Empirical Evaluation of Different Portfolio Optimization Strategies Using State-of-The-Art Covariance Matrix Estimation Methods

Luiz Fernando F. Medeiros

March 1, 2020

Abstract

This work evaluates the performance of different covariance matrix estimators proposed in the domain of portfolio optimization. Three different estimators, namely sample based (SCE), Ledoit-Wolf (LWE) [1], and Rotationally Invariant Estimators (RIE) [2] are evaluated on their ability to estimate covariance matrix, as well as how they are able to maintain the desired goals of different portfolio allocation strategies. The strategies considered were as follows: Global Minimum Variance (GMVP), where the goal is produce a portfolio with the minimum risk [3]. Most-Diversified Portfolio (MDP), whose goal is to produce a portfolio of with a maximum diversification ratio [4]. And finally, a variant of MDP called the Modified Most-Diversified Portfolio (MMDP), a new strategy proposed by the authors, where diversification is achieved, while attempting to minimize risk. The general conclusion is that LWE outperforms SCE and RIE under most conditions, while RIE outperforms in a few specific scenarios, while MMDP produces the highest returns in general.

Contents

1 Introduction

Covariance Matrix Estimation is an important topic in a number of fields. The covariance matrix may be computed from a time window containing a time varying signals, data composed of different medical information from patient/s over time, or from the returns of different assets over a specific time window. Additionally, covariance matrices must be computed in order to apply Principal Component Analysis (PCA), perform Exploratory Data Analysis (EDA), and in portfolio optimization problems such as when one wants to minimize the variance of a set of asset returns.

The general question surrounding covariance matrices is as follows: is it possible to produce an estimate $\mathbf{S}$ that is as close as possible to the true covariance matrix $\mathbf{\Sigma}$? The scenario is generally composed of a situation where there is limited knowledge or access to data. Moreover, the input data matrix, say $\mathbf{X}$, is composed of T observations and N variables. The goal becomes difficult as the number of observations T approaches the number of variables N.

In this work, the focus is Portfolio Optimization (PO) [3]. This means that the information taken into account is matrix $\mathbf{X}$, composed of columns containing the returns $\mathbf{r}_i$, whereby each vector $\mathbf{r}_i$ (for $i \in [1, ..., T]$) represents an asset, and each row (or item in $\mathbf{r}_i$) represents an observed daily return (vector length $|\mathbf{r}_i| = T$). Finally, the general goal is to find an effective allocation strategy defined by weight vector $\mathbf{w}$, where $\mathbf{w} = (w_1, ..., w_N)^\top, w_k \in [0, 1]$, and $\sum_{k=1}^{N}\{w_k\} = 1$, indicates to the investor or portfolio manager what portion of the capital should be allocated to a specific asset [1].

As such, the work produced here evaluates three different covariance matrix estimators: Sample Covariance Estimator (SCE), Ledoit-Wolf (LWE) [1], and Rotationally Invariant Estimators (RIEs) [2]. The evaluation is done mainly on the PO performance. Namely, how well are these different estimators fulfilling the goals of different portfolio optimization strategies. For example, how well is the LWE helping GMVP minimize the risk. Or how is it aiding in finding the MDP? More specifically, we backtest the strategies which contain these estimators(further details will be provided in section 2) and evaluate the results.

[1]This asset allocation definition indicates that investors can only produce long-only portfolios $w_k > 0$, and may not borrow money $\sum_{k=1}^{N}\{w_k\} = 1$ [3]

2 High-Dimensional Covariance Matrix Estimators

In this section we present the covaraiance matrix estimators, along with the perceived logic behind each of the estimators. SCE, probably the most commonly used covariance matrix estimator. Moreover, LWE [1] and RIEs [2] are also highlighted and explored herein due to their relevance and state-of-the-art in high-dimensional setting where the number of observations is smaller or not much larger than the data dimensionality. See Definition 2.0.2. However, as mentioned by the authors of both works, there exists other estimators that have been developed over time. The scope of the current work focuses on the three estimators mentioned above and applied to the problem definitions described below:

Definition 2.0.1. *A data matrix* $\mathbf{X} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_T \end{bmatrix}$ *where* $\mathbf{X} \in \mathbb{R}^{T \times N}$, *$N$ is the number of assets, T is number of observations to be used for portfolio optimization,* $\mathbf{r}_t = [r_1, ..., r_N]$, *is the vector containing observation t for each of the N assets.*

Definition 2.0.2. *Considering Definition 2.0.1, T and N are conditioned as follows: $T \approx N$. So, $q = \frac{N}{T}$ is not sufficiently small, so $T = O(N)$ [2].*

In addition to the above definitions, we can also define the true covariance between assets $\mathbf{v}$ and $\mathbf{z}$ as follows:

$$\mathrm{Cov}(\mathbf{v}, \mathbf{z}) = \mathbb{E}\left[\mathbf{v}^\top \mathbf{z}\right] - \mathbb{E}[\mathbf{v}]^\top \mathbb{E}[\mathbf{z}], \tag{1}$$

and the true covariance matrix of data matrix $\mathbf{X}$ as:

$$\mathbf{\Sigma} = \mathbb{E}\left[\mathbf{X}^\top \mathbf{X}\right] - \mathbb{E}[\mathbf{X}]^\top \mathbb{E}[\mathbf{X}], \tag{2}$$

where $\mathbf{\Sigma} \in \mathbb{R}^{N \times N}$, symmetric, and positive semi-definite. Furthermore, $\mathbb{E}[\cdot]$ is the expectation of argument as defined in [5]. Finally, $\mathbf{z}$ and $\mathbf{v}$ are entries of $\mathbf{X}$, analogous to entries $\mathbf{r}_i$ and $\mathbf{r}_j \in \mathbf{X}$.

The conditions imposed by Definition 2.0.2 indicates that the classical SCE will not produce an estimate that is necessarily close to the true covariance matrix $\mathbf{\Sigma}$, due to insufficient sample size [1]. In PO, such situations emerge often when a strategy which uses N assets,say $N = 50$, and T observations, say $T = 60$ trading-days (considering the three previous months) of asset returns.

Definition 2.0.3. *With Definition 2.0.2, the problem becomes to find an estimator that produces an estimate $\mathbf{\Sigma}^\star$ of the true covariance matrix $\mathbf{\Sigma}$, with the smallest expected mean-squared-error. More specifically, we want to:*

$$\text{minimize } \mathbb{E}\Big\{\|\mathbf{\Sigma}^\star - \mathbf{\Sigma}\|_F^2\Big\}, \tag{3}$$

where $\|\cdot\|_F$ *is the Frobenius norm* [2]

Therefore, in order to investigate the problem described in Definition 2.0.3, the performance of SCE, along with LWE and RIE will be explored in the following sections.

2.1 Sample Covariance Estimator

A sample covariance matrix $\mathbf{S}$ is defined as follows:

$$\mathbf{S} = \frac{1}{T} \cdot \mathbf{X}^\top \mathbf{X} - \bar{\mathbf{r}} \cdot \bar{\mathbf{r}}^\top, \tag{4}$$

where $\bar{\mathbf{r}} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{r}_t, \bar{\mathbf{r}} \in \mathbb{R}^N$, and $\mathbf{r}_t$ is the row vector at row t in $\mathbf{X}$.

Equation 4 describes SCE [5]. As discussed in [1] and [2], and justified by the results presented in section 3, SCE is a good estimator as q approaches 0. In such cases, one has enough data such that the method is able to produce a good estimate of the underlying covariance matrix.

However, when this is not the case, SCE becomes unreliable thereby leading users to produce erroneous conclusions. In fact, Bun et al [6] points this out as one reason for unfortunate portfolio choices.

2.2 Ledoit-Wolf Large-Dimensional Covariance Estimator

LWE builds on the SCE, and the general idea being to minimize the difference between a true covariance matrix $\mathbf{\Sigma}$ and an estimate $\mathbf{\Sigma}^\star$. In fact, it aims at solving the optimization problem mentioned in Equation 3. Namely, minimizing the mean squared error (MSE):

$$\begin{aligned} &\underset{\zeta_1,\zeta_2}{\text{minimize}}\ \mathbb{E}\Big\{\|\mathbf{\Sigma}^\star - \mathbf{\Sigma}\|_F^2\Big\} \\ &\text{s.t. } \mathbf{\Sigma}^\star = \zeta_1 \mathbf{I} + \zeta_2 \mathbf{S}, \end{aligned} \tag{5}$$

[2]The Frobenius norm is defined as: $\|\mathbf{A}\|_F = \sqrt{\text{tr}\Big(\mathbf{A}^\top\mathbf{A}\Big)/N}$, for $\mathbf{A} \in \mathbb{R}^{T\times N}$

where $\mathbf{I}$ is the identity matrix, $\mathbf{S}$ is the sample covariance estimate, and ζ_1 and ζ_2 are constants.

Now, the aim in Equation 5 is to find an optimal linear combination of the SCM and the identity matrix that is able to provide an optimum estimation for the covariance matrix. Let us define constants α, β, and δ verifying:

$$\begin{aligned} &\alpha^2 + \beta^2 = \delta^2, \\ &\text{where: } \alpha^2 = ||\mathbf{\Sigma} - \mu\mathbf{I}||_F^2, \\ &\qquad \mu = \mathrm{tr}\Big(\mathbf{\Sigma}\Big)/N, \\ &\qquad \beta^2 = \mathbb{E}[||\mathbf{S} - \mathbf{\Sigma}||_F^2], \\ &\qquad \delta^2 = \mathbb{E}[||\mathbf{S} - \mu\mathbf{I}||_F^2]. \end{aligned} \tag{6}$$

The solution to Equation 5 is given in Lemma 2.1 provided in [1]. Moreover, the MSE:

$$\mathbf{\Sigma}^\star = \frac{\beta^2}{\delta^2}\mu\mathbf{I} + \frac{\alpha^2}{\delta^2}\mathbf{S}, \tag{7}$$

$$\mathbb{E}[||\mathbf{\Sigma}^\star - \mathbf{\Sigma}||_F^2] = \frac{\alpha^2\beta^2}{\delta^2}. \tag{8}$$

One way of interpret the work of Ledoit & Wolf [1] is by means of considering a balancing between bias and variance for our objective. Looking at the shrinkage term to be purely bias makes sense from the perspective of having an established baseline. Then, based on how we sample the data, we have an arbitrary, uncontrolled variance, deriving from the sample covariance matrix. In other words, when pursuing to minimize the objective in Equation 5, we are essentially minimizing the mean squared error, in which the shrinkage target $\mu\mathbf{I}$ is purely bias, and the sample term $\mathbf{S}$ is purely variance.

2.3 Rotationally Invariant Estimators (RIEs)

From this point on, let's assume that $T > N$. In addition, during this section only, we assume that $\mathbf{X}$ is a normalized data matrix, which means that a specific data point in $\mathbf{X}$ for this section will be defined as: $\mathbf{X}[t,i] = \frac{r_{t,i} - \bar{\mathbf{r}}_i}{\sigma_{\mathbf{r}_i}}$, whereby $\sigma_{\mathbf{r}_i} = \sqrt{\mathbb{E}[\mathbf{r}_i^\top \mathbf{r}_i] - \mathbb{E}[\mathbf{r}_i]^2}$, and $\bar{\mathbf{r}}_i = \mathbb{E}[\mathbf{r}_i]$. With the latter in mind, we shall also define the true correlation matrix:

$$\mathbf{C} = \mathbb{E}[\mathbf{X}^\top \mathbf{X}]. \tag{9}$$

In contrast to the previous two methods, RIE builds on the idea of attempting to estimate a true correlation matrix $\mathbf{C}$ [2]. However, in a similar fashion to LWE, RIE also uses the sample based estimation. This is done in order to attempt to reach an estimate of the true correlation matrix $\mathbf{C}$ that produces the smallest error. Bun et al define the sample based estimate for the correlation matrix as follows:

$$\begin{aligned}\mathbf{E} &:= \frac{1}{T}\mathbf{X}^\top\mathbf{X} \\ &:= \sum_{k=1}^{N} \lambda_k \mathbf{u}_k \mathbf{u}_k^* \in \mathbb{R}^{N\times N},\end{aligned} \tag{10}$$

where $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_N \geq 0$ are the eigenvalues of $\mathbf{E}$, and $\mathbf{u}_1, ..., \mathbf{u}_k$ are the eigenvectors. In addition, vectors $\mathbf{u}_k$ represent the eigenvectors of $\mathbf{E}$. With that efforts are then focused on understanding the eigenvalues that compose $\mathbf{E}$. In the case of interest, Definition 2.0.2, the eigenvalues of the SCE become distorted (compared to the eigenvalues of $\mathbf{C}$). Again, the idea is to construct an estimate $\mathbf{\Xi}$ of the correlation matrix that is closer to the true correlation matrix $\mathbf{C}$.

With the above in mind, the following may be stated:

$$\mathbf{\Xi}^{\mathbf{RIE}} := \sum_{\mathbf{k=1}}^{\mathbf{N}} \xi_{\mathbf{k}}^{\mathbf{RIE}} \mathbf{u_k} \mathbf{u_k^*}. \tag{11}$$

Equation 11 suggests an estimate $\mathbf{\Xi}^{\mathbf{RIE}}$ which depends on a modified eigenvalue ξ_k^{RIE}. The new eigenvalue is further defined as:

$$\xi_k^{RIE} = \frac{\lambda_k}{|1 - q + q z_k s_k(z_k)|^2}, \tag{12}$$

where $q = N/T$ (as in Definition 2.0.2), $z_k = \lambda_k - \frac{i}{\sqrt{N}}$, $i^2 = -1$, $s_k(\cdot)$ is defined in Equation 13, and $k \in [1, N]$.

Now, Equation 12 suggests that one should re-scale the eigenvalues. The work of Marcenko & Pastur [7] propose that $\mathbf{E}$ contains a spectrum that is broader than the spectrum of $\mathbf{C}$. Due to this scaling, one should expect small eigenvalues to be smaller than those of the true matrix $\mathbf{C}$ (and large eigenvalues to be larger).

Next, one of the main weighing components which considers the difference between eigenvalues is defined:

$$\begin{aligned} s_k(z_k) &= \frac{1}{N}\sum_{\substack{j=1\\ j\neq k}}^{N} \frac{1}{z_k - \lambda_j} \\ &= \frac{1}{N}\sum_{\substack{j=1\\ j\neq k}}^{N} \frac{1}{\lambda_k - \lambda_j - \frac{i}{\sqrt{N}}}, \end{aligned} \tag{13}$$

where is again defined as $i^2 = -1$.

Although the achieved estimate in Equation 12 may work for a set of cases, and therefore can be the final RIE, it generally adds a bias to the final solution. This bias can be significant if the number of variables in being considered are not larger than 400 [6]. As such, Bun et al proposed a de-biasing term Γ_k for eigenvalue estimate ξ_k^{RIE}. It is defined as follows:

$$\Gamma_k = \sigma^2 \frac{|1 - q + qz_k g_{mp}(z_k)|^2}{\lambda_k}, \tag{14}$$

where $\sigma^2 = \frac{\lambda_N}{(1-\sqrt{q})^2}$, and λ_N is the smallest eigenvalue of $\mathbf{E}$.

Now, the de-biasing term also depends on another variable that has not yet been defined: $g_{mp}(\cdot)$. This is the Stieltjes transform of the Marcenko-Pastur density. This term brings information about how the current eigenvalue λ_k measures in the underlying distribution. This highlights once again the authors emphasis on the idea of re-balancing eigenvalues of $\mathbf{E}$ in order to reach a better estimation accuracy. Article [6] defines g_{mp} as:

$$g_{mp}(z) = \frac{z + \sigma^2(q-1) - \sqrt{z - \lambda_N}\sqrt{z - \lambda_+}}{2qz\sigma^2}, \tag{15}$$

where $\lambda_+ = \lambda_N \left(\frac{1+\sqrt{q}}{1-\sqrt{q}}^2\right)$

With the de-biasing term defined, the final de-biased estimator of the eigenvalue is defined as:

$$\hat{\xi_k} := \xi_k^{RIE} \times \max(1, \Gamma_k). \tag{16}$$

The final de-biased estimation of the true correlation matrix $\mathbf{C}$ becomes:

$$\mathbf{\Xi^{RIE}} := \sum_{\mathbf{k=1}}^{\mathbf{N}} \hat{\xi}_{\mathbf{k}} \mathbf{u_k} \mathbf{u_k^*}. \tag{17}$$

It should be noted, however, that in order to reach an estimation $\mathbf{\Xi}^{\mathbf{RIE}}$ which produces good results, one need to perform an extensive deal or pre-processing for data $\mathbf{X}$. Namely, one must remove the sample mean of the assets and normalize the result by the sample volatility estimate. Further details about this pre-processing procedure is presented in [2] and [6].

3 Test Results

This section will present first present information about the data used. Then, we will go through the results achieved while testing the different covariance estimators.

3.1 Data

The data used 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. Date values were then matched so each i asset has a value for the same day t. Therefore, a table containing the original data matrix $\mathbf{X}_{raw}$ would look as described Figure 1.

date	F	L	MSFT	S	NVDA	ANSS	VZ	WMT	MRK	EQV1V_HE
2000-11-02 00:00:00+02:00	15.048508	11.346896	22.8218	17.588865	10.2844	2.9075	21.8678	34.0712	43.3611	0.963153
2000-11-03 00:00:00+02:00	14.615457	10.939939	22.1531	18.663744	10.9564	2.9075	21.5838	33.3256	42.7529	1.016662
2000-11-06 00:00:00+02:00	15.265035	11.083570	22.5589	18.321737	11.8021	2.9375	21.0385	34.4650	43.7843	1.177187
2000-11-07 00:00:00+02:00	14.435014	11.330938	22.8835	19.298891	11.0040	2.8125	20.8038	34.4228	42.2665	1.230696
2000-11-08 00:00:00+02:00	14.326763	11.474570	22.5394	19.885189	10.6780	2.8750	20.8492	34.2892	44.1784	1.123679

Figure 1: Data table view. Every column is a different asset. Each row contains the closing value for that respective data.

Next, the values presented in Figure 1 are translated to daily return values, meaning:

$$\mathbf{X}[t,i] = \frac{\mathbf{X}_{raw}[t+1,i] - \mathbf{X}_{raw}[t,i]}{\mathbf{X}_{raw}[t,i]}. \tag{18}$$

Now, the total number of observations T and assets N in $\mathbf{X}_{raw}$ is 4409 and 207. This means: $\mathbf{X}_{raw} \in \mathbb{R}^{4409\times 207}$.

For the experiments that follow we vary the different estimators, along with the length of look-back windows, as well as the number of assets. Now, a look-back window is the number of data points considered at every re-balancing point [3]. As such, we consider the last month of observations $[t-22, t]$, or the last three months $[t-66, t]$ (assuming a month contains 22 trading days). Finally, we also consider the effects of the different number of assets. For example, for a given experiment we may evaluate how a specific trading strategy using LWE will perform with only 10 assets. This means that we pick ten assets at random such that $\mathbf{X} = \mathbf{X}_{10}, \mathbf{X}_{10} \in \mathbf{X}$ and $\mathbf{X}_{10} \in \mathbb{R}^{T \times 10}$.

3.2 Global Minimum Variance Portfolio

In this section, the performance achieved by the multiple estimators will be presented in the context of a Global Minimum Variance Portfolio (GMVP) strategy [3]. The goal with GMVP is to minimize the risk of a portfolio. A GMVP which follows a long-only strategy, along with with no borrowed margin may be described as the following optimization problem:

$$\begin{aligned} \underset{\mathbf{w} \in \mathbb{R}^T}{\text{minimize}} \ & \mathbf{w}^\top \mathbf{\Sigma} \mathbf{w} \\ \text{s. t.} \ & \mathbf{1} \cdot \mathbf{w} = 1 \\ & \mathbf{w} \geq \mathbf{0}. \end{aligned} \tag{19}$$

Equation 19 contains a $\mathbf{\Sigma}$ which indicates the covariance matrix for a group of asset returns, and $\mathbf{w}$, the portion of capital allocated to a specific set of assets. The goal with the constraints is to simplify the problem, since adding ideas such as shorting and margin investing would further complicate the experiment. However, it should be mentioned that the unconstrained case of Equation 19 has a direct solution: $\mathbf{w} = \frac{\mathbf{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \mathbf{\Sigma} \mathbf{1}}$.

The overall performance of the estimators is presented in the images below. The first image demonstrates the results for total capital return of the GMVP strategy, re-balanced monthly. This means that every month, the portfolio was analyzed, a covariance matrix was computed (on the basis of a 6 month lookback period), and finally re-weighted according the newly achieved weights. In addition, the graphs demonstrate results from varying number of covariance estimation methods, and number of assets. This was specifically done to analyze the different covariance estimations under a chosen number of assets (10, 50, 100, 150, and 200). The expected insight is a view on how total asset returns change with varying covariance matrix

[3] Re-balancing is when we re-compute the covariance matrix to evaluate the strategies objective. Every re-balancing point means we either buy or sell, since the weights $\mathbf{w}$ change. Re-balancing is kept constant at once per month, throughout this experiment

estimators and number of assets (dimensions N), given a constant number of observations T, and optimization strategy.

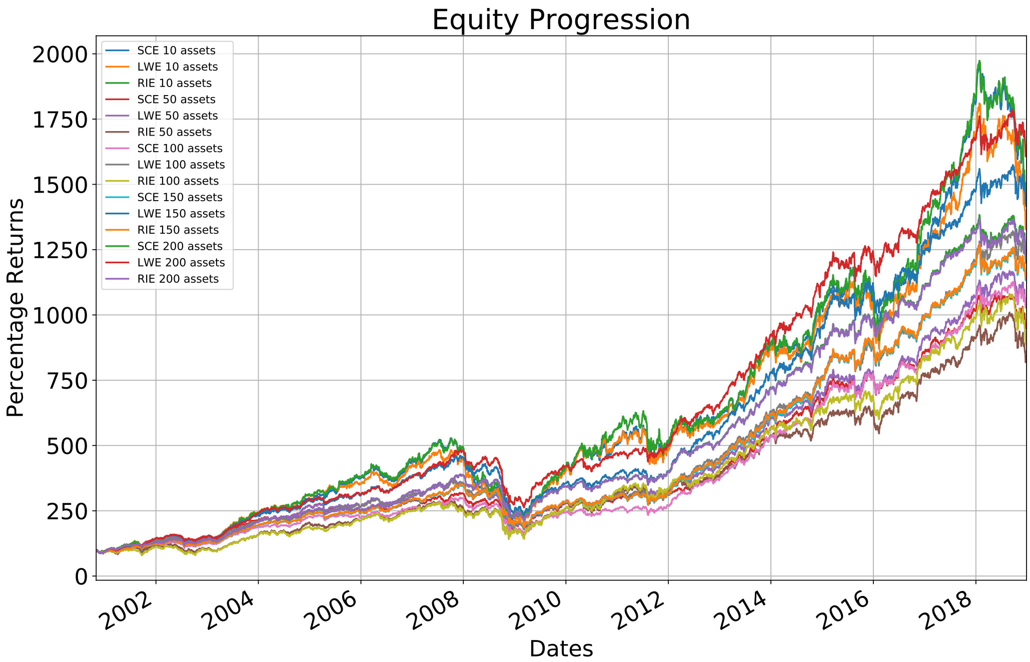

Figure 2: Percentage returns based on initially invested capital for the GMVP strategy on a 6 months look-back window.

The general message from Figure 2 is that there is a positive trend. The methods, when used in the strategy to produce the covariance matrices, produce a sufficient level of estimation such to maintain positive returns.

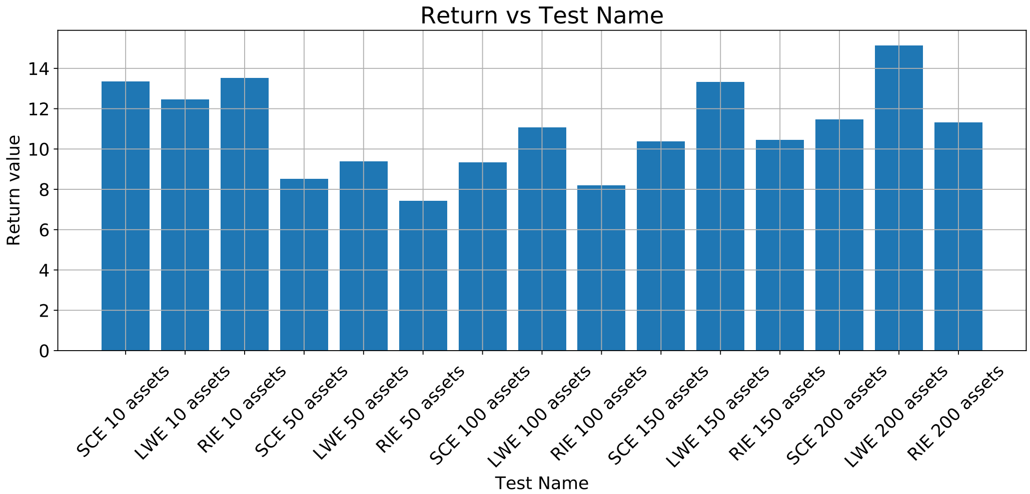

Figure 3: On the y axis we have how many times our return on equity was for a GMVP strategy and a 6 months look-back window. Assets are chosen randomly from the data return matrix $\mathbf{X}$.

Figure 3 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. The general trend shows that the larger the number of assets, the larger the returns.

Finally, by building on the above findings, its then possible to investigate which lookback period would produce the smallest risk (since that is what GMVP aims to minimize).

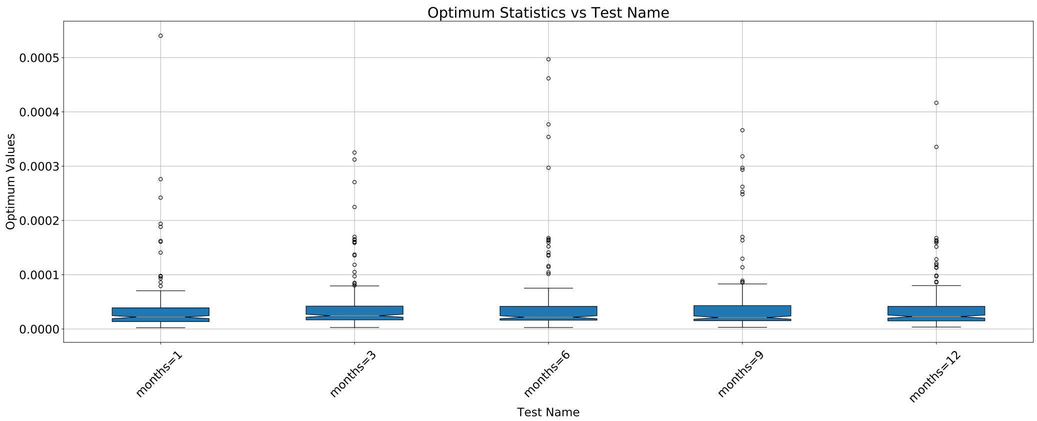

Figure 4: The optimum statistics in this case implies risk. This boxplot shows how risk varies when we change the number of look-back months to consider with the GMVP strategy, LWE, and a 150 asset portfolio.

The goal with Figure 4 is to provide insight on how risk varies when the only variable we change is the look-back period. The reason why we chose 150 assets and the LWE is because it performed well on a random set of assets in previous experiments (see Figure 2). Now, considering the results in Figure 4, it's possible to note that for three (3) months look-back, we have smaller outlying risk, while most risk values achieved are within the same 25-75 percentile margin.

3.3 Most Diversified Portfolio

This section discusses the results achieved for the Most Diversified Portfolio (MDP) strategy [4]. MDP builds on the idea that it is possible to design a measure that indicates how diversified a portfolio is based on its volatility vector and covariance matrix description. By maximizing this measure, one achieves a portfolio that assures the maximum diversification. The diversification ratio $D(\mathbf{w})$ for a portfolio $\mathbf{w}$ can be described as follows:

$$D(\mathbf{w}) = \frac{\mathbf{w}^T \text{diag}(\mathbf{S})}{\sqrt{\mathbf{w}^T \mathbf{S} \mathbf{w}}}. \tag{20}$$

It is not possible to directly compute the diversification ratio with the constraints imposed for the GMVP strategy. As a consequence, for this strategy the constraints are dropped, and therefore attempt to maximize $D(\mathbf{w})$ without weight constraints. This means that it is possible to have negative weights. The results which follow reflect the performance of different covariance estimators applied to the MDP problem.

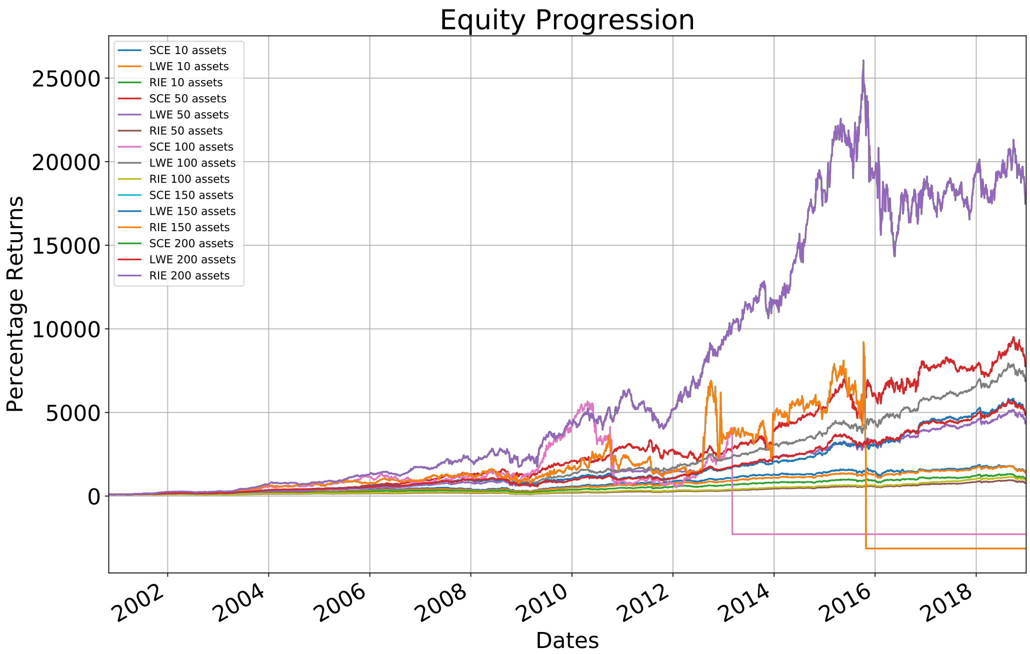

Figure 5: The Equity Progression plot for the MDP strategy, with a 6 months look-back period.

When considering Figure 5, for some cases MDP provides superior returns compared to the GMVP. However, for a number of tests, namely SCE and 100 assets, SCE and 150 assets, and RIE 150 assets. The returns turned sufficiently negative, such that the backtesting framework terminated the strategy (bankruptcy). Since the assets are chosen randomly, and there are no constraints to the MDP strategy, this demonstrates the added risk in the strategy when paired with particular estimators. The results below demonstrates the achieved returns in more details:

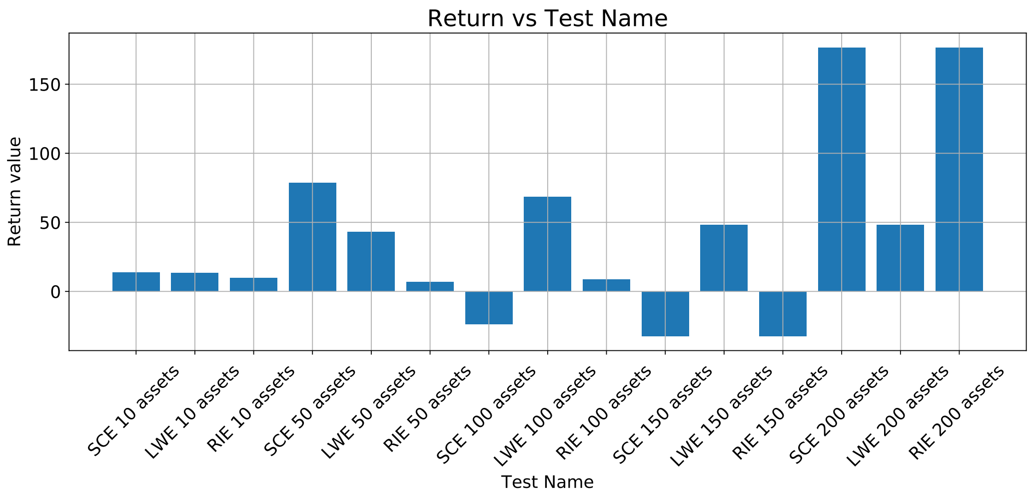

Figure 6: Detailed returns for MDP strategy, given 6 months look-back window.

Figure 6 shows that for 150 assets, SCE and RIE provided negative returns when applied to the MDP strategy. Moreover, SCE also provided negative returns for 100 asset strategy. However, in comparison, for 200 assets, RIE provided returns that surpassed 150 times the initial investment. That is significantly better than the best GMVP return.

Now, let us consider how the diversification ratio changed over different lookback periods. Figure 7 shows that most of the diversification ratios achieved are around within the same 25-75 percentile range. However, 12 months demonstrates outliers which are generally higher than other look-back windows. This is an interesting insight because it demonstrates that the diversification ratio does change significantly with different look-back windows.

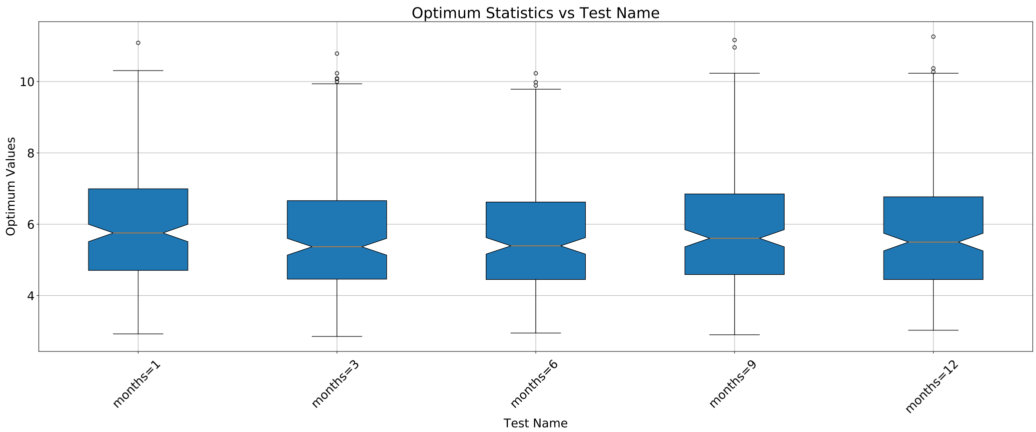

Figure 7: The optimum statistics in this case implies diversification ratio. This boxplot shows how the diversification ratio varies when we change the number of look-back months to consider with the MDP strategy, LWE, and a 150 asset portfolio.

3.4 Modified MDP Strategy

During this experiment, a third strategy was devised and tested in order to evaluate the covariance estimation methods. This strategy was developed based on the ideas proposed by the MDP. We call it Modified Maximum Diversification Portfolio strategy (MMDP). The idea is similar to MDP, in the sense of maximizing the diversification ratio, however, in this strategy the risk is constrained to be less than a certain established criteria. The logic behind the strategy is as follows:

$$\begin{aligned} \underset{\mathbf{w}\in\mathbb{R}^T}{\text{maximize}} \; & \mathbf{w}^\top \text{diag}(\mathbf{S}) \\ \text{s. t.} \; & \mathbf{1}^\top \cdot \mathbf{w} = 1 \\ & \mathbf{w} \geq \mathbf{0} \\ & \mathbf{w}^\top \mathbf{S} \mathbf{w} \leq \frac{1}{1+c}, \forall c > -1. \end{aligned} \tag{21}$$

For this experiment, the above variable c was chosen to be 0.50. In short, the constant c controls how much the user is willing to deviate from the optimum $D(\mathbf{w})$, with the constraint of risk, and an intuition that the higher returns may not necessarily lie in the optimum $D(\mathbf{w})$. Finally, if a solution to Equation 21 exists, it is used, otherwise, $\mathbf{w} = 1/N$. The equity progression graph presented on Figure 8 demonstrates the results achieved for this strategy.

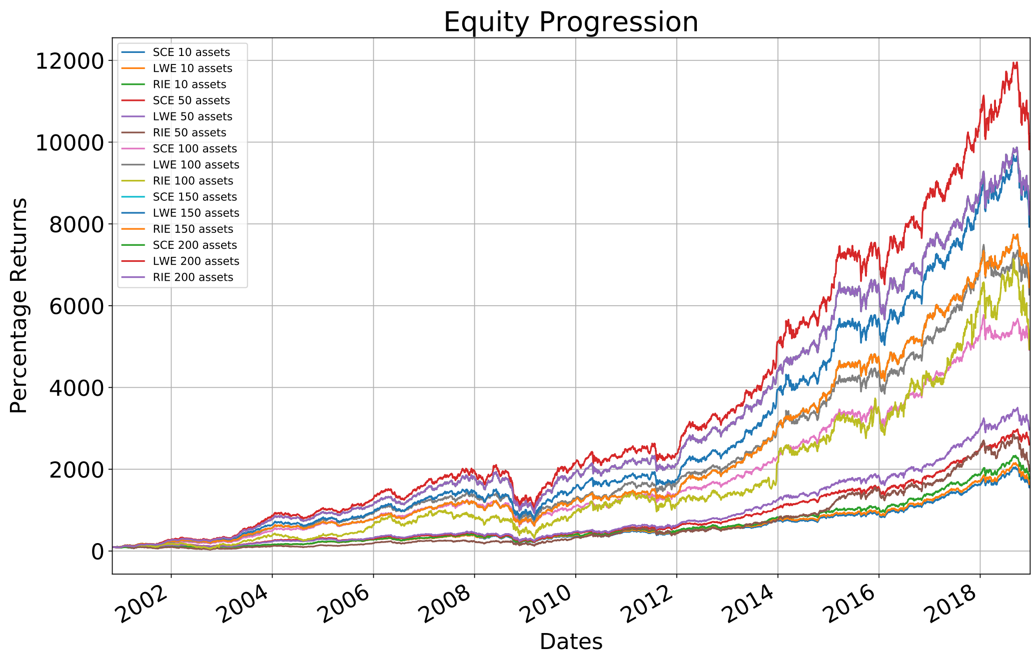

Figure 8: The Equity Progression plot for the MMDP strategy, with a 6 months look-back period.

MMDP produces positive returns, while maintaining a rising equity value trend. Figure 9 provides further details into the overall returns:

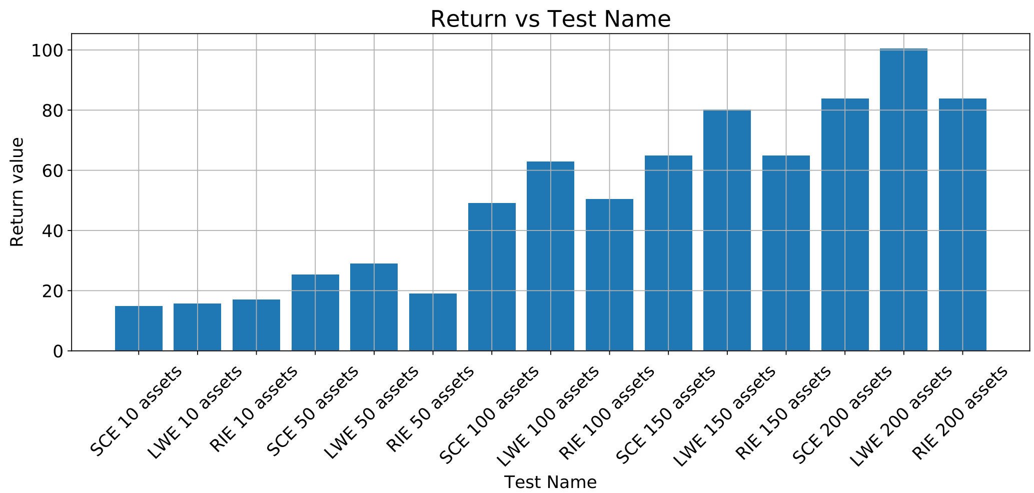

Figure 9: Detailed returns for MMDP strategy, given 6 months look-back window.

It's possible to note that once again LWE provided the largest returns for a 150 asset portfolio, while maintaining consistency in all other asset numbers. RIE followed closely, while producing the largest returns for 100 asset portfolio. It's also possible to note that asset numbers below 100 provide significantly lower returns. Finally, not only doe MMDP provide higher returns than MDP, it also manages to bring about positive returns for all covariance matrix estimators.

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

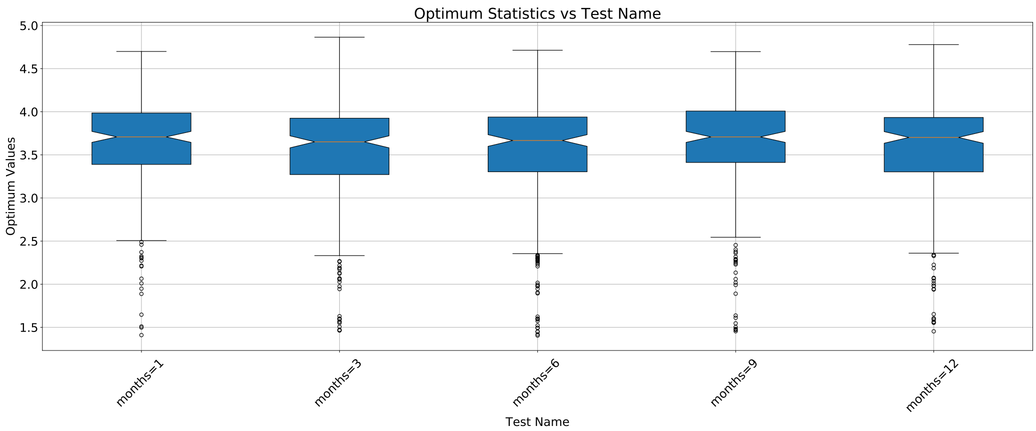

Figure 10: The optimum statistics in this case implies diversification ratio. This boxplot shows how the diversification ratio varies when we change the number of look-back months to consider with the MMDP strategy, LWE, and a 150 asset portfolio.

Figure 10 shows that the highest diversification ratio achieved was at a look-back period of 3 months, while outliers are generally on the lower end (when compared to MDP which had outliers generally on the higher end). This may be due to the fact that if the strategy is not able to find a suitable optima for the problem in Equation 21, it resorts to a 1/N strategy. This means that $\mathbf{w} = 1/N$, 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 not significantly askew the bulk of the results.

4 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] B. Hajek, *Random Processes for Engineers*. Cambridge University Press, 2015.

[6] 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.

[7] 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.