ADVANCED METHODS IN MATRIX ANALYSIS

PLAMEN KOEV

Notation

We consider $n$-by-$n$ square matrices, e.g.,

$$A = [a_{ij}]_{i,j=1}^n = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$$

The submatrices consisting of rows $i_1, i_2, \ldots, i_k$ and columns $j_1, j_2, \ldots, j_k$ will be denoted using MATLAB notation as

$$A([i_1, i_2, \ldots, i_k], [j_1, j_2, \ldots, j_k]) = \begin{bmatrix} a_{i_1,j_1} & a_{i_1,j_2} & \ldots & a_{i_1,j_k} \\ a_{i_2,j_1} & a_{i_2,j_2} & \ldots & a_{i_2,j_k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_k,j_1} & a_{i_k,j_2} & \ldots & a_{i_k,j_k} \end{bmatrix}.$$ 

The determinants

$$\det A([i_1, i_2, \ldots, i_k], [j_1, j_2, \ldots, j_k]) = \det A(i_1:k, j_1:k)$$

are called minors of $A$. In particular,

$$\det A(1:k, 1:k), k = 1, 2, \ldots, n,$$

are called leading principal minors.

1. Gaussian Elimination

The process of Gaussian elimination is a fundamental tool in solving linear systems of equations.

Example 1.1. Consider the linear system:

$$\begin{align*}
  x_1 + x_2 + x_3 &= 3 \\
  x_1 + 2x_2 + 4x_3 &= 7 \\
  x_1 + 3x_2 + 9x_3 &= 13.
\end{align*}$$

(1)

The traditional way of solving this system is to subtract the first equation from the second and the third to obtain

$$\begin{align*}
  x_1 + x_2 + x_3 &= 3 \\
  x_2 + 3x_3 &= 4 \\
  2x_2 + 8x_3 &= 12.
\end{align*}$$

Date: November 7, 2006.
Now subtract 2 times the second equation from the third to obtain
\[ x_1 + x_2 + x_3 = 3 \]
\[ x_2 + 3x_3 = 4 \]
\[ 2x_3 = 2. \]

Now we can perform back substitution:
\[ x_3 = 1 \]
\[ x_2 = 3 - 3x_3 = 4 - 3 \cdot 1 = 1 \]
\[ x_1 = 3 - x_2 - x_3 = 3 - 1 - 1 = 1. \]

Instead of performing the same process for every right hand side, it is more advantageous to use matrix factorizations instead. Write the system (1) as
\[
\begin{bmatrix}
1 & 1 & 1 \\
1 & 2 & 4 \\
1 & 3 & 9 \\
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
\end{bmatrix}
= 
\begin{bmatrix}
3 \\
7 \\
13 \\
\end{bmatrix}.
\]

In general linear systems are written as
\[ Ax = b \]
or
\[
\begin{bmatrix}
a_{11} & a_{12} & \ldots & a_{1n} \\
a_{21} & a_{22} & \ldots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \ldots & a_{nn} \\
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n \\
\end{bmatrix}
= 
\begin{bmatrix}
b_1 \\
b_2 \\
\vdots \\
b_n \\
\end{bmatrix},
\]
where we assume that leading principal minors \( A(1 : k, 1 : k), k = 1, 2, \ldots, n \), of the \( n \)-by-\( n \) matrix \( A = [a_{ij}]_{i,j=1}^n \) are nonzero.

Some linear systems are easy to solve. For example if \( A \) is triangular or diagonal. If \( A \) is (lower or upper) triangular nonsingular matrix, then \( Ax = b \) can be solved via back substitution. The system
\[
\begin{bmatrix}
a_{11} & a_{12} & \ldots & a_{1n} \\
a_{21} & a_{22} & \ldots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \ldots & a_{nn} \\
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n \\
\end{bmatrix}
= 
\begin{bmatrix}
b_1 \\
b_2 \\
\vdots \\
b_n \\
\end{bmatrix},
\]
(the zero entries of the upper triangular part have been omitted) is equivalent to
\[ a_{11}x_1 = b_1 \]
\[ a_{21}x_1 + a_{22}x_2 = b_2 \]
\[ \ldots \]
\[ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \]
and is solved by computing \( x_1 \) from the first equation, substituting into the second and so on:
\[ x_1 = \frac{b_1}{a_{11}} \]
\[ x_2 = \frac{b_2 - a_{21}x_1}{a_{22}} \]
\[ \ldots \]
\[ x_n = \frac{b_n - a_{n1}x_1 - a_{n2}x_2 - \ldots - a_{n,n-1}x_{n-1}}{a_{nn}}. \]
The solution to a diagonal linear system is trivial:
\[
\begin{bmatrix}
d_1 \\
d_2 \\
\vdots \\
d_n \\
\end{bmatrix} 
\cdot 
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n \\
\end{bmatrix} 
= 
\begin{bmatrix}
b_1 \\
b_2 \\
\vdots \\
b_n \\
\end{bmatrix}
\]
implies \( x_i = \frac{b_i}{d_i} \), \( i = 1, 2, \ldots, n \).

**Definition 1.2.** A matrix \( A \) is called **unit lower triangular** if \( a_{ij} = 0 \), \( 1 \leq i < j \leq n \) and \( a_{ii} = 1 \), \( 1 \leq i \leq n \).

An **unit upper triangular** matrix is defined analogously.

**Example 1.3.** The following 4-by-4 matrix is unit lower triangular
\[
\begin{bmatrix}
1 & 2 & 1 & 4 \\
3 & 5 & 1 & 6 \\
7 & 9 & 27 & 1 \\
\end{bmatrix}
\]

**Exercise 1.4.** Prove that if \( A \) and \( B \) are unit lower triangular matrices, then so are \( A^{-1} \) and \( AB \).

**Definition 1.5.** Let \( A \) be a nonsingular matrix. A decomposition of \( A \) as a product of a unit lower triangular matrix \( L \), a diagonal matrix \( D \), and a unit upper triangular matrix \( U \):
\[
A = LDU
\]
is called an **LDU decomposition** of \( A \).

The main idea in what follows is to use Gaussian elimination to compute the LDU decomposition of \( A \).

Once we have the LDU decomposition of \( A \), the equation \( Ax = b \) becomes \( LDUx = b \), which is easy to solve. First compute the solution \( y \) to the lower triangular system \( Ly = b \), then the solution \( z \) to the diagonal system \( Dy = y \), and finally the solution \( x \) to the upper triangular system \( Ux = z \). Finally,
\[
Ax = LDUx = LD(Ux) = L(Dy) = Ly = b,
\]
as desired.

So how does one compute the LDU decomposition of a nonsingular matrix \( A \)?

First we represent a subtraction of a multiple of one row from another in matrix form. Consider the matrix:
\[
\begin{bmatrix}
1 & 1 & 1 \\
1 & 2 & 4 \\
3 & 9 & 27 \\
\end{bmatrix}
\]

In order to introduce a zero in position \((3,1)\) we need to subtract 3 times the first row from the third. This is equivalent to multiplication by the matrix
\[
\begin{bmatrix}
1 & 0 & 1 \\
0 & 1 & 0 \\
-3 & 0 & 1 \\
\end{bmatrix}
\]
namely

\[
\begin{bmatrix}
1 & 1 & 1 \\
0 & 1 & 1 \\
-3 & 0 & 1
\end{bmatrix} \cdot \begin{bmatrix}
1 & 1 & 1 \\
1 & 2 & 1 \\
3 & 9 & 27
\end{bmatrix} = \begin{bmatrix}
1 & 1 & 1 \\
1 & 2 & 1 \\
0 & 6 & 24
\end{bmatrix}.
\]

Since

\[
\begin{bmatrix}
1 & 0 & 1 \\
-3 & 0 & 1
\end{bmatrix}^{-1} = \begin{bmatrix}
1 & 0 & 1 \\
0 & 1 & 3
\end{bmatrix},
\]

the equality (2) implies

\[
\begin{bmatrix}
1 & 1 & 1 \\
1 & 2 & 1 \\
3 & 9 & 27
\end{bmatrix} = \begin{bmatrix}
1 & 1 & 1 \\
0 & 1 & 3 \\
0 & 6 & 24
\end{bmatrix} \cdot \begin{bmatrix}
1 & 1 & 1 \\
1 & 2 & 1 \\
3 & 0 & 1
\end{bmatrix}.
\]

Next, subtract the first row from the second to analogously obtain

\[
\begin{bmatrix}
1 & 1 & 1 \\
1 & 2 & 1 \\
3 & 9 & 27
\end{bmatrix} = \begin{bmatrix}
1 & 1 & 1 \\
0 & 1 & 3 \\
0 & 6 & 24
\end{bmatrix} \cdot \begin{bmatrix}
1 & 1 & 1 \\
1 & 1 & 1 \\
3 & 0 & 1
\end{bmatrix}.
\]

Now observe that the matrices used for elimination combine very nicely:

\[
\begin{bmatrix}
1 & 1 & 1 \\
0 & 1 & 3 \\
3 & 0 & 1
\end{bmatrix} \cdot \begin{bmatrix}
1 & 1 & 1 \\
0 & 1 & 3 \\
0 & 6 & 24
\end{bmatrix} = \begin{bmatrix}
1 & 1 & 1 \\
1 & 1 & 1 \\
3 & 0 & 1
\end{bmatrix},
\]

therefore

\[
\begin{bmatrix}
1 & 1 & 1 \\
1 & 2 & 1 \\
3 & 9 & 27
\end{bmatrix} = \begin{bmatrix}
1 & 1 & 1 \\
1 & 1 & 1 \\
3 & 0 & 1
\end{bmatrix} \cdot \begin{bmatrix}
1 & 1 & 1 \\
0 & 1 & 3 \\
0 & 6 & 24
\end{bmatrix}.
\]

Then continue by induction—subtract 6 times the second row from the third, obtaining the decomposition

\[
\begin{bmatrix}
1 & 1 & 1 \\
0 & 1 & 3 \\
0 & 6 & 24
\end{bmatrix} = \begin{bmatrix}
1 & 1 & 1 \\
0 & 1 & 3 \\
0 & 6 & 24
\end{bmatrix} \cdot \begin{bmatrix}
1 & 1 & 1 \\
0 & 1 & 3 \\
0 & 6 & 24
\end{bmatrix}.
\]

Therefore

\[
\begin{bmatrix}
1 & 1 & 1 \\
1 & 2 & 1 \\
3 & 9 & 27
\end{bmatrix} = \begin{bmatrix}
1 & 1 & 1 \\
0 & 1 & 3 \\
0 & 6 & 24
\end{bmatrix} \cdot \begin{bmatrix}
1 & 1 & 1 \\
0 & 1 & 3 \\
0 & 6 & 24
\end{bmatrix}.
\]
Algorithm 1.6 (Gaussian Elimination). The following algorithm computes the LDU decomposition of a matrix $A$ whose leading principal minors are nonzero.

$U = A$, $L = I$, $D = I$

for $i = 1 : n - 1$

for $j = i + 1 : n$

\[ l_{ji} = u_{ji}/u_{ii} \]

\[ u_{j,i:n} = u_{j,i:n} - l_{ji}u_{i,n} \]

endfor

\[ d_{ii} = u_{ii} \]

\[ u_{i,n} = u_{i,n}/d_{ii} \]

endfor

Let the entries of $L$, $D$, and $U$ be $l_{ij}$, $d_i$, $1 \leq i \leq n$, and $u_{ij}$, $i < j$. Next, we obtain formulas for $l_{ij}$, $d_i$, and $u_{ij}$ in terms of the minors of $A$.

If we performed Gaussian elimination on the leading submatrix $A(1 : k, 1 : k)$ only, we would obtain

\[ A(1 : k, 1 : k) = L(1 : k, 1 : k) \cdot D(1 : k, 1 : k) \cdot U(1 : k, 1 : k). \]

Since $L$ and $U$ are both unit lower and upper triangular, respectively, we have

\[ \det A(1 : k, 1 : k) = d_1d_2 \cdots d_k. \]

Therefore

\[ d_k = \frac{\det A(1 : k, 1 : k)}{\det A(1 : k - 1, 1 : k - 1)}. \]

Since subtracting a multiple of a row from another does not change the value of a determinant, we conclude that the values of all minors of $A$ and $DU$ are the same, in particular

\[ \det A(1 : k, [1 : k - 1, j]) = d_1d_2 \cdots d_k u_{kj}. \]

Therefore

\[ u_{kj} = \frac{\det A(1 : k, [1 : k - 1, j])}{\det A(1 : k, 1 : k)}, \quad k < j. \]

Analogously,

\[ l_{kj} = \frac{\det A([1 : k - 1, j], 1 : k)}{\det A(1 : k, 1 : k)}, \quad k > j. \]

Theorem 1.7. If the leading principal minors $A(1 : k, 1 : k)$, $k = 1, 2, \ldots, n$, of $A$ are nonzero, then its LDU decomposition exists and is unique.

Proof. Existence: Algorithm 1.6 completes successfully and computes an LDU decomposition of $A$ as long as there are no divisions by zero. Divisions by zero cannot occur because $d_i = \frac{\det A(1 : k, 1 : k)}{\det A(1 : k - 1, 1 : k - 1)} \neq 0$.

Uniqueness: Suppose $A = LDU = L_1D_1U_1$. Then

\[ L_1^{-1}L = D_1U_1U^{-1}D^{-1}. \]

The matrix $L_1^{-1}L$ is unit lower triangular (see Exercise 1.4) and $D_1U_1U^{-1}D^{-1}$ is upper triangular. This is only possible if they are both diagonal. Since $L_1^{-1}L$ has ones on the main diagonal, we must have $L_1^{-1}L = I$, i.e., $L = L_1$. Analogously, $U = U_1$. Finally, $D_1D^{-1} = I$ implies $D = D_1$. \qed
2. Symmetric Positive Definite Matrices

A matrix $A$ is symmetric positive definite (s.p.d.) if it is symmetric, $A^T = A$, $x^T Ax \geq 0$ for every $x$, and $x^T Ax = 0$ only when $x = 0$.

If $A$ is s.p.d., then all eigenvalues of $A$ are positive and all leading principal minors $A(1 : k, 1 : k) > 0$, $k = 1, 2, \ldots, n$.

Let $A = LDU$ be the LDU decomposition of an s.p.d. matrix $A$. Then

$U^T D^T L^T = U^T D L^T = A^T = A = LDU$.

Since $U^T$ and $L^T$ are unit lower and upper triangular matrices, respectively, we obtain two LDU decompositions of $A$:

$A = U^T DL^T$ and $A = LDU$.

From Theorem 1 these decompositions must be the same, therefore $U^T = L$. Finally,

$A = LDL^T$.

We can also write $A = LDL = (LD^{1/2})(D^{1/2}L^T) = CCT$, where $C = LD^{1/2}$ is lower triangular (all elements of $D$, $d_i = \frac{\det A(1:k,1:k)}{\det A(1:k-1,k-1)} > 0$, so we can safely form $D^{1/2}$).

**Definition 2.1.** The decomposition

$A = CCT$

of an s.p.d. matrix as a product of a nonsingular lower triangular matrix and its transpose is called Cholesky decomposition.

**Theorem 2.2.** A matrix $A$ is s.p.d. if and only if it has a Cholesky decomposition.

**Proof.** If $A$ is s.p.d., then it has a Cholesky decomposition as we described above.

If $A = CCT$, where $C$ is nonsingular, let $y = CTx$. Then

$x^T Ax = x^T CCT x = (CTx)^T C^T x = y^T y = y_1^2 + y_2^2 + \cdots + y_n^2 \geq 0$,

with equality only when $y = 0$, i.e., only when $x = 0$ (since $y = CTx$ and $C^T$ is nonsingular).

**Example 2.3.**

\[
\begin{bmatrix}
1 & 2 & 3 \\
2 & 13 & 18 \\
3 & 18 & 50
\end{bmatrix} = \begin{bmatrix}
1 & 2 & 3 \\
2 & 3 & \\
3 & 4 & 5
\end{bmatrix} \cdot \begin{bmatrix}
1 & 2 & 3 \\
 & 4 & \\
 & & 5
\end{bmatrix}.
\]

**Exercise 2.4.** If $A$ is s.p.d., then $a_{ii}a_{jj} > |a_{ij}^2|$, $i \neq j$. 
3. POSITIVE AND NONNEGATIVE MATRICES

**Definition 3.1.** A matrix $A$ is called positive (nonnegative) if $a_{ij} > 0$ ($a_{ij} \geq 0$).

**Definition 3.2.** The set of eigenvalues $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ of an $n$-by-$n$ matrix $A$ is called spectrum of $A$ and denoted $\sigma(A)$.

**Definition 3.3.** The spectral radius of $A$ is the nonnegative number

$$\rho(A) = \max\{|\lambda_i|, \lambda_i \in \sigma(A)\}.$$

**Definition 3.4.** A matrix norm $\|\cdot\|$ is called an operator norm if $\|AB\| \leq \|A\| \cdot \|B\|$.

**Theorem 3.5.** If $\|\cdot\|$ is an operator norm, then $\rho(A) \leq \|A\|$.

*Proof.* Let $\lambda$ be an eigenvalue such that $|\lambda| = \rho(A)$. Let $x$ be an eigenvector corresponding to the eigenvalue $\lambda$ and let $X$ be a matrix consisting of $n$ copies of the vector $x$. Then $AX = \lambda X$ and

$$|\lambda| \cdot \|X\| = \|AX\| = \|A \cdot X\| \leq \|A\| \cdot \|X\|.$$

Thus $\rho(A) \leq \|A\|$.

**Theorem 3.6.** $\lim_{k \to \infty} A^k = 0$ if and only if $\rho(A) < 1$.

*Proof.* If $\lim_{k \to \infty} \rho(A) = 0$, then

$$\lim_{k \to \infty} (\rho(A))^k = \lim_{k \to \infty} (\rho(A^k)) = \lim_{k \to \infty} \|A^k\| = 0,$$

thus $\rho(A) < 1$.

Conversely, if $\rho(A) < 1$, then there exists an $\epsilon > 0$ such that $\rho(A) + \epsilon < 1$. Also, there exists a similarity transformation $S$ such that $S^{-1}AS = A$ and $A_{i,i+1} \leq \epsilon$.

Finally,

$$\lim_{k \to \infty} \|A^k\|_{\infty} \leq \|S\| \cdot \|S^{-1}\| \cdot \lim_{k \to \infty} \|A^k\|_{\infty} = \|S\| \cdot \|S^{-1}\| \cdot \lim_{k \to \infty} (\rho(A) + \epsilon)^k = 0.$$  

**Theorem 3.7.** Let $\|\cdot\|$ be an operator norm. Then

$$\rho(A) = \lim_{k \to \infty} \|A^k\|^{1/k}.$$

*Proof.* We have $\rho(A)^k \leq \rho(A^k) \leq \|A^k\|$. In the other direction, let $\epsilon > 0$ be given, then $B \equiv (\rho(A) + \epsilon)^{-1}A$ has spectral radius $\rho(B) = (\rho(A) + \epsilon)^{-1}\rho(A) < 1$, thus $\lim_{k \to \infty} \|B^k\| = 0$. Therefore there exists a number $N$ such that $\|B^k\| < 1$ for $k > N$, which means that for $k > N$, $\|A^k\| \leq (\rho(A) + \epsilon)^k$. Finally, for $k > N$ and arbitrary $\epsilon$ we have $\|A^k\|^{1/k} \leq \rho(A) + \epsilon \leq \|A^k\|^{1/k} + \epsilon$ and we are done.

**Definition 3.8.** The matrix $C = |A|$ is defined as $c_{ij} = |a_{ij}|, i, j = 1, 2, \ldots, n$. The notation $A \leq B$ means $a_{ij} \leq b_{ij}, i, j = 1, 2, \ldots, n$.

**Exercise 3.9.** The following are true:

1. $|Ax| \leq |A| \cdot |x|$
2. $|AB| \leq |A| \cdot |B|$
3. $|A^m| \leq |A|^m$
4. $0 \leq A \leq B$ and $0 \leq C \leq D$ imply $0 \leq AC \leq BD$
5. $A \geq 0$ implies $A^m \geq 0$ and $A > 0$ implies $A^m > 0$, $m > 0$.
6. $A > 0, x \geq 0, x \neq 0$ imply $Ax > 0$
(7) $|A| \leq |B|$ imply $\|A\|_\infty \leq \|B\|_\infty$.
(8) $\|A\|_\infty = \|\|A\|_\infty$

**Theorem 3.10.** If $|A| \leq B$, then $\rho(A) \leq \rho(|A|) \leq \rho(B)$.

*Proof.* From the above exercise we have $|A|^m \leq |A|^m \leq B^m$,
\[
\|A^m\|_\infty \leq \|\|A^m\|_\infty \leq \|B^m\|_\infty \text{ and } \|A^m\|^{1/m}_\infty \leq \|\|A^m\|^{1/m}_\infty \leq \|B^m\|^{1/m}_\infty.
\]
Now we let $m \to \infty$ and use Theorem 3.7. □

**Corollary 3.11.** $0 \leq A \leq B$ imply $\rho(A) \leq \rho(B)$.

**Corollary 3.12.** If $B$ is any principal submatrix of $A > 0$, then $\rho(B) \leq \rho(A)$.

**Lemma 3.13.** If the row sums of $A$ are constant, then $\rho(A) = \|A\|_\infty$. If the column sums of $A$ are constant, then $\rho(A) = \|A\|_1$.

*Proof.* Say the row sums of $A$ are constant (and thus they equal $\|A\|_\infty$). Then $\rho(A) \leq \|A\|_\infty$. On the other side the vector $(1,1,\ldots,1)^T$ is an eigenvector with eigenvalue $\|A\|_\infty$. If the column sums of $A$ are constant, we apply the same argument to $A^T$. □

**Theorem 3.14.** If $A \geq 0$, then
\[
\min_i \sum_{j=1}^n a_{ij} \leq \rho(A) \leq \max_i \sum_{j=1}^n a_{ij}
\]
and
\[
\min_j \sum_{i=1}^n a_{ij} \leq \rho(A) \leq \max_j \sum_{i=1}^n a_{ij}.
\]

*Proof.* Construct a matrix $B \leq A$ with row sums $\alpha \equiv \min_j \sum_{j=1}^n a_{ij}$. For example take $b_{ij} = a_{ij} \alpha / \sum_{j=1}^n a_{ij}$. From the previous lemma $\rho(B) = \alpha$ and from $0 \leq B \leq A$ we have $\alpha = \rho(B) \leq \rho(A)$. The rest is proven analogously. □

**Corollary 3.15.** If $A > 0$, then $\rho(A) > 0$.

We now generalize Theorem 3.14.

**Theorem 3.16.** If $A \geq 0$ and $x > 0$
\[
\min_i \frac{1}{x_i} \sum_{j=1}^n a_{ij} x_j \leq \rho(A) \leq \max_i \frac{1}{x_i} \sum_{j=1}^n a_{ij} x_j
\]
and
\[
\min_j x_j \sum_{i=1}^n \frac{a_{ij}}{x_i} \leq \rho(A) \leq \max_j x_j \sum_{i=1}^n \frac{a_{ij}}{x_i}.
\]

*Proof.* We have $\rho(S^{-1}AS) = \rho(A)$. If $A \geq 0$ and $S = \text{diag}(x_1, \ldots, x_n)$, $x_i > 0$, $i = 1, 2, \ldots, n$, then $S^{-1}AS \geq 0$. We apply Theorem 3.14 to $S^{-1}AS = [a_{ij}x_jx_i^{-1}]_{i,j=1}^n$ to finish the proof. □

**Corollary 3.17.** Let $A \geq 0$ and $x > 0$. If $\alpha, \beta > 0$ are such that $\alpha x \leq Ax \leq \beta x$, then $\alpha \leq \rho(A) \leq \beta$. Moreover if $\alpha x < Ax$, then $\alpha < \rho(A)$. If $Ax < \beta x$, then $\rho(A) < \beta$. 

Corollary 3.23. \( \beta \leq \alpha \). If \( \alpha x \leq Ax \), then \( \alpha \leq \min_{i \leq n} x_i^{-1} \sum_{j=1}^{n} a_{ij} x_j \) and by the above theorem, \( \alpha \leq \rho(A) \). If \( \alpha x < Ax \), then there is \( \alpha' > \alpha \) such that \( \alpha' x \leq Ax \) for which \( \alpha' \leq \rho(A) \), thus \( \alpha < \rho(A) \). \( \square \)

**Lemma 3.18.** Let \( A > 0 \), \( Ax = \lambda x \) and \( |\lambda| = \rho(A) \). Then \( A|x| = \rho(A)|x| \) and \( |x| > 0 \).

*Proof.*

\( \rho(A)|x| = |\lambda||x| = |\lambda x| = |Ax| \leq |A||x| = A|x| \)

therefore \( y = A|x| - \rho(A)|x| \geq 0 \). Since \( |x| \geq 0, x \neq 0 \) we have \( A|x| > 0 \). Since \( A > 0 \) we have \( \rho(A) > 0 \). If \( y = 0 \), then \( A|x| = \rho(A)|x| \) and \( |x| = (\rho(A))^{-1}A|x| > 0 \).

If \( y \neq 0 \), let \( z = A|x| > 0 \). Then

\[ 0 < Ay = Az - \rho(A)z \]

and \( Az > \rho(A)z \). Therefore \( \rho(A) > \rho(A) \), which is a contradiction. \( \square \)

**Theorem 3.19.** Let \( A > 0 \). Then \( \rho(A) > 0 \), \( \rho(A) \) is an eigenvalue of \( A \), and there is a positive eigenvector \( x \) corresponding to \( \rho(A) \).

*Proof.* Let \( Ax = \lambda x \), where \( |\lambda| = \rho(A) \). Then according to Lemma 3.18, \( A|x| = \rho(A)|x| \), i.e., \( \rho(A) \) is an eigenvalue with an eigenvector \( |x| > 0 \). \( \square \)

**Lemma 3.20.** Let \( A > 0, Ax = \lambda x, x \neq 0 \), and \( |\lambda| = \rho(A) \). Then \( |x| = e^{-i\theta}x \) for some real number \( \theta \).

*Proof.* We have \( |Ax| = |\lambda x| = \rho(A)|x| \). From Lemma 3.18 we have \( A|x| = \rho(A)|x| \).

Therefore \( A|x| = |Ax| \), i.e.,

\[ \left| \sum_{j=1}^{n} a_{ij} x_j \right| = \sum_{j=1}^{n} a_{ij} |x_j| \]

This is only possible if all \( x_j \) lie on the same ray in the complex plane. We denote the common argument by \( \theta \). Therefore \( e^{-i\theta}a_{ij}x_j > 0 \) for all \( j = 1, 2, \ldots, n \). Since \( a_{ij} > 0 \), we have \( e^{-i\theta}x > 0 \). \( \square \)

**Theorem 3.21.** Let \( A > 0 \), then \( |\lambda| < \rho(A) \) for every eigenvalue \( \lambda \neq \rho(A) \).

*Proof.* We have \( |\lambda| \leq \rho(A) \) for every eigenvalue \( \lambda \) of \( A \). If \( |\lambda| = \rho(A) \) and \( Ax = \lambda x, x \neq 0 \), then according to Lemma 3.20, \( w \equiv e^{-i\theta}x > 0 \) for some real argument \( \theta \), so \( Aw = \lambda w \). Now \( A > 0 \) and \( w > 0 \) imply \( \lambda > 0 \), i.e., \( \lambda = \rho(A) \). \( \square \)

**Theorem 3.22.** Let \( A > 0, Aw = \rho(A)w \) and \( Az = \rho(A)z \). Then \( w = az \) for some real number \( a \).

*Proof.* By Lemma 3.20, there exist \( \theta \) and \( \eta \) such that \( p \equiv e^{-i\theta}z > 0 \) and \( q \equiv e^{-i\eta}w > 0 \). Let

\[ \beta = \min_{1 \leq i \leq n} q_i \]

(say \( \beta = q_k/p_k \)), and define \( r \equiv q - \beta p \geq 0 \). Then \( Ar = Aq - \beta Ap = \rho(A)(q - \beta p) = \rho(A)r \). Now \( r \geq 0 \) and \( A > 0 \) imply \( r = (\rho(A))^{-1}Ar > 0 \), which is a contradiction with \( r_k = 0 \). Therefore \( r = 0 \), hence \( q = \beta p \) and \( w = \beta e^{i(\theta - \eta)}z \). \( \square \)

**Corollary 3.23.** The geometric multiplicity of \( \rho(A) \) as an eigenvalue of \( A \) is one.
Theorem 3.24. The algebraic multiplicity of $\rho(A)$ as an eigenvalue of $A$ is also one.

Proof. (Due to Froilán Dopico) Let $x$ and $y$ be left and right eigenvectors of $A$ corresponding to $\rho(A)$. If the algebraic multiplicity of $\rho(A)$ is greater than one, then $x^T y = 0$ (since the geometric multiplicity of $\rho(A)$ is one). This is a contradiction with $|x| > 0$ and $|y| > 0$. □

Corollary 3.25. If $A > 0$, then there exists a unique vector $x > 0$ such that $Ax = \rho(A)x$, and $\sum_{i=1}^n x_i = 1$.

Definition 3.26. The unique vector of Corollary 3.25 is called Perron vector of $A > 0$.

Theorem 3.27 (Perron). If $A > 0$, then

1. $\rho(A) > 0$;
2. $\rho(A)$ is an eigenvalue of $A$;
3. There is an $x > 0$ such that $Ax = \rho(A)x$;
4. $\rho(A)$ is a simple eigenvalue of $A$;
5. $|\lambda| < \rho(A)$ for every eigenvalue $\lambda \neq \rho(A)$. 
4. Formula of Cauchy–Binet

Let \( C = AB \) where \( A \) and \( B \) are \( m \)-by-\( n \) and \( n \)-by-\( m \), respectively, \( m \leq n \), i.e.,

\[
(c_{ij}) = \sum_{k=1}^{n} a_{ik} b_{kj}.
\]

Theorem 4.1 (Formula of Cauchy–Binet).

\[
\det C = \sum_{1 \leq k_1 < \cdots < k_m \leq n} \det A(1 : m, k_1 : m) \cdot \det B(k_1 \ldots m, 1 : m).
\]

Proof. Using (3) we get

\[
\det C = \det \begin{bmatrix}
\sum_{\alpha_1=1}^{n} a_{1\alpha_1} b_{\alpha_11} & \cdots & \sum_{\alpha_m=1}^{n} a_{1\alpha_m} b_{\alpha_m m} \\
\vdots & \ddots & \vdots \\
\sum_{\alpha_1=1}^{n} a_{m\alpha_1} b_{\alpha_11} & \cdots & \sum_{\alpha_m=1}^{n} a_{m\alpha_m} b_{\alpha_m m}
\end{bmatrix}
\]

\[
= \sum_{\alpha_1, \ldots, \alpha_m=1}^{n} \det \begin{bmatrix}
a_{1\alpha_1} b_{\alpha_11} & \cdots & a_{1\alpha_m} b_{\alpha_m m} \\
\vdots & \ddots & \vdots \\
a_{m\alpha_1} b_{\alpha_11} & \cdots & a_{m\alpha_m} b_{\alpha_m m}
\end{bmatrix}
\]

(4)

A summand in (4) is nonzero only if all \( \alpha_i \) are different. Therefore

\[
\det C = \sum_{1 \leq k_1 < \cdots < k_m \leq n} \sum_{\alpha_{1 : m} = \text{permutation of } k_{1 : m}} \det A(1 : m, \alpha_1 : m) \cdot b_{\alpha_11} b_{\alpha_2 1} \cdots b_{\alpha_m m}.
\]

where \((-1)^{(\alpha_{1 : m})}\) is the sign of the permutation \( \{\alpha_1, \ldots, \alpha_m\} \) of \( \{k_1, \ldots, k_m\} \). □

Corollary 4.2. If \( C = AB \) (\( A \) and \( B \) are \( n \)-by-\( n \)), then

\[
\det C(i_1 : p, j_1 : p) = \sum_{1 \leq i_1 < \cdots < i_p \leq n} \det A(i_1 : p, k_1 : m) \cdot \det B(k_1 : m, j_1 : p)
\]

Proof. \( C(i_1 : p, j_1 : p) = A(i_1 : p, 1 : n) B(1 : n, j_1 : p) \). □
5. Compound Matrices

Let $A$ be $n$-by-$n$. Consider all $p$th order minors of $A$:

$$A(i_1:p, k_1:p),$$

where $1 \leq i_1 < \cdots < i_p \leq n$ and $1 \leq k_1 < \cdots < k_p \leq n$. There are $N^2$ such minors, where $N \equiv \binom{n}{p}$. We arrange the minors (5) in a square array by enumerating all $N$ combinations of $p$ numbers in (say) lexicographic order.

If we assign the numbers $\alpha$ and $\beta$ to the combinations of indexes $i_1 < i_2 < \cdots < i_p$ and $k_1 < k_2 < \cdots < k_p$, then we denote

$$a_{\alpha \beta}^{(p)} = A(i_1:p, k_1:p).$$

Definition 5.1. The matrix

$$A^{(p)} = [a_{\alpha \beta}^{(p)}]_{\alpha, \beta = 1}^N$$

is called $p$th compound matrix of the matrix $A$.

Example 5.2. If $A = [a_{ij}]_{i,j=1}^4$, then we can order the combinations of 2 numbers from $\{1, 2, 3, 4\}$ as follows:

$$(12), (13), (14), (23), (24), (34).$$

Then

$$A^{(2)} = \begin{bmatrix}
A(1, 2, 1, 2) & A(1, 2, 1, 3) & A(1, 2, 1, 4) & A(1, 2, 2, 3) & A(1, 2, 2, 4) & A(1, 2, 3, 4) \\
A(1, 3, 1, 2) & A(1, 3, 1, 3) & A(1, 3, 1, 4) & A(1, 3, 2, 3) & A(1, 3, 2, 4) & A(1, 3, 3, 4) \\
A(1, 4, 1, 2) & A(1, 4, 1, 3) & A(1, 4, 1, 4) & A(1, 4, 2, 3) & A(1, 4, 2, 4) & A(1, 4, 3, 4) \\
A(2, 3, 1, 2) & A(2, 3, 1, 3) & A(2, 3, 1, 4) & A(2, 3, 2, 3) & A(2, 3, 2, 4) & A(2, 3, 3, 4) \\
A(2, 4, 1, 2) & A(2, 4, 1, 3) & A(2, 4, 1, 4) & A(2, 4, 2, 3) & A(2, 4, 2, 4) & A(2, 4, 3, 4) \\
A(3, 4, 1, 2) & A(3, 4, 1, 3) & A(3, 4, 1, 4) & A(3, 4, 2, 3) & A(3, 4, 2, 4) & A(3, 4, 3, 4)
\end{bmatrix}.$$}

The entries of the compound matrix are meant as determinants.

Theorem 5.3. The following properties of compound matrices hold.

1. If $T$ is upper triangular, then $T^{(p)}$ is also upper triangular;
2. If $C = AB$, then $C^{(p)} = A^{(p)}B^{(p)}$;
3. If $B = A^{-1}$, then $B^{(p)} = (A^{(p)})^{-1}$.

Proof. We use the definition of a compound matrix and the Theorem of Cauchy–Binet.

1. Assume the minors in the compound matrix are in lexicographic order. Let

$$i_{lm}^{(k)} = T(i_{1:k}, j_{1:k})$$

and $l > m$. Therefore $i_1 = j_1, \ldots, i_r = j_r$, and $i_{r+1} > j_{r+1}$. Then

$$T(i_{1:k}, j_{1:k}) = \det \begin{bmatrix}
\lambda_{i_1} & * & \cdots & * & \cdots & * \\
0 & \lambda_{i_2} & \cdots & * & \cdots & * \\
& & \ddots & & & \\
0 & 0 & \cdots & \lambda_{i_r} & \cdots & * \\
0 & 0 & \cdots & 0 & \cdots & * \\
& & & & \ddots & \\
0 & 0 & \cdots & 0 & \cdots & *
\end{bmatrix} = 0.$$
From the Cauchy–Binet formula we have
\[
\det C(i_1:p, k_1:p) = \sum_{1 \leq l_1 < \cdots < l_p \leq n} \det A(i_1:p, l_1:p) \det B(l_1:p, k_1:p),
\]
where \(1 \leq i_1 < \cdots < i_p \leq n\) and \(1 \leq k_1 < \cdots < l_p \leq n\). We can rewrite
the above as
\[
c_{p}^{(\alpha, \beta)} = \sum_{\lambda=1}^{N} a_{\alpha \lambda}^{(p)} b_{\lambda \beta}^{(p)},
\]
\(\alpha, \beta = 1, 2, \ldots, N\), where \(\alpha, \beta,\) and \(\lambda\) are the indexes of the combinations
\(i_1:p, k_1:p,\) and \(l_1:p,\) respectively. Hence
\[
C^{(p)} = A^{(p)} B^{(p)}.
\]
(3) From \(AB = I\) we have \(A^{(p)} B^{(p)} = T^{(p)}\).

**Theorem 5.4.** Let the eigenvalues of a matrix \(A\) be \(\lambda_1, \lambda_2, \ldots, \lambda_n\). Then the
eigenvalues of \(A^{(k)}\) are all possible products of \(\lambda_1, \lambda_2, \ldots, \lambda_n\) taken \(k\) at a time.

**Proof.** Consider the Schur normal form of \(A\):
\[
A = QTQ^{-1},
\]
where \(T\) is upper triangular. Then
\[
A^{(k)} = Q^{(k)} T^{(k)} (Q^{(k)})^{-1}.
\]
The matrix \(T^{(k)} = [t_{lm}^{(k)}]\) is upper triangular. The elements on the main diagonal
of \(T^{(k)}\) are
\[
t_{ii}^{(k)} = T(i_{1:k}, i_{1:k}) = \lambda_{i_1} \lambda_{i_2} \cdots \lambda_{i_k},
\]
all products of \(\lambda_1, \lambda_2, \ldots, \lambda_n\) taken \(k\) at a time. \(\square\)
6. Totally Nonnegative and Totally Positive Matrices

**Definition 6.1.** A matrix is called *totally nonnegative* [TN] (totally positive [TP]) if all its minors of any order are nonnegative (positive).

**Example 6.2.** The Cauchy matrix

\[
C = \left[ \frac{1}{x_i + y_j} \right]_{i,j=1}^n,
\]

\[0 < x_1 < x_2 < \cdots < x_n, \quad 0 < y_1 < y_2 < \cdots < y_n \] is totally positive.

Any submatrix of \(C\) is also Cauchy and satisfies the same condition on the nodes, therefore it suffices to establish that \(\det C > 0\). We have

\[
\det C = \prod_{i<j} (x_j - x_i)(y_j - y_i) \prod_{i,j} (x_i + y_j) > 0.
\]

**Example 6.3.** A tridiagonal matrix \(T\) with positive elements and positive principal minors is TN.

Indeed, let \(i_1, \ldots, i_p\) and \(k_1, \ldots, k_p\) be such that

\[1 \leq i_1 < i_2 < \cdots < i_p \leq n, \quad 1 \leq k_1 < k_2 < \cdots < k_p \leq n\]

and

\[i_{j+1} = k_1, i_{j+1} \neq k_{j+1}, \ldots, i_{j_2} \neq k_{j_2}, i_{j_2+1} = k_{j_2+1}, \ldots\]

Then

\[
\det T(i_1, k_1, \ldots, i_p, k_p) = \det T(i_1, k_1, i_{j_1+1}, k_{j_1+1}, \ldots, i_{j_2}, k_{j_2}, i_{j_2+1}, k_{j_2+1}, \ldots).
\]

Therefore \(T(i_1, k_1, \ldots, i_p, k_p) \geq 0\).

**Example 6.4.** A bidiagonal matrix with positive diagonal and only one nonzero offdiagonal entry \(x > 0\) is TN.

\[
A = \begin{bmatrix}
  a_1 & x & & \\
  & a_2 & & \\
  & & \ddots & \\
  & & & a_n
\end{bmatrix}.
\]

As we will see later any nonsingular TN matrix can be represented as a product of these bidiagonal matrices and their transposes. A matrix \(A\) of this type is the most elementary building block of a nonsingular TN matrix.

**Theorem 6.5.** If \(A\) and \(B\) are TN (TP), then \(AB\) is also TN (TP).

*Proof.* Cauchy–Binet. \(\square\)

**Theorem 6.6.** The eigenvalues of a TP matrix \(A\) are real, positive, and distinct.

*Proof.* We number the eigenvalues of \(A\) as \(|\lambda_1| \geq |\lambda_2| \geq \cdots\).

The \(k\)th compound matrix \(A^{(k)}\) is positive and its eigenvalues are

\[\lambda_1 \lambda_2 \cdots \lambda_k, \quad \lambda_1 \lambda_2 \cdots \lambda_{k-1} \lambda_{k+1}, \ldots\]

Since \(A^{(k)} > 0\) we have

\[\lambda_1 \lambda_2 \cdots \lambda_k > 0, \quad (k = 1, 2, \ldots, n)\]
and

\[ \lambda_1 \lambda_2 \cdots \lambda_k > \lambda_1 \lambda_2 \cdots \lambda_{k-1} \lambda_{k+1}, \quad (k = 1, 2, \ldots, n - 1). \]

The result follows. \qed

**Definition 6.7.** A matrix \( A \) is called oscillatory if \( A \) is TN and \( A^k \) is TP for some \( k \geq 1 \).

**Exercise 6.8.** Prove that the eigenvalues of an oscillatory matrix are real, positive, and distinct.
7. Neville Elimination

A nonsingular TN matrix $A$ can be reduced to upper triangular form using only adjacent rows or columns for elimination. This process is called Neville elimination. For example the $(i,j)$th entry $a_{ij} \neq 0$ is eliminated by subtracting a multiple $m_{ij} = a_{ij}/a_{i-1,j}$ of the $(i-1)$st row from the $i$th. In matrix form this results in the decomposition $A = E_i(m_{ij}) \cdot A'$, where $E_i$ differs from the identity only in its $(i,i-1)$ entry:

$$E_i(m_{ij}) \equiv \begin{bmatrix}
  1 & & & \\
  & 1 & & \\
  & m_{ij} & 1 & \\
  & & & \ddots \\
  & & & m_{ij} & 1 \\
  & & & & \ddots \\
  & & & & & \ddots \\
  & & & & & & 1
\end{bmatrix}.$$  \hspace{1cm} (6)

Once $A$ is reduced to upper triangular form $U$, the same elimination process is applied to $U$ using only adjacent columns. As a result $A$ is factored as a product of $(n^2-n)/2$ matrices of type (6), a diagonal matrix $D$, and $(n^2-n)/2$ transposes of matrices of type (6).

One may have noticed that we bravely divided by $a_{i-1,j}$ in forming $m_{ij}$. This will never lead to division by zero for nonsingular TN matrices. Indeed, if $a_{i-1,k}$ is any nonzero (thus positive) entry in the $(i-1)$st row of $A$, then the $2 \times 2$ minor of rows $i-1$ and $i$ and columns $j$ and $k$ ($j < k$) must be nonnegative. Therefore $a_{ij} > 0$ implies $a_{i-1,j} > 0$ and in turn $a_{i-1,j} = 0$ implies $a_{ij} = 0$. Equivalently, $m_{ij} = 0$ implies $m_{i+1,j} = \ldots = m_{nj} = 0$ and thus

$$m_{ij} = 0 \implies \begin{cases}
  m_{kj} = 0 \text{ for all } k > i, \text{ if } i > j; \\
  m_{ik} = 0 \text{ for all } k > j, \text{ if } i < j.
\end{cases} \hspace{1cm} (7)$$

For example:

$$\begin{bmatrix}
  1 & 2 & 6 \\
  4 & 13 & 69 \\
  28 & 131 & 852
\end{bmatrix}
= \begin{bmatrix}
  1 & & \\
  & 1 & \\
  & 4 & 1
\end{bmatrix}
\begin{bmatrix}
  1 & & \\
  & 5 & \\
  & 8 & 1
\end{bmatrix}
\begin{bmatrix}
  1 & 2 & \\
  & 1 & 6 \\
  & 9 & 1
\end{bmatrix}
\begin{bmatrix}
  1 & \\
  & 1 \\
  & 3
\end{bmatrix}.$$  

The matrices $E_i$ have the following important properties

$$E_i^{-1}(x) = E_i(-x) \hspace{1cm} (8)$$

$$E_i(x)E_j(y) = E_j(y)E_i(x), \text{ unless } |i-j| = 1 \text{ and } xy \neq 0$$

$$E_i(x)E_i(y) = E_i(x + y)$$

If we apply Neville elimination to $A$, eliminating one subdiagonal at a time, starting with the $(n,1)$ entry we obtain

$$A = (E_n(m_{n1})) \cdot (E_{n-1}(m_{n-1,1})E_n(m_{n2})) \cdot \ldots \cdot (E_2(m_{21})E_3(m_{32}) \ldots E_n(m_{n,n-1})) \cdot D \cdot (E_n^T(m_{n-1,n}) \ldots E_3^T(m_{23})E_2^T(m_{12}) \ldots \cdot (E_n^T(m_{2n})E_{n-1}^T(m_{1,n-1})) \cdot (E_n^T(m_{1n})). \hspace{1cm} (9)$$
Written another way (9) becomes:

\[ A = \prod_{k=1}^{n-1} \prod_{j=n-k+1}^{n} E_j(m_{j,k+j-n}) \cdot D \cdot \prod_{j=1}^{k=n-1} \prod_{i=n-k+1}^{n} E_j(m_{k+j-n,j}) , \]

We “assemble” the \( E_i \)’s for each \( k \) (inside the parentheses of (9)) into bidiagonal matrices \( L^{(k)} \):

\[ L^{(k)} = \begin{cases} 1 & \cdots & a_{1k} \cdots a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{k-1,k-1} & \cdots & a_{k-1,k-n} & a_k(k-2) & \cdots & a_k(k-1) \\ 0 & \cdots & 0 & a_{k+1,k-1} & \cdots & a_{k+1,k-n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{nk} & \cdots & a_{nk} \end{cases} \]

where

\[ (L^{(k)})_{i-1,1} \equiv (L^{(k)})_{i,i-1} = m_{i,k+i-n} \]

We analogously form unit upper bidiagonal matrices \( U^{(k)} \). Finally,

\[ A = L^{(1)} \cdots L^{(n-1)} \cdot D \cdot U^{(n-1)} \cdots U^{(1)}, \]

We can obtain simple determinantal formulas for the entries of \( L^{(i)} , D \), and \( U^{(i)} \).

Since \( L = L^{(1)} \cdots L^{(n-1)} \) is lower triangular and \( U = U^{(n-1)} \cdots U^{(1)} \) is upper triangular, \( A = LDU \) is the usual LDU decomposition of \( A \) and

\[ D_{ii} = \frac{\det A(1:i,1:i)}{A(1:i-1,1:i-1)} \]

Consider the matrix \( A \) after \( k - 1 \) steps of Neville elimination:

The minors of \( A \) do not change in this process, therefore

\[ \det A^{(k-1)}(1:k,1:k) = a_{11}a_{22}^{(1)} \cdots a_{kk-1,k-1}^{(k-2)}a_{kk}^{(k-1)} \]

and

\[ a_{kk}^{(k-1)} = \frac{\det A(1:k,1:k)}{\det A(1:k-1,1:k-1)} \]

We will now obtain similar formulas for \( a_{ij}^{(k-1)} \). The crucial observation is that we will still get the entry \( a_{ij}^{(k-1)} \) whether we run Neville elimination on rows \( 1:n, \)
2 : n, . . . , or i − k + 1 : n, since \( a_{ij}^{(k-1)} \) will be obtained as a result of the very same operations.

Consider applying \( k - 1 \) steps of Neville elimination to rows \( i - k + 1 \) through \( n \) of \( A \) to obtain:

\[
\begin{bmatrix}
    a_{i+k-1,1} & a_{i+k-1,2} & \cdots & a_{i+k-1,k-1} & a_{i+k-1,k} & \cdots & a_{i+k-1,n} \\
    0 & a_{i+k,2}^{(1)} & \cdots & a_{i+k,k-1}^{(1)} & a_{i+k,k}^{(1)} & \cdots & a_{i+k,n}^{(1)} \\
    \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & a_{i-1,k-1}^{(k-2)} & a_{i-1,k}^{(k-2)} & \cdots & a_{i-1,n}^{(k-2)} \\
    0 & 0 & \cdots & 0 & a_{i,k}^{(k-1)} & \cdots & a_{i,n}^{(k-1)} \\
    \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & 0 & a_{nk}^{(k-1)} & \cdots & a_{nn}^{(k-1)} \\
\end{bmatrix}
\]

Analogously to (13) we get:

\[
a_{ik}^{(k-1)} = \frac{\det A(i - k + 1 : i, 1 : k)}{\det A(i - k + 1 : i - 1, 1 : k - 1)}.
\]

Therefore the multiplier used to set an entry \((i, j)\) to zero in the process of Neville elimination equals:

\[
m_{ik} = \frac{a_{ik}}{a_{i,k-1}} = \frac{\det A(i - k + 1 : i, 1 : k)}{\det A(i - k + 1 : i - 1, 1 : k - 1)} \cdot \frac{\det A(i - k : i - 2, 1 : k - 1)}{\det A(i - k : i - 1, 1 : k)}.
\]

The connection between \( l_i^{(r)} \) and \( m_{ik} \) is given by (11).

**Theorem 7.1.** The following expressions are valid for the entries of the bidiagonal decomposition (12) of \( A \):

\[
(14) \quad l_1^{(k)} = \frac{\det A(q + 1 : i + 1, 1 : q + 1)}{\det A(q + 1 : i, 1 : q)} \cdot \frac{\det A(q : i - 1, 1 : i - q)}{\det A(q : i - 1, 1 : i)},
\]

\[
(15) \quad u_i^{(k)} = \frac{\det A(1 : i - q + 1, q + 1 : i + 1)}{\det A(1 : i - q + 1, q + 1 : i)} \cdot \frac{\det A(1 : i - q, q : i - 1)}{\det A(1 : i - q, q : i)},
\]

\[
(16) \quad d_i = \frac{\det A(1 : i, 1 : i)}{\det A(1 : i - 1, 1 : i - 1)},
\]

where \( q = n - k; i \geq q = n - k \) in (14) and (15); and \( l_1^{(k)} = u_1^{(k)} = 0 \) for \( i < n - k \).

Therefore, if \( A \) is TN, then \( l_i^{(k)} \geq 0, d_i \geq 0, \) and \( u_i^{(k)} \geq 0 \). Conversely, if \( l_i^{(k)} \geq 0, d_i \geq 0, \) and \( u_i^{(k)} \geq 0 \), then \( A \) is TN as a product of TN matrices, see (10).

**Theorem 7.2** (Gasca, Peña). A nonsingular matrix \( A \) is TN if and only if it can be uniquely factored as

\[
A = L^{(1)} \cdots L^{(n-1)} \cdot D \cdot U^{(n-1)} \cdots U^{(1)}
\]

where \( D = \text{diag}(d_1, d_2, \ldots, d_n) \), and \( L^{(k)} \) and \( U^{(k)} \) are lower and upper unit bidiagonal matrices, respectively, such that

1. \( d_i > 0 \) for all \( i \);
2. \( l_1^{(k)} = u_1^{(k)} = 0 \) for \( i < n - k \);
3. \( l_i^{(k)} \geq 0, u_i^{(k)} \geq 0 \) for \( i \geq n - k \), \( l_i^{(k)} > 0, u_i^{(k)} > 0 \) for \( i \geq n - k \) if \( A \) is TP.
(4) \( i^{(k)} = 0 \) implies \( i^{(k-s)} = 0 \) for \( s = 1, \ldots, k - 1 \); and \( u_i^{(k)} = 0 \) implies 
\( u_{i+s}^{(k-s)} = 0 \) for \( s = 1, \ldots, k - 1 \). This is equivalent to (7) and is automatically 
satisfied if \( A \) is TP.

**Definition 7.3.** A minor of a matrix \( A \) is called initial if it is contiguous and 
includes the first row or the first column. Namely, it looks like \( A(i : j, k : l) \), 
\( j - i = k - l \), where \( i = 1 \) or \( k = 1 \).

Clearly the expressions (14), (15), and (16) include only (and all) initial minors 
of \( A \).

Therefore in order to test if a matrix is TP, it suffices to verify that its initial 
minors are positive. In practice, one computes the Neville elimination of a matrix. 
If it completes successfully with all pivots and multipliers being positive, then \( A \) is 
TP.

8. **Properties of Nonsingular Totally Nonnegative Matrices**

The total nonnegativity is preserved under a number of matrix operations: multiplication, 
Schur complementation, taking a converse, etc. It is easiest to see these properties 
starting from the bidiagonal decomposition.

There are four elementary transformations, which we call **Elementary Elimination 
Transformations** (EET), each of which preserve the total nonnegativity. Most 
matrix operations that preserve the total nonnegativity can typically be represented 
as a sequence of EET. The EET are:

- **EET1:** Subtracting a multiple of a row from the next in order to create a zero in 
a process of reducing a matrix to upper triangular form;
- **EET2:** Adding a multiple of a row (column) to the previous one;
- **EET3:** Adding a multiple of a row (column) to the next one;
- **EET4:** Scaling by a positive diagonal matrix.

**Theorem 8.1.** Each EET preserves the total nonnegativity of a nonsingular TN 
matrix.

**Proof.** EET 2-4 clearly preserve the TN—they a multiplication of a TN matrix by 
another TN matrix—\( E_i(x), E^T_i(x) \), or \( D > 0 \).

For EET1 we will only prove that the very first step in an elimination process 
preserves the TN, the rest is analogous.

Subtracting a multiple of the \((n - 1)\)st row from the \(n\)th in order to create a zero in 
position \((n, 1)\) in \( A \) is equivalent to forming

\[
E_n \left( -\frac{a_{n1}}{a_{n-1,1}} \right) A \\
= E_n(-m_{n1})A \\
= \left( \prod_{k=1}^{n-1} \prod_{j=n-k+1}^{n} E_j(n_{j,k+j-n}) \right) \left( \prod_{k=1}^{n-1} \prod_{j=n-k+1}^{n} E^T_j(m_{k+j-n,j}) \right)
\]

which is a product of TN matrices, and is thus TN. 

\( \square \)
A Givens rotation \( G \) applied the process of reducing a matrix to upper triangular form preserves the TN.

The trick is to represent \( G \) as a sequence of two EETs. Indeed, applying a Givens rotation to create a zero in position (say) \((n,1)\) is equivalent to (1) subtracting a (positive) multiple of the \((n-1)\)st row from the \(n\)th in order to create a zero in position \((n,1)\), followed by (2) adding a positive multiple of the \(n\)th row to the \((n-1)\)st and finally (3) scaling the last two rows.

Let
\[
G = \begin{bmatrix} c & s \\ -s & c \end{bmatrix},
\]
where \( s^2 + c^2 = 1 \), be a Givens rotation used to set the \((n,1)\) entry of \( A \) to zero. Then
\[
G \cdot \begin{bmatrix} a_{n-1,1} & a_{n-1,2} \\ a_{n1} & a_{n2} \end{bmatrix} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \cdot \begin{bmatrix} a_{n-1,1} & a_{n-1,2} \\ a_{n1} & a_{n2} \end{bmatrix} = \begin{bmatrix} a_{(1)n-1,1}^{(1)} & a_{(1)n-1,2}^{(1)} \\ 0 & a_{n2}^{(1)} \end{bmatrix}.
\]

A simple calculation shows that \( c = 1/\sqrt{1 + x^2} \) and \( s = x/\sqrt{1 + x^2} \), where
\[
x = a_{n1}/a_{n-1,1} = m_{n1} = l_{n-1}^{(1)}
\]
is the only nonzero offdiagonal entry in \( L^{(1)} \). Now write \( G \) as
\[
G = \begin{bmatrix} 1/c & cx \\ 0 & c \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ -x & 1 \end{bmatrix}
\]
\[
G = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1/c & 0 \\ -x & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ -x & 1 \end{bmatrix} = E_n(m_{n1}) \cdot \begin{bmatrix} 1/c & 0 \\ -x & 1 \end{bmatrix} \cdot E_n(-m_{n1}).
\]

This is a product of three EET, each of which preserve the total nonnegativity.

**Theorem 8.2.** Let \( A \) and \( B \) be nonsingular, square TN matrices. The following matrices are also nonsingular and TN.

1. \( A \cdot B \)
2. The Schur complement of the \((1,1)\) entry in \( A \).
3. The J-inverse of \( A: A^* = JAJ \), where \( J = \text{diag}((-1)^i)_{i=1}^n \)
4. The converse of \( A: A^# = [a_{n+1-j,n+1-i}]_{i,j=1}^n = YAY \), where \( Y \) is the reverse identity, \( Y_{n+1-i,j} = 1 \).
5. The matrix \( R \), where \( A = QR \) is the QR decomposition of \( A \)

**Proof.** We use the bidiagonal decomposition of \( A \) (10) to prove that any of the matrices in question are products of positive diagonal matrices, \( E_t(x) \) or \( E_t^T(x) \), \( x > 0 \).

1. Cauchy–Binet
(2) Let $A'$ be obtained from $A$ after one step of Gaussian elimination. We have $A' = KA$, where

\[
K = \begin{bmatrix}
\frac{a_{21}}{a_{11}} & 1 & 1 \\
\frac{a_{31}}{a_{11}} & 1 & 1 \\
\vdots & \ddots & \ddots \\
\frac{a_{n1}}{a_{11}} & 1 & \frac{a_{n-11}}{a_{n-1-1}} \\
1 & \frac{a_{n1}}{a_{21}} & 1 \\
\end{bmatrix}
\]

\[
= \begin{bmatrix}
1 & 1 & \frac{a_{11}}{a_{11}} \\
1 & \frac{a_{11}}{a_{11}} & 1 \\
\vdots & \ddots & \ddots \\
1 & \frac{a_{n1}}{a_{n1}} & 1 \\
\end{bmatrix}
\]

\[
\times \begin{bmatrix}
1 & 1 & \frac{a_{11}}{a_{11}} \\
\frac{a_{21}}{a_{21}} & 1 & 1 \\
\vdots & \ddots & \ddots \\
\frac{a_{n1}}{a_{n1}} & 1 & \frac{a_{n-11}}{a_{n-1-1}} \\
1 & \frac{a_{n1}}{a_{n1}} & 1 \\
\end{bmatrix}
\]

\[
= \prod_{i=3}^{n} E_i(m_{i1}) \times \prod_{i=2}^{n} E_i(-m_{i1}).
\]

Using (10) we get

\[
KA = \prod_{i=3}^{n} E_i(m_{i1}) \times \prod_{i=2}^{n} E_i(-m_{i1}) \times A
\]

\[
= \prod_{i=3}^{n} E_i(m_{i1}) \times \prod_{i=2}^{n} E_i(-m_{i1})
\]

\[
\times \left( \prod_{k=1}^{n-1} \prod_{j=n-k+1}^{n} E_j(n_{j,k+j-n}) \right) D \left( \prod_{k=1}^{n-1} \prod_{j=n-k+1}^{n} E_j^T(m_{k+j-n,j}) \right)
\]

\[
= \prod_{i=3}^{n} E_i(m_{i1})
\]

\[
\times \left( \prod_{k=1}^{n-1} \prod_{j=n-k+2}^{n} E_j(n_{j,k+j-n}) \right) D \left( \prod_{k=1}^{n-1} \prod_{j=n-k+1}^{n} E_j^T(m_{k+j-n,j}) \right).
\]
(3) $A^* \equiv ((-1)^{i+j}a_{ij})^{-1} = (JAJ)^{-1} = JA^{-1}J$.

We have $J^2 = I$, $(E_i(x))^{-1} = E_i(-x)$, and $JE_i(-x)J = E_i(x)$. We use (10) to write

$$A^* = J \left( \prod_{i=1}^{n-1} \prod_{j=n-i+1}^n E_j^T (-b_{i+j-n,j}) \right) D^{-1} \left( \prod_{i=1}^{n-1} \prod_{j=n-i+1}^n E_j (-b_{i+j-n,j}) \right) J$$

$$= \left( \prod_{i=1}^{n-1} \prod_{j=n-i+1}^n (JE_j^T (-b_{i+j-n,j})J) \right) J D^{-1} \left( \prod_{i=1}^{n-1} \prod_{j=n-i+1}^n (JE_j (-b_{i+j-n,j})J) \right)$$

$$= \left( \prod_{i=1}^{n-1} \prod_{j=n-i+1}^n E_j^T (b_{i+j-n,j}) \right) D^{-1} \left( \prod_{i=1}^{n-1} \prod_{j=n-i+1}^n E_j (b_{i+j-n,j}) \right).$$

(4) $A^\# \equiv (a_{n+1-i,n+1-j})_{i,j=1}^n$.

Let $Y_k \equiv (\delta_{n+1-i,j})_{i,j=1}^n$ be the reverse identity. Using $Y^2 = I$ and $E_i^\#(x) = YE_i(x)Y = E_i^{n+2-i}(x)$ we obtain

$$A^\# = YAY$$

$$= Y \left( \prod_{i=1}^{n-1} \prod_{j=n-i+1}^n E_j (m_{j,i+j-n}) \right) D \left( \prod_{i=1}^{n-1} \prod_{j=n-i+1}^n E_j^T (m_{j,i+j-n}) \right) Y$$

$$= \left( \prod_{i=1}^{n-1} \prod_{j=n-i+1}^n YE_j (m_{j,i+j-n})Y \right) YDY \left( \prod_{i=1}^{n-1} \prod_{j=n-i+1}^n YE_j^T (m_{j,i+j-n})Y \right)$$

$$= \left( \prod_{i=1}^{n-1} \prod_{j=n-i+1}^n E_j^{n+2-j} (m_{j,i+j-n}) \right) D^\# \left( \prod_{i=1}^{n-1} \prod_{j=n-i+1}^n E_j^{n+2-j} (m_{j,i+j-n}) \right).$$

where $D^\# = YDY$ is diagonal, $D_{ii}^\# = D_{n+1-i,n+1-i}$, $i = 1, 2, \ldots, n$.

(5) A single Givens rotation, designed to reduce $A$ to upper triangular form preserves the total nonnegativity. Therefore the resulting $R$ matrix will still be TN.