

第四章 特征值的估计与摄动

在学习线性代数的时候，我们了解过如何求一个矩阵的特征值，并得到与该特征值对应的特征向量，如，对一个 n 阶方阵 \mathbf{A} ，求特征值的方法是构建特征方程

$$|\lambda \mathbf{E} - \mathbf{A}| = 0$$

所得到的 λ_i 便是求解出的特征值。

这种构建特征方程求解特征值的方法求出来的解尽管精确，但有很大的局限性，如果在面对阶数较大的矩阵中，该方法需要庞大的计算量才可以计算出矩阵的特征值，同时，面对高阶行列式，对其进行处理，判断其解的情况也是一件需要耗费功夫的事情。

幸运的是，在工程中，很多情况下我们往往并不需要了解矩阵特征值精确解，而只需要了解该特征值的一个大概范围即可，这就是本章要讨论的问题——特征值的估计。

由于考试，本章只介绍 Gerschgorin(盖尔) 圆盘定理的知识，本章的其他内容如果有机会会在后面补上。

4.1 Gerschgorin(盖尔) 圆盘定理

本节所涉及的圆盘定理可以帮助确定一个矩阵的特征值的大致范围。

4.1.1 第一圆盘定理

行盖尔圆盘和列盖尔圆盘

在介绍第一圆盘定理前，需要了解何为行盖尔圆盘和列盖尔圆盘，这两个概念会在本节内容中反复提到，因此需要掌握。

定义 4.1.1: 行盖尔圆盘和列盖尔圆盘

设 $A \in \mathbb{C}^{n \times n}$ ，则

$$S_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq R_i\}$$

为矩阵 A 在复平面上的第 i 个行盖尔圆，其中

$$R_i = R_i(A) = \sum_{j \neq i} |a_{ij}| \quad (i = 1, 2, \dots, n)$$

叫做 S_i 的半径。

类似的

$$G_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq C_i\}, C_i = \sum_{i \neq j} |a_{ji}|$$

叫做列盖尔圆，半径为 C_i

可能一下子看到这个定义不是很理解，下面给出通俗的说法

行盖尔圆：以矩阵对角线元素为圆心，当前行其余元素的绝对值的和为半径的一个圆。

列盖尔圆：以矩阵对角线元素为圆心，当前列其余元素的绝对值的和为半径的一个圆。

有了这两个概念，才有下面的第一圆盘定理：

定理 4.1.1: 第一圆盘定理

设 $A \in \mathbb{C}^{n \times n}$ ，则 A 的任一特征值

$$\lambda_i \in S = \bigcup_{j=1}^n S_j \quad (i = 1, 2, \dots, n)$$

定理1.1.1表明：矩阵 \mathbf{A} 的任意特征值都在行盖尔圆中，证明过程就不写在这里了，感兴趣的可以翻阅教材。

下面通过一个例题具体感受一下第一圆盘定理的应用：

例 4.1.1. 估计矩阵

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

的特征值的分布范围 (i 为虚数单位)

解. 根据第一圆盘定理，可以画出如下的 4 个行盖尔圆：

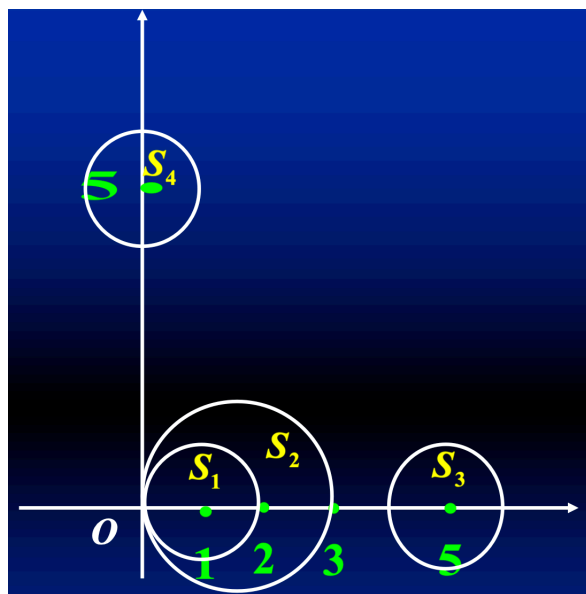
$$S_1 : |z - 1| \leq 1$$

$$S_2 : |z - \frac{3}{2}| \leq \frac{3}{2}$$

$$S_3 : |z - 5| \leq 1$$

$$S_4 : |z - 5i| \leq 1$$

将这四个盖尔圆画出来就如下图所示：



这就是如何画出盖尔圆，这个矩阵的特征值就会分布在盖尔圆内。

□

同样的，我们还有如下推论：

推论 4.1.2

设 $\mathbf{A} \in \mathbb{C}^{n \times n}$, 则 \mathbf{A} 的任一特征值

$$\lambda_i \in \bigcup_{j=1}^n G_i$$

也就是说, 矩阵的特征值同样分布在列盖尔圆内。

4.1.2 第二圆盘定理

第一圆盘定理只说明了矩阵 \mathbf{A} 的特征值都在盖尔圆中, 但并没有说哪个连通部分有哪些特征值, 要让特征值的分布更明确, 就需要第二圆盘定理。

定理 4.1.3: 第二圆盘定理

设 n 阶方阵 \mathbf{A} 的 n 个盖尔圆盘中有 k 个圆盘的并形成连通区域 \mathbf{G} , 且它与余下的 $n - k$ 个圆盘都不相交, 则在该区域 \mathbf{G} 中恰好有 \mathbf{A} 的 k 个特征值。

这里仅仅解释一下何为“连通”, 关于该定理的证明就不赘述了, 详细可以看 PPT 或者课本。

以例1.1.1画出的盖尔圆为例, S_1, S_2 重叠在一起, 它们的并集是一个连通区域, 交结在一起的盖尔圆所构成的连通区域为并集 S 的一个连通部分, 孤立的盖尔圆就是一个连通部分, 所以, S_1 与 S_2 的并集, 即 S_2, S_3, S_4 各是一个连通部分。

4.1.3 圆盘定理的其他推论

接下来介绍的圆盘定理的其他推论更有助于借助圆盘来判断矩阵特征值的情况:

推论 4.1.4

设 n 阶方阵 \mathbf{A} 的 n 个盖尔圆盘两两互不相交, 则 \mathbf{A} 相似于对角阵。

如果 n 个盖尔圆盘两两互不相交, 则说明每一个盖尔圆盘就包含一个矩阵的特征值, 这也就是说 n 个盖尔圆盘就会有 n 个互不相同的特征值, 由线性代数的知识可知, 矩阵 \mathbf{A} 可以相似对角化。

推论 4.1.5

设 n 阶实矩阵 \mathbf{A} 的 n 个盖尔圆盘两两互不相交, 则 \mathbf{A} 的特征值全为实数。

如果 \mathbf{A} 不是实数而是复数的话, 圆盘里面会有两个特征值, 这与第二圆盘定理相矛盾。

4.1.4 缩小特征值的范围

圆盘定理虽然可以帮助确定特征值的范围, 但这个范围我们还可以通过一些处理进行缩小, 这就要利用到下面在线性代数中学到的性质: 对于一个矩阵 \mathbf{A} , 乘上一个可逆矩阵, 不改变矩阵特征值。

对于一个矩阵 \mathbf{A} , 我们做出如下处理: 左乘一个可逆矩阵 \mathbf{S}^{-1} , 右乘一个 \mathbf{S} , 于是, 对下面的一个新的矩阵画盖尔圆盘:

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$$

如果对于 \mathbf{S} 的选择较为恰当, 是能够得到更精确的特征值范围的, 通常, 一个比较特别的选择是选取

$$\mathbf{S} = \mathbf{D} = \text{diag}(p_1, p_2, \dots, p_n) \quad p_i > 0 \quad (i = 1, 2, \dots, n)$$

这样的一个对角矩阵, 于是, 新的矩阵就成了

$$\begin{aligned} \mathbf{S}^{-1}\mathbf{A}\mathbf{S} &= \begin{bmatrix} p_1^{-1} & & & \\ & p_2^{-1} & & \\ & & \ddots & \\ & & & p_n^{-1} \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} p_1 & & & \\ & p_2 & & \\ & & \ddots & \\ & & & p_n \end{bmatrix} \\ &= \begin{bmatrix} \frac{a_{11}}{p_1} & \frac{a_{12}}{p_1} & \cdots & \frac{a_{1n}}{p_1} \\ \frac{a_{21}}{p_2} & \frac{a_{22}}{p_2} & \cdots & \frac{a_{2n}}{p_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{n1}}{p_n} & \frac{a_{n2}}{p_n} & \cdots & \frac{a_{nn}}{p_n} \end{bmatrix} \cdot \begin{bmatrix} p_1 & & & \\ & p_2 & & \\ & & \ddots & \\ & & & p_n \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & \frac{p_2}{p_1}a_{12} & \cdots & \frac{p_n}{p_1}a_{1n} \\ \frac{p_1}{p_2}a_{21} & a_{22} & \cdots & \frac{p_n}{p_2}a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{p_1}{p_n}a_{n1} & \frac{p_2}{p_n}a_{n2} & \cdots & a_{nn} \end{bmatrix} \end{aligned}$$

这个时候下的盖尔圆的半径可以受 \mathbf{S} 矩阵的对角元素来影响, 可以看到, 以第一行为例, 可以增大 p_1 的值, 减小 p_2, \dots, p_n 的值, 这个时候盖尔圆 S_1 就会变小, 可相对应的会发现, S_2, \dots, S_n 的半径就会变大。

采用这种方式和正常求盖尔圆的对于特征值影响到底有多大呢? 请看下面的例子:

例 4.1.2. 估计矩阵

$$\mathbf{A} = \begin{bmatrix} 0.9 & 0.01 & 0.12 \\ 0.01 & 0.8 & 0.13 \\ 0.01 & 0.02 & 0.4 \end{bmatrix}$$

的特征值的范围。

解.

如果采用原始方法, 求解的盖尔圆的信息如下:

$$S_1 : |z - 0.9| \leq 0.13$$

$$S_2 : |z - 0.8| \leq 0.14$$

$$S_3 : |z - 0.4| \leq 0.03$$

若取 $p_1 = p_2 = 1, p_3 = 0.1$, 则

$$\mathbf{D} = \text{diag}(1, 1, 0.1)$$

新的盖尔圆盘信息就成了:

$$S'_1 : |z - 0.9| \leq 0.022$$

$$S'_2 : |z - 0.8| \leq 0.023$$

$$S'_3 : |z - 0.4| \leq 0.3$$

对比一下就会发现, 前面两个盖尔圆盘 (S_1, S_2) 的半径的确变小了, 代表特征值的范围更加精确, 但相对应的, S_3 的半径变大, 特征值的范围更加模糊。□

4.1.5 盖尔圆盘的其他性质

在本节的最后, 介绍一下其他盖尔圆盘的性质

行对角占优和列对角占优

定义 4.1.2: 行对角占优和列对角占优

设 $\mathbf{A} \in \mathbb{C}^{n \times n}$, 则

1. 若 $|a_{ii}| \geq R_i$, 则称该矩阵为行对角占优。
2. 若 $|a_{ii}| \geq C_i$, 则称该矩阵为列对角占优。

如果去掉等于号, 上面的两个就成了行严格对角占优和列严格对角占优。

这两个概念是为了下面的定理服务的, 这里不给出证明, 感兴趣的可以自己翻课本

定理 4.1.6

设 $\mathbf{A} \in \mathbb{C}^{n \times n}$ 行(或列)严格对角占优, 则

1. \mathbf{A} 可逆, 且 $\lambda_i \in \bigcup_{i=1}^n S_i$ ($S_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq |a_{ii}|\}$)
2. 若 \mathbf{A} 的所有主对角元都为正数, 则 \mathbf{A} 的特征值都有正实部
3. 若 \mathbf{A} 为 Hermite 矩阵, 且所有主对角元都为正数, 则 \mathbf{A} 的特征值都为正数。