[Volume 4 issue 04 April 2016]
PageNo.1304-1316 ISSN : 2320-7167

INTERNATIONAL JOURNAL OF MATHEMATICS AND COMPUTER RESEARCH

Results on Collapsing of Non Homogeneous Markov Chains

Agnish Dey
University of Florida

Abstract

Let $X(0)$, $X(1)$, $X(2)$,... be a discrete Markov chain with state space $S = \{1,2,...,m\}$. Let S be the disjoint union of sets S_1, S_2, ..., S_r which form a partition of S . De ne $Y(n) = i$ if and only if $X(n) \in S_i$ for $i = 1,2,...,r$. Is the $Y(n)$ chain Markov? Such questions come up in learning theory and in other contexts, when the experimenter observes the derived chain $Y(n)$ rather than the original chain $X(n)$. In the homogeneous case, this problem has been studied in details. In this note this problem is studied when the $X(n)$ chain is non-homogeneous and Markov.

Keywords: Markov chains, Reversibility, Lumpability.

Introduction

In this paper we study finite state space non-homogeneous Markov chains in the context of collapsibility. This is an old problem first addressed by Burke and Rosenblatt in [2]. The problem can be described as follows: let $X(n)$ be a non-homogeneous Markov chain with state space $S = \{1,2,...,m\}$. Let S_1, S_2,..., S_r be r, $1 \le r \le m$ pairwise disjoint subsets of S each containing more than one state so that $S = S_1 \cup S_2 \cup ... \cup S_r \cup A$, where $A = S - \cup_{i=1}^{r} S_i$.

Then the partition of S, given by S_1, S_2,..., S_r and the singletons in A defines a collapsed chain

$Y(n)$ given by:

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

where $n \ge 0$, $1 \le i \le r$, and $u \in A$. In this article we explore conditions under which the collapsed chain $Y(n)$ will be Markov again, see [3] for further discussion on the problem. In [9], one can find motivating examples in this regard.

Here we would like to state some definitions:

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

Note that for $1 \leq i \leq m$ and $1 \leq j \leq m$,

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

Let us consider an $m \times m$ diagonal matrix D such that $D_{ii} = p_i$, $p_i > 0$, $1 \leq i \leq m$.

Definition 2. The NHMC $X(n)$ is called reversible if and only if $DP_n = P_n^T D$ for each $n \geq 1$. Note that if NHMC $X(n)$ is reversible, then its initial distribution vector p must be left invariant (see [3]).

Definition 3. We shall say that a Markov chain is strongly lumpable with respect to a partition $\{A_1, A_2, ..., A_r\}$ of the chain's state space if for every initial vector p, the collapsed chain corresponding to this partition is a Markov chain and the transition probabilities do

not depend upon the choice of p.

Definition 4. A Markov chain is weakly lumpable with respect to a certain partition whenever the markovian property of the corresponding collapsed chain depends upon the choice of the initial vector p. That is in this case, the collapsed chain will be Markov only when some particular initial vectors (it may be just one initial vector) are chosen.

The paper is arranged as follows: In the following section, we will present an alternative (inductive) proof of the main result in [3]. In [7], Kemeny and Snell came up with certain results regarding weak lumpability of homogeneous Markov chains. In the last section, we shall explore similar results in the non-homogeneous context.

1 Alternative Proof of the Main result in [3]

Here we are going to present an alternative proof of the main result in [3] and in the process we will address the case when there are only two sets that collapse more than one state of the original

Markov chain (this is interesting enough in its own right). To make this paper more readable we would like to restate the main result in [3] here (in a slightly different form):

Let $X(n)$ be a non-homogeneous Markov chain (NHMC) with a finite state space S. Suppose there are r pairwise disjoint subsets $S_1, S_2, \ldots, S_r$ of S such that they are the only subsets of S collapsing more than one state. Consider the following conditions:

(1) $\sum_{l \in S_i} P_n(k,l)P_{n+1}(l,u) = P_n(k,S_i)C^{(i)}_{n+1}(u)$
where $C^{(i)}_{n+1}(u) = P(X(n+1) = u | X(n) \in S_i),\ i = 1, 2, \ldots, r$;

(2) For k *not in* S_i, $n \geq 1$, $P_n(k,S_i) = 0$, $i = 1,2,\ldots,r$. Then when $Y(n)$ is Markov, for each quadruple (k,n,i,u), where k *not in* $\cup^r_{j=1}S_j$, $n \geq 1$, $1 \leq i \leq r$, and u *not in* $\cup^r_{j=1}S_j$, either condition (1) or condition (2) holds. Conversely, if condition (2) holds, or if for all k and for $u \notin \cup^r_{i=1} S_i$, condition (1) holds, then $Y(n)$ is Markov.

Here we would like to mention that the homogeneous analogs of this result for $r > 1$ are not given by Burke and Rosenblatt or elsewhere.

First we prove the case $r = 2$ of the result stated above.

Proof: We prove it in several steps.

Step 1: Here we assume that $Y(n)$ chain is Markov and show that for each triple (k,n,S_1) and (k,n,S_2), k *not in* $S_1 \cup S_2$, $n \geq 1$, we must have either $P_n(k,S_i) = 0$ or

$$\sum_{l \in S_i} P_n(k,l)P_{n+1}(l,u) = P_n(k,S_i)C^{i}_{n+1}(u), \quad (1)$$

where $C^{i}_{n+1}(u) = P(X(n+1) = u | X(n) \in S_i),\ i = 1, 2,\ u \notin S_1 \cup S_2$.

Let us now assume that $P_n(k,S_1) > 0$. Then we have: for u *not in* $S_1 \cup S_2$,

$\frac{\sum_{l\in S_1} P_n(k,l)P_{n+1}(l,u)}{P_n(k,S_1)}$

$$= \frac{P(X(n-1)=k)\sum_{l\in S_1} P_n(k,l)P_{n+1}(l,u)}{P(X(n-1)=k)P_n(k,S_1)}$$

$$= \frac{\sum_{l\in S_1} P(X(n+1)=u|X(n)=l)P(X(n)=l,X(n-1)=k)}{P(X(n)\in S_1,X(n-1)=k)}$$

$$= \frac{\sum_{l\in S_1} P(X(n+1)=u|X(n)=l,X(n-1)=k)P(X(n)=l,X(n-1)=k)}{P(X(n)\in S_1,X(n-1)=k)}$$

$$= P(X(n+1)=u|X(n)\in S_1,X(n-1)=k)$$

$$= P(Y(n+1)=u|Y(n)=S_1,Y(n-1)=k)$$

$$= P(Y(n+1)=u|Y(n)=S_1)$$

$= P(X(n+1)=u|X(n)\in S_1) = C^1_{n+1}(u)$. Equation (1) follows for S_1. Similarly, when $P_n(k,S_2) > 0$, equation (1) follows for S_2. This completes step 1.

Step 2: We now assume that condition (2) holds. In this step, we show that then $Y(n)$ chain must be Markov. Let us now consider the states i_0 ,..., i_n of the $Y(n)$ chain. Often, while switching from $Y(n)$ to $X(n)$, we will use $X(n) \in \{i_n\}$ to mean $Y(n) = i_n$, when i_n may not be a single state. Suppose that $i_k = S_1$ only when $k = 0$ and for $k > 0$, i_k is a singleton.

Then we have:

$$P(Y(n)=i_n|Y(n)=i_{n-1},...,Y(0)=i_0)$$

$$= \frac{\sum_{\alpha\in S_1} P(X(n)=i_n|X(n-1)=i_{n-1},...,X(0)=\alpha)P(X(n-1)=i_{n-1},...,X(0)=\alpha)}{\sum_{\alpha\in S_1} P(X(n-1)=i_{n-1},...,X(0)=\alpha)}$$

$$= \frac{P(X(n)=i_n|X(n-1)=i_{n-1})\sum_{\alpha\in S_1} P(X(n-1)=i_{n-1},...,X(0)=\alpha)}{\sum_{\alpha\in S_1} P(X(n-1)=i_{n-1},...,X(0)=\alpha)}$$

$$= P(X(n)=i_n|X(n-1)=i_{n-1})$$

$= P(Y(n) = i_n|Y(n-1) = i_{n-1})$. Similar is the situation when $i_k = S_2$ only when $k = 0$.

Now let us consider the possibility: $m = max\{k|0 \le k \le n, i_k = S_1 \text{ or } S_2\} > 0$. If $i_n = S_2$, then all the states $i_{n-1}, i_{n-2},...,i_0$ are necessarily equal to S_2, as, otherwise, we shall have, $P(Y(n) = i_n|Y(n-1) = i_{n-1},...,Y(0) = i_0) = 0 = P(Y(n) = i_n|Y(n-1) = i_{n-1})$. Also since $P(Y(n) = S_2, Y(n-1) = S_2,...,Y(1) = S_2, Y(0) \text{ not in } S_2) = 0 = P(Y(n-1) = S_2,...,Y(1) = S_2, Y(0) \text{ not in } S_2)$, we have, when $i_n = S_2$, the following:

$P(Y(n) = i_n|Y(n-1) = i_{n-1},...,Y(0) = i_0) = P(Y(n) = S_2|Y(n-1) = S_2,...,Y(0) = S_2)$

$= P(Y(n) = S_2|Y(n-1) = S_2,...,Y(1) = S_2)$

$= P(Y(n) = S_2|Y(n-1) = S_2) = P(Y(n) = i_n|Y(n-1) = i_{n-1})$.

Similar is the situation when $i_n = S_1$.

In case $m = n - 1$ and $i_{n-1} = S_2$, again, as before, $i_0 = i_1 = \ldots = i_{n-2} = S_2$, as, otherwise, $P(Y(n-1) = i_{n-1},...,Y(0) = i_0) = 0$. Thus, it follows, as before, $P(Y(n) = i_n|Y(n-1) = i_{n-1},...,Y(0) = i_0) = P(Y(n) = i_n|Y(n-1) = i_{n-1})$. Finally let us consider the possibility $0 < m < n - 1$ and $i_m = S_2$. In this case, as before, we have necessarily $i_0 = i_1 = \ldots = i_{m-1} = S_2$, and i_{n-1} is either S_2 or a singleton element not in $S_1 \cup S_2$. Thus, in this case, when $i_m = i_{n-1} = S_2$, as before, we have:

$P(Y(n) = i_n|Y(n-1) = i_{n-1},...,Y(m) = i_m,...,Y(0) = i_0) = P(Y(n) = i_n|Y(n-1) = i_n)$.

In case $i_m = S_2$ and i_{n-1} not in $S_1 \cup S_2$, i_{n-1} is a singleton, and so is each i_k, $m < k < n - 1$, and as before,

$$P(Y(n) = i_n|Y(n-1) = i_{n-1},...,Y(m) = i_m,...,Y(0) = i_0)$$

$$= P(Y(n) = i_n|Y(n-1) = i_{n-1},...,Y(m+1) = i_{m+1},Y(m) = S_2,...,Y(0) = S_2)$$

$$= P(Y(n) = i_n|Y(n-1) = i_{n-1},...,Y(m+1) = i_{m+1})$$

$$= P(X(n) \in \{i_n\}|X(n-1) = i_{n-1},...,X(m+1) = i_{m+1}) = P(X(n) \in \{i_n\}|X(n-1) = i_{n-1})$$

$$= P(Y(n) = i_n|Y(n-1) = i_{n-1}),$$

using the fact that $X(n)$ is Markov and that each of the elements $\{i_{m+1},...,i_{n-1}\}$ is a singleton. Step 2 is now complete.

Step 3: In this step we assume that condition (1) holds for all k and n, and show that then $Y(n)$ must be Markov. We consider again the states $i_0, i_1,...,i_n$ of the $Y(n)$ chain. The state i_k of $Y(n)$ may be either S_1 or S_2, or simply a singleton element not in $S_1 \cup S_2$. Thus, $\{i_k\} = i_k$, when it is a singleton, and $\{i_k\}$ denotes the set S_1 or S_2 otherwise. We have the following possibilities:

(i) $i_{n-1} = S_j$, $i_n \neq S_j$, $j = 1,2$; (ii) $i_n = S_j$, $i_{n-1} \neq S_j$, $j = 1,2$; (Note that i_n in (i) and i_{n-1} in (ii) are both singleton elements as no transition is possible from S_1 to S_2.) (iii) $i_n = i_{n-1} = S_j$, $j = 1,2$. Let us now assume (i), and $i_{n-1} = S_1$ and i_n, a singleton not in $S_1 \cup S_2$. Then we have:

$$P(Y(n) = i_n, Y(n-1) = S_1, Y(n-2) = i_{n-2}, \ldots, Y(0) = i_0)$$

$$= P(X(n) = i_n, X(n-1) \in S_1, X(n-2) \in \{i_{n-2}\}, \ldots, X(0) \in \{i_0\})$$

$$= C_n^1(i_n) P(X(n-1) \in S_1, X(n-2) \in \{i_{n-2}\}, \ldots, X(0) \in \{i_0\}) \text{ (using condition (1) and simplifying).}$$

Thus, it is clear that

$$P(Y(n) = i_n | Y(n-1) = i_{n-1}, \ldots, Y(0) = i_0) = C_n^1(i_n) = P(X(n) = i_n | X(n-1) \in S_1)$$

$$= P(Y(n) = i_n | Y(n-1) = i_{n-1}).$$

Now let us observe that case (ii) is simple, and here we do not have to use condition (1) to prove the Markov property of $Y(n)$.

The last case is similar to case (i). Using condition (1), as in case (i), we can show here that

$P(Y(n) = S_1 | Y(n-1) = S_1, Y(n-2) = i_{n-2}, \ldots, Y(0) = i_0)$

$= P(Y(n) = S_1 | Y(n-1) = S_1)$. Step 3 is thus taken care of. This concludes the proof.

Now we shall prove the result stated in the introduction of this paper using induction.

Proof: Let r be a positive integer and $\{X(n) : n \geq 0\}$ be a NHMC with finite state space S. Let S_1, $S_2, \ldots, S_r$ be r pairwise disjoint subsets of S such that each of these contains more than one state. Define the chain $Z(n)$ such that,

$$Z(n) = i \text{ if and only if } X(n) \in S_i,\ 1 \leq i \leq r;$$

$$Z(n) = x \text{ if and only } X(n) = x, \text{ when } x \notin \cup_{i=1}^{r} S_i.$$

Note that the main result holds for $r = 2$. We assume $r > 2$ and that the main result holds for all chains $Y(n)$ which collapse $p(\leq r)$ pairwise disjoint subsets of S, each containing more than one state. To prove the result by induction, we need to prove it for a chain $Y(n)$

such that

$$Y(n) = i \text{ if and only if } X(n) \in S_i,\ 1 \leq i \leq r + 1;$$
$$Y(n) = x \text{ if and only if } X(n) = x \text{ when } x \notin \cup_{i=1}^{r+1} S_i,$$

where $S_1, S_2, \ldots, S_{r+1}$ are pairwise disjoint subsets of S. To this end, let us assume first that the $Y(n)$ chain is Markov. We must prove that either condition (2) or condition (1) must then hold, for all k and u in $(\cup_{i=1}^{r+1} S_i)^c$. Let $k \notin \cup_{i=1}^{r+1} S_i$. If for this k and for some $n \geq 1$, $P_n(k,S_i) > 0$ for some i, $1 \leq i \leq r + 1$, then just like in step 1 of the above result, we can show that condition (1) holds for each quadruple (k,n,i,u) as required. We do not need to use induction for this part.

Conversely, let us assume that condition (2) holds. We need to show that the $Y(n)$ chain is Markov. Let $i_0, i_1, \ldots, i_n$ be $(n + 1)$ states of the $Y(n)$ chain. We assume that $i_0 = S_{r+1}$, and for $0 < j \leq n$, $\{i_j\} \cap S_{r+1} = \emptyset$. By induction hypothesis, the $Z(n)$ chain (defined earlier) is

Markov. Thus, we can write:

$P(Y(n) = i_n, Y(n-1) = i_{n-1}, \ldots, Y(0) = i_0)$

$=P(Y(n) = i_n | Y(n-1) = i_{n-1}) P(Y(n-1) = i_{n-1} = i_{n-1}, \ldots, Y(0) = i_0)$

Now let $m = max\{j | 0 \leq j \leq n, i_j = S_{r+1}\} > 0$. If $i_n = S_{r+1}$, and $\{i_{n-1}\} \cap S_{r+1} = \emptyset$, then we have, $P(Y(n) = i_n | Y(n-1) = i_{n-1}, \ldots, Y(0) = i_0) = P(Y(n) = i_n | Y(n-1) = i_{n-1}) = 0)$. If $i_n = i_{n+1} = S_{r+1}$, then we must have by condition (2), $i_0 = i_1 = \ldots = i_{n-1} = S_{r+1}$; otherwise,

$P(Y(n-1) = i_{n-1}, \ldots, Y(0) = i_0) = 0$. Thus, when $i_n = i_{n-1} = S_{r+1}$,

$P(Y(n) = S_{r+1}, Y(n-1) = S_{r+1}, \ldots, Y(0) = S_{r+1})$

$=P(Y(n) = S_{r+1}, Y(n-1) = S_{r+1}, \dots, Y(1) = S_{r+1})$

$=P(Y(n) = S_{r+1}, Y(n-1) = S_{r+1}, \dots, Y(2) = S_{r+1}) = P(Y(n) = S_{r+1})$, since $P(Y(n) = S_{r+1}, Y(n-1) = S_{r+1}, \dots, Y(n-s) = S_{r+1})$

$=P(Y(n) = S_{r+1}, Y(n-1) = S_{r+1}, \dots, Y(n-s) = S_{r+1}) + P(Y(n) = S_{r+1}, Y(n-1) = S_{r+1}, \dots, Y(n-s)$ not in $S_{r+1})$

$=P(Y(n) = S_{r+1}, \dots, Y(n-s+1) = S_{r+1}) = P(Y(n) = S_{r+1})$. Now suppose $n-1 = m$ so that $i_{n-1} = S_{r+1}$, and $\{i_n\} \cap S_{r+1} = \emptyset$. Then, as before, as no transition is possible from S_i to S_j $(i \neq j)$, we must have $i_0 = i_1 = \dots = i_{n-1} = S_{r+1}$, and as before, $P(Y(n) = i_n | Y(n-1) = S_{r+1}, \dots, Y(0) = S_{r+1}) = P(Y(n) = i_n | Y(n-1) = i_{n-1})$. Finally, we assume that $0 < m < n-1$. Then again we must have: $i_0 = i_1 = \dots = i_m = S_{r+1}$. It is also clear, as we showed earlier,

$P(Y(n) = i_n | Y(n-1) = i_{n-1}, \dots, Y(m+1) = i_{m+1}, Y(m) = S_{r+1}, \dots, Y(0) = S_{r+1})$

$=P(Y(n) = i_n | Y(n-1) = i_n, \dots, Y(m+1) = i_{m+1})$, where i_{m+1} can only be either S_{r+1} or a singleton not in $\cup_{i=1}^{r+1} S_i$. If $i_{m+1} = S_{r+1}$, then continuing the same reasoning, we must have i_{n-1} equal to either S_{r+1} or a singleton not in $\cup_{i=1}^{r+1} S_i$. Thus, it is no loss of generality to consider each i_j, $0 \le j \le n$, as a singleton not in $\cup_{i=1}^{r+1} S_i$. Then the Markov property follows immediately if we replace $Y(j)$ by $Z(j)$, $o \le j \le n$, and observe that the $Z(n)$ chain is Markov, by induction hypothesis.

Finally, we assume that condition (1) holds for the $Y(n)$ chain, and then we need to show that $Y(n)$ chain is Markov. By induction hypothesis, the $Z(n)$ chain is Markov. Let i_o, $i_1, \dots, i_n$ be $(n+1)$ states of the $Y(n)$ chain. Then each i_j is either the set S_i, $1 \le i \le r+1$, or just a singleton element not in $\cup_{m=1}^{r-1} S_m$. There are five possibilities: (i) $i_{n-1} = S_t$, $i_n = S_j$, $t \neq j$, $1 \le t \le r+1$, $1 \le j \le r+1$; (ii) $i_{n-1} \notin \cup_{t=1}^{r+1} S_t,\ i_n = S_j,\ 1 \le j \le r+1$; (iii) $i_{n-1} = i_n = S_j$, $1 \le j \le r+1$; (iv) $i_{n-1} = S_j,\ 1 \le j \le r+1,\ i_n \notin \cup_{t=1}^{r+1} S_t$; (v) $i_{n-1} \notin \cup_{t=1}^{r+1} S_t,\ i_n \notin \cup_{t=1}^{r+1} S_t$. Note that we must prove: $P(Y(n) = i_n | Y(n-1) = i_{n-1}, \dots, Y(0) = i_0) = P(Y(n) = i_n | Y(n-1) = i_{n-1})$. When (v) occurs since then in both sides of the above equation $Y(m)$ can be replaced by $Z(n)$, $0 \le m \le n$, and $Z(n)$ is Markov by induction hypothesis. Also, both sides are equal to zero when (i) occurs, since, by hypothesis in the statement of the main result, transition from S_i to S_j is not possible when $i \neq j$.

We now assume (ii). In this case, i_{n-1} is a singleton and $i_n = S_j$, for some j, $1 \leq j \leq r + 1$. Thus we can write:

$$P(Y(n) = i_n | y(n-1) = i_{n-1}, ..., Y(0) = i_0)$$
$$= \frac{P(X(n) \in S_j | X(n-1) = i_{n-1}) P(Y(n-1) = i_{n-1}, ..., Y(0) = i_0)}{P(Y(n-1) = i_{n-1}, ..., Y(0) = i_0)}$$

$=P(Y(n) = i_n | Y(n-1) = i_{n-1})$.

Let us now assume that (iii). Let $i_{n-1} = i_n = S_j$, $1 \leq j \leq r + 1$. If $1 \leq j \leq r$, then replacing each $Y(m)$ by $Z(m)$ and noting that the $Z(n)$ chain is Markov, it is easy to see that

$P(Z(n) = i_n, Z(n-1) = i_{n-1}, ..., Z(0) = i_0)$

$=P(Z(n) = j | Z(n-1) = j) P(Z(n-1) = i_{n-1}, ..., Z(0) = i_0)$ and this implies that the Markov property holds.

Let us now assume that $i_{n-1} = i_n = S_{r+1}$. In this case, using condition (1) we can show that $Y(n)$ is Markov. The case (iv) can be taken care of in the same way as (iii).

New Results

Theorem 1: Let $X(n)$, $n \geq 0$, be a reversible NHMC with state space $S = \{1,2,...,m\}$ having $P_n = P_{n+1}$ for each odd n. Let $S_1, S_2, ..., S_r$ be a partition of S, $r \leq m$. We de ne $Y(n) = i$ if and only if $X(n) \in S_i$, $i = 1,2,...,r$. Then weak lumpability of $X(n)$ with respect to uniform initial probability vector P(X(0)=i)=1/m implies strong lumpability.

Proof: We de ne the $m \times r$ matrix B and the $r \times m$ matrix A as they are done in [3].

Since $X(n)$ is weakly lumpable with respect to uniform initial probability vector, we claim $Q_n Q_{n+1} = A P_n P_{n+1} B$ [where Q_n is the transition probability matrix of the lumped chain $Y(n)$].

Proof of the claim (for any left invariant initial probability vector p, proof for uniform initial probability vector is very similar):

$$
\begin{aligned}
(Q_nQ_{n+1})_{ij} &= P(Y(n+1)=j|Y(n-1)=i)\\
&= \frac{P(X(n+1)\in S_j, X(n-1)\in S_i)}{P(X(n-1)\in S_i)}\\
&= \frac{1}{\sum_{j\in S_i} p_j}\sum_{k\in S_i} P(X(n+1)\in S_j, X(n-1)=k)\\
&= \frac{1}{\sum_{j\in S_i} p_j}\sum_{k\in S_i}\sum_{l} P(X(n+1)\in S_j, X(n)=l, X(n-1)=k)\\
&= \frac{1}{\sum_{j\in S_i} p_j}\sum_{k\in S_i}\sum_{l} P(X(n+1)\in S_j|X(n)=l)P(X(n)=l|X(n-1)=k)P(X(n-1)=k)\\
&= \frac{1}{\sum_{j\in S_i} p_j}\sum_{k\in S_i}\sum_{l,t\in S_j} P(X(n+1)=t|X(n)=l)P(X(n)=l|X(n-1)=k)p_k\\
&= \sum_{k\in S_i, t\in S_j, l} \frac{p_k}{\sum_{j\in S_k} p_j} P_n(k,l)P_{n+1}(l,t)B_{tj}\\
&= \sum_{k\in S_i, t\in S_j, l} A_{ik}P_n(k,l)P_{n+1}(l,t)B_{tj}\\
&= (AP_nP_{n+1}B)_{ij}.
\end{aligned}
$$

Hence the claim follows.

Thus from Theorem 2 in [3] we have : $P_nB = BAP_nB$. Now,

$$(P_nB)_{ij} = \textstyle\sum_k P_n(i,k)B_{kj} = \sum_{k\in S_j} P_n(i,k).$$

And,

$$
\begin{aligned}
(BAP_nB)_{ij} &= \sum_{k,l,t} B_{ik}A_{kl}P_n(l,t)B_{tj}\\
&= \sum_{i,l\in S_k}\sum_{t\in S_j}\frac{1}{|S_k|}P_n(l,t)\\
&= \sum_{i,l\in S_k}\frac{1}{|S_k|}P_n(l,S_j).
\end{aligned}
$$

Thus from Theorem 1 in [3], we have the sufficient condition for $X(n)$ to be strongly lumpable.

In the following we shall use the notion of reversibility developed in [7] which coincides with our definition here if initial probability vector of the original chain is uniform (this can be verified pretty easily).

Theorem 2: Let $X(n)$ be a reversible NHMC with state space $S = \{1,2,\ldots,m\}$ and uniform initial probability vector. Let $\{S_1,S_2,\ldots,S_r\}$ be a partition of S, $r \leq m$. Then the collapsed chain with respect to this partition is also reversible given that the collapsed chain is Markov.

Proof: Since the collapsed chain is Markov, from Theorem 2 in [3] we have: $Q_n = AP_nB$. Since $X(n)$ is reversible, we obtain: $Q_n=AP_nB=ADP_n^TD^{-1}B$, where $D_{ii}=1/m$.

We define $\hat{D}_{ii} = \frac{D_{jj}}{|S_i|}, j \in S_i$. Now using the same A,B matrices we get:

$$(\hat{D}B^T)_{ik} = \hat{D}_{ii}B_{ki} = \frac{1}{|S_i|}D_{kk}, \;\; k \in S_i (AD)_{ik} = A_{ik}D_{kk} = \frac{1}{|S_i|}D_{kk}, \;\; k \in S_i$$

Thus we have $DB^T = AD$. Similarly,

$$(A^T\hat{D^{-1}})_{ki} = A_{ik}\hat{D_{ii}^{-1}} = \frac{1}{|S_i|}\frac{|S_i|}{D_{kk}} = \frac{1}{D_{kk}}, \;\; k \in S_i$$

and

$$(D^{-1}B)_{ki} = D_{kk}^{-1}B_{ki} = \frac{1}{D_kk}, \;\; k \in S_i$$

This gives us $A^TD^{-1} = D^{-1}B$. Hence we have $Q_n = ADP_n^TD^{-1}B = \hat{D}B^TP_n^TA^T\hat{D^{-1}} = \hat{D}(AP_nB)^T\hat{D^{-1}} = \hat{D}Q_n^T\hat{D^{-1}}$, that is the collapsed chain is also reversible.

Reverse Markov chains: A Markov chain observed in the reversed order is also Markov because of the following:

$$P(X(n-1)=i_{n-1}|X(n)=i_n, X(n+1)=i_{n+1}, ..., X(n+p)=i_{n+p})$$

$$=\frac{P(X(n+p)=i_{n+p}, ..., X(n-1)=i_{n-1})}{P(X(n+p)=i_{n+p}, ..., X(n)=i_n)}$$

$$=\frac{P(X(n+p)=i_{n+p}|X(n+p-1)=i_{n+p-1})...P(X(n)=i_n|X(n-1)=i_{n-1})P(X(n-1)=i_{n-1})}{P(X(n+p)=i_{n+p}|X(n+p-1)=i_{n+p-1})...P(X(n+1)=i_{n+1}|X(n)=i_n)P(X(n)=i_n)}$$

$$=\frac{P(X(n)=i_n, X(n-1)=i_{n-1})}{P(X(n)=i_n)}$$

$$=P(X(n-1)=i_{n-1}|X(n)=i_n).$$

We shall use this property to establish our last result:

Theorem 3: If a given NHMC is weakly lumpable with respect to partition $A = \{A_1, A_2, ..., A_n\}$,

then so is the reverse chain.

Proof: Let $X(n)$ be a non homogeneous Markov chain which is weakly lumpable with respect to the partition $A = \{A_1, ..., A_n\}$. We need to show that all probabilities of the form $P_\beta(X(1) \in A_i|X(2) \in A_j, ..., X(n) \in A_t)$ depend only upon A_i and A_j where β is the initial vector with respect to which the collapsed chain is Markov.

$P(Y(1) = i|Y(2) = j, ..., Y(n) = t)$

$=P(Y(1) = i|Y(2) = j)$ from the above discussion on reverse Markov chains

$=P_\beta(X(1) \in A_i|X(2) \in A_j)$.

Conclusion

In this article we have dealt with Markov chains with finite state space only. Along with the results in [3], now we have a sound understanding about markovian property of collapsed Markov chains with finitely many states. But this particular problem is still open when the original Markov chain has countable or uncountably many states.

Conflict of Interest: None declared.

Reference

[1] A.M.Abdel-Moniem; F. Leysie er. (1982) Weak lumpability in nite markov chains, J. App. Prob. 19, 685-691.

[2] C. J. Burke; M Rosenblatt. (1958) A markovian function of a markov chain, Ann. Math. Stat. 29, 1112-1122.

[3] A. Dey; A. Mukherjea. (2014) Collapsing of non-homogeneous markov chains, Statistics and Probability Letters, 84, 140-148.

[4] Joe Glover; Joanna Mitro. (1990) Symmetries and functions of markov processes, Ann. Prob. 18, 655-658.

[5] N. Higham. (1990) Analysis of cholesky decomposition of a semi-de nite matrix. Reliable Numerical Computations , 161-185.

[6] M. Iosifescu. (1979) Finite markov processes and their applications, John Wiley and Sons.

[7] J.G. Kemeny; J.L. Snell. (1960) Finite markov chains, Springer-Verlag.

[8] L.C.G Rogers; J.W. Pitman. (1981) Markov Functions, Ann. Prob. 9, 573-582.

[9] M. Rosenblatt. (1971) Markov process structure and asymptotic behavior, Springer-Verlag.

[10] M. Rosenblatt. (1973) Random processes(second edition) , SpringerVerlag (New York).

[11] G. Rubino; B. Sericola. (1989) On weak lumpability in markov chains, J. App. Prob. 26, 446-457.