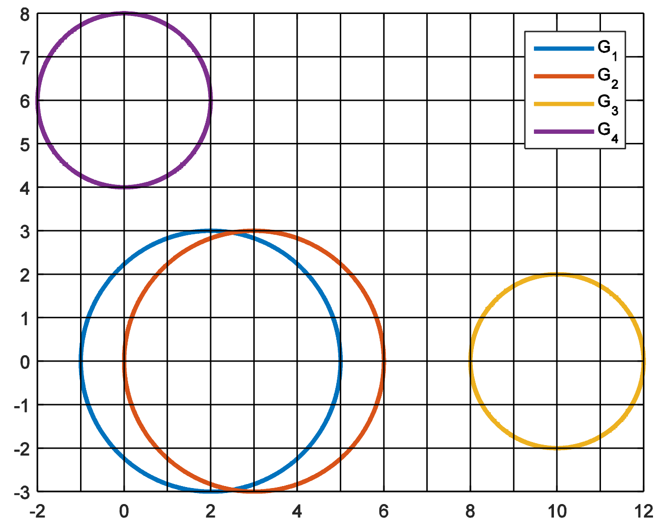
例4.5.3 计算矩阵$A=\begin{bmatrix}1 & 0.02 & 0.11\\ 0.01 & \mathrm{i} & 0.14\\ 0.02 & 0.01 & 0.5\end{bmatrix}$的盖尔圆盘.



定理4.5.3（盖尔圆盘定理） 设$A=(a_{ij})\in\mathbb{C}^{n\times n}$的$n$个盖尔圆盘$G_1,\cdots,G_n$，则矩阵$A$的任一特征值$\lambda\in\bigcup_{i=1}^{n}G_i$.

例4.5.4 试估计矩阵$A=\begin{bmatrix} 2 & -1 & -2 & 0 \\ -1 & 3 & 2\mathrm{i} & 0 \\ 0 & -\mathrm{i} & 10 & \mathrm{i} \\ -2 & 0 & 0 & 6\mathrm{i} \end{bmatrix}$的特征值分布.

例4.5.4 试估计矩阵$A=\begin{bmatrix}2 & -1 & -2 & 0\\ -1 & 3 & 2i & 0\\ 0 & -i & 10 & i\\ -2 & 0 & 0 & 6i\end{bmatrix}$的特征值分布.



A^T的4个盖尔圆为:

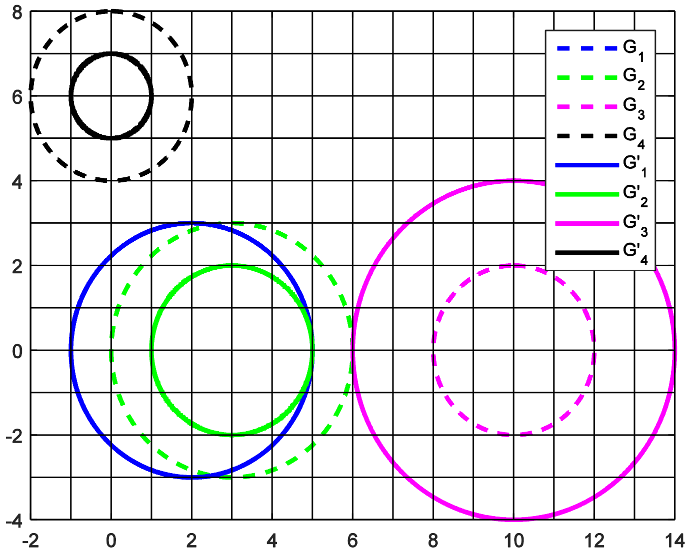
G_1': $|z-2| \le 3$;

G_2': $|z-3| \le 2$;

G_3': $|z-10| \le 4$;

G_4': $|z-6i| \le 1$.

A的4个特征值都在$\bigcup_{i=1}^{4} G_i'$中.



注2: 设A^T的盖尔圆为$G_1', \cdots, G_n'$, 则G_i与G_i'有相同的圆心. 因此, 矩阵A的特征值必须满足

$$\lambda \in \left(\bigcup_{i=1}^{n} G_i\right) \cap \left(\bigcup_{i=1}^{n} G_i'\right)$$

思考: n阶复方阵A的n个特征值是否恰好好落入它的n个盖尔圆内?

思考: n阶复方阵A的n个特征值是否恰好好落入它的n个盖尔圆内?

例4.5.5 考查矩阵$A=\begin{bmatrix} 0 & -0.4 \\ 0.9 & 1 \end{bmatrix}$, 其盖尔圆为

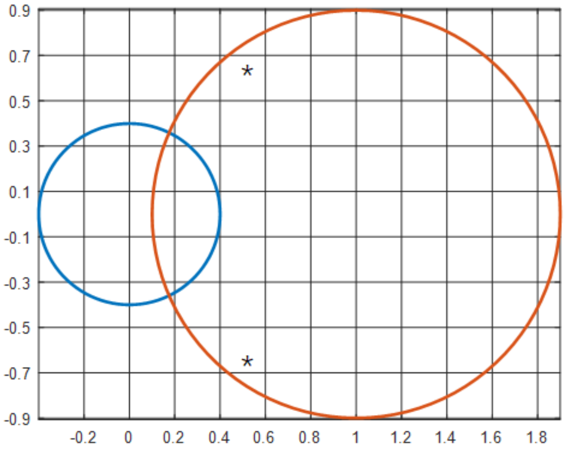
$$G_1: |z| \le 0.4, \qquad G_2: |z-1| \le 0.9$$

例4.5.5 考查矩阵$A=\begin{bmatrix}0 & -0.4\\ 0.9 & 1\end{bmatrix}$, 其盖尔圆为

$$G_1: |z| \le 0.4, \qquad G_2: |z-1| \le 0.9$$

经解析计算矩阵A的特征值为$\lambda_{1,2}=\frac{1\pm\sqrt{0.44i}}{2}$.

所以, 并非每个盖尔圆内都恰好有一个特征值.



定理4.5.4（盖尔圆盘定理续） 设$A=(a_{ij})\in\mathbb{C}^{n\times n}$的$n$个盖尔圆盘为$G_1,\cdots,G_n$, 若其中的$k$个盖尔圆盘的并集形成一个连通的区域, 且该区域与其余$n-k$个圆盘都不相交，则此连通域内恰有k个特征值. 特别地, 孤立盖尔圆内有且仅有一个特征值.

推论4.5.1 设矩阵$A=(a_{ij})\in\mathbb{C}^{n\times n}$的$n$个盖尔圆盘为$G_1,\cdots,G_n$, 若原点$0\notin\bigcup_{i=1}^{n}G_i$, 则矩阵$A$为非奇异矩阵.

推论4.5.2 设矩阵$A=(a_{ij})\in\mathbb{C}^{n\times n}$是对角占优矩阵，对即$i=1,\cdots,n$, 有

$$|a_{ii}|>\sum_{i=1,j\neq i}^{n}|a_{ij}|,$$（列对角占优）

或$$|a_{ii}|>\sum_{j=1,j\neq i}^{n}|a_{ij}|,$$（行对角占优）

则矩阵A非奇异.

推论4.5.3 若复方阵A有k个孤立的盖尔圆, 则它至少有k个互异的特征值. 特别地, 若矩阵A的所有盖尔圆两两互不相交, 则A是单纯矩阵.

推论4.5.4 若实方阵A有k个孤立的盖尔圆, 则它至少有k个互异的实特征值. 特别地,若矩阵A的所有盖尔圆两两互不相交,则它有至少n个互异的是特征值.

例4.5.6 证明n阶矩阵A是单纯矩阵, 其中

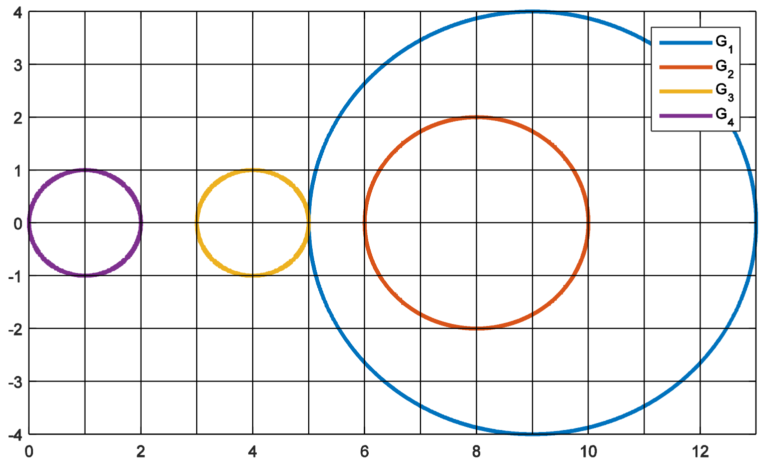
$$A=\begin{bmatrix} 2 & \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \frac{1}{n} & 4 & \frac{1}{n} & \cdots & \frac{1}{n} \\ & \cdots & & \cdots & \\ \frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \cdots & 2n \end{bmatrix}$$

例4.5.7 证明矩阵A至少有两个特征值, 其中

$$A=\begin{bmatrix}9 & 1 & -2 & 1\\ 0 & 8 & 1 & 1\\ -1 & 0 & 4 & 0\\ 1 & 0 & 0 & 1\end{bmatrix}$$

例4.5.7 证明矩阵A至少有两个特征值，其中

$$A=\begin{bmatrix} 9 & 1 & -2 & 1 \\ 0 & 8 & 1 & 1 \\ -1 & 0 & 4 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$



注3: 在使用盖尔圆估计矩阵A的特征值时, 我们总希望获得更多的孤立圆, 这时可采取如下方法:

取合适的非零实数$d_1,\cdots,d_n$, 并令

$$D=\operatorname{diag}(d_1,\cdots,d_n)$$

则

$$B=DAD^{-1}=\left(a_{ij}\frac{d_i}{d_j}\right)_{n\times n}$$

显然, 矩阵A与B相似,它们具有相同的特征值.

注3（续）：

通常d_i的选取办法为:

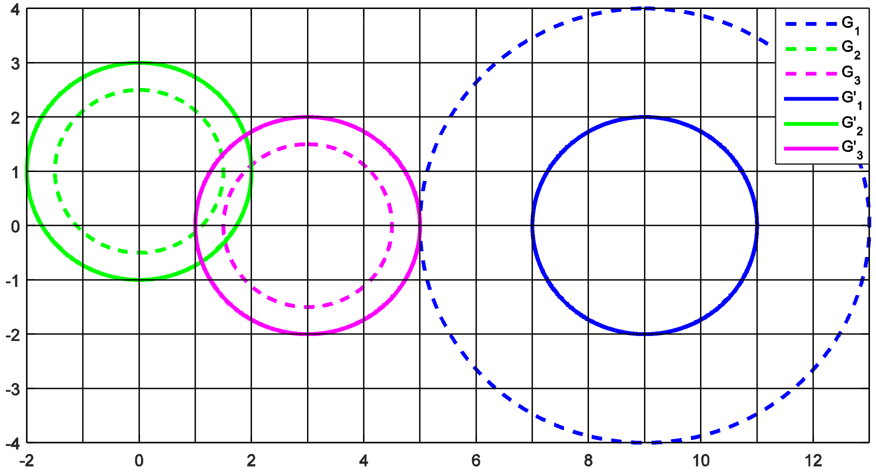
（1）若取$d_i < 1$, 其余元素为1, 则第i个盖尔圆G_i会缩小,其余所有盖尔圆会放大;

（2）若取$d_i > 1$, 其余元素为1, 则第i个盖尔圆G_i会放大,其余所有盖尔圆会缩小;

例4.5.8 估计复方阵$A=\begin{bmatrix}9 & 1 & 1\\ 1 & \mathrm{i} & 1\\ 1 & 1 & 3\end{bmatrix}$的特征值分布范围.

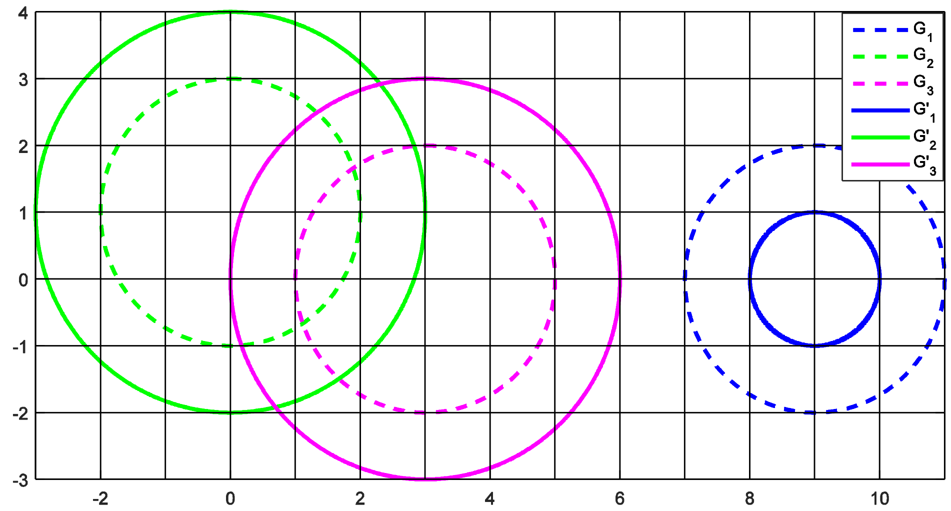
例4.5.8 估计复方阵$A=\begin{bmatrix}9 & 1 & 1\\1 & i & 1\\1 & 1 & 3\end{bmatrix}$的特征值分布范围.

$$D=\begin{bmatrix}2 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$$



例4.5.8 估计复方阵$A=\begin{bmatrix}9&1&1\\1&\mathrm{i}&1\\1&1&3\end{bmatrix}$的特征值分布范围.

$$D=\begin{bmatrix}0.5&0&0\\0&1&0\\0&0&1\end{bmatrix}$$



例4.5.9 利用盖尔圆定理考察矩阵$A=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$的特征值分布.

例4.5.9 利用盖尔圆定理考察矩阵$A=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$的特征值分布.

取$D=\begin{bmatrix}k & 0\\ 0 & 1\end{bmatrix}$，则$B=\begin{bmatrix}0 & -k\\ 1/k & 0\end{bmatrix}$

第四章 矩阵分析

矩 阵 级 数

定义4.6.1（向量序列按范数收敛） 设$(V, \|\cdot\|_\alpha)$是n维赋范线性空间, $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k, \cdots$是$V$中一个向量序列, 记为$\{\boldsymbol{x}_k\}$. 若存在$V$的向量$\boldsymbol{x}$满足

$$\lim_{k\to\infty} \|\boldsymbol{x}_k - \boldsymbol{x}\|_\alpha = 0$$

则称向量序列$\{\boldsymbol{x}_k\}$按范数$\|\cdot\|_\alpha$收敛于$\boldsymbol{x}$, 记作

$$\lim_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x} \text{ 或 } \boldsymbol{x}_k \xrightarrow{\alpha} \boldsymbol{x}$$

不收敛的向量称为是发散的.

例4.6.1 设$V=\mathbb{R}$, $\|\cdot\|_\alpha$取为绝对值, 则定义4.6.1中敛散性与高等数学中的相关概念一致.

例4.6.1 设$V=\mathbb{R}$, $\|\cdot\|_\alpha$取为绝对值, 则定义4.6.1中敛散性与高等数学中的相关概念一致.

定理4.6.1 设$(V,\|\cdot\|)$是n维赋范线性空间, $\{\boldsymbol{x}_k\}$是V的一个向量序列. 若序列$\{\boldsymbol{x}_k\}$按某种范数收敛于$\boldsymbol{x}$, 则序列$\{\boldsymbol{x}_k\}$按任意范数收敛于$\boldsymbol{x}$, 即有限维空间中按范数收敛是等价的.

定义4.6.2（向量序列按坐标收敛）设$(V,\|\cdot\|_\alpha)$是n维赋范线性空间, $\boldsymbol{\epsilon}_1,\cdots,\boldsymbol{\epsilon}_n$是$V$中一组基, $\{\boldsymbol{x}_k\}$是V中一个向量序列, 并记向量序列$\{\boldsymbol{x}_k\}$中的任一向量$\boldsymbol{x}_k$在$\boldsymbol{\epsilon}_1,\cdots,\boldsymbol{\epsilon}_n$下坐标为

$$\xi_k=\left[\xi_1^{(k)},\cdots,\xi_n^{(k)}\right]^T\in F^n$$

若存在V的向量$\boldsymbol{x}$满足$\lim\limits_{k\to\infty}\xi_i^{(k)}=\xi_i,\ i=1,\cdots,n$

则称向量序列$\{\boldsymbol{x}_k\}$按坐标收敛于向量$\boldsymbol{x}$, 其中ξ是向量$\boldsymbol{x}$在基$\boldsymbol{\epsilon}_1,\cdots,\boldsymbol{\epsilon}_n$下坐标.

定理4.6.2 设$(V,\|\cdot\|)$是n维赋范线性空间，$\{\boldsymbol{x}_k\}$是V的一个向量序列且$\boldsymbol{x}\in V$. 向量序列$\{\boldsymbol{x}_k\}$按范数收敛于向量$\boldsymbol{x}$当且仅当它按坐标收敛于$\boldsymbol{x}$.

定义4.6.3（矩阵序列按坐标收敛）设矩阵序列$\{A_k\}$，其中矩阵$A_k = (a_{ij}{}^{(k)}) \in \mathbb{C}^{m\times n}$, $k = 1,2,\cdots$, 如果当$k \to \infty$时, 矩阵A_k的每一个元素$a_{ij}{}^{(k)}$都有极限$a_{ij}{}^{(0)}$，即$\lim\limits_{k\to\infty} a_{ij}{}^{(k)} = a_{ij}{}^{(0)}$, $1 \le i \le m$, $1 \le j \le n$

则称矩阵序列按元素收敛或按坐标收敛（或简称为矩阵序列$\{A_k\}$收敛），$A_0 = a_{ij}{}^{(0)}$称为序列的极限，记为$\lim\limits_{k\to\infty} A_k = A_0$.

例4.6.2 设$A_k = \begin{bmatrix} \frac{2k^2+k+1}{k^2} & \frac{\sin k}{k} \\ e^{-k}\sin k & (1+\frac{1}{2k})^k \end{bmatrix}$，求极限 $\lim\limits_{k\to\infty} A_k$.

命题4.6.1 设矩阵序列$\{A_k\}$和$\{B_k\}$分别收敛于矩阵A和B, 则对任意$c_1, c_2 \in \mathbb{C}$, 有

（1）$\lim\limits_{k\to\infty}(c_1 A_k + c_2 B_k) = c_1 A + c_2 B$, 其中$A_k$, $B_k \in \mathbb{C}^{m\times n}$;

（2）$\lim\limits_{k\to\infty}(A_k B_k) = AB$, 其中$A_k \in \mathbb{C}^{m\times n}$, $B_k \in \mathbb{C}^{n\times p}$;

（3）若A_k和A为可逆矩阵, 则$\lim\limits_{k\to\infty} A_k^{-1} = A^{-1}$.

推论4.6.1 设$\|\cdot\|$是$\mathbb{C}^{m\times n}$上任一矩阵范数，$\mathbb{C}^{m\times n}$中矩阵序列$\{A_k\}$收敛于矩阵A_0的充分必要条件是

$$\lim_{k\to\infty}\|A_k - A_0\| = 0.$$

推论4.6.2 复方阵A的某一范数满足$\|A\| < 1$, 则$\lim_{k\to\infty} A^k = 0$.

思考: 若$\lim_{k\to\infty} A^k = 0$, 则$\|A\| < 1$成立吗?

推论4.6.2 复方阵A的某一范数满足$\|A\| < 1$, 则 $\lim\limits_{k\to\infty} A^k = 0$.

思考:若$\lim\limits_{k\to\infty} A^k = 0$, 则$\|A\| < 1$成立吗?

$$A = \begin{bmatrix} 0.1 & 0 \\ 1 & 0.1 \end{bmatrix} \Longrightarrow A^k = \begin{bmatrix} 0.1^k & 0 \\ k0.1^{k-1} & 0.1^k \end{bmatrix}$$

定理4.6.3 设$A=(a_{ij})\in\mathbb{C}^{n\times n}$, $\lim\limits_{k\to\infty}A^k=0$的充分必要条件是$\rho(A)<1$.

定理4.6.3 设$A=(a_{ij})\in\mathbb{C}^{n\times n}$, $\lim\limits_{k\to\infty}A^k=0$的充分必要条件是$\rho(A)<1$.

$$J_i^k=\begin{bmatrix}\lambda_i^k & C_k^1\lambda_i^{k-1} & C_k^2\lambda_i^{k-2} & \cdots & C_k^{n_i-1}\lambda_i^{k-n_i+1}\\ & \lambda_i^k & C_k^1\lambda_i^{k-1} & \cdots & C_k^{n_i-2}\lambda_i^{k-n_i+2}\\ & & \ddots & \ddots & \vdots\\ & & & \lambda_i^k & C_k^1\lambda_i^{k-1}\\ & & & & \lambda_i^k\end{bmatrix}$$

例4.6.3 设$A=\begin{bmatrix} c & 2c \\ 3c & 4c \end{bmatrix}$, 其中$c$为实数. 求$\lim\limits_{k\to\infty} A^k=0$的充分必要条件.

推论4.6.3 若复方阵A的某一矩阵范数满足$\|A\|<1$，则矩阵序列$\{A_k\}$收敛于零矩阵.

定义4.6.4（矩阵级数） 设矩阵序列$\{A_k \in \mathbb{C}^{m\times n}\}$，称$\sum_{k=1}^{\infty} A_k$为**矩阵级数**. 令$S_N = \sum_{k=1}^{N} A_k$，称$S_N$为**矩阵级数的部分和**. 若矩阵序列$\{S_N\}$收敛且有极限$S$，即$\lim\limits_{N\to\infty} S_N = S$，则称矩阵级数$\sum_{k=1}^{\infty} A_k$**收敛且有和$S$**. 不收敛的矩阵级数称为**发散级数**.

例4.6.4 设$A_k=\begin{bmatrix}\frac{1}{k(k+1)} & 0 \\ 0 & \frac{2k}{3^k}\end{bmatrix}$, $k=1,2,\cdots$, 求级数 $\sum_{k=1}^{\infty} A_k$.

例4.6.4 设$A_k=\begin{bmatrix}\frac{1}{k(k+1)} & 0\\ 0 & \frac{2k}{3^k}\end{bmatrix}$, $k=1,2,\cdots$, 求级数$\sum_{k=1}^{\infty}A_k$.

$$\sum_{k=1}^{\infty}\frac{1}{k(k+1)}=1$$

当$|x|<1$时,

$$1+x+x^2+\cdots+x^k+\cdots=\frac{1}{1-x}$$

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

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

$$\sum_{k=1}^{\infty}\frac{2k}{3^k}=2\sum_{k=1}^{\infty}k\left(\frac{1}{3}\right)^k=\frac{3}{2}$$

命题4.6.2 设$A_k \in \mathbb{C}^{m\times n}$, 则以下命题成立:

（1）$\sum_{k=1}^{\infty} A_k$ 收敛 $\Leftrightarrow$ mn个数值级数$\sum_{k=1}^{\infty} a_{ij}^{(k)}$ 收敛;

（2） $\sum_{k=1}^{\infty} A_k$ 发散 $\Leftrightarrow$ mn个数值级数$\sum_{k=1}^{\infty} a_{ij}^{(k)}$ 中至少有一个发散;

（3） $\sum_{k=1}^{\infty} A_k$ 收敛 $\Rightarrow \lim\limits_{k\to\infty} A_k = 0.$

定义4.6.5（矩阵级数绝对收敛） 设$A_k \in \mathbb{C}^{m\times n}$, 若矩阵级数$\sum_{k=1}^{\infty} A_k$所对应的$mn$个数值级数$\sum_{k=1}^{\infty} a_{ij}^{(k)}$, $1 \le i \le m, 1 \le j \le n$ 均绝对收敛, 则称$\sum_{k=1}^{\infty} A_k$绝对收敛.

命题4.6.3 设$A_k \in \mathbb{C}^{m\times n}$, 则以下命题成立:

（1）若$\sum_{k=1}^{\infty} A_k$绝对收敛, 则$\sum_{k=1}^{\infty} A_k$收敛. 但$\sum_{k=1}^{\infty} A_k$收敛并不蕴含$\sum_{k=1}^{\infty} A_k$绝对收敛;

（2）若$\sum_{k=1}^{\infty} A_k$绝对收敛于S, 对$\sum_{k=1}^{\infty} A_k$任意重组重排得矩阵级数$\sum_{k=1}^{\infty} B_k$, 则$\sum_{k=1}^{\infty} B_k$绝对收敛于S;

（3）对任意常矩阵$P \in \mathbb{C}^{p\times m}$和$Q \in \mathbb{C}^{n\times q}$, 若矩阵级数$\sum_{k=1}^{\infty} A_k$（绝对）收敛, 则矩阵级数$\sum_{k=1}^{\infty} PA_kQ$（绝对）收敛. 反之则不然.

例4.6.5 考查矩阵级数$\sum_{k=1}^{\infty} A_k$的敛散性，其中

$$A_k = \begin{bmatrix} (-1)^{k+1}\dfrac{1}{k} & 0 \\ 0 & 0 \end{bmatrix}$$

定理4.6.4 设$A_k=\left({a_{ij}}^{(k)}\right)\in\mathbb{C}^{m\times n}$, 矩阵级数$\sum_{k=1}^{\infty}A_k$绝对收敛当且仅当对任意矩阵范数, 数值级数$\sum_{k=1}^{\infty}\|A_k\|$收敛.

例4.6.6 考查矩阵级数$\sum_{k=1}^{\infty} A_k$的敛散性，其中

$$A_k = \begin{bmatrix} \frac{1}{3^k} & 0 \\ 0 & \frac{1}{k(k+1)} \end{bmatrix}$$

定义4.6.6（矩阵幂级数）设$A \in \mathbb{C}^{n\times n}$, 定义矩阵级数$\sum_{m=0}^{\infty} c_m A^m$, 其中$A^0 = I$, 则称该级数为**矩阵幂级数**.

引理4.6.1 设$z \in \mathbb{C}$, 若幂级数$\sum_{m=0}^{\infty} c_m z^m$在$z = z_0$点收敛, 则对满足不等式$|z| < |z_0|$的幂级数都绝对收敛. 反之, 若幂级数$\sum_{m=0}^{\infty} c_m z^m$在$z = z_0$点发散, 则对于$|z| > |z_0|$的幂级数都发散.

若存在非负实数或无穷大数R满足$|z| < r$时,幂级数$\sum_{m=0}^{\infty} c_m z^m$收敛; 而$|z| > r$时,幂级数$\sum_{m=0}^{\infty} c_m z^m$都发散, 则称$r$为收敛半径.

定理4.6.5（Abel型定理） 设复变量幂级数 $\sum_{m=0}^{\infty} c_m z^m$ 的收敛半径为 r, 矩阵 $A \in \mathbb{C}^{n\times n}$ 的谱半径为 $\rho(A)$, 则

（1）当 $\rho(A) < r$, $\sum_{m=0}^{\infty} c_m A^m$ 绝对收敛;

（2）当 $\rho(A) > r$, $\sum_{m=0}^{\infty} c_m A^m$ 发散.

定理4.6.5（Abel型定理） 设复变量幂级数$\sum_{m=0}^{\infty} c_m z^m$的收敛半径为$r$, 矩阵$A \in \mathbb{C}^{n\times n}$的谱半径为$\rho(A)$, 则

（1）当$\rho(A) < r$, $\sum_{m=0}^{\infty} c_m A^m$绝对收敛;

（2）当$\rho(A) > r$, $\sum_{m=0}^{\infty} c_m A^m$发散.

注1: 当$\rho(A) = r$时, 矩阵幂级数$\sum_{m=0}^{\infty} c_m A^m$的敛散性无法由Abel型定理判断, 只能根据矩阵级数的定义进行判断.

例4.6.7 设$A=\begin{bmatrix}-1 & 1\\ 0 & -1\end{bmatrix}$, 判断$\sum_{k=1}^{\infty}\frac{1}{k}A^k$的敛散性.

推论4.6.4 若幂级数$\sum_{m=0}^{\infty} c_m z^m$在整个复平面上都收敛, 则对任意复方阵$A$, 有幂级数$\sum_{m=0}^{\infty} c_m A^m$收敛.

例4.6.8 已知$e^z = \sum_{k=0}^{\infty} \frac{1}{k!} z^k$对任意复数$z$都收敛，则对任意复方阵$A$，矩阵幂级数$\sum_{k=0}^{\infty} \frac{1}{k!} A^k$收敛.

推论4.6.5（Neumann级数） 矩阵幂级数$\sum_{m=0}^{\infty} A^m$收敛当且仅当$\rho(A) < 1$, 此时

$$\sum_{m=0}^{\infty} A^m = (I - A)^{-1}.$$

例4.6.9 判断幂级数$\sum_{m=1}^{\infty} A^m$的敛散性, 其中

$$A = \begin{bmatrix} 0.2 & 0.5 & 0.2 \\ 0.1 & 0.5 & 0.3 \\ 0.1 & 0.4 & 0.2 \end{bmatrix}$$

注3: 可逆矩阵$(I-A)^{-1}$与矩阵A的乘积可交换,

$$(I-A)^{-1}A = A(I-A)^{-1}$$

这是基于幂级数的定义推导的. 实际上, 所有可展开成幂级数的矩阵都具有这一性质.

例如, 当$|z|<1$时, 有$\sum_{k=0}^{\infty} kz^k = \frac{z}{(1-z)^2}$. 相应的, 令$\rho(A)<1$, 则有矩阵幂级数

$$\sum_{k=0}^{\infty} kA^k = A(I-A)^{-2} = (I-A)^{-2}A$$

$$= (I-A)^{-1}A(I-A)^{-1}$$