A Block Jacobi-like Method for Eigenvalue Decomposition of a Real Normal Matrix Using Real Arithmetic*

B. B. Zhou and R. P. Brent†
Computer Sciences Laboratory
The Australian National University
Canberra, ACT 0200, Australia

Abstract

In this paper we discuss how to design efficient Jacobi-like algorithms for eigenvalue decomposition of a real normal matrix. We introduce a block Jacobi-like method. This method uses only real arithmetic and orthogonal similarity transformations and achieves ultimate quadratic convergence. A theoretical analysis is conducted and some experimental results are presented.

Keywords Eigenvalue decomposition, normal matrix, Jacobi algorithm and QR algorithm, parallel computing.

1 Introduction

A real matrix A is said to be normal if it satisfies the equation

$$AA^T = A^T A$$

where A^T is the transpose of matrix A. A normal matrix can be reduced to a diagonal form using unitary similarity transformations

$$QAQ^H = D$$

where D is diagonal, Q is unitary and Q^H is the conjugate transpose of Q. The standard sequential method for eigenvalue decomposition of this kind of matrix is the QR algorithm. However, when massively parallel computation is considered, the parallel version of the QR-based algorithms for solving unsymmetric eigenvalue problems may not be very efficient because the algorithms are sequential in nature and not scalable.

One alternative to the QR method is a Jacobi method. Jacobi-based algorithms have recently attracted a lot of attention as they have a higher degree of potential parallelism. The Jacobi method, though originally designed for symmetric eigenvalue problems, can be extended to solve eigenvalue problems for unsymmetric normal matrices [3, 9]. A problem is that we have to use complex arithmetic even for real-valued normal matrices. Complex operations are expensive and should be avoided if possible. A quaternion-Jacobi method was recently introduced [5]. In this method a 4×4 symmetric matrix can be reduced to a 2×2 block diagonal form using one

 rpb163r
†Current address: Oxford University Computing Laboratory, Parks Road, Oxford OX1 3QD, UK.

orthogonal similarity transformation. This method can also be extended to compute eigenvalues of a general normal matrix. However, problems are that the original matrix has to be divided into a sum of a symmetric matrix and a skew-symmetric matrix, and that the algorithm cannot be used to solve the eigenvalue problem of near-normal matrices. Another parallel Jacobi-like algorithm, named the RTZ (Real Two-Zero) algorithm, was also proposed recently [6]. This method uses real arithmetic and orthogonal similarity transformations. It is claimed that quadratic convergence can be obtained when computing eigenvalues of a real near-normal matrix with real distinct eigenvalues. However, a serious problem with this algorithm is that the process may fail to converge if the matrix has complex eigenvalues.

In this paper we discuss a block Jacobi-like method for computing the eigenvalue decomposition of a real normal matrix. The method uses only real arithmetic and orthogonal similarity transformations. The theoretical analysis and experimental results show that ultimate quadratic convergence can be achieved for general real normal matrices with distinct eigenvalues.

Our aim is to design scalable algorithms which can efficiently be implemented on parallel machines. The key to parallelize a Jacobi (or Jacobi-like) algorithm is to design a parallel Jacobi ordering. Ordering schemes may affect the overall performance. However, most existing parallel Jacobi orderings are proved to be equivalent to the well-known sequential cyclic ordering [4, 12]. To simplify our discussion, therefore, we only use the cyclic Jacobi ordering.

Real normal matrices are generalisations of real symmetric matrices. A real symmetric matrix is normal, but a real normal matrix is not necessarily symmetric. We shall focus our attention on the unsymmetric case although the method to be described applies to both cases.

Since the RTZ algorithm uses a similar idea to ours, an analysis of the RTZ algorithm is presented in Section 2. Our block method is described in Section 3. In that section we give a theoretical analysis which indicates how to choose orthogonal similarity transformations so that ultimate quadratic convergence can be obtained. Some experimental results are presented in Section 4, and conclusions are in Section 5.

2 An Analysis of the RTZ Algorithm

The experimental results presented in [6] show that the convergence rate of the RTZ algorithm is not quadratic if all eigenvalues of a given matrix are complex. However, the paper [6] did not give any theoretical explanations. We show here that, when applying the RTZ algorithm, we can obtain at best a linear convergence rate if the matrix has complex eigenvalues.

In the following a 4×4 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}, \tag{1}$$

or its block form (with each block being of size 2×2)

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

is used to show how the convergence rate may be affected when applying the RTZ algorithm.

The basic idea for computing the eigenvalue decomposition of a 4×4 matrix using the RTZ algorithm is as follows: The first leading diagonal element a_{11} in A_{11} is chosen together with the first row in A_{12}, the first column in A_{21} and the whole of A_{22} to form a 3×3 matrix, that is,

$$A_1 = \begin{pmatrix} a_{11} & a_{13} & a_{14} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{pmatrix}.$$

It is known that any real 3×3 matrix has at least one real eigenvalue. An eigenvector associated with a real eigenvalue of the above matrix can be obtained and used to generate a Householder matrix which is then applied to update the matrix A so that two off-diagonal elements a_{31} and a_{41} are annihilated. After that the second leading diagonal element a_{22} in A_{11} is chosen (together with A_{22}, the second row in A_{12} and the second column in A_{21}) and another 3×3 matrix

$$A_2 = \begin{pmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{pmatrix}$$

is formed. A Householder matrix is generated from an eigenvector (associated with a real eigenvalue) of this 3×3 matrix so that two other lower triangular off-diagonal elements a_{32} and a_{42} are eliminated through a similarity transformation. However, this destroys the zeros introduced previously. The two leading diagonal elements are thus chosen alternately and the process continues until all the elements in A_{21} are small enough to be considered as zero.

In the following discussions ϵ denotes a small positive number close to zero. If certain elements of a matrix are written as ϵ, we mean that the values of these elements are small and of the same order of ϵ, but they are not necessarily the same.

If the above RTZ procedure converges, all the elements in A_{21} and A_{12} become ϵ after a few iterations and each 3×3 matrix will thus have the same form as

$$B = \begin{pmatrix} b_{11} & \epsilon & \epsilon \\ \epsilon & b_{22} & b_{23} \\ \epsilon & b_{32} & b_{33} \end{pmatrix}. \tag{2}$$

Assume that B is close to a normal block-diagonal matrix and written as

$$B = D + \epsilon F \tag{3}$$

where the elements in F satisfy $|f_{ij}| < 1$ and D has a form

$$D = \begin{pmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & d_{23} \\ 0 & d_{32} & d_{33} \end{pmatrix}.$$

If the eigenvalues of D are well separated, it is easy to prove using perturbation theory [11] that there exists an eigenvalue of B which satisfies the equation

$$|d_{11} - \lambda| = O(\epsilon) \tag{4}$$

and that an eigenvector v associated with λ will also satisfy

$$\|v - v'\| = O(\epsilon) \tag{5}$$

for v' an eigenvector associated with d_{11}, a real eigenvalue of D. Therefore, we have the following lemma.

Lemma 1 *Let a 3×3 matrix B have a form as in (2) and be written as in (3). There exists a real eigenvalue λ of B , such that*

$$|d_{11} - \lambda| = O(\epsilon).$$

An eigenvector v associated with λ has a form

$$v = (v_1 \;\; \epsilon \;\; \epsilon)^T . \tag{6}$$

Proof. The first part of the lemma is directly obtained from (4). Let $v' = (v_1', v_2', v_3')^T$ be an eigenvector associated with d_{11}, a real eigenvalue of D. Then

$$(D - d_{11}I)v' = 0,$$

or

$$\left\{ \begin{array}{ccccccc} (d_{11} - d_{11})v_1' & + & 0v_2' & + & 0v_3' & = & 0 \\ 0v_1' & + & (d_{22} - d_{11})v_2' & + & d_{23}v_3' & = & 0 \\ 0v_1' & + & d_{32}v_2' & + & (d_{33} - d_{11})v_3' & = & 0 \end{array} \right. .$$

Since the eigenvalues of D are well separated, the determinant of the coefficient matrix from the second and the third equations will not be equal to zero. Thus v_2' and v_3' must be zero. We then have

$$v' = (v_1', 0, 0)^T .$$

From (5), therefore, v will have the form as in (6). □

Since the RTZ algorithm uses Householder transformations, it can be seen in the following that it is crucial for eigenvectors generated at each step to have the form as in (6) in order to obtain ultimate quadratic convergence.

We now show how the lower triangular off-diagonal norm is affected when the generated Householder matrix is applied to update certain rows and columns of the matrix.

Assume that a Householder vector v is chosen based on vector $b = (b_1 \; b_2 \; b_3)^T$, that is, $v = (v_1 \; b_2 \; b_3)^T$ for $v_1 = b_1 + sign(b_1)\|b\|$ and $\|b\| = \sqrt{b_1^2 + b_2^2 + b_3^2}$. The Householder matrix is then obtained as $H = I - 2vv^T/v^Tv$.

Suppose that the values of b_2 and b_3 are of order ϵ, but b_1 is "large". We have

$$\begin{aligned} \frac{1}{\|b\|} &= \frac{1}{sign(b_1)b_1\sqrt{1 + (b_2^2 + b_3^2)/b_1^2}} \\ &\approx \frac{1}{sign(b_1)b_1\sqrt{1 + \epsilon^2}} \\ &\approx \frac{1}{sign(b_1)b_1}(1 - \frac{1}{2}\epsilon^2). \end{aligned}$$

It is easy to verify that the Householder matrix has the structure

$$H \approx \begin{pmatrix} -1 + \epsilon^2 & \epsilon & \epsilon \\ \epsilon & 1 - \epsilon^2 & \epsilon^2 \\ \epsilon & \epsilon^2 & 1 - \epsilon^2 \end{pmatrix} . \tag{7}$$

When using this matrix to update (left multiply) a column vector $(x_1 \; 0 \; 0)^T$, it is easy to see that, if x_1 is of order ϵ, the zero elements will become $O(\epsilon^2)$. If x_1 is "large", however, after the updating, the zero elements will become $O(\epsilon)$, rather than $O(\epsilon^2)$.

Sweep	Lower Block Triangular Norm	
	4 real eigenvalues	2 real and 2 complex eigenvalues
0	0.2620215927345935	0.6198994950456397
1	3.400727964203065e-02	3.649800212444496e-02
2	1.112102911733364e-06	6.950946615948856e-06
3	5.364457308382537e-25	1.705643339710924e-21

Table 1: The 4×4 matrices have at most two complex eigenvalues.

Suppose that all the elements in the off-diagonal blocks become small after a few sweeps and that the last two elements in the first column have been eliminated by the immediately preceding iteration in which a_{11} and the corresponding 3×3 matrix are chosen. The updated matrix then has a form

$$A = \begin{pmatrix} a_{11} & a_{12} & \epsilon & \epsilon \\ a_{21} & a_{22} & \epsilon & \epsilon \\ 0 & \epsilon & a_{33} & a_{34} \\ 0 & \epsilon & a_{43} & a_{44} \end{pmatrix}.$$

Now we choose a_{22} and the corresponding 3×3 matrix as

$$\begin{pmatrix} a_{22} & \epsilon & \epsilon \\ \epsilon & a_{33} & a_{34} \\ \epsilon & a_{43} & a_{44} \end{pmatrix}$$

and apply the RTZ procedure again to annihilate the last two elements in the second column of matrix A. Assume that this 3×3 matrix has the same properties as the matrix described in Lemma 1. The eigenvector will be in the same form as in (6). Therefore, the generated Householder matrix is of the form like (7).

The last three elements in the first column of matrix A will be affected in the updating procedure using this Householder matrix. If A_{11} has two real eigenvalues, we may reasonably assume that a_{21} and a_{12} are of order ϵ. (If not, an orthogonal similarity transformation can be applied to annihilate a_{21} without using complex arithmetic as A_{11} has two real eigenvalues.) When this Householder matrix is applied, the values of the zero elements in the first column will be of order ϵ^2. Quadratic convergence may then be achieved. If A_{11} (as well as A_{22}) has two complex eigenvalues, however, the value of a_{21} may no longer be small. The values of the zero elements in the first column will be increased to $O(\epsilon)$ as we discussed above. The convergence rate is thus only linear at best.

To verify our analysis we implemented the RTZ algorithm. Some experimental results are given in Table 1 and Table 2. In the experiment the RTZ algorithm is used to decompose three different kinds of 4×4 normal matrices. The first matrix has four real eigenvalues and the second one has two real and two complex eigenvalues. The RTZ algorithm works efficiently for these two matrices. As indicated in Table 1, a better than quadratic convergence rate is achieved.

The problem occurs when all four eigenvalues of a normal matrix are complex. The lower block triangular norm converges to zero very slowly when the RTZ algorithm is used. Let the two pairs of complex eigenvalues have the form as $\lambda \pm i\mu$. In our experiment λ is set to be a positive random number smaller than one, while specific values are assigned to μ. When μ is set to be very small, the eigenvalues will be close to real. Our experimental results show that the convergence is very slow when μ is great. As shown in Table 2, for example, the matrix does

Sweep	Lower Block Triangular Norm		
	$\mu = 0.5$	$\mu = 0.05$	$\mu = 0.005$
0	0.6711408155892714	0.2333023567461740	0.2246552719474016
1	0.2716123691993152	1.199223022243009e-02	2.135320869285008e-02
2	9.755104834822810e-02	5.917326295608719e-05	1.059463164110422e-06
3	3.263905677008454e-02	2.918965666220263e-07	5.251887781007312e-11
4	1.082900741978114e-02	1.439900400735424e-09	2.603424651608092e-15
5	3.589505732485564e-03	7.102903566815823e-12	1.832988980675934e-17
6	1.189696302360603e-03	3.503800613515733e-14	
7	3.943053592985449e-04	1.728394398251085e-16	
8	1.306858932316573e-04	1.213931761292142e-17	
27	1.008676811946900e-13		
28	3.343089487481607e-14		
29	1.108010728487409e-14		
30	3.672315027649536e-15		

Table 2: The matrix has four complex eigenvalues.

not converge to block triangular form even after thirty sweeps when $\mu = 0.5$. It converges faster only if μ becomes smaller. This evidence confirms the correctness of our analysis.

In the above we only considered a very simple problem, that is, the 4×4 case. When the matrix size is much larger, the slow local convergence can significantly affect the global convergence. In certain cases the process may not even converge.

3 Block Jacobi-Like Method

To simplify our discussion, we assume that the matrix size is even. The basic idea of our method is described as follows: A real normal matrix is first divided into blocks. To avoid using complex arithmetic the size of each block is chosen to be 2×2 so that each pair of conjugate complex eigenvalues can be grouped into the same block. A sequence of orthogonal similarity transformations is then applied to annihilate the off-diagonal blocks in the lower triangle and this process continues until all the lower triangular off-diagonal blocks are considered as zero. The basic structure of the method is depicted in Fig. 1.

In Fig. 1 NITN denotes the number of lower triangular off-diagonal blocks, or number of iterations in a sweep, which is equal to $n(n-2)/8$ for n the size of the problem. The 2×2 block A_{ij} has the form

$$A_{ij} = \begin{pmatrix} a_{2i-1,2j-1} & a_{2i-1,ij} \\ a_{2i,2j-1} & a_{2i,2j} \end{pmatrix}$$

and therefore B is of size 4×4. A counter NZCONT is used to check how many off-diagonal blocks in the lower triangle are zero. If this is equal to NITN, the process stops. The cyclic ordering is used for sequential computation. If we consider each block as a single element, The action of the designed algorithms is just the same as that of a nonblocked Jacobi algorithm. Therefore, any existing efficient parallel orderings, for example, those in [1, 2, 4, 10, 12, 13], can be adopted to form an efficient parallel algorithm for solving the problem.

The structure in Fig. 1 is similar to any standard block Jacobi schemes. The crux is how to choose the orthogonal transformation matrices Q during the computation such that the quadratic

```
NITN = n(n-2)/8
NZCONT = 0
REPEAT UNTIL NZCONT = NITN
    DO i = 1, n/2
        DO j = i+1, n/2
            B = ( A_ii  A_ij )
                ( A_ji  A_jj )
            IF (||A_ji|| not zero) THEN
                1. Find an orthogonal matrix Q such that B is reduced to a block
                triangular form through a similarity transformation.
                2. Update the corresponding rows and columns of A using Q.
            ELSE
                NZCONT = NZCONT + 1
            END IF
        END DO
    END DO
    IF (NZCONT ≠ NITN) NZCONT = 0
END REPEAT
```

Figure 1: The basic structure of the block Jacobi-like method.

convergence can be achieved. The remaining paper will deal with this very important issue. In the following $\|A\|$ denotes the Frobenius norm of A.

Lemma 2 *Let a normal matrix A be divided into blocks B_{ij}. If A is nearly block triangular, that is,*

$$\sum_{k>i} \|B_{ki}\|^2 = O(\epsilon), \quad \textit{for all } i,$$

then A is nearly block diagonal.

Proof. Let $C = AA^T$ and $C' = A^T A$. We have $C_{ii} = \sum_k B_{ik} B_{ik}^T$ and $C'_{ii} = \sum_k B_{ki}^T B_{ki}$. Since the matrix A is normal,

$$C_{ii} - C'_{ii} = \sum_k B_{ik} B_{ik}^T - \sum_k B_{ki}^T B_{ki} = 0,$$

or

$$B_{ii} B_{ii}^T - B_{ii}^T B_{ii} = \sum_{k \neq i} (B_{ki}^T B_{ki} - B_{ik} B_{ik}^T). \tag{8}$$

Taking the sum of the diagonal elements on both sides of the above equation, we have the following relation for all i:

$$0 = \sum_{k \neq i} (\|B_{ki}\|^2 - \|B_{ik}\|^2). \tag{9}$$

When $i = 1$, the above implies that

$$\sum_{k>1} \|B_{1k}\|^2 = \sum_{k>1} \|B_{k1}\|^2 = O(\epsilon).$$

When $i = 2$, we have

$$\sum_{k>2} \|B_{2k}\|^2 + \|B_{21}\|^2 = \sum_{k>2} \|B_{k2}\|^2 + \|B_{12}\|^2,$$

or

$$\begin{aligned} \sum_{k>2} \|B_{2k}\|^2 &= \sum_{k>2} \|B_{k2}\|^2 + \|B_{12}\|^2 - \|B_{21}\|^2 \\ &\leq \sum_{k>2} \|B_{k2}\|^2 + \sum_{k>1} \|B_{1k}\|^2 \\ &= O(\epsilon). \end{aligned}$$

Therefore, by induction we can easily prove

$$\sum_{k>i} \|B_{ik}\|^2 = O(\epsilon).$$

□

From Lemma 2 we see that, if we can reduce a real normal matrix to a block upper triangular form using orthogonal similarity transformations, the matrix will become block diagonal.

We show in the following three lemmas that, if the orthogonal transformation matrices are chosen properly and if the process converges for a normal matrix with distinct eigenvalues, the convergence rate will ultimately be quadratic.

Lemma 3 *If a normal matrix A is divided into blocks B_{ij}, then the main diagonal blocks have the following property:*

$$\|B_{ii}B_{ii}^T - B_{ii}^T B_{ii}\| \leq \sum_{k \neq i} (\|B_{ki}\|^2 + \|B_{ik}\|^2). \tag{10}$$

Proof. From equation (8) we have

$$\begin{aligned} \|B_{ii}B_{ii}^T - B_{ii}^T B_{ii}\| &\leq \sum_{k \neq i} \|B_{ki}^T B_{ki} - B_{ik} B_{ik}^T\| \\ &\leq \sum_{k \neq i} (\|B_{ki}^T B_{ki}\| + \|B_{ik} B_{ik}^T\|). \end{aligned}$$

Since $\|B_{ki}^T B_{ki}\| \leq \|B_{ki}^T\| \|B_{ki}\| = \|B_{ki}\|^2$, we thus obtain

$$\|B_{ii}B_{ii}^T - B_{ii}^T B_{ii}\| \leq \sum_{k \neq i} (\|B_{ki}\|^2 + \|B_{ik}\|^2).$$

□

It can be seen from this lemma that each block on the main diagonal will be very close to a normal matrix if the norm of each off-diagonal block is small, that is,

$$\|B_{ii}B_{ii}^T - B_{ii}^T B_{ii}\| = O(\epsilon^2)$$

if $\max(\|B_{ij}\|) = \epsilon$ for $i \neq j$.

The next lemma shows that, when the lower off-diagonal block of a 2×2 block matrix is annihilated through an orthogonal similarity transformation, the norm of its upper off-diagonal block will also be decreased if this block matrix is close to normal.

Lemma 4 *Assume that a* 2×2 *non-normal block matrix* B*, with the size of each block being* 2×2*, is close to a normal matrix and has the property*

$$\|BB^T - B^T B\| = O(\eta) \tag{11}$$

where η *is a small positive number, and that* B *has four nonzero eigenvalues and can be reduced to a block triangular form through an orthogonal similarity transformation, that is,*

$$Q^T B Q = D \tag{12}$$

where Q *is a real orthogonal matrix and*

$$D = \begin{pmatrix} D_{11} & D_{12} \\ 0 & D_{22} \end{pmatrix} \tag{13}$$

for each block D_{ij} *being of size* 2×2*.*

Let the eigenvalues of D_{11} *and* D_{22} *be* λ_{1i} *and* λ_{2i} *for* $i = 1, 2$*, respectively. (These are also the eigenvalues of* B*.) If the eigenvalues of* D_{11} *are separated from those of* D_{22} *and they satisfy the following two inequalities:*

$$c_1 \le |\lambda_{ij}| \le c_2 \tag{14}$$

and

$$\left| 1 - \frac{\lambda_{1i}}{\lambda_{2j}} \right| > c_3 \tag{15}$$

for c_1*,* c_2 *and* c_3 *constants greater than zero, then we have*

$$\|D_{12}\| = O(\eta).$$

Proof. Since $B = QDQ^T$ and $B^T = QD^TQ^T$, we have

$$\|BB^T - B^T B\| = \|Q(DD^T - D^T D)Q^T\|.$$

Thus

$$\|DD^T - D^T D\| = O(\eta). \tag{16}$$

It is easy to see that the second block of the first column in $DD^T - D^T D$ is $D_{22}D_{12}^T - D_{12}^T D_{11}$. We know that $\|D_{12}\| \ne 0$. Otherwise, B is a normal matrix which is contrary to our assumption. Since $\|D_{12}\| \ne 0$ and the eigenvalues are nonzero and distinct, $\|D_{22}D_{12}^T - D_{12}^T D_{11}\|$ may be nonzero. However, it is easy to see from (16) that it should be of order η at most, that is,

$$\|D_{22}D_{12}^T - D_{12}^T D_{11}\| = O(\eta). \tag{17}$$

Since B has four nonzero eigenvalues, both D_{11} and D_{22} have full rank. The following inequality holds:

$$\|D_{12}^T - D_{22}^{-1} D_{12}^T D_{11}\| \le \|D_{22}^{-1}\| \|D_{22}D_{12}^T - D_{12}^T D_{11}\|.$$

According to one of our assumptions, that is, the eigenvalues of matrix B are bounded by two positive constants c_1 and c_2, $\|D_{22}^{-1}\|$ should also be bounded by two positive constants, that is,

$$\frac{1}{c_2} \le \|D_{22}^{-1}\| \le \frac{1}{c_1}.$$

We thus have

$$\|D_{12}^T - D_{22}^{-1} D_{12}^T D_{11}\| = O(\eta). \tag{18}$$

Let

$$D_{11} = Q_1 R_1 Q_1^H$$

and

$$D_{22}^{-1} = Q_2 R_2 Q_2^H$$

be the eigenvalue decomposition of D_{11} and D_{22}^{-1} for Q_1 and Q_2 unitary and R_1 and R_2 upper triangular, and define

$$Q_2^H D_{12}^T Q_1 = E.$$

we then have

$$\begin{aligned} \|E - R_2 E R_1\| &= \|Q_2(E - R_2 E R_1) Q_1^H\| \\ &= \|D_{12}^T - D_{22}^{-1} D_{12}^T D_{11}\| \\ &= O(\eta). \end{aligned} \tag{19}$$

Let

$$E = \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix},$$

$$R_1 = \begin{pmatrix} r_{11}^{(1)} & r_{12}^{(1)} \\ 0 & r_{22}^{(1)} \end{pmatrix}$$

and

$$R_2 = \begin{pmatrix} r_{11}^{(2)} & r_{12}^{(2)} \\ 0 & r_{22}^{(2)} \end{pmatrix}.$$

Expanding $G = E - R_2 E R_1$, we obtain

$$G = \begin{pmatrix} s_{11}e_{11} - r_{12}^{(2)} r_{11}^{(1)} e_{21} & s_{12}e_{12} - r_{11}^{(2)} r_{12}^{(1)} e_{11} - r_{12}^{(2)} r_{12}^{(1)} e_{21} - r_{12}^{(2)} r_{22}^{(1)} e_{22} \\ s_{21}e_{21} & s_{22}e_{22} - r_{22}^{(2)} r_{12}^{(1)} e_{21} \end{pmatrix}$$

where $s_{ji} = 1 - r_{jj}^{(2)} r_{ii}^{(1)}$.

Since $r_{ii}^{(1)}$ is an eigenvalue of D_{11} and $1/r_{jj}^{(2)}$ is an eigenvalue of D_{22}, then

$$\begin{aligned} |s_{ji}| &= \left|1 - r_{jj}^{(2)} r_{ii}^{(1)}\right| \\ &= \left|1 - \frac{\lambda_{1i}}{\lambda_{2j}}\right| \\ &> c_3. \end{aligned}$$

According to equation (19) all the elements in G must be of order η, that is,

$$\begin{aligned} |s_{21}e_{21}| &= O(\eta); \\ \left|s_{11}e_{11} - r_{12}^{(2)} r_{11}^{(1)} e_{21}\right| &= O(\eta); \\ \left|s_{22}e_{22} - r_{22}^{(2)} r_{12}^{(1)} e_{21}\right| &= O(\eta); \\ \left|s_{12}e_{12} - r_{11}^{(2)} r_{12}^{(1)} e_{11} - r_{12}^{(2)} r_{12}^{(1)} e_{21} - r_{12}^{(2)} r_{22}^{(1)} e_{22}\right| &= O(\eta). \end{aligned}$$

It is then easy to verify from the above four equations that all the elements in E must be of order η. From the first equation we can obtain $|e_{21}| = O(\eta)$ because s_{21} is greater than a constant c_3. After we know that e_{21} is of order η, we can easily verify that e_{11} (and e_{22}) is of order η from the second equation (and the third equation). Because e_{11}, e_{21} and e_{22} are all of order η, we obtain $|e_{12}| = O(\eta)$ from the fourth equation. Therefore, we have

$$\begin{aligned} \|D_{12}\| &= \|Q_2 E Q_1^H\| \\ &= \|E\| \\ &= O(\eta). \end{aligned}$$

□

Assume that a normal matrix is divided into blocks of size 2×2 and that the norm of each off-diagonal blocks is of order ϵ. Further we assume that each B_{ij} in (10) is a 2×2 block (i.e., 4×4) submatrix. Thus the norm of each off-diagonal submatrix B_{ij} for $i \neq j$ must also have an order of ϵ. From Lemma 3 we obtain

$$\|B_{ii}B_{ii}^T - B_{ii}^T B_{ii}\| = O(\epsilon^2)$$

where B_{ii} is of the same form as B in Fig. 1.

After an orthogonal similarity transformation which reduces B_{ii} to a block upper triangular matrix, the norm of the upper off-diagonal block in B_{ii} should have the same order as $\|B_{ii}B_{ii}^T - B_{ii}^T B_{ii}\|$ according to the above lemma. Therefore, it will be of order ϵ^2.

Since the off-diagonal norm of B_{ii} is reduced from $O(\epsilon)$ to $O(\epsilon^2)$ and the norms of other B_{ij}s are not affected during the updating procedure, we thus obtain a steady decrease in off-diagonal norm during the computation. In the following we show that, if the orthogonal transformation matrices are chosen properly, ultimate quadratic convergence can be achieved.

Explicitly write matrices B and Q in (12) as 2×2 block matrices, that is,

$$B = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

and

$$Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}$$

and assume that $\|A_{12}\| = O(\epsilon)$ and $\|A_{21}\| = O(\epsilon)$. Then we have the following lemma.

Lemma 5 *Assume that the eigenvalues of A_{11} and A_{22} are γ_{1i}, γ_{2i} for $i = 1, 2$, respectively, and the eigenvalues of D_{11} and D_{22} (which are also the eigenvalues of B) are defined as those in Lemma 4. If we can find an orthogonal matrix Q such that, after the similarity transformation $Q^T BQ = D$, the eigenvalues γ_{ij} of A_{ii} and the eigenvalues λ_{ij} of D_{ii} satisfy the following inequality*

$$\left| 1 - \frac{\gamma_{ik}}{\lambda_{jl}} \right| > c_3 \tag{20}$$

for $i \neq j$, and c_3 being a constant greater than zero, then the norms of both Q_{12} and Q_{21} in the generated orthogonal matrix must be of order ϵ, that is,

$$\|Q_{12}\| = O(\epsilon)$$

and

$$\|Q_{21}\| = O(\epsilon).$$

Proof. From (12) we have

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \begin{pmatrix} D_{11} & D_{12} \\ 0 & D_{22} \end{pmatrix}. \tag{21}$$

Thus

$$A_{21}Q_{11} + A_{22}Q_{21} = Q_{21}D_{11},$$

or

$$A_{21}Q_{11} = Q_{21}D_{11} - A_{22}Q_{21},$$

or

$$A_{21}Q_{11}D_{11}^{-1} = Q_{21} - A_{22}Q_{21}D_{11}^{-1} \tag{22}$$

and

$$A_{11}Q_{12} + A_{12}Q_{22} = Q_{11}D_{12} + Q_{12}D_{22},$$

or

$$(A_{12}Q_{22} - Q_{11}D_{12})D_{22}^{-1} = Q_{12} - A_{11}Q_{12}D_{22}^{-1}. \tag{23}$$

Since $\|A_{21}\| = O(\epsilon)$, $\|A_{12}\| = O(\epsilon)$ and $\|D_{12}\| = O(\epsilon^2)$ which can be obtained by combining the results of Lemma 2 and Lemma 4, the norms of the left-side of the equations in (22) and (23) must be of order ϵ. We thus have

$$\|Q_{21} - A_{22}Q_{21}D_{11}^{-1}\| = O(\epsilon) \tag{24}$$

and

$$\|Q_{12} - A_{11}Q_{12}D_{22}^{-1}\| = O(\epsilon). \tag{25}$$

The above two equations have the same form as in (18). Adopting the same technique as that used in Lemma 4, therefore, we can easily obtain $\|Q_{21}\| = O(\epsilon)$ and $\|Q_{12}\| = O(\epsilon)$. □

Using an orthogonal transformation matrix with $\|Q_{21}\| = O(\epsilon)$ and $\|Q_{12}\| = O(\epsilon)$ to update (premultiply) a vector $(0\ 0\ v_3\ v_4)^T$ (or $(v_1\ v_2\ v_3\ v_4)^T$ for v_1 and v_2 being of order ϵ^2), the zero elements in the vector will become $O(\epsilon^2)$ (or v_1 and v_2 will remain their original order) if both v_3 and v_4 are of order ϵ. This is one of the key factors for obtaining ultimate quadratic convergence.

Consider a 4×4 block matrix A for each block being of size 2×2 and that after k sweeps the norm of each off-diagonal block becomes the order of ϵ, that is,

$$A^{(k)} = \begin{pmatrix} A_{11} & \epsilon_{12} & \epsilon_{13} & \epsilon_{14} \\ \epsilon_{21} & A_{22} & \epsilon_{23} & \epsilon_{24} \\ \epsilon_{31} & \epsilon_{32} & A_{33} & \epsilon_{34} \\ \epsilon_{41} & \epsilon_{42} & \epsilon_{43} & A_{44} \end{pmatrix}$$

where the off-diagonal block A_{ij} is represented by ϵ_{ij} which means $\|A_{ij}\|$ is of order ϵ.

Assume that the cyclic ordering is adopted in the computation. At the $(k+1)^{th}$ sweep A_{21} will be annihilated first. According to the basic structure of the block Jacobi-lick method depicted in Fig. 1, a submatrix B is formed, that is,

$$B = \begin{pmatrix} A_{11} & \epsilon_{12} \\ \epsilon_{21} & A_{22} \end{pmatrix}.$$

Since all off-diagonal blocks A_{ij} are of order ϵ, from Lemma 3 we have

$$\|BB^T - B^TB\| = O(\epsilon^2).$$

After an orthogonal similarity transformation on B to annihilate A_{21}, the norm of A_{12} will become $O(\epsilon^2)$ according to Lemma 4.

Next A_{31} is annihilated. For the same reason we have $\|A_{13}\| = O(\epsilon^2)$ after the orthogonal similarity transformation. However, A_{21} will become nonzero again. If the norms of the off-diagonal blocks of the generated orthogonal matrix are of order ϵ, we can have $\|A_{21}\| = O(\epsilon^2)$ and $\|A_{12}\| = O(\epsilon^2)$ after the updating.

It can be seen from above discussion that $\|A_{ji}\|$ will become $O(\epsilon^2)$ after A_{ij} is annihilated for $i > j$ and during the successive transformations both $\|A_{ji}\|$ and $\|A_{ij}\|$ can remain to be of order ϵ^2 if at each step the orthogonal transformation matrix is chosen properly such as that described in Lemma 5. After the $(k+1)^{th}$ sweep the matrix will become

$$A^{(k+1)} = \begin{pmatrix} A_{11} & \epsilon_{12}^2 & \epsilon_{13}^2 & \epsilon_{14}^2 \\ \epsilon_{21}^2 & A_{22} & \epsilon_{23}^2 & \epsilon_{24}^2 \\ \epsilon_{31}^2 & \epsilon_{32}^2 & A_{33} & \epsilon_{34}^2 \\ \epsilon_{41}^2 & \epsilon_{42}^2 & \epsilon_{43}^2 & A_{44} \end{pmatrix}.$$

Therefore, the asymptotic quadratic convergence is obtained.

Now the problem is how to choose the orthogonal transformation matrix so that the norms of its off-diagonal blocks are of order ϵ. The inequality in (20) plays a crucial role in proving Lemma 5. When B is close to normal, the eigenvalues of A_{11} and A_{22}, the diagonal blocks of B, will be very close to the true eigenvalues of B. If an orthogonal transformation also involves (either implicitly or explicitly) a permutation on rows and columns between blocks, the eigenvalues of A_{11} can be very close to those of D_{22}. In that case the inequality in (20) will not be satisfied. We do not know if the generated orthogonal matrix still has the desired form when the inequality in (20) cannot be satisfied.

A natural way to alleviate this problem is to incorporate sorting with Jacobi ordering. With sorting diagonal blocks can be settled down more quickly. After the diagonal blocks being settled down, there will be no permutation on rows and columns between blocks and thus the inequality in (20) is guaranteed. Experiments (e.g. those presented in [7, 12]) have shown that to incorporate sorting with Jacobi ordering can always improve the performance for symmetric matrices. In the next section we give some experimental results which show that the ultimate quadratic convergence can be obtained for general normal matrices by incorporating a sorting procedure with the cyclic ordering when the QR algorithm with double implicit shift is used in the block Jacobi-like method for the local block submatrix reduction.

4 Experimental Results

In our experiments the basic algorithm for reducing B in Fig. 1 to a block triangular matrix at each step is the QR algorithm with double implicit shift and relevant subroutines in EISPACK [8] are used. (Note diagonal eigenvalues have sometimes to be swapped to ensure the norm of the lower off-diagonal block $\|A_{ji}\| = 0$.)

The stopping criteria used in our experiment is the same as that in EISPACK, that is, an off-diagonal element a_{ij} is considered as zero if $|a_{ij}| \leq (|a_{ii}| + |a_{jj}|) * \epsilon_{mach}$ for ϵ_{mach} the machine precision. The norm of a lower triangular off-diagonal block $\|A_{ji}\|$ is considered as zero if all the elements in it are considered as zero. The computation stops if all the lower triangular off-diagonal blocks are considered as zero.

The test matrices used in the experiment are generated by computing QDQ^T, where Q is

Sweep	Lower Block Triangular Norm		
	Matrix 1	Matrix 2	Matrix 3
0	1.216919467218377	1.942474075382786	2.496954343408182
1	0.8439242118214305	1.444070030260214	1.849869669358216
2	0.6715976208966092	0.9099848524494509	1.085259770981389
3	0.5694434773575591	0.5781929865448607	0.4617255166054944
4	0.4741511123441828	0.3448206046877936	0.1168896666930765
5	0.4158622482882098	0.2100276271297201	2.394464637809306e-02
6	0.3396277313047820	0.1199726773598173	1.947540032820058e-04
7	0.2952900301463802	9.980169372632480e-02	9.808465071791773e-09
8	0.2670511113785617	9.139422480738207e-02	1.177560482223073e-15
9	0.1976341619796229	6.635400959295680e-02	3.205392919010865e-16
21	4.058256009664995e-02	4.109644046234354e-07	
22	3.774124341065204e-02	4.109430061021018e-07	
23	2.982459710120542e-02	2.170532650662510e-09	
24	1.472832625703867e-02	2.183222832088858e-09	
25	1.366728263813406e-02	9.386963954187879e-11	
26	1.125831279655028e-02	3.310682565962637e-16	
27	4.490752073166923e-03		
28	4.213903185722646e-03		
29	3.239742746922371e-03		
30	2.037809355066387e-03		

Table 3: Sweeps and lower blocks triangular norms for 40×40 matrices obtained by using the block algorithm without sorting.

an orthogonal matrix and D is a block diagonal matrix. Each block in D is of size 2×2, that is,

$$D_{ii} = \begin{pmatrix} d_1 & d_2 \\ d_3 & d_4 \end{pmatrix}$$

where the four elements are positive random numbers smaller than one. When $d_2 = -d_3$ and $d_1 = d_4$, we have two complex eigenvalues $d_1 \pm i d_2$. Otherwise, we set $d_2 = d_3 = 0$ for two real eigenvalues. We use three different matrices in the experiments. The first matrix has distinct real eigenvalues, the second one has half of its eigenvalues real and the other half complex. and the third contains distinct complex eigenvalues.

In our first experiment we did not adopt any special sorting procedure. Table 3 gives the lower block triangular norms after each sweep for computing eigenvalues of the matrices of size 40×40. It can be seen from the table that our block algorithm perform well for the matrix which has only complex eigenvalues, but not so well for the matrices which contain real eigenvalues. The algorithm for block submatrix reduction is the QR with double implicit shift. It does not sort real eigenvalues. Though it does not order complex eigenvalues either, the lower 2×2 block will converge to the true eigenvalues rapidly when it is close to a pair of conjugate complex eigenvalues because of the double shift. Therefore, the rows and columns of the block submatrix will not be exchanged during the orthogonal similarity transformation. This can be considered as sorting although it does not sort the eigenvalues in a normal way. That may be the reason why the matrix with only complex eigenvalues can converge to the block diagonal form quadratically.

In our second experiment the same QR algorithm was used for block submatrix reduction. In

Sweep	Lower Block Triangular Norm		
	Matrix 1	Matrix 2	Matrix 3
0	1.216919467218377	1.942474075382786	2.496954343408182
1	0.6750246824159247	1.461835195736259	1.839162340923474
2	0.2199706803702323	0.9192624818874845	1.094029846622824
3	4.559208067945802e-02	0.4212417065916218	0.3932978785936760
4	2.000134475041065e-03	7.792134902151260e-02	7.739108958000541e-02
5	5.069396887509214e-06	4.672120103780485e-03	2.036895300210892e-03
6	1.708986452402760e-11	8.243231692930333e-06	8.707250423915960e-07
7	3.322903255620598e-16	1.606329703756845e-11	1.568197532065648e-13
8		1.764324938059070e-16	1.246328515799187e-16

Table 4: Sweeps and lower blocks triangular norms for 40×40 matrices obtained by using the block algorithm with sorting.

Matrix size	40	80	120	160	200
Matrix 1	7	8	9	9	10
Matrix 2	8	10	11	12	13
Matrix 3	8	10	11	12	13

Table 5: Sweeps taken for matrices of various sizes.

each step, however, we added a sorting procedure after the QR algorithm to sort real eigenvalues (if any) into nonincreasing order. When the real eigenvalues in each block submatrix are sorted in a nonincreasing order, the cyclic ordering can guarantee that all the real eigenvalues will be sorted in a nonincreasing order. The experimental results are presented in Table 4 for the same matrices as in our first experiment. It can be seen that quadratic convergence is obtained for all the matrices after the sorting procedure is incorporated. This is a firm evidence showing the importance of sorting for the block Jacobi-like method.

Some experimental results for matrices of different sizes are given in Table 5. We can see from this table that it takes a few more sweeps to converge for a matrix with complex eigenvalues than for a matrix with all real eigenvalues. We tried sorting the complex eigenvalues to improve the performance, but no significant improvement was achieved. It seems that the problem is harder to solve by the block Jacobi-like method when the matrix has some complex eigenvalues.

5 Conclusions

In this paper we first gave an analysis of the RTZ algorithm and showed that at best a linear convergence rate can be obtained when this algorithm is applied to decompose a normal matrix which has complex eigenvalues. However, our analysis and experimental results indicate that quadratic convergence may be achieved if the given matrix has only real eigenvalues. Thus the algorithm is still useful for eigenvalue decomposition of a normal or near normal matrix if the matrix has only real eigenvalues.

We then discussed a block Jacobi-like method for eigenvalue decomposition of a real normal matrix using real arithmetic. The theoretical analysis shows that ultimate quadratic convergence can be achieved for matrices with distinct eigenvalues if the orthogonal transformation is chosen properly. Some experimental results are also given to show the importance of incorporating a sorting procedure when the QR algorithm with double implicit shift is used for local block submatrix reduction. It is expected that ultimate quadratic, or near-quadratic convergence is achievable when the algorithms are applied to compute the eigenvalue decomposition of a near-normal matrix.

It should be noted that the algorithm used in our experiment is not the only candidate. For example, similar results were obtained when we used a scheme which combines the RTZ and the QR algorithms for local block submatrix reduction [14]. The key to the success is that the generated orthogonal transformation matrices at each step should satisfy the inequality in (20).

References

[1] R. P. Brent, "Parallel algorithms for digital signal processing", *Proceedings of NATO Advanced Study Institute on Numerical Linear Algebra, Digital Signal Processing and Parallel Algorithms*, Leuven, Belgium, August, 1988.

[2] R. P. Brent and F. T. Luk, "The solution of singular-value and symmetric eigenvalue problems on multiprocessor arrays", *SIAM J. Sci. and Statist. Comput.*, 6, 1985, pp. 69-84.

[3] P. J. Eberlein, "On the Schur decomposition of a matrix for parallel computation", *IEEE Trans. on Computers*, Vol. C-36, No. 2, 1987, pp. 167-174.

[4] F. T. Luk and H. Park, "On parallel Jacobi orderings", *SIAM J. Sci. and Statist. Comput.*, 10, 1989, pp. 18-26.

[5] N. Mackey, "Hamilton and Jacobi meet again: quaternions and the eigenvalue problem", *SIAM J. Matrix Anal. Appl.*, Vol. 16, No. 2, pp. 421-435.

[6] M. Mantharam and P. J. Eberlein, "The real two-zero algorithm: a parallel algorithm to reduce a real matrix to a real Schur form", *IEEE Trans. on Parallel and Distributed Systems*, Vol. 6, No. 1, 1995, pp. 48-62.

[7] W. F. Mascarenhas, "On the convergence of the Jacobi method for arbitrary orderings", *SIAM J. Matrix Anal. Appl.*, Vol. 16, No. 4, pp. 1197-1209.

[8] B. T. Smith, J. M. Boyle, J. J. Dongarra, B. S. Garbow, Y. Ikebe, V. C. Klema and C. B. Moler, *Matrix Eigensystem Routines - EISPACK Guide*, Vol. 6 of Lecture Notes in Computer Science, Springer-verlag, Berlin, 2 ed., 1976.

[9] G. W. Stewart, "A Jacobi-like algorithm for computing the Schur decomposition of a non-hermitian matrix", *SIAM J. Sci. and Statist. Comput.*, Vol. 6, No. 4, 1985, pp. 853-864.

[10] R. Schreiber, "Solving eigenvalue and singular value problems on an undersized systolic array", *SIAM. J. Sci. Stat. Comput.*, 7, 1986, pp. 441-451.

[11] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, England, 1965.

[12] B. B. Zhou and R. P. Brent, "A parallel ordering algorithm for efficient one-sided Jacobi SVD computations", Journal of Parallel and Distributed Computing, Vol. 42, 1997, pp.1-10.

[13] B. B. Zhou and R. P. Brent, "On the parallel implementation of the one-sided Jacobi algorithm for singular value decompositions", *Proc. of 3rd Euromicro Workshop on Parallel and Distributed Processing*, San Remo, Italy, January 1995, pp. 401-408.

[14] B. B. Zhou and R. P. Brent, "Jacobi-like algorithms for eigenvalue decomposition of a real normal matrix using real arithmetic", *Proceedings of the 10th International Parallel Processing Symposium*, Honolulu, Hawaii, April, 1996, pp.593-600. [A preliminary version of the present paper.]