JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 41, 439–459 (1973)

Markov Chains and M-Matrices: Inequalities and Equalities

JULIAN KEILSON

The University of Rochester, Rochester, New York

AND

GEORGE P. H. STYAN

McGill University, Montreal, Quebec, Canada

Submitted by Ky Fan

1. INTRODUCTION

We use the theory of Markov chains to establish several new inequalities for M-matrices. If $\mathbf{A}$ is an $n \times n$ matrix with nonnegative elements and dominant characteristic root r, then for any $s > r$, $\mathbf{M}_s = s\mathbf{I} - \mathbf{A}$ is called an *M-matrix*, as introduced by Ostrowski [10] and treated by Varga [11]. Apart from their basic role in matrix theory, M-matrices are of substantial interest to probability theory in connection with finite Markov chains in discrete and continuous time.

Our study has had a dual motivation. First, we have been interested in the structure of permissible sets of complex characteristic roots of stochastic matrices and more generally of matrices with nonnegative elements. Such questions were raised, probably for the first time, by Kolmogorov [5], and discussed by Mirsky [8, 9]. Second, we have sought insight into related convexity properties of scalar functions of such matrices.

Our principal results are the following.

THEOREM 1. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix with nonnegative elements and let* r *be the dominant characteristic root. Then for all* $s > r$

$$ns^{n-1}/(s^n - r^n) \leqslant \operatorname{tr} \mathbf{M}_s^{-1} \leqslant n/(s - r), \tag{1.1}$$

where the M-matrix

$$\mathbf{M}_s = s\mathbf{I} - \mathbf{A}. \tag{1.2}$$

For $r > 0$, equality on the left of (1.1) *holds if and only if there exists a positive definite diagonal matrix* $\mathbf{D}$ *such that*

$$\mathbf{A} = r\mathbf{D}\mathbf{R}\mathbf{D}^{-1}, \tag{1.3}$$

where $\mathbf{R}$ *is an irreducible permutation matrix. Equality on the right of* (1.1) *holds if and only if*

$$\alpha_1 = \alpha_2 = \cdots = \alpha_n = r, \tag{1.4}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_n$ *denote the characteristic roots of* $\mathbf{A}$.

Our proof of Theorem 1 employs elementary probabilistic reasoning (based on Markov chain theory) as well as simple notions of convexity. A purely matrix-theoretic demonstration would be of interest.

THEOREM 2. *Let* $\mathbf{M}_s$ *be defined as in* (1.2). *Then for all* $s > r$,

$$(s - r)^n \leqslant \det \mathbf{M}_s \leqslant s^n - r^n. \tag{1.5}$$

Equality on the left holds if and only if (1.4) *is true; equality on the right holds if and only if* (1.3) *is true.*

When $\mathbf{A} = \mathbf{P}$, a substochastic matrix, $r \leqslant 1$. Hence from (1.5) we obtain

$$(1 - r)^n \leqslant \det(\mathbf{I} - \mathbf{P}) \leqslant 1 - r^n. \tag{1.6}$$

If $\mathbf{P}$ is stochastic, $r = 1$ and we find:

$$\prod_{j=2}^{n} (1 - \lambda_j) \leqslant n, \tag{1.7}$$

where $\lambda_1 = 1, \lambda_2, \ldots, \lambda_n$ are the characteristic roots of the stochastic matrix $\mathbf{P}$. Equality in (1.7) holds if and only if $\mathbf{P} = \mathbf{R}$, an irreducible permutation matrix. If the characteristic polynomial is written as

$$\det(\lambda\mathbf{I} - \mathbf{P}) = \sum_{s=1}^{n} a_s^*(\lambda - 1)^s \tag{1.8}$$

then

$$0 \leqslant a_s^* \leqslant \binom{n}{s} = \frac{n!}{s!(n-s)!}; \qquad s = 1, 2, \ldots, n. \tag{1.9}$$

If an n-state irreducible Markov chain governed by the stochastic matrix $\mathbf{P}$ has fundamental matrix $\mathbf{Z}$ [cf. (6.14) and Corollary 6.1], then

$$\det \mathbf{Z} \geqslant 1/n; \qquad \operatorname{tr} \mathbf{Z} \geqslant \tfrac{1}{2}(n + 1). \tag{1.10}$$

Proof of (1.1) is provided by Theorems 3.1, 4.1, and 4.2 with Theorem 3.2 establishing the condition (1.3). Corollary 4.3 proves condition (1.4), and Corollaries 3.1, 4.1, and 4.2 establish (1.5). Proof of (1.7) is given by Theorem 6.1 while Theorem 6.3 and Corollary 6.1 establish (1.9) and (1.10), respectively.

2. Some Preliminary Results

Our proof of Theorem 1 depends on a probabilistic argument based on Markov chain theory. We also need the notion of reducibility:

DEFINITION 2.1. For $n \geqslant 2$, an $n \times n$ matrix $\mathbf{A}$ is *reducible* if there exists an $n \times n$ permutation matrix $\mathbf{R}$ such that

$$\mathbf{R}'\mathbf{A}\mathbf{R} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_2 \end{pmatrix}, \tag{2.1}$$

with $\mathbf{A}_1$ and $\mathbf{A}_2$ both square. Otherwise $\mathbf{A}$ is *irreducible*. When $n = 1$ $\mathbf{A}$ is irreducible if its single entry is nonzero and reducible otherwise.

Some writers prefer the term *indecomposable*. We follow Varga [11, pp. 18–19].

LEMMA 2.1. *Let X_k be a temporally homogeneous finite Markov chain in discrete time $k = 0, 1,\ldots$ governed by an irreducible $n \times n$ stochastic matrix $\mathbf{P}$. Let T_m be the random elapsed time till the first return to state m $(= 1, 2,\ldots, n)$ after $X_0 = m$. Then*

$$\sum_{k=0}^{\infty} w^k \operatorname{tr} \mathbf{P}^k = \sum_{m=1}^{n} [1 - \mathscr{E}(w^{T_m})]^{-1}; \qquad 0 \leqslant w < 1, \tag{2.2}$$

where $\mathscr{E}$ denotes mathematical expectation.

Proof. Let the indicator function

$$\begin{aligned} I_m(k) &= 1 \qquad \text{whenever} \qquad X_k = m \\ &= 0 \qquad \text{otherwise.} \end{aligned} \tag{2.3}$$

Then $\mathscr{E}[I_m(k) \mid X_0 = m] = \mathscr{P}(X_k = m \mid X_0 = m) = p_{mm}^{(k)}$, the mth diagonal element of $\mathbf{P}^k$. Hence for $0 \leqslant w < 1$,

$$\sum_{k=0}^{\infty} w^k \operatorname{tr} \mathbf{P}^k = \sum_{m=1}^{n} \sum_{k=0}^{\infty} w^k p_{mm}^{(k)} = \sum_{m=1}^{n} \mathscr{E}\left[\sum_{k=0}^{\infty} w^k I_m(k) \mid X_0 = m\right]. \tag{2.4}$$

If T_{mj} denotes the random elapsed time between the $(j-1)$th and jth return to state m after $X_0 = m$ $(j = 1, 2,...; m = 1, 2,..., n)$, then

$$T_{m1} = T_m, T_{m2}, ..., T_{mk}, ...$$

are independently and identically distributed. Let $R_{mj} = T_{m1} + \cdots + T_{mj}$ be the time of the jth return to state m. Then for $0 \leqslant w < 1$

$$\begin{aligned}\mathscr{E}\left[\sum_{k=0}^{\infty} w^k I_m(k) \mid X_0 = m\right] &= \mathscr{E}[1 + w^{R_{m1}} + w^{R_{m2}} + \cdots] \\ &= 1 + \mathscr{E}(w^{T_{m1}}) + \mathscr{E}(w^{T_{m1}+T_{m2}}) + \cdots \qquad (2.5) \\ &= \sum_{k=0}^{\infty} [\mathscr{E}(w^{T_m})]^k = [1 - \mathscr{E}(w^{T_m})]^{-1},\end{aligned}$$

which substituted in (2.4) yields (2.2). (Q.E.D.)

LEMMA 2.2. *Let* $\mathbf{P}$ *and* T_m *be defined as in Lemma* 2.1. *Then there exists an* $n \times 1$ *column vector* $\boldsymbol{l} = \{l_m\}$ *with positive elements such that* $\boldsymbol{l}'\mathbf{P} = \boldsymbol{l}'$, $\mathscr{E}(T_m) = 1/l_m$, *and*

$$\sum_{m=1}^{n} l_m = \sum_{m=1}^{n} [\mathscr{E}(T_m)]^{-1} = 1. \qquad (2.6)$$

Proof. See, e.g., Karlin [3, pp. 63–65]. (Q.E.D.)

LEMMA 2.3. *The function*

$$f(x) = [1 - e^{-1/x}]^{-1} \qquad (2.7)$$

is strictly convex for all $x > 0$.

Proof. The second derivative is

$$f'' = x^{-4} f(f-1)(2f - 2x - 1). \qquad (2.8)$$

As $f > 1$ it suffices to prove $2f > 2x + 1$. If $t = 1/x$, this is equivalent to $2 + t > (2 - t)\, e^t$ which is so since

$$2 + t - (2 - t)\, e^t = \sum_{k=3}^{\infty} (k-2)\, t^k/k! > 0. \qquad \text{(Q.E.D.)}$$

LEMMA 2.4. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix and* s *a real scalar not equal to a characteristic roof of* $\mathbf{A}$. *Then*

$$\left\{\frac{d}{ds}[\det(s\mathbf{I} - \mathbf{A})]\right\} \Big/ \det(s\mathbf{I} - \mathbf{A}) = \operatorname{tr}(s\mathbf{I} - \mathbf{A})^{-1}. \qquad (2.9)$$

If $\det(s\mathbf{I} - \mathbf{A}) > 0$ *then*

$$\frac{d}{ds}\{\log[\det(s\mathbf{I} - \mathbf{A})]\} = \operatorname{tr}(s\mathbf{I} - \mathbf{A})^{-1}. \tag{2.10}$$

Proof. The left-hand side of (2.9) equals

$$\begin{aligned}\left\{\frac{d}{ds}\left[\prod_{j=1}^{n}(s-\alpha_j)\right]\right\}\Big/\prod_{j=1}^{n}(s-\alpha_j) &= \sum_{i=1}^{n}\prod_{\substack{j=1\\ j\neq i}}^{n}(s-\alpha_j)\Big/\prod_{j=1}^{n}(s-\alpha_j)\\ &= \sum_{j=1}^{n}(s-\alpha_j)^{-1}\end{aligned} \tag{2.11}$$

and (2.9) is established. We have used $\alpha_1, \alpha_2, \ldots, \alpha_n$ to denote the characteristic roots of $\mathbf{A}$ and

$$\operatorname{tr}(s\mathbf{I} - \mathbf{A})^{-1} = \sum_{j=1}^{n}(s-\alpha_j)^{-1}.$$

If $\det(s\mathbf{I} - \mathbf{A})$ is positive, its logarithm is defined and (2.10) follows directly from (2.9). (Q.E.D.)

LEMMA 2.5. *If* $\mathbf{A}$ *is an* $n \times n$ *stochastic matrix with characteristic polynomial*

$$\det(\lambda\mathbf{I} - \mathbf{A}) = \lambda^n - 1 \tag{2.12}$$

then $\mathbf{A}$ *is an irreducible permutation matrix.*

Proof. If $\mathbf{A}$ is reducible its rows and columns may be rearranged into the form (2.1). Then $\mathbf{A}_2$ is stochastic and its dominant root of unity is a root of $\mathbf{A}$. Moreover, $\mathbf{A}_1$ has dominant root $r_1 \leqslant 1$. If $r_1 = 1$ $\mathbf{A}$ has a double root of unity contradicting (2.12); if $r_1 < 1$ (2.12) is also contradicted. Hence $\mathbf{A}$ is irreducible and the lemma follows from the general representation of irreducible stochastic matrices with k roots of absolute value 1 {cf. Varga [11, p. 38]}. (Q.E.D.)

3. SOME INTERMEDIATE RESULTS FOR IRREDUCIBLE MATRICES

We proceed to establish the left-hand side of (1.1) when $\mathbf{A}$ is irreducible.

THEOREM 3.1. *Let* $\mathbf{A}$ *be an* $n \times n$ *irreducible matrix with nonnegative elements and dominant characteristic root* r. *Then for all* $s > r$ *the* M-*matrix* $\mathbf{M}_s = s\mathbf{I} - \mathbf{A}$ *satisfies*

$$\operatorname{tr}\mathbf{M}_s^{-1} \geqslant ns^{n-1}/(s^n - r^n). \tag{3.1}$$

Proof. As $s > r$, $\mathbf{M}_s$ is nonsingular and a geometric power series expansion gives

$$s \operatorname{tr} \mathbf{M}_s^{-1} = \sum_{k=0}^{\infty} \operatorname{tr} \mathbf{A}^k / s^k. \tag{3.2}$$

Since $\mathbf{A}$ is irreducible, there exists a stochastic matrix $\mathbf{P}$ and a positive definite diagonal matrix $\mathbf{D}$ such that $\mathbf{A} = r\mathbf{DPD}^{-1}$. With $w = r/s$, substitution in (3.2) yields

$$s \operatorname{tr} \mathbf{M}_s^{-1} = \sum_{k=0}^{\infty} w^k \operatorname{tr} \mathbf{P}^k; \qquad 0 < w < 1. \tag{3.3}$$

From (2.2),

$$s \operatorname{tr} \mathbf{M}_s^{-1} = \sum_{m=1}^{n} [1 - \mathscr{E}(w^{T_m})]^{-1}.$$

For $x > 0$ and $0 < w < 1$, w^x is convex. Jensen's inequality {cf. e.g., Feller [1, pp. 153–4]} then implies $\mathscr{E}(w^{T_m}) \geqslant w^{\mathscr{E}(T_m)}$ and so

$$s \operatorname{tr} \mathbf{M}_s^{-1} \geqslant \sum_{m=1}^{n} [1 - w^{\mathscr{E}(T_m)}]^{-} = \sum_{m=1}^{n} [1 - w^{1/l_m}]^{-1} \tag{3.4}$$

using Lemma 2.2. Consider the function

$$f(x_m) = [1 - e^{-1/x_m}]^{-1} = [1 - w^{1/l_m}]^{-1}, \tag{3.5}$$

where $x_m = -\, l_m / \log w > 0$. By Lemma 2.3 $f(x_m)$ is strictly convex in x_m and hence using (3.4) and Lemma 2.2,

$$\begin{aligned} (s/n) \operatorname{tr} \mathbf{M}_s^{-1} &\geqslant (1/n) \sum_{m=1}^{n} f(x_m) \geqslant f\left(\sum_{m=1}^{n} x_m / n\right) \\ &= f(-\,[n \log w]^{-1}) \\ &= (1 - w^n)^{-1} = s^n/(s^n - r^n), \end{aligned} \tag{3.6}$$

as $w = r/s$. Thus (3.1) is proven. (Q.E.D.)

COROLLARY 3.1. *Let* $\mathbf{M}_s = s\mathbf{I} - \mathbf{A}$ *be defined as in Theorem* 3.1. *Then*

$$\det \mathbf{M}_s \leqslant s^n - r^n, \qquad s > r. \tag{3.7}$$

Proof. If $s_0 > r$, (3.1) and (2.10) imply

$$0 \leqslant \int_{s_0}^{\infty} \left[\operatorname{tr} \mathbf{M}_s^{-1} - \frac{n s^{n-1}}{s^n - r^n}\right] ds = \log \left[\frac{s_0^{\,n} - r^n}{\det(s_0 \mathbf{I} - \mathbf{A})}\right], \tag{3.8}$$

since

$$\log[\det(t\mathbf{I} - \mathbf{A})] - \log(t^n - r^n) = \log[\det(\mathbf{I} - \mathbf{A}/t)] - \log[1 - (r/t)^n]$$

converges to 0 as $t \to \infty$. Hence (3.7) follows. (Q.E.D.)

THEOREM 3.2. *Let* $\mathbf{M}_s = s\mathbf{I} - \mathbf{A}$ *be defined as in Theorem* 3.1. *If there exists an* $s = s_0 > r$ *such that*

$$\operatorname{tr} \mathbf{M}_s^{-1} = ns^{n-1}/(s^n - r^n) \tag{3.9}$$

or such that

$$\det \mathbf{M}_s = s^n - r^n, \tag{3.10}$$

then simultaneously (3.9) *holds for all* s *not equal to a characteristic root of* $\mathbf{A}$ *and* (3.10) *holds for all* s. *Moreover,* (3.9) *or* (3.10) *holds if and only if there exists a positive definite diagonal matrix* $\mathbf{D}$ *such that*

$$\mathbf{A} = r\mathbf{D}\mathbf{R}\mathbf{D}^{-1}, \tag{3.11}$$

where $\mathbf{R}$ *is an irreducible permutation matrix.*

Proof. Let (3.9) hold for some $s = s_0 > r$. Then in the proof of Theorem 3.1 equality must hold in (3.4) so that $\mathscr{E}(w^{T_m}) = w^{\mathscr{E}(T_m)}$ and $T_m = 1/l_m$ with probability one. As equality must also hold throughout (3.6), $x_m = -l_m/\log w$ must be independent of m and hence $T_m = n$ with probability one ($\sum_{m=1}^{n} l_m = 1$). Thus $\mathbf{P}$ is an irreducible permutation matrix and (3.11) is true. Therefore (3.10) holds for all s (characteristic polynomial for $\mathbf{A}$); (3.9) follows by applying (2.9) to (3.10). Now suppose (3.10) holds for some $s = s_0 > r$. For any $s > r$,

$$\det \mathbf{M}_s = \left[\left(\sum_{i=1}^{n} \det \mathbf{M}_i\right)\Big/ \operatorname{tr} \mathbf{M}_s^{-1}\right] \leqslant \left[\left(\sum_{i=1}^{n} \det \mathbf{M}_i\right)\Big/ ns^{n-1}\right] (s^n - r^n), \tag{3.12}$$

using (3.1), where $\mathbf{M}_i$ is the $(n-1) \times (n-1)$ principal submatrix formed from $\mathbf{M}$ by removing its ith row and column. Then $\mathbf{M}_i$ is also an M-matrix since the dominant characteristic root r_i of $s\mathbf{I} - \mathbf{M}_i$ cannot exceed r {cf. Gantmacher [2, pp. 83–84]}. Thus, using (3.7),

$$\det \mathbf{M}_s \leqslant \left[\left(\sum_{i=1}^{n} (s^{n-1} - r_i^{n-1})\Big/ ns^{n-1}\right)\right] (s^n - r^n) \leqslant s^n - r^n. \tag{3.13}$$

For $s = s_0$, however, equality must hold throughout (3.12) and (3.13), which implies that (3.9) holds for $s = s_0$, completing the proof. (Q.E.D.)

4. Extensions to Reducible Matrices

We now show that inequality (3.1) is strict when $\mathbf{A}$ is reducible ($n > 1$).

Theorem 4.1. *Let* $\mathbf{A}$ *be an* $n \times n$ ($n > 1$) *reducible matrix with non-negative elements and dominant characteristic root* r. *Then for all* $s > r$ *the M-matrix* $\mathbf{M}_s = s\mathbf{I} - \mathbf{A}$ *satisfies*

$$\operatorname{tr} \mathbf{M}_s^{-1} > ns^{n-1}/(s^n - r^n). \tag{4.1}$$

Proof. There exists a permutation matrix $\mathbf{R}$ such that {cf. Varga [11, p. 46]}

$$\mathbf{R}'\mathbf{A}\mathbf{R} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1h} \\ \mathbf{0} & \mathbf{A}_2 & \cdots & \mathbf{A}_{2h} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_h \end{bmatrix}, \tag{4.2}$$

where each submatrix $\mathbf{A}_i$ is

$$n_i \times n_i \qquad \left(i = 1, 2,..., h; h > 1; \sum_{i=1}^{h} n_i = n\right)$$

and irreducible ($n_i > 1$) or a scalar (cf. Definition 2.1). Thus, using (3.2)

$$\operatorname{tr} \mathbf{M}_s^{-1} = \sum_{k=0}^{\infty} \operatorname{tr} \mathbf{A}_m{}^k/s^{k+1} + \sum_{\substack{i=1 \\ i \neq m}}^{h} \sum_{k=0}^{\infty} \operatorname{tr} \mathbf{A}_i{}^k/s^{k+1}, \tag{4.3}$$

where m identifies a diagonal submatrix in (4.2) whose dominant characteristic root is r. As (3.1) holds for $n = n_m = 1$ as well as $n_m > 1$ and $\mathbf{A} = \mathbf{A}_m$ irreducible, we obtain after truncation of the last term in (4.3) at $k = 0$,

$$\begin{aligned} \operatorname{tr} \mathbf{M}_s^{-1} &\geqslant (n_m s^{n_m-1}/(s^{n_m} - r^{n_m})) + (n - n_m)/s \\ &= (1/s)\{[n_m r^{n_m}/(s^{n_m} - r^{n_m})] + n\} \\ &> (1/s)\{[nr^n/(s^n - r^n)] + n\} = ns^{n-1}/(s^n - r^n), \end{aligned} \tag{4.4}$$

as $n_m < n$ ($h > 1$) and $xr^x/(s^x - r^x) = x/(e^{\alpha x} - 1)$ is strictly decreasing in x $[\alpha = \log(s/r) > 0]$. (Q.E.D.)

Corollary 4.1. *Let* $\mathbf{M}_s = s\mathbf{I} - \mathbf{A}$ *be defined as in Theorem* 4.1. *Then for all* $s > r$,

$$\det \mathbf{M}_s < s^n - r^n. \tag{4.5}$$

Proof. Following the proof of Corollary 3.1, use of (4.1) instead of (3.1) gives strict inequality on the left of (3.8), and (4.5) follows. (Q.E.D.)

We now complete the proofs of Theorems 1 and 2.

THEOREM 4.2. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix with nonnegative elements and dominant characteristic root* r. *Then for all* $s > r$ *the M-matrix* $\mathbf{M}_s = s\mathbf{I} - \mathbf{A}$ *satisfies*

$$\operatorname{tr} \mathbf{M}_s^{-1} \leqslant n/(s-r). \tag{4.6}$$

Proof. Using (3.2) we have

$$\operatorname{tr} \mathbf{M}_s^{-1} = \sum_{k=0}^{\infty} \operatorname{tr} \mathbf{A}^k/s^{k+1} = \sum_{k=0}^{\infty} \sum_{m=1}^{n} a_{mm}^{(k)}/s^{k+1} \leqslant \sum_{k=0}^{\infty} nr^k/s^{k+1} = n/(s-r), \tag{4.7}$$

as $a_{mm}^{(k)}$, the mth diagonal element of $\mathbf{A}^k$, is not greater than r^k, the dominant root of $\mathbf{A}^k$ (cf. Gantmacher [2, p. 84]). (Q.E.D.)

COROLLARY 4.2. *Let* $\mathbf{M}_s = s\mathbf{I} - \mathbf{A}$ *be defined as in Theorem* 4.2. *Then for all* $s > r$

$$\det \mathbf{M}_s \geqslant (s-r)^n. \tag{4.8}$$

Proof. If $s_0 > r$, (4.6) and (2.10) imply

$$0 \leqslant \int_{s_0}^{\infty} \left[\frac{n}{s-r} - \operatorname{tr} \mathbf{M}_s^{-1}\right] ds = \log \left[\frac{\det(s_0\mathbf{I} - \mathbf{A})}{(s_0 - r)^n}\right], \tag{4.9}$$

since

$$\log[\det(t\mathbf{I} - \mathbf{A})] - \log[(t-r)^n] = \log[\det(\mathbf{I} - \mathbf{A}/t)] - \log[(1 - r/t)^n]$$

converges to 0 as $t \to \infty$. Hence (4.8) follows. (Q.E.D.)

COROLLARY 4.3. *Let* $\mathbf{M}_s = s\mathbf{I} - \mathbf{A}$ *be defined as in Theorem* 4.2. *Then if there exists an* $s = s_0 > r$ *such that*

$$\operatorname{tr} \mathbf{M}_s^{-1} = n/(s-r) \tag{4.10}$$

or such that

$$\det \mathbf{M}_s = (s-r)^n, \tag{4.11}$$

then simultaneously (4.10) *holds for all* s *not equal to a characteristic root of* $\mathbf{A}$ *and* (4.11) *holds for all* s. *Moreover,* (4.10) *or* (4.11) *holds if and only if there exists a permutation matrix* $\mathbf{R}$ *such that* $\mathbf{R}'\mathbf{A}\mathbf{R}$ *is triangular with all its diagonal elements equal to* r. *In this case all characteristic roots of* $\mathbf{A}$ *are equal to* r.

Proof. For $n = 1$ there is nothing to prove. Suppose $n > 1$ and (4.10) holds for some $s = s_0$. Then in the proof of Theorem 4.2 equality holds throughout for $s = s_0$ and hence $a_{mm} = r$, $m = 1,..., n$. Thus $\mathbf{A}$ must be reducible and there exists a permutation matrix $\mathbf{R}$ such that $\mathbf{R}'\mathbf{A}\mathbf{R}$ has the form (4.2). Let $a_{mi}^{(k)}$ be the mth diagonal element of $\mathbf{A}_i^k$; then using (4.7) we have for all $s > r$ and $\mathbf{A}$ reducible,

$$\begin{aligned} \operatorname{tr} \mathbf{M}_s^{-1} &= \sum_{k=0}^{\infty} \sum_{i=1}^{h} \operatorname{tr} \mathbf{A}_i^k / s^{k+1} = \sum_{k=0}^{\infty} \sum_{i=1}^{h} \sum_{m=1}^{n_i} a_{mi}^{(k)} / s^{k+1} \\ &\leqslant \sum_{k=0}^{\infty} \sum_{i=1}^{h} n_i r_i^k / s^{k+1} \leqslant \sum_{k=0}^{\infty} n r^k / s^{k+1} = n/(s - r). \end{aligned} \tag{4.12}$$

For $s = s_0$ (4.10) implies equality throughout (4.12) and hence $n_i = 1$, $i = 1,..., h$ as otherwise $a_{mi}^{(k)} < r_i^k$ for at least one i. We also require $r_i = r$, $i = 1,..., h = n$, so that $\mathbf{R}'\mathbf{A}\mathbf{R}$ is triangular. Hence (4.11) holds for all s (characteristic polynomial of $\mathbf{A}$). Applying (2.9) to (4.11) then shows (4.10) valid for all s not equal to a characteristic root of $\mathbf{A}$. Now suppose (4.11) holds for some $s = s_0$. For all $s > r$, (4.6) implies

$$\det \mathbf{M}_s = \left(\sum_{i=1}^{n} \det \mathbf{M}_i\right) \Big/ (\operatorname{tr} \mathbf{M}_s^{-1}) \geqslant \left[\left(\sum_{i=1}^{n} \det \mathbf{M}_i\right) \Big/ n/(s - r)\right], \tag{4.13}$$

where $\mathbf{M}_i$ is defined as in (3.12). Since $\mathbf{M}_i$ is also an M-matrix, (4.8) applied to the numerator in (4.13) implies

$$\det \mathbf{M}_s \geqslant \left[\sum_{i=1}^{n} (s - r_i)^{n-1}\right] \Big/ n/(s - r) \geqslant (s - r)^n, \tag{4.14}$$

where $r_i (\leqslant r)$ is the dominant characteristic root of $\mathbf{M}_i$. For $s = s_0$, however, equality must hold throughout (4.13) and (4.14) which implies that (4.10) holds for $s = s_0$, completing the proof. (Q.E.D.)

COROLLARY 4.4. *If in Corollary* 4.3 $\mathbf{A}$ *is a stochastic matrix then* (4.10) *or* (4.11) *holds if and only if* $\mathbf{A} = \mathbf{I}$.

Proof. When $\mathbf{A}$ is stochastic $r = 1$ and $\mathbf{R}'\mathbf{A}\mathbf{R}$ is stochastic with each diagonal element equal to 1. Hence $\mathbf{R}'\mathbf{A}\mathbf{R} = \mathbf{I} = \mathbf{A}$. (Q.E.D.)

A geometric interpretation of the bounds for $\det \mathbf{M}_s$ may be helpful. The characteristic roots $\alpha_1 = r$, α_2,..., α_n of $\mathbf{A}$ lie in the circle of radius r in the complex plane with center at the origin. For $s > r$,

$$\det \mathbf{M}_s = \prod_{j=1}^{n} (s - \alpha_j) = \prod_{j=1}^{n} | s - \alpha_j |$$

is the product of the distances from $\alpha_0 = s$ to $\alpha_1 = r$, $\alpha_2, \ldots, \alpha_n$. As

$$(s - r)^n \leqslant \det \mathbf{M}_s \leqslant s^n - r^n, \tag{4.15}$$

the minimum occurs when $\alpha_1 = r = \alpha_2 = \cdots = \alpha_n$ and the maximum when $\alpha_2, \alpha_3, \ldots, \alpha_n$ complete the regular n-gon inscribed in the circle anchored at $\alpha_1 = r$; e.g., with $n = 5$, we have the configuration shown in Fig. 1.

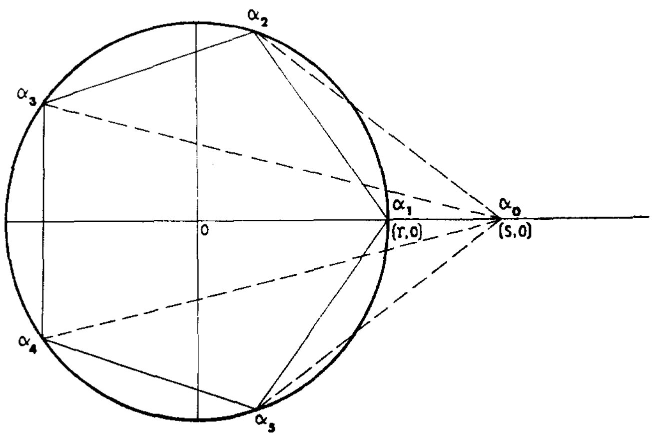

FIGURE 1

5. RESULTS RELATED TO HADAMARD'S INEQUALITY

A special case of Hadamard's inequality {cf. Mirsky [7, p. 419]} is

$$\det \mathbf{B} \leqslant \prod_{i=1}^{n} b_{ii}, \tag{5.1}$$

where the $n \times n$ matrix $\mathbf{B}$ is positive semidefinite. A parallel inequality for M-matrices can be proven from the results of Ostrowski [10]; we give a new proof which we use later on.

THEOREM 5.1. *Let $\mathbf{A}$ be an $n \times n$ matrix with nonnegative elements and dominant characteristic root r. Then for all $s > r$ the M-matrix $\mathbf{M}_s = s\mathbf{I} - \mathbf{A}$ satisfies*

$$\det(s\mathbf{I} - \mathbf{A}) = \det \mathbf{M}_s \leqslant \prod_{i=1}^{n} m_{ii} = \prod_{i=1}^{n} (s - a_{ii}), \tag{5.2}$$

where $m_{ii} = s - a_{ii}$ *is the* ith *diagonal element of* $\mathbf{M}_s$. *Equality holds in* (5.2) *if and only if there exists a permutation matrix* $\mathbf{R}$ *such that* $\mathbf{R}'\mathbf{A}\mathbf{R}$ *is triangular.*

Proof. Let $\mathbf{M}_s = \{m_{ij}\}$ and $\mathbf{M}_s^{-1} = \{m^{ij}\}$. Then

$$m_{11}m^{11} + m_{12}m^{21} + \cdots + m_{1n}m^{n1} = 1. \tag{5.3}$$

As $\mathbf{M}_s^{-1} = \sum_{k=0}^{\infty} \mathbf{A}^k/s^{k+1}$, $m^{ij} \geqslant 0$ for all i, j. Also $m_{ij} \leqslant 0$ for all $i \neq j$ as such $m_{ij} = -a_{ij}$. Hence $m_{11}m^{11} \geqslant 1$, and by the same reasoning,

$$m^{ii} \geqslant 1/m_{ii}, \qquad i = 1, 2,..., n. \tag{5.4}$$

If $\mathbf{M}_i$ denotes the principal submatrix of $\mathbf{M}_s$ with row and column i removed, then $m^{ii} = \det \mathbf{M}_i / \det \mathbf{M}$ and so

$$\det \mathbf{M} \leqslant m_{ii}(\det \mathbf{M}_i), \qquad i = 1, 2,..., n. \tag{5.5}$$

Applying induction, (5.5) yields (5.2), as every principal submatrix of an M-matrix is itself an M-matrix [cf. (3.12)].

Equality in (5.2) holds if and only if equality holds in (5.4) for all i. Suppose first $\mathbf{A}$ is irreducible. Then $m^{ij} > 0$ for all i, j. Thus equality holds in (5.4) for some i if and only if $m_{ij} = 0 = a_{ij}$ for $j \neq i$. But in this case $\mathbf{A}$ must be reducible. Suppose then that $\mathbf{A}$ is reducible; there exists a permutation matrix $\mathbf{R}$ such that $\mathbf{R}'\mathbf{A}\mathbf{R}$ has the form (4.2) and so

$$\det \mathbf{M}_s = \prod_{g=1}^{h} \det(s\mathbf{I} - \mathbf{A}_g) = \prod_{g=1}^{h} \prod_{l=1}^{n_g} (s - a_{ll}) \tag{5.6}$$

and hence

$$\det(s\mathbf{I} - \mathbf{A}_g) = \prod_{l=1}^{n_g} (s - a_{ll}), \qquad g = 1,..., h.$$

But this is impossible unless $n_g = 1, g = 1,..., h$ and hence $\mathbf{R}'\mathbf{A}\mathbf{R}$ is triangular and the stated condition for equality in (5.2) is established. (Q.E.D.)

THEOREM 5.2. *Let* $\mathbf{M}_s = s\mathbf{I} - \mathbf{A}$ *be defined as in Theorem* 5.1. *Then*

$$\operatorname{tr}(s\mathbf{I} - \mathbf{A})^{-1} = \operatorname{tr} \mathbf{M}_s^{-1} \geqslant \sum_{i=1}^{n} \frac{1}{m_{ii}} = \sum_{i=1}^{n} \frac{1}{s - a_{ii}}, \tag{5.7}$$

where $m_{ii} = s - a_{ii}$ *is the* ith *diagonal element of* $\mathbf{M}_s$. *Equality holds in* (5.7) *if and only if equality holds in* (5.2).

Proof. Summing (5.4) from $i = 1$ to n yields (5.7) directly. For equality in (5.7) equality must hold in (5.4) for all $i = 1, 2,..., n$. Hence equality holds in (5.2). Conversely equality in (5.2) implies equality in (5.5) for all $i = 1, 2,..., n$ and hence equality in (5.7). (Q.E.D.)

Marcus and Minc [6, p. 127] present a similar proof of Theorem 5.1 for $r < 1$. The inequalities (5.2) and (5.7) give alternate bounds to those in (3.7) and (3.1), respectively. For $s - r$ large, the bounds (5.2) and (5.7) appear tighter, while for $s - r$ small (3.7) and (3.1) seem superior.

6. Special Results for Stochastic Matrices

Let $\mathbf{P}$ be an $n \times n$ stochastic matrix with a simple characteristic root of unity. (This is implied by, but does not imply irreducibility.) Then there exists a unique $n \times 1$ vector $\boldsymbol{l} = \{l_m\}$, $l_m \geqslant 0$, such that $\boldsymbol{l}'\mathbf{P} = \boldsymbol{l}'$ and $\boldsymbol{l}'\mathbf{1} = 1$, where $\mathbf{1}$ is an $n \times 1$ vector with each component unity. If adj$(\cdot)$ denotes the adjugate matrix of first cofactors transposed,

$$[\text{adj}(\mathbf{I} - \mathbf{P})]\,(\mathbf{I} - \mathbf{P}) = (\mathbf{I} - \mathbf{P})\,[\text{adj}(\mathbf{I} - \mathbf{P})] = [\det(\mathbf{I} - \mathbf{P})]\,\mathbf{I} = \mathbf{0}, \tag{6.1}$$

since $\det(\mathbf{I} - \mathbf{P}) = 0$ {cf. Mirsky [7, p. 88]}. When $\mathbf{P}$ has a simple root of unity, $\mathbf{I} - \mathbf{P}$ has rank $n - 1$; thus every row of adj$(\mathbf{I} - \mathbf{P})$ must be proportional to $\boldsymbol{l}'$ and every column proportional to $\mathbf{1}$. Hence

$$\text{adj}(\mathbf{I} - \mathbf{P}) = k\mathbf{1}\boldsymbol{l}' \tag{6.2}$$

where

$$k = \text{tr}[\text{adj}(\mathbf{I} - \mathbf{P})] = \sum_{i=1}^{n} \det(\mathbf{I} - \mathbf{P}_i) = \prod_{j=2}^{n} (1 - \lambda_j). \tag{6.3}$$

Here, following (3.11), $\mathbf{P}_i$ denotes the $(n - 1) \times (n - 1)$ principal submatrix of $\mathbf{P}$ after deletion of its ith row and ith column, and $\lambda_1 = 1, \lambda_2, \ldots, \lambda_n$ are the characteristic roots of $\mathbf{P}$. We use the results developed in the preceding sections to obtain an upper bound for k. Each row of (6.2) gives the stationary distribution $\boldsymbol{l}'$ (after division by k) of the Markov chain governed by $\mathbf{P}$.

The roots $\lambda_2, \lambda_3, \ldots, \lambda_n$ must lie in a subset of the unit circle in the complex plane; the problem of finding necessary and sufficient conditions for a *set* of $n - 1$ points in the unit circle to be the nonunit roots of a stochastic matrix is "very difficult" and "has not yet been solved or even discussed in the literature" {Mirsky [8, 9]}. We note that $\prod_{j=2}^{n} (1 - \lambda_j)$ is the product of the distances from the $n - 1$ nonunit roots to the anchor point $(1, 0)$ [cf. Fig. 1]. We obtain the following.

Theorem 6.1. *Let the $n \times n$ stochastic matrix $\mathbf{P}$ have characteristic roots $\lambda_1 = 1, \lambda_2, \ldots, \lambda_n$. Then*

$$0 \leqslant \prod_{j=2}^{n} (1 - \lambda_j) \leqslant n, \tag{6.4}$$

with equality on the right if and only if $\mathbf{P}$ *is an irreducible permutation matrix* ($\lambda_2, \lambda_3, \ldots, \lambda_n$ *are then the* $n-1$ *nonunit* nth *roots of unity*).

Proof. Let $\mathbf{P}_i$ be defined as in (6.3) and let r_i be its dominant characteristic root ($i = 1, 2, \ldots, n$). Then $0 \leqslant r_i \leqslant 1$. Since (3.7) remains valid by continuity for $s = 1$, we obtain $\det(\mathbf{I} - \mathbf{P}_i) \leqslant 1 - r_i^{n-1}$; hence from (6.3) it follows that

$$0 \leqslant \prod_{j=2}^{n} (1 - \lambda_j) = \sum_{i=1}^{n} \det(\mathbf{I} - \mathbf{P}_i) \leqslant \sum_{i=1}^{n} (1 - r_i^{n-1}) \leqslant n, \tag{6.5}$$

and (6.4) is established. Equality on the right of (6.5) holds if and only if $r_1 = r_2 = \cdots = r_n = 0$. Thus, all principal submatrices of order less than n are singular, and so the characteristic polynomial of $\mathbf{P}$ is

$$\det(\lambda \mathbf{I} - \mathbf{P}) = \lambda^n - 1, \tag{6.6}$$

since the coefficients of $\lambda^{n-1}, \lambda^{n-2}, \ldots, \lambda^2, \lambda$ are sums of principal minors of $\mathbf{P}$ of order $1, 2, \ldots, n-1$, respectively. Hence, using Lemma 2.5, equality on the right of (6.4) holds if and only if $\mathbf{P}$ is an irreducible permutation matrix. (Q.E.D.)

When $\prod_{j=2}^{n} (1 - \lambda_j) = n$ the points $\lambda_2, \lambda_3, \ldots, \lambda_n$ complete the regular n-gon in the unit circle. The product of the distances to these $n - 1$ vertices from the anchor point at $(1, 0)$ is n. When the λ_j may only lie on the real line the upper bound for $\prod_{j=2}^{n} (1 - \lambda_j)$ is much smaller, for then $1 - \lambda_j > 0$ and so

$$\begin{aligned} \prod_{j=2}^{n} (1 - \lambda_j) &\leqslant \left[\frac{\sum_{j=2}^{n} (1 - \lambda_j)}{n - 1}\right]^{n-1} \\ &= \left[1 - \frac{(\operatorname{tr} \mathbf{P} - 1)}{n - 1}\right]^{n-1} \leqslant \left(1 + \frac{1}{n - 1}\right)^{n-1}, \end{aligned} \tag{6.7}$$

which converges to e as $n \to \infty$ and for quite small n is close to e. In (6.7) we have used the arithmetic/geometric mean inequality and $\operatorname{tr} \mathbf{P} \geqslant 0$. An open problem is to find a bound for $\prod_{j=2}^{n} (1 - \lambda_j)$ when exactly c pairs of λ_j's are complex. For $c = 0$ the solution is given by (6.7) and for $c = [(n-1)/2]$ the bound is n. We expect the upper bound to be monotonically increasing in c, and like (6.4) and (6.7) to require $p_{11} = \cdots = p_{nn} = 0$.

Equality is attained in (6.7) when $\lambda_2 = \cdots = \lambda_n = -1/(n-1)$. This does not, however, determine $\mathbf{P}$ uniquely. Let p be the number of linearly independent characteristic vectors of $\mathbf{P}$ corresponding to the multiple root $-1/(n-1)$. Then $1 \leqslant p \leqslant n-1$. If $p = n-1$, $\mathbf{P} + \mathbf{I}/(n-1)$ has rank one and we may write $\mathbf{P} = (\mathbf{uv}' - \mathbf{I})/(n-1)$ where $\mathbf{u}$ and $\mathbf{v}$ are $n \times 1$ vectors. Since $\mathbf{P1} = \mathbf{1}$, $\mathbf{u} = \mathbf{1}$ and as $p_{11} = p_{22} = \cdots = p_{nn} = 0$, $\mathbf{v} = \mathbf{1}$ and so

$$\mathbf{P} = (\mathbf{11}' - \mathbf{I})/(n-1) = \frac{1}{n-1}\begin{bmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ 1 & 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 0 \end{bmatrix}. \tag{6.8}$$

But if $p < n-1$, $\mathbf{P}$ does not need to be of this form; for example, let $n = 3$ and $p = 1$. Then

$$\mathbf{P} = \tfrac{1}{2}\begin{bmatrix} 0 & 2 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \tag{6.9}$$

has roots, $1, -\frac{1}{2}, -\frac{1}{2}$; but (6.8) with $n = 3$ gives

$$\mathbf{P} = \tfrac{1}{2}\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}. \tag{6.10}$$

THEOREM 6.2. *Let the $n \times n$ stochastic matrix $\mathbf{P}$ have characteristic roots $\lambda_1 = 1, \lambda_2, \ldots, \lambda_n$. Then if the root $\lambda_1 = 1$ is simple*

$$\sum_{j=2}^{n} \frac{1}{1-\lambda_j} \geqslant \frac{1}{2}(n-1), \tag{6.11}$$

with equality if and only if $\mathbf{P}$ is an irreducible permutation matrix ($\lambda_2, \lambda_3, \ldots, \lambda_n$ are then the $n-1$ nonunit nth roots of unity).

Proof. Let $\mathbf{P}_{ij}$ be the $(n-2) \times (n-2)$ principal submatrix of $\mathbf{P}$ after deletion of its ith and jth row and column ($i \neq j$). Then

$$\sum_{j=2}^{n} \frac{1}{1-\lambda_j} = \frac{\sum_{i=2}^{n} \prod_{j=2, j\neq i}^{n} (1-\lambda_j)}{\prod_{j=2}^{n} (1-\lambda_j)} = \frac{\frac{1}{2}\sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \det(\mathbf{I} - \mathbf{P}_{ij})}{\sum_{i=1}^{n} \det(\mathbf{I} - \mathbf{P}_i)}, \tag{6.12}$$

where $\mathbf{P}_i$ is defined as in (6.3). The numerators on the right of (6.12) each equal the coefficient of $(-1)^n\lambda^2$ in the expansion of the characteristic polynomial $\det(\lambda\mathbf{I} - [\mathbf{I} - \mathbf{P}])$ {cf. e.g., Mirsky [7, p. 198]}. Suppose $r_i < 1$, $i = 1, 2,..., n$. Then from (3.1) and (4.1),

$$\sum_{\substack{j=1\\j\neq i}}^{n} \det(\mathbf{I} - \mathbf{P}_{ij}) = \det(\mathbf{I} - \mathbf{P}_i)\,\mathrm{tr}(\mathbf{I} - \mathbf{P}_i)^{-1}$$
$$\geqslant \frac{(n-1)\det(\mathbf{I} - \mathbf{P}_i)}{1 - r_i^{n-1}} \geqslant (n-1)\det(\mathbf{I} - \mathbf{P}_i). \tag{6.13}$$

Substitution in (6.12) yields (6.11) directly. If some r_i are equal to 1 then the above proof goes through similarly but with the summations on i in (6.12) excluding the unit r_i's. Not all $r_i = 1$ as then

$$\sum_{i=1}^{n} \det(\mathbf{I} - \mathbf{P}_i) = \prod_{j=2}^{n} (1 - \lambda_j) = 0$$

and at least one $\lambda_j = 1$; $j = 2,..., n$. Equality in (6.11) holds if and only if equality holds throughout (6.13) and so $r_1 = r_2 = \cdots = r_n = 0$. Thus (6.6) holds and $\mathbf{P}$ must be an irreducible permutation matrix by the same reasoning as used in Theorem 6.1. (Q.E.D.)

For an irreducible Markov chain, Kemeny and Snell [4] study the *fundamental matrix*

$$\mathbf{Z} = (\mathbf{I} - \mathbf{P} + \mathbf{1}\mathbf{l}')^{-1}, \tag{6.14}$$

where $\mathbf{l}$ is defined as in (6.2). Since $\mathbf{P}$ is irreducible, $\mathbf{I} - \mathbf{P}$ has a simple characteristic root of 0 and since $\mathbf{1}$ is the corresponding characteristic vector it follows that the characteristic roots of $\mathbf{Z}^{-1}$ are $1, 1 - \lambda_2 ,..., 1 - \lambda_n$ where $\lambda_1 = 1, \lambda_2 ,..., \lambda_n$ are the roots of $\mathbf{P}$ as before. Thus

$$\det \mathbf{Z} = \prod_{j=2}^{n} (1 - \lambda_j)^{-1}; \qquad \mathrm{tr}\, \mathbf{Z} = 1 + \sum_{j=2}^{n} (1 - \lambda_j)^{-1}. \tag{6.15}$$

Theorems 6.1 and 6.2 establish the following.

COROLLARY 6.1. *Let the n-state irreducible Markov chain governed by the stochastic matrix* $\mathbf{P}$ *have fundamental matrix* $\mathbf{Z}$ *as in* (6.14). *Then*

$$\det \mathbf{Z} \geqslant 1/n; \qquad \mathrm{tr}\, \mathbf{Z} \geqslant \tfrac{1}{2}(n+1). \tag{6.16}$$

Equality holds in either part of (6.16) *if and only if the Markov chain is cyclic with period n.*

Kemeny and Snell [4, p. 81] study $\operatorname{tr}\mathbf{Z}$ but give no bounds; they show, however, that $\mathbf{M}l = (\operatorname{tr}\mathbf{Z})\,\mathbf{1}$, where $\mathbf{M} = \{m_{ij}\}$ and m_{ij} is the mean first passage time from state i to state j.

We conclude with bounds for the coefficients and for ratios of the coefficients in the characteristic polynomial of $\mathbf{I} - \mathbf{P}$, where $\mathbf{P}$ is a stochastic matrix. If

$$\begin{aligned}\det[\lambda\mathbf{I} - (\mathbf{I} - \mathbf{P})] &= \lambda^n - a_1\lambda^{n-1} + \cdots + (-1)^{n-1} a_{n-1}\lambda \\ &= \sum_{t=0}^{n-1} (-1)^t a_t\lambda^{n-t},\end{aligned} \tag{6.17}$$

then the coefficient of $(-1)^t \lambda^{n-t}$ is {cf. Mirsky [7, pp. 197–8]}

$$\begin{aligned}a_t &= \sum\sum_{1\leqslant i_1<\cdots<i_{n-t}\leqslant n}\cdots\sum \det(\mathbf{I} - \mathbf{P}_{i_1\cdots i_{n-t}}) \\ &= \sum\sum_{2\leqslant j_1<\cdots<j_t\leqslant n}\cdots\sum (1-\lambda_{j_1})\cdots(1-\lambda_{j_t}); \qquad t = 1,\ldots, n-1,\end{aligned} \tag{6.18}$$

where $\mathbf{P}_{i_1\cdots i_{n-t}}$ is the $t \times t$ principal submatrix of $\mathbf{P}$ after deletion of its i_1th,..., i_{n-t}th rows and columns. In (6.17) $a_0 = 1$ and the constant term (a_n) is zero as $\mathbf{I} - \mathbf{P}$ is singular. We may also write

$$\det(\lambda\mathbf{I} - \mathbf{P}) = \sum_{q=0}^{n} (-1)^q b_q\lambda^{n-q} = \sum_{s=1}^{n} a_s^*(\lambda - 1)^s, \tag{6.19}$$

where $a_s^* = a_{n-s}$. To see this, we have from (6.17) that

$$\begin{aligned}\det[\lambda\mathbf{I} - (\mathbf{I} - \mathbf{P})] &= \sum_{t=0}^{n-1} (-1)^t a_t\lambda^{n-t} \\ &= (-1)^n \det[(1-\lambda)\,\mathbf{I} - \mathbf{P}].\end{aligned} \tag{6.20}$$

Hence

$$\det[(1-\lambda)\,\mathbf{I} - \mathbf{P}] = \sum_{t=0}^{n-1} a_t(-\lambda)^{n-t} \tag{6.21}$$

and (6.19) follows upon replacing $1 - \lambda$ by λ and t by $n - s$. A generalization of Theorem 6.1 is as follows.

THEOREM 6.3. *Let $a_s^* = a_{n-s}$ be defined as in* (6.17) *and* (6.19). *Then*

$$0 \leqslant a_s^* = a_{n-s} \leqslant \binom{n}{s}; \qquad s = 1, 2,\ldots, n-1. \tag{6.22}$$

Equality on the left holds for $s = n - 1$ if and only if $\mathbf{P} = \mathbf{I}$; if equality holds on the right for some s then

$$b_1 = b_2 = \cdots = b_{n-s} = 0, \tag{6.23}$$

where b_q is as defined in (6.19).

Proof. Following (6.5), we have using (6.18) that

$$\begin{aligned} a_t &= \sum\sum_{1 \leqslant i_1 < \cdots < i_{n-t} \leqslant n} \cdots \sum \det(\mathbf{I} - \mathbf{P}_{i_1 \cdots i_{n-t}}) \\ &\leqslant \sum\sum_{1 \leqslant i_1 < \cdots < i_{n-t} \leqslant n} \cdots \sum (1 - r_0^t) \leqslant \binom{n}{t}; \qquad t = 1, 2, \ldots, n-1, \end{aligned} \tag{6.24}$$

where r_0 is the dominant root of $\mathbf{P}_{i_1 \cdots i_{n-t}}$. Setting $t = n - s$ gives (6.22). Equality on the left of (6.22) holds for $s = n - 1$ if and only if all 1×1 principal minors of $\mathbf{I} - \mathbf{P}$ are 0 [cf. (6.18)]. Thus $\mathbf{P} = \mathbf{I}$. The converse is evident. If equality on the right of (6.22) holds for some $s = s_0$ then all principal minors of $\mathbf{P}$ of order $q \leqslant n - s_0$ are zero and (6.23) is true. (Q.E.D.)

When $a_{n-1} = a_1^* = n$ it follows from Theorem 6.1 that

$$\det(\lambda \mathbf{I} - \mathbf{P}) = \lambda^n - 1$$

and hence from (6.19) that

$$a_{n-t} = a_t^* = \binom{n}{t} \qquad \text{for all } t = 1, 2, \ldots, n-1.$$

Also for $s = 1$ (6.23) is a necessary and sufficient condition for equality on the right of (6.22). A generalization of Theorem 6.2 is as follows.

THEOREM 6.4. *Let the $n \times n$ stochastic matirx $\mathbf{P}$ have a simple characteristic root of unity and let $a_s^* = a_{n-s}$ be defined as in* (6.17) *and* (6.19). *Then*

$$a_{s+1}^*/a_s^* = [a_{n-(s+1)}]/a_{n-s} \geqslant (n-s)/(s+1); \qquad s = 1, 2, \ldots, n-2 \tag{6.25}$$

with equality for some s if and only if (6.23) *is true.*

Proof. Since $\mathbf{P}$ has a simple root of unity,

$$a_s^* = a_{n-s} > 0, \qquad s = 1, 2, \ldots, n-1.$$

From (6.18), following (6.12),

$$\frac{a_{n-(s+1)}}{a_{n-s}} = \frac{\sum\sum\cdots\sum_{1\leqslant i_1<\cdots<i_{s+1}\leqslant n} \det(\mathbf{I}-\mathbf{P}_{i_1i_2\cdots i_{s+1}})}{\sum\sum\cdots\sum_{1\leqslant i_1<\cdots<i_s\leqslant n} \det(\mathbf{I}-\mathbf{P}_{i_1i_2\cdots i_s})}$$
$$= \frac{\sum\sum\cdots\sum_{1\leqslant i_1<\cdots<i_s\leqslant n} \sum\sum\cdots\sum_{i_{s+1}\neq i_1,\ldots,i_s} \det(\mathbf{I}-\mathbf{P}_{i_1i_2\cdots i_{s+1}})}{(s+1)\sum\sum\cdots\sum_{1\leqslant i_1<\cdots<i_s\leqslant n} \det(\mathbf{I}-\mathbf{P}_{i_1i_2\cdots i_s})} \geqslant \frac{n-s}{s+1}, \tag{6.26}$$

since form (3.1) and (4.1), following (6.13),

$$\sum\sum\cdots\sum_{i_{s+1}\neq i_1,\ldots,i_s} \det(\mathbf{I}-\mathbf{P}_{i_1i_2\cdots i_{s+1}}) \geqslant \{[\det(\mathbf{I}-\mathbf{P}_{i_1i_2\cdots i_s})]/(1-r_0^{n-s})\}\,(n-s), \tag{6.27}$$

where r_0 is the dominant root of $\mathbf{P}_{i_1i_2\cdots i_s}$. Equality holds in (6.25) if and only if equality holds in (6.26) and so $r_0 = 0$; hence (6.23) holds by reasoning similar to that used in Theorem 6.3. (Q.E.D.)

From (6.25) it follows that

$$a_{s+t}^*/a_s^* = a_{n-(s+t)}/a_{n-s} \geqslant \binom{n}{s+t}\Big/\binom{n}{s};$$
$$t = 1, 2,\ldots, n-s-1, \qquad s = 1, 2,\ldots, n-2 \tag{6.28}$$

with equality for some s if and only if (6.23) is true.

Addendum

Robert Vermes has proved (November 1971) the main results in Theorems 1 and 2 by induction. Let $f(s) = \det \mathbf{M}_s$, where $\mathbf{M}_s = s\mathbf{I} - \mathbf{A}$ as in (1.2); let $\mathbf{A}_i$ be the $(n-1)\times(n-1)$ principal submatrix of $\mathbf{A}$ with row and column i removed and let r_i be the dominant characteristic root of $\mathbf{A}_i$.

The right side of (1.5) holds for $n = 1$. Suppose it is valid for $(n-1)\times(n-1)$ matrices. Then for all $s > r$,

$$f'(s) = d(\det \mathbf{M}_s)/ds$$
$$= \sum_{i=1}^{n} \det(s\mathbf{I}-\mathbf{A}_i) \leqslant \sum_{i=1}^{n} (s^{n-1} - r_i^{n-1}) \leqslant ns^{n-1}, \tag{A.1}$$

as $s\mathbf{I} - \mathbf{A}_i$ is an M-matrix and $r \geqslant r_i \geqslant 0$. Therefore, since $f(r) = 0$,

$$\det \mathbf{M}_s = f(s) = f(s) - f(r) = \int_r^s f'(t)\,dt \leqslant \int_r^s nt^{n-1}\,dt = s^n - r^n, \qquad \text{(A.2)}$$

and the right side of (1.5) holds for $n \times n$ matrices.

The left side of (1.1) may be established using the result that

$$g(s) = \det \mathbf{M}_s/s^n = f(s)/s^n \qquad \text{(A.3)}$$

is nondecreasing in s for all $s > r$. Since $dg(s)/ds = f'(s)/s^n - nf(s)/s^{n+1}$, $g(s)$ is nondecreasing provided $\operatorname{tr} \mathbf{M}_s^{-1} = f'(s)/f(s) \geqslant n/s$. This, however, follows directly from (5.7), or by induction from (A.6) below. Since $r \geqslant r_i$, $s\mathbf{I} - \mathbf{A}_i$ is an M-matrix and

$$\det(s\mathbf{I} - \mathbf{A}_i)/s^{n-1} \geqslant \det(t\mathbf{I} - \mathbf{A}_i)/t^{n-1} \qquad \text{(A.4)}$$

for any t such that $s > t > r$. By the generalized mean value theorem we may choose t such that

$$f(s)/(s^n - r^n) = [f(s) - f(r)]/(s^n - r^n) = f'(t)/(nt^{n-1}), \qquad \text{(A.5)}$$

as $f(r) = 0$. Hence, using (A.4) and (A.5),

$$\begin{aligned} \operatorname{tr} \mathbf{M}_s^{-1} &= \frac{f'(s)}{f(s)} = \frac{\sum_{i=1}^n \det(s\mathbf{I} - \mathbf{A}_i)}{f(s)} \geqslant \frac{s^{n-1} \sum_{i=1}^n \det(t\mathbf{I} - \mathbf{A}_i)}{t^{n-1} f(s)} \\ &= \frac{s^{n-1} f'(t)}{t^{n-1} f(s)} = \frac{ns^{n-1}}{s^n - r^n}. \end{aligned} \qquad \text{(A.6)}$$

Acknowledgments

The authors are grateful to Arthur Albert, Herman Chernoff, Teun Ott, and Robert Vermes for their very generous assistance and helpful discussions concerning this research.

References

1. William Feller, "An Introduction to Probability Theory and its Applications," Vol. II, Wiley, New York, 1966.
2. F. R. Gantmacher, "Applications of the Theory of Matrices," Interscience, New York, 1959.
3. Samuel Karlin, "A First Course in Stochastic Processes," Academic Press, New York, 1966.
4. John G. Kemeny and J. Laurie Snell, "Finite Markov Chains," Van Nostrand, Princeton, NJ, 1960.

5. A. N. Kolmogorov, Lectures on Markov Chains, Presented at a meeting of the Moscow Mathematical Society, 1938.
6. Marvin Marcus and Henryk Minc, "A Survey of Matrix Theory and Matrix Inequalities," Allyn and Bacon, Boston, MA, 1964. (Reprinted with corrections: Prindle, Weber & Schmidt, Boston, MA, 1969.)
7. L. Mirsky, "An Introduction to Linear Algebra," Oxford University Press, Oxford, England, 1955.
8. L. Mirsky, Results and problems in the theory of doubly-stochastic matrices, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 1 (1963), 319–334.
9. L. Mirsky, Inequalities and existence theorems in the theory of matrices, *J. Math. Anal. Appl.* **9** (1964), 99–118.
10. Alexander Ostrowski, Über die Determinanten mit überwiegender Hauptdiagonale, *Comment. Math. Helv.* **10** (1937), 69–96.
11. Richard S. Varga, "Matrix Iterative Analysis," Prentice–Hall, Englewood Cliffs, NJ, 1962.