NON-STATIONARY MARKOV CHAINS

A THESIS
SUBMITTED TO THE DEPARTMENT OF MATHEMATICS
AND THE INSTITUTE OF ENGINEERING AND SCIENCES
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

By
Saed Mallak
July, 1996

NON-STATIONARY MARKOV CHAINS

A THESIS
SUBMITTED TO THE DEPARTMENT OF MATHEMATICS
AND THE INSTITUTE OF ENGINEERING AND SCIENCES
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

By
Saed Mallak
July, 1996

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Asst. Prof. Dr. Azer Kerimov (Principal Advisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof. Dr. Mefharet Kocatepe

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Farhad Hüseyinov

Approved for the Institute of Engineering and Sciences:

Prof. Dr. Mehmet Baray
Director of Institute of Engineering and Sciences

ABSTRACT

NON-STATIONARY MARKOV CHAINS

Saed Mallak
M.S. in Mathematics
Supervisor: Asst. Prof. Dr. Azer Kerimov
July, 1996

In this work, we studied the Ergodicity of Non-Stationary Markov chains. We gave several examples with different cases. We proved that given a sequence of Markov chains such that the limit of this sequence is an Ergodic Markov chain, then the limit of the combination of the elements of this sequence is again Ergodic (under some condition if the state space is infinite). We also proved that the limit of the combination of an arbitrary sequence of Markov chains on a finite state space is Weak Ergodic if it satisfies some condition. Under the same condition, the limit of the combination of a doubly stochastic sequence of Markov chains is Ergodic.

Keywords : Markov chain, Stochastic, Doubly stochastic, Irreducible, Aperiodic matrix, Persistent, Transient, Ergodic, Ergodic Theorem.

ÖZET

DURAĞAN OLMAYAN MARKOV ZİNCİRLERİ

Saed Mallak
Matematik Bölümü Yüksek Lisans
Tez Yöneticisi: Asst. Prof. Dr. Azer Kerimov
Temmuz, 1996

Bu çalışmada durağan olmayan Markov zincirlerinin Ergodikliğini inceledik. Bazı farklı durumlardan örnekler verdik. Limiti Ergodik Markov zinciri olmak üzere verilen bir Markov zincirleri dizisi için bu dizinin kombinasyonunun limitinin de gene bir Ergodik Markov zinciri olduğunu ispatladık (Ancak durum uzayının sonsuz olması halinde, bu durum bazı koşullar altında mümkündür). Ayrıca sonlu bir durum uzayında verilen herhangi bir Markov zincirleri dizisi için bu dizinin bazı koşulları sağlaması halinde kombinasyonunun limitinin zayıf Ergodik olduğunu gösterdik. Gene bazı koşullar altında bu dizinin çifte stokastik Markov zincirilerinde oluşması halinde limitinin Ergodik olduğunu gördük.

Anahter Kelimeler: Markov zinciri, Stokastik, Çifte stokastik, İndigenemez, Periyodik olmayan matris, Devamlı, Geçici, Ergodik, Ergodik Teorem.

ACKNOWLEDGMENT

As soon as I had started my work in this thesis, I was shocked by the sudden death of my father to whom I owe everything in my life. To his spirit I pray and to his love memory I dedicate this thesis.

It is my pleasure to thank my supervisor Asst. Prof. Dr. Azer Kerimov for his supervision, guidance and encouragement during my research in this thesis.

Words are not enough to express my thanks to my family who are on my side for good and bad times.

Lastly, I would like to thank all my friends for their help and support.

TABLE OF CONTENTS

Chapter 1

Introduction

The basic property characterizing Markov chains is a probabilistic analogue of a familiar property of dynamical systems. If one has a system of particles and the positions and velocities of all particles are given at time t, the equations of motions can be completely solved for the future development of the system. Therefore, any other information given concerning the past of the process up to time t is superfluons as far as future development is concerned. The present state of the system contains all relevant information concerning the future.

Markov chains are stochastic processes which are ways of quantifying the dynamic relationships of sequences of random variables. Stochastic models play an important role in many areas of the natural and engineering sciences.

Indeed, if we have a sequence of random variables with values in a discrete set, a countable set, then any such a sequence can form a Markov chain, which is conditional probabilities relating the elements of this sequence.

The most interesting object of the theory of Markov chains is the asymptotic behavior of these probabilities. The most interesting case when we have independence of the initial state, that is starting from any state, the particle reaches the desired state almost with the same probability. A Markov chain satisfying this is called an Ergodic Markov chain. We may characterize Ergodic Markov chains by the saying:"All The Ways Lead To Rome".

In the second chapter, we give a general review of the theory of Stationary Markov chains, definitions, classifications of the chains and main theorems.

In the third chapter, we introduce another situation, that is combinations from transition probabilities of different Markov chains. That is, we have different transition matrix in each transition (step). We give several examples to illustrate and explain this idea. Let us call such a situation Non-Stationary Markov chains.

In the fourth chapter, we state and prove some facts about Non-Stationary Markov chains.

Chapter 2

Stationary Markov Chains

2.1 Definitions

Definition 1 *Let S be a finite or countable set, suppose that to each pair $i, j \in S$, there is assigned a non-negative number p_{ij} such that these numbers satisfy the constraint:*

$$\sum_{j \in S} p_{ij} = 1, i \in S. \tag{2.1}$$

Let $X_0, X_1, X_2, \cdots$ be a sequence of random variables whose ranges are contained in S, the sequence is a Markov chain if:

$$P[X_{n+1} = j | X_0 = i_0, \cdots, X_n = i_n] = P[X_{n+1} = j | X_n = i_n] = p_{i_n j} \tag{2.2}$$

for all n and every sequence $i_0, \cdots, i_n$ in S for which $P[X_0 = i_0, \cdots, X_n = i_n] > 0$, this property is called Markov Property.

Definition 2 *S is called the state space or the phase space of the Markov chain.*

Definition 3 *A square matrix P is called a stochastic matrix if all its entries are non-negative and the summation of the elements of each row is 1.*

Definition 4 *Let $P = [p_{ij}]_{i,j \in S}$, P is called the one step transition (probability) matrix of this Markov chain.*

$P^2 = [p_{ij}^{(2)}]_{i,j \in S}$. $p_{ij}^{(2)}$ means starting from i reaching j in two steps, P^2 is the second step transition matrix.

And for any positive integer n. $P^n = [p_{ij}^{(n)}]_{i,j\in S}$. $p_{ij}^{(n)}$ *means starting from i reaching j in n steps.* P^n *is the n-th step transition matrix.*

Definition 5 *A sequence of random variables* $(X_n)_{n\geq 1}$ *is called a stationary sequence if for each natural numbers k and n,* $(X_1,\cdots,X_n)$ *and* $(X_{k+1},\cdots,X_{k+n})$ *have the same distribution.*

Remark 1 *Since we have the same transition matrix in each step, it is clear that the sequence of random variables which forms a Markov chain is stationary.*

2.2 Classifications Of The Chains

Let $f_{ii}^{(n)} := P_i(X_1 \neq i,\cdots,X_{n-1}\neq i, X_n = i)$.

$f_{ii} := \sum_{n=1}^{\infty} f_{ii}^{(n)} = P_i(\bigcup_{n=1}^{\infty}(X_n = i))$.

$\mu_i := \sum_{n=1}^{\infty} n f_{ii}^{(n)}$

Definition 6

1. *A state* $i\in S$ *is called persistent if* $f_{ii}=1$, *transient if* $f_{ii}<1$.
2. *A state* $i\in S$ *is called null persistent if the mean recurrence time* $\mu_i=\infty$.
3. *A state* $i\in S$ *is called periodic if* $\exists t>1$ *such that* $p_{ii}^{(n)}=0$ *unless* $n=rt$, *otherwise it is called aperiodic.*
4. *A Markov chain is called irreducible if* $\exists n$ *such that* $p_{ij}^{(n)}>0$, $\forall i,j\in S$, *otherwise it is called reducible .*
5. *A Markov chain is called Ergodic if all its states are persistent, aperiodic and non-null persistent states, there exists a stationary distribution.*
6. *A set of probabilities* $(\pi_j)_{j\in S}$ *satisfying* $\sum_{i\in S}\pi_i p_{ij} = \pi_j$ *is callad a stationary distribution.*

Theorem 1 *A state j is persistent if and only if* $P_j[X_n = j \quad i.o.] = 1$, *and* $\sum_n p_{jj}^{(n)} = \infty$.

A state j is transient if and only if $P_j[X_n = j \ \ i.o.] = 0$, *and* $\sum_n p_{jj}^{(n)} < \infty$.

Where i.o. stands for infinitely often.

Proof:

See, for example, [3], [4].

Lemma 1 $P_j[X_n = j \ \ i.o.]$ *is either 0 or 1.*

Proof:

See [3].

Theorem 2 *If a Markov chain is irreducible, then either all states are transient,* $P_i[\bigcup_j (X_n = j \ \ i.o.)] = 0$, $\forall i, j \in S$ *and* $\sum_n p_{ij}^{(n)} < \infty$.

Or, all states are persistent, $P_i[\bigcap_j (X_n = j \ \ i.o.)] = 1$, $\forall i, j \in S$ *and* $\sum_n p_{ij}^{(n)} = \infty$.

Proof:

See [3], [6], [9].

Remark 2 *Since* $\sum_{j \in S} p_{ij}^{(n)} = 1$, *the first alternative above is impossible if S is a finite set, that is a finite irreducible Markov chain is persistent.*

2.3 Convergence Theorems (Ergodic Theorems)

Theorem 3 *Suppose of an irreducible, aperiodic Markov chain that there exists a stationary distribution, that is a solution of* $\sum_{i \in S} \pi_i p_{ij}^{(n)} = \pi_j$, $\forall j \in S, n = 1, 2, \cdots$ *satisfying* $\pi_i > 0$ *and* $\sum_{i \in S} \pi_i = 1$, $\forall i \in S$, *then the Chain is persistent and* $\lim_{n \to \infty} p_{ij}^{(n)} = \pi_j$, $\forall i, j \in S$.

Theorem 4 *In the previous theorem if the state space is finite, then:*

$|p_{ij}^{(n)} - \pi_j| < A\rho^n$, *where A is a constant and* $0 \leq \rho < 1$.

Proof:

See [3], [6],[9].

Remark 3 *The main point of the conclusion is that since $p_{ij}^{(n)}$ reaches π_j for large n, the effect of the initial state wears off, that is the chain is very stable.*

Corollary 1 *Let $\xi_1, \xi_2, \cdots$ be a sequence of random variables which forms an Ergodic Markov chain.*

Let

$$I_j(\xi_n) = \begin{cases} 1, & if\ \xi_n = j, \\ 0, & otherwise. \end{cases}$$

$$V_j(n) = \frac{I_j(\xi_1) + \cdots + I_j(\xi_n)}{n}.$$

Then $V_j(n) \to \pi_j$ a.s., where a.s. stands for almost sure.

Proof:

See [5].

2.4 Summary

For an irreducible, aperiodic Markov chain there exist three possibilities:

1. The chain is transient, $\forall i, j \in S$, $\lim_{n\to\infty} p_{ij}^{(n)} = 0$ and $\sum_n p_{ij}^{(n)} < \infty$.

   If the state space is finite, then this case is impossible.

2. The chain is persistent, there exists no stationary distribution, $\forall i, j \in S$, $\lim_{n\to\infty} p_{ij}^{(n)} = 0$ and $\sum_n p_{ij}^{(n)} = \infty$, $\mu_j = \infty$.

   The null persistent case, if the state space is finite again this case is impossible.

3. The chain is Ergodic, there exists a stationary distribution, the chain is non-null persistent, $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j > 0$ and $\mu_j = \frac{1}{\pi_j} < \infty$, $\forall j \in S$.

Chapter 3

Non-Stationary Markov Chains And Examples

3.1 Introduction to Non-Stationary Markov Chains

Assume we have different Markov chains with different transition matrices, we will consider combinations of the probabilities of these chains. In other words, to get the higher probabilities of these combinations, we will use different transition matrices.

So, in one step, if we denote the probability of starting from state i reaching state j in one step by $q_{ij}^{(1)}$. then $q_{ij}^{(1)}$ is the same as the one step transition probability of the first chain. denote it by $p_{ij}^{(1)}$ and the transition matrix by P_1.

In two steps, denote it by $q_{ij}^{(2)}$, then $q_{ij}^{(2)} = \sum_{k\in S} p_{ik}^{(1)} p_{kj}^{(2)}$: where $[p_{ij}^{(1)}]_{i,j\in S}$, denote it by P_1, is the one step transition matrix of the first chain and $[p_{ij}^{(2)}]_{i,j\in S}$, denote it by P_2, is the one step transition matrix of the second chain.

And in general, for any positive integer n. the n-th step probability of the combination, $q_{ij}^{(n)}$, is $\sum_{k\in S} q_{ik}^{(n-1)} p_{kj}^{(n)}$, where $q_{ij}^{(n-1)}$ is the (n-1)th step probability of the combination and $p_{ij}^{(n)}$ is the one step probability of the n-th chain, denote its matrix by P_n. In matrix form, $Q_n = P_1 P_2 \cdots P_n$.

Obviously, since we consider combinations to find the higher order transition probability of a particular state. this means that we are in the same state space, that is all the chains have the same state space.

We will use the same definitions of the original case for irreducible, reducible, periodic, aperiodic, transient, persistent, null persistent and Ergodic state (chain).

The main question will be about Ergodicity of such combinations, that is whether the limit of $q_{ij}^{(n)}$ exists or not and the effect of the initial state whether it wears off or not for large n. In particular we will consider a sequence of Markov chains which tends to some Markov chain.

3.2 Examples of Stationary Markov Chains

Example 1 *Consider a Markov chain whose transition matrix is:*

$$P = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}.$$

$p_{ij} > 0$, $\forall i, j = 0, 1$.

Such a chain is Ergodic with the unique stationary distribution:

$\pi_0 = \frac{1-p_{11}}{2-p_{00}-p_{11}}$, $\pi_1 = \frac{1-p_{00}}{2-p_{00}-p_{11}}$.

Example 2 *A stochastic matrix is called doubly stochastic if all its columns sum to 1, i.e.* $\sum_{i \in S} p_{ij} = 1$, $\forall j \in S$. *It is clear that the n-th power of a doubly stochastic matrix is again doubly stochastic.*

If an irreducible, aperiodic finite Markov chain has a doubly stochastic transition matrix then this chain is Ergodic with the unique stationary distribution:

$\pi_j = \frac{1}{N}$, $\forall j \in S$, *where* N *is the cardinality of* S.

If the state space is infinite, then either all states are null persistent or all of them are transient.

Example 3 *Assume we have an irreducible, aperiodic Markov chain on an infinite state space. Assume* $\exists j_0 \in S$ *such that* $p_{ij_0} \geq \delta > 0$, $\forall i \in S$, *then such a chain is Ergodic, indeed:*

$p_{ij_0}^{(2)} = \sum_{k \in S} p_{ik} p_{kj_0} \geq \sum_{k \in S} p_{ik} \delta \geq \delta > 0$

and by induction $p_{ij_0}^{(n)} \geq \delta$, $\forall n = 1, 2, \cdots$

$\Rightarrow \lim_{n \to \infty} p_{ij_0}^{(n)} \geq \delta > 0$.

Since the chain is aperiodic the limit exists, thus j_0 is Ergodic. Hence, since the chain is irreducible, all states are Ergodic: that is this Markov chain is Ergodic.

3.3 Examples of Non-Stationary Markov Chains

Example 4 *Let A and B be two transition matrices of two different Markov chains, where:*

$$A = \begin{pmatrix} 0 & 1 \\ p & q \end{pmatrix}$$

$$B = \begin{pmatrix} p_1 & q_1 \\ 1 & 0 \end{pmatrix}.$$

Notice that both A^2 and B^2 are with non-zero entries, so both of them are transition matrices of Ergodic chains.

Now, if we consider the trivial combination of A and B, that is $ABAB\cdots$, then this combination is not Ergodic: indeed:

$$AB = \begin{pmatrix} 0 & 1 \\ p & q \end{pmatrix} \begin{pmatrix} p_1 & q_1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ pp_1 + q & pq_1 \end{pmatrix}.$$

According to this combination, the first state is absorbent, that is once the particle is in the first state it does not leave it.

Indeed, the limit of $ABAB\cdots$ does not exist, since:

$$\lim_n ABAB\cdots A = \begin{pmatrix} 0 & 1 \\ s & t \end{pmatrix}$$

$$\lim_n ABAB\cdots B = \begin{pmatrix} 1 & 0 \\ s_1 & t_1 \end{pmatrix}$$

This is an example such that the chains used are Ergodic while the combination is not Ergodic.

Example 5 *Let A and B be two transition matrices of two Markov chains, where:*

$$A = \begin{pmatrix} 1/2 & 0 & 0 & 1/2 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 & 1/2 \end{pmatrix}$$

$$B = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}.$$

Notice that $A^2 = A$ and in general for any n, $A^n = A$, also the same for B. So, both A and B are transition matrices of Non-Ergodic chains. Now:

$$AB = BA = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}.$$

Moreover, $ABA = ABB = AB = BA = ABAB = ABBA$, so if we consider any combination of these chains (providing that using both A and B at least one time) is Ergodic, indeed:

$$\lim_n (any \;\; combination) = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}.$$

This is an example such that the chains used are not Ergodic while any combination is Ergodic.

Example 6 *(a) Let A be the identity matrix, the transition matrix of a Markov chain whose all states are absorbent.*

Let B be any transition matrix of any Markov chain which is not Ergodic.

It is obvious that any combination of these chains is not Ergodic.

(b) Again let A be the identity matrix.

Let B be the transition matrix of any Markov chain which is Ergodic.

Consider any combination of these chains such that using B is infinitely often. Again it is obvious that any combination is Ergodic. Moreover, the limit of any is the same as the limit of B^n.

This is an example of a combination of an Ergodic chain with a Non-Ergodic one such that the combination is Ergodic.

Example 7 *Let A and B be two transition matrices of two Markov chains, where:*

$$A = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

$$B = \begin{pmatrix} 1/4 & 3/4 \\ 3/4 & 1/4 \end{pmatrix}.$$

•

Both A and B are with non-zero entries, so both A and B are transition matrices of Ergodic Markov chains.

Consider any combination of A and B, then it is Ergodic and the limit of any combination is the same as the limit of A^n the same as the limit of B^n.

Indeed this is not strange since both A and B are doubly stochastic, so they have the same unique stationary distribution $(1/2 \;\; 1/2)$

This an example of a combination of two Ergodic chains such that the combination is Ergodic and has the same stationary distribution as the first chain which is the same as the second chain.

Example 8 *let A and B be two transition matrices of two Markov chains, where:*

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$B = \begin{pmatrix} 0 & 1 \\ p & q \end{pmatrix}.$$

Notice that A is a transition matrix of a Non-Ergodic Markov chain. It is periodic with period 2. While B is a transition matrix of an Ergodic Markov chain (since B^2 is with non-zero entries).

Consider the trivial combination $ABAB\cdots$, then it is not Ergodic; indeed the limit of $ABAB\cdots$ does not exist.

This is an example of a combination of a Non-Ergodic chain with an Ergodic one such that the combination is not Ergodic.

Example 9 *Let A and B be two transition matrices of two Markov chains, where:*

$$A = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & & \\ 1/2 & 0 & 1/2 & 0 & & \\ 1/2 & 0 & 0 & 1/2 & & \\ & & & & & \\ 1/2 & 0 & 0 & 0 & & 1/2\cdots \\ & & & & & \end{pmatrix}$$

$$B = \begin{pmatrix} 1/4 & 3/4 & 0 & 0 & & \cdots \\ 1/9 & 0 & 8/9 & 0 & & \\ 1/16 & 0 & 0 & 15/16 & & \\ & & & & & \\ \frac{1}{(i+1)^2} & 0 & 0 & & \frac{(i+1)^2-1}{(i+1)^2} & \\ \cdots & & & & & \cdots \end{pmatrix}, \text{ } i \text{ is the number of the row.}$$

A is a transition matrix of an Ergodic Markov chain (the first column is bounded). B is a transition matrix of a Non-Ergodic Markov chain. Since:

$\lim_n \prod_{k=2}^{n} \frac{k^2-1}{k^2} = 1/2 \Rightarrow f_{11} < 1.$

So the first state is transient. Since the chain is irreducible, all states are transient.

Now, any combination of A and B is irreducible. It is obvious since:

$a_{ij} > 0 \Leftrightarrow b_{ij} > 0, \forall i, j \in S$ *and both A and B are irreducible.*

Next, any combination of A and B such that using A infinitely often is non-null persistent. Since we are using A infinitely often, once we use A, $q_{i1}^{(n)} = 1/2$ ($q_{ij}^{(n)}$ is the n-th step probability of the combination), so:

$\sum_{n=1}^{\infty} q_{i1}^{(n)} \geq \sum_{k=1}^{\infty} 1/2 = \infty.$

So, the first state is non-null persistent. The combination is irreducible, so all states are non-null persistent.

Next, any combination of A and B such that using both A and B infinitely often is not Ergodic, indeed the limit does not exist. Let C_1 be the class of any combination such that in the n-th step we have A and C_2 be the class of any combination such that in the n-th step we have B, both C_1 and C_2 have probability 1/2. Now , for C_1, $q_{i1}^{(n)} = 1/2$, for C_2, $q_{i1}^{(n)} \leq 1/4$, $\forall i \in S$. Hence the limit of any such a combination does not exist.

This is an example of a combination of Ergodic chain with Non-Ergodic one such that the combination is not Ergodic, moreover the limits of both A^n and B^n exist while the limit of the combination does not exist.

Example 10 *Let A and B be two transition matrices of two Markov chains, where:*

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & \\ 0 & 1/2 & 1/2 & 0 & \\ 0 & 1/2 & 1/2 & 0 & \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1/2 & 1/2 \end{pmatrix}$$

$$B = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/2 \end{pmatrix}$$

both A and B are transition matrices of reducible chains, so they are not Ergodic ($A^n = A$, $B^n = B$, $\forall n = 1, 2, \cdots$). Indeed, any combination of A and B is not Ergodic. If we use one of them just finitely often, then the combination will be reducible, that is, it is not Ergodic. If we use both of them infinitely often, then the combination will be irreducible but the limit of the combination will tend to the zero matrix (both A and B are doubley stochastic), that is, it is not Ergodic.

This is an example of a combination of Non-Ergodic chains such that the combination is not Ergodic.

Remark 4 *From the previous examples, if we consider arbitrary transition matrices of arbitrary Markov Chains and we consider combinations of these* *chains, then we have all the possibilities. We can find combinations of Ergodic chains which are not Ergodic, other combinations which are Ergodic. We can find combinations of Non-Ergodic chains which are Ergodic, other combinations which are not Ergodic. We can find combinations of Non-Ergodic chains with Ergodic ones which are Ergodic, other combinations which are not Ergodic.*

So, for such a case we did not reach any conclusion about the limit of the combination.

3.4 Examples of Sequences of Markov Chains

Example 11 *Let $(A_n)_{n=1}^{\infty}$ be a sequence of transition matrices of Markov chains, where:*

$$A_n = \begin{pmatrix} 1/2 - 1/n & 1/2 + 1/n & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Such a sequence tends to a Markov chain whose transition matrix is:

$$A = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

The limit of this sequence is not Ergodic, the second and third states are periodic with period 2. Thus, $\lim_n A^n$ *does not exist. Moreover,* $\lim_n A_1 A_2 \cdots A_n$ *does not exist.*

Example 12 *Let $(A_n)_{n=1}^{\infty}$ be a sequence of transition matrices of Markov chains, where:*

$$A_n = \begin{pmatrix} 1/n & 1 - 1/n \\ 1 - 1/n & 1/n \end{pmatrix}.$$

This sequence tends to a Markov chain whose transition matrix is:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The limit of this sequence is not Ergodic, it is periodic with period 2, thus $\lim_n A^n$ *does not exist while* $\lim_n A_1 A_2 \cdots A_n$ *exists and Ergodic. Indeed:*

$$\lim_n A_1 A_2 \cdots A_n = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

Example 13 *Let* $(A_n)_{n=1}^{\infty}$ *be a sequence of transition matrices of Markov chains, where:*

$$A_1 = \begin{pmatrix} 1/2 & 0 & 0 & 1/2 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 & 1/2 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}$$

$$A_n = \begin{pmatrix} 1/n & 0 & 0 & 1-1/n \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 1-1/n & 0 & 0 & 1/n \end{pmatrix}, \quad \forall n = 3, 4, \cdots.$$

Such a sequence tends to a Markov chain whose transition matrix is:

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Notice that all the elements of the sequence are not Ergodic, also the limit of the sequence is not Ergodic (and $\lim_n A^n$ *does not exist). while* $\lim_n A_1 A_2 \cdots A_n$ *exists and Ergodic, indeed:*

$$\lim_n A_1 A_2 \cdots A_n = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}.$$

Example 14 *Let* $(A_n)_{n=2}^{\infty}$ *be a sequence of transition matrices of Markov chains, where:*

$$A_n = \begin{pmatrix} 1/3 + 1/n & 2/3 - 1/n \\ 2/3 - 1/n & 1/3 + 1/n \end{pmatrix}.$$

Such a sequence tends to a Markov chain whose transition matrix is:

$$A = \begin{pmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{pmatrix}.$$

All the chains used are Ergodic, the limit is Ergodic, moreover:

$$\lim_n A^n = \lim_n A_1 A_2 \cdots A_n = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

Example 15 *Let* $(A_n)_{n=1}^{\infty}$ *be a sequence of transition matrices of Markov chains, where:*

$$A_n = \begin{pmatrix} 1 - 1/n^2 & 1/n^2 \\ 1 - 1/n^2 & 1/n^2 \end{pmatrix}.$$

Such a sequence tends to a Markov chain whose transition matrix is:

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

For each fixed n, A_n *is a transition matrix of an Ergodic chain, while the limit of the sequence is not Ergodic and the limit of the combination is not Ergodic, it is the same as A, that is:*

$$\lim_n A_1 A_2 \cdots A_n = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$$

Example 16 *Let* $(A_n)_{n=1}^{\infty}$ *be a sequence of transition matrices of Markov chains, where:*

$$A_n = \begin{pmatrix} 1/2 & (1/2)^2 & & (1/2)^{n-1} & 0 & (1/2)^n & \cdots \\ 1/2 & (1/2)^2 & & (1/2)^{n-1} & 0 & (1/2)^n & \\ 1/2 & (1/2)^2 & & (1/2)^{n-1} & 0 & (1/2)^n & \\ 1/2 & (1/2)^2 & & (1/2)^{n-1} & 0 & (1/2)^n & \\ & & & & & & \\ 1/2 & (1/2)^2 & & (1/2)^{n-1} & 0 & (1/2)^n & \\ & & & & & & \cdots \end{pmatrix}.$$

That is, in the n-th chain, the n-th state is isolated, it is not reached from any state. For each fixed n, the n-th chain is not Ergodic, while the limit of the *sequence is Ergodic. The limit of the sequence has the transition matrix A, where:*

$$A = \begin{pmatrix} 1/2 & (1/2)^2 & (1/2)^3 & \\ 1/2 & (1/2)^2 & (1/2)^3 & \\ 1/2 & (1/2)^2 & (1/2)^3 & \\ & & & \cdots \end{pmatrix}.$$

Moreover, $A_1 A_2 = A_2$, *and* $A_1 A_2 \cdots A_n = A_n$, *thus:*

$\lim_n A_1 A_2 \cdots A_n = \lim_n A_n = A$, *which is Ergodic.*

Example 17 *Let* $(A_n)_{n=1}^{\infty}$ *be a sequence of transition matrices of Markov chains, where:*

$$A_n = \begin{pmatrix} (1/2)^n & (1/2)^{n-1} & 1/2 & (1/2)^{n+1} & \cdots \\ (1/2)^n & (1/2)^{n-1} & 1/2 & (1/2)^{n+1} & \\ (1/2)^n & (1/2)^{n-1} & 1/2 & (1/2)^{n+1} & \\ & & & & \\ (1/2)^n & (1/2)^{n-1} & 1/2 & (1/2)^{n+1} & \\ & & & & \end{pmatrix}$$

That is,

$$A_1 = \begin{pmatrix} 1/2 & (1/2)^2 & (1/2)^3 & \\ 1/2 & (1/2)^2 & (1/2)^3 & \\ 1/2 & (1/2)^2 & (1/2)^3 & \\ & & & \cdots \end{pmatrix}$$

$$A_2 = \begin{pmatrix} (1/2)^2 & 1/2 & (1/2)^3 & \cdots \\ (1/2)^2 & 1/2 & (1/2)^3 & \\ (1/2)^2 & 1/2 & (1/2)^3 & \cdots \\ & & & \cdots \end{pmatrix}$$

and so on. For each fixed n, the n-th chain is Ergodic while the limit of the sequence is not Ergodic (the sequence tends to the zero matrix).

Moreover, $A_1 A_2 = A_2$ *and* $A_1 A_2 \cdots A_n = A_n$.

Thus, $\lim_n A_1 A_2 \cdots A_n = \lim_n A_n$, *that is the combination is not Ergodic.*

Remark 5 *From the examples of this section, if the limit of the sequence is not Ergodic, then the limit of the combination may not exist, may exist and not Ergodic, may exist and Ergodic.*

Chapter 4

Convergence Theorems Related With Non-Stationary Markov Chains

Theorem 5 *Assume we have a finite state space S. Assume we have a sequence of Markov chains such that the limit of this sequence is an Ergodic Markov chain. Then the limit of the combination of the elements of this sequence exists and Ergodic.*

Proof:

Let $(A_n)_{n=1}^{\infty}$ be the transition matrices of this sequence.

Denote the transition matrix of the n-th chain by A_n and its entries by $[p_{ij}^{(n)}]_{i,j\in S}$.

Let $\lim_n A_n = A$, denote its entries by $[a_{ij}]_{i,j\in S}$, let δ be the minimum over all the entries of A, let N be the cardinality of S and $\delta_n = min_{i,j}\ p_{ij}^{(n)}$.

Let $Q_n = A_1 A_2 \cdots A_n$, denote its entries by $[q_{ij}^{(n)}]_{i,j\in S}$.

Assume without loss of generality that A is with non-zero entries (otherwise $\exists n_0$ such that A^n is with non-zero entries $\forall n \geq n_0$).

Assume without loss of generality that $\forall n$, A_n is with non-zero entries (otherwise $\exists k_0$ such that A_n is with non-zero entries $\forall n \geq k_0$).

Now, for any stochastic matrix A with entries $[p_{ij}]_{i,j\in S}$ and minimum over all its entries δ, the following relations are valid:

Denote the summation over j in S satisfying $p_{uj} \geq p_{vj}$ by $\sum^+$ and the summation over j in S satisfying $p_{uj} < p_{vj}$ by $\sum^-$ for arbitrary states u and v.

Then:

$$\sum^{+}(p_{uj} - p_{vj}) + \sum^{-}(p_{uj} - p_{vj}) = 1 - 1 = 0 \tag{4.1}$$

and since $\sum^{+} p_{vj} + \sum^{-} p_{uj} \geq N\delta$, then:

$$\sum^{+}(p_{uj} - p_{vj}) = 1 - \sum^{-} p_{uj} - \sum^{+} p_{vj} \leq (1 - N\delta). \tag{4.2}$$

Next, we will use induction on n to prove that:

$(max_i\, q_{ij}^{(n)} - min_i\, q_{ij}^{(n)}) \leq \prod_{i=1}^{n}(1 - N\delta_i)$.

For $n = 1$:

$$max_{i,j}\, p_{ij}^{(1)} \leq (1 - (N-1)\delta_1)$$

$$min_{i,j}\, p_{ij}^{(1)} \geq \delta_1$$

$$\Rightarrow\ (max_i\, p_{ij}^{(1)} - min_i\, p_{ij}^{(1)}) \leq (1 - N\delta_1).$$

and since $Q_1 = A_1$

$$\Rightarrow\ (max_i\, q_{ij}^{(1)} - min_i\, q_{ij}^{(1)}) \leq (1 - N\delta_1) \tag{4.3}$$

For $n = 2$:

$$(q_{uj}^{(2)} - q_{vj}^{(2)}) = \sum_{k \in S}(q_{uk}^{(1)} - q_{vk}^{(1)})p_{kj}^{(2)}$$

$$\Rightarrow\ (q_{uj}^{(2)} - q_{vj}^{(2)}) \leq \sum^{+}(q_{uk}^{(1)} - q_{vk}^{(1)})M_j^{(2)} + \sum^{-}(q_{uk}^{(1)} - q_{vk}^{(1)})m_j^{(2)}$$

where, $M_j^{(2)} = max_i\, p_{ij}^{(2)}$, $m_j^{(2)} = min_i\, p_{ij}^{(2)}$

$$\Rightarrow\ (q_{uj}^{(2)} - q_{vj}^{(2)}) \leq \sum^{+}(q_{uk}^{(1)} - q_{vk}^{(1)})(M_j^{(2)} - m_j^{(2)})$$

$$\Rightarrow\ (q_{uj}^{(2)} - q_{vj}^{(2)}) \leq (1 - N\delta_1)(M_j^{(2)} - m_j^{(2)})$$

$$\Rightarrow\ (q_{uj}^{(2)} - q_{vj}^{(2)}) \leq (1 - N\delta_1)(1 - N\delta_2)$$

$$\Rightarrow\ (R_j^{(2)} - r_j^{(2)}) \leq (1 - N\delta_1)(1 - N\delta_2) \tag{4.4}$$

where $R_j^{(2)} = max_i\, q_{ij}^{(2)}$, $r_j^{(2)} = min_i\, q_{ij}^{(2)}$.

Next, assume that:

$$R_j^{(n-1)} - r_j^{(n-1)} \leq \prod_{i=1}^{n-1}(1 - N\delta_i) \tag{4.5}$$

where $R_j^{(n-1)} = max_i\ q_{ij}^{(n-1)}$, $r_j^{(n-1)} = min_i\ q_{ij}^{(n-1)}$.

Now, we want to prove that it is correct for n.

$$(q_{uj}^{(n)} - q_{vj}^{(n)}) = \sum_{k \in S}(q_{uk}^{(1)} - q_{vk}^{(1)})b_{kj}^{(n-1)}$$

where $[b_{ij}^{(n-1)}]_{i,j \in S}$ are the entries of $A_2 A_3 \cdots A_n$.

$$\Rightarrow\ (q_{uj}^{(n)} - q_{vj}^{(n)}) \leq \sum^{+}(q_{uk}^{(1)} - q_{vk}^{(1)})B_j^{(n-1)} + \sum^{-}(q_{uk}^{(1)} - q_{vk}^{(1)})b_j^{(n-1)}$$

$$= \sum^{+}(q_{uk}^{(1)} - q_{vk}^{(1)})(B_j^{(n-1)} - b_j^{(n-1)})$$

$$\Rightarrow\ (q_{uj}^{(n)} - q_{vj}^{(n)}) \leq (1 - N\delta_1)(B_j^{(n-1)} - b_j^{(n-1)})$$

where $B_j^{(n-1)} = max_i\ b_{ij}^{(n-1)}$, $b_j^{(n-1)} = min_i\ b_{ij}^{(n-1)}$.

$$\Rightarrow\ (q_{uj}^{(n)} - q_{vj}^{(n)}) \leq (1 - N\delta_1)\prod_{i=2}^{n}(1 - N\delta_i) = \prod_{i=1}^{n}(1 - N\delta_i)$$

$$\Rightarrow\ (R_j^{(n)} - r_j^{(n)}) \leq \prod_{i=1}^{n}(1 - N\delta_i) \tag{4.6}$$

where $R_j^{(n)} = max_i\ q_{ij}^{(n)}$, $r_j^{(n)} = min_i\ q_{ij}^{(n)}$.

Thus, for all n , $(R_j^{(n)} - r_j^{(n)}) \leq \prod_{i=1}^{n}(1 - N\delta_i)$.

Hence, since $\delta_n \to \delta$. $\Rightarrow \prod_{i=1}^{n}(1 - N\delta_i) \to 0$ as $n \to \infty$.

Thus, if $R_j^{(n)} \to \pi_j$, then $r_j^{(n)} \to \pi_j$, indeed:

$|q_{ij}^{(n)} - q_{kj}^{(n)}| \leq (R_j^{(n)} - r_j^{(n)}) < C\alpha^n$, where C is constant and $0 \leq \alpha < 1$. $\forall i, j, k \in S$.

That is, $|q_{ij}^{(n)} - \pi_j| \leq C\alpha^n$, for each $j \in S$ and for any $i \in S$.

Notes:

If A has zero entries, then $\exists n_0$ such that A^n is with non-zero entries $\forall n \geq n_0$, so $\exists l$ such that $A_1 A_2 \cdots A_l$ is with non-zero entries, so we may consider $A_1 A_2 \cdots A_{kl}$ and take the limit as $k \to \infty$.

The limit of the combination does not depend on n, that is, the limit exists. If we have the same transition matrix in each step, then $R_j^{(n)}$ is non-increasing (with respect to n) and $r_j^{(n)}$ is non-decreasing. If we do not have the same transition matrix, then they are almost monotonic (since we have a convergent sequence of transition matrices).

Corollary 2 *Assume we have a sequence of Markov chains on a finite state space S. Assume that the limit of this sequence is Ergodic. Let $V_j(n)$ be the average number of staying in state j. Let $(\pi_j)_{j\in S}$ be the stationary distribution of the combination of this sequence. Then:*

$P[|V_j(n) - \pi_j| > \epsilon] \to 0$ *as* $n \to \infty$. $\forall \epsilon > 0$.

Proof:

Let $\xi_0, \xi_1, \cdots, \xi_n, \cdots$ be the random variables which form these Markov chains (the values of them are in S).

Let

$$I_j(\xi_n) = \begin{cases} 1 & if\ \xi_n = j. \\ 0 & otherwise. \end{cases}$$

$V_j(n) = \frac{I_j(\xi_0)+\cdots+I_j(\xi_n)}{n+1}$.

Let $q_{ij}^{(n)}$ be the n-th step probability of the combination.

Let i, j be two states in S.

We want to prove that $P[|V_j(n) - \pi_j| > \epsilon|\xi_0 = i] \to 0,\ as\ n \to \infty$.

Notice that $EV_j(n) = \frac{1}{n+1}\sum_{m=0}^{n} q_{ij}^{(m)}$, which tends to π_j as n tends to ∞.

By Chebyshev's Inequality :

$P[|V_j(n) - \pi_j| > \epsilon|\xi_0 = i] < \frac{E[(V_j(n)-\pi_j)^2|\xi_0=i]}{\epsilon^2}$.

Thus, we have to prove that $E[(V_j(n) - \pi_j)^2|\xi_0 = i] \to 0$.

$E[(V_j(n) - \pi_j)^2|\xi_0 = i] = \frac{1}{(n+1)^2}E[[\sum_{k=0}^{n}(I_j(\xi_k) - \pi_j)]^2|\xi_0 = i]$

$= \frac{1}{(n+1)^2}\sum_{k=0}^{n}\sum_{l=0}^{n} m_{ij}^{(k,l)}$, where

$m_{ij}^{(k,l)} = E[(I_j(\xi_k)I_j(\xi_l))|\xi_0 = i] - \pi_j E[I_j(\xi_k)|\xi_0 = i] - \pi_j E[I_j(\xi_l)|\xi_0 = i] + \pi_j^2$

$\Rightarrow \; m_{ij}^{(k,l)} = q_{ij}^{(s)} q_{jj}^{(t)} - \pi_j q_{ij}^{(k)} - \pi_j q_{ij}^{(l)} + \pi_j^2$, where $s = min(k,l)$, $t = |k-l|$ and since we do not have the same transition matrix in each step, $q_{ij}^{(t)}$ stands for the probability of the combination from $s+1$ to $max(k,l)$.

But we have $q_{ij}^{(n)} = \pi_j + \epsilon_{ij}^{(n)}$, $\epsilon_{ij}^{(n)} \leq C\alpha^n$

$\Rightarrow \; m_{ij}^{(k,l)} \leq M[\alpha^s + \alpha^t + \alpha^k + \alpha^l]$. M is constant.

$\Rightarrow \; \frac{1}{(n+1)^2} \sum_{k=0}^{n} \sum_{l=0}^{n} m_{ij}^{(k,l)} \leq \frac{M}{(n+1)^2} \sum_{k=0}^{n} \sum_{l=0}^{n} [\alpha^s + \alpha^t + \alpha^k + \alpha^l]$

$\leq \frac{4M}{(n+1)^2} \frac{2(n+1)}{(1-\alpha)}$

$= \frac{8M}{(1-\alpha)} \frac{1}{n+1} \rightarrow 0$ as $n \rightarrow \infty$

$\Rightarrow \; P[|V_j(n) - \pi_j| > \epsilon | \xi_0 = i] \rightarrow o$ as $n \rightarrow \infty$.

Theorem 6 *Assume we have an arbitrary sequence of Markov chains on a finite state space S with corresponding transition matrices $(A_n)_{n=1}^{\infty}$.*

Denote the entries of A_n by $[a_{ij}^{(n)}]_{i,j \in S}$.

Assume that: $min_{i,j}\, a_{ij}^{(n)} = \delta_n \geq \epsilon > 0$ for infinitely many n's.

Then, $\lim_n A_1 A_2 \cdots A_n$ is Weak Ergodic.

That is: $\forall j \in S$, $\lim_n |q_{ij}^{(n)} - q_{kj}^{(n)}| = 0$, $\forall i, k \in S$, where $[q_{ij}^{(n)}]_{i,j \in S}$ are the entries of $Q_n = A_1 A_2 \cdots A_n$.

Proof:

Let N be the cardinality of the state space S.

Recall that for any stochastic matrix P with entries $[p_{ij}]_{i,j \in S}$, if we denote the summation over $j \in S$ satisfying $p_{uj} \geq p_{vj}$ by $\sum^{+}$ and the summation over $j \in S$ satisfying $p_{uj} < p_{vj}$ by $\sum^{-}$, for arbitrary states u and $v \in S$, then the following relations are valid:

$$\sum^{+}(p_{uj} - p_{vj}) + \sum^{-}(p_{uj} - p_{vj}) = 0 \tag{4.7}$$

$$\sum^{+}(p_{uj} - p_{vj}) \leq (1 - N\delta) \tag{4.8}$$

where $\delta = min_{i,j}\, p_{ij}$, $i, j \in S$.

Next, we will use induction on n to prove that:

$(max_i\, q_{ij}^{(n)} - min_i\, q_{ij}^{(n)}) \leq \prod_{i=1}^{n}(1 - N\delta_i)$.

For $n = 1$:

$(max_i\, q_{ij}^{(1)} - min_i\, q_{ij}^{(1)}) = (max_i\, a_{ij}^{(1)} - min_i\, a_{ij}^{(1)}) \leq (1 - (N-1)\delta_1 - \delta_1)$

$\Rightarrow$

$$(max_i\, q_{ij}^{(1)} - min_i\, q_{ij}^{(1)}) \leq (1 - N\delta_1). \tag{4.9}$$

For $n = 2$:

$(q_{uj}^{(2)} - q_{vj}^{(2)}) = \sum_{k \in S}(q_{uk}^{(1)} - q_{vk}^{(1)})a_{kj}^{(2)}$

$\leq \sum^{+}(q_{uk}^{(1)} - q_{vk}^{(1)})max_k\, a_{kj}^{(2)} \;+\; \sum^{-}(q_{uk}^{(1)} - q_{vk}^{(1)})min_k\, a_{kj}^{(2)}$.

Applying the first two equations (4.7,4.8) we get:

$(q_{uj}^{(2)} - q_{vj}^{(2)}) \leq \sum^{+}(q_{uk}^{(1)} - q_{vk}^{(1)})(max_k\, a_{kj}^{(2)} - min_k\, a_{kj}^{(2)})$

$\leq (1 - N\delta_1)(1 - N\delta_2)$

$\Rightarrow$

$$(max_i\, q_{ij}^{(2)} - min_i\, q_{ij}^{(2)}) \leq (1 - N\delta_1)(1 - N\delta_2). \tag{4.10}$$

Now, assume that for $n = m - 1$ it is correct that:

$$(max_i\, q_{ij}^{(m-1)} - min_i\, q_{ij}^{(m-1)}) \leq \prod_{i=1}^{m-1}(1 - N\delta_i). \tag{4.11}$$

Next, we want to prove it for $n = m$.

$(q_{uj}^{(m)} - q_{vj}^{(m)}) = \sum_{k \in S}(q_{uk}^{(1)} - q_{vk}^{(1)})b_{kj}^{(m-1)}$, where $[b_{ij}^{(m-1)}]_{i,j \in S}$ are the entries of $A_2 A_3 \cdots A_m$.

$\Rightarrow \; (q_{uj}^{(m)} - q_{vj}^{(m)}) \leq \sum^{+}(q_{uk}^{(1)} - q_{vk}^{(1)})max_k\, b_{kj}^{(m-1)} + \sum^{-}(q_{uk}^{(1)} - q_{vk}^{(1)})min_k\, b_{kj}^{(m-1)}$.

Again, applying the first two equations (4.7,4.8), we get:

$(q_{uj}^{(m)} - q_{vj}^{(m)}) \leq \sum^{+}(q_{uk}^{(1)} - q_{vk}^{(1)})(max_k\, b_{kj}^{(m-1)} - min_k\, b_{kj}^{(m-1)})$

$\leq (1 - N\delta_1)\prod_{i=2}^{m}(1 - N\delta_i) = \prod_{i=1}^{m}(1 - N\delta_i)$.

Thus,

$$(max_i\, q_{ij}^{(m)} - min_i\, q_{ij}^{(m)}) \leq \prod_{i=1}^{m}(1 - N\delta_i). \tag{4.12}$$

Hence, for each natural number n, we have:

$(max_i\ q_{ij}^{(n)} - min_i\ q_{ij}^{(n)}) \leq \prod_{i=1}^{n}(1 - N\delta_i)$.

Next, notice that:

$|q_{ij}^{(n)} - q_{kj}^{(n)}| \leq (max_i\ q_{ij}^{(n)} - min_i\ q_{ij}^{(n)}) \leq \prod_{i=1}^{n}(1 - N\delta_i)$.

Since $\delta_i \geq \epsilon > 0$ infinitely many, passing to the limit as n tends to ∞, the product tends to zero.

Hence, for each $j \in S$ and arbitrary $i, k \in S$, we have:

$\lim_{n\to\infty} |q_{ij}^{(n)} - q_{kj}^{(n)}| = 0$.

That is $\lim_n q_{ij}^{(n)}$ is Weak Ergodic.

Corollary 3 *In the previous theorem, if the sequence of the transition matrices is doubly stochastic; that is:*

$\sum_{i\in S} a_{ij}^{(n)} = 1$, $\forall j \in S$ *and* $\forall n = 1, 2, \cdots$.

Then, the limit of the combination exists and Ergodic.

Indeed, $\lim_{n\to\infty} q_{ij}^{(n)} = 1/N$, $\forall i, j \in S$, *where* N *is the cardinality of* S.

Proof:

By the previous theorem we have Weak Ergodicity, so:

$\forall j \in S$, $q_{ij}^{(n)} = q_j^{(n)} + \epsilon_{ij}^{(n)}$, where $\epsilon_{ij}^{(n)} \to 0$ *as* $n \to \infty$.

Next, $\sum_{i\in S} q_{ij}^{(n)} = \sum_{i\in S}(q_j^{(n)} + \epsilon_{ij}^{(n)}) = 1$

$\Rightarrow\ Nq_j^{(n)} + \sum_{i\in S} \epsilon_{ij}^{(n)} = 1$

$\Rightarrow\ q_j^{(n)} + \epsilon_{ij}^{(n)} = 1/N$.

Thus, $\forall i, j \in S$, $q_{ij}^{(n)} = 1/N + \epsilon_{ij}^{(n)}$.

Now, $q_{ij}^{(n+m)} = \sum_{k\in S} q_{ik}^{(n)} b_{kj}^{(m)}$, where $[b_{ij}^{(m)}]_{i,j\in S}$ are the entries of $A_{n+1}A_{n+2}\cdots A_{n+m}$.

$\Rightarrow\ q_{ij}^{(m+n)} = \sum_{k\in S}(1/N + \epsilon_{ik}^{(n)}) b_{kj}^{(m)}$

$\leq \sum_{k\in S}(max_k\ (1/N + \epsilon_{ik}^{(n)})) b_{kj}^{(m)}$

$\leq (1/N + max_k\ |\epsilon_{ik}^{(n)}|) \sum_{k\in S} b_{kj}^{(m)}$

$= 1/N + max_k\ |\epsilon_{ik}^{(n)}|.$

On the other hand,

$q_{ij}^{(n+m)} \geq \sum_{k \in S}(min_k\ (1/N + \epsilon_{ik}^{(n)}))b_{kj}^{(m)}$

$\geq (1/N - max_k\ |\epsilon_{ik}^{(n)}|) \sum_{k \in S} b_{kj}^{(m)}$

$= 1/N - max_k\ |\epsilon_{ik}^{(n)}|.$

Thus, $|q_{ij}^{(m+n)} - q_{ij}^{(n)}| \leq max_j\ |\epsilon_{ij}^{(n)}| \to 0$ *as* $n \to \infty$, $\forall j \in S$.

Hence, $\lim_{n \to \infty} q_{ij}^{(n)} = 1/N,\ \forall i, j \in S$.

Theorem 7 *Assume we have a sequence of Markov chains on a countable state space, S, whose limit is Ergodic. Assume that:*

$\exists j_0 \in S$ *such that* $\forall i \in S$, $a_{ij_0} \geq \delta > 0$, *where* $[a_{ij}]_{i,j \in S}$ *are the entries of the transition matrix of the limit of the sequence, say* A.

Then, the limit of the combination of this sequence exists and it is Ergodic.

Proof:

Let $(A_n)_{n=1}^{\infty}$ be the transition matrices of the sequence.

Denote the entries of $Q_n = A_1 A_2 \cdots A_n$ by $[q_{ij}^{(n)}]_{i,j \in S}$.

Denote the entries of $A_{m+1} A_{m+2} \cdots A_n$ by $[q_{ij}^{(n-m)}]_{i,j \in S}$.

Denote the entries of A_n by $[p_{ij}^{(n)}]_{i,j \in S}$.

Assume without loss of generality that A is with non-zero entries and that all $A_n's$ are with non-zero entries. Otherwise, the same argument of the finite case.

Define a coupled chain on the state space (S, S) with transition probabilities:

$$P[(X_{n+1}, Y_{n+1}) = (k, l) | (X_n, Y_n) = (i, j)] = p^{(n)}(ij, kl) = p_{ik}^{(n)} p_{jl}^{(n)}.$$

Notice that this coupled chain is irreducible (it is with non-zeros).

Since $p_{ik}^{(n)} \to a_{ik}$, $p_{jl}^{(n)} \to a_{jl} \Rightarrow p_{ik}^{(n)} p_{jl}^{(n)} \to a_{ik} a_{jl}$.

In particular, $p_{ij_0}^{(n)} p_{kj_0}^{(n)} \to a_{ij_0} a_{kj_0} \geq \delta^2 > 0$.

So, if we consider the state (j_0, j_0) then:

$q_{ik}[(X_n, Y_n) = (j_0, j_0) \ \ i.o.] = 1$ and by irreducibility of the combination it is correct for any state (i_0, i_0).

Thus, if T is the first time such that $X_T = Y_T = i_0$, then T is finite with probability 1.

Next,

$q_{ij}[(X_n, Y_n) = (k, l). T = m]=$

$q_{ij}[(X_t, Y_t) \neq (i_0, i_0), t < m, (X_m, Y_m) = (i_0, i_0)] q_{i_0 i_0}[(X_{n-m}, Y_{n-m}) = (k, l)]$

$= q_{ij}[T = m] q_{i_0 k}^{(n-m)} q_{i_0 l}^{(n-m)}$

adding out l gives $q_{ij}[X_n = k, T = m] = q_{ij}[T = m] q_{i_0 k}^{(n-m)}$

adding out k gives $q_{ij}[Y_n = l, T = m] = q_{ij}[T = m] q_{i_0 l}^{(n-m)}$

take $k = l$, add over $m = 1, 2, \cdots, n$

$\Rightarrow \ q_{ij}[X_n = k, T \leq n] = q_{ij}[Y_n = k, T \leq n]$

$\Rightarrow \ q_{ij}[X_n = k] \leq q_{ij}[X_n = k, T \leq n] + q_{ij}[T > n]$

$= q_{ij}[Y_n = k, T \leq n] + q_{ij}[T > n]$

$$\Rightarrow \ q_{ij}[X_n = k] \leq q_{ij}[Y_n = k] + q_{ij}[T > n]. \tag{4.13}$$

Similarly,

$$q_{ij}[Y_n = k] \leq q_{ij}[X_n = k] + q_{ij}[T > n]. \tag{4.14}$$

The above two inequlities $\Rightarrow$

$(q_{ik}^{(n)} - q_{jk}^{(n)}) \leq q_{ij}[T > n]$.

Since T is finite with probability 1, $\Rightarrow \ \lim_{n \to \infty} |q_{ik}^{(n)} - q_{jk}^{(n)}| = 0$.

This means that the combination is independent of the initial state.

Next,

$q_{ij_0}^{(n)} \geq \delta_n > 0$,($\delta_n \to \delta$),thus, by the irreducibility of the combination, $\lim_n q_{ik}^{(n)} > 0, \ \forall k \in S$.

Hence, the limit of the combination exists and Ergodic (by the same argument of the finite case, replacing minimum by infimum and maximum by supremum, the limit is independent of n).

Corollary 4 *In theorems (5) and (7), if the limit of the sequence of the transition matrices A is stable from the first step; that is:*

$\lim_n A^n = A$ *and* $\forall i, j \in S,\ a_{ij} = a_j$.

Then, $\lim_n Q_n = \lim_n A_1 A_2 \cdots A_n = \lim_n A^n = \lim_n A_n = A$.

Proof:

$q_{ij}^{(n)} = \sum_{k \in S} q_{ik}^{(n-1)} p_{kj}^{(n)} \leq \sum_{k \in S} q_{ik}^{(n-1)} (a_j + max_k\, |\epsilon_{kj}^{(n)}|) = a_j + max_k\, |\epsilon_{kj}^{(n)}|$.

On the other hand,

$q_{ij}^{(n)} = \sum_{k \in S} q_{ik}^{(n-1)} p_{kj}^{(n)} \geq \sum_{k \in S} q_{ik}^{(n-1)} (a_j - max_k\, |\epsilon_{kj}^{(n)}|) = a_j - max_k\, |\epsilon_{kj}^{(n)}|$.

Thus,

$|q_{ij}^{(n)} - a_j| \leq max_k\, |\epsilon_{kj}^{(n)}| \to 0$ *as* $n \to \infty$.

That is, $\lim_n q_{ij}^{(n)} = \lim_n p_{ij}^{(n)} = a_j$.

Chapter 5

Conclusion And Comments

In this work, we classified all the possibilities of Non-Stationary Markov chains on a finite state space S.

In theorem (6), we have the condition:

$min_{i,j}\ a_{ij}^{(n)} = \delta_n \geq \epsilon > 0$ for infinitely many n's.

This condition is essential for our proof. We can restate this condition in an equivalent form; that is:

$\exists$ a sequence of integers $(r_i)_{i=1}^{\infty}$ such that $A_{r_i+1}A_{r_i+2}\cdots A_{r_{i+1}}$ is with non-zero entries and the minimum over all the entries is bounded from below (for infinitely many i's).

When the state space, S, is countable, we gave theorem (7) with the condition:

$\exists j_0 \in S$ such that $\forall i \in S,\ a_{ij_0} \geq \delta > 0$.

This condition probably can be weakened.

REFERENCES

[1] M.A Berger, An Introduction To Probability And Stochastic Processes, 1993, Springer-Verlag, New York.

[2] R.N Bhattacharya, E.C Waymire. Stochastic Processes With Applications, 1990, Wiley, New York.

[3] P. Billingsley, Probability And Measure, 1986, Wiley, New York.

[4] L. Breiman, Probability, 1992, Society For Industrial And Applied Mathematics, Philadelphia.

[5] R. Durrett, Probability: Theory And Examples, 1991, Wadsworth And Brooks, Pacific Grove, Calif.

[6] W. Feller, An Introduction To Probability Theory And Its Applications, Vol.1, 1968, Wiley, New York.

[7] R.G Gallager, Discrete Stochastic Processes, 1996, Kluwer Academic Publishers, Boston.

[8] Resnick, I. Sidney, Adventures In Stochastic Processes, 1992, Birkhauser, Boston.

[9] A.N Shiryayev, Probability, 1984, Springer-Verlag, New York.