NORTH-HOLLAND

Total Positivity and the QR Algorithm

G. M. L. Gladwell*
Solid Mechanics Division
Faculty of Engineering
University of Waterloo
Waterloo, Ontario, Canada N2L 3G1

Submitted by Robert E. Hartwig

ABSTRACT

An $n \times n$ real matrix $\mathbf{A}$ is TP (totally positive) if all its minors are positive or zero; NTP, if it is nonsingular and TP; STP, if it is strictly TP; O (oscillatory), if it is TP and a power $\mathbf{A}^m$ is STP. Let P be one of NTP, O, STP. We prove that if $\mathbf{A}$ is symmetric and has property P, μ is not an eigenvalue of $\mathbf{A}$, and $\mathbf{A} - \mu\mathbf{I} = \mathbf{QR}$ and $\mathbf{A}' - \mu\mathbf{I} = \mathbf{RQ}$ with $\mathbf{R}$ having positive diagonal, then $\mathbf{A}'$ has property P, and vice versa. The analysis includes a new criterion for $\mathbf{A}$ to be STP. © 1998 Elsevier Science Inc.

1. INTRODUCTION

Totally positive (TP) and oscillatory (O) matrices play an important role in the study of vibratory systems, as discussed by Gantmacher and Krein (1950). The results in the present paper arose from the problem of constructing a finite element model of a vibrating system with specified natural frequencies $(\omega_i)_1^n$: construct symmetric positive definite (PD) tridiagonal matrices $\mathbf{K}, \mathbf{M}$, with respectively negative and positive codiagonals, such that

$$(\mathbf{K} - \omega^2\mathbf{M})\mathbf{x} = \mathbf{0}$$

* E-mail: gladwell@sail.uwaterloo.ca.

LINEAR ALGEBRA AND ITS APPLICATIONS 271:257–272 (1998)

655 Avenue of the Americas, New York, NY 10010

0024-3795/98/$19.00
PII S0024-3795(97)00275-9

has eigenvalues $(\omega_i^2)_1^n$. The factorization $\mathbf{K} = \mathbf{L}\mathbf{L}^T$, $\mathbf{L}^T\mathbf{x} = \mathbf{u}$ leads to the standard form

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{u} = \mathbf{0},$$

where $\mathbf{A} = \mathbf{L}^{-1}\mathbf{M}\mathbf{L}^{-T}$, $\lambda = 1/\omega^2$; the matrix $\mathbf{A}$ is oscillatory, as shown by Gladwell (1986). The technical details of the inverse problem are discussed in Gladwell (1997); here we focus on the fundamental problem of constructing an isospectral family of oscillatory matrices with eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0$, by using shifted **QR** factorization followed by the reversal **QR** $\to$ **RQ**.

2. NOTATION AND PRELIMINARIES

If $1 \leqslant k \leqslant n$, $Q_{k,n}$ will denote the set of strictly increasing sequences $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_k\}$ chosen from $\{1, 2, \dots, n\}$. We write $d(\alpha) = \sum_{i=1}^{k-1}(\alpha_{i+1} - \alpha_i - 1)$, and note that if $\alpha \in Q_{k,n}$, then $d(\alpha) = 0$ iff $\alpha_{i+1} = \alpha_i + 1$. The elements of $Q_{k,n}$ are partially ordered as follows: If $\alpha, \beta \in Q_{k,n}$ then $\alpha \leqslant \beta$ if $\alpha_i \leqslant \beta_i$ for $1 \leqslant i \leqslant k$.

Let $\mathbf{A} \in \mathfrak{R}^{n \times n}$ and $\alpha, \beta \in Q_{k,n}$. The minor of $\mathbf{A}$ formed from rows α and columns β is denoted by $A(\alpha; \beta)$. Following Karlin (1968), we say that $\mathbf{A}$ is

(1) TP (*totally positive*) if all the minors of A are nonnegative;
(2) NTP if $\mathbf{A}$ is nonsingular and TP;
(3) STP (*strictly TP*) if all minors are strictly positive;
(4) O (*oscillatory*) of $\mathbf{A}$ is TP, and $\mathbf{A}^m$ is STP for some positive integer m.

It is known that $\mathbf{A}$ is O iff $\mathbf{A}$ is NTP and $a_{i,i+1} > 0$, $a_{i+1,i} > 0$ for $i = 1, 2, \dots, n-1$.

We use SY to denote *symmetric*.

Throughout the paper we assume that $\mathbf{A} \in \mathfrak{R}^{n \times n}$, and we will use P to denote one of the properties NTP, O, STP.

If $\mu \in \mathfrak{R}$ is not an eigenvalue of $\mathbf{A}$, there is a unique factorization

$$\mathbf{A} - \mu\mathbf{I} = \mathbf{Q}\mathbf{R}, \tag{2.1}$$

with $\mathbf{Q}$ orthogonal, and $\mathbf{R}$ upper triangular with *positive* diagonal. We can therefore define an operator $\mathcal{G}_\mu : \mathbf{A} \to \mathbf{A}' = \mathbf{A}'(\mu)$, where

$$\mathbf{A}' - \mu \mathbf{I} = \mathbf{RQ}. \tag{2.2}$$

We will prove

THEOREM 2.1. *$\mathbf{A}'$ has property SYP iff $\mathbf{A}$ has property SYP.*

The paper runs as follows. In Section 3 we show that the theorem holds for $\mu = 0$ by using known results on LU factorization. In Section 4 we obtain the two new criteria for **A** to be STP. We combine these with some relations between the so-called corner minors of **A** and **A′**, to prove the theorem when **A** is STP. Finally we prove it when **A** is NTP or O.

3. LU FACTORIZATION

The foundations of the theory regarding **LU** factorization were laid by Gantmacher and Krein (1950) and Karlin (1968). Cryer (1973) proved a theorem which includes

THEOREM 3.1. ***A** has property P iff **A** has an **LU** factorization such that **L** and **U** have property $\triangle$ P. Also, **A** has property P iff **A** has a **UL** factorization such that **L** and **U** have property $\triangle$ P.*

We need to explain the symbol $\triangle$P. If **A** is a lower (upper) triangular matrix, the minors $A(\alpha; \beta)$ for which $\beta \leqslant \alpha$ ($\beta \geqslant \alpha$) will be called the *nontrivial* minors of **A**. The remaining minors of **A**, the trivial minors, are identically zero. We say that **A** has property $\triangle$P if **A** is a triangular matrix and the *nontrivial* minors of **A** satisfy the required inequalities of P.

Regarding the property $\triangle$O, Cryer proves

THEOREM 3.2. *Let **A** be a $\triangle$TP lower (upper) triangular matrix. Then **A** is ΔO (i.e., $\mathbf{A}^m$ is $\triangle$STP for some m) iff* (i) ***A** is nonsingular and* (ii) $a_{i+1,i} > 0$ ($a_{i,i+1} > 0$) *for* $i = 1, 2, \ldots, n-1$.

Theorems 3.1, 3.2 have the following corollary. Let **A** have property P, and $\mathbf{A} = \mathbf{LU}$ where $\mathbf{L}, \mathbf{U}$ have property $\triangle$P; then $\mathbf{B} = \mathbf{UL}$ has property P.

We are concerned only with symmetric matrices having the property P, i.e. having property SYP. Such matrices are PD, and therefore have a unique Cholesky factorization $\mathbf{A} = \mathbf{L}\mathbf{L}^T$ with positive diagonal. Theorems 3.1, 3.2 imply that if $\mathbf{A}$ has property SYP then $\mathbf{L}$ will have property $\triangle$P, and $\mathbf{B} = \mathbf{L}^T\mathbf{L}$ will have property SYP.

For matrices with property SYP and $\mu = 0$, the matrix $\mathbf{A}'$ of Equation (2.2) may be obtained from $\mathbf{A}$ by using two Cholesky factorizations and reversals:

$$\mathbf{A} = \mathbf{L}_1\mathbf{L}_1^T, \qquad \mathbf{B} = \mathbf{L}_1^T\mathbf{L}_1 = \mathbf{L}_2\mathbf{L}_2^T, \qquad \mathbf{A}'(0) = \mathbf{L}_2^T\mathbf{L}_2.$$

We write $\mathbf{Q} = \mathbf{L}_1\mathbf{L}_2^{-T} = \mathbf{L}_1^{-T}\mathbf{L}_2$ and $\mathbf{R} = \mathbf{L}_2^T\mathbf{L}_1^T$, and note that $\mathbf{Q}\mathbf{Q}^T = \mathbf{L}_1\mathbf{L}_2^{-T}(\mathbf{L}_1^{-T}\mathbf{L}_2)^T = \mathbf{I}$, so that $\mathbf{Q}$ is orthogonal. Now $\mathbf{A} = \mathbf{L}_1\mathbf{L}_2^{-T}\cdot\mathbf{L}_2^T\mathbf{L}_1^T = \mathbf{Q}\mathbf{R}$, $\mathbf{A}'(0) = \mathbf{L}_2^T\mathbf{L}_1^T\cdot\mathbf{L}_1^{-T}\mathbf{L}_2 = \mathbf{R}\mathbf{Q}$. If $\mathbf{A}$ has property SYP, so do $\mathbf{B}$ and $\mathbf{A}'(0)$, and vice versa. This proves Theorem 2.1 for $\mu = 0$.

Symmetry is essential for the permanence of property P under $\mathcal{S}_\mu$, as is shown by the counterexample for $\mu = 0$:

$$\mathbf{A} = \begin{bmatrix} 2 & a \\ 1 & 2 \end{bmatrix}, \qquad \mathbf{Q} = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix},$$

$$\mathbf{A}'(0) = \frac{1}{5}\begin{bmatrix} 12 + 2a & 4a - 1 \\ 4 - a & 2(4 - a) \end{bmatrix}. \tag{3.1}$$

When $a = \frac{1}{5}$, $\mathbf{A}$ is O and STP, and $\mathbf{A}'$ is not TP; when $a = 0$, $\mathbf{A}$ is NTP, and $\mathbf{A}'$ is not TP. When $\mathbf{A}$ is not symmetric, its $\mathbf{QR}$ factorization may *not* be formed from two $\mathbf{LU}$ factorizations.

The condition that μ is not an eigenvalue of $\mathbf{A}$ is essential. The matrix $\mathbf{A}$ in (3.1) is SYO and SYSTP when $a = 1$. When $\mu = 1$, μ is an eigenvalue of $\mathbf{A}$, and

$$\mathbf{A} - \mu I = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} c & -c \\ c & c \end{bmatrix}\begin{bmatrix} c^{-1} & 2c \\ & 0 \end{bmatrix}, \qquad c = \frac{1}{\sqrt{2}},$$

$$\mathbf{A}'(1) - \mu\mathbf{I} = \begin{bmatrix} c^{-1} & 2c \\ & 0 \end{bmatrix}\begin{bmatrix} c & -c \\ c & c \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \mathbf{A}'(1) = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}.$$

$\mathbf{A}'(1)$ is not oscillatory. When $\mathbf{A}$ is SYO or SYSTP, its eigenvalues are distinct. When $\mu = \lambda_q$ for some q, $\mathbf{A} - \mu\mathbf{I}$ has rank $n - 1$, and $r_{nn} = 0$ (but no other r_{ii} is zero). Thus the last row and column of $\mathbf{A}' - \mu\mathbf{I}$ will be identically zero; in particular, $a'_{n,n-1} = 0$, so that $\mathbf{A}$ cannot be SYO.

4. CRITERIA FOR STP

We define a *successive* minor as an $A(\alpha; \beta)$ with $d(\alpha) = 0 = d(\beta)$. Fekete (1913) established the fundamental

LEMMA 4.1. *Let* $\mathbf{A} \in \Re^{m \times p}$ $(m \geqslant p)$. *If*

(1) *the last* $p - 1$ *columns of* $\mathbf{A}$
(2) *all minors of order* $p - 1$ *taken from the last* $p - 1$ *columns*
(3) *all* successive *minors of order* p

are positive, then all p *th order minors are positive.*

(Note that positive here means strictly positive, not nonnegative as in the definition of TP.)

Induction on p leads to the criteiron which, following Ando (1987), we state as

THEOREM 4.1. $\mathbf{A}$ *is STP if* $A(\alpha; \beta) > 0$ *whenever* $\alpha, \beta \in Q_{k,n}$ *and* $d(\alpha) = 0 = d(\beta)$, $k = 1, 2, \dots, n$.

In other words, for $\mathbf{A}$ to be STP it is sufficient that the successive minors be positive.

Markham (1970) proved

LEMMA 4.2. *Let* $\mathbf{A}$ *be* O. *If* $a_{n1} > 0$ *and* $a_{1n} > 0$, *then* $\mathbf{A}$ *is (strictly) positive.*

The lemma states that $\mathbf{A}$ is positive, i.e., $a_{ij} > 0$ for all $i, j = 1, 2, \dots, n$, *not* that $\mathbf{A}$ is STP. Its proof does not require that $\mathbf{A}$ be O, only that the elements of $\mathbf{A}$ and the 2×2 minors of $\mathbf{A}$ be nonnegative, and the principal diagonal be positive. He uses, as we do, the statements:

(Θ) if $a \geqslant 0$ or $d \geqslant 0$, $b \geqslant 0$ and $c \geqslant 0$, and

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} > 0,$$

then $a > 0$ and $d > 0$;

(Φ) if $a \geqslant 0$ or $d \geqslant 0$, $b > 0$ and $c > 0$, and

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \geqslant 0,$$

then $a > 0$ and $d > 0$.

We denote the pth compound matrix of $\mathbf{A}$, i.e. the matrix of order

$$N = \binom{n}{p},$$

composed of the pth order minors of $\mathbf{A}$ arranged in lexical order, by $\mathbf{A}_p$. We may rephrase Lemma 4.2 as

LEMMA 4.3. *If*

(i) $\mathbf{A}$ *and* $\mathbf{A}_2$ *are nonnegative (we write this* $\mathbf{A} \geqslant 0$, $\mathbf{A}_2 \geqslant 0$),
(ii) *the principal diagonal of* $\mathbf{A}$ *is positive, and*
(iii) $a_{n1} > 0$ *and* $a_{1n} > 0$,

then $\mathbf{A}$ *is positive, i.e.,* $\mathbf{A} > 0$.

Proof. We consider the lower triangle of $\mathbf{A}$. The first column and last row are positive, because statement (Φ) applied to

$$A(i, n; 1, i) = a_{i1}a_{ni} - a_{ii}a_{n1} \geqslant 0$$

implies $a_{i1} > 0$ and $a_{ni} > 0$. If now $j < i < n$, then statement (Φ) applied to

$$A(i, n; j, i) = a_{ij}a_{ni} - a_{ii}a_{nj} \geqslant 0$$

implies $a_{ij} > 0$. We can consider the upper triangle likewise. ■

We generalize this to give a new criterion for $\mathbf{A}$ to be STP.

THEOREM 4.2. *If*

(i) $\mathbf{A}$ *is TP, i.e.,* $\mathbf{A}_p \geqslant 0$, $p = 1, 2, \ldots, n$, *and*
(ii) $A(\alpha; \beta) > 0$ *and* $A(\beta; \alpha) > 0$ *for*

$$\alpha = \{n - p + 1, n - p + 2, \ldots, n\}, \quad \beta = \{1, 2, \ldots, p\}, \quad p = 1, 2, \ldots, n,$$

then **A** *is STP.*

Note that the minors $A(\alpha;\beta)$ and $A(\beta;\alpha)$ are the southwest (SW) and northeast (NE) corner elements of $\mathbf{A}_p$, i.e., $A(\alpha;\beta) = (A_p)_{N,1}$, $A(\beta;\alpha) = (A_p)_{1,N}$. We cannot prove this simply by applying Lemma 4.2 to each compound matrix $\mathbf{A}_p$, for if $\mathbf{A}$ is TP, then $(\mathbf{A}_p)_2$ is not necessarily nonnegative, as shown by the counterexample

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}, \qquad \mathbf{A}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & \boxed{\begin{matrix} 1 & 1 \\ 1 & 0 \end{matrix}} & 0 & 0 & 0 \\ 0 & & & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 \end{bmatrix}. \tag{4.1}$$

A is TP but $(\mathbf{A}_2)_2$ is not nonnegative; the boxed minor is negative. Fekete's Theorem 4.1 shows that for **A** to be STP it is sufficient to prove that all successive minors of **A** are positive. We could therefore consider proving Theorem 4.2 by applying Lemma 4.3 to the submatrices $\mathbf{A}_{ps}$ made up of successive minors of order p. But if **A** is TP, then $(\mathbf{A}_{ps})_2$ is not necessarily nonnegative. For the **A** in (4.1),

$$\mathbf{A}_{2s} = \begin{bmatrix} 1 & 0 & 0 \\ \boxed{\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}} & \begin{matrix} 0 \\ 1 \end{matrix} \end{bmatrix},$$

and $(\mathbf{A}_{2s})_2$ is not nonnegative; the boxed minor is negative.

Note a difference between Theorem 4.2 and Fekete's Theorem 4.1. In Theorem 4.2 we presuppose that **A** is TP, and then have to check that the $2n-1$ corner minors are positive; in Theorem 4.1, we do not presuppose that **A** is TP, and have to check *all* the successive minors; there are $n(n+1)(2n+1)/6$ in all. Condition (i) of Theorem 4.2 is essential;

$$\mathbf{A} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$$

satisfies (ii), but is not STP.

We will prove the theorem by applying the 2×2 version of Sylvester's identity (Gantmacher, 1959) for bordered determinants: let

$$c_{ij} = B(1,2,\ldots,p,i;1,2,\ldots,p,j), \qquad i,j = p+1, p+2;$$

then

$$C(p+1,p+2;p+1,p+2)$$
$$= B(1,2,\ldots,p;1,2,\ldots p)B(1,2,\ldots,p+2;1,2,\ldots,p+2). \tag{4.2}$$

We will apply this to 2×2 minors taken from the compound matrices $\mathbf{A}_p$. We need some abbreviations; we write

$$A_p\{i,j\} = \begin{cases} A(i-p+1, i-p+2,\ldots,i;j,j+1,\ldots,j+p-1) \\ \quad \text{for} \quad i=p,p+1,\ldots,n, \quad j=1,2,\ldots,n-p+1, \\ \qquad \text{and} \quad i-p+1 \geqslant j, \\ A(i,i+1,\ldots,i-p+1;j-p+1,j-p+2,\ldots,j) \\ \quad \text{for} \quad i=1,2,\ldots,n-p+1, \quad j=p,p+1,\ldots,n \\ \qquad \text{and} \quad i-p+1 \leqslant j, \end{cases}$$

$$B_p\{i;j\} = \begin{cases} A(i-p+1,i-p+2,\ldots,i;j,j+1,\ldots,j+p-2,j+p) \\ \quad \text{for} \quad i=p,p+1,\ldots,n, \quad j=1,2,\ldots,n-p, \\ \qquad \text{and} \quad i-p+1 \geqslant j, \\ A(i,i+1,\ldots,i+p-2,i+p;j-p+1,j-p+2,\ldots,j) \\ \quad \text{for} \quad i=1,2,\ldots,n-p, \quad j=p,p+1,\ldots,n, \\ \qquad \text{and} \quad i-p+1 \leqslant j. \end{cases}$$

Note that the corner minor $A(\alpha;\beta)$, $A(\beta;\alpha)$ in the condition of the theorem are $A(\alpha;\beta) = A_p\{n;1\}$, $A(\beta;\alpha) = A_p\{1;n\}$ respectively.

Proof of Theorem 4.2. We apply induction on n. The theorem is trivially true for $n = 1, 2$. Assume that it is true for matrices of order $n-1$ ($n \geqslant 3$) and that $\mathbf{A}$ satisfies conditions (i) and (ii).

Condition (ii) states that $A_p\{n;1\} > 0$ for $p = 1,2,\ldots,n$, so that $A_{p+1}\{n;1\} > 0$ for $p = 0,1,2,\ldots,n-1$; $p = 0$ gives $A_1\{n;1\} \equiv a_{n,1} > 0$; $p = 1$ gives

$$A_2\{n;1\} = a_{n-1,1}a_{n,2} - a_{n,1}a_{n-1,2} > 0$$

Since all four terms in this inequality are nonnegative, statement (Θ) yields $a_{n-1,1} \equiv A_1\{n-1;1\} > 0$.

We examine the logic: the positivity of the SW corner terms of $\mathbf{A}$ and $\mathbf{A}_2$ implies the positivity of the SW corner term ($a_{n-1,1}$) in $\mathbf{A}_L$, the leading principal minors of $\mathbf{A}$. If we can show that conditions (i) and (ii) for $\mathbf{A}$ imply conditions (i) and (ii) for $\mathbf{A}_L$, then our induction hypothesis will imply that $\mathbf{A}_L$ is STP. This is the first step. Then all we will have to do, in the second step, to prove that $\mathbf{A}$ itself is STP, is to show that all the minors involving the last row and/or last column are positive.

Step I: Sylvester's identity (4.2) and appropriate row and column interchanges yields the identity

$$\begin{vmatrix} A_p\{n-1;1\} & B_p\{n-1;1\} \\ A_p\{n;1\} & B_p\{n;1\} \end{vmatrix} = A_{p-1}\{n-1;1\}A_{p+1}\{n;1\} \quad (4.3)$$

linking four terms in $\mathbf{A}_p$. Consider (4.3) for $p = 2, 3, \ldots, n-1$. Since $A_{p+1}\{n;1\} > 0$ and all four terms in the determinant, being minors of $\mathbf{A}$, are nonnegative, statement (Θ) states that

$$\text{if} \quad A_{p-1}\{n-1;1\} > 0, \quad \text{then} \quad A_p\{n-1;1\} > 0.$$

But $A_1\{n-1;1\} > 0$, so that $A_p\{n-1;1\} > 0$, $p = 1, 2, \ldots, n-1$. But $A_p\{n-1;1\} \equiv A(n-p, n-p+1, \ldots, n-1; 1, 2, \ldots, p)$ is the SW corner pth order minor of $\mathbf{A}_L$; it is positive. By considering the NE corner in a similar fashion, we can show that $A_p\{1;n-1\} > 0$, $p = 1, 2, \ldots, n-1$. This proves step I. Our induction hypothesis now states that $\mathbf{A}_L$ is STP.

Step II: We now show that all the minors involving the last row and/or last column are positive. Lemma 4.1 states that we need consider only successive minors involving the last row or column. We consider those involving the last row. Condition (ii) gives

$$\mathscr{E}_1 \equiv \{A_p\{n;1\},\, p = 1, 2, \ldots, n\} > 0.$$

Sylvester's indeitity yields

$$\begin{vmatrix} A_p\{n-1;j\} & A_p\{n-1;j+1\} \\ A_p\{n;j\} & A_p\{n;j+1\} \end{vmatrix} = A_{p-1}\{n-1;j+1\}A_{p+1}\{n;j\}. \quad (4.4)$$

Since we have shown that $\mathbf{A}_L$ is STP, $A_{p-1}\{n-1; j+1\} > 0$ for $(j+1) + (p-1) \leqslant n-1$, i.e. $j + p \leqslant n$. Since all four terms on the left of (4.4) are nonnegative, statement (Θ) gives

$$\text{if} \quad j + p \leqslant n \ \text{ and } \ A_{p+1}\{n; j\} > 0, \quad \text{then} \quad A_p\{n; j+1\} > 0.$$

This gives the pattern of implications shown in Figure 1.

This means that if

$$\mathscr{E}_j \equiv \left\{A_p\{n; j\},\ p = 1, 2, \ldots, n - j + 1\right\},$$

then $\mathscr{E}_j > 0$ implies $\mathscr{E}_{j+1} > 0$. Since $\mathscr{E}_1 > 0$, we conclude that $\mathscr{E}_j > 0$ for $j = 1, 2, \ldots, n$. We can prove similarly that all the successive minors involving the last *column* of $\mathbf{A}$ are positive. Hence $\mathbf{A}$ is STP. ■

COROLLARY 1. *Under the conditions of the theorem,* $(\mathbf{A}_{ps})_2$ *is strictly positive.*

An application of Sylvester's identity similar to (4.4) shows that successive 2×2 minors of $\mathbf{A}_{ps}$ are positive. Lemma 4.1 implies that all are positive.

COROLLARY 2. *Condition* (i) *of the theorem may be replaced by*

(i a) $\mathbf{A}_L$ *is TP*;
(i b) *the principal minors of* $\mathbf{A}$ *are positive.*

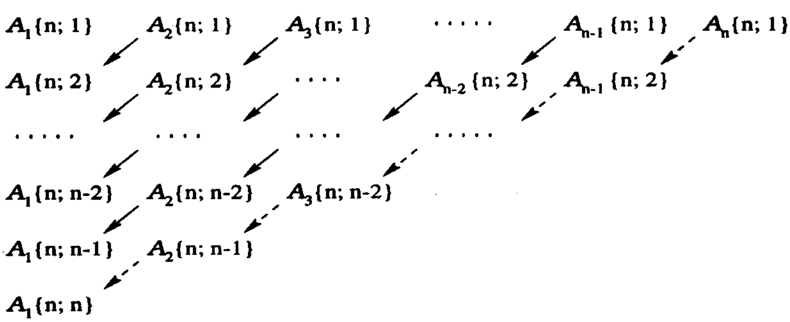

FIG. 1. The pattern of implication of positivity.

Proof. We use induction on n as before. The corollary holds when $n = 1, 2$. Assume that it is true for matrices of order $n - 1$ ($n \geqslant 3$). The argument used in the proof of Theorem 4.2 has two parts. For the lower triangle, they are

(1) $A_p\{n; 1\} > 0$, $p = 1, 2, \dots, n$, implies $A_p\{n - 1; 1\} > 0$, $p = 1, 2, \dots, n - 1$;

(2) $\mathscr{E}_j > 0$ implies $\mathscr{E}_{j+1} > 0$, $j = 1, 2, \dots n - 1$.

When only $\mathbf{A}_L$, not $\mathbf{A}$, is TP, the argument leading to (1) fails when $p = n - 1$; for $B_{n-1}\{n - 1; 1\}$ in (4.2) is not a minor of $\mathbf{A}_L$, since it involves the last column of $\mathbf{A}$. Thus we cannot deduce that $A_{n-1}\{n - 1; 1\} > 0$. However, this is a principal minor of $\mathbf{A}$, and is positive by the hypothesis (i b); we have therefore established (1).

The argument leading to (2) fails in a similar way. When $j + p = n$, i.e. $p = n - j$, $A_{p-1}\{n - 1; j + 1\}$ is not a minor of $\mathbf{A}_L$, because it involves the last column of $\mathbf{A}_L$. From (4.3) we can deduce only

$$\text{if} \quad j + p \leqslant n - 1 \text{ and } A_{p+1}\{n; j\} > 0, \quad \text{then} \quad A_p\{n; j + 1\} > 0.$$

This means that the pattern of implications excludes those shown as broken lines in Figure 1. But the final diagonal of principal minors of $\mathbf{A}$ is positive by hypothesis. Thus (2) holds, and $\mathbf{A}$ is STP. ■

5. SOME FUNDAMENTAL PROPERTIES OF QR

The basic equations are

$$\mathbf{A} - \mu\mathbf{I} = \mathbf{QR}, \qquad \mathbf{A}' - \mu\mathbf{I} = \mathbf{RQ}. \tag{5.1}$$

These yield

$$\mathbf{A}' = \mu\mathbf{I} + \mathbf{Q}^T\mathbf{QRQ} = \mu\mathbf{I} + \mathbf{Q}^T(\mathbf{Q} - \mu\mathbf{I})\mathbf{Q},$$

so that

$$\mathbf{A}' = \mathbf{Q}^T\mathbf{AQ}.$$

This shows that $\mathbf{A}'$ is SY iff $\mathbf{A}$ is SY. If $\mathbf{A}$ is SYP, then it is PD, and hence $\mathbf{A}'$ is PD.

The equations (5.1) yield

$$\mathbf{A}'\mathbf{R} = \mu\mathbf{R} + \mathbf{RQR} = \mu\mathbf{R} + \mathbf{R}(\mathbf{A} - \mu\mathbf{I}) = \mathbf{RA}. \tag{5.2}$$

The Binet-Chauchy theorem ($\mathbf{AB} = \mathbf{C} \Rightarrow \mathbf{A}_p\mathbf{B}_p = \mathbf{C}_p$) now gives

$$(\mathbf{A}')_p\mathbf{R}_p = \mathbf{R}_p\mathbf{A}_p. \tag{5.3}$$

Now we can prove

$$(\mathbf{A}'^m)_p\mathbf{R}_p = \mathbf{R}_p\mathbf{A}_p^m \tag{5.4}$$

by induction. The Binet-Cauchy theorem gives

$$(\mathbf{A}^m)_p = (\mathbf{A}_p)^m = \mathbf{A}_p^m \tag{5.5}$$

and similarly

$$(\mathbf{A}'^m)_p = (\mathbf{A}'_p)^m = \mathbf{AA}'^m_p. \tag{5.6}$$

The result holds for $m = 1$. If it holds for m, then

$$\begin{aligned}(\mathbf{A}'^{m+1})_p\mathbf{R}_p &= \mathbf{A}'_p(\mathbf{A}'^m)_p\mathbf{R}_p = \mathbf{A}'_p(\mathbf{R}_p\mathbf{A}_p^m)\\ &= (\mathbf{A}'_p\mathbf{R}_p)\mathbf{A}_p^m = (\mathbf{R}_p\mathbf{A}_p)\mathbf{A}_p^m = \mathbf{R}_p\mathbf{A}_p^{m+1}.\end{aligned}$$

Equating the lower left-hand corner elements on both sides of (5.4), we find

$$A_p'^m\{n;1\}(R_p)_{11} = (R_p)_{NN}A_p^m\{n;1\} \tag{5.7}$$

where, as before, $N = \binom{n}{p}$ is the order of $\mathbf{A}_p$.

When μ is not an eigenvalue of $\mathbf{A}$, so that all the r_{jj} are positive, we may write (5.7) as

$$A_p'^m\{n;1\} = \left(\prod_{j=1}^{p}\frac{r_{n-j+1,n-j+1}}{r_{jj}}\right)A_p^m\{n;1\}. \tag{5.8}$$

This holds for $p = 1, 2, \ldots, n$ and $m = 1, 2, \ldots$.

6. PROOF OF THEOREM 2.1 FOR $\mu \neq 0$

Suppose first that **A** is SYSTP. The analysis of Section 3 shows that $\mathbf{A}'(0)$ is STP. Equation (5.6) for $m = 1$ shows that, for all $\mu \neq \lambda_q$, $q = 1, 2, \ldots, n$,

$$A'_p\{n;1\} > 0, \qquad p = 1, 2, \ldots, n. \tag{6.1}$$

The matrix $\mathbf{A}'(\mu)$, and all the minors of $\mathbf{A}'(\mu)$, are continuous functions of μ. Therefore there is an open interval around zero in which $\mathbf{A}'(\mu)$ is STP. We first show that $\mathbf{A}'(\mu)$ is STP for all $\mu < 0$. For if that were not so, there would be a $\mu_0 < 0$ for which some of the minors were zero while the remainder were positive. But then $\mathbf{A}'(\mu_0)$ would be TP, while the corner minors $A'_p\{n;1\}$ would be positive, so that Theorem 4.2 would show that $\mathbf{A}'(\mu_0)$ is STP. This contradiction implies that $\mathbf{A}'(\mu)$ is STP for all $\mu < 0$. An exactly similar argument shows that $\mathbf{A}'(\mu)$ is STP for all μ satisfying $0 \leqslant \mu < \lambda_n$, where λ_n is the least eigenvalue of **A**.

At $\mu = \lambda_n$, as we showed in Section 3, the last row and column of $\mathbf{A}'(\mu) - \mu\mathbf{I}$ are zero. The leading principal submatrix $\mathbf{A}'_L(\lambda_n)$ is PD and, by continuity as $\mu \to \lambda_n$, TP.

We cannot immediately extend this argument to $\mu > \lambda_n$ by continuity, as we know only that $\mathbf{A}'_L(\lambda_n)$ is TP, not STP. But $\mathbf{A}'_L(\lambda_n)$ is SYNTP, and an SYNTP matrix **A** may be approximated arbitrarily closely by the SYSTP matrix **GAG**, where

$$\mathbf{G} = \mathbf{G}(k) = (g_{ij}) = \exp\left[-k(i-j)^2\right], \qquad i, j = 1, 2, \ldots, n.$$

The matrix **G** is SYSTP and $\mathbf{G}(k) \to \mathbf{I}_n$ as $k \to \infty$. [Note that the parameter p in (c) on p. 213 of Ando (1987) should be in the numerator, not the denominator.]

Write

$$\mathbf{A}'(\mu) = \begin{bmatrix} \mathbf{A}'_L & \mathbf{b} \\ \mathbf{b}^T & c \end{bmatrix},$$

let $\mathbf{G}_L$ be the principal leading submatrix of **G**, and form the matrix

$$\mathbf{A}'(\mu, k) = \begin{bmatrix} \mathbf{G}_L & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{A}'_L & \mathbf{b} \\ \mathbf{b}^T & c \end{bmatrix} \begin{bmatrix} \mathbf{G}_L & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{G}_L\mathbf{A}'_L\mathbf{G}_L & \mathbf{G}_L\mathbf{b} \\ \mathbf{b}^T\mathbf{G}_L & c \end{bmatrix}.$$

At $\mu = \lambda_n$, $\mathbf{A}'_L$ is NTP and $\mathbf{b}(\mu) = \mathbf{0}$; $\mathbf{G}_L\mathbf{A}'_L\mathbf{G}_L$ is STP. By continuity there is an open interval around $\mu = \lambda_n$ in which $\mathbf{G}_L\mathbf{A}'_L\mathbf{G}_L$ is STP. Either $\mathbf{G}_L\mathbf{A}'_L(\mu)\mathbf{G}_L$ is STP for all μ satisfying $\lambda_n \leqslant \mu < \lambda_{n-1}$, or there is a μ_1 satisfying $\lambda_n < \mu_1 < \lambda_{n-1}$ at which $\mathbf{G}_L\mathbf{A}'_L(\mu_1)\mathbf{G}_L$ is only TP. In the former case, on taking the limit $k \to \infty$ we deduce that $\mathbf{A}'_L(\mu)$ is TP for all $\lambda_n < \mu < \lambda_{n-1}$. Then we can apply Corollary 2 of Theorem 4.2 to deduce that $\mathbf{A}'(\mu)$ is STP. In the latter case, we show that $\mathbf{A}'(\mu_1, k)$ is STP by again applying Corollary 2 of Theorem 4.2; the leading principal submatrix $\mathbf{G}_L\mathbf{A}'_L(\mu_1)\mathbf{G}_L$ of $\mathbf{A}'(\mu_1, k)$ is TP, and $\mathbf{A}'(\mu_1, k)$ is PD; we need to ensure that the corner minors of $\mathbf{A}'(\mu_1, k)$ are positive. Since $\mathbf{G}_L$ becomes increasingly diagonally dominant as k increases, we may state

$$A'_p(\mu_1, k)\{n;1\} = A'(\mu_1)_p\{n;1\} + O(e^{-k})$$

so that for sufficiently large k, i.e. $k > K$,

$$A'_p(\mu_1, k)\{n;1\} > 0.$$

Thus Corollary 2 of Theorem 4.2 states that, for $k > K$, $\mathbf{A}'(\mu_1, k)$ is STP. Now, letting $k \to \infty$, we deduce that $\mathbf{A}'(\mu_1)$ is TP. Now we apply Theorem 4.2 and deduce that $\mathbf{A}'(\mu_1)$ is STP. Hence $\mathbf{A}'(\mu)$ is STP for all μ satisfying $\lambda_n < \mu < \lambda_{n-1}$, and by the same argument for all $\mu \neq \lambda_q$, $q = 1, 2, \dots, n$.

Now suppose that $\mathbf{A}$ is SYNTP. We may approximate it arbitrarily closely by the SYSTP matrix $\mathbf{A}(k) = \mathbf{GAG}$, and use the theorem we have proved for an SYSTP matrix. We know that the matrix $\mathbf{A}(k)'(\mu)$ derived from $\mathbf{A}(k)$ is STP for all $\mu \neq \lambda(k)_q$, where $\lambda(k)_q$ denotes the qth eigenvalue of $\mathbf{A}(k)$. Since the eigenvalues of a matrix are continuous functions of its entries, each sequence $\lambda(k)_q$, $k = 1, 2, \dots,$ will converge to λ_q. Choose $\epsilon > 0$; then we may find K such that for all $k > K$ and all $q = 1, 2, \dots, n$, $|\lambda(k)_q - \lambda_q| < \epsilon$. This means that if $k > K$, $\mathbf{A}'(k)(\mu)$ will be STP for all μ satisfying $|\lambda_q - \mu| > \epsilon$, $q = 1, 2, \dots, n$. This means that for all $k = 1, 2, \dots,$ we may find K depending on k such that for all $k > K$ and all $q = 1, 2, \dots, n$, we have $|\lambda_q(k) - \lambda_q| < 1/k$. Thus if $k \geqslant K$, $\mathbf{A}(k)'(\mu)$ is STP for all μ satisfying $|\mu - \lambda_q| > 1/k$ for $q = 1, 2, \dots, n$. Taking the limit $k \to \infty$ and noting that the elements of $\mathbf{A}'$ are continuous functions of those of $\mathbf{A}$, we deduce that

$$\mathbf{A}'(\mu) = \lim_{k \to \infty} \mathbf{A}(k)'(\mu)$$

is NTP for $\mu \neq \lambda_q$, $q = 1, 2, \dots, n$.

If $\mathbf{A}$ is SYO, then it is SYNTP. Therefore by the preceding conclusion $\mathbf{A}'(\mu)$ is NTP. Now we use (5.8). Since $\mathbf{A}$ is SYO, there is a power $\mathbf{A}^m$ that is

SYSTP. For that power, $A_p^m\{n;1\} > 0$ for all p, so that $A_p'^m(\mu)\{n;1\} > 0$ for all $\mu \neq \lambda_q$. Thereom 4.2 applied to $A'^m(\mu)$ shows that $\mathbf{A}'^m(\mu)$ is STP, hence $\mathbf{A}'(\mu)$ is O for all $\mu \neq \lambda_q$.

We have now proved the "if" part of Theorem 2.1. The "only if" part follows in exactly the same way, for if the equations (5.1) hold, then since **A** and **A**′ are symmetric, we may write

$$\mathbf{A}' - \mu\mathbf{I} = \mathbf{Q}^T\mathbf{R}^T, \qquad \mathbf{A} - \mu\mathbf{I} = \mathbf{R}^T\mathbf{Q}^T.$$

We now argue from **A**′ to **A**. The change from the upper triangle **R** to the lower triangle $\mathbf{R}^T$ has no effect on any of the arguments we have used.

7. CONCLUDING REMARKS

If **A** has property SYP, then the operator $\mathscr{S}_\mu$ maps **A** into an orthogonally similar matrix $\mathbf{A}'(\mu)$ with property SYP. Equation (5.2) and its generalizations (5.3) and (5.4) show that $\mathscr{S}_\mu$ will preserve the band structure **A**, $\mathbf{A}_p$, and $\mathbf{A}_p^m$. Thus from any one **A** with property SYP, we may generate an orthogonally similar family of matrices by repeated application of $\mathscr{S}_\mu$, with the same or different values of μ; all of them will have the same structure as **A**. Note that this does not contradict the possible convergence of symmetric **A** to its diagonal under repeated **QR** factorizations and reversals. If **A** has property SYP, then all the iterates will have property SYP even though they may converge to the diagonal under the appropriate conditions.

There are two simple, but important special cases of our analysis. If **A** is symmetric and tridiagonal, it will be SYO iff it is PD and has positive codiagonal. Equation (5.2) shows that **A**′ is symmetric and tridiagonal. This case was considered in more detail in Gladwell (1995). If **A** is the inverse of a symmetric PD tridiagonal matrix with *negative* codiagonal, then it is SYO. The equation $\mathbf{R}\mathbf{A}^{-1} = (\mathbf{A}')^{-1}\mathbf{R}$ derived from (5.2) shows that **A**′ will be the inverse of a PD tridiagonal matrix with negative codiagonal. The family of such matrices includes the D-matrices considered by Markham (1970). We note however that if **A** is a D-matrix, then **A**′ is not necessarily a D-matrix, as shown by the counterexample

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \qquad \mu = \frac{1}{4}, \qquad \mathbf{A}' = \begin{bmatrix} \frac{13}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{2}{5} \end{bmatrix}.$$

A′ is not a D-matrix.

Provided that μ is less (greater) than the lowest (highest) eigenvalues of $\mathbf{A}$, there is an analogue of Theorem 2.1 for $\mathbf{A}'$ derived from $\mathbf{A}$ by the operations $\mathbf{A} - \mu\mathbf{I} = \pm\mathbf{L}\mathbf{L}^T$, $\mathbf{A}' - \mu\mathbf{I} = \pm\mathbf{L}^T\mathbf{L}$, with $\mathbf{L}$ having positive diagonal.

I thank Thomas Markham for telling me about Ando's paper. Hongmei Zhu helped me by finding counterexamples.

REFERENCES

Ando, T. 1987. Totally positive matrices, *Linear Algebra Appl.* 90:165–219.

Cryer, C. W. 1973. The LU-factorization of totally positive matrices, *Linear Algebra Appl.* 7:83–92.

Fekete, M. 1913. Über ein Problem von Laguerre, *Rend. Circ. Math. Palermo* 34:89–100, 110–120.

Gantmacher, F. R. 1959. *The Theory of Matrices*, Vol. 1, Chelsea, New York, p. 32.

Gantmacher, F. R. and Krein, M. G. 1950. *Oscillation Matrices and Kernels, and Small Vibrations of Mechanical Systems* (English transl., 1961), Dept. of Commerce, Washington.

Gladwell, G. M. L. 1986. *Inverse Problems in Vibration*, Kluwer, Dordrecht.

Gladwell, G. M. L. 1995. On isospectral spring-mass systems, *Inverse Problems* 11:591–602.

Gladwell, G. M. L. 1997. Inverse vibration problems for finite element models, *Inverse Problems* 13:1–12.

Karlin, S. 1968, *Total Positivity*, Vol. 1, Stanford U.P.

Markham, T. 1970. On oscillatory matrices, *Linear Algebra Appl.* 3:143–158.

Received 12 August 1996; final manuscript accepted 24 April 1997