8-2011

Collapsing of Non Homogeneous Markov Chains

Agnish Dey
University of Texas-Pan American

COLLAPSING OF NON HOMOGENEOUS MARKOV CHAINS

A Thesis

by

AGNISH DEY

Submitted to the Graduate School of the
University of Texas Pan American
In partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2011

Major Subject : Mathematics

COLLAPSING OF NON HOMOGENEOUS MARKOV CHAINS

A Thesis
by
AGNISH DEY

COMMITTEE MEMBERS

Arunava Mukherjea
Chair of Committee

Paul Bracken
Committee Member

Santanu Chakraborty
Committee Member

George Yanev
Committee Member

August 2011

ABSTRACT

Dey, Agnish, Collapsing of Non Homogeneous Markov Chains. Master of Science (MS), August , 2011 , 45 pp, references, 11 titles.

In this thesis we have considered questions about lumpability of a non homogeneous markov chain. A number of results similar to those obtained in [2] in the case of homogeneous markov chains have been presented here. We have also considered a few results along the lines of those considered in [8] .

DEDICATION

This thesis is dedicated to my parents, my high school teachers Mrs Anubha Chatterjee, Mr Sidhhartha Chatterjee , Mr Arun Roy , my college professor Dr. Sankar Chandra Ghosh and of course to my advisor Dr. Arunava Mukherjea. Thank you all for your unconditional love , guidance and invaluable inspiration that you have always given me. Your immense faith in my abilities was the central driving force behind the successful completion of this thesis. Love you all.

ACKNOWLEDGEMENTS

I am grateful to Dr. Arunava Mukherjea for introducing me to this problem . Without his guidance this thesis would not have been possible. As I was always interested in probability theory , stochastic processes to be more precise , he suggested that I could try to work on this problem following the classical results of Burke and Rosenblatt in the context of collapsed homogeneous markov chains. I am also thankful to Dr. Bhatta , Dr. Bracken , Dr. Chakrabarty and to Dr. Yanev for their constant support and encouragement. Finally my sincere gratitude goes to our department chair Dr. Lokenath Debnath who has made the first two years of my life as a pupil of mathematics extremely memorable.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

This thesis is essentially concerned with the markovian property of the collapsed non homogeneous markov chains. Let us assume that $X(n)$ is a finite markov chain with state space $\{1,2,...,m\}$. Suppose that $\cup_{i=1}^{r} S_i$ gives a partition of the state space of $X(n)$. Clearly $r \leq m$. We define the collapsed chain $Y(n)$ as follows: $Y(n) = i$ if and only if $X(n) \in S_i$, where $i = 1,2,...,r$. The question that we try to explore in this thesis is whether the collapsed chain $Y(n)$ is markov again. It is clear this is not the case in general .

It is, of course, well known (see [2]) that significant work has been done in this problem assuming the homogeneity property of the original markov chain. In this thesis we shall examine the conditions under which the collapsed chain will be markovian given that the original chain is non-homogeneous markov. To the best of our knowledge , this work in the non-homogeneous context is new and not available in published literature yet. However, we have heard about one unpublished preliminary work stated for bistochastic non-homogeneous chains done in the present context by C. Lo many years ago. We believe our theorems 1 and 2 (see chapter iv) altogether are different in some sense even from the homogeneous version in [2]. Also our results in theorem 3 and its corollary, our reversibility results, and our examples (see chapter iv) are all new.

CHAPTER II

INTRODUCTION TO MARKOV PROPERTY

Stochastic Process

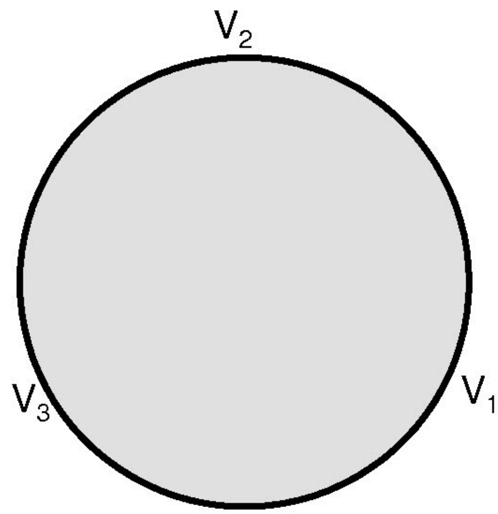
A stochastic process is a family of random variables defined on some sample space Ω. If there are countably many members of the family, then the process is denoted by $X_1, X_2, X_3, \ldots$. If there are uncountably many members of the family, then the process is denoted by $\{X(t) : t \geq 0\}$. In the first case the process is called a discrete-time process while in the second case it is called a continuous-time process.

State Space

The set of distinct values assumed by a stochastic process is called the state space of the process. If the state space of a stochastic process is countable or finite then the process is called a chain.

Markov Property

Let us assume that we have a circular track with three landmarks given by v_1, v_2, v_3 as shown in the following figure.



At time 0 a man stands at v_1. At time 1, he flips a fair coin and moves immediately to v_2 or v_3 according to whether the coin comes up with a head or a tail. At time 2 he flips the coin again to decide which of the two adjacent landmarks to move to, with the decision rule that if the coin comes up with a head , he moves one step clockwise and anti-clockwise in case the coin comes up with a tail .

For each n, let $X(n)$ denote the index of the track landmark at which the walker stands at time n . Hence $X_0, X_1, \ldots$ is a random process taking values in $\{v_1, v_2, v_3\}$. Since the man starts at time 0 in v_1, we have $P(X(0) = v_1) = 1$. Next he will move to v_2 or v_3 with probability $\frac{1}{2}$ each , so that $P(X(1) = v_2) = \frac{1}{2}$ and $P(X(1) = v_3) = \frac{1}{2}$. To compute the distribution of $X(n)$ for $n \geq 2$ requires a little more thought . Suppose at time n, the man stands at v_2. Then we have ,

$$P(X(n+1) = v_1 | X(n) = v_2) = \frac{1}{2} \text{ and}$$
$$P(X(n+1) = v_3 | X(n) = v_2) = \frac{1}{2}$$

Therefore,

$$P(X(n+1) = v_1 | X(n) = v_2, X(n-1) = i_{n-1}, \ldots, X(1) = i_1, X(0) = i_0) = \frac{1}{2} \text{ and}$$

$$P(X(n+1)=v_3|X(n)=v_2, X(n-1)=i_{n-1},...,X(1)=i_1, X(0)=i_0)=\frac{1}{2}$$

for any choice of $i_0, i_1, ..., i_{n-1}$ for which the conditional probabilities are defined . This property is known as the markov property.

Let us now assume that at a particular instant the man moves along the clockwise direction with probability p and along the anti-clockwise direction with probability $(1-p)$, i.e we have

$P(X(n+1)=v_2|X(n)=v_1)=p$ and

$P(X(n+1)=v_3|X(n)=v_1)=1-p$

In this markov chain the probabilities at different time instants are computed as follows:

We have $P(X(0)=v_1)=1$. Then clearly

$P(X(1)=v_2)=p$ and $P(X(1)=v_3)=1-p$

Then $P(X(2)=v_3|X(1)=v_2)=p$ and $P(X(2)=v_1|X(1)=v_2)=1-p$

Then $P(X(2)=v_3)$

$=P(X(2)=v_3|X(1)=v_2)\times P(X(1)=v_2)$

$=p^2$

Again $P(X(2)=v_1)$

$=P(X(2)=v_1|X(1)=v_2)\times P(X(1)=v_2)+P(X(2)=v_1|X(1)=v_3)\times P(X(1)=v_3)$

$=(1-p)p+p(1-p)$ [using markovian property]

$=2p(1-p)$

And $P(X(2)=v_2)$

$=P(X(2)=v_2|X(1)=v_3)\times P(X(1)=v_3)$

$=(1-p)^2$

Similarly , $P(X(3) = v_1) = p^3 + (1-p)^3$, $P(X(3) = v_2) = 3p^2(1-p)$ and $P(X(3) = v_3) = 3p(1-p)^2$. In general one can find the distribution of $X(n)$ for general n .

Formal Definition

A stochastic process $\{X(k)\}$, $k = 0,1,2,...$ with state space $S = \{1,2.3,...\}$ is said to satisfy the markov property if for every n and all states $i_0, i_1, i_2, ..., i_n$ it is true that

$P(X(n) = i_n | X(n-1) = i_{n-1}, ...X(1) = i_1, X(0) = i_0)$

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

In the above random walk example , the conditional distribution of $X(n+1)$ given that $X(n) = v_2$ (say) is the same for all n. This property is known as homogeneity.

Homogeneous Markov Chain

A discrete time markov chain is said to be stationary or homogeneous in time if the probability of going from one state to another is independent of the time at which the step is being made. That is , for all states i and j,

$$P(X(n) = j | X(n-1) = i) = P(X(n+k) = j | X(n+k-1) = i) \text{ for all integers } k \text{ , } n+k \geq 1$$

.

The markov chain is said to be non-homogeneous if the condition for homogeneity fails.

Examples

(A) Assume a machine is producing items independently at the rate of one a minute. Let $X(n)$ denote the number of defectives produced by time n. If the probability of producing a defective item remains constant throughout the life of the machine, then $X(n)$ would be a stationary markov chain. However , if the probability of producing a defective item changes as the machine grows older , then the markov chain will be non-homogeneous .

(B) Consider a drunk man walking along a long straight road and let $X(n)$ denote his position at time n relative to some fixed position O on the road. Let

$P(X(n+1)=i+1|X(n)=i)=a_n>0$ and

$P(X(n+1)=i-1|X(n)=i)=1-a_n>0$

Here we observe that the probability at a particular time instant n depends upon n . Hence this $X(n)$ is a non-homogeneous markov chain.

Transition Probability Matrix

Let $\{X(k)\}$ denote a discrete time homogeneous markov chain with a finite state space $S=\{1,2,...,n\}$. For this chain there are n^2 transition probabilities , $\{p_{ij}\}$ $i=1,2,...,n$; $j=1,2,...,n$. The most convenient way of recording these values is in the form of a matrix P. Associate the i^{th} row and j^{th} column element of P with the transition probability

$p_{ij}=P(X(n+1)=j|X(n)=i)$. Then we have,

$$P=\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ : & : & : & : \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

This matrix is called the transition probability matrix corresponding to the markov chain $\{X(k)\}$. Every transition matrix has the following properties:

(1) All the entries are non negative.

(2) The sum of the entries in each row is one.

Both the properties follow easily. Any square matrix with these two properties is called a $\underline{stochastic \quad matrix}$.

Thus the transition probability matrix of a homogeneous markov chain is always a stochastic matrix . Note that product of stochastic matrices is again a stochastic matrix .

Initial Distribution

Initial distribution tells us about the commencement of the markov chain . The initial distribution is represented as a row vector $w^{(0)}$ given by

$$w^{(0)} = (w_1^{(0)}, w_2^{(0)}, ..., w_m^{(0)})$$

$= (P(X(0) = 1), P(X(0) = 2), ..., P(X(0) = m)$ where $\{1, 2, ..., m\}$ is the state space of the markov chain.

Since $w^{(0)}$ represents a probability distribution , we have $\sum_{i=1}^{m} w_i^{(0)} = 1$ and $w_i \geq 0$ for all $i = 1, 2, ..., m$.

Chapman Kolmogorov Equation

$P_{ij}^{(n)} = \sum_{k=0}^{m} P_{ik}^{(r)} P_{kj}^{(n-r)}$ for all $0 < r < n$ where m is the total number of states.

Consequences of Markov Property

(1) Given the past $\{X(0) = i_0, ..., X(n) = i_n\}$ the markov property suggests that the current state $X(n) = i_n$ is enough to determine all distributions of the future. To see this , the chain rule of the conditional probability yields

$$\begin{aligned} &P[X(n+m) = i_{n+m}, ..., X(n+1) = i_{n+1} | X(n) = i_n, ..., X(0) = i_0] \\ &= P(X(n+m) = i_{n+m} | X(n+m-1) = i_{n+m-1}, ..., X(0) = i_0) \quad \times \\ &P[X(n+m-1) = i_{n+m-1} | X(n+m-2) = i_{n+m-2}, ..., X(0) = i_0] \quad \times \quad \cdots \\ &\times P(X(n+1) = i_{n+1} | X(n) = i_n, ..., X(0) = i_0) \end{aligned}$$

for all $m = 1, 2, ...$. But from the markov property the right hand side of the above equation becomes

$$P[X(n+m) = i_{n+m}|X(n+m-1) = i_{n+m-1}]\times$$
$$P[X(n+m-1) = i_{n+m-1}|X(n+m-2) = i_{n+m-2}] \times ... \times$$
$$P[X(n+1) = i_{n+1}|X(n) = i_n]$$

Therefore we have

$$P[X(n+m) = i_{n+m}, ..., X(n+1) = i_{n+1}X(n) = i_n, ..., X(0) = i_0] = P[X(n+m) = i_{n+m}, ..., X(n+1) = i_{n+1}|X(n) = i_n]$$

From here it follows that

$$P(X(n+m) = i_{n+m}|X(n) = i_n, ..., X(0) = i_0) = P(X(n+m) = i_{n+m}|X(n) = i_n) \;\; for \;\; all \;\; m \geq 1.$$

Thus , for a markov chain once the present state is known , prediction of future distributions cannot be improved by adding any knowledge of the past. (2) The process is completely determined once the initial probabaility distribution and the transition probabilities are known . We shall now prove this fact.

$$P(X(n) = i_n, X(n-1) = i_{n-1}, ..., X(1) = i_1, X(0) = i_0)$$
$$= P(X(n) = i_n|X(n-1) = i_{n-1}, ..., X(0) = i_0) \times P(X(n-1) = i_{n-1}, ..., X(0) = i_0)$$
$$= P_{i_{n-1}, i_n} \times P(X(n-1) = i_{n-1}, ..., X(0) = i_0) \;\; [using \;\; markov \;\; property]$$

Thus,

$$P(X(n) = i_n, X(n-1) = i_{n-1}, ..., X(0) = i_0) = P(X(0) = i_0)P_{i_0,i_1}P_{i_1,i_2} ... \times P_{i_{n-1},i_n}.$$

Notice that for a non homogeneous markov chain $X(n)$, $n = 0, 1, 2, ...$, if we write

$$P(X(n+1)=j|X(n)=i)=P_{n+1}(i,j)$$

where P_{n+1} is a $m\times m$ stochastic matrix , and also , is the transition probability matrix of the chain $X(n)$ at time n , then

$$\begin{aligned}&P(X(n)=i_n,X(n-1)=i_{n-1},...,X(0)=i_0)\\&=P(X(0)=i_0)P_1(i_0,i_1)P_2(i_1,i_2)...\times P_n(i_{n-1},i_n).\end{aligned}$$

CHAPTER III

RESULTS IN HOMOGENEOUS CASE

Motivation of Burke and Rosenblatt

In [2], Burke and Rosenblatt gave a number of theorems regarding collapsed markov chains. In this chapter we shall briefly discuss some of their results.

Let us assume that $X(n)$, $n = 0,1,2,...$ be a markov chain with a finite number of states $1,2,...,m$ and stationary transition probability matrix $P = p_{ij}$ such that

$p_{ij} = P(X(n+1) = j|X(n) = i) \geq 0$ for $i, j = 1,2,...,m$

where for each $i \geq 1$, we have $\sum_j p_{ij} = 1$

Suppose the experimenter does not observe the process $X(n)$ but rather a derived process $Y(n) = f(X(n))$ where f is a given function on $1,2,...,m$. The states i of the original process $X(n)$ on which f equals some fixed constant are collapsed into a single state of the new process $Y(n)$. Let us call these collapsed sets of states S_i , $i = 1,2,...,r$, $r \leq m$. Thus if $f(S_i) = i$, $1 \leq i \leq r$, then $Y(n) = i$ if and only if $X(n) \in S_i$. The obvious question that needs to be addressed is whether or not the new chain is markovian. Clearly this new chain is not markovian in general.

Let us restrict ourselves to a process $X(n)$ with its initial probability distribution a left invariant vector p of the matrix P, that is $pP = p$. We further assume that all the components of p are positive . Let D be the diagonal matrix with i^{th} diagonal entry p_i . The process is said to be reversible if

$$DP = P^T D$$

Results of Burke and Rosenblatt

Theorem 1

Let $X(n)$ be a stationary reversible process with $p_i > 0$ for all i . Then $Y(n)$ is markovian if and only if for any fixed $\beta = 1, 2, ..., r$

$$\sum_{j \in S_\beta} p_{ij} = P(X(n+1) \in S_\beta | X(n) = i) = C_{S_\alpha, S_\beta}$$

has the same value for all i in any given collapsed set of states S_α , $\alpha = 1, 2, ..., r$.

Theroem 2

Let f be a function that collapses only one class of states S containing more than one state . Then $Y(n)$ is markovian for all initial distributions w of $X(n)$ if and only if one of the following two conditions is satisfied:

(1) $\sum_{l \in S} p_{kl} p_{lu} = p_{k,S} C_u$ for all $u \notin S$ and all k

(2) $p_{i,S} = 0$ for all $i \notin S$

where

$$p_{k,S} = \sum_{j \in S} p_{kj} = P[X(n+1) \in S | X(n) = k)$$

Let us now consider the class of stationary markov chains $X(n)$ with $p_i > 0$, $i = 1, 2, ..., m$ such that $Y(n) = f(X(n))$ is markovian for any many to one transformation f .

Theorem 3

Let $X(n)$ be a stationary markov chain with $p_i > 0$, $i = 1, 2, ..., m$. $f(X(n))$ is markovian for every many to one transformation f if and only if the transition probability matrix P of $X(n)$ is of the form

$$P = \alpha I + (1 - \alpha) U$$

where U is a matrix with identical rows and α is a real number.

Discussion on The First Theorem

We observe that the statement of the first theorem is both sufficient and necessary. Here we shall discuss the proof of the necessary part of the result.

Proof

Burke and Rosenblatt constructed matrices A and B as follows :

$$A = (B^T DB)^{-1} B^T D$$

and

$$B_{ij} \quad = \begin{cases} 1 & if \quad i \in S_j \\ 0 & if \quad i \notin S_j \end{cases}$$

where $(D)_{ii} = p_i$ for all i .

Follwing is the proof of the equality $Q = APB$ used in [2], where P and Q are the transition probability matrices of $X(n)$ and $Y(n)$ respectively.

First we shall write the matrix A elementwise.

$A_{ik} = (B^T DB)^{-1}_{ii} B^T_{ik} D_{kk}$ (since $B^T DB$ is diagonal)

$(B^T DB)_{ii} = \sum_u B^T_{iu} D_{uu} B_{ui} = \sum_{u \in S_i} D_{uu} = \sum_{u \in S_i} p_u$

Hence $A_{ik} = \dfrac{p_k}{\sum_{u \in S_i} p_u}$ (where $k \in S_i$)

Now, $(APB)_{ij} = \sum_{k,l} A_{ik} P_{kl} B_{lj} = \sum_{k \in S_i} \sum_{l \in S_j} \dfrac{p_k P_{kl}}{(\sum_{u \in S_i} p_u)}$ (A)

Again,

$$Q_{ij} = P(Y(n) = j | Y(n-1) = i)$$
$$= \frac{\sum_{u \in S_i} P(X(n) \in S_j, X(n-1) = u)}{\sum_{u \in S_i} P(X(n-1) = u)}$$

Let $p_n(j) = P(X(n) = j) = \sum_k P(X(n) = j | X(n-1) = k) P(X(n-1) = k)$

Let $n = 1$. Then we have ,

$p_1(j) = \sum_k p_0(k) P(k, j) = p_0(j)$ (since p is left invariant of the transition matrix P)

Clearly by induction we have , $p_n(j) = p_0(j)$ $\forall j$ and $\forall n$.

Then,

$$Q_{ij} = \frac{\sum_{k \in S_i} P(k, S_j) p_k}{\sum_{u \in S_i} p_u} = \frac{1}{\sum_{u \in S_i} p_u} \sum_{k \in S_i} \sum_{l \in S_j} p_k P(k, l) \quad \text{(B)}$$

Hence from equations (A) and (B) , we have , $Q = APB$.

Here we shall prove another equality given by $Q^2 = AP^2B$

$$Q^2_{ij} = P(Y(2) = j | Y(0) = i)$$
$$= \frac{P(X(2) \in S_j, X(0) \in S_i)}{P(X(0) \in S_i)}$$
$$= \frac{\sum_{k \in S_i} \sum_l P(X(2) \in S_j, X(1) = l, X(0) = k)}{\sum_{k \in S_i} P(X(0) = k)}$$
$$= \frac{1}{\sum_{k \in S_i} p_k} \sum_{k \in S_i} \sum_l p_k P(k, l) P(l, S_j)$$
$$= \frac{\sum_{k \in S_i} \sum_l \sum_{t \in S_j} p_k P(k, l) P(l, t) B_{tj}}{\sum_{k \in S_i} p_k} = (AP^2B)_{ij}$$

Let $Y(n)$ be markovian , hence the n- step transtition probability matrix of $Y(n)$ given by $Q^{(n)}$ must satisfy Chapman Kolmogorov Equation. Therefore we have,

$$Q^{(2)} = \left[Q^{(1)}\right]^2$$

$$\implies AP^2B = APBAPB$$

$$\implies AP^2B - APBAPB = 0$$

$$\implies AP[I - BA]PB = 0$$

$$\implies \left(B^T DB\right)^{-1} B^T DP\left(I - BA\right)PB = 0$$

$$\implies B^T P^T D\left(I - BA\right)PB = 0 \text{ [because of reversibility]}$$

Burke and Rosenblatt stated in their paper (see [2]) that $D(I - BA)$is positive definite. Here we shall give an example which will show that the matrix may not be positive definite.

Example

Let $X(n)$ be the original markov chain with state space $S = \{1,2,3,4\}$. Let $\{S_1, S_2\}$ be a partition of S. Let $S_1 = \{1,2\}$ and $S_2 = \{3,4\}$. Hence following the construction given by the authors,

$$B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } A = \begin{pmatrix} \frac{p_1}{p_1+p_2} & \frac{p_2}{p_1+p_2} & 0 & 0 \\ 0 & 0 & \frac{p_3}{p_3+p_4} & \frac{p_4}{p_3+p_4} \end{pmatrix}$$

$$\text{Hence } BA = \begin{pmatrix} \frac{p_1}{p_1+p_2} & \frac{p_2}{p_1+p_2} & 0 & 0 \\ \frac{p_1}{p_1+p_2} & \frac{p_2}{p_1+p_2} & 0 & 0 \\ 0 & 0 & \frac{p_3}{p_3+p_4} & \frac{p_4}{p_3+p_4} \\ 0 & 0 & \frac{p_3}{p_3+p_4} & \frac{p_4}{p_3+p_4} \end{pmatrix}$$

$$I - BA = \begin{pmatrix} \frac{p_2}{p_1+p_2} & \frac{-p_2}{p_1+p_2} & 0 & 0 \\ \frac{-p_1}{p_1+p_2} & \frac{p_1}{p_1+p_2} & 0 & 0 \\ 0 & 0 & \frac{p_4}{p_3+p_4} & \frac{-p_4}{p_3+p_4} \\ 0 & 0 & \frac{-p_3}{p_3+p_4} & \frac{p_3}{p_3+p_4} \end{pmatrix}$$

$$\implies D(I-BA) = \begin{pmatrix} p_1 & 0 & 0 & 0 \\ 0 & p_2 & 0 & 0 \\ 0 & 0 & p_3 & 0 \\ 0 & 0 & 0 & p_4 \end{pmatrix} \times (I-BA) = \begin{pmatrix} \frac{p_1p_2}{p_1+p_2} & \frac{-p_1p_2}{p_1+p_2} & 0 & 0 \\ \frac{-p_1p_2}{p_1+p_2} & \frac{p_1p_2}{p_1+p_2} & 0 & 0 \\ 0 & 0 & \frac{p_3p_4}{p_3+p_4} & \frac{-p_3p_4}{p_3+p_4} \\ 0 & 0 & \frac{-p_3p_4}{p_3+p_4} & \frac{p_3p_4}{p_3+p_4} \end{pmatrix}$$

Clearly the top left corner 2×2 block of the above matrix has determinant 0, i.e the matrix is not positive-definite.

Here we will show that $D(I-BA)$ is semidefinite instead. And essentially the semidefiniteness property is sufficient for Burke-Rosenblatt's proof to go through.

$[D(I-BA)]_{ij}$

$= p_i(1-\frac{p_i}{P_\alpha})$ if $i = j$ and both i , j in S_α

$= -\frac{p_ip_j}{P_\alpha}$ if $i \neq j$, and both i , j in S_α

$= 0$ if $i \in S_\alpha$, $j \in S_\beta$ and $\alpha \neq \beta$ [where $P_\alpha = \sum_{k\in S_\alpha} p_k$]

Thus ,

$x[D(I-BA)]x^T$ [where $x \neq 0$ is a non zero vector]

$= \sum_{\alpha=1}^{r} [\sum_{i\in S_\alpha} x_i^2 p_i(1-\frac{p_i}{P_\alpha}) + \sum_{i\neq j, i\in S_\alpha, j\in S_\alpha} -x_ix_j\frac{p_ip_j}{P_\alpha}]$

$= \sum_{\alpha=1}^{r} \frac{1}{P_\alpha} \sum_{i<j, i\in S_\alpha, j\in S_\alpha} p_ip_j(x_i-x_j)^2 \geq 0$, hence semi definite .

Since $D(I-BA)$is semi-definite hence there exist matrix $R\,(m\times m)$such that $D(I-BA) = R^TR$.

Therefore we have,

$$B^TP^TD(I-BA)PB = 0$$

$$\implies B^TP^TR^TRPB = 0$$

$$\implies (RPB)^T RPB = 0$$

$$\implies RPB = 0$$

$\implies R^T RPB = 0$

$\implies D(I - BA)PB = 0$

$\implies PB = BAPB$ - this is essentially what Burke and Rosenblatt need to show for the statement of the necessary part of their first theorem to be true.

Introduction to Some Resutls of Kemeny and Snell

In their book “Finite Markov Chains”, Kemeny and Snell gave certain results regarding lumpability of markov chains in the homogeneous context. Let us first discuss some of the terms necessary to have a grasp of their results:

Strong Lumpability

We shall say that a markov chain is strongly lumpable with respect to a partition $A = \{A_1, A_2, ..., A_r\}$ of the chain's state space if for every starting vector π the collapsed chain corresponding to this partition is a markov chain and the transition probabilities do not depend upon the choice of π .

Weak Lumpability

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 π . That is , in this case the collapsed chain will be markov only when some particular initial vectors (it may even be just one initial vector) are chosen.

Results of Kemeny and Snell Which Will be Needed Later

Theorem 1

For a reversible regular markov chain weak lumpability implies strong lumpability.

Theorem 2

A reversible regular markov chain is reversible when lumped.

Theorem 3

If a given process is weakly lumpable with respect to a partition A of the state space , then so is the reverse process.

CHAPTER IV

NON HOMOGENEOUS CONTEXT

Suppose $X(n)$ $n = 1, 2, ...$ is a non homogeneous markov chain with state space $S = \{1, 2, ..., m\}$. Let $S = S_1 \cup S_2 \cup ... \cup S_r$, $r \leq m$, i.e $\{S_1, S_2, ..., S_r\}$ gives a partition of S . We define $Y(n) = i$ if and only if $X(n) \in S_i$ where $1 \leq i \leq r$.

Burke-Rosenblatt Analogues

In this section we shall present results similar to those of Burke-Rosenblatt in the non-homogeneous setting .

Theorem 1(A)

Let X(n), n= $1, 2, ...$ be a non-homogeneous markov chain with finite state-space S= $\{1, 2, ..., m\}$ and with transition probability matrices P_n such that

$$P(X(n) = j | X(n-1) = i) = P_n(i, j) \ \forall i, j \in S \ .$$

Suppose $S_1 \cup S_2 \cup ... \cup S_r = S$ $(r \leq m)$, that is $\{S_1, S_2, ..., S_r\}$is a partition of S.

We define $Y(n) = i$ if and only if $X(n) \in S_i$ $1 \leq i \leq r$.

Then a sufficient condition for $Y(n)$ to be a non-homogeneous markov chain is that for all α, β

$$\sum_{j \in S_\beta} P_n(i, j) = Q_n(\alpha, \beta) \ \forall i \in S_\alpha \ ,$$

i.e $\sum_{j \in S_\beta} P_n(i, j)$ is independent of $i \in S_\alpha$, where Q_n's are the transition probability matrices of $Y(n)$, i.e

$$Q_n(\alpha, \beta) = P(Y(n) = \beta | Y(n-1) = \alpha) \ , n = 1, 2, ...$$

Proof

Let $\{i_0, i_1, ..., i_n\} \subset \{1,2,...,r\}$and $n \leq r$. Then we have,

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

$= \sum_{\alpha_{i_0} \in S_{i_0}} \cdots \sum_{\alpha_{i_{n-1}} \in S_{i_{n-1}}} \sum_{\alpha_{i_n} \in S_{i_n}} P(X(n) = \alpha_{i_n}, ..., X(0) = \alpha_{i_0})$

$= \sum_{\alpha_{i_0} \in S_{i_0}} \cdots \sum_{\alpha_{i_{n-1}} \in S_{i_{n-1}}} \sum_{\alpha_{i_n} \in S_{i_n}} P(X(0) = \alpha_{i_0}) P_1(\alpha_{i_0}, \alpha_{i_1}) ... P_n(\alpha_{i_{n-1}}, \alpha_{i_n})$

$= \sum_{\alpha_{i_0} \in S_{i_0}} \cdots \sum_{\alpha_{i_{n-1}} \in S_{i_{n-1}}} P(X(0) = \alpha_{i_0}) ... P_{n-1}(\alpha_{i_{n-2}}, \alpha_{i_{n-1}}) P_n(\alpha_{i_{n-1}}, S_{i_n})$

$= \sum_{\alpha_{i_0} \in S_{i_0}} \cdots \sum_{\alpha_{i_{n-1}} \in S_{i_{n-1}}} P(X(0) = \alpha_{i_0}) ... P_{n-1}(\alpha_{i_{n-2}}, \alpha_{i_{n-1}}) Q_n(i_{n-1}, i_n)$

=continuing the process

$$= P(X(0) \in S_{i_0}) Q_1(i_0, i_1) ... Q_n(i_{n-1}, i_n) \qquad (1)$$

Now,

$$\begin{aligned} & P(Y(n) = i_n | Y(n-1) = i_{n-1}, ..., Y(0) = i_0) \\ &= \frac{P(Y(n) = i_n, Y(n-1) = i_{n-1}, ..., Y(0) = i_0)}{P(Y(n-1) = i_{n-1}, ..., Y(0) = i_0)} \\ &= \frac{P(X(0) \in S_{i_0}) Q_1(i_0, i_1) ... Q_n(i_{n-1}, i_n)}{P(X(0) \in S_{i_0}) Q_1(i_0, i_1) ... Q_{n-1}(i_{n-2}, i_{n-1})} \text{ [from equation (1)]} \\ &= Q_n(i_{n-1}, i_n) \end{aligned}$$

Hence the condition mentioned in the theorem is sufficient for $Y(n)$ to be markov.

Theorem 1(B)

Let $X(n)$ be a non homogeneous markov chain with state space $S = \{1,2,...,m\}$ and let the left invariant positive vector p be the distribution of $X(0)$. We also assume that $X(n)$ is reversible , i.e $DP_n = P_{n+1}^T D \; \forall n \geq 1$, where D is a diagonal matrix with i^{th} diagonal entry p_i and P_n is the transition probability matrix of $X(n)$. Let $S = S_1 \cup S_2 \cup ... \cup S_r$, where $r \leq m$. We define $Y(n) = i$ if and only if $X(n) \in S_i$. Let $Y(n)$ be markovian . Then the following condition holds:

If S_u and S_v are any two partitioning sets of S , then for each $n \geq 1$,

$P_n(i,S_v) = Q_n(u,v) = P(Y(n) = v|Y(n-1) = u)$ which is independent of i in S_u .

Notice that the reversiblity condition implies that for all i , j in S and for all $n \geq 1$,

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

i.e $p_i(P_n)_{ij} = p_j(P_{n+1})_{ji} = p_i(P_{n+2})_{ij}$, i.e $P_n = P_{n+2}$.

We prove the theorem in several steps .

Step 1 :

Lemma 1

$Q_n = AP_nB \ \forall n$, where $A = (B^T D_n B)^{-1} B^T D_n$ and

$$B_{ij} = \begin{cases} 1 & if \quad i \in S_j \\ 0 & if \quad i \notin S_j \end{cases}$$

where $(D_n)_{ii} = p_i \ \forall n$.

Proof

Since $Y(n)$is markov, hence its transition probability matrices are given by:

$$Q_n(i,j) = P(Y(n) = j|Y(n-1) = i)$$

$$\implies Q_n(i,j) = P(X(n) \in S_j|X(n-1) \in S_i)$$

$$\implies Q_n(i,j) = \frac{P(X(n) \in S_j, X(n-1) \in S_i)}{P(X(n-1) \in S_i} \tag{2}$$

Let $\pi_n(j) = P(X(n) = j)$

Then,

$$\pi_n(j) = \sum_{k=1}^{m} P(X(n) = j|X(n-1) = k)P(X(n-1) = k)$$

$$\implies \pi_n(j) = \sum_{k=1}^{m} \pi_{n-1}(k)P_n(k,j)$$

Let $n = 1$.

Then $\pi_1(j) = \sum_{k=1}^{m} \pi_0(k) P_1(k,j)$

$$= \sum_{k=1}^{m} p_k P_1(k,j) = p_j \ \forall j$$

Clearly by induction $\forall n \geq 1$, $\pi_n(j) = p_j \ \forall j$.

Thus from (2)

$$Q_n(i,j) = \frac{P(X(n) \in S_j, X(n-1) \in S_i)}{P(X(n-1) \in S_i)}$$

$$\implies Q_n(i,j) = \frac{\sum_{k \in S_i} P(X(n) \in S_j, X(n-1) = k)}{\sum_{k \in S_i} P(X(n-1) = k)}$$

$$\implies Q_n(i,j) = \frac{\sum_{k \in S_i} P(X(n) \in S_j | X(n-1) = k) P(X(n-1) = k)}{\sum_{k \in S_i} P(X(n-1) = k)}$$

$$\implies Q_n(i,j) = \frac{\sum_{k \in S_i} P(X(n) \in S_j | X(n-1) = k) p_k}{\sum_{k \in S_i} p_k} \tag{3}$$

Now,

$A_{ik} = (B^T D_n B)^{-1}_{ii} B^T_{ik} (D_n)_{kk} = \dfrac{p_k}{\sum_{j \in S_i} p_j} \ k \in S_i$

Hence we have

$$(AP_nB)_{ij} = \sum_{k,l} A_{ik} (P_n)_{kl} B_{lj}$$

$$\implies (AP_nB)_{ij} = \sum_{k \in S_i, l \in S_j} \frac{p_k}{(\sum_{j \in S_i} p_j)} (P_n)_{kl}$$

$$\implies (AP_nB)_{ij} = \frac{1}{\sum_{j \in S_i} p_j} \sum_{k \in S_i} p_k P_n(k, S_j)$$

$$\implies (AP_nB)_{ij} = \frac{\sum_{k \in S_i} P(X(n) \in S_j | X(n-1) = k) p_k}{\sum_{j \in S_i} p_j}$$

$$\implies (AP_nB)_{ij} = Q_n(i,j) \text{ [from (3)]}$$

Lemma 2

$Q_nQ_{n+1} = AP_nP_{n+1}B$

Proof

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

Step 2 :

Now we are ready to prove our theorem $1(B)$ .Let $Y(n)$ be non-homogeneous Markov Chain. We know,

$AP_nP_{n+1}B = Q_nQ_{n+1} = AP_nBAP_{n+1}B$ (from lemma 1 and lemma 2)

$\implies AP_n(I-BA)P_{n+1}B = 0$

$\implies (B^TDB)^{-1}B^TDP_n(I-BA)P_{n+1}B = 0$

$\implies B^TP_{n+1}^TD(I-BA)P_{n+1}B = 0$ [from reversibility]

$\implies B^TP_{n+1}^TR^TRP_{n+1}B = 0$ [since $D(I-BA)$ is semidefinite]

$$\implies RP_{n+1}B = 0$$

$$\implies R^T RP_{n+1}B = 0$$

$$\implies D(I - BA)P_{n+1}B = 0$$

$$\implies P_{n+1}B = BAP_{n+1}B \implies P_{n+1}B = BQ_{n+1}$$

Now, BQ_{n+1} is an $m \times r$ matrix and for any two partitioning sets S_u and S_v with $i \in S_u$, the $i - vth$ element of BQ_{n+1} is simply $Q_{n+1}(u,v)$; on the other hand , the $i - vth$ element of $P_{n+1}B$ is $P_{n+1}(i,S_v)$. Thus for each $n \geq 1$, $P_n(i,S_v) = Q_n(u,v)$ for each $i \in S_u$. Thus the proof of the theorem is complete .

Theorem 2

Suppose that there is only one set S which collapses more than one state. Then $Y(n)$ is Markovian if one of the following two conditions hold:

1. for all $n \geq 1$,

$\sum_{l \in S} P_n(k,l)P_{n+1}(l,u) = P_n(k,S)C_{n+1}(u)$ $\forall k$ and $\forall u \notin S$.

2. $P_n(i,S) = 0$ $\forall i \notin S$, $\forall n \geq 1$

where $C_{n+1}(u) = P(X(n+1) = u | X(n) \in S)$.

Conversely , when $Y(n)$ is markovian at least one of the two conditions (condition (2) and condition (1) , not for every k , but for $k \notin S$) must hold. [Here we remark that when $k \notin S$, k is a singleton and as such , $X(n) = k$ if and only if $Y(n) = k$. This is though simple , a useful observation in the proof of the theorem] .

Proof

Let $Y(n)$ be Markovian and there exists $k \notin S$ such that $P_n(k,S) \neq 0$ for some $n \geq 1$. We prove below that condition (1) holds for this n and for this k. This proof works whenever $P_n(k,S)$ is non zero $\forall k \notin S$. In particular, for all n and all states $k \notin S$ such that $P_n(k,S) = 0$ condition (1) holds trivially as both sides of (1) are then zero.

Now,

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

Then,

$$\frac{\sum_{l\in S} P_n(k,l)P_{n+1}(l,u)}{P_n(k,S)}$$
$$=\frac{P(X(n-1)=k)\sum_{l\in S} P_n(k,l)P_{n+1}(l,u)}{P(X(n-1)=k)P_n(k,S)}$$
$$=\frac{\sum_{l\in S} P(X(n+1)=u, X(n)=l, X(n-1)=k)}{P(X(n)\in S, X(n-1)=k)}$$
$$=P(X(n+1)=u|X(n)\in S, X(n-1)=k)$$
$$=P(Y(n+1)=u|Y(n)=S, Y(n-1)=k)$$
$$=P(Y(n+1)=u|Y(n)=S) \text{ [since } Y(n) \text{ is markov]}$$
$$=P(X(n+2)=u|X(n+1)\in S)$$
$$=C_{n+1}(u)$$

Now conversely, let us assume that $\forall i \notin S$, $P_n(i,S)=0$ $\forall n$. Here we have two possibilities:

Case 1

Let $i_k = S$ for $k=0$ only. Then,

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

Case 2

Let $i_k = S$ for largest k such that $0 < k \leq n$. Then i_j must also be in S $\forall j$ where $0 \leq j < k$, otherwise $P_n(i_j, i_k)$ will be 0 from our assumption $\forall n$. In this case,

$P(Y(n) = i_n, Y(n-1) = i_{n-1}, ..., Y(k+1) = i_{k+1}, Y(k) = S, ..., Y(0) = S)$

$= P(X(n) = i_n, X(n-1) = i_{n-1}, ..., X(k+1) = i_{k+1}, X(k) \in S, ..., X(0) \in S) +$
$P(X(n) = i_n, X(n-1) = i_{n-1}, ..., X(k+1) = i_{k+1}, X(k) \in S, ..., X(0) \notin S)$

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

=continuing the process

$$= P(X(n) = i_n, X(n-1) = i_{n-1}, ..., X(k+1) = i_{k+1}, X(k) \in S) \quad (4)$$

Now,

$P(Y(n) = i_n | Y(n-1) = i_{n-1}, ..., Y(k+1) = i_{k+1}, Y(k) = S, ..., Y(0) = S)$
$= \dfrac{P(X(n) = i_n, X(n-1) = i_{n-1}, ..., X(0) \in S)}{P(X(n-1) = i_{n-1}, ..., X(0) \in S)}$
$= \dfrac{P(X(n) = i_n, X(n-1) = i_{n-1}, ..., X(k) \in S)}{P(X(n-1) = i_{n-1}, ..., X(k) \in S)}$ [from (4)]
$= P(X(n) = i_n | X(n-1) = i_{n-1}, ..., X(k) \in S)$ (Thereby reduced to the previous case)

Hence $Y(n)$ is Markovian.

Now let us assume that for all $n \geq 0$, for all k and for all $u \notin S$, condition (1) holds, i.e

$\sum_{l \in S} P_n(k,l) P_{n+1}(l,u) = P_n(k,S) C_{n+1}(u)$ holds.

We need to prove when (1) holds ,

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

We write $S_{i_k} = i_k$ when i_k is a singleton and $S_{i_k} = S$ when $i_k = S$.

For each i_j above, there are two possibilities : 1) $i_j \notin S$ and 2) $i_j \in S$. We have three possible cases :

Case 1

$i_{n-1} = S$, $i_n \notin S$ and hence i_n is a singleton.

In this case,

$$P(Y(n) = i_n | Y(n-1) = i_{n-1}, ..., Y(0) = i_0)$$
$$= \frac{\sum_{\alpha_{i_n} \in S_{i_n}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P(X(n) = \alpha_{i_n}, ..., X(0) = \alpha_{i_0})}{\sum_{\alpha_{i_{n-1}} \in S_{i_{n-1}}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P(X(n-1) = \alpha_{i_{n-1}}, ..., X(0) = \alpha_{i_0})}$$
$$= \frac{\sum_{\alpha_{i_n} \in S_{i_n}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P_n(\alpha_{i_{n-1}}, \alpha_{i_n}) P_{n-1}(\alpha_{i_{n-2}}, \alpha_{i_{n-1}}) P(X(n-2) = \alpha_{i_{n-2}}, ..., X(0) = \alpha_{i_0})}{\sum_{\alpha_{i_{n-1}} \in S_{i_{n-1}}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P_{n-1}(\alpha_{i_{n-2}}, \alpha_{i_{n-1}}) P(X(n-2) = \alpha_{i_{n-2}}, ..., X(0) = \alpha_{i_0})}$$
$$= \frac{\sum_{\alpha_{i_{n-1}} \in S_{i_{n-1}}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P_n(\alpha_{i_{n-1}}, i_n) P_{n-1}(\alpha_{i_{n-2}}, \alpha_{i_{n-1}}) P(X(n-2) = \alpha_{i_{n-2}}, ..., X(0) = \alpha_{i_0})}{\sum_{\alpha_{i_{n-1}} \in S_{i_{n-1}}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P_{n-1}(\alpha_{i_{n-2}}, \alpha_{i_{n-1}}) P(X(n-2) = \alpha_{i_{n-2}}, ..., X(0) = \alpha_{i_0})}$$
$$= \frac{C_n(i_n) \sum_{\alpha_{i_{n-2}} \in S_{i_{n-2}}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P_{n-1}(\alpha_{i_{n-2}}, S) P(X(n-2) = \alpha_{i_{n-2}}, ..., X(0) = \alpha_{i_0})}{\sum_{\alpha_{i_{n-2}} \in S_{i_{n-2}}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P_{n-1}(\alpha_{i_{n-2}}, S) P(X(n-2) = \alpha_{i_{n-2}}, ..., X(0) = \alpha_{i_0})}$$
$$= C_n(i_n) = P(X(n) = i_n | X(n-1) \in S)$$

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

Case 2

$i_{n-1} \notin S$, $i_n \in S$.

$$P(Y(n) = i_n | Y(n-1) = i_{n-1}, ..., Y(0) = i_0)$$
$$= \frac{\sum_{\alpha_{i_n} \in S_{i_n}} \sum_{\alpha_{i_{n-2}} \in S_{i_{n-2}}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P(X(n) = \alpha_{i_n}, X(n-1) = i_{n-1}, ..., X(0) = \alpha_{i_0})}{\sum_{\alpha_{i_{n-2}} \in S_{i_{n-2}}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P(X(n-1) = i_{n-1}, ..., X(0) = \alpha_{i_0})}$$
$$= \frac{\sum_{\alpha_{i_n} \in S_{i_n}} \sum .. \sum_{\alpha_{i_0} \in S_{i_0}} P(X(n) = \alpha_{i_n} | X(n-1) = i_{n-1}) P(X(n-1) = i_{n-1}, .., X(0) = \alpha_{i_0})}{\sum_{\alpha_{i_{n-2}} \in S_{i_{n-2}}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P(X(n-1) = i_{n-1}, ..., X(0) = \alpha_{i_0})}$$
$$= \frac{\displaystyle\sum_{\alpha_{i_n} \in S_{i_n}} P(X(n) = \alpha_{i_n} | X(n-1) = i_{n-1})}{\sum_{\alpha_{i_{n-2}} \in S_{i_{n-2}}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P(X(n-1) = i_{n-1}, ..., X(0) = \alpha_{i_0})} \times$$
$$\sum_{\alpha_{i_{n-2}} \in S_{i_{n-2}}} .. \sum_{\alpha_{i_0} \in S_{i_0}} P(X(n-1) = i_{n-1}, .., X(0) = \alpha_{i_0})$$
$$= \sum_{\alpha_{i_n} \in S_{i_n}} P(X(n) = \alpha_{i_n} | X(n-1) = i_{n-1})$$

$$= P(X(n) \in S | X(n-1) = i_{n-1})$$

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

Case 3

$i_n = S$ and $i_{n-1} = S$.

Let us first observe that

$$\sum_{j \in S} P_n(k,j) P_{n+1}(j,S)$$

$= \sum_{j \in S} P_n(k,j) - \sum_{j \in S} P_n(k,j) P_{n+1}(j,S^c)$

$= \sum_{j \in S} P_n(k,j) - P_n(k,S) C_{n+1}(S^c)$

$= P_n(k,S)\,(1 - C_{n+1}(S^c))$

$= P_n(k,S) C_{n+1}(S)$

Then we have ,

$$P(Y(n) = i_n | Y(n-1) = i_{n-1}, ..., Y(0) = i_0)$$
$$= \frac{\sum_{\alpha_{i_n} \in S_{i_n}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P(X(n) = \alpha_{i_n}, ..., X(0) = \alpha_{i_0})}{\sum_{\alpha_{i_{n-1}} \in S_{i_{n-1}}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P(X(n-1) = \alpha_{i_{n-1}}, ..., X(0) = \alpha_{i_0})}$$
$$= \frac{\sum_{\alpha_{i_n} \in S_{i_n}} .. \sum_{\alpha_{i_0} \in S_{i_0}} P_n(\alpha_{i_{n-1}}, \alpha_{i_n}) P_{n-1}(\alpha_{i_{n-2}}, \alpha_{i_{n-1}}) P(X(n-2) = \alpha_{i_{n-2}}, .., X(0) = \alpha_{i_0})}{\sum_{\alpha_{i_{n-1}} \in S_{i_{n-1}}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P_{n-1}(\alpha_{i_{n-2}}, \alpha_{i_{n-1}}) P(X(n-2) = \alpha_{i_{n-2}}, ..., X(0) = \alpha_{i_0})}$$
$$= \frac{C_n(S) \sum_{\alpha_{i_{n-2}} \in S_{i_{n-2}}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P_{n-1}(\alpha_{i_{n-2}}, S) P(X(n-2) = \alpha_{i_{n-2}}, .., X(0) = \alpha_{i_0})}{\sum_{\alpha_{i_{n-2}} \in S_{i_{n-2}}} \cdots \sum_{\alpha_{i_0} \in S_{i_0}} P_{n-1}(\alpha_{i_{n-2}}, S) P(X(n-2) = \alpha_{i_{n-2}}, .., X(0) = \alpha_{i_0})}$$
$$= C_n(S)$$

$= P(X(n) \in S | X(n-1) \in S)$

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

Here we choose not to consider the last remaining case , namely when $i_n \notin S$ and $i_{n-1} \notin S$, as this case is very simple . Thus in all the possible cases $Y(n)$ is Markovian. This completes the proof of theorem 2 .

In the above thereom there was only one set which had more than one states collapsed into it (analogous to Burke-Rosenblatt second theroem) . In the following theorem we shall explore a more general case where there are two disjoint sets , each of which collapse more than one state.

Theorem 3

Let $X(n)$ be a finite non homogeneous markov chain. Suppose that there are only two disjoint sets S_1 and S_2 which collapse more than one state. Suppose that for each $n \geq 1$,

$P_n(x,S_1) = 0$ for all $x \in S_2$ and

$P_n(x,S_2) = 0$ for all $x \in S_1$

The following conditions are also assumed:

1. $\sum_{l \in S_1} P_n(k,l)P_{n+1}(l,u) = P_n(k,S_1)C^{(1)}_{n+1}(u)$ for all k and for all $u \notin S_1$where

$C^{(1)}_{n+1}(u) = P(X(n+1) = u|X(n) \in S_1)$ for all $n \geq 1$

2. $\sum_{l \in S_2} P_n(k,l)P_{n+1}(l,u) = P_n(k,S_2)C^{(2)}_{n+1}(u)$ for all k and for all $u \notin S_2$ where

$C^{(2)}_{n+1}(u) = P(X(n+1) = u|X(n) \in S_2)$ for all $n \geq 1$

3. For $k \notin S_1$, $P_n(k,S_1) = 0$ for all $n \geq 1$

And for $k \notin S_2$, $P_n(k,S_2) = 0$ for all $n \geq 1$

1\`. This is condition (1) , when $k \notin S_1 \cup S_2$

2\`. This is condition (2) , when $k \notin S_1 \cup S_2$

Then $Y(n)$ is markovian if either condition (3) holds or condition (1) and (2) hold. Conversely , if $Y(n)$ is markovian , then either (3) holds or condition $(1\`)$ and $(2\`)$ hold.

Proof

We prove the theorem in several steps.

Step 1: We assume that $Y(n)$ is markovian. Let $k \notin S_1 \cup S_2$. If condition (3) does not hold , then there is $n \geq 1$ such that $P_n(k, S_1 \cup S_2) > 0$. Thus either $P_n(k, S_1) > 0$ or $P_n(k, S_2) > 0$. Let $P_n(k, S_1) > 0$. Let $u \notin S_1$. Notice that whenever $P_n(k, S_1) = 0$, condition (1`) holds for this $k$ and this $n$ and whenever $P_n(k, S_2) = 0$ , condition (2`) holds for this n and this k.

Now we show that if $P_n(k, S_1) > 0$ and $k \notin S_1 \cup S_2$, then condition (1`) holds with $u \notin S_1$.

We have,

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

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

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

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

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

$= P(Y(n+1)=u | Y(n)=S_1, Y(n-1)=k)$ [$k \notin S_1 \cup S_2 \implies X(n-1)=k$ iff $Y(n-1)=k$. Also if $u \in S_2$, condition (1`) holds trivially , so $u \notin S_1 \cup S_2$]

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

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

$= C_{n+1}^{(1)}(u)$ [as $u \notin S_1 \cup S_2$ and thus $X(n+1)=u$ iff $Y(n+1)=u$]

This establishes condition (1`). Exactly similarly we handle the case when $k \notin S_2 \cup S_1$ , $u \notin S_2$ and $P_n(k, S_2) > 0$ , in which case condition (2`) holds.

Step 2 :

Let us now assume that condition (3) holds. Then we shall prove that $Y(n)$ is markovian.

Let us now consider the states i_0 ,..., i_n of the $Y(n)$ chain. Suppose that $i_k = S_1$ only when $k = 0$ and for $k > 0$, i_k is a singleton. Then we have:

$$P(Y(n)=i_n|Y(n)=i_{n-1},...,Y(0)=i_0)$$
$$= \frac{\sum_{\alpha\in S_1} P(X(n)=i_n|X(n-1)=i_{n-1},...,X(0)=\alpha)P(X(n-1)=i_{n-1},...,X(0)=\alpha)}{\sum_{\alpha\in S_1} P(X(n-1)=i_{n-1},...,X(0)=\alpha)}$$
$$= \frac{P(X(n)=i_n|X(n-1)=i_{n-1})\sum_{\alpha\in S_1} P(X(n-1)=i_{n-1},...,X(0)=\alpha)}{\sum\limits_{\alpha\in S_1} P(X(n-1)=i_{n-1},...,X(0)=\alpha)}$$
$$= P(X(n)=i_n|X(n-1)=i_{n-1})$$
$$= P(Y(n)=i_n|Y(n-1)=i_{n-1})$$

Now we consider the states $i_0,...,i_n$ of $Y(n)$ and let

$$m = max\{k|0<k\leq n, i_k = S_1 \text{ or } S_2\}$$

Then we notice that $m \leq n$ and

$P(Y(n)=i_n|Y(n-1)=i_{n-1},Y(n-2)=i_{n-2},...,Y(0)=i_0)$ must be of the form

$P(Y(n)=i_n|Y(n-1)=i_{n-1},...,Y(m)=S_k,Y(m-1)=S_k,...,Y(0)=S_k)$ where $k=1$ throughout or $k=2$ throughout where we have used condition (3).

However we have:

$P(Y(n)=i_n,Y(n-1)=i_{n-1},...,Y(m)=S_1,...,Y(1)=S_1,Y(0)=S_1)$

$=P(Y(n)=i_n,Y(n-1)=i_{n-1},...,Y(m)=S_1,...,Y(1)\in S_1,Y(0)\in S_1)+$
$P(Y(n)=i_n,Y(n-1)=i_{n-1},...,Y(m)=S_1,...,Y(1)\in S_1,Y(0)\neq S_1)$

$=P(Y(n)=i_n,Y(n-1)=i_{n-1},...,Y(m)=S_1,...,Y(1)\in S_1)$

=continuing the process

$=P(Y(n)=i_n,Y(n-1)=i_{n-1},...,Y(m)\in S_1)$

$=P(Y(n)=i_n|Y(n-1)=i_{n-1})$ [by the argument used in the beginning of this step, because $i_{m+1},i_{m+2},...,i_n$ are all singletons]

In case $m=n$ and $i_n=i_{n-1}=S_1$ or $m=n$ and $i_n=i_{n-1}=S_2$, the same will be true.

Step 3 :

Here we assume that conditions (1) and (2) hold. We shall prove that :

$P(Y(n) = i_n | Y(n-1) = i_{n-1}, ..., Y(0) = i_0) = P(Y(n) = i_n | Y(n-1) = i_{n-1})$ We have the following possibilities:

Case 1

$i_{n-1} = S_1$, $i_n \neq S_1$

Since we have assumed that no transition is possible between S_1 and S_2 , in this case i_n must be a singleton (i.e not in S_2).

Case 2

$i_{n-1} = S_2$, $i_n \neq S_2$

Here i_n must a singleton as before.

Case 3

$i_n = S_1$, $i_{n-1} \neq S_1$

Here i_{n-1} must be a singleton set as before and $i_{n-1} \neq S_2$.

Case 4

$i_n = S_2$, $i_{n-1} \neq S_2$

Here i_{n-1} is a singleton.

Case 5

$i_n = S_1$ and $i_{n-1} = S_1$

Case 6

$i_n = S_2$ and $i_{n-1} = S_2$

We shall address all the six cases one by one.

Case 1

Writing $S_{i_k} = i_k$ when i_k is a singleton and $S_{i_k} = S_j$ when $i_k = S_j (j = 1 \text{ or } 2)$, we have:

$$P(Y(n) = i_n | Y(n-1) = S_1, Y(n-2) = i_{n-2}, ..., Y(0) = i_0)$$
$$= \frac{\sum_{\alpha_{n-1} \in S_1} \cdots \sum_{\alpha_0 \in S_{i_0}} P(X(n) = i_n, ..., X(0) = \alpha_0)}{\sum_{\alpha_{n-1} \in S_1} \cdots \sum_{\alpha_0 \in S_{i_0}} P(X(n-1) = \alpha_{n-1}, ..., X(0) = \alpha_0)}$$
$$= \frac{\sum_{\alpha_{n-1} \in S_1} \cdots \sum_{\alpha_0 \in S_{i_0}} P_n(\alpha_{n-1}, i_n) P_{n-1}(\alpha_{n-2}, \alpha_{n-1}) P(X(n-2) = \alpha_{n-2}, ..., X(0) = \alpha_0)}{\sum_{\alpha_{n-2} \in S_{i_{n-2}}} \cdots \sum_{\alpha_0 \in S_{i_0}} P_{n-1}(\alpha_{n-2}, S_1) P(X(n-2) = \alpha_{n-2}, ..., X(0) = \alpha_0)}$$

$= C_n^{(1)}(i_n)$ [using condition (1)]

$= P(X(n) = i_n | X(n-1) \in S_1)$

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

Case 2

It is similar to case 1.

Case 3

Here we have :

$$P(Y(n) = i_n | Y(n-1) = i_{n-1}, Y(n-2) = i_{n-2}, ..., Y(0) = i_0)$$
$$= \frac{P(X(n) \in S_1 | X(n-1) = i_{n-1}) \sum_{\alpha_{n-1} \in S_{i_{n-2}}} \cdots \sum_{\alpha_0 \in S_{i_0}} P(X(n-1) = i_{n-1}, ..., X(0) = \alpha_0)}{\sum_{\alpha_{n-2} \in S_{i_{n-2}}} \cdots \sum_{\alpha_0 \in S_{i_0}} P(X(n-1) = i_{n-1}, ..., X(0) = \alpha_0)}$$

$= P(X(n) \in S_1 | X(n-1) = i_{n-1})$

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

Case 4

It is similar to case 3 .

Case 5

Let us assume that $i_n = i_{n-1} = S_1$. Then we have:

$$P(Y(n) = S_1 | Y(n-1) = S_1, Y(n-2) = i_{n-2}, ..., Y(0) = i_0)$$
$$= \frac{\sum_{\alpha_n \in S_1} \cdots \sum_{\alpha_0 \in S_{i_0}} P(X(n) = \alpha_n, ..., X(0) = \alpha_0)}{\sum_{\alpha_{n-1} \in S_1} \cdots \sum_{\alpha_0 \in S_{i_0}} P(X(n-1) = \alpha_{n-1}, ..., X(0) = \alpha_0)}$$

$$= \frac{\sum_{\alpha_n \in S_{i_n}} \cdot\cdot \sum_{\alpha_0 \in S_{i_0}} P_n(\alpha_{n-1}, \alpha_n) P_{n-1}(\alpha_{n-2}, \alpha_{n-1}) P(X(n-2) = \alpha_{n-2}, .., X(0) = \alpha_0)}{\sum_{\alpha_{n-1} \in S_1} \cdots \sum_{\alpha_0 \in S_{i_0}} P_{n-1}(\alpha_{n-2}, \alpha_{n-1}) P(X(n-2) = \alpha_{n-2}, ..., X(0) = \alpha_0)}$$

$= 1 - C_n^{(1)}(S_1^c)$ [using condition (1)]

$= C_n^{(1)}(S_1) = P(X(n) \in S_1 | X(n-1) \in S_1)$

$= P(Y(n) = S_1 | Y(n-1) = S_1)$

Case 6

Again it is similar to case 5 .

Clearly the above result can be generalised even further in the following manner:

Corollary. Let $X(n)$ be a finite non homogeneous markov chain . Suppose that there are r many disjoint sets $S_1, S_2, ..., S_r$ each of which collapses more than one state. Suppose that for each $n \geq 1$,

$$P_n(x, S_i) = 0 \text{ for all } x \in S_j \ i \neq j \ , i = 1, 2, ..., r \text{ and } j = 1, 2, ..., r$$

The following conditions are also assumed :

(A) $\sum_{l \in S_i} P_n(k, l) P_{n+1}(l, u) = P_n(k, S_i) C_{n+1}^{(i)}(u)$ for all k and for all $u \notin S_i$ where

$C_{n+1}^{(i)}(u) = P(X(n+1) = u | X(n) \in S_i)$ $i = 1, 2, ..., r$

(B) For $k \notin S_i$ $P_n(k, S_i) = 0$ for all $n \geq 1$ $i = 1, 2, ..., r$

$(A^{\backprime})$ This is condition (A) when $k \notin \cup_{i=1}^{r} S_i$

Then $Y(n)$ is markovian if either condition (B) holds for all i or condition (A) holds for all $i = 1, 2, ..., r$. Conversely if $Y(n)$ is markovian , then either condition (B) holds for all $i = 1, 2, ..., r$ or condition $(A^{\backprime})$ holds for all i.

The following result is the analogue of the Burke-Rosenblatt third theorem.

Theorem 4

Let $X(n)$ be a non homogeneous markov chain with state space $S = \{1,2,...,m\}$, initial probability vector $(\frac{1}{m},...,\frac{1}{m})$, symmetric transition probability matrices P_n and $P_n = P_{n+1}$ for every odd n. Then $f(X(n))$ is Markovian for every many to one transformations f if and only if the transition probability matrices P_n of $X(n)$ is of the form

$$P_n = \begin{pmatrix} \beta_n & \frac{1-\beta_n}{m-1} & \frac{1-\beta_n}{m-1} & \cdots & \frac{1-\beta_n}{m-1} \\ \frac{1-\beta_n}{m-1} & \beta_n & \frac{1-\beta_n}{m-1} & \cdots & \frac{1-\beta_n}{m-1} \\ .. & .. & .. & \cdots & .. \\ .. & .. & .. & \cdots & .. \\ \frac{1-\beta_n}{m-1} & \frac{1-\beta_n}{m-1} & .. & \cdots & \beta_n \end{pmatrix}$$

where β_n is a real number in $(0,1)$.

Proof

Suppose $f(X(n))$ is markovian for every many to one transformation f . We have also assumed that the initial probability vector is uniform, the transition probability matrices of $X(n)$ are symmetric and $P_n = P_{n+1}$ for every odd n.Then by theorem $(1B)$,

$P_n(i,S_j) = \sum_{k \in S_j} P_n(i,k)$ is independent of $i \notin S_j$ where $S_j \subset \{1,2,...,m\}$.

For $1 \leq i \leq m$, $P_n(i,i) = 1 - \sum_{k \neq i} P_n(i,k)$ (5)

For $k \neq i$, $l \neq i$, $P_n(i,k) = P_n(k,i)$ and $P_n(i,l) = P_n(l,i)$ [because of symmetry]

Let $S = \{i\}$, $k \notin S$, $l \notin S$, then we have $P_n(k,i) = P_n(l,i)$.

Hence for $k \neq i$ and $l \neq i$ and taking $S = \{i\}$ we have

$P_n(i,k) = P_n(k,i) = P_n(l,i) = P_n(i,l)$.

Then from (5) we have,

$P_n(i,i) = 1 - (m-1)P_n(i,k)$ where $k \neq i$ and $1 \leq k \leq m$.

If we write $\beta_n^i = P_n(i,i)$, then $P_n(i,k) = \dfrac{1-\beta_n^i}{m-1}$ for every $k \neq i$.

But by symmetry, $P_n(i,k) = P_n(k,i)$ for $k \neq i$ and as such β_n^i is independent of i. Thus the only if part has been proven.

Conversely, if P_n is of the form mentioned in the theorem , then clearly for any set $S \subset \{1,2,...,m\}$ and $i \notin S$ and $j \notin S$, we have

$$\sum_{k\in S} P_n(i,k) = \sum_{k\in S} P_n(j,k)$$

and from theorem1 , this is a sufficient condition for the process , $f(X(n))$ to be Markovian.

Remark :

The condition $P_n = P_{n+1}$ for every odd n is not required for the sufficiency part of the above theorem.

Now, We shall construct an example to show that if we drop the assumption of reversibility of $X(n)$ from theorem 1(B) then the condition

$$\sum_{j\in S_\beta} P_n(i,j) \text{ is independent of } i \in S_\alpha$$

will not be necessary any more for $Y(n)$ to be markov.

Example

Let $X(n)$ be a non homogeneous markov chain with state space $S = \{1,2,3\}$ and transition probability matrices

$$P_1 = \begin{pmatrix} \frac{1}{8} & \frac{1}{4} & \frac{5}{8} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{5}{8} & \frac{1}{4} & \frac{1}{8} \end{pmatrix} \qquad P_n = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} n = 2,3,...$$

Let the initial probabaility distribution of $X(n)$ be uniform . Let $S = S_1 \cup S_2$ where $S_1 = \{1,2\}$ and $S_2 = \{3\}$ and $Y(n) = i$ if and only if $X(n) \in S_i$.

Now let $i \in S_1$ and $i = 1$.

Then, $\sum_{j \in S_2} P_1(1,j) = P_1(1,3) = \frac{5}{8}$. Again let $i \in S_1$ and $i = 2$.
Then, $\sum_{j \in S_2} P_1(2,j) = P_1(2,3) = \frac{1}{4}$. Clearly $\sum_{j \in S_2} P_1(i,j)$ is not independent of $i \in S_1$.
If we can show that for all $n \geq 1$,

$$\sum_{l \in S_1} P_n(k,l)P_{n+1}(l,u) = P_n(k,S_1)C_{n+1}(u) \ \forall k \text{ and } \forall u \notin S_1,$$

then from theorem (2) we know that $Y(n)$ will be Markov.

Case 1. $n = 1$, $k = 3$, $u = 3$.

$$\frac{\sum_{l \in S_1} P_n(k,l)P_{n+1}(l,u)}{P_n(k,S_1)}$$
$$= \frac{P_1(3,1)P_2(1,3) + P_1(3,2)P_2(2,3)}{P_1(3,1) + P_1(3,2)}$$
$$= \frac{1}{3}$$

$$C_{n+1}(u) = P(X(n+1) = u | X(n) \in S_1)$$

$$= P(X(2) = 3 | X(1) \in S_1)$$
$$= \frac{P(X(2) = 3, X(1) \in S_1)}{P(X(1) \in S_1}$$
$$= \frac{P(X(2) = 3, X(1) = 1) + P(X(2) = 3, X(1) = 2)}{P(X(1) = 1) + P(X(1) = 2)}$$
$$= \frac{P(X(2) = 3 | X(1) = 1)P(X(1) = 1) + P(X(2) = 3 | X(1) = 2)P(X(1) = 2)}{P(X(1) = 1) + P(X(1) = 2)}$$
$$= \frac{\frac{1}{3}[P_2(1,3) + P_2(2,3)]}{\frac{1}{3} + \frac{1}{3}}$$
$$= \frac{1}{3}$$

Case 2.$n = 1$, $k = 2$, $u = 3$.

$$\frac{\sum_{l \in S_1} P_n(k,l)P_{n+1}(l,u)}{P_n(k,S_1)}$$
$$= \frac{P_1(2,1)P_2(1,3) + P_1(2,2)P_2(2,3)}{P_1(2,1) + P_1(2,2)}$$

$$= \frac{1}{3} = C_{n+1}(u) \text{ [from previous case]}$$

Case 3. $n = 1$, $k = 1$, $u = 3$.

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

$$= \frac{P_1(1,1)P_2(1,3) + P_1(1,2)P_2(2,3)}{P_1(1,2) + P_1(1,1)}$$

$$= \frac{1}{3} = C_{n+1}(u) \text{ [from case (1)]}$$

Case 4. $n = 2$, $k = 3$, $u = 3$.

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

$$= \frac{P_2(3,1)P_3(1,3) + P_2(3,2)P_3(2,3)}{P_2(3,1) + P_2(3,2)}$$

$$= \frac{1}{3}$$

$$C_{n+1}(u) = P(X(n+1) = u | X(n) \in S_1)$$

$$= \frac{P(X(3) = u, X(2) \in S_1)}{P(X(2) \in S_1)}$$

$$= \frac{P(X(3) = 3, X(2) = 1) + P(X(3) = 3, X(2) = 2)}{P(X(2) = 1) + P(X(2) = 2)}$$

$$= \frac{P_3(1,3)P(X(2) = 1) + P_3(2,3)P(X(2) = 2)}{P(X(2) = 1) + P(X(2) = 2)}$$

$$= \frac{1}{3}$$

Since P_n's are same for all $n \geq 2$, hence we do not need to consider any more case.

Hence $Y(n)$ is Markov but the condition mentioned in theorem 1(B) is not necessary.

Kemeny-Snell Analogues

In their book "Finite Markov Chains", Kemeny and Snell gave certain results regarding lumpability of markov chains in the homogeneous context. In this section we will reproduce analogous results in the non homogeneous scenario.

Theorem 1

Let $X(n)$ be a non homogeneous markov chain with

1) State space $S = \{1, 2, ..., m\}$

2) Symmetric transition probability matrices P_n such that

$$P(X(n) = j|X(n-1) = i) = P_n(i, j) \; \forall i, j \in S \; n = 1, 2, ...$$

3)$P_n = P_{n+1}$ for every odd n.

Let $\{S_1, S_2, ..., S_r\}$ be a partition of S , $r \leq m$. We define $Y(n) = i$ if and only if $X(n) \in S_i$. Then weak lumpability of $X(n)$ with respect to uniform initial probability vector ($P(X(0) = i) = \frac{1}{m}$) implies strong lumpability.

Proof

Since $X(n)$ is weakly lumpable with respect to uniform initial probability distribution, hence from theorem $1(B)$ of chapter (3) we have,

$$Q_n Q_{n+1} = AP_n P_{n+1} B = AP_n BAP_{n+1} B$$

$$\implies AP_n(I - BA)P_{n+1}B = 0$$

$$\implies P_n B = BAP_n B \text{ (From theorem } 1(B)) \tag{1}$$

Now, $$(P_n B)_{ij} = \sum_k P_n(i,k) B_{kj} = \sum_{k \in S_j} P_n(i,k) \tag{2}$$

Again, $$(BAP_n B)_{ij} = \sum_{l,t} B_{ik} A_{kl} P_n(l,t) B_{tj}$$

$$= \sum_{i,l \in S_k} \sum_{t \in S_j} \frac{1}{|S_k|} P_n(l,t)$$

$$= \sum_{i,l \in S_k} \frac{1}{|S_k|} P_n(l, S_j)$$

$$= Q_n(k, j) \tag{3}$$

Hence from (1), (2) and (3) we have

$\sum_{k \in S_j} P_n(i,k) = Q_n(k, j) \; \forall n$ which is exactly the sufficient condition for $X(n)$ to be strongly lumpable from theorem $1(A)$ of chapter (3).

Reversibility in Non Homogeneous Context

Here we define reversibility in a slightly different manner . Let π be a left invariant probability vector and $X(n)$ be a non homogeneous markov chain. We say that $X(n)$ is a reversible markov chain if

$$\pi_i P(X(n+1) = j|X(n) = i) = \pi_j P(X(n+1) = i|X(n) = j)$$

Suppose π is the initial probability vector. Then we observe that

$$\pi_i \frac{P(X(n+1) = j, X(n) = i)}{P(X(n) = i)} = \pi_j \frac{P(X(n+1) = i, X(n) = j)}{P(X(n) = j)}$$

$\implies P(X(n+1) = j, X(n) = i) = P(X(n+1) = i, X(n) = j)$ [since π is left invariant]

$$\implies \frac{P(X(n+1) = j, X(n) = i)}{P(X(n) = i)} = \frac{P(X(n) = j, X(n+1) = i}{P(X(n+1) = i)}$$

$$\implies P(X(n+1) = j|X(n) = i) = P(X(n) = j|X(n+1) = i)$$

Here we used the fact that $P(X(n) = i) = \pi_n(i) = \pi_{n+1}(i) = P(X(n+1) = i)$ because of left invariance of π.

Theorem 2

Let $X(n)$ be a reversible non-homogeneous markov chain with state space $S = \{1, 2, ..., m\}$ and uniform initial probability distribution, i.e $P(X(0) = i) = \frac{1}{m}\ \forall i \in S$. Let $\{S_1, S_2, ..., S_r\}$ be a partition of S , $r \leq m$. Then the lumped chain with respect to this partition is also reversible. We assume that the collapsed chain is Markov here.

Proof

By reversibility we have, $P_n = D_n P_n^T D_n^{-1}\ \forall n$ where $(D_n)_{ii} = \dfrac{1}{m}\ \forall n$.

From theorem $1(B)$, we know that the transition probability matrix $Y(n)$ is of the form

$$Q_n = AP_nB\ \forall n$$

$$\implies Q_n = AD_nP_n^T D_n^{-1}B \tag{4}$$

We define $(\hat{D}_n)_{ii} = \dfrac{(D_n)_{jj}}{|S_i|}\ j \in S_i$

Then we have,

$$(\hat{D}_n B^T)_{ij} = (\hat{D}_n)_{ii} B^T_{ij}$$

$$= (\hat{D}_n)_{ii} B_{ji}$$

$$= (\hat{D}_n)_{ii}\ j \in S_i$$

$$= \frac{(D_n)_{jj}}{|S_i|}\ j \in S_i$$

Again,

$$(AD_n)_{ij} = A_{ij}(D_n)_{jj}$$

$$=A_{ij}(D_n)_{jj}$$

$$= \frac{1}{|S_i|}(D_n)_{jj}\ j \in S_i$$

Hence $AD_n = \hat{D}_n B^T$ (5)

Again we observe that,

$$(D_n^{-1}B)_{ij} = (D_n^{-1})_{ii} B_{ij} = \frac{1}{(D_n)_{ii}}\ i \in S_j$$

And, $$(A^T \hat{D}_n^{-1})_{ij} = A^T_{ij}(\hat{D}_n^{-1})_{jj}$$

$$= A_{ji}(\hat{D}_n^{-1})_{jj}$$

$$= \frac{1}{|S_j|}(\hat{D}_n^{-1})_{jj}\ i \in S_j$$

$$= \frac{1}{|S_j|}\frac{|S_j|}{(D_n)_{ii}}\ i \in S_j$$

$$= \frac{1}{(D_n)_{ii}}\ i \in S_j$$

Hence $D_n^{-1}B = A^T \hat{D}_n^{-1}$ (6)

Therefore from $(4),(5),(6)$ we have,

$$Q_n = \hat{D}_n B^T P_n^T A^T \hat{D}_n^{-1}$$

$= \hat{D}_n (AP_nB)^T \hat{D}_n^{-1}$

$= \hat{D}_n Q_n^T \hat{D}_n^{-1}$

Hence the lumped chain is also reversible.

Reverse Markov Chain

A markov chain observed in the reverse order is also markov because of the following:

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

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

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

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

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

Theorem 3

If a given non homogeneous markov chain is weakly lumpable with respect to a partition $A = \{A_1, ..., A_n\}$, then so is the reverse chain.

Proof

Let $X(n)$ be a non homogeneous markov chain which is weakly lumpable with respect to partition $A = \{A_1, ..., A_n\}$.

Now, we must prove that all the 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 still markov.

$P(Y(1) = i|Y(2) = j, Y(3) = h, ..., Y(n) = t)$

$= P(Y(1) = i|Y(2) = j)$ [from the above discussion on reverse markov chain]

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

CHAPTER V

CONCLUDING REMARKS

In this thesis we have considered only discrete time finite state space markov chains. We would like to consider the same problem in the parlance of continuous time markov chain with countably many and uncountably many states in near future.

REFERENCES

[1] Abdel-Moniem, A., & Leysieffer, F. (1982). Weak lumpability in finite markov chains. *Journal of Applied Probability*, 19(3), 685-691.

[2] Burke, C., & Rosenblatt, M. (1958). A markovian function of a markov chain. *Annals of Mathematical Statistics*, 29(4), 1112-1122.

[3] Franklin, J. (1968). Matrix Theory. *Prentice Hall Inc*.

[4] Higham, N. (1990). Analysis of cholesky decomposition of a semi-definite matrix. *Reliable Numerical Computations*, 161-185.

[5] Iosifescu, M. (1979). Finite Markov Processes and Their Applications. *John Wiley and Sons*.

[6] Issacson, D., & Madsen, R. (1976). Markov Chains Theory and Applications. *John Wiley and Sons*.

[7] Karlin, T., & Taylor, H. (1975). A First Course in Stochastic Process. *Academic Press*.

[8] Kemeny, G., & Snell, L. (1960). Finite Markov Chains. *Springer-Verlag*.

[9] Kijima, M. (1997). Markov Processes for Stochastic Modeling. *Chapman and Hall*.

[10] Rosenblatt M. (1971). Markov Process Structure and Asymptotic Behavior. *Springer-Verlag*.

[11] Rubinho, G., & Sericola, B. (1989). On weak lumpability in markov chains. *Journal of Applied Probability*, 26(3), 446-457.

BIOGRAPHICAL SKETCH

Agnish Dey was born in 1986, in Kolkata, West Bengal, India. He did his schooling from a local high school. Then in his bachelor program he did engineering in Information Technology from West Bengal University of Technology in 2008. After that he became a pupil of mathematics and earned his master's degree in mathematics from UT Pan American in August, 2011. During the master's program in UT Pan American, he worked as a graduate teaching assistant in the department of mathematics. He will be joining the PhD program in mathematics at University of Florida, Gainesville as a teaching assistant from the fall of 2011. His permanent address is 47, Baguiati second lane, Kolkata - 700028, West Bengal, India.