International Research Journal of MMC (IRJMMC)
ISSN 2717-4999 (Online) | 2717-4980 (Print)
Volume 6, Issue 5 | December 2025

# The Application of the Markov Chain Model to Sales Forecasting

**Nand Kishor Kumar**[1]

[1]Lecturer, Tri-Chandra Campus
Tribhuvan University, Nepal

**Corresponding Author**
**Nand Kishor Kumar**
Email: nandkishorkumar2025@gmail.com

**To Cite this article:** Kumar, N. K. (2025). The application of the Markov Chain model to sales forecasting. *International Research Journal of MMC, 6*(5), 111–116. https://doi.org/10.3126/irjmmc.v6i5.89065

**Submitted:** 13 October 2025 **Accepted:** 15 December 2025 **Published**: 31 December 2025

## Abstract

The Markov Chain model, rooted in stochastic processes, provides a robust analytical framework by modeling transitions between states over time. Unlike regression-based models that focus on continuous data or trends, Markov Chains capture the likelihood that current sales levels will shift to different levels in the future. Sales forecasting plays a pivotal role in strategic planning, resource allocation, and decision-making within businesses. Traditional forecasting methods often rely on historical sales data and trend analysis; however, they may fall short in capturing dynamic transitions between states of customer behavior or product performance. The Markov Chain model offers a probabilistic and state-based approach that accounts for such transitions. This article explores the principles of Markov Chains, their relevance to sales forecasting, modeling approaches, advantages, limitations, and practical applications in modern business environments. For companies looking to obtain a competitive edge in sales forecasting, utilizing sophisticated analytical approaches like Markov Chain analysis is essential at a time where decisions are mostly based on data. Utilizing probabilistic modeling, companies may predict market trends, adjust to shifting consumer tastes, and maximize sales tactics for long-term success and profitability. Effective business planning and execution will surely continue to rely on the use of Markov Chain analysis into sales forecasting procedures as technology advances.

***Keywords**:* Markov chain, transition matrix, probability vector
**Mathematical Subject Classification:** 05B20,11C20,11J06,11J83

## 1. Introduction

The Markov chain idea was created by Russian mathematician Andrei A. Markov (1856–1922). In 1907, A. A. Markov started studying a crucial new type of random process. A stochastic model that depicts a series of likely events or changes in the state of a system is called a Markov chain. The current state of the system is the only factor that determines the likelihood of transitions between states. With just the present state of a stochastic process known, future states may be inferred using a Markov Chain (Kumar, 2024).

Accurate sales forecasting is essential for inventory management, cash-flow planning, marketing strategy, and organizational growth. In many real-world scenarios, sales behavior evolves in discrete states—for example, periods of high, moderate, or low demand. Moreover, customer purchasing behavior often depends on prior experiences or states, making it inherently probabilistic.

"Forgetting" past circumstances is referred to as the memoryless quality. To simulate a stochastic process using a Markov chain, a collection of states—that is, potential outcomes and the likelihood of moving from one state to another—is needed. The state space, or collection of all conceivable events, can include letters, numbers, weather, baseball scores, or stock performances. The process of changing from one state to another is called a transition. Transition probabilities are the likelihood of transitioning from one situation to another in a single step. We use a transition matrix to list the transition probabilities (Kumar,2024).

The transition matrix will be a n ×n matrix if the chain has n possible states. The probability of going from state I to state J is represented by the (i, j) item. To specify the transitions of a Markov Chain, we will use a directed graph in addition to the matrix. Vertices and the edges that link them make up a directed graph. The probability of getting from m to n in a single step is represented by the weight assigned to each graph edge (m, n). In a directed graph, a random walk is a set of vertices that starts at one vertex and moves over an edge to another vertex. The purpose of this research is to use historical data to forecast and evaluate sales using a Markov chain model.

## 2. Markov Chain

Markov chain is a stochastic model relating a sequence of possible events in which the probability of each event depends only on the state attained in the previous event (Klacksell, & Sundberg, 2013).

Suppose, we have a set of states, S = $(j, i_{n-1}, i_{n-2}, \cdots\cdots\cdots\cdots i_0),\ n \geq i$

Markov chains track Markov property which can be written as: For any n ≥ 1,

P$(X_n = j|\ X_0 = i_0 \cdots\cdots X_{n-1} = i_{n-1}\ )$ = P$(X_n = j| = i_{n-1})$

In other words, a Markov chain model is one in which the likelihood of an event is determined by what happened previously. While the theory of it is important precisely because so many "everyday" processes satisfy their property. However, there are many common examples of stochastic properties that do not gratify its property (Klacksell, & Sundberg, 2013).

### 2.1 Representation of Markov chain

Markov chain is a directed graph as $+- + - +$ in two state process. So, we can signify it with an adjacency matrix. |A| E| has each element represent probability weight of the edge. If the Markov chain has N possible state, the matrix will be an N× $N$ matrix. Here each row of this matrix should sum to be unity. It has also an initial state vector of order N× $1$, these are two entities Markov chain (Kumar,2024).

### 2.2 Types of Markov chain

(a) Discrete- time Markov chains:

The countable index set T undergoes modifications at certain beginnings. DTMC is the shorthand for it. There are only discrete values for the index variable in the discrete temporal distribution. It is a finite process as well, with a predetermined step interval that may be measured in ticks. Only when there is a change in the index is the index updated when ticks are used. Then, the time between each stage of the procedure might be changed. The probability of something occurring is often computed in discrete time (Satorra, 2016).

(b) Continuous- time Markov chains:

The index collection in CTMC, this continuum undergoes constant modifications. Understanding the continuous time distribution is more challenging. Due to the infinitesimal step between each measurement, the index in the continuous distribution has an unlimited number of values, even over a finite interval. Consequently, the change is similarly negligible. This technique will always produce a constant if we believe that a tiny change is equivalent to no change. Naturally, this isn't the case for all continuous distributions. Every state will emerge because a continuous process can occur in an endless number of ways. Therefore, seeing how often it occurs rather than how likely it is to occur is more informative (Satorra, 2016).

### 2.3 Properties

(a) It is irreducible- it processes from one state to another in a single or more than one step.
(b) Periodic if it needs, a minute of same integer large than one,
(c) Transient- it there is a non- zero probability that the chain will never return to the same state, otherwise, it is recurrent
(d) Absorbing -if there is no possible way to leave that state. It has no any outgoing transitions from it.

### 2.4 Application

Markov chains is applied to study to many real-world processes in such more same useful results as stationary distribution and many more.

(a) MCMC (Markov Chain Monte Carlo). This gives a solution to the problems that come from the normalization factor, which is blessed on Markov chin.
(b) They are used in information theory search engines, speech recognition etc.
(c) They have huge probabilities, in field of data science.

## 3. Discussion

### 3.1 Markov chain transition matrix

A Markov chain transition matrix is a stochastic matrix whose (i, j) entry specifies the probability that an element will move from state $s_i$ to state $s_j$ during the next phase of the process. The probability is denoted by $P_{ij}$ and it is independent of which states the chain was in prior to the current state. Transition probabilities are the probability $P_{ij}$.

$(P_t)_{i,j} = \mathbb{P}\ (X_{t+1} = j | X_t = i)$.

$$P= \begin{bmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,j} & \cdots \\ \vdots & & \ddots & \vdots & \\ P_{i,1} & P_{i,2} & \cdots & P_{i,j} & \cdots \end{bmatrix}$$

This shows each row of the matrices is a probability vector and the sum of the its entries is 1. The probability for going from state I to j is now calculated to be

$P_r(j|i) = P_{i,j}$

One important comment is that the row probability is summing up to 1, since the sum of the probabilities of going from states i to any state is 1 (100 %). The probability of going from any state to another in k- steps for a Markov chain can be calculated by using the matrix P. The probability is then given by $P^k$ (Satorra, 2016).

### 3.2 N- step Transition matrix

These higher order transition matrices determine the likelihood that the transition will take place over a number of steps. The transition matrix is raised to the power of N in order to compute the N-step matrix.

### 3.3 Properties of Transition Matrix

The product of subsequent ones describes a transition along the time interval spanned by the transition matrices (Encyclopedia, 2012).

$P_0.P_1$ has it's $(i,j)^{th}$ position, the probability $X_2$ = j given that $X_0$ = 1. In general, $(i,j)^{th}$ position of the $P_t.P_{t+1}$ is the probability $\mathbb{P}(X_{t+k+1} = j|X_t = i)$.

**Theorem1.** Prove that for any natural numbers t and states i ,j∈ $S$, the matrix entry $(P_t.P_{t+1})_{i,j} = \mathbb{P}(X_{t+2} = j = i)$.

Proof. Let M = $P_t.P_{t+1}$ , for matrix multiplication,

$M_{i,j} = \sum_{k=1}^{n}(P_t)_{i,k}(P_{t+1})_{k,j} = \sum_{k=1}^{n}\mathbb{P}(X_{t+1} = k|X_t = i)\mathbb{P}(X_{t+2} = j|X_{t+1} = k)$

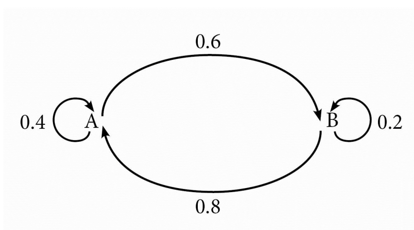
$= \mathbb{P}(X_{t+2} = j|X_t = i)$ is the expression of conditional probability.

The $\mathcal{K}$-step transition matrix is

$P_t^k = P_t.P_{t+1} \cdots\cdots\cdots P_{t+k+1}$ and by the above expression satisfies,

$$P_t^k = \begin{pmatrix} \mathbb{P}(X_{t+k} = 1|X_t = 1) & \mathbb{P}(X_{t+k} = 2|X_t = 1) & \mathbb{P}(X_{t+k} = n|X_t = 1) \\ \mathbb{P}(X_{t+k} = 1|X_t = 2) & \mathbb{P}(X_{t+k} = 2|X_t = 2) & \mathbb{P}(X_{t+k} = n|X_t = 2) \\ \vdots & \vdots \quad\ddots & \vdots \\ \mathbb{P}(X_{t+k} = 1|X_t = n) & \mathbb{P}(X_{t+k} = 2|X_t = n) & \mathbb{P}(X_{t+k} = n|X_t = n) \end{pmatrix}$$

**Example.** For the time-independent Markov chain described by the picture below. What is it's 2-step transition matrix?




The matrix is P = $\begin{pmatrix} .4 & .6 \\ .8 & .2 \end{pmatrix}$

Now from two- steps transition matrix is

$P^2 = \begin{pmatrix} .4 & .6 \\ .8 & .2 \end{pmatrix} * \begin{pmatrix} .4 & .6 \\ .8 & .2 \end{pmatrix}$ where $*$ is multiplication.

$\begin{pmatrix} .4 \times .4 + .8 \times .6 & .4 \times .6 + .4 \times .2 \\ .8 \times .4 + .2 \times .8 & .8 \times .6 + .2 \times .2 \end{pmatrix} =$

$\begin{pmatrix} .64 & .32 \\ .48 & .52 \end{pmatrix}$ process changing from one state to another state.

An illustration of a Markov process with two states. In this case, the arrows point to a future state from the present state, and the values that correspond to the arrows show the likelihood that the Markov process E will change states (Medhi,2009 & Zhang, 2009).

### 3.4 Probability Vector

A probability vector with non-negative entries is a row vector element that are non-negative and sum to be 1. Let clarify with some example. Suppose Tata car control 50% of the car market. They decided to hire a research team to see the effects after an ad campaign (Mahfuz,2021). Here each time step is a week. The team finds that:

(a) Someone using Tata car will stay with 85% probability.

(b)Someone not using Tata car will switch to Tata with 60% probability.
Let T be the state "use Tata "and T' be the state "Does not use Tata". The initial distribution matrix, A indicates the initial probability for each state. In this example $A_0$ is given by

$$A_0 = \begin{matrix} T & T' \\ [.50 & .60] \end{matrix}$$

From the transition matrix, P is the matrix of transition probabilities. In this example, probability of
T→ $T$ is 0.8 5. We have T→ $T'$ is 1-0.85= 0.15.
Similarly, T'→ $T$ is 0.60. So, T'→ $T$ is 1-0.60 = 0.40. The matrix formed by this information's as

$$P = \begin{matrix} & T & T' & \\ & \begin{bmatrix} .85 & .15 \\ .60 & .50 \end{bmatrix} & & \begin{matrix} T \\ T' \end{matrix} \end{matrix} \qquad \text{(i)}$$

**Theorem2.** Let P be the transition matrix of a Markov chain. The (i,j) entry $P_{i,j}^n$ of the matrix $P^n$ gives the probability where Markov chain starting in state $S_i$ will be in state $S_j$ after n steps. With the help of theorem 2, we can answer the following tropical questions.

(a) What is the market share after 1 week?
(b) After two weeks?
(c) After n weeks?
(d) How does the market share change in the long run?

Continuing, after two weeks can be performed from

$$P^2 = \begin{bmatrix} .85 & .15 \\ .60 & .50 \end{bmatrix} \times \begin{bmatrix} .85 & .15 \\ .60 & .50 \end{bmatrix} = \begin{bmatrix} .85 \times .85 + .15 \times .60 & .85 \times .15 + .15 \times .50 \\ .60 \times .85 + .50 \times .60 & .60 \times .15 + .50 \times .50 \end{bmatrix}$$
$$= \begin{bmatrix} .7225 + .09 & .1275 + .o6 \\ .51 + .24 & .09 + .16 \end{bmatrix} = \begin{bmatrix} .8125 & .1875 \\ .75 & .25 \end{bmatrix}$$

The probability that someone using Tata car will be stay with Tata in two weeks is calculated to be 81.25% and after 10 weeks

$P^{10} \approx \begin{bmatrix} .80 & .20 \\ .80 & .20 \end{bmatrix}$ and also, for $P^{100} \approx \begin{bmatrix} .80 & .20 \\ .80 & .20 \end{bmatrix}$

In last $P^n \approx \begin{bmatrix} .80 & .20 \\ .80 & .20 \end{bmatrix}$ for larger n.
The theorem 3 will help to answer the question (a).

**Theorem 3.** Let P be the transition matrix of a Markov chain, and let $A_0$ be the probability vector which represents the starting distribution. Then the probability that the chain is in state $S_i$ after n steps is the $i^{th}$ entry in the vector $A_n = A_{n-1}P$.
With the help of theorem 3, the market share after one week can be predict as

$A_1 = A_0 \times P$ = [.4 .6] $\begin{bmatrix} .85 & .15 \\ .60 & .40 \end{bmatrix}$ = [.7 .4]

After one-week Tata car to have 70% of the market share. Similarly, the market share after two weeks

$A_2 = A_0 \times P^2$ = [,7 .3] $\begin{bmatrix} .85 & .15 \\ .60 & .40 \end{bmatrix}$ = [.775 .225]

After two weeks, Tata have 77.5% to market share. After 10 weeks,

$A_{10} = A_0 \times P^{10}$ = [.4 .6] $\begin{bmatrix} .85 & .15 \\ .60 & .40 \end{bmatrix}$ = [.799999 .2000001]

After 10 weeks, Tata have almost 80% of market share. Similarly continuing this process to stabilize the market with time.

## 4. Conclusion

Markov Chain is a stochastic model that depicts a series of occurrences in which the likelihood of changing states only depends on the present state and not on the events that came before it. Put more simply, it's a process that alternates between states, with the current state and a few probabilistic rules being the only factors that decide the future state.

This Chain model provides a powerful, structured, and probabilistic approach for sales forecasting. By modeling transitions between discrete states of sales performance or customer behavior, it offers insights that traditional forecasting techniques may overlook. While it has limitations, especially in rapidly changing environments, its simplicity, adaptability, and ability to capture state-based dynamics make it a valuable tool for businesses seeking data-driven forecasting solutions.

As organizations increasingly rely on predictive analytics, integrating Markov Chain models with other forecasting techniques—such as time-series models, machine learning, or regression analysis—can yield robust and comprehensive forecasting systems.

## Ethical Approval for the Research

I declare that this paper work has been conducted ethically.

## Conflict of Interest

There is no conflict of interest with any institutions and individuals.

## References

1. Brilliant. https:|brilliant.org$> wiki > markov - chain.$
2. Encyclopedia (2012). "probability theory." Encyclopedia Britannica. Encyclopedia Britannica Online Academic Edition. Encyclopædia Britannica Inc., 2012. Web. 10 April. https://en.wikipedia.org/wiki/PageRank
3. Klacksell, G. & Sundberg, J. (2013). Markov Chain. Thesis submitted to Royal Institute of Technology.
4. Kumar, N. K. (2024). Effects of unemployment and inflation on the economy of Nepal. *Rajarshi Janak University Research Journal*, *2*(1-2), 42–49. https://doi.org/10.3126/rjurj.v2i1-2.72303
5. Kumar, N. K. (2024). Index of Social Vulnerability to Natural Disasters in Nepal. *BMC Journal of Scientific Research*, *7*(1), 135-146. https:/doi.org/10.3126/bmcjsr.v7i1.72950
6. Kumar, N.K. (2024). An In-Depth Analysis of Macroeconomic Factors Influencing Nepal's Economic Growth. NPRC Journal of Multidisciplinary Research, 1(6), 13-22. https://doi.org/10.3126/nprcjmr.v1i6.71739
7. Kumar, N.K. (2024). Application of Mathematics in Economics. NCCS Research Journal, 3(1). DOI: https://doi.org/10.3126/nccsrj.v3i1.72344
8. Mahfuz, F. (2021). Markov Chains and their Applications. Paper 10, thesis submitted to the university of Texas at Tyler.
9. Medhi, J. (2009). Stochastic Processes. 3rd Edition, New Age International Publishers, New Delhi, India.
10. Satorra, A.A (2016). Markov Chains. A thesis submitted to the University of Barcelona.
11. Zhang D, & Zhang X. (2009). Study on Forecasting the Stock Market Trend Based on Stochastic Analysis Method: International Journal of Business and Management. 4(6).