

实规范矩阵的特征值问题*

曹志浩

(复旦大学)

实对称矩阵的特征值问题,无论是低阶稠密矩阵的全部特征值问题,或高阶稀疏矩阵的部分特征值问题,都已有许多有效的计算方法,迄今最重要的一些成果已总结在[5]中.本文利用规范矩阵的一些重要性质将对于 Hermite 矩阵(特别是对称矩阵)特征值问题的一些有效算法推广到规范矩阵的特征值问题,由于对复规范阵的推广是简单的,而且实际上常遇到的是实矩阵(这时常要求只用实运算),因此我们着重讨论实规范矩阵的特征值问题.

一、实规范矩阵的拟 Jacobi 算法

设 A 是 $n \times n$ 实规范阵,则 A 可通过正交矩阵 P , 相似变换为块对角的

$$P^T A P = \text{diag} \left\{ \begin{bmatrix} \mu_1 & \nu_1 \\ -\nu_1 & \mu_1 \end{bmatrix}, \dots, \begin{bmatrix} \mu_q & \nu_q \\ -\nu_q & \mu_q \end{bmatrix}, \mu_{2q+1}, \dots, \mu_n \right\}, \quad (1)$$

其中 μ_i, ν_i 是实数, 且 $\nu_i \neq 0 (i = 1, \dots, q)$.^[1]

显然, 对应于二阶块 $\begin{bmatrix} \mu_j & \nu_j \\ -\nu_j & \mu_j \end{bmatrix}$ 的是 A 的共轭复特征值 $\mu_j \pm i\nu_j$, 而相应的规范化特征向量是

$$\frac{1}{\sqrt{2}} (p_{2j-1} \pm i p_{2j}),$$

其中 p_l 表示正交变换阵 P 的第 l 列. 因此, 为求解实规范特征值问题, 只要能计算(1)中的正交阵 P 和右端的块对角阵即可.

令

$$A_+ = \frac{A + A^T}{2}, \quad A_- = \frac{A - A^T}{2}$$

分别是实规范阵 A 的对称部分和反对称部分.

定理1 设 A 为实规范阵, 则 A_+ 和 A_- 有公共的完全正交特征向量组, A_+ 的 n 个实特征值为 A 的特征值的实部, A_- 的 n 个纯虚的或零特征值为 A 的特征值的虚部.

证 因 A 规范, $A^T A = A A^T$, 因此 A_+ 和 A_- 可交换 $A_+ A_- = A_- A_+$, A_+ 和 A_- 都具有(正交的)完全特征向量组, 由熟知定理^[1], 存在对 A_+ 和 A_- 公共的完全正交特征向量组, 若以这些特征向量为列组成的正交阵为 P , 则

* 1982年4月30日收到.

$$P^T A_+ P = \Lambda_+, \quad P^T A_- P = i \Lambda_-,$$

其中 Λ_+ 和 Λ_- 都是实对角阵, 故

$$P^T A P = \Lambda_+ + i \Lambda_-. \quad \text{证毕}$$

对于低阶稠密的对称矩阵 A 的特征值问题, 特别当 A 近似对角时, **Jacobi** 方法是很有效的, 此时 A 可迭代地约化为对角阵

$$R^T A R = \text{diag}\{\mu_1, \dots, \mu_n\},$$

其中 R 为一系列 **Jacobi** 旋转矩阵的乘积。

对于低阶稠密的反对称矩阵 A' , [4] 提出了 **Jacobi** 法的推广, 即可迭代地将 A' 约化为块对角的

$$R'^T A' R' = \text{diag}\left\{\begin{bmatrix} 0 & v_1 \\ -v_1 & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & v_q \\ -v_q & 0 \end{bmatrix}, 0, \dots, 0\right\},$$

其中 R' 为一系列 **Jacobi** 零化子的乘积, 一个 **Jacobi** 零化子的作用类似于一个 **Jacobi** 旋转, 但经一个零化子作用后, 使在矩阵对称位置上的一对二阶块化为零, 一个 **Jacobi** 零化子一般为四个平面旋转阵的乘积。

结合这两个算法, 应用定理 1 就可导出下列

实规范矩阵的拟 **Jacobi** 算法:

1) 对 A_+ 实施 **Jacobi** 算法, 并应用行列同时交换, 使所得对角阵 (其元素是 A_+ 的特征值) 中, 元素相等的排列在一起, 若共有 m 个相异特征值 $\mu_1, \dots, \mu_m$, 则结果可表为

$$R^T A_+ R = \text{diag}\{\mu_1 I_{n_1}, \dots, \mu_m I_{n_m}\}, \quad (2)$$

其中 R 是一系列 **Jacobi** 旋转阵和排列阵的乘积。

2) 将 1) 中的 **Jacobi** 算法同时作用在 A_- 上, 由定理 1 可知 R 的第 $(n_{i-1} + 1)$ 到第 n_i 列 ($i = 1, \dots, m, n_0 = 0$) 也是 A_- 的不变子空间的正交基, 故

$$R^T A_- R = \text{diag}\{S_1, \dots, S_m\},$$

其中 S_i 是 n_i 阶反对称矩阵 ($i = 1, \dots, m$)。

3) 对每个 S_i 实施反对称阵的拟 **Jacobi** 算法, 结果可表为

$$R_i'^T S_i R_i' = \text{diag}\left\{\begin{bmatrix} 0 & v_{l_{i-1}+1} \\ -v_{l_{i-1}+1} & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & v_{l_i} \\ -v_{l_i} & 0 \end{bmatrix}, 0, \dots, 0\right\}, \quad (3)$$

其中 R_i' 为一系列 **Jacobi** 零化子的乘积, $i = 1, \dots, m$, $\sum_{i=1}^m l_i = q$, $l_0 = 0$, 显然,

$$\text{diag}\{R_1'^T, \dots, R_m'^T\} (R^T A_+ R) \text{diag}\{R_1', \dots, R_m'\} = R^T A_+ R.$$

4) 将 (2) 和 (3) 中对应角块相加, 且计算

$$P = R \text{diag}\{R_1', \dots, R_m'\}$$

过程结束就实现了 (1)。

二、实规范矩阵的子空间迭代法

对高阶稀疏矩阵, 常常需要求部分特征值和相应的特征向量。首先建立下述

定理 2 若 A 是实规范阵, 其特征值排列次序只要满足条件: 模相等的排在一起。对称非负定矩阵 $A^T A$ 的特征值排列次序与 A 的一致, 设 Q_1 是 A 的对应于一组或几组模相等特征值的不变子空间的正交基为列组成的矩阵, 则 Q_1 也是 $A^T A$ 的相应不变子空间的正交基矩阵, 且 $Q_1^T A Q_1$ 仍是实规范矩阵。

证 因 A 是规范的, $A^T A = A A^T$, 故 A 和 $A^T A$ 是可交换的, 又因为 A 和 $A^T A$ 都有(正交)完全特征向量组, 由熟知定理^[1], 存在 A 和 $A^T A$ 的公共正交完全特征向量组, 因而有公共的不变子空间的正交基矩阵 Q_1 。

令 Q_2 为 A 的与 Q_1 互补的不变子空间的正交基矩阵, 因 A 是规范阵, Q_2 是存在的, 显然 $Q = [Q_1, Q_2]$ 是正交阵, 且 $Q_1^T A Q_1 = 0$, $Q_2^T A Q_2 = 0$, 故

$$Q^T A Q = \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix} A [Q_1 \ Q_2] = \begin{bmatrix} Q_1^T A Q_1 & \\ & Q_2^T A Q_2 \end{bmatrix}, \quad (4)$$

显然,

$$(Q^T A Q)^T (Q^T A Q) = Q^T A^T A Q = Q^T A A^T Q = (Q^T A Q) (Q^T A Q)^T,$$

结合(4)就得到

$$(Q_1^T A Q_1)^T (Q_1^T A Q_1) = (Q_1^T A Q_1) (Q_1^T A Q_1)^T,$$

即 $Q_1^T A Q_1$ 是规范矩阵。证毕。

由定理 2 可知, 求实规范阵 A 的模最大的一些特征值和对应的特征向量可以约化为: 首先, 求对称非负矩阵 $A^T A$ 的相应控制特征值和对应的特征向量, 设这些正交规范特征向量为列组成矩阵 X ; 然后, 对低阶规范矩阵 $X^T A X$ 求解全部特征值问题。由此导出下述

规范阵子空间迭代法 I: (设要求的特征值个数为 r)

(1) 随机地取 p ($r \leq p < n$) 个初始向量, 然后正交规范化, 记为 $n \times p$ 矩阵 $X^{(0)}$, $X^{(0)T} X^{(0)} = I_p$ 。

(2) 若已有第 $(v-1)$ 次矩阵 $X^{(v-1)}$, 求 $X^{(v)}$ 的迭代过程如下:

i) 计算 $Z^{(v)} = A X^{(v-1)}$,

ii) 计算 $p \times p$ 正定阵 $S^{(v)} = Z^{(v)T} Z^{(v)}$ (这里假定 $Z^{(v)}$ 是列满秩的)

iii) 用 Jacobi 算法求 $S^{(v)}$ 的全部特征值和特征向量

$$Q^{(v)T} S^{(v)} Q^{(v)} = (D^{(v)})^2,$$

其中 $Q^{(v)}$ 是 $p \times p$ 正交阵, $D^{(v)}$ 是对角元为正的 diagonal 阵, 它的对角元下降排列,

iv) 计算 $X^{(v)} = Z^{(v)} Q^{(v)} (D^{(v)})^{-1}$,

v) 检验收敛性。

(3) 将收敛的 $X^{(v)}$ 记为 X , 且将 X 分块为

$$X = [X_1, \dots, X_m],$$

其中 X_l , $l = 1, \dots, m$, 对应于 $A^T A$ 的相等特征值, 对每个 l 计算 $X_l^T A X_l$, 并对这低阶规范阵实施拟 Jacobi 算法, 表为

$$P^T(X_1^T A X_1)P = \text{diag}\left\{\begin{bmatrix} \mu_1^{(1)} & \nu_1^{(1)} \\ -\nu_1^{(1)} & \mu_1^{(1)} \end{bmatrix}, \dots, \begin{bmatrix} \mu_{n_i}^{(1)} & \nu_{n_i}^{(1)} \\ -\nu_{n_i}^{(1)} & \mu_{n_i}^{(1)} \end{bmatrix}, \mu_{2n_i+1}^{(1)}, \dots, \mu_{m_i}^{(1)}\right\},$$

算法过程结束即可得特征值 $\mu_j^{(1)} \pm i\nu_j^{(1)}$, $\mu_i^{(1)}$ 等, 由 $X_1 P$ 的相应列即得相应特征向量.

注. (1) 由上述算法可知

$$X^{(v)} T X^{(v)} = (D^{(v)})^{-T} Q^{(v)T} Z^{(v)T} Z^{(v)} Q^{(v)} (D^{(v)})^{-1} = I_p,$$

$$S^{(v)} = X^{(v-1)T} (A^T A) X^{(v-1)},$$

因此算法过程(1), (2)实际上是对非负对称矩阵 $A^T A$ 的子空间迭代法, [6]中对 A 是对称正定矩阵的情况提出了这算法, 我们把它推广应用于 A 为实规范矩阵的情况. 注意在具体迭代过程中不必计算 $A^T A$, 甚至 A^T 也不出现.

(2) 对算法的收敛性分析直接可应用[9]中的方法, 从而得到同 A 为对称时相同的收敛结果.

对某些特殊的规范矩阵(只要能满足条件: 对任意列正交的 $n \times p$ 矩阵 Q , $Q^T A Q$ 同 A 一样是规范的, 例如 A 为反对称就满足这条件)可直接对 A 作子空间迭代, 从而导出下述实规范阵子空间迭代法 II

(1) 随机地取 p 个初始向量, 然后正交规范化, 记为 $n \times p$ 矩阵 $X^{(0)}$, $X^{(0)T} X^{(0)} = I_p$

(2) 若已有第 $(v-1)$ 次矩阵 $X^{(v-1)}$, 求 $X^{(v)}$ 的迭代过程如下:

i) 计算 $Z^{(v)} = A X^{(v-1)}$,

ii) 将 $Z^{(v)}$ 的列正交规范化, 得 $Q^{(v)}$:

$$Z^{(v)} = Q^{(v)} R^{(v)},$$

其中 $Q^{(v)}$ 具正交规范列, $R^{(v)}$ 为 $p \times p$ 上三角阵.

iii) 计算由 $Q^{(v)}$ 确定的 A 的投影

$$B^{(v)} = Q^{(v)T} A Q^{(v)}.$$

iv) 用实规范矩阵的拟 Jacobi 算法求解 $B^{(v)}$ 的全部特征值问题

$$P^{(v)T} B^{(v)} P^{(v)} = \text{diag}\left\{\dots, \begin{bmatrix} \mu_j^{(v)} & \nu_j^{(v)} \\ -\nu_j^{(v)} & \mu_j^{(v)} \end{bmatrix}, \dots, \mu_i^{(v)}, \dots\right\} \quad (5)$$

且一阶或二阶对角块按模的大小下降排列.

v) 计算 $X^{(v)} = Q^{(v)} P^{(v)}$.

vi) 检验收敛性.

算法过程结束, 即可从(5)的对角阵中得到要求的特征值, 和从 $X^{(v)}$ 的相应列得到特征向量.

算法 II 的收敛性可类似于[10]中定理 2.5 (对证明作某些明显的修改) 导出. 我们甚至也能导出与 A 对称时相同的收敛结果:

定理 3 设实规范阵的特征值按模下降排列: $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$, 若 $|\lambda_p| > |\lambda_{p+1}|$, 则子空间迭代法 II 的对应于 (实或共轭复的) 重特征值的特征不变子空间的收敛速度为 $O(|\lambda_{p+1}|^n / |\lambda_1|^n)$, 其中 $|\lambda_i|$ 表示相应特征不变子空间中特征值的 (相等) 模.

定理 3 的证明类似于[7]中定理 3.2, 只要注意到对实规范阵的拟 Schur 典则形式可取为(1), 当 A 规范时[7]中定理 2.1 可加强为

定理 4 令 $S_1 \in R^{m \times m}$, $K \in R^{m \times n}$, $S_2 \in R^{n \times n}$, 若 S_1 和 S_2 是规范的, 且

$$\mu \equiv \min_i |\lambda_i(S_2)| > \max_i |\lambda_i(S_1)| \equiv \lambda$$

则

$$\|S_1^* K S_2^{-*}\| \leq \|K\| \lambda / \mu^{1/2}.$$

注 (1) [3], [8]中对 A 为任意实矩阵提出了求右特征向量的单侧同时迭代法, 并宣称有同对称情况相同的收敛速度, 但并未给出证明, 由于他们的算法与子空间迭代法 II 有很多类似之处, 定理 3 说明当 A 是实规范矩阵时他们的结论是可以证明的。但对一般矩阵这是一个尚未解决的问题。

三、实规范矩阵的 Lanczos 算法

对称 Lanczos 算法是求解大型稀疏对称矩阵的部分特征值问题的最有效方法之一^[6,10], 对于实规范矩阵 A 的 Lanczos 方法, 可类似于子空间迭代法 I 那样, 先对 $A^T A$ 或 $(A^T A - \sigma I)^{-1}$ 实施对称 Lanczos 算法, 求得部分极端特征值, 或最接近 σ 的特征值后, 利用反迭代或共轭斜量法求得相应特征向量, 再利用 A 的由这些特征不变子空间的正交基确定的投影矩阵求得规范阵 A 的特征值和相应特征向量。对于 $A^T A$ 的 Lanczos 算法可采用如下格式产生 Lanczos 向量 $\{v_1, v_2, \dots, v_i\}$ 和对称三对角阵 T_i :

$$T_i = \begin{pmatrix} \alpha_1 & \beta_2 & & & \\ \beta_2 & \alpha_2 & \beta_3 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_i & \alpha_i & \beta_{i+1} \\ & & & \beta_{i+1} & \alpha_{i+1} \end{pmatrix}$$

任取初始向量, $v_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, 然后进行计算如下:

$$\left. \begin{aligned} w_1 &= A v_1 \\ u_1 &= A^T w_1 \\ \alpha_i &= v_i^T u_i \\ r_i &= u_i - \alpha_i v_i \\ \beta_{i+1} &= \|r_i\| \\ v_{i+1} &= \beta_{i+1}^{-1} r_i \\ w_{i+1} &= A v_{i+1} \\ n_{i+1} &= A^T w_{i+1} - \beta_{i+1} v_i \end{aligned} \right\} \quad i = 1, 2, \dots$$

参 考 文 献

- [1] Гантмахер, Ф. Р. Теория Матриц (中译本), 1957.
- [2] Householder, A. S., The Theory of Matrices in Numerical Analysis, 1964.
- [3] Jennings, A. & W. J. Stewart, Simultaneous Iteration for Partial Eigensolution of Real Matrix, *J. Inst. Math. Appl.* 15(1975).
- [4] Paardekooper, M. H. C., An Eigenvalue Algorithm for Skew-Symmetric Matrices, *Numer. Math.* 17:3(1971).
- [5] Parlett, B. N., The Symmetric Eigenvalue Problem, Prentice Hall Inc., New York, 1980.
- [6] Rutishauser, H., Computation Aspects of F. L. Bauer's Simultaneous Iteration Method, *Numer. Math.* 13:1, 1969.
- [7] Stewart, S. W., Simultaneous Iteration for Computing Invariant Subspaces of Non-Hermitian Matrices, *Numer. Math.* 25:2, 1978.
- [8] Stewart, W. J. & A. Jennings, A Simultaneous Iteration Algorithm for Real Matrices, *acm Trans. Math. Soft.*, 7:2, 1981.
- [9] 曹志浩, 分块子空间迭代法及其在有限元计算中的应用, 高校计算数学学报, 1:2, 1979.
- [10] 曹志浩, 矩阵特征值问题, 上海科技出版社, 1980.

On the Eigenvalue Problem for Real Normal Matrices

Cao Zhi-Hao

(Fudan University)

Abstract

Using the important properties of normal matrices we generalize some algorithms which are efficient for solving symmetric eigenvalue problem.

For small full normal matrices we generalize Jacobi algorithm to get Jacobi-like algorithm. For large sparse normal matrices A 's we present two types of subspace iterative algorithm, one is derived from using symmetric subspace iterative algorithm to $A^T A$. We also show that the convergence behavior of both types are same as that of symmetric subspace iterative algorithm. In addition, the generalization of symmetric Lanczos algorithm is briefly discussed.

All algorithms presented here for real matrices can be used by real arithmetic.