International Journal of Mathematics Trends and Technology
ISSN: 2231-5373/ https://doi.org/10.14445/22315373/IJMTT-V72I5P105

Volume 72 Issue 5, 64-73, May 2026

Original Article

A Probabilistic Markov Chain Framework for Portfolio Optimization in the Vietnamese Stock Market

Doan Thanh Son

Faculty of Management Information, University of Finance and Business Administration, Hung Yen, Viet Nam.

Corresponding Author : ngay060779@gmail.com

Received: 27 March 2026 Revised: 30 April 2026 Accepted: 15 May 2026 Published: 29 May 2026

Abstract - *The purpose of this paper is to investigate the theoretical foundations of Markov chains and to develop a model for forecasting stock price volatility based on return series. Building on this framework, the study constructs an optimal investment portfolio composed of securities with superior expected performance. The proposed model is then applied to real-world data from stocks listed on the Ho Chi Minh City Stock Exchange (HOSE). The findings of this research provide a quantitative basis for supporting investors in making more informed and effective capital allocation decisions.*

Keywords - *Markov chain theory, Stochastic processes, Financial markets, Forecasting stock price, Optimal investment portfolio.*

1. Introduction

Vietnam's economy has become increasingly integrated into the global system and continues to experience strong growth. In this context, the stock market has established itself as a vital channel for capital mobilization, contributing significantly to economic development, while also serving as an attractive investment avenue for the public. However, financial markets are inherently characterized by uncertainty, volatility, and substantial risk.

For investors, a fundamental challenge lies in forecasting the trends of stock price movements in order to construct well-balanced portfolios that maximize returns while minimizing risk. To address this issue, alongside traditional fundamental and technical analysis methods, quantitative approaches based on mathematical modeling have gained growing prominence. Among these, Markov chain theory-grounded in probability theory and stochastic processes-has emerged as a powerful tool, demonstrating effectiveness in capturing and modeling the dynamic evolution of financial variables over time [8,10,11].

Building on this practical foundation, this paper investigates the theoretical underpinnings of Markov chains and develops a model for forecasting stock price volatility based on return series. On this basis, an investment portfolio is constructed, comprising securities with superior expected returns. The proposed approach is then applied to empirical data from stocks listed on the Ho Chi Minh City Stock Exchange (HOSE). The results of this study provide a quantitative framework to support investors in making more informed and rational capital allocation decisions.

The paper is organized as follows. Section 2 recalls the fundamental concepts of Markov chains, including the definition of Markov processes, transition probability matrices, the Chapman-Kolmogorov equation, and, in particular, the theory of stationary distributions, which plays a crucial role in long-term forecasting. Section 3 presents the development of a Markov chain-based model for optimal stock portfolio construction. Section 4 applies the proposed model to empirical data from stocks listed on the Ho Chi Minh City Stock Exchange (HOSE). Using historical data on the adjusted closing prices of ten selected stocks, this section conducts analysis, evaluation, and ultimately proposes an optimal investment portfolio.

2. Some knowledge about Markov chains

Consider the time evolution of a given physical system. Let X_t denote the state of the system at time $t \in T$, where X_t is a random variable. The set of all possible states is called the state space of the system and is denoted by E. The collection $X = \{X_t : t \in T\}$ is then referred to as a stochastic process with state space E. If $T = \{0,1,2,\dots\}$, the process X is called a discrete-time stochastic process, whereas if $T = [0,\infty)$, it is called a continuous-time stochastic process (more details, see [1,2,7,9,12]).

In this paper, we focus on a discrete-time stochastic process. $\{X_n\}_{n\geq 0}$, where the state space E is finite or countable, and its elements are denoted by $i, j, k, \ldots$

The stochastic process $\{X_n\}$ is called a Markov chain if

$$P\{X_{n_{k+1}} = i_{k+1} | X_{n_1} = i_1, X_{n_2} = i_2, \ldots, X_{n_k} = i_k\} = P\{X_{n_{k+1}} = i_{k+1} | X_{n_k} = i_k\}$$

For all $n_1 < n_2 < \cdots < n_k < n_{k+1}$ where $i_1, i_2, \ldots, i_k, i_{k+1} \in E$.

The one-step transition probability is defined as the conditional probability that the system, being in state i at time n, moves to state j at time $n+1$. It is denoted by

$$p_{ij} = P(X_{n+1} = j | X_n = i),$$

and the transition probability matrix $P = (p_{ij})$ satisfies the following properties:

- $0 \leq p_{ij} \leq 1, \forall i, j \in E,$
- $\sum_{j\in E} p_{ij} = 1.$

If the state space is $E = \{0,1,2,\ldots\}$, then the transition probability matrix takes the form

$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

The n-step transition probability is defined as

$$p_{ij}(n) = P(X_{n+m} = j | X_m = i).$$

From this, we have

$$p_{ij}(1) = p_{ij}.$$

Additionally, according to the convention

$$p_{ij}(0) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

The matrix $P(n) = (p_{ij}(n))$ is the transition probability matrix after n steps. The elements of this matrix have the following properties:

$$0 \leq p_{ij}(n) \leq 1; \sum_{j\in E} p_{ij}(n) = 1; \forall i, j \in E.$$

Theorem 2.1 ([1]). The joint distribution of $(X_0, X_1, \ldots, X_n)$ is completely determined by the initial distribution and the transition probabilities. Specifically, we have

$$P(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = u_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n},$$

where $u_{i_0} = P(X_0 = i_0)$ is the initial distribution and p_{ij} are the one-step transition probabilities.

Theorem 2.2 (Chapman-Kolmogorov). For all $i, j \in E$ and $n, m > 0$,

$$p_{ij}(n+m) = \sum_{k\in E} p_{ik}(n) p_{kj}(m).$$

In matrix form, this can be written as

$$P(n+m) = P(n)P(m).$$

Since $P = P(1)$, it follows by induction that.

$$P(n) = P^n.$$

Definition 2.3 (Stationary Distribution).
Let $U = (u_i), i \in E$ be an initial distribution. It is called a stationary distribution if

$$U(n) = U, \forall n \geq 0,$$

that is,

$$u_i(n) = u_i, \forall i \in E, \forall n \geq 0.$$

In this case, the sequence (X_n) has the same distribution at all times.

From the Chapman–Kolmogorov theorem, it follows that U is a stationary distribution if and only if

$$UP = U.$$

Equivalently, $U = (u_i), i \in E$ is a stationary distribution if and only if:

- $u_i \geq 0, \sum_{i \in E} u_i = 1$,
- $u_j = \sum_{i \in E} u_i p_{ij}, \forall j \in E$.

In other words, a stationary distribution is a probability distribution that remains unchanged under the action of the transition matrix of the Markov chain.

Theorem 2.4.
Let (X_n) be a Markov chain with state space $E = \{1,2, \dots\}$, transition matrix $P = \left(p_{\mathrm{i}}\right)$, and n-step transition matrix $P(n) = \left(p_{ij}(n)\right)$. Assume that for all $i, j \in E$, the limit.

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

exists and is independent of i. Then

1. The vector $\pi = (\pi_j)$ satisfies
   $\sum_{j \in E} \pi_j \leq 1$ and $\pi_j = \sum_{i \in E} \pi_i p_{ij}$,
2. Exactly one of the following holds
   - $\sum_{j \in E} \pi_j = 1$;
   - $\pi_j = 0$ for all $j \in E$,
3. If $\sum_{j \in E} \pi_j = 1$, then π is the unique stationary distribution of the chain. If $\pi_j = 0$ for all $j \in E$, then the chain has no stationary distribution of this form.

Definition 2.5 (Limiting Distribution)
Let (X_n) be a Markov chain with state space $E = \{1,2, \dots\}$, transition matrix $P = \left(p_{ij}\right)$, and n-step transition matrix $P(n) = \left(p_{ij}(n)\right)$. We say that the chain has a limiting distribution if, for all $i, j \in E$, the limit.

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

exists and is independent of the initial state i. Moreover, the vector $\pi = \left(\pi_j\right), j \in E$ satisfies

$$\pi_j \geq 0; \sum_{j \in E} \pi_j = 1,$$

so that π is a probability distribution on E. In this case, π is called the limiting distribution of the Markov chain.

Remark

- If at some time n_0, the Markov chain has a stationary distribution π, i.e., $P\left(X_{n_0} = i\right) = \pi_i, \ \forall i \in E$, then for all $m \geq n_0$, $P(X_m = i) = \pi_i, \forall i \in E$. In other words, once the chain reaches a stationary distribution, its distribution remains unchanged over time.
- If a limiting distribution π exists, then a stationary distribution exists and coincides with it, i.e., $\pi P = \pi$. However, the converse is not true: there exist Markov chains that admit stationary distributions but do not possess a limiting distribution (for example, periodic chains).

Theorem 2.6.

Let (X_n) be a Markov chain with finite state space $E = \{1,2,\dots,d\}$ and transition matrix P. Then the following statements are equivalent

(i) There exists a limiting distribution $\pi = (\pi_1, \pi_2, \dots, \pi_d)$ such that

$$\lim_{n\to\infty} p_{ij}(n) = \pi_j, \forall i, j \in E,$$

with $\pi_j > 0$ for all $j \in E$.

(ii) The Markov chain is regular, i.e., there exists an integer $n_0 \geq 1$ such that

$$p_{ij}(n) > 0, \qquad \forall i, j \in E.$$

In this case, the limiting distribution π is the unique stationary distribution of the chain and satisfies.

$$\pi = \pi P, \qquad \sum_{j=1}^{d} \pi_j = 1.$$

Equivalently, π is obtained as the (normalized) solution of

$$\pi(I - P) = 0.$$

3. Optimal Markov Chain Application for Stock Investment Portfolio

In financial markets, accurately forecasting stock prices is a major challenge due to their stochastic nature and volatility. However, studies by McQueen and Thorley [5] and Doubleday and Esunge [4] suggest that stock returns are not entirely random but exhibit dependence on previous states.

The model in this project is based on the Markov assumption, which states that:

- The state of daily returns at time $n+1$, denoted by X_{n+1}, depends only on the state at time n, X_n, and is independent of earlier states $(X_{n-1}, X_{n-2}, \dots)$.
- Formally, the conditional probability satisfies.

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j | X_n = i).$$

Under this assumption, the complex problem of forecasting stock price movements can be simplified to estimating transition probabilities between states. This, in turn, allows us to analyze the long-term (steady-state) behavior of the system through tools such as transition matrices and stationary distributions.

3.1. Constructing the State Space from the Rate of Return

The project uses the continuous rate of return (log-return) instead of the simple rate of return, defined by [3,6]

$$r_n = \ln\left(\frac{S_n}{S_{n-1}}\right) = \ln(S_n) - \ln(S_{n-1})$$

where S_n and S_{n-1} are the closing prices of the stock on day n and day $n-1$, respectively.

The return rate r_n is a continuous random variable. To apply a Markov chain with a finite state space, we discretize the range of r_n. We partition the state space into three states representing market trends $E = \{S_1, S_2, S_3\}$. The classification is defined as follows:

- State S_1 (Decrease): $r_n < -0.005$,
- State S_2 (Stable): $-0.005 \leq r_n \leq 0.005$,
- State S_3 (Increase): $r_n > 0.005$.

Thus, each observed return r_n is mapped to a discrete state, allowing the evolution of stock returns to be modeled as a Markov chain on a finite state space.

3.2. Transition Probability Matrix

To construct the transition probability matrix, the first step is to count the frequencies of state transitions from historical data. Let N be the total number of observations. We define the frequency matrix $F = \left(f_{ij}\right)_{3\times 3}$, where f_{ij} denotes the number of observed transitions from state i to state j in one time step.

The transition probability p_{ij} is estimated based on observed frequencies

$$p_{ij} = P(X_{n+1} = j|X_n = i) \approx \frac{f_{ij}}{\sum_{k=1}^{3} f_{ik}}.$$

The transition probability matrix P is constructed as follows.

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix},$$

where $p_{ij} \geq 0, \forall i, j \in E$ and $\sum_{j=1}^{3} p_{ij} = 1, \forall i \in E$.

3.3. Stationary Distribution and Long-Term Forecasting

The stationary distribution $\pi = (\pi_1, \pi_2, \pi_3)$ is a probability vector satisfying the balance equation $\pi = \pi P$. From an economic perspective, the stationary distribution represents the long-run probabilities that the stock return is in the Declining (S_1), Stable (S_2), or Increasing (S_3) state, regardless of the initial market condition. Therefore, it provides an important basis for assessing the intrinsic trend of the stock, reflecting its long-term behavior under the underlying stochastic dynamics.

To find the stationary distribution $\pi = (\pi_1, \pi_2, \pi_3)$, we solve the following system of linear equations

$$\begin{cases} \pi_1 p_{11} + \pi_2 p_{21} + \pi_3 p_{31} = \pi_1 \\ \pi_1 p_{12} + \pi_2 p_{22} + \pi_3 p_{32} = \pi_2 \\ \pi_1 p_{13} + \pi_2 p_{23} + \pi_3 p_{33} = \pi_3 \\ \quad \pi_1 + \pi_2 + \pi_3 = 1 \end{cases}$$

Equivalently, this system can be written in matrix form as

$$\pi(I - P) = 0 \text{ subject to } \sum_{j=1}^{3} \pi_j = 1.$$

3.4. Long-Term Portfolio Strategy Using Stationary Distribution

To quantify and compare the potential among different stock codes, we define a score function $Score_i$ for each stock i:

$$Score_i = \pi_3^{(i)} - \omega . \pi_1^{(i)}$$

where

- $\pi_3^{(i)}$: The long-term probability that stock i is in the Increasing state S_3. This reflects the profit potential of the stock,
- $\pi_1^{(i)}$: The long-term probability that stock i is in the Decreasing state S_1. This represents the risk level of the stock,
- ω: The risk aversion parameter, which balances the trade-off between return and risk. In the baseline model, we set $\omega = 1$, implying that risk and return are weighted equally.

3.4.1. Decision-Making Criteria

Based on the score values, stocks are classified and ranked as follows [6,13,16]:

- *Potential Group*: $Score_i > 0$, and the larger the score, the more attractive the stock. This implies $\pi_3 > \pi_1$, meaning that, in the long run, the stock is more likely to be in the Increasing state than in the Decreasing state.
- *Risk Group*: $Score_i < 0$. In this case $\pi_3 < \pi_1$, indicating that the stock tends to decline more frequently than it increases over time.
- *Neutral Group*: $Score_i \approx 0$. This suggests $\pi_3 \approx \pi_1$, indicating a balanced behavior between upward and downward movements.

4. Applying the Proposed Model to Empirical Data from the Vietnamese Stock Market

4.1. Description of Data and Initial Probability Distribution

To ensure that the research results are highly representative of the Vietnamese stock market, the study selected a sample of 10 large-cap, highly liquid stocks representing key sectors of the economy, all listed on the Ho Chi Minh City Stock Exchange (HOSE). The input data consists of a time series of daily adjusted closing prices.

Data source: The data were collected and extracted from the financial information website CafeF (https://s.cafef.vn).

Collection period: From October 1, 2024, to October 1, 2025.

Sample size: A total of 249 adjusted closing price observations for each stock, ensuring sufficient time series length for reliable estimation of Markov chain parameters.

List of stocks:

No.	Stock	Company
1	VHM	Vinhomes Joint Stock Company
2	FPT	FPT Corporation
3	GVR	Vietnam Rubber Industry Group
4	HPG	Hoa Phat Group Joint Stock Company
5	SAB	Saigon Beer-Alcohol-Beverage Corporation
6	SSI	SSI Securities Corporation
7	VIC	Vingroup Group
8	VJC	VietJet Air Joint Stock Company
9	VNM	Vietnam Milk Joint Stock Company
10	VPB	VietinBank (Vietnam Prosperity Joint Stock Commercial Bank)

The detailed statistical results on the frequency distribution and initial probability distribution for the 10 selected stock codes are presented in the following table.

No.	Stock	Frequency S_1	Frequency S_2	Frequency S_3	Initial Distribution $U(0)$
1	VHM	69	87	91	(0.2794, 0.3522, 0.3684)
2	FPT	90	81	77	(0.3629, 0.3266, 0.3105)
3	GVR	101	48	99	(0.4073, 0.1935, 0.3992)
4	HPG	80	81	87	(0.3226, 0.3266, 0.3508)
5	SAB	70	117	61	(0.2823, 0.4718, 0.2459)
6	SSI	86	71	91	(0.3468, 0.2863, 0.3669)
7	VIC	58	103	86	(0.2348, 0.4170, 0.3482)
8	VJC	75	110	62	(0.3036, 0.4453, 0.2510)
9	VNM	70	120	57	(0.2834, 0.4858, 0.2308)
10	VPB	83	80	84	(0.3360, 0.3239, 0.3401)

Remark: Based on the above table, a clear differentiation in the dynamic characteristics of the selected stocks can be observed [14,15]:

High Stability Group: Stocks such as VNM ($u_2 = 48.58\%$), SAB ($u_2 = 47.18\%$), and VJC ($u_2 = 44.53\%$) exhibit a dominant proportion of returns in the state S_2. This indicates a relatively stable price behavior with low volatility, reflecting the defensive nature of these stocks.

High Volatility Group: The GVR stock shows the lowest level of stability ($u_2 = 19.35\%$), while the probabilities of the Decrease (S_1) and Increase (S_3) states are both notably high (40.73% and 39.92%, respectively). This suggests that the stock experiences strong fluctuations and therefore carries a higher level of risk.

Positive Trend Group: Stocks such as VHM and VIC demonstrate a significantly higher probability of being in the Increase state (S_3) compared to the Decrease state (S_1). In particular, VIC has $u_3 = 34.82\%$ exceeding $u_1 = 23.48\%$, indicating a prevailing upward tendency during the study period.

4.2. Construct the Frequency Matrix and the Corresponding Transition Probability Matrix

Let f_{ij} denote the number of observations in which the stock transitions from state i at time t to state j at time $t + 1$ during the study period. The transition frequency matrix F is defined as follows.

$$F = [f_{ij}]_{3\times 3} = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$$

Based on the historical data of the 10 selected stocks, the transition frequency matrices are constructed as follows.

Stock	Frequency matrix ***F***
VHM	$\begin{bmatrix} 17 & 30 & 21 \\ 24 & 37 & 26 \\ 28 & 19 & 44 \end{bmatrix}$
FPT	$\begin{bmatrix} 31 & 32 & 27 \\ 34 & 20 & 26 \\ 25 & 28 & 24 \end{bmatrix}$
GVR	$\begin{bmatrix} 40 & 17 & 44 \\ 22 & 10 & 16 \\ 39 & 20 & 39 \end{bmatrix}$
HPG	$\begin{bmatrix} 25 & 29 & 26 \\ 23 & 30 & 28 \\ 31 & 22 & 33 \end{bmatrix}$
SAB	$\begin{bmatrix} 22 & 33 & 15 \\ 33 & 56 & 27 \\ 14 & 28 & 19 \end{bmatrix}$
SSI	$\begin{bmatrix} 30 & 21 & 35 \\ 26 & 22 & 22 \\ 29 & 28 & 34 \end{bmatrix}$
VIC	$\begin{bmatrix} 18 & 18 & 21 \\ 15 & 60 & 28 \\ 24 & 25 & 37 \end{bmatrix}$
VJC	$\begin{bmatrix} 29 & 24 & 21 \\ 25 & 59 & 26 \\ 21 & 26 & 15 \end{bmatrix}$
VNM	$\begin{bmatrix} 17 & 35 & 18 \\ 30 & 63 & 27 \\ 23 & 21 & 12 \end{bmatrix}$
VPB	$\begin{bmatrix} 34 & 21 & 27 \\ 23 & 33 & 24 \\ 25 & 26 & 33 \end{bmatrix}$

And the transition probability matrices

Stock	Transition probability matrix ***P***
VHM	$\begin{bmatrix} 0.2500 & 0.4412 & 0.3088 \\ 0.2759 & 0.4253 & 0.2989 \\ 0.3077 & 0.2088 & 0.4835 \end{bmatrix}$
FPT	$\begin{bmatrix} 0.3444 & 0.3556 & 0.3000 \\ 0.4250 & 0.2500 & 0.3250 \\ 0.3247 & 0.3636 & 0.3117 \end{bmatrix}$
GVR	$\begin{bmatrix} 0.3960 & 0.1683 & 0.4356 \\ 0.4583 & 0.2083 & 0.3333 \\ 0.3980 & 0.2041 & 0.3980 \end{bmatrix}$
HPG	$\begin{bmatrix} 0.3125 & 0.3625 & 0.3250 \\ 0.2840 & 0.3704 & 0.3457 \\ 0.3605 & 0.2558 & 0.3837 \end{bmatrix}$

SAB	$\begin{bmatrix} 0.3143 & 0.4714 & 0.2143 \\ 0.2845 & 0.4828 & 0.2328 \\ 0.2295 & 0.4590 & 0.3115 \end{bmatrix}$
SSI	$\begin{bmatrix} 0.3488 & 0.2442 & 0.4070 \\ 0.3714 & 0.3143 & 0.3143 \\ 0.3187 & 0.3077 & 0.3736 \end{bmatrix}$
VIC	$\begin{bmatrix} 0.3158 & 0.3158 & 0.3684 \\ 0.1456 & 0.5825 & 0.2718 \\ 0.2791 & 0.2907 & 0.4302 \end{bmatrix}$
VJC	$\begin{bmatrix} 0.3919 & 0.3243 & 0.2838 \\ 0.2273 & 0.5364 & 0.2364 \\ 0.3387 & 0.4194 & 0.2419 \end{bmatrix}$
VNM	$\begin{bmatrix} 0.2429 & 0.5000 & 0.2571 \\ 0.2500 & 0.5250 & 0.2250 \\ 0.4107 & 0.3750 & 0.2143 \end{bmatrix}$
VPB	$\begin{bmatrix} 0.4146 & 0.2561 & 0.3293 \\ 0.2875 & 0.4125 & 0.3000 \\ 0.2976 & 0.3095 & 0.3929 \end{bmatrix}$

4.3. Estimate the Limiting (stationary) distribution and analyze the long-term trend of the system

The distribution of the limit $\pi = (\pi_1, \pi_2, \pi_3)$ is obtained by solving the system

$$\begin{cases} \pi P = \pi \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases}$$

The limiting distributions for all 10 stock codes are obtained by applying the above system of equations to each stock. Based on the difference $\pi_3 - \pi_1$, which reflects the balance between upward and downward movements, appropriate investment recommendations are proposed.

No.	**Stock**	**π_1(Decrease)**	**π_1(Stable)**	**π_3(Increase)**	**Score**	**Conclusion**
1	VIC	0.2317	0.4187	0.3496	+0.1179	Strong Buy
2	VHM	0.2804	0.3496	0.3700	+0.0896	Buy
3	HPG	0.3201	0.3275	0.3525	+0.0324	Buy
4	SSI	0.3442	0.2877	0.3680	+0.0238	Buy
5	VPB	0.3333	0.3252	0.3415	+0.0082	Reminder
6	GVR	0.4087	0.1903	0.4011	-0.0076	Avoid
7	SAB	0.2792	0.4737	0.2470	-0.0322	Avoid
8	FPT	0.3644	0.3239	0.3117	-0.0527	Avoid
9	VJC	0.3057	0.4420	0.2523	-0.0534	Avoid
10	VNM	0.2852	0.4831	0.2317	-0.0535	Avoid

4.4. Analyze the obtained Results and Construct an Optimal Investment Portfolio

Based on the quantitative results, we proceed to analyze and propose a specific investment portfolio allocation strategy as follows:

4.4.1. Potential Stock Group

The primary recommendation focuses on stocks for which the probability of long-term growth (π_3) exceeds the probability of decline (π_1).

- VIC (Vingroup) and VHM (Vinhomes): These two stocks exhibit the most favorable signals in the model. In particular, VIC has $\pi_3 \approx 35\%$ while $\pi_1 \approx 23\%$, resulting in a substantial positive margin. This suggests strong underlying fundamentals and a sustained upward tendency during the study period.
- SSI and HPG: These leading stocks in the securities and steel sectors also demonstrate a positive balance, with a higher likelihood of price increases relative to declines. They represent suitable candidates for portfolio diversification, contributing to the mitigation of systematic risk.

4.4.2. Stocks Requiring Caution

Conversely, the model provides warning signals for stocks where the probability of decline (π_1) exceeds the probability of Increase (π_3).

- FPT and GVR: Although these are large-cap companies, the historical data during the study period indicate considerable downward pressure. For instance, GVR exhibits a relatively high π_1 (approximately 40.87%), suggesting elevated risk. Investors are therefore advised to remain cautious, either waiting for clearer reversal signals or maintaining only a limited allocation in these stocks.
- VNM, SAB, VJC: These stocks are characterized by a high probability of the Stable state (π_2) or a tendency toward weaker price movements. As such, they are more suitable for defensive investment strategies rather than for short-term profit seeking.

4.4.3. Proposed Investment Portfolio

Based on the model's quantitative indicators, the recommended portfolio allocation concentrates on four stocks with the most favorable characteristics:

- VIC: 30%,
- VHM: 30%,
- SSI: 20%,
- HPG: 20%.

This allocation emphasizes stocks with strong upward tendencies ($\pi_3 > \pi_1$) while maintaining a degree of diversification across sectors, thereby balancing return potential and risk exposure.

5. Conclusion

This paper develops a modeling framework for stock price fluctuations by transforming adjusted closing prices into return series. It is assumed that the daily return series of each stock follows a three-state Markov chain. Based on this assumption, a transition probability matrix is constructed, and the corresponding stationary distribution is derived to support stock selection for portfolio construction.

The proposed model is applied to empirical data from 10 stocks listed on the HOSE exchange over the period from October 1, 2024, to October 1, 2025. The estimated stationary distributions provide insights into the long-run probabilities of upward and downward movements for each stock. Based on a predefined objective function, the study classifies and ranks the stocks, and subsequently proposes an optimal investment portfolio aimed at maximizing expected returns while minimizing risk.

References

[1] D.H. Thang, *Random Processes and Stochastic Calculations*, Hanoi National University Publishing House, 2006.

[2] N.D. Tien, *Probability Models and their Applications. Part I: Markov Chains and Applications*, Hanoi National University Publishing House, 2000.

[3] S.O.N. Agwuegbo, A.P. Adewole, and A.N. Maduegbuna, "A Random Walk Model for Stock Market Prices," *Journal of Mathematics and Statistics*, vol. 6, no. 3, pp. 342-346, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[4] Kevin J. Doubleday, and Julius N. Esunge, "Application of Markov Chains to Stock Trends," *Journal of Mathematics and Statistics*, vol. 7, no. 2, pp. 103-106, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[5] Grant Mcqueen, and Steven Thorley, "Are Stock Returns Predictable? A Test Using Markov Chains," *The Journal of Finance*, vol. 46, no. 1, pp. 239-263, 1991. [CrossRef] [Google Scholar] [Publisher Link]

[6] Milan Svoboda, and Ladislav Lukáš, "Application of Markov Chain Analysis to Trend Prediction of Stock Indices," *Proceedings of 30th International Conference Mathematical Methods in Economics*, pp. 848-853, 2012. [Google Scholar] [Publisher Link]

[7] Athanasios Papoulis, and S. Unnikrishna Pillai, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill Science, 2001. [Google Scholar] [Publisher Link]

[8] Sean P. Meyn, and Richard L. Tweedie, *Markov Chains and Stochastic Stability*, Cambridge University Press, 2009. [Google Scholar] [Publisher Link]

[9] Emanuel Parzen, *Stochastic Processes*, Courier Dover Publications, 2015. [Google Scholar] [Publisher Link]

[10] John Lamperti, *Stochastic Processes: A Survey of the Mathematical Theory*, Springer-Verlag, 1977. [Google Scholar] [Publisher Link]

[11] Sean P. Meyn, and Richard L. Tweedie, *Markov Chains and Stochastic Stability*, Cambridge University Press, 2009. [Google Scholar] [Publisher Link]

[12] Donald L. Snyder, and Michael I. Miller, *Random Point Processes in Time and Space*, Springer Science & Business Media, 2012. [Google Scholar] [Publisher Link]

[13] Ravi Bansal, Dana Kiku, and Amir Yaron, "Risks for the Long Run: Estimation with Time Aggregation," *Journal of Monetary Economics*, vol. 82, pp. 52-69, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[14] Nicholas Barberis, Andrei Shleifer, and Robert Vishny, "A Model of Investor Sentiment," *Journal of Financial Economics*, vol. 49, no. 3, pp. 307-343, 1998. [CrossRef] [Google Scholar] [Publisher Link]

[15] Lars Peter Hansen, John C. Heaton, and Nan Li, "Consumption Strikes Back? Measuring Long-Run Risk," *Journal of Political Economy*, vol. 116, no. 2, pp. 260-302, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[16] Lars Peter Hansen, and Ravi Jagannathan, "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, vol. 99, pp. 225-262, 1991. [CrossRef] [Google Scholar] [Publisher Link]