Indian J. pure appl. Math., 21(12): 1049-1054, December 1990

COMPUTABLE ERROR ESTIMATES FOR NEWTON'S ITERATIONS FOR REFINING INVARIANT SUBSPACES

M. THAMBAN NAIR

Department of Mathematics, Goa University, Taleigao Plateau Santa-Cruz, P.O. Goa 403005

(*Received* 6 *February* 1990; *accepted* 13 *June* 1990)

Chatelin[1] has established, under assumptions on the unknown invariant subspace, the quadratic convergence of Newton's iterative refinements. We modify the procedure in line with Demmel's[2] suggestions and obtain a criterion for quadratic convergence in terms of the known initial approximation. Our procedure enables computable error estimates to be obtained for the iterations.

Let A be an $n \times n$ matrix with entries in the complex field $\mathbb{C}$. Our aim is to determine computationally one of the maximal invariant subspaces (or generalized eigenspaces) M associated with a set of m eigenvalues of A, counting their algebraic multiplicities, assuming that an m-dimensional initial approximation M_0 is available. Let $X_0 = [u_1,\ldots,u_m]$ be an $n \times m$ matrix the columns of which span M_0, and let $Z = [z_1,\ldots,z_m]$ be an $n \times m$ matrix of adjoint base $\{z_1,\ldots,z_m\}$ to $\{u_1,\ldots,u_m\}$. Then

$$z_i^* u_j = \delta_{ij} \text{ or } Z^* X_0 = I_m$$

where Z^* denotes the conjugate transpose of Z and I_m the $m \times m$ identity matrix.

We seek to construct computationally an $n \times m$ matrix $X = [x_1,\ldots,x_m]$ the columns of which span one of the m-dimensional invariant subspaces M. It is therefore necessary to introduce the following additional condition on X:

$$Z^* X = I_m. \qquad \ldots(1)$$

Consequently, the invariance of M implies

$$AX = XB$$

where $B = (b_{ij})$ is the $m \times m$ matrix defined by

$$Ax_j = \sum_{i=1}^{m} b_{ij}\, x_i, \quad j = 1,\ldots,m.$$

As a consequence of (1) it follows that $B = Z^*AX$ and hence that $AX = XB$ takes the form

$$AX = X(Z^*AX). \qquad \ldots(2)$$

To solve (2), Chatelin[1] proposed the use of Newton's method; namely,

$$X_{k+1} = X_k - F'(X_k)^{-1} F(X_k) \qquad \ldots(3)$$

where

$$F : K \mapsto AK - K(Z^*AK)$$

and F' is the Frechét derivative of F:

$$F'(K)Y = (1 - KZ^*)\,AY - Y\,(Z^*\,AK).$$

The following is the main result of Chatelin[1].

Theorem (Chatelin[1]) – Let the columns of an $n \times m$ matrix X span a generalized eigenspace of A associated with m non-zero eigenvalues of A, counting algebraic multiplicities, and $Z^*X = I_m$. Then there exists $\rho > 0$ such that for any $n \times m$ matrix X_0 satisfying $\|X - X_0\| \leq \rho$, the Newton's iterations (3) are defined and converge quadratically to X as $k \to \infty$.

The above theorem of Chatelin is essentially an existence result based on the assumptions imposed on the unknown matrix X; and consequently the convergence of the method cannot be verified for an initial approximation X_0. The purpose of this note is to remedy this by giving a criterion based on the known matrix X_0 which guarantees the existence of an X satisfying (1) and (2) as well as convergence of (3) to X. Moreover our procedure enables a computable estimate for the errors $\|X - X_k\|$. The motivation for the analysis is the work of Demmel[2]. He obtained results similar to those obtained here by tranforming (2) into a Riccati equation. In fact, Demmel[2] (p. 46) states that "the quadratic convergence criterion does not appear in Chatelin's paper and seems to be stronger than her results". In this paper, we wish how to fill this gap in Chatelin's paper without requiring a transformation.

For this purpose we first rewrite (3) as

$$F'(X_k)\ (X_{k+1} - X_k) = -F'(X_k)$$

which yields the Sylvester equation

$$(1 - X_k\,Z^*)\,AX_{k+1} - X_{k+1}\,(Z^*AX_k) = -X_k\,(Z^*AX_k). \qquad \ldots(4)$$

It is important to note that, if (X_k) converges to $\tilde{X}$, say, then $\tilde{X}$ satisfies (2).

For our analysis we recall the following result. It's proof is given by Nair[3] in a more general context. (c.f. Stewart[5]).

Theorem 1 – Let $\mathfrak{X}_1$ and $\mathfrak{X}_2$ be Banach spaces, and $A_1 : \mathfrak{X}_1 \to \mathfrak{X}_1$ and $A_2 : \mathfrak{X}_2 \to \mathfrak{X}_2$ be bounded linear operators. Then we have the following :

(i) For any bounded linear operator $H : \mathfrak{X}_1 \to \mathfrak{X}_2$, the operator equation

$$A_2\,K - KA_1 = H$$

has a unique (bounded linear operator) solution $K : \mathfrak{X}_1 \to \mathfrak{X}_2$ if and only if

$$\sigma(A_1) \cap \sigma(A_2) = \phi.$$

(ii) Denoting

$$T\,(A_1, A_2) : K \mapsto A_2\,K - KA_1$$

and

$$\operatorname{sep}\,(A_1, A_2) := \begin{cases} \|T(A_1, A_2)^{-1} & \text{if } \sigma(A_1) \cap \sigma(A_2) = \phi \\ 0 & \text{otherwise} \end{cases}$$

we have

$$\text{sep}\ (A_1 + E_1, A_2 + E_2) \geq \text{sep}\ (A_1, A_2) - \|E_1\| - \|E_2\|$$

for any bounded linear operators $E_1 : \mathfrak{X}_1 \to \mathfrak{X}_2$, $E_2 : \mathfrak{X}_2 \to \mathfrak{X}_2$. Here $\sigma(K)$ denotes the spectrum of the operator K and $\|K\|$ denotes an operator norm induced by the space norm.

Let X_0 and Z be $n \times m$ matrices satisfying $Z^* X_0 = I_m$. Then $P_0 = X_0 Z^*$ is a projection matrix of rank m. We introduce the following notation :

$\mathfrak{X}_1 = \mathbb{C}^m$, the space of m - vectors

$\mathfrak{X}_2 =$ the space spanned by the columns of $I - P_0$

$B_0 : \mathfrak{X}_1 \to \mathfrak{X}_1$ defined by $B_0\, x = (Z^* A X_0) x,\ x \in \mathfrak{X}_1$

$C_0 : \mathfrak{X}_2 \to \mathfrak{X}_2$ defined by $C_0\, y = (1 - P_0) A y,\ y \in \mathfrak{X}_2$.

At the outset we assume that

$$\sigma\ (B_0) \cap \sigma\ (C_0) = \phi. \qquad \ldots(5)$$

Then by Theorem 1, it follows that

$$\delta = \text{sep}\ (B_0, C_0) > 0.$$

Let

$$\alpha = \frac{1}{\delta}\ \|(1 - X_0\, Z^*)\, A X_0\|,\ \beta = \frac{1}{\delta}\ \|Z^* A\|$$

$$\epsilon = \alpha\beta \text{ and } g(t) = \begin{cases} (1 - \sqrt{1-4t})/2t, & 0 < t \leq 1/4 \\ 1, & t = 0. \end{cases}$$

It follows that $s = g(t)$ satisfies the equation $ts^2 - s + 1 = 0$, and $1 \leq s \leq 2$ and $g(t_1) \leq g(t_2)$ whenever $0 \leq t_1 \leq t_2 \leq 1/4$.

Proposition 2 — If $\epsilon < 1/4$, then there exists an $n \times m$ matrix X satisfying (1) and (2). Morever

$$\|X - X_0\| \leq \alpha g(\epsilon).$$

PROOF : Let $Y_0 = 0$, and Y_k be the unique solution of the equation

$$C_0\, Y_k - Y_k\, B_0 = -\ (1 - X_0\, Z^*)\, A X_0 + Y_{k-1}\, Z^* A\, Y_{k-1}.$$

Now, exactly as in Nair[4], the following results can be proved by induction on k:

(i) $\|Y_k\| \leq \alpha g(\epsilon)$

(ii) (Y_k) is a Cauchy sequence in the space of $n \times m$ matrices.

Thus, (Y_k) converges to an $n \times m$ matrix Y, say, which also satisfies

$$C_0\, Y - Y B_0 = -\ (1 - X_0\, Z^*)\, A X_0 + Y Z^*\, A Y.$$

Now, taking $X = X_0 + Y$, we see that the above equation is the same as (2). Moreover, since Y_k is an operator from $\mathfrak{X}_1$ to $\mathfrak{X}_2$, we have $Z^* Y_k = 0$ for all $k = 1, 2, \ldots$, so that Y satisfies $Z^* Y = 0$, equivalently, $Z^* X = I_m$. Taking limit as $k \to \infty$ in (i), we obtain the estimate in the proposition.

Proposition 3 — If $\epsilon < 1/12$ then equation (4) is uniquely solvable for X_k, and

$$\|X_k - X_0\| \leq \alpha g\ (3\epsilon). \qquad \ldots(6)$$

Proof : Equation (4) can be written as

$$[(1 - X_0 Z^*) A - D_k Z^*A]\ D_{k+1} - D_{k+1}\ [Z^*AX_0 + Z^*AD_k]$$
$$= -(1 - X_0 Z^*)\ AX_0 - D_k\ Z^*AD_k \qquad \ldots(7)$$

with $D_i = X_i - X_0$, $i = 0, 1, 2, \ldots$. We prove the proposition by induction on k :

For $k = 0$, (7) is

$$C_0 D_1 - D_1 B_0 = -(1 - X_0 Z^*)\ AX_0.$$

Since $\sigma(B_0) \cap \sigma(C_0) = \phi$, by Theorem 1 (i), this equation has a unique solution $D_1 : \mathfrak{X}_1 \to \mathfrak{X}_2$, and it satisfies

$$\|D_1\| \leq \frac{1}{\delta}\ \|(1 - X_0 Z^*)\ AX_0\| = \alpha \leq \alpha\ g(\epsilon).$$

Now, assume that $D_k : \mathfrak{X}_1 \to \mathfrak{X}_2$ exists uniquely and satisfies

$$\|D_k\| \leq \alpha\ g(3\epsilon).$$

By Theorem 1 (ii), we have

$$\text{sep}\ (B_0 + Z^*AD_k, C_0 - D_k Z^*A) \geq \delta - 2\|Z^*A\|\ \|D_k\| \geq \delta\,(1 - 2\,\epsilon\, g(3\epsilon)) > 0. \ \ldots(8)$$

Hence, again by Theorem 1 (ii), (7) has a unique solution D_{k+1}, and it satisfies

$$\|D_{k+1}\| \leq \frac{\|(1 - X_0 Z^*)\ AX_0\| + \|Z^*A\|\ \|D_k\|^2}{\delta\ (1 - 2\ \epsilon\ g\ (3\epsilon)}$$

$$\leq \frac{\alpha + \alpha\ \epsilon\ g(3\epsilon)^2}{1 - 2\ \epsilon\ g(3\epsilon)} = \alpha g(3\epsilon).$$

Next we prove the main theorem of this paper.

Theorem 4 — If $\epsilon < 1/12$, then the equation (4) is uniquely solvable for X_{k+1}, and the sequence (X_k) converges quadratically to X satisfying (1) and (2). Moreover, we have the following error estimates:

$$\|X_{k+1} - X\| \leq \eta\|X_k - X\|^2 \leq \alpha g(\epsilon)\mu^{2^{k+1}-1} \qquad \ldots(9)$$

and

$$\|X_{k+1} - X\| \leq \frac{\alpha\ \epsilon\ g(\epsilon)^2}{1 - 2\beta\|X_k - X_0\|}\ \mu^{2^{k+1}-1} \leq \alpha g(\epsilon)\ \mu^{2^{k+1}-1} \qquad \ldots(10)$$

where

$$\eta = \frac{\beta}{1 - 2\ \epsilon\ g(3\epsilon)} \text{ and } \mu = \frac{2\epsilon}{1 - 4\epsilon} < 1/4.$$

Proof : By Proposition 2, there exists an $n \times m$ matrix X satisfying (1) and (2). Now, using (2) the equations (4) can be written as

$$(1 - X_k Z^*) AE_{k+1} - E_{k+1} (Z^*AX_k) = - E_k (Z^*A)E_k \qquad ...(11)$$

with $E_i = X_i - X$, $i = 0, 1, 2, \ldots$. Using the relation (8) equation (11) is uniquely solvable for E_{k+1}, and

$$\|E_{k+1}\| \leq \frac{\|Z^*A\| \, \|E_k\|^2}{\delta - 2\|Z^*A\| \, \|D_k\|} = \frac{\beta}{(1 - 2\beta\|D_k\|)} \|E_k\|^2. \qquad ...(12)$$

Now, using the bound for $\|D_k\|$ from (6), we obtain

$$\|X_{k+1} - X\| \leq \eta\|X_k - X\|^2.$$

Now, the relation $\|X_0 - X\| \leq \alpha g(\epsilon)$ from Proposition 2 gives,

$$\eta\|X_0 - X\| \leq \eta\alpha g(\epsilon) \leq \frac{2\epsilon}{1 - 4\epsilon} = \mu < 1/4$$

so that inequalities in (9) follow. Also, $\|E_k\| \to 0$ as $k \to 0$. From (9) and (12), we also have

$$\|X_{k+1} - X\| \leq \frac{\beta}{(1 - 2\beta\|X_k - X_0\|)} (\alpha g(\epsilon)\mu^{2^k - 1})^2$$

$$= \frac{\alpha\epsilon g(\epsilon)^2}{(1 - 2\beta\|X_k - X_0\|)} \mu^{2^{k+1} - 2} \leq \alpha g(\epsilon)\mu^{2^{k+1} - 1}.$$

This completes the proof.

Remark 5 : Since columns of X span an invariant subspace of A we see that

$$\sigma(A) = \sigma(B) \cup \sigma(C)$$

where

$$Bx = (Z^*AX)x \text{ for } x \in \mathbb{C}^m$$

and

$$Cy = (1 - XZ^*) Ay \quad \text{for} \quad y \in \text{range of } (1 - XZ^*).$$

Note also that,

$$Bx = B_0 x + Z^*A(X - X_0)x$$

and

$$Cy = C_0 y - (X - X_0) Z^*Ay.$$

Hence,

$$\begin{aligned} \text{sep } (B, C) &\geq \text{sep } (B_0, C_0) - \|Z^*A \ (X - X_0)\| - \|(X - X_0)Z^*A\| \\ &\geq \delta - 2\|Z^*A\| \, \|X - X_0\| \\ &\geq \delta - 2\delta\beta\alpha g(\epsilon) > 0 \end{aligned}$$

for $\epsilon < 1/4$. Hence by Theorem 1, $\sigma(B) \cap \sigma(C) = \phi$. Now, by a characterization result for generalized eigenspaces (see Theorem 3.2 in Nair[4]), the columns of X span a generalized eigenspace of A associated with m eigenvalues, counting algebraic multiplicities; and these eigenvalues are precisely the eigenvalues of B. Thus, the assumption on X of Chatelin is a consequence of our result.

Acknowledgement

Author thanks Professor Robert Anderssen for carefully reading the manuscript and suggesting many improvements in the presentation, and the Centre for Mathematical Analysis, Australian National University, for their support..

References

1. F. Chatelin, *Computing Suppl.* **5** (1984), 67-74.
2. J. W. Demmel, *Computing* **38** (1987), 43-57.
3. M. T. Nair, Ph.D. Thesis, Indian Institute of Technology, Bombay, 1984.
4. M. T. Nair, *J. Indian Math. Soc.* **54** (1989), 1-14.
5. G. W. Stewart, *SIAM Review* **15** (1973), 752-64.