# Markov Chains and Transition Probabilities

**Ko Ko Oo**


## Abstract

Markov chain is widely applicable to the study of many real-world phenomene. We represent the probability vectors, stochastic matrices, Markov chains, higher transition probabilities, stationary distribution, and absorbing states in this paper.


**Keywords**: Stochastic matrices, Markov chains, higher transition probabilities.

## Introduction

We come into contact with our first random process (stochastic process), known as a Markov chain, which is widely applicable to the study of many real-world phenomena. Applications to genetics and production processes are presented.

## Probability Vectors

A row vector $\underline{u} = (u_1, u_2, \dots, u_n)$ is called a **probability vector** if its components are non-negative and their sum is 1.

Consider the following vectors:

$$\underline{u} = (\tfrac{3}{4}, 0, -\tfrac{1}{4}, \tfrac{1}{2}),\ \ \underline{v} = (\tfrac{3}{4}, \tfrac{1}{2}, 0, \tfrac{1}{4}), \text{ and } \ \underline{w} = (\tfrac{1}{4}, \tfrac{1}{4}, 0, \tfrac{1}{2}).$$

Then: $\underline{u}$ is not a probability vector since its third component is negative;

$\underline{v}$ is not a probability vector since the sum of its components is greater than 1;

$\underline{w}$ is a probability vector.

Remark: Since the sum of the components of a probability vector is one, an arbitrary probability vector with n components can be represented in terms of n-1 unknowns as follows:

$$(x_1, x_2, \dots, x_{n-1}, 1-x_1-x_2 \dots -x_{n-1}).$$

In particular, arbitrary probability vectors with 2 and 3 components can be represented, respectively, in the form

$$(x, 1-x) \ \text{ and } \ (x, y, 1-x-y).$$

## Stochastic and Regular Stochastic Matrices

A square matrix $P = (p_{ij})$ is called a **stochastic matrix** (Hus, 1997) if each of its rows is a probability vector, *that is*, if each entry of $P$ is non-negative and the sum of the entries in every row is 1.

**Theorem 1** If $A$ and $B$ are stochastic matrices, then the product $AB$ is a stochastic matrix. Therefore, in particular, all powers $A^n$ are stochastic matrix.

Lecturer, Department of Mathematics, Hinthada University

**Definition:** **A** stochastic matrix $P$ is said to be *regular* if all the entries of some power $P^m$ are positive.

The stochastic matrix $A = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ is regular since

$$A^2 = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$ is positive in every entry.

Consider the stochastic matrix $B = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

Here $B^2 = \begin{pmatrix} 1 & 0 \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}$, $B^3 = \begin{pmatrix} 1 & 0 \\ \frac{7}{8} & \frac{1}{8} \end{pmatrix}$, $B^4 = \begin{pmatrix} 1 & 0 \\ \frac{15}{16} & \frac{1}{16} \end{pmatrix}$, … .

In fact every power $B^m$ will have 1 and 0 in the first row, hence $B$ is not regular.

## Fixed Points of Square Matrices

A non-zero row vector $\underline{u} = (u_1, u_2, \dots, u_n)$ is called a **fixed points of a square matrix** (Kandasamy et al., 2010). $A$ if $\underline{u}$ is left fixed, *that is*, is not changed, when multiplied by $A$ : $\underline{u}\,A = \underline{u}$.

We do not include the zero vector $\underline{0}$ as a fixed point of a matrix since it is always left fixed by every matrix $A$: $\underline{0}\,A = \underline{0}$.

**Theorem 2** If $\underline{u}$ is a fixed vector of a matrix $A$, then for any real number $\lambda \neq 0$ the scalar multiple $\lambda\underline{u}$ is also a fixed vector of $A$.

## Fixed Points and Regular Stochastic Matrices

The main relationship between regular stochastic matrices and fixed points is contained in the following theorem.

**Theorem 3** Let $P$ be a regular stochastic matrix. Then :

(i) $P$ has a unique fixed probability vector $\underline{t}$, and the components of $\underline{t}$ are all positive;

(ii) the sequence $P, P^2, P^3, \dots$ of powers of $P$ approaches the matrix $T$ whose row are each the fixed point $\underline{t}$;

(iii) if $\underline{p}$ is any probability vector, then the sequence of vectors $\underline{p}P, \underline{p}P^2, \underline{p}P^3, \dots$ approaches the fixed point $\underline{t}$;

**Note:** $P^n$ approaches $T$ means that each entry of $P^n$ approaches the corresponding entry of $T$, and $\underline{p}P^n$ approaches $\underline{t}$ means that each component of $\underline{p}P^n$ approaches the corresponding component of $\underline{t}$ .

Stochastic matrix $P$ is said to be *regular* if all the entries of some power $P^m$ are positive.

## Markov Chains

We now consider a sequence of trials whose outcomes, *say*, $X_1$, $X_2,\ldots$, satisfy the following two properties :

(i) Each outcome belongs to a finite set of outcomes ($a_1, a_2,\ldots, a_m$) called the **state space** of the system; if the outcomes on the $n^{th}$ trial is $a_i$, then we say that the system is in state $a_i$ at time n or at the $n^{th}$ step.

(ii) The outcome of any trial depends at most upon the outcome of the immediately preceding trial and not upon any other previous outcomes; with each pair of states ($a_i$, $a_j$) there is given the probability $p_{ij}$ that $a_j$ occurs immediately after $a_i$ occurs.

Such a stochastic process is called a (finite) *Markov chain* (Natarajan & Tamilarsi, 2007). The numbers $p_{ij}$, called the *transition probabilities*, can be arranged in a matrix

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ p_{m1} & p_{m1} & \cdots & p_{mm} \end{pmatrix}$$

called the **transition matrix**.

Thus with each state $a_i$ there corresponds the $i^{th}$ row ($p_{i1}, p_{i2}, \ldots , p_{im}$) of the transition matrix $P$; if the system is in state $a_i$, then this row vector represents the probabilities of all the possible outcomes of the next trial and so it is a probability vector.

**Theorem 4** The transition matrix $P$ of a Markov chain is a stochastic matrix.

## Higher Transition Probabilities

The entry $p_{ij}$ in the transition matrix $P$ of a Markov chain is the probability that the system changes from the state $a_i$ to the state $a_j$ in one step: $a_i \rightarrow a_j$.

The entry $p_{ij}^{(n)}$ in the transition matrix $P^{(n)}$ of a Markov chain is the probability that the system changes from the state $a_i$ to the state $a_j$ in exactly n steps : $a_i \rightarrow a_k \rightarrow a_l \rightarrow \ldots \rightarrow a_j$.

**Theorem 5** Let $P$ be the transition matrix of a Markov chain process. Then the n-step transition matrix is equal to the $n^{th}$ power of $P$ : $P^{(n)} = P^n$.

Now suppose that, at some arbitrary time, the probability that the system is in state $a_i$ is $p_i$; we denote these probabilities by the probability vector we denote these probabilities by the probability vector $\underline{p} = (p_1, p_2, \ldots , p_m)$ which is called the **probability distribution** of the system at that time. In particular, we shall let

$$\underline{p}^0 = (p_1^0, p_2^0, \ldots , p_m^0)$$

denote the *initial probability distribution*, *that is,* the distribution when the process begins, and we shall let

$$\underline{p}^n = (p_1^n, p_2^n, \ldots , p_m^n)$$

denote the $n^{th}$ step probability distribution, *that is,* the distribution after the first n steps.

**Theorem 6** Let $P$ be the transition matrix of a Markov chain process. If $\underline{p} = (p_i)$ is the probability distribution of the system at some arbitrary time, then $\underline{p}P$ is the probability distribution of the system one step later and $\underline{p}P^n$ is the probability distribution of the system n steps later. In particular,

$$\underline{p}^1 = \underline{p}^0 P,\ \ \underline{p}^2 = \underline{p}^1 P,\ \ \underline{p}^3 = \underline{p}^2 P,\ \ \ldots\ ,\ \ \underline{p}^n = \underline{p}^0 P^n.$$

## Stationary Distribution of Regular Markov Chains

Suppose that a Markov chain is regular, *that is*, that its transition matrix $P$ is regular. By *Theorem* 3, the sequence of n-step transition matrices $P^n$ approaches the matrix $T$ whose rows are each the unique fixed probability vector $\underline{t}$ of $P$; hence the probability $p_{ij}^n$ that $a_j$ occurs for sufficiently large n is independent of the original state $a_i$ and it approaches the component $t_j$ of $\underline{t}$.

**Theorem 7** Let the transition matrix $P$ of a Markov chain be regular. Then, in the long run, the probability that any state $a_j$ occurs is approximately equal to the component $t_j$ of the unique fixed probability vector $\underline{t}$ of $P$.

Thus we see that the effect of the initial state or the initial probability distribution of the process wears off as the number of steps of the process increase. Furthermore, every sequence of probability distribution approaches the fixed probability vector $\underline{t}$ of $P$, called the *stationary distribution* of the Markov chain.

## Absorbing States

A state $a_i$ of a Markov chain is called *absorbing* if the system remains in the state $a_i$ once it enters there. Thus a state $a_i$ is absorbing if and only if the $i^{th}$ row of the transition matrix $P$ has a 1 on the main diagonal and zeros everywhere else.

Suppose the following matrix is the transition matrix of a Markov chain:

$$P = \begin{array}{c} \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{array} \begin{array}{c} \begin{array}{ccccc} a_1 & a_2 & a_3 & a_4 & a_5 \end{array} \\ \begin{pmatrix} \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

The states $a_2$ and $a_5$ are each absorbing, since each of the second and fifth rows has a 1 on the main diagonal.

## Acknowledgement

I would like to express my sincere gratitude to Dr. Aung Win, Acting Rector, Hinthada University, and Dr. Theingi Shwe, Pro-Rector, Hithada University, for their kind permission to submit this paper. Special thanks are also due to Daw Aye Aye, Professor and Head of the Department of Mathematics, Hinthada University, for her encouragement and comments.

## References

Hsu, H. P., 1997. *Probability, Random Variables* & *Random Processes, (S.O.S),* Mc-Graw-Hill Company, Inc., New York, USA.

Kandasamy, P., Thilagavathi, K., and K. Gunavathi, 2010. *Probability*, *Random Variables*, & *Random Processes*, Reprint, S-Chand & Company Limited, New Delhi.

Natarajan, A. M., and A. Tamilarsi, 2007. *Probability Random Processes and Queuing Theory*, Reprint, New Age International Publishing Company, New Delhi.