NOTE ON M-MATRICES

By KY FAN (*Notre Dame, Indiana*)

[Received 14 August 1958]

1. BY an *M-matrix* we understand a square matrix A of the form $A = \rho I - B$, where B is a matrix with non-negative elements, I denotes the identity matrix, and ρ is a positive number greater than the absolute value of every characteristic root of B. Alternatively, an M-matrix $A = (a_{ij})$ of order n may be defined as a real matrix with $a_{ij} \leqslant 0$ $(i \neq j)$ and possessing one of the following three properties [see e.g. (4) 338–40, (3)]:

(i) there exist n positive numbers $x_j > 0$ $(1 \leqslant j \leqslant n)$ such that

$$\sum_{j=1}^{n} a_{ij} x_j > 0 \quad (1 \leqslant i \leqslant n);$$

(ii) A is non-singular and all elements of A^{-1} are non-negative;

(iii) all principal minors of A are positive.

M-matrices were first introduced and studied by Ostrowski (8), (9) under the names of 'eigentliche M-Determinanten' and 'eigentliche M-Matrizen'.

For a square matrix A of order n, and for indices

$$1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n,$$

we denote by $A(i_1, i_2,\ldots, i_p)$ the principal minor of A formed by the rows and columns with indices i_1, i_2,..., i_p. Thus, if α denotes a subset of the set $\{1, 2,\ldots, n\}$, $A(\alpha)$ will denote the principal minor of A formed by the rows and columns with indices contained in α. We denote the empty set by $\varnothing$ and define $A(\varnothing) = 1$. For two real matrices $A = (a_{ij})$, $B = (b_{ij})$ of same order, we use the notation $A \leqslant B$ to signify $a_{ij} \leqslant b_{ij}$ for all i, j.

In (8), Ostrowski proved the theorem: *Let $A = (a_{ij})$ be an M-matrix of order n, and let $B = (b_{ij})$ be a matrix of order n with real or complex elements. If $a_{ii} \leqslant |b_{ii}|$ for every i and $|b_{ij}| \leqslant |a_{ij}|$ for $i \neq j$, then $A(1, 2,\ldots, n) \leqslant |B(1, 2,\ldots, n)|$, and every element of A^{-1} is at least equal to the absolute value of the corresponding element of B^{-1}.*

For two M-matrices A, B of order n such that $A \leqslant B$, Ostrowski's theorem asserts that $A(1, 2,\ldots, n) \leqslant B(1, 2,\ldots, n)$ and $B^{-1} \leqslant A^{-1}$. Since the hypothesis remains fulfilled by any two corresponding principal submatrices of A, B, it follows that

$$\frac{A(\alpha \cup \beta)}{A(\alpha)} \leqslant \frac{B(\alpha \cup \beta)}{B(\alpha)}$$

holds for any two subsets α, β of $\{1, 2, \ldots, n\}$.

The purpose of the present note is to prove the theorem:

THEOREM. *Let A, B be two M-matrices of order n, and let α, β, γ be subsets of the set $\{1, 2, \ldots, n\}$. If $A \leqslant B$, then*

$$\Phi(A; \alpha, \beta, \gamma) \leqslant \Phi(B; \alpha, \beta, \gamma), \tag{1}$$

where
$$\Phi(A; \alpha, \beta, \gamma) = \frac{A(\alpha \cap \beta)A(\alpha \cap \gamma)A(\beta \cap \gamma)A(\alpha \cup \beta \cup \gamma)}{A(\alpha)A(\beta)A(\gamma)A(\alpha \cap \beta \cap \gamma)}. \tag{2}$$

2. The proof of the theorem requires the following lemmas:

LEMMA 1. *Let $A = (a_{ij})$, $B = (b_{ij})$ be two M-matrices of order n such that $A \leqslant B$. If $C = (c_{ij})$, $D = (d_{ij})$ are matrices of order $n-1$ defined by*

$$c_{ij} = \frac{a_{ij}a_{nn} - a_{in}a_{nj}}{a_{nn}}, \qquad d_{ij} = \frac{b_{ij}b_{nn} - b_{in}b_{nj}}{b_{nn}} \qquad (i, j = 1, 2, \ldots, n-1), \tag{3}$$

then C, D are M-matrices and $C \leqslant D$.

Proof. From (3), it is clear that $c_{ij} \leqslant 0$ $(i \neq j)$. By Sylvester's identity, we have, for any $\alpha \subset \{1, 2, \ldots, n-1\}$,

$$C(\alpha) = \frac{A(\alpha \cup \{n\})}{A(n)}. \tag{4}$$

Thus all principal minors of C are positive, and C is an M-matrix. Similarly D is an M-matrix. That $C \leqslant D$ can be easily verified.

LEMMA 2. *If A, B are two M-matrices of order n such that $A \leqslant B$, then*

$$\frac{A(1, \ldots, n)A(3, \ldots, n)}{A(2, 3, \ldots, n)A(1, 3, \ldots, n)} \leqslant \frac{B(1, \ldots, n)B(3, \ldots, n)}{B(2, 3, \ldots, n)B(1, 3, \ldots, n)}. \tag{5}$$

Proof. In the case $n = 2$, (5) becomes

$$\frac{A(1, 2)}{A(2)A(1)} \leqslant \frac{B(1, 2)}{B(2)B(1)}, \quad \text{i.e.} \quad \frac{a_{12}a_{21}}{a_{11}a_{22}} \geqslant \frac{b_{12}b_{21}}{b_{11}b_{22}},$$

which is obviously true. For the general value n $(\geqslant 3)$, we define matrices C, D of order $n-1$ by (3). Then, by the inductive assumption, we have

$$\frac{C(1, \ldots, n-1)C(3, \ldots, n-1)}{C(2, 3, \ldots, n-1)C(1, 3, \ldots, n-1)} \leqslant \frac{D(1, \ldots, n-1)D(3, \ldots, n-1)}{D(2, 3, \ldots, n-1)D(1, 3, \ldots, n-1)},$$

which, according to Sylvester's identity (4), is precisely (5).

LEMMA 3. *Let A, B be two M-matrices of order n such that $A \leqslant B$. For any two subsets α, β of $\{1, 2,..., n\}$, we have*

$$\frac{A(\alpha\cap\beta)A(\alpha\cup\beta)}{A(\alpha)A(\beta)} \leqslant \frac{B(\alpha\cap\beta)B(\alpha\cup\beta)}{B(\alpha)B(\beta)}. \tag{6}$$

Proof. We may assume that none of the sets α, β contains the other, for otherwise (6) becomes trivial. Because the hypothesis remains fulfilled by any two corresponding principal submatrices of A, B, we may assume

$$\alpha\cup\beta = \{1, 2,..., n\}.$$

Furthermore, by a simultaneous permutation of the rows and columns, it suffices to consider the following two cases.

Case 1. $\alpha = \{1,..., p\}$, $\beta = \{p+1,..., n\}$, where $1 \leqslant p < n$.

For any two fixed indices i, j such that $1 \leqslant i \leqslant p$, $1 \leqslant j \leqslant n-p$, we have, by Lemma 2,

$$\frac{A(1,..., i-1, i, p+j, p+j+1,..., n)A(1,..., i-1, p+j+1,..., n)}{A(1,..., i-1, p+j, p+j+1,..., n)A(1,..., i-1, i, p+j+1,..., n)}$$
$$\leqslant \frac{B(1,..., i-1, i, p+j, p+j+1,..., n)B(1,..., i-1, p+j+1,..., n)}{B(1,..., i-1, p+j, p+j+1,..., n)B(1,..., i-1, i, p+j+1,..., n)}.$$

Multiplying these inequalities for $j = 1, 2,..., n-p$, we get

$$\frac{A(1,..., i-1, i, p+1,..., n)A(1,..., i-1)}{A(1,..., i-1, p+1,..., n)A(1,..., i-1, i)}$$
$$\leqslant \frac{B(1,..., i-1, i, p+1,..., n)B(1,..., i-1)}{B(1,..., i-1, p+1,..., n)B(1,..., i-1, i)} \quad (1 \leqslant i \leqslant p). \tag{7}$$

Again, if we multiply these inequalities for $i = 1, 2,..., p$, the resulting relation,

$$\frac{A(1,..., p, p+1,..., n)}{A(1,..., p)A(p+1,..., n)} \leqslant \frac{B(1,..., p, p+1,..., n)}{B(1,..., p)B(p+1,..., n)},$$

is precisely (6).

Case 2. $\alpha = \{1,..., p, p+1,..., q\}$, $\beta = \{1,..., p, q+1,..., n\}$, where

$$1 \leqslant p < q < n.$$

Since (7) is already proved, we have

$$\frac{A(1,..., r-1, r, q+1,..., n)A(1,..., r-1)}{A(1,..., r-1, q+1,..., n)A(1,..., r-1, r)}$$
$$\leqslant \frac{B(1,..., r-1, r, q+1,..., n)B(1,..., r-1)}{B(1,..., r-1, q+1,..., n)B(1,..., r-1, r)}$$

for $1 \leqslant r \leqslant q$. If we multiply these inequalities for $r = p+1,..., q$, we obtain the desired inequality

$$\frac{A(1,...,p)A(1,...,n)}{A(1,...,p,p+1,...,q)A(1,...,p,q+1,...,n)} \leqslant \frac{B(1,...,p)B(1,...,n)}{B(1,...,p,p+1,...,q)B(1,...,p,q+1,...,n)}.$$

3. We proceed to prove the theorem stated in § 1.

If one of the sets α, β, γ is empty, or if two of the sets are the same, then the theorem reduces to Lemma 3. If one of α, β, γ is the entire set $\{1, 2,..., n\}$, then $\Phi(A; \alpha, \beta, \gamma) = 1$ is independent of A, and (1) becomes trivial. These facts imply that the theorem is valid for $n = 2$.

To prove the theorem for matrices $A = (a_{ij})$, $B = (b_{ij})$ of order n, observe first that we may assume $a_{ii} = b_{ii}$ for every i. In fact, let

$$\theta_i = \frac{b_{ii}}{a_{ii}} \geqslant 1$$

and define $B' = (b'_{ij})$ by $b'_{ij} = b_{ij}/\theta_i$. Then

$$a_{ii} = b'_{ii} \quad (1 \leqslant i \leqslant n), \qquad A \leqslant B'.$$

B' is again an M-matrix, and $\Phi(B'; \alpha, \beta, \gamma) = \Phi(B; \alpha, \beta, \gamma)$.

Furthermore, given two M-matrices A, B of order n with $A \leqslant B$ and $a_{ii} = b_{ii}$ $(1 \leqslant i \leqslant n)$, one can change A into B through a chain of matrices

$$A = A_0 \leqslant A_1 \leqslant ... \leqslant A_{n^2-n} = B$$

such that each A_{k+1} is obtained from A_k by replacing only a single off-diagonal element by the corresponding element of B. Each A_k is an M-matrix, because the maximum absolute value of the characteristic roots of a matrix with non-negative elements decreases when the elements of the matrix decrease and remain non-negative.

Assume now that the theorem is proved for matrices of order not exceeding $n-1$. Consider two M-matrices $A = (a_{ij})$, $B = (b_{ij})$ of order n such that $a_{n-1,n} < b_{n-1,n}$ and $a_{ij} = b_{ij}$ for all other ordered pairs of i, j. For these two matrices A, B, we are to prove that (1) holds for any three subsets α, β, γ of $\{1, 2,..., n\}$. There are two possible cases.

Case 1. At least one of α, β, γ does not contain both indices $n-1$, n; say $\{n-1, n\} \not\subset \gamma$.

In this case, we have $A(\gamma') = B(\gamma')$ for every subset γ' of γ. Therefore

$$\frac{A(\alpha \cap \gamma)A(\beta \cap \gamma)}{A(\gamma)A(\alpha \cap \beta \cap \gamma)} = \frac{B(\alpha \cap \gamma)B(\beta \cap \gamma)}{B(\gamma)B(\alpha \cap \beta \cap \gamma)}. \tag{8}$$

By Lemma 3,

$$\frac{A(\alpha \cap \beta)A(\alpha \cup \beta)}{A(\alpha)A(\beta)} \leqslant \frac{B(\alpha \cap \beta)B(\alpha \cup \beta)}{B(\alpha)B(\beta)}. \tag{6}$$

We have also

$$\frac{A(\alpha \cup \beta \cup \gamma)}{A(\alpha \cup \beta)} \leqslant \frac{B(\alpha \cup \beta \cup \gamma)}{B(\alpha \cup \beta)} \tag{9}$$

by Ostrowski's theorem (§ 1). Then (1) is obtained by multiplying (8), (6), (9).

Case 2. $\{n-1, n\} \subset \alpha \cap \beta \cap \gamma$.

Consider matrices $C = (c_{ij})$, $D = (d_{ij})$ of order $n-1$ defined from A, B by (3). Let

$$\alpha' = \alpha - \{n\}, \quad \beta' = \beta - \{n\}, \quad \gamma' = \gamma - \{n\}.$$

By Lemma 1, C, D are M-matrices and $C \leqslant D$. Then, by the inductive assumption,

$$\Phi(C;\, \alpha', \beta', \gamma') \leqslant \Phi(D;\, \alpha', \beta', \gamma'). \tag{10}$$

But, in view of Sylvester's identity (4) and the fact that $n \in \alpha \cap \beta \cap \gamma$, we have

$$\Phi(C;\, \alpha', \beta', \gamma') = \Phi(A;\, \alpha, \beta, \gamma), \qquad \Phi(D;\, \alpha', \beta', \gamma') = \Phi(B;\, \alpha, \beta, \gamma),$$

so that (10) is precisely (1). This completes the proof.

4. When the matrix $B = (b_{ij})$ is given by

$$b_{ii} = a_{ii} \quad (1 \leqslant i \leqslant n), \qquad b_{ij} = 0 \quad (i \neq j),$$

the theorem just proved reduces to the following proposition:

PROPOSITION 1. *If A is an M-matrix of order n, then for any subsets α, β, γ of $\{1, 2,..., n\}$, we have*

$$\Phi(A;\, \alpha, \beta, \gamma) \leqslant 1. \tag{11}$$

In particular, when $\gamma = \varnothing$, (11) becomes

$$A(\alpha \cap \beta)A(\alpha \cup \beta) \leqslant A(\alpha)A(\beta). \tag{12}$$

Because of the simple relation between the principal minors of A and those of A^{-1}, we also have

$$A^{-1}(\alpha \cap \beta)A^{-1}(\alpha \cup \beta) \leqslant A^{-1}(\alpha)A^{-1}(\beta) \tag{13}$$

for any M-matrix A.

According to a theorem given by Gantmaher and Krein [see (5) 117 with a later correction (4) 363, footnote 1], inequality (12) is also valid for positive-definite Hermitian matrices and completely non-negative matrices (i.e. matrices whose minors of all orders are non-negative).

For positive-definite Hermitian matrices, inequality (12) was rediscovered by the present author [(**2**) Theorem 1 and Remark, 416] and recently again by Krull (**6**).

For the stronger inequality (11), the situation is quite different. Neither for positive-definite Hermitian matrices nor for completely non-negative matrices is (11) valid. This can be seen from the simple example

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 3 \end{pmatrix}, \quad \alpha = \{1,2\}, \quad \beta = \{1,3\}, \quad \gamma = \{2,3\}.$$

Szász (**10**) proved the following inequalities for any positive-definite Hermitian matrix A of order n,

$$P_1 \geqslant P_2^{1/\binom{n-1}{1}} \geqslant \ldots \geqslant P_k^{1/\binom{n-1}{k-1}} \geqslant \ldots \geqslant P_n, \tag{14}$$

where P_k denotes the product of all k-rowed principal minors of A. This was rediscovered by Faguet (**1**). Recently, Mirsky (**7**) has given still another proof of (14), again for positive definite Hermitian matrices. Since an M-matrix A has the properties (12) and (13), both Faguet's and Mirsky's arguments can be used to prove the following proposition:

PROPOSITION 2. *The inequalities* (14) *hold, if either A or A^{-1} is an M-matrix of order n.*

5. Actually *the theorem in* § 1 (*and therefore also Proposition* 1) *can be extended to any finite number of subsets of* $\{1, 2, \ldots, n\}$. For instance, in the case of four subsets α, β, γ, δ of $\{1, 2, \ldots, n\}$, $\Phi(A; \alpha, \beta, \gamma, \delta)$ is to be defined by

$$\Phi(A; \alpha, \beta, \gamma, \delta) = \frac{A(\alpha \cap \beta) \ldots A(\gamma \cap \delta) A(\alpha \cap \beta \cap \gamma \cap \delta) A(\alpha \cup \beta \cup \gamma \cup \delta)}{A(\alpha) \ldots A(\delta) A(\alpha \cap \beta \cap \gamma) \ldots A(\beta \cap \gamma \cap \delta)}. \tag{15}$$

The inductive proof given in § 3 is applicable to the general case of any finite number of sets.

The proofs given in § 2 for Lemmas 1, 2, and 3 can be easily modified to establish the proposition:

PROPOSITION 3. *Let $A = (a_{ij})$ be an M-matrix of order n, and let $B = (b_{ij})$ be a matrix of order n with real or complex elements. If $a_{ii} \leqslant |b_{ii}|$ for every i and $|b_{ij}| \leqslant |a_{ij}|$ for $i \neq j$, then*

$$\frac{A(\alpha \cap \beta) A(\alpha \cup \beta)}{A(\alpha) A(\beta)} \leqslant \left| \frac{B(\alpha \cap \beta) B(\alpha \cup \beta)}{B(\alpha) B(\beta)} \right| \tag{16}$$

holds for any two subsets α, β of $\{1, 2, \ldots, n\}$.

It is reasonable to expect that this result could be extended to three or more sets, but it would require a proof different from that given in § 3.

REFERENCES

1. M. Faguet, 'A generalization of the Hadamard determinant inequality', *C.R. (Doklady) Acad. Sci. URSS*, 54 (1946) 761–4.
2. K. Fan, 'Some inequalities concerning positive definite Hermitian matrices', *Proc. Cambridge Phil. Soc.* 51 (1955) 414–21.
3. —— 'Topological proofs for certain theorems on matrices with non-negative elements', *Monatsh. für Math.* 62 (1958) 219–37.
4. F. R. Gantmaher, *Theory of matrices* (Russian) (Moscow, 1954).
5. —— and M. G. Krein, *Oscillatory matrices and kernels and small vibrations of mechanical systems* (Russian) (2nd ed. Moscow, 1950).
6. W. Krull, 'Über eine Verallgemeinerung der Hadamardschen Ungleichung', *Arch. der Math.* 9 (1958) 42–45.
7. L. Mirsky, 'On a generalization of Hadamard's determinantal inequality due to Szász'. ibid. 8 (1957) 274–5.
8. A. Ostrowski, 'Über die Determinanten mit überwiegender Hauptdiagonale', *Comm. Math. Helvetici*, 10 (1937) 69–96.
9. —— 'Determinanten mit überwiegender Hauptdiagonale und die absolute Konvergenz von linearen Iterationsprozessen', ibid. 30 (1956) 175–210.
10. O. Szász, 'Über eine Verallgemeinerung des Hadamardschen Determinantensatzes', *Monatsh. für Math. u. Phys.* 28 (1917) 253–7.