Statistics and Probability Letters 84 (2014) 140–148

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Collapsing of non-homogeneous Markov chains

Agnish Dey [a], Arunava Mukherjea [b,*]

[a] Department of Mathematics, 1400 Stadium Road, Gainesville, FL 32611, United States
[b] 1216 Labrad Lane, Tampa, FL 33613, United States

ARTICLE INFO

Article history:
Received 30 August 2012
Received in revised form 18 September 2013
Accepted 2 October 2013
Available online 8 October 2013

Keywords:
Non-homogeneous Markov chains
Collapsing of Markov chains
Reversibility
Left invariant initial distribution

ABSTRACT

Let $X(n)$, $n \geq 0$, be a (homogeneous) Markov chain with a finite state space $S = \{1, 2, \ldots, m\}$. Let S be the union of disjoint sets $S_1, S_2, \ldots, S_k$ which form a partition of S. Define $Y(n) = i$ if and only if $X(n) \in S_i$ for $i = 1, 2, \ldots, k$. Is the collapsed chain $Y(n)$ Markov? This problem was considered by Burke and Rosenblatt in 1958 and in this note this problem is studied when the $X(n)$ chain is non-homogeneous and Markov. To the best of our knowledge, the results here are new.

1. Introduction

In this paper we study non-homogeneous Markov chains (NHMCs) $X(n)$, $n \geq 0$, with finite state space $S = \{1, 2, \ldots, m\}$, an initial distribution $p = (p_1, p_2, \ldots, p_m)$, $P(X(0) = i) = p_i$, and the transition probability matrices P_n, $n \geq 1$, given by

$$(P_n)_{ij} = P(X(n) = j | X(n-1) = i),$$

in the context of collapsibility. This is an old problem first studied by Burke and Rosenblatt (1958) for homogeneous Markov chains.

The problem can be stated as follows. Let $S_1, S_2, \ldots, S_r$ be r, $1 \leq r \leq m$, pairwise disjoint subsets of S each containing more than one state so that $S = S_1 \cup S_2 \cup \cdots \cup S_r \cup A$, where $A = S - \cup_{i=1}^{r} S_i$. Then the partition of S, given by $S_1, S_2, \ldots, S_r$, and the singletons in A defines a collapsed chain $Y(n)$ given by

$$Y(n) = i \quad \text{if and only if} \quad X(n) \in S_i,$$

$$\text{and} \quad Y(n) = u \quad \text{if and only if} \quad X(n) = u,$$

where $n \geq 0$, $1 \leq i \leq r$, and $u \in A$. The problem we study here is when the collapsed chain $Y(n)$ is Markov. In his book, Rosenblatt (1971), Chapter III, Section 2, Rosenblatt presented some motivating examples in this context. One simple model motivated by these examples can be described as follows.

An experimenter leads a guinea pig into a maze consisting of a straight line path OA, forking at A into three different paths AB, AC, and AD; each path takes the guinea pig back to OA. A positive stimulus (an appealing food) is left on AB, while a negative stimulus (a mild electric shock) is left on AC, and another negative stimulus (a slightly less mild shock) on AD. After observing the guinea pig's journey along this maze a large number of times, the experimenter decides on a 3-state

* Corresponding author.
E-mail address: arunava.mukherjea@gmail.com (A. Mukherjea).

0167-7152/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.spl.2013.10.002

Markov chain model to describe the "learning" behavior of the guinea pig. To simplify the model even further, he now has the problem of deciding if collapsing the paths with negative stimuli can give him a 2-state Markov chain.

Many papers on functions of Markov chains are available in the printed literature. Some of them are included in the references here. However, all such articles are on homogeneous Markov chains. Similar interesting and non-trivial questions can be asked for NHMCs. In this paper, we address some of them. While we are able to solve some of these problems, some interesting problems still remain. For example, how do the present results generalize in the case of a Markov process with state space either denumerable or continuous? Each of these references in the list: Abdel-Moniem and Leysieffer (1982), Glover and Mitro (1990), Iosifescu (1979), Kemeny and Snell (1960), Rogers and Pitman (1981), Rosenblatt (1973) and Rubino and Sericola (1989) is relevant in some respect with the present paper, though not discussed here directly.

In Section 2, we discuss a few relevant examples and some necessary definitions. Section 3 is the main section, where our results are presented.

2. Definitions and examples

Definition 1. The initial distribution vector $p = (p_1, p_2, \ldots, p_m)$, $\sum_{i=1}^{m} p_i = 1$, $0 \le p_i \le 1$ for each i, $P(X(0) = i) = p_i$, where $X(n)$ is a NHMC, is called left invariant if for each $n \ge 1$, $pP_n = p$. □

Note that for $1 \le i \le m$, $1 \le j \le m$,

$$(P_n)_{ij} = P(X(n) = j | X(n-1) = i), \quad n \ge 1.$$

Thus if p is left invariant, then for $n \ge 1$, $p^{(n)} = p$, where $p^{(n)}(i) = P(X(n) = i)$. Notice that the uniform distribution vector $p=(\frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m})$ is left invariant if and only if each P_n is an $m \times m$ bi-stochastic matrix (that is, a matrix for which each row sum is 1 and each column sum is 1).

Consider the $m \times m$ diagonal matrix D such that $D_{ii} = p_i > 0$, $1 \le i \le m$.

Definition 2. The NHMC $X(n)$ is called reversible if and only if $DP_n = P_n^T D$ for each $n \ge 1$. □

If a NHMC $X(n) n \ge 0$ is reversible, then its initial distribution p must be left invariant. The reason is the following. Reversibility implies that for all i, j, $1 \le i \le m$, $1 \le j \le m$, and $n \ge 1$,

$$(DP_n)_{ij} = (P_n^T D)_{ij},$$

or

$$p_i (P_n)_{ij} = p_j (P_n)_{ji}.$$

Thus,

$$\sum_{j=1}^{m} p_i (P_n)_{ij} = \sum_{j=1}^{m} p_j (P_n)_{ji},$$

or

$$p_i = (pP_n)_i, \quad 1 \le i \le m,$$

implying $p = pP_n$ for $n \ge 1$.

Thus, left invariance is relevant when we deal with reversibility and recall that $p^{(n)} = p$ for $n \ge 1$, if p is left invariant.

We should also mention that it follows immediately that for a reversible NHMC $X(n)$,

$$P(X(n) = i, X(n-1) = j) = P(X(n) = j, X(n-1) = i)$$

for $n \ge 1$, $1 \le i \le m$, $1 \le j \le m$. The converse also holds.

Let us also mention that there may not exist any left invariant distribution vector for a NHMC $X(n)$. A very simple example is the following: suppose that for a particular 2-state NHMC,

$$P_n = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} \quad \text{for } n \text{ odd;}$$

$$P_n = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix} \quad \text{for } n \text{ even.}$$