AGGREGATION OF STATES IN A MARKOV CHAIN

1965

SovietRxiv

Abstract

Full Text

MATHEMATICS

V. K. ZAKHAROV, O. V. SARMANOV

AGGREGATION OF STATES IN A MARKOV CHAIN

AND STATIONARY CHANGE OF THE SPECTRUM

(Presented by Academician S. N. Bernstein on 21 VII 1964)

1. We consider two dependent finite schemes of events $\{A_i\}$ and $\{B_j\}$, $i, j = 1, 2, \dots, n$. The probabilities of joint occurrences of the events A_i and B_j are given by a symmetric matrix (correlation table) $\{p_{ij}\}$

$$p_{ij} = p_{ji} = \mathbf{P}(A_{iB}j) \geqslant 0, \tag{1}$$

where

$$p_i = \sum_{j=1}^{n} p_{ij} = \mathbf{P}(A_i) = \mathbf{P}(B_i) > 0, \qquad \sum_{i=1}^{n} p_i = 1. \tag{2}$$

By aggregation of schemes of events we mean the union of several events A_i (and simultaneously B_j with the same indices) into one.

Let

$$\bar{x}_0 = (1, 1, \dots, 1), \qquad \bar{x}_k = (x_{k1}, x_{k2}, \dots, x_{kn}), \qquad k = 1, 2, \dots, n-1;$$

$1, \lambda_1^{-1}, \lambda_2^{-1}, \dots, \lambda_{n-1}^{-1}$, where $1 \leqslant |\lambda_1| \leqslant |\lambda_2| \leqslant \dots \leqslant |\lambda_{n-1}|$, be a system of eigenvectors and their corresponding eigenvalues of the stochastic matrix $\{p_{ij}/p_i\}$. It is known that the p_{ij} have the spectral decomposition

$$p_{ij} = p_{ip}j \left[1 + \sum_{k=1}^{n-1} \frac{x_{ki} x_{kj}}{\lambda_k} \right]. \tag{3}$$

As was shown in ([1]), when a correlation table of order n is aggregated into a table of order $m < n$, the spectrum of eigenvalues $1, 1/\lambda'_1, 1/\lambda'_2, \dots, 1/\lambda'_{m-1}$ of the corresponding stochastic matrix, which we shall call the aggregated one, is expressed in the following way through the spectrum of the original matrix $\{p_{ij}/p_i\}$:

$$\frac{1}{\lambda_i'} = \sum_{j=1}^{n-1} \frac{a_{ij}^2}{\lambda_j}, \qquad i = 1, 2, \dots, m-1, \tag{4}$$

where

$$\sum_{j=1}^{n-1} a_{ij}^2 = 1, \qquad \sum_{i=1}^{m-1} a_{ij}^2 \leqslant 1.$$

In this paper we study the change of the spectrum of the eigenvalues and eigenvectors of stochastic matrices under aggregations of the original correlation table, and also the conditions for preservation of chain dependence under aggregations of states in a Markov chain.

As was shown in ([1]), the nonnegative coefficients a_{ij}^2 from the equalities (4) depend both on the coordinates of the eigenvectors $\bar{x}_k$, $k = 1, 2, \dots, n-1$, of the original matrix $\{p_{ij}/p_i\}$, and on the coordinates of the eigenvectors $\bar{X}_l$, $l = 1, 2, \dots, m-1$, of the aggregated matrix. In the general case the coordinates $\bar{X}_l$ depend both on the coordinates $\bar{x}_k$ and on the spectrum of eigenvalues λ_k^{-1}, $k = 1, 2, \dots, n-1$.

Definition. If the coordinates of the vectors $\overline{X}_l$ do not depend on the eigenvalues λ_k^{-1}, we shall say that, under the given lumping, a **stationary change of the spectrum** occurs.

Under a stationary change of the spectrum, the coefficients a_{ij}^2 do not depend on λ_k^{-1}. From the results of (1) it follows immediately:

Theorem 1. *In order that the change of the spectrum be stationary, it is necessary and sufficient that the conditions $a_{ij} = 0$ for $i > j$ be satisfied.*

Remark. When a stochastic matrix is lumped to order two ($m = 2$), the coefficients a_{ij}^2 have the following expression:

$$a_{ij}^2 = \frac{p_n}{1 - p_n} x_{jn}^2, \qquad j = 1, 2, \dots, n-1, \tag{5}$$

and thus here a stationary change of the spectrum always takes place.

3. Let $\{p_{ij}/p_j\}$ be the matrix of transition probabilities in a simple reversible and stationary Markov chain with n states. Under lumping of states, the chain dependence– “Markovianity” –is, generally speaking, not preserved. It turns out that the conditions for preservation of Markovianity under lumping are a special case of the conditions for a stationary change of the spectrum.

Lemma. *Suppose the n states A_i, $i = 1, 2, \dots, n$, of a simple Markov chain are combined into $m < n$ states $B_1, B_2, \dots, B_m$. If the transition probabilities from all states A_i belonging to B_j to the state B_k are identical, then after lumping one again obtains a simple Markov chain with the lumped states $B_1, B_2, \dots, B_m$.*

This lemma for homogeneous chains is given in ([2]).

Theorem 2. *If all eigenvalues λ_k^{-1} of the original transition-probability matrix $\{p_{ij}/p_j\}$ are distinct, then a necessary and sufficient condition for preservation of Markovianity under the corresponding lumping of states is that all coefficients a_{ij}^2 be equal to zero or to one.*

Necessity. If Markovianity is preserved, then the lumped matrix obtained from the transition-probability matrix for k steps is the k-th power of the lumped transition-probability matrix for one step. In particular, for the eigenvalues of the lumped transition-probability matrix for two steps, we obtain the following expression in terms of the eigenvalues of the original transition-probability matrix for two steps:

$$\left(\frac{1}{\lambda_i'}\right)^2 = \sum_{j=1}^{n-1} \frac{a_{ij}^2}{\lambda_j^2}, \qquad i = 1, 2, \dots, m-1. \tag{6}$$

Writing the expression for the left-hand side of (6) by formula (4), we obtain

$$\left(\sum_{j=1}^{n-1} \frac{a_{ij}^2}{\lambda_j}\right)^2 = \sum_{j=1}^{n-1} \frac{a_{ij}^2}{\lambda_j^2}, \qquad i = 1, 2, \dots, m-1, \tag{7}$$

where $\sum_{j=1}^{n-1} a_{ij}^2 = 1$, i.e. equality holds in the Cauchy inequality, which, for distinct λ_j, can occur only under the condition that, for each fixed i, $i = 1, 2, \dots, m-1$, one coefficient a_{ij}^2 is equal to one and all the others are equal to zero.

Sufficiency is easy to prove by using the spectral decomposition (3), the expression for the coefficients a_{ij}^2 from (1), and the lemma stated above.

For simplicity of notation let us consider the case of a single lumping, in which the first s states $A_1, A_2, \dots, A_s$ of the Markov chain are combined into one state. In this case, using the orthogonality of $\overline{x}_k$ and the results-

states ([1]), we obtain

$$\sum_{i=1}^{n-s} a_{ij}^2 = \frac{(p_1 x_{j1} + \dots + p_s x_{js})^2}{p_1 + \dots + p_s} + \sum_{l=s+1}^{n} p_l x_{jl}^2 = 1 - \sum_{1 \le l < i \le s} \frac{p_l p_i}{p_1 + \dots + p_s} (x_{jl} - x_{ji})^2. \tag{8}$$

Since, by assumption, all a_{ij}^2 are equal to zero or one, and in each row and each column of the matrix of coefficients $\{a_{ij}^2\}$ there can be no more than one unit, it follows that in the matrix $\{a_{ij}^2\}$, $i = 1, 2, \dots, m-1$; $j = 1, 2, \dots, n-1$, there are exactly $n-m$ columns consisting solely of zeros, and $m-1$ columns containing one unit; here $m = n-s+1$. Let the zero columns have numbers $k_1, k_2, \dots, k_{s-1}$; then from formula (8), which gives expressions for the sums of the elements of the columns of the matrix $\{a_{ij}^2\}$, the equalities follow

$$p_1 x_{k_i 1} + \dots + p_s x_{k_i s} = 0, \qquad x_{k_i l} = 0, \quad l = s+1, \dots, n; \quad i = 1, 2, \dots, s-1; \tag{9}$$

$$x_{j1} = x_{j2} = \dots = x_{js}, \qquad j \neq k_1, k_2, \dots, k_{s-1}; \tag{10}$$

substituting (9) and (10) into (3), we obtain the equalities

$$\frac{p_{1l}}{p_1} = \frac{p_{2l}}{p_2} = \dots = \frac{p_{sl}}{p_s}, \qquad l = s+1, \dots, n, \tag{11}$$

which, by virtue of the lemma, guarantee preservation of Markovity.

Remark. If the original matrix $\{p_{ij}/p_i\}$ has equal eigenvalues, for example $1/\lambda_j = 1/\lambda_{j+1} = \dots = 1/\lambda_{j+r}$, then the sum $a_{ij}^2 + a_{i,j+1}^2 + \dots + a_{i,j+r}^2$ must be equal to zero or one for all $i = 1, 2, \dots, m-1$. In particular, if the spectrum of eigenvalues consists of 1 and the number $1/\lambda$ of multiplicity $n-1$, where $|\lambda| > 1$, then, since the sum

$$\sum_{j=1}^{n-1} a_{ij}^2$$

of the coefficients of formulas (4) over the rows is always equal to one, for any lumpings of such a matrix Markovity is always preserved.

In paper ([1]) matrices with a spectrum of eigenvalues satisfying the condition $1/\lambda_i = 1/\lambda$, $i = 1, 2, \dots, n-1$, were called **arithmetic-free**; there the general form of the elements of such a matrix was also found, namely

$$\frac{p_{ii}}{p_i} = p_i \left(1 - \frac{1}{\lambda}\right) + \frac{1}{\lambda}; \qquad \frac{p_{ij}}{p_i} = p_j \left(1 - \frac{1}{\lambda}\right), \qquad i \neq j. \tag{12}$$

From Theorem 2 and the last remark there follows directly

Theorem 3. *In order that, under any lumpings, Markovity be preserved, it is necessary and sufficient that the spectrum of eigenvalues of the transition probability matrix consist of* 1 *and one further number* $1/\lambda$ *of multiplicity* $n-1$*, where* $|\lambda| > 1$.

A statement equivalent to the last theorem is given in ([2]).

Remark. The equality of the coefficients a_{ij}^2 to zero or one, which guarantees preservation of Markovity under lumping, is a special case of a stationary change of the spectrum, since, in general, under a stationary change of the spectrum a_{ij}^2 (or sums of coefficients corresponding to equal original eigenvalues) do not depend on λ_k^{-1}, $k = 1, 2, \dots, n-1$, and when Markovity is preserved they are simply constants.

V. A. Steklov Mathematical Institute
Academy of Sciences of the USSR

Received
21 VII 1964

CITED LITERATURE

1. O. V. **Sarmanov**, V. K. **Zakharov**, *Matem. sborn.*, **52**, no. 4, 953 (1960).

2. C. **Burke**, M. **Rosenblatt**, *Ann. Math. Stat.*, **29**, 1112 (1958).

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