第二章 矩阵的标准形与特征值计算（2）盖尔圆定理，幂迭代与逆幂迭代法

盖尔圆定理

盖尔圆定理的定义

对于一个 $n \times n$ 的复数矩阵 $A = [a_{ij}]$，盖尔圆定理定义了多个圆盘（称为盖尔圆盘），**每个特征值至少位于其中一个圆盘内。**

- **行盖尔圆盘**：对于矩阵 A 的每一行 i（其中 $i = 1, 2, \ldots, n$），定义第 i 个盖尔圆盘的中心为对角线元素 a_{ii}，半径 R_i 为该行非对角线元素的绝对值之和：

  $$R_i = \sum_{j \neq i} |a_{ij}|$$

  对应的圆盘为：

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

- **列盖尔圆盘**：类似地，基于列定义盖尔圆盘，中心为 a_{jj}，半径 C_j 为该列非对角线元素的绝对值之和：

  $$C_j = \sum_{i \neq j} |a_{ij}|$$

  对应的圆盘为：

  $$G'_j = \{z \in \mathbb{C} : |z - a_{jj}| \leq C_j\}$$

用盖尔圆定理进行特征值的隔离

用盖尔圆定理证明矩阵 A 有互异的特征值，即用盖尔圆定理确定每个特征值在什么范围内（通过圆盘约束）。

- 计算每个行盖尔圆，得到每个特征值范围；
- 也可以计算每个列盖尔圆，得到每个特征值范围；
- 有时通过行盖尔圆和列盖尔圆的约束，可以使得特征值分离；
- 但通常情况下，需要利用 **相似性变换**，使得特征值分离。
  1. 选择一个正定对角矩阵 $D = \operatorname{diag}(d_1, d_2, \ldots, d_n)$，其中 $d_i > 0$；
  2. 构造一个相似性矩阵 $B = D^{-1}AD$，矩阵 B 和 A 具有相同的特征值；
  3. B 的元素为 $b_{ij} = \frac{d_i}{d_j} a_{ij}$；

4. B 的行盖尔圆盘半径为 $R'_i = \sum_{j \neq i} |b_{ij}| = \sum_{j \neq i} |a_{ij}| \frac{d_i}{d_j}$;

5. 通过优化选择 d_i（例如，使 R'_i 最小化），可以使盖尔圆盘更小，从而更精确地隔离特征值。

   例如当某个盖尔圆 R_i 较大时，可以考虑缩小这个盖尔圆的半径，取 $d_i = \frac{1}{2}$ 尝试一下。

例题 1

> $A = \begin{pmatrix} 2 & 2 & -1 \\ 1 & 10 & -1 \\ 8 & 2 & 20 \end{pmatrix}$，用盖尔圆定理隔离矩阵 A 的特征值。

解答:

- 首先可以求矩阵 A 的行盖尔圆:

$$G_1: \ |z-2| \leq |2| + |-1| = 3$$

$$G_2: \ |z-10| \leq |1| + |-1| = 2$$

$$G_3: \ |z-20| \leq |8| + |2| = 10$$

- 由此得到当前每个特征值的范围，第 1 个圆盘范围 $[-1,5]$，第 2 个圆盘范围 $[8,12]$，第 3 个圆盘范围 $[10,30]$。发现第 2 个和第 3 个范围重叠，且第 3 个盖尔圆半径较大。
- 考虑缩小第 3 个行盖尔圆半径:

  取 $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$，并构造相似矩阵:

$$B = D^{-1}AD = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 & 2 & -1 \\ 1 & 10 & -1 \\ 8 & 2 & 20 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 2 & -2 \\ 1 & 10 & -2 \\ 4 & 1 & 20 \end{pmatrix}$$

- 得到 B 的行盖尔圆:

$$G_1^{'}: \ |z-2| \leq |2| + |-2| = 4$$

$$G_2^{'}: \ |z-10| \leq |2| + |1| = 3$$

$$G_3^{'}: \ |z-20| \leq |4| + |1| = 5$$

- 由此得到每个特征值相似变换后的范围，第 1 个圆盘范围 $[-2,6]$，第 2 个圆盘范围 $[7,13]$，第 3 个圆盘范围 $[15,25]$。**显然，矩阵 A 的特征值相互隔离。**

幂迭代法

特征值和特征向量

对一个 $n \times n$ 矩阵 A，如果存在**非零向量** x 和**标量** λ，使得

$$Ax = \lambda x$$

- λ 叫做矩阵 A 的一个 **特征值**；
- 对应的非零向量 x 叫做 A 的一个 **特征向量。

直观理解：**乘上矩阵前后，长度被乘了一个倍数** λ. 可以将矩阵 A 想成一个“线性变换机器”，它可以把平面或空间拉伸、压缩、旋转等。

幂迭代法公式

幂迭代法：已知矩阵 A，用迭代的方法求 **模最大特征值** $\lambda_{\max}$ 及其对应的特征向量 $x_{\max}$。

迭代公式：

$$\begin{cases} u_0 = v_0 \neq 0 \\ v_k = Au_{k-1} \\ m_k = \max(v_k) \\ u_k = \dfrac{v_k}{m_k} \end{cases}$$

通常，默认 v_0 向量的最大分量为 1，即已经进行了一次归一化。如果 v_0 最大分量不为 1，需要 $u_0 = \frac{v_0}{m_0}$ （m_0 为 v_0 最大分量 $\max(v_0)$）

公式解析：

1. **初始化：**
   选一个非零初始向量 $u_0 = v_0$。
2. **迭代乘矩阵：**

   $$v_k = Au_{k-1}$$

   这一步就是不断做

   $$u_0, ; Au_0, ; A^2u_0, ; \dots$$

   的效果。假设 A 的最大特征值 $|\lambda_1|$ 最大，所以 $A^k u_0$ 会越来越“靠近” λ_1 对应的特征向量。
3. **用最大分量归一化：**

   $$m_k = \max(v_k), \qquad u_k = \frac{v_k}{m_k}$$

   - 这是把向量的 **最大分量缩放到 1**，避免数值发散；
   - 归一化后的 u_k 作为下一步的迭代向量；
   - m_k 就接近特征值 $\lambda_{\max}$，而 u_k 接近它的特征向量。
4. **收敛判据：**
   当两次迭代的向量几乎不变时：

   $$u_k \approx u_{k-1}$$

   就认为已经收敛，此时

$$\lambda_{\max} \approx m_k, \quad x \approx u_k.$$

公式总结：从一个非零初始向量出发，不断计算 $v_k = Au_{k-1}$，再用其最大分量 $m_k = \max(v_k)$ 把向量归一化 $u_k = v_k / m_k$。在最大特征值（按模）唯一且对应特征向量分量为正的条件下，u_k 收敛到最大特征值的特征向量，而 m_k 收敛到该特征值。

例题 1

> $A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}, \quad v_0 = \begin{pmatrix} 0.75 \\ 1 \end{pmatrix}$，用幂迭代法求 A 的最大特征值对应的特征向量

解答：

$$u_0 = v_0 = \begin{pmatrix} 0.75 \\ 1 \end{pmatrix}$$

- 第 1 次迭代

$$v_1 = Au_0 = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 0.75 \\ 1 \end{pmatrix} = \begin{pmatrix} 2.75 \\ 3.75 \end{pmatrix}$$

$$m_1 = 3.75, \quad u_1 = \frac{v_1}{m_1} \approx \begin{pmatrix} 0.7333 \\ 1 \end{pmatrix}$$

- 第 2 次迭代

$$v_2 = Au_1 = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 0.7333 \\ 1 \end{pmatrix} = \begin{pmatrix} 2.7333 \\ 3.7333 \end{pmatrix}$$

$$m_2 = 3.7333, \quad u_2 = \frac{v_2}{m_2} \approx \begin{pmatrix} 0.7321 \\ 1 \end{pmatrix}$$

- 第 3 次迭代

$$v_3 = Au_2 = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 0.7321 \\ 1 \end{pmatrix} = \begin{pmatrix} 2.7321 \\ 3.7321 \end{pmatrix}$$

$$m_3 = 3.7321, \quad u_3 = \frac{v_3}{m_3} \approx \begin{pmatrix} 0.7321 \\ 1 \end{pmatrix}$$

到第 3 次迭代时，u_k 基本不再变化。

近似最大特征值和对应特征向量：

$$\lambda_{\max} \approx 3.7321, \quad x_{\max} = \begin{pmatrix} 0.7321 \\ 1 \end{pmatrix}$$

逆幂迭代法

逆幂迭代法：求矩阵的 **模最小特征值** $\lambda_{\min}$ 及其对应特征向量 $x_{\min}$ 。

普通幂迭代找的是**最大特征值**，而逆幂迭代通过“反过来处理矩阵”，让迭代过程自动偏向最小特征值。

迭代公式：

$$
\begin{cases}
u_0 = v_0 \neq 0 \\
Av_k = u_{k-1} \quad (\text{即 } v_k = A^{-1}u_{k-1}) \\
m_k = \max(v_k) \\
u_k = \dfrac{v_k}{m_k}
\end{cases}
$$

方法	计算公式核心	找到的特征值	和 (m_k) 的关系
幂迭代法	$v_k = Au_{k-1}$	最大特征值 $\lambda_{\max}$	$m_k \approx \lambda_{\max}$
逆幂迭代法	$Av_k = u_{k-1} \quad (\text{即 } v_k = A^{-1}u_{k-1})$	**最小特征值** $\lambda_{\min}$	**求得** m_k **的倒数才是特征值：** $\lambda_{\min} \approx \frac{1}{m_k}$

带原点位移的逆幂迭代法：

已经**大致知道矩阵某个特征值的位置**（记为 p），但它不是最大也不是最小特征值，此时：

> 可以通过"带位移的逆幂迭代法"，让迭代过程 **快速逼近这个特定的特征值及其特征向量。**

适用于：

- 想求 **中间的特征值**（不是最大或最小）
- 特征值可能是 **无理数**
- 想提高迭代收敛速度

迭代公式：

$$
\begin{cases}
u_0 = v_0 \neq 0 \\
(A - pI)v_k = u_{k-1} \quad (\text{即 } v_k = (A - pI)^{-1}u_{k-1}) \\
m_k = \max(v_k) \\
u_k = \dfrac{v_k}{m_k}
\end{cases}
$$

最后求得的近似特征值：$\lambda \approx p + \frac{1}{m_k}$

例题 2

> $A = \begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix}$，$p = 3.5$，$v_0 = \begin{pmatrix} 0.95 \\ 0.25 \end{pmatrix}$，用带原点位移的逆幂迭代法，求经过 2 次迭代后，得到的特征值及对应特征向量。

解答：

$$
A - pI = \begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix} - 3.5 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -0.5 & 2 \\ 1 & -3.5 \end{pmatrix}
$$

$$
B = (A - pI)^{-1} = \begin{pmatrix} 14 & 8 \\ 4 & 2 \end{pmatrix}
$$

$$
m_0 = 0.95, \quad u_0 = \frac{v_0}{m_0} = \begin{pmatrix} 1 \\ 0.2632 \end{pmatrix}
$$

- 第 1 次迭代

$$v_1 = Bu_0 = \begin{pmatrix} 14 & 8 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0.2632 \end{pmatrix} = \begin{pmatrix} 16.1056 \\ 4.5264 \end{pmatrix}$$

$$m_1 = 16.1056, \quad u_1 = \frac{v_1}{m_1} = \begin{pmatrix} 1 \\ 0.2810 \end{pmatrix}$$

- 第 2 次迭代

$$v_2 = Bu_1 = \begin{pmatrix} 14 & 8 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0.2810 \end{pmatrix} = \begin{pmatrix} 16.248 \\ 4.562 \end{pmatrix}$$

$$m_2 = 16.248, \quad u_2 = \frac{v_2}{m_2} = \begin{pmatrix} 1 \\ 0.2808 \end{pmatrix}$$

得到近似特征值和特征向量:

$$\lambda \approx p + \frac{1}{m_2} = 3.5615, \quad x \approx \begin{pmatrix} 1 \\ 0.2808 \end{pmatrix}$$
