Tridiagonalization of Completely Nonnegative Matrices*

By J. W. Rainey and G. J. Habetler

Abstract. Let \( M = [m_{ij}]_{i,j=1}^{n} \) be completely nonnegative (CNN), i.e., every minor of \( M \) is nonnegative. Two methods for reducing the eigenvalue problem for \( M \) to that of a CNN, tridiagonal matrix, \( T = [t_{ij}] \) \( (t_{ij} = 0 \) when \( |i - j| > 1 \)), are presented in this paper. In the particular case that \( M \) is nonsingular it is shown for one of the methods that there exists a CNN nonsingular \( S \) such that \( SM = TS \).

1. Introduction. It is well known that if \( M = [m_{ij}]_{i,j=1}^{n} \) is Hermitian, there exists an orthogonal \( Q \) such that \( QMQ^* = T \) is tridiagonal, i.e., \( t_{ij} = 0 \) when \( |i - j| > 1 \). Moreover, for \( \lambda > 0 \) sufficiently large and some nonsingular, diagonal \( D \), \( D(T + \lambda I)D^{-1} \) is completely nonnegative (CNN), i.e., every minor of \( D(T + \lambda I)D^{-1} \) is nonnegative. (See [2], [3] for a discussion and applications of CNN matrices.) We want to show that an analogous result can be obtained when \( M \) is CNN. Namely, we will show that given any arbitrary CNN matrix, \( M \), one can easily construct a CNN tridiagonal matrix, \( T \), which has the same eigenvalues as \( M \). Two methods for obtaining \( T \) are described in Section 2, both methods being based upon a result derived in Section 3.

2. Outline of the Methods. (a) First Method. If for some \( k \) \( (2 \leq k \leq n - 1) \),

\[
(2.1) \quad m_{ii} = 0 \quad (m_{ii} = 0), \quad i = 1, \ldots, k - 1, \quad j = i + 2, \ldots, n,
\]

we will say that \( M \) is "lower (upper) Hessenberg through its first \( k \) rows (columns)." For convenience, we will say that any matrix is Hessenberg through its first row or column. A matrix is Hessenberg in the case \( k = n - 1 \).

In Section 3, we prove the

Basic Lemma. Let \( M \) be lower Hessenberg through its first \( k \) rows. Then, there exists a CNN matrix, \( M' \), which has the same eigenvalues as \( M \) and which is lower Hessenberg through its first \( k + 1 \) rows. If \( M \) is nonsingular, then there exists a CNN nonsingular \( S' \) such that \( S'M = M'S' \).

By a sequential application of the Basic Lemma, it follows that we can find a CNN lower Hessenberg matrix, \( H \), which has the same eigenvalues as \( M \). We note that if \( M \) is nonsingular then \( H = S''M(S'')^{-1} \), where \( S'' \) is CNN (from, e.g., the Cauchy-Binet theorem [2, I]).

Let \( P \) be the matrix obtained by reversing the order of the rows of the \( n \times n \) identity, \( I \); trivially, \( P^{-1} = P \).

Received January 11, 1971, revised May 24, 1971.


Key words and phrases. Tridiagonalization, tridiagonal matrices, completely nonnegative matrices, Hessenberg matrices.

* Part of this paper appeared in J. W. Rainey's dissertation (Rensselaer, 1967). Research on the dissertation was supported in part by funds from the National Science Foundation under Grant NSF-GP-6339.

Copyright © 1972, American Mathematical Society

121
Define $\hat{H} = PHP$. $\hat{H}$ is similar to $H$ and therefore has the same eigenvalues as $M$. $\hat{H}$ is obtained by reversing the order of the rows and columns of $H$ and therefore is upper Hessenberg; since the value of a minor is not changed by reversing the order of the rows and columns of its array form, $\hat{H}$ must be CNN.

As we indicate in Section 3, a sequential application of our method of proof of the Basic Lemma to $\hat{H}$ maintains the upper Hessenberg form of $\hat{H}$ and therefore yields a CNN tridiagonal matrix, $\hat{T}$. In general, we could take $T = \hat{T}$. In the particular case that $M$, and therefore $H$, is nonsingular, we note as before that there exists a nonsingular CNN $S$ such that $\hat{T} = SH(S)^{-1}$; defining

$$T = P\hat{T}P = P\hat{H}(\hat{S})^{-1}P = P\hat{S}PH(\hat{S})^{-1}P = P\hat{S}PS''M(S'')^{-1}P(\hat{S})^{-1}P = SM\hat{S}^{-1}$$

where $S = P\hat{S}PS''$, it is easily verified that $T$ is tridiagonal, CNN, and that $S$ (the product of the CNN matrices, $P\hat{S}$ and $S''$) is CNN.

(b) Second Method. If, for some $k (2 \leq k \leq n - 1)$,

$$m_{ii} = m_{jj} = 0, \quad i = 1, \cdots, k - 1, \quad j = i + 2, \cdots, n,$$

we will say that $M$ is “tridiagonal through its first $k$ rows and columns.” For convenience, we will say that any square matrix is tridiagonal through its first row and column. A matrix is tridiagonal in the case $k = n - 1$. We want to prove the

**SEQUENTIAL LEMMA.** Let $M$ be tridiagonal through its first $k (<n - 1)$ rows and columns. Then there exists a CNN matrix $\bar{M}$ which has the same eigenvalues as $M$ and which is tridiagonal through its first $k + 1$ rows and columns.

**Proof.** Applying the method of proof of the Basic Lemma to $M$ yields $M'$ which is tridiagonal through its first $k$ rows and columns and lower Hessenberg through its first $k + 1$ rows.

Since every minor of the transpose $(M')^t$ of $M'$ will be the transpose of some minor of $M'$, we note that $(M')^t$ is CNN. Moreover, $(M')^t$ has the same eigenvalues as $M$, is tridiagonal through its first $k$ rows and columns and is upper Hessenberg through its first $k + 1$ columns. Applying the method of proof of the Basic Lemma to $(M')^t$ would now yield $\bar{M}$.

The proof of the preceding lemma indicates a method of “sequentially tridiagonalizing” (a term introduced in [1]) $M$ with, as we will show, the desirable property that each intermediate result of the procedure is CNN.

Let $M^{(k)} = \{m_{ij}^{(k)}\}_{i,j=1}^{n}$ be the $(k - 1)$th result of applying the sequential tridiagonalization procedure to $M$ (in general, $M^{(1)} = M$, $M^{(n-1)} = T$). In analogy with (2.2), we can assume that

$$m_{ii}^{(k)} = m_{jj}^{(k)} = 0, \quad i = 1, \cdots, k - 1, \quad j = i + 2, \cdots, n.$$

As shown in [4, p. 399 ff.], a measure of the stability of the procedure (but by no means the most important measure) is the growth of the quantities

$$\rho_k = \sum_{i=k+1}^{n} m_{ii}^{(k)} m_{ij}^{(k)},$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
where the \( \rho_k \) (see, e.g., [1]) also satisfy
\[
(2.5) \quad \rho_k = t_{k+1,k} t_{k,k+1}, \quad k = 1, \ldots, n - 1.
\]

We want to show that the \( \rho_k \) cannot become arbitrarily large.

First of all, we note that \( M^{(k)} (k > 1) \) is obtained by similarity transformations performed on either \( M^{(k-1)} \) or on a "reduced" form of \( M^{(k-1)} \); in either case, \( \text{trace}(M^{(k)}) = \text{trace}(M^{(k-1)}) \) and, therefore, \( \text{trace}(M) = \text{trace}(M^{(k)}) \) for all \( k \).

Now, since \( M^{(k)} \) is CNN,
\[
\begin{array}{c}
m_{kk}^{(k)} m_{ij}^{(k)} \geq m_{kj}^{(k)} m_{ik}^{(k)} \geq 0, \quad j \geq k + 1, \\
\end{array}
\]
and, therefore,
\[
m_{kk}^{(k)} \sum_{j=k+1}^{n} m_{ij}^{(k)} \geq \rho_k \geq 0,
\]
or, since \( \text{trace}(M) = \text{trace}(M^{(k)}) = \sum_{i=1}^{n} m_{ii}^{(k)} \),
\[
0 \leq \rho_k \leq (\text{trace}(M))^2.
\]

By maintaining the CNN property in our procedure, we are assured that the \( \rho_k \) remain uniformly bounded with respect to \( k \).

We note that if \( M^{(1)} \) is nonsingular, then \( M^{(n-1)} = T \) is similar to \( M^{(1)} \); letting "\( \sim \)" indicate similarity, we have, in the notation of the Sequential Lemma, \( M \sim M' \sim (M')' \sim \tilde{M} \) (since any square matrix is similar to its transpose) and by induction, \( M^{(1)} \sim T \). Thus, \( T = SM^{(1)} S^{-1} \) for some \( S \) but the \( S \) "constructed" as in the proof of the Basic and Sequential Lemmas is not, in general, CNN. For example, if
\[
M = \begin{bmatrix}
3 & 1 & 1 \\
2 & 2 & 2 \\
1 & 3 & 4
\end{bmatrix},
\]
then, following the procedure indicated on the proof of the Sequential Lemma, one obtains
\[
T = \begin{bmatrix}
3 & 2 & 0 \\
1 & 5.5 & .75 \\
0 & 1 & .5 \\
1 & 0 & 0 \\
0 & .5 & .5 \\
0 & -1 & 1
\end{bmatrix},
\]
\[
S = \begin{bmatrix}
1 & 0 & 0 \\
0 & .5 & .5 \\
0 & 1 & 1
\end{bmatrix}.
\]
The question of whether or not there exists, in the general case, some CNN \( S \) such that \( T = SMS^{-1} \) remains open.

3. Proof of the Basic Lemma. Let \( M \) be CNN and lower Hessenberg through its first \( k \) rows but not through its first \( k + 1 \) rows. Then, there exists \( p \geq k + 1 \) such that
\[ M = \begin{bmatrix}
X & \cdots & X & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & u & v & 0 & \cdots & 0 \\
X & \cdots & \cdots & X
\end{bmatrix},\]

where the \( X \)'s indicate possibly nonzero elements, \( u = m_{k,p} \) and \( v = m_{k,p+1} \neq 0 \).

Note. We indicate in (3.1) that \( p + 1 < n \) and \( k > 1 \); whether or not this is true will make no difference in our argument.

We assert that we can verify our primary statement in the lemma by showing that there exists a CNN matrix, say \( \tilde{M} \), which has precisely the same form as \( M \) in (3.1) and the same eigenvalues, but \( \xi = 0 \), and then calling on finite induction. We proceed with the proof of the latter.

Consider first the case when \( u = m_{kp} = 0 \). From the latter assumption, (3.2) and the fact that \( u \cdot m_{i,p+1} - v \cdot m_{i,p} \geq 0 \) when \( i \geq k \), it follows that the \( p \)th column of \( M \) must be null. By a similarity transformation involving elementary permutation matrices, one can therefore obtain

\[ M' = \begin{bmatrix}
M'_1 & 0 \\
m'_1 & 0
\end{bmatrix},\]

where \( M'_1 \) is obtained by deleting the \( p \)th row and column of \( M \) while \( m'_1 \) is obtained by deleting the \( p \)th column of the \( p \)th row of \( M \). Now, \( M'_1 \) would not, in general, be CNN but

\[ \tilde{M} = \begin{bmatrix}
M'_1 & 0 \\
0 & 0
\end{bmatrix} \]

is easily shown to be CNN; moreover, since \( M' \) and \( M \) are similar, \( \tilde{M} \) must have the same eigenvalues as \( M \). Finally, from our description of \( M'_1, \tilde{M} \) evidently has the desired form.

Now, suppose that \( u \neq 0 \). We can, therefore, use \( u \) to eliminate \( v \) by an “elementary column operation”; in particular, let

\[ S^{-1} = I - \frac{v}{u}E_pE_{p+1}^\dagger, \]

where \( I \) is the \( n \times n \) identity and \( E_i \) is the \( i \)th column of \( I \). We want to show that we may choose

\[ \tilde{M} = SMS^{-1}, \]

where

\[ S = I + \frac{v}{u}E_pE_{p+1}^\dagger. \]

Since \( p \geq k + 1 \), it is evident that \( \tilde{M} \) has the desired form; it remains now to show
that \( \tilde{M} \) is CNN. Since \( S \) is evidently CNN, we can and will verify the latter by showing that \( M' = MS^{-1} \) is CNN. Note that if \( M_i \) is the \( i \)th column of \( M \), then

\[
M' = [M_1 \cdots M_p M_{p+1} - (v/u)M_p M_{p+2} \cdots M_n].
\]

In showing that \( M' \) is CNN, we assert that we need only consider those minors, of which, say, \( \mu \) is an example, which satisfy the following conditions:

(a) \( \mu \) depends upon elements of the \((p + 1)\)th column of \( M' \) but not upon elements of the \( p \)th column.

(b) If \( \mu \) depends upon elements of the first \( k - 1 \) rows of \( M' \), then \( \mu \) depends upon elements of the first \( p - 1 \) columns.

If \( \mu \) did not satisfy (a), then by inspection of (3.1) and (3.4), \( \mu \) would be numerically equal to a minor of \( M' \); if \( \mu \) did not satisfy (b), then by inspection, \( \mu \) depends upon a null row of \( M \). In either of the latter cases, \( \mu \) would be nonnegative.

For brevity in the following, we introduce the Gantmacher notation: \( A(\alpha \beta \cdots) \) is that submatrix of the matrix \( A \) composed of elements from rows \( \alpha, \beta, \cdots \) and columns \( a, b, \cdots \) while \( \tilde{A}(\alpha \beta \cdots) \) is obtained by deleting row \( \alpha, \beta, \cdots \) and column \( a, b, \cdots \) from \( A \). Also, \( A(\cdot \cdot \cdot) = \det \{A(\cdot \cdot \cdot)\} \) and \( \tilde{A}(\cdot \cdot \cdot) = \det \{\tilde{A}(\cdot \cdot \cdot)\} \).

Let \( \mu \) be a minor of \( M' \) satisfying conditions (a) and (b), e.g.,

\[
\mu = M'[\alpha \beta \cdots c p + 1 d \cdots],
\]

where \( \alpha < \beta < \cdots \) and \( a < b < \cdots < c < p + 1 < d < \cdots \) and \( c \neq p \).

Note. Those minors of \( M' \) which depend only upon the columns, \( M'_i, i \geq p + 1 \), will be simple special cases of the following.

Now, from (3.4), (3.5) and a well-known determinantal property,

\[
\mu = M' \begin{bmatrix} \alpha \beta & \cdots & c & p & + & 1 & d & \cdots \\ a & b & \cdots & c & p & + & 1 & d & \cdots \end{bmatrix} - (v/u)M' \begin{bmatrix} \alpha & \beta & \cdots & c & p & + & 1 & d & \cdots \\ a & \cdots & c & p & + & 1 & d & \cdots \end{bmatrix},
\]

(3.6)

\[
= u^{-1}vM' \begin{bmatrix} \alpha & \beta & \cdots & c & p & + & 1 & d & \cdots \\ a & \cdots & c & p & + & 1 & d & \cdots \end{bmatrix} - vM' \begin{bmatrix} \alpha & \beta & \cdots & c & p & + & 1 & d & \cdots \\ a & \cdots & c & p & + & 1 & d & \cdots \end{bmatrix}.
\]

Let

\[
A = M' \begin{bmatrix} \alpha & \beta & \cdots & \gamma & k & \delta & \cdots \\ a & \cdots & c & p & + & 1 & d & \cdots \end{bmatrix},
\]

(3.7)

where, say, \( \gamma < k \leq \delta \).

Note. If \( \alpha > k \), then the first row of \( A \) would be composed of elements from the \( k \)th row of \( M' \); as will be seen, we lose no generality by supposing \( k > \alpha \).

For reference, we suppose that \( a_{i \cdot} = m_{i \cdot p} \). Then, from (3.6) and (3.7),

\[
\mu = v^{-1} \{a_{i \cdot} \tilde{A} \begin{bmatrix} t \\ i \end{bmatrix} - a_{i \cdot i+1} \tilde{A} \begin{bmatrix} t \\ i + 1 \end{bmatrix} \}.
\]

Thus, we must show that the quantity in brackets is nonnegative.

From (3.1) and (3.7),
where \( u' = a_{i,i} \), \( v' = a_{i,i+1} \). Since, with the possible exception of a “repeated” row, \( \mathbf{A} \) is a submatrix of \( \mathbf{M} \), \( \mathbf{A} \) is evidently \( \text{CNN} \). We require two lemmas, the second of which will readily imply that \( \mu \), as defined by (3.8), must be nonnegative when \( v > 0 \) and \( \mathbf{A} \) is \( \text{CNN} \) and has the form noted in (3.9).

The following lemma was proved in [3, p. 309]; for completeness, we offer a proof which does not require certain special results derived in [3].

**Lemma 1.** Let \( \mathbf{A} \) be \( \text{CNN} \). Then, for \( 1 \leq p \leq n \),

\[
(-1)^{p+1} \sum_{i=2}^{n} (-1)^{i+1} a_{i,i} \mathbf{A}_{i}^{1} \geq 0.
\]

**Proof.** In the case \( p = 1 \), the left-hand side of (3.10) is just \( \det(\mathbf{A}) \); in the case \( p = n \), the left-hand side reduces to \( a_{n} \mathbf{A}_{n}^{1} \). Since \( \mathbf{A} \) is \( \text{CNN} \), (3.1) is evidently valid for these cases.

Assume now that \( 1 < p < n \). Let \( s \) and \( i \) be chosen such that \( 1 \leq s < p < i \leq n \) and suppose that, for all such pairs \((s, i)\) and all \( k \) such that \( 2 \leq k \leq n \),

\[
\mathbf{A}_{i}^{k} = 0.
\]

Then, for all \( i > p \),

\[
\mathbf{A}_{i}^{k} = \sum_{s=2}^{n} (-1)^{k+s-1} a_{s,i} \mathbf{A}_{s}^{1} = 0
\]

and (3.10) would reduce to the known inequality, \( a_{s} \mathbf{A}_{s}^{1} \geq 0 \).

Assume that for some choice of \( s \), \( i \), and \( k \), restricted as above, that

\[
\mathbf{A}_{s}^{k} = 0.
\]

Let \( \mathbf{N} \) be the matrix obtained when the elements \( m_{1}, m_{2}, \cdots, m_{p-1} \) of \( \mathbf{A} \) are replaced by zeros. (3.10) is then equivalent to the assertion that

\[
(-1)^{p+1} \det(\mathbf{N}) \geq 0.
\]

(3.12) is evidently true when \( n = 2 \); we make the usual inductive hypothesis that (3.12) is valid for all \( \mathbf{N} \) of dimension less than \( n \). Now, from Sylvester’s identity (see, e.g., [2, p. 33]),

\[
\det(\mathbf{N}) \mathbf{R}_{i}^{k} = \mathbf{R}_{i}^{k} \mathbf{R}_{s}^{k} - \mathbf{R}_{i}^{k} \mathbf{R}_{s}^{k}.
\]
or since all rows, except the first, of $A$ and $N$ are identical and noting (3.11),

$$\det(N) = \left( A \begin{bmatrix} 1 & k \\ s & i \end{bmatrix} \right)^{-1} \left( A \begin{bmatrix} 1 & k \\ s & i \end{bmatrix} N \begin{bmatrix} k \\ s \end{bmatrix} - A \begin{bmatrix} 1 \\ s \end{bmatrix} N \begin{bmatrix} k \\ s \end{bmatrix} \right).$$

Now $\tilde{N}_1^i$ (and $\tilde{N}_2^i$) can be obtained by replacing the first $p-1$ (and $p-2$) elements of the first row of $A_1^i$ (and $A_2^i$) with zeros; by the inductive hypothesis

$$(-1)^{p+1} N_i^k \geq 0, \quad (-1)^{(p-1)+1} N_s^k \geq 0.$$

(3.12), and therefore (3.10), now follow readily from (3.13) and (3.14) and the fact that $A$ is CNN which completes our proof.

The following lemma now generalizes the result of Lemma 1 for the case that $A$ has a form such as in (3.9).

**Lemma 2.** Suppose that $A$ is CNN and that

$$a_{ii} = 0, \quad i = 1, \ldots, t - 1, j = s, \ldots, n,$$

for some $s$ and $t$ satisfying $1 \leq s$, $t \leq n$. Then, for $p \geq s$,

$$(-1)^{t+p} \sum_{i=p}^{n} (-1)^{t+i} a_{ii} \tilde{A}_i^t \geq 0.$$

**Proof.** Assume initially that $s > t - 1$. Define

$$r_{ii} = A \begin{bmatrix} 1 & \cdots & t - 1 & i + t - 2 \\ 1 & \cdots & t - 1 & j + t - 1 \end{bmatrix}$$

and let $R = [r_{ii}]^{t-1}_{i=t-1}$.

Again utilizing Sylvester's identity,

$$R \begin{bmatrix} e & \eta & \cdots & \rho \\ e & f & \cdots & g \end{bmatrix} = \Delta^{t-1} A \begin{bmatrix} 1 & \cdots & t - 1 & e + t - 1 & \cdots & \rho + t - 1 \\ 1 & \cdots & t - 1 & e + t - 1 & \cdots & g + t - 1 \end{bmatrix},$$

presuming that the latter minor is $q$th order and

$$\Delta = A \begin{bmatrix} 1 & \cdots & t - 1 \\ 1 & \cdots & t - 1 \end{bmatrix}.$$

Evidently, $R$ is CNN. From Lemma 1, then follows

$$(-1)^{t+1} \sum_{j=s}^{n-(t-1)} (-1)^{t+j} r_{ii} R \begin{bmatrix} 1 \\ j \end{bmatrix} \geq 0,$$

whenever $1 \leq q \leq n - (t - 1)$.

Now, from (3.15) and (3.17),

$$r_{ii} = A \begin{bmatrix} 1 & \cdots & t - 1 & t \\ 1 & \cdots & t - 1 & j + t - 1 \end{bmatrix} = a_{i,i+t-1} \Delta,$$

whenever $j \geq s - (t - 1)$; from (3.18),

$$R \begin{bmatrix} 1 \\ j \end{bmatrix} = \Delta^{t-1} A \begin{bmatrix} 1 \\ j + t - 1 \end{bmatrix}.$$
Utilizing these last two relations and (3.19) yields, after some simplification,
\[(3.20) \quad A'(-t')c'(-t') + i(t')c'(-t') = 0,\]
whenever \(p \geq s > t - 1\). If \(\Delta > 0\), then (3.20) reduces to (3.16). Suppose, however, that \(\Delta = 0\); the inequality
\[(3.21) \quad A[t - j] \leq \Delta A[1 \ldots t - 1 \ldots j], \quad j > t - 1,\]
is a special case of a result due to [2, II, p. 100]. Evidently, \(\Delta = 0\) would imply the equality in (3.16) for the case \(j \geq p \geq s > t - 1\).

Finally, suppose that \(s \leq t - 1\). Then, \(A[i^t] = 0\), since \(A[i^t] = 0\) has a column of zeros. Then, as in (3.21),
\[A[t - j] \leq A[1 \ldots s]A[1 \ldots s \ldots j] = 0, \quad j > s.\]
Therefore, (3.16) either reduces to an equality (when \(p > s\)) or to the known inequality, \(A[i^t] \geq 0\) (when \(p = s\)). This completes our proof of the lemma.

From (3.9) and (3.16), then follows
\[0 \leq (-1)^{t+1} \sum_{k=1}^{n} (-1)^{t+i} a_{t, i} A[t], \]
and therefore, from (3.8), \(\mu \geq 0\), which completes our proof of the primary assertion of the Basic Lemma.

Noting that we choose \(\tilde{M}\) similar to \(M\) as long as \(M\) does not have a column of zeros, the second assertion of the Basic Lemma is now obvious.

Finally, as in all elementary similarity transformations of the form (3.3), \(\tilde{M}\) will be upper Hessenberg as long as \(M\) is upper Hessenberg.

Department of Mathematics
University of Tulsa
Tulsa, Oklahoma 74104

Department of Mathematics
Rensselaer Polytechnic Institute
Troy, New York 12181