EMPIRICAL EVALUATION OF PORTFOLIO OPTIMIZATION STRATEGIES

Luiz Fernando Medeiros *
Maasalo-Medeiros Oy
luiz.medeiros@maasalo-medeiros.com

Esa Ollila
Department of Signal Processing and Acoustics
Aalto University, Finland
esa.ollila@aalto.fi

August 1, 2020

ABSTRACT

This work evaluates the performance of two state-of-the-art covariance matrix estimators proposed in the domain of portfolio optimization, Ledoit-Wolf (shrinkage) estimator (LWE) [1] and Rotationally Invariant Estimators (RIE) [2]. They are evaluated on their ability to maintain the desired goals of different portfolio allocation strategies considered and results are compared to using conventional sample covariance matrix (SCM). The investment strategies considered are the Global Minimum Variance (GMVP), where the goal is produce a portfolio with the minimum risk [3] and Most-Diversified Portfolio (MDP), where the goal is to produce a portfolio of with a maximum diversification ratio [4]. We also propose a variant of MDP called the Modified Most-Diversified Portfolio (MMDP), where diversification is maintained while attempting to control an empirical risk. The general conclusion is that the LWE outperforms SCE and RIE under most conditions, while RIE outperforms in a few specific scenarios, while the proposed modified MDP produces the highest returns in general.

1 Introduction

Covariance matrix estimation is an important topic in machine learning and statistics. The covariance matrix is often computed from a set of signals within a time window. The signals can be different medical information from patients over time or returns of different assets over a specific time window as in this paper. Covariance matrices are also needed in Principal Component Analysis (PCA), Exploratory Data Analysis (EDA), or in portfolio optimization strategies [3], where one often wants to minimize the standard deviation (i.e., volatility) of asset returns.

We consider a scenario where the time window for covariance matrix estimation is limited due to non-stationarity. The input data matrix $(\mathbf{r}_1 \ \cdots \mathbf{r}_T)^\top$ is composed of T observations on N variables, where the number of observations T is only slightly larger than the number of variables N. In portfolio optimization, the vector $\mathbf{r}_t = (r_{t1}, \dots, r_{tN})^\top$ contains the net returns of N different assets at time t and the sequence of measurements, $\{\mathbf{r}_t\}_{t=1}^T$, is thus a financial times series over some fixed time interval or time window. Typically, the frequency of t are (trading) days as in this paper. In this work, we explore different long-only allocation strategies defined by portfolio weight vector

$$\mathbf{w} = (w_1, \dots, w_N)^\top, \quad w_k \in [0,1], \quad \sum_{i=1}^N w_i = 1$$

determining the portion of capital that is allocated to each specific asset i, $i = 1, \dots, N$.

In the portfolio optimization problem, one assumes that $\mathbf{r}_t$-s are i.i.d. N-variate random vectors with symmetric positive definite $N \times N$ covariance matrix

$$\boldsymbol{\Sigma} = \mathbb{E}[(\mathbf{r}_t - \mathbb{E}[\mathbf{r}_t])(\mathbf{r}_t - \mathbb{E}[\mathbf{r}_t])^\top], \tag{1}$$

*Part of the work of L. F. Medeiros was conducted while attending Aalto University where $\mathbb{E}[\cdot]$ denotes the expectation. The covariance matrix is then estimated from the available data and a conventional estimate is the *sample covariance matrix (SCM)*,

$$\mathbf{S} = \frac{1}{T}\sum_{t=1}^{T}(\mathbf{r}_t - \bar{\mathbf{r}})(\mathbf{r}_t - \bar{\mathbf{r}})^\top \tag{2}$$

where $\bar{\mathbf{r}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{r}_t$ is the sample mean of the returns. In this paper, the performance of the SCM is compared in porfolio optimization framework against the state-of-the-art covariance matrix estimators introduced in the financial literature: the Ledoit-Wolf estimator (LWE) [1] and the Rotationally Invariant Estimator (RIE) [2]. When this is not the case, as in portfolio optimization problem, the usage of SCM can lead to very unfortunate portfolio choices [5].

We investigate how well these different covariance matrix estimators fulfill the goals of different portfolio optimization strategies. That is, we aim to answer questions such as how well the LWE is helping the Global Minimum Variance Portfolio (GMVP) strategy to minimize the risk. More specifically, we backtest the strategies using these estimators on historical stock data and evaluate the results.

2 High-dimensional covariance matrix estimators

The LWE [1] and RIE [2] estimators were developed for high-dimensional data in low sample support scenarios (HDLSS), where the number of observations T is of similar magnitude as the data dimensionality N, i.e., $T = O(N)$ [2]. This is a situation that emerges in portfolio optimization, where T, the number of historical returns, is rarely larger than N, the number of assets in the portfolio. Due to $T = O(N)$ settings, the SCM $\mathbf{S}$ will not produce an estimate that is necessarily close to the true covariance matrix $\mathbf{\Sigma}$ [1].

2.1 Shrinkage SCM á la Ledoit and Wolf

In the approach developed by Ledoit and Wolf [1], one aims to find an optimal shrinkage SCM,

$$\mathbf{S}(\alpha,\beta) = \beta\mathbf{S} + \alpha\mathbf{I}, \quad \alpha,\beta \geq 0$$

that attains the minimum mean-squared-error (MMSE), $\mathbb{E}\big[\|\mathbf{S}(\alpha,\beta) - \mathbf{\Sigma}\|_{\mathrm{F}}^2\big]$, where $\|\cdot\|_{\mathrm{F}}$ denotes (normalized) Frobenius matrix norm ($\|\mathbf{A}\|_{\mathrm{F}}^2 = \mathrm{tr}\big(\mathbf{A}^\top\mathbf{A}\big)/N$ for any matrix $\mathbf{A}$). The authors in [1] show that the optimal MMSE shrinkage constants α and β can be estimated from the available data, and the resulting shrinkage SCM estimator build based on these, which we refer to as the LW-estimator (LWE), is defined as

$$\mathbf{S}_{\mathrm{LW}} = \hat{\beta}\mathbf{S} + \hat{\alpha}\mathbf{I}, \tag{3}$$

where

$$\hat{\alpha} = (1-\hat{\beta})\frac{\mathrm{tr}\big(\mathbf{S}\big)}{N} \text{ and } \hat{\beta} = \max\left(\frac{\frac{1}{T^2}\sum_{t=1}^{T}\|\mathbf{S} - \mathbf{r}_t\mathbf{r}_t\|_F^2}{\|\mathbf{S} - (\mathrm{tr}\big(\mathbf{S}\big)/N)\mathbf{I}\|_F^2}\right). \tag{4}$$

Note that LWE estimator in (3) does not need to assume that $T > N$, and thus it can be applied also in the cases that the number of stocks in the portfolio is larger than the number of historical returns.

2.2 Rotationally invariant estimator (RIE)

Assume that $T > N$ and that return vectors $\mathbf{r}_t$ are standardized, that is, they are centered to have mean zero and unit standard deviation, i.e., $\bar{\mathbf{r}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{r}_t = \mathbf{0}$ and $\frac{1}{T}\sum_{t=1}^{T}r_{ti}^2 = 1$ for $i = 1,\ldots,N$. RIE provides an estimate of the correlation matrix $\mathbf{C}$ of the net returns, defined as

$$\mathbf{C} = \mathrm{diag}(\mathbf{\Sigma})^{-1/2}\mathbf{\Sigma}\mathrm{diag}(\mathbf{\Sigma})^{-1/2}.$$

Let the eigenvalue decomposition (EVD) of the sample correlation matrix $\mathbf{E} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{r}_t\mathbf{r}_t^\top$ be denoted by

$$\mathbf{E} = \sum_{k=1}^{N}\lambda_k\mathbf{u}_k\mathbf{u}_k^\top \tag{5}$$

where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N \geq 0$ are the eigenvalues of $\mathbf{E}$, and $\mathbf{u}_1,\ldots,\mathbf{u}_N$ are the eigenvectors. The modified eigenvalues of RIE [2,5] are defined as

$$\hat{\xi}_k = \frac{\lambda_k}{|1-q+qz_ks_k|^2} \times \max(1,\Gamma_k), \quad \text{for } k = 1,\ldots,N, \tag{6}$$

where $q = N/T$, $z_k = \lambda_k - \frac{\jmath}{\sqrt{N}}$, and

$$s_k = \frac{1}{N} \sum_{\substack{j=1 \\ j \neq k}}^{N} \frac{1}{\lambda_k - \lambda_j - \frac{\jmath}{\sqrt{N}}}. \tag{7}$$

Above $\jmath$ denotes the imaginary unit ($\jmath^2 = -1$). The reason for this eigenvalue modifications stems from Marcenko-Pastur law [6] which states that the eigenvalue spectrum of $\mathbf{E}$ is broader than the spectrum of $\mathbf{C}$ when $T = O(N)$. Commonly, the small eigenvalues of $\mathbf{E}$ are smaller than those of $\mathbf{C}$ while large eigenvalues are overestimating the true eigenvalues. The constant Γ_k is a debiasing term, defined as

$$\Gamma_k = \sigma^2 \frac{|1 - q + q z_k g_{mp}(z_k)|^2}{\lambda_k}, \tag{8}$$

where $g_{mp}(\cdot)$ is the Stieltjes transform of the Marcenko-Pastur density,

$$g_{mp}(z) = \frac{z + \sigma^2(q-1) - \sqrt{z - \lambda_N}\sqrt{z - \lambda_+}}{2qz\sigma^2} \tag{9}$$

with $\lambda_+ = \lambda_N \left(\frac{1+\sqrt{q}}{1-\sqrt{q}}\right)^2$ and $\sigma^2 = \frac{\lambda_N}{(1-\sqrt{q})^2}$. The RIE estimator of the correlation matrix $\mathbf{C}$ is then defined as

$$\mathbf{\Xi} = \sum_{k=1}^{N} \hat{\xi}_k \mathbf{u}_k \mathbf{u}_k^\top. \tag{10}$$

Thus RIE makes nonlinear modifcation to the eigenvalues of the SCM but keeps the eigenvectors intact.

3 Portfolio optimization strategies

Next we define the portfolio optimization strategies studied in this work: The *Global Minimum Variance Portfolio (GMVP)* strategy [3], the Most Diversified Portfolio (MDP) strategy [4] and then our modification of it, called the Modified Maximum Diversification Portfolio strategy (MMDP). Note that all strategies depend upon the covariance matrix $\mathbf{\Sigma}$ of the returns which is unknown in practice and needs to be estimated from the past historical returns. This is where the different covariance matrix estimators comes into play.

GMVP aims at minimizing the risk of the portfolio. The portfolio net return is $R_P = \mathbf{w}^\top \mathbf{r}_t$ and the *risk* of the portfolio is its variance

$$\mathsf{risk}(\mathbf{w}) = \mathsf{var}(R_P) = \mathbf{w}^\top \mathbf{\Sigma} \mathbf{w}.$$

A GMVP that adheres to long-only strategy is obtained by solving the following optimization problem:

$$\begin{aligned} \underset{\mathbf{w}}{\text{minimize}} \quad & \mathbf{w}^\top \mathbf{\Sigma} \mathbf{w} \\ \text{subject to} \quad & \mathbf{1}^\top \mathbf{w} = 1 \\ & \mathbf{w} \geq \mathbf{0}. \end{aligned} \tag{11}$$

It should be mentioned that if one relaxes the problem by ignoring the long-only constraint ($\mathbf{w} \geq \mathbf{0}$), then the optimization problem has a closed-form solution:

$$\mathbf{w} = \frac{\mathbf{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \mathbf{\Sigma} \mathbf{1}}. \tag{12}$$

This is useful due to the fact that if in the unconstrained case the optimal coefficients are nonnegative, then the solution in (12) will also be the optimal solution for the constrained long-only case (11) as well. However, it should noted that (11) is a strictly convex quadratic programming problem (QP) which can be solved efficiently by any standard QP solver.

In MDP, one constructs a measure that quantifies how diversified the portfolio is. By maximizing this measure, one achieves a portfolio that assures maximum diversification. Let

$$\sigma_i = \sqrt{\Sigma_{ii}} = \sqrt{\mathsf{var}(r_{ti})}$$

denote the volatility of the ith asset, $i = 1, \ldots, N$. The *diversification ratio* $\mathrm{DR}(\mathbf{w})$ for a portfolio is defined as

$$\mathrm{DR}(\mathbf{w}) = \frac{\mathbf{w}^\top \boldsymbol{\sigma}}{\sqrt{\mathbf{w}^\top \mathbf{\Sigma} \mathbf{w}}}, \tag{13}$$

where $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N)^\top$ denotes the vector of volatilities of the assets. DR is thus the ratio of the portfolio's weighted average volatility to its overall volatility. It is intuitive that portfolios with concentrated weights and/or highly correlated holdings would be poorly diversified, and hence be characterized by relatively low DR-s. The MPD portfolio solves the following problem

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{maximize}} & \mathrm{DR}(\mathbf{w}) \\ \text{subject to} & \mathbf{1}^\top \mathbf{w} = 1 \\ & \mathbf{w} \geq \mathbf{0}. \end{array} \tag{14}$$

It should be mentioned that again ignoring the long-only constraint ($\mathbf{w} \geq \mathbf{0}$), it is possible to obtain a closed-form solution to (14) which is given by

$$\mathbf{w} = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\sigma}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\sigma}}. \tag{15}$$

Next we propose a modification of MDP. The idea is similar to MDP, i.e., to maximize the diversification ratio, but the risk is constrained to be less than a threshold c. Our modified MPD portfolio solves the following problem

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{maximize}} & \mathbf{w}^\top \boldsymbol{\sigma} \\ \text{subject to} & \mathbf{1}^\top \mathbf{w} = 1 \\ & \mathbf{w} \geq \mathbf{0} \\ & \mathbf{w}^\top \mathbf{S} \mathbf{w} \leq c. \end{array} \tag{16}$$

For this experiment, the above variable c was chosen to be:

$$c = 1.5 \cdot (\mathsf{risk}(\mathbf{w}_{MDP}) - \mathsf{risk}(\mathbf{w}_{GMVP})) + \mathsf{risk}(\mathbf{w}_{GMVP}). \tag{17}$$

In short, the constant c controls how much the user is willing to deviate from the optimum $\mathrm{DR}(\mathbf{w})$, with the constraint of risk, and an intuition that the higher returns may not necessarily lie in the optimum $\mathrm{DR}(\mathbf{w})$. Finally, if a solution to Equation 16 exists, it is used, otherwise, $\mathbf{w}_{MDP}$.

4 Backtesting results on real-world stock data

The data was retrieved from Yahoo Finance. The data includes a range of assets pertaining to American and European exchanges. Some of the assets included in this dataset range from Amazon and Microsoft to Daimler and LVMH Group. The assets were generally selected on a random basis. The total number of daily net returns T and assets N is $T = 4409$ and $N = 207$. Finally, the time period for the dataset used was 01/2001 - 12/2019.

In our experiment, different lengths of look-back windows are used as well as number of assets in the portfolio. A look-back window is the number of data points considered at every re-balancing point. Re-balancing is when we compute the portfolio weights (and thus the covariance matrix estimate). Thus we either buy or sell since the weights $\mathbf{w}$ change at this point. Rebalancing occurs once per month throughout the study, i.e., after 22 trading days. At given rebalancing point t, we use the past historical daily returns from the past k months. Thus we use observations from $\{t - 22 \cdot k, \ldots, t\}$ to estimate the covariance matrix. The number of assets in the portfolio is also allowed to vary. For example, for a given experiment we may evaluate how a specific trading strategy using LWE will perform with only 10 randomly selected assets. This means that we pick ten assets at random from the whole universe of 207 assets. Thus every month, a covariance matrix estimate is computed based on the data of look-back period, and the portfolio weights are updated according to the newly achieved weights.

The overall performance of the estimators is presented next. Figure 1 demonstrates the results for percentage of total capital return of the GMVP strategy. The graphs demonstrate results from different covariance estimation methods and number of assets (either using $10, 50, 100, 150$, or 200 assets) included in the portfolio. Figure 1 illustrate that there is a positive trend where the percentage of returns at time t_{i+1} is greater than the initial invested capital at time t_0. Moreover, all covariance matrix estimation methods when used in the strategy are able to maintain positive returns.

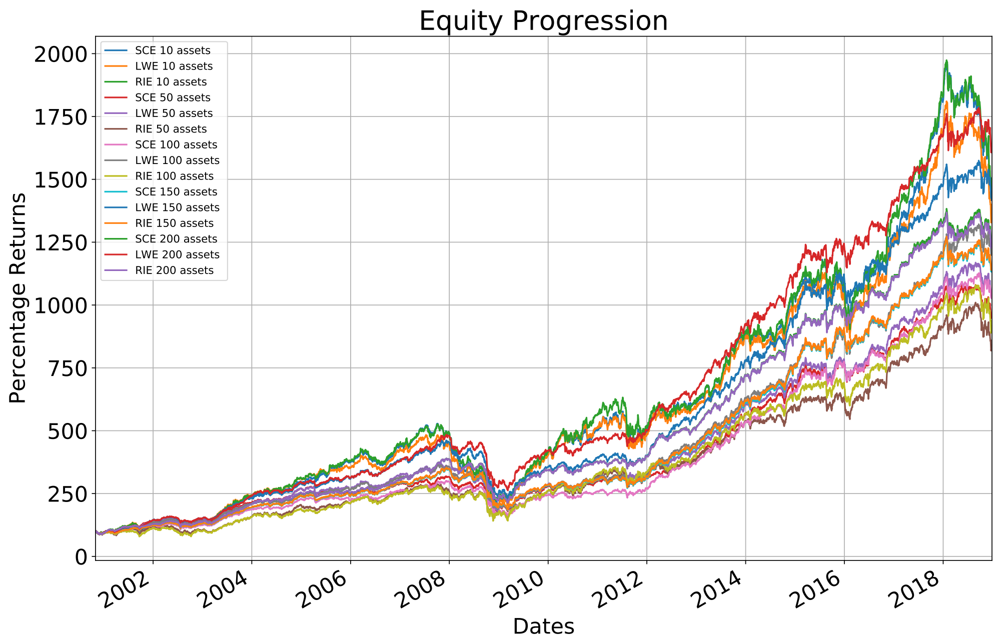

Figure 1: Percentage returns based on initially invested capital for the GMVP strategy on a 6 months look-back window.

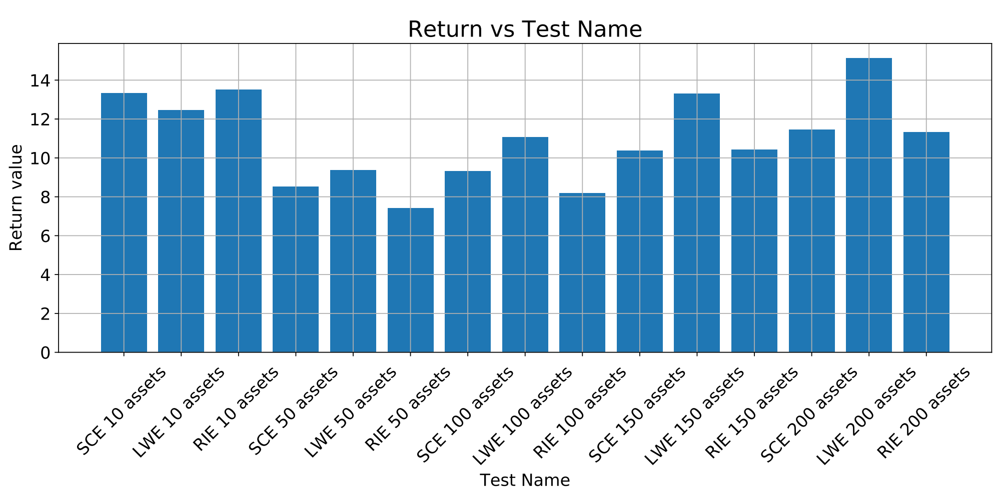

Figure 2: Returns of a GMVP strategy and a 6 months look-back window..

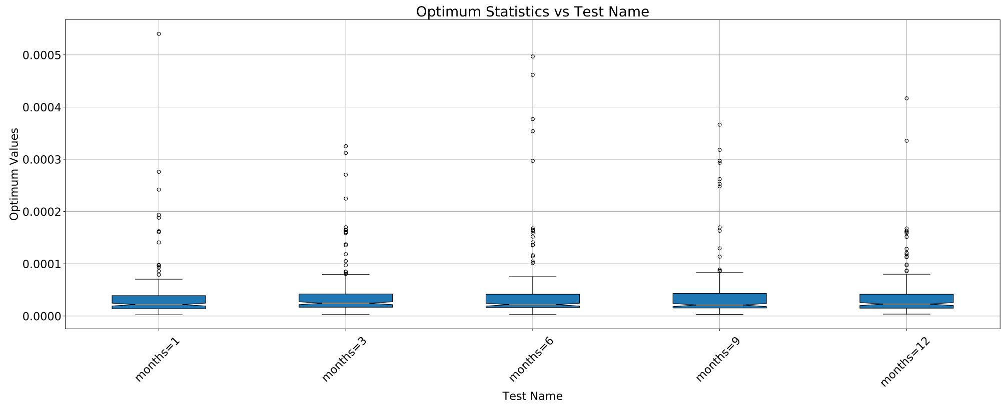

Figure 3: The boxplot shows how risk varies w.r.t. look-back months considered using the GMVP strategy. Results are for LWE in the case of 150 asset portfolio.

Figure 2 shows in more detail the added value between the different covariance method estimators. The returns are calculated by considering the capital achieved at the end of the investment period, with the initial capital as reference. The largest return was achieved by the LWE, with 200 assets. Although the largest return was achieved with a large number of assets, the general trend for GMVP does not provide a clear answer as to whether or not a larger number of assets provide a larger return when attempting to minimize risk.

Next we investigate which look-back period produced the smallest empirical risk. Since GMVP aims to minimize the risk, one can measure the performance of different covariance matrix estimators in terms of how well it helped the GMVP strategy to achieve this goal. Figure 3 illustrate how risk varies in terms of look-back period in the LWE and 150 assets. It is possible to note from Figure 3 that in the case of three (3) months look-back, we have smaller empirical risk, while most risk values achieved are within the same 25-75 percentile margin.

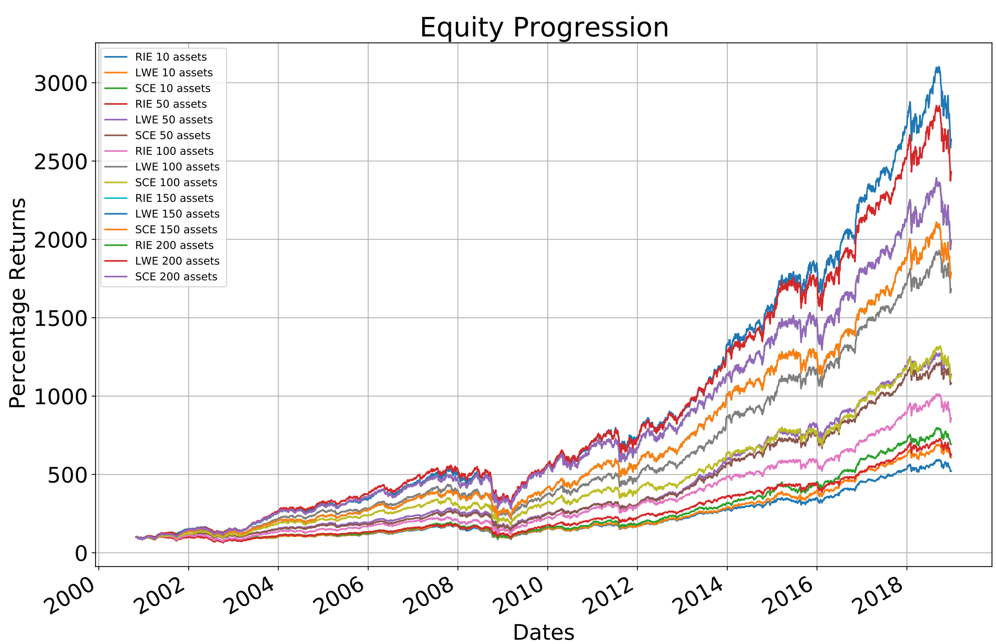

Figure 4: The Equity Progression plot for the MDP strategy, with a 6 months look-back period.

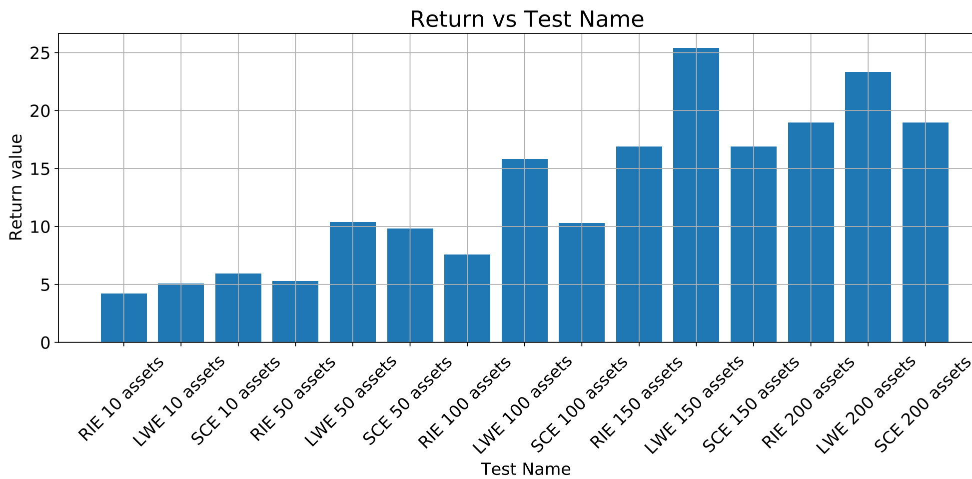

Figure 5: Detailed returns for MDP strategy, given 6 months look-back window.

Next, Figure 4 shows equity progression curves in the case of MDP strategy. Figure 5 shows that for 150 assets LWE provides superior returns. It also indicates a comparable performance by RIE and 200 assets. Additionally, it is possible to note that MDP provides a clear trend indicating that larger number of assets provide larger returns when the goal is to maximize the diversification ratio.

Now, let us consider how the diversification ratio changed over different look-back periods. Figure 6 shows that most of the diversification ratios achieved reach a value of 6. However, a 1 month look-back window demonstrates generally higher diversification ratios, where some values surpass 10. This is an interesting insight because it demonstrates that the diversification ratio does change significantly with different look-back windows.

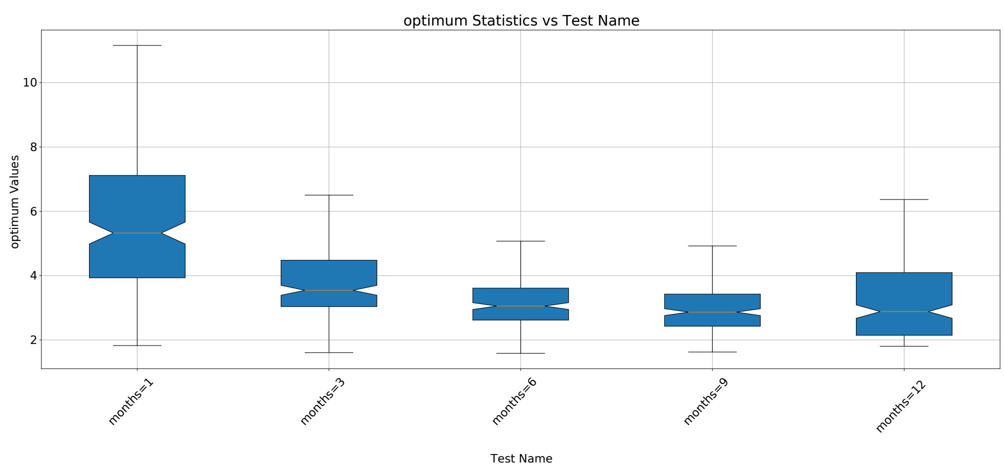

Figure 6: The optimum statistics in this case implies diversification ratio. This boxplot shows how the diversification ratio varies when we change the number of look-back months to consider with the MDP strategy, LWE, and a 150 asset portfolio.

Next we consider the proposed modified MDP (MMDP) strategy. The equity progression graph presented on Figure 7 demonstrates the results achieved for this strategy where c was as defined in Equation 17. MMDP produces positive returns, while maintaining a rising equity value trend. Figure 8 provides further details into the overall returns:

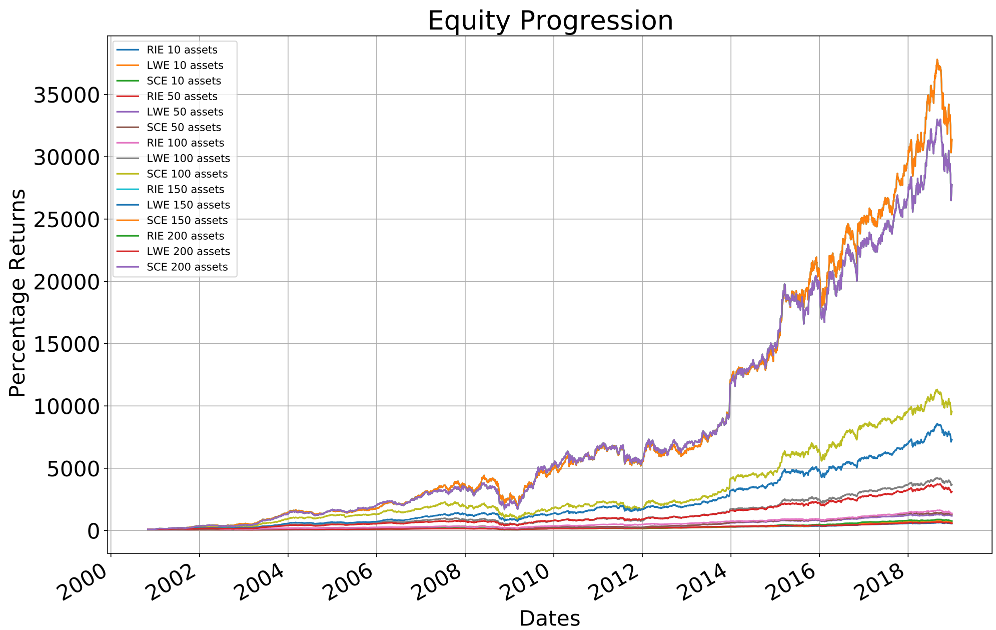

Figure 7: The equity progression plot for the MMDP strategy when a 6 months look-back period is used to estimate the covariance matrix.

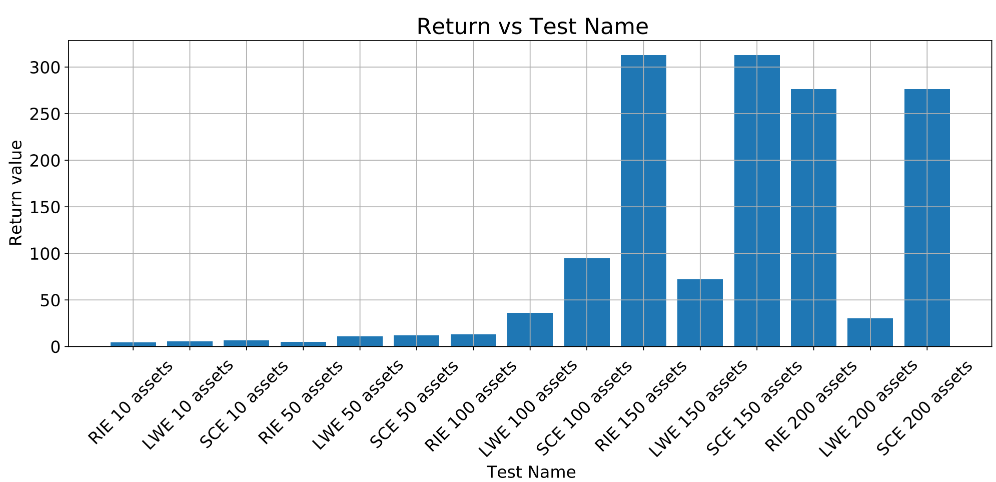

Figure 8: Detailed returns for MMDP strategy, given 6 months look-back window.

It's possible to note that SCM and RIE provided the largest returns for a 150 and 100 assets portfolio respectively, while maintaining consistency in all other asset numbers. Interestingly, LWE did not provide the best results in this class of optimization strategy. It's also possible to note that asset numbers below 100 provide significantly lower returns. Finally, not only does MMDP provide higher returns than MDP, it also manages to bring about returns that are more than ten (10) times higher than MDP in some test cases.

Now, although this strategy does not aim to reach the maximum $\mathrm{DR}(\mathbf{w})$ or the minimum risk, it's still interesting to analyze the achieved diversification ratio for different look-back periods.

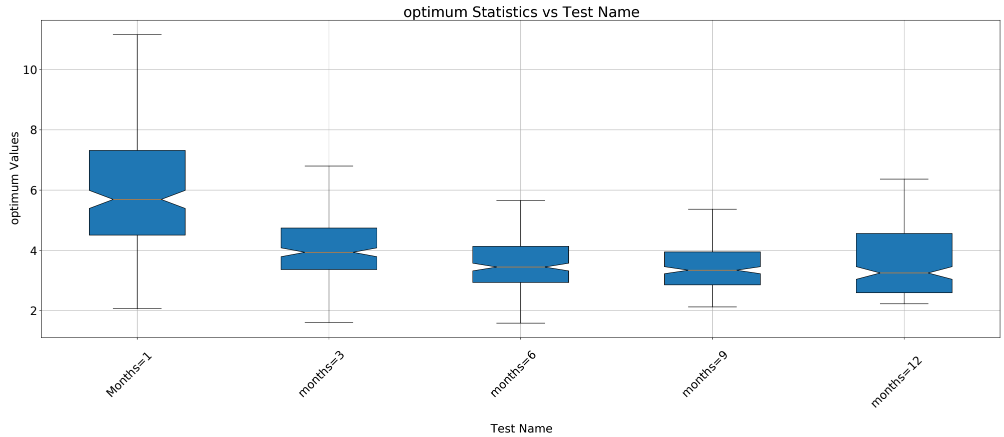

Figure 9: The optimum statistics in this case implies diversification ratio. This boxplot shows how the diversification ratio varies when we change the number of look-back months to consider with the MMDP strategy, LWE, and a 150 asset portfolio.

Figure 9 shows that the highest diversification ratio achieved was at a look-back period of 1 month, while maxima and minima is compared to MDP. This may be due to the fact that if the strategy is not able to find a suitable optima for the problem in Equation 16, it resorts to a MDP strategy. This means $\mathbf{w}$ is computed according to MDP, and the diversification $D(\mathbf{w})$ is computed accordingly. Finally, we can note a similar situation to the MDP strategy, where changes in look-back windows between 1 to 12 months does seem to significantly affect the diversification ratio range.

5 Conclusion and Future Work

The initial aim of the project was to apply and evaluate the performance of SCE, LWE, and RIE on GMVP and MDP. To this end, the conclusion was that LWE produces the better performance for most cases, and is therefore the recommended estimator to be used when producing returns based covariance matrix estimation.

On a different note, while learning and testing the different estimators and portfolio optimization strategies we stumbled upon MMDP. A different strategy based on the original MDP work that managed to produce superior returns for most tested cases. This leads the way to the aim for future work: further evaluation of MMDP, including under the different covariance matrix estimators, while formalizing the results for a possible publication.

References

[1] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of Multivariate Analysis*, vol. 88, no. 2, pp. 365–411, 2004.

[2] J. Bun and J. P. Bouchaud, "Cleaning correlation matrices," *Risk.net*, no. April, 2016.

[3] A. Meucci, *Risk and Asset Allocation*. Springer Finance, 2005.

[4] Y. Choueifaty and Y. Coignard, "Toward maximum diversification," *Journal of Portfolio Management*, vol. 35, no. 1, pp. 40–51, 2008.

[5] J. Bun, J. P. Bouchaud, and M. Potters, "Cleaning large correlation matrices: Tools from Random Matrix Theory," *Physics Reports*, vol. 666, pp. 1–109, 2017.

[6] V. A. Marčenko and L. A. Pastur, "Distribution of Eigenvalues for Some Sets of Random Matrices," *Mathematics of the USSR-Sbornik*, vol. 1, no. 4, pp. 457–483, 1967.