第五讲 对称特征值问题

设 $A \in \mathbb{R}^{n\times n}$ 是对称矩阵. 在计算 A 的特征值和特征值向量时, 我们可以充分利用 A 的对称结构, 一方面尽可能地减少运算量, 另一方面也能构造出更加快速高效的算法.

关于对称矩阵的特征值和特征值向量, 目前常用算法有:

- Jacobi 迭代: 最古老的方法, 收敛速度较慢, 但精度较高, 且很适合并行计算.
- Rayleigh 商迭代: 利用 Rayleigh 商作为位移的反迭代算法, 一般具有三次收敛性.
- 对称 QR 迭代: 计算对称矩阵的特征值和特征向量的 QR 算法. 如果只需计算对称三对角矩阵的所有特征值, 则该算法是目前最快的方法 (运算量为 $O(n^2)$). 如果需要计算所有的特征值和特征向量, 则运算量约为 $6n^3$.
- 分而治之法 (Divide-and-Conquer) : 计算对称三对角矩阵的特征值和特征值向量的一种快速算法. 基本思想是将大矩阵分解成小矩阵, 然后利用递推思想求特征值和特征向量. 在最坏的情形下, 运算量为 $O(n^3)$, 但在实际应用中, 平均为 $O(n^{2.3})$. 如果使用快速多极子算法 (FMM) 后, 理论上的运算量可降低到 $O(n\log^p n)$, 其中 p 是一个较小的整数, 这使得分而治之算法成为目前求解对称三对角矩阵的所有特征值和特征向量的最快方法之一.
- 对分法和反迭代: 对分法主要用于求解对称三对角矩阵在某个区间中的特征值, 运算量约为 $O(kn)$, 其中 k 为所需计算的特征值的个数. 反迭代用于计算特征向量, 在最佳情况下, 即特征值 “适当分离” 时, 运算量约为 $O(kn)$, 但在最差情况下, 即特征值成串地紧靠在一起时, 运算量约为 $O(k^2n)$, 而且不能保证特征向量的精度 (虽然实际上它几乎是精确的).

† 除了 Jacobi 迭代和 Rayleigh 商迭代外, 其余算法都需要先将对称矩阵三对角化. 这个过程大约需花费 $\frac{4}{3}n^3$ 的工作量, 如果需要计算特征向量的话, 则运算量约为 $\frac{8}{3}n^3$.

5.1 Jacobi 迭代

该算法的基本思想是通过一系列的 Jacobi 旋转 J_k 将 A 正交相似于一个对角矩阵, 即

$$A^{(0)} = A, \quad A^{(k+1)} = J_k^{\intercal} A^{(k)} J_k, \quad k = 0, 1, \dots,$$

且 $A^{(k)}$ 收敛到一个对角矩阵, 其中 J_k 为正交矩阵.

Jacobi 旋转 J_k 的构造

我们通常选取 J_k 为 Givens 变换, 即

$$J_k = G(i_k, j_k, \theta_k) = \begin{array}{c} \\ \\ \\ i_k \\ \\ j_k \\ \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccccccc} & & & i_k & & j_k & & & \end{array} \\ \begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & \cos\theta_k & & -\sin\theta_k & & & \\ & & & & \ddots & & & & \\ & & & \sin\theta_k & & \cos\theta_k & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{bmatrix} \end{array}.$$

由于 $A^{(k)}$ 是对称矩阵, 所以可以选取适当的 θ_k, 将 $A^{(k)}(i,j)$ 和 $A^{(k)}(j,i)$ 化为 0.

引理 5.1 设 $A \in \mathbb{R}^{2\times 2}$ 是对称矩阵, 则存在 Givens 变换 $G \in \mathbb{R}^{2\times 2}$ 使得 $G^\intercal AG$ 为对角阵.

证明. 设

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}, \quad G = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

则

$$\begin{aligned} G^\intercal AG &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^\intercal \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} a\cos^2\theta + c\sin^2\theta + b\sin 2\theta & \frac{1}{2}(c-a)\sin 2\theta + b\cos 2\theta \\ \frac{1}{2}(c-a)\sin 2\theta + b\cos 2\theta & a\sin^2\theta + c\cos^2\theta - b\sin 2\theta \end{bmatrix} \end{aligned}$$

令 $\frac{1}{2}(c-a)\sin 2\theta + b\cos 2\theta = 0$ 即得

$$\frac{a-c}{2b} = \cot 2\theta = \frac{1-\tan^2\theta}{2\tan\theta}.$$

解得

$$\tan\theta = \frac{\mathrm{sign}(\tau)}{|\tau| + \sqrt{1+\tau^2}}, \qquad \tau = \frac{a-c}{2b}.$$

故引理结论成立. □

为了使得 $A^{(k)}$ 收敛到一个对角矩阵, 其非对角元素必须趋向于 0. 记 $\mathrm{off}(A)$ 为所有非对角元素的平方和, 即

$$\mathrm{off}(A) = \sum_{i\neq j} a_{ij}^2 = \|A\|_F^2 - \sum_{i=1}^n a_{ii}^2,$$

我们的目标就是使得 $\mathrm{off}(A)$ 尽快趋向于 0.

引理 5.2 设 $A=[a_{ij}]_{n\times n}\in\mathbb{R}^{n\times n}$ 是对称矩阵, $\hat{A}=[\hat{a}_{ij}]_{n\times n}=J^{\intercal}AJ$, $J=G(i,j,\theta)$, 其中 θ 的选取使得 $\hat{a}_{ij}=\hat{a}_{ji}=0$, 则

$$\mathrm{off}(\hat{A})=\mathrm{off}(A)-2a_{ij}^2.$$

证明. 设 $A=[a_1,a_2,\ldots,a_n]$. 令 $\tilde{A}=J^{\intercal}A=[\tilde{a}_{ij}]_{n\times n}$. 由于 J 是正交阵, 故

$$\|J^{\intercal}a_k\|_2=\|a_k\|_2,\quad k=1,2,\ldots,n.$$

又 $J^{\intercal}$ 左乘 a_k 时, 只影响其第 i,j 个元素的值, 故由 $\|J^{\intercal}a_i\|_2=\|a_i\|_2$ 和 $\|J^{\intercal}a_j\|_2=\|a_j\|_2$ 可得

$$\tilde{a}_{ii}^2+\tilde{a}_{ji}^2=a_{ii}^2+a_{ji}^2,\quad \tilde{a}_{ij}^2+\tilde{a}_{jj}^2=a_{ij}^2+a_{jj}^2. \tag{5.1}$$

同理, 由 $\hat{A}=\tilde{A}J$ 可得

$$\hat{a}_{ii}^2+\hat{a}_{ij}^2=\tilde{a}_{ii}^2+\tilde{a}_{ij}^2,\quad \hat{a}_{ji}^2+\hat{a}_{jj}^2=\tilde{a}_{ji}^2+\tilde{a}_{jj}^2. \tag{5.2}$$

又 $\hat{a}_{ij}=\hat{a}_{ji}=0$, 故

$$\hat{a}_{ii}^2+\hat{a}_{jj}^2=a_{ii}^2+a_{jj}^2+a_{ij}^2+a_{ji}^2=a_{ii}^2+a_{jj}^2+2a_{ij}^2.$$

由于 $J^{\intercal}AJ$ 只影响 A 的第 i,j 行和第 i,j 列, 故对角线元素中只有 a_{ii} 和 a_{jj} 受影响. 所以

$$\sum_{k=1}^{n}\hat{a}_{kk}^2=\sum_{k=1}^{n}a_{kk}^2+2a_{ij}^2,$$

故

$$\mathrm{off}(\hat{A})=\|\hat{A}\|_F^2-\sum_{k=1}^{n}\hat{a}_{kk}^2=\|A\|_F^2-\sum_{k=1}^{n}a_{kk}^2-2a_{ij}^2=\mathrm{off}(A)-2a_{ij}^2,$$

即引理结论成立. □

由此可知, $\mathrm{off}(A^{(k)})$ 总是不断减小的. 下面给出 Jacobi 迭代算法.

算法 5.1. Jacobi 迭代算法

Given a symmetric matrix $A\in\mathbb{R}^{n\times n}$
if eigenvectors are desired **then**
  set $J=I$ and $shift=1$
end if
while not converge **do**
  choose an index pair (i,j) such that $a_{ij}\neq 0$
  $\tau=(a_{ii}-a_{jj})/(2a_{ij})$
  $t=\mathrm{sign}(\tau)/(|\tau|+\sqrt{1+\tau^2})$ % 计算 $\tan\theta$
  $c=1/\sqrt{1+t^2}$ % 计算 $\cos\theta$
  $s=c\cdot t$ % 计算 $\sin\theta$
  $A=G(i,j,\theta)^{\intercal}AG(i,j,\theta)$ % 实际计算时不需要做矩阵乘积
  if $shift=1$ **then**
    $J=J\cdot G(i,j,\theta)$

end if
end while

该算法涉及到 a_{ij} 的选取问题, 一种直观的选取方法就是使得 a_{ij} 为所有非对角元素中绝对值最大的一个, 于是我们就得到下面的经典 Jacobi 算法.

算法 5.2. 经典 Jacobi 迭代算法

Given a symmetric matrix $A \in \mathbb{R}^{n\times n}$
if eigenvectors are desired **then**
  set $J = I$ and $shift = 1$
end if
while off(A)> tol **do**
  choose (i, j) such that $|a_{ij}| = \max_{k\neq l} |a_{kl}|$
  $\tau = (a_{ii} - a_{jj})/(2a_{ij})$
  $t = \mathrm{sign}(\tau)/(|\tau| + \sqrt{1+\tau^2})$
  $c = 1/\sqrt{1+t^2}$
  $s = c \cdot t$
  $A = G(i, j, \theta)^{\intercal} AG(i, j, \theta)$
  if $shift = 1$ **then**
    $J = J \cdot G(i, j, \theta)$
  end if
end while

可以证明, 经典 Jacobi 算法至少是线性收敛的.

定理 5.1 对于经典 Jacobi 算法 5.2, 有

$$\mathrm{off}(A^{(k+1)}) \leq \left(1 - \frac{1}{N}\right) \mathrm{off}(A^{(k)}), \quad N = \frac{n(n-1)}{2}.$$

故 k 步迭代后, 有

$$\mathrm{off}(A^{(k)}) \leq \left(1 - \frac{1}{N}\right)^k \mathrm{off}(A^{(0)}) = \left(1 - \frac{1}{N}\right)^k \mathrm{off}(A).$$

证明. 由于在经典 Jacobi 算法 5.2 中, $|a_{ij}| = \max_{k\neq l} |a_{kl}|$, 故 $\mathrm{off}(A^{(k)}) \leq n(n+1)\left(a_{ij}^{(k)}\right)^2$, 即

$$2\left(a_{ij}^{(k)}\right)^2 \geq \frac{1}{N}\mathrm{off}(A^{(k)}), \quad N = \frac{n(n-1)}{2}.$$

所以由引理 5.2 可知

$$\mathrm{off}(A^{(k+1)}) = \mathrm{off}(A^{(k)}) - \left(a_{ij}^{(k)}\right)^2 \leq \left(1 - \frac{1}{N}\right) \mathrm{off}(A^{(k)}).$$

□

事实上, 经典 Jacobi 算法最终是二次局部收敛的.