On Sturm Sequences for Tridiagonal Matrices*

J. M. ORTEGA†

Stanford University, Stanford, California

1. *Introduction*

One of the leading methods for computing the eigenvalues of a real symmetric matrix is that of Givens [1]. In that method, after transforming the matrix to a triple-diagonal form S, one isolates the eigenvalues by using the fact that, roughly speaking, the leading principal minors of $S - \lambda I$ form a Sturm sequence. However, the classical theory of a Sturm sequence, expounded in [2], needs some extension to give signs to zero values in the sequence. We have noticed that the extension of Givens in the text of [1] is not quite correct. The difficulty is a purely algebraic one and has nothing to do with the digital realization on a computer. Professor Givens [personal statement] concurs in this, but states that the machine codes in [1] are correct.

The theorem given below gives a correct extension of a Sturm sequence.

2. *Extension of the Sturm Sequence*

DEFINITION 1. To any expression $f_i(\lambda)$ defined below ($i = 0, 1, \cdots, n$; λ a real number), we attach a unique sign, $\mathrm{SIG}[f_i(\lambda)]$, defined recursively as follows:

$$\mathrm{SIG}[f_i(\lambda)] = \begin{cases} +1 & \text{if } f_i(\lambda) > 0, \\ -1 & \text{if } f_i(\lambda) < 0, \\ \mathrm{SIG}[f_{i-1}(\lambda)] & \text{if } f_i(\lambda) = 0. \end{cases}$$

Thus, if $f_0(\lambda) \neq 0$, $\mathrm{SIG}[f_i(\lambda)]$ is well-defined.

DEFINITION 2. We denote by $A(\lambda)$ the number of agreements in sign of the sequence $\{f_i(\lambda)\}$. I.e., $A(\lambda)$ is the number of values of i ($i = 1, \cdots, n$) for which $\mathrm{SIG}[f_i(\lambda)] \cdot \mathrm{SIG}[f_{i-1}(\lambda)] = +1$.

THEOREM. *Let $S = (s_{i,j})$ be a real $n \times n$ symmetric matrix in triple-diagonal form; i.e., $s_{i,i} = a_i$, $s_{i,i+1} = s_{i+1,i} = b_i$, and $s_{i,j} = 0$ for $|i - j| > 1$. Let $b_0 = 0$ and define a sequence f_i ($i = 0, 1, \cdots, n$) of functions of a real variable λ as follows: $f_0(\lambda) \equiv 1$ and, for $i = 1, \cdots, n$,*

$$(1) \quad f_i(\lambda) = \begin{cases} (a_i - \lambda)\mathrm{SIG}[f_{i-1}(\lambda)], & \textit{if } \; b_{i-1} = 0; \\ (a_i - \lambda)f_{i-1}(\lambda) - b_{i-1}^2\mathrm{SIG}[f_{i-2}(\lambda)], & \textit{if } \; b_{i-2} = 0 \\ & \quad \textit{and } \; b_{i-1} \neq 0; \\ (a_i - \lambda)f_{i-1}(\lambda) - b_{i-1}^2 f_{i-2}(\lambda), & \textit{otherwise}. \end{cases}$$

* Received February, 1960.

† The author wishes to thank Prof. George E. Forsythe of Stanford University for his helpful criticism during the preparation of this paper. This work was supported by the Office of Naval Research under Contract Nonr-225(37).

Then, for any λ, the number of eigenvalues of S that are greater than or equal to λ is given by $A(\lambda)$.

NOTE. The statement and conclusion of the theorem are those of Givens ([1], p. 16), except for the following:

I. Givens makes $\mathrm{SIG}[f_i(\lambda)] = +1$ whenever $f_i(\lambda) = 0$.

II. Our algorithm (1) for computing the sequence $\{f_i(\lambda)\}$ has replaced Givens' algorithm:

$$(2)\quad f_i(\lambda) = \begin{cases} (a_i - \lambda)f_{i-1}(\lambda) - b_{i-1}^2 f_{i-2}(\lambda), & \text{if } f_{i-1}(\lambda) \neq 0; \\ -b_{i-1}^2 f_{i-2}(\lambda), & \text{if } f_{i-1}(\lambda) = 0, f_{i-2}(\lambda) \neq 0 \\ & \quad \text{and } b_{i-1} \neq 0; \\ -b_{i-1}^2, & \text{if } f_{i-1}(\lambda) = 0, f_{i-2}(\lambda) = 0, \\ & \quad \text{and } b_{i-1} \neq 0; \\ a_i - \lambda, & \text{if } f_{i-1}(\lambda) = 0 \text{ and } b_{i-1} = 0. \end{cases}$$

A restatement of the theorem was rendered necessary by the discovery of two classes of counterexamples, of which the following are illustrative.

1. With Givens' sign convention, the theorem fails at $\lambda = 2$ for the matrix

$$S = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

since $f_0(2) = 1$, $f_1(2) = -1$ and $f_2(2) = 0$. Thus $A(2) = 0$ although $\lambda = 2$ is an eigenvalue of S.

2. Applying the algorithm (2) to the matrix

$$S = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

gives the sequence

$$f_0(2) = 1, \quad f_1(2) = -1, \quad f_2(2) = 0, \quad f_3(2) = -1, \quad f_4(2) = 0.$$

If Givens' sign convention is used we have $A(2) = 0$, whereas with our sign convention we have $A(2) = 3$. In either case the result is incorrect, since the eigenvalues of S are 0, 0, 2, 2.

3. *Proof of the Theorem*

CASE I. *No $b_i = 0$.*

We first prove the following properties of the sequence $\{f_i(\lambda)\}$:

(a) Two consecutive $f_i(\lambda)$ cannot both vanish for the same value of λ.

(b) If $f_i(\lambda) = 0$, then $f_{i-1}(\lambda)f_{i+1}(\lambda) < 0$ $(i = 1, \cdots, n - 1)$.

(c) f_n has no multiple root.

Since, by definition, $f_0(\lambda)$ and $f_1(\lambda)$ are not both zero, property (a) follows immediately by induction, using (1).

Property (b) follows from property (a), since, if $f_i(\lambda) = 0$, by (1) we have $f_{i+1}(\lambda) = -b_i^2 f_{i-1}(\lambda)$ and neither $f_{i+1}(\lambda)$ nor $f_{i-1}(\lambda)$ is zero.

To prove property (c), we note that $f_i(\lambda)$ is the determinant of the ith leading principal minor of the matrix $S - \lambda I$. (Since no b_i is zero, all of the f_i are polynomials in λ.) If $f_n(\lambda) = 0$ then, by (a), $f_{n-1}(\lambda) \neq 0$. Thus $S - \lambda I$ is of rank $n - 1$ and, since S is symmetric, this guarantees that λ is a simple eigenvalue of S and hence a simple root of f_n.

By properties (a), (b), and (c) we have for any λ not a root of f_n that

$$(-1)^n f_n(\lambda), (-1)^{n-1} f_{n-1}(\lambda), \cdots, f_0(\lambda) \tag{3}$$

is a classical Sturm sequence [2]. Hence the number of roots of f_n greater than λ is the number of variations in sign of the sequence (3). But this is identical to the number of agreements in sign of the sequence computed by (1), i.e., to $A(\lambda)$.

If λ_0 is a root of f_n, then by (a) there is an ϵ-neighborhood of λ_0, $N(\lambda_0, \epsilon)$, such that $f_{n-1}(\lambda) \neq 0$ for $\lambda \in N(\lambda_0 ; \epsilon)$. Thus $\mathrm{SIG}[f_{n-1}(\lambda)]$ is constant for $\lambda \in N(\lambda_0 ; \epsilon)$; also, the number of sign agreements of the sequence $f_{n-1}(\lambda), \cdots, f_0(\lambda)$ remains constant for $\lambda \in N(\lambda_0 ; \epsilon)$. But, since λ_0 is a simple root of f_n, we must have $A(\lambda_0 - \delta) - A(\lambda_0 + \delta) = 1$ for any δ with $0 < \delta < \epsilon$. This in turn implies that $\mathrm{SIG}[f_{n-1}(\lambda_0 - \delta)] = \mathrm{SIG}[f_n(\lambda_0 - \delta)]$ and, by Definition 1, this holds in the limit as $\delta \to 0$. Thus $A(\lambda_0) = A(\lambda_0 - \delta)$, and the proof is complete when no $b_i = 0$.

CASE II. *Exactly one* $b_i = 0$.

Assume $b_\nu = 0$. Then S decomposes into the direct sum of two matrices S_1 and S_2 of orders ν and $n - \nu$, respectively, such that each has no $b_i = 0$. Let $f_0^{(1)}(\lambda), \cdots, f_\nu^{(1)}(\lambda)$ and $f_0^{(2)}(\lambda), \cdots, f_{n-\nu}^{(2)}(\lambda)$ be the sequences which would be computed by the algorithm for S_1 and S_2 separately, and let $A^{(1)}(\lambda)$ and $A^{(2)}(\lambda)$ be their respective sign agreements. Then, for any λ, the number of eigenvalues of S greater than or equal to λ is $A^{(1)}(\lambda) + A^{(2)}(\lambda)$ and we must prove that the number of sign agreements of the sequence $f_0(\lambda), \cdots, f_n(\lambda)$ computed for S by (1) is $A^{(1)}(\lambda) + A^{(2)}(\lambda)$.

We note first that $f_i(\lambda) \equiv f_i^{(1)}(\lambda)$ for $i = 0, \cdots, \nu$. Thus it suffices to show that the number of sign agreements of the sequence $f_\nu(\lambda), \cdots, f_n(\lambda)$ is $A^{(2)}(\lambda)$. Since $b_\nu = 0$, we have by (1) that $f_{\nu+1}(\lambda) = (a_{\nu+1} - \lambda)\mathrm{SIG}[f_\nu(\lambda)]$. But

$$f_1^{(2)}(\lambda) = (a_{\nu+1} - \lambda)$$

and therefore $f_{\nu+1}(\lambda) = f_1^{(2)}(\lambda)\mathrm{SIG}[f_\nu(\lambda)]$. Likewise, if $n - \nu > 1$, we have that $f_{\nu+2}(\lambda) = f_2^{(2)}(\lambda)\mathrm{SIG}[f_\nu(\lambda)]$, since $f_2^{(2)}(\lambda) = (a_{\nu+2} - \lambda)f_1^{(2)}(\lambda) - b_{\nu+1}^2$. In carrying out the algorithm for the (possibly) remaining functions $f_{\nu+3}(\lambda), \cdots, f_n(\lambda)$ we obtain $f_{\nu+i}(\lambda) = f_i^{(2)}(\lambda)\mathrm{SIG}[f_\nu(\lambda)]$ for $i = 3, \cdots, n - \nu$. Now multiplication of the sequence $f_1^{(2)}(\lambda), \cdots, f_{n-\nu}^{(2)}(\lambda)$ by the constant factor $\mathrm{SIG}[f_\nu(\lambda)]$ does not change its number of sign agreements. Moreover, there is a sign agreement between $f_\nu(\lambda)$ and $f_{\nu+1}(\lambda)$ if and only if there is one between $f_0^{(2)}(\lambda) = 1$ and

$f_1^{(2)}(\lambda)$. Hence the number of sign agreements of $f_\nu(\lambda), \cdots, f_n(\lambda)$ is $A^{(2)}(\lambda)$, and the proof is complete when exactly one $b_i = 0$.

The extension to the case of an arbitrary number σ of b_i equal to zero can be immediately proved by induction on σ.

4. *An ALGOL Subroutine for Computing $A(\lambda)$*

The following subroutine is the equivalent of the algorithm described in section 2, and is formulated in the version of the algorithmic language ALGOL defined in reference [3].

```
procedure agree (lambda, a[ ], b[ ], N) =: (A)
array (a, f[1:N] , b[0:N − 1] , SIGf[0:N])  ;
integer (N)  ;
comment: This procedure computes the number of agreements in sign of the sequence
  {f_i} defined above for a tridiagonal real symmetric N × N matrix whose diagonal ele-
  ments are given by the array a and off-diagonal elements by b. No attempt is made to
  minimize temporary storage;
begin agree: b[0] := 0  ;  SIGf[0] := +1  ;  A := 0  ;
for I := 1(1)N  ;  begin 1: alpha := a[I] − lambda  ;
beta := b[I − 1] × b[I − 1]  ;  if either (b[I − 1] = 0)  ;
f[I] := alpha × SIGf[I − 1]  ;  or if (b[I − 2] = 0)  ;
f[I] := alpha × f[I − 1] − beta × SIGf[I − 2]  ;  or if (1 = 1)  ;
f[I] := alpha × f[I − 1] − beta × f[I − 2] end  ;
if (f[I] ≠ 0)  ;  SIGf[I] := sign(f[I])  ;  SIGf[I] := SIGf[I − 1]  ;
if ((SIGf[I] × SIGf[I − 1]) > 0)  ;  A := A + 1  ;  A := A end 1  ;
return
end agree
```

REFERENCES

1. W. GIVENS, Numerical computation of the characteristic values of a real symmetric matrix. Oak Ridge National Laboratory, Report 1574, February 19, 1954.
2. J. V. USPENSKY, *Theory of Equations*. New York: McGraw-Hill, 1958.
3. A. J. PERLIS AND K. SAMELSON, Preliminary Report—International Algebraic Language, *Comm. ACM 1*, No. 12 (Dec. 1958), 8–22.