Covariance Matrix Estimation

Zhehan Shi, Shubo Xu, Xiaobin Ou

New York University

April 21, 2022

Abstract

In this report, I evaluated and compared the performance of three common covariance estimators by taking different training and testing scenarios and using three predictors $\mathbf{g}$.

1 Covariance Matrix Estimation

1.1 Introduction

The performance of any mean-variance optimization scheme closely depends on the variance-covariance matrices. Therefore, it is of utmost importance to select a suitable metric against which candidate methods of covariance matrix estimators can be tested. In this section, I compare three commonly used covariance matrix estimators – the empirical covariance, the "eigenvalue clipping" covariance estimator, and the "optimal shrinkage" estimator by the out-of-sample volatility of the Markowitz portfolios built from each estimator.

1.2 Methodology

The methodology I have adopted to test the candidate covariance estimators is the out-of-sample average volatility of the optimal portfolio built from each covariance. In a general sense, I follow BBP-Risk2016[1] for the construction of this metric. This procedure is as follows:

1. I standardize the 5-min returns of all assets by firstly taking out the sample mean of each asset's returns, and secondly normalizing the results by the cross-sectional daily volatility $\hat{\sigma}$, where $\hat{\sigma}_{it} = \sqrt{\sum_j r_{jt}^2}$;

2. To achieve approximate stationarity, I further standardize the returns by the division of the sample estimator of the volatility. I follow BBP's notation and denote this final return matrix as $\mathbf{X}$;

3. I find three estimator covariance matrices of $\mathbf{X}$: 1) the empirical covariance matrix by implementing scikit-learn; 2) the eigenvalue-clipped covariance by implementing GGiecold/pyRMT[2]; 3) the optimal shrinkage by implementing jduarte00/rie_estimator[3]. I denote the three estimators by Σ_{emp}, Σ_{ev_clip} and Σ_{os};

4. I thereby build optimal Markowitz portfolios given each covariance estimator Σ:

$$\mathbf{w} = \frac{\Sigma^{-1}\mathbf{g}}{\mathbf{g}^T\Sigma^{-1}\mathbf{g}}$$

   with choices of predictor vectors $\mathbf{g}$. I use three choices of predictor vectors $\mathbf{g}$ where

   (a) The mean variance $\mathbf{g}_{mvp} = \mathbf{e}$;

(b) The omniscient $\mathbf{g}_{omni} = \sqrt{N} \cdot \frac{r_t}{\hat{\sigma}}$, where N is the number of assets, r_t is a size N vector with each entry $r_{it} = \frac{P_{t+\tau} - P_t}{P_t}$. $\hat{\sigma}$ is as defined in 1). τ is the length of the testing period to be discussed later;

(c) The random long-short $\mathbf{g}_{rand} = \sqrt{N} \cdot \mathbf{v}$ where $\mathbf{v}$ is a unit-length vector drawn from a uniform distribution.

5. I use the out-of-sample variance of the optimal Markowitz portfolio, $\mathcal{R}^2_{t,\mathbf{w}}$ built at each time step t as the performance metric:

$$\mathcal{R}^2_{t,\mathbf{w}} = \frac{1}{T_{out}} \sum_{\tau=t+1}^{t+T_{out}} (\sum_{i=1}^{N} \mathbf{w}_i X_{i\tau})^2$$

6. I calculate the array of $\mathcal{R}^2$ values across time for each choice of predictor $\mathbf{g}$ and covariance estimator Σ. I find the average and standard deviation of these $\mathcal{R}^2$ values.

1.3 Unit Tests

I have built a set of unit tests for our implementation in code. Please refer to the code submission.

1.4 Analysis of the Covariance Matrix Estimators and Their Performance

While performing data analysis, I have observed that the performance of the covariance estimators – specifically the comparative performance of the empirical and the optimal shrinkage estimators vary for different choices of time intervals for training and testing. These choices can be easily implemented by choosing different values of $q = N/T$. Here, I present results from two choices of time intervals for training and testing. The notations I have adopted denote $T_{total} = 5070$ to be the total number of 5-min returns buckets across three months, T to be the number of returns buckets used for training, T_{out} to be the number of returns buckets used for testing. There are 78 5-min returns buckets in a trading day.

1. I choose $T = 650$, which is roughly 8 days and $T_{out} = 39$ which is half a day. $q = 505/650 = 0.777$. Please refer to Table 1 for the results.

Annualized out-of-sample average volatility (in %) of the different strategies (standard errors are given in brackets) with $T = 650, T_{out} = 39$	
Minimum variance portfolio	
Empirical Covariance	0.3696 (0.2589)
Eigenvalue Clipping	22.3156 (3.4306)
RIE	0.3311 (0.0481)
Omniscient predictor	
Empirical Covariance	0.0791 (0.0469)
Eigenvalue Clipping	2.0698 (1.3001)
RIE	0.0787 (0.0195)
Uniformly random predictor	
Empirical Covariance	0.0204 (0.0155)
Eigenvalue Clipping	0.6806 (0.2052)
RIE	0.0161 (0.0031)

Table 1: Performance of covariance estimators with $T = 650, T_{out} = 39$

2. I choose $T = 1010$, which is roughly 13 days and $T_{out} = 60$ which is short of a day. $q = 505/1010 = 0.5$. This ratio is approximately the same as the one used by BBP-Risk2016[1]. Please refer to Table 2 for the results.

Annualized out-of-sample average volatility (in %) of the different strategies (standard errors are given in brackets) with $T = 1010, T_{out} = 60$	
Minimum variance portfolio	
Empirical Covariance	0.2388 (0.0515)
Eigenvalue Clipping	22.739 (3.1210)
RIE	0.2731 (0.0287)
Omniscient predictor	
Empirical Covariance	0.0471 (0.0122)
Eigenvalue Clipping	1.6927 (1.0455)
RIE	0.0491 (0.0121)
Uniformly random predictor	
Empirical Covariance	0.0137 (0.0036)
Eigenvalue Clipping	0.0711 (0.2243)
RIE	0.0138 (0.0017)

Table 2: Performance of covariance estimators with $T = 1010, T_{out} = 60$

3. To confirm the correctness of our model and the validity of findings, I apply the same methodology to **daily** returns of 470 stocks. I obtained the stock data from Kaggle[4] from 2013-02-08 to 2018-02-07. I filtered out stocks that do have returns every day, leaving us with 470 stocks. I then randomly selected 300 stocks among them.
   I choose $T = 600$, which is roughly two and half years and $T_{out} = 36$ which is roughly one month. Please refer to Table 3 for the results.

Annualized out-of-sample average volatility (in %) of the different strategies (standard errors are given in brackets) with $T = 600, T_{out} = 36$	
Minimum variance portfolio	
Empirical Covariance	698.544 (85.867)
Eigenvalue Clipping	628.157 (55.397)
RIE	546.644 (62.478)
Omniscient predictor	
Empirical Covariance	44.983 (10.876)
Eigenvalue Clipping	43.429 (11.377)
RIE	41.710 (10.718)
Uniformly random predictor	
Empirical Covariance	8.319 (1.538)
Eigenvalue Clipping	7.613 (1.256)
RIE	6.626 (0.801)

Table 3: Performance of covariance estimators with $T = 600, T_{out} = 36$

Analysis:
The results our model finds with daily return data is in accordance with our expectation, and confirms its validity. Therefore, I move on to analyze the results I have obtained for the high-frequency data.

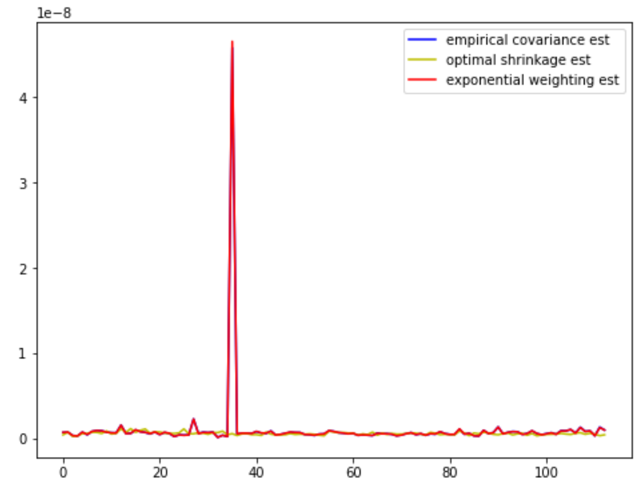
I observe that controlling for the covariance estimator, the uniformly random predictor performs the best, followed by the omniscient predictor. The minimum variance predictor incurs the most out-of-sample average volatility. This behavior is consistent with the results from BBP-Risk2016[1]

With either choice of time interval length, controlling for the predictor vector, the eigenvalue clipped covariance estimator incurs the highest out-of-sample average volatility. (See Table 1 and 2). The relative performance of the empirical covariance and the optimal shrinkage estimator vary.

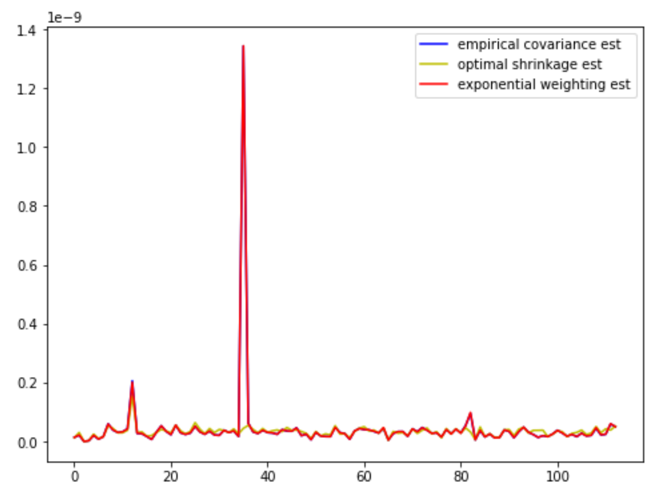
Graphs 1-3 show the out-of-sample average volatility of the two estimators with approximately 8-day training period and half-day testing period. In this case, the optimal shrinkage estimator performs superiorly.

Graphs 4-6 show the out-of-sample average volatility of the two estimators with approximately 13-day training period and less than a day testing period. In this case, the empirical covariance performs superiorly.

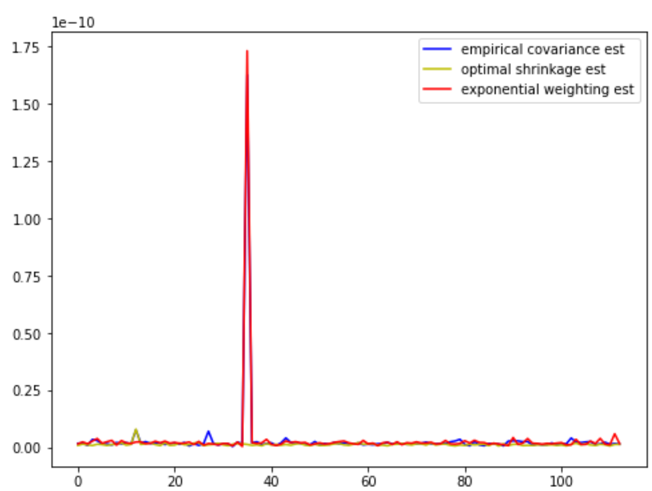
The results I have attained could not allow us to draw the conclusion that optimal shrinkage is the



(a) Graph 1: Minimum variance portfolio

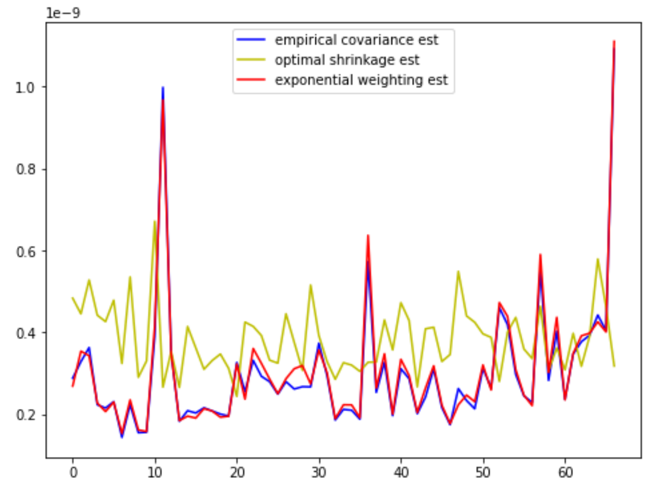


(b) Graph 2: Omniscient predictor

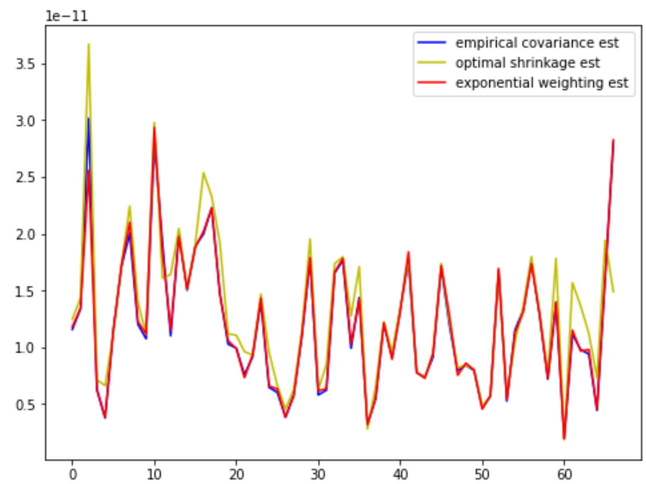


(c) Graph 3: Uniformly random predictor

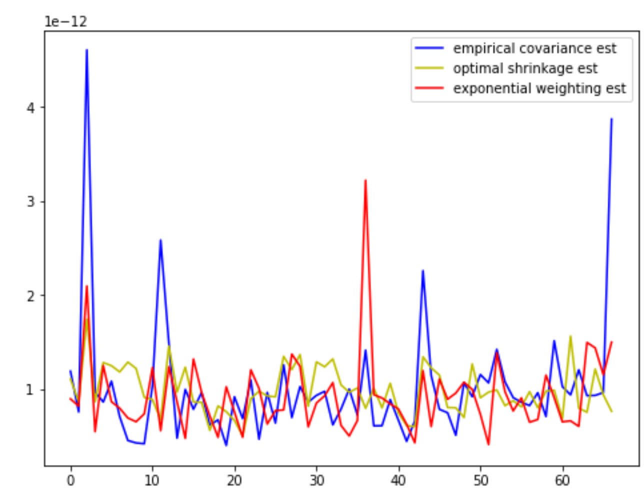
$$T = 650, T_{out} = 39$$



(a) Graph 4: Minimum variance portfolio

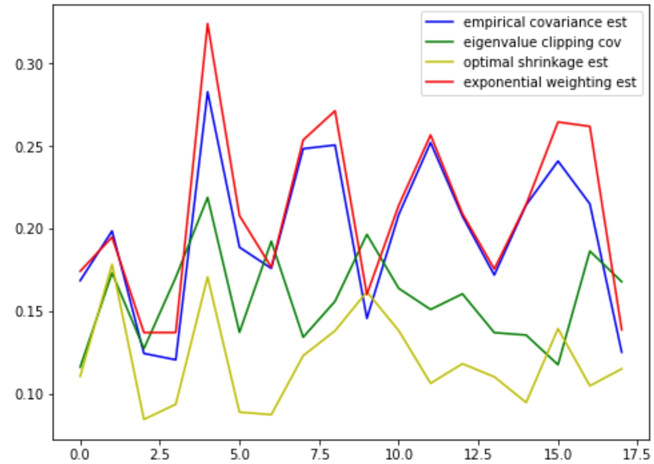


(b) Graph 5: Omniscient predictor

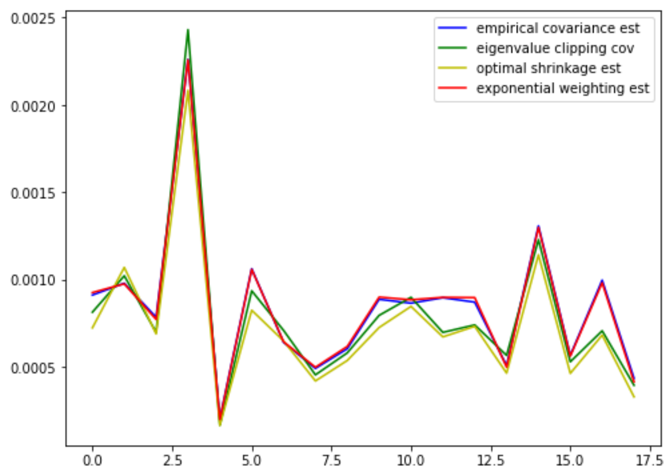


(c) Graph 6: Uniformly random predictor

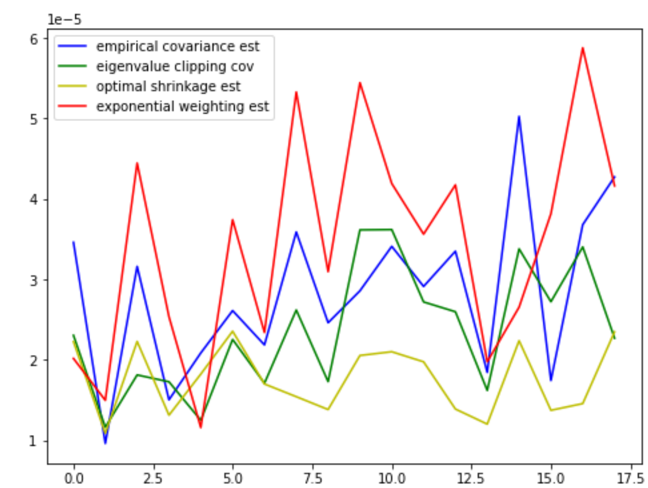
$$T = 1010, T_{out} = 60$$



(a) Graph 7: Minimum variance portfolio



(b) Graph 8: Omniscient predictor



(c) Graph 9: Uniformly random predictor

$$T = 600, T_{out} = 36$$

absolute superior of the three, which is the conclusion of BBP-Risk2016[1]. The Eigenvalue clipping estimator performing the worst is also disparate from the findings of BBP-Risk2016[1]. It is our belief that these differences in findings are due to the different frequencies of the data. Indeed, our framework reaches the same conclusion as BBP-Risk2016[1] when ran on daily return data (see Table 3, Graphs 7-9).

1.5 Summary of Findings

To summarize, our results align with BBP-16 in recommending the use of optimal shrinkage estimators when handling daily return data. However, the superiority of the optimal shrinkage estimator over the empirical covariance is no longer absolute in the high-frequency scenario. For high-frequency data, the eigenvalues clipping method can be ruled out based on our findings, however, I cannot solely advise for the usage of the optional shrinkage estimator, as the success of a rotationally invariant estimator is highly dependent on the appropriate selection of $q = N/T$. A naive value of $q = 0.5$ is no longer valid in the high-frequency realm. This gives us the direction of future research which evolves calibrating q from empirical data. Our selection of $q = 0.777$ allowed us to realize the potential of the rotationally invariant estimator. This selection of q allows us to arrive at an optimal shrinkage estimator that avoids catastrophic risk appearing in empirical covariance estimators in the Graphs 4-6.

1.6 Exponential Weighting Covariance Estimator (Extra Credit)

In this section, I consider an additional covariance estimator developed by Marc Potters in 2004[5]. The estimator is calculated as below:

$$C_{ij} = \frac{1-\alpha}{1-\alpha^T} \sum_{k=0}^{T-1} \alpha^k x_{ik} x_{jk}$$

where α satisfies

$$\frac{1}{N(1-\alpha)} = 2$$

In summary, its results are mediocre. It has the second and third lowest risk value in high frequency data depending on the choice of T, T_{out} and the use of predictors, and the highest risk value in daily data. It is very computationally expensive because it involves much more floating point operations than others. The overwhelming amount of floating point computation makes it the slowest of all 4 covariance estimators. The mediocrity of its risk performance is not justifiable for its computational expensiveness. I would not recommend this method for covariance estimation as the result of this average performance and high computational cost.

Performance of Exponential weighting covariance (standard errors are given in brackets)	
$T = 650, T_{out} = 39$	
Minimum variance portfolio	0.3702 (0.2609)
Omniscient predictor	0.0792 (0.0469)
Uniformly random predictor	0.0211 (0.0159)
$T = 1010, T_{out} = 60$	
Minimum variance portfolio	0.2414 (0.0525)
Omniscient predictor	0.0472 (0.0120)
Uniformly random predictor	0.0133 (0.0026)
$T = 600, T_{out} = 36$	
Minimum variance portfolio	720.990 (89.828)
Omniscient predictor	44.989 (10.940)
Uniformly random predictor	9.125 (1.889)

Table 4: Performance of Exponential covariance estimator

References

[1] https://www.risk.net/risk-magazine/technical-paper/2452666

[2] https://github.com/GGiecold/pyRMT

[3] https://github.com/jduarte00/rie_estimator

[4] https://www.kaggle.com/datasets/camnugent/sandp500

[5] https://doi.org/10.48550/arXiv.cond-mat/0402573