NONNEGATIVE MATRICES WITH STOCHASTIC POWERS*

BY

DAVID LONDON

ABSTRACT

Matrices with nonnegative elements, which are nonstochastic but have stochastic powers, are considered. These matrices are characterized in the irreducible case and in the symmetric one.

1. Introduction. In this paper we consider square matrices with nonnegative elements which themselves are not stochastic, but for which a certain power is stochastic. In §2 we deal with nonnegative irreducible matrices, and in §3 with nonnegative symmetric matrices. In each of these cases we obtain a characterization of the nonstochastic matrices of the corresponding class which have stochastic powers. Our characterizations are constructive and enable us to build effectively the corresponding matrices. A very special case of our second result, the characterization of all 3×3 nonnegative symmetric matrices A which are nonstochastic, but for which A^2 is stochastic, was obtained earlier as a byproduct of the proof of a certain matrix inequality [2, Remark 3 following Theorem 1].

The main tool used in this paper is the Perron-Frobenius theorem [1, p. 53]. Let $A = (a_{ij})$ be a $n \times n$ nonnegative irreducible matrix. By the Perron-Frobenius theorem, A has a dominant simple positive characteristic value $\alpha = \alpha(A)$. If $\alpha_1, \cdots, \alpha_h = \alpha$ are all the characteristic values of A with modulus α, then $\alpha_k = \alpha\omega^k$, $k = 1, \cdots, h$, where $\omega = e^{2\pi i/h}$. If $h = 1$ A is *primitive*. If $h > 1$ A is *cyclic of index h*. If A is cyclic of index h, then there exists a permutation matrix P such that

$$(1.1) \qquad PAP^T = \begin{bmatrix} 0 & A_1 & 0 & - & - & - & 0 \\ 0 & 0 & A_2 & 0 & - & - & 0 \\ \vdots & \vdots & & \ddots & & & \vdots \\ \vdots & \vdots & - & \ddots & - & - & \vdots \\ \vdots & \vdots & & & \ddots & & \vdots \\ 0 & 0 & - & - & - & 0 & A_{h-1} \\ A_h & 0 & - & - & - & - & 0 \end{bmatrix}.$$

Received December 15, 1964.

* This paper represents part of a thesis submitted to the Senate of the Technion–Israel Institute of Technology in partial fulfillment of the requirements for the degree of Doctor of Science. The author wishes to thank Professor B. Schwarz for his guidance in the preparation of this paper.

The null matrices in the main diagonal are squares of orders n_k, $k = 1, \cdots, h$. (1.1) is *the Frobenius normal form of A*. Let $r_1, \cdots, r_h$ be the characteristic values of PAP^T corresponding respectively to $\alpha_1, \cdots, \alpha_h$. r_h is positive ($r_h > 0$). Write

$$r_h = z_1 \dotplus \cdots \dotplus z_h,$$

where z_k is a vector of order n_k, and the symbol $\dotplus$ indicates direct sum. (If $u = (u_1, \cdots u_m)$ and $v = (v_1, \cdots, v_n)$, then $u \dotplus v = (u_1, \cdots, u_m, v_1 \cdots, v_n)$). We have

$$r_k = z_1 \dotplus \omega^k z_2 \dotplus \omega^{2k} z_3 \dotplus \cdots \dotplus \omega^{(h-1)k} z_h, \quad k = 1, \cdots, h.$$

We end this introduction by a definition. Let $B = (b_{ij})$ be a nonnegative $m \times n$ matrix. If

$$\sum_{j=1}^{n} b_{ij} = \beta, \qquad i = 1, \cdots, m,$$

then B is *β stochastic* or *generalized stochastic*. If $\beta = 1$ then B is *stochastic*. We remark that usually this definition is given only for square matrices. However, for our purpose it is convenient to use it for rectangular matrices.

2. **Nonnegative irreducible matrices.** Let A be a nonnegative irreducible matrix which is not stochastic. In this section we obtain a necessary and sufficient condition for some power of A to be stochastic.

THEOREM 1. *Let A be a nonnegative irreducible square matrix and let $m > 1$ be a positive integer. Let H be the cyclic permutation*

$$H = (1\,2, \cdots h),$$

and let

$$H^m = C_1 C_2 \cdots C_r \tag{2.1}$$

be the representation of H^m as the product of disjoint cycles. A is not a stochastic matrix while A^m is stochastic if and only if

(I) *A is cyclic of index h, where $(h, m) > 1$.*

(II) *There exist positive numbers β_i, $i = 1, \cdots, h$, such that the matrices A_i appearing in the Frobenius normal form* (1.1) *of A are respectively β_i/β_{i+1} stochastic.*(*) *The numbers β_i fulfill the following two conditions*:

(A) *They are not all equal.*

(B) *Every two numbers with indices belonging to the same cycle in* (2.1) *are equal.*

Proof. First we prove that the conditions (I) and (II) are necessary. Let A be a nonnegative irreducible matrix which is not stochastic while A^m is stochastic.

* Here and in the sequel the indices are taken modulo h.

As A^m is stochastic, 1 is the dominant characteristic value of A^m and $e=(1,\cdots,1)$ is a corresponding characteristic vecor. Returning to A, it follows that 1 is the dominant characteristic value of A. As A is not stochastic, e is not a characteristic vector of A. Assume A is primitive. Then 1 is a simple characteristic value of A^m and the only characteristic vector of A^m corresponding to the characteristic value 1 is the characteristic vector of A corresponding to 1. But as this vector is different from e, it follows that A cannot be primitive. Hence, A is cyclic and can be represented by the Frobenius normal form (1.1). As PAP^T is only a cogredient permutation of the rows and columus of A, we may change in the above considerations A and A^m respectively with PAP^T and PA^mP^T.

Let $\alpha_1,\cdots,\alpha_h = 1$ be all the characteristic values of PAP^T(or of A) with modulus 1, and let $r_1,\cdots,r_h$ be the corresponding characteristic vectors. As quoted in §1 we have

$$\alpha_k = \omega^k, \quad \omega = e^{2\pi i/h}, \qquad k = 1,\cdots,h, \tag{2.2}$$

$$r_k = z_1 + \omega^k z_2 + \omega^{2k} z_3 + \cdots + \omega^{(h-1)k} z_h, \qquad k = 1,\cdots,h. \tag{2.3}$$

As e is a characteristic vector of PA^mP^Tcorresponding to 1 while it is not a characteristic vector of PAP^T, there exist integers $k_1,\cdots,k_l$; $1 \leqq k_1 < k_2 \cdots < k_l \leqq h$, $l > 1$, such that

$$\omega^{mk_1} = \omega^{mk_2} = \cdots = \omega^{mk_l} = 1, \tag{2.4}$$

and also numbers $d_1,\cdots,d_l$ such that

$$d_1 r_{k_1} + d_2 r_{k_2} + \cdots + d_l r_{k_l} = e. \tag{2.5}$$

(2.3) and (2.5) imply

$$z_1(d_1 + \cdots + d_l) + \cdots + z_h(d_1\omega^{(h-1)k_1} + \cdots + d_l\omega^{(h-1)k_l}) = e. \tag{2.6}$$

Let $e_i = (1,\cdots,1)$, $i = 1,\cdots,h$, be a vector of order n_i. From (2.3), (2.6) and the fact that $r_h > 0$ it follows that there exist positive numbers $\beta_1,\cdots,\beta_h$ such that

$$r_h = \beta_1 e_1 + \beta_2 e_2 + \cdots + \beta_h e_h. \tag{2.7}$$

As r_h is a chracteristic vector of PAP^T corresponding to the characteristic value 1, we obtain from (1.1) and (2.7)

$$PAP^T r_h = \beta_2 A_1 e_2 + \cdots + \beta_h A_{h-1} e_h + \beta_1 A_h e_1 = \beta_1 e_1 + \cdots + \beta_h e_h.$$

Hence,

$$A_i e_{i+1} = \frac{\beta_i}{\beta_{i+1}} e_i, \qquad i = 1,\cdots,h. \tag{2.8}$$

From (2.8) follows that A_i is a β_i/β_{i+1} stochastic matrix.

We have now to show that β_i fulfill the conditions (A) and (B). PAP^T is not stochastic and therefore not all the matrices A_i are stochastic. As A_i is β_i/β_{i+1} stochastic, it follows that not all the numbers β_i are equal. (A) is thus proved. To prove (B) denote the blocks of PAP^T in the partitioning (1.1) by A_{ij}, $i,j = 1,\cdots,h$, and the blocks of PA^mP^T in the same partitioning by $A_{i,j}^{(m)}$. We have

$$A_{ij} = \begin{cases} A_i, & j \equiv i+1 \pmod{h} \\ 0, & j \not\equiv i+1 \pmod{h} \end{cases} \tag{2.9}$$

and

$$A_{ij}^{(m)} = \sum_{k_1,\ldots,k_{m-1}=1}^{h} A_{ik_1}A_{k_1k_2}\cdot\cdots\cdot A_{k_{m-1},j}. \tag{2.10}$$

From (2.9) and (2.10) follows

$$A_{ij}^{(m)} = \begin{cases} A_iA_{i+1}\cdot\cdots\cdot A_{i+m-1}, & j \equiv i+m \pmod{h} \\ 0, & j \not\equiv i+m \pmod{h}. \end{cases} \tag{2.11}$$

As PA^mP^T is stochastic, all the matrices $A_{i,i+m}^{(m)}$ are stochastic. As A_i is β_i/β_{i+1} stochastic, it follows from (2.11) that $A_{i,i+m}^{(m)}$ is

$$\frac{\beta_i}{\beta_{i+1}}\cdot\frac{\beta_{i+1}}{\beta_{i+2}}\cdot\ \cdots\ \cdot\frac{\beta_{i+m-1}}{\beta_{i+m}} = \frac{\beta_i}{\beta_{i+m}} \text{ stochastic.}$$

Hence,

$$\beta_i = \beta_{i+m}, \qquad i = 1,\cdots,h. \tag{2.12}$$

The permutation H^m carries i into $i+m$ and therefore i and $i+m$ belong to the same cycle in (2.1), and so (2.12) is equivalent to (B). (B) is thus proved, and the proof of (II) is established.

We have already proved that A is cyclic of index h. To complete the proof of (I), we have to show that $(h,m) > 1$. This fact follows easily from (2.2) and (2.4). The proof of the necessity part of the theorem is completed.

We now prove that the conditions (I) and (II) are sufficient. Let A be a matrix which fulfills the conditions (I) and (II). From (A) follows that PAP^T, and therefore A too, is not stochastic. According to (2.11) the matrix $A_{i,i+m}^{(m)}$ is β_i/β_{i+m} stochastic. (B) is equivalent to $\beta_i = \beta_{i+m}$, and so $A_{i,i+m}^{(m)}$ is stochastic. Hence, PA^mP^T, and therefore A^m is stochastic. The proof of Theorem 1 is thus completed.

REMARK 1. We have

$$H^h = (1)(2)\cdots(h),$$

and so for $m = h$ the condition (B) holds for any β_i. In this case it is thus sufficient that the condition (A) holds. From this we conclude that if A is a nonstochastic cyclic matrix of index h and if A^m is stochastic, then A^h is also stochastic and so any power A^{m1}, where $m_1 \equiv m \pmod{h}$.

REMARK 2. Let m and h be positive integers $(m,h) > 1$. By the sufficient conditions of Theorem 1 we can construct all the matrices A which are non-stochastic and cyclic of order h, and for which A^m is stochastic. As $(m,h) > 1$, there is more than one cycle in the representation (2.1), and so we can find positive numbers β_i, $i = 1,\cdots,h$, for which both (A) and (B) hold. Let A_i, $i = 1,\cdots,h$, be β_i/β_{i+1} stochastic matrices chosen so that their dimensions fit the structure of (1.1) and so that A is cyclic of index h. Using the A_j's, we construct A according to (1.1).

3. Nonnegative symmetric matrices. Let A be a nonnegative symmetric matrix which is not stochastic. In this section we obtain a necessary and sufficient condition for some powers of A to be stochastic.

Let us first define a class of matrices $\mathfrak{A}_n$. *A matrix A belongs to the class $\mathfrak{A}_n$ if and only if A is a $n \times n$ nonnegative symmetric matrix, A is not stochastic and there exists a natural number m for which A^m is stochastic.*

Let $A \in \mathfrak{A}_n$. A^m is thus stochastic while A is not stochastic. It is necessary that the multiplicity of the dominant characteristic value of A^m is greater than the multiplicity of the dominant characteristic value of A. Hence, m is even. As the multiplicity of the dominant characteristic value of A^m is equal for all the even m's, it follows that if $A \in \mathfrak{A}_n$ then A^m is stochastic if and only if m is even.

In the following theorem we characterize the classes $\mathfrak{A}_n$ by a recursive procedure. The structure of the class $\mathfrak{A}_n$ is determined by the structure of the classes $\mathfrak{A}_m$, $m < n$. As the class $\mathfrak{A}_1$ is void, we can by this procedure determine the structure of $\mathfrak{A}_n$ for any n.

THEOREM 2. *Let A be a $n \times n$ matrix.*

(1) *If A is reducible, then $A \in \mathfrak{A}_n$ if and only if there exists a permutation matrix P such that*

$$A = P^T \begin{bmatrix} B_k & 0 \\ 0 & B_{n-k} \end{bmatrix} P. \tag{3.1}$$

k is an integer for which the inequality

$$\frac{n}{2} \leqq k < n \tag{3.2}$$

holds. B_k and B_{n-k} are respectively $k \times k$ and $(n-k)\times(n-k)$ matrices, and at least one of the following two conditions

$$B_k \in \mathfrak{A}_k\,; \quad B_{n-k} \in \mathfrak{A}_{n-k} \tag{3.3}$$

holds. If only one of these conditions holds, then the matrix for which the condition does not hold is symmetric and stochastic.

(2) *If A is irreducible, then $A \in \mathfrak{A}_n$ if and only if there exists a permutation matrix P such that*

$$A = P^T \begin{bmatrix} 0 & A_1 \\ A_1^T & 0 \end{bmatrix} P. \tag{3.4}$$

0 indicates square null matrices. k is an integer for which the inequality

$$\frac{n}{2} < k < n \tag{3.5}$$

holds. A_1 is a $k \times (n-k)$ matrix $[(n-k)/k]^{1/2}$ stochastic and its transposed A_1^T is $[k/(n-k)]^{1/2}$ stochastic.

Proof of (1). First we prove the necessity part. Let A be a reducible matrix belonging to $\mathfrak{A}_n$. As A is reducible and symmetric, there exists a permutation matrix P for which (3.1) holds, where B_k and B_{n-k} are symmetric matrices. It is obvious that P can be chosen so that (3.2) holds. We have

$$A^2 = P^T \begin{bmatrix} B_k^2 & 0 \\ 0 & B_{n-k}^2 \end{bmatrix} P.$$

As $A \in \mathfrak{A}_n$, A^2 is stochastic and therefore B_k^2 and B_{n-k}^2 are both stochastic. As A is nonstochastic, at least one of the two matrices B_k and B_{n-k} is nonstochastic. For the matrix which is nonstochastic the corresponding condition in (3.3) holds. If the other matrix is also nonstochastic, then (3.3) holds for this matrix too. If the other matrix is stochastic, then it is symmetric and stochastic.

It is easy to verify that the conditions are also sufficient.

Proof of (2). Let us begin with the necessity part. Let A be an irreducible matrix belonging to $\mathfrak{A}_n$. According to Theorem 1, A is cyclic. As A is symmetric, it is cyclic of index 2 and so there exists a permutation matrix P for which (3.4) holds. It is obvious that P can be chosen so that (3.2) holds. According to Theorem 1 there exist positive numbers β_1 and β_2, $\beta_1 \neq \beta_2$, such that A_1 is β_1/β_2 stochastic and A_1^T is β_2/β_1 stochastic. As A_1 is a $k \times (n-k)$ matrix, we obtain

$$k \frac{\beta_1}{\beta_2} = (n-k) \frac{\beta_2}{\beta_1}.$$

Hence,

$$\frac{\beta_1}{\beta_2} = \left(\frac{n-k}{k}\right)^{1/2}.$$

A_1 is thus $[(n-k)/k]^{1/2}$ stochastic and A_1^T is $[k/(n-k)]^{1/2}$stochastic. As $\beta_1 \neq \beta_2$, it follows that the sign of equality in the lefthand side of (3.2) does not hold, and so (3.5) holds.

The sufficiency part follows by direct computation of PA^2P^T. The proof of Theorem 2 is completed.

We shall now discuss the structure of the classes $\mathfrak{A}_n$ for n up to 4.

$n = 1.$

As already mentioned $\mathfrak{A}_1$ is void.

$n = 2.$

(1) A reducible. (3.2) implies $k = 1$, $n - k = 1$. As $\mathfrak{A}_1$ is void, the condition (3.3) cannot be fulfilled, and so there are no reducible matrices in $\mathfrak{A}_2$.

(2) A irreducible. No natural k exists for which (3.5) holds, and so there are no irreducible matrices in $\mathfrak{A}_2$.

Conclusion: $\mathfrak{A}_2$ is void.

$n = 3.$

(1) A reducible. (3.2) implies $k = 2$, $n - k = 1$. As the classes $\mathfrak{A}_1$ and $\mathfrak{A}_2$ are void, the condition (3.3) can not be fulfilled, and therefore there are no reducible matrices in $\mathfrak{A}_3$.

(2) A irreducible. (3.5) implies $k = 2$, $n - k = 1$. A_1 is a 2×1, $1/\sqrt{2}$ stochastic matrix, and A has the following form

$$A = P^T \begin{bmatrix} 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} P. \tag{3.6}$$

There exists 3 distinct matrices of the the form (3.6).

Conclusion: $\mathfrak{A}_3$ includes precisely the three matrices given by (3.6).

From this conclusion follows the result mentioned in the introduction.

$n = 4.$

(1) A reducible. (3.2) implies $k = 2; 3$ and so $n - k = 2; 1$ respectively. As $\mathfrak{A}_2$ is void, there remains only the possibility $k = 3$, $n - k = 1$. Let B_3 be one of the three matrices belonging to $\mathfrak{A}_3$. A has the following form

$$A = P^T \begin{bmatrix} & B_3 & & 0 \\ & & & 0 \\ & & & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} P. \tag{3.7}$$

There are 12 distinct matrices of the form (3.7).

(2) A irreducible. (3.5) implies $k=3$, $n-k=1$. A_1 is a 3×1 matrix, $1/\sqrt{3}$ stochastic, and A has the following form

$$A = P^T \begin{bmatrix} 0 & 0 & 0 & \frac{1}{\sqrt{3}} \\ 0 & 0 & 0 & \frac{1}{\sqrt{3}} \\ 0 & 0 & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 0 \end{bmatrix} P. \tag{3.8}$$

There are 4 distinct matrices of the form (3.8).

Conclusion: $\mathfrak{A}_4$ includes 12 reducible matrices given by (3.7) and 4 irreducible matrices given by (3.8).

References

1. F. R. Gantmacher, *The Theory of Matrices*, Vol. 2, Chelsea, New-York, 1959.
2. D. London, *Two inequalities in nonnegative symmetric matrices*, (to appear in the Pacific J. Math.)

Technion-Israel Institute of Technology,
Haifa, Israel