The Principle of the Markov Chain Prediction Model and Its Application in Stock Price Forecasting

Yuzhe Yan
International School Altdorf, Altdorf, Uri, CH-6460, Switzerland
yuzhe.yan@lisa.swiss

Abstract. Markov chains, due to their "memoryless" property, align well with the characteristics of many phenomena in life, making Markov prediction models an important tool for predicting random events in daily life. This paper aims to summarize the rich findings obtained by researchers on Markov prediction models, and to specifically demonstrate the powerful capabilities of Markov prediction models in forecasting stock prices, allowing readers to intuitively experience the differences in prediction accuracy brought about by continuously improved models. Specifically, the paper first introduces the origin and development history of Markov chains. It then briefly covers the relevant knowledge of Markov chains and the basic steps for using them in predictions. Following this, it presents the effects of three progressively advanced Markov chain prediction methods in stock price forecasting. The study finds that when more careful consideration is given to the relationships between variables and more complex and precise algorithms are integrated, the prediction results are more accurate, and the predictability extends over longer periods. Overall, although there are difficulties in data collection and state classification, its advantages in handling stochastic processes make it one of the important choices for practical applications in various fields. With the advancement of technology, these models will become more efficient and accurate, allowing them to be applied in a wider range of fields. Moreover, attempts to combine various algorithms with Markov chain models will be a direction for improving prediction results.

Keywords: Markov chain; stochastic process prediction models; stock prices.

1. Introduction

Markov models are a mathematical tool widely used to describe stochastic processes. After a century of development, the Markov chain's property of "the future being independent of the past given the present" has led to its widespread applications. Its accuracy, adaptability, and predictive power make it a key tool in modern forecasting, with uses in economics, education, medicine, and public health. In the study of practical problems, it is observed that many phenomena arise from continuous changes over time. Some phenomena or processes can be described as follows: when the "present" is known, the "future" of the process is entirely disconnected from the "past". These types of processes with the aforementioned properties as Markov processes.

Markov processes can describe many phenomena in real life. For example, Brownian motion of particles in liquids, machine failures, stock price fluctuations, video signals, and so on. In 1931, Kolmogorov introduced key concepts in probability theory, and three years later, Khintchine developed the theoretical framework for stationary processes. These two influential works laid the foundation for Markov processes and stationary processes. In 1980, Researchers introduced environmental factors into the classical Markov chains for the first time, providing formulas for Markov chains in random environments, and listed cases under special circumstances such as branching processes, queues, birth-death chains, and random walks in random environments [1]. Continuing with in-depth studies, based on Chung's research on Doeblin's limit theorems for Markov processes under general conditions, Steven conducted and organized research on the limit theorems of transition probabilities for Markov chains [2, 3].

The forecast of the Markov chain in the economy and other aspects is the main research content of scientists, and there are also many research results in this field. Turner and Hull theoretically analyses the price fluctuations of related assets, and the model obtained through prediction gives an effective option pricing method for some time in the future [4, 5]. The fluctuation of the stock market is random, and the fluctuation is large. After testing, the researchers found that the stock market price change in the entire securities market was a Markov process. Therefore, the Markov chain can be introduced to discuss the market's strategy for securities investment.

The Markov chain calculation method has application value in the prediction and forecast of the epidemic of hemorrhagic fever. According to the collected data, the researchers determined the value range and a state division of each state and found the first-order transfer high-rate matrix. The Markov chain forecasts and the result of the prediction is interval estimation. Although the relevant error of the prediction is reduced, the accuracy of the prediction is improved, which is a practical value in EHF prevention and control and epidemic prediction. Otherwise, Hidden Markov models are widely applied in ecology, from individual-level to ecosystem-level modelling. Their intuitive structure often corresponds with our conceptual models for ecological systems: there is a hidden process unfolding over time from which we obtain noisy, multivariate observations.

The purpose of this paper is to discuss the Applied research of the Markov model in various fields. The main part of this paper will be focused on the functions of the Markov chain and how it applies to various fields.

2. Basic Descriptions

Definition 2.1: A stochastic process $\{X_n, n \in T\}$ is called a Markov chain if, for any integer $n \in T$ and time $i_0, i_1 ..., i_{n+1} \in I$,the conditional probability satisfies:

$$P\{X_{n+1} = i_{n+1} \mid X_0 = i_0, X_1 = i_1, ..., X_n = i_n\} = P\{X_{n+1} = i_{n+1} \mid X_n = i_n\} \tag{1}$$

Definition 2.2: The conditional probability $p_{ij}^{(n)} = p\{X_{n+1} = j \mid X_n = i\}$,$i, j \in I$ is called the one-step transition probability of the Markov chain $\{X_n, n \in T\}$ at time n ,abbreviated as the transition probability.

When the transition probabilities $p_{ij}^{(n)}$ do not depend on time n ,it indicates that the Markov chain has stationary transition probabilities. If, for any $i, j \in I$, the transition probabilities of the Markov chain are independent of n ,then the Markov chain $\{X_n, n \in T\}$ is said to be homogeneous.

Definition 2.3: Let P denote the matrix of one-step transition probabilities $p_{ij}^{(n)}$, and let the state space be $I = \{1,2,...\}$. Then,

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} & \cdots \\ p_{21} & p_{22} & \cdots & p_{2n} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \tag{2}$$

Where $p_{ij} = p\{X_{n+m} = j \mid X_n = i\}$, is called the one-step transition matrix of the Markov chain. It has the properties: (1) $p_{ij} \geq 0, i, j \in I$; (2) $\sum_{j \in I} p_{ij} = 1, i \in I$.

The conditional probability $p\{X_{n+m} = j \mid X_n = i\}$ is called the n-step transition probability for the Markov chain, which is independent of m, for a time-homogeneous Markov chain, denoted as $p_{ij}^{(n)}$.

For a Markov chain, for any integer $n \geq 0$,$0 \leq l < n$,$i, j \in I$, the n-step transition probabilities satisfy the Chapman-Kolmogorov equations:

$$p_{ij}^{(n)} = \sum_{k \in I} p_{ik}^{(l)} p_{kl}^{n-l} \tag{3}$$

Basic Steps of Markov Chain Prediction Method:

1. Enumerate all possible states of the target variable to be predicted. Divide the data into different states, denoted respectively by $I_1, I_2, ..., I_n$.

2. Calculate the initial state distribution vector.

3. Determine the one-step transition matrix, specifically the transition probability matrix

$$P^{(m)} = \begin{pmatrix} p_{11}^{(m)} & p_{12}^{(m)} & \cdots & p_{1n}^{(m)} \\ p_{21}^{(m)} & \cdots & \cdots & p_{2n}^{(m)} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1}^{(m)} & \cdots & \cdots & p_{nn}^{(m)} \end{pmatrix} \tag{4}$$

Where $p_{ij}^{(m)} = \frac{m_{ij}}{M_i}$ represents the probability of transitioning.

from state I_i to state I_j in m-step, m_{ij} is the number of samples transitioning from state I_i to state I_j in m-step, and M_i is the number of samples in the original data that fall into state I_i based on certain probabilities.

4. Compute the state vectors for each state $\pi(i)$:

$$\pi(1) = \pi(0)p, \pi(2) = \pi(1)p, ..., \pi(i) = \pi(i-1)p (i = 1, 2, ..., n) \tag{5}$$

3. Application of Markov Chains in Stock Price Prediction

Stocks, as one of the most important financial instruments in the financial market, have always been a subject of significant research interest regarding the analysis and prediction of their price fluctuations. Although stock data constitutes a complex nonlinear system structure, and due to its high volatility, past data serves merely as a reference, indicating potential directions of development. Investors often refer to historical stock prices when selecting stocks, but stock prices are also frequently influenced by factors such as national policies, international situations, and social unrest, making current stock prices more referential compared to historical ones. In light of these characteristics, numerous methods for predicting stock prices have emerged. They can be broadly categorized into three types: stock price fluctuation prediction models based on statistical methods, machine learning prediction models, and prediction models based on stochastic processes.

Among stochastic process prediction models, Markov chains, as typical stochastic process analysis methods, naturally incorporate non-factorial and "memoryless" characteristics, making them naturally suited for analyzing stock prices.

Below, this paper introduces three specific methods within the Markov chain approach to examine their practical applications in stock price prediction.

(1) Markov Chain Prediction Model:

Unlike previous applications of the Markov chain method to individual stocks, Yang focuses on the CSI 300 Index (comprising the 300 largest market capitalization stocks in the Shanghai and Shenzhen markets, which has certain predictive significance), validating the short-term predictive effectiveness of the Markov chain and demonstrating the validity of this method [6]. This study collected 43 time series data points from February 19, 2024, to April 19, 2024, to predict daily closing price fluctuations. The basic steps were: classifying the price time series data; establishing the daily closing price fluctuation transition matrix for the stock; conducting a Markov test; and predicting the stock price ranges based on equations. Table 1 presents a comparison table of the prediction results.

Table 1. Comparison Table of Predictive Results

Date	S1	S2	S3	S4	S5	S6	Forecast Interval	Actual Growth Rate	Relative Error
2024/4/22	0	0	0.428	0.286	0.143	0.143	[-0.5%, 0%]	-0.30%	0
2024/4/23	0.153	0.332	0.135	0.186	0.099	0.095	[-1%, -0.5%]	-0.70%	0
2024/4/24	0.080	0.165	0.211	0.200	0.143	0.201	[-0.5%, 0%]	0.44%	0.44%
2024/4/25	0.126	0.205	0.172	0.224	0.134	0.138	[0, 0.5%]	0.25%	0
2024/4/26	0.106	0.196	0.177	0.202	0.139	0.180	[0, 0.5%]	1.53%	1.03%

According to Table 1, the Markov chain has a higher probability of accurately predicting price fluctuation rates for the short term of 1-2 days. However, as the number of days increases, the predictions deviate from actual outcomes, and the probability of successful predictions decreases.

(2) Weighted Markov Chain Prediction Model:

The weighted Markov chain prediction method introduces weights to reflect the relationship between transition probability matrices at various lag times and their impact on prediction, thereby extracting as much information as possible from the initial data [7]. This method represents an improvement over the original Markov chain approach. The steps of the weighted Markov chain prediction method are as Fig. 1.

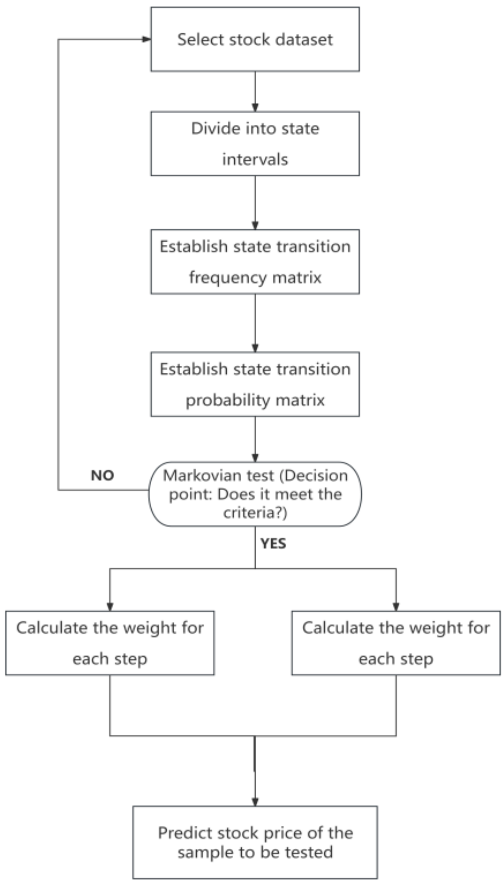

Fig. 1 Flowchart of stock price prediction based on Markov chain (Picture credit: Original).

Chen closes stock prices of China Merchants Port from January 1, 2022, to June 15, 2022, were selected as experimental samples. The comparison of stock state predictions using the Markov chain and the weighted Markov chain revealed notable differences in performance. While the Markov chain consistently predicted a state of 4 for all cases, only 2 of these predictions were correct. In contrast, the weighted Markov chain had only 2 incorrect predictions. Furthermore, when comparing the stock price prediction results, the weighted Markov chain model generally exhibited smaller absolute errors in stock prices, except for the last date. The predicted stock prices from the weighted Markov chain also aligned more closely with the actual stock price trends. To assess the models' generalizability, both the Markov chain and weighted Markov chain were applied to predict 10 identical stocks. The relative errors for these stocks on June 10, along with the average absolute errors from June 8 to June 10, further demonstrated the weighted Markov chain's superior accuracy.

Overall, the weighted Markov chain model achieves higher precision in stock price predictions, indicating its superior ability to assist investors in analyzing multiple stocks.

(3) Hidden Markov Chain Model:

The Hidden Markov Model (HMM), first proposed by Baum and Petrie, is an extension of the mixture model and stands as one of the most popular and widely applied statistical models to date [8]. Its flexibility has enabled applications across various fields, including economic finance, speech recognition, DNA analysis, fault detection, and bioinformatics. The structure of an HMM involves an underlying Markov chain that generates different hidden state sequences at various time points based on initial probability distributions and state transition matrices. At each time point, the hidden state influences and randomly generates the next hidden state through the transition matrix, describing the transition process between different states. Additionally, the hidden states influence and randomly generate observation sequences through the observation matrix, detailing the relationship between hidden states and observations under those states.

The basic steps for prediction using this model are as follows:

1. Select appropriate variables related to stock price fluctuations as the observation sequences for the HMM, collecting and organizing the original data.
2. Determine the number of hidden states in the HMM.
3. Estimate the parameters of the HMM.
4. Proceed to stock price prediction.

Hassan and Nath applied the HMM to airline stock prices, innovatively using multiple variables such as opening and closing prices as parameters [9]. They validated the feasibility of using HMMs for stock price prediction. In another study, they achieved a combination of stock price predictions using a model that integrates the HMM, genetic algorithms, and artificial neural networks [10]. The process involved using an artificial neural network to convert daily stock prices into a set of independent values as input for the HMM, optimizing the initial parameters of the HMM with a genetic algorithm, and, during the prediction phase, assigning different weights to dates with similar patterns, with weights increasing as the dates approached the prediction date. This approach yielded a weighted average of price differences from similar patterns, facilitating the prediction of stock prices for the following day. The results indicated that the integrated model based on HMM, ANN(Artificial Neural Network), and GA(Genetic Algorithm) achieved higher prediction accuracy than the HMM model alone. Jahan and Majid, under the assumption of a normally functioning market and excluding other non-human factors, utilized the HMM to study stock prices at the Tehran Stock Exchange, selecting variables such as closing prices, industry indices, and trading volumes [11]. Their research demonstrated that the precision of parameters obtained from training on the dataset significantly influenced the prediction outcomes. Venugopal employed the HMM for trend analysis in the stock market, categorizing the market into four states and calculating the probabilities of each state occurring based on observation and hidden sequences [12]. These probabilities were used as percentage probabilities for stock price trend changes, aiding investors in making decisions under uncertainty.

In recent years, a dynamic Bayesian network containing hidden variables—the Hidden Semi-Markov Model (HSMM)—has become a powerful tool in the fields of anomaly detection and automatic reasoning. Building upon the HMM, the HSMM incorporates state duration times to more effectively describe the transition relationships of hidden states, providing a more objective description of network states. Chinese scholars have applied this model to market environments, capturing the structural relationships between market sentiments, market states, and state durations, dubbing it the Sentiment Vector-Based Hidden Semi-Markov Model (SV-HSMM). Researchers introduced a sentiment vector to smooth fluctuations in market sentiment and used the Kullback-Leibler (KL) divergence to quantify the intensity of sentiment changes, thereby accurately capturing variations in market sentiment [13]. Utilizing stock market data from 2018 to 2020, including market sentiment indicators and the closing prices of the Shanghai Composite Index, the results showed that the SV-HSMM method, by integrating market sentiment features and HSMM modeling, could more accurately predict turning points in the stock market.

4. Limitations and Prospects

The results of the first case demonstrate that the predictive effectiveness of Markov chains is significantly influenced by historical data. For short-term stock price fluctuations, Markov chains perform exceptionally well, achieving relatively high prediction accuracy. However, as time progresses, although the accuracy of predictions may decrease, the discrepancies between predicted and actual outcomes remain manageable, preserving their referential value. Additionally, stock prices are influenced by various uncertainties, such as national policies, natural disasters, and investor sentiments, which complicate and vary the prediction process. Therefore, Markov chain models are more appropriately applied to sequences that are relatively stable, exhibit low volatility, and span short time periods.

The second case indicates that compared to traditional Markov chain models, the weighted model significantly enhances prediction accuracy, validating its applicability in stock price forecasting. Specifically, the weighted model calculates an average absolute error in stock prices that is significantly lower than that of the non-weighted model. Furthermore, extended experiments demonstrate that the weighted Markov chain model exhibits stronger generalization capabilities across multiple stock price predictions. However, for longer-term predictions, the model's accuracy still requires improvement, which will be a focal point for future research.

The traditional Hidden Markov Model primarily faces the following limitations in practical applications: 1. The observation sequences directly utilize the four types of daily stock price data. 2 [14]. In the prediction methods for stock prices, traditional approaches assume that dates with similar likelihood values correspond to similar price changes, lacking the necessary explanation and theoretical foundation. However, the trend of integrating Hidden Markov Models with other algorithms is becoming increasingly evident, and the Hidden Semi-Markov Model (HSMM) in recent years represents an improved approach with a promising future, deserving of thorough research. Yet, from another perspective, despite the emergence of various combination prediction methods, almost all fail to explain why a particular combination method is preferred over others. Moreover, the correctness of these prediction outcomes often carries significant randomness, leading to a prediction process that lacks systematicness and integrity. These issues will be key areas for future research efforts.

Regardless of the type of Markov chain prediction model employed, the crux of making accurate predictions lies in constructing a reliable transition probability matrix. This requires that the input raw data be extremely accurate to ensure the credibility of the prediction results. However, in reality, obtaining such precise data is often challenging and requires substantial effort, which stands as a major drawback of this method. Additionally, the classification of states is integral to the prediction model. To enhance the rationality, scientific nature, and accuracy of Markov chain models, extensive research is still needed on the scientific rationality of state classification criteria. While common methods like ordered clustering and classification based on sample means and variances are clear and have some scientific basis, there remains room for exploring more scientific and precise classification methods.

5. Conclusion

The core of the Markov chain prediction model lies in its memorylessness, meaning that future states depend solely on the current state, independent of past states. This characteristic makes it particularly effective in handling systems characterized by randomness and dynamic changes.

However, like all models, the Markov chain prediction model has its limitations. The main issue lies in its dependence on accurate transition probability matrices, which require high-quality input data. In the real world, obtaining such precise data is often challenging due to various uncertainties, such as policy changes, market fluctuations, or unforeseen events.

Additionally, the subjectivity in state classification can introduce biases, affecting the accuracy of predictions. Despite these limitations, the Markov chain prediction model remains advantageous because it can handle uncertainty and randomness in complex systems. With advancements in computational power and data science, these models have become more efficient in adapting to new data and improving prediction accuracy. Therefore, it is expected that Markov chain prediction models will continue to play a significant role in the future, especially in addressing complex prediction problems in natural science, social science, and practical applications.

In summary, the Markov chain prediction model is an extremely valuable mathematical tool for analyzing and forecasting uncertain situations. Although challenges exist in data acquisition and state classification, its strengths in handling stochastic processes make it a preferred method in scientific and industrial applications. As technology continues to progress, these models are expected to become more efficient and precise, further expanding their application across various domains.

References

[1] Cogburn R. Markov chains in random environments. The case of Markovian environments. The Annals of Probability, 1980, 8: 908-916.

[2] Chung K. L. The general theory of Markov processes according to Doeblin. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 1964, 2(3): 230-254.

[3] Steven O. Limit Theorems for Markov Chain Transition Probabilities. Van Nostrand Reinhold Co, 1971.

[4] Turner C. M., Starz R., Nelson C. R. A Markov Model of Heteroskedasticity Risk and Learning in the Stock Market. Journal of Finance Economics, 1989, 25(1): 3-22.

[5] Hull J., White A. The Pricing of Option on Assets with Stochastic Volatilities. Journal of Finance, 1987, 42(2): 281-300.

[6] Mu Y., Li D. Study on Stock Price Prediction Based on Markov Chain Method. Financial Development Research, 2024, (08): 89-92.

[7] Zhang C. Integrated Prediction Model Based on Markov Chain. Liaoning Technical University, 2012.

[8] Chen Z., Huang X. Stock Price Prediction Research Based on Weighted Markov Chains. Journal of Dongguan University of Technology, 2023, 30(03): 1-8.

[9] Hassan M. R., Nath B. Stock market forecasting using hidden Markov model: a new approach. 5th International Conference on Intelligent Systems Design and Applications (ISDA'05). IEEE, 2005: 192-196.

[10] Hassan R., Nath B., Kirley M. A fusion model of HMM, ANN and GA for stock market forecasting. Expert Systems with Applications, 2007, 33(1): 171-180.

[11] Jahan M. V. Stock Market Prediction With Hidden Markov Model. ICTCK 2015.

[12] Venugopal D., Kannan Kaliyaperumal S., Muthu Niraikulathan S. Stock market trend prediction using hidden Markov model. In Forecasting in Mathematics - Recent Advances, New Perspectives and Applications. IntechOpen, 2021.

[13] Yao H., Jiang Y., Yang J., Yu K. Stock Market Turning Point Prediction Method Based on Sentiment Vector and Hidden Semi-Markov Model. Journal of Hefei University of Technology (Natural Science Edition), 2024, 47(10): 1335-1340.

[14] Yao P. Stock Price Prediction Research from the Perspective of Improving Hidden Markov Models. Zhejiang University of Finance and Economics, 2022.