Trend Analysis and Prediction of Indian Stock Market using Markov Chain Model

Dr. RAJKUMARI SONI  
Assistant Professor, Department of Accounting and Financial Management  
Faculty of Commerce, The Maharaja Sayajirao University of Baroda,  
Vadodara, Gujarat

SHRI. BHUMIT SHAH  
Assistant Professor, Department of Banking and Insurance  
Faculty of Commerce  
The Maharaja Sayajirao University of Baroda  
Vadodara, Gujarat

Abstract

The purpose of the study is to analyze and predict the National Stock Exchange (NSE) Index. For Investors, who are seekers of Capital Appreciation, trend analysis and prediction are very essential. Therefore, investors, whether retail or institutional, must predict the trend of the stock market where the investments have been made to create wealth. The study uses the Markov Chain Model over a period from 3rd February 2020 to 29th January forming, 247 days of trading data from the official website of the NSE. The NSE Index shows three different states - Increase, Unchanged, and Decrease. The Markov Chain model is a probability model based on the transition probability matrix and initial state vector. The transition probability matrix has been obtained by observing the number of transitions from one state to another. The study explores that in the long run, regardless of the current state of the NSE index, the probability of index increase, unchanged, and decrease are 42.29%, 32.87%, and 24.84%, respectively. If the closing value of the NSE index is in the state Increase on day one, then it can be expected to return to the state increase for the first time in two days or on the third day.

Keywords: markov chain; trend analysis; prediction; transition probability matrix; indian stock market.

Introduction

The Stock Market is a barometer of any economy. The Stock market plays a significant role in the life of Investors with the sole objective of making a profit by investing in different shares of listed companies. The Stock market is the legal platform where such shares are traded. And according to the performance of the company, the price of the shares grows up or down in the stock market. Apart from the performance of the company, worldwide trend of business, natural calamities, the sociopolitical policy, global market conditions also affect the stock market. Therefore, there lies the significance of stock market prediction, which may benefit an investor. Various scientific methods are implemented for stock market prediction through the analysis of past information. Investors in the stock market are interested to know the future occurrence of the market as they are motivated by the desire for capital appreciation. However, the fluctuating nature of the stock market makes the analysis and prediction a complicated phenomenon and making accurate prediction of the stock market by any single method a very difficult task. Different methods such as Machine Learning, Neural Network, Moving Average, Regression Analysis, ARIMA, Data Mining, Markov Chain analysis, Hidden Markov model are used by different researchers to forecast the stock market.

Review of Related Work

Prediction is the method where we find the future value on a specific field with the help of the past data records and conclude the specific result. There are several fields where prediction is used, and various prediction methods have been proposed and implemented. These methods range from Machine Learning (Jigar Patel, Sahil Shah, Priyank Thakkar, and K Kotecha; David Enke and Suraphan Thawornwong), Neural Network (Manna Majumder and MD Anwar Hussien; Akintola K.G., Alese B.K. and Thompson A.F.; Sneha Soni; Tiffany Hui-Kuang and Kun-Huang Huang; R.K. and Pawar D.D.; Halbert white; Jing Tao Yao and chew Lim tan; Riki Herliansyah and Jamilatuzzahro),

Bodhi International Journal of Research in Humanities, Arts and Science

The Markov Chain Model has been used by many researchers to predict and analyze the trend of the stock market at different times. Styan, George P. H. and Smith, Harry Jr. (1962) used market behavioral analysis data provided by the transitional or switching, habits of the consumer. They had taken types of laundry powder purchased by a housewife to define the state space of a Markov chain and predicted the future purchase behavior and statistical inferences on the switching habits. Zhang, Deju and Zhang, Xiaomin (2009) applied the Markov process model for the Chinese stock market trend forecasting and concluded that the operational status of the stock market is subject to the impact of various factors from the market; hence, no single method can accurately predict changes in the stock market every day. Doubleday, Kevin J., and Esunge, Julius N. (2011) determined the relationship between a diverse portfolio of stocks and the market. The DJIA and portfolio of five stocks were analyzed using a Markov Chain. They concluded that when treated as a Markov process, the entire market was useful in gauging how a diverse portfolio of stocks might behave. Vasanthi Sei et al. (2011) tried to predict the stock index trend of various global stock indices using Markov Chain Analysis and found that majority of the time, the Markov model outperformed the traditional trend prediction methods. Badge, Jyoti (2012) used selected technical indicators of macro-economic factors of the Indian stock market as an input variable. Future prices have been found through Hidden Markov Model (HMM). It has been observed that HMM with PCA performed well and provided a Mean Absolute Percentage Error (MAPE) of 1.77%. Choji, Davu Nyapet et al. (2013) used the Markov chain model to analyze and to make predictions of the two top banks in the three states. They found that regardless of a bank’s current share price, in the long run, its share price will depreciate with a probability of 0.4229, remain unchanged with a probability of 0.2072, and appreciate with a probability of 0.3699. Otieno, Simeyo et al. (2015) used secondary quantitative data on the daily closing share prices of Safaricom from the NSE website. It has been observed that irrespective of the current state of share price, the model predicted that the Safaricom share prices would depreciate, maintain value, or appreciate with a probability of 0.3, 0.1, and 0.5, respectively. Bairagi, Aparna and Kakaty, Sarat C. (2016) used the Markov Chain model to analyze and predict the stock behavior considering three different states, ‘up’, ‘down’, and ‘remain same’. The study revealed that regardless of the bank’s current share price, steady-state probabilities of share ‘up’, ‘down’ and ‘remain same’ for SBI were 46.99%, 49.81%, and 3.19%, respectively. If the closing value of SBI share was on the state ‘up’ on day one, then it can be expected to return to the state ‘up’ for the first time on the third day. Singh, Waikhom Rojen et al. (2017) presented a paper with objectives to predict the prompt future change in opening stock price and to find a steady-state transition probability matrix. The opening stock price of National Stock Exchange NIFTY 50 of India has been studied and the opening price of Stock was examined as following Markov chain. Bhusal, Madhav Kumar (2017) attempted to apply a Markov Chain model to forecast the behaviors of the Nepal Stock Exchange (NEPSE) index. The study explored that regardless of the present status of the NEPSE index, in the long run, the index will increase with a probability of 0.3855, remains in the same state with probability 0.1707, and decrease with a probability of 0.4436.

The study is an attempt to analyze the trend and prediction of the National Stock Exchange (NSE) Index, India, using the Markov Chain Model in line with the above-mentioned work of the authors.

**Objectives of the Study**

1. To study the long-run behavior of the NSE Index
2. To find out the expected number of visits to a particular state and
3. To find out the expected first reaching time of various states.

Markov Chains and their respective diagrams can be used for predicting the likelihood of future market conditions. These conditions, also known as trends, are:

Increase (Bull Markets): periods where prices generally are rising due to the investors having optimistic hopes of the future, i.e., the closing value is greater than the closing value of the previous day.

Unchanged (Stagnant Markets): periods, where the market is characterized by neither a decline nor rise in general price, i.e., closing value, is equal to the closing value of the previous day.

Decrease (Bear Markets) periods where prices generally are declining due to the investors having pessimistic views of the future, i.e., Closing value is less than the closing value of the previous day.

Data and Methodology

For the study, secondly quantitative data on the daily closing prices of the NSE Index was obtained over a period from 3rd February 2020 to 29th January, forming 247 days of trading data from the official website of the NSEwww.nseindia.com. Here, the number of observations for Index value increase, unchanged, and decrease has been found as 105, 81, and 61, respectively. Since the last trading day is recorded as a decrease and there is no information regarding the next day’s transition, the total number of decreases should be recorded as (61-1) = 60.

The basic theory of the Markov chain model

In 1907, A. A. Markov began the study of an important new type of chance process. In this process, the outcome of a given experiment can affect the outcome of the next experiment. This type of process is called a Markov chain.

Markov process is a special type of stochastic process for which the future occurrence of any event depends only on the present state. The set of values that the Markov process takes is known as its state-space. Depending on the values of state-space, the process may be discrete or continuous. The process with discrete state space is known as the Markov chain. Markov chain is one of the most well-developed theories of a stochastic process with wider applications.

Definition of Markov chain

The sequence \( \{X_n, n \geq 0\} \) is said to be a Markov chain if for all state values \( i_0, i_1, i_2, \ldots, i_n \in \mathcal{I} \), then

\[
P\{X_{n+1} = j / X_0 = i_0, X_1 = i_1, \ldots, X_n = \cdots = i \} = p_{ij}
\]

Where, \( i_0, i_1, i_2, \ldots, i_n \) are the states in the state space \( \mathcal{I} \).

This kind of probability is called Markov chain probability. This implies that regardless of its history before time \( n \), the probability that it will make a transition to another state \( j \) depends only on state \( i \).

Transition probability and transition probability matrix

The transition probability, as defined by the Markov chain, is called transition, or jump probability from state \( i \) to state \( j \). Then,

\[
P \{X_{n+1} = j / X_n = i\} = p_{ij}
\]

This is also termed as one-step transition probability. If the transition probabilities defined above are independent of time \( n \), then such an assumption is called a homogenous or stationary Markov chain. Thus,

\[
P \{X_{n+1} = j / X_n = i\} = P(X_1 = j / X_0 = i) = p_{ij}
\]

The transition probabilities \( p_{ij} \) can be written or arranged in a matrix form as,

\[
P = [p_{ij}], i, j \in \mathcal{I}
\]

Here, the matrix \( P \) is called the transition probability matrix (TPM) or stochastic matrix. The matrix \( P \) consists of non-negative elements with row sum unity. Hence

\[
0 \leq p_{ij} \leq 1 \text{ and } \sum_{j=1}^{\mathcal{I}} P_{ij} = 1, \forall i \in \mathcal{I}
\]

The k-step transition probability from state \( i \) to state \( j \) in k steps is,

\[
P(k) = P (X_{n+k} = j / X_n = i), \sqrt{k > 0}, n \geq 0, i, j \in \mathcal{I}
\]

The Transition matrix \( P \) has the following property.

\[
P(n) = P^{n-1} \ast P = P^n
\]

State Probability Matrix

The average transition process of the Markov chain depends on the system’s initial state and the transition probability matrix. The initial state is a line matrix called initial probability vector defined as;
Thus, to state j becomes sufficiently large, then the transition probability from state i evolves, when the number of transition steps is the initial state of the system, how does the stochastic process evolve, when the number of transition steps is sufficiently large, then the transition probability from state i to state j becomes settle down to some constant value. Thus, 

\[ \lim_{n \to \infty} p_{ij}(n) = \pi_j \]

Such quantities are referred to as steady-state probabilities.

If the limits \( \pi_i = \lim_{n \to \infty} p_{ij}(n) = \lim_{n \to \infty} p_{ij}(n) \) exists and does not depend on the initial state, then \( p_{ij}(n) = \sum_k P_k(n-1) p_{ik} \) becomes \( \pi_j = \sum_k \pi_i p_{ik} \), as \( n \to \infty \) for \( j = 0,1,2,........ \)

This is equivalent to \( \pi = \pi \star P \)

The probability distribution \( \{\pi_j, i \in I\} \) is called stationary or invariant for the given chain if

\[ \pi_i = \sum_{j \in I} \pi_i P_{ij} \] such that \( \pi_i \geq 0 \) and \( \sum_i \pi_i = 1 \)

This property of the Markov Chain helps to determine the long-run behavior of the chain.

**Expected number of Visits**

The expected number of visits made by the chain to state \( j \) starting from state \( i \) is given by

\[ \mu_{ij}(n) = E(N_i(n)) \]

Where \( N_i(n) \) denotes the number of visits to state \( j \) starting from state \( i \) in \( n \)-steps.

Were,

\[ N_i(n) = \sum_{k=1}^{n} Y_{ij}(k) \] with \( Y_{ij}(0) = \delta_{ij} \), the Kronecker delta.

And \( Y_{ij}(k) = \begin{cases} 1, & \text{if } X_k = j/X_0 = i \\ 0, & \text{otherwise} \end{cases} \)

Then, \( \mu_{ij}(n) = E[\sum_{k=1}^{n} Y_{ij}(k)] \)

\[ = \sum_{k=1}^{n} E(Y_{ij}(k)) \]

\[ = \sum_{k=1}^{n} P[Y_{ij}(k) = 1] \]

\[ \therefore \mu_{ij}(n) = \sum_{k=1}^{n} P_{ij}(k) \]

Also, the expected number of visits to state \( j \) from state \( i \) after long-run is:

\[ \mu_{ij}(n) = \lim_{n \to \infty} E(N_{ij}(n)) \]

**Expected Return Time**

For a finite irreducible Markov chain, the expected return time to state \( j, i \in I \) can be obtained by taking the reciprocal of limiting probability \( p_{ij}(n) \).

**Determination of Initial State Vector**

The closing price shows three different states Increase, Unchanged, and Decrease, as each closing price index is taken as a discrete-time unit. The Initial state vector gives the probabilities of the three different states.
The state vector is denoted by \( \pi^{(0)} = (\pi_1, \pi_2, \pi_3) \) then \( \pi_1, \pi_2 \) and \( \pi_3 \) gives the probability of the NSE index increase, unchanged and decrease as

\[
\pi_1 = 105/246 = 0.4268 \\
\pi_2 = 81/246 = 0.3293 \\
\pi_3 = 60/246 = 0.2439
\]

So, the initial state vector for the NSE index is - \( \pi^{(0)} = [0.4268 \ 0.3293 \ 0.2439] \)

### Determination of Three state transition probability metrics

The states are the chances that the NSE index increases, that is unchanged, and that it decreases. That transitions from one state to another observed from the data panel and were compiled in Table 1 as follows:

<table>
<thead>
<tr>
<th></th>
<th>Increase</th>
<th>Unchanged</th>
<th>Decrease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase</td>
<td>50</td>
<td>38</td>
<td>105</td>
</tr>
<tr>
<td>Unchanged</td>
<td>25</td>
<td>29</td>
<td>81</td>
</tr>
<tr>
<td>Decrease</td>
<td>29</td>
<td>14</td>
<td>51</td>
</tr>
</tbody>
</table>

The transition probability matrix of the NSE index using the above information can be constructed as;

\[
P^{(1)}_{NSE} = \begin{bmatrix}
0.4762 & 0.3619 & 0.1619 \\
0.3086 & 0.3580 & 0.3333 \\
0.3922 & 0.2745 & 0.3333
\end{bmatrix}
\]

The transition digraph for the explicit presentation of transition probability of the NSE index is shown below -

![Transition Probability Digraph of the NSE Index](image)

Calculating state probability for forecasting the next day NSE index price

By applying initial state vector and transition probability metrics, it is possible to find out the state probability of various closing days in the future. The state probability of closing price of the NSE index for the 248\(^{th}\) day will be:

\[
\pi^{(1)} = \pi^{(0)} \times P_{NSE} = \begin{bmatrix}
0.4268 & 0.3293 & 0.2439
\end{bmatrix} \times \begin{bmatrix}
0.4762 & 0.3619 & 0.1619 \\
0.3086 & 0.3580 & 0.3333 \\
0.3922 & 0.2745 & 0.3333
\end{bmatrix}
\]

\[
248^{th} \text{ Day} = \pi^{(1)} = [0.4228 \ 0.3293 \ 0.2480]
\]

This indicates that on the 248\(^{th}\) day the state increases have maximum probability, and it can be predicted that the NSE index will increase on this day with a probability of 0.4228. Similarly, the state probabilities of closing index price for 249\(^{th}\) day will be:

\[
\pi^{(2)} = \pi^{(1)} \times P_{NSE} = [0.4228 \ 0.3293 \ 0.2480] \times \begin{bmatrix}
0.4762 & 0.3619 & 0.1619 \\
0.3086 & 0.3580 & 0.3333 \\
0.3922 & 0.2745 & 0.3333
\end{bmatrix}
\]

\[
249^{th} \text{ Day} = \pi^{(2)} = [0.4228 \ 0.3287 \ 0.2485]
\]

Thus, it can be said that there is a possibility of the NSE index price increases with a probability of 0.4228, unchanged with a probability of 0.3287, and decreases with a probability of 0.2485. These predictions are the same as the actual data.

### The n-step Transition Metrics

The prediction of the long-run behavior of the NSE index is very meaningful and important for investors. The long-run behavior of the NSE index can be determined by using the n-step transition probability matrix. The n-step metrics \(P^{(n)}_{NSE}\) shows the behavior of stock index n-step later. If the number of steps increases then the \(P^{(n)}_{NSE}\) converges to limiting transition metrics \(P^{(\infty)}_{NSE}\), where each row of the matrix is identical, and it is said that the chain has attained a steady-state or state of equilibrium. The steady-state metrics provide the probability of the NSE index increase, unchanged, and decrease in the future. In other words, the \(n^{th}\) power transition probability metrics will provide the probability that the NSE index price will be in a particular state in \(n\) days, given that it is currently in some specified state. To examine the long-term behavior of the NSE index, the higher-order transition probability metrics are calculated with the help of Microsoft Excel.
This result indicates the long-run probability of the NSE index being in increasing, unchanged, or decreasing states. The probability that the NSE index increases at the state of equilibrium are 0.4229, unchanged is 0.3287, and the probability that the NSE index decreases are 0.2484.

Expected number of visits

Here an attempt has been made to find out the expected number of visits, the NSE index makes to a particular state from another state in different steps.

For the NSE index, the number of visits the chain makes in five trading days is given by the following metrics.

\[ \mu_i(5) = \begin{bmatrix} 2.1596 & 1.6880 & 1.1523 \\ 1.9982 & 1.6599 & 1.3419 \\ 2.1914 & 1.5458 & 1.2628 \end{bmatrix} \]

From the matrix obtained above, it may be revealed that if the NSE index starts from the increasing state, the expected number of visits the chain for the NSE index makes to the increasing state out of five trading days is 2.1596, to the state unchanged is 1.6880 and to the state, the decrease is 1.1523.

Likewise, if the NSE index starts from an unchanged state, the expected number of visits the chain makes to the state increase is 1.9982, to the state unchanged is 1.6599 and to the state, a decrease is 1.3419.

Similarly, if the NSE index starts from decreasing state, the expected number of visits the chain makes to the state increase is 2.1914, to the state unchanged is 1.5458 and to the state, a decrease is 1.2628.

Expected return time

It will be meaningful to understand the expected duration the NSE index will stay in the increase, decrease, or unchanged state. The steady-state transition probability is used to determine the expected return time to a state starting from the same state. For a finite irreducible Markov Chain, the expected return time to the same state is a reciprocal of the steady-state probabilities. Here for the NSE index the expected return time to the increasing state, starting from the same increasing state is 1/0.4229 = 2.3646. This result shows that the chain for the NSE index should visit the increasing state on an average in two days. Similarly, the expected return time to remain unchanged, starting from the unchanged state is 1/0.3287 = 3.0423.

\[
\begin{align*}
\text{Vol. 5} & \quad \text{No. 4} & \quad \text{July 2021} & \quad \text{E-ISSN: 2456-5571} \\
\end{align*}
\]
This means the chain for the NSE index should visit the state unchanged on an average of three days. The expected return time to the decreasing state, starting from the decreasing state is $1/0.2484 = 4.0258$. This result helps to conclude that the chain should visit the decreasing state on an average of four days.

Conclusion
To predict the stock market behavior, the Markov Chain model assumes that the performance of the stock market is completely affected by stochastic factors. The movement of a stock index to various states in a particular trading day is independent of the index of initial trading days but depends only on the index of the most recent day. The prediction of the behavior of the stock market is very complex because the operational status of the stock market is subject to the influence of various factors from the market; therefore, there is no single method available that can accurately predict changes in the stock market every day. Hence, investors can combine the results of forecasts from using the Markov chain to predict with other factors and achieve relatively good results.

In this paper, the Markov Chain Model is applied to predict the behavior of the NSE Index. The predicted results are stated in terms of the probability of a certain state of the NSE Index in the future. The model does not provide the prediction results in an absolute state. The initial state vector and the transition probability matrices are used to estimate the probability of the NSE Index being in different states in the future days. The steady-state probabilities are obtained from the n-step transition probability matrices. The result of the steady-state probability matrix shows that the chance of the NSE Index will increase in the future is 0.4229, will be unchanged soon is 0.3287, and will decrease soon is 0.2484. The expected number of visits to a particular state from other states is also computed. The result shows that out of five trading days, the expected number of visits for the NSE Index made to the increasing state starting from the increasing state is 2.1596. The expected number of visits to the decreasing state starting from the decreasing state is out of five trading days for the chain is 1.1523. The result of the expected return time shows that the chain for the NSE index should visit the increasing state on an average in two days and the expected return time to the decreasing state, starting from the decreasing state is four days.

References


