An Estimate for the Norms of Certain Cyclic Jacobi Operators*

PETER HENRICI AND KATHARINA ZIMMERMANN
Eidgenössische Technische Hochschule
Zürich, Switzerland

1. INTRODUCTION

Let $A = (a_{pq})$ be a real symmetric matrix of order n. A Jacobi method [1] for determining the eigenvalues of A consists in constructing a sequence of matrices $A_k = (a_{pq}^{(k)})$ where $A_0 = A$,

$$A_{k+1} = U_k A_k U_k^T, \qquad k = 0, 1, 2, \ldots, \tag{1.1}$$

and where U_k is an orthogonal matrix which up to a similarity transformation by a permutation matrix is equal to

$$\begin{pmatrix} \cos\phi & \sin\phi & & & \\ -\sin\phi & \cos\phi & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}.$$

The matrix U_k is said to define a rotation of A. The pair of indices (i, j) $(i < j)$ of the nontrivial superdiagonal element of U_k is said to form the pivot of the rotation, and ϕ is called its angle. The pivots and angles of the rotations U_k depend on k. Generally they are selected in such a fashion that the sequence $\{A_k\}$ converges to a diagonal matrix Λ, for if this is the case, then the diagonal elements of Λ are the eigenvalues of A. The closeness of A_k to Λ is measured by the quantity

$$s_k = \left[\sum_{1 \leqslant p < q \leqslant n} (a_{pq}^{(k)})^2 \right]^{1/2}. \tag{1.2}$$

* Dedicated to Professor A. M. Ostrowski on his 75th birthday.

Linear Algebra and Its Applications 1, 489–501 (1968)

A Jacobi method such that $s_k \to 0$ $(k \to \infty)$ for all matrices A_0 is called convergent.

In this paper we shall be concerned only with cyclic Jacobi methods. Here the $N = \frac{1}{2}n(n-1)$ different pivots are selected in some fixed cyclic order. If (i, j) is the pivot of the kth rotation, its angle ϕ_{ij} is always chosen such that $a_{ij}^{(k+1)} = 0$ (the "rotated" element is annihilated). The product of the N rotations belonging to a full cycle is said to form a sweep of the cyclic method. Many cyclic Jacobi methods are known to be convergent [1, 2, 6]. If the eigenvalues of A are distinct, and also in some other situations, the convergence is even known to be quadratic [4–7], i.e., there exists a constant K such that, if s_0 is sufficiently small,

$$s_N \leqslant K s_0^2. \tag{1.3}$$

In the present paper we shall study the convergence of certain cyclic Jacobi methods by scrutinizing the linear transformation inflicted upon the off-diagonal elements by a sweep. Let

$$\mathbf{a}^T = (a_{12}; a_{13}, a_{23}; \ldots; a_{1n}, a_{2n}, \ldots, a_{n-1,n})$$

be the vector of the N superdiagonal elements of A (it is convenient to retain the double indexing of the elements of A). Using the euclidean norm, $s_k = \|\mathbf{a}^{(k)}\|$. Let (i, j) denote the pivot of the kth rotation, and let ϕ_{ij} be its angle. The well-known formulas for any Jacobi method annihilating the rotated element can then be written

$$\mathbf{a}^{(k+1)} = \mathbf{R}^{(i,j)}\mathbf{a}^{(k)},$$

where $\mathbf{R}^{(i,j)}$ is a certain matrix of order N (see Section 2 for a complete listing of its elements). It follows that

$$\mathbf{a}^{(N)} = \mathbf{R}\mathbf{a}^{(0)}, \tag{1.4}$$

where

$$\mathbf{R} = \prod_{(i,j)} \mathbf{R}^{(i,j)},$$

the factors appearing in the reverse order of the pivots. The matrix $\mathbf{R}$ will be called the Jacobi operator associated with the particular ordering under consideration. In addition to the ordering, $\mathbf{R}$ also depends on the angles of rotation ϕ_{ij}, and thus on the matrix A, $\mathbf{R} = \mathbf{R}_A$. $\mathbf{R}$ is not, of course, a linear operator in the sense that $\mathbf{R}_{A+B} = \mathbf{R}_A + \mathbf{R}_B$. On the other hand, it can happen that the Jacobi operators $\mathbf{R}_A$ and $\mathbf{R}'_A$ belonging to two different orderings satisfy $\mathbf{R}_A = \mathbf{R}'_A$ for all A. Two such orderings are called equivalent. It is shown in [2] and [5] that the orderings by columns $\{(1, 2); (1, 3), (2, 3); \ldots; (1, n), (2, n), \ldots, (n-1, n)\}$ and by rows $\{(1, 2), (1, 3), \ldots, (1, n); (2, 3), \ldots, (2, n); \ldots; (n-1, n)\}$ are equivalent.

Denoting by $\|\mathbf{R}\|$ the spectral norm of $\mathbf{R}$, we shall prove:

THEOREM 1. *For all orderings equivalent to the ordering by columns and for all matrices* A,

$$\|\mathbf{R}\| \leqslant C, \tag{1.5}$$

where

$$C^2 = 1 - \prod_{j=3}^{n} \prod_{i=1}^{j-2} \cos^2 \phi_{ij} \tag{1.6}$$

(*empty products are* 1).

If ϕ_{ij} is chosen in the interval $[-\pi/4, \pi/4]$ (this is always possible [1]), then $\cos \phi_{ij} \geqslant 2^{-1/2}$, and

$$C^2 \leqslant 1 - 2^{-\frac{1}{2}(n-2)(n-1)} < 1.$$

Theorem 1 implies

$$\|\mathbf{a}^{(N)}\| = \|\mathbf{R}\mathbf{a}^{(0)}\| \leqslant C\|\mathbf{a}^{(0)}\|; \tag{1.7}$$

hence all cyclic Jacobi methods whose orderings are equivalent to the ordering by columns converge at least linearly.

The inequality (1.7) means the same as the result

$$s_N \leqslant C s_0 \tag{1.8}$$

established by a different method in [6]. However, Theorem 1 is more general than (1.8), for it shows, in the more explicit notation used above, that $\|\mathbf{R}_A\mathbf{b}\| \leqslant C\|\mathbf{b}\|$ for all vectors $\mathbf{b}$ and not merely for the special vector $\mathbf{b} = \mathbf{a}^{(0)}$ of the off-diagonal elements of A. We intend to make use of this observation in a subsequent paper.

As shown in [6], the result (1.8) implies the quadratic convergence in the case of separated eigenvalues. This can be seen directly as follows.

By a variant of Bernoulli's inequality,

$$\prod \cos^2 \phi_{ij} = \prod (1 - \sin^2 \phi_{ij}) \geqslant 1 - \sum \sin^2 \phi_{ij};$$

hence $C^2 \leqslant \sum \sin^2 \phi_{ij}$, where products and sums are extended as in (1.6). If the eigenvalues λ_i of A satisfy $|\lambda_i - \lambda_j| \geqslant 2\delta$ $(i \neq j)$, then for s_0 sufficiently small it follows as in [6] that

$$\sum \sin^2 \phi_{ij} \leqslant \delta^{-2} \sum_{k=0}^{N-1} (a_{ij}^{(k)})^2 \leqslant \delta^{-2} s_0{}^2.$$

Hence (1.5) implies $s_N{}^2 \leqslant \delta^{-2} s_0{}^4$, which is Wilkinson's form of the estimate (1.3) [7].

2. THE ROTATION MATRIX

If $\mathbf{R}^{(i,j)} = (r_{pq,st})$, then it is easily shown that

$$r_{ij,ij} = 0, \tag{2.1}$$

$$r_{pq,pq} = 1, \qquad p \neq i, j \qquad \text{and} \qquad q \neq i, j; \tag{2.2}$$

furthermore, if $c = \cos \phi_{ij}$, $\quad s = \sin \phi_{ij}$,

$$\begin{aligned}
&r_{pi,pi} = c, \quad && r_{pi,pj} = s \\
&r_{pj,pi} = -s, && r_{pj,pj} = c
\end{aligned}\Bigg\} \quad 1 \leqslant p < i,$$

$$\begin{aligned}
&r_{ip,ip} = c, \quad && r_{ip,pj} = s \\
&r_{pj,ip} = -s, && r_{pj,pj} = c
\end{aligned}\Bigg\} \quad i < p < j, \tag{2.3}$$

$$\begin{aligned}
&r_{ip,ip} = c, \quad && r_{ip,jp} = s \\
&r_{jp,ip} = -s, && r_{jp,jp} = c
\end{aligned}\Bigg\} \quad j < p \leqslant n.$$

All remaining elements of $\mathbf{R}^{(i,j)}$ are zero. If $r_{ij,ij}$ were 1 instead of 0, the matrix $\mathbf{R}^{(i,j)}$ would be orthogonal. This can also be expressed as follows. We denote by $\mathbf{e}_{pq}$ $(1 \leqslant p < q \leqslant n)$ the unit coordinate vectors in the space of vectors $\mathbf{a}$ and put

$$\mathbf{E}^{(p,q)} = \mathbf{E} - \mathbf{e}_{pq}\mathbf{e}_{pq}^T,$$

where $\mathbf{E}$ is the unit matrix. ($\mathbf{E}^{(p,q)}$ is the diagonal matrix having a zero in the (p, q) position and ones in all other diagonal positions.) Then we have

LEMMA 1. *For all* (i, j),

$$\mathbf{R}^{(i,j)} = \mathbf{U}^{(i,j)}\mathbf{E}^{(i,j)},$$

where $\mathbf{U}^{(i,j)}$ *is an orthogonal matrix.*

3. THE INDUCTION STEP

We now begin the proof of Theorem 1. It suffices to prove the theorem for the special cyclic ordering by columns. For $n = 2$, $\mathbf{R}$ is the zero matrix of order 1, and $C = 0$. Hence (1.5) is true for matrices of order 2. In order to step from $n - 1$ to n, where $n > 2$, let, for $m = 2, 3, \ldots, n$,

$$\mathbf{R}^{(m)} = \mathbf{R}^{(m-1,m)}\mathbf{R}^{(m-2,m)} \cdots \mathbf{R}^{(1,m)}.$$

Furthermore, let $c \geqslant 0$,

$$c^2 = 1 - \prod_{j=3}^{n-1} \prod_{i=1}^{j-2} \cos^2 \phi_{ij}, \tag{3.1}$$

and denote by $\mathbf{D}$ the diagonal matrix whose first $N - n + 1$ diagonal elements are c and whose remaining diagonal elements are 1.

LEMMA 2. *If Theorem* 1 *is true for matrices of order* $n - 1$, *then*

$$||\mathbf{R}|| \leqslant ||\mathbf{R}^{(n)}\mathbf{D}||. \tag{3.2}$$

Proof. We have

$$\mathbf{R} = \mathbf{R}^{(n)}\mathbf{S}, \tag{3.3}$$

where the matrix

$$\mathbf{S} = \mathbf{R}^{(n-1)}\mathbf{R}^{(n-2)} \cdots \mathbf{R}^{(2)}$$

describes the state of the matrix A before the rotations of the elements in the last column. During the rotations of the columns 2 through $n - 1$ of A, the elements in these columns are transformed exactly as they would be in a complete sweep of the Jacobi method as applied to the matrix A^* obtained from A by deleting its last row and column. The elements in the last column of A, on the other hand, are coupled [1] only among themselves. Consequently, the $(n - 1)$-dimensional subspace corresponding to the elements in the last column is invariant during the first $N^* = N - n + 1$ rotations, and any vector in it is transformed orthogonally. It follows that the matrix $\mathbf{S}$ has the form

$$\mathbf{S} = \begin{pmatrix} \mathbf{R}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{U}^\dagger \end{pmatrix}, \tag{3.4}$$

where $\mathbf{0}$ is a zero matrix, $\mathbf{R}^*$ denotes the Jacobi operator (of order N^*) associated with the matrix A^*, and $\mathbf{U}^\dagger$ is an orthogonal matrix of order $n - 1$.

By the definition of norm,

$$||\mathbf{R}||^2 = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{||\mathbf{R}\mathbf{x}||^2}{||\mathbf{x}||^2} = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{||\mathbf{R}^{(n)}\mathbf{y}||^2}{||\mathbf{x}||^2}, \tag{3.5}$$

where $\mathbf{y} = \mathbf{S}\mathbf{x}$. We partition $\mathbf{x}$ in the form

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}^* \\ \mathbf{x}^\dagger \end{pmatrix},$$

where $\mathbf{x}^*$ comprises the first N^* components of $\mathbf{x}$. If $\mathbf{y}$ is partitioned similarly, then by (3.4)

$$||\mathbf{y}^*|| = ||\mathbf{R}^*\mathbf{x}^*||, \qquad ||\mathbf{y}^\dagger|| = ||\mathbf{x}^\dagger||.$$

If Theorem 1 is true for matrices of order $n - 1$, then

$$||\mathbf{y}^*|| \leqslant C^* ||\mathbf{x}^*||,$$

where $C^* = c$ as defined by (3.1).

We first consider the case where $c \neq 0$. Then

$$\begin{aligned} ||\mathbf{x}||^2 &= ||\mathbf{x}^*||^2 + ||\mathbf{x}^\dagger||^2 \\ &\geqslant c^{-2}||\mathbf{y}^*||^2 + ||\mathbf{y}^\dagger||^2, \end{aligned}$$

and from (3.5) we get

$$||\mathbf{R}||^2 \leqslant \sup_{\mathbf{x} \neq \mathbf{0}} \frac{||\mathbf{R}^{(n)}\mathbf{y}||^2}{c^{-2}||\mathbf{y}^*||^2 + ||\mathbf{y}^\dagger||^2}.$$

The supremum can only be enlarged if $\mathbf{y}$ runs through all vectors $\neq \mathbf{0}$ and not merely those of the form $\mathbf{S}\mathbf{x}$ where $\mathbf{x} \neq \mathbf{0}$. Letting $\mathbf{y} = \mathbf{D}\mathbf{z}$, we have

$$c^{-2}||\mathbf{y}^*||^2 + ||\mathbf{y}^\dagger||^2 = ||\mathbf{z}||^2,$$

and the set of all $\mathbf{y} \neq \mathbf{0}$ is obtained by letting $\mathbf{z}$ run through all nonzero vectors. Thus

$$||\mathbf{R}||^2 \leqslant \sup_{\mathbf{z}\neq\mathbf{0}} \frac{||\mathbf{R}^{(n)}\mathbf{D}\mathbf{z}||^2}{||\mathbf{z}||^2},$$

proving (3.2) for $c \neq 0$.

If $c = 0$, then $\mathbf{y}^* = \mathbf{0}$ for every $\mathbf{x}$, and hence $\mathbf{y} = \mathbf{D}\mathbf{y}$, where the diagonal matrix $\mathbf{D}$ now has zeros in its first N^* positions. Hence (3.5) now may be written

$$||\mathbf{R}||^2 = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{||\mathbf{R}^{(n)}\mathbf{D}\mathbf{y}||^2}{||\mathbf{x}||^2}.$$

Again, $||\mathbf{x}||^2 = ||\mathbf{y}||^2$, and the supremum is enlarged by letting $\mathbf{y}$ run through all vectors $\neq \mathbf{0}$. Hence

$$\mathbf{R}^2 \leqslant \sup_{\mathbf{y}\neq\mathbf{0}} \frac{||\mathbf{R}^{(n)}\mathbf{D}\mathbf{y}||^2}{||\mathbf{y}||^2},$$

whence Lemma 1 again follows.

4. THE MATRIX $\mathbf{R}^{(n)}$

The norm of $\mathbf{R}^{(n)}\mathbf{D}$ will be found by calculating explicitly the eigenvalues of $\mathbf{D}^T\mathbf{R}^{(n)T}\mathbf{R}^{(n)}\mathbf{D}$. In this section we shall determine the matrix $\mathbf{R}^{(n)T}\mathbf{R}^{(n)}$ by expressing the quadratic form

$$Q(\mathbf{x}) = \mathbf{x}^T\mathbf{R}^{(n)T}\mathbf{R}^{(n)}\mathbf{x} \tag{4.1}$$

in terms of the elements of $\mathbf{x}$.

Let V denote the space of vectors $\mathbf{x}$ with components x_{pq} $(1 \leqslant p < q \leqslant n)$. If $\mathbf{x}$ is any vector in V, we let

$$\mathbf{x}^{(m)} = \mathbf{R}^{(m,n)}\mathbf{R}^{(m-1,n)} \cdots \mathbf{R}^{(1,n)}\mathbf{x}.$$

In particular, $\mathbf{R}^{(n)}\mathbf{x} = \mathbf{x}^{(n-1)}$, and hence

$$Q(\mathbf{x}) = ||\mathbf{x}^{(n-1)}||^2. \tag{4.2}$$

For $k = 1, 2, \ldots, n-1$, let V_k denote the k-dimensional subspace of V spanned by the coordinate vectors $\mathbf{e}_{1,k}, \mathbf{e}_{2,k}, \ldots, \mathbf{e}_{k-1,k}; \mathbf{e}_{k,n}$. Clearly, the spaces V_k are mutually orthogonal and together span the whole space V. For any $\mathbf{x} \in V$, let

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_{n-1},$$

where $\mathbf{x}_k \in V_k$ $(k = 1, 2, \ldots, n-1)$. We define

$$\mathbf{x}_k^{(m)} = \mathbf{R}^{(m,n)}\mathbf{R}^{(m-1,n)} \cdots \mathbf{R}^{(1,n)}\mathbf{x}_k,$$

where $m = 0, 1, \ldots, n-1$. It is not asserted that $\mathbf{x}_k^{(m)} \in V_k$. However, we have

Lemma 3. (i) *For every m, $0 \leqslant m \leqslant n-1$, the $n-1$ vectors $\mathbf{x}_k^{(m)}$, $1 \leqslant k \leqslant n-1$, are mutually orthogonal.* (ii) *For $k = 1, 2, \ldots, n-1$,*

$$||\mathbf{x}_k^{(n-1)}||^2 = ||\mathbf{x}_k^{(0)}||^2 - [x_{kn}^{(k-1)}]^2, \tag{4.3}$$

where $x_{kn}^{(k-1)}$ denotes the (k, n) component of $\mathbf{x}^{(k-1)}$.

Proof. In the notation of Lemma 1, let

$$\mathbf{z}_k^{(m)} = \mathbf{E}^{(m+1,n)}\mathbf{x}_k^{(m)}. \tag{4.4}$$

Then for $1 \leqslant k \leqslant n-1$, $0 \leqslant m \leqslant n-1$,

$$\mathbf{x}_k^{(m+1)} = \mathbf{U}^{(m+1,n)}\mathbf{z}_k^{(m)}, \tag{4.5}$$

where $\mathbf{U}^{(m+1,n)}$ is orthogonal.

We now prove (i) by induction with respect to m. Clearly, the vectors $\mathbf{x}_k^{(0)} = \mathbf{x}_k$, $1 \leqslant k \leqslant n-1$, are mutually orthogonal. Assume now that the vectors $x_k^{(m)}$ are mutually orthogonal for some $m \geqslant 0$. It follows from the definition of $\mathbf{R}^{(p,n)}$ that

$$\mathbf{x}_1^{(m)}, \mathbf{x}_2^{(m)}, \ldots, \mathbf{x}_m^{(m)} \in V_1 \oplus V_2 \oplus \cdots \oplus V_m, \tag{4.6}$$

$$\mathbf{x}_k^{(m)} \in V_k, \qquad m+1 \leqslant k \leqslant n-1. \tag{4.7}$$

Of all vectors $\mathbf{x}_k^{(m)}$, $\mathbf{x}_{m+1}^{(m)}$ is the only one having a nonzero component in the direction of $\mathbf{e}_{m+1,n}$. Hence by (4.4),

$$\mathbf{z}_k^{(m)} = \mathbf{E}^{(m+1,n)}\mathbf{x}_k^{(m)} = \mathbf{x}_k^{(m)}, \qquad k \neq m+1. \tag{4.8}$$

Since

$$\mathbf{x}^{(m)} = \mathbf{x}_1^{(m)} + \mathbf{x}_2^{(m)} + \cdots + \mathbf{x}_{n-1}^{(m)}, \tag{4.9}$$

the remaining relation can be written

$$\mathbf{z}^{(m)}_{m+1} = \mathbf{x}^{(m)}_{m+1} - x^{(m)}_{m+1,n}\mathbf{e}_{m+1,n}. \tag{4.10}$$

By (4.8), the $\mathbf{z}_k^{(m)}$ are orthogonal for $k \neq m+1$. By (4.7), $\mathbf{x}^{(m)}_{m+1} \in V_{m+1}$, and by (4.9) the same is true for $\mathbf{z}^{(m)}_{m+1}$. Thus all $n-1$ vectors $\mathbf{z}_k^{(m)}$, $1 \leqslant k \leqslant n-1$, belong to $n-1$ different orthogonal subspaces, and hence are orthogonal. The vectors $\mathbf{x}_k^{(m+1)}$ result from the $\mathbf{z}_k^{(m)}$ by the orthogonal transformation (4.5) and thus are likewise orthogonal, which proves (i).

To prove (ii), we apply (4.8) and (4.5) for fixed k and all admissible m. There follows

$$||\mathbf{x}_k^{(m+1)}||^2 = ||\mathbf{x}_k^{(m)}||^2 \qquad \text{for} \quad 0 \leqslant m \leqslant k-1 \quad \text{and} \quad k \leqslant m \leqslant n-1.$$

On the other hand, by (4.10) and (4.5),

$$||\mathbf{x}_k^{(k)}||^2 = ||\mathbf{x}_k^{(k-1)}||^2 - (x^{(k-1)}_{kn})^2.$$

The last two relations imply (4.3). The proof of Lemma 3 is complete.

By assertion (i) and by (4.9),

$$||\mathbf{x}^{(m)}||^2 = \sum_{k=1}^{n-1} ||\mathbf{x}_k^{(m)}||^2, \qquad m = 0, 1, \ldots, n-1.$$

By summing (4.3) with respect to k we thus get

$$||\mathbf{x}^{(n-1)}||^2 - ||\mathbf{x}^{(0)}||^2 = -\sum_{k=1}^{n-1} (x^{(k-1)}_{kn})^2. \tag{4.11}$$

In order to find $Q(\mathbf{x})$, it thus remains only to express $x^{(k-1)}_{kn}$ in terms of the components of $\mathbf{x}^{(0)} = \mathbf{x}$.

We write $\cos\phi_{kn} = c_k$, $\sin\phi_{kn} = s_k$, $k = 1, 2, \ldots, n-1$. According to Section 2, we then have

$$x^{(0)}_{pn} = x_{pn}, \qquad 1 \leqslant p \leqslant n-1,$$

and generally for $1 \leqslant k \leqslant n-1$

$$x^{(k)}_{pn} = c_k x^{(k-1)}_{pn} - s_k x_{kp}, \qquad k < p \leqslant n-1.$$

Hence it follows easily by induction that

$$x_{kn}^{(k-1)} = c_1 c_2 \cdots c_{k-1} x_{kn} - \sum_{p=1}^{k-1} s_p c_{p+1} c_{p+2} \cdots c_{k-1} x_{pk}, \tag{4.12}$$

$$k = 2, 3, \ldots, n-1$$

(empty products are 1).

From (4.2) and (4.11) we now have

$$Q(\mathbf{x}) = ||\mathbf{x}||^2 - \sum_{k=1}^{n-1} (x_{kn}^{(k-1)})^2, \tag{4.13}$$

where $x_{kn}^{(k-1)}$ is given by (4.12). Our final task is to determine the eigenvalues of the matrix of the quadratic form

$$Q(\mathbf{Dx}) = \mathbf{x}^T \mathbf{D}^T \mathbf{R}^{(n)T} \mathbf{R}^{(n)} \mathbf{Dx},$$

which is obtained from $Q(\mathbf{x})$ by multiplying all x_{pq} where $q < n$ by the constant c.

5. CALCULATION OF EIGENVALUES

The representation (4.13) together with (4.12) shows that

$$Q(\mathbf{Dx}) = \sum_{k=1}^{n-1} Q_k(\mathbf{x}),$$

where

$$Q_k(\mathbf{x}) = x_{kn}^2 + c^2 \sum_{p=1}^{k-1} x_{pk}^2 - (c_1 c_2 \cdots c_{k-1} x_{kn} - \sum_{p=1}^{k-1} c s_p c_{p+1} \cdots c_{k-1} x_{pk})^2,$$

$$k = 1, \ldots, n-1. \tag{5.1}$$

The form Q_k depends only on the variables $x_{1k}, \ldots, x_{k-1,k}, x_{kn}$. Since each of the Q_k depends on a different set of variables, the set of eigenvalues of $Q(\mathbf{Dx})$ is the union of the sets of eigenvalues of the forms Q_k.

Let $Q^{(k)} = (q_{st}^{(k)})$ be the matrix of Q_k. (We return to simple indexing.) $Q^{(k)}$ is a symmetric matrix of order k. It follows from (5.1) that for a suitable numbering of the elements,

$$q_{tt}^{(k)} = c^2(1 - s_t^2 c_{t+1}^2 c_{t+2}^2 \cdots c_{k-1}^2), \qquad 1 \leqslant t \leqslant k-1,$$

$$q_{kk}^{(k)} = 1 - c_1^2 c_2^2 \cdots c_{k-1}^2,$$

$$q_{st}^{(k)} = -c^2 s_s c_{s+1} \cdots c_{k-1} s_t c_{t+1} \cdots c_{k-1}, \qquad 1 \leqslant s < t \leqslant k-1,$$

$$q_{tk}^{(k)} = c c_1 c_2 \cdots c_{k-1} s_t c_{t+1} \cdots c_{k-1}, \qquad 1 \leqslant t \leqslant k-1.$$

LEMMA 4. *The eigenvalues of* Q_k *are*

$$\lambda_1 = 0;$$

$$\lambda_2 = \lambda_3 = \cdots = \lambda_{k-1} = c^2;$$

$$\lambda_k = 1 - (1 - c^2) c_1{}^2 c_2{}^2 \cdots c_{k-1}^2.$$

Proof. To show that 0 is an eigenvalue, we show that the rows of $Q^{(k)}$ are linearly dependent. To this end we multiply the tth column by $s_t c_{t+1} \cdots c_{k-1}$ $(1 \leqslant t \leqslant k-1)$, and the kth column by $-c c_1 c_2 \cdots c_{k-1}$. The nondiagonal elements of the sth row are then

(a) for $1 \leqslant s \leqslant k-1$,

$$-c^2 s_s c_{s+1} \cdots c_{k-1} (s_t c_{t+1} \cdots c_{k-1})^2, \qquad 1 \leqslant t \leqslant k-1, \quad t \neq s,$$

$$-c^2 s_s c_{s+1} \cdots c_{k-1} (c_1 c_2 \cdots c_{k-1})^2, \qquad t = k;$$

(b) for $s = k$,

$$c c_1 c_2 \cdots c_k (s_t c_{t+1} \cdots c_{k-1})^2, \qquad 1 \leqslant t \leqslant k-1.$$

Using the identity

$$(c_1 c_2 \cdots c_{k-1})^2 + \sum_{t=1}^{k-1} (s_t c_{t+1} \cdots c_{k-1})^2 = 1, \tag{5.2}$$

we thus find for the sum of the nondiagonal elements in the sth row

$$-c^2 s_s c_{s+1} \cdots c_{k-1} [(s_s c_{s+1} \cdots c_{k-1})^2 - 1], \qquad 1 \leqslant s \leqslant k-1,$$

and

$$c^2 c_1 c_2 \cdots c_{k-1} [1 - (c_1 c_2 \cdots c_{k-1})^2], \qquad s = k.$$

These are just the negatives of the diagonal elements of the modified matrix, proving that the sum of its rows is zero.

To show that c^2 is an eigenvalue of multiplicity $\geqslant k-2$, we show that the rank of the matrix $Q_k - c^2 I_k$ (I_k = k-dimensional unit matrix) is at most 2. Indeed the tth column of this matrix $(1 \leqslant t \leqslant k-1)$ is $s_t c_{t+1} \cdots c_{k-1}$ times the vector

$$\begin{pmatrix} -c^2 s_1 c_2 \cdots c_{k-1} \\ -c^2 s_2 c_3 \cdots c_{k-1} \\ \cdots \\ -c^2 s_{k-1} \\ c c_1 c_2 \cdots c_{k-1} \end{pmatrix};$$

thus the first $k-1$ columns are all proportional to the same vector, and the matrix contains at most two linearly independent columns. By a familiar fact from linear algebra (see, e.g., Theorem 7.6.1 of [4]) it follows that c^2 is an eigenvalue of multiplicity $\geqslant k-2$ of $Q^{(k)}$.

To determine the remaining eigenvalue, we use the fact that the trace of a matrix equals the sum of its eigenvalues. Using (5.2), the trace of $Q^{(k)}$ is

$$\begin{aligned}(k-1)c^2 - c^2[1-(c_1c_2\cdots c_{k-1})^2] + 1 - (c_1c_2\cdots c_{k-1})^2 \\ = (k-2)c^2 + 1 - (1-c^2)(c_1c_2\cdots c_{k-1})^2.\end{aligned}$$

The sum of the $k-1$ eigenvalues already found is $(k-2)c^2$; hence it follows that the remaining eigenvalue is $1-(1-c^2)(c_1c_2\cdots c_{k-1})^2$, proving Lemma 4.

In view of the definitions (3.1) of c and of the c_s, the largest eigenvalue of all $Q^{(k)}$ $(k = 1, 2, \ldots, n-1)$ is

$$\begin{aligned}\lambda &= 1-(1-c^2)c_1{}^2c_2{}^2\cdots c_{n-2}^2 \\ &= 1 - \prod_{j=3}^{n}\prod_{i=1}^{j-2}\cos^2\varphi_{ij}.\end{aligned}$$

The theorem now follows in view of Lemma 2 and the fact that $\|\mathbf{R}^{(m)}\mathbf{D}\| = \lambda^{1/2}$.

REFERENCES

1 G. E. Forsythe and P. Henrici, The cyclic Jacobi method for computing the principal values of a complex matrix. *Trans. Amer. Math. Soc.* **94**(1960), 1–23.

2 E. R. Hansen, On cyclic Jacobi methods, *J. Soc. Indust. Appl. Math.* **11**(1963), 448–459.

3 P. Henrici, On the speed of convergence of cyclic and quasicyclic Jacobi methods for computing the eigenvalues of hermitian matrices, *J. Soc. Indust. Appl. Math.* **6**(1958), 144–162.
4 L. Mirsky, *An Introduction to Linear Algebra*, Clarendon Press, Oxford, 1961.
5 A. Schoenhage, Zur Konvergenz des Jacobi-Verfahrens, *Num. Math.* **3**(1961), 374–380.
6 G. Schröder, Über die Konvergenz einiger Jacobi-Verfahren zur Bestimmung der Eigenwerte symmetrischer Matrizen, *Forschungsberichte des Landes Nordrhein-Westfalen* No. 1291, Westdeutscher Verlag, Köln und Opladen, 1964.
7 J. H. Wilkinson, Note on the quadratic convergence of the cyclic Jacobi process, *Num. Math.* **4**(1962), 296–300.

Received April 15, 1968