Stochastic Processes and Markov Chain

Jean Martin Houag

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Applied Mathematics and Computer Sciences

Eastern Mediterranean University
December 2016
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Mustafa Tümer
Director

I certify that this thesis satisfies the requirements as a thesis for the degree of Master of Science in Applied Mathematics and Computer Science.

Prof. Dr. Nazım Mahmudov
Chair, Department of Mathematics

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Applied Mathematics and Computer Science.

Asst. Prof. Dr. Nidai Şemi
Supervisor

Examining Committee

1. Assoc. Prof. Dr. Hüseyin Aktuğlu
2. Asst. Prof. Dr. Mustafa Kara
3. Asst. Prof. Dr. Nidai Şemi

ABSTRACT

Andrey Andreyevich Markov is the founder of the Markov Chain. The Markov Chain is a stochastic process involving modeling over time and space. In sciences or randomize sciences in particular, it is usually important to predict an outcome based on the acquired or previous knowledge of a process. There exits various random processes. The Markov Chain appears as a key technique to deal and model such processes.

Keywords: Stochastic Matrix, Probability Vector, Markov Chain.

ÖZ

Bu çalışmada, öncelikle ıstokastik süreçler tanımlanarak özellikleri verilmiş, sonrasında da örneklerle ve uygulamalarla konu pekiştirilmeye çalışılmıştır. Daha sonra da, Markov Zinciri tanımlanmış ve uygulama alanları verilmiş ve örneklerle desteklenerek konu anlatılmıştır.

Anahtar kelimeler: Istokastik Matris, Olasılık Vektörü, Markov Zinciri.

To my family

ACKNOWLEDGMENT

I would like to thank my supervisor Asst. Prof. Dr. Nidai Şemi, who supported, advised and helped me to prepare this thesis.

A special thank goes to my mother Ngo Njeck Rosalie Charlotte for financial support and advices that have been useful during my Master Program.

I would also say thank to my siblings, Potga Alphonce Marcel, Houag Moise Aurelien, Houag Nicolas Desire and Houag Makong Jeanne Babielle for the love and support they have given to me.

Also special thanks to Ngo Nyobe Marie-michelle, to her love and support she has given to me over the past three years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

A^{-1}	The inverse of the matrix A
A^T	The transpose of the matrix A
$\det(A)$	The determinant of the matrix A
a_{ij}	The ij -th entry of the matrix A
$\lim_{n\to\infty} M_n$	The limit of a sequence of matrices
δ_{ij}	The kronecker delta
$\mathbb{R}$	The field of real numbers
$\mathbb{C}$	The field of complex numbers
$\mathbb{N}$	The field of natural numbers
$M_{n\times n}(F)$	The set of $n\times n$ matrices with entries in F
I_n or I	The $n\times n$ identity matrix
$P(A)$	The probability of the event A
P_{ij}	The probability to move from state i to state j
Ŧ	The field of probability space
Ω	The sample space
$P(A\cap B)$	The probability of the intersection of A and B
$P(X_{n+1} \mid X_n)$	The conditional probability of an event X_{n+1} when X_n is occurred.

Chapter 1

INTRODUCTION PRELEMINARIES
AND SOME REVIEWS

Probability and statistics sciences are usually called the uncertain sciences. The aim in those sciences (probability and statistics) is usually to find a good estimation or to define a process which is a suitable model to the data. The observed variables are usually random. A special case of random processes called Markov Chain is of our interest in this work. The Markov chain plays an important role in various fields of sciences from social sciences to computer sciences.

1.1 Definition

A sequence of experiments is called stochastic process. A stochastic process is a mathematical model that evolves over time in a probabilistic manner. If the outcomes of an experiment depend on only outcomes of previous experiment, then such a process is called **Markov Chain** or **Markov Model** or **Markov Process**. In other words, the next state of a Markov Chain (Markov Model or Markov Process), the system depends only on present state, not on preceding states.

We will clarify this definition with theorems, properties and some examples.

1.2 History

Markov Chain was initially introduced by Russian Mathematician called Andrey Markov 1906. Since then it has had many fields of applications. Below are some keys dates when the Markov Chain has efficiently affecting some particular topics in sciences. The keys dates are mostly considered from the application to health sciences.

In the year 1986, Hillis et Al., and Jain, show that the Markov Chain was a perfect alternative of evaluating a time-event data set. This developed the idea of the application of the Markov chain to many others sciences. Health sciences researchers and practitioners also got interested in the Markov Chain or Markov process. Explicitly, Marshall and Jones applied the techniques for the study of diabetic retinopathy in 1995. Whereas Silverstein, Shaubel applied in the studies of renal disease and papillona virus respectively in 1998. In the year 1997, Norris defined the Markov properties. The state space under measurement is effective to classify Markov Chain. Therefore there exists finite space or discrete Markov process, which is defined under the assumption that there is a finite number of states to be reached by the process. In the either case, the process is described as an infinite or continuous process. The mentioned classification was introduced by Bard and Jesen in 2002. In a similar way, a classification based on time intervals leads to the name discrete interval and continuous interval respectively. In many references, the term Markov process is used for continuous – time process whereas Markov Chain is used for discrete – time process. This means, the name Markov process may eventually refer to all chains and processes.

1.3 Plan

We briefly defined above the Markov Chain and we all gave a little review about the Markov Chain’s history. In the remaining of our work, we discussed more deeply about the topic of our interest called Markov Chain. To do so, our work is divided into several chapters. Some chapters are considered as preliminaries to others chapters. Our discussion started with a chapter based on reviews of probability and algebraic theories which are absolute necessities to discuss about the Markov Chain. It follows by a chapter on probability vectors and stochastic matrices. After the latest mentioned chapter, we move to the heart of our task which is the main chapter focusing on the Markov Chain. We finally conclude our work by given a briefly review of what we did so far.

Chapter 2

REVIEW OF PROBABILITY AND ALGEBRAIC THEORY

In this part we shall focus on some important notations and basic concepts of probability theory such as probability space, Ŧ-field, conditional probability and matrix theory such as matrix diagonalization and matrix limits [7, 11].

2.1 Definitions of Probability Space and Ŧ-fields

The probability space will be explained by using the system language of measure theory.

Definition 2.1. (Sample Space (Ω))

The set of all possible outcomes of a random experiment is called sample space.

Example 1

The possible outcomes of the experiment to a toss a die are 1, 2, 3, 4, 5 or 6. Therefore the sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Definition 2.2. (Event Space (E))

The outcomes of an experiment are called events of the experiment.

Example 2

We can define an event as the die shows an odd number. In this case the space event is D = {1, 3, 5}.

Definition 2.3. (Probability Measure (P))

Probability measure P is a function defined as P: $\Omega \rightarrow [0, 1]$ such that the following axioms are satisfied.

1. $P(\Omega) = 1$

2. $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$. When E_1 and E_2 are not disjoint.

3. For events $E_1, E_2 \in \Omega$, where $E_1 \cap E_2 = \varnothing$ then

$$P(E_1 \cup E_2) = P(E_1) + P(E_2)$$

More generally,

$$\mathrm{P}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E)$$

Where $E_i \cap E_j = \varnothing$ and $i \neq j$. [11,14]

2.2 Conditional Probability

Definition 2.4. The probability of an event A under a condition that an event B has already occurred is called the conditional probability of A under B [11]. This conditional probability of A under the condition B, is denoted by $P(A|B)$ and it is defined by

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Properties (Conditional Probability)

1) For some B fixed, A_1 and A_2 are mutually exclusive, then

2) $P\left((A_1 \cup A_2)|B\right) = P(A_1|B) + P(A_2|B)$

3) In general, $P(\bigcup_{i=1}^{n} A|B_i) = \sum_{i=1}^{n} P(A_i|B)$ where $A_i \cap A_j \neq \emptyset$ when $i \neq j$.

Note: $P(A|B \cup C) \neq P(A|B) + P(A|C)$ and also $P(A|B) \neq P(B|A)$.

Example 3

An amphitheater in Eastern Mediterranean University we have regrouped the following data. [7]

Table 1: Smokers data

	Male	Female	Total
Smoke	82	38	120
No smoke	26	54	80
Total	108	92	200

What is the probability that an student chosen at randomly,

1. smoke cigarette?
2. is male and smoke cigarette?
3. is female and does not smoke?

Solutions:

1. $P(Smoke) = \frac{120}{200}$ = 0.6 or 60%

2. $P(Smoke|Male) = \frac{P(smoke \cap male)}{p(male)} = \frac{82}{108} \approx$ 0.76 or 76%

3. $P(\,No\,Smoke \mid Female\,) = \dfrac{p(nosmoke \cap female)}{p(\,female)} = \dfrac{54}{92} \approx 0.59$ or 59%

2.3 Independence of an Event

Definition 2.5. Two events; A and $B \in \Omega$ are independent, if

$$P(A \cap B) = P(A).P(B)$$

We may also define that A and B are independent if

$$P(A \mid B) = P(A) \text{ and } P(B \mid A) = P(B).$$

In general, if $E_1, E_2, \ldots E_n \in \Omega$ are mutually exclusive, then

$$P(E_1 \cap E_2 \cap \ldots E_n) = \prod_{i=1}^{n} P(E_i) = P(E_1).P(E_2)\ldots.P(E_n).$$

Example 4

Your supervisor invites you to a restaurant, saying it open sometime on weekend between 4 in afternoon and midnight, but won't say more. What is the probability that it starts on Saturday between 6 and 8 at night?

Solution: Time between 4 and midnight we have 8 hours, but we want between 6 and 8 which are 2 hours.

$P(\text{time}) = \dfrac{2}{8} = 0.25$

Day: we have 2 days on the weekend, so

$P(\text{Saturday}) = \dfrac{1}{2} = 0.5$

Therefore, P(Saturday and your time) = P(Saturday) . P(your time) = 0.5x0.25 = 0.125 or 12.5%

2.4 Elementary Matrices Operations

2.4.1Matrix Multiplication

Definition 2.6. A matrix $A = (a_{ij})$ is said to have dimension $m_A \times n_A$ if and only if it has m_A rows and n_A columns [4,6].

Definition 2.7. Let matrix $A = (a_{ij})$ having dimension $m_A \times n_A$ and $B = (b_{ij})$ be $m_B \times n_B$ matrix. Then if $n_A = m_A$ the matrix product $A \times B$ is defined by

$$C = A \times B = c_{ij} = \sum_{k=1}^{n} a_{ik} . b_{kj}$$

Properties of Matrix Product

P1) In general the product of two matrices is not commutative, i.e. in general $AB \neq BA$.

P2) The matrix product AB is defined if and only if the number of columns of A equals the number of rows of B, i.e. if $n_A = m_B$

P3) If the multiplication can be performed (that is $n_A = m_B$), the matrix product C will be a matrix having dimension $m_A \times n_B$. [4,6].

Example 5

Let

$$A = \begin{pmatrix} 0 & 3 & 1 \\ 1 & 2 & 0 \\ 2 & 1 & 3 \end{pmatrix}, \text{ and } B = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}.$$

Then

$$A\times B=\begin{pmatrix}0&3&1\\1&2&0\\2&1&3\end{pmatrix}\times\begin{pmatrix}1\\0\\2\end{pmatrix}=\begin{pmatrix}0\times1+0\times3+2\times1\\1\times1+0\times2+2\times0\\2\times1+0\times1+2\times3\end{pmatrix}=\begin{pmatrix}2\\1\\8\end{pmatrix}.$$

But $B\times A$ is not defined.

2.4.2 Determinant of Order 2

Definition 2.8. Let

$$\mathrm{A}=\begin{pmatrix}a&b\\c&d\end{pmatrix}\in M_{2\times2}(\mathbb{R}),$$

then the **determinant** of A is denoted by $|A|$ or $\det(A)$ and it is defined by

$$ad-bc\ [6].$$

Example 6

For a 2×2 matrix

$$A=\begin{pmatrix}2&4\\5&3\end{pmatrix}\in M_{2\times2}(\mathbb{R}),$$

we have

$$\det(A)=2\times3-4\times5=-14.$$

2.4.3 Determinants of Order *n*

In this section, we extend the definition of the determinant to $n\times n$ matrices for $n\geq3$. It is convenient to introduce the following definition:

Definition 2.9. Let $A\in M_{n\times n}(F)$ be a square matrix with $n\geq2$ and let B_{ij} denote the $(n-1)\times(n-1)$ matrix obtained from A by deleting row i and column j. The scalar value

$$\mathrm{C}_{ij}=(-1)^{i+j}\det(B_{ij})$$

is called a cofactor of $A\in M_{n\times n}(F)$, in row i, column j.

Definition 2.10. Let $A \in M_{n\times n}(F)$ be a square matrix then the matrix defined by $\mathrm{C} = (\mathrm{C}_{ij})$ where C_{ij} is the cofactor of $A \in M_{n\times n}(F)$, in row i, column j, is called the **cofactor matrix** of $A \in M_{n\times n}(F)$.

Definition 2.11. (Determinant Order *n*)

Let $A \in M_{n\times n}(F)$. If $n = 1$, so that $A = (A_{11})$, we define $\det(A) = A_{11}$.

For, $n \geq 2$, the scalar value $\det(A)$ is defined by;

$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j} A_{1j}.\det(B_{1j}).$$

or

$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j} A_{2j}.\det(B_{2j})$$

$$\vdots$$

$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j} A_{nj}.\det(B_{nj}).$$

Example 7

Compute the determinant of the matrix A

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix} \in M_{3\times 3}(\square).$$

Using cofactor expansion along the first row, we obtain

$$\det(A) = (-1)^{1+1} A_{11} \det(B_{11}) + (-1)^{1+2} A_{12} \det(B_{12}) + (-1)^{1+3} A_{13} \det(B_{13})$$

$$= (-1)^{2}(0).\det\begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix} + (-1)^{3}(1).\det\begin{pmatrix} 3 & 5 \\ 6 & 8 \end{pmatrix} + (-1)^{4}(2).\det\begin{pmatrix} 3 & 4 \\ 6 & 7 \end{pmatrix}$$

$$= 0 + (-1)(-6) + (2)(-3)$$

$$= 6 - 6$$

$$= 0.$$

2.4.4 Transpose of Matrix

Definition 2.12. Let $A \in M_{n\times n}(F)$ be any matrix and let B be the matrix obtained from A by interchanging rows by columns. The matrix B is called **transpose** of A and denoted $B = A^T$. [4,6].

Example 8

Find the transpose of the matrix

$$K = \begin{pmatrix} 1 & 0 & 2 \\ 4 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} \in M_{3\times 3}(\mathbb{R})$$

Solution.

$$K^T = \begin{pmatrix} 1 & 4 & 0 \\ 0 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix}$$

2.4.5 Adjoint of Matrix

Definition 2.13. Let A be $n \times n$ matrix and let $C = (C_{ij})$ be the cofactor matrix of A then the transpose of $C = (C_{ij})$ is called the adjoint matrix of A and denoted by $AdjA$.

Example 9

Compute $Adj(A)$, where

$$A=\begin{pmatrix}1 & 0 & 2\\ 1 & 2 & -1\\ 3 & 0 & 2\end{pmatrix}\in M_{3\times 3}(\square)$$

Solution: It is easy to see that,

$$C=\begin{pmatrix}4 & -5 & -6\\ 0 & -4 & 0\\ -4 & 3 & 2\end{pmatrix} \text{ and } adj(A)=C^T=\begin{pmatrix}4 & 0 & -4\\ -5 & -4 & 3\\ -6 & 0 & 2\end{pmatrix}.$$

2.4.6 Inverse of a Matrix

Definition 2.14. Let A be a square matrix which is non singular (i.e. $\det(A)\neq 0$), then the matrix denoted by A^{-1} which satisfies $A.A^{-1}=A^{-1}.A=I$, where I is the identity matrix, is called inverse of A.

Properties of Inverse Matrix

P1. Let $A=\begin{bmatrix}a & b\\ c & d\end{bmatrix}$ be 2×2 matrix with $\det(A)\neq 0$, where a, b, c and d are real or complex numbers then the **inverse** of A is

$$A^{-1}=\frac{1}{ad-bc}\begin{bmatrix}d & -b\\ -c & a\end{bmatrix}=\frac{1}{ad-bc}AdjA. \text{ [6]}$$

P2. In general, if A is $n\times n$ matrix with $n\geq 3$ and $\det(A)\neq 0$ then

$$A^{-1}=\frac{adj(A)}{\det(A)}$$

Example 10

Compute the inverse A^{-1} of the following matrix

$$A=\begin{pmatrix}1 & 0 & 2\\ 1 & 2 & -1\\ 3 & 0 & 2\end{pmatrix}$$

Solution:

By the property P2,

$$A^{-1} = \frac{adj(A)}{\det(A)}$$

$$= -\frac{1}{8}\begin{pmatrix} 4 & 0 & -4 \\ -5 & -4 & 3 \\ -6 & 0 & 2 \end{pmatrix}.$$

2.4.7 Power of a Matrix

Definition 2.15. Let A be a square matrix then the power A^n of A where n is a non-negative integer, is defined as matrix product of copies of A.

$$A^n = \underbrace{A \times A \times \ldots \times A}_{n}.$$

In particular, the matrix to the zeroth power is **identity** matrix denoted $A^0 = I$.

Example 11

Compute A, A^2 and A^3 for the matrix A given below:

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

Solution :

$$A^2 = A.A = \begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -5 & 4 \\ -6 & -5 \end{pmatrix}$$

$$A^3 = A.A.A = A.A^2 = \begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} -5 & 4 \\ -6 & -5 \end{pmatrix} = \begin{pmatrix} -17 & -6 \\ 9 & -17 \end{pmatrix}$$

$$A^4 = A.A.A.A = A.A^3 = \begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} -17 & -6 \\ 9 & -17 \end{pmatrix} = \begin{pmatrix} 1 & -40 \\ 60 & 1 \end{pmatrix}.$$

In the next paragraph, we will consider diagonalization method which is a useful method to compute the large numbers of powers of a matrix.

2.5 Diagonalization of Matrix

The diagonalization problem of a square matrix is directly related with the concept of eigenvalue and eigenvector. Therefore, in the first part of this section we will focus on eigenvalues and eigenvectors.

2.5.1 Eigenvalues and Eigenvectors

Definition 2.16. Let A be a matrix in $M_{n\times n}(F)$. A non zero vector $x \in F^n$ is called an **eigenvector** of A if A $x = \lambda x$ for some scalar λ.The scalar λ is called **eigenvalue** corresponding to the eigenvector x.

Theorem 2.5.1: Let A $\in M_{n\times n}(F)$. Then a scalar λ is an eigenvalue of A if and only if

$$\det(A - \lambda I_n) = 0$$

Definition 2.17. Let A $\in M_{n\times n}(F)$. Then the polynomial $f(\lambda) = \det(A - \lambda I_n)$ is called **characteristic polynomial** of A.

Definition 2.18. Let A $\in M_{n\times n}(F)$. Then the zeros of the characteristic polynomial are called the **eigenvalues** of the matrix *A*.

Example 12

Let

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix} \in M_{3\times 3}(\mathbb{R}).$$

Find the eigenvalues and the eigenvectors of the matrix A.

Solution:

The characteristic polynomial of A is the following equation,

$$f(\lambda) = \det\left(\begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right).$$

Thus,

$$\Leftrightarrow \begin{vmatrix} 1-\lambda & 1 & 0 \\ 0 & 2-\lambda & 2 \\ 0 & 0 & 3-\lambda \end{vmatrix} = 0 \Rightarrow (1-\lambda)(2-\lambda)(3-\lambda) = 0.$$

Then

$$\lambda_1 = 1 \,;\, \lambda_2 = 2 \text{ and } \lambda_3 = 3$$

are the eigenvalues of A. Let us find corresponding eigenvectors.

To find the eigenvectors x, corresponding to the eigenvalue we will replace λ by eigenvalues in A $x = \lambda x$.

Let $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. For $\lambda = 1$, we have

$$A\,x = x \Rightarrow (A - I)\,x = 0$$

$$\Rightarrow \left(\begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

$$\Rightarrow \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

Then

$$x_1 = p,\ x_2 = 0 \text{ and } x_3 = 0 \text{ where } p \text{ is the parameter}$$

$$\Rightarrow \text{x} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \text{ when we assign p = 1.}$$

Similarly, for λ = 2, we have

$$x_1 = p,\ \ x_2 = p \text{ and } x_3 = 0$$

$$\Rightarrow \text{x=} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \text{ when p = 1.}$$

For λ = 3, we have

$$x_1 = p,\ x_2 = 2p \text{ and } x_3 = p$$

$$\Rightarrow \text{x} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

Hence, the set of eigenvectors is $S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}$.

2.5.2 Diagonalizability

We presented the diagonalization problem and we can observe that not all matrices are diagonalizable. Although we are able to diagonalize matrices and even to obtain necessary and sufficient condition for diagonalizability of a matrix *A*.

2.5.2.1 Diagonal Matrix

Definition 2.19. Let $D = (c_{ij})$ be a square matrix. If D is of the form

$$D = \begin{pmatrix} c_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_n \end{pmatrix},$$

then, it is called a diagonal matrix.

Note that, a diagonal matrix *D*, is also denoted by $D = diag(c_1, c_2, \ldots, c_n)$.

Properties (Diagonal Matrices)

P1) The determinant of a diagonal matrix is the product of elements of diagonal. i.e.

$$\text{if } D = \begin{pmatrix} c_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_n \end{pmatrix} \text{ then } \det(D) = c_1.c_2.\ldots.c_n .$$

P2) Let D be the diagonal matrix and n be a positive integer. The n^{th} power of diagonal matrix D equals to

$$D^n = \begin{pmatrix} c_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_n \end{pmatrix}^n = \begin{pmatrix} c_1{}^n & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_n{}^n \end{pmatrix}.$$

2.5.2.2 Diagonalizable Matrix

Definition 2.20. Let A be a $n \times n$ matrix. A is **diagonalizable** if it can be written as $A = P.D.P^{-1}$, where D is diagonal matrix, with entries eigenvalues of A and P is the $n \times n$ matrix consisting of the eigenvectors corresponding to the eigenvalues in D i.e.

$$D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}, \text{ and } P = (v_1, v_2, \ldots, v_n)$$

Where $v_1, v_2, \ldots, v_n$ are eigenvectors of A (written as the column vectors) and P^{-1} is the inverse of P.

Theorem 2.1. Let A be an $n \times n$ matrix. A is diagonalizable if and only if A has n linearly independent eigenvectors, i.e. if the matrix rank of the matrix formed by eigenvectors is n. [6]

Example 13

Consider the matrix A given by,

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix} \in M_{3\times3}(\square).$$

We can rewrite A as

$$A = P.D.P^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} . \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} . \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

where

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \ P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \text{and}\ P^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Remark: When the size of the matrix is too high, it will be difficult to write the matrix by using these three parts P, D and P^{-1}, in this case, we will use the applications as **Matlab, Scilab,** etc. to find the eigenvalues and eigenvectors.

Why it's interesting to know a diagonalization of a matrix A?

In the next chapters, some of time it will be necessary to compute the great power of matrix , for instance, we will need to evaluate the A^n, where n is a large natural number. It is not applicable to evaluate A^n. If the matrix is diagonalizable, we will use the transformation of matrix A as $A = P.D.P^{-1}$ then $A^n = (P.D.P^{-1})^n = P.D^n.P^{-1}$. Since D is diagonal matrix it is easy to evaluate D^n.

2.6 Matrix Limit

In this section we will study the limit of a sequence of matrices $M, M^2, \cdots, M^n$ where M is a square matrix with complex entries. The limit of sequence of complex $\{z_n : n = 1, 2, 3, \ldots\}$ can be defined in terms of limits of the sequences of real and imaginary numbers. Let $z_n = a_n + ib_n$ with a_n and b_n are real numbers and i is the complex number such that $i = \sqrt{-1}$ ($i \in$). Then

$$\lim_{n\to\infty} z_n = \lim_{n\to\infty} a_n + i \lim_{n\to\infty} b_n$$

Provide that $\lim_{n\to\infty} a_n$ and $\lim_{n\to\infty} b_n$ exist.

Definition 2.21. Let $L, M, M^2, \cdots, M^n \cdots$, be $n \times n$ matrices with the complex entries. The sequence $M_1, M_2, \ldots$ is said to **converge** to the matrix L, if

$$\lim_{n\to\infty} (M_n)_{ij} = L_{ij}, \text{ for all } 1 \le i, j \le n.$$

In this case, we write

$$\lim_{n\to\infty} M_n = L$$

and L is called the **limit** of the sequence. [1,6]

Example 14

Let M_n be the sequence

$$M_n = \begin{pmatrix} \left(1+\frac{1}{n}\right)^n & \left(\frac{1}{3}\right)^n \\ 3-\frac{2}{n} & \frac{n^2}{3n^2-1}+2i \end{pmatrix} \in M_{2\times 2}(\mathbb{C}),$$

then

$$\lim_{n\to\infty} M_n = \begin{pmatrix} \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n & \lim_{n\to\infty}\left(\frac{1}{3}\right)^n \\ \lim_{n\to\infty}\left(3-\frac{2}{n}\right) & \lim_{n\to\infty}\left(\frac{n^2}{3n^2-1}+2i\right) \end{pmatrix}$$

Hence,

$$\lim_{n\to\infty} M_n = \begin{pmatrix} e & 0 \\ 3 & \frac{1}{3}+2i \end{pmatrix} = L .$$

Where **e** is the base of the natural logarithm.

Theorem 2.2. Let $M_1, M_2, \ldots$ be a sequence of $n\times n$ matrices with complex entries and L be its limit. Then, for any $r\times n$ matrix P and $p\times s$ matrix Q,

we have

$$\lim_{n\to\infty} PM_n = PL \text{ and } \lim_{n\to\infty} M_n Q = LQ .$$

Proof. By the definition of limit and properties of matrix multiplication we have,

$$\lim_{n\to\infty}(PM_n) = \lim_{n\to\infty}\sum_{k=1}^{n} P_{ik}(M_n)_{kj} \text{ where } 1\le i\le r \text{ and } 1\le j\le p$$

$$=\sum_{k=1}^{n} P_{ik} \lim M_{kj} = \sum_{k=1}^{n} P_{ik} L_{kj} = (PL)_{ij}$$

Hence,

$$\lim_{n\to\infty} PM_n = PL \, .$$

Similarly, we can prove that

$$\lim_{n\to\infty} M_n Q = LQ \, .$$

Corollary 2.1. Let M be a $n \times n$ matrix with complex entries where

$$\lim_{n\to\infty} M^n = L \, .$$

Then for any invertible matrix T with complex entries ,

$$\lim_{n\to\infty} (TMT^{-1})^n = TLT^{-1}.$$

Proof. By definitions of power of matrix and matrix limit we have,

$$(TMT^{-1})^n = (TMT^{-1})(TMT^{-1})...(TMT^{-1}) = TM^nT^{-1}$$

$$\Rightarrow \lim_{n\to\infty} (TMT^{-1})^n = \lim_{n\to\infty} TM^nT^{-1} = T\,(\lim_{n\to\infty} M^n)T^{-1} = TLT^{-1} \, .$$

End of proof.

Chapter 3

PROBABILITY VECTORS AND STOCHASTIC MATRICES

In this chapter, we are going to give a new concept to the vectors and matrices which are related to **Markov Chain**. These feature vector sand matrices allow us to model the socio-economic and scientific problems in the context to understanding, predict, solve and anticipate. [1,2,9].

3.1 Probability Vector

Definition 3.1. Let $v=(v_1,v_2,...,v_n)$ be a vector. In mathematics, especially in statistics, a vector v is called **probability vector** or **stochastic vector** if the entries are non-negative and their sum equals to 1. i.e. $\sum_{i=1}^{n} v_i = 1$, and each individual component v_i must have a probability value which is $0 \le v_i \le 1$ for all $i = 1, 2, \ldots, n$. [2,12,14].

Example 1

The vectors; *u*, *v*, *w* and *t* given below are all probability vectors.

$$u=\begin{bmatrix}0.15 & 0.25 & 0.6\end{bmatrix}, v=\begin{bmatrix}0.20\\0.30\\0.50\end{bmatrix}, w=\begin{bmatrix}0.23 & 0.77\end{bmatrix},$$

$$t=\begin{bmatrix}0.12 & 0 & 0.28 & 0.6\end{bmatrix}.$$

Properties (Probability Vector)

Let p be a probability vector of the form; $p = [p_1, p_2, \ldots, p_n]$ where p has n components, then it satisfies the following;

- The mean of vector p is $\frac{1}{n}$. [2,9]

  (The mean of probability vector does not depend on the values of the components but with the number of entries.)
- The longest probability vector has the value 1 in a single component and 0 in all others and its length is 1.
- The shortest probability vector has the value $\frac{1}{n}$ as each component of the vector and its length is $\frac{1}{\sqrt{n}}$. [13,17]
- The length of a stochastic vector to $\sqrt{n\sigma^2 + \frac{1}{n}}$ where σ^2 is the variance of the probability vector.

Example 2

i) Let t be the following vector;

$$t = \begin{bmatrix} 0.12 & 0 & 0.28 & 0.6 \end{bmatrix},$$

then the mean of the vector t is equals to $\frac{1}{4}$.

ii) Given the vector k of the form

$$k = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

then k is an longest probability vector.

iii) Given the vector b of the form

$$b=\begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix},$$

then b is an shortest probability vector.

3.2 Transition Matrix

A **stochastic matrix** or **transition matrix** describes a **Markov Chain** X_n over a finite state space ***S***, then there are several different definitions and types of transition matrix or **probability matrix.**

Definition 3.2. A square matrix is called **Right Transition Matrix** if all entries are non-negative and the sum of each row equals to 1. [1,15]

Definition 3.3. A square matrix is called **Left Transition Matrix** if all entries are non-negative and the sum of each column equals to 1. [15,16]

Definition 3.4. A square matrix is called **Double Transition Matrix** if all entries are non-negative and each row and column sums equal to 1. [1,10]

Example 3

Consider the following matrices

$$M_1=\begin{bmatrix} 0 & 0.25 & 0.25 & 0.5 \\ 0 & 0 & 0 & 1 \\ 0.1 & 0 & 0 & 0.9 \\ 0 & 0.28 & 0.62 & 0.1 \end{bmatrix} M_2=\begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0.25 & 0.4 & 0.3 & 0 & 0 \\ 0 & 0.6 & 0.3 & 0 & 0 \\ 0 & 0 & 0.2 & 1 & 0 \\ 0.75 & 0 & 0.2 & 0 & 0 \end{bmatrix} M_3=\begin{bmatrix} 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \\ 0.5 & 0 & 0.5 \end{bmatrix}$$

M_1, M_2 and M_3 are right, left and doubly transition matrices, respectively.

We may also represent the transition matrix by the graph which called **transition diagram.**

Example 4

Given the left transition matrix $T = \begin{pmatrix} 0.3 & 0.4 & 0.5 \\ 0.3 & 0.4 & 0.3 \\ 0.4 & 0.2 & 0.2 \end{pmatrix}$, then it can also be represented by the follow graph:

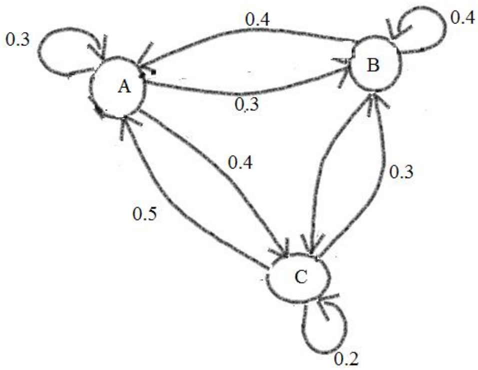


Figure 1: Transition diagram

Definition 3.5. The graph given above is called **transition diagram.**

Theorem 3.1. Let A be an $n \times n$ matrix having real non-negative entries and let v be a column vector in $\mathbb{R}^n$ having non-negative coordinates, and $u \in \mathbb{R}^n$ be the column vector in which each coordinate equals to 1, i.e. $u = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$, [6]

then

1. v is probability vector if and only if $u^T v = (1)$

2. A is transition matrix if and only if $A^T u = u$.

Proof.

1. $\Rightarrow$) Let $v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ where $\sum_{i=1}^{n} v_i = 1$ and let $u = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$ be an nx1 column vector, then

$$u^T v = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = (v_1 + v_2 + ... + v_n) = (1)$$

$\Leftarrow$) let $u^T v = (1)$. We will prove that v is probability vector i.e. we will show that

$$\sum_{i=1}^{n} v_i = 1 \text{ for all } i = 1, 2, \ldots, n$$

$$u^T v = (1) \Rightarrow (u^T v)^T = (1)^T$$

$$\Rightarrow v^T u = (1)$$

$$\Rightarrow \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = (v_1 + v_2 + ... + v_n) = (1)$$

Therefore, v is probability vector.

2. $\Rightarrow$) Let A be transition matrix. We will prove that $A^T u = u$

Just make a precision in this case. We will consider A as a double transition matrix, i.e. sum of each row and sum of each column is equal to 1.

Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \text{ and } u = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$A^T u = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \cdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} + a_{21} + \cdots + a_{n1} \\ a_{12} + a_{22} + \cdots + a_{n2} \\ \vdots \\ a_{1n} + a_{2n} + \cdots + a_{nn} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = u$$

$\Leftarrow$) let

$$A^T u = u$$

We will prove that M is transition matrix.

$$A^T u = u$$

$$\Rightarrow (A^T u)^T = u^T \Rightarrow u^T A = u^T$$

$$\Rightarrow \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} a_{11} + a_{21} + \cdots a_{n1} & a_{12} + a_{22} + \cdots + a_{n2} \cdots & a_{1n} + a_{2n} + \cdots + a_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}$$

$$\Rightarrow a_{11} + a_{21} + \ldots + a_{n1} = 1;\ a_{12} + a_{22} + \ldots + a_{n2} = 1;\ a_{1n} + a_{2n} + \ldots + a_{nn} = 1.$$

i.e.

$$\sum_{i=1}^{n} a_{ij} = 1.$$

Therefore, A is transition matrix.

Corollaries 3.1

A) The product of two transitions matrices is a transition matrix. In particular, any power of transition matrix is a transition matrix (but error can appear because of truncation.)

B) The product of a transition matrix and probability vector is a probability vector.

Proof. To prove the corollary we will use an algebraic definition of **endomorphism function** and the previous theorem.

A1) The order matrix n expresses an endomorphism f in the canonical basis, and we know that the coefficients of the product matrix are positive; more f_1, f_2 being endomorphisms of these matrices

$f_1 \circ f_2(u) = f_1[f_2(u)] = f_1(u) = u$ by the previous theorem, where u is a column vector in which each coordinate equals to 1.

A2) Let A be a transition matrix. We will use proof by **induction** to show that A^n is also an transition matrix.

For $n = 0$, we have

$$A^0 = I_n \text{. Where } I_n = \begin{cases} 1 & if \quad i = j \\ 0 & if \quad i \neq j \end{cases} \text{ by convention. } A^0 \text{ is transition matrix.}$$

For n=1, $A^1 = A$ is stochastic by hypothesis.

We assume that it's true for A^{n-1} and we will prove that it also true for $(A^n)_{ij}$.

For all i, j and for i fixed we have

$\sum_j (A^n)_{ij} = \sum_j (\sum_k (A^{n-1})_{ik} \times A_{kj}) = \sum_j [(A^{n-1})_{ik} A_{kj}] = 1$. End of proof.

B1) Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \text{ and } v = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix},$$

be a transition matrix and a transition vector, respectively. We will prove that $v.A$ is a probability vector.

$$v\,.\,A=\begin{pmatrix}v_1 & v_2 & \cdots & v_n\end{pmatrix}.\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \cdots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}=$$

$$=\begin{pmatrix}\mathrm{a}_{11}v_1+a_{21}v_2+...+a_{1n}v_n & a_{12}v_1+a_{22}v_2+...+a_{n2}v_n... & a_{1n}v_1+a_{2n}v_2+...+a_{nn}v_n\end{pmatrix}.$$

When we put each v_i in factor, we obtain

$$\left(v_1[a_{11}+a_{12}+...+a_{1n}]+v_2[a_{21}+a_{22}+...+a_{2n}]+...+v_n[a_{n1}+a_{n2}+...+a_{nn}]\right).$$

We know that

$\sum_{i=1}^{n} v_i=1$ and $\sum_{j=1}^{n} a_{ij}=1$ hence the result.

Example 5

Let

$$M=\begin{pmatrix}0 & \frac{2}{3} & \frac{1}{3}\\ 0.5 & 0 & 0.5\\ \frac{3}{4} & \frac{1}{4} & 0\end{pmatrix};\quad N=\begin{pmatrix}0.65 & 0.28 & 0.07\\ 0.15 & 0.67 & 0.18\\ 0.12 & 0.36 & 0.52\end{pmatrix}\text{ and } v=\begin{pmatrix}0.5 & 0.5 & 0\end{pmatrix},$$

where M and N are transition matrices and v is a probability vector.

1. $$M.N=\begin{pmatrix}0 & \frac{2}{3} & \frac{1}{3}\\ 0.5 & 0 & 0.5\\ \frac{3}{4} & \frac{1}{4} & 0\end{pmatrix}.\begin{pmatrix}0.65 & 0.28 & 0.07\\ 0.15 & 0.67 & 0.18\\ 0.12 & 0.36 & 0.52\end{pmatrix}=\begin{pmatrix}0.14 & 0.5733 & 0.2933\\ 0.385 & 0.33 & 0.295\\ 0.525 & 0.377 & 0.0975\end{pmatrix}.$$

We can verify that the sum of each row is equal to 1 so the matrix $M.N$ is also a transition matrix.

2. $v.M = \begin{pmatrix} 0.5 & 0.5 & 0 \end{pmatrix} . \begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{3} \\ 0.5 & 0 & 0.5 \\ \frac{3}{4} & \frac{1}{4} & 0 \end{pmatrix} = \begin{pmatrix} 0.25 & 0.3333 & 0.4167 \end{pmatrix}$

0.25+0.3333+0.4167=1.

Therefore $v.M$ is also a transition vector.

3. $M^2 = M.M = \begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{3} \\ 0.5 & 0 & 0.5 \\ \frac{3}{4} & \frac{1}{4} & 0 \end{pmatrix} . \begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{3} \\ 0.5 & 0 & 0.5 \\ \frac{3}{4} & \frac{1}{4} & 0 \end{pmatrix} = \begin{pmatrix} 0.5833 & 0.0833 & 0.3333 \\ 0.375 & 0.4583 & 0.1667 \\ 0.125 & 0.5 & 0.375 \end{pmatrix}$

$$M^3 = M.M^2 = \begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{3} \\ 0.5 & 0 & 0.5 \\ \frac{3}{4} & \frac{1}{4} & 0 \end{pmatrix} . \begin{pmatrix} 0.5833 & 0.0833 & 0.3333 \\ 0.375 & 0.4583 & 0.1667 \\ 0.125 & 0.5 & 0.375 \end{pmatrix}$$

$$= \begin{pmatrix} 0.2917 & 0.4722 & 0.2361 \\ 0.3542 & 0.2917 & 0.3542 \\ 0.5313 & 0.1771 & 0.2917 \end{pmatrix}$$

$$M^4 = M.M^3 = \begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{3} \\ 0.5 & 0 & 0.5 \\ \frac{3}{4} & \frac{1}{4} & 0 \end{pmatrix} . \begin{pmatrix} 0.2917 & 0.4722 & 0.2361 \\ 0.3542 & 0.2917 & 0.3542 \\ 0.5313 & 0.1771 & 0.2917 \end{pmatrix}$$

$$= \begin{pmatrix} 0.4132 & 0.2535 & 0.3333 \\ 0.4115 & 0.3247 & 0.2639 \\ 0.3073 & 0.4271 & 0.2656 \end{pmatrix} \ldots$$

We can check that all the powers of matrix M are transition matrices.

3.3 Regular Transition Matrix

Definition 3.6. A transition matrix P is regular if some integer power of it has all positive entries, i.e. for some $n \in \mathbb{N}$, the entries of P^n are positive. [7,9]

i.e. if $P^n = (p_{ij})$ then $p_{ij} > 0$ for all $i, j = 1, 2, \ldots, n$.

Example 6

The transition matrix;

$$M = \begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{3} \\ 0.5 & 0 & 0.5 \\ \frac{3}{4} & \frac{1}{4} & 0 \end{pmatrix},$$

of the previous example is regular. In fact when we compute the different power of this transition matrix, we obtain

$$M^2 = \begin{pmatrix} 0.5833 & 0.0833 & 0.3333 \\ 0.375 & 0.4583 & 0.1667 \\ 0.125 & 0.5 & 0.375 \end{pmatrix}, M^3 = \begin{pmatrix} 0.2917 & 0.4722 & 0.2361 \\ 0.3542 & 0.2917 & 0.3542 \\ 0.5313 & 0.1771 & 0.2917 \end{pmatrix},$$

$$M^4 = \begin{pmatrix} 0.4132 & 0.2535 & 0.3333 \\ 0.4115 & 0.3247 & 0.2639 \\ 0.3073 & 0.4271 & 0.2656 \end{pmatrix}, \ldots$$

All the entries of M^2 are positives we can stop the proof.

On the other hand, the matrix $Q = \begin{pmatrix} 1 & 0 \\ 0.5 & 0.5 \end{pmatrix}$ is not regular.

In fact

$$Q^2 = \begin{pmatrix} 1 & 0 \\ 0.75 & 0.25 \end{pmatrix}, \quad Q^3 = \begin{pmatrix} 1 & 0 \\ 0.875 & 0.125 \end{pmatrix}, \; Q^4 = \begin{pmatrix} 1 & 0 \\ 0.9375 & 0.0625 \end{pmatrix}, \ldots$$

If we continue we will see that every time we have at least an entry which equal to zero for all power Q.

Theorem 3.2. Let P be a regular transition matrix, then

(i) There exists a unique **stationary vector** or **fixed probability vector** S.

(ii) Given any initial stable matrix S_0**,** the state matrix S_k approach the stationary matrix S. [8,11]

(iii)The matrix P^k approach a limiting $\overline{p}$, where each row of $\overline{p}$ is equal to the stationary matrix S.

Proof. Let matrix P be regular.

(i) Consider there is two stationary vectors S_1 and S_2 and we will prove that $S_1 = S_2$.

S_1 an stationary vector of P then

$$S_1P = S_1 \Rightarrow S_1P - S_1 = 0_M \Rightarrow S_1(P-I) = 0_M \ (1)$$

S_2 an stationary vector of P then

$$S_2P = S_2 \Rightarrow S_2P - S_2 = 0_M \Rightarrow S_2(P-I) = 0_M \ (2)$$

Where I is identity matrix and 0_M is a zero matrix.

From equations (1) and (2),

$$0_M = 0_M \Leftrightarrow S_1(P-I) = S_2(P-I) \Rightarrow S_1 = S_2.$$

Therefore, **t**he stationary vector is unique.

(ii) If S_0 is initial stable matrix, then recursively we have:

$$
\begin{aligned}
S_1 &= S_0 P \\
S_2 &= S_1 P \\
S_3 &= S_2 P \\
&\cdots \\
S_k &= S_{k-1} P
\end{aligned}
$$

When we make multiplication member by member, i.e. the left side and the right side we obtain

$$S_1 \times S_2 \times \cdots \times S_k = S_0 \times S_1 \times \cdots S_{k-1} P^k$$

$$\Rightarrow S_k = S_0 P^k \Rightarrow \lim_{k\to\infty} S_k = \lim_{k\to\infty} S_0 P^k = S_0 \lim_{k\to\infty} P^k$$

Take

$$\text{S} = S_0 \lim_{k\to\infty} P^k \ .$$

Remark It does not mean that every stochastic matrix have a **unique** stationary matrix except a regular stochastic matrix and the successive state matrices always approach this stationary matrix.

Example 7

Let

$$P = \begin{pmatrix} 0.1 & 0.1 & 0.8 \\ 0.4 & 0.4 & 0.2 \\ 0.1 & 0.2 & 0.7 \end{pmatrix},$$

be a regular transition matrix.

Then let’s find a stationary matrix S where $S = \begin{bmatrix} s_1 & s_2 & s_3 \end{bmatrix}$

Solution:

The matrix P is regular that means there exist a unique stationary matrix such that,

$$SP = S \Leftrightarrow \begin{bmatrix} s_1 & s_2 & s_3 \end{bmatrix} \begin{pmatrix} 0.1 & 0.1 & 0.8 \\ 0.4 & 0.4 & 0.2 \\ 0.1 & 0.2 & 0.7 \end{pmatrix} = \begin{bmatrix} s_1 & s_2 & s_3 \end{bmatrix}$$

$$\Rightarrow \begin{cases} 0.1s_1 + 0.4s_2 + 0.1s_3 = s_1 \\ 0.1s_1 + 0.4s_2 + 0.2s_3 = s_2 \\ 0.8s_1 + 0.2s_2 + 0.7s_3 = s_3 \end{cases} \text{ and we can add } s_1 + s_2 + s_3 = 1$$

By substitution we obtain

$$s_1 \approx 0.1688,\ s_2 \approx 0.2289 \text{ and } s_3 \approx 0.6024,$$

hence $S = [0.1688\ 0.2289\ 0.6024]$.

Chapter 4

MARKOV CHAINS

There are many stochastic processes in mathematics. In this chapter, we will study a special kind of stochastic process, called **Markov Chain**, where the next state of the system depends only on the present state. Before to start, just recall that Markov Chain where introduced in **1906** by the Russian mathematician **Andrei Andreyevich Markov** (**1856 – 1922**) and were named in his honor.

4.1 Some Definitions

Definition 4.1. Let $I = (i_1, i_2, ..., i_k)$ be a countable set and each $i_n \in I$ is a **state**, then I is called a **state-space.**

In this chapter, we will work in the **probability space (**Ω ,Ŧ, P**)** where Ω is a set of outcomes, Ŧ the set of subsets of Ω and for any $A \in$ Ŧ **,** $\mathrm{P}(A)$ is a probability of A .Our goal is to study a sequence $\left(X_n\right)_{n\geq 0}$ where $X_1, X_2, \ldots$ are taking from the set I .

Definition 4.2. The function $X: \ \Omega \to I$ is called a **random variable**, where the values of X belong the state-space I . [1,9]

Definition 4.3. Let $\lambda = (\lambda_i : i \in \mathrm{I})$ be a row vector. Then λ is called **measure** if for all $i \in I$, $\lambda_i \geq 0$.If $\sum_i \lambda_i = 1$ then λ is a **probability measure** or **probability vector** given in Chapter 3. In the special case, when $\lambda = (0,0,...,1,...,0)$, it is **longest probability vector** given in Chapter 3.We will denote $\lambda = \delta_i$. [5,7]

4.2 Markov Chain

Definition 4.4. Let $P = (P_{ij} : i, j \in I)$ be a transition matrix. Then the sequence $\left(X_n\right)_{n\geq 0}$ is called **Markov Chain** with transition matrix P and initial distribution λ, if for all $n \geq 0$

a) $i_0, i_1, \ldots, i_n, i_{n+1} \in I$ $P = P(X_0 = i_0) = \lambda_{i_0}$

b) $\mathrm{P}(X_{\mathrm{n}+1} = i_{n+1} \left| X_0 = i_0, X_1 = i_1, ..., X_n = i_n\right.) = P(X_{n+1} = i_{n+1} \left| X_n = i_n\right.) = P_{i_n i_{n+1}} \lambda$.

On the order hand, we may also say that a sequence $\left(X_n\right)_{n\geq 0}$ is Markov (λ, P).

Theorem 4.1. A sequence $\left(X_n\right)_{n\geq 0}$ is a Markov chain if for any $i_0, i_1, ..., i_n \in \mathrm{I}$,

$P(X_0 = i_0, X_1 = i_1, ..., X_n = i_n) = \lambda_{i_0} P_{i_0 i1} ... P_{i_n i_{n-1}}$.

Proof. Suppose $\left(X_n\right)_{n\geq 0}$ is Markov (λ, P). Then

$$\mathrm{P}(X_n = i_n, X_{n\text{-}1} = i_{n-1}, ..., X_0 = i_0)$$

$$= P(X_n = i_n \left| X_{n-1} = i_{n-1}, ..., X_0 = i_0\right.) P(X_{n-1} = i_{n-1}, ..., X_0 = i_0)$$

$$= P(X_0 = i_0) P(X_1 = i_1 \left| X_0 = i_0\right.) ... P(X_n = i_n \left| X_0 = i_0, ..., X_{n-1} = i_{n-1}\right.)$$

$$= \lambda_{i_0} P_{i_0 i_1} ... P_{i_{n-1} i_n}.$$

4.3 Homogeneous Markov Chain

There are several Markov Chains. In this section we will consider Markov Chains that do not evolve in time.

Definition 4.5. A Markov chain is called **homogeneous** if its one-step transition probability does not depend on n. In other words,

$$\forall n, m \in \square, \text{and } i, j \in \square, \quad P_{ij}^{(n)} = P_{ij}^{(m)}$$

Then we define the n steps transition probabilities of homogeneous Markov Chain by

$$P_{ij}^{(m)} = P(X_{n+m} = j \mid X_n = i),$$

which means that each row of P defines a conditional probability distribution on the state space. By convention

$$P_{ij}^{(0)} = \begin{cases} 1 & if \quad i = j \\ 0 & if \quad i \neq j \end{cases}.$$

Remark If $E = \{x_1, x_2, \ldots, x_n\}$ and $(X_n)_{n\geq 0}$ is homogeneous Markov Chain, then the transition matrix P_{ij} is given by:

$$P = \begin{pmatrix} p(X_{n+1} = x_1 \mid X_n = x_1) & p(X_{n+1} = x_2 \mid X_n = x_1) & \cdots & p(X_{n+1} = x_n \mid X_n = x_1 \\ p(X_{n+1} = x_1 \mid X_n = x_2) & p(X_{n+1} = x_2 \mid X_n = x_2) & \cdots & p(X_{n+1} = x_n \mid X_n = x_2 \\ \vdots & \vdots & \vdots & \vdots \\ p(X_{n+1} = x_1 \mid X_n = x_n) & p(X_{n+1} = x_2 \mid X_n = x_n) & \cdots & p(X_{n+1} = x_n \mid X_n = x_n) \end{pmatrix}$$

$$= \begin{pmatrix} p(x_1, x_1) & p(x_1, x_2) & \cdots & p(x_1, x_n) \\ p(x_2, x_1) & p(x_2, x_2) & \cdots & p(x_2, x_n) \\ \vdots & \vdots & \vdots & \vdots \\ p(x_n, x_1) & p(x_n, x_2) & \cdots & p(x_n, x_n) \end{pmatrix}.$$

Example 1 (Predicting the Weather (Finite State-Space))

In Cameroon, there are only 3 types of weather: sunny, foggy and rainy (a state-space takes three discrete values.) the weather patterns are very stable there, so a Cameroonians weatherman can predict the weather next week based on the weather today with the transition rules:

If it is sunny today, then

-probability it will be sunny next week is

$$* P(X_{(week)} = sunny \mid X_{(today)} = sunny) = 0.7$$

-probability somewhat it will be foggy next week is

$$* P(X_{(week)} = foggy | X_{(today)} = sunny) = 0.25$$

- it is very unlikely that it will be rainy next week

$$* P(X_{(week)} = rainy | X_{(today)} = sunny) = 0.05$$

If it is foggy today then

-likely that it will be sunny next week

$$* P(X_{(week)} = sunny | X_{(today)} = foggy) = 0.35$$

-less likely it will be foggy next week

$$* P(X_{(week)} = foggy | X_{(today)} = foggy) = 0.55$$

-fairly unlikely it will be raining next week is

$$* P(X_{(week)} = rainy | X_{(today)} = foggy) = 0.1$$

If it is rainy today then

-unlikely that it will sunny next week is

$$* P(X_{(week)} = sunny | X_{(today)} = rainy) = 0.1$$

-probability somewhat it will foggy next week is

$$* P(X_{(week)} = foggy | X_{(today)} = rainy) = 0.2$$

-fairly likely that it will rainy next week is

$$* P(X_{(week)} = rainy | X_{(today)} = rainy) = 0.7$$

If S=sunny, F=foggy and R= rainy, the we can model this example by the following transition matrix

$$P= \begin{array}{c} \\ S \\ F \\ R \end{array} \begin{array}{c} \begin{array}{ccc} S & F & R \end{array} \\ \begin{pmatrix} 0.7 & 0.25 & 0.05 \\ 0.35 & 0.55 & 0.1 \\ 0.1 & 0.2 & 0.7 \end{pmatrix} \end{array}$$

Note that each row of the matrix P above corresponds to the weather of today, and each column corresponds to the weather of the next week.

Question: Assume that it is sunny today what can be the probability it will rainy next week, in two next weeks or after 8 months?

We will answer these kinds of questions after we will study the next paragraphs.

4.4 Global Markov Property

Definition 4.6. Let A, B and C three sets where $A \cup B \cup C$ be a partition of V and B separates A from C as shown the graph above; i.e. starting in A and terminate in C. [11]

Then distribution μ over X^V satisfies the **global Markov property** if for any partition (A, B, C),

$$\mu(X_A, X_C \mid X_B) = \mu(X_A \mid X_B)\mu(X_C \mid X_B).$$

These previous definitions can introduce a new theorem.

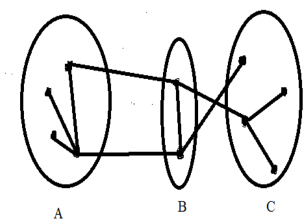


Figure 2: Global Markov Property

Theorem 4.1. (Chapman Kolmogorov Equations)

$$P_{ij}^{(m)} = \sum_{k\in\square} p_{ik}^{(r)} P_{kj}^{(m-r)}, r \in \square \cup \{0\}$$

Proof. To prove it we will use a total probability rule and global Markov property.

$$P_{ij} = P(X_m = j | X_0 = i) = \sum_{k\in\square} P(X_m = j, X_r = k | X_0 = i)$$

$$= \sum_{k\in\square} P(X_m = j | X_r = k\, X_r, X_0 = i) P(X_r = k | X_0 = i)$$

$$= \sum_{k\in\square} P_{ik}^{(r)} p_{kj}^{(m-r)} \text{ by Markov property.}$$

4.5 Asymptotic Behavior of Homogeneous Markov Chains

The study of the long-term behavior of Markov Chain seeks to respond to diverse questions as λ^n distribution does converge when $n \to \infty$?

If λ^n distribution converge when $n \to \infty$ what is a limit λ^*? And this limit it is independent to a initial distribution λ?

4.5.1 Stationary Chain

Definition 4.7. The Markov Chain whose evolution does not evolved over time is called Stationary Markov Chain. [3,5,9]

4.5.2 Distribution Invariant

Definition 4.8. λ is a probability **distribution invariant** to the transition matrix P if $\lambda P = \lambda$ in this case $(X_n)_{n\geq 1}$ be Markov (P, λ) is a stationary Markov Chain. We say λ is invariant if the terms **equilibrium** and stationary are also used to mean the same.

Theorem 4.2. Let I be a finite set. Then for some $i \in I$ such that

$$p^{(n)}_{ij} \to \pi_j \text{ as } n \to \infty \text{ for all } j \in I.$$

Then π=$(\pi_{\mathrm{j}} : j \in I)$ is an invariant distribution.

Proof. We have

$$\sum_{\mathrm{j}\in\mathrm{I}} \pi_{\mathrm{j}} = \sum_{j\in I} \lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty} \sum_{j\in I} p^{(n)}{}_{ij} = 1$$

and

$$\pi_j = \lim_{n\to\infty} p^{(n)}{}_{ij} = \lim_{n\to\infty} \sum_{k\in I} p^{(n)}{}_{ik}\ p_{kj} = \sum_{k\in I} \lim_{n\to\infty} p^{(n)}{}_{ik}\, p_{kj} = \sum_{k\in I} \pi_k\, p_{kj}\ .$$

We used here finiteness of I to justify interchange of summation and limit operations. Therefore, π an invariant distribution.

Example 2

Find the invariant distribution π according to regular transition matrix P where

$$\mathrm{P}=\begin{pmatrix} 0.1 & 0.1 & 0.8 \\ 0.4 & 0.4 & 0.2 \\ 0.1 & 0.2 & 0.7 \end{pmatrix}$$

Solution See Example 7 in Chapter 3.

4.6 Absorbing Markov Chains

Definition 4.9. An state x_j is called **Absorbing Markov Chain**, if

$$P(X_{n+1} = x_j \,|\, X_n = x_j) = 1.$$

Properties: A Markov Chain is **absorbing** if

-it has at least one absorbing state; and

-it is possible to go from any **non-absorbing** state to an absorbing state. [12,13,14]

Example 3

Between the two matrices below, identify all absorbing states in the Markov chain and decide whether the Markov chain is absorbing.

$$A = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array}\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.8 & 0.2 & 0 \\ 0 & 0 & 1 & 0 \\ 0.7 & 0 & 0.3 & 0 \end{pmatrix} \end{array}; \quad \text{and } B = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array}\begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0.6 & 0.2 & 0.2 \\ 0 & 0 & 1 \end{pmatrix} \end{array}$$

Solution

From matrix A, we have

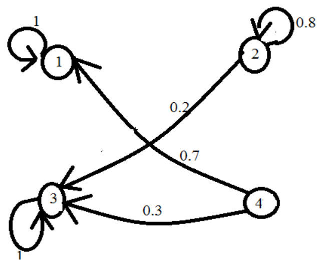


Figure 3: Transition Diagram Absorbing Markov Chain 1

States 1 and 3 are absorbing, with states 2 and 4 non-absorbing. From state 2 it is only possible to go state 3. From state 4 it is only possible to go state 3 and state 1 the transition diagram above shows it.

Conclusion: At least an non-absorbing state go to an absorbing state.

Hence the matrix A is a absorbing Markov chain.

From matrix B, we have:

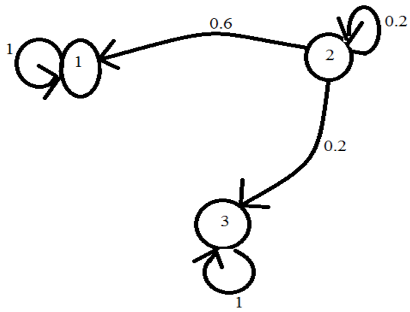


Figure 4: Transition Diagram Absorbing Markov Chain 2

$P_{11} = 1$ and $P_{33} = 1$both state 1 and state 3 are absorbing state. State 2 is only non-absorbing state. From state 2, it is possible to go to state 1 with a 0.6 probability and 0.2 probability from state 2 to state 3.

Conclusion: It possible to go from non-absorbing state to absorbing state as shown in figure then matrix B is also absorbing Markov Chain.

4.7 Irreducible Markov Chain

Definition 4.10. A Markov Chain is **irreducible** if every state is accessible from any other state with non-zero probability.

To detect an chain irreducible, we just have to check that $i \to j$ for every i, j.

Note. Any chain possessing an **absorbing** state is not irreducible.

4.8 Simulative Study of Homogeneous Markov Chain at Infinity

Given a Markov Chain $(X_n)_{n \geq 0}$ and a transition matrix P. We seek to study the behavior of the distribution of X_n when $n \to \infty$ which come to study a sequence of a matrix $(P^n)_{n \in \mathbb{N}}$ when $n \to \infty$with $P^0 = Id$. [16]

4.8.1 Markov Chain at Two-State $P^{(n)}$

Example 4

Consider the state of a phone line where $X_n = 0$ if the line is free at the time n. $X_n = 1$ if the line is busy at the time n. Also assume each time interval there is probability p when the call comes in (a call for more). If the line is already busy, the call is lost. Suppose again that if the call is busy at the time n there is probability q it is released at the time $n+1$.

What is the transition matrix of this stochastic process?

We can model this process by an homogeneous Markov Chain with values of E. Where E is the set of state $E = \{0,1\}$.

So the transition matrix

$$\mathrm{P} = \begin{pmatrix} p(0,0) & p(0,1) \\ p(1,0) & p(1,1) \end{pmatrix} = \begin{pmatrix} p(X_{n+1}=0|X_n=0) & p(X_{n+1}=1|X_n=0) \\ p(X_{n+1}=0|X_n=1) & p(X_{n+1}=1|X_n=1) \end{pmatrix} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

We then seek a simplified expression for P^n easily to calculate its limit.

We may diagonalize P because its spectral is $\{1, 1-p-q\}$

Then we can write

$$\mathrm{P} = QDQ^{-1}$$

where

$$Q = \begin{pmatrix} 1 & -p \\ 1 & q \end{pmatrix}$$

and

$$Q^{-1} = \begin{pmatrix} \frac{q}{p+q} & \frac{1}{p+q} \\ \frac{-1}{p+q} & \frac{1}{p+q} \end{pmatrix}$$

We should show that $\mathrm{P}^n = QD^nQ^{-1}$.

So

$$\mathrm{P}^n = \frac{1}{p+q}\begin{pmatrix} -q + p(1-p-q)^n & p - p(1-p-q)^n \\ q - q(1-p-q)^n & p + q(1-p-q)^n \end{pmatrix}$$

from where

$$\lim_{n\to\infty} \mathrm{P}^n = \frac{1}{p+q}\begin{pmatrix} q & p \\ q & p \end{pmatrix} = \mathrm{P}^\infty .$$

In general, to get $p_{11}{}^{(n)}$ for instance, we have $p_{11}{}^{(n)} = A + B(1-p-q)^n$ for some A and B. But

$$p^{(0)}{}_{11} = 1 = A + B \text{ and } p^{(1)}{}_{11} = 1 - p = A + B(1-p-q) .$$

Then

$$(A, B) = (q, p)/(p+q),$$

Thus

$$p_{11}{}^{(n)} = p_1(X_n = 1) = \frac{q}{p+q} + \frac{p}{p+q}(1-p-q)^n$$

and the linear recurrence relation is

$$p_{11}{}^{(n)} = q + (1-p-q)\, p_{11}{}^{(n-1)} .$$

Remark. As P^n converges to P^∞ it means that the homogeneous Markov chain approaches an equilibrium system or (stable), i.e. the distribution of this chain is stationary at a certain rank.

In other words,

$$\exists n_o \in \mathbb{N} \,/\, \mathrm{P}^\infty = \mathrm{P}^{n_0}, \mathrm{P}^{n_0+1} = \mathrm{P}^{n_0} \text{ and } \Pi_{n_0+1} = \Pi_{n_0}\mathrm{P}^{n_0+1} = \Pi_0 \mathrm{P}^{n_0} = \Pi_{n_0} .$$

4.9 Application of Markov Chain at Two-State

4.9.1 Gambler's Ruin Problem

This problem is motivated by trying to determine the success or failure of a gambler who bet with some amount money initially and wants to leave some larger money at the evening.

Let $Z_1, Z_2, \ldots$ be a sequence of Bernoulli random variable where $Z_i = +1$ with probability α and $Z_i = -1$ with probability β. We start with initial value $V_0 > 0$. We define the sequence of sum $V_n = \sum_{i=0}^{n} Z_i$ we are interested in the sequence $V_1, V_2, V_3 \ldots$ which is the stochastic process.

Now consider a Gambler who wins or loses **Turkish lira** (TL) on each turn of game with probabilities α and β respectively where $\beta = 1-\alpha$. Let V_0 the initial capital with $V_0 > 0$. This capital can increase to θ or be reduced to 0. A probability of the Gambler's ruin is βV_0 and αV_0 his probability he wins. We may also show that $\alpha V_0 + \beta V_0 = 1$.

Theorem 4.3. The probability of the Gambler's ruin is

$$\beta V_0 = \begin{cases} \dfrac{(\beta/\alpha)^{\theta} - (\beta/\alpha)^{V_0}}{(\beta/\alpha)^{\theta} - 1} & if \quad \alpha \neq \beta \\ 1 - V_0/\theta & if \quad \alpha = \beta = 1/2 \end{cases}$$

4.9.2 Birth and Death Chain

We define in this section some of notations we will use to compute the transition probabilities of Birth and Death chain.

Let b_i and d_i the probabilities of birth rate and death rate respectively with $b_i + d_i \leq 0$ and b_i , $d_i \geq 0$. We will work in the state space $S = \{0,1,2,3,\ldots\}$ or $S = \{0,1,2,3,\ldots,k\}$ with zero boundary condition usually we use $d_0 = 0$ and b_0>0 .Then the transition probability of birth and death chain is:

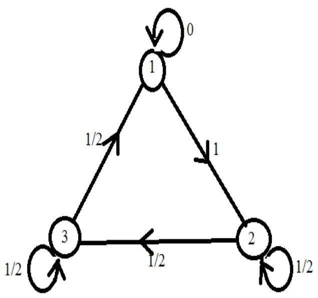
$$p_{ij} = \begin{cases} b_i & if \quad i = j-1 \\ d_i & if \quad i = j+1 \\ 1-b_i-d_i & if \quad i = j \\ 0 & else \end{cases}$$

Remark Gambler's ruin is an example of Birth and Death chain.

4.10 Markov Chain at *n*-steps Transition Probabilities

Example 5

Consider the following diagram related to a three-state Markov Chain.



$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

Figure 5: *n*-steps transition diagram

Using the characteristic polynomial to find the eigenvalues we obtain

$$f(\lambda) = \det(\lambda I - P) = \lambda(\lambda - 1/2)^2 - 1/4$$

$$= \frac{1}{4}(\lambda - 1)(4\lambda^2 + 1) = 0$$

So, the eigenvalues are obtained as 1, $\pm i/2$.

This means that

$$p_{11}^{(n)} = U + V(i/2)^n + W(-i/2)^n. \qquad (1)$$

We know that trigonometry and complex exponentials of

$$(\pm i/2)^n = (1/2)^n \left(\cos(n\pi/2) \pm \sin(n\pi/2)\right) = (1/2)^n e^{\pm in\pi/2}.$$

Substituting it in equation (1) for some V' and W' we have

$$p_{11}^{(n)} = U + (1/2)^n \left(V'\cos(n\pi/2) \pm W'\sin(n\pi/2)\right).$$

To find U, V' and W' we can assign n the values 0, 1 and 2 we get

$p_{11}^{(0)} = 1,\ p_{11}^{(1)} = 0,\ p_{11}^{(2)} = 0$ then we have U=$\frac{1}{5}$, V' =$\frac{4}{5}$ and W' =$\frac{2}{5}$.

Therefore

$$p_{11}^{(n)} = \frac{1}{5} + \left(\frac{1}{2}\right)^n \left[\frac{4}{5}\cos\frac{n\pi}{2} - \frac{2}{5}\sin\frac{n\pi}{2}\right]. \qquad (2)$$

Recall that this is for some of fix U, V' and W'. Using this equation (2), we can predict different values of probability $p_{11}^{(n)}$.

For instance, for $n = 5$ we have $p_{11}^{(5)} = 0.1875$.

In general, for chain with *n* states to get probability $p_{ij}^{(n)}$, from states *i* and *j* we use these different steps:

(a) Compute the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ of the square matrix $n \times n$ of P.

(b) If the eigenvalues are distinct, for some constants $c_1, c_2, ..., c_n$, $p_{ij}^{(n)}$ has the form $p_{ij}^{(n)} = c_1 + c_2\lambda_2^n + ... + c_n\lambda_n^n$.

(Recall that λ_1=1). If there is repeat eigenvalue λ *k*-times we will write the term in this form $\left(a_0 + a_1 n + ... + a_{k-1} n^{k-1}\right)\lambda^n$.

(c) Each complex eigenvalues come in conjugate pairs these can be written in trigonometry and complex exponentials form as the previous example.

Chapter 5

CONCLUSION

In the probability theory, the Markov Chain plays an important role. For instance being able to forecast an outcome based on the current state of a process has various applications in the life and sciences. The Gamblers may used the Markov change to forecast whether or not there can win or lose their bet.

The Markov Chain has also wide application areas; especially, in health sciences, biology and genetics sciences. It helps to predict the evolution of a disease and also to forecast in the future the outcome of a drug being used currently to fight against an epidemic. The Markov Chain is the key of epidemiology sciences.

The financial mathematics is also a field of study where the Markov Chain is widely used. It is actually used by insurances companies to simulate the risk before setting the insurance prime on a disaster. The current state of a financial market is used to predict the next state based on the Markov Chain. Brownian motion are defined as phenomenon with random outcome, but Markov Chain helps to predict outcome of processes.

REFERENCES

[1] E. B. Dynkin. (2006). *theory of Markov Processes.* new york: dover publication, inc.

[2] Blumenthal, R. M., & Getoor, R. K. (2007). *Markov Processes and Potential Theory.* New York: Dover Publication.

[3] Bowerman, B. L., O'Connell, R. T., & Murphree, E. S. (2014). *Business Statistics in Pratice, Seventh Edition.* New York: McGraw-Hill.

[4] Budnick, F. S. (1993). *Applied Mathematics for Business, Economics,and the Social Sciences, Fourth Edition.* Singapore: McGraw-Hill.

[5] Cinlar, E. (2013). *Introduction to Stochastic Processes.* New Jersey: Prentice-Hall.

[6] Friedberg, S. H., Insel, A. J., & Spence, L. E. (2003). *Linear Algebra, Fourth Edition.* Pearson.

[7] Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Techniques, Second Edition.* San Francisco: Morgan Kaufmann.

[8] Norris, J. R. (1998). *Markov Chains.* London: Cambridge University Press.

[9] Privault, N. (2013). *Understanding Markov Chains: Examples and Applications.* Springer.

[10] Stroock, D. W. (2005). *An Introduction to Markov Processes.* Cambridge: Springer.

[11] Walpole, R. E., Myers, R. H., & Myers, S. L. (2012). *Probability and Statistics for Engineers and Scientists, Ninth Edition* . Boston: Prentice Hall.

[12] Dynkin. (2006). *Theory of Markov Processes.*New York: Dover Publication, inc.

[13] Walpole Ronald E., Myers Raymond H., Myers Sharon L., *Probability and Statistics for Engineers and Scientists*, 6th Ed., Prentice Hall, New Jersey, 1998.

[14] Lıpschutz Seymour, *Theory and Problems of Probability*, Schaum's Outline Series, McGraw-Hill, 1974.

[15] Markov Chain. (2016, september 13). Retrived from http://setosa.io/ev/markov-chains

[16] Stochastic Matrix. (2016, september 13). Retrived from http://math.tutorvista.com/algebra/stochastic-matrix.html

[17] Finite Markov Chain.(2016, september 13). Retrived from http://quant-econ.net/py/finite_markov.html