Upper Bounds for the Determinant of a Row Stochastic Matrix

Karl Goldberg

Institute for Basic Standards, National Bureau of Standards, Washington, D.C.

(February 16, 1966)

If \( A = (a_{ij}) \) is a row stochastic matrix of order \( n \) then \( |\det A| \leq (1 - \Sigma_{j=1}^{n} \min_{i} a_{ij})^{n-1} \). If \( \alpha \) is a 1 by \( n \) real vector such that every element of \( \alpha A \) is nonnegative then \( |\det A| \leq s(\alpha)/|g(\alpha)| \), where \( s(\alpha) \) is the sum of the elements of \( \alpha \) and \( g(\alpha) \) is the element of \( \alpha \) of greatest absolute value. The conditions for equality are determined in both cases.

Key Words: Determinants, matrices (row stochastic), upper bounds.

It is known that if \( A \) is a row stochastic matrix then \( |\det A| \leq 1 \) with equality if and only if \( A \) is a permutation matrix. We will use this result to extend itself.

The result is equivalent to the following statement about nonnegative matrices.\(^1\) Let \( A \) be a nonnegative matrix of order \( n \), let \( r_{A}^{(i)} \) denote the sum of the elements in the \( i \)th row, let

\[
r_{A} = r_{A}^{(1)} r_{A}^{(2)} \ldots r_{A}^{(n)}
\]

and

\[
R_{A} = \text{diag} (r_{A}^{(1)}, r_{A}^{(2)}, \ldots, r_{A}^{(n)}).
\]

Note that \( r_{1} = 0 \) if and only if some row of \( A \) is all 0's. Otherwise \( R_{A}^{0} A \) is row stochastic. Thus \( |\det A| \leq r_{A}^{0} \) with equality if and only if either \( r_{A} = 0 \) or \( A = R_{A} P \) for some permutation matrix \( P \).

We use the following procedure. Given a row stochastic matrix \( A \), let \( B \) be a nonsingular matrix such that \( AB \geq 0 \) (i.e., \( AB \) is nonnegative). Then the above result implies that

\[
|\det A| \leq \frac{r_{AB}}{|\det B|}
\]

with equality if and only if either \( r_{AB} = 0 \) or \( A = R_{AB} PB^{-1} \). The same procedure with \( CA \geq 0 \) yields a similar result. Since \( B \) (or \( C \)) need not be nonnegative we have room for improvement.

From this point on \( A \) will denote a row stochastic matrix of order \( n \).

Lower case Greek letters will denote 1 by \( n \) real vectors including

\( \epsilon = (1 \ 1 \ \ldots \ 1) \) and \( \delta_{i} = (0 \ \ldots \ 1 \ 0 \ \ldots \ 0) = (\delta_{ij}) \).

For a vector \( \alpha \), \( s(\alpha) \) will denote the sum of its elements, \( g(\alpha) \) will denote its element of greatest absolute value, and \( M_{\alpha}(\alpha) \) will denote the matrix

\[
M_{\alpha}(\alpha) = \begin{bmatrix}
\delta_{1} \\
\vdots \\
\delta_{i-1} \\
\alpha \\
\delta_{i+1} \\
\vdots \\
\delta_{n}
\end{bmatrix}
\]

i.e., the identity matrix with the \( i \)th row replaced by \( \alpha \).

For future clarity we note now that

\[
s(\alpha) = \alpha \epsilon^{T},
\]

while \( \epsilon^{T} \alpha \) is the matrix

\[
\epsilon^{T} \alpha = \begin{bmatrix}
\alpha \\
\alpha \\
\vdots \\
\alpha
\end{bmatrix}
\]

Finally \( A \epsilon^{T} = \epsilon^{T} \). We shall prove:

**Theorem 1.** Let \( A \) be a row stochastic matrix of order \( n \), and let

\[
c = \sum_{j=1}^{n} \min_{i} a_{ij}.
\]

Then

\[
|\det A| \leq (1 - c)^{n-1}
\]

\(^{1}\) A matrix is nonnegative if each element is real and nonnegative. It is row stochastic if, in addition, each row sum equals 1.
with equality if and only if
\[ A = (1 - s(\beta))P + \varepsilon^T\beta \]
for some vector \( \beta \geq 0 \) with \( 1 \geq s(\beta) \), and some permutation matrix \( P \).

**Theorem 2.** Let \( A \) be a row stochastic matrix and \( \alpha \) a real, nonzero vector such that \( \alpha A \geq 0 \). Then
\[ |\det A| \leq \frac{s(\alpha)}{g(\alpha)} \]
with equality if and only if either \( s(\alpha) = 0 \), or \( \alpha \) has exactly one positive element (in the \( i \)-th position) and
\[ A = M_i(\gamma)P \]
for some permutation matrix \( P \), with
\[ \gamma = \frac{1}{g(\alpha)} \{ (g(\alpha) + s(\alpha))\delta_i - \alpha \}. \]
To prove Theorem 1 let \( c_j = \min_i a_{ij} \) and let
\[ B = I - \sum_{j=1}^n c_j \varepsilon^T \delta_j. \]
The matrix \( \varepsilon^T \delta_j \) has 1's in the \( j \)-th column and 0's elsewhere, so that \( AB = (a_{ij} - c_j) \). By the definition of the \( c_j \), \( AB \) is nonnegative.
The row sums of \( AB \) are the elements of the vector
\[ ABe^T = Ae^T - \sum_{j=1}^n c_j e^T \delta_j = (1-c)e^T. \]
Thus \( r_{AB} = 1-c \) for all \( i \), so that
\[ r_{AB} = (1-c)^n. \]
It is a simple exercise (left to the reader) to prove that
\[ \det B = 1-c \]
and
\[ B^{-1} = I + (1-c)^{-1} \sum_{j=1}^n c_j e^T \delta_j \quad c \neq 1. \]
If \( c \neq 1 \) then, from our previous discussion, the necessity case of Theorem 1 follows with
\[ \beta = (c_1, c_2 \ldots c_n). \]
In the case of sufficiency, if \( A = (1 - s(\beta))P + \varepsilon^T\beta \) then all the row sums of \( A \) equal \( 1(Ae^T = e^T) \), \( A \) is nonnegative if \( \beta \geq 0 \), and \( 1 \geq s(\beta), \beta = (c_1, c_2 \ldots c_n) \) with the \( c_j \) defined as above, and \( |\det A| = (1-s(\beta))^n-1 \).
A short discussion of the value of \( c \) will complete our proof. Since \( A \) is row stochastic the sum of all of its elements is \( n \). Thus
\[ n = \sum_{j=1}^n \sum_{i=1}^n a_{ij} \geq \sum_{j=1}^n \sum_{i=1}^n c_j = nc \]
with equality if and only if only if \( a_{ij} = c_j \) for all \( i, j \), in which case \( A \) is of rank 1. Then \( \det A = 1 \) for \( n = 1 \), and \( \det A = 0 \) with \( n > 1 \), and \( A = \varepsilon^T\beta \). This checks with the statement of Theorem 1 and completes our proof.
To prove Theorem 2 let \( C = M_i(\alpha) \). The rows of \( CA \) are the rows of \( A \) except for the \( i \)-th row which is \( \alpha A \). By hypothesis \( \alpha A \geq 0 \) so that \( CA \) is nonnegative.
The row sums of \( CA \) are 1 except for the \( i \)-th row whose sum is
\[ s(\alpha A) = (\alpha A)e^T = \alpha(Ae^T) = \alpha e^T = s(\alpha). \]
Thus \( r_{CA} = s(\alpha) \) and \( R_{CA} \) is the identity matrix transformed by replacing the 1 in the \( i \)-th position on the diagonal by \( s(\alpha) \). Note that, if \( s(\alpha) \neq 0 \), we have
\[ R_{CA}M_i(\alpha) = M_i(\alpha/s(\alpha)). \]
The determinant of \( C \) is the product of its diagonal elements. Let
\[ \alpha = (a_1 a_2 \ldots a_n). \]
Then \( \det C = a_i \). By our previous discussion \( |\det A| \leq \frac{s(\alpha)}{g(\alpha)} \).
Thus
\[ |\det A| \leq \frac{s(\alpha)}{|a_i|}. \]
This inequality holds for all \( i \); the best of the inequalities is
\[ |\det A| \leq \frac{s(\alpha)}{g(\alpha)}. \]
Suppose \( g(\alpha) = a_i \). We have equality if and only if either \( s(\alpha) = 0 \), (whence \( \alpha A = 0 \), \( \alpha \neq 0 \), so that \( \det A = 0 \)) or, if \( s(\alpha) = 0 \);
\[ A = C^{-1}R_{CA}P = (R_{CA}M_i(\alpha))^{-1}P = (M_i(\alpha/s(\alpha)))^{-1}P \]
for some permutation matrix \( P \).
But \( (M_i(\alpha/s(\alpha)))^{-1} = M_i(\gamma) \) with
\[ \gamma = a_i^{-1}(-a_i \ldots -a_{i-1} s(\alpha) - a_{i+1} \ldots - a_n). \]
Since \( A = M_i(\gamma)P \) is nonnegative we must have \( a_j \leq 0 \) for \( j \neq i \). This completes the proof of Theorem 2.