对称特征值问题的计算方法

教师: 胡俊

北京大学数学科学学院

December 16, 2021

目录

目录

重要定理

Theorem 1.1 (极小极大定理)

设$A \in \mathbf{R}^{n\times n}$ 是对称矩阵，并假定A 的特征值为$\lambda_1 \geqslant \cdots \lambda_n$，则有

$$
\begin{aligned}
\lambda_i &= \max_{\chi \in \mathcal{G}_i^n} \min_{0 \neq u \in \chi} \frac{u^T A u}{u^T u} \\
&= \min_{\chi \in \mathcal{G}_{n-i+1}^n} \max_{0 \neq u \in \chi} \frac{u^T A u}{u^T u}
\end{aligned}
$$

其中$\mathcal{G}_k^n$ 表示$\mathbf{R}^n$ 中所有k 维子空间的全体.

定理(Weyl 定理)： 设n 阶对称矩阵A 和B 的特征值分别为：

$$\lambda_1 \geq \cdots \geq \lambda_n \quad \text{和} \quad \mu_1 \geq \cdots \geq \mu_n$$

则有

$$|\lambda_i - \mu_i| \leq \|A - B\|_2, i = 1, \cdots, n$$

证明： 令$C = A - B$, 则有

$$A = B + C$$

$$\lambda_s = \min\max(u^T Au)$$

其中

$$u^T u = 1, P_i^T u = 0, P_1, \cdots, P_{s-1} \quad \text{为} s - 1 \text{ 个线性无关的向量}$$

因此，若取一组特殊的P_i, 则对相应的u, 有

$$\lambda_s \leq \max(u^T Au) = \max(u^T Bu + u^T Cu)$$

若R 为正交矩阵, 且使得

$$R^TBR = \mathrm{diag}(\mu_1, \cdots, \mu_n)$$

则取$P_i = Re_i$, 这样

$$\lambda_s \leq \max(u^TBu + u^TCu) = \max(u^TRR^TBRR^Tu + u^TCu)$$

因为R^Tu 的前$s-1$ 个分量为零, 因此

$$\lambda_s \leq \mu_s + \max(u^TCu)$$

因而

$$|\lambda_s - \mu_s| \leq \|u\|_2^2\|C\|_2 = \|A-B\|_2$$

证毕.

关于特征向量的敏感性，有如下结论：

Theorem 1.2

设A 和$A+E$ 是两个n 阶实对称矩阵，并假定q_1 是A 的一个单位特征向量，$Q=[q_1,Q_2]$ 是n 阶正交矩阵，Q^TAQ 和Q^TEQ 分块如下：

$$Q^TAQ=\begin{array}{cc} \begin{bmatrix} \lambda & 0 \\ 0 & D_2 \end{bmatrix} & \begin{array}{c} 1 \\ n-1 \end{array} \\ \begin{array}{cc} 1 & n-1 \end{array} & \end{array}, \qquad Q^TEQ=\begin{array}{cc} \begin{bmatrix} \varepsilon & e^T \\ e & E_{22} \end{bmatrix} & \begin{array}{c} 1 \\ n-1 \end{array} \\ \begin{array}{cc} 1 & n-1 \end{array} & \end{array}$$

若

$$d=\min_{\mu\in\lambda(D_2)}|\lambda-\mu|>0, \quad \|E\|_2\leqslant\frac{1}{4}d,$$

则存在$A+E$ 的一个单位特征向量$\hat{q}_1$，使得

$$\sin\theta=\sqrt{1-|q_1^T\hat{q}_1|^2}\leqslant\frac{4}{d}\|e\|_2\leqslant\frac{4}{d}\|E\|_2,$$

定理(SVD 分解定理)： 设$A \in R^{m\times n}$, 则存在正交矩阵$U \in R^{n\times n}$ 和$V \in R^{m\times m}$ 使得

$$V^T AU = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix}$$

其中$\Sigma_r = \mathrm{diag}(\sigma_1, \cdots, \sigma_r)$, $\sigma_1 \geq \sigma_2 \geq \sigma_r > 0$. 称$V$ 为左奇异向量矩阵, U 为右奇异向量矩阵

证明： 矩阵$A^T A$ 为半正定矩阵, 其特征值为$\sigma_1^2, \sigma_2^2, \cdots, \sigma_n^2$, 其中

$$\sigma_1 \geq \sigma_2 \cdots \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_n$$

设$u_1, \cdots, u_n$ 为$\sigma_1^2, \sigma_2^2, \cdots, \sigma_n^2$ 对应的特征向量(标准正交向量). 令$U = [U_1 U_2]$, 其中

$$U_1 = [u_1, u_2, \cdots, u_r], U_2 = [u_{r+1}, \cdots, u_n]$$

于是

$$U_1^T A^T A U_1 = \Sigma_r^2.$$

因此

$$\Sigma_r^{-1} U_1^T A^T A U_1 \Sigma_r^{-1} = I$$

令$V_1 = AU_1\Sigma_r^{-1}$, 得

$$V_1^T V_1 = I$$

故V_1 的列向量彼此两两正交. 可以选择$m-r$ 个向量$v_{r+1}, \cdots, v_m$ 和V_1 的列向量构成R^m 的一组标准正交基, 记

$$V_2 = [v_{r+1}, \cdots, v_m]$$

则$V = [V_1 V_2]$ 为一正交矩阵.
又根据U_2 的定义可得

$$U_2^T A^T A U_2 = 0$$

即

$$AU_2 = 0$$

下面来验证$U=[U_1U_2]$ 和$V=[V_1V_2]$ 即为定理所需的正交矩阵.
因为

$$V^TAU=\begin{pmatrix} V_1^TAU_1 & V_1^TAU_2 \\ V_2^TAU_1 & V_2^TAU_2 \end{pmatrix}$$

由$AU_2=0$, 得$V_2^TAU_2=0$, $V_1^TAU_2=0$. 由V_1 的定义

$$V_1^TAU_1=(AU_1\Sigma_r^{-1})^TAU_1=\Sigma_r^{-1}U_1^TA^TAU_1=\Sigma_r^{-1}\Sigma_r^2=\Sigma_r$$

$$V_2^TAU_1=V_2^T\boldsymbol{(AU_1\Sigma_r^{-1})}\Sigma_r=V_2^T\mathbf{V_1}\Sigma_r=0$$

证毕.

作为上面定理的简单推论，有如下结论：

定理： 设$A, B \in \mathbf{R}^{m\times n}$, 并假定它们的奇异值分别为

$$\sigma_1 \geqslant \cdots \geqslant \sigma_n, \qquad \tau_1 \geqslant \cdots \geqslant \tau_n,$$

则有

$$|\sigma_i - \tau_i| \leqslant \|A - B\|_2, \quad i = 1, \cdots, n.$$

这一结果表明奇异值亦是十分良态的.

目录

对称 QR 方法

对称 QR 方法 就是求解对称特征值问题的 QR 方法, 是将 QR 方法应用于对称矩阵, 并且充分利用其对称性得到的.

三对角化:

若 A 是 n 阶实对称矩阵, 并假定 A 上的 Hessenberg 分解为

$$Q^T AQ = T \tag{2.1}$$

其中 Q 是正交矩阵, T 是上 Hessenberg 矩阵, 则容易验证 T 必是对称三角阵. 因此, 上 Hessenberg 化实质上就是三对角化. 如果在约化过程中充分利用其对称性, 可减少运算量. 将A 作如下分块

$$A = \begin{array}{cc} \begin{bmatrix} \alpha_1 & v_0^T \\ v_0 & A_0 \end{bmatrix} & \begin{array}{c} 1 \\ n-1 \end{array} \\ \begin{array}{cc} 1 & \quad n-1 \end{array} & \end{array}$$

从约化一个矩阵为上 Hessenberg 矩阵的 Householder 变化法不难推出, 利用 Householder 变换将 A 约化为对称三对角阵的第 k 步为:

(1) 计算 Householder 变换 $\tilde{H}_k \in \mathbf{R}^{(n-k)\times(n-k)}$, 使得

$$\tilde{H}_k v_{k-1} = \beta_k e_1, \ \beta_k \in \mathbf{R} \tag{2.2}$$

(2) 计算

$$\begin{array}{cc} & \begin{array}{cc} \\ \end{array} \\ \begin{array}{c} 1 \\ n-k-1 \end{array} & \begin{bmatrix} \alpha_{k+1} & v_k^T \\ v_k & A_k \end{bmatrix} \\ & \begin{array}{cc} 1 & \quad n-k-1 \end{array} \end{array} = \tilde{H}_k A_{k-1} \tilde{H}_k.$$

如果用上述约化过程所产生的 α_k, β_k 和 $\tilde{H}_k$ 定义

$$T = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_{n-1} \\ & & \beta_{n-1} & \alpha_n \end{bmatrix},$$

$$Q = H_1 H_2 \cdots H_{n-2}, \quad H_k = \mathrm{diag}(I_k, \tilde{H}_k),$$

其中

$$\begin{bmatrix} \alpha_{n-1} & \beta_{n-1} \\ \beta_{n-1} & \alpha_n \end{bmatrix} = \tilde{H}_{n-2} A_{n-3} \tilde{H}_{n-2},$$

则有

$$Q^T A Q = T,$$

即

$$A = Q T Q^T,$$

并称之为 A 的三对角分解.

从上述约化过程容易看出, 第 k 步约化的主要工作量是计算 $\tilde{H}_k A_{k-1} \tilde{H}_k$. 设

$$\tilde{H}_k = I - \beta v v^T, \; v \in \mathbf{R}^{n-k},$$

则利用 A_{k-1} 的对称性, 易得

$$\tilde{H}_k A_{k-1} \tilde{H}_k = A_{k-1} - v w^T - w v^T,$$

其中

$$w = u - \frac{1}{2}\beta(v^T u)v, \quad u = \beta A_{k-1} v.$$

不难算出, 利用这一等式来计算, 其运算量仅为 $4(n-k)^2$.

Algorithm 1 (计算三对角分解: Householder 变换法)

for $k = 1 : n-2$
  $[v, \beta] = \textbf{house}(A(k+1:n, k))$
  $u = \beta A(k+1:n, k+1:n)v$
  $w = u - (\beta u^T v/2)v$
  $A(k+1, k) = \|A(k+1:n, k)\|_2$
  $A(k, k+1) = A(k+1, k)$
  $A(k+1:n, k+1:n) = A(k+1:n, k+1:n) - vw^T - wv^T$
end

该算法的额运算量为 $4n^3/3$ 次乘法运算, 而非对称矩阵的上 Hessenberg 化需要的运算量为 $10n^3/3$; 如果需要将变换的矩阵累积起来, 则还需要增加运算量 $4n^3/3$.

隐式对称 QR 迭代

完成了把 A 约化为三对角阵 T 的任务之后，下一个任务就是选取适当的位移进行 QR 迭代. 由于此时 A 的特征值均为实数，因而再使用双重步位移就完全没有必要了，只需要进行单步位移即可.

考虑带原点位移的 QR 迭代格式

$$
\begin{aligned}
T_k - \mu_k I &= Q_k R_k, \quad (QR\ )) \\
T_{k+1} &= R_k Q_k + \mu_k I, \quad k = 0, 1, \cdots,
\end{aligned}
$$

其中 $T_0 = T$ 是三对角阵. 由于 QR 迭代保持上 Hessenberg 形和对称性的特点，上述迭代格式产生的 T_k 都是对称三对角阵.

与非对称 QR 方法不一样，假定迭代中所出现的 T_k 均是不可约的，即其次对角元不为零.

Wilkinson 位移

先来讨论如何选取位移 μ_k 的问题. 从非对称 QR 迭代的讨论可知, 最简单的做法是取 $\mu_k = T_k(n,n)$. 然而, 更好的做法是取 μ_k 为矩阵

$$T_k(n-1:n, n-1,n) = \begin{bmatrix} \alpha_{n-1} & \beta_{n-1} \\ \beta_{n-1} & \alpha_n \end{bmatrix}$$

的两个特征值之中靠近 α_n 的那一个, 即取

$$\mu_k = \alpha_n + \delta - \operatorname{sgn}(\delta)\sqrt{\delta^2 + \beta_{n-1}^2}, \tag{2.3}$$

其中 $\delta = (\alpha_{n-1} - \alpha_n)/2$. 这就是 **Wilkinson 位移**. Wilkinson 曾经证明了这两种位移最终都是三次收敛的, 并且说明了为什么后者比前者好.

再来考虑如何具体实现一次对称 QR 迭代:

$$T - \mu I = QR, \quad \tilde{T} = RQ + \mu I. \tag{2.4}$$

当然可以利用 Givens 变换来直接实现 $T - \mu I$ 的 QR 分解, 进而完成一步迭代. 但是, 更漂亮的做法是以隐含的方式来实现 T 到 $\tilde{T}$ 的变换.

迭代 (2.4) 的实质是用正交变换将 T 变为 $\tilde{T}$, 即 $\tilde{T} = Q^T T Q$. 因此, 根据 Hessenberg 分解定理, $\tilde{T}$ 本质上是由 Q 的第一列完全确定的. 从利用 Givens 比那换实现 $T - \mu I$ 的 QR 分解的计算过程可知, $Qe_1 = G_1 e_1$, 其中 $G_1 = G(1, 2, \theta)$ 是通过 (1,2) 平面的旋转将 $T - \mu I$ 的第一列第二个元素变为零的, 即 θ 满足

$$\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \alpha_1 - \mu \\ \beta_1 \end{bmatrix} = \begin{bmatrix} * \\ 0 \end{bmatrix}.$$

令

$$B = G_1 T G_1^T,$$

则 B （例如 $n=4$ ）有如下形状:

$$B = \begin{bmatrix} \times & \times & + & 0 \\ \times & \times & \times & 0 \\ + & \times & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix},$$

仅比对称三对角阵多两个非零元素“+”. 根据前面的定理, 只需将 B 用 Givens 变换约化为三对角阵, 即可得到所需的三对角阵 $\tilde{T}$. 这一约化过程不难从下面 $n=4$ 的例子中明白:

$$B \xrightarrow{G_2} \begin{bmatrix} \times & \times & 0 & 0 \\ \times & \times & \times & + \\ 0 & \times & \times & \times \\ 0 & + & \times & \times \end{bmatrix} \xrightarrow{G_3} \begin{bmatrix} \times & \times & 0 & 0 \\ \times & \times & \times & 0 \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix},$$

其中 $G_i = G(i, i+1, \theta_i), \ i = 2, 3.$

Algorithm 2 (带 Wilkinson 位移的隐式对称 QR 迭代)

$d = (T(n-1,n-1) - T(n,n))/2$

$\mu = T(n,n) - T(n,n-1)^2/(d + sgn(d)\sqrt{d^2 + T(n,n-1)^2})$

$x = T(1,1) - \mu; z = T(2,1)$

for $k = 1 : n-1$

$\quad [c,s] =$ **givens**(x,z)

$\quad T = G_k T G_k^T$，其中 $G_k = G(k,k+1,\theta)$

$\quad$ **if** $k < n-1$

$\quad\quad x = T(k+1,k);\ z = T(k+2,k)$

$\quad$ **end**

end

隐式对称 QR 算法

类比于非对称 QR 算法, 综合上面的讨论, 可得如下算法:

(1) 输入 A (实对称矩阵)

(2) 三对角化: 用算法 1 计算 A 的三对角分解, 得

$$T = U_0^T A U_0; \quad Q = U_0.$$

(3) 收敛性判定:

(i) 把所有满足条件

$$|t_{i+1,i}| = |t_{i,i+1}| \leq (|t_{ii}| + |t_{i+1,i+1}|)\mathbf{u}$$

的 $t_{i+1,i}$ 和 $t_{i,i+1}$ 置零.

(ii) 确定最大的非负整整 m 和最小的非负整数 l , 使得

$$T = \begin{array}{cc} \begin{bmatrix} T_{11} & 0 & 0 \\ 0 & T_{22} & 0 \\ 0 & 0 & T_{33} \end{bmatrix} & \begin{array}{c} l \\ n-l-m \\ m \end{array} \\ \begin{array}{ccc} l & n-l-m & m \end{array} & \end{array},$$

其中 T_{33} 为对角阵, 而 T_{22} 为不可约的三对角阵.

(iii) 如果 $m = n$, 则输出有关信息, 结束; 否则, 进行下一步.

(4) QR 迭代: 对 T_{22} 用算法 迭代一次得

$$T_{22} = GT_{22}G^T, \ G = G_1G_2 \cdots G_{n-m-l-1}.$$

(5) $Q = Q\text{diag}(I_l, G, I_m)$, 然后转步 (3) .

如果只计算特征值, 则该算法的运算量平均约为 $4n^3/3$; 如果特征值和特征向量都需要, 则运算量平均约为 $9n^3$.

这一算法是矩阵计算中最漂亮的算法之一. 误差分析结果表明, 该算法计算得到的特征值 $\tilde{\lambda}_1, \cdots, \tilde{\lambda}_n$ 满足

$$Q^T(A+E)Q = \text{diag}(\tilde{\lambda}_1, \cdots, \tilde{\lambda}_n),$$

其中 $Q \in \mathbf{R}^{n\times n}$ 是正交矩阵, $\|E\|_2 \approx \|A\|_2\mathbf{u}$. 再由定理 7.1.3 知

$$|\lambda_i - \tilde{\lambda}_i| \approx \|A\|_2\mathbf{u}, \ i = 1, \cdots, n,$$

其中 $\lambda_i \in \lambda(A)$.

目录

经典Jacobi 方法

设$A = [\alpha_{ij}]$ 是$n \times n$实对称矩阵. Jacobi 方法的目标就是将A 的非对角“范数”

$$E(A) = (\|A\|_F^2 - \sum_{i=1}^{n} \alpha_{ii}^2)^{\frac{1}{2}} = (\sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \alpha_{ij}^2)^{\frac{1}{2}} \tag{3.1}$$

逐步化为零.

给定如下的旋转变换:

$$J(p,q,\theta) = I + (\cos\theta - 1)(e_p e_p^T + e_q e_q^T) + \sin\theta(e_p e_q^T - e_q e_p^T) \tag{3.2}$$

其中假定$p < q$, e_k 表示单位矩阵的第k 列. Jacobi 方法基本步骤:

(1) 选择旋转平面(p,q), $1 \leq p < q \leq n$;

(2) 确定旋转平面θ, 使得:

$$\begin{pmatrix} \beta_{pp} & \beta_{pq} \\ \beta_{qp} & \beta_{qq} \end{pmatrix} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} \alpha_{pp} & \alpha_{pq} \\ \alpha_{qp} & \alpha_{qq} \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix}$$

(3) 对A 作相似变换: $B = [\beta_{ij}] = J^T A J$, 其中$J = J(p,q,\theta)$.

注意A 与B 只在第p 行(列)和第q 行(列)不同, 他们之间有如下关系:

$$\begin{aligned}
\beta_{ip} &= \beta_{pi} = c\alpha_{ip} - s\alpha_{iq}, i \neq p, q \\
\beta_{iq} &= \beta_{qi} = -s\alpha_{ip} + c\alpha_{iq}, i \neq p, q \\
\beta_{pp} &= c^2\alpha_{pp} - 2sc\alpha_{pq} + s^2\alpha_{qq} \\
\beta_{qq} &= s^2\alpha_{pp} + 2sc\alpha_{pq} + c^2\alpha_{qq} \\
\beta_{pq} &= \beta_{qp} = (c^2 - s^2)\alpha_{pq} + sc(\alpha_{pp} - \alpha_{qq})
\end{aligned}$$

为了使$\beta_{pq} = \beta_{qp} = 0$, 等价于计算$s$, c 使得:

$$\alpha_{pq}(c^2 - s^2) + (\alpha_{pp} - \alpha_{qq})cs = 0 \tag{3.3}$$

如果$\alpha_{pq}=0$, 则只需取$c=1, s=0$. 若$\alpha_{pq}\neq 0$, 令:

$$\tau=\frac{\alpha_{qq}-\alpha_{pp}}{2\alpha_{pq}}, \quad t=\tan\theta=\frac{s}{c}$$

代入上式, 得t 为:

$$t^2+2\tau t-1=0$$

的解. 这样的t 有两种选择, 我们选取其绝对值较小的一个. 即:

$$t=\frac{sgn(\tau)}{|\tau|+\sqrt{1+\tau^2}} \tag{3.4}$$

这样选取可以保证$\theta\leq\frac{\pi}{4}$. 根据上述讨论, 确定c, s 如下:

$$c=\frac{1}{\sqrt{1+t^2}}, \quad s=tc. \tag{3.5}$$

下面讨论怎样选取旋转平面. 由于F 范数为酉不变范数, 即$\|A\|_F = \|B\|_F$. 另一方面, 有

$$\alpha_{pp}^2 + \alpha_{qq}^2 + 2\alpha_{pq}^2 = \beta_{pp}^2 + \beta_{qq}^2 + 2\beta_{pq}^2 = \beta_{pp}^2 + \beta_{qq}^2$$

于是

$$\begin{aligned} E(B)^2 &= \|B\|_F^2 - \sum_{i=1}^n \beta_{ii}^2 \\ &= \|A\|_F^2 - \sum_{i=1}^n \alpha_{ii}^2 + (\alpha_{\mathbf{pp}}^{\mathbf{2}} + \alpha_{\mathbf{qq}}^{\mathbf{2}} - \beta_{\mathbf{pp}}^{\mathbf{2}} - \beta_{\mathbf{qq}}^{\mathbf{2}}) \\ &= E(A)^2 - \mathbf{2}\alpha_{\mathbf{pq}}^{\mathbf{2}} \end{aligned}$$

这样, 取(p, q) 为:

$$|\alpha_{pq}| = max_{1\leq i<j\leq n}|\alpha_{ij}| \tag{3.6}$$

使得$E(B)$ 尽可能的小.

经典Jacobi 的基本迭代格式如下:

$$A_k = [\alpha_{ij}^{(k)}] = J_k^T A_{k-1} J_k, \quad k = 1, 2, \cdots, \tag{3.7}$$

其中$A_0 = A$,J_k 是对A_{k-1} 用之前的算法确定的Jacobi 变换.

定理(Jacobi 算法的收敛性): 存在A 的特征值的一个排列$\lambda_1, \lambda_2, \cdots, \lambda_n$, 使得:

$$\lim_{k\to\infty} A_k = diag(\lambda_1, \lambda_2, \cdots, \lambda_n) \tag{3.8}$$

证明: 先证明$\lim_{k\to\infty} E(A_k) = 0$. 由前面的讨论, 知:

$$E(A_k)^2 = E(A_{k-1})^2 - \mathbf{2(\alpha_{pq}^{(k-1)})^2} \tag{3.9}$$

这里$\alpha_{pq}^{(k-1)}$是A_{k-1}的非对角元之中绝对值最大者.又有:

$$E(A_{k-1})^2 \leq n(n-1)(\alpha_{pq}^{(k-1)})^2 \tag{3.10}$$

因此:

$$E(A_k)^2 \leq (1-\frac{1}{N})E(A_{k-1})^2$$

其中$N=\frac{1}{2}n(n-1)$, 由此可知$\lim_{k\to\infty}E(A_k)=0$.

再证明存在A的特征值的一个排列$\lambda_1,\lambda_2,\cdots,\lambda_n$, 使得:

$$\lim_{k\to\infty}\alpha_{ii}^{(k)}=\lambda_i, \quad i=1,2,\cdots,n \tag{3.11}$$

假定A互不相同的特征值之间最小的间距为δ, 即:

$$\delta = min\{|\mu-\lambda| : \lambda,\mu\in\lambda(A),\lambda\neq\mu\}$$

任取正数$\epsilon < \delta/4$, 则由$\lim_{k\to\infty} E(A_k) = 0$ 知, 存在k_0 使$k \geq k_0$ 后有:

$$E(A_k) < \epsilon < \delta$$

注意到$\lambda(A_{k_0}) = \lambda(A)$, 对矩阵$A_{k_0}$ 与其对角元做成的对角阵

$$D_{k_0} = diag(\alpha_{11}^{(k_0)}, \alpha_{22}^{k_0}, \cdots, \alpha_{nn}^{(k_0)})$$

应用Weyl 定理: 存在A 的特征值的一个排列$\lambda_1, \lambda_2, \cdots, \lambda_n$, 使得:

$$|\lambda_i - \alpha_{ii}^{(k_0)}| \leq \|A_{k_0} - D_{k_0}\|_2 \leq E(A_{k_0}) < \epsilon < \delta/4 \tag{3.12}$$

$$i = 1, 2, \cdots, n$$

这样只要能证明上式蕴含着

$$|\lambda_i - \alpha_{ii}^{(k_0+1)}| < \epsilon, \quad i = 1, 2, \cdots, n \tag{3.13}$$

则由归纳法原理知, 对一切的$k \geq k_0$ 有

$$|\lambda_i - \alpha_{ii}^{(k)}| < \epsilon, \quad i = 1, 2, \cdots, n$$

从而证明了(3.11).

由于A_{k_0+1} 与A_{k_0} 的对角元只可能有两个不同($\alpha_{pp}^{(k_0+1)}$ 和$\alpha_{qq}^{(k_0+1)}$), 故只需证明(3.13) 式对$i = p, q$ 成立即可. 由()()() 式,有

$$\alpha_{pp}^{(k_0+1)} = \alpha_{pp}^{(k_0)} + c^2(-2t\alpha_{pq}^{(k_0)} + t^2(\alpha_{qq}^{(k_0)} - \alpha_{pp}^{(k_0)}))$$

$$\left(\alpha_{\mathbf{qq}}^{(\mathbf{k_0})} - \alpha_{\mathbf{pp}}^{(\mathbf{k_0})} = \mathbf{2}\tau\alpha_{\mathbf{pq}}^{(\mathbf{k_0})} = \frac{\mathbf{2}\tau\mathbf{t}\alpha_{\mathbf{pq}}^{(\mathbf{k_0})}}{\mathbf{t}} = \frac{(\mathbf{1} - \mathbf{t^2})\alpha_{\mathbf{pq}}^{(\mathbf{k_0})}}{\mathbf{t}}\right)$$

$$= \alpha_{pp}^{(k_0)} + c^2(-2t\alpha_{pq}^{(k_0)} + t(1 - t^2)\alpha_{pq}^{(k_0)})$$

$$= \alpha_{pp}^{(k_0)} - t\alpha_{pq}^{(k_0)}$$

同理可证:

$$\alpha_{qq}^{(k_0+1)} = \alpha_{qq}^{(k_0)} + t\alpha_{pq}^{(k_0)}$$

这样, 对任何$\lambda_j \neq \lambda_p$, 有:

$$
\begin{aligned}
|\alpha_{pp}^{(k_0+1)} - \lambda_j| &= |\alpha_{pp}^{(k_0)} - \lambda_p + \lambda_p - \lambda_j + t\alpha_{pq}^{(k_0)}| \\
&\geq |\lambda_p - \lambda_j| - |\alpha_{pp}^{(k_0)} - \lambda_p| - |t|E(A_{k_0}) \\
&\geq \delta - \epsilon - \epsilon \geq 2\epsilon
\end{aligned}
$$

这里用到了$|t| \leq 1$. 此外,由于$\lambda(A_{k_0}) = \lambda(A)$ 和$E(A_{k_0+1}) < \epsilon$, 因此$\alpha_{pp}^{(k_0+1)}$ 必须与A 的某个特征值之间的距离小于ϵ. 这样(3.13) 对$i = p$ 成立. 类似的有$i = q$ 成立. 证毕.

上述定理给出了经典Jacobi 方法的收敛速度的一个粗略估计:

$$
E(A_k)^2 \leq (1 - \frac{1}{N})^k E(A_0)^2.
$$

循环Jacobi 方法及其变形

按照某种预先指定的顺序对每个非对角元恰好消去一次的方法称为循环Jacobi 方法. 最自然的循环Jacobi 方法按照如下顺序来扫描:

$$(p,q)=(1,2),\cdots,(1,n);(2,3),\cdots,(2,n);\cdots;(n-1,n)$$

对上述方法设置一个阈值,即只有绝对值大于该值的非对角元所在平面才需要进行Jacobi 变换,然后通过不停减少阈值达到收敛的过程.这种方法我们称为过关Jacobi 方法. 常用的选取方法有:

$$\delta_0=E(A),\quad \delta_k=\frac{\delta_{k-1}}{\sigma},\quad k=1,2,\cdots,n$$

其中$\sigma\geq n$是一个固定常数. 可以证明: 按照这样选取阈值的过关Jacobi 方法是收敛的.

Jacobi 方法有如下优点:

(1) 计算特征向量方便. 如果经过k 次变换后迭代停止了, 则有:

$$A_k = J_k^T J_{k-1}^T \cdots J_1^T A J_1 J_2 \cdots J_k$$

记:

$$Q_k = J_1 J_2 \cdots J_k$$

则有:

$$AQ_k = Q_k A_k$$

上述表明Q_k 的列向量就是A 的很好的近似特征向量.

(2) Jacobi 方法十分利于并行. 因为当$p_1 \neq p_2, q_1 \neq q_2$ 时, 对于(p_1, q_1) 做变换和(p_2, q_2) 做变换是互不影响的.

目录

二分法

二分法主要是用来求解实对称三对角矩阵的特征值的一种方法.
设实对称三对角矩阵:

$$T = \begin{pmatrix} \alpha_1 & \beta_2 & & & \\ \beta_2 & \alpha_2 & \beta_3 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_n \\ & & & \beta_n & \alpha_n \end{pmatrix}$$

考虑T 的特征值的计算. 不失一般性, 假定$\beta_i \neq 0(i = 2, \cdots, n)$, 即假定$T$ 是不可约的对称三对角阵. 否则可以将A 化为更小的不可约对称三对角矩阵.

记$p_i(\lambda)$ 为$T-\lambda I$的i 阶顺序主子式, 可以得到:

$$p_0(\lambda)\equiv 1,\quad p_1(\lambda)=\alpha_1-\lambda,$$
$$p_i(\lambda)=(\alpha_i-\lambda)p_{i-1}(\lambda)-\beta_i^2 p_{i-2}(\lambda) \tag{4.1}$$
$$i=2,\cdots,n$$

由于T 是实对称, 所以$p_i(\lambda)(i=1,\cdots,n)$ 的根都是实的.

定理： $p_i(\lambda)$ 有如下性质:

(1) 存在正数M, 使得当$\lambda > M$ 时,有$p_i(-\lambda) > 0$, 而$p_i(\lambda)$ 的符号为$(-1)^i$;

(2) 相邻两个多项式没有公共根

(3) 若$p_i(\mu) = 0$, 则$p_{i-1}(\mu)p_{i+1}(\mu) < 0$

(4) $p_i(\lambda)$ 的根全是单重的, 并且$p_i(\lambda)$ 的根严格分隔$p_{i+1}(\lambda)$ 的根

证明： 由于$p_i\lambda$ 是$T - \lambda I$的第i 阶顺序主子式, 故其首项为$(-1)^i\lambda^i$. 由此知(1) 成立.

(2) 用反证法. 假设存在某个i, 使得$p_{i-1}(\lambda)$ 与$p_i(\lambda)$ 有公共根μ, 即: $p_{i-1}(\mu) = p_i(\mu) = 0$, 则:

$$0 = p_i(\mu) = (\alpha_i - \mu)p_{i-1}(\mu) - \beta_i^2(\mu)p_{i-2}(\mu) = -\beta_i^2 \mathbf{p_{i-2}}(\mu)$$

但已假定$\beta_i \neq 0$, 故上式蕴含着$p_{i-2}(\mu) = 0$. 这样, 由

$$p_{i-2}(\mu) = p_{i-1}(\mu) = 0$$

又可推出$p_{i-3}(\mu) = 0$. 以此类推有$p_0(\mu) = 0$ 矛盾. 所以(2) 得证.

(3) 可由本定理的结论(2) 和(4.1) 式立即推出. 设$p_i(\mu)=0$, 则

$$p_{i+1}(\mu)p_{i-1}(\mu)=-\beta_{i+1}^2(p_{i-1}(\mu))^2<0$$

(4) 用数学归纳法来证明. 当$i=1$, $p_1(\lambda)=\alpha_1-\lambda$, 即$\alpha_1$ 是$p_1(\lambda)$ 的根. 另一方面, $p_2(\alpha_1)=-\beta_2^2<0$, 而本定理的结论(1) 蕴含当$\lambda$ 充分大时$p_2(-\lambda)$ 和$p_2(\lambda)$ 均大于零, 因此在$(-\infty,\alpha_1)$ 与$(\alpha_1,+\infty)$ 之内各有$p_2(\lambda)$ 的一个根. 这样, 对$i=2$ 已证明(4) 成立.

假设已经证明了(4) 对$i=k$ 成立, 即假定已经证明了$p_{k-1}(\lambda)$ 和$p_k(\lambda)$ 的根都是单根, 并且$p_{k-1}(\lambda)$ 的根严格分隔$p_k(\lambda)$ 的根. 设$p_{k-1}(\lambda)$ 和$p_k(\lambda)$ 的根分别为

$$\nu_1<\nu_2<\cdots<\nu_{k-1} \quad \text{和} \quad \mu_1<\mu_2<\cdots<\mu_k$$

则由归纳假设有

$$\mu_1<\nu_1<\mu_2<\nu_2<\cdots<\nu_{k-1}<\mu_k \tag{4.2}$$

应用递推式(4.1), 有

$$p_{k+1}(\mu_j) = -\beta_{k+1}^2 p_{k-1}(\mu_j), \qquad j = 1, \cdots, k \tag{4.3}$$

由结论(1) 和$p_{k-1}(\nu_j) = 0$ $(1 \le j \le k-1)$ 容易推出, (4.2) 蕴含着

$$(-1)^{j-1} p_{k-1}(\mu_j) > 0, \qquad j = 1, \cdots, k$$

于是, 由式得

$$(-1)^{j} p_{k+1}(\mu_j) > 0, \qquad j = 1, \cdots, k$$

再注意到对充分大的正数μ 有

$$p_{k+1}(-\mu) > 0 \quad \text{和} \quad (-1)^{k+1} p_{k+1}(\mu) > 0$$

即知在区间

$$(-\infty, \mu_1), (\mu_1, \mu_2), \cdots, (\mu_{k-1}, \mu_k), (\mu_k, +\infty)$$

内都有$p_{k+1}(\lambda)$ 的根. 这里共有$k+1$ 个区间, 而$p_{k+1}(\lambda)$ 只有$k+1$ 个根, 因此, 在每个区间内有且仅有$p_{k+1}(\lambda)$ 的一个根. 因此(4) 得证.

注解： 上述定理的第四条性质表明, n 不可约实对称三对角矩阵有n 个互不相同的特征值.

对于任意给定的实数μ, 定义$s_k(\mu)$ 是数列$p_0(\mu)$, $\cdots$, $p_k(\mu)$ 的变号数. 这里规定: 若$p_i(\mu)=0$, 则$p_i(\mu)$ 与$p_{i-1}(\mu)$ 同号(根据之前的讨论知, $p_{i-1}(\mu)$ 不可能亦为零).

定理： 假设T 为不可约实对称三对角矩阵, 则$s_k(\mu)$ $(1 \le k \le n)$ 恰好是$p_k(\lambda)$ 在区间$(-\infty,\ \mu)$ 内的根数.

证明： 用数学归纳法来证明. 当$k=1$ 时, 结论显然成立. 现假设当$k=l$ 时结论成立. 设$p_l(\lambda)$ 和$p_{l+1}(\lambda)$ 的根分别为

$$\mu_1 < \mu_2 < \cdots < \mu_l \quad 和 \quad \lambda_1 < \lambda_2 < \cdots < \lambda_{l+1}$$

则由定理的结论(4) 知

$$\lambda_1 < \mu_1 < \lambda_2 < \mu_2 < \cdots < \mu_l < \lambda_{l+1} \tag{4.4}$$

再设$s_l(\mu) = m$, 则由归纳法假设有

$$\mu_{\mathbf{m}} < \mu \le \mu_{\mathbf{m+1}}$$

这样由上式知μ 所在的位置有两种可能性:

$$\lambda_{\mathbf{m}} < \mu_{\mathbf{m}} < \mu \le \lambda_{\mathbf{m+1}} \quad 或 \quad \lambda_{\mathbf{m+1}} < \mu \le \mu_{\mathbf{m+1}}$$

注意到

$$\mathbf{p_l}(\mu) = \mathbf{\Pi_{i=1}^{l}}(\mu_{\mathbf{i}} - \mu), \qquad \mathbf{p_{l+1}}(\mu) = \mathbf{\Pi_{i=1}^{l+1}}(\lambda_{\mathbf{i}} - \mu) \tag{4.5}$$

则当$\lambda_m < \mu_m < \mu \le \lambda_{m+1}$ 成立时，知$p_l(\mu)$ 与$p_{l+1}(\mu)$ 同号(即使$\mu = \lambda_{m+1}$，按规定亦有此两数同号).
这样$s_{l+1}(\mu) = s_l(\mu) = m$，正好是$p_{l+1}(\lambda)$ 在区间$(-\infty, \mu)$ 内根的个数.

当$\lambda_{m+1} < \mu \le \mu_{m+1}$ 成立时，分两种情况来证$s_{l+1}(\mu) = m + 1$:

(1) 若$\mu < \mu_{m+1}$，则由(4.5) 式知，此时$p_l(\mu)$ 与$p_{l+1}(\mu)$ 异号，因而$s_{l+1}(\mu) = s_l(\mu) + 1 = m + 1$.

(2) 若$\mu = \mu_{m+1}$，则此时有$p_l(\mu) = 0$. 于是按规定此时$p_l(\mu)$ 与$p_{l-1}(\mu)$ 同号. 而定理的结论(3) 表明，此时$p_{l-1}(\mu)$ 与$p_{l+1}(\mu)$ 异号，因此$p_{l+1}(\mu)$ 与$p_l(\mu)$ 异号，即$s_{l+1}(\mu) = s_l(\mu) + 1 = m + 1$.

这样，就证明了无论哪种情况，都有$s_{l+1}(\mu)$ 正好等于$p_{l+1}(\mu)$ 在区间$(-\infty, \mu)$ 内根的个数. 由归纳法原理知定理得证.

在上述定理中令$k = n$ 即得如下结论:

推论： 若T 是不可约的实对称三对角矩阵, 则$s_n(\mu)$ 正好是该矩阵在区间$(-\infty, \mu)$ 内特征值的个数.

利用这一推论, 可以用二分法来求T 的任何一个指定的特征值.
设T 的特征值为

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n$$

则必有

$$|\lambda_i| < \rho(T) \leq \|T\|_\infty$$

假定希望求T 的第m 个特征值λ_m. 先取

$$l_0 = -\|T\|_\infty, \qquad u_0 = \|T\|_\infty$$

则λ_m 必在区间$[l_0, u_0]$ 内. 取$[l_0, u_0]$ 的中点$r_1 = (l_0 + u_0)/2$, 并计算$s_n(r_1)$.

(1) 若$s_n(r_1) \geq m$, 则$\lambda_m \in [l_0, r_1]$, 于是取$l_1 = l_0, u_1 = r_1$;
(2) 否则, $\lambda_m \in [r_1, u_0]$, 于是取$l_1 = r_1, u_1 = u_0$;
(3) 依此做下去, 当区间足够小时可以用该区间的中点代替λ_m.

为了避免高阶多项式的计算容易发生溢出, 定义

$$q_i(\lambda) = \frac{p_i(\lambda)}{p_{i-1}(\lambda)}, \quad i = 1, \cdots, n$$

利用公式(4.1) 可得:

$$q_1(\lambda) = p_1(\lambda) = \alpha_1 - \lambda$$

$$q_i(\lambda) = \alpha_i - \lambda - \frac{\beta_i^2}{q_{i-1}(\lambda)}, \quad i = 2, \cdots, n$$

由此知, $s_n(\mu)$ 正好是数列$q_1(\mu), \cdots, q_n(\mu)$ 中负数的个数.

计算变号数算法

$x = [\alpha_1, \alpha_2, \cdots, \alpha_n]$

$y = [0, \beta_2, \cdots, \beta_n]$

$s = 0; \quad q = x(1) - \mu$

for $k = 1 : n$

    if $q < 0$

        $s = s + 1$

    end

    if $k < n$

        if $q = 0$

            $q = |y(k+1)|\mathbf{u}$

        end

        $q = x(k+1) - \mu - y(k+1)^2/q$

    end

end

目录

分而治之法

分而治之法是求实对称三对角阵的全部特征值和特征向量的一种数值方法, 是由 Dongarra 和 Sorensen 于 1987 年首先提出的.

基本思想:

(1) 先将给定的对称三对角阵“分割”成 2^k 个低阶的对称三对角阵;

(2) 然后分别求出每个低阶的对称三对角阵的谱分解;

(3) 最后将这些低阶谱分解“胶合”在一起得到原矩阵的谱分解.

因此, 这一方法特别适用于并行计算. 设

$$T = \begin{bmatrix} \alpha_1 & \beta_2 & & & & \\ \beta_2 & \alpha_2 & \beta_3 & & & \\ & \beta_3 & \alpha_3 & \beta_4 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \beta_{n-1} & \alpha_{n-1} & \beta_n \\ & & & & \beta_n & \alpha_n \end{bmatrix} \tag{5.1}$$

是一个给定的实对称三对角阵, 并假定 $n = 2^m$.

分割

定义 $v \in \mathbf{R}^n$ 为

$$v = (0, \ \cdots, \ 0, \ \underset{m}{1}, \ \underset{m+1}{\theta}, \ 0, \ \cdots, \ 0)^T, \tag{5.2}$$

并考虑矩阵 $\tilde{T} = T - \rho v v^T$, 其中 θ 和 ρ 是待定实数. 易知, $\tilde{T}$ 除了中间四个元素为

$$\begin{bmatrix} \alpha_m - \rho & \beta_{m+1} - \rho\theta \\ \beta_{m+1} - \rho\theta & \alpha_{m+1} - \rho\theta^2 \end{bmatrix}$$

外, 其余元素与 T 的完全一样. 因此, 假如我们取 $\rho\theta = \beta_{m+1}$, 则有

$$T = \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix} + \rho v v^T, \tag{5.3}$$

$$T_1 = \begin{bmatrix} \alpha_1 & \beta_2 & & & & & \\ \beta_2 & \alpha_2 & \beta_3 & & & & \\ & \beta_3 & \alpha_3 & \beta_4 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \beta_{m-1} & \alpha_{m-1} & \beta_m \\ & & & & \beta_m & \tilde{\alpha}_m \end{bmatrix}$$

$$T_2 = \begin{bmatrix} \tilde{\alpha}_{m+1} & \beta_{m+2} & & & \\ \beta_{m+2} & \alpha_{m+2} & \beta_{m+3} & & \\ & \beta_{m+3} & \ddots & \ddots & \\ & & \ddots & \alpha_{n-1} & \beta_n \\ & & & \beta_n & \alpha_n \end{bmatrix}$$

这里 $\tilde{\alpha}_m - \rho$, $\tilde{\alpha}_{m+1} = \alpha_{m+1} - \rho\theta^2$. 这样, 将 T 分割为一个分块矩阵和一个秩 1 矩阵的和. 如果需要, 还可以对 T_1 和 T_2 进行形如 (5.3) 式的分割. 如此下去, 就可将 T 分割为 2^k 块.

胶合

假定已经求得 T_1 和 T_2 的谱分解

$$Q_1^T T_1 Q_1 = D_1, \quad Q_2^T T_2 Q_2 = D_2,$$

其中 Q_1 和 Q_2 是 m 阶正交矩阵, D_1 和 D_2 是对角阵.

下面利用 T_1 和 T_2 的谱分解来求出 T 的谱分解, 即求正交矩阵 V, 使得

$$V^T T V = \mathrm{diag}(\lambda_1, \cdots, \lambda_n). \tag{5.4}$$

令

$$U=\begin{bmatrix}Q_1 & 0\\ 0 & Q_2\end{bmatrix},$$

则

$$\begin{aligned}U^TTU&=\begin{bmatrix}Q_1 & 0\\ 0 & Q_2\end{bmatrix}^T\left(\begin{bmatrix}T_1 & 0\\ 0 & T_2\end{bmatrix}+\rho vv^T\right)\begin{bmatrix}Q_1 & 0\\ 0 & Q_2\end{bmatrix}\\&=D+\rho zz^T,\end{aligned}\tag{5.5}$$

其中,

$$z=U^Tv,\quad D=\mathrm{diag}(D_1,D_2).$$

这样一来, 欲求 T 的谱分解的问题就归结为求 $D+\rho zz^T$ 的谱分解的问题.

引理： 设$D = \mathrm{diag}(d_1, \ldots, d_n) \in \mathbf{R}^{n\times n}$ 满足

$$d_1 > d_2 > \cdots > d_n$$

并假定ρ 是一个非零实数，$z \in \mathbf{R}^n$ 的分量均不为零. 如果$u \in \mathbf{R}^n$ 和$\lambda \in \mathbf{R}$ 满足

$$(D + \rho zz^T)u = \lambda u, \quad u \neq 0$$

则$z^Tu \neq 0$，而且$D - \lambda I$非奇异.

证明： 若$z^Tu = 0$，则$Du = \lambda u$，$u \neq 0$，即λ 是D 的特征值，u 是属于λ 的特征向量. 而已知D 是对角元互不相同的对角阵，故必有某个i，使得$d_i = \lambda$，而且$u = \alpha e_i$，$\alpha \neq 0$. 这样便有

$$0 = z^Tu = \alpha z^T e_i$$

这与z 的分量均不为零的假定矛盾. 因此，必有$z^Tu \neq 0$.

此外，若$D-\lambda I$奇异，则必有某个i，使得$e_i^T(D-\lambda I)=0$，从而有

$$0=e_i^T(D-\lambda I)u=-\rho z^T u e_i^T z$$

但$\rho z^T u\neq 0$，故必有$e_i^T z=0$. 这亦与z 的分量均不为零矛盾.

注解： 在上述引理的条件下，u 不是D 的特征向量. 否则, $u=\alpha e_i$. 于是

$$(D+\rho zz^T)e_i=\lambda e_i, z=(\lambda e_i-De_i)/(\rho z^T e_i).$$

这和z 的分量均不为零矛盾.

定理： 在上一个引理的假设条件下，再假定$D+\rho zz^T$ 的谱分解为

$$V^T(D+\rho zz^T)V = \mathrm{diag}(\lambda_1,\ldots,\lambda_n)$$

其中$V=[v_1,\ldots,v_n]$ 为正交阵，$\lambda_1\geq\cdots\geq\lambda_n$，则有

(1) $\lambda_1,\ldots,\lambda_n$ 正好是函数

$$f(\lambda)=1+\rho z^T(D-\lambda I)^{-1}z$$

的n 个零点.

(2) 当$\rho>0$ 时，有$\lambda_1>d_1>\lambda_2>d_2>\cdots>\lambda_n>d_n$; 而当$\rho<0$ 时，有$d_1>\lambda_1>d_2>\lambda_2>\cdots>d_n>\lambda_n$.

(3) 存在常数$\alpha_i\neq 0$，使得$v_i=\alpha_i(D-\lambda_i I)^{-1}z$，$i=1,\ldots,n$.

证明： 由已知条件知

$$(D+\rho zz^T)v_i=\lambda_i v_i,\quad \|v_i\|_2=1$$

于是从前述引理知，$D-\lambda_i I$非奇异，从而有

$$v_i=-\rho z^T v_i(D-\lambda_i I)^{-1}z,\quad i=1,\ldots,n \tag{5.6}$$

这就证明了(3) 成立，同时也证明了$D+\rho zz^T$ 的特征值互不相同(否则，如果$\lambda_i=\lambda_j$，则有v_i 与v_j 线性相关，这与v_i 与v_j 相互正交矛盾)

在 (5.6) 式两边左乘z^T，并注意到$z^T v_i\neq 0$，即有

$$1=-\rho z^T(D-\lambda_i I)^{-1}z$$

即

$$f(\lambda_i)=0,\quad i=1,\ldots,n$$

这说明λ_i $(i=1,\ldots,n)$ 均是$f(\lambda)$ 的零点.

下面来证$f(\lambda)$ 正好有n 个零点.
设$z = (\xi_1, \ldots, \xi_n)^T$，则

$$\mathbf{f}(\lambda) = \mathbf{1} + \rho\left(\frac{\xi_1^2}{\mathbf{d_1} - \lambda} + \cdots + \frac{\xi_n^2}{\mathbf{d_n} - \lambda}\right)$$

于是有

$$\mathbf{f}'(\lambda) = \rho\left[\frac{\xi_1^2}{(\mathbf{d_1} - \lambda)^2} + \cdots + \frac{\xi_n^2}{(\mathbf{d_n} - \lambda)^2}\right]$$

因此，$f(\lambda)$ 在任意相邻两极点d_i 和d_{i+1} 之间是严格单调的：
(1) $\rho > 0$ 时，严格增加；
(2) $\rho < 0$ 时，严格减少.
由此易知，$f(\lambda)$ 正好有n 个零点，而且当$\rho > 0$ 时，它们正好分别位于如下的n 个区间：

$$(\mathbf{d_n}, \mathbf{d_{n-1}}), \ldots, (\mathbf{d_2}, \mathbf{d_1}), (\mathbf{d_1}, +\infty);$$

当$\rho < 0$ 时，它们正好分别位于如下的n 个区间：

$$(-\infty, d_n), (d_n, d_{n-1}), \ldots, (d_2, d_1);$$

由此立即知定理的结论(1) 和(2) 成立.

在前面的条件下，可以按照如下两步快速、稳定地求出$D+\rho zz^T$ 的谱分解：

第一步，求$f(\lambda)$ 的零点$\lambda_1,\ldots,\lambda_n$. 由于每一区间内仅有$f(\lambda)$ 的唯一零点，而且$f(\lambda)$ 在区间内严格单调，因此这一步可以应用Newton 类型的算法快速、稳定的实现.

第二步，计算

$$v_i = \frac{(D-\lambda_i I)^{-1} z}{\|(D-\lambda_i I)^{-1} z\|_2}, \quad i = 1,\ldots,n$$

这一步当然亦可以快速、稳定地实现.

对于一般$D+\rho zz^T$ 的谱分解亦可以归结为前述定理的情形. 为此，构造性的证明下面的结果.

定理： 设$D=\mathrm{diag}(d_1,\ldots,d_n)\in\mathbf{R}^{n\times n}$, $z\in\mathbf{R}^n$, 则存在正交矩阵V 和$\{1,\ldots,n\}$ 的一个排列π，使得

(1) $V^Tz=(\xi_1,\ldots,\xi_r,\underbrace{0,\ldots,0}_{n-r\text{个}})^T$ 满足$\xi_i\neq 0\ (i=1,\ldots,r)$;

(2) $V^TDV=\ \mathrm{diag}(d_{\pi(1)},\ldots,d_{\pi(n)})$ 满足

$$d_{\pi(1)}>d_{\pi(2)}>\cdots>d_{\pi(r)}$$

证明： 如果有两个指标$i < j$，使得$d_i = d_j$，则可取(i, j) 坐标平面内的平面旋转变换$P_{ij} = G(i, j, \theta)$，使得$P_{ij}z$ 的第j 个分量为零，而且

$$P_{ij}^T D P_{ij} = D$$

这样进行若干步之后，就可以找到一个由一些平面旋转变换的乘积构成的正交矩阵V_1，使得

$$V_1^T D V_1 = D$$

而且

$$V_1^T z = (\xi_1, \ldots, \xi_n)^T$$

满足：若$\xi_i \xi_j \neq 0 \ (i \neq j)$，则必有$d_i \neq d_j$.

然后，对$V_1^T z$的分量进行若干次两两对换，可使其所有不为零的分量都位于它的前面，即可找到一个排列方阵P_1，使得

$$P_1 V_1^T z = (\xi_{\pi_1(1)}, \ldots, \xi_{\pi_1(n)})^T$$

$$\xi_{\pi_1(i)} \neq 0, \quad i = 1, \ldots, r$$

$$\xi_{\pi_1(i)} = 0, \quad i = r+1, \ldots, n$$

其中π_1 是$\{1, \ldots, n\}$ 的一个排列. 由P_1 的取法知，矩阵

$$P_1^T V_1^T D V_1 P_1 = P_1^T D P_1 = \mathrm{diag}(d_{\pi_1(1)}, \ldots, d_{\pi_1(n)})$$

的前r 个对角元$d_{\pi_1(1)}, \ldots, d_{\pi_1(r)}$ 互不相同.

最后，对$d_{\pi_1(1)},\ldots,d_{\pi_1(r)}$ 进行若干次对换，使它们按从大到小的次序排列，即可找到一个r 阶排列方阵P_2，使得

$$P_2^T \mathrm{diag}(d_{\pi_1(1)},\ldots,d_{\pi_1(r)})P_2 = \mathrm{diag}(\mu_1,\ldots,\mu_r)$$

其中$\mu_1,\mu_2,\ldots,\mu_r$ 是由$d_{\pi_1(1)},\ldots,d_{\pi_1(r)}$ 从大到小排列而得到的，即

$$\mu_1 > \mu_2 > \cdots > \mu_r$$

现令

$$\begin{aligned} V &= V_1 P_1 \mathrm{diag}(P_2, I_{n-r}) \\ V^T z &= (\xi_1,\ldots,\xi_n)^T \\ V^T D V &= \mathrm{diag}(d_{\pi(1)},\ldots,d_{\pi(n)}) \end{aligned} \tag{5.7}$$

其中π 是由P_1 和P_2 决定的$\{1,\ldots,n\}$ 的一个排列，则有

$$\xi_i \neq 0, i = 1,\ldots,r \quad \xi_{r+1} = \cdots = \xi_n = 0$$

$$d_{\pi(1)} > d_{\pi(2)} > \cdots > d_{\pi(r)}$$

即定理得证.

由上述定理可知, 对于任意$D = \mathrm{diag}(d_1, \ldots, d_n) \in \mathbf{R}^n$ 和$z \in \mathbf{R}^n$ 可构造出一个正交矩阵V，使得

$$V^T(D + \rho zz^T)V = \begin{array}{cc} \left[\begin{array}{cc} D_1 + \rho ww^T & 0 \\ 0 & D_2 \end{array}\right] & \begin{array}{c} r \\ n-r \end{array} \\ \begin{array}{cc} \quad r \quad & \quad n-r \end{array} & \end{array}$$

其中

$$\begin{aligned} D_1 &= \mathrm{diag}(d_{\pi(1)}, \ldots, d_{\pi(r)}), \quad d_{\pi(1)} > d_{\pi(2)} > \cdots > d_{\pi(r)} \\ D_2 &= \mathrm{diag}(d_{\pi(r+1)}, \ldots, d_{\pi(n)}) \\ w &= (\xi_1, \ldots, \xi_r)^T, \xi \neq 0, \quad i = 1, \ldots, r \end{aligned}$$

因此，要求$D+\rho zz^T$ 的谱分解只需求出$D_1+\rho ww^T$ 的谱分解即可，而后者已是再前面的定理所述的情形，可以快速稳定的求出.

在实际计算时，需要事先给定一个准则，来判定何时两数视为相等，何时一个数视为零. 例如，可取

$$\epsilon_1 = (\|D\|_2 + |\rho|\|z\|_2)\mathbf{u}$$

来作为误差限，当$|d_i - d_j| < \epsilon_1$ 时，就认为d_i 和d_j 相等；而当$|\xi_i| < \epsilon_1$ 时，就认为ξ_i 为零

并行计算

下面简要的说明一下如何将这一方法应用于并行计算. 为了叙述简单起见，假定是在四个处理器的并行机上计算一个$4N$ 阶的对称三对角阵T 的谱分解. 整个计算过程可分为如下四步：

(1) 分割：

$$T = \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix} + \rho v v^T, \quad T_1 \in \mathbf{R}^{2N\times 2N}, v \in \mathbf{R}^{4N}$$

$$T_i = \begin{bmatrix} T_{i1} & 0 \\ 0 & T_{i2} \end{bmatrix} + \rho_i w_i w_i^T, \quad T_{ij} \in \mathbf{R}^{N\times N}, w_i \in \mathbf{R}^{2N}, i = 1, 2$$

这步仅需计算少数几个数.

(2) 将$T_{11}, T_{12}, T_{21}, T_{22}$ 分别分配給四个处理器，去求其谱分解(例如可以用隐式对称QR 算法实现).

(3) 将T_{11} 和T_{12} 的谱分解以及T_{21} 和T_{22} 的谱分解分别胶合成T_1 和T_2 的谱分解. 这一步，由于胶合过程主要是求形如$D+\rho zz^T$ 的矩阵的特征值和特征向量，而这些特征值的计算基本上是相互独立的，因此亦可分配给四个处理器同时进行.

(4) 将T_1 和T_2 的谱分解胶合成T 的谱分解. 这一步亦可分配给四个处理器同时进行.

从上面的讨论可以看出，分而治之法并行的效率是很高的. 因此，它适用于在并行机上求解大型对称三对角矩阵的全部特征值和特征向量.

目录

二对角化

对应于将A^TA 三对角化，这里是将A 二对角化，即计算两个正交矩阵U 和V，使得

$$U^TAV = \begin{bmatrix} B \\ 0 \end{bmatrix}, \quad B = \begin{bmatrix} \delta_1 & \gamma_1 & & \\ & \ddots & \ddots & \\ & & \ddots & \gamma_{n-1} \\ & & & \delta_n \end{bmatrix} \tag{6.1}$$

一旦分解式 (6.1) 已经实现，则有$V^TA^TAV = B^TB$ 是一个对称三对角阵. 这就相当于已经把A^TA 三对角化，正好对应于应用对称QR 方法于A^TA 上的第一步. **但这里并不需要将A^TA 真正计算出来.**

分解式 (6.1) 可以借助Householder 变换来实现. 首先确定一个n阶Householder 变换P_1，使得

$$P_1^T A = \begin{bmatrix} \times & \times & \times & \dots & \times \\ 0 & \times & \times & \dots & \times \\ 0 & \times & \times & \dots & \times \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & \times & \times & \dots & \times \end{bmatrix}$$

然后确定一个$n-1$ 阶Householder 变换H_1，使得

$$P_1^T A \begin{bmatrix} 1 & 0 \\ 0 & H_1 \end{bmatrix} = \begin{bmatrix} \times & \times & 0 & \dots & 0 \\ 0 & \times & \times & \dots & \times \\ 0 & \times & \times & \dots & \times \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & \times & \times & \dots & \times \end{bmatrix}$$

如此继续下去，就可以通过一系列Householder 变换把A 化成二对角矩阵.

Algorithm 3 (二对角化: Householder 变换法)

for $k = 1 : n$

  $[v, \beta] =$ **house**$(A(k : m, k))$

  $u^T = (\beta v^T)A(k : m, k : n)$

  $A(k : m, k : n) = A(k : m, k : n) - vu^T$

  $A(k + 1 : m, k) = v(2 : m - k + 1); b(k) = \beta$

  if $k < n - 1$

    $[v, \beta] =$ **house**$(A(k, k + 1 : n)^T)$

    $u = A(k : m, k + 1 : n)(\beta v)$

    $A(k : m, k + 1 : n) = A(k : m, k + 1 : n) - vu^T$

    $A(k, k + 2 : n) = v(2 : n - k)^T; c(k) = \beta$

  end

end

这一算法所需运算量为$4mn^2 - 4n^3/3$. 如果需要积累U 和V，则还需要增加的运算量分别为$4m^2n - 4n^3/3$ 和$4n^3/3$.

SVD 迭代

应用对称QR方法的第二步就是对对称三对角阵$T \equiv B^T B$进行带位移的隐式QR 迭代. 这里也不用明确地把T 算出就可以完成这一计算任务.

应用带位移的对称QR迭代于$T = B^T B$ 上，首先要选取位移μ. 矩阵T 的右下角的2 阶矩阵为

$$\begin{bmatrix} \delta_{n-1}^2 + \gamma_{n-2}^2 & \delta_{n-1}\gamma_{n-1} \\ \delta_{n-1}\gamma_{n-1} & \delta_n^2 + \gamma_{n-1}^2 \end{bmatrix}$$

可以选取该矩阵的两个特征值中靠近$\delta_n^2 + \gamma_{n-1}^2$ 的那个作为位移μ. 当然这一步我们亦不需要把T 计算出来即可进行.

迭代的下一步就是确定一个Givens 变换$G_1 = G(1,2,\theta_1)$，即确定$s_1 = sin\theta_1$ 和$c_1 = cos\theta_1$ ，满足

$$\begin{bmatrix} c_1 & s_1 \\ -s_1 & c_1 \end{bmatrix}^T \begin{bmatrix} \delta_1^2 - \mu \\ \delta_1\gamma_1 \end{bmatrix} = \begin{bmatrix} * \\ 0 \end{bmatrix}$$

这里$(T - \mu I)e_1 = (\delta_1^2 - \mu, \delta_1\gamma_1, \underbrace{0, \ldots, 0}_{n-r个})^T$. 显然，这一步也不需要把$T$ 明确的计算出来.

迭代的最后一步就是确定一个正交矩阵Q，使得

$$Q^T(G_1^T T G_1)Q$$

是对称三对角矩阵，且

$$Qe_1 = e_1$$

这一步，为了避免$T = B^T B$的计算，只需计算正交矩阵P和Q，使得$P^T(BG_1)Q$ 是二对角阵，且$Qe_1 = e_1$ 即可. 这一步可利用Givens 变换来完成，具体约化过程可以从如下$n = 4$ 的例子中明白：

设$n = 4$. 注意，此时$C = BG_1$ 有如下形状：

$$C = \begin{bmatrix} \times & \times & 0 & 0 \\ + & \times & \times & 0 \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \end{bmatrix}$$

因此，接下来要做的就是，先给C 左乘一个(1,2) 平面的Givens 变换$G_2 = G(1, 2, \theta)$ 将其(2,1) 位置的元素化为零，而这样同时也会在(1,3) 位置上引进一个可能的非零元，即

$$C \leftarrow G_2C = \begin{bmatrix} \times & \times & + & 0 \\ 0 & \times & \times & 0 \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \end{bmatrix}$$

然后对C 右乘一个Givens 变换$G_3 = G(2,3,\theta_3)$ 将其(1,3) 位置的元素化为零，同时在(3,2) 位置上又会引进一个可能的非零元，即

$$C \leftarrow CG_3 = \begin{bmatrix} \times & \times & 0 & 0 \\ 0 & \times & \times & 0 \\ 0 & + & \times & \times \\ 0 & 0 & 0 & \times \end{bmatrix}$$

类似地，对C 左乘一个Givens 变换$G_4 = G(2,3,\theta_4)$ 将其(3,2) 位置的元素化为零，同时在(2,4) 位置上又会引进一个可能的非零元，即

$$C \leftarrow G_4C = \begin{bmatrix} \times & \times & 0 & 0 \\ 0 & \times & \times & + \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \end{bmatrix}$$

再对C 右乘一个Givens 变换$G_5 = G(3,4,\theta_5)$ 将其(2,4) 位置的元素化为零，同时在(4,3) 位置上又会引进一个可能的非零元，即

$$C \leftarrow CG_5 = \begin{bmatrix} \times & \times & 0 & 0 \\ 0 & \times & \times & 0 \\ 0 & 0 & \times & \times \\ 0 & 0 & + & \times \end{bmatrix}$$

最后，对C 左乘一个Givens 变换$G_6 = G(3,4,\theta_6)$ 将其(4,3) 位置的元素化为零，即有

$$C \leftarrow G_6 C = \begin{bmatrix} \times & \times & 0 & 0 \\ 0 & \times & \times & 0 \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \end{bmatrix}$$

是二对角阵.

由前面对对称QR迭代的讨论，采用隐式QR方法的前提是$T = BB^T$ 必须是不可约的(即T的次对角元均不为零). 而T 的次对角线为$\delta_j\gamma_j$，因此T 不可约的充分必要条件是

$$\delta_j\gamma_j \neq 0, \quad j = 1, \ldots, n-1 \tag{6.2}$$

当某个$\gamma_j = 0$ 时，B 具有如下形状：

$$B = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}$$

因此，此时可以把B 的奇异值分解的计算问题分解为两个低阶二对角阵B_1 和B_2 的奇异值分解的计算问题.

当某个$\delta_j = 0$ 而$\gamma_j \neq 0$ 时，可以通过适当的Givens 变换把B 的第j 行元素都变成零，而保持其二对角形式不变. 这一过程可以从如下$n = 5, j = 2$ 的例子中明白：假定$n = 5, j = 2$. 此时B 有如下形状:

$$B = \begin{bmatrix} \times & \times & 0 & 0 & 0 \\ 0 & 0 & \times & 0 & 0 \\ 0 & 0 & \times & \times & 0 \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & \times \end{bmatrix}$$

首先对B 左乘一个Givens 变换$G_1 = G(2, 3, \theta_1)$ 将其(2,3) 位置的元素化为零，同时在(2,4) 位置上又会引进一个可能的非零元，即

$$B \leftarrow G_1 B = \begin{bmatrix} \times & \times & 0 & 0 & 0 \\ 0 & 0 & 0 & + & 0 \\ 0 & 0 & \times & \times & 0 \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & \times \end{bmatrix}$$

再左乘一个Givens 变换$G_2 = G(2,4,\theta_2)$ 将其(2, 4)位置的元素化为零，同时在(2,5) 位置上又会引进一个可能的非零元，即

$$B \leftarrow G_2 B = \begin{bmatrix} \times & \times & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & + \\ 0 & 0 & \times & \times & 0 \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & \times \end{bmatrix}$$

最后，左乘一个Givens 变换$G_3 = G(2,5,\theta_3)$ 将其(2,5) 位置的元素化为零，即有

$$B \leftarrow G_3 B = \begin{bmatrix} \times & \times & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \times & \times & 0 \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & \times \end{bmatrix}$$

这样就又可以把B 的奇异值分解的计算问题分解为两个低阶的二对角阵的奇异值分解的计算问题.

Algorithm 4 (带 Wilkinson 位移的SVD 迭代)

$\alpha = \delta_n^2 + \gamma_{n-1}^2; \delta = (\delta_{n-1}^2 + \gamma_{n-2}^2 - \alpha)/2; \beta = \delta_{n-1}\gamma_{n-1}$
$\mu = \alpha - \beta^2/(\delta + \textbf{sign}(\delta)\sqrt{\delta^2 + \beta^2})$
$y = \delta_1^2 - \mu; z = \delta_1\gamma_1$
for $k = 1 : n-1$
  确定$c = \cos\theta$ 和$s = \sin\theta$，使得

$$\begin{bmatrix} y & z \end{bmatrix} \begin{bmatrix} c & s \\ -s & c \end{bmatrix} = \begin{bmatrix} \gamma_{k-1} & 0 \end{bmatrix}$$

$Q_c(k) = c; Q_s(k) = s$(记录形成Q 的Givens 变换)

$$\begin{bmatrix} y & \gamma_k \\ z & \delta_{k+1} \end{bmatrix} = \begin{bmatrix} \delta_k & \gamma_k \\ 0 & \delta_{k+1} \end{bmatrix} \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$$

Algorithm 5 (带 Wilkinson 位移的SVD 迭代: 续)

确定$c = \cos\theta$ 和$s = \sin\theta$，使得

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix}^T \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} \delta_k \\ 0 \end{bmatrix}$$

$P_c(k) = c; P_s(k) = s$(记录形成P 的Givens 变换)

if $k < n-1$

$$\begin{bmatrix} y & z \\ \delta_{k+1} & \gamma_{k+1} \end{bmatrix} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}^T \begin{bmatrix} \gamma_k & 0 \\ \delta_{k+1} & \gamma_{k+1} \end{bmatrix}$$

else

$$\begin{bmatrix} \gamma_k \\ \delta_{k+1} \end{bmatrix} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}^T \begin{bmatrix} \gamma_k \\ \delta_{k+1} \end{bmatrix}$$

end

end

这一算法需要$30n$ 次四则运算和$2n$ 次开方运算.

SVD 算法

在实际运算中，当δ_j 或γ_j 很小时，就可以把B 分解为两个低阶的二对角阵. 通常使用的准则是：当

$$|\delta_j| \leq \epsilon \|B\|_\infty \quad 或 \quad |\gamma_j| \leq \epsilon(|\delta_j| + |\delta_{j+1}|)$$

时，就将δ_j 或γ_j 视为零，其中ϵ 是一个略大于机器精度的正数.

Algorithm 6 (SVD 算法)

(1) 输入$A \in \mathbf{R}^{m\times n}$，$m \geq n$
(2) 二对角化：应用上述算法到A 上，产生二对角阵B 和正交矩阵U 和V，使得$U^TAV = \begin{bmatrix} B \\ 0 \end{bmatrix}$
(3) 收敛性判定：
(i) 把所有满足条件

$$|b_{i,i+1}| \leq (|b_{i,i}| + |b_{i+1,i+1}|)\epsilon$$

的$b_{i,i+1}$ 置为零，即$b_{i,i+1} = 0$
(ii) 把所有满足条件

$$|b_{ii}| \leq \|B\|_\infty \epsilon$$

的b_{ii} 置为零，即$b_{i,i} = 0$

Algorithm 7 (SVD 算法: 续)

(iii) 确定最大的非负整数p 和最小的非负整数q 使得

$$B = \text{diag}(B_{11}, B_{22}, B_{33})$$

其中$B_{11} \in \mathbf{R}^{p\times p}$，$B_{33}$是$q$ 阶对角阵，而B_{22} 的对角元之上的次对角元均不为零.
(iv) 如果$q = n$，则输出有关信息，结束；否则进行下一步.
(4) SVD 迭代：
(i) 若B_{22} 有对角元为零(最后一个对角元除外)，则利用前面所介绍的方法将其对应行的元素均化为零，并且将相应的变换矩阵都累积到U 上，然后转步(2); 否则，进行下一步.
(ii) 应用上述算法到B_{22} 上，产生正交矩阵P 和Q 以及二对角矩阵$B_{22} = P^T B_{22} Q$，并且计算

$$U = U diag(I_p, P, I_{q+m-n}), \quad V = V diag(I_p, Q, I_q)$$

然后转步(3)