Jiawei Lu

Research Summary

Title: Comparing performances of Nested Clustered and Convex Portfolio Optimization Solution machine learning portfolio optimization algorithms with two traditional methods.

1, summary

The new machine learning portfolio optimization methods are much different from traditional models, Mean-Variance[4][5][6][7] and Black-Litterman optimization[4][5][6][7], because they are far from the optimal computational model of securities market information data. Nested Clustered Optimization (NCO)[1][2][3][6] and Convex Optimization Solution (CVO) [1][2][3][6] are two representative models of new machine learning portfolio optimization computation algorithms.

Nested clustering optimization is an algorithm that estimates the optimal weight assignment to maximize the Sharpe ratio or minimize the variance of a portfolio. The convex optimization solution is a convex optimization result when using the true covariance matrix and the true mean vector of the portfolio to solve the problem of computing the optimal weight assignment. The objective can be the maximum Sharpe ratio or the minimum variance of the portfolio.

A Monte Carlo method that estimates the allocation error produced by various optimization methods for a specific set of input variables. The result is a precise determination of which method is most robust for a given situation. Thus, MCOS does not always rely on a particular method, but allows the user to randomly apply the most appropriate optimization method in a given environment.

Mean-variance model[4][5][6][7] is a technique in which investors decide which financial instruments to invest in based on the amount of risk they are willing to accept (risk tolerance). Ideally, investors expect to earn higher returns when investing in riskier assets. In measuring the level of risk, investors consider the potential variance (i.e., the volatility of the return generated by the asset) versus the expected return on that asset. Meanvariance analysis essentially examines the average variance of an investment's expected returns.

The Black-Litterman (BL) model [4][5][6][7] is an analytical tool for portfolio managers to optimize asset allocation within the risk tolerance and market views of investors. Global investors, such as pension funds and insurance companies, need to decide how to allocate their investments across different asset classes and countries. the BL model helps them to do this by generating expected returns for hypothetical portfolios.

In this paper, I present two traditional optimization methods, Mean-Variance optimization and the Black-Litterman model, optimizing 92 stocks in the U.S market and compare these models with the new methods of machine learning optimization, NCO, CVO and MCOS by calculating the Return, Volatility and Sharpe Ratio of each portfolio.

First, this thesis use SP 100 INDEX Historical Data with 92 components without undefined or unrepresentable values, splitting a approximately 3300-day training dataset and a 180-day testing dataset, and all test results are based on the results of the test dataset. This paper split two testing period 180-days, before and after the COVID happening, the first training period will be from June 16, 2010 to July 4, 2019, and the testing period will be from July 5, 2019 to January 1 2020 to test the general situation in U.S. stock market.The second training period will be from June 16 2010 to December 18,

2019. The testing period will be from December 19, 2019 to June 16, 2020 to test the performance from the beginning of COVID affect period.

Second, I define the risk-free rate using the yields of United States 52 Week Treasury Bill.The risk-free rate being used as a parameter for the Mean-Variance model and BlackLitterman model. And then I introduce S&P 500 INDEX as market price also being used as a parameter for the Mean-Variance model and Black-Litterman model. I set the parameters of Black-Litterman to be the same as the implied expected market return in testing dataset in order to achieve the best performance of the model.

Third, I calculate the Return, Volatility and Sharpe Ratio of Mean-Variance model , Black-Litterman model, NCO and CVO from both training dataset and testing dataset.

In the result, this paper compare the portfolios performance before the beginning of COVID affect periods.(Period 1). In this case the first training period will be from June 16, 2010 to July 4, 2019, and the testing period will be from July 5, 2019 to January 1 2020 to test the general situation in U.S. stock market.(Period 1). The goal of all models is to maximize the sharpe ratio, except for the Black-litterman model, which is adapted to the returns of test dataset. figure 1 demonstrates that they all have very close sharpe ratios. Covex Optimisation method is able to achieve the highest return and sharpe ratio with 132.79% and 335.07%, and Mean-Variance model either with 131.78% and 331.63%. This indicates that the Mean-Variance model does not lose out to the Black-litterman model in terms of return and Sharpe index. NCO, a machine learning model equipped with MCOS, has increased returns. But in the case of COV is decreasing.

In test dataset, the MeanVariance model gives the best results with 67.12%return and 230.67% sharpe index. it has the second highest return with the lowest volatility and the highest sharpe ratio. it has exactly double the volatility compared to the COV model with similar returns. This suggests that the COV model, while gaining stability in data learning, may deliberately select riskier stocks because of the lack of market intelligence like risk free rate in training. Finally, this paper compare the portfolios performance in the beginning of COVID affect periods.(Period 2) In this case the training period will be from June 16, 2010 to December 18, 2019. The testing period is from December 19, 2019 to June, 16, 2020. In this case, Mean-Variance model still brings the most superior results. It also has the highest sharpe ratio with 401.2%, about twice as high as the other models and the lowest volatility with 56.44%. Same as the result above, it indicates that the Mean-Variance model is also able to maximize the sharpe ratio while selecting less risky commodities according to market conditions despite the uncertainty of COVID, although its learning performance is vulnerable to market movements.

This paper shows that through this research, Mean-Variance optimization is still the best performing portfolio optimization model if the risk free rate of the market is accurately defined. Moreover, This result is independent from the effect of whether presence or absence of COVID. Mean-Variance optimization is able to select lower risk stocks than the machine learning models, Nested Clustered, Convex Portfolio Optimization Solution and Monte Carlo Optimization Selection. Of course, if more market intelligence could be defined, such as future stock price trends, the Black-Litterman model would be a better choice for asset managers.

However, if the market information is missing, or to maximize the exclusion of data noise, or to pursue optimal stability, Nested Clustered, Convex Portfolio Optimization Solution and Monte Carlo Optimization Selection will be the best choices.

2, reference list

[1] de Prado, L. Advances in financial machine learning. Oxford University Press (Feb. 2018), 221–245.

[2] de Prado, M. L. A robust estimator of the efficient frontier. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3469961, October 15, 2016.

[3] de Prado Machine, L. Machine learning for asset managers (elements in quantitative finance). Cambridge University Press (Apr. 2020), 92–104.

[4] Komatsu, T. Optimal investment strategies: Theory and practice of portfolio technology. Asakura Shoten (Dec. 2018), 50–67.

[5] Martin, R. Pyportfolioopt. https://pyportfolioopt.readthedocs.io/en/ latest/, 2021.

[6] Oleksandr, P. Machine learning financial laboratory (mlfinlab). https://mlfinlab.readthedocs.io/en/latest/index.html, 2021.

[7] Ruenberger, D. G. Investment science. Oxford University Press; 2nd edition (June 2013), 40–101.