An Algorithm for Finding Generalized Eigenpairs of a Symmetric Definite Matrix Pencil

Jiang Erxiong

Department of Mathematics
Fudan University
Shanghai, People's Republic of China

Submitted by Hans Schneider

ABSTRACT

A new method for finding eigenpairs of any symmetric definite matrix pencil is proposed. It is a modification of the well-known Rayleigh quotient iteration. It circumvents a drawback of that method, i.e., its nonconvergence or slow convergence for poor initial vectors. The global convergence and cubic convergence rate of the new method are proved. Its efficiency is also demonstrated by numerical results.

1. INTRODUCTION

The symmetric positive definite generalized eigenvalue problems

\[ Ax = \lambda Bx, \]

where \( A \) is an \( n \times n \) real symmetric matrix and \( B \) is an \( n \times n \) real symmetric positive definite matrix, often arise from practical problems, e.g., vibration problems. We call \( \lambda \) the eigenvalue of the matrix pencil \((A, B)\), and \( x \) the eigenvector corresponding to \( \lambda \). A well-known method to solve this problem, which is referred to as Rayleigh quotient iteration, is listed below; see [1, p. 317; 5, p. 319].

1. Choose an initial vector \( x_1, \| x_1 \| = 1; 1 \rightarrow k \).
2. Calculate \( \rho_k = (Ax_k, x_k)/(Bx_k, x_k) \). Solve the system of linear equations for \( z_{k+1} \)

\[
( A - \rho_k B ) z_{k+1} = Bx_k.
\]
3. If $\|z_{k+1}\|$ is large enough, go to step 4; otherwise

$$x_{k+1} = z_{k+1}/\|z_{k+1}\|,$$

$$k + 1 \rightarrow k; \quad \text{go to step 2}.$$

4. $(\rho_k, x_k)$ is the approximate eigenpair of $(A, B)$. 

Hereafter the inner product $(\cdot, \cdot)$ and norm $\| \|$ are both Euclidean.

The Rayleigh quotient iteration retains the structure of matrices $A$ and $B$. Therefore sparsity can be utilized. When the initial vector $x_1$ is good, this method converges very quickly. The convergence rate is cubic. But there do exist cases where it does not converge or converges very slowly. We can see this in an example in Section 2. There is no convergence theorem for Rayleigh quotient iteration. In this paper we modify Rayleigh quotient iteration so that the modified methods converge globally and at the same time the convergence rates are at least cubic. We refer to these modified methods as MRQI. We also obtain a convergence theorem and an error estimate for the approximate eigenpair $(\rho_k, x_k)$.

In Section 2, we propose the algorithm MRQI and a basic theorem. Some numerical results are presented in Section 2 to show the efficiency of MRQI. We prove the basic theorem in Section 3.

2. MRQI AND THE BASIC THEOREM

Let $A$ be an $n \times n$ real symmetric matrix, and $B$ be an $n \times n$ real symmetric positive definite matrix. Consider the generalized eigenvalue problem

$$Ax = \lambda Bx. \quad (1)$$

It is well known that $(1)$ has $n$ real eigenvalues $\lambda_i$, $i = 1, \ldots, n$. Here we assume $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. The corresponding eigenvectors are denoted as $z_1, z_2, \ldots, z_n$. In addition, we assume

$$(Bz_i, z_j) = \delta_{ij},$$

i.e., $\{z_i\}$ forms a $B$-orthogonal basis of $R^n$. 
In order to find the eigenpair \((\lambda_i, z_i)\), we propose the following MRQI algorithm.

**Algorithm I.**

1. Choose an initial vector \(x_1\), \((Bx_1, x_1) = 1; 1 \rightarrow k\).
2. Calculate \(\rho_k = (Ax_k, x_k)\). Solve \(Bd_k = Ax_k - \rho_k Bx_k\).
3. Calculate \(a_k = (Ad_k, d_k)/(Bd_k, d_k)\).
4. If \((Bd_k, d_k) < \epsilon\), go to step 5.
5. Solve \((A - \omega_k B)x_{k+1} = \tau Bx_k\)

where \(\tau\) is a number chosen so as to make
\[ (Bx_{k+1} + x_{k+1}) = 1 \quad \text{and} \quad (Bx_k, x_{k+1}) > 0. \]

\(k + 1 \rightarrow k; \) go to 2.

5. \((\rho_k, x_k)\) is an approximate eigenpair; stop.

If we choose \(\epsilon = 0\) and assume \(\omega_k\), \(k = 1, 2, \ldots\), are not eigenvalues of \((A, B)\), we can compute the sequence \(\{(\rho_k, x_k)\}_1^\infty\) by Algorithm I.

**Basic Theorem.** For any initial vector \(x_1\), \((Bx_1, x_1) = 1\). If \(\omega_k\), \(k = 1, 2, \ldots\), are not eigenvalues of \((A, B)\), then:

1. \(\lim_{k \to \infty}(Bd_k, d_k) = 0\), and

\[ (Bd_{k+1}, d_{k+1}) = M_k \frac{(Bd_k, d_k)}{(a_k - \omega_k)^2}. \]

where \(|M_k| \leq \text{const.}\)
2. \(\lim_{k \to \infty} \rho_k = \lambda_{i_1}\), where \(\lambda_{i_1}\) is one eigenvalue of \((A, B)\). If we denote by \(\lambda_{j(k)}\) the eigenvalue of \((A, B)\) nearest to \(\rho_k\), then

\[
|\lambda_{j(k)} - \rho_k| \leq \min \left( (Bd_k, d_k)^{1/2}, \frac{(Bd_k, d_k)}{\sigma_k} \right)
\]

where \(\sigma_k = \min_{\lambda_i \neq \lambda_{i_1}} |\lambda_i - \rho_k|\). When \(k\) is large enough, \(\lambda_{i_1} = \lambda_{j(k)}\).

3. \(\lim_{k \to \infty} x_k = z_{i_1}\), where \(z_{i_1}\) is the eigenvector corresponding to \(\lambda_{i_1}\). Moreover \((Bz_{i_1}, z_{i_1}) = 1\). We have

\[
\|x_k - z_{i_1}\| \leq \sqrt{\frac{2}{1 + |\cos \theta_k|} \mu_1} \left( \frac{(Bd_k, d_k)^{1/2}}{\eta_k} \right),
\]

[\[
|\sin \theta_k| \leq \frac{(Bd_k, d_k)^{1/2}}{\eta_k},
\]

where \(\eta_k = \min_{\lambda_i \neq \lambda_{i_1}} |\lambda_i - \rho_k|\) and \(\mu_1\) is the smallest eigenvalue of \(B\).

This theorem will be proved in next section.

**Remark 1.** If \(\omega_k\) is an eigenvalue of \((A, B)\), then \(x_{k+1}\) is the approximate eigenvector in practical computation. In this case, \((Bd_{k+1}, d_{k+1})\) must be very small. Therefore, \((\rho_{k+1}, x_{k+1})\) is the approximate eigenpair.

**Remark 2.** Step 4 in Algorithm I is the inverse iteration. Instead of taking \(\rho_k\) as the shift as in Rayleigh quotient iteration, we choose \(\omega_k\) as the shift for MRQI. This is the reason why we call our method modified Rayleigh quotient iteration.

**Remark 3.** To reduce the cost of MRQI, we store the factorized form of \(B\), i.e., the lower triangle and upper triangle of the LU decomposition of \(B\). The extra cost for the solution of the linear systems in step 2 is negligible compared with that for the solution of linear systems in step 4, which is also needed in Rayleigh quotient iteration.
When we need to compute more than one eigenpair of \((A, B)\), we have the following algorithm.

**Algorithm II.**

1. Choose \(x_1, (Bx_1, x_1) = 1; 1 \rightarrow i.\)
2. \(1 \rightarrow k.\)
3. Calculate \(\rho_k = (Ax_k, x_k).\) Solve
   \[
   Bd_k = Ax_k - \rho_k Bx_k.
   \]
   Calculate \(a_k = (Ad_k, d_k)/(Bd_k, d_k).\)
4. If \((Bd_k, d_k) < \varepsilon\) go to step 6.
5. Calculate \(\Delta = (a_k - \rho_k)/2\) and
   \[
   \omega_k = \rho_k - \text{sign}(\Delta) \frac{(Bd_k, d_k)}{|\Delta| + \sqrt{\Delta^2 + (Bd_k, d_k)}}.
   \]
   Solve
   \[
   (A - \omega_k B)x_{k+1} = \tau Bx_k,
   \]
   where \(\tau\) is chosen so as to make
   \[
   (Bx_k, x_k) = 1 \quad \text{and} \quad (Bx_{k+1}, x_k) \geq 0.
   \]
   \(k + 1 \rightarrow k;\) go to step 3.
6. \((\rho_k, y_i) = (\rho_k, x_k)\) is the \(i\)th approximate eigenpair of \((A, B)\). If \(i = m,\) stop; otherwise, choose a new vector \(y,\) and compute the new initial vector \(x_1\) as
   \[
   \tilde{x}_1 = y - \sum_{j=1}^{i} c_j y_j
   \]
   where \(c_j = (By, y_j), j = 1, \ldots, i.\) If \(\tilde{x}_1 \neq 0,\) take \(x_1 = c\tilde{x}_1\) which satisfies \((Bx_1, x_1) = 1; i + 1 \rightarrow i;\) go to step 2. Otherwise, choose another vector \(y;\) again compute the initial vector \(x_1\) by the same process.

**Remark 4.** Algorithm II computes \(m\) eigenpairs of \((A, B).\)
EXAMPLE 1.

\[
A = \begin{bmatrix}
1 & 3 & 4 & 1 & 11 \\
2.1 & 7 & 3 & 6.3 \\
3 & 3.5 & 6 \\
4 & 8 & \\
\text{Symm.} & 1
\end{bmatrix},
\]

\[
B = \begin{bmatrix}
4 & -1 & -1 & 0 & -1 \\
7 & 0 & -2 & -1 \\
3 & -1 & 0 \\
9 & -3 & \\
\text{Symm.} & 6.5
\end{bmatrix}
\]

We compute all eigenpairs of \((A, B)\) by Algorithm II and list the results below:

\[
\lambda_1 = -1.7412895, \quad \lambda_2 = -0.9853912,
\]

\[
\lambda_3 = 0.2267455, \quad \lambda_4 = 0.2697041,
\]

\[
\lambda_5 = 13.5032461.
\]

The corresponding eigenvectors are

\[
z_1 = ( -0.2892533, -0.0081512, -0.1143412, -0.0147079, 0.279432)^T,
\]

\[
z_2 = (0.0470529, 0.2732705, -0.3993868, 0.0375245, 0.0382863)^T,
\]

\[
z_3 = ( -0.1376874, 0.1914548, 0.2426654, -0.1408566, -0.1327007)^T,
\]

\[
z_4 = ( -0.2648449, 0.0541929, 0.0408549, 0.2657377, -0.0566337)^T,
\]

\[
z_5 = (0.3855991, 0.2646692, 0.4103834, 0.2753487, 0.3516049)^T.
\]

In many software packages (for example the routine ecczs in the IMSL package), this definite generalized eigenvalue problem is transformed into standard eigenvalue problem by means of Cholesky factorization of the positive definite matrix \(B\). Methods based on factorization destroy the structure of the matrices \(A\) and \(B\), so they can not utilize sparsity of \(A\) and \(B\). On the contrary, MRQI does not destroy the structure of \(A\) and \(B\). It is
even more efficient when both $A$ and $B$ are sparse matrices. We can see this in the following example.

**Example 2.**

\[
A = \begin{bmatrix}
1 & 2 & 3 \\
2 & 3 & 4 \\
3 & \ddots & \ddots \\
& \ddots & \ddots & \ddots \\
& & 100 & 100 \\
\end{bmatrix},
\]

\[
B = \begin{bmatrix}
6 & -1 & -1 \\
7 & -1 & \ddots \\
8 & \ddots & -1 & -1 \\
\end{bmatrix},
\]

We computed all eigenpairs of this problem by MRQI and EICZS respectively on the Honeywell DPS-8 computer. It cost $0.2647766 \times 10^{-1}$ CPU hours to run a code for MRQI, whereas it cost $0.33972 \times 10^{-1}$ CPU hours to run EICZS. The difference of CPU time between MRQI and EICZS increases with $n$.

**Example 3.** In Reference [4], Li and Rhee proposed a homotopy algorithm to compute all eigenvalues of the symmetric tridiagonal matrix as listed in Table 1 (when $B = I$). For a homotopy algorithm one needs to solve linear systems such as that appearing at step 3 of MRQI a number of times. In Table 2, we compare the numbers of times these systems had to be solved for computing all the eigenvalues by various methods.

**Example 4.** In Table 3 we give an example to show that the Rayleigh quotient iteration converges very slowly for certain initial vectors. The symmetric tridiagonal matrix is the same as in Example 3. The initial vector is $x_1 = (0.11482345 - 0.06534155 0.04742620 - 0.04767902 - 0.01355686 0.07597307 - 0.22031045 - 0.13380648 0.48749217 0.73066570 0.30980320 0.19241876 0.03222235 - 0.00804191)$ for both RQI and MRQI.
### TABLE 1

<table>
<thead>
<tr>
<th>$i$</th>
<th>$a_i$</th>
<th>$\beta_i$</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.25000000</td>
<td>0.00000000</td>
<td>0.06437991</td>
</tr>
<tr>
<td>2</td>
<td>0.768491173</td>
<td>0.153660746</td>
<td>0.07359712</td>
</tr>
<tr>
<td>3</td>
<td>0.919556756</td>
<td>0.467260328</td>
<td>0.08422527</td>
</tr>
<tr>
<td>4</td>
<td>0.230938895</td>
<td>0.119256498</td>
<td>0.09720922</td>
</tr>
<tr>
<td>5</td>
<td>0.133053788</td>
<td>0.080763539</td>
<td>0.10321576</td>
</tr>
<tr>
<td>6</td>
<td>0.222549575</td>
<td>0.033947196</td>
<td>0.12278752</td>
</tr>
<tr>
<td>7</td>
<td>0.116127856</td>
<td>0.036090904</td>
<td>0.14342288</td>
</tr>
<tr>
<td>8</td>
<td>0.120339373</td>
<td>0.035022375</td>
<td>0.16632460</td>
</tr>
<tr>
<td>9</td>
<td>0.123719912</td>
<td>0.029157561</td>
<td>0.17130756</td>
</tr>
<tr>
<td>10</td>
<td>0.128561407</td>
<td>0.037453705</td>
<td>0.17735634</td>
</tr>
<tr>
<td>11</td>
<td>0.107668089</td>
<td>0.016090599</td>
<td>0.23163948</td>
</tr>
<tr>
<td>12</td>
<td>0.137032923</td>
<td>0.023826407</td>
<td>0.26773329</td>
</tr>
<tr>
<td>13</td>
<td>0.138057030</td>
<td>0.029468449</td>
<td>0.46276620</td>
</tr>
<tr>
<td>14</td>
<td>0.103796943</td>
<td>0.007646394</td>
<td>1.33403484</td>
</tr>
</tbody>
</table>

### TABLE 2

<table>
<thead>
<tr>
<th>No. of times</th>
<th>Homotopy</th>
<th>RQI</th>
<th>MRQI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>7</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>78</td>
<td>53</td>
<td>39</td>
</tr>
</tbody>
</table>

### TABLE 3

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Steps of iter.</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>RQI</td>
<td>17</td>
<td>0.14342288</td>
</tr>
<tr>
<td>MRQI</td>
<td>2</td>
<td>0.14342288</td>
</tr>
</tbody>
</table>
3. THE PROOF OF THE BASIC THEOREM

Consider the generalized eigenvalue problem

\[ Ax = \lambda Bx. \]

Since \( B \) is a symmetric positive definite matrix, \( B^{1/2} \) exists. The generalized eigenvalue problem is equivalent to

\[ B^{-1/2}AB^{-1/2}B^{1/2}x = \lambda B^{1/2}x. \]

Denoting \( \tilde{x} = B^{1/2}x \), \( \tilde{A} = B^{-1/2}AB^{-1/2} \), we obtain the equivalent ordinary eigenvalue problem

\[ \tilde{A}\tilde{x} = \lambda \tilde{x}. \]

In (2), let the eigenvalues by \( \lambda_1, \lambda_2, \ldots, \lambda_n \), which are the same as those in (1), and correspond to unit eigenvectors \( \tilde{z}_1, \tilde{z}_2, \ldots, \tilde{z}_n \) respectively. We consider the following algorithm for the problem (2):

**Algorithm III.**

1. Choose the initial vector \( \tilde{x}_1, \|\tilde{x}_1\| = 1; 1 \rightarrow k. \)
2. Calculate

\[
\tilde{\rho}_k = (\tilde{A}\tilde{x}_k, \tilde{x}_k), \quad r_k = \tilde{A}\tilde{x}_k - \tilde{\rho}_k\tilde{x}_k, \\
b_k = \|r_k\|, \quad y_k = r_k/b_k, \\
\tilde{\alpha}_k = (\tilde{A}y_k, y_k), \quad \tilde{\Delta} = \frac{\tilde{\alpha}_k - \tilde{\rho}_k}{2}, \\
\tilde{\omega}_k = \tilde{\rho}_k - \text{sign}(\tilde{\Delta}) \frac{b_k^2}{|\tilde{\Delta}| + \sqrt{\tilde{\Delta}^2 + b_k^2}}.
\]
3. Solve

\[
(\tilde{A} - \tilde{\omega}_k I)\tilde{x}_{k+1} = \tilde{\tau}\tilde{x}_k,
\]
where $\tau$ is a number chosen so as to make

$$(\bar{x}_{k+1}, \bar{x}_k) \geq 0 \quad \text{and} \quad \|\bar{x}_{k+1}\| = 1.$$

4. If $b_k < \epsilon$, go to set 5; otherwise

$$k + 1 \rightarrow k; \quad \text{go to step 2.}$$

5. $(\tilde{\rho}_k, \bar{x}_k)$ is the approximate eigenpair of $\tilde{A}$; stop.

Given $x_1$, if no $\omega_k$ is an eigenvalue of $(A, B)$, we can compute a sequence \{$(\rho_k, x_k)$\} by Algorithm I for the generalized problem (1). If no $\bar{\omega}_k$ is an eigenvalue of $\tilde{A}$ and $\epsilon = 0$, then from Algorithm III we have a sequence \{$(\tilde{\rho}_k, \bar{x}_k)$\}. There are some relations between \{$(\rho_k, x_k)$\} and \{$(\tilde{\rho}_k, \bar{x}_k)$\}. We give the following lemma without proof, for brevity.

**Lemma 1.** In Algorithm III, if $\bar{x}_1 = B^{1/2}x_1$, assume that $\omega_k$ are not eigenvalues of $(A, B)$, Then

$$d_k = B^{-1/2}r_k, \quad \tilde{\rho}_k = \rho_k, \quad \bar{x}_k = B^{1/2}x_k. \quad (3)$$

Therefore, proving the convergence of the sequence \{$(\rho_k, x_k)$\} produced by Algorithm I reduces to proving the convergence of the sequence \{$(\tilde{\rho}_k, \bar{x}_k)$\} produced by Algorithm III.

Now let us prove the convergence of \{$(\tilde{\rho}_k, \bar{x}_k)$\}.

For the initial vector $\bar{x}_1$ in Algorithm III, there exists an orthogonal matrix

$$W = [\bar{x}_1, s_2, \ldots, s_n], \quad (4)$$

with $\bar{x}_1$ as its first column, such that

$$W^T\tilde{A}W = T_1,$$

where $T_1$ is a symmetric tridiagonal matrix. In fact, this is the Lanczos tridiagonalization. To explore the relations between Algorithm III and the QL algorithm, we compute $T_k$ by the following QL algorithm with shifts $\bar{\omega}_k$ (see [5, p. 140]):

$$T_k - \bar{\omega}_k I = Q_k L_k,$$

$$T_{k+1} = L_k Q_k + \bar{\omega}_k I, \quad (5)$$
where $Q_k$ is an orthonormal matrix and $L_k$ is a lower triangular matrix. Denote

$$T_k = \begin{bmatrix}
\alpha_1^{(k)} & \beta_1^{(k)} & & \\
\beta_1^{(k)} & \alpha_2^{(k)} & \beta_2^{(k)} & \\
& \ddots & \ddots & \ddots \\
& & \ddots & \ddots & \ddots \\
& & & \ddots & \beta_{n-1}^{(k)} \\
& & & & \beta_{n-1}^{(k)} & \alpha_n^{(k)}
\end{bmatrix},$$

$$Q_k = [q_1^{(k)}, q_2^{(k)}, \ldots, q_n^{(k)}],$$

$$L_k = (l_{ij}^{(k)}).$$

Let $l_{ii}^{(k)} > 0$, $i = 2, 3, \ldots, n$, and the sign of $l_{11}^{(k)}$ be such that $q_{11}^{(k)} > 0$, i.e., the first component of

$$t_1^{(k)} = \frac{\sum_{j=2}^{n} l_{j1}^{(k)} q_j^{(k)}}{l_{11}^{(k)}}$$

is nonnegative; here $t_1^{(k)}$ is the first column of $T_k - \tilde{\omega}_k I$. Under this rule, if $T_k$ is irreducible, then the decomposition of (5) is unique too.

**Lemma 2.** If we denote

$$P_0 = W,$$

$$P_k = WQ_1Q_2 \cdots Q_k, \quad k = 1, 2, \ldots,$$

and no $\tilde{\omega}_k$ is an eigenvalue of $\tilde{A}$, then:

1. $\tilde{A}P_k = P_k T_{k+1}.$
2. Let $P_k = [s_1^{(k+1)}, s_2^{(k+1)}, \ldots, s_n^{(k+1)}];$ then

   $$s_1^{(k+1)} = \tilde{x}_{k+1}, \quad s_2^{(k+1)} = \pm y_{k+1}.$$

3. $\tilde{\beta}_k = \alpha_1^{(k)}; \quad \tilde{\alpha}_k = \alpha_2^{(k)}; \quad b_k = |\beta_1^{(k)}|.$
Proof. 1: Since \( T_1 = W^T \tilde{A} W \), we have \( \tilde{A} W = W T \), i.e.,

\[
\tilde{A} P_0 = P_0 T.
\]

It is easy to see

\[
T_{k+1} = Q_k^T Q_{k-1}^T \cdots Q_1^T Q_1 Q_2 \cdots Q_k
\]

\[
= P_k^T \tilde{A} P_k.
\]

Hence, we have

\[
\tilde{A} P_k = P_k T_{k+1}.
\]

3: From the QL process, we get

\[
(T_1 - \tilde{\omega}_1 I) q_1^{(1)} = l_{11}^{(1)} e_1,
\]

i.e.,

\[
(\tilde{A} - \tilde{\omega}_1 I) W q_1^{(1)} = l_{11}^{(1)} W e_1 = l_{11}^{(1)} x_1.
\]

The sign of \( l_{11}^{(1)} \) makes \( q_{11}^{(1)} \geq 0 \), so

\[
(W q_1^{(1)}, x_1) = (q_1^{(1)}, W^T x_1) = (q_1^{(1)}, e_1) = q_{11}^{(1)} \geq 0.
\]

Since \( \tilde{\omega}_1 \) is not an eigenvalue of \( \tilde{A} \), we have

\[
W q_1^{(1)} = \tilde{x}_2,
\]

i.e.,

\[
P_1 e_1 = \tilde{x}_2.
\]

By induction, we can prove

\[
P_k e_1 = s_1^{(k+1)} = \tilde{x}_{k+1}.
\]
From

$$\tilde{A}P_{k-1} = P_{k-1}T_k$$

(6)

Comparing the first columns of the two sides of (6), we obtain

$$\tilde{A}\tilde{x}_k = \tilde{x}_k\alpha_1^{(k)} + \delta_2^{(k)}\beta_1^{(k)}.$$  

(7)

Taking the inner product with $\tilde{x}_k$, we have

$$(\tilde{A}\tilde{x}_k, \tilde{x}_k) = \alpha_1^{(k)},$$

i.e.,

$$\tilde{\rho}_k = \alpha_1^{(k)},$$

and

$$\beta_1^{(k)}\delta_2^{(k)} = \tilde{A}\tilde{x}_k - \alpha_1^{(k)}\tilde{x}_k$$

$$= \tilde{A}\tilde{x}_k - \tilde{\rho}_k\tilde{x}_k = r_k.$$ 

Hence

$$b_k = ||r_k|| = ||\beta_1^{(k)}||,$$

$$\delta_2^{(k)} = \frac{r_k}{\beta_1^{(k)}} = \begin{cases} y_k & \text{if } \beta_1^{(k)} > 0, \\ -y_k & \text{if } \beta_1^{(k)} < 0. \end{cases}$$

Comparing the second columns of the two sides of (6), we get

$$\tilde{A}\delta_2^{(k)} = \tilde{\beta}_1^{(k)}\tilde{x}_k + \alpha_2^{(k)}\delta_2^{(k)} + \beta_2^{(k)}\delta_3^{(k)}.$$ 

Taking the inner product with $\delta_2^{(k)}$, we have

$$\alpha_2^{(k)} = (\tilde{A}\delta_2^{(k)}, \delta_2^{(k)})$$

$$= (\tilde{A}y_k, y_k) = \tilde{a}_k.$$ 

$\blacksquare$
Theorem 1. In Algorithm III, if no $\tilde{\omega}_k$ is an eigenvalue of $\tilde{A}$, then:

1. $b_k = \|r_k\| \to 0$ and $b_{k+1} = M_k b_k c_k^2 / |a_k - \tilde{\omega}_k|$, where $|M_k| \leq \text{const.}$

2. $\lim_{k \to \infty} \tilde{p}_k = \lambda_{i1}$, where $\lambda_{i1}$ is one eigenvalue of $A$. If $\lambda_j(k)$ is the eigenvalue of $\tilde{A}$ nearest to $\tilde{p}_k$, then

$$|\lambda_j(k) - \tilde{p}_k| \leq \min \left( b_k, \frac{\sigma_k b_k^2}{\sigma_k} \right),$$

where $\sigma_k = \min_{\lambda \neq \lambda_{i1}} |\lambda - \tilde{p}_k|$. When $k$ is large enough,

$$\lambda_j(k) = \lambda_{i1}.$$

3. $\lim_{k \to \infty} \bar{x}_k = \tilde{z}_{i1}$, where $\tilde{z}_{i1}$ is the unit eigenvector of $\tilde{A}$ corresponding to $\lambda_{i1}$. Denote $(\tilde{z}_{i1}, \bar{x}_k) = \cos \theta_k$; then

$$||\bar{x}_k - \tilde{z}_{i1}|| \leq \sqrt{\frac{2}{1 + |\cos \theta_k|}} \frac{b_k}{\eta_k},$$

$$|\sin \theta_k| \leq b_k / \eta_k,$$

where

$$\eta_k = \min_{\lambda \neq \lambda_{i1}} |\tilde{p}_k - \lambda_i|.$$

Proof. 1: By Lemma 2, we know $\tilde{p}_k = \alpha_1^{(k)}$, $\bar{a}_k = \alpha_2^{(k)}$, $b_k = |\beta_1^{(k)}|$, and so

$$\tilde{\omega}_k = \alpha_1^{(k)} - \text{sign}(\Delta_k) \frac{\beta_1^{(k)} \beta_2^{(k)}}{|\Delta_k| + \sqrt{\Delta_k^2 + \beta_2^{(k)}}},$$

where $\Delta_k = (\alpha_2^{(k)} - \alpha_1^{(k)})/2$. The shift $\tilde{\omega}_k$ is exactly Wilkinson's shift for the shifted QL algorithm. By the results of Wilkinson [6], the element $\beta_1^{(k)}$ of $T_k$ converges to zero for any $T_1$, i.e., $b_k$ converges to zero. It can be shown that

$$\hat{\beta}_1 = M_k \frac{\beta_1^3 \beta_2^2}{|\alpha_2 - \omega_k|},$$
where \( \hat{\beta}_1 = \beta_1^{(k+1)} \), \( \beta_1 = \beta_1^{(k)} \), \( \beta_2 = \beta_2^{(k)} \), and \( \alpha_2 = \alpha_2^{(k)} \), and where \( |M_k| \leq \text{const} \); see [2]. Therefore we obtain the result

\[
b_{k+1} = M_k \frac{\beta_{k+1}^2}{|\tilde{a}_k - \omega|}, \quad c_k = \beta_2.
\]

2: According to Theorem 2 of Reference [2], we know that for any shift sequence \( \{ \epsilon_k \} \), if \( \beta_1 \to 0 \) and \( \alpha_1 - \epsilon_k \to 0 \), then \( \alpha_1^{(k)} \to \lambda_{i1} \), where \( \lambda_{i1} \) is the eigenvalue of \( T_{\lambda} \). For Wilkinson’s shift, it has been shown that \( \beta_1 \to 0 \) and \( |\alpha_1 - \tilde{\alpha}_k| \leq \beta_k^{(k)} \to 0 \). Therefore we have

\[
\alpha_1^{(k)} \to \lambda_{i1},
\]
i.e.,

\[
\tilde{\rho}_k \to \lambda_{i1}.
\]

On the other hand, \( \tilde{\rho}_k \) is the Rayleigh quotient of \( \tilde{A} \) with \( \tilde{x}_k \). From the equality

\[
\tilde{A} \tilde{x}_k - \tilde{\rho}_k \tilde{x}_k = r_k
\]

we know there exists at least one eigenvalue \( \lambda_{j(k)} \) such that

\[
|\lambda_{j(k)} - \tilde{\rho}_k| \leq \min \left( \|r_k\|, \frac{\|r_k\|^2}{\sigma_k} \right);
\]

see [5, p. 222].

Since \( \tilde{\rho}_k \to \lambda_{i1} \), we know that \( \lambda_{i1} \) is the eigenvalue nearest to \( \tilde{\rho}_k \), provided \( k \) is large enough. Hence \( \lambda_{i1} = \lambda_{j(k)} \).

3: Denote

\[
\tilde{x}_1 = g_1 \tilde{z}_{i1} + \sum_{\lambda_j \neq \lambda_{i1}} h^{(1)}_{i} \tilde{z}_j.
\]

If the multiplicity of \( \lambda_{i1} \) is larger than one, the unit vector \( \tilde{z}_{i1} \) is the linear combination of the eigenvectors corresponding to \( \lambda_{i1} \).
It is easy to see

\[ \tilde{x}_2 = (\tilde{A} - \tilde{\omega}_1 I)^{-1} \tau \tilde{x}_1 \]
\[ = g_2 \tilde{z}_{t1} + \sum_{\lambda_i \neq \lambda_{t1}} h_i(2) \tilde{z}_i. \]

In general we have

\[ \tilde{x}_k = g_k \tilde{z}_{t1} + \sum_{\lambda_i \neq \lambda_{t1}} h_i(k) \tilde{z}_i. \] (8)

On the other hand,

\[ r_k = \tilde{A} \tilde{x}_k - \tilde{\rho}_k \tilde{x}_k \]
\[ = g_k (\lambda_{t1} - \tilde{\rho}_k) \tilde{z}_{t1} + \sum_{\lambda_i \neq \lambda_{t1}} h_i(k) (\lambda_i - \tilde{\rho}_k) \tilde{z}_i, \]
\[ \|r_k\|^2 = g_k^2 (\lambda_{t1} - \tilde{\rho}_k)^2 + \sum_{\lambda_i \neq \lambda_{t1}} h_i(k)^2 (\lambda_i - \tilde{\rho}_k)^2. \]

Denote \( \eta_k = \min_{\lambda_i \neq \lambda_{t1}} |\tilde{\rho}_k - \lambda_i|; \) then
\[ \lim_{k \to \infty} \eta_k = |\lambda_{t1} - \lambda_{t1}| = \sigma_0 > 0. \]

For the sake of convenience, we denote
\[ \Sigma^* = \sum_{\lambda_i \neq \lambda_{t1}} h_i(k)^2 = (\sin \theta_k)^2. \]

Since \( \|r_k\|^2 \geq \eta_k^2 \Sigma^* \), we have
\[ |\sin \theta_k| \leq \frac{b_k}{\eta_k}. \]

From \( \|r_k\| \to 0 \) and \( g_k = \cos \theta_k \), we know
\[ \Sigma^* = (1 - g_k^2) \to 0, \]
i.e.,
\[ g_k^2 \to 1. \]
By (8), we have
\[ x_k - \tilde{z}_{i1} = (g_k - 1)\tilde{z}_{i1} + \sum_{\lambda_i \neq \lambda_{i1}} h^{(k)}_i \tilde{z}_i, \]
\[ \|x_k - \tilde{z}_{i1}\|^2 = (1 - g_k)^2 + \Sigma^* \]
\[ = 2(1 - g_k). \]

In order to prove the convergence of \( x_k \), we notice
\[ (x_{k+1}, x_k) = g_{k+1}g_k + \sum_{\lambda_i \neq \lambda_{i1}} h^{(k+1)}_i h^{(k)}_i. \]

The first term of the right side of the above equality dominates when \( k \) is large enough. Since \( (x_{k+1}, x_k) > 0 \), \( g_k \) and \( g_{k+1} \) must be of the same sign. Without the loss of generality, we assume \( g_k > 0 \). From
\[ (1 - g_k)(1 + g_k) = 1 - g_k^2 \]
\[ = \Sigma^* \leq \frac{b_k^2}{\eta_k^2}, \]
we get
\[ 0 \leq 1 - g_k \leq \frac{b_k^2}{(1 + g_k)\eta_k^2}, \]
i.e.,
\[ \|x_k - \tilde{z}_{i1}\| = \sqrt{2}(1 - g_k)^{1/2} \]
\[ \leq \sqrt{\frac{2}{1 + g_k}} \frac{b_k}{\eta_k}. \]
Since \( g_k = \cos \theta_k = (\bar{x}_k, \bar{z}_{i_1}) \), we have

\[
\|\bar{z}_{i_1} - \bar{x}_k\| \leq \sqrt{\frac{2}{1 + \cos \theta_k}} \frac{b_k}{\eta_k}.
\]

At the same time

\[
|\sin \theta_k| = \left( \Sigma_{**} \right)^{1/2} \frac{b_k}{\eta_k}.
\]

As \( b_k \to 0 \), we have \( \sin \theta_k \to 0 \), i.e., \( \cos \theta_k \to 1 \); hence we deduce \( \bar{x}_k \to \bar{z}_{i_1} \).

**COROLLARY 1.** If for some \( k \)

\[
\|r_k\| + \frac{\|r_k\|}{\sqrt{1 - (\|r_k\|/\eta_k)^2}} < \sigma_0,
\]

where \( \sigma_0 = \min(|\lambda_{i_1} - \lambda_{i_1}|, |\lambda_{i_1 + 1} - \lambda_{i_1}|) \), then

\[
\lambda_{i_1} = \lambda_{j(k)} \quad \text{and} \quad \eta_k = \sigma_k.
\]

**Proof.** Since \( |\bar{\rho}_k - \lambda_{j(k)}| \leq \|r_k\| \) and \( g_k^2(|\bar{\rho}_k - \lambda_{i_1})^2 \leq \|r_k\|^2 \), we know

\[
|\bar{\rho}_k - \lambda_{i_1}| \leq \frac{\|r_k\|}{|g_k|} \leq \frac{\|r_k\|}{\sqrt{1 - (\|r_k\|/\eta_k)^2}}.
\]

Therefore

\[
|\lambda_{i_1} - \lambda_{j(k)}| \leq |\lambda_{i_1} - \bar{\rho}_k| + |\bar{\rho}_k - \lambda_{j(k)}|
\leq \|r_k\| + \frac{\|r_k\|}{\sqrt{1 - (\|r_k\|/\eta_k)^2}} < \sigma_0.
\]

The last inequality indicates

\[
\lambda_{j(k)} = \lambda_{i_1}.
\]
For the QL algorithm with Wilkinson’s shift, [7, 8] gave a proof of that in almost any case \( \beta_2 \to 0 \). We give the results as follows:

**Lemma 3.** Let

\[
T = \begin{bmatrix}
\alpha_1 & \beta_1 & 0 \\
\beta_1 & \alpha_2 & \beta_2 \\
& \ddots & \ddots \\
0 & \cdots & \beta_{n-1} & \alpha_n
\end{bmatrix}
\]

be an \( n \times n \) irreducible symmetric tridiagonal matrix, and \( \lambda_1, \lambda_2, \ldots, \lambda_n \) be its eigenvalues corresponding to eigenvectors \( s_1, s_2, \ldots, s_n \), respectively. Let \((i_1, i_2, \ldots, i_n)\) be an arbitrary permutation of \((1, 2, \ldots, n)\). Then every principal minor determinant of the matrix \([s_{i_1}, s_{i_2}, \ldots, s_{i_n}]\) is nonzero.

**Proof.** Let \( \Omega_n = [s_{i_1}, s_{i_2}, \ldots, s_{i_n}] \), and \( \Omega_k \) be its \( k \times k \) principal submatrix, \( k = 1, 2, \ldots, n \). By Paige’s formula [5], \( \Omega_1 = s_{i_1} \neq 0 \). Therefore \( \det(\Omega_1) \neq 0 \). By induction, if \( \det(\Omega_k) \neq 0 \), \( k = 1, 2, \ldots, m \), we prove \( \det(\Omega_{m+1}) \neq 0 \) as follows. If on the contrary \( \det(\Omega_{m+1}) = 0 \), there exists a nonzero vector \( p = (p_1, p_2, \ldots, p_{m+1})^T \) such that

\[
\Omega_{m+1} p = 0. \tag{9}
\]

Without loss of generality, let \( p_{m+1} = -1 \), and \( s_j^{(k)} \) be a \( k \)-vector obtained by cutting off \( s_j \)'s last \( n - k \) components.

Now, (9) can be written as

\[
\sum_{j=1}^{m} p_j s_j^{(m+1)} = s_{i_{m+1}}^{(m+1)}. \tag{10}
\]
Obviously,

$$\sum_{j=1}^{m} p_j s_{ij}^{(m)} = s_{i_{m+1} i}^{(m)}.$$  \hfill (11)

Let $T_m$ be an $m \times m$ principal submatrix of $T$. Left-multiplying (10) by an $m \times (m + 1)$ matrix,

$$\begin{bmatrix} T_m & 0 \\ \end{bmatrix} \beta_m = G_m,$$

we have

$$\sum_{j=1}^{m} p_j G_m s_{ij}^{(m+1)} = G_m s_{i_{m+1} i}^{(m+1)}.$$  \hfill (12)

Since

$$G_m s_{k}^{(m+1)} = \lambda_k s_{k}^{(m)},$$

by (12)

$$\sum_{j=1}^{m} p_j \lambda_j s_{ij}^{(m)} = \lambda_{i_{m+1}} s_{i_{m+1} i}^{(m)}.$$  \hfill (13)

From (11) and (13),

$$\sum_{j=1}^{m} p_j (\lambda_{ij} - \lambda_{i_{m+1}}) s_{ij}^{(m)} = 0.$$  \hfill (14)

But $\lambda_{ij} - \lambda_{i_{m+1}} \neq 0$, so (14) contradicts $\det(\Omega_m) \neq 0$, and so proves $\det(\Omega_{m+1}) \neq 0$.

**Theorem 2.** In the QL process (5), suppose the shift $\tilde{\omega}_k$ is Wilkinson's shift and no $\tilde{\omega}_k$ is an eigenvalue of $T_1$. Let $\lambda_{\infty}$ be the limit of $\alpha_{\infty}^{(k)}$, and the permutation ($t_1, t_2, \ldots, t_n$) be defined as follows:

$$0 = |\lambda_{t_1} - \lambda_{t_1}| < |\lambda_{t_2} - \lambda_{t_1}| < |\lambda_{t_3} - \lambda_{t_1}| < \cdots < |\lambda_{t_n} - \lambda_{t_1}|.$$
If the condition

$$|\lambda_{i_2} - \lambda_{i_1}| < |\lambda_{i_3} - \lambda_{i_1}|$$  \tag{15}$$
holds, then $\beta_2^{(k)} \to 0$.

Proof. From (5),

$$T_{k+1} - \tilde{Q}_kT_k\tilde{Q}_k \quad \text{and} \quad \prod_{j=1}^{k} (T_1 - \tilde{\omega}_j) = \tilde{Q}_k\tilde{L}_k,$$

where

$$\tilde{Q}_k = Q_1Q_2\cdots Q_k, \quad \tilde{L}_k = L_kL_{k-1}\cdots L_1.$$ 

Every diagonal element of $\tilde{L}_k$, except the first one $\tilde{l}_{11}$, is positive. When $\tilde{l}_{11} < 0$, set $D = \text{diag}( -1 \cdot I)$; when $\tilde{l}_{11} > 0$, set $D = I$. So we get

$$\prod_{j=1}^{k} (T_1 - \tilde{\omega}_j) = (\tilde{Q}_kD)(DL\tilde{L}_k),$$

which is $QL$ decomposition in the common sense, that is, every diagonal element of $DL\tilde{L}_k$ is positive. We know that such a decomposition is unique.

Let the Jordan decomposition of the matrix $T_1$ be

$$T_1 = S\text{diag}(\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_n})S^T.$$ 

Therefore,

$$\prod_{j=1}^{k} (T_1 - \tilde{\omega}_j) = S\text{diag} \left( \prod_{j=1}^{k} (\lambda_{i_1} - \tilde{\omega}_j), \ldots, \prod_{j=1}^{k} (\lambda_{i_n} - \tilde{\omega}_j) \right) S^T$$

$$= S\Lambda_kS^T,$$

where

$$\Lambda_k = \text{diag} \left( \prod_{j=1}^{k} (\lambda_{i_1} - \tilde{\omega}_j), \ldots, \prod_{j=1}^{k} (\lambda_{i_n} - \tilde{\omega}_j) \right).$$
If \( \tilde{\omega}_j \) is Wilkinson's shift, we have \( \tilde{\omega}_j \to \lambda_i \), so when \( j \) is large enough,

\[
|\lambda_{i_1} - \tilde{\omega}_j| < |\lambda_{i_2} - \tilde{\omega}_j| < |\lambda_{i_3} - \tilde{\omega}_j| \leq \cdots \leq |\lambda_{i_m} - \tilde{\omega}_j|
\]

and

\[
\frac{|\lambda_{i_1} - \tilde{\omega}_j|}{|\lambda_{i_2} - \tilde{\omega}_j|} < \frac{|\lambda_{i_2} - \tilde{\omega}_j|}{|\lambda_{i_3} - \tilde{\omega}_j|} \leq \eta < 1.
\]

Set \( \eta^{(k)}_m = \prod_{j=1}^{k}(\lambda_{i_m} - \tilde{\omega}_j) \), \( m = 1, 2, \ldots, n \). We have

\[
\frac{\eta^{(k)}_1}{\eta^{(k)}_2} \to 0, \quad \frac{\eta^{(k)}_2}{\eta^{(k)}_3} \to 0,
\]

and generally,

\[
\frac{\eta^{(k)}_{m-2}}{\eta^{(k)}_m} \to 0, \quad m > 4.
\]

By Lemma 3, the Crout decomposition of the matrix \( S \) exists:

\[
S = LR,
\]

where \( L \) is a lower triangular matrix and \( R \) is a upper unit triangular matrix. Therefore, \( S^T = S^{-1} = R^{-1}L^{-1} \) and

\[
\prod_{j=1}^{k} (T_j - \tilde{\omega}_j) = S\Lambda_k S^{-1} = S\Lambda_k R^{-1}\Lambda_k^{-1} \Lambda_k L^{-1}.
\]

Obviously,

\[
\Lambda_k R^{-1}\Lambda_k^{-1} \to \begin{bmatrix} I_2 & 0 \\ 0 & I_{n-2} + E \end{bmatrix},
\]

where \( E \) is a strict upper triangular matrix whose nonzero elements appear
only in the upper diagonal part. Denote the QL decomposition of the matrix \( \Lambda_k R^{-1} \Lambda_k^{-1} \) by

\[
\Lambda_k R^{-1} \Lambda_k^{-1} = \hat{Q}_k \hat{L}_k,
\]

where every diagonal element of \( \hat{L}_k \) is positive. Then,

\[
\hat{Q}_k \rightarrow \begin{bmatrix} I & 0 \\ 0 & Q \end{bmatrix}, \quad \hat{L}_k \rightarrow \begin{bmatrix} I & 0 \\ 0 & \hat{L} \end{bmatrix},
\]

and

\[
\prod_{j=1}^{k} (T_1 - \omega_j) = S \hat{Q}_k \hat{L}_k \Lambda_k L^{-1},
\]

where \( \hat{L}_k \Lambda_k L^{-1} \) is a lower triangular matrix. If not all the diagonal elements of \( \hat{L}_k \Lambda_k L^{-1} \) are positive, left-multiply it by a diagonal matrix \( D_k \) whose norm of every diagonal element is equal to 1, so that every diagonal element of matrix \( D_k \hat{L}_k \Lambda_k L^{-1} \) is positive. Then

\[
\prod_{j=1}^{k} (T_1 - \omega_j) = (S \hat{Q}_k D_k^{-1})(D_k \hat{L}_k \Lambda_k L^{-1})
\]

is a QL decomposition, and

\[
S \hat{Q}_k D_k^{-1} = \tilde{Q}_k D, \quad \tilde{Q}_k = S \hat{Q}_k D_k^{-1} D,
\]

\[
\beta_2^{(k+1)} = e_2^T T_{k+1} e_3 = e_2^T \tilde{Q}_k T_{k+1} \tilde{Q}_k e_3,
\]

\[
\tilde{Q}_k e_3 = S \hat{Q}_k D_k^{-1} D e_3 \rightarrow S \begin{pmatrix} 0 \\ 0 \\ \pm 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \rightarrow \pm s_{i_3},
\]

Similarly,

\[
e_2^T \tilde{Q}_k \rightarrow \pm s_{i_3}^T.
\]
We conclude that

\[ \beta_2^{(k+1)} \rightarrow \pm s_{i_2}^T T_1 s_{i_3} = 0. \]

By Theorem 2, we know that in almost every case \( \beta_2 \) tends to zero.

In Reference [2], the author gave the following result.

**Lemma 4.** In the QL algorithm with shift \( \{ \varepsilon_k \} \), if \( \beta_1^{(k)} \rightarrow 0 \), \( \alpha_1^{(k)} - \varepsilon_k \rightarrow 0 \), and \( \beta_2^{(k)} \rightarrow 0 \), then

\[ \alpha_1^{(k)} \rightarrow \lambda_{i_1} \quad \text{and} \quad \alpha_2^{(k)} \rightarrow \lambda_{i_2} \]

where \( \lambda_{i_1} \) and \( \lambda_{i_2} \) are two eigenvalues of \( T_1 \). If \( \beta_1^{(1)} \neq 0 \), then they are different.

According to the above results, we have the following corollary for MRQI.

**Corollary 2.** If \( \tilde{x}_1 \) is not an eigenvector of \( \tilde{A} \) and the eigenvalues of \( \tilde{A} \) satisfy

\[ |\lambda_{i_{1-1}} - \lambda_{i_1}| \neq |\lambda_{i_{1+1}} - \lambda_{i_1}|, \]

then \( b_k \) in Theorem 1 satisfy

\[ b_{k+1} = \tilde{M}_k b_k^3 c_k^2 \]

and \( c_k \rightarrow 0 \), where \( |\tilde{M}| \leq \text{const.} \)

**Proof.** By the assumptions and Theorem 2, we have \( \beta_2 \rightarrow 0 \); hence \( c_k \rightarrow 0 \). Since \( \tilde{x}_1 \) is not an eigenvector of \( \tilde{A} \), we have \( \beta_1^{(1)} \neq 0 \) and \( \alpha_2^{(k)} \rightarrow \lambda_{i_2} \), where \( \lambda_{i_2} \neq \lambda_{i_1} \). Therefore

\[ \tilde{a}_k - \omega_k \rightarrow \lambda_{i_2} - \lambda_{i_1} \neq 0. \]

It follows that

\[ b_{k+1} - M_k \frac{b_k^3 c_k^2}{|\tilde{a}_k - \omega_k|} = \tilde{M}_k b_k^3 c_k^2, \]

where \( |\tilde{M}_k| \leq \text{const.} \). 

This corollary indicates that the convergence rate of Algorithm III is at least cubic in most cases.

If the eigenvalue \( \lambda_{i1} \) is the smallest or the largest eigenvalue of \( \tilde{A} \), then (15) is trivially satisfied. For some matrices whose eigenvalues do not satisfy (15), whether the convergence rate is still cubic is an unsolved problem.

Now we write out the proof of the basic theorem.

1: By Theorem 1, for the sequence \( \{ (\tilde{\rho}_k, \tilde{x}_k) \} \), we have

\[
h_k = \| r_k \| \to 0 \quad \text{and} \quad h_{k+1} = M_k \frac{b_k^3 c_k^2}{|a_k - \omega_k|}.
\]

Equivalently,

\[
(Bd_k, d_k) \to 0,
\]

\[
(Bd_{k+1}, d_{k+1}) = M_k \left( \frac{(Bd_k, d_k)^3}{|a_k - \omega_k|^2} \right),
\]

where \( |M_k| \leq \text{const.} \)

2: \( \lim_{k \to \infty} \tilde{\rho}_k = \lim_{k \to \infty} \rho_k = \lambda_{i1} \), where \( \lambda_{i1} \) is an eigenvalue of \( \tilde{A} \), i.e., \( \lambda_{i1} \) is an eigenvalue of \( (A, B) \). If \( \lambda_{j(k)} \) is the eigenvalue of \( \tilde{A} \) nearest to \( \tilde{\rho}_k \), from

\[
\tilde{A} \tilde{x} - \tilde{\rho}_k \tilde{x}_k = r_k
\]

we have

\[
|\lambda_{j(k)} - \rho_k| \leq \min \left( b_k \frac{b_k^2}{\sigma_k} \right)
\]

\[
= \min \left( (Bd_k, d_k)^{1/2}, \frac{(Bd_k, d_k)}{\sigma_k} \right),
\]

where \( \sigma_k = \min_{\lambda_i \neq \lambda_{j(k)}} |\lambda_i - \rho_k| \). When \( k \) is large enough, \( \lambda_{j(k)} = \lambda_{i1} \).

3: \( \lim_{k \to \infty} \tilde{x}_k = \tilde{z}_{i1} \), where \( \tilde{z}_{i1} \) is the corresponding eigenvector of \( \tilde{A} \).

Since \( \tilde{x}_k = B^{1/2} x_k \), we have

\[
\lim_{k \to \infty} x_k = B^{-1/2} \tilde{z}_{i1} = z_{i1}.
\]
From

\[ \tilde{A} \tilde{z}_{i1} = \lambda_{i1} \tilde{z}_{i1}, \]

we deduce

\[ B^{-1/2} \tilde{A} B^{-1/2} \tilde{z}_{i1} = \lambda_{i1} B^{1/2} \tilde{z}_{i1}, \]

i.e.,

\[ A \tilde{z}_{i1} = \lambda_{i1} B \tilde{z}_{i1}. \]

Therefore \( \tilde{z}_{i1} \) is the eigenvector of \((A, B)\) corresponding to \( \lambda_{i1} \).

Since \( (\tilde{z}_{i1}, \tilde{z}_{i1}) = 1, \) i.e., \( (B \tilde{z}_{i1}, \tilde{z}_{i1}) = 1, \) denoting \( (\tilde{z}_{i1}, \tilde{x}_k) = \cos \theta_k \) \( = (B x_k, z_{i1}), \) by Theorem 1 we have

\[ \| \tilde{x}_k - \tilde{z}_{i1} \| \leq \sqrt{\frac{2}{1 + |\cos \theta_k|}} \frac{b_k}{\eta_k}, \quad (16) \]

where \( \eta_k = \min_{\lambda} \| \rho_k - \lambda \|. \) Rewrite (16) as

\[ \| B^{1/2} x_k - B^{1/2} z_{i1} \| \leq \sqrt{\frac{2}{1 + |\cos \theta_k|}} \frac{b_k}{\eta_k}. \]

Since

\[ \| B^{1/2} x_k - B^{1/2} z_{i1} \|^2 = (B(x_k - z_{i1}), x_k - z_{i1}) \]

\[ \geq \mu_1 \| x_k - z_{i1} \|^2, \]

we have

\[ \| x_k - z_{i1} \| \leq \sqrt{\frac{2}{(1 + |\cos \theta_k|) \mu_1}} \frac{(B d_k, d_k)^{1/2}}{\eta_k}. \]
Finally, by Theorem 2, we obtain

\[ |\sin \theta_k| \leq \left| \frac{r_k}{\eta_k} \right| = \frac{b_k}{\eta_k} = \frac{(Bd_k, d_k)^{1/2}}{\eta_k}. \]

We have finished the proof of the basic theorem.

*The author would like to thank Mr. Zhang Guodong for his help in a lot of numerical tests.*

REFERENCES


Received 14 July 1988; final manuscript accepted 27 February 1989