DIAGONAL EQUIVALENCE TO MATRICES WITH PRESCRIBED ROW AND COLUMN Sums. II

RICHARD SINKHORN

ABSTRACT. Let $A$ be a nonnegative $m \times n$ matrix and let $r = (r_1, \ldots, r_m)$ and $c = (c_1, \ldots, c_n)$ be positive vectors such that
$$\sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j.$$ It is well known that if there exists a nonnegative $m \times n$ matrix $B$ with the same zero pattern as $A$ having the $i$th row sum $r_i$ and $j$th column sum $c_j$, there exist diagonal matrices $D_1$ and $D_2$ with positive main diagonals such that $D_1 A D_2$ has $i$th row sum $r_i$ and $j$th column sum $c_j$. However the known proofs are at best cumbersome.

It is shown here that this result can be obtained by considering the minimum of a certain real-valued function of $n$ positive variables.

It has been shown originally by Sinkhorn and Knopp [8] and Brualdi, Parter, and Schneider [3] that if $A$ is a nonnegative fully indecomposable matrix, i.e. $A$ contains no $s \times (n-s)$ zero submatrix, then there exists a doubly stochastic matrix of the form $D_1 A D_2$ where $D_1$ and $D_2$ are diagonal matrices with positive main diagonals. Later Djoković [4], and independently, London [5], proved the same theorem by considering the minimum of

$$f(x) = \prod_{i=1}^{n} \left( \sum_{j=1}^{n} a_{ij} x_j \right) / \prod_{j=1}^{n} x_j$$

for vectors $x = (x_1, \ldots, x_n)$ with positive coordinates.

In the meantime Menon [6] had obtained the following modification of this result.

Theorem 1. Let $A$ be a nonnegative $m \times n$ matrix and let $r = (r_1, \ldots, r_m)$ and $c = (c_1, \ldots, c_n)$ be positive vectors such that $\sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j$. If there exists a nonnegative $m \times n$ matrix $B$ with the same zero pattern as $A$, i.e. $b_{ij} = 0 \iff a_{ij} = 0$, having $i$th row sum $r_i$ and $j$th column sum $c_j$, then

Received by the editors September 17, 1973 and, in revised form, October 31, 1973.


Key words and phrases. Nonnegative matrix, diagonal equivalence, fully indecomposable matrix, zero pattern.

Copyright © 1974, American Mathematical Society
there exist diagonal matrices $D_1$ and $D_2$ with positive main diagonals such that $D_1AD_2$ has $i$th row sum $r_i$ and $j$th column sum $c_j$.

Brualdi [2] showed that the existence of $B$ in Theorem 1 is equivalent to the conditions that

1. $A[E|F] = 0$, $A(E|F) \neq 0 \implies \Sigma_{i \in E} r_i < \Sigma_{j \in F} c_j$, and
2. $A[E|F] = 0$, $A(E|F) = 0 \implies \Sigma_{i \in E} r_i = \Sigma_{j \in F} c_j$,

if $\Sigma_{i=1}^m r_i = \Sigma_{j=1}^n c_j$ holds.

The notation used has the following meaning. If $E$ is a proper nonvoid subset of $M = \{1, \ldots, m\}$ and $F$ is a proper nonvoid subset of $N = \{1, \ldots, n\}$, then $A(E|F)$ is that submatrix of $A$ obtained by deleting from $A$ those rows whose indices do not belong to $E$ and those columns whose indices belong to $F$. The rows and columns of this submatrix appear in the same order as in $A$: rows are counted from top to bottom; columns are counted from left to right. $A(E|F)$ is that submatrix of $A$ obtained by deleting from $A$ those rows whose indices belong to $E$ and those columns whose indices do not belong to $F$, where, as before, the rows and columns of this submatrix appear in the same order as in $A$. Observe that the submatrix $A(E|F)$ is the same as the submatrix $A[M - E|N - F]$. In the course of the paper two other submatrix notations are used. $A[E|F]$ is used to denote the submatrix $A[E|N - F] = A(M - E|F)$ in $A$; $A(E|F)$ is used to denote the submatrix $A[M - E|F] = A(E|N - F)$ in $A$.

Menon and Schneider [7] have given another proof of the Menon-Brualdi results.

It is the intent of this paper to show how the Djoković-London formula can be modified to yield the Menon-Brualdi-Schneider results.

We shall require the following lemma which follows at once from the concavity of the logarithm function. See [1, p. 7].

**Lemma.** Let $x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_n$ be nonnegative real numbers and put $\lambda_1 + \cdots + \lambda_n = \lambda$. Then if $0 \lambda_0$ is taken to be 1,

$$
\left( \sum_{k=1}^n \lambda_k x_k \right) \geq \lambda \left( \prod_{k=1}^n x_k^{\lambda_k} \right).
$$

We now prove the intended result. We shall assume that whenever there is a submatrix $A[E|F] = 0$ in $A$, $A(E|F) \neq 0$, for otherwise we could establish the result for the submatrices $A[E|F]$ and $A(E|F)$. We assume that Brualdi’s condition (1) holds and that $\Sigma_{i=1}^m r_i = \Sigma_{j=1}^n c_j$.

Put
\[
\phi(x) = \prod_{i=1}^{m} \left( \sum_{j=1}^{n} a_{ij} x_j \right)^{r_i} / \prod_{j=1}^{n} x_j^{c_j},
\]
where \( x = (x_1, \cdots, x_n) \) is positive, i.e. \( x \in (\mathbb{R}^n)^+ \). We shall consider the problem of determining the minimum of \( \phi \) on \( (\mathbb{R}^n)^+ \). Since \( \sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j \), \( \phi(\lambda x) = \phi(x) \) for all \( \lambda > 0 \) and thus we can restrict our attention to the set \( K \) of \( x \in (\mathbb{R}^n)^+ \) for which \( \|x\| = (x_1^2 + \cdots + x_n^2)^{1/2} = 1 \).

Suppose on \( K, x \to \Delta \), the boundary of \( (\mathbb{R}^n)^+ \). Let \( E = \{j | x_j \to 0\} \) and then set \( E = \{i | a_{ij} = 0 \text{ for all } j \notin F\} \). Since \( x \to \Delta \) on \( K, F \) is a nonvoid proper subset of \( \{1, \cdots, n\} \). Since every \( c_j > 0 \), \( E \) is a proper subset of \( \{1, \cdots, m\} \). If \( E = \emptyset \), \( \phi(x) \to \infty \) as \( x \to \Delta \). If \( E \neq \emptyset \), we write \( \phi(x) = \phi_1(x) \phi_2(x) \) where

\[
(3) \quad \phi_1(x) = \prod_{i \in E} \left( \sum_{j=1}^{n} a_{ij} x_j \right)^{r_i} / \prod_{j \in F} x_j^{c_j} = \prod_{i \in E} \left( \sum_{j \in F} a_{ij} x_j \right)^{r_i} / \prod_{j \in F} x_j^{c_j}
\]

and

\[
(4) \quad \phi_2(x) = \prod_{i \notin F} \left( \sum_{j=1}^{n} a_{ij} x_j \right)^{r_i} / \prod_{j \notin F} x_j^{c_j}.
\]

Since \( \phi_2 \) has a positive limit as \( x \to \Delta \), we concentrate on \( \phi_1 \).

Let \( B \) be as in Theorem 1. Then \( B[E|F] = 0 \) and therefore \( \sum_{j \in F} b_{ij} = r_i \) for each \( i \in E \) and hence from the Lemma

\[
(5) \quad \left( \sum_{j \in F} a_{ij} x_j \right)^{r_i} \geq r_i \prod_{j \in F} a_{ij}^{b_{ij}} / \prod_{j \in F} b_{ij}^{b_{ij}} \prod_{j \in F} x_j^{b_{ij}} = \theta_i \prod_{j \in F} x_j^{b_{ij}}
\]

for all \( i \in E \), where \( 0^0 \) is taken to be 1. Whence

\[
(6) \quad \phi_1(x) \geq \prod_{i \in E} \theta_i / \prod_{j \in F} x_j^{c_j} - \sum_{i \in E} b_{ij}.
\]

Since \( \sum_{i=1}^{m} b_{ij} = c_j, j = 1, \cdots, n \), certainly \( \sum_{i \in E} b_{ij} \leq c_j \) for every \( j \in F \).

However since \( \sum_{j \in F} c_j > \sum_{i \in E} r_i = \sum_{i \in E} \sum_{j \in F} b_{ij}, \sum_{i \in E} b_{ij} < c_j \) for at least one \( j_0 \in F \). Thus \( \phi_1(x) \to \infty \) and so \( \phi(x) \to \infty \) as \( x \to \Delta \).

It follows that \( \phi \) achieves a minimum on \( (\mathbb{R}^n)^+ \). At such a point \( \vec{x} = (\vec{x}_1, \cdots, \vec{x}_n) \), \( \partial \ln \phi(x) / \partial x_k = 0 \) for \( k = 1, \cdots, n \). Whence

\[
(7) \quad \sum_{i=1}^{m} r_i \left( a_{ik} / \sum_{j=1}^{n} a_{ij} \vec{x}_j \right) - c_k / \vec{x}_k = 0,
\]
\[
k = 1, \ldots, n. \text{ Put } y_i = r_i / \sum_{j=1}^{m} a_{ij}, i = 1, \ldots, m, \text{ and then set } D_1 = \text{diag}(y_1, \ldots, y_m), D_2 = \text{diag}(x_1, \ldots, x_n). \text{ Then } D_1AD_2 \text{ satisfies the conclusion of Theorem 1.}
\]

REFERENCES


DEPARTMENT OF MATHEMATICS, UNIVERSITY OF HOUSTON, HOUSTON, TEXAS 77004