

Bounds for Eigenvalues of Tridiagonal Symmetric Matrices Computed by the LR Method

By Gene H. Golub

1. Introduction. In recent years, a number of methods have been proposed for finding the eigenvalues of real, symmetric matrices. The methods of Lanczos [8], Givens [3], and Householder [5, 14] reduce the original matrix to a tridiagonal matrix whose eigenvalues are the same as those of the original matrix. The problem reduces then to finding the eigenvalues of a tridiagonal form. Givens has suggested the use of Sturm sequences, and others have used Muller's method [9].

In this paper, Rutishauser's LR algorithm and its variants [10, 11] will be considered for tridiagonal symmetric matrices. Henrici [4] has shown that for this case the LR algorithm is equivalent to the QD algorithm. It will be shown that *during* the iteration procedure, it is possible to determine bounds on the eigenvalues. Wilkinson [12] has recently considered the problem of determining rigorous error bounds after the eigensystem has been computed.

2. An Inclusion Theorem. An inclusion theorem is one which exhibits a set known to contain at least one eigenvalue. A general discussion of such theorems is given by Bauer and Householder [1]. We shall now derive an inclusion theorem with the aid of the Lanczos algorithm.

The Lanczos algorithm for reducing a symmetric matrix A to tridiagonal form is as follows:

Let

$$\mathbf{x}_0 = \mathbf{0}, \quad \text{the null vector,}$$

and

$$\mathbf{x}_1 = \mathbf{x}, \quad \text{a non-null arbitrary vector,}$$

then

$$(1) \quad \mathbf{x}_{k+1} = A\mathbf{x}_k - \alpha_k\mathbf{x}_k - \beta_{k-1}\mathbf{x}_{k-1} \quad k = 1, 2, \dots, n-1$$

with

$$(2) \quad \alpha_k = \frac{\mathbf{x}_k' A \mathbf{x}_k}{\mathbf{x}_k' \mathbf{x}_k}$$

and

$$(3) \quad \beta_{k-1} = \frac{\mathbf{x}_k' A \mathbf{x}_{k-1}}{\mathbf{x}_{k-1}' \mathbf{x}_{k-1}}.$$

It is not difficult to show [cf. 12] that

$$(4) \quad \mathbf{x}_i' \mathbf{x}_j = 0 \quad \text{for } i \neq j$$

Received February 22, 1962.

and hence

$$\mathbf{x}_{n+1} = \mathbf{0}.$$

If \mathbf{x} lies in a subspace generated by p eigenvectors of A , then

$$\mathbf{x}_{p+1} = \mathbf{0}.$$

Then since

$$\begin{aligned} \beta_{k-1} &= \frac{\mathbf{x}_k' A \mathbf{x}_{k-1}}{\mathbf{x}_{k-1}' \mathbf{x}_{k-1}} \\ (5) \quad &= \frac{\mathbf{x}_k' (\mathbf{x}_k + \alpha_{k-1} \mathbf{x}_{k-1} + \beta_{k-2} \mathbf{x}_{k-2})}{\mathbf{x}_{k-1}' \mathbf{x}_{k-1}} \\ &= \frac{\mathbf{x}_k' \mathbf{x}_k}{\mathbf{x}_{k-1}' \mathbf{x}_{k-1}} \geq 0. \end{aligned}$$

Now

$$(A - \alpha_k I) \mathbf{x}_k = \mathbf{x}_{k+1} + \beta_{k-1} \mathbf{x}_{k-1}$$

and hence by (4) and (5),

$$\mathbf{x}_k' (A - \alpha_k I)^2 \mathbf{x}_k = \mathbf{x}_{k+1}' \mathbf{x}_{k+1} + \frac{(\mathbf{x}_k' \mathbf{x}_k)^2}{\mathbf{x}_{k-1}' \mathbf{x}_{k-1}}.$$

Thus

$$(6) \quad \frac{\mathbf{x}_k' (A - \alpha_k I)^2 \mathbf{x}_k}{\mathbf{x}_k' \mathbf{x}_k} = \beta_k + \beta_{k-1} = \sigma_k^2$$

with

$$\beta_0 = 0, \quad \beta_n = 0.$$

The left hand side of (6) is simply a Rayleigh quotient and as such must lie between the smallest and largest eigenvalue of $(A - \alpha_k I)^2$. Consequently,

$$\min_i (\lambda_i - \alpha_k)^2 \leq \sigma_k^2 \leq \max_i (\lambda_i - \alpha_k)^2$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A .

THEOREM 1. *Let A be a symmetric matrix. Then there is an eigenvalue of A in the interval*

$$(7) \quad \alpha_k - \sigma_k \leq \lambda \leq \alpha_k + \sigma_k.$$

If A is a symmetric tridiagonal matrix, then for

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

it is not difficult to compute the coefficients α_k and β_k .

COROLLARY 1.1. *Let*

$$\begin{aligned} a_{ij} &= a_i \quad \text{for } i = j \\ &= b_m \quad \text{for } |i - j| = 1, \quad m = \min(i, j) \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

Then for the Lanczos algorithm,

$$\alpha_k = a_k,$$

and

$$\beta_k = b_k^2$$

and consequently, the interval

$$(8) \quad a_k - \sigma_k \leq \lambda \leq a_k + \sigma_k$$

where

$$\sigma_k^2 = b_k^2 + b_{k-1}^2$$

contains at least one eigenvalue.

Corollary 1.1 is a special case of a known theorem [cf. 2, p. 128]. Note that if the intervals are non-overlapping then the bounds given by (8) are smaller than those obtained by the Gershgorin theorem.

If the intervals are non-overlapping, it is possible to obtain improved bounds on the eigenvalues by using the bounds of Kohn [7] and Kato [6].

THEOREM 2. *Let*

$$\lambda_j^l \leq \lambda_j \leq \lambda_j^u \quad j = 1, \dots, n.$$

Then if

$$\lambda_{j-1} < \lambda_j < \lambda_{j+1}$$

and if

$$\lambda_{j-1}^u < \alpha_k < \lambda_{j+1}^l,$$

then

$$(9) \quad \alpha_k - \frac{\sigma_k^2}{\lambda_{j+1}^l - \alpha_k} \leq \lambda_j \leq \alpha_k + \frac{\sigma_k^2}{\alpha_k - \lambda_{j-1}^u} \quad j = 2, \dots, n - 1.$$

Since the coefficients α_k are Rayleigh quotients,

$$\lambda_1 \leq \min_k \alpha_k$$

and

$$\lambda_n \geq \max_k \alpha_k$$

In Section 3, the bounds of Theorem 1 will be used. However, the bounds of Theorem 2 will be equally applicable.

3. Application to the LR Algorithm. The LR algorithm proceeds as follows: Begin with the given matrix $A = A_0$, compute a triangular decomposition

$$A_0 = L_0 R_0,$$

and then multiply the matrices in reverse order so that

$$A_1 = R_0 L_0.$$

Now one treats the matrix A_1 in the same fashion as the matrix A_0 , and a sequence of matrices is obtained by continuing *ad infinitum*.

Under certain conditions [10], it is known that A_i converges for $i \rightarrow \infty$ to a lower triangular matrix with the eigenvalues on the diagonal.

Let

$$A_i = \begin{pmatrix} \alpha_1^{(i)} & \beta_1^{(i)} & & & & \circ \\ 1 & \alpha_2^{(i)} & \beta_2^{(i)} & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \beta_{n-1}^{(i)} \\ \circ & & & & 1 & \alpha_n^{(i)} \end{pmatrix}.$$

If the Gauss-Banachiewicz procedure is used to decompose A_i , then

$$A_i = L_i R_i$$

with

$$L_i = \begin{pmatrix} q_1^{(i)} & & & & & \circ \\ 1 & q_2^{(i)} & & & & \\ & \cdot & \cdot & & & \\ & & \cdot & \cdot & & \\ & & & \cdot & \cdot & \\ \circ & & & & 1 & q_n^{(i)} \end{pmatrix},$$

$$R_i = \begin{pmatrix} 1 & e_1^{(i)} & & & & \circ \\ & 1 & \cdot & & & \\ & & \cdot & \cdot & & \\ & & & \cdot & \cdot & \\ & & & & \cdot & \\ & & & & & e_{n-1}^{(i)} \\ \circ & & & & & 1 \end{pmatrix}$$

where

$$q_1^{(i)} = \alpha_1^{(i)}$$

$$e_k^{(i)} = \beta_k^{(i)} / q_k^{(i)}$$

$$q_k^{(i)} = \alpha_k^{(i)} - e_{k-1}^{(i)}.$$

Then since

$$A_{i+1} = R_i \cdot L_i$$

$$(10) \quad \begin{aligned} \alpha_k^{(i+1)} &= e_k^{(i)} + q_k^{(i)} & k = 1, \dots, n-1 \\ \alpha_n^{(i+1)} &= q_n^{(i)} \\ \beta_k^{(i+1)} &= e_k^{(i)} \cdot q_k^{(i)}. \end{aligned}$$

If the Lanczos algorithm is performed not with the single vector $\mathbf{x}_1 = \mathbf{x}$ but with the infinite set of initial vectors $\mathbf{x}_1^{(i)} = A^i \mathbf{x} (i = 0, 1, \dots)$, then it has been shown by Henrici [4] that the coefficients α_k and β_k may be replaced by $\alpha_k^{(i)}$ and $\beta_k^{(i)}$ and that these are the same as those given by (10). Consequently, Theorem 1 may be applied to the LR algorithm.

THEOREM 3. *If the LR algorithm is used with the Gauss-Banachiewicz decomposition then there is an eigenvalue in the interval*

$$\alpha_k^{(i)} - \sigma_k^{(i)} \leq \lambda \leq \alpha_k^{(i)} + \sigma_k^{(i)}$$

where

$$\begin{aligned} (\sigma_k^{(i)})^2 &= \beta_k^{(i)} + \beta_{k-1}^{(i)} \\ \beta_0^{(i)} &= \beta_n^{(i)} = 0. \end{aligned}$$

Since the coefficients $\beta_k^{(i)}$ will become very small as i increases, it will also be possible to use the bounds of Theorem 2. Theorem 3 enables one to stop the iteration procedure when the eigenvalues have attained a predetermined accuracy.

The LR algorithm converges linearly. Rutishauser [11] has shown it is possible to attain cubic convergence if A is positive definite and symmetric. The increase in the rate of convergence comes about by working with the matrix

$$A_i - y_i I$$

and using the Choleski decomposition. In what follows $y_i = 0$ for all i since for $y_i \neq 0$ the eigenvalues are translated, and this in no way changes the argument.

Let

$$A_i = \begin{pmatrix} a_1^{(i)} & b_1^{(i)} & & & & & \circ \\ b_1^{(i)} & a_2^{(i)} & & & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & \cdot & \cdot & b_{n-1}^{(i)} \\ \circ & & & & & b_{n-1}^{(i)} & a_n^{(i)} \end{pmatrix}.$$

Then for the Choleski decomposition,

$$A_i = R_i^T R_i$$

with

$$R_i = \begin{pmatrix} p_1^{(i)} & d_1^{(i)} & & & & & \circ \\ & p_2^{(i)} & \cdot & & & & \\ & & \cdot & \cdot & & & \\ & & & \cdot & \cdot & & \\ & & & & \cdot & \cdot & \\ & & & & & \cdot & d_{n-1}^{(i)} \\ \circ & & & & & & p_n^{(i)} \end{pmatrix}$$

where

$$\begin{aligned}(p_1^{(i)})^2 &= a_1^{(i)} \\ (d_k^{(i)})^2 &= (b_k^{(i)})^2 / (p_k^{(i)})^2 \\ (p_k^{(i)})^2 &= a_k^{(i)} - (d_{k-1}^{(i)})^2.\end{aligned}$$

Then since

$$\begin{aligned}A_{i+1} &= R_i R_i^T \\ \alpha_k^{(i+1)} &= (p_k^{(i)})^2 + (d_k^{(i)})^2 \quad k = 1, \dots, n-1 \\ \alpha_n^{(i+1)} &= (p_n^{(i)})^2, \\ (b_k^{(i+1)})^2 &= (d_k^{(i)})^2 \cdot (p_{k+1}^{(i)})^2.\end{aligned}$$

Note that although a Choleski decomposition of A_i was made, it was *not* necessary to compute any square roots. This observation is due to Professor H. Kaiser* of the University of Illinois. Since each A_i is a tridiagonal matrix, Corollary 1.1 is applicable.

THEOREM 4. *If the LR algorithm is used with the Choleski decomposition then there is an eigenvalue in the interval*

$$a_k^{(i)} - \sigma_k^{(i)} \leq \lambda \leq a_k^{(i)} + \sigma_k^{(i)}$$

where

$$\begin{aligned}(\sigma_k^{(i)})^2 &= (b_k^{(i)})^2 + (b_{k-1}^{(i)})^2 \\ b_0^{(i)} &= b_n^{(i)} = 0.\end{aligned}$$

4. A Numerical Example. Consider the tridiagonal symmetric matrix A with

$$\begin{aligned}a_k &= 1 \quad k = 1, \dots, 5 \\ b_k &= \frac{1}{2} \quad k = 1, \dots, 4.\end{aligned}$$

When the Lanczos scheme is applied to this matrix,

$$\begin{aligned}\alpha_k &= 1, \quad k = 1, \dots, 5 \\ \beta_k &= \frac{1}{4} \quad k = 1, \dots, 4.\end{aligned}$$

The eigenvalues of A are

$$\begin{aligned}\lambda_5 &= 1 + \frac{\sqrt{3}}{2} \doteq 1.8660254038 \\ \lambda_4 &= 1.5 \\ \lambda_3 &= 1.0 \\ \lambda_2 &= 0.5 \\ \lambda_1 &= 1 - \frac{\sqrt{3}}{2} \doteq 0.1339745962.\end{aligned}$$

* Personal communication.

TABLE 1

$i \backslash k$	5		15	
	$\alpha_k^{(i)}$	$\sigma_k^{(i)}$	$\alpha_k^{(i)}$	$\sigma_k^{(i)}$
1	.13566009	2.676×10^{-2}	.13397460	3.255×10^{-5}
2	.64318606	2.895×10^{-1}	.50011486	7.611×10^{-3}
3	1.2211538	4.520×10^{-1}	1.0088091	6.813×10^{-2}
4	1.3750000	4.346×10^{-1}	1.5286215	1.306×10^{-1}
5	1.6250000	2.602×10^{-1}	1.8284798	1.117×10^{-1}

TABLE 2

k	Lower Bounds for $\lambda_k - \alpha_k^{(i)}$	Upper Bounds for $\lambda_k - \alpha_k^{(i)}$
1	-2.955×10^{-9}	0
2	-1.315×10^{-4}	1.582×10^{-4}
3	-1.192×10^{-2}	9.263×10^{-3}
4	-9.060×10^{-2}	3.775×10^{-2}
5	0	7.365×10^{-2}

The LR algorithm with the Gauss-Banachiewicz decomposition was applied to A. The results after five and 15 iterations are given in Table 1. Since the intervals still overlapped after five iterations, it was not possible to apply the bounds of Theorem 2. However, after fifteen iterations the intervals were non-overlapping, and the bounds for $\lambda_k - \alpha_k^{(i)}$ are given in Table 2. Note that

$$-2.955 \times 10^{-9} \leq \lambda_1 - \alpha_1^{(15)} \leq 0.$$

and thus at least seven places of $\alpha_1^{(15)}$ are correct.

5. Acknowledgement. The author wishes to thank Professor P. Henrici of the University of California, Los Angeles, for his comments and Mr. James Ayers of Space Technology Laboratories, Inc. for performing the calculations given in Section 4.

Space Technology Laboratories, Inc.
Redondo Beach, California

1. F. L. BAUER & A. S. HOUSEHOLDER, "Moments and characteristic roots," *Numer. Math.*, v. 2, 1960, p. 42-53.
2. KY FAN & A. J. HOFFMAN, "Lower bounds for the rank and location of the eigenvalues of a matrix," *Nat. Bur. Standards Appl. Math. Ser.* 39, 1954, p. 117-130.
3. W. GIVENS, "Numerical computation of the characteristic values of a real symmetric matrix," Oak Ridge National Laboratory, *Report No.* 1574, 1954.
4. P. HENRICI, "The quotient difference algorithm," *Nat. Bur. Standards Appl. Math. Ser.* 49, 1958, p. 23-46.
5. A. S. HOUSEHOLDER & F. L. BAUER, "On certain methods for expanding the characteristic polynomial," *Numer. Math.*, v. 1, 1959, p. 29-37.
6. T. KATO, "On the upper and lower bounds of eigenvalues," *J. Phys. Soc. Japan*, v. 4, 1949, p. 334-339.
7. W. KOHN, "A note on Weinstein's variational method," *Phys. Rev.*, v. 71, 1947, p. 902-904.
8. C. LANCZOS, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators," *J. Res. Nat. Bur. Standards*, v. 45, 1950, p. 255-282.

9. D. MULLER, "A method for solving algebraic equations using an automatic computer," *MTAC*, v. 10, 1956, p. 208-215.
10. H. RUTISHAUSER, "Solution of eigenvalue problems with the LR-transformation," *Nat. Bur. Standards Appl. Math. Ser.* 49, 1958, p. 47-81.
11. H. RUTISHAUSER "Über eine kubisch konvergente Variante der LR-Transformation," *Z. Angew. Math. Mech.*, v. 40, 1960, p. 49-54.
12. J. WILKINSON, "The calculation of eigenvectors by the method of Lanczos," *Comput. J.*, v. 1, 1958, p. 148-152.
13. J. WILKINSON, "Rigorous error bounds for computed eigensystems," *Comput. J.*, v. 4, 1961, p. 230-241.
14. J. WILKINSON, "Householder's method for the solution of the algebraic eigenproblem," *Comput. J.*, v. 3, 1960, p. 23-28.