Portfolio Optimization of Ten Assets in US Stock Market

Jiahua Liu

College of art and science, New York University, New York, USA

jl10453@nyu.edu

Abstract. Recent years, the Covid-19 epidemic, Russia-Ukraine war and the raising of interest rate in US led to a significant fluctuate in the U.S. stock market, which then led investors to focus more on portfolio optimization to reduce their invest risk, and this paper selected ten U.S. stocks with high market capitalization in several different industries. All five strategies, AR (1) and AR (2), Cross-Sectional, LASSO reg, GBDT, KNN(Correlation) and (Cosine) in this paper produced different return rate, volatility and sharpe ration with different portfolio. The result shows that assets in different industries are more likely to be collect in the portfolio than assets in same or similar industry. This paper compared their results, returns and their sharpe ratio with two most important index in the U.S. stock market, IXIC, SPX. and found the relative more effective portfolio optimization method. The results in this paper benefit the related investors in financial markets.

Keywords: Portfolio Optimization, Sharpe Ratio, US Stock.

1. Introduction

In today's financial markets, where investment markets have become extreme volatile due to epidemics, Russo-Ukrainian War and inflation. Portfolio optimization has become a important strategy for both individual investors and institutions alike [1]. As the beginning of epidemic in 2020, the U.S. stock market suffered a devastating blow. The Federal Reserve took a series of market rescue activities to recover the U.S. stock market in a short period of time. But as more us dollars poured into the market; U.S. inflation reached a nearly 40-year new high on the annually increase rate [2]. In this paper, the Sharpe Ratio is the most important measurement which is created by Professor William Sharp in 1966 [3], and since then it has been one of the most important and popular measures for risk/return in the financial markets. Professor Sharpe won the Nobel Memorial Prize in 1990. After that, Sharpe Ratio became extremely helpful in all the stock risk/return studies [4]. In the case of doing portfolio optimization, the most important thing that investors want is to be able to make as much profit as possible while taking relatively less risk. Therefore, to balance the risks in different industries, outstanding investors usually diversify their money into different assets [5]. What is needed to build an ideal portfolio is a good asset allocation. Most investment returns rely on effective asset allocation. However, a current fact is that in the investment field, many investors don't realize this. Many single investments can yield high returns in the short term, but the fact is that in value investing, long-term stable income is more important, because in the business world there are often uncontrollable risks, such as financial crises, epidemics, wars, etc. that have devastated many investors and even investment institutions. In such cases, only a good asset allocation and portfolio can help investors to overcome the difficulties. Most investors just randomly choose a good-looking portfolio based on their perception alone, and few investors put various strategies together to calculate and compare the results to determine the optimal portfolio approach. In that case, so in this research, 5 different portfolio optimization strategy brought together [6], AR (1) and (2), Cross- Sectional, LASSO reg [7], GBDT, KNN(Correlation) and (Cosine). This paper compared and analysis the difference of their return, volatility [8] and the sharpe ratio, then analysis their efficiency level by their differences and compare their returns with NASDAQ and S&P 500. Because NASDAQ and S&P 500 could be considered as 2 large portfolios since they all contain many stocks [9-10]. This paper is constructed as follows. Section 2 shows the data, Section 3 shows the methods and related results, Section 4 concludes the paper.

2. Data

In this paper, ten US stocks across different industries are included and considered. There are AAL, AAPL, AMC, CSCO, CVS, GM, JPM, MSFT, SBUX and TSLA. This paper use daily price date for each individual asset and the training data is from 2017 to 2021, and the testing data is from 2021 to 2022. This paper assume there are 252 trading days each year for consistency. And also, All the Returns, covariance matrix and Sharpe Ratios in this process are annualized. All the data and sources are collected from Yahoo Finance. Details are shown in Table 1.

Table 1. Selected assets

Stock	market value / million
AAL (American Airline)	7,800
AAPL (Apple)	2,220,000
AMC (AMC Theatres)	3,600
CSCO (Cisco)	164,400
CVS (CVS Health)	125,200
GM (Gnenral Motors)	46,700
JPM (JPMorgan)	307,400
MSFT (Microsoft)	1,740,000
SBUX (Starbucks)	96,600
TSLA (Tesla)	831,200

The ten selected US stocks that is located in many different industries. All these ten stocks have relatively high market value to ensure that the stock price will not change significantly due to the special behavior of some individual investors or institutions. The figure below is the price changing of 10 selected stocks in selected period

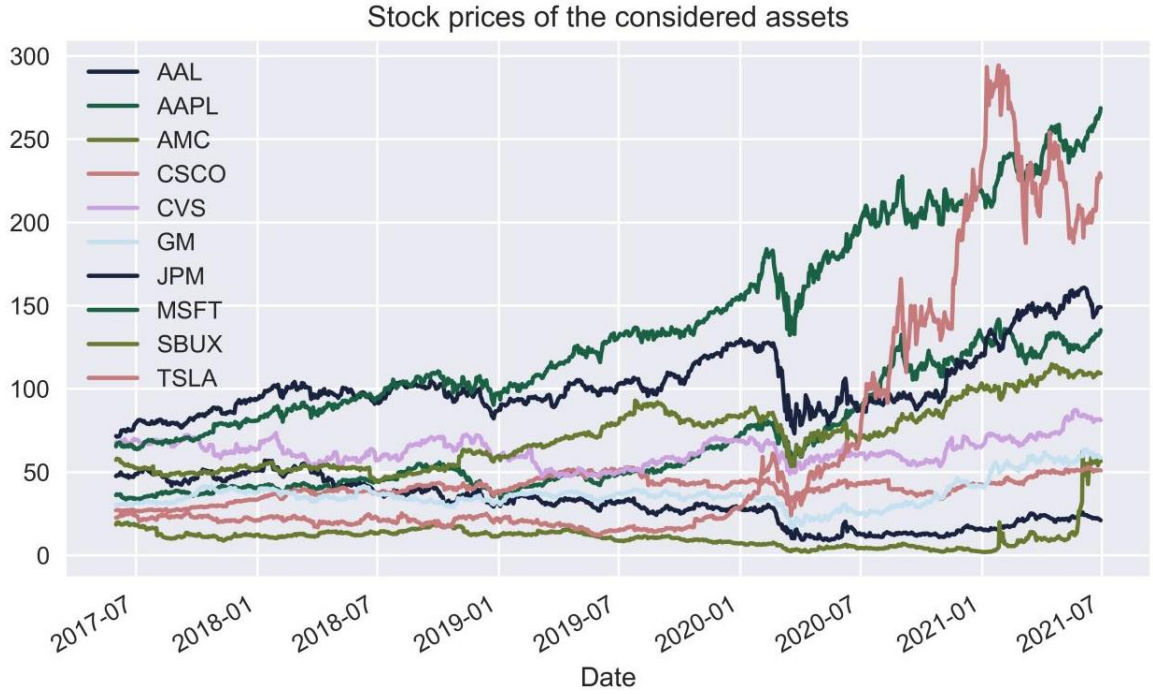


Figure 1. Price trends of the selected assets

3. Method and results

In this article, the method this paper use is that first used the historical average return to be the benchmark. After that, this paper use the following five different strategies with different calculations method to get different portfolios, and this paper can also calculate different returns and volatility for each strategy. The way to determine the best strategy is to compare the final sharpe ratios. The sharpe

ratio the difference bet this paper expected return in the strategy and the risk-free return, the divide by the volatility. So, the higher sharpe ratio always indicate the investment is worthy.

Strategy 1: Using AR (Autoregressive) model to predict expected return

$$x_t = b_0 + b_1 x_{t-1} + \varepsilon_t \tag{1}$$

x is the value of time series, b0 is the intercept with the y-axis, b1 is the slope coefficient, ε is the error term, t is the time period

Strategy 2: Using cross-sectional analysis to predict expected return.

$$x_t = \frac{Z^2 P(1-P)}{d^2} \tag{2}$$

Where n is the sample size, Z is the statistic corresponding to level of confidence, P is expected prevalence.

Strategy 3: Using LASSO Regression based on cross-sec analysis to predict expected return.

$$\sum_{i=1}^{M}(y_i - \hat{y})^2 = \sum_{i=1}^{M}(y_i - \sum_{j=0}^{P}\omega_j \cdot x_{ij})^2 \tag{3}$$

Strategy 4: Using GBDT to predict expected return. In GBDT, this paper used different layer to compare the results

Strategy 5: Using KNN to predict expected return. In KNN, there are to distance measure method this paper used.

$$\text{Cosine}: \ \mathrm{D} = \frac{\sum_i x_1 y_1}{\sqrt{\sum_i x_i^2} \cdot \sqrt{\sum_i y_i^2}} \tag{4}$$

$$\text{Correlation}: \ \mathrm{D} = \frac{N\sum xy - (\sum x)(\sum y)}{\sqrt{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]}} \tag{5}$$

All the strategies this paper used could be used to predict stock's future trend and prices based on its past performance from the data collected from Yahoo finance. And they can be used to calculate the optimal portfolio, asset allocation, sharp ratio, volatility and return of assets under the strategy from the historical data of the selected stocks in this paper. It is important to note that due to the uniqueness and contingency of the selected stocks, the volatility and Sharpe ratio calculated in this paper cannot be applied to other stocks, in other period, and if the strategy is used for subsequent studies on other stocks, it needs to be recalculated with their historical data. Individual expected return are below in Table 2.

Table 2. Expected return

AAL	-0.018492
AAPL	0.376729
AMC	1.314737
CSCO	0.196999
CVS	0.089961
GM	0.239809
JPM	0.230725
MSFT	0.390005
SBUX	0.201022
TSLA	0.768178

The average data above is based on 252 trading days per year. Before everything is the strategy 0, In this paper, used the historical average return to be the benchmark. Then, the maximum sharp ratio portfolio is found. This paper then calculated the maximum Sharpe ratio portfolio; the performance is below in Table 3.

Table 3. Performance and this weight of assets in the maximum sharpe ratio

Return	53.52%
Volatility	33.34%
Sharpe Ratio	160.53%
AAL	0.00%
AAPL	13.56%
AMC	6.14%
CSCO	0.00%
CVS	0.00%
GM	0.00%
JPM	0.00%
MSFT	56.44%
SBUX	0.00%
TSLA	23.86%

The analysis of the data above shows that not every stock in the optimal asset allocation of our selection of substantial stocks is worthy of being purchased, and that most of the stocks in the optimal asset allocation that are heavily held are those of U.S. high-tech companies. In this strategy, the back test result this paper got a Sharp ratio of 0.4064, it could be considered as a fair result, which is not bad. So, this paper considered to use this as a benchmark for all the other strategies.

The Strategy 1 is to use an AR model to predict and forecast the expected return. this paper used the array (-0.48547214, 0.01672497, 2.05001869, 0.24181765, 0.33154023, -0.23417368, 0.27573669, 0.22241723, 0.12830943, 0.821449561) as an input and predicted all the expected returns. this paper got the maximum Sharpe Ratio portfolio; the performance is below in Table 4.

Table 4. Maximum sharpe ratio under AR1 model

Return	66.82%
Volatility	37.85%
Sharpe Ratio	176.55%
AAL	0.00%
AAPL	0.00%
AMC	10.43%
CSCO	19.93%
CVS	0.00%
GM	0.00%
JPM	35.22%
MSFT	56.44%
SBUX	0.00%
TSLA	34.42%

Here this paper can find that the portfolio of stocks this paper got in different strategy could be totally different. The stock this paper hold the most is Microsoft in strategy 0, but this paper hold 0% Microsoft here. Instead, this paper hold a lot CSCO and JP Morgan. The final Sharp Ratio is about -0.03. In this paper, also tried the AR2, which in similar way, which the input array become (-0.14562862, -0.09048613, 0.05146929, 4.33454171, 0.34446619, -0.40864925, 0.27972219, 0.16083995, 0.12773586, 1.002478341). Here is the new Maximum Sharpe Ratio Portfolio Performance in Table 5.

Table 5. Maximum sharpe ratio under AR2 model

Return	153.08%
Volatility	55.86%
Sharpe Ratio	274.06%
AAL	0.00%
AAPL	0.00%
AMC	23.02%
CSCO	7.01%
CVS	0.00%
GM	0.00%
JPM	28.57%
MSFT	0.00%
SBUX	0.00%
TSLA	41.40%

And the sharp ratio is about -0.0834. this paper can clearly find the historical average strategy works better. The next strategy this paper used is the strategy 2, the cross-sectional analysis, here is the expected return that array (-0.59019777, 0.16035285, -1.81363406, -0.68612559, -0.73050416, -0.22554394 -0.68830008. -0.14041986. -0.59026888. 0.22588721). In this way, the maximum Sharpe Ratio and the performance are below in Table 6.

Table 6. Maximum sharpe ratio under cross-sectional analysis

Return	17.19%
Volatility	33.06%
Sharpe Ratio	52.00%
AAL	0.00%
AAPL	82.35%
AMC	0.00%
CSCO	0.00%
CVS	0.00%
GM	0.00%
JPM	0.00%
MSFT	0.00%
SBUX	0.00%
TSLA	17.66%

The final sharp ratio this paper got here is about 0.77, which is the highest so far. Because there are 10 Independent variables, so this paper suspect there is a large possibility of existing of Multicollinearity, so this paper penalized it, having a very large coefficients for the independent variables. So, this paper will use LASSO regression. By using LASSO regression, this paper can define different alpha.

This paper used 3 Alpha as below: Alpha=1e-5; Alpha=1e-6; Alpha=1e-7. In this case, array (-0.58457095, -0.2199157 0.15264845, -1.8062476 -0.68667343, -0.72478105, -0.68804984, -0.14693678, -0.58642155, 0.23082834). Here the maximum Sharpe Ratio portfolio is below in Table 7.

Table 7. Maximum sharpe ratio under LASSO regression

Return	16.88%
Volatility	33.49%
Sharpe Ratio	50.41%
AAL	0.00%
AAPL	79.29%
AMC	0.00%
CSCO	0.00%
CVS	0.00%
GM	0.00%
JPM	0.00%
MSFT	0.00%
SBUX	0.00%
TSLA	20.71%

Here the sharp ratio is also about 0.77 which is very similar to the LASSO reg. Next strategy is the self-implemented GBDT, which this paper can set different layers, this paper did both 2-layers, 3-layers and 4-layers, because they performed best here. Array ([-0.11698811, 0.24472986, -0.17072631, -0.3807882 1.36197483, 0.13666586, -0.2683696, 0.06100249, -0.36481011, 1.0899847]). Maximum Sharpe Ratio portfolio is below in Table 8.

Table 8. Maximum sharpe ratio under self-implemented GBDT

Return	111.18%
Volatility	61.25%
Sharpe Ratio	181.51%
AAL	0.00%
AAPL	0.92%
AMC	10.87%
CSCO	0.00%
CVS	0.00%
GM	0.00%
JPM	0.00%
MSFT	0.00%
SBUX	0.00%
TSLA	88.22%

The Sharpe ratio is about 0.52, which performed better than the benchmark but still lo this paper than the previous 2 strategies. The last strategy this paper did here is the self-implemented KNN. By using the correlation coefficient distance, the array is here (-1.02708373, 0.53757864, -0.37787011, -0.08527204, -0.60851625, 0.29700285, 0.08110772, -0.13764994. -0.21153308, 0.9216698). The maximum Sharpe Ratio portfolio is shown in Table 9.

Table 9. Maximum sharpe ratio under self-implemented KNN

Return	63.96%
Volatility	34.56%
Sharpe Ratio	185.09%
AAL	0.00%
AAPL	73.43%
AMC	0.00%
CSCO	0.00%
CVS	0.00%
GM	0.00%
JPM	0.00%
MSFT	0.00%
SBUX	0.00%
TSLA	26.57%

In this way the sharp ratio is about 0.787. The second is to use the cosine as distance measurement. The new array is here (array (-0.81801568, -0.14430413, 0.9619993 0.08868782, -0.47228829, 0.72401006, 0.15275085, -0.22580774, 0.07469863, 1.34054018). Maximum Sharpe Ratio portfolio performance is here below in Table 10.

Table 10. Maximum sharpe ratio under cosine distance measurement

Return	100.18%
Volatility	32.15%
Sharpe Ratio	311.61%
AAL	0.00%
AAPL	74.89%
AMC	0.00%
CSCO	0.00%
CVS	0.00%
GM	0.00%
JPM	0.00%
MSFT	8.96%
SBUX	0.00%
TSLA	16.15%

The final sharpe ratio here is about 0.744. All the Sharpe Ratio results are listed in the Table 11 below. As can be seen, for almost all the cases, the strategies beat the historical average benchmark model.

Table 11. Summarize of the above results

Strategy	Sharp Ratio	Return	Volatility
Historical Average	0.406	0.5352	0.3334
AR(1)	-0.032	0.6682	0.3785
AR(2)	-0.083	1.53	0.5586
Cross- Sectional	0.772	0.1719	0.3306
LASSO reg	0.779	0.1688	0.3349
GBDT	0.529	1.111	0.6125
KNN(Correlation)	0.787	0.6396	0.3456
KNN(Cosine)	0.744	1.001	0.3215

4. Conclusion

Portfolio construction is always a research interest in the financial area. In this paper, after comparing the results of five different strategies by their returns, all of them successfully beat S&P 500 in the selected period and 3 of them beat NASDAQ, AR, GBDT and KNN. With their sharpe ratio, this paper conclude that self-implemented KNN strategy with correlation coefficient distance is relatively the best method is our judging criteria. but relatively surprisingly is that self-implemented KNN strategy with correlation coefficient distance strategy only selected two stocks, AAPL and TSLA, both are U.S. high-tech companies. This means that despite the fact that in 2020, the entire U.S. stock market is effected by the epidemic, investing in Apple and Tesla from 2017 to 2021 as a whole period still has the highest return with relative less risk. And this paper can clearly see that there is a clear difference in returns and volatility among the different strategies, no two strategies have chosen exactly the same portfolio. This is due to the uncertainty of the stock market. Based on the historical stock movement information, different strategy approaches analyze different situations and strategies of capital flows behind the stock market, which is why different strategies, based on the same historical data information can get completely different portfolio choices. The results of this paper will hopefully help investors in the financial markets. It is also hoped that it can provide technical support for later studies.

References

[1] What are the benefits of having an investment portfolio? URL: https://finance.zacks.com/benefits-having-investment-portfolio-3985.html (Accessed on 2022-9-15).

[2] U.S. inflation reached a new 40-year high in June of 9.1 percent. URL: https://www.politico.com/news/2022/07/13/us-inflation-new-40-year-high-june-00045541 (Accessed on 2022-9-30).

[3] Investopedia. URL: https://www.investopedia.com/articles/07/sharpe_ratio.asp (Accessed on 2022-10-1).

[4] Understanding the Sharpe ratio. URL: https://www.forbes.com/advisor/investing/sharpe-ratio/ (Accessed on 2022-9-30).

[5] Getting your money right: How to diversify an investment portfolio to reduce risk and potential losses. URL: https://www.cnbc.com/select/diversifying-a-portfolio-to-reduce-market-risk-and-potential-losses/ (Accessed on 2022-9-30).

[6] A guide to portfolio optimization strategies. URL: https://smartasset.com/investing/guide-portfolio-optimization-strategies (Accessed on 2022-9-30).

[7] Sparse Group LASSO in python. URL: https://towardsdatascience.com/sparse-group-lasso-in-python-255e379ab892 (Accessed on 2022-9-30).

[8] Volatility and the Sharpe ratio: The COVID-19 ERA's equity market surprises. URL: https://www.refinitiv.com/perspectives/market-insights/the-volatility-and-the-sharpe-ratio-equity-market-surprises-in-the-covid-19-era/ (Accessed on 2022-9-30).

[9] NASDAQ-100 portfolio. URL: https://portfolioslab.com/portfolio/nasdaq-100 (Accessed on 2022-9-30).

[10] S&P 500 portfolio. URL: https://portfolioslab.com/portfolio/sp-500 (Accessed on 2022-9-30).