第六章 非对称特征值问题的计算方法

杜磊
dulei@dlut.edu.cn

大连理工大学 数学科学学院
创新园大厦 B1207

2019 年 11 月 5 日

内容提要

矩阵特征值问题

线性特征值问题

- 标准特征值问题: $Ax = \lambda x,\ \ A \in \mathbb{C}^{n\times n}$;
- 广义特征值问题: $Ax = \lambda Bx,\ \ A, B \in \mathbb{C}^{n\times n}$;

G.H. Golub & H.A. van der Vorst
Eigenvalue computation in the 20th century
J. Comput. Appl. Math., 123:1-2(2000), 35-65.

矩阵特征值问题

线性特征值问题

- 标准特征值问题: $Ax = \lambda x, \ \ A \in \mathbb{C}^{n\times n}$;
- 广义特征值问题: $Ax = \lambda Bx, \ \ A, B \in \mathbb{C}^{n\times n}$;

G.H. Golub & H.A. van der Vorst
Eigenvalue computation in the 20th century
J. Comput. Appl. Math., 123:1-2(2000), 35-65.

非线性特征值问题 ($T(\lambda)x = 0$)

- $(\lambda^m A_m + \lambda^{m-1} A_{m-1} + \cdots + \lambda A_1 + A_0)x = 0$;
- $\left(K - \lambda M + i\sqrt{\lambda - \omega_1^2}\, W_1 + i\sqrt{\lambda - \omega_2^2}\, W_2\right)x = 0$;
- $(\lambda I + A + Be^{-\lambda})x = 0$;
- $\cdots\cdots$

S. Güttel & F. Tisseur
The nonlinear eigenvalue problem
Acta Numerica, 26(2017), 1-94.

标准特征值问题

设 $A \in \mathbb{C}^{n\times n}$, 称使得

$$Ax = \lambda x$$

成立的复数 λ 为矩阵 A 的**特征值**, 非零向量 x 为 λ 对应的**右特征向量**.

标准特征值问题

设 $A \in \mathbb{C}^{n\times n}$, 称使得

$$Ax = \lambda x$$

成立的复数 λ 为矩阵 A 的特征值, 非零向量 x 为 λ 对应的右特征向量. 记 A 的特征值的全体为 $\lambda(A)$, 通常称之为 A 的谱集.

标准特征值问题

设 $A \in \mathbb{C}^{n\times n}$, 称使得

$$Ax = \lambda x$$

成立的复数 λ 为矩阵 A 的特征值, 非零向量 x 为 λ 对应的右特征向量. 记 A 的特征值的全体为 $\lambda(A)$, 通常称之为 A 的谱集. 易知, λ 是 A 的一个特征值的充分必要条件是 $\det(\lambda I - A) = 0$, 称 $P_A(\lambda) = \det(\lambda I - A)$ 为 A 的特征多项式, 其 n 个根即为 A 的 n 个特征值.

标准特征值问题

假定 $P_A(\lambda) = \det(\lambda I - A)$ 有如下分解:

$$P_A(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2}\cdots(\lambda - \lambda_p)^{n_p},$$

其中 $n_1 + n_2 + \cdots + n_p = n, \lambda_i \neq \lambda_j (i \neq j)$,

标准特征值问题

假定 $P_A(\lambda) = \det(\lambda I - A)$ 有如下分解:

$$P_A(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_p)^{n_p},$$

其中 $n_1 + n_2 + \cdots + n_p = n, \lambda_i \neq \lambda_j (i \neq j)$, 则称 n_i 为 λ_i 的代数重数(简称重数), 称

$$m_i = n - \operatorname{rank}(\lambda_i I - A)$$

为 λ_i 的几何重数. 易知 $m_i \leq n_i \, (i = 1, 2, \cdots, p)$.

标准特征值问题

假定 $P_A(\lambda)=\det(\lambda I-A)$ 有如下分解:

$$P_A(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_p)^{n_p},$$

其中 $n_1+n_2+\cdots+n_p=n, \lambda_i\neq\lambda_j(i\neq j)$, 则称 n_i 为 λ_i 的代数重数(简称重数), 称

$$m_i=n-\text{rank}(\lambda_i I-A)$$

为 λ_i 的几何重数. 易知 $m_i\leq n_i\,(i=1,2,\cdots,p)$. 如果 $n_i=1$, 则称 λ_i 为 A 的一个单特征值; 否则, 称 λ_i 是 A 的一个重特征值.

标准特征值问题

假定 $P_A(\lambda)=\det(\lambda I-A)$ 有如下分解:

$$P_A(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_p)^{n_p},$$

其中 $n_1+n_2+\cdots+n_p=n, \lambda_i\neq\lambda_j(i\neq j)$, 则称 n_i 为 λ_i 的代数重数(简称重数), 称

$$m_i=n-\operatorname{rank}(\lambda_i I-A)$$

为 λ_i 的几何重数. 易知 $m_i\leq n_i\,(i=1,2,\cdots,p)$. 如果 $n_i=1$, 则称 λ_i 为 A 的一个单特征值; 否则, 称 λ_i 是 A 的一个重特征值. 如果 $n_i=m_i$, 则称 λ_i 是 A 的半单特征值. 如果 A 的所有特征值都是半单的, 则称 A 是非亏损的.

标准特征值问题

假定 $P_A(\lambda)=\det(\lambda I-A)$ 有如下分解:

$$P_A(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_p)^{n_p},$$

其中 $n_1+n_2+\cdots+n_p=n, \lambda_i\neq\lambda_j (i\neq j)$, 则称 n_i 为 λ_i 的**代数重数**(简称重数), 称

$$m_i=n-\operatorname{rank}(\lambda_i I-A)$$

为 λ_i 的**几何重数**. 易知 $m_i\leq n_i\,(i=1,2,\cdots,p)$. 如果 $n_i=1$, 则称 λ_i 为 A 的一个**单特征值**; 否则, 称 λ_i 是 A 的一个**重特征值**. 如果 $n_i=m_i$, 则称 λ_i 是 A 的**半单特征值**. 如果 A 的所有特征值都是半单的, 则称 A 是**非亏损**的.

易证, A 是非亏损的充分必要条件是 A 有 n 个线性无关的特征向量 (即 A 是可对角化矩阵).

Jordan 分解定理

定理

设 $A \in \mathbb{C}^{n\times n}$ 有 r 个互不相同的特征值 $\lambda_1, \cdots, \lambda_r$, 其代数重数分别为 $n(\lambda_1), \cdots, n(\lambda_r)$, 则必存在一个非奇异矩阵 $P \in \mathbb{C}^{n\times n}$, 使得

$$J := P^{-1}AP = \begin{bmatrix} J(\lambda_1) & & & \\ & J(\lambda_2) & & \\ & & \ddots & \\ & & & J(\lambda_r) \end{bmatrix},$$

其中 $J(\lambda_i)$ 的维数等于 λ_i 的代数重数 $n(\lambda_i)$, 且具有下面的结构

$$J(\lambda_i) = \begin{bmatrix} J_1(\lambda_i) & & \\ & \ddots & \\ & & J_{k_i}(\lambda_i) \end{bmatrix}, \quad J_j(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}.$$

Schur 分解定理

定理

设 $A \in \mathbb{C}^{n\times n}$, 则存在酉矩阵 $U \in \mathbb{C}^{n\times n}$, 使得

$$U^* A U = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \cdots & * \\ & & \ddots & \vdots \\ & & & \lambda_n \end{bmatrix},$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 是 A 的特征值.

证明.

归纳法证明. □

Schur 补

设 A, B, C, D 分别为 $m \times m, m \times n, n \times m, n \times n$ 矩阵, 记

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

若 A 非奇异, 则称 $D - CA^{-1}B$ 为矩阵 M 块 A 的 Schur 补; 若 D 非奇异, 则称 $A - BD^{-1}C$ 为矩阵 M 块 D 的 Schur 补.

Fuzhen Zhang
The Schur Complement and Its Applications
Springer, 2005.

Gerschgorin 圆盘定理

设 $A = [a_{ij}] \in \mathbb{C}^{n\times n}$, 令

$$G_i(A) = \left\{ z \in \mathbb{C} : |z - a_{ii}| \le \sum_{j\neq i} |a_{ij}| \right\}, \quad i = 1, \cdots, n,$$

则有

$$\lambda(A) \subset G_1(A) \cup G_2(A) \cup \cdots \cup G_n(A).$$

R.S. Varga
Geršgorin and His Circles
Springer, 2004.

扰动矩阵特征值实例

设

$$A_\epsilon = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \epsilon & & & 0 \end{bmatrix}.$$

其特征多项式为:

$$\det(\lambda I - A_\epsilon) = \lambda^n - \epsilon = 0.$$

取 $\epsilon = 10^{-16}$, $n = 16$, 则 $\lambda = 0.1$ 为其特征值.

内容提要

幂法

幂法是计算矩阵**模最大特征值**及其对应特征向量的一种迭代方法.

Algorithm 1 幂法

设定初始向量 u_0, $k = 0$;
while 不收敛 do
    $y_k = Au_k$;
    $\mu_k = \|y_k\|_\infty$
    $u_{k+1} = y_k/\mu_k$;
    $k = k + 1$;
end while

收敛性定理

定理

设 $A \in \mathbb{C}^{n\times n}$ 有 p 个互不相同的特征值满足 $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_p|$, 并且模最大特征值 λ_1 是半单的. 如果初始向量 u_0 在 λ_1 的特征值空间上的投影不为零, 则由幂迭代算法*1*产生的向量序列 $\{u_k\}$ 收敛到 λ_1 的一个特征向量 x_1, 数值序列 $\{\mu_k\}$ 收敛到 λ_1.

证明.

略. □

- 遇到的其它问题:
  - 模最大特征值相同;
  - 求第二个模最大特征值.

内容提要

反幂法

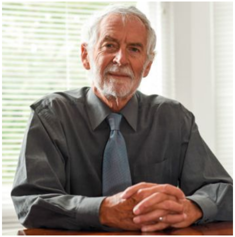
反幂法又称反迭代法，是计算矩阵**模最小特征值**及其对应特征向量的一种迭代方法.

Algorithm 2 反幂法

> 设定初始向量 u_0，$k = 0$;
> **while** 不收敛 **do**
> $y_k = A^{-1} u_k$;
> $\mu_k = \|y_k\|_\infty$
> $u_{k+1} = y_k / \mu_k$;
> $k = k + 1$;
> **end while**

内容提要

Inventor of QR algorithm



John G.F. Francis (born 1934) is an English computer scientist, who in 1961 published the QR algorithm, one of the top 10 algorithms of the 20th century. The algorithm was also proposed independently by Vera N. Kublanovskaya of the Soviet Union in the same year.

In 1955-1956 he attended Cambridge University, but did not complete a degree. He had positions with various industrial organizations and consultancies, but never returned to the field of numerical computation.

He had no idea of the impact his work on the QR algorithm had had, until re-contacted by Gene Golub and Frank Uhlig in 2007, by which time he was retired and living in Hove, England (near Brighton). Francis was awarded a University of Sussex honorary doctorate in July 2015.

相关文献

J. Francis
The QR transformation, a unitary analogue to the LR transformation-part 1
Comput. J., 4:(1961), 265–271.

V.N. Kublanovskaya
On some algorithms for the solution of the complete eigenvalue problem
USSR Comput. Math. Math. Phys.,, 1:4(1961), 555–570.

D.S. Watkins
The QR Algorithm Revisited
SIAM Review 50:1(2008), 133–145.

F. Uhlig
Finding John Francis who found QR fifty years ago
IMAGE Bull. Int. Lin. Alg. Soc., 43(2009), 19–21.

G. Golub & F. Uhlig
The QR algorithm: 50 years later its genesis by John Francis and Vera Kublanovskaya and subsequent developments
IMA J. Numer. Anal., 29(2009), 467–485.

基本迭代与收敛性

Algorithm 3 QR 算法

1: $A_0 = A$, $k = 1$;
2: **while** 不收敛 **do**
3: $\quad A_{k-1} = Q_k R_k$;
4: $\quad A_k = R_k Q_k$
5: $\quad k = k + 1$;
6: **end while**

定理

设 A 的 n 个特征值满足 $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0$, 并设 n 阶方阵 Y 的第 i 行是 A 对应于 λ_i 的左特征向量. 如果 Y 有 LU 分解, 则由 QR 算法 3 产生的矩阵 $A_m = [\alpha_{ij}^{(m)}]$ 的对角线以下的元素趋向于零, 同时对角元 $\alpha_{ii}^{(m)}$ 趋向于 $\lambda_i \, (i = 1, \ldots, n)$.

实 Schur 分解

定理 (实 Schur 分解)

设 $A \in \mathbb{R}^{n\times n}$, 则存在正交矩阵 $Q \in \mathbb{R}^{n\times n}$, 使得

$$Q^{\mathrm{T}} A Q = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ & R_{22} & \ddots & R_{2m} \\ & & \ddots & \vdots \\ & & & R_{mm} \end{bmatrix},$$

其中 R_{ii} 或者是一个实数, 或者是一个具有一对复共轭特征值的 2 阶方阵.

证明.

归纳法证明. □

上 Hessenberg 化

为减少每次 QR 迭代的运算量，通过相似变换将原矩阵 A 约化为上 Hessenberg 矩阵，然后再对约化后的矩阵进行 QR 迭代.

约化方式:

- Householder 变换;
- Givens 变换;
- 列主元 Gauss 消去法.

定理

设 $A \in \mathbb{R}^{n\times n}$ 有如下两个上 Hessenberg 分解:

$$U^{\mathrm{T}} A U = H, \;\; V^{\mathrm{T}} A V = G,$$

其中 $U = [u_1, u_2, \cdots, u_n]$ 和 $V = [v_1, v_2, \cdots, v_n]$ 是 n 阶正交矩阵, $H = [h_{ij}$ 和 $G = [g_{ij}]$ 是上 Hessenberg 矩阵. 若 $u_1 = v_1$, 而且 H 的次对角元 $H_{i+1,i}$ 均不为零 (H 不可约), 则存在对角元均为 1 或 -1 的对角阵 D, 使得

$$U = VD, \;\; H = DGD.$$

带原点位移的 QR 迭代

Algorithm 4 带原点位移的 QR 迭代

$k = 1$;
while 不收敛 **do**
    选取 μ_k;
    $H_k - \mu_k I = Q_k R_k$;
    $H_{k+1} = R_k Q_k + \mu_k I$;
    $k = k + 1$;
end while

双重步位移的 QR 迭代

Algorithm 5 带原点位移的 QR 迭代

1: **while** 不收敛 **do**
2: 选取 μ_1, μ_2;
3: $H - \mu_1 I = U_1 R_1$; $H_1 = R_1 U_1 + \mu_1 I$;
4: $H_1 - \mu_2 I = U_2 R_2$; $H_2 = R_2 U_2 + \mu_2 I$;
5: $H = H_2$;
6: **end while**

定理

若 H 是不可约的上 Hessenberg 矩阵，且 μ_1 和 μ_2 均非 H 的特征值，则 H_2 也是不可约的上 Hessenberg 矩阵.