Bachelor's Thesis
Academic Year 2020

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

Faculty of Law, Department of Political Science
Keio University

Jiawei Lu

A Bachelor's Thesis
submitted to Faculty of Economics, Keio University

Jiawei Lu

Thesis Committee:

Keio University Faculty of Economics Professor Teruo Nakatsuma (Supervisor)

Abstract of Bachelor's Thesis of Academic Year 2020

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

Category: Science / Engineering

Summary

This thesis investigates the optimization models in stock portfolios. This thesis finds that the traditional Mean-Variance model and Black-Litterman model can achieve higher return, Sharpe ratio, and lower risk comparing to four new machine learning models, Nested Clustered Optimization(NCO)[1][2][3][6], Convex Portfolio Optimization Solution (CVO)[1][2][3][6], NCO with Monte Carlo Optimization Selection (MCOS)[1][2][3][6], and CVO with MCOS[1][2][3][6]. With more market intelligence, the Mean-Variance model[6] is able to select lower risk stocks and maximize Sharpe ratio than the above four models.

This paper present two traditional Optimisation methods, Mean-Variance Optimisation[4][5][7] and the Black-Litterman model[4][5][6], optimizing 92 stocks in the U.S market and compare these models with four new methods of machine learning Optimisation, Nested Clustered Optimization(NCO), Convex Optimization(CVO), NCO with Monte Carlo Optimization Selection (MCOS) and CVO with MCOS by calculating the Return, Volatility and Sharpe Ratio of each portfolio.

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.

This paper 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 Black-Litterman model. And then this paper introduce SP 500 INDEX as market price also being used as a parameter for the Mean-Variance model and Black-Litterman model, setting 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.

Finally, this paper calculate the Return, Volatility and Sharpe Ratio of Mean-Variance model , Black-Litterman model, NCO and CVO from both training dataset and both testing dataset before and after the COVID period.

Keywords:

Portfolio Optimization, Mean-Variance model, Black-Litterman model, Nested Clustered Optimization, Convex Optimization Solution, Monte Carlo Optimization Selection

Faculty of Law, Department of Political Science, Keio University

Jiawei Lu

Contents

Chapter 1

Introduction

1.1. Background

The new machine learning portfolio Optimisation approach differs significantly from the traditional models Mean-Variance and Black-Litterman Optimisation in that they do not rely on risk-free rates and market indices to calculate the weights of individual assets. In the traditional models Mean-variance Optimisation 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. Mean-variance analysis essentially examines the average variance of an investment's expected returns.[2][6] The Black-Litterman (BL) model 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 contrast to two of these models, Nested Clustered Optimization (NCO)[6] and Convex Optimization Solution (CVO)[6] are two representative models of new machine learning portfolio Optimisation computation algorithms.[6] Nested clustering Optimisation is an algorithm that estimates the optimal weight assignment to maximize the Sharpe ratio or minimize the variance of a portfolio. The convex Optimisation solution is a convex Optimisation 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.[6] A Monte Carlo method that estimates the allocation error produced by various Optimisation 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 Optimisation method in a given environment.[6]

Chapter 2

Overview of Mean-Variance and Black-Litterman Optimisation

2.1. Mean-Variance Optimisation

Traditionally, portfolio optimization is nothing more than a simple mathematical optimization problem, where your objective is to achieve optimal portfolio allocation bounded by some constraints. It can be mathematically expressed as follows:[4][5]

$$max \qquad f(\mathbf{x})$$

$$s.t. \qquad g(\mathbf{x}) \leq 0 \qquad h(\mathbf{x}) = 0$$

where $x \in R^n$ and h(x), g(x) represent convex functions correlating to the equality and inequality constraints respectively. Based on the mean-variance framework first developed by Harry Markowitz, a portfolio optimization problem can be formulated as follows,

$$min \qquad w^T \sum w$$

$$s.t. \qquad \sum_{i=1}^{n} w_i = 1 \qquad \mu^T w = \mu_t$$

where w refers to the set of weights for the portfolio assets, $\sum$ is the covariance matrix of the assets, μ is the expected asset returns and μ_t represents the target portfolio return of the investor. Note that this represents a very basic (and a specific) use-case of portfolio allocation where the investor wants to minimse the portfolio risk for a given target return. As the needs of an investor increase, the complexity of the problem also changes with different objective functions and multitude of constraints governing the optimal set of weights.

2.1.1 Maximum Sharpe Ratio

For this solution, the objective is (as the name suggests) to maximise the Sharpe Ratio of your portfolio.[5][6]

$$maximise \qquad \frac{\mu^T w - R_f}{(w^T \sum w)^{1/2}}$$

$$s.t. \qquad \sum_{j=1}^{n} w_j = 1 w_j \geq 0, j = 1, .., N$$

A major problem with the above formulation is that the objective function is not convex and this presents a problem for cvxpy which only accepts convex optimization problems. As a result, the problem can be transformed into an equivalent one, but with a convex quadratic objective function.

$$minimise \qquad y^T \sum y$$

$$s.t. \qquad (\mu^T w - R_f)^T y = 1$$

$$\sum_{j=1}^{N} y_j = \kappa, \kappa \geq 0, w_j = \frac{y_j}{\kappa}, j = 1, .., N$$

where y refer to the set of unscaled weights, κ is the scaling factor and the other symbols refer to their usual meanings.

2.2. Black-Litterman Optimisation

The Black-Litterman (BL) Model[4][5] is one of the most widely used portfolio allocation models in the industry. The mean-variance framework developed by Harry Markowitz requires investors to specify the mean and covariance of the portfolio assets. Before BL, people used historical sample estimates of expected returns and covariance matrix for this purpose but it is commonly known these samples estimates of returns and risk are not robust to market changes, highly sensitive and hence prove to be bad estimators of true market parameters. Instead of relying on these inefficient sample estimates, BL uses Bayesian theory to combine historical market data (prior) with additional investor views (likelihood) to generate a posterior estimate of expected returns and covariance in sync with an investor's intuitions.

2.2.1 Overview of the Algorithm

Calculating the Prior: Implied Excess Equilibrium Returns

The first step is to start by formulating the prior distribution of returns for our Bayesian BL model. Because there is a complex mapping between the returns, and no natural starting point for assumptions on expected returns, users of the standard portfolio optimizers often find their specification of expected returns produces output portfolio weights which do not seem to make sense. It all starts with the Capital Asset Pricing Model (CAPM) market portfolio. The typical quadratic utility for portfolio optimisation is:

$$U = \pi^T w_m - \frac{\delta}{2} w_m^T \Sigma w_m$$

where π is the vector of implied excess equilibrium weights, Σ is the covariance matrix of asset returns, w_m is the set of weights of the market portfolio and δ is the risk aversion coefficient. Under the CAPM assumptions and no investment constraints, an investor will start by holding the market portfolio. Differentiating the above utility equation w.r.t. w_m and equating to 0, can get the following,

$$\nabla U = \pi - \delta \Sigma w_m = 0$$

$$w_m = (\delta \Sigma)^{-1} \pi$$

In this case, this paper already know the market weights and instead want to find π which is done via reverse optimisation trick that gives us the following:

$$\pi = \delta \Sigma w_m$$

w_m is typically calculated by taking the market-caps of assets in our portfolio and normalising them to sum to 1. The empirical covariance of returns is normally used as an estimate of Σ. Finding the value of δ is a tricky part and can either be assumed based on prior knowledge or calculated using any index portfolio like SP500. This gives us the final prior distribution of market returns:

$$r \sim N(\pi, \tau\Sigma)$$

It is a multivariate normal distribution parameterised by the mean - π - and the covariance - $\tau\Sigma$ - where τ is the constant of proportionality with values around 0.005-0.01.

Calculating the Likelihood: Market Views The second step is to take into account an investor's views on the market. It is a very important step which highlights the ingenuity of the BL model. In many scenarios, an asset manager or investor has his/her own views about assets in a portfolio. For example, one might have an opinion that some assets in the portfolio will yield a higher expected return while some may underperform. There are three important inputs of interest here - P, Q and Ω. is a vector of means that contains specific values denoting the portfolio views and P is a matrix denoting which assets in our portfolio are involved in the views. It is often called a *picking matrix*. Ω matrix is a very important parameter because it quantifies the variance/errors in the investor's views. Note that the investor's beliefs are based on the prior distribution mentioned before and combining this information gives us the investor's distribution of the returns:

$$r \sim N(Q, \Omega)$$

Generating the Posterior Market Distribution

The BL model generates the final posterior distribution from the above described prior and likelihood distributions,

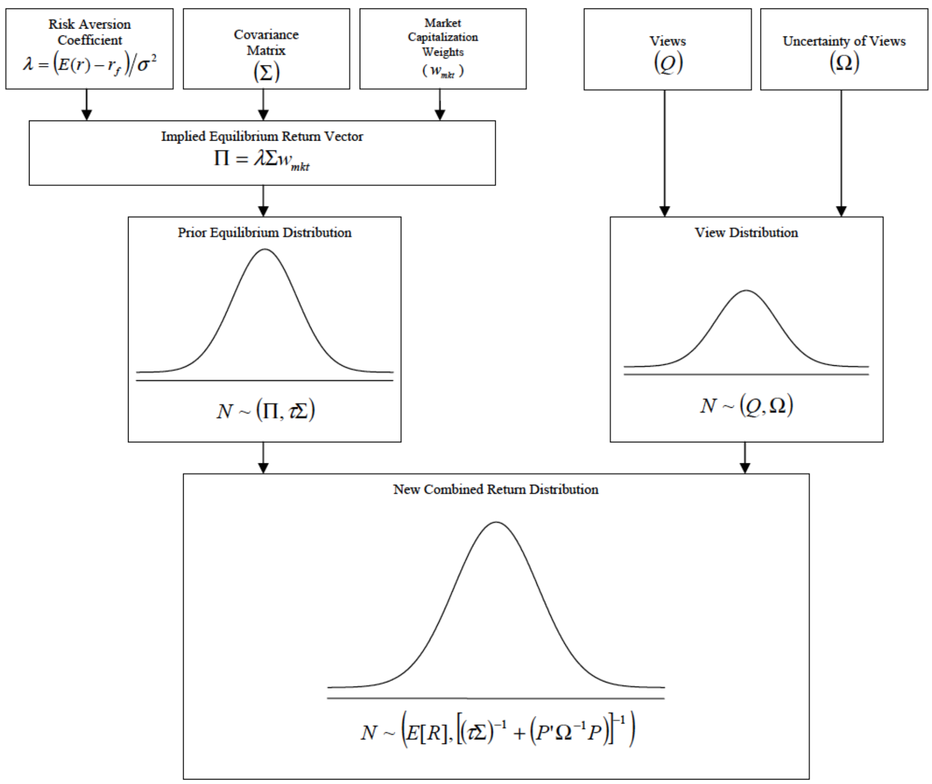
$$r \sim N(\mu_p, \Sigma_p)$$

where,

$$\mu_p = [(\tau\Sigma)^{-1} + P^T\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\pi + P^T\Omega^{-1}Q] \qquad \Sigma_p = [(\tau\Sigma)^{-1} + P^T\Omega^{-1}P]^{-1}$$

where μ_p and Σ_p are the mean and covariance of the new posterior returns distribution. For a detailed derivation of the above results, please refer to the links in the Underlying Literature section. The following image (Figure 1)by Thomas Idzorek summarises the entire Black-Litterman model in a good way,

Figure 1: Black-Litterman model



The parameters of Black-Litterman were set to be the same as the implied expected market return in the testing dataset in order to achieve and test the best Return and Sharpe Ratio of the dataset as a reference to other optimization methods.

Chapter 3

Nested Clustered Optimization (NCO) and Convex Optimization Solution(CVO)

3.1. Nested Clustered Optimization (NCO)

The NCO class includes functions related to:[1][6]

1, Weight allocation using the Nested Clustered Optimization (NCO) algorithm.

2, Weight allocation using the Convex Optimization Solution (CVO).

3, Multiple simulations for the NCO and CVO algorithms using Monte Carlo Optimization Selection (MCOS) algorithm.

4, Sample data generation to use in the above functions.

3.1.1 The Algorithm

The Nested Clustered Optimization algorithm estimates optimal weight allocation to either maximize the Sharpe ratio or minimize the variance of a portfolio. The steps of the NCO algorithm are:

1, Get the covariance matrix of the outcomes as an input (and the vector of means if the target is to maximize the Sharpe ratio).

2, Transform the covariance matrix to the correlation matrix and calculate the distance matrix based on it.

3, Cluster the covariance matrix into subsets of highly-correlated variables.

4, Compute the optimal weights allocation (Convex Optimization Solution) for every cluster.

5, Reduce the original covariance matrix to a reduced one - where each cluster is represented by a single variable.

6, Compute the optimal weights allocation (Convex Optimization Solution) for the reduced covariance matrix.

7, Compute the final allocations as a dot-product of the allocations between the clusters and inside the clusters.

3.2. Convex Optimization Solution(CVO)

The Convex Optimization Solution is the result of convex optimization when solving a problem of calculating the optimal weight allocation using the true covariance matrix and the true vector of means for a portfolio. The goal can be either the maximum Sharpe ratio or minimum variance of a portfolio. If the problem of portfolio optimization is:[2][3][6]

$$min_w \frac{1}{2} w'Vw$$

$$s.t. : w'a = 1$$

Where is the covariance matrix of elements in a portfolio, is the vector of weights that minimizes the variance or maximizes the Sharpe ratio, is an optimal solution that defines the goal of optimization. Then the Convex Optimization Solution to the problem is:

$$w^* = \frac{V^{-1}a}{a'V^{-1}a}$$

3.3. Monte Carlo Optimization Selection (MCOS)

The Monte Carlo Optimization Selection algorithm calculates the NCO allocations and a simple optimal allocation for multiple simulated pairs of mean vector and the covariance matrix to determine the most robust method for weight allocations for a given pair of means vector and a covariance vector. The steps of the MCOS algorithm are:[2][3][6]

1, Get the covariance matrix and the means vector of the outcomes as an input (along with the simulation parameters to use).

2, Drawing the empirical covariance matrix and the empirical means vector based on the true ones.

3, If the bandwidth of the KDE kernel parameter is given, the empirical covariance matrix is de-noised.

4, Based on the Minimum Variance Portfolio parameter, either the minimum variance or the maximum Sharpe ratio is targeted in weights allocation.

5, CVO is applied to the empirical data to obtain the weights allocation.

6, NCO is applied to the empirical data to obtain the weights allocation.

7, Based on the original covariance matrix and the means vector a true optimal allocation is calculated.

8, For each weights estimation in a method, a standard deviation between the true weights and the obtained weights is calculated.

9, The error associated with each method is calculated as the mean of the standard deviation across all estimations for the method.

3.3.1 Sample Data Generating

This method allows creating a random vector of means and a random covariance matrix that has the characteristics of securities. The elements are divided into clusters. The elements in clusters have a given level of correlation. The correlation between the clusters is set at another level. This structure is created in order to test the NCO and MCOS algorithms.[6]

Chapter 4

Data description and Portfolio Performance Measure

4.1. Data description

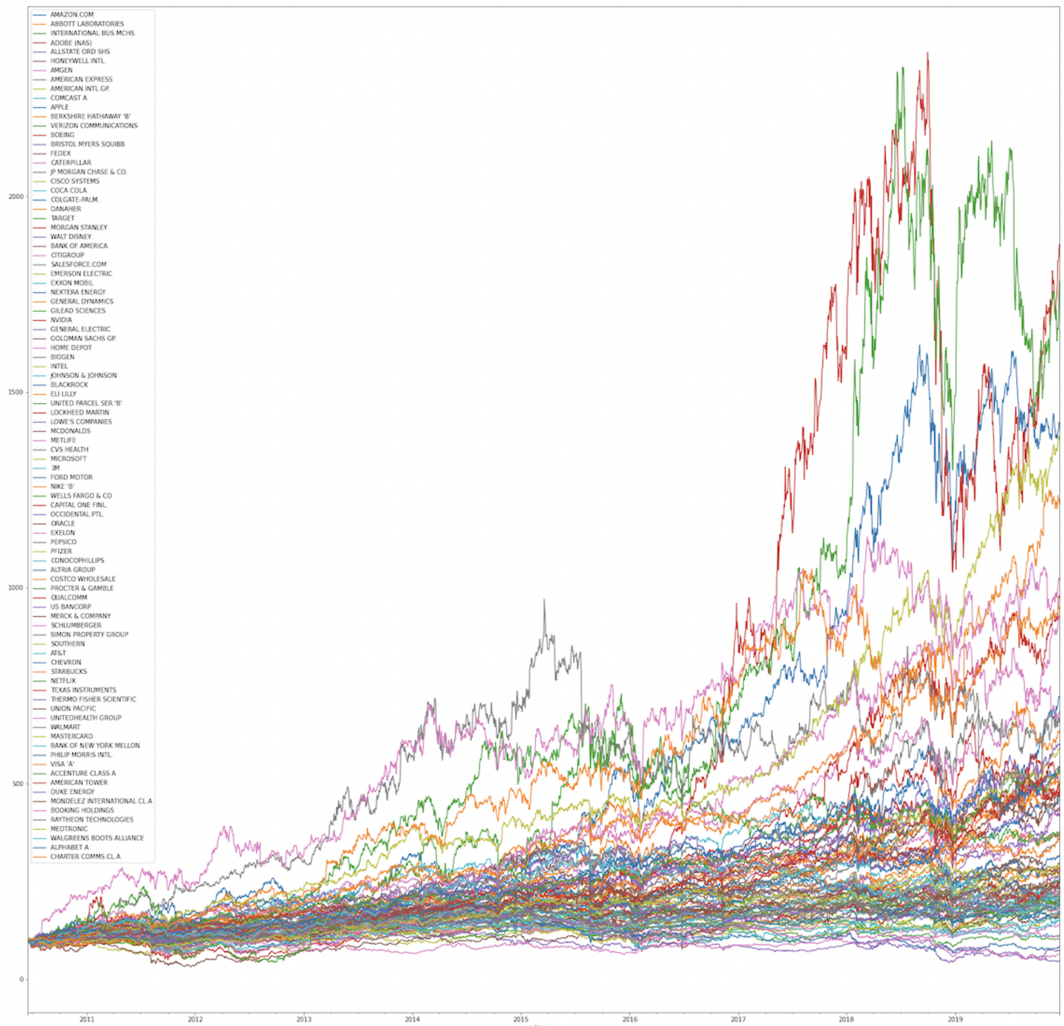
All stock information used in this paper come from the SP 100 Index constituents, a sub-set of the SP 500®, is designed to measure the performance of large-cap companies in the United States. Figure 2 shows the the distribution of the log returns of the stocks selected. Constituents of the SP 100 are selected for sector balance and represent about 67% of the market capitalization of the SP 500 and almost 54% of the market capitalization of the U.S. equity markets. The stocks in the SP 100 tend to be the largest and most established companies in the SP 500.

The data usage will be closing prices for the 92 assets without missing values in the period of SP 100 at a daily resolution.

This research splits two training and testing period 180-days, before and after the COVID happening to test the Portfolio Performance in different optimization methods. 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 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.(Period 2)

The U.S. 52 Week Treasury Bill and SP 500 index was used to calculate the risk-free interest rate and market price for Mean-Variance model and Black-Litterman model.[4][5][7]

Figure 2:The distribution of the log returns



4.2. Portfolio Performance Measure

4.2.1 The Annualized mean historical returns

The mean historical returns function calculating a mean annual return for every element in a dataframe of prices.[6] The calculation is done in the following way:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

AnnualizedMeanReturn $= \sum_{t=0}^{T} R_{t\frac{}{T*N}}$ Where R_t is the return for t -th observation, and P_t is the price for t -th observation, T is the total number of observations, N is an average number of observations in a year.

4.2.2 Variance

This measure can be used to compare portfolios based on estimations of the volatility of returns. The Variance of a portfolio is calculated as follows:[6]

$$\sigma^2 = w^T \sum w$$

where w is the vector of weights for instruments in a portfolio, and $\sum$ is a covariance matrix of assets in a portfolio. Result σ^2 is a scalar.

4.2.3 Annualized Sharpe Ratio

Calculates Annualized Sharpe Ratio for pd.Series of normal or log returns. A usual metric of returns in relation to risk. Also takes into account number of return entries per year and risk-free rate.[6] Risk-free rate should be given for the same period the returns are given. For example, if the input returns are observed in 3 months, the risk-free rate given should be the 3-month risk-free rate. Calculated as:

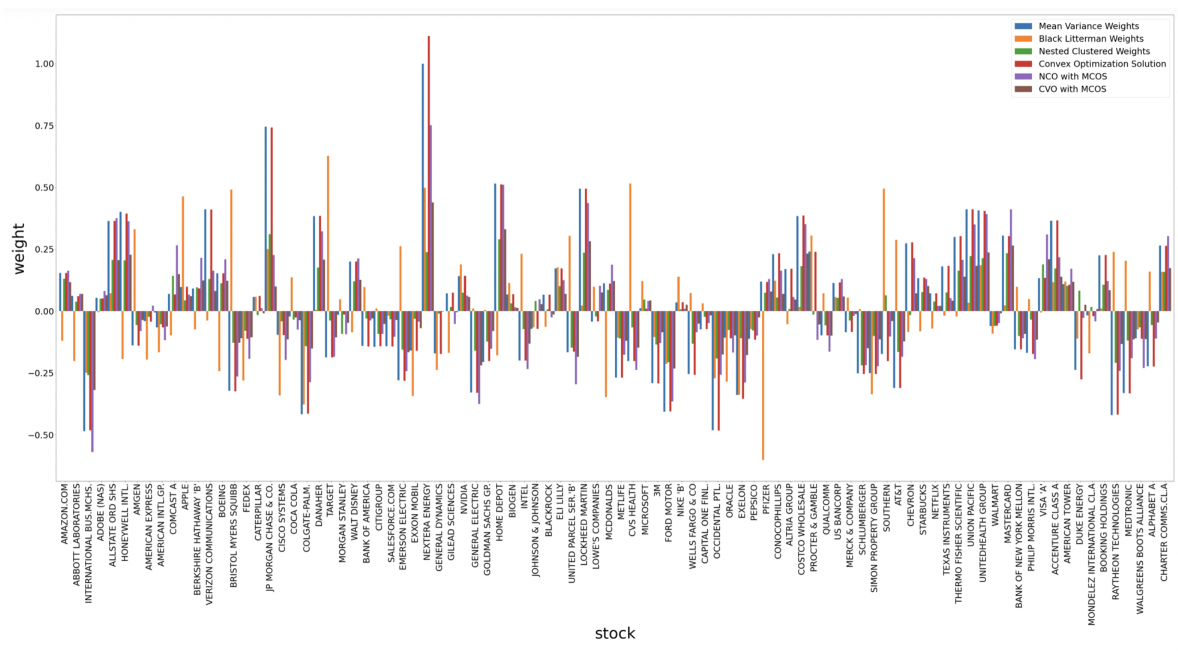
$$SharpeRatio = \frac{E[Returns] - RiskFreeRate}{\sqrt{V[Returns]}} * \sqrt{n}$$

Chapter 5

Result

Figure 3 shows all methods of the portfolio optimisation weight comparison in 92 stocks. Each optimisation method allows Short selling.

Figure 3:Portfolio weight comparison(Period 1)



First, 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.

Table 1:Comparison of portfolios performance in Training Dataset(Period 1)

	Mean Variance simulated value	Black Litterman simulated value	Nested Clustered simulated value	Convex Optimization simulated value	NCO with MCOS simulated value	COV with MCOS simulated value
Return	1.317799	0.799154	0.789909	1.327950	1.237470	0.837969
Volatility	0.391099	0.341644	0.253938	0.393990	0.397157	0.270803
Portfolio Sharpe Ratio	3.318338	2.280601	3.079948	3.350733	3.096195	3.065601

Table 2 shows that test dataset, the Mean-Variance 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.

Table 2:Comparison of portfolios performance in Testing Dataset(Period 1)

	Mean Variance simulated value	Black Litterman simulated value	Nested Clustered simulated value	Convex Optimization simulated value	NCO with MCOS simulated value	COV with MCOS simulated value
Return	0.671215	4.934892	0.367944	0.700696	0.310424	0.295841
Volatility	0.287613	0.433405	0.359631	0.586424	0.602040	0.398368
Portfolio Sharpe Ratio	2.306650	11.368343	1.001448	1.181573	0.502677	0.723068

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.

Table 3:Comparison of portfolios performance in the beginning of COVID affect periods in Testing Dataset(Period 2)

	Mean Variance simulated value	Nested Clustered simulated value	Convex Optimization simulated value	NCO with MCOS simulated value	COV with MCOS simulated value
Return	2.272084	1.408660	2.233404	1.996221	1.344646
Volatility	0.564385	0.692469	1.118165	0.939512	0.667084
Portfolio Sharpe Ratio	4.011962	2.023003	1.990413	2.116447	2.004024

Chapter 6

Conclusion

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.

References

[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.