ON SETS OF TUPLES OF CHARACTERISTIC NUMBERS OF STOCHASTIC MATRICES

MATHEMATICS

1968

SovietRxiv

Abstract

Full Text

UDC 519.21

MATHEMATICS

R. G. BUKHARAEV

ON SETS OF TUPLES OF CHARACTERISTIC NUMBERS OF STOCHASTIC MATRICES

(Presented by Academician V. M. Glushkov, 13 III 1967)

The study of the distribution in the number plane of the characteristic numbers of stochastic matrices was the subject of works by N. A. Dmitriev and E. V. Dynkin (^{1,2}) and by F. I. Karpelevich (^{3}). Later, the clarification of conditions under which a given tuple of numbers can serve as a tuple of characteristic numbers of a stochastic matrix was studied by Kh. R. Suleimanova (^{4}), N. Perfect (^{5}). Necessary and sufficient conditions have not yet been obtained even in the simplest case, when the stochastic matrix has a simple structure and tuples consisting only of real numbers are considered.

The aspect in which we solve the problem arose in connection with the problem of obtaining a criterion for the representability of events in finite probabilistic automata of a special kind. In this connection the notion of a matrix spectrum, introduced below, proved fruitful.

Let A be an arbitrary $n \times n$-matrix with real elements; let $\lambda_s,\ s = 1, 2, \dots, \mu$, be the distinct nonzero characteristic numbers of this matrix; let $m_{sj},\ s = 1, 2, \dots, \mu,\ j = 1, 2, \dots, k_s$, be the degrees of all elementary divisors corresponding to the characteristic root λ_s, and let $m_j,\ j = 1, 2, \dots, t$, be the degrees of the elementary divisors corresponding to the zero root. Let the $n \times n$-matrices E'_{sj} and J'_j have block-diagonal form, all diagonal blocks being zero matrices except the block corresponding to the elementary divisor of index sj, of multiplicity m_{sj}, defined by the root λ_s, which is equal to the $m_{sj} \times m_{sj}$ identity matrix, and the elementary divisor of index j, of multiplicity m_j, defined by the zero root, which is equal to the $m_j \times m_j$ matrix of the (left) shift; $\widetilde{A}$ is the Jordan normal form of the matrix A, and T is a nonsingular matrix reducing it to the form $\widetilde{A}$. Then for any positive integer k the expansion is valid

$$A^k = \sum_{s=1}^{\mu} \sum_{i=1}^{k_s} \sum_{j=1}^{m_{si}-1} C_k^j T^{-1} \left[\frac{\widetilde{A} E'_{si}}{\lambda_s} - E'_{si} \right]^j T \lambda_s^k + \sum_{j=1}^{t} T^{-1} J_j'^k T, \qquad (1)$$

where C_k^j is the number of combinations of k elements taken j at a time if

$k \geq j$, and is equal to zero otherwise. This expansion, valid for nonsingular matrices and for negative k, can also be extended to the case $k = 0$, if one sets $J_j'^0 = E_j'$, $j = 1, 2, \dots, t$. Hence, in particular, we obtain the consequence that for any square matrix with real elements there exists a system of polynomial matrices $P_s(k)$, $s = 1, 2, \dots, \mu$, respectively of degrees not exceeding the maximal degree of the elementary divisors, defined by the root λ_s, minus one, and a system of constant matrices J_s, $s = 1, 2, \dots, t$, which turn into zero in the degree equal to the degree of the s-th elementary divisor defined by the zero root, and are nonzero in smaller degrees, such that for any positive integer-

of a positive k, we have

$$A^k = \sum_{s=1}^{\mu} P_s(k)\lambda_s^k + \sum_{s=1}^{t} J_s^k.$$

Further, in this note we restrict ourselves to considering the case of matrices of simple structure and tuples of real characteristic numbers. If a nonsingular $n \times n$ matrix with real entries A has simple structure, then there exists a system of $n \times n$ matrices $(A_1, A_2, \dots, A_n)$ such that, for any integer power k, the decomposition

$$A^k = \sum_{s=1}^{n} A_s \lambda_s^k, \qquad k = 0,\ \pm 1,\ \pm 2, \dots \tag{2}$$

is valid.

Let $\bar{e}_1, \bar{e}_2, \dots, \bar{e}_n$ be the "right" eigenvectors (columns) of the matrix A, and let $\bar{\varepsilon}_1, \bar{\varepsilon}_2, \dots, \bar{\varepsilon}_n$ be the "left" eigenvectors (rows), so that if T brings A to diagonal form, then

$$T = \begin{pmatrix} \bar{e}_1 \\ \bar{e}_2 \\ \cdot \\ \cdot \\ \bar{e}_n \end{pmatrix}, \qquad T^{-1} = (\bar{\varepsilon}_1 \bar{\varepsilon}_2 \dots \bar{\varepsilon}_n). \tag{3}$$

Then

$$A_s = \bar{e}_s \cdot \bar{\varepsilon}_s, \qquad s = 1, 2, \dots, n. \tag{4}$$

We shall call a system $(A_1 A_2 \dots A_n)$ of $n \times n$ matrices with real entries a **matrix spectrum** (of simple structure) if it satisfies the following system of conditions:

$$\begin{array}{llll} 1. & A_s \neq 0,\ s = 1, 2, \dots, n. & 2. & A_s^2 = A_s,\ s = 1, 2, \dots, n. \\ 3. & A_s \cdot A_k = 0,\ s \neq k. & 4. & \sum_{s=1}^{n} A_s = E. \end{array} \tag{5}$$

A matrix A belongs to the matrix spectrum $(A_1 A_2 \dots A_n)$ if there is a tuple of real numbers $(\lambda_1, \lambda_2, \dots, \lambda_n)$ such that the decomposition (2) holds for nonnegative k. We shall call the matrix spectrum $(A_1 A_2 \dots A_n)$ **stochastic** if there is a tuple of characteristic numbers $(1, \lambda_2, \dots, \lambda_n)$ such that the matrix

$$A = \sum_{s=1}^{n} A_s \lambda_s$$

is stochastic, without being the identity.

Theorem 1. In order that the matrix spectrum $(A_1 A_2 \dots A_n)$ be stochastic, it is necessary and sufficient that the sum of the matrices corresponding to the characteristic root 1 be stochastic and have equal rows.

Let us note that, for a regular matrix, the rows of the matrix A_1 are the vectors of the limiting distribution of the corresponding homogeneous Markov chain.

Theorem 2. In order that a nonsingular $n \times n$ matrix with real entries define, by formulas (3) and (4), a stochastic matrix spectrum, it is sufficient that its first row be a stochastic vector, and that the sum of the entries in each of the remaining rows be equal to zero.

Denote by $\Xi(A_1 A_2 \dots A_n)$ the set of all stochastic matrices belonging to the stochastic matrix spectrum $(A_1 A_2 \dots A_n)$, and by $\Sigma(A_1 A_2 \dots A_n)$ the set of all tuples of real numbers that determine stochastic matrices with spectrum $(A_1 A_2 \dots A_n)$. If, in the set $\Sigma(A_1 A_2 \dots A_n)$, one defines the operations of stochastic linear combination of tuples

$$\nu = L_\alpha(\lambda, \bar{\mu}) = \alpha\lambda + \beta\bar{\mu}, \quad \alpha + \beta = 1,$$

$\alpha, \beta \geq 0$, and the operation of componentwise multiplication of tuples

$$\nu = (\nu_1 \nu_2 \dots \nu_n) = \Pi(\lambda, \bar{\mu}) = (\lambda_1 \mu_1, \lambda_2 \mu_2, \dots, \lambda_n \mu_n),$$

then the set $\Sigma(A_1 A_2 \dots A_n)$, with the system of operations $L_\alpha(\bar{\lambda}, \bar{\mu})$, $0 \leq \alpha \leq 1$, $\Pi(\bar{\lambda}, \bar{\mu})$, forms an algebra; moreover, formula (2) with k equal to one induces an isomorphic algebra on the set $\Xi(A_1 A_2 \dots A_n)$ with the operations of stochastic linear combination and multiplication of matrices. It is clear that multiplication of matrices from Ξ is commutative.

Theorem 3. Let $\varphi(x_1, x_2, \dots, x_N)$ be an arbitrary nonnegative function possessing nonnegative finite derivatives of any order with respect to any combination of arguments at the point $(0, 0, \dots, 0)$, and equal to one at the point $(1, 1, \dots, 1)$. Let $B_1, B_2, \dots, B_N$ be stochastic matrices with tuples of characteristic numbers $\mu_{i1}, \mu_{i2}, \dots, \mu_{iN}$, $i = 1, 2, \dots, n$, belonging to $\Xi(A_1 A_2 \dots A_n)$. Then the stochastic matrix

$$A = \varphi(B_1, B_2, \dots, B_N)$$

is defined, with tuple of characteristic numbers

$$\lambda_i = \varphi(\mu_{i1}, \mu_{i2}, \dots, \mu_{iN}), \quad i = 1, 2, \dots, n,$$

which also belongs to $\Xi(A_1 A_2 \dots A_n)$.

Theorem 4. Let a nonsingular matrix T with real entries determine the matrix spectrum $(A_1 A_2 \dots A_n)$ by formulas (3) and (4). Let

$$d_s = \max_{\substack{k \\ \varepsilon_s^{(k)} > 0}} \left(-\varepsilon_1^{(k)} / \varepsilon_s^{(k)}\right), \quad D_s = \max_{\substack{k \\ \varepsilon_s^{(k)} < 0}} \left(-\varepsilon_1^{(k)} / \varepsilon_s^{(k)}\right),$$

$$q_s = \min_k e_s^{(k)}, \quad Q_s = \max_k e_s^{(k)}, \qquad s = 2, 3, \dots, n;$$

$$\max(d_s/Q_s, D_s/q_s) \le \gamma_s \le \min(d_s/q_s, D_s/Q_s), \qquad s = 2, 3, \dots, n.$$

Then any tuple of real numbers $(1, \lambda_2, \lambda_3, \dots, \lambda_n)$, where

$$\lambda_s = \alpha_1 + \alpha_s \gamma_s, \quad \alpha_s \ge 0, \quad s = 1, 2, \dots, n, \quad \sum_{s=1}^{n} \alpha_s = 1,$$

is the tuple of characteristic numbers of some stochastic matrix belonging to the spectrum $(A_1 A_2 \dots A_n)$.

In the general case, the conditions of Theorem 4 do not completely determine the entire set $\Sigma(A_1 A_2 \dots A_n)$. An example of a tuple that does not satisfy the conditions of Theorem 4, but nevertheless belongs to the set Σ, is provided by the tuple $(1, -\frac{1}{2}, \frac{1}{4})$ for the matrix spectrum

$$\begin{pmatrix} \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \end{pmatrix}, \qquad \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ -\frac{1}{3} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}, \qquad \begin{pmatrix} 0 & -\frac{1}{6} & \frac{1}{6} \\ 0 & \frac{1}{3} & -\frac{1}{3} \\ 0 & -\frac{2}{3} & \frac{2}{3} \end{pmatrix}.$$

We shall call a set of tuples $\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_N, \dots$ belonging to $\Sigma(A_1 A_2 \dots A_n)$ a **basis** if it has the property that any tuple from Σ can be represented in the form of a stochastic linear combination of a finite set of tuples taken from the basis.

Theorem 5. The set $\Sigma(A_1 A_2 \dots A_n)$ has a finite basis. There exists a finite procedure that makes it possible to construct this basis.

One way of constructing this basis is as follows. Let $\bar{n}_{ij}$, $i, j = 1, 2, \dots, n$, be vectors with coordinates

$$\bar{n}_{ij} = \left(e_1^{(i)} \cdot e_1^{(j)},\ e_2^{(i)} \cdot e_2^{(j)}, \dots,\ e_n^{(i)} \cdot e_n^{(j)}\right), \qquad i, j = 1, 2, \dots, n.$$

Let $H = \{\mu\}$ be the set of solutions of the system of linear algebraic equations

$$\mu_1 = 1, \qquad \bar{\mu} \bar{n}_{ij} = 0, \qquad (ij) \in J_N,$$

where J_N is the set of distinct pairs of indices of elements of an $n \times n$ matrix, taken $(n-1)$ at a time, for all possible combinations. The basis of the set $\Sigma(A_1 A_2 \dots A_n)$ is formed by those tuples belonging to H for which the matrix

$$A = \sum_{s=1}^{n} A_s \mu_s$$

is stochastic. By applying methods of linear programming, one can describe the basis without exhaustive enumeration.

Therefore the set Σ_n^* of all tuples of real characteristic numbers of stochastic matrices of simple structure can be written in the form $\Sigma_n^* = \bigcup_T \Sigma(T)$, where $\Sigma(T) = \Sigma(A_1 A_2 \dots A_n)$, and T is an arbitrary nonsingular matrix with real elements satisfying the conditions of Theorem 2.

I take this opportunity to express my gratitude to Prof. V. V. Morozov for valuable consultations.

Kazan State University
named after V. I. Ulyanov-Lenin

Received
9 III 1967

REFERENCES

[1] N. A. Dmitriev, E. B. Dynkin, DAN, **49**, No. 3, 159 (1945).
[2] N. A. Dmitriev, E. B. Dynkin, Izv. AN SSSR, Ser. Mat., **10**, 167 (1946).
[3] F. I. Karpelevich, Izv. AN SSSR, Ser. Mat., **15**, 361 (1951).
[4] Kh. R. Suleimanova, DAN, **66**, No. 3, 343 (1949).
[5] H. Perfect, Proc. Cambridge Phil. Soc., **48**, 271 (1952).

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.