第 22 卷 第 1 期 桂 林 工 学 院 学 报 Vol. 22 No. 1
2002 年 1 月 JOURNAL OF GUILIN INSTITUTE OF TECHNOLOGY Jan. 2002

文章编号：1006－544X(2002)01－0085－04

利用非线性方程组求解矩阵特征值特征向量

郑敏玲，武 坤

（中南大学 应用数学与应用软件系，湖南 长沙 410083）

摘 要：矩阵特征值问题已成为数值计算中的一个重要组成部分，为有效求解此类问题，提出了一种求解特征值的新方法：利用非线性方程组的 Newton 迭代法求解特征向量，为提高迭代的收敛速度，引入同伦思想，利用插值方法，得到近似特征向量 $\boldsymbol{Y}^{(N)}$，以 $\boldsymbol{Y}^{(N)}$ 作为迭代初值，从而快速求出问题的具有较高精度的解．该算法稳定性好，可并行运算．

关键词：特征值；特征向量；同伦；牛顿迭代；插值法；盖尔圆

中图分类号：O177.91 **文献标识码**：A①

计算特征值、特征向量，不管在理论上还是在实际应用中，都是一个非常重要的问题．经典的 Jacobi 算法是解对称矩阵问题的一个稳定的算法，且精度高，特征向量的正交性有保证，但是它有两个主要缺点：一是收敛慢，二是串行运算．目前已有快速算法和并行算法．无论是经典的或是新的快速并行 Jacobi 算法都是用正交变换对原 n 阶矩阵中的二阶主子矩阵运算，目的是消去它的两个非对角元素．二分法[1]前最常用的解对称矩阵特征值问题的并行算法，其思想是计算 A 的特征矩阵 $A-\lambda I$ 的各级前主子式 $P_i(\lambda)$，$i=1,2,\cdots,n$．利用序列 $\{P_i(\lambda)\}$ 的 Sturm 序列性质，确定特征值的分布情况，通过逐步计算可得到只含有一个特征值的区间 $[c,d]$，将区间逐次二等分确定满足精度要求的特征值．分治法[2]（简称 DC 法）也是一个比较新的适合于并行运算的矩阵特征值问题算法，其基本思想是把矩阵分成两个子矩阵，先解子矩阵特征值问题，然后构成一个比原问题容易解的特征值问题．本文是在经典 Jacobi 算法的基础上，通过构造特征向量的非线性方程组而提出的一种新方法．利用同伦的思想求 Newton 迭代初值，使迭代算法快速收敛．并对该算法进行了稳定性分析．

1 推导过程及算法

定理 1[3] 设 $\boldsymbol{A}=(a_{ij})\in R^{n\times n}$，$\boldsymbol{A}_k=\boldsymbol{D}+t_k\boldsymbol{P}$，$0\leqslant t_k\leqslant 1$，$t_k=hk$，$k=0,1,2,\cdots,N$，$N=1/h$，其中

$$\boldsymbol{D}=\begin{bmatrix} a_{11} & \cdots & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix},\quad t_k\boldsymbol{P}=\begin{bmatrix} 0 & t_ka_{12} & \cdots & t_ka_{1n} \\ t_ka_{21} & 0 & \cdots & t_ka_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ t_ka_{n1} & t_ka_{n2} & \cdots & 0 \end{bmatrix},$$

则 当 $h=2^{-s}$ 时，A_1 的所有盖尔圆不相交．

其中 $s=[\dfrac{\ln(r/d)}{\ln 2}+2]$，$r=\max\limits_i\{\sum\limits_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}|\}$，$d=\min\limits_{\substack{1\leqslant i,j\leqslant n\\ i\neq j}}\{|a_{ii}-a_{jj}|\}$．

① 收稿日期：2001－03－12；修订日期：2001－05－14
基金项目：湖南邮电系统基金资助项目
作者简介：郑敏玲（1969－），男，湖南隆回人，硕士研究生，研究方向：计算数学与应用软件．

推导过程及算法：

设 $\boldsymbol{A}\in R^{n\times n}$，对应于 λ 的特征向量 $\boldsymbol{X}_i=\begin{bmatrix}\vdots\\1\\\vdots\end{bmatrix}$（第 i 个分量为 1），则有：

$$\boldsymbol{A}\boldsymbol{X}_i=\lambda_i\boldsymbol{X}_i, \tag{1}$$

它的第 i 个方程

$$a_{ii}+\sum_{p=1}^{i-1}a_{ip}x_p+\sum_{p=i+1}^{n}a_{ip}x_p=\lambda_i, \tag{2}$$

当 $j\neq i$ 时，第 j 个方程

$$a_{ji}+\sum_{p=1}^{i-1}a_{jp}x_p+\sum_{p=i+1}^{n}a_{jp}x_p=\lambda_i x_j, \tag{3}$$

由式（2），（3）可以得到：

$$a_{ji}+\sum_{p=1}^{i-1}a_{jp}x_p+\sum_{p=i+1}^{n}a_{jp}x_p=\left[a_{ii}+\sum_{p=1}^{i-1}a_{ip}x_p+\sum_{p=i+1}^{n}a_{ip}x_p\right]x_j,\ j=1,2,\cdots,i-1,i+1,\cdots,n \tag{4}$$

所以，

$$f_j=a_{ij}x_j^2+\left[a_{ii}-a_{jj}+\sum_{p=1}^{j-1}a_{jp}x_p+\sum_{p=j+1}^{i-1}a_{ip}x_p+\sum_{p=i+1}^{n}a_{ip}x_p\right]x_j-\left[a_{ji}+\sum_{p=1}^{j-1}a_{jp}x_p+\sum_{p=j+1}^{i-1}a_{jp}x_p+\sum_{p=i+1}^{n}a_{jp}x_p\right]=0,\quad j=1,2,\cdots,i-1 \tag{5}$$

$$f_j=a_{ij}x_j^2+\left[a_{ii}-a_{jj}+\sum_{p=1}^{i-1}a_{ip}x_p+\sum_{p=i+1}^{j-1}a_{ip}x_p+\sum_{p=j+1}^{n}a_{ip}x_p\right]x_j-\left[a_{ji}+\sum_{p=1}^{i-1}a_{jp}x_p+\sum_{p=i+1}^{j-1}a_{jp}x_p+\sum_{p=j+1}^{n}a_{jp}x_p\right]=0,$$
$$j=i+1,\cdots,n \tag{6}$$

设方程（5），（6）中 X_j 的系数为 b_j，则它的 Jacobi 矩阵

$$\boldsymbol{J}=(q_{lm})_{(n-1)\times(n-1)},\quad q_{lm}=\begin{cases}2a_{il}x_l+b_l, & (l=m)\\ a_{im}x_l-a_{lm}, & (l\neq m)\end{cases}$$

由牛顿迭代法可得到

$$X^{(k)}=X^{(k-1)}+\boldsymbol{J}(X^{(k-1)})^{-1}\boldsymbol{F}(X^{(k-1)}),$$
$$\boldsymbol{F}=-(f_1,\cdots,f_{i-1},f_{i+1},\cdots,f_n)^{\mathrm{T}} \tag{7}$$

当 $\boldsymbol{J}(X^{(k-1)})^{-1}$ 存在和选取合适的初值时，Newton 迭代至少二阶收敛.

下面来考虑初始值 $X^{(0)}$ 的选取. 由前面定理知，$\boldsymbol{A}_0=\boldsymbol{D}$，$\boldsymbol{A}_N=\boldsymbol{A}$，对给定的 h，可以得到 $N+1$ 个矩阵 $\boldsymbol{A}_0$，$\boldsymbol{A}_1$，…，$\boldsymbol{A}_N$. 而 $\boldsymbol{A}_0$ 的特征向量为

$$\boldsymbol{Y}^{(0)}=\left(\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix},\begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix},\cdots,\begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}\right),$$

以 $\boldsymbol{Y}^{(0)}$ 作为迭代初值，应用牛顿法于矩阵 $\boldsymbol{A}_1$ 求出 $Y^{(1)}$. 再由两点插值法，求得 $Y^{(2)}=2Y^{(1)}-Y^{(0)}$. 对于 $k\geqslant 3$，则由三点插值法，求得 $Y^{(k)}=3Y^{(k-1)}-3Y^{(k-2)}+3Y^{(k-3)}$（$k\geqslant 3$），从而最后求 $Y^{(N)}$. 以 $Y^{(N)}$ 作为牛顿迭代的初值用于（7）式便可求解.

根据以上的讨论可以得到如下算法：

①求初值算法 1：（Newton 迭代求 $\boldsymbol{A}_1$ 的特征向量 $\boldsymbol{Y}^{(1)}=(X_1,X_2,\cdots,X_n)$

步 1 给出计算精度 ε_1，迭代初值为 $\boldsymbol{Y}^{(0)}$；

步 2 计算 $s=\dfrac{\ln(r/d)}{\ln 2}+2]$，$\boldsymbol{A}_1=(a_{ij})=\boldsymbol{D}+t_1\boldsymbol{P}$，$r=\max\limits_i\left\{\sum\limits_{\substack{j=1\\i\neq j}}^{n}|a_{ij}|\right\}$，$d=\min\limits_{i\neq j}|a_{ii}-a_{jj}|$.

步 3 假设已进行了 k 次迭代，已求出 X^k，计算 $F_{A_1}(X^{(k)})$，$J_{A_1}(X^{(k)})$，设 $b_k=F_{A_1}(X^{(k)})$

步 4 解线性方程组 $J_{A_1}(X^{(k)})\Delta X^k=-b_k$，得到 ΔX^k，$X^{(k+1)}=X^{(k)}+\Delta X^k$；

步 5 若 $\| F(X^{(k+1)}) \| \leqslant \varepsilon_2$,则置 $X^* = X^{(k+1)}$,打印 X^*,否则 $k+1 \to k$,$X^{(k+1)} \to X^{(k)}$,$F(X^{(k+1)}) \to F(X^{(k)})$,转 2;

步 6:结束.

②算法 2:(利用插值法求 $Y^{(N)}$)

步 1:计算 $N = 1/h$,$h = 2^{-s}$;

步 2:由 $Y^{(0)}$,$Y^{(1)}$,根据两点插值法求 $Y^{(2)} = 2Y^{(1)} - Y^{(0)}$;

步 3:当 $3 \leqslant k \leqslant N$ 时,$Y^{(k)} = 3Y^{(k-1)} - 3Y^{(k-2)} + 3Y^{(k-3)}$.

③算法 3:(迭代求 A 的特征值特征向量)

步 1:给出计算精度 ε_2,迭代初值为 $Y^{(N)}$;

步 2:假设已进行了 k 次迭代,已求出 $X^{(k)}$,计算$F(X^{(k)})$,$J(X^{(k)})$,记 $b_k = F(X^{(k)})$;

步 3:解线性方程组 $J(X^{(k)})\Delta X^k = -b_k$,得到 ΔX^k;

步 4:求 $X^{(k+1)} = X^{(k)} + \Delta X^k$ 及 $F(X^{(k+1)})$;

步 5 若 $\| F(X^{(k)}) \| \leqslant \varepsilon_2$,则置 $X^* = X^{(k+1)}$,打印 X^*,$\| F(X^*) \|$ 及 $\| \Delta X^k \|$,否则 $k+1 \to k$,$X^{(k+1)} \to X^{(k)}$,$F(X^{(k+1)}) \to F(X^{(k)})$,转 2;A

步 6:结束.

2　算法稳定性分析

定理 2　$Y^{(n)} = nY^{(1)} - (n-1)Y^{(0)}$　$(k \geqslant 2)$

证明:(数学归纳法)　当 $n = 2$ 时,结论显然成立.

假设当 $2 < n < k-1$ 结论成立,则由算法

$$\begin{aligned} Y^{(k)} &= 3Y^{(k-1)} - 3Y^{(k-2)} + Y^{(k-3)} \\ &= 3[(k-1)Y^{(1)} - (k-2)Y^{(0)}] - 3[(k-2)Y^{(1)} - (k-3)Y^{(0)}] + (k-3)Y^{(1)} - (k-4)Y^{(0)} \\ &= kY^{(1)} - (k-1)Y^{(0)} \end{aligned}$$

由 Ostrowski 定理,可以得到:

定理 3　设 $\lambda(A_1) = \{\lambda_i\}$,$\lambda(A_2) = \{\mu_i\}$,$m_\alpha = \max\limits_{\substack{i,j \\ i \neq j}} \{a_{ii} + t_2 |a_{ij}|\}$,对任一 $\mu \in \lambda(A_2)$,必有 $\lambda_i(\mu) \in \lambda(A_1)$,使

$$|\lambda_i(\mu) - \mu| < (n+2) m_\alpha^{1-1/n} \| A_2 - A_1 \|_\alpha^{1/n},\ \text{其中} \| A_2 - A_1 \|_\alpha = \frac{h}{n} \sum_{\substack{i,j=1 \\ j \neq i}}^{n} |a_{ij}|.$$

由定理 2 知道,算法的稳定性依赖于 A_1 的特征值问题的稳定性. 根据定理 1,不妨设 A_1 的特征值为 λ_1,λ_2,…,λ_n.

设 X_k,Y_k 分别是 A_1 的属于 λ_k 的特征向量,即

$$A_1 X_k = \lambda_k X_k,\quad Y_k^H A = \lambda Y_k^H$$

令 $S(\lambda_k) = Y_k^H X_k / \| X_k \|_2 \| Y_k \|_2$,容易证明[4],$|S(\lambda_k)|$ 是唯一确定的,且 $S(\lambda_k) \neq 0$.

设 $E \in C^{n \times n}$($\| E \|_2 = 1$),则利用函数论的有关知识可证,存在 $\varepsilon_0 > 0$ 和在 $|\varepsilon| < \varepsilon_0$ 上的解析函数 $\lambda_i(\varepsilon)$和向量函数 $X_i(\varepsilon)$,$i = 1,2,\cdots,n$,使得

$$(A_1 + \varepsilon E) X_i(\varepsilon) = \lambda_i(\varepsilon) X_i(\varepsilon) \tag{8}$$

并满足

$$\lambda_i(0) = \lambda_i, X_i(0) = X_i$$

其中 X_i 是对应于λ_i 的特征向量. 可以证明,

$$|\lambda_i(\varepsilon) - \lambda_i| \leqslant \frac{\| \varepsilon E \|_2}{|S(\lambda_i)|} + 0(|\varepsilon|^2),$$

假定 $X_i(\varepsilon)$ 的泰勒展开式为

$$X_i(\varepsilon) = X_i + \varepsilon Z_i + 0(\varepsilon^2) \tag{9}$$

由于 $X_1,\cdots,X_n$ 构成 C_n 的一组基,故有

$$Z_i = \sum_{j=1}^{n} \alpha_{ij} X_j \quad , \quad \alpha_{ij} \in C, \tag{10}$$

把(10)代入(9)得

$$X_i(\varepsilon) = (1 + \varepsilon\alpha_{ii})\boldsymbol{X}_i + \varepsilon \sum_{\substack{j=1 \\ j\neq i}}^{n} \alpha_{ij} X_j + 0(\varepsilon^2) \tag{11}$$

设 $X_i(\varepsilon)$ 已选好使得 $\alpha_{ii} = 0$,因为

$$\lambda_i(\varepsilon) = \lambda_i + \lambda'_i(\varepsilon) + 0(\varepsilon^2) \tag{12}$$

(11)、(12)代入(8),得

$$\sum_{j\neq i} \alpha_{ij}(\lambda_j - \lambda_i) X_j = (\lambda_i(0) I - E)\boldsymbol{X}_i, \tag{13}$$

在(13)两边左乘 Y_j^H,注意到 $Y_j^H \boldsymbol{X}_i = 0$, $i \neq j$,有

$$\alpha_{ij}(\lambda_j - \lambda_i) Y_j^H \boldsymbol{X}_i = - Y_j^H E \boldsymbol{X}_i, \quad i \neq j,$$

不妨设 $S(\lambda_j) = | Y_j^H \boldsymbol{X}_i | = Y_j^H \boldsymbol{X}_i$,于是

$$X_i(\varepsilon) = \boldsymbol{X}_i + \varepsilon \sum_{j\neq i} \frac{Y_j^H E X_i}{(\lambda_i - \lambda_j) S(\lambda_j)} \cdot X_j + 0(\varepsilon^2)$$

3 结束语

利用非线性方程组求矩阵特征值特征向量的方法需要进行$(2n-2)$次 Newton 法来解非线性方程组. Newton 法虽然有收敛快和自校正等优点,但应用到实际计算中仍有不少的问题,例如,迭代初值 $X^{(0)}$要求与解 X^* 很靠近迭代才收敛,还有每步要计算 Jacobi 矩阵$F'(X^k)$及解线性方程组工作量较大. 然而本文提出的方法可以在计算机上并行运行,而且由于在处理初值时利用了同伦思想,根据稳定性分析的结果,该算法是可以稳定执行的. 必须指出,牛顿迭代在虚特征值的情况下不收敛,所以这种方法对于非对称矩阵,只能求它的实特征值.

参考文献:

[1] Barlow B, Evans D. A parallel organization of the bisection algorithm [J]. The Computer Journal, 1977, (22): 267~269.
[2] Ming Gu, Stanley C Eisenstat. A divide-and-conquer algorithm for the symmetric tridiagonal eigenproblem [J]. SIAM J. MatriX Anal. Appl. 1995, (16): 172~191.
[3] Elgindi M B. The quadratic method for computing the eigenpairs of a matrix [J]. Intern. J. Computer Math, 2000, (73): 517~523.
[4] 孙继广. 矩阵扰动分析 [M]. 北京:科学出版社, 1987. 175~176.

A new algorithm for computing the eigenpairs of a matrix by Newton iteration

ZHENG Min-ling, WU Kun
(*Department of Mathematics and Software Application, Central South University, Changsha* 410083, *China*)

Abstract: Matrix 's eigenvalue problem, concerned with morden computational physic computational chemics, computational biology, many scientific research and engineer computation, hasbeen a main component of Numerial computation. This paper works on solving the problem effectively. It presents a new method of computing eigenpairs by Newton iterationfor nonlinear equations. In order to improve iteration convergence rate, it introduces a homotopy idea, and employs interpolation to attain approximate eigenvector $\boldsymbol{Y}^{(N)}$. As a result, we can get more accurate solution rapidly when employing $\boldsymbol{Y}^{(N)}$ as an iteration starting value. The paper includes some discussion about algorithm stability. The algorithm has the advantage of parallel working.

Key words: eigenvalue; eigenvetor; homotopy ; Newton iteration; polynomical interpolation; Gerschgorin circle