第四章 矩阵特征值的估计与计算

§3.1 盖尔圆定理

1. 定理 4.1 设 $A=(a_{ij})_{n\times n}$, $a_{ij}\in C$, 则 A 的特征值有 $\lambda\in\bigcup_{i=1}^{n}D_i$。其中

$$D_i=\{z:|z-a_{ii}|\le\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|\}\quad(i=1,2,\cdots n)$$

为复平面上的圆盘。

1. 定理 4.1 设 $A=(a_{ij})_{n\times n}$, $a_{ij}\in C$, 则 A 的特征值有 $\lambda\in\bigcup_{i=1}^{n}D_i$。其中

$$D_i=\{z:|z-a_{ii}|\leq\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|\}\quad(i=1,2,\cdots n)$$

为复平面上的圆盘。

注： 定理中的圆 D_i 也称为矩阵 A 的第 i 个gershgorin圆盘(盖尔圆)。 $R_i=\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|$ 为盖尔圆 D_i 的半径。

证明： 设 λ 是 A 的任意特征值，x 为 A 对应于 λ 的特征向量， 且单位化后使其模为1，假定 i 是使 $|x_i| = 1$ 的一个指标。 由于 $(Ax)_i = \lambda x_i$, 即 $\lambda x_i = \sum_{j=1}^{n} a_{ij} x_j$. 则我们有

$$(\lambda - a_{ii})x_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij} x_j.$$

注意到 $|x_j| \leq |x_i| = 1 (j \neq i)$, 则由上式得

$$|\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}||x_j| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|.$$

于是 $\lambda \in D_i$。

注： A 的全体特征值也都在 A^T 的 n 个盖尔圆构成的并集中，称 A^T 的盖尔圆为 A 的列盖尔圆。

注： A 的全体特征值也都在 A^T 的 n 个盖尔圆构成的并集中，称 A^T 的盖尔圆为 A 的列盖尔圆。

2. 定理 4.2 设定理4.1中的 n 个圆盘中，有 m 个圆盘相交构成一个连通域 S，且 S 与其余 $n-m$ 个圆盘严格分离，则 S 中恰有 A 的 m 个特征值，其中重特征值按其重数计算。

例：估计矩阵

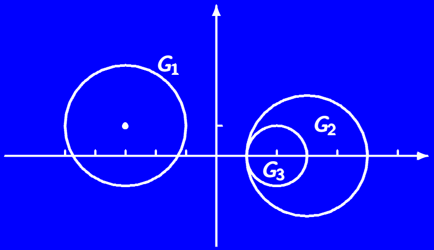
$$A = \begin{pmatrix} -3+i & 0 & 2 \\ -1 & 3 & i \\ 0 & -1 & 2 \end{pmatrix}$$

特征值范围。

解： 矩阵 A 的3个盖尔圆为

$$G_1 : |z+3-i| \leq 2, \qquad G_2 : |z-3| \leq 2, \qquad G_3 : |z-2| \leq 1,$$

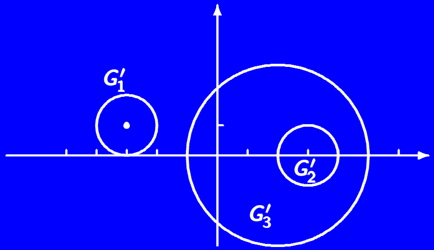
故 A 的3个特征值都在 $\bigcup_{i=1}^{3} G_i$ 之中。



矩阵 A 的3个列盖尔圆为

$G_1' : |z+3-i| \leq 1,\qquad G_2' : |z-3| \leq 1,\qquad G_3' : |z-2| \leq$

故 A 的3个特征值都在 $\bigcup_{i=1}^{3} G_i'$ 之中。



由定理4.2，矩阵 A 的特征值在 G_1 中有一个，在 $G_2 \cup G_3$ 中有两个；或者一个 在 G_1'中，两个在 $G_2' \cup G_3'$中。

注：　特征值在两个或两个以上的盖尔圆构成的连通部分中的分布不一定是平均的。

$$A = \begin{pmatrix} 10 & -8 \\ 5 & 0 \end{pmatrix}$$

其特征值为 $\lambda_1 = 5 - i\sqrt{15}, \lambda_2 = 5 + i\sqrt{15}$，两个盖尔圆为

$$G_1 : |z - 10| \leq 8 \quad G_2 : |z| \leq 5$$

A 的特征值都在 G_1 中。

§4.2 特征值的隔离

⋆ 结合矩阵 A 的列盖尔圆。

§4.2 特征值的隔离

⋆ 结合矩阵 A 的列盖尔圆。

例：应用盖尔圆定理隔离

$$A = \begin{pmatrix} 20 & 3 & 2 \\ 2 & 10 & 4 \\ 4 & 0.5 & 0 \end{pmatrix}$$

的特征值。

解： A 的三个盖尔圆为

$$G_1 : |z-20| \leq 5,\ G_2 : |z-10| \leq 6,\ G_3 : |z| \leq 4.5.$$

G_2, G_2, G_3 相交。

而 A 的三个列盖尔圆为

$$\overline{G}_1 : |z-20| \leq 6,\ \overline{G}_2 :\ |z-10| \leq 3.5,\ \overline{G}_3 : |z-6| \leq 6,$$

它们相互分离。故在 $\overline{G}_1$, $\overline{G}_2$, $\overline{G}_3$ 中各有 A 的一个特征值。

⋆ 利用相似变换。

⋆ 利用相似变换。

- 设 $A=(a_{ij})_{n\times n}$，
  $D=\mathrm{diag}(d_1,d_2,\cdots,d_n)$，$(d_i>0,\ i=1,2,\cdots,n)$。
  则
  $$A\sim B=D^{-1}AD,$$
  有相同特征值;

⋆ 利用相似变换。

- 设 $A = (a_{ij})_{n\times n}$，

  $D = \mathrm{diag}(d_1, d_2, \cdots, d_n)$，$(d_i > 0, i = 1, 2, \cdots, n)$。

  则

  $$A \sim B = D^{-1}AD,$$

  有相同特征值;

- B 的圆盘半径为 $\sum\limits_{\substack{j=1\\ j\neq i}}^{n} \left|\frac{a_{ij}d_j}{d_i}\right|$

- 适当选取 D 的元素值，可以使某个圆盘半径相对减少。一般地，

- 适当选取 D 的元素值，可以使某个圆盘半径相对减少。一般地，
  - 要使 A 的第 i 个盖尔圆半径放大，其余盖尔圆适量缩小（相对于 A 的同序号盖尔圆），我们可取 $d_i < 1$, 其它元素值取 1;
  - 相反，若使 A 的第 i 个盖尔圆半径缩小，其余盖尔圆适量放大（相对于 A 的同序号盖尔圆），我们可取 $d_i > 1$, 其它元素值取 1;

例： 应用盖尔圆定理隔离

$$A = \begin{pmatrix} 2 & 2 & -1 \\ 1 & 10 & -1 \\ 8 & 2 & 20 \end{pmatrix}$$

的特征值。

例： 应用盖尔圆定理隔离

$$A = \begin{pmatrix} 2 & 2 & -1 \\ 1 & 10 & -1 \\ 8 & 2 & 20 \end{pmatrix}$$

的特征值。

解： A 的三个盖尔圆为

$$G_1 : |z-2| \leq 3, \ G_2 : |z-10| \leq 2, \ G_3 : |z-20| \leq 10.$$

G_2 与 G_3 相交。

取 $D=\mathrm{diag}(\frac{1}{2},1,1)$，则

$$B=D^{-1}AD=\begin{pmatrix} 2 & 4 & -2 \\ 0.5 & 10 & -1 \\ 4 & 2 & 20 \end{pmatrix}.$$

则 B 的三个盖尔圆为

$$\overline{G}_1:|z-2|\le 6,\ \overline{G}_2:\ |z-10|\le 1.5,\ \overline{G}_3:|z-20|\le 6,$$

它们相互分离。若实矩阵 A 的特征值 λ 为复数，则 $\overline{\lambda}$ 也为特征值，但 $\overline{G}_1,\overline{G}_2,\overline{G}_3$ 关于实轴对称，则 A 特征值必皆为实数，故在区间 $[-1,5],[8.5,11.5],[14,26]$ 中各有 A 的一个特征值。

§4.3 幂迭代法与逆幂迭代法

一. 幂迭代法

§4.3 幂迭代法与逆幂迭代法

一. 幂迭代法

设矩阵 A 可对角化， 其特征值满足

$$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$$

且 $x_1, x_2, \cdots, x_n$ 为相应线性无关的特征向量。

§4.3 幂迭代法与逆幂迭代法

一. 幂迭代法

设矩阵 A 可对角化， 其特征值满足

$$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$$

且 $x_1, x_2, \cdots, x_n$ 为相应线性无关的特征向量。

A按模最大的特征值 λ_1 称为主特征值，对应 x_1 称为主特征向量。

算法思想

设 v 是 R^n 的任意元素，则有

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

得

$$A^k v = \lambda_1^k[\alpha_1 x_1 + \sum_{i=2}^{n} \alpha_i (\frac{\lambda_i}{\lambda_i})^k x_i].$$

- 当 $k \to \infty$，$(\frac{\lambda_i}{\lambda_1}) \to 0$。且当 $\alpha_1 \neq 0$ 时， $A^k v$ 除去一个数量因子外，趋向于向量 x_1。

- 当 $k \to \infty$，$(\frac{\lambda_i}{\lambda_1}) \to 0$。且当 $\alpha_1 \neq 0$ 时， $A^k v$ 除去一个数量因子外，趋向于向量 x_1。
- 当 $|\lambda_1| > 1$ 时，$|\lambda_1|^k$ 趋向于无穷；而 $|\lambda_1| < 1$ 时，$|\lambda_1|^k$ 趋向于 0。

- 当 $k \to \infty$，$(\frac{\lambda_i}{\lambda_1}) \to 0$。且当 $\alpha_1 \neq 0$ 时， $A^k v$ 除去一个数量因子外，趋向于向量 x_1。
- 当 $|\lambda_1| > 1$ 时，$|\lambda_1|^k$ 趋向于无穷；而 $|\lambda_1| < 1$ 时，$|\lambda_1|^k$ 趋向于 0。
- 迭代中加入规范化步骤。令 $x = (\xi_1, \xi_2, \cdots, \xi_n)^T \in R^n$, 记

$$\max(x) = \xi_i, \quad |\xi_i| = \|x\|_\infty.$$

迭代格式：选取向量 $v_0 \neq 0$, 做迭代

$$\begin{cases} u_k = Av_{k-1} \\ m_k = max(u_k) \quad k = 1, 2, \cdots \\ v_k = \frac{u_k}{m_k} \end{cases}$$

收敛分析：

- $v_k \to \frac{x_1}{max(x_1)}$.
- $m_k = \lambda_1[1 + O(|\frac{\lambda_2}{\lambda_1}|^k)] \to \lambda_1$.
- 收敛速度取决于 $|\frac{\lambda_2}{\lambda_1}|$

例: 设矩阵

$$A = \begin{pmatrix} 2 & 3 & 2 \\ 10 & 3 & 4 \\ 3 & 6 & 1 \end{pmatrix}$$

的按模最大的特征值及特征向量。当 $\|v_k - v_{k+1}\|_\infty \leq 10^{-4}$ 时停止计算。

取 $v_0 = (0,0,1)^T$， 计算结果见表

k	v_k^T			m_k
0	0	0	1	1
1	0.5	1.0	0.25	4
2	0.5	1.0	0.8611	9
3	0.5	1.0	0.7306	11.44
4	0.5	1.0	0.7535	10.9224
5	0.5	1.0	0.7493	11.0140
6	0.5	1.0	0.7501	10.9972
7	0.5	1.0	0.7500	11.0004
8	0.5	1.0	0.7500	11.0000

所以按模最大特征值为 $\lambda_1 = 11$, 特征向量为 $x_1 = (0.5, 1, 0.75)^T$.

二. 逆幂迭代法

⋆ 应用1：计算非奇异矩阵 A 按模最小特征值及相应特征向量。

二. 逆幂迭代法

⋆ 应用1：计算非奇异矩阵 A 按模最小特征值及相应特征向量。

- A 非奇异，且 $0 < |\lambda_n| < |\lambda_{n-1}| \leq \cdots |\lambda_2| \leq |\lambda_1|$，则 A 的按模最小特征值 λ_n 的倒数恰好为 A^{-1} 的按模最大特征值。
- 对 A^{-1} 应用乘幂法

算法结构:

$$\begin{cases} Au_k = v_{k-1} \\ m_k = max(u_k) \quad k = 1, 2, \cdots \\ v_k = \frac{u_k}{m_k} \end{cases}$$

算法结构:

$$\begin{cases} Au_k = v_{k-1} \\ m_k = max(u_k) \quad k = 1, 2, \cdots \\ v_k = \frac{u_k}{m_k} \end{cases}$$

- $m_k \to \frac{1}{\lambda_n}$, $v_k \to \frac{x_n}{max(x_n)}$。
- 将 A 进行 LU 分解，每次迭代求解两个三角形方程组。

⋆ 应用2： 已知矩阵 A 的特征值 λ_i 的近似值 $\tilde{\lambda}_i$， 对 $A-\tilde{\lambda}_i I$ 实行逆幂迭代法计算 λ_i。

$$\begin{cases} (A-\tilde{\lambda}_i I)u_k = v_{k-1} \\ m_k = max(u_k) \quad (k=1,2,\cdots) \\ v_k = \frac{u_k}{m_k} \end{cases}$$

- 收敛性：

$$v_k \to \frac{x_i}{\max(x_i)},$$

$$m_k \to \frac{1}{\lambda_i - \tilde{\lambda}_i} = \bar{\lambda},$$

得 $\lambda_i = \frac{1}{\bar{\lambda}} + \tilde{\lambda}_i$。

- $A-\tilde{\lambda}_i I$ 接近奇异，采用选列主元三角分解 $P(A-\tilde{\lambda}_i I)=LU$ 有必要。

改进：

$$\begin{cases} Ly_k = Pv_{k-1} \\ Uu_k = y_k \\ m_k = max(u_k) \quad (k=1,2,\cdots) \\ v_k = \frac{u_k}{m_k} \end{cases}$$

例：用逆幂法求矩阵 $A=\begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 0 \\ 1 & 1 & 4 \end{pmatrix}$ 最接近 3.5 的特征值和相应特征向量。初始向量设为 $(1,1,1)^T$，停机准则设为 $\|v_{k+1}-v_k\| \le 10^{-4}$。(精确值 $\frac{5+\sqrt{5}}{2}$)。

解:

$$A - 3.5\mathrm{I}_3 = \begin{pmatrix} -1.5 & 1 & 0 \\ 1 & -0.5 & 0 \\ 1 & 1 & 0.5 \end{pmatrix},$$

列主元三角分解 $P(A - 3.5\mathrm{I}_3) = LU$，其中

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

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -0.6667 & 1 & 0 \\ -0.6667 & 0.1 & 1 \end{pmatrix} U = \begin{pmatrix} -1.5 & 1 & 0 \\ 0 & 1.6667 & 0.5 \\ 0 & 0 & -0.05 \end{pmatrix}.$$

k	v_k^T			m_k
1	-0.2	-0.3333	1.0000	-30.0006
2	-0.1566	-0.2530	1.0000	11.0667
3	-0.1482	-0.2399	1.0000	8.9398
4	-0.1464	-0.2370	1.0000	8.5768
5	-0.1460	-0.2363	1.0000	8.4965
6	-0.1459	-0.2361	1.0000	8.4779
7	-0.1459	-0.2361	1.0000	8.4735

则

$$\lambda \approx = 3.5 + \frac{1}{8.4735} = 3.6180,$$

近似特征向量为

$$x \approx v_7 = (-0.1459, -0.2361, 1.0000)^T.$$