求解非线性矩阵特征值问题的一个三阶收敛的算法*

陈广义　　薛彦才
(中国科学院沈阳计算所)

A CUBICALLY CONVERGENT ALGORITHM FOR SOLVING NONLINEAR EIGENVALUE PROBLEMS

Chen Guang-yi　　Xue Yan-cai
(Shenyang Institute of Computing Technology, Academia Sinica)

Abstract

In this paper we give a cubically convergent algorithm for solving the nonlinear eigenvalue problem $A(\lambda)X=0$, $X\neq 0$, where $A(\lambda)$ is a functional λ-matrix. The amount of computation is much less than that of Lancaster's cubically convergent algorithm. Several numerical examples are also presented.

§1. 引　言

考虑矩阵 $A(\lambda)\in C^{n\times n}$，它的元素 $a_{ij}(\lambda)$ 是复变量 λ 的解析函数，我们称这样的矩阵为泛函 λ-矩阵[3]．如果 $a_{ij}(\lambda)$是 λ 的多项式，则称它为 λ-矩阵[1]．

我们把满足

$$\begin{cases}A(\lambda)x=0\\ y^HA(\lambda)=0\end{cases}\tag{1.1}$$

的值 λ 和向量 y、x 分别称为泛函 λ-矩阵 $A(\lambda)$ 的特征值和与之对应的特征向量．满足(1.1)的特征值 λ 当然也满足

$$\det A(\lambda)=0.\tag{1.2}$$

求解(1.1)或(1.2)称为求解非线性矩阵特征值问题．关于这类问题的求解已有很多有效的二阶收敛的算法(见[1]—[7])．Lancaster[1] 对一般的 λ-矩阵曾经给出一个三阶收敛的算法．该方法每步迭代，需求出常数矩阵$A(\lambda_k)$的全部特征向量．利用常用的QR算法求常数矩阵 $A(\lambda_k)$的全部特征向量约需 $15n^3$ 个运算量(见[9]，pp. 235)，即使 $A(\lambda_k)$

* 1992 年 3 月25日收到.

为对称矩阵也需 $5n^3$ 个运算量(见[9],pp. 282)，显然运算量太大. 另外，Lancaster 是对 $A(\lambda)$ 的特征值 $\mu_n(\lambda)$ 求零点给出的，但当 λ^* 为重特征值时，$\mu_n(\lambda)$ 是不一定可微的[8]. 因此,从理论上讲，用 Lancaster 方法求重特征值是不适宜的. 本文利用列选主元 QR 分解,对解析函数 $\phi(\lambda)=\dfrac{r_{nn}(\lambda)}{\bar{q}_{nn}(\lambda)}$ 求零点给出了一个三阶收敛的算法.本算法每步迭代的运算量为 $\dfrac{2}{3}n^3+7n^2$，远远小于 Lancaster 方法的运算量. 本算法对重特征值亦收敛,特别地,对具有线性初等因子的重特征值三阶收敛.

§2. 算法的导出

对某点 λ，我们可将矩阵 $A(\lambda)$ 进行列选主元 QR 分解(见[9]，pp. 162—165)

$$A(\lambda)P(\lambda)=Q(\lambda)R(\lambda),$$

其中 $P(\lambda)$ 是在进行列选主元中由列互换确定的置换阵，$Q(\lambda)$ 和 $R(\lambda)$ 分别是 $n\times n$ 阶酉矩阵和上三角阵.

$$R(\lambda)=\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix},$$

由列选主元可知，$R(\lambda)$ 的元素满足

$$\begin{cases} |r_{11}|\geqslant|r_{22}|\geqslant\cdots\geqslant|r_{nn}|, \\ |r_{ii}|\geqslant|r_{ij}| \qquad (1\leqslant i\leqslant j\leqslant n). \end{cases} \tag{2.1}$$

Kublanovskaya[2] 1970 年对 $r_{nn}(\lambda)$ 求零点给出一个算法,但该方法的推导有错误. Jain 和 Singhal[3] 1983 年给出一个改进的方法，它们不是对 $r_{nn}(\lambda)$ 求零点,而是对解析函数 $\phi(\lambda)=\dfrac{r_{nn}(\lambda)}{\bar{q}_{nn}(\lambda)}$ 求零点，从而给出一个二阶收敛的算法. 本文的目的是在 Jain 和 Singhal 工作的基础上给出一个三阶收敛的算法,作为对 Lancaster 给出方法的改进.

为了保证 $q_{nn}\neq 0$，我们采取如下方法:

记 $q_n=Qe_n=(q_{1n},q_{2n},\cdots,q_{nn})^T$，如果 $|q_{nn}|>\varepsilon$（事先给定的一个正常数）无需处理;反之,可取初等置换矩阵 T(由单位矩阵交换两列而得到的矩阵),使 Tq_n 的按模最大元素位于矩阵 TQ 的 (n,n) 位置. 设

$$|q_{pn}|=\max\{|q_{1n}|,|q_{2n}|,\cdots,|q_{nn}|\},$$

则 T 由单位矩阵交换第 P 列和第 n 列得到,显然 $T^H=T^{-1}=T$. 记

$$\tilde{Q}=TQ,$$
$$\tilde{q}_n=Tq_n=(\tilde{q}_{1n},\tilde{q}_{2n},\cdots,\tilde{q}_{nn})^T,$$

于是

$$|\tilde{q}_{nn}|\geqslant\frac{\sqrt{\sum_{i=1}^{n}|\tilde{q}_{in}|^2}}{\sqrt{n}}=\frac{1}{\sqrt{n}}\ (>\varepsilon).$$

引入初等置换矩阵T后，矩阵分解式变为

$$T\cdot A(\lambda)P(\lambda)=\tilde{Q}(\lambda)R(\lambda), \tag{2.2}$$

记 $B=TAP$. Jain 和 Singhal[3] 直接对分解式 $A(\lambda)P(\lambda)=Q(\lambda)R(\lambda)$ 求导并给出了 ψ' 的公式。本文从另外一个角度出发给出了 ψ' 和 ψ'' 的公式。对解析函数 $\psi(\lambda)$ 的一、二阶导数为

$$\begin{cases}\psi'(\lambda)=-e_n^T(B^{-1})'e_n/(e_n^TB^{-1}e_n)^2,\\ \psi''(\lambda)=-e_n^T(B^{-1})''e_n/(e_n^TB^{-1}e_n)^2+2[e_n^T(B^{-1})'e_n]^2/(e_n^TB^{-1}e_n)^3.\end{cases} \tag{2.3}$$

由于

$$(B^{-1})'=-B^{-1}B'B^{-1},$$
$$(B^{-1})''=2B^{-1}B'B^{-1}B'B^{-1}-B^{-1}B''B^{-1},$$

于是

$$\begin{cases}\psi'(\lambda)=e_n^TB^{-1}B'B^{-1}e_n/(e_n^TB^{-1}e_n)^2,\\ \psi''(\lambda)=e_n^T(B^{-1}B''B^{-1}-2B^{-1}B'B^{-1}B'B^{-1})e_n/(e_n^TB^{-1}e_n)^2\\ \qquad\qquad+2(e_n^TB^{-1}B'B^{-1}e_n)^2/(e_n^TB^{-1}e_n)^3.\end{cases}$$

当λ充分接近λ^*时，矩阵T和P可视为常数[2]，对矩阵 $B=TAP$ 分别求一、二阶导数得

$$B'=TA'P,\quad B''=TA''P.$$

代入 ψ'、ψ'' 的表达式得

$$\begin{cases}\psi(\lambda)=r_{nn}/\bar{\tilde{q}}_{nn},\\ \psi'(\lambda)=\dfrac{r_{nn}}{\bar{\tilde{q}}_{nn}^2}e_n^T\tilde{Q}^HTA'PR^{-1}\tilde{Q}^He_n,\\ \psi''(\lambda)=\dfrac{2(\psi')^2}{\psi}+\dfrac{r_{nn}}{\bar{\tilde{q}}_{nn}^2}[e_n^T\tilde{Q}^HTA''PR^{-1}\tilde{Q}^He_n\\ \qquad\qquad-2e_n^T\tilde{Q}^HTA'PR^{-1}\tilde{Q}^HTA'PR^{-1}\tilde{Q}^He_n].\end{cases} \tag{2.4}$$

设向量 u、v 并定义为

$$\begin{cases}Ru=\dfrac{r_{nn}}{\bar{\tilde{q}}_{nn}}\tilde{Q}^He_n, & (2.5)\\ \tilde{Q}^HTv=\dfrac{1}{\tilde{q}_{nn}}e_n. & (2.6)\end{cases}$$

引入向量 u、v 之后，式(2.4)可写为

$$\begin{cases}\psi(\lambda)=r_{nn}/\bar{\tilde{q}}_{nn},\\ \psi'(\lambda)=v^HA'Pu,\\ \psi''(\lambda)=\dfrac{2(\psi')^2}{\psi}+v^HA''Pu-2v^HA'PR^{-1}\tilde{Q}^HTA'Pu.\end{cases} \tag{2.7}$$

记

$$R=\begin{array}{c}\left[\begin{array}{cc}R_1 & r\\ 0 & r_{nn}\end{array}\right]\begin{array}{c}{\scriptstyle n-1}\\ {\scriptstyle 1}\end{array}\\ \begin{array}{cc}{\scriptstyle n-1} & {\scriptstyle 1}\end{array}\end{array},\quad \tilde{Q}_{e_n}^H=\begin{pmatrix}\bar{\tilde{q}}_{n1}\\ \vdots\\ \bar{\tilde{q}}_{nn}\end{pmatrix}=\begin{pmatrix}q\\ \\ \bar{\tilde{q}}_{nn}\end{pmatrix}\begin{array}{c}{\scriptstyle n-1}\\ {\scriptstyle 1}\end{array},$$

于是

$$R^{-1}=\begin{bmatrix}R_1^{-1} & -R_1^{-1}r/r_{nn}\\ 0 & 1/r_{nn}\end{bmatrix}=\begin{bmatrix}R_1^{-1} & 0\\ 0 & 0\end{bmatrix}+\begin{pmatrix}-R_1^{-1}r/r_{nn}\\ 1/r_{nn}\end{pmatrix}e_n^H,$$

$$\begin{aligned}&2v^HA'PR^{-1}\widetilde{Q}^HTA'Pu\\
&=2v^HA'P\left[\begin{pmatrix}R_1^{-1} & 0\\ 0 & 0\end{pmatrix}+\begin{pmatrix}-R_1^{-1}r/r_{nn}\\ 1/r_{nn}\end{pmatrix}e_n^H\right]\widetilde{Q}^HTA'Pu\\
&=2v^HA'P\begin{pmatrix}R_1^{-1} & 0\\ 0 & 0\end{pmatrix}\widetilde{Q}^HTA'Pu+2v^HA'P\begin{bmatrix}-R_1^{-1}r/r_{nn}\\ 1/r_{nn}\end{bmatrix}e_n^H\widetilde{Q}^HTA'Pu\\
&=2v^HA'P\begin{pmatrix}R_1^{-1} & 0\\ 0 & 0\end{pmatrix}\widetilde{Q}^HTA'Pu+2\bar{\tilde{q}}_{nn}\cdot v^HA'P\begin{bmatrix}-R_1^{-1}r/r_{nn}\\ 1/r_{nn}\end{bmatrix}\cdot v^HA'Pu\\
&=2v^HA'P\begin{pmatrix}R_1^{-1} & 0\\ 0 & 0\end{pmatrix}\widetilde{Q}^HTA'Pu+2\psi'\bar{\tilde{q}}_{nn}\cdot v^HA'P\begin{bmatrix}-R_1^{-1}r/r_{nn}\\ 1/r_{nn}\end{bmatrix}. \qquad (2.8)\end{aligned}$$

又因为

$$\begin{aligned}u&=\frac{r_{nn}}{\bar{\tilde{q}}_{nn}}R^{-1}\widetilde{Q}^He_n=r_{nn}\begin{bmatrix}R_1^{-1} & -R_1^{-1}r/r_{nn}\\ 0 & 1/r_{nn}\end{bmatrix}\begin{bmatrix}q/\bar{\tilde{q}}_{nn}\\ 1\end{bmatrix}\\
&=r_{nn}\left[\begin{pmatrix}R_1^{-1}q/\bar{\tilde{q}}_{nn}\\ 0\end{pmatrix}+\begin{pmatrix}-R_1^{-1}r/r_{nn}\\ 1/r_{nn}\end{pmatrix}\right],\end{aligned}$$

代入(2.8)式的后半部分

$$\begin{aligned}2\psi'\bar{\tilde{q}}_{nn}\cdot v^HA'P\begin{pmatrix}-R_1^{-1}r/r_{nn}\\ 1/r_{nn}\end{pmatrix}&=2\psi'\bar{\tilde{q}}_{nn}\cdot v^HA'P\left[\frac{u}{r_{nn}}-\begin{pmatrix}R_1^{-1}q/\bar{\tilde{q}}_{nn}\\ 0\end{pmatrix}\right]\\
&=\frac{2(\psi')^2}{\psi}-2\psi'\cdot v^HA'P\begin{pmatrix}R_1^{-1}q\\ 0\end{pmatrix}.\end{aligned}$$

因此

$$\psi''=v^HA''Pu+2v^HA'P\begin{pmatrix}R_1^{-1} & 0\\ 0 & 0\end{pmatrix}\left[\psi'\begin{pmatrix}q\\ 0\end{pmatrix}-\widetilde{Q}^HTA'Pu\right],$$

于是,(2.7)式可写为

$$\begin{cases}\psi(\lambda)=r_{nn}/\bar{\tilde{q}}_{nn},\\ \psi'(\lambda)=v^HA'Pu,\\ \psi''(\lambda)=v^HA''Pu+2v^HA'P\begin{pmatrix}R_1^{-1} & 0\\ 0 & 0\end{pmatrix}\left[\psi'\begin{pmatrix}q\\ 0\end{pmatrix}-Q^HA'Pu\right].\end{cases}\qquad (2.9)$$

由(2.9)式便可构造三阶收敛的算法. 实际上,我们还可以构造更高阶的算法.

定理 2.1. 将矩阵 $A(\lambda)$ 按(2.2)式进行分解,则 $\det A(\lambda)$ 的零点与 $\psi(\lambda)=r_{nn}(\lambda)/\bar{\tilde{q}}_{nn}(\lambda)$ 的零点相同,这里 $\bar{\tilde{q}}_{nn}(\lambda)$ 是酉阵 $\widetilde{Q}$ 的 $(n、n)$ 元素的共轭.

定理 2.2. 设向量 u、v 是由(2.5)、(2.6)定义的两个向量. 当 λ^* 为 $A(\lambda)$ 的特征值时, v 和 Pu 分别是对应 λ^* 的左、右特征向量.

算法. 本算法适用于求解一般非线性矩阵特征值问题,具体步骤如下:

1. 给定初始值 λ_0 和正常数 $\varepsilon(<1)$.
2. 形成矩阵 $A(\lambda_k)$, $k=0、1、2、\cdots$.
3. 对矩阵 $A(\lambda_k)$ 进行列选主元 QR 分解

$$A(\lambda_k)P_k = Q_k R_k.$$

4. 如果 $|q_{nn}| > \varepsilon$、取 $T = I$；否则，按上面给出的方法选取矩阵 T.
5. 按(2.5)给出向量 u.
6. 求向量 $v = \frac{1}{\tilde{q}_{nn}} Qe_n$.
7. 按(2.9)计算 ψ_k、ψ'_k、ψ''_k.
8. 实行 Halley 迭代

$$\lambda_{k+1} = \lambda_k - \psi_k \Big/ \left(\psi'_k - \frac{\psi_k \psi''_k}{2\psi'_k} \right).$$

9. 重复步骤 2—8，直到满足精度要求为止.

本算法能同时给出特征值 λ^* 和与之对应的特征向量. 当 $A(\lambda)$ 为实阵时，若取 λ_1 为实数，则由迭代产生的序列 $\{\lambda_k\}$ 均为实数，不可能收敛到复根. 因此，我们推荐取初始值 λ_0 为复数.

§3. 运算量和收敛性分析

在计算每步迭代的运算量时，由于形成 $A(\lambda_k)$、$A'(\lambda_k)$ 和 $A''(\lambda_k)$ 的运算量依赖于 $A(\lambda)$ 本身. 因此，我们不考虑形成 $A(\lambda_k)$、$A'(\lambda_k)$、$A''(\lambda_k)$ 的运算量. 列选主元 QR 分解约需 $\frac{2}{3}n^3$ 个运算量，计算向量 u、v 共需 $\frac{5}{2}n^2$ 个运算量，计算 ψ'_k 需 n^2 个运算量，计算 ψ''_k 需 $\frac{7}{2}n^2$ 个运算量. 因此，本文给出的算法每步迭代需 $\frac{2}{3}n^3 + 7n^2$ 个运算量.

Jain 和 Singhal 改进的 Kublanovskaya 算法，每步迭代需 $\frac{2}{3}n^3 + \frac{7}{2}n^2$ 个运算量. 因此，本算法比 Jain 和 Singhal 的二阶收敛的方法每步迭代仅多 $\frac{7}{2}n^2$ 个运算量，但收敛阶提高一阶.

Lancaster[1] 曾给出一个三阶收敛的算法，该算法每步迭代的运算量为 $15n^3$. 显然，本文给出的算法的运算量少于 Lancaster 方法.

为了说明本算法对重特征值的收敛性，我们定义：

设 λ^* 是泛函 λ—矩阵 $A(\lambda)$ 的 $\alpha(\geq 2)$ 重特征值，即

$$\det A(\lambda) = O(|\lambda - \lambda^*|^\alpha).$$

如果 $\operatorname{rank} A(\lambda^*) = n - \alpha$，则称 λ^* 是具有线性初等因子的重特征值.

定理 3.1. 设 λ^* 是具有线性初等因子的重特征值，则 $\psi'(\lambda^*) \neq 0$.

证明. 不妨假设 λ^* 为 $\alpha(\geq 2)$ 重特征值，我们用反证法证明本命题：

若 $\psi'(\lambda^*) = 0$，由于 $\psi(\lambda)$ 在 λ^* 点的某个小邻域内是解析的[3]，因此，可对 $\psi(\lambda)$ 在 λ^* 点泰勒展开

$$\psi(\lambda) = \psi(\lambda^*) + \psi'(\lambda^*)(\lambda - \lambda^*) + O(|\lambda - \lambda^*|^2)$$

$$= O(|\lambda - \lambda^*|^2).$$

但 $\psi(\lambda) = r_{nn}/\bar{\bar{q}}_{nn}, |\bar{\bar{q}}_{nn}| \geqslant \varepsilon$，于是

$$r_{nn}(\lambda) = O(|\lambda - \lambda^*|^2).$$

由于 $A(\lambda)$ 是泛函 λ-矩阵且 $\mathrm{rank}A(\lambda^*) = n - \alpha$，故可假设

$$r_{ii}(\lambda) = O(|\lambda - \lambda^*|),\ n - \alpha + 1 \leqslant i \leqslant n - 1.$$

于是

$$\det A(\lambda) = \prod_{i=1}^{n} r_{ii}(\lambda) = O(|\lambda - \lambda^*|^{\alpha+1})$$

这与 $\det A(\lambda) = O(|\lambda - \lambda^*|^{\alpha})$ 产生矛盾．命题得证．

推论． 设 λ^* 是泛函 λ—矩阵 $A(\lambda)$ 的单重特征值,则 $\psi'(\lambda^*) \neq 0$.

定理 3.2. 设 λ^* 是解析函数 $\psi(\lambda)$ 的零点，λ_0 为给定的初值，$\{\lambda_k\}$ 为由 Halley 迭代

$$\lambda_{k+1} = \lambda_k - \psi_k \Big/ \left(\psi_k' - \frac{\psi_k \psi_k''}{2\psi_k'} \right)$$

产生的收敛于 λ^* 的序列,则

1）如果 $\psi'(\lambda^*) = 0$，则 $\{\lambda_k\}$ 三阶收敛．

2）如果 $\psi'(\lambda^*) \neq 0$，则 $\{\lambda_k\}$ 二阶收敛．

证明．记 $e_k = \lambda_k - \lambda^*$

1）$e_{k+1} = e_k - \psi_k \Big/ \left(\psi_k' - \dfrac{\psi_k \psi_k''}{2\psi_k'} \right)$,

$$\frac{\psi_k \psi_k'}{(\psi_k')^2 - \dfrac{\psi_k \psi_k''}{2}} = \left\{ \left[\psi'(\lambda^*) e_k + \frac{\psi''(\lambda^*)}{2} e_k^2 + O(e_k^3) \right] \right.$$

$$\left. \cdot \left[\psi'(\lambda^*) + \psi''(\lambda^*) e_k + \psi'''(\lambda^*) \frac{e_k^2}{2} + O(e_k^3) \right] \right\} \Big/$$

$$\left\{ \left[\psi'(\lambda^*) + \psi''(\lambda^*) e_k + \frac{\psi'''(\lambda^*)}{2} e_k^2 + O(e_k^3) \right]^2 \right.$$

$$- \frac{1}{2} \left[\psi'(\lambda^*) e_k + \frac{\psi''(\lambda^*)}{2} e_k^2 + O(e_k^3) \right]$$

$$\left. \cdot \left[\psi''(\lambda^*) + \psi'''(\lambda^*) e_k + \frac{\psi^{(4)}(\lambda^*)}{2} e_k^2 + O(e_k^3) \right] \right\}$$

$$= \left\{ (\psi'(\lambda^*))^2 e_k + \frac{3}{2} \psi'(\lambda^*) \psi''(\lambda^*) e_k^2 + O(e_k^3) \right\} \Big/$$

$$\left\{ (\psi'(\lambda^*))^2 + \frac{3}{2} \psi'(\lambda^*) \psi''(\lambda^*) e_k \right.$$

$$\left. + \frac{e_k^2}{2} [(\psi''(\lambda^*))^2 + \psi'(\lambda^*) \psi'''(\lambda^*)] + O(e_k^3) \right\},$$

于是

$$
\begin{aligned}
e_{k+1} &= e_k - \frac{\psi_k \psi'_k}{(\psi'_k)^2 - \psi_k \psi''_k/2} \\
&= \Big\{[\psi'(\lambda^*)]^2 e_k + \frac{3}{2}\psi'(\lambda^*)\psi''(\lambda^*)e_k^2 - [\psi'(\lambda^*)]^2 e_k \\
&\quad - \frac{3}{2}\psi'(\lambda^*)\psi''(\lambda^*)e_k^2 + O(e_k^3)\Big\} \Big/ \Big\{[\psi'(\lambda^*)]^2 + \frac{3}{2}\psi'(\lambda^*)\psi''(\lambda^*)e_k \\
&\quad + \frac{e_k^2}{2}\{[\psi''(\lambda^*)]^2 + \psi'(\lambda^*)\psi'''(\lambda^*)\} + O(e_k^3)\Big\} \\
&= \frac{O(e_k^3)}{(\psi'(\lambda^*))^2 + O(e_k)}
\end{aligned}
$$

即 $e_{k+1} = O(e_k^3)\cdot\dfrac{1}{(\psi'(\lambda^*))^2 + O(e_k)}$，三阶收敛.

2) 设 $\psi'(\lambda^*) = \cdots = \psi^{(t)}(\lambda^*) = 0$，$\psi^{(t+1)}(\lambda^*) \neq 0$，$t$ 是大于或等于 1 的正整数. 将 $\psi(\lambda_k)$、$\psi'(\lambda_k)$ 和 $\psi''(\lambda_k)$ 分别在 λ^* 点泰勒展开

$$
\begin{aligned}
\psi(\lambda_k) &= \frac{\psi^{(t+1)}(\lambda^*)}{(t+1)!}\cdot e_k^{t+1} + O(e_k^{t+2}), \\
\psi'(\lambda_k) &= \frac{\psi^{(t+1)}(\lambda^*)}{t!}\, e_k^t + O(e_k^{t+1}), \\
\psi''(\lambda_k) &= \frac{\psi^{(t+1)}(\lambda^*)}{(t-1)!}\, e_k^{t-1} + O(e_k^t).
\end{aligned}
$$

于是

$$
\begin{aligned}
e_{k+1} &= e_k - \frac{2\psi_k \psi'_k}{2(\psi'_k)^2 - \psi_k \psi''_k} \\
&= e_k - \Big\{2\left[\frac{\psi^{(t+1)}(\lambda^*)}{(t+1)!} e_k^{t+1} + O(e_k^{t+2})\right]\left[\frac{\psi^{(t+1)}(\lambda^*)}{t!}\cdot e_k^t + O(e_k^{t+1})\right]\Big\} \Big/ \\
&\quad \Big\{2\left[\frac{\psi_k^{(t+1)}(\lambda^*)}{t!} e_k^t + O(e_k^{t+1})\right]^2 - \left[\frac{\psi^{(t+1)}(\lambda^*)}{(t+1)!} e_k^{t+1} + O(e_k^{t+2})\right] \\
&\quad \cdot\left[\frac{\psi^{(t+1)}(\lambda^*)}{(t-1)!} e_k^{t-1} + O(e_k^t)\right]\Big\} \\
&= \frac{t}{t+2} e_k + O(e_k^2).
\end{aligned}
$$

证毕.

综上所述，本算法的收敛性归纳为:

本算法对单特征值和具有线性初等因子的重特征值三阶收敛，对其它情形的特征值线性收敛.

§4. 数值例子

为了验证算法的收敛性，我们在 VAX 2000 系列计算机上对三个问题进行了数值实

验，计算使用双精度.

例 1. 考虑如下二次特征值问题([3,6]):

$$A(\lambda)=A\lambda^2+B\lambda+C,$$

其中

$$A=\begin{bmatrix}1.76 & 1.28 & 2.89\\ 1.28 & 0.824 & 0.413\\ 2.89 & 0.413 & 0.725\end{bmatrix},\quad B=\begin{bmatrix}7.66 & 2.45 & 2.1\\ 0.23 & 1.04 & 0.223\\ 0.6 & 0.756 & 0.658\end{bmatrix},$$

$$C=\begin{bmatrix}121.0 & 18.9 & 15.9\\ 0.0 & 2.7 & 0.145\\ 11.9 & 3.64 & 15.5\end{bmatrix}.$$

这个问题有三对分离得很好的非线性特征值(只给出 10 位有效数字)，即

$$\lambda_1=-0.9179981715+i1.760584204,\quad \lambda_4=\bar{\lambda}_1,$$

$$\lambda_2=0.09472172578+i2.522876588,\quad \lambda_5=\bar{\lambda}_2,$$

$$\lambda_3=-0.8848302463+i8.441512159,\quad \lambda_6=\bar{\lambda}_3.$$

表 1 给出了用本算法与[3]中给出的方法求解本问题的收敛结果、迭代次数和 e_k (迭代结果与精确特征值的误差). 两种方法迭代的终止条件相同. 可以看出，本算法的迭代次数少于[3]的迭代次数.

表 1

初始值	JAIN 和 SINGHAL 的方法		本文的算法		e_k
	迭代次数	收敛结果	迭代次数	收敛结果	
$-0.9+1.7i$	4	$\lambda1$	2	$\lambda1$	0.65E-13
$-1.0+1.5i$	5	$\lambda1$	3	$\lambda1$	0.25E-16
$2.0i$	8	$\lambda2$	5	$\lambda2$	0.38E-16
$0.1+2.5i$	4	$\lambda2$	2	$\lambda2$	0.24E-14
$2.4i$	5	$\lambda2$	3	$\lambda2$	0.22E-16
$2.5i$	5	$\lambda2$	3	$\lambda2$	0.26E-16
$3.0i$	8	$\lambda2$	4	$\lambda2$	0.24E-16
$4.0i$	6	$\lambda3$	4	$\lambda1$	0.71E-10
$5.0i$	5	$\lambda3$	3	$\lambda3$	0.85E-16
$10.0i$	5	$\lambda3$	3	$\lambda3$	0.53E-16

注 1. 表 1 第二列中的迭代次数有几个比[3]中的迭代次数多一，这是由于迭代终止时要求的精度较高造成的.

注 2. 对于不同的算法，相同的初值可能收敛到不同的特征值.

例 2. 考虑二次对称特征值问题[7]

$$A(\lambda)=\begin{bmatrix}-\lambda^2-3\lambda+1 & & & \\ \lambda^2-1 & -2\lambda^2-3\lambda+5 & & \text{对称}\\ -\lambda^2-3\lambda+1 & \lambda^2-1 & -2\lambda^2-5\lambda+2 & \\ -2\lambda^2-6\lambda+2 & 2\lambda^2-2 & -4\lambda & -9\lambda^2-19\lambda+14\end{bmatrix},$$

这个矩阵有两个二重特征值 1 和 -2，还有四个单特征值 $-4\pm\sqrt{19}$ 和 $-4\pm\sqrt{18}$. 由于二

表 2

次数 \ 数据 \ 初值	1.2+0.1i		−2.01+0.2i	
	收敛结果	$\lvert\lambda_k-\lambda^*\rvert$	收敛结果	$\lvert\lambda_k-\lambda^*\rvert$
1	0.99339215593515 − i0.11E−01	0.13E−01	−1.9865024457538 − i0.27E−01	0.30E−01
2	0.99999458939738 + i0.16E−06	0.54E−05	−2.0001052281430 + i0.24E−04	0.11E−03
3	1.00000000000000 + i0.37E−16	0.54E−15	−1.9999999999999 − i0.10E−12	0.17E−12
4			−2.0000000000000 − i0.95E−29	0.95E−29
5				

重特征值 1 和 −2 都是具有线性初等因子的重特征值，因此，由算法的收敛性结果可知迭代应具有三阶收敛性．表 2 中的数据也说明了这一点，$|\lambda_{k+1}-\lambda^*| \approx c\cdot|\lambda_k-\lambda^*|^3$，$c$ 为某正常数．

例 3. 为了考察算法对其它情形的重特征值的收敛性，我们给出了下面的二次特征值问题([4,6])：

$$A(\lambda)=A\lambda^2+B\lambda+C,$$

其中

$$A=I,$$

$$B=\begin{bmatrix}3\alpha & -(1+\alpha^2+2\beta^2) & \alpha(1+2\beta^2) & -\beta^2(\alpha^2+\beta^2)\\ 2 & 0 & 0 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & 2 & 0\end{bmatrix},$$

$$C=\begin{bmatrix}-1+2\alpha^2 & \alpha(1-\alpha^2-2\beta^2) & 2\alpha^2\beta^2 & -\alpha\beta^2(\alpha^2+\beta^2)\\ 2\alpha & -(\alpha^2+2\beta^2) & 2\alpha\beta^2 & -\beta^2(\alpha^2+\beta^2)\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\end{bmatrix},$$

这里 $\alpha\geqslant 0$ 是参数，且 $\beta=1+\alpha$. $\det A(\lambda)$ 的零点之间的分离程序随 α 的改变而改变．特别当 $\alpha=0$ 时，它有三重特征值 $\pm i$ 和二重特征值 0.

由于 $\operatorname{rank}A(\pm i)=2$，$\operatorname{rank}A(0)=3$，因此 $\pm i$ 和 0 都不是具有线性初等因子的重特征值．表 3 给出了 $\alpha=0$ 时用本算法求解此问题的数值结果．

表 3

初始值	迭代次数	收敛结果
0.1	3	−0.12E−15 − 0.0000000i
−0.01−1.01i	7	−0.46E−05 − 1.0000047i
2+i	11	−0.36E−05 − 0.9999956i
2+2i	12	0.44E−05 − 1.0000008i

对于初值 $2+i$ 和 $2+2i$，[4]给出的迭代次数分别为 22 和 25 次，而对于初值 0.1，[6]迭代了 30 次才精确到 0.5×10^{-5}。可见，本算法对不具有线性初等因子的重特征值也是有效的。

参 考 文 献

[1] P. Lancaster, Lambda matrices and vibrating system, Pergamon Press, Oxford, 1966.

[2] V. N. Kublanovskaya, On an aproach to the solution of the generalized latent value problem for lambda-matrices. SIAM J. Numer. Anal., 7(1970), 532—537.

[3] N. K. Jain, K. Singhal, On Kublanovskaya's approach to the solution of the generalized latent value problem for functional lambda-matrices, SIAM J. Numer. Anal., 20(1983) 1062—1070.

[4] 李仁仓，QR 分解与非线性特征值问题，计算数学，**11**: 4(1989),374—385.

[5] P. Lancaster, A generalized Rayleigh quotient iteration for lambda-matrices, Arch. Rat. Mech. Anal., 8(1961), 309—322.

[6] A. Ruhe, Algorithms for the nonlinear eigenvalue problem, SIAM J. Numer. Anal., 10(1973), 674—689.

[7] Neumaier, Residual inverse iteration for the nonlinear eigenvalue problem, SIAM J. Numer. Anal., 22(1985), 914—923.

[8] Sun Ji-guang, Eigenvalues and eigenvectors of a matrix dependent on several parameters, Journal of computational mathematics, 3(1985), 351—364.

[9] G.H. Golub, C.F. Van Loan, Matrix Computations, The Johns Hopkins University Press, 1983.