Markov Chain based Stochastic Processes for Portfolio Optimization

MATH 171

Peter Yim, Arjun Parikh, Takao Oba

https://github.com/toba717/financial_engineering_projects/tree/main/markov_chain_based_stochastic_processes_for_portfolio_optimization

I. Introduction

Stochastic processes are mathematical models that are used to describe random events or behavior over time. In finance, stochastic processes are particularly useful for modeling the random fluctuations in financial markets. One popular type of stochastic process used in finance is Markov Chains, which are a mathematical framework that allows us to model the transition probabilities between different states of a system.

The objective of our research paper is to explore the applications of stochastic processes, particularly Markov Chains, in portfolio optimization. Portfolio optimization is the process of constructing a portfolio of assets that maximizes expected returns while minimizing the risk of the investor. To achieve this objective, we plan to build a model using Markov Chains to quantify three states (bullish, neutral, or bearish) and ultimately determine which stocks have the highest probability of positive returns.

II. Research Review

The research article, "Portfolio Optimization by Applying Markov Chains" by Nina Petkovic, Milan Bozinovic, and Sanja Stojanovi explores the use of the Markov chains method in portfolio optimization in the Belgrade Stock Exchange. Petkovic et al. (2018) highlights the lack of sufficient research done in this area and how the Markov chains method has been widely used in financial market analyses worldwide, despite the level of their development. The Markov chains method is non-parametric, less complex, and has been found to produce similar results to the Harry Markowitz model but faster and easier to obtain. This research is unique as it is the first to employ the Markov chains method in the returns analysis on the Belgrade Stock Exchange. The study provides insight into the use of the Markov chains method in portfolio optimization, especially in emerging markets, and its potential benefits over traditional methods.

III. Proposal of Stochastic Process

Stochastic processes have been widely used in finance to model the randomness of financial markets. We introduce Markov Chains, which is a type of a stochastic process that is characterized by the Markov property, which states that the probability of transitioning from one state to another only depends on the current state of the system.

We propose to use Markov Chains to model the long term behavior of stock prices which can provide invaluable information when optimizing one's portfolio. Markov Chains involve state and transitions. We characterize the states to be the following: 1. The stock is most likely going to rise in price, 2. The stock is most likely going to remain at its current price, and 3. The stock is most likely going to decline in price. These three states will be unique to all stocks that we will be examining. However, the transition probabilities will be unique for all stocks and will be calculated through assessing past values. In other words, by using Markov Chains, we will consider historical price changes to determine the probabilities of going from one state to another, which will help us in building a probability transition matrix. Through manipulation of this chain, we can then find long-term probabilities through calculating the stationary distribution of each state.

IV. Computational Analysis of the Model

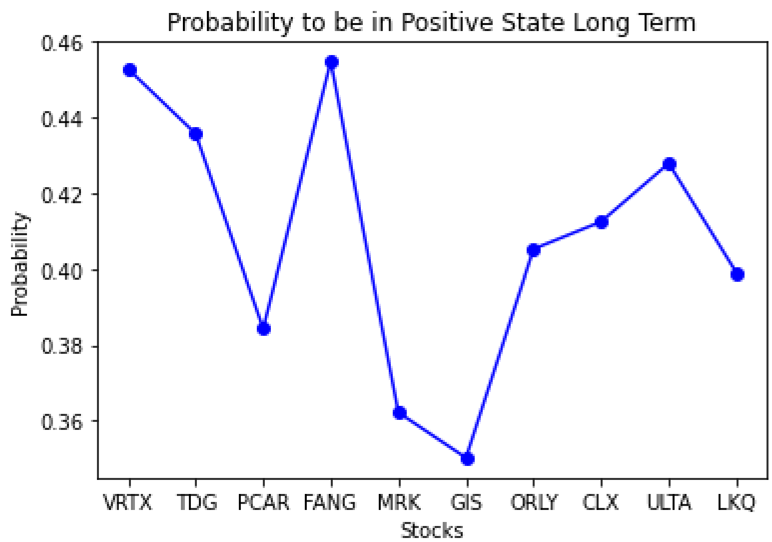
We pulled data on daily returns for the 500 stocks making up the S&P500 from Yahoo Finance. We then computed the Probability Transition matrix for each stock based on a three state model, where the states were a daily return less than -0.005, a daily return in the range -0.005 to 0.005, and a daily return greater than 0.005. We built the Probability Transition Matrix by finding the number of transitions from one state to another, then dividing that value by the total number of days that the stock was in each state (e.g. To find p_{1,2}, we took the number of transitions from a return less than -0.005 to a return in the range -0.005 to 0.005, and divided this by the total number of days the stock had a return less than -0.005). We restricted our computations of these matrices to the ten stocks we found to have the most number of positive trading days over the past year. These stocks have tickers VRTX, TDG, PCAR, FANG, MRK, GIS, ORLY, CLX, ULTA, LKQ. We then calculated the stationary distributions for each of these stocks to see their long term probabilities of being negative, neutral, and positive. The results were as follows:

```
stationary distribution of  VRTX [0.34278249 0.2043344  0.45288311]
stationary distribution of  TDG [0.30977301 0.25448453 0.43574246]
stationary distribution of  PCAR [0.31993319 0.29569762 0.38436919]
stationary distribution of  FANG [0.2843297  0.26087662 0.45479368]
stationary distribution of  MRK [0.31542873 0.32238054 0.36219073]
stationary distribution of  GIS [0.32446809 0.32532187 0.35021004]
stationary distribution of  ORLY [0.29874676 0.29599459 0.40525865]
stationary distribution of  CLX [0.30036404 0.2871515  0.41248446]
stationary distribution of  ULTA [0.28810679 0.28400384 0.42788937]
stationary distribution of  LKQ [0.33465502 0.26629158 0.3990534 ]
```

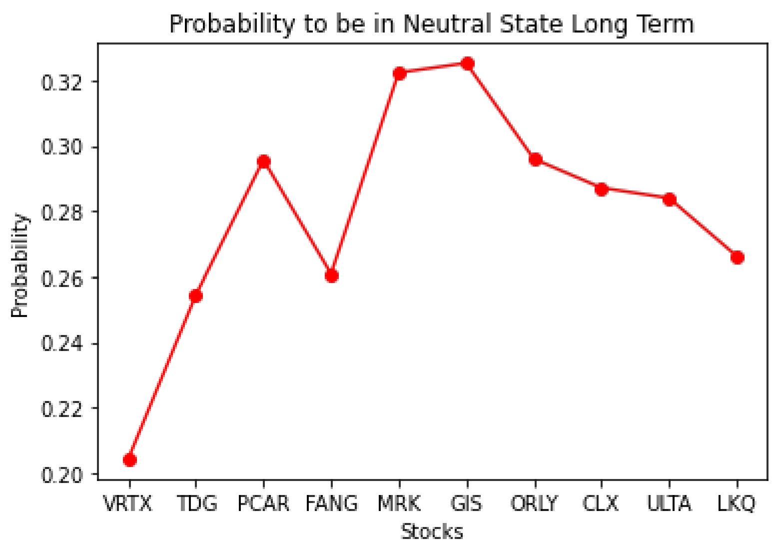
V. Conclusion

In this research paper, we have explored the applications of stochastic processes, particularly Markov Chains, in the realm of finance and specifically to optimize one's portfolio. We have built a model using Markov Chains to determine the probabilities of stocks moving between negative, neutral, and positive states. Through this model, we were able to select the ten most profitable stocks to invest in based on previous data as a way of "optimizing a portfolio".

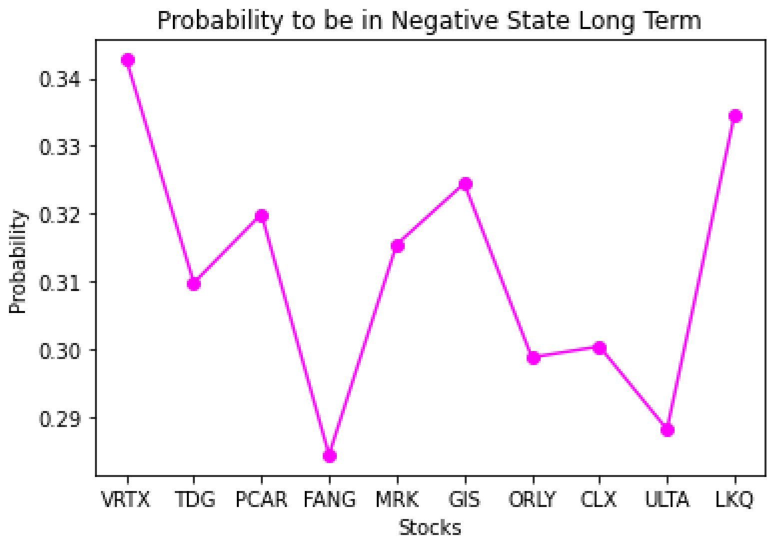
The following are charts representing the stationary distribution, or the long term probabilities for each of the stocks to be in our specified state space. Our model suggested that the stocks VRTX, FANG, and TDG have the highest probability of staying in a positive state over the long term, that is, these are the stocks that have the highest probability of returning daily returns greater than 0.005.



Our model suggested that the stocks MRK, GRS, ORLY, and PCAR have the highest probability of staying in the neutral state over the long term, that is, these are the stocks that have the highest probability of returning daily returns in the range -0.005 to 0.005.



Our model suggested that the stocks VRTX, GIS, and LKQ have the highest probability of staying in a negative state over the long term, that is, these are the stocks that have the highest probability of returning daily returns less than -0.005.



It was interesting to see VRTX as the highest in likelihood to be both in the positive and negative states, while having the lowest likelihood of being in the neutral state, thus suggesting heightened risk with investing in this stock. This model can be helpful in informing investing decisions based on risk tolerance. Those investors looking for lower risk investments would likely invest more in the stocks with higher probability of being in a neutral state, while those with more of a risk appetite may consider investing in stocks like VRTX.

While our model has shown promising results, there can be limitations to our approach. It is important to note that for the purpose of this project we chose to focus on the stocks from the S&P500 that had the most number of positive trading days over the past trading year. This said, this process can be used to analyze any stock with sufficient available data to aid in investing decisions and risk tolerance. In addition to historic price, there could be various factors involved in the price movement of a stock and unexpected world-wide events may occur such as the COVID-19 pandemic. Furthermore, the chosen stocks that had the most days with positive returns can be correlated with each other in terms of company-specific factors and sector-specific factors and thus we may have to be careful when investing collectively in a portfolio.

In future research, we plan to expand our model to include other factors that may affect the behavior of financial markets, such as macroeconomic indicators and geopolitical events. It may also be worth testing models based on various time intervals of data. In addition, we plan to test the robustness of our model by applying it to different markets and time periods.