17. Inverse Iteration Method with a Complex Parameter

By Toshio SUZUKI
Department of Mathematics, Yamanashi University

(Communicated by Shokichi IYANAGA, M. J. A., March 12, 1992)

1. Introduction. The inverse iteration method has been used as one of the most powerful ways for computing eigenvectors. And the theory of error estimates has been established almost completely as is seen in Wilkinson [2]. Let A be a symmetric (n,n) matrix and let $\{\lambda_k, \phi_k\}$, $k=1, \cdots, n$, be pairs of eigenvalues and the corresponding eigenvectors of A. The inverse iterarion process for the eigenvector ϕ_j is to solve the following linear equations with initial data $z^{(0)}$ under the conditions $|\lambda_j-\lambda| \ll c < |\lambda_k-\lambda|$, $(k\neq j)$:

$$(1.1) \qquad (A-\lambda I)z^{(m+1)}=z^{(m)}, \quad m=0,1,2,\cdots.$$

In this paper, we propose to introduce into this method a new technique, which is simple but effective in practical computations. Our method is to solve the same linear equation, but with a complex parameter $\lambda+\sqrt{-1}\tau$ instead of real λ in (1.1) and to carry out the next iteration process after substituting the imaginary part of the solution for the initial vector. We can show that the imaginary part y of the solution of the linear equation contains the component of the aimed eigenvector far more than the real part x. The ratio of the l^2 norms $\|x\|/\|y\|$ can be used to derive a sharp error estimate for the computed eigenvector. It may be emphasized that the error bound given by (2.6) in Theorem 2.2 is rather effective so that one can judge how many digits in actual computations are correct in significant decimals by estimating the right hand side of (2.6). It is also emphasized that in our method the efficiency of enriching the component of the aimed eigenvector is almost doubled compared with the standard traditional method.

In § 2, we explain our method and state the theorems. In § 3, we show some propositions which describe how our method works. When we refer to the traditional method based on (1.1), we call it, for brevity, the standard method. Our main purpose here is to present the idea of our method. So throughout this paper, we state our theory as if rounding errors were zero.

2. A new method with a complex parameter and the theorems. Let A be a real (n,n) matrix which is symmetric and has n different eigenvalues. Let $\{\lambda_k, \phi_k\}$, $k=1,2,3,\cdots,n$, be pairs of eigenvalues and the corresponding normalized real eigenvectors of A. First we describe our method for computing the eigenvector ϕ_j corresponding to the eigenvalue λ_j under the following assumption.

Assumption H. *Eigenvalues λ_k, $k=1,2,\cdots,n$, of A are known with the following accuracy: There are three numerical constants c, ε and λ such that* $\inf_{k\neq j}|\lambda_j-\lambda_k|>2c$, $|\lambda_j-\lambda|<\varepsilon$ *and* $0<2\varepsilon<c$.

Let ξ be an initial vector and let τ be a positive number smaller than ε. Our iteration process consists of the following three steps where $u^{(m)}$ and $v^{(m)}$ are real vectors.

$$(2.1)\qquad (A-\lambda^{(m)}I-\sqrt{-1}\,\tau I)w^{(m)}=z^{(m)} \qquad \text{where } z^{(0)}=\xi,\ \lambda^{(0)}=\lambda,$$

$$(2.2)\qquad z^{(m+1)}=\frac{v^{(m)}}{\|v^{(m)}\|} \qquad \text{where } w^{(m)}=u^{(m)}+\sqrt{-1}\,v^{(m)},$$

$$(2.3)\qquad \lambda^{(m+1)}=\begin{cases}(Az^{(m+1)},z^{(m+1)}), & \text{if } \|v^{(m)}\|>\|u^{(m)}\|,\\ \lambda^{(m)}, & \text{otherwise.}\end{cases}$$

The most essential and characteristic feature of our process is the second step (2.2) where the imaginary part of the solution of the linear equation (2.1) is taken as an approximating eigenvector. In the third step (2.3), we change the value λ to a better approximating value obtained by Rayleigh-Ritz formula where the inequality $|\lambda_j-\lambda^{(m+1)}|<\varepsilon$ also holds as is seen in Proposition 3.1 later. The following theorem guarantees that this iteration process works well.

Theorem 2.1. *If the assumption* H *is satisfied, the iteration process* (2.1)–(2.3) *excites the component of the eigenvector ϕ_j, namely $z^{(m)}\to\pm\phi_j$ as $m\to\infty$, provided $|\tau|<\varepsilon$.*

Before the proof of Theorem 2.1, we need some preparations to simplify the notations. Consider the following equation with $\|z\|=1$:

$$(2.4)\qquad (A-\lambda I-\sqrt{-1}\,\tau I)w=z.$$

Let $z=\sum_{k=1}^n a_k\phi_k$. Then we have

$$(2.5)\qquad w=\sum_{k=1}^n\frac{1}{\lambda_k-\lambda-\sqrt{-1}\,\tau}a_k\phi_k$$

$$=\sum_{k=1}^n\frac{\lambda_k-\lambda}{(\lambda_k-\lambda)^2+\tau^2}a_k\phi_k+\sqrt{-1}\sum_{k=1}^n\frac{\tau}{(\lambda_k-\lambda)^2+\tau^2}a_k\phi_k.$$

Put $x_k=(\lambda_k-\lambda)/((\lambda_k-\lambda)^2+\tau^2)a_k\phi_k$ and $y_k=\tau/((\lambda_k-\lambda)^2+\tau^2)a_k\phi_k$. Let $x=\sum_{k=1}^n x_k$ and $y=\sum_{k=1}^n y_k$.

Proof of Theorem 2.1. For the proof we observe that

$$\frac{\|y_j\|}{\|y_k\|}=\frac{(|\lambda_k-\lambda|^2+\tau^2)}{(|\lambda_j-\lambda|^2+\tau^2)}\frac{|a_j|}{|a_k|}>\frac{|\lambda_k-\lambda|^2}{2\varepsilon^2}\frac{|a_j|}{|a_k|}>\frac{c^2}{2\varepsilon^2}\frac{|a_j|}{|a_k|}.$$

Thus the component of the aimed eigenvector ϕ_j is more excited than that of any other eigenvector ϕ_k with the relative order greater than $2(c/2\varepsilon)^2>1$ because $c>2\varepsilon$. This means that the components of ϕ_k's for $k\neq j$ in $z^{(m)}$ dump down exponentially through iterations. So $z^{(m)}$ approaches to $\pm\phi_j$ as $m\to\infty$.

Next, we consider the error estimates. As is seen from (2.2) and (2.3) of our process, y is our approximation and y_j is the true eigenvector, so that we have only to estimate $\|y-y_j\|$.

Theorem 2.2. *Put $\delta=\|x\|/\|y\|$. Under the assumption H, the rela-*

tive error $\|y-y_j\|/\|y\|$ *is estimated as*

(2.6) $$\frac{\|y-y_j\|}{\|y\|}\leq\frac{\tau}{c}\delta.$$

Proof. Since

(2.7) $$\|x_k\|=\frac{|\lambda_k-\lambda|}{|\lambda_k-\lambda|^2+\tau^2}|a_k|=\frac{|\lambda_k-\lambda|}{|\tau|}\ \frac{|\tau|}{|\lambda_k-\lambda|^2+\tau^2}|a_k|\geq\frac{c}{|\tau|}\|y_k\|,$$

we have

(2.8) $$\|y-y_j\|^2=\sum_{k\neq j}\|y_k\|^2\leq\sum_{k\neq j}\frac{\tau^2}{c^2}\|x\|^2\leq\frac{\tau^2}{c^2}\|x\|^2$$

Thus (2.6) follows.

Error estimate for the standard method is as follows. Let $(A-\lambda I+E)\tilde{w}=z^{(m)}$ with the rounding error term E. The estimate $\|E\|\leq K\sqrt{n}2^{-t}$ is known for t-digits floating point computers. Put $\tilde{z}=\tilde{w}/\|\tilde{w}\|$ and $\eta=(A-\lambda I)\tilde{z}$. Then $\|\eta\|\leq\|E\|+1/\|\tilde{w}\|$ is satisfied. Let $\tilde{z}=\sum a_k\phi_k$ with $\|\tilde{z}\|^2=\sum a_k^2=1$. The known error estimate for $\tilde{z}-a_j\phi_j$ is as follows (see Atkinson [1]):

(2.9) $$\|\tilde{z}-a_j\phi_j\|\leq\frac{1}{c}\|\eta\|.$$

In the corresponding inequality (2.6) of our method, the estimating value for the value $\|\tilde{z}-a_j\phi_j\|$ is smaller than the right hand sipe of (2.9) about by the factor δ. (Note that $\|\tilde{w}\|$ is about as large as $|\lambda-\lambda_j|^{-1}$ so that $\|\eta\|$ is about as small as $|\lambda-\lambda_j|\sim\tau$.)

3. Propositions and remarks. In this section we state supplementary propositions and remarks. The proofs will be published elsewhere with some generalizations. The first proposition shows that in the iteration process (2.1)–(2.3) the inequality $|\lambda_j-\lambda|<\varepsilon$ in the assumption H continues to hold after the approximating eigenvalues are replaced in (2.3).

Proposition 3.1. *Let x and y be the real and the imaginary part of the solution of the equation* (2.4) *with $|\tau|<\varepsilon$ under the assumption H in which the inequality $|\lambda_j-\lambda|<\varepsilon$ is assumed. Put $\tilde{\lambda}=(Ay,y)/\|y\|^2$. If $\|y\|\geq\|x\|$, then $|\lambda_j-\tilde{\lambda}|<\varepsilon$.*

In order to compare the efficiencies of our method and the standard one, the following proposition is important, where λ and λ' should be considered to be the parameters used in our method and in the standard one, respectively.

Proposition 3.2. *Choose τ as $\tau<\varepsilon$ and suppose that $|\lambda_j-\lambda|<\tau$ and $|\lambda_j-\lambda'|<\tau$. Let λ_l be the eigenvalue which attains $\inf_{k\neq j}|\lambda_k-\lambda|$, that is, $|\lambda_l-\lambda|=\inf_{k\neq j}|\lambda_k-\lambda|$. If*

(3.1) $$\frac{|\lambda_j-\lambda|^2+\tau^2}{|\lambda_l-\lambda|^2+\tau^2}\leq\frac{|\lambda_j-\lambda'|}{|\lambda_l-\lambda'|}$$

then, for any $k\neq j$,

(3.2) $$\frac{|\lambda_j-\lambda|^2+\tau^2}{|\lambda_k-\lambda|^2+\tau^2}\leq\frac{|\lambda_j-\lambda'|}{|\lambda_k-\lambda'|}\left(\frac{c+\varepsilon}{c-\varepsilon}\right).$$

Remark 3.3. It is seen from the proof of Proposition 3.2 that, if $\lambda=\lambda'$, the inequality (3.2) can be rewritten in the following form:

$$\frac{|\lambda_j-\lambda|^2+\tau^2}{|\lambda_k-\lambda|^2+\tau^2}\leq\frac{|\lambda_j-\lambda'|}{|\lambda_k-\lambda'|}. \tag{3.3}$$

Since the terms of the left and the right hand sides of (3.3) represent the relative exciting rates of the aimed eigenvectors of two methods, we can derive the range of τ from the inequality (3.1) where our method works better than the standard one. Put $J=|\lambda_j-\lambda|$, $L=|\lambda_l-\lambda|$, $J'=|\lambda_j-\lambda'|$ and $L'=|\lambda_l-\lambda'|$. Then the inequality (3.1) is equivalent to the following:

$$\frac{\tau^2}{J'^2}\leq\frac{1}{L'-J'}\left[\frac{L^2}{J'}-\frac{J^2}{J'^2}L'\right]. \tag{3.4}$$

As one sufficient condition for τ/J' to satisfy (3.4), we can give the following one provided that $4\varepsilon<c$ and $J\leq J'$:

$$\frac{\tau^2}{J'^2}\leq\frac{c}{\varepsilon}. \tag{3.5}$$

As far as the condition numbers of the linear operators in equations (2.1) and (1.1) are concerned, it can be said that our method is of advantage to the standard one, since the value c/ε in (3.5) is taken far greater than 1 in many cases.

Remark 3.4. The inequality (3.5) can be interpreted as asserting that with the value of τ determined by $|\lambda_j-\lambda'|=\tau\sqrt{\varepsilon/c}$ our method works as effective as the standard one with the parameter λ'.

Remark 3.5. In the case $\tau\ll c$, the inequality (3.3) shows that the relative excitation rate of our method is nearly of order two while that of the standard one is of order one.

4. Numerical examples. We carried out some numerical experiments for symmetric matrices and had satisfactory results which indicate that our method works well as theory indicates. Here we first present a simple but non-symmetric example which describes our theory more clearly. It can be easily seen that, when we know almost exact eigenvalues, our method can be applied to the non symmetric matrix with different real simple eigenvalues if (2.3) in the iteration process is omitted.

Example 4.1. We consider the case $n=2$. Let matrix $A=(a_{p,q})$ with $a_{1,1}=1$, $a_{1,2}=1$, $a_{2,1}=10^{-10}$, $a_{2,2}=1$. We know the eigenvalues and eigenvectors: $\lambda_1=1.00001$, $\lambda_2=0.99999$, $\phi_1=(1, 10^{-5})$ and $\phi_2=(1, -10^{-5})$. (As for this matrix A, see Peters and Wilkinson [3].) We denote the supremum norm by $\|*\|_\infty$ and the normalized computed eigenvector by $\phi^{(m)}$. The results are shown in Tables (1–1) and (1–2).

Table (1–1) The results of our method

λ_j	τ	m	$\|\phi_j-\phi^{(m)}\|_\infty$
1. 00001	10^{-10}	0	$.100\times10^{-14}$
. 99999	10^{-10}	0	$.100\times10^{-14}$

Table (1–2) The results of the standard method

λ_j	$\|\lambda_j-\lambda\|$	m	$\|\phi_j-\phi^{(m)}\|_\infty$
1.00001	10^{-15}	0	$.100\times10^{-14}$
.99999	10^{-15}	0	$.100\times10^{-14}$

Though eigenvectors are not orthogonal, it is easily seen that Proposition 3.2 is valid for l^∞ norm instead of l^2 norm and so is the inequality (3.5). The values of τ and $|\lambda_j-\lambda|$ are chosen, following Remark 3.4, so that our method brings better excitation than the standard one.

Next example is the case of a symmetric matrix with a greater size.

Example 4.2. We tried the case $n=20$ and $\lambda_k=2.0+(k-1)\times10^{-5}$, $k=1,2,\cdots,20$. The matrix A is given by $A=\sum_{k=1}^{20}\lambda_k\phi_k\phi_k^t$ where ϕ_k's are orthonormal vectors previously constructed from randomly selected column vectors. We show here in Tables (2–1) and (2–2) the results of only one case $\lambda_j=2.00009$ which lies in the midst of λ_k's.

Table (2–1) The results of our method

$\|\lambda_j-\lambda^{(0)}\|$	τ	m	$\|\phi_j-\phi^{(m)}\|$	δ	$\tau\delta/c$
10^{-13}	10^{-9}	0	$.643\times10^{-8}$	$.130\times10^{-3}$	$.261\times10^{-7}$
10^{-13}	10^{-9}	1	$.613\times10^{-16}$	$.679\times10^{-12}$	$.136\times10^{-15}$
10^{-9}	10^{-9}	0	$.129\times10^{-7}$	$.100\times10^{1}$	$.200\times10^{-3}$
10^{-9}	10^{-9}	1	$.123\times10^{-15}$	$.168\times10^{-11}$	$.335\times10^{-15}$
10^{-15}	10^{-11}	0	$.643\times10^{-12}$	$.100\times10^{-3}$	$.200\times10^{-9}$
10^{-15}	10^{-11}	1	$.613\times10^{-24}$	$.679\times10^{-18}$	$.136\times10^{-23}$

Table (2–2) The results of the standard method

$\|\lambda_j-\lambda\|$	m	$\|\phi_j-\phi^{(m)}\|$	μ/c
10^{-13}	0	$.835\times10^{-8}$	$.544\times10^{-7}$
10^{-13}	1	$.643\times10^{-16}$	$.200\times10^{-7}$
10^{-15}	0	$.835\times10^{-10}$	$.544\times10^{-9}$
10^{-15}	1	$.643\times10^{-20}$	$.200\times10^{-9}$

These results show that the iterations (2.1)–(2.3) works well and Remark 3.4 is also true here. Especially, it should be emphasized that the value $(\tau\delta/c)$ of the error bound is very close to the computated value of $\|\phi_j-\phi^{(m)}\|$.

Remark 4.3. Our numerical tests were done by HITAC M-682H (at Computer Centre, University of Tokyo) and computations were carried out in quadruple precision to avoid the influences of rounding errors as much as possible.

References

[1] Atkinson, K. E.: An Introduction to Numerical Analysis. 2nd ed., John Wiley and Sons Inc., New York (1989).
[2] Wilkinson, J. H.: The Algebraic Eigenvalue Problem. Oxford Univ. Press, London (1965).
[3] Peters, G. and Wilkinson, J. H.: Inverse iteration, ill-conditioned equations and Newton's method. SIAM Review, **21**, 339–360 (1979).