Eigenvalues of $Ax = \lambda Bx$ for Real Symmetric Matrices A and B Computed by Reduction to a Pseudosymmetric Form and the HR Process*

M. A. Brebner
Department of Computer Science
University of Calgary
Alberta, Canada

and

J. Grad
Faculty of Economics and University Computer Centre
University of Ljubljana, Yugoslavia

Submitted by Richard A. Brualdi

ABSTRACT

The paper presents a method for solving the eigenvalue problem $Ax = \lambda Bx$, where A and B are real symmetric but not necessarily positive definite matrices, and B is nonsingular. The method reduces the general case into a form $Cz = \lambda z$ where C is a pseudosymmetric matrix. A further reduction of C produces a tridiagonal pseudosymmetric form to which the iterative HR process is applied. The tridiagonal pseudosymmetric form is invariant under the HR transformations. The amount of computation is significantly less than in treating the problem by a general method.

1. INTRODUCTION

This paper describes some numerical procedures which can be combined to form a possible method for solving the eigenvalue problem

$$Ax = \lambda Bx, \tag{1}$$

where A and B are n by n real symmetric matrices and at least one is nonsingular. We shall assume B is nonsingular.

*Work for this paper was partially supported by the National Research Council of Canada and by Raziskovalna skupnost Slovenije-Sklad "Borisa Kidriča" Ljubljana.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/82/020099+20$02.75

Before continuing it is necessary to define the term pseudosymmetric matrix.

DEFINITION 1. A matrix G is pseudosymmetric if $G = HJ$ where $H = H^T$ and $J = \text{diag}\{\pm 1\}$.

However, any matrix of the form $G = J_L S J_R$ with $S = S^T$ is pseudosymmetric, as $G = J_L S J_L J_L J_R = HJ$, where $H = J_L S J_L = H^T$, $J = J_L J_R$, and J_L and J_R are of the same form as the matrix J.

The method consists of three stages.

(i) Firstly, the system (1) is reduced to

$$(C - \lambda I)\mathbf{y} = \mathbf{0}, \tag{2a}$$

where C is a pseudosymmetric matrix.

(ii) Secondly, the system (2) is reduced to

$$(T - \lambda I)\mathbf{z} = \mathbf{0}, \tag{2b}$$

where T is a tridiagonal pseudosymmetric matrix.

(iii) Finally, an analogue of the QR algorithm for pseudosymmetric matrices is applied to find the eigenvalues of T. This is called the HR algorithm.

It should be noted that each stage accomplishes a specific task, and any one or more of the stages might be replaced by an alternative method, or, if appropriate, some of the stages might be used as part of some other algorithm. For example, stage (ii) could be replaced by an adaptation of the Lanczos method as suggested in [3]. Also, stages (ii) and (iii) preserve pseudosymmetry.

2. REDUCTION TO THE FORM $(C - \lambda I)\mathbf{y} = \mathbf{0}$

We describe three alternative approaches to this stage and discuss the associated problems.

2.1. *A Straightforward Reduction*

A straightforward reduction can be achieved by

(i) decomposing a matrix similar to B to a form LJL^T:

$$P^{-1}BP = LJL^T, \tag{3}$$

where P is a product of orthogonal plane rotation matrices and L is a lower triangular matrix (Jacobi type plane rotations are introduced to avoid a possible breakdown of the process and to improve the numerical stability as discussed in [2] and [6]), and then

(ii) premultiplying (1) by P^{-1} and using (3) to obtain

$$P^{-1}APP^{-1}\mathbf{x}=\lambda LJL^{T}P^{-1}\mathbf{x},$$

which can be arranged to give

$$JC'\mathbf{y}=\lambda I\mathbf{y},$$

or

$$C\mathbf{y}=\lambda I\mathbf{y},$$

where $C'=L^{-1}P^{-1}APL^{-T}$ is symmetric, $\mathbf{y}=L^{T}P^{-1}\mathbf{x}$, and $C=JC'$.

The matrix L^{-1} is a full lower triangular matrix and is computed explicitly. This is the weak point of the process and should if possible be avoided. In the next subsection we show how the reduction can be achieved in a stable and efficient manner.

2.2. *A Stepwise Reduction*

After each main step r, $r=1,2,\ldots,n-1$, the algorithm produces a new matrix form B_r, where

$$B_r=(W_rD_rP_rI_{rk}R_r)B_{r-1}(W_rD_rP_rI_{rk}R_r)^{T},$$

$$B_0=B,$$

$$B_n=D_nB_{n-1}D_n=J.$$

The form of B_{r-1} is

$$B_{r-1}=\begin{array}{c} \\ r-1 \\ n-r+1 \end{array}\begin{array}{c} \overset{r-1}{} \\ \left[\begin{array}{c|ccc} J & & & \\ \hline & b_{r,r} & b_{r,r+1} & \cdots & b_{r,n} \\ 0 & b_{r+1,r} & b_{r+1,r+1} & \cdots & b_{r+1,n} \\ & \vdots & & & \\ & b_{n,r} & & \cdots & b_{n,n} \end{array}\right] \end{array}.$$

The same transformations are performed simultaneously on the matrix A to produce essentially C' after $n-1$ steps.

Each step r consists of three substeps:

2.2.1. Two similarity transformations using a Jacobi plane rotation matrix R_r in the (p, q) plane [8, p. 266] and an elementary matrix I_{rk} [8, p. 44] are chosen so that the elements b'_{ij} of

$$B'_{r-1} = I_{rk} R_r B_{r-1} R_r^T I_{rk}^T$$

satisfy the condition

$$|b'_{rr}| \geqslant \frac{\max\limits_{r \leqslant i,j \leqslant n} |b'_{ij}|}{\sqrt{2}} \tag{4}$$

The relation (4) was proved in [6]. For some matrices this transformation improves the stability of substep 2.2.3 below, and allows an estimation of $\|B_r\|$. When no improvement is expected, this stage can be bypassed. The (p, q) plane and index k in I_{rk} are selected so that

$$|b_{pq}| = \max_{r \leqslant i,j \leqslant n} |b_{ij}|$$

and

$$\tilde{b}_{kk} = \max\left(|\tilde{b}_{pp}|, |\tilde{b}_{qq}|\right),$$

where b_{ij} and $\tilde{b}_{ij}$ are the elements in the (i, j) positions of B_{r-1} and $R_r B_{r-1} R_r^T$, respectively.

2.2.2. A Householder transformation [8, p. 290]

$$B''_{r-1} = P_r B'_{r-1} P_r$$

is then applied to reduce B'_{r-1} to the form

$$B''_{r-1} = \begin{matrix} \\ r-1 \\ n-(r-1) \end{matrix} \begin{matrix} \begin{matrix} r-1 & n-(r-1) \end{matrix} \\ \left[\begin{array}{c|c} J & 0 \\ \hline 0 & H_r \end{array}\right] \end{matrix},$$

where

$$H_r=\begin{bmatrix} h_{rr} & h_{r,r+1} & 0 & \cdots & 0 \\ h_{r+1,r} & h_{r+1,r+1} & h_{r+1,r+2} & \cdots & h_{r+1,n} \\ 0 & h_{r+2,r+1} & h_{r+2,r+2} & \cdots & h_{r+2,n} \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & h_{n,r+1} & h_{n,r+2} & \cdots & h_{n,n} \end{bmatrix}$$

As a result of the transformations in substeps 2.2.1 and 2.2.2 we have

$$\begin{aligned} |h_{rr}| &= |b'_{rr}| \geq b'_M/\sqrt{2} \\ |h_{r+1,r}| &\leq b'_M\sqrt{n-r-1} \end{aligned} \tag{5}$$

where $b'_M=\max_{r\leq i,j\leq n}(|b'_{ij}|)$.

2.2.3. Finally the transformation

$$B_r=W_rD_rB''_{r-1}D_rW_r^T$$

reduces B''_{r-1} to the form

$$B_r=\begin{array}{c} \\ r \\ n-r \end{array}\begin{array}{c} \begin{array}{cc} r & \quad n-r \end{array} \\ \left[\begin{array}{c|ccc} J & & 0 & \\ \hline & h'_{r+1,r+1} & h_{r+1,r+2} & \cdots & h_{r+1,n} \\ 0 & h_{r+2,r+1} & h_{r+2,r+2} & \cdots & h_{r+2,n} \\ & \vdots & \vdots & & \vdots \\ & h_{n,r+1} & h_{n,r+2} & \cdots & h_{n,n} \end{array}\right] \end{array}.$$

The elements d_{ij} and w_{ij} of matrices D_r and W_r which produce this reduction are

$$\begin{aligned} d_{ii} &= 1 \qquad \text{for} \quad i\neq r, \\ d_{rr} &= 1/\sqrt{|h_{rr}|}\,, \\ d_{ij} &= 0 \qquad \text{for} \quad i\neq j, \end{aligned}$$

and

$$w_{ii} = 1,$$

$$w_{ij} = 0 \qquad \text{for} \quad i \neq j,$$

except $w_{r+1,r}$, which is defined by

$$w_{r+1,r} = \pm h_{r+1,r} / \sqrt{|h_{rr}|}\,.$$

The minus sign in the definition of $w_{r+1,r}$ is used when $h_{rr} > 0$ and therefore $j_{rr} = 1$, while the plus sign is used when $h_{rr} < 0$ and therefore $j_{rr} = -1$.

From (5)

$$|w_{r+1,r}| \leq \sqrt{\sqrt{2}\,(n-r-1)b'_M}\,.$$

Considering the properties of the transformations used in substeps 2.2.1 to 2.2.3, the following relations are obtained:

$$\|B''_{r-1}\|_2 = \|B'_{r-1}\|_2 = \|B_{r-1}\|_2,$$

$$\|W_r\|_2^2 \leq \max\left(1, 2 + h_{r+1,r}^2/|h_{rr}|\right) = 2 + h_{r+1,r}^2/|h_{rr}|,$$

$$\|D_r\|_2^2 = \max(1, 1/|h_{rr}|),$$

and

$$\begin{aligned}\|B_r\|_2 &= \|W_r D_r B''_{r-1} D_r W_r^T\|_2 \\ &\leq \|W_r\|_2^2 \|D_r\|_2^2 \|B_{r-1}\|_2 \\ &\leq \max\left(1, \frac{1}{|h_{rr}|}\right)\left(2 + \frac{h_{r+1,r}^2}{|h_{rr}|}\right)\|B_{r-1}\|_2.\end{aligned}$$

This indicates the important role of the values h_{rr} and $h_{r+1,r}$ in the stability of the reduction process.

Let us now consider the effect of the simultaneous transformation of A to C'. The transformations using the matrices R_r, I_{rk}, and P_r in substeps 2.2.1 and 2.2.2 are stable. Also after all $n-1$ transformations in substep 2.2.3, each

individual element a_{rs} will be altered at most twice by the transformations using the D_r matrices and at most twice by the transformations using the W_r matrices. As a result these steps produce an estimated amplification factor for a typical element a_{rs} of

$$2\sqrt{2b_M^{(r-1)}(n-r)b_M^{(s-1)}(n-s)/b_M^{(s)}b_M^{(r)}}$$

where $b_M^{(l)}$ is the value of b'_M at the lth step. This amplification factor is unlikely to be attained in practice.

The number of operations involved in the computation of J is approximately $\frac{5}{6}n^3$ additions (comparisons), $\frac{2}{3}n^3$ multiplications, and $4n$ square roots, while the computation of the elements of C' involves approximately n^3 additions and n^3 multiplications. The distribution of the approximate number of operations is presented in the following table for the transformation of A to C', and the reduction of B to J, where full advantage has been taken of symmetry:

	A to C'			B to J		
Substep	+	*	SQRT	+	*	SQRT
2.2.1	$2n^2$	$4n^2$	0	$\frac{1}{6}n^3$	$2n^2$	$2n-2$
2.2.2	n^3	n^3	0	$\frac{2}{3}n^3$	$\frac{2}{3}n^3$	$n-2$
2.2.3	n^2	$2n^2$	0	n	$2n$	$n-1$

2.3. *Reduction by Orthogonal Transformations and a Diagonal Transformation*

The third possible approach is reduction by orthogonal transformations and a diagonal transformation. In this case B is reduced firstly to a symmetric tridiagonal form H by Householder transformations [8, p. 290]. Then H is reduced to $\operatorname{diag}\{\beta_i\}$ by the QR iterative process [8, p. 515] to give

$$QBQ^T=\operatorname{diag}\{\beta_i\}=J\operatorname{diag}\{|\beta_i|\},$$

where

$$Q^{-1}=Q^T$$

and β_i are the eigenvalues of B. Correspondingly A is transformed to

$$C''=QAQ^T.$$

Finally

$$C = J\,\mathrm{diag}\{1/|\beta_i|^{1/2}\}\, C''\, \mathrm{diag}\{1/|\beta_i|^{1/2}\}.$$

Stability is assured throughout the process, while the number of operations required is generally much larger than in the method described in Section 2.2. Assuming that there are $2n$ iteration steps in the QR process and that deflation of the matrix is used, the following gives an estimate of the number of operations:

+	*	SQRT
$3n^3$	$4n^3$	$2n^2$

A possible danger is in the final multiplications by $\mathrm{diag}\{1/|\beta_i|^{1/2}\}$, but here at least the potential danger is obvious.

3. TRANSFORMATION TO $(T-\lambda I)\mathbf{z}=\mathbf{0}$

The problem of finding the eigenvalues of a pseudosymmetric matrix can be successfully solved by means of the well-known algorithms for general unsymmetric matrices [8, Chapter 8; 5]. These algorithms do not preserve pseudosymmetry, and require more arithmetic operations than methods which do, particularly in the later iterative steps.

The transformation to a tridiagonal pseudosymmetric matrix can be performed by an analogue of either Givens's or Householder's method.

3.1. *Pseudosymmetric Givens Method*

An analogue of Givens's method [8, p. 282] which preserves pseudosymmetry is now described. The form of the individual transformations is

$$QJC'Q^{-1} = (QJQ^T)(Q^{-T}C'Q^{-1}),$$

where QJQ^T produces J or J permuted, and $Q^{-T}C'Q^{-1}$ is obviously symmetric.

The transformation of C to T consists of $n-2$ general steps. In the rth step $(r=1,2,\ldots,n-2)$ the elements in positions (i,r) and (r,i), for $r+2 \leqslant i \leqslant n$, of the intermediate matrix, say $C_{r,i}$, are reduced to zero by means of the similarity transformation

$$Q_{r+1,i} C_{r,i} Q_{r+1,i}^{-1},$$

where $Q_{r+1,i}$ produces a similarity transformation in the $(r+1,i)$ plane.

We shall now describe $Q_{r+1,i}$ and $Q_{r+1,i}^{-1}$ in detail. The elements of $Q_{r+1,i}$ and its inverse in the positions $(r+1,r+1)$, $(r+1,i)$, $(i,r+1)$, and (i,i) are functions of elements of $C_{r,i}$. All other diagonal elements of the two matrices are equal to 1, and all other off-diagonal elements are equal to 0.

Denote the elements of $C_{r,i}$, $Q_{r+1,i}$, and $Q_{r+1,i}^{-1}$ by c_{st}, q_{st}, and q_{st}^{-1} respectively. Assume further that $c_{r+1,r}\neq 0$ and $c_{i,r}\neq 0$. The case when this is not true is also explained below as a special case. Then the following two possibilities can occur:

(i) $\operatorname{sign}(c_{r+1,r}\cdot c_{i,r})=\operatorname{sign}(c_{r,r+1}\cdot c_{r,i})$. For this case

$$q_{r+1,r+1}=-q_{i,i}=c_{r+1,r}/\left(c_{r+1,r}^2+c_{i,r}^2\right)^{1/2},$$

$$q_{r+1,i}=q_{i,r+1}=c_{i,r}/\left(c_{r+1,r}^2+c_{i,r}^2\right)^{1/2},$$

and

$$Q_{r+1,i}^{-1}=Q_{r+1,i}^{T}$$

(ii) $\operatorname{sign}(c_{r+1,r}\cdot c_{i,r})=-\operatorname{sign}(c_{r,r+1}\cdot c_{r,i})$. For this case we have

(a) If $|c_{r+1,r}|>|c_{i,r}|$, then

$$q_{r+1,r+1}=q_{i,i}=q_{r+1,r+1}^{-1}=q_{i,i}^{-1}=\frac{c_{r+1,r}}{\left(c_{r+1,r}^2-c_{i,r}^2\right)^{1/2}},$$

$$q_{r+1,i}=q_{i,r+1}=-q_{r+1,i}^{-1}=-q_{i,r+1}^{-1}=-\frac{c_{i,r}}{\left(c_{r+1,r}^2-c_{i,r}^2\right)^{1/2}}.$$

(b) If $|c_{r+1,r}|<|c_{i,r}|$, then

$$q_{r+1,r+1}=q_{i,i}=-q_{r+1,r+1}^{-1}=-q_{i,i}^{-1}=\frac{c_{r+1,r}}{\left(c_{i,r}^2-c_{r+1,r}^2\right)^{1/2}},$$

$$q_{r+1,i}=q_{i,r+1}=q_{r+1,i}^{-1}=q_{i,r+1}^{-1}=-\frac{c_{i,r}}{\left(c_{i,r}^2-c_{r+1,r}^2\right)^{1/2}}.$$

Three special cases can occur in the similarity transformation:

(1) *When* $c_{i,r}=0$. If this happens then no similarity transformation is necessary for this i.

(2) *When* $c_{r+1,r}=0$ *and* $c_{i,r}\neq 0$. In this case rows and columns $r+1$ and i have to be exchanged. No other similarity transformation is necessary for this i.

(3) *In case* (ii) *when* $c_{r+1,r}=c_{i,r}\neq 0$. In this case an index s, $i<s\leqslant n$, has to be found for which

$$c_{r+1,r}\neq c_{s,r}.$$

Then the rows and columns i and s have to be exchanged and the similarity transformation performed. But if $i=n$ this transformation fails to work. However, there is only a small possibility that case (3) will occur in practice. If it does occur, then some rows and columns of C must be exchanged by means of similarity transformations and the whole process repeated.

3.2. *Pseudosymmetric Householder Method*

An alternative is to use an analogue of Householder's method for pseudosymmetric matrices.

The process consists of $n-2$ general steps r, $r=1,2,\ldots,n-2$. In each step the similarity transformation is performed, giving

$$\begin{aligned} C_{r+1} &= \left(I-2J_r v_r v_r^T\right) C_r \left(I-2J_r v_r v_r^T\right) \\ &= P_r C_r P_r = P_r J_r C_r' P_r, \end{aligned} \tag{6}$$

where $C_1=C$, $C_{n-2}=T$, C_r' is symmetric, and J_r belongs to the class of J matrices. The elements of the column vector v_r of length n are

$$v_i^{(r)}=0 \qquad \text{for} \quad 1\leqslant i\leqslant r,$$

$$\left[v_{r+1}^{(r)}\right]^2 = j_{r+1,r+1}^{(r)} \frac{1+\operatorname{sign}\left(c_{r+1,r}^{(r)}\right) c_{r+1,r}^{(r)}/S_r}{2},$$

and

$$v_i^{(r)} = j_{ii}^{(r)} \operatorname{sign}\left(c_{r+1,r}^{(r)}\right) c_{i,r}^{(r)} / \left[2 v_{r+1}^{(r)} S_r\right] \qquad \text{for} \quad r+2\leqslant i\leqslant n,$$

where $j_{ii}^{(r)}$ are the elements of J_r, $c_{ik}^{(r)}$ are the elements of C_r, and

$$S_r = \left\{ j_{r+1,r+1}^{(r)} \sum_{i=r+1}^{n} j_{ii}^{(r)} \left[c_{i,r}^{(r)}\right]^2 \right\}^{1/2}. \tag{7}$$

It should be noted that $v_r^T J_r v_r = 1$ [1]. It was proved in [1] that the transformation (6): (i) preserves a pseudosymmetric matrix form, (ii) is a similarity transformation, and (iii) reduces to zero the elements of C_r in the positions $(r+2, r),(r+3, r),\ldots,(n, r)$ and $(r, r+2),(r, r+3),\ldots,(r, n)$. The transformation preserves all the zeros obtained in previous steps.

The value of $v_{r+1}^{(r)}$ can be imaginary, but in practice essentially only the value of $[v_{r+1}^{(r)}]^2$ is required when computing the elements of P_r.

The definition (7) of S_r can also admit an imaginary value. In order to avoid complex arithmetic the transformation matrices must be modified when S_r^2 is negative. The sign of S_r^2 can be changed by swapping rows and columns $r+1$ and t, where $j_{r+1,r+1}^{(r)} = -j_{tt}^{(r)}$ and $r+2 \leqslant t \leqslant n$.

LEMMA. *If* $S_r^2 < 0$, *there exists a* t *such that* $j_{r+1,r+1}^{(r)} = -j_{tt}^{(r)}$ *and* $r+2 \leqslant t \leqslant n$.

Proof. Let $S_r^2 < 0$. Then $j_{ii}^{(r)}$, $i = r+1,\ldots,n$, can be neither all negative nor all positive. Hence there exists a t such that $j_{r+1,r+1}^{(r)} = -j_{tt}^{(r)}$ and $r+2 \leqslant t \leqslant n$.

The similarity permutation of rows and columns $r+1$ and t swaps $j_{r+1,r+1}^{(r)}$ and $c_{r+1,r}^{(r)}$ with $j_{tt}^{(r)}$ and $c_{tr}^{(r)}$, for

$$\begin{aligned} I_{r+1,t} C_r I_{r+1,t} &= I_{r+1,t} J_r C_r' I_{r+1,t} \\ &= (I_{r+1,t} J_r I_{r+1,t})(I_{r+1,t} C_r' I_{r+1,t}) \\ &= J_r^* C_r'^* = C_r^*. \end{aligned}$$

The value of the sum in S^2 is not changed by the similarity permutation, but the sign of the new $j_{r+1,r+1}^{(r)}$ changes. Hence the sign of S_r^2 changes from negative to positive.

Possibly some other method could be used instead of the pseudosymmetric Givens or Householder method. The reduction of a general matrix to a tridiagonal form is analyzed in detail in [8, Chapter 5]. It is shown that the process can break down at some stage. When this occurs the process must be repeated from the beginning with a transformed starting matrix.

Clearly a potential source of numerical breakdown or severe cancellation errors occurs for the case where S_r^2 is zero or very small. If S_r^2 is not small, the grouping of positive and negative terms removes the danger of cancellation errors. The calculation of S_r^2 should always use double precision accumulation.

4. REDUCTION TO $(S-\lambda I)v=0$ BY THE ITERATIVE *HR* PROCESS.

The system (2b) is reduced to

$$Sv=\lambda v, \tag{8}$$

where for a matrix T with k, $0\leqslant k\leqslant n/2$, pairs of complex conjugate eigenvalues $\lambda_r \pm i\mu_r$, $r=1,2,\ldots,k$, and real eigenvalues $\lambda_{2k+1},\ldots,\lambda_n$. Here S is of the form

$$PSP^T=\begin{bmatrix} X_1 & & & & & \\ & X_2 & & 0 & & \\ & & X_k & & & \\ & 0 & & \lambda_{2k+1} & & \\ & & & & \ddots & \\ & & & & & \lambda_n \end{bmatrix},$$

where X_r is a 2×2 matrix on the diagonal of PSP^T having the eigenvalues $\lambda_r \pm i\mu_r$ [8, p. 486]. P is the product of matrices I_{ij}, $1\leqslant i,j\leqslant n$, where a similarity transformation with the I_{ij} interchanges rows i and j and columns i and j [8, p. 44]. The eigenvalues of S which are also the computed eigenvalues of (1) can be extracted from S directly.

The *HR* process is very similar to the *QR* process, which is described in detail in [8], the *H* standing for potentially hyperbolic transformations. Shifts of origin are introduced to increase the rate of convergence. In fact these shifts are essential to insure the *HR* decomposition is constructable, as remarked in [3]. Shifts are obtained from the two roots of the bottom right-hand corner 2×2 submatrix of the matrix on which the *HR* transformation is performed. The convergence of the *HR* process is discussed in [3], and the suitability for numerical algorithms of the decomposition used is justified in Elsner [4].

For the case when T has real eigenvalues the single step *HR* process can be applied with the shifts being real numbers. All calculations can be performed with real numbers, providing that T is real. For T having some complex conjugate pairs of eigenvalues the shifts may be complex numbers as well. In order to avoid complex arithmetic the double step *HR* process must be applied in this case. Both types of *HR* process preserve the pseudosymmetric tridiagonal matrix form.

We now give a short description of each of the two *HR* processes mentioned above. Those details which make *HR* different from *QR* will be pointed out precisely.

4.1. The Single Step HR Process

The process is defined by

$$T_s - k_s I = H_s R_s, \qquad R_s H_s + k_s I = T_{s+1},$$

giving

$$T_{s+1} = H_s^{-1} T_s H_s \qquad \text{for} \quad s = 1,2,3,\ldots,$$

where $T_1 = T$, R_s is an upper triangular matrix, $H_s^{-1} = H_{n,n-1}^{-1} \cdots H_{3,2}^{-1} H_{2,1}^{-1}$ with $H_{i+1,i}^{-1}$ as described later, and k_s is the shift of origin. In the computer algorithm starting with $T_1' = T$ a sequence of matrices T_s', instead of T_s, is computed as follows:

Firstly $T_s' - k_s I$ is premultiplied by $H_{n,n-1}^{-1} \cdots H_{3,2}^{-1} H_{2,1}^{-1}$, giving R_s, where k_s is the shift of origin computed from T_s'. Secondly R_s is postmultiplied by $H_{2,1} H_{3,2} \cdots H_{n,n-1}$, giving T_{s+1}'. A total shift $k_t = \Sigma k_s$ is accumulated. When T_{s+1}' is of the form of the matrix S of (8) to working accuracy, the total shift k_t is added to the diagonal elements of T_{s+1}', giving T_{s+1}.

$H_{i+1,i}^{-1}$ is defined so that:

(I) Premultiplication of $H_{i,i-1}^{-1} \cdots H_{2,1}^{-1}(T_s' - k_s I)$—which is, for example for $n = 5$, $i = 3$, of the form

$$H_{3,2}^{-1} H_{2,1}^{-1}(T_s' - R_s I) = \begin{bmatrix} x & x & x & 0 & 0 \\ 0 & x & x & x & 0 \\ 0 & 0 & x_{i,i} & x & 0 \\ 0 & 0 & x_{i+1,i} & x & x \\ 0 & 0 & 0 & x & x \end{bmatrix} \tag{9}$$

—by $H_{i+1,i}^{-1}$ reduces to zero the element in the position $(i+1, i)$.

(II) $H_s^{-1}(T_s' - k_s I)$ is upper triangular.

(III) T_{s+1}' is a tridiagonal pseudosymmetric matrix.

For the case when $H_{i+1,i}^{-1}$ is a rotation in the plane $(i, i+1)$, conditions (I) and (II) are fulfilled.

To fulfill (III) the sign patterns of the off diagonal elements t_{ij} of T_s' have to be taken into account when defining $H_{i+1,i}^{-1}$. Suppose that

$$\text{sign}(i) = \text{sign}(t_{i+1,i} t_{i,i+1}) \qquad \text{for} \quad i = 1,2,\ldots,n-1,$$

and suppose that in the formulae defining cases (i) and (ii) of Section 3.1 the element $c_{r+1,r}$ has been replaced by $x_{r+1,r+1}$ and the element $c_{i,r}$ has been replaced by $x_{i,r+1}$ of (9). Then for sign(i) positive, $H_{i+1,i}^{-1}$ is of Givens's form,

i.e. $H_{i+1,i}^{-1} = Q_{i,i+1}$, where $Q_{i,i+1}$ is defined by (i) in Section 3.1. For sign(i) negative, then $H_{i+1,i}^{-1} = Q_{i,i+1}$, where $Q_{i,i+1}$ and its inverse are defined by (ii) in Section 3.1. In case (b), when $|x_{i,i}| < |x_{i+1,i}|$, the sign patterns will be modified in the matrix

$$\left[H_{i+1,i}^{-1} \cdots H_{3,2}^{-1} H_{2,1}^{-1} (T_s' - k_s I) H_{2,1} H_{3,2} \cdots H_{i,i-1}\right] H_{i+1,i},$$

and the computations must take account of these sign changes when determining the type of the matrix $H_{i+2,i+1}^{-1}$. The computer algorithm must therefore modify the sign pattern indicators as follows.

$$\begin{aligned} \text{sign}(i+1) &= -\text{sign}(i+1), \\ \text{sign}(i-1) &= -\text{sign}(i-1), \end{aligned} \tag{10}$$

before $H_{i+2,i+1}^{-1}$ is determined. It should again be noted that the numerical values of the elements of $H_{i+2,i+1}^{-1}$ are determined from the matrix

$$H_{i+1,i}^{-1} \cdots H_{2,1}^{-1} (T_s' - k_s I).$$

If the sign indicators are modified as in (10) in step s of the HR process, then they can be used as a starting set of indicators for step $s+1$ of HR.

The single step HR process was applied to a set of pseudosymmetric tridiagonal matrices with real eigenvalues, and in each case the sequence of matrices T_s' rapidly tended to a diagonal form.

4.2. The Double Step HR Process

The process is defined by

$$T_s - k_1 I = H_s R_s, \qquad R_s H_s + k_1 I = T_{s+1},$$

$$T_{s+1} - k_2 I = H_{s+1} R_{s+1}, \qquad R_{s+1} H_{s+1} + k_2 I = T_{s+2},$$

giving

$$(H_s H_{s+1})(R_{s+1} R_s) = (T_s - k_1 I)(T_s - k_2 I)$$

and

$$T_{s+2} = (H_s H_{s+1})^{-1} T_s (H_s H_{s+1}) \tag{11}$$

for $s=1,3,5,\ldots$, where k_1 and k_2 are either two real or a pair of complex conjugate eigenvalues for the bottom right-hand corner 2×2 submatrix of T_s; R_s and R_{s+1} are upper triangular matrices; and the matrix $(H_sH_{s+1})^{-1}$ is defined so that

$$(H_sH_{s+1})^{-1}\left[(T_s-k_1I)(T_s-k_2I)\right]=R_{s+1}R_s \tag{12}$$

where $R_{s+1}R_s$ is an upper triangular matrix.

Following [8], we can see that $(H_sH_{s+1})^{-1}$, which satisfies (12) and performs (11), can be obtained in practice as a product of elementary matrices $Q_{r,i}$ of type (i) and (ii) respectively defined in Section 3.1. In fact we have

$$\begin{aligned}(H_sH_{s+1})^{-1}&=(Q_{2,3}Q_{2,4}Q_{3,4}Q_{3,5}\cdots Q_{n-1,n}Q_{n-1,n+1}Q_{n,n+1})^{-1}\\&=(P_2P_3\cdots P_{n-1}P_n)^{-1},\end{aligned}$$

where

$$P_i=Q_{i,i+1}Q_{i,i+2}\quad\text{for } i=2,3,\ldots,n-1\quad\text{and}\quad P_n=Q_{n,n+1}.$$

Essentially as in [8], we find

$$P_n^{-1}\Big(P_{n-1}^{-1}\Big(\cdots\Big(P_3^{-1}\big(P_2^{-1}T_s'P_2\big)P_3\Big)\cdots\Big)P_{n-1}\Big)P_n=T_{s+2}', \tag{13}$$

where

$$T_s'=\left[\begin{array}{c|ccccc} 0 & x & y' & z' & 0 & \cdots & 0\\ \hline x & & & & & & \\ y & & & & & & \\ z & & & & & & \\ 0 & & & T_s & & & \\ \vdots & & & & & & \\ 0 & & & & & & \end{array}\right]$$

and

$$T_{s+2}'=\left[\begin{array}{c|cccc} 0 & x & 0 & \cdots & 0\\ \hline x & & & & \\ 0 & & & & \\ \vdots & & & T_{s+2} & \\ 0 & & & & \end{array}\right].$$

The elements x, y, z are computed as in QR from the elements of T_s:

$$x = t_{11}^2 + t_{12}t_{21} - t_{11}(k_1 + k_2) + k_1k_2,$$

$$y = t_{21}(t_{11} + t_{22} - k_1 - k_2),$$

$$z = t_{32}t_{21}.$$

In addition, for HR two more elements y' and z' must be computed from

$$y' = t_{12}(t_{11} + t_{22} - k_1 - k_2),$$

$$z' = t_{23}t_{12},$$

where $|y'| = |y|$ and $|z'| = |z|$. The parentheses in (13) show the order in which the similarity transformations by P_i^{-1} and P_i are performed.

Let

$$T'_{s,i} = P_i^{-1}T'_{s,i-1}P_i \tag{14}$$

for $i = 2,3,\ldots,n$, where $T'_{s,1} = T'_s$ and $T'_{s,n} = T'_{s+2}$. In the similarity transformation (14), first the premultiplication of $T'_{s,i-1}$ by P_i^{-1} is performed, followed by the postmultplication of $P_i^{-1}T'_{s,i-1}$ by P_i. The matrix $P_i^{-1} = Q_{i,i+2}^{-1}Q_{i,i+1}^{-1}$ is defined so that premultiplication of $T'_{s,i-1}$ by $Q_{i,i+1}^{-1}$ reduces to zero the element in the position $(i+1, i-1)$ of $Q_{i,i+1}^{-1}T'_{s,i-1}$, and premultiplication of $Q_{i,i+1}^{-1}T'_{s,i-1}$ by $Q_{i,i+2}^{-1}$ reduces to zero the element in the position $(i+2, i-1)$ of $P_i^{-1}T'_{s,i-1}$.

In order to preserve the pseudosymmetric tridiagonal form of T_{s+2}, the sign patterns of the elements of $T'_{s,i-1}$ must be considered when defining $Q_{i,i+1}^{-1}$ and $Q_{i,i+2}^{-1}$. Suppose that the (i,j) element of $T'_{s,i-1}$ is t'_{ij}, that in the formulae for (i) and (ii) in Section 3.1 the elements c_{ij} have been replaced by t'_{ij}, and that

$$\text{sign}(k) = \text{sign}(t'_{i,i-1}t'_{i-1,i}t'_{k,i-1}t'_{i-1,k}) \qquad \text{for} \quad k = i+1, i+2.$$

Then for sign(k) positive $Q_{i,k}^{-1} = Q_{i,k}$, where $Q_{i,k}$ is defined by (i) in Section 3.1. For sign(k) negative $Q_{i,k}^{-1} = Q_{i,k}$, where $Q_{i,k}$ is defined by (ii) in Section 3.1. In case (b) for $k = i+1$, when $|t'_{i,i-1}| < |t'_{i+1,i-1}|$, the sign pattern indicator sign$(i+2)$ must be modified as follows:

$$\text{sign}(i+2) = -\text{sign}(i+2)$$

before $Q_{i,i+2}^{-1}$ is determined.

The numerical values of the elements of $Q_{i,i+1}^{-1}$ are determined from $T'_{s,i-1}$, while the numerical values of the elements of $Q_{i,i+2}^{-1}$ are determined from $Q_{i,i+1}^{-1}T'_{s,i-1}$.

For the purpose of programming it is necessary to know the form of matrices $T'_{s,i}$. We present two examples for $n=8$ and $i=2,3$, where we have marked by X those elements which are in general different from zero. They are

$$T'_{s,2}=\left[\begin{array}{c|ccccccc} 0 & X & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline X & X & X & X & X & 0 & 0 & 0 \\ 0 & X & X & X & 0 & 0 & 0 & 0 \\ 0 & X & X & X & X & 0 & 0 & 0 \\ 0 & X & 0 & X & X & X & 0 & 0 \\ 0 & 0 & 0 & 0 & X & X & X & 0 \\ 0 & 0 & 0 & 0 & 0 & X & X & X \\ 0 & 0 & 0 & 0 & 0 & 0 & X & X \end{array}\right],$$

$$T'_{s,3}=\left[\begin{array}{cc|cccccc} 0 & X & 0 & 0 & 0 & 0 & 0 & 0 \\ X & X & X & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & X & X & X & X & X & 0 & 0 \\ 0 & 0 & X & X & X & 0 & 0 & 0 \\ 0 & 0 & X & X & X & X & 0 & 0 \\ 0 & 0 & X & 0 & X & X & X & 0 \\ 0 & 0 & 0 & 0 & 0 & X & X & X \\ 0 & 0 & 0 & 0 & 0 & 0 & X & X \end{array}\right].$$

The double step *HR* process was applied to a set of pseudosymmetric tridiagonal matrices with real and complex conjugate eigenvalues, and in each case the sequence of matrices T_s rapidly tended to the form of the matrix S in (8). It should be noted that the double step process also converged rapidly for the set of matrices with only real eigenvalues used to test the convergence of the single step process.

5. EXAMPLES

To test the method described in the previous sections a program has been written in FORTRAN and applied to some examples of form (1). The program is subdivided into three subroutines. Each subroutine solves an independent stage of the general problem:

Subroutine (i) transforms the system $Ax=\lambda Bx$ by the method described in Section 2.2 into the form $Cy=\lambda y$, where C is in general a full pseudosymmetric matrix.

Subroutine (*ii*) transforms C into a pseudosymmetric tridiagonal form T, by the method described in Section 3.2.

Subroutine (*iii*) finds the eigenvalues of T by the double step HR method.

Firstly the subroutines were tested on a CDC 6400 (60 bit word) using double precision floating point arithmetic in subroutines (i) and (ii) and single precision in subroutine (iii). For examples 1 and 2 given below the results matched those computed by using the program EIGENP [5] on $B^{-1}A\mathbf{x}=\lambda\mathbf{x}$ to 12 decimal digits, and with those reported in [7] to 11 decimal digits. Several other tests on the CDC with $n=5$ and 6 produce comparable accuracy. The number of iterations required ranged from 5 to 9.

EXAMPLE 1. For

$$A=\begin{bmatrix} -1 & -3 & -3 & -3 & -3 & -3 \\ -3 & -4 & -3.1 & -3.1 & -3.1 & -3.1 \\ -3 & -3.1 & 2.8 & 3.8 & 3.8 & 3.8 \\ -3 & -3.1 & 3.8 & 9.8 & 10.7 & 10.7 \\ -3 & -3.1 & 3.8 & 10.7 & 12.6 & 14.6 \\ -3 & -3.1 & 3.8 & 10.7 & 14.6 & 15.6 \end{bmatrix},$$

$$B=\begin{bmatrix} -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 & 2 & 2 \\ -1 & 0 & 1 & 2 & 3 & 3 \\ -1 & 0 & 1 & 2 & 3 & 2 \end{bmatrix},$$

the computed eigenvalues $\lambda_i=\xi_i+i\eta_i$ are given by

i	ξ_i	η_i
1	4.18245919165$\underline{37}$	0
2	6.13692605088$\underline{57}$	0
3,4	0.90877040417$\underline{29}$	$\pm$1.93967680102$\underline{2}$
5,6	0.93153697455$\underline{65}$	$\pm$1.97197662562$\underline{0}$

EXAMPLE 2. For

$$A=\begin{bmatrix} 10 & 2 & 3 & 1 & 1 \\ 2 & 12 & 1 & 2 & 1 \\ 3 & 1 & 11 & 1 & -1 \\ 1 & 2 & 1 & 9 & 1 \\ 1 & 1 & -1 & 1 & 15 \end{bmatrix}, \qquad B=\begin{bmatrix} 12 & 1 & -1 & 2 & 1 \\ 1 & 14 & 1 & -1 & 1 \\ -1 & 1 & 16 & -1 & 1 \\ 2 & -1 & -1 & 12 & -1 \\ 1 & 1 & 1 & -1 & 11 \end{bmatrix},$$

the computed eigenvalues λ_i, which are all real, are

$$0.43278721101\underline{7}$$
$$0.66366274839\underline{2}$$
$$0.94385900466\underline{8}$$
$$1.10928454001\underline{7}$$
$$1.49235323254\underline{3}$$

Later a larger set of eighty 10×10, 14×14, and 19×19 problems were tested on a Honeywell level 2 DPS 68 processor (36 bit word) using double precision floating point arithmetic in subroutines (i), (ii), and (iii). Several of these matrices possessed 3 or more multiple roots, and the multiplicity of the roots ranged from 2 to 5. The computed eigenvalues were correct to at least 11 decimal digits, and in many cases the precision was 14 or more decimal digits. In two cases a pivoting strategy was essential in stage (i) for the computation of accurate results. Generally less than $1.3n$ iterations were required in stage (iii).

6. REMARKS

Although the authors have successfully applied the methods described in this paper to many problems, the reader is warned that the process can be unstable or break down for some problems, particularly during the reduction of C to a tridiagonal form. However, the single and double step HR processes never broke down and converged rapidly to the correct eigenvalues for a large set of test problems.

The authors wish to express their gratitude to Dr. J. H. Wilkinson (NPL) *and E. Zakrajsek (University of Ljubljana) for their very helpful comments and suggestions, and to J. Vrecko (University of Ljubljana) and W. Wiggins (University of Calgary) for help with some of the programming. They would also like to thank the referee for constructive comments.*

REFERENCES

1 M. A. Brebner and J. Grad, An analogue of Householder's method for pseudo-symmetric Matrices. Part 1—Basic formulae and method, Research Report No. 280, Department of Mathematics, Statistics and Computer Science, Univ. of Calgary, June 1975.

2 M. A. Brebner and J. Grad, A reduction of an indefinite symmetric matrix A to the form LJL^T by rotation and decomposition. Part 1—A plane rotation strategy, Research Report No. 263, Department of Mathematics, Statistics and Computer Science, Univ. of Calgary, Dec. 1974.
3 A. Bunse-Gerstner, An analysis of the *HR* algorithm for computing the eigenvalues of a matrix, *Linear Algebra Appl.* 35:155–173 (1981).
4 L. Elsner, On some algebraic problems in connection with general eigenvalue algorithms, *Linear Algebra Appl.* 26:123–138 (1979).
5 J. Grad and M. A. Brebner, Eigenvalues and eigenvectors of a real general matrix, *Comm. ACM* 11:820–825 (1968).
6 A. Kmet, Metoda Choleskega s kompletnim pivotiranjem, unpublished paper, Ljubljana, Yugoslavia, 1977.
7 R. S. Martin and J. H. Wilkinson, Reduction of the symmetric eigenproblem $Ax = \lambda Bx$ and related problems to standard form, *Numer. Math.* 11:99–110 (1968).
8 J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon, Oxford, 1965.

Received 16 July 1980; revised 27 July 1981