

Article

Improved Large Covariance Matrix Estimation Based on Efficient Convex Combination and Its Application in Portfolio Optimization

Yan Zhang ¹, Jiyuan Tao ², Zhixiang Yin ¹ and Guoqiang Wang ^{1,*} 

¹ School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science, Shanghai 201620, China

² Department of Mathematics and Statistics, Loyola University Maryland, Baltimore, MD 21210, USA

* Correspondence: guoq_wang@hotmail.com

Abstract: The estimation of the covariance matrix is an important topic in the field of multivariate statistical analysis. In this paper, we propose a new estimator, which is a convex combination of the linear shrinkage estimation and the rotation-invariant estimator under the Frobenius norm. We first obtain the optimal parameters by using grid search and cross-validation, and then, we use these optimal parameters to demonstrate the effectiveness and robustness of the proposed estimation in the numerical simulations. Finally, in empirical research, we apply the covariance matrix estimation to the portfolio optimization. Compared to the existing estimators, we show that the proposed estimator has better performance and lower out-of-sample risk in portfolio optimization.

Keywords: covariance matrix estimation; shrinkage transformations; rotation-invariant estimator; portfolio optimization

MSC: 90C25; 62P05; 62P20



Citation: Zhang, Y.; Tao, J.; Yin, Z.; Wang, G. Improved Large Covariance Matrix Estimation Based on Efficient Convex Combination and Its Application in Portfolio Optimization. *Mathematics* **2022**, *10*, 4282. <https://doi.org/10.3390/math10224282>

Academic Editors: Xiang Li, Shuo Zhang and Wei Zhang

Received: 24 October 2022

Accepted: 13 November 2022

Published: 16 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the development of information technology, the covariance matrix estimation plays a crucial role in multivariate statistics analysis, and it is used widely in many fields, such as finance, wireless communications, biology, chemometrics, social networks, health sciences, etc. [1–5]. In particular, due to the high noise of the sample covariance matrix, the properties of financial data are not characterized by multivariate normality and stationarity [6]. As an essential input to many financial models, it is vital to remove the sample noise to improve the estimation accuracy of the covariance matrix in asset allocation and risk management [4,7,8]. It is known that the sample covariance matrix is no longer a good estimator of the population covariance matrix when the dimension of the matrix is close to or larger than the sample size. In fact, the sample covariance matrix becomes a singular matrix in high-dimensional data. The so-called “high dimensions” mainly include large orders of 30 in magnitude and high data dimensions [1–3]. So far, some popular ways used to obtain a good estimator are the shrinkage estimation methods without prior information, sparse estimation methods with prior information, the factor model [9,10], the rank model [11], etc.

The shrinkage method, proposed by Stein [12], is one without prior information for estimating the covariance matrix. The essential idea of this method is to pull extreme eigenvalues of the sample covariance matrix toward the mean of the eigenvalues by shrinking the eigenvalues when the dimension of the matrix is close to the sample size. Ledoit and Wolf showed that the shrinkage estimation methods have an improvement over the sample covariance matrix. Specifically, they proposed the shrinkage estimation

methods provide good solutions to deal with the overfitting of the sample covariance matrix [8,13,14].

Since the linear shrinkage method is the first-order approximation to a nonlinear problem, as the dimension of the matrix becomes high, it is no longer suitable for the improvement of the sample covariance matrix. Thus, they proposed the nonlinear shrinkage method [15], which has better performance for high-dimensional asymptotics. Recently, Ledoit and Wolf proposed optimal nonlinear shrinkage estimators [16], which are decision-theoretically optimal within a class of nonlinear shrinkage estimators. For more details on the shrinkage methods, refer to [2,3,17,18].

The sparse estimation with prior information is another one for estimating the covariance matrix, which estimates the sparse matrix directly and its inverse indirectly. In the case of direct estimation, Bickel and Levina [19] showed that the estimation can be obtained by the threshold methods under the hypothesis of the sparseness of the true covariance matrix. In the case of no assumption of the sparse pattern, Rothman et al. [20] proposed a new class of generalized threshold estimators to obtain the sparse estimation by inducing sparsity and imposing the norm penalty. Theoretically, these methods are shown to be superior, and the generalized thresholding estimators are consistent with a large class of approximate sparse covariance matrices. In fact, the resulting estimators are not always positive-definite. In order to guarantee the positive definiteness of the covariance matrix estimation, Rothman et al. [21] built a convex optimization model based on the quadratic loss function under the Frobenius norm (F -norm) and studied the estimation of the high-dimensional covariance matrix. Subsequently, some convex optimization models with penalty functions such as the L_1 function were proposed [22,23], and some nonconvex penalty functions were used to achieve both sparsity and positive definiteness [24,25]. However, since the changes of the variance and covariance over time are not considered, they are affected by dimensional disasters and large noise problems. For more details about optimization algorithms and inverse matrix estimation methods, refer to the literature [7,26–31] and the references therein.

In addition, Bun et al. [32] introduced the rotation-invariant estimation in which they assumed that the estimator of the population correlation matrix shares the same eigenvectors as the sample covariance matrix itself. The experiments' results demonstrated that the rotation-invariant estimator is more suitable for dealing with large dimension datasets than the eigenvalue clipping methods and can be significantly improved over the sample covariance matrix as the data size grows, but it did not perform well on a small sample data. In a recent paper [33], Deshmukh et al. combined the shrinkage transformation with the eigenvalue clipping to obtain the estimator of the covariance matrix for the convex combination of the optimal parameters. This estimator can achieve less out-of-sample risk in portfolio optimization for small datasets.

The research in this paper was mainly motivated by [32,33], and the novelties of this study are as follows:

1. A new large covariance matrix estimator is proposed by constructing a convex combination of the linear shrinkage estimation and the rotation-invariant estimator under the Frobenius norm.
2. Our new covariance matrix estimator improves the impact of the sample noise on the covariance matrix by adjusting the parameters of the convex combination in financial data.
3. The proposed estimator has better performance and lower out-of-sample risk in portfolio optimization.

The rest of this paper is organized as follows: Section 2 describes the related work of covariance matrix estimation. Section 3 introduces our new proposed estimator and its application. Section 4 implements the numerical simulation and empirical research. Section 5 gives the conclusions.

2. Preliminaries

2.1. The Rotation-Invariant Estimator

First, we briefly introduce the basic idea of the rotation-invariant estimator. For more details, we refer to [32].

Let $r = (r_1, r_2, \dots, r_N)$ denote a $T \times N$ matrix of T independent and identically distributed (iid) observations on a system of N random variables with mean vector μ . N and T denote the number of variable and the size of the variable, respectively. In this case, the sample covariance matrix is given by

$$\Sigma_{SCM} = (\sigma_{ij}) = \frac{1}{N-1} \sum_{i=1}^N (r_i - \mu)(r_i - \mu)', \tag{1}$$

Let N and T be asymptotic in the high-dimensional regime, i.e.,

$$N \asymp T. \tag{2}$$

In addition, the concentration ratio is given by

$$c = \frac{N}{T}. \tag{3}$$

The construction steps of the rotation-invariant estimator are as follows:

- **Step 1:** Calculate the Stieltjes transform of the empirical spectral measure of S_1 from

$$s(z) = \frac{1}{T} \text{Tr}(S_1 - z)^{-1} \tag{4}$$

where z denotes the spectral parameter and

$$S_1 = \frac{1}{N-1} \sum_{i=1}^N (r_i - \mu)'(r_i - \mu), \tag{5}$$

The function (4) contains all the information about the eigenvalues of the matrix S_1 , which has the same nonzero eigenvalues as Σ_{SCM} .

- **Step 2:** Update (4) based on the nonzero eigenvalues of $Y'Y$ and YY' , i.e.,

$$s(z) = \frac{1}{T} \text{Tr}(S_1 - z)^{-1}, \tag{6}$$

where $y_i = r_i - \mu$, $Y = (y_1, y_2, \dots, y_N)$, and λ_i denotes the i th eigenvalue of the sample covariance matrix Σ_{SCM} .

- **Step 3:** Calculate the function $\hat{\delta}_i$ of the i th eigenvalue of S from

$$\hat{\delta}_i = \frac{1}{\lambda_i |s(\lambda_i + i\eta)|^2}. \tag{7}$$

where $s(\cdot)$ is the empirical Stieltjes transform from (6) and parameter $\eta = T^{-\frac{1}{2}}$.

- **Step 4:** Output the resulting covariance matrix estimator Σ_{RIE} from

$$\Sigma_{RIE} = U_N \hat{D}_N U_N', \tag{8}$$

where

$$\hat{D}_N = \text{Diag}(\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_N), \tag{9}$$

U_N is an orthogonal matrix, whose columns $[u_1, u_2, \dots, u_N]$ are the corresponding eigenvectors, with the eigenvalue of the rotation-invariant estimator defined by

$$\hat{\lambda}_i = \frac{\sum_{i=1}^N \lambda_i}{\sum_{i=1}^N \hat{\delta}_i}. \tag{10}$$

One can easily verify that

$$\sum_{i=1}^N \hat{\lambda}_i = \sum_{i=1}^N \lambda_i. \tag{11}$$

This implies that the estimator Σ_{RIE} has the same trace as the sample covariance matrix. More literature reviews on rotation-invariant estimators are presented in the Table 1.

Table 1. The related literature review.

Author	Brief Introduction	Ref.
Ledoit, O., Wolf, M.	Under the assumption of the large dimension asymptotic, Ledoit and Wolf kept the eigenvectors of the sample covariance matrix and shrunk the inverse sample eigenvalues to construct a rotation-invariant estimator of the large covariance matrix.	[34]
Donoho et al.	Based on spiked covariance and the rotation-invariant estimator, Donoho et al. demonstrated that the optimal estimation of the population covariance matrix is related to the best shrinker, which acts as an element of the sample eigenvalues.	[35]
J. Bun et al.	J. Bun et al. established the asymptotic global law estimate model for three general classes of noisy matrices using the replica method and introduced how to “clean” the noisy eigenvalues of the noisy observation matrix.	[36]
Debashis Paul, Alexander Aue	Debashis Paul and Alexander Aue summarized the random matrix theory (RMT) and described how the development of high-dimensional statistical inference theory and practice is affected by the corresponding development in the RMT field.	[37]

2.2. Improved Covariance Estimator Based on Eigenvalue Clipping

Deshmukh et al. [33] introduced an improved estimation based on eigenvalue clipping, which takes the optimal parameters in the convex combination of the sample covariance matrix Σ_{SCM} , the shrinkage target Σ_F

$$\Sigma_F = (f_{ij}), \tag{12}$$

with

$$f_{ij} = \begin{cases} \frac{2\sqrt{\sigma_{ii}\sigma_{jj}}}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^T \frac{\sigma_{ij}}{\sigma_{ii}\sigma_{jj}}, & i \neq j, \\ \sigma_{ii}, & i = j, \end{cases} \tag{13}$$

and the matrix Σ_{MP} obtained by applying eigenvalue clipping.

Let $y_i = r_i - \bar{r}_i$ be independent, identically distributed, random variables with finite variance σ . The Marchenko–Pastur density $\rho_{\Sigma_{SCM}}(\lambda)$ of the eigenvalues of Σ_{SCM} is defined by

$$\rho_{\Sigma_{SCM}}(\lambda) = \frac{1}{N} \frac{dn(\lambda)}{d\lambda}. \tag{14}$$

where $n(\lambda)$ is the number of eigenvalues of the sample covariance matrix Σ_{SCM} less than λ .

In the condition of the limit $N \rightarrow \infty, T \rightarrow \infty$, and $\frac{1}{c} \geq 1$, the density follows from (14):

$$\rho_{\Sigma_{SCM}}(\lambda) = \frac{1}{2\pi c\sigma^2} \frac{\sqrt{(\lambda_{max} - \lambda)(\lambda - \lambda_{min})}}{\lambda}, \tag{15}$$

where

$$\lambda_{max} = \sigma^2(1 + c + 2\sqrt{c}), \quad \lambda_{min} = \sigma^2(1 + c - 2\sqrt{c}). \tag{16}$$

$[\lambda_{min}, \lambda_{max}]$ represents the MP law bounds. In this case, the covariance matrix can be cleaned by scaling the eigenvectors of Σ_{SCM} with these new eigenvalues. Σ_{MP} is obtained by this method.

Let Σ be the population covariance matrix; the optimal parameters in convex combination can be found from the following optimization problem [33].

$$\min_{\theta, \phi} \quad \|\Sigma - \Sigma_{est}\|_F \tag{17}$$

$$\text{s.t.} \begin{cases} \Sigma_{est} = \phi(\theta\Sigma_F + (1 - \theta)\Sigma_{MP}) + (1 - \phi)\Sigma_{SCM}, \\ 0 \leq \theta \leq 1, 0 \leq \phi \leq 1. \end{cases} \tag{18}$$

where θ and ϕ are the parameters of the convex combination.

Usually, the eigenvectors of the sample covariance matrix deviate from those of the population covariance matrix under large-dimensional asymptotics. Correcting the deviation of the eigenvalues of the sample covariance matrix can improve the performance of the large covariance matrix. Although the estimation can adapt to changing the sampling noise conditions by performing parameter optimization, the performance of the estimation outperforms other estimations only for small-dimensional problems.

3. Proposed Estimator and Application in Portfolio Optimization

3.1. Proposed Estimator

For further improve the performance of the large covariance matrix, we replaced the eigenvalues falling inside Marchenko–Pastur (MP) law bounds with the rotation-invariant estimator Σ_{RIE} and applied the linear shrinkage estimation to shrink the eigenvalues falling outside the MP law bounds in this paper. Our new estimation is presented below.

$$\min_{\theta, \phi} \quad \|\Sigma - \Sigma_{est}\|_F \tag{19}$$

$$\text{s.t.} \begin{cases} \Sigma_{est} = \phi(\theta\Sigma_F + (1 - \theta)\Sigma_{RIE}) + (1 - \phi)\Sigma_{SCM}, \\ 0 \leq \theta \leq 1, 0 \leq \phi \leq 1, \end{cases} \tag{20}$$

Thus, the estimation of the covariance matrix is given by

$$\Sigma^* = \phi^*(\theta^*\Sigma_F + (1 - \theta^*)\Sigma_{RIE}) + (1 - \phi^*)\Sigma_{SCM}. \tag{21}$$

where θ^* and ϕ^* are the optimal parameters of the optimization problem given by (19) and (20).

It is well known that the financial data are heavy-tailed and non-normal [7,33]. However, the existing covariance matrix estimation methods generally requires the assumption of normality [32,38]. For overcoming this drawback, we propose a new estimator, which is a convex combination of the linear shrinkage estimation and the rotation-invariant estimator

under the Frobenius norm. One advantage of the new estimation is that we can remove the noise caused by the bulk eigenvalues and the extreme eigenvalues in the financial data. Furthermore, we set five-fold cross-validation $k = 5$ to implement the simulation and empirical research for improving the accurate estimation of the covariance matrix.

The detailed steps of our new estimation are as follows.

- **Step 0:** Input the sample data $r = (r_1, r_2, \dots, r_N)$, and set $k = 1$.
- **Step 1:** Calculate Σ_{SCM} , Σ_F , and Σ_{RIE} from (5), (12), and (8), respectively.
- **Step 2:** Calculate Σ_{est} from (20), and denote $\Sigma_{est}^{(\theta_i, \phi_j)}$ for $\theta_i, \phi_j, i = 1, 2, \dots, M, j = 1, \dots, P$, where M and P are the numbers of θ and ϕ taken between 0 and 1, respectively.
- **Step 3:** Calculate the error:

$$\Delta_k^{(\theta_i, \phi_j)} = \|\Sigma_{est}^{(\theta_i, \phi_j)} - \Sigma\|_F,$$

for $\theta_i, \phi_j, i = 1, 2, \dots, M, j = 1, \dots, P$, and let $k = k + 1$.

- **Step 4:** Repeat Steps 1–3 in the cross-validation until the folds $k = 5$.
- **Step 5:** Find the optimal parameters θ^* and ϕ^* in the convex combination from the corresponding smallest error sum Δ_{sum} given by

$$\Delta_{sum} = \sum_{k=1}^5 \sum_{i=1}^M \sum_{j=1}^P \Delta_k^{(\theta_i, \phi_j)}.$$

- **Step 6:** Output the proposed estimation Σ^* from (21).

3.2. Minimum Variance Portfolio Optimization

According to Markowitz’s theory [39], we included an additional return constraint in the portfolio because even a risk-averse investor would expect a minimal positive return. The classic portfolio optimization model that satisfies the minimum expected return is defined by

$$\begin{aligned} \min_x \quad & x' \Sigma x \\ \text{s.t.} \quad & \begin{cases} 1' x = 1, \\ r' x \geq r_{min}, \\ x \geq 0, \end{cases} \end{aligned} \tag{22}$$

where x , r , and r_{min} represent the weight of portfolio optimization, daily return, and the minimum daily expected return, respectively. It is well known that the portfolio selection is widely used in the financial field, which is a convex quadratic programming problem [39].

In the portfolio optimization, the weight of each asset is closely related to the covariance matrix. An accurate covariance matrix can achieve a more reasonable weight distribution and better portfolio effect. Due to the heavy-tailed nature of financial data and the availability of limited samples [8], many studies started concentrating on the global minimum variance (GMV) portfolio. To improve the performance of the sample covariance matrix in the portfolio optimization, DeMiguel [40] added the additional constraint and regularizing asset weight vector into the minimum variance portfolio and showed that the estimator always leads the constructed portfolio to achieve a smaller variance and a higher Sharpe ratio than other portfolios. Furthermore, Ledoit and Wolf [18] applied the estimation to the portfolio optimization to overcome the dimension and noise problems of a high-dimensional covariance matrix, and the results were better than the linear shrinkage estimation [13]. Moreover, due to the influence of financial market information on covariance matrix estimation, the time-varying covariance matrix or the correlation matrix also have practical significance in portfolio optimization. For more details, refer to [41] and the references therein.

In this paper, we divided the sample data into in-sample data and out-of-sample data, which were used for the estimation and prediction of the covariance matrix, respectively. To measure the out-of-sample performance of the estimation of the covariance matrix in portfolio optimization, we used the out-of-sample risk, the average return, and the Sharpe ratio as the criteria of the measurement. The average return was annualized by multiplying it by 252 (252 trading days per year), and the standard deviation was annualized by multiplying it by $\sqrt{252}$. The out-of-sample performance of the portfolio model was evaluated through the following procedure.

- **Step 0:** Input the returns of the current in-sample r_{in} and out-of-sample data r_{out} , the expected return r_{min} , and the estimation of the covariance matrix Σ^* .
- **Step 1:** Set $r := r_{in}$, and solve the optimal weight vector x^* from Model (22) by the quadratic optimizer called quadprog in Matlab.
- **Step 2:** Calculate the out-of-sample $\hat{\Sigma}^*$ from (21), and obtain the out-of-sample standard deviation:

$$\sigma_{out} = var((x^*)' r_{out}),$$

the average return:

$$r_{ave} = E((x^*)' r_{out}),$$

and the Sharpe Ratio:

$$SR = \frac{r_{ave} - r_f}{\sigma_{out}}$$

where x^* and r_f represent the optimal weight vector and the risk-free interest, respectively.

4. Numerical Simulation and Empirical Research

4.1. Numerical Simulation

In the simulation, we used the simulation data of Engle et al. [41], and the mean return ranged between -0.0031 and 0.0036 . We divided the dataset into in-sample data and out-of-sample data, and both the sample sizes were $T = 500$. In pursuit of accuracy, we implemented the five-fold cross-validation, and the parameter selection criterion was the F -norm of the estimator and the population covariance matrix. We set three dimensions for the return series, which were $N = 100, 200, \text{ and } 400$, respectively. The maximum concentration ratio is

$$c = \frac{N}{T} = \frac{400}{500} = 0.80. \tag{23}$$

To measure the performance of the estimators, we compared the error between each estimator and the population covariance matrix. The six estimators are shown in Table 2.

Table 2. The estimators for comparison.

The Formulation of Estimation	Ref.
$\Sigma_{SCM} = \frac{1}{N-1} \sum_{i=1}^N (r_i - \mu)(r_i - \mu)'$	[42]
$\Sigma_{Identity} = I$	[8,38]
$\Sigma_L = \hat{\rho}\Sigma_{SCM} + (1 - \hat{\rho})\Sigma_F$	[13]
$\Sigma_{NL} = U_N D_N^{or} U_N^*, \lambda_i^{or} = \frac{\lambda_i}{ 1 - c - c\lambda_i \hat{m}_F(\lambda_i) }$	[15,18,43]
$\Sigma_D = \phi(\theta\Sigma_F + (1 - \theta)\Sigma_{MP}) + (1 - \phi)\Sigma_{SCM}$	[33]
$\Sigma^* = \phi(\theta\Sigma_F + (1 - \theta)\Sigma_{RIE}) + (1 - \phi)\Sigma_{SCM}$	/

In the five-fold cross-validation, we obtained the error between the proposed estimation and the population covariance matrix for the different parameters θ and ϕ under

three asset dimensions. In Figures 1–3, the horizontal and longitudinal axis represent the different values of θ and ϕ , respectively, and the vertical axis represents the sum of the error. It is obvious that there are two optimal parameters to minimizing the error between the proposed estimator and the population covariance matrix for all θ and ϕ . The results are shown in Table 3. To some degree, this ensures the effectiveness of the proposed estimation.

Table 4 shows that the F -norm error of our new estimation is the smallest in the ones of the six estimations. Under this premise, we calculated the portfolio variance in the minimum variance portfolio that satisfies the minimum 0.0015 expected return. The results are shown in Table 5. Figures 4–6 show the mean return of out-of-sample data ranging from the 1st asset to the 400th asset. We mark individual points on the graph. The horizontal and longitudinal axis represent the order of assets in the total assets and the mean return of this asset, respectively. We can see that the mean return of the point that is marked is relatively high. Generally speaking, higher asset returns will also face relatively large investment risks. To understand the following description, we divided the return into three asset risk grades: high ($r_{min} \geq 0.001$), middle ($0.0005 \leq r_{min} \leq 0.001$), and low ($r_{min} < 0.0005$), respectively. In Table 5, it is obvious that the variance of $\Sigma_{Identity}$ is the largest in all asset dimensions. In the case of $N = 200$, the asset weights of the portfolio model corresponding to $\Sigma_{Identity}$ are distributed on the 15th, 51st, 71st, 75th, and 138th assets in Figure 7, respectively, with 87% high-risk assets and 13% medium assets. However, in the case of $N = 400$, the asset weights of the portfolio model corresponding to $\Sigma_{Identity}$ are distributed on twenty assets, with 60% high-risk assets, 39.5% medium assets, and only 0.5% low-risk assets, and we can see that the high-risk and medium-risk assets account for the vast majority of the 20 assets. Instead, the portfolio model corresponding to our new estimator Σ^* distributed the weights on high-risk assets and medium-risk assets as 69% and 26.12%, respectively, to achieve the 0.0015 expected return. The remaining 5% was distributed on low-risk assets to reduce investment risk. The corresponding results are shown in Figures 8 and 9. Overall, the reasonable distribution of asset weights on high-, medium, and low-risk assets can appropriately decrease investment risks. As the number of the assets increased, the performance of our new estimator Σ^* became better. At the same time, the proposed estimation in the minimum variance portfolio was more dispersed on the allocation of the assets.

Table 3. The optimal parameters θ and ϕ in convex combination and the sum of the corresponding error.

N	θ	ϕ	Error
100	0.3333	0.3636	0.0273
200	0.3434	0.3939	0.0515
400	0.3939	0.4040	0.1080

Table 4. The error between each estimator and the population covariance matrix under the optimal parameter.

N	Σ_{SCM}	$\Sigma_{Identity}$	Σ_L	Σ_{NL}	Σ_D	Σ^*
100	0.0132	9.9927	0.0121	0.0128	0.0117	0.0055
200	0.0291	14.1313	0.0261	0.0282	0.0256	0.0103
400	0.0580	19.9842	0.0508	0.0560	0.0498	0.0216

Table 5. The variance comparison of six estimations in the minimum variance portfolio.

N	Σ_{SCM}	$\Sigma_{Identity}$	Σ_L	Σ_{NL}	Σ_D	Σ^*
100	0.0011	0.9694	0.0011	0.0010	8.0298 *	7.0942 *
200	7.8576 *	0.3956	7.9096 *	7.7615 *	6.7529 *	5.9007 *
400	3.8828 *	0.1077	4.2676 *	3.8435 *	3.7395 *	3.1368 *

* denotes the unit is 10^{-4} .

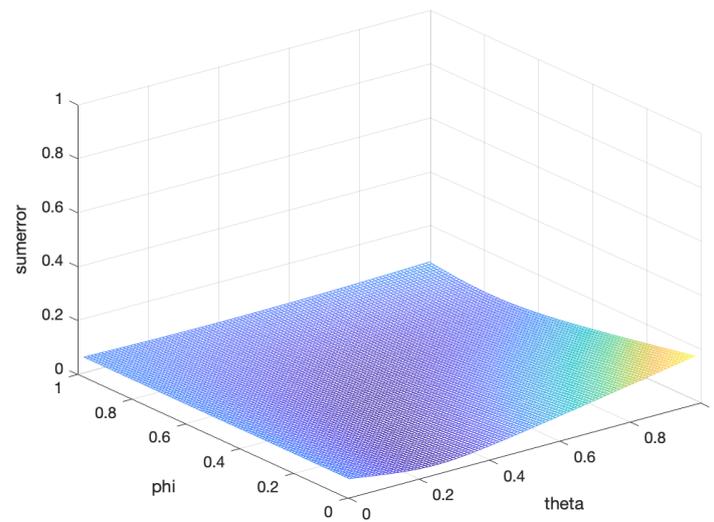


Figure 1. The sum of the error of five-fold cross-validation between the proposed estimation and the population covariance matrix under the different θ and ϕ for $N = 100$.

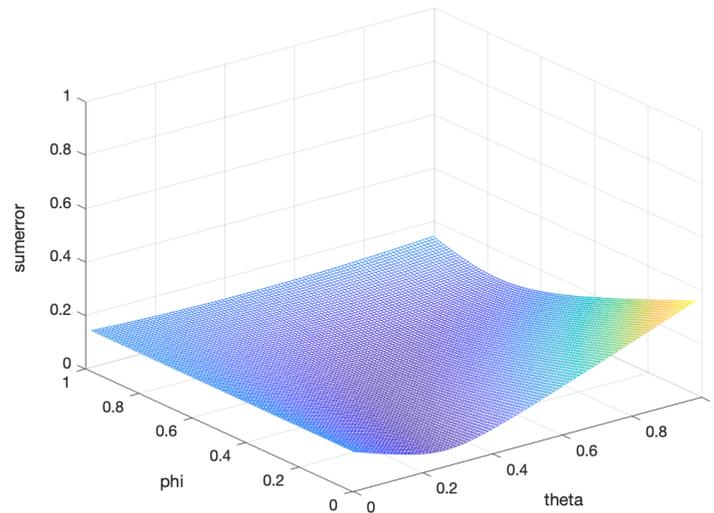


Figure 2. The sum of the error of the five-fold cross-validation between the proposed estimation and the population covariance matrix under the different θ and ϕ for $N = 200$.

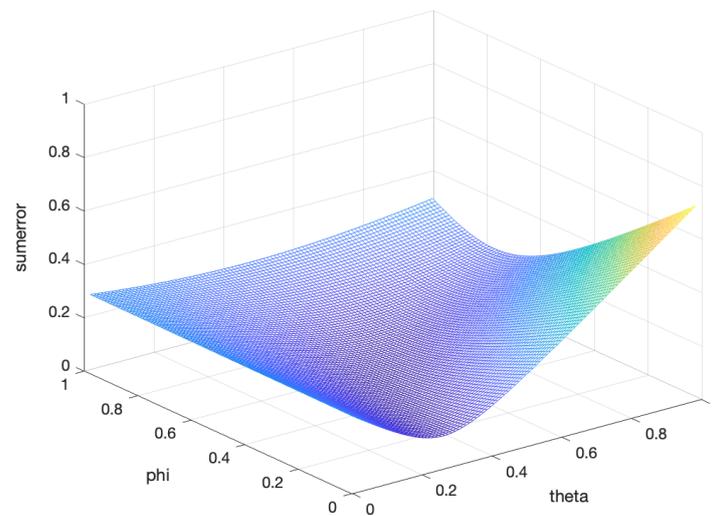


Figure 3. The sum of the error of the five-fold cross-validation between the proposed estimation and the population covariance matrix under the different θ and ϕ for $N = 400$.

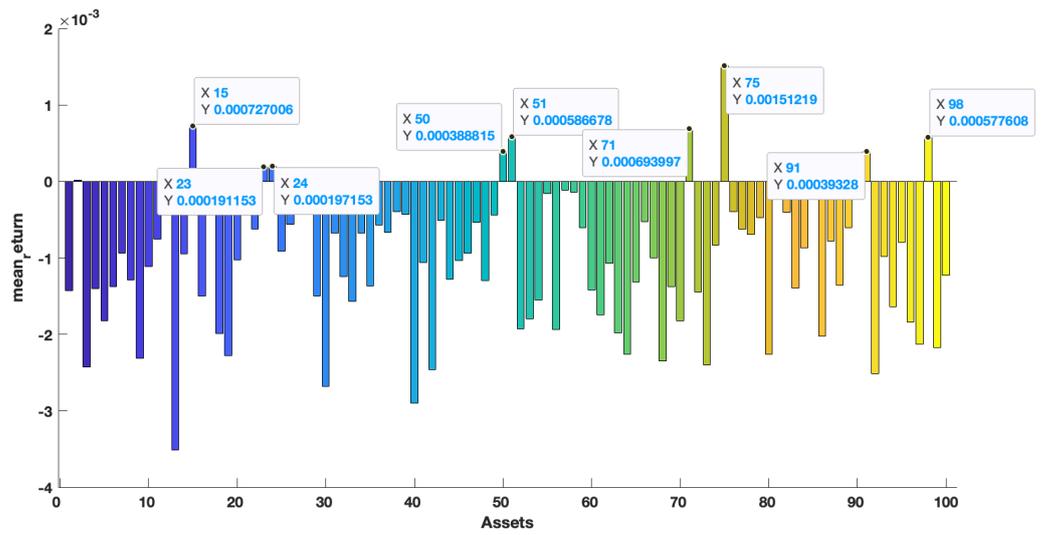


Figure 4. The mean return of the out-of-sample data for $N = 100$.

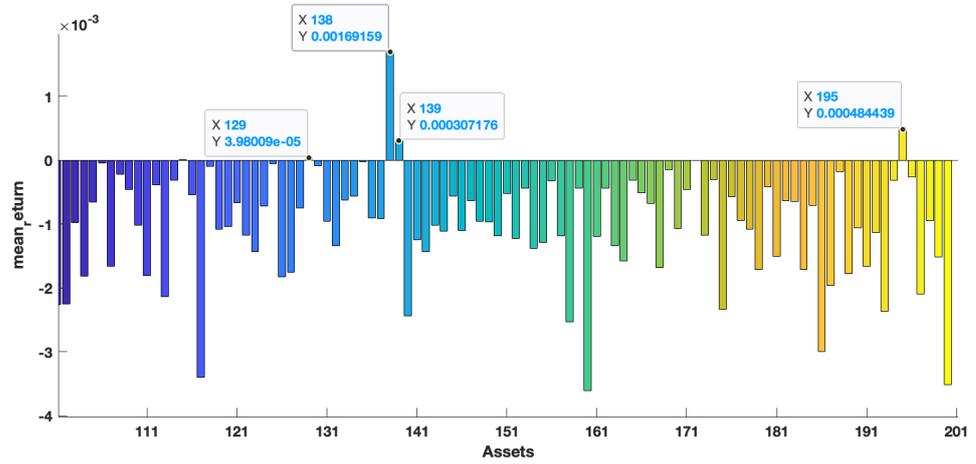


Figure 5. The mean return of the out-of-sample data range from the 101st asset to the 200th asset.

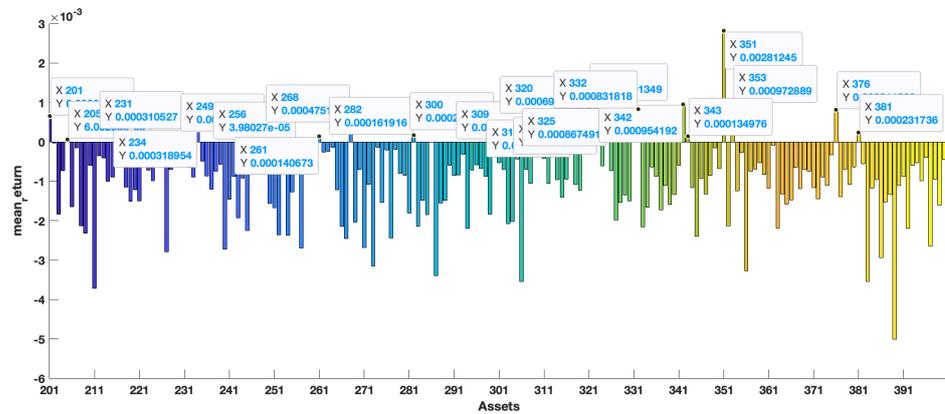


Figure 6. The mean return of the out-of-sample data range from the 201st asset to the 400th asset.

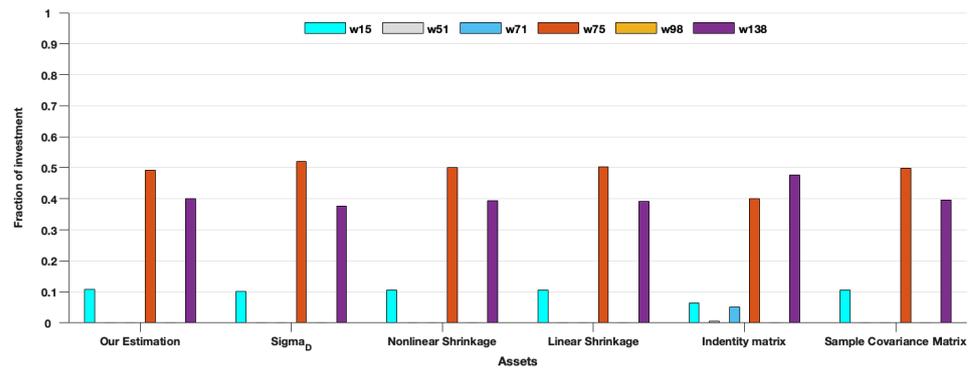


Figure 7. The assets’ weights of each estimation under the out-of-sample data $N = 200$.

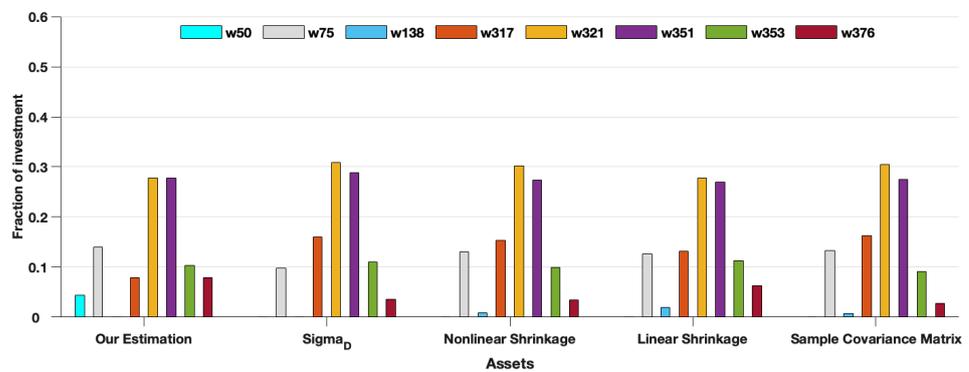


Figure 8. The assets’ weights of each estimation under the out-of-sample data for $N = 400$.

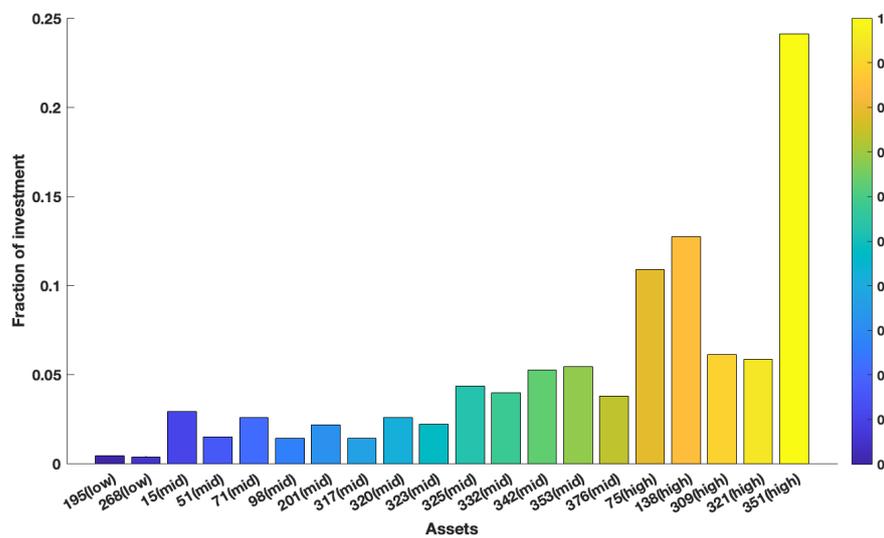


Figure 9. The assets’ weights of the identity matrix under the out-of-sample data $N = 400$.

4.2. Empirical Research

The data of this paper came from the component stock of CSI500, HS300, and SSE50 on the tushare financial website. The whole period of the samples was from 24 May 2017 to 1 July 2021. Removing the missing data of the samples from transaction, we finally obtained 426 component stocks of CSI500, 218 component stocks of HS300, and 41 component stocks of SSE50.

In this paper, we set $T_1 = 500$ and $T_2 = 500$ as the window of estimation and prediction, respectively. The maximum concentration ratio is

$$c = \frac{N}{T_1} = \frac{426}{500} = 0.852. \quad (24)$$

We used the log return as we studied the object and divide all samples into two parts for estimation and prediction. We constructed six portfolio optimization models by using the estimator in Table 2. The procedures of estimation and prediction are as follows.

- **Step 0:** Input the sample data; divide the data into in-sample data $T_1 = 500$ and out-of-sample data $T_2 = 500$.
- **Step 1:** Calculate Σ_F , Σ_{RIE} , and Σ_{SCM} based on the in-sample data.
- **Step 2:** Set five-fold cross-validation; calculate the Σ_{est} from (20) for parameters between 0 and 1; implement the minimum variance portfolio (22), where r_{min} takes a value from the minimum to maximum mean return; solve the corresponding weight vector x by multiple times of iteration over the values of both θ and ϕ in the convex optimization.
- **Step 3:** Calculate the standard deviation based on the out-of-sample data, and record the out-of-sample risk σ_{out} in each iteration.
- **Step 4:** Calculate the optimal parameters θ^* , ϕ^* when the sum of σ_{out} of the five-fold cross-validation achieves the minimum.
- **Step 5:** Implement the minimum variance portfolio (22) to obtain the assets' weights, the average return r_{ave} , and the out-of-sample risk σ_{risk} when satisfying the minimum 0.002 return constraints under the θ^* and ϕ^* .
- **Step 6:** Calculate the Sharpe ratio, where the risk-free interest is set as 1.75%.

In portfolio optimization, the reduction of volatility at the first decimal place is also considered to be quite significant [13,33]. Tables 6 and 7 show the performance of the six portfolio optimization model based on the different asset dimensions. For out-of-sample data, we used the standard deviation as the performance metric. Furthermore, we calculated the average return and the Sharpe ratio in the portfolio optimization model (22) and the risk-free interest was set to 1.75%.

Tables 6–8 show that the average returns of each estimator are equal. Table 6 shows that the standard deviation of the portfolio optimization model corresponding to $\Sigma_{Identity}$ is only 27.53%, which is the smallest in each estimator. The standard deviation of the portfolio optimization model corresponding to our new estimator Σ^* is 27.97%, and its performance ranks fourth among all estimators.

Table 7 shows the performance of the portfolio optimization model corresponding to each estimator with 218 assets. It can be seen that the performance of the estimator $\Sigma_{Identity}$ becomes weak. In this case, $\Sigma_{Identity}$ is the worst estimator, which is expected as it assumes zero correlations among stocks, and Σ_{SCM} is the second-worst. The Sharpe ratio of the portfolio optimization corresponding to Σ_{NL} is the highest in each estimator followed by the one of Σ_D . At the same time, the performance of our new estimator Σ^* ranks third among all estimators. Comparing to the case of $N = 218$, the performance of our new estimator improved as the asset dimension increased.

Table 8 compares the performance of each model with the number of assets of 426. The result shows that the portfolio optimization model corresponding to our new estimator Σ^* obtaining the smallest standard deviation leads to the highest Sharpe ratio. Obviously, compared with other estimators, especially with $\Sigma_{Identity}$, our new estimator has a significant decrease in the out-of-sample standard deviation. Meanwhile, this also implies that the performance of our new estimator Σ^* is superior to the ones of the other five estimators as the asset dimension increases.

The homogeneity of the variance test is a metric to measure whether the variances of the two investment strategies are equal, and we used the improved bootstrap inference [44] to test the significant variance difference between Σ^* and other estimations excess returns.

Table 9 shows that the test between Σ^* and the alternative methods all reject the null hypothesis of equal variances. Moreover, the sample variance of the excess return generated corresponding to our new estimator Σ^* is significantly lower than the other portfolio optimization models as the number of assets increases. Therefore, our new estimation Σ^* is superior to other estimations.

Table 6. The out-of-sample performance comparison between each estimator of the 41 assets in SSE50.

Index	Average Return *	Standard Deviation *	Sharpe Ratio
Σ_{SCM}	50.40	27.98	1.7390
$\Sigma_{Identity}$	50.40	27.53	1.7399
Σ_L	50.40	28.00	1.7375
Σ_{NL}	50.40	27.95	1.7400
Σ_D	50.40	27.96	2.7399
Σ^*	50.40	27.97	1.7391

* denotes the unit is %.

Table 7. The out-of-sample performance comparison between each estimations of the 218 assets in HS300.

Index	Average Return *	Standard Deviation *	Sharpe Ratio
Σ_{SCM}	50.40	18.69	2.6031
$\Sigma_{Identity}$	50.40	20.28	2.3990
Σ_L	50.40	18.70	2.6017
Σ_{NL}	50.40	17.79	2.7340
Σ_D	50.40	17.81	2.7314
Σ^*	50.40	18.68	2.6044

* denotes the unit is %.

Table 8. The out-of-sample performance comparison between each estimations of the 426 assets in CSI500.

Index	Average Return *	Standard Deviation *	Sharpe Ratio
Σ_{SCM}	50.40	21.25	2.2899
$\Sigma_{Identity}$	50.40	23.04	2.1115
Σ_L	50.40	21.24	2.2909
Σ_{NL}	50.40	21.22	2.2927
Σ_D	50.40	21.18	2.2970
Σ^*	50.40	20.87	2.3315

* denotes the unit is %.

Table 9. The difference in the out-of-sample variance between Σ^* and the alternative estimation with all assets (the significance level is 5%).

Number of Assets	Σ_{SCM}	$\Sigma_{Identity}$	Σ_L	Σ_{NL}	Σ_D
41	-0.0001	0.0350	-0.0019	0.0011	0.0010
218	-0.0000	-0.1607	-0.0010	0.0986	0.0966
426	-0.0363	-0.1979	-0.0350	-0.0342	-0.0303

5. Conclusions

In this paper, we proposed a new estimator for the covariance matrix, which is a convex combination of the linear shrinkage estimation and the rotation-invariant estimator under the F -norm. We first obtained the optimal parameters through considerable numerical operations, and then, we focused on the accuracy of the model and ignored the complexity of the calculation. Moreover, we demonstrated the effectiveness of the model in the simulation. Finally, we applied our estimation to the minimum variance portfolio

optimization and showed that the performance of the proposed estimator is superior to the other five existing estimators in the portfolio optimization for high-dimensional data.

In addition, we only considered the sample noise on the covariance matrix in this article, but in the financial field, the covariance matrix estimation will vary with time due to the influence of market information, and the covariance matrix estimation will be affected by the market information. Therefore, the performance of our new estimation in the dynamic conditional correlation model [41] can be investigated as part of future work for a large-dimensional covariance matrix.

Author Contributions: Conceptualization, Y.Z.; Fund acquisition, G.W. and Z.Y.; methodology, Y.Z. and J.T.; supervision, G.W. and Z.Y.; writing—original draft preparation, Y.Z.; writing—review and editing, Y.Z., J.T. and G.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Nos. 11971302 and 12171307).

Informed Consent Statement: Not applicable.

Data Availability Statement: The data of this paper came from the component stock of CSI500, HS300, and SSE50 on the tushare financial website accessed on 1 July 2021: <https://www.tushare.pro>.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Hu, J.; Bai, Z.D. A review of 20 years of naive tests of significance for high-dimensional mean vectors and covariance matrices. *Sci. China Math.* **2016**, *59*, 2281–2300. [[CrossRef](#)]
- Tong, T.; Wang, C.; Wang, T. Estimation of variances and covariances for high-dimensional data: A selective review. *Wiley Interdiscip. Rev. Comput. Stat.* **2014**, *6*, 255–264. [[CrossRef](#)]
- Engel, J.; Buydens, L.; Blanchet, L. An overview of large-dimensional covariance and precision matrix estimators with applications in chemometrics. *J. Chemom.* **2017**, *31*, e2880. [[CrossRef](#)]
- Sun, R.; Ma, T.; Liu, S.; Sathye, M. Improved covariance matrix estimation for portfolio risk measurement: A review. *J. Risk Financ. Manag.* **2019**, *12*, 48. [[CrossRef](#)]
- Fan, J.; Liao, Y.; Liu, H. An overview of the estimation of large covariance and precision matrices. *Econom. J.* **2016**, *19*, C1–C32. [[CrossRef](#)]
- Rachev, S.T. *Handbook of Heavy Tailed Distributions in Finance*; Elsevier: North Holland, The Netherlands, 2003.
- Yuan, X.; Yu, W.Q.; Yin, Z.X.; Wang, G.Q. Improved large dynamic covariance matrix estimation with graphical lasso and its application in portfolio selection. *IEEE Access* **2020**, *8*, 189179–189188. [[CrossRef](#)]
- Ledoit, O.; Wolf, M. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J. Empir. Financ.* **2003**, *10*, 603–621. [[CrossRef](#)]
- Fan, J.; Fan, Y.; Lv, J. High dimensional covariance matrix estimation using a factor model. *J. Econ.* **2008**, *147*, 186–197. [[CrossRef](#)]
- Xu, F.F.; Huang, J.C.; Wen, Z.W. High dimensional covariance matrix estimation using multi-factor models from incomplete information. *Sci. China Math.* **2015**, *58*, 829–844. [[CrossRef](#)]
- Liu, H.; Han, F.; Yuan, M.; Lafferty, J.; Wasserman, L. High-dimensional semiparametric Gaussian copula graphical models. *Ann. Stat.* **2012**, *40*, 2293–2326. [[CrossRef](#)]
- Stein, C. Lectures on the theory of estimation of many parameters. *J. Sov. Math.* **1986**, *34*, 1373–1403. [[CrossRef](#)]
- Ledoit, O.; Wolf, M. Honey, I shrunk the sample covariance matrix. *J. Portfolio Manag.* **2004**, *30*, 110–119. [[CrossRef](#)]
- Ledoit, O.; Wolf, M. A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivar. Anal.* **2004**, *88*, 365–411. [[CrossRef](#)]
- Ledoit, O.; Wolf, M. Nonlinear shrinkage estimation of large-dimensional covariance matrices. *Ann. Stat.* **2012**, *40*, 1024–1060. [[CrossRef](#)]
- Ledoit, O.; Wolf, M. Optimal estimation of a large-dimensional covariance matrix under stein’s loss. *Bernoulli* **2018**, *24*, 3791–3832. [[CrossRef](#)]
- Ledoit, O.; Wolf, M. The power of (non-) linear shrinking: A review and guide to covariance matrix estimation. *J. Financ. Econ.* **2022**, *20*, 187–218. [[CrossRef](#)]
- Ledoit, O.; Wolf, M. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks. *Rev. Financ. Stud.* **2017**, *30*, 4349–4388. [[CrossRef](#)]
- Bickel, P.J.; Levina, E. Covariance regularization by thresholding. *Ann. Stat.* **2008**, *36*, 2577–2604. [[CrossRef](#)]
- Rothman, A.; Levina, E.; Zhu, J. Generalized thresholding of large covariance matrices. *J. Am. Stat. Assoc.* **2009**, *104*, 177–186. [[CrossRef](#)]

21. Rothman, A.J. Positive definite estimators of large covariance matrices. *Biometrika* **2012**, *99*, 733–740. [[CrossRef](#)]
22. Zhou, S.L.; Xiu, N.H.; Luo, Z.Y.; Kong, L.C. Sparse and low-rank covariance matrix estimation. *J. Oper. Res. Soc. China* **2015**, *3*, 231–250. [[CrossRef](#)]
23. Xue, L.Z.; Ma, S.Q.; Zou, H. Positive-definite l1-penalized estimation of large covariance matrices. *J. Am. Stat. Assoc.* **2012**, *107*, 1480–1491. [[CrossRef](#)]
24. Liu, H.; Wang, L.; Zhao, T. Sparse covariance matrix estimation with eigenvalue constraints. *Comput. Graph. Stat.* **2014**, *23*, 439–459. [[CrossRef](#)] [[PubMed](#)]
25. Wen, F.; Yang, Y.; Liu, P.; Qiu, R.C. Positive definite estimation of large covariance matrix using generalized non-convex penalties. *IEEE Access* **2016**, *4*, 4168–4182. [[CrossRef](#)]
26. Friedman, J.; Hastie, T.; Tibshirani, R. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* **2008**, *9*, 432–441. [[CrossRef](#)]
27. Finegold, M.; Drton, M. Robust graphical modeling of gene networks using classical and alternative t-distributions. *Ann. Appl. Stat.* **2011**, *5*, 1057–1080. [[CrossRef](#)]
28. Yuan, X.M. Alternating direction method for covariance selection model. *J. Sci. Comput.* **2012**, *51*, 261–273. [[CrossRef](#)]
29. Li, P.L.; Xiao, Y.H. An efficient algorithm for sparse inverse covariance matrix estimation based on dual formulation. *Comput. Stat. Data Anal.* **2018**, *128*, 292–307. [[CrossRef](#)]
30. Yuan, M.; Lin, Y. Model selection and estimation in the Gaussian graphical model. *Biometrika* **2007**, *94*, 19–35. [[CrossRef](#)]
31. Yang, J.F.; Sun, D.F.; Toh, K.C. A proximal point algorithm for log-determinant optimization with group lasso regularization. *SIAM J. Optim.* **2013**, *23*, 857–893. [[CrossRef](#)]
32. Bun, J.; Bouchaud, J.-P.; Potters, M. Cleaning large correlation matrices: Tools from random matrix theory. *Phys. Rep.* **2017**, *666*, 1–9. [[CrossRef](#)]
33. Deshmukh, S.; Dubey, A. Improved covariance matrix estimation with an application in portfolio optimization. *IEEE Signal Process. Lett.* **2020**, *27*, 985–989. [[CrossRef](#)]
34. Ledoit, O.; Wolf, M. Quadratic shrinkage for large covariance matrices. *Bernoulli* **2022**, *28*, 1519–1547. [[CrossRef](#)]
35. Donoho, D.; Gavish, M.; Johnstone, I. Optimal shrinkage of eigenvalues in the spiked covariance model. *Ann. Stat.* **2018**, *46*, 1742–1778. [[CrossRef](#)] [[PubMed](#)]
36. Bu, J.; Allez, R.; Bouchaud, J.P.; Potters, M. Rotational invariant estimator for general noisy matrices. *IEEE Trans. Inf. Theory* **2016**, *62*, 7475–7490.
37. Paul, D.; Aue, A. Random matrix theory in statistics: A review. *J. Stat. Plan. Inference* **2014**, *150*, 1–29. [[CrossRef](#)]
38. Haff, L.R. Empirical Bayes estimation of the multivariate normal covariance matrix. *Ann. Stat.* **1980**, *8*, 586–597. [[CrossRef](#)]
39. Markowitz, H. Portfolio Selection. *J. Finance* **1952**, *7*, 77–91.
40. DeMiguel, V.; Garlappi, L.; Nogales, F.J.; Uppal, R. A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Manag. Sci.* **2009**, *55*, 798–812. [[CrossRef](#)]
41. Engle, R.F.; Ledoit, O.; Wolf, M. Large dynamic covariance matrices. *J. Bus. Econom. Stat.* **2019**, *37*, 363–375. [[CrossRef](#)]
42. Bollerslev, T.R.; Engle, R.; Nelson, D. A capital asset pricing model with time varying covariances. *J. Polit. Econ.* **1988**, *96*, 116–131. [[CrossRef](#)]
43. Ledoit, O.; Peche, S. Eigenvectors of some large sample covariance matrix ensembles. *Probab. Theory Relat. Fields* **2011**, *151*, 233–264. [[CrossRef](#)]
44. Ledoit, O.; Wolf, M. Robust performances hypothesis testing with the variance. *Wilmott* **2011**, *2011*, 86–89. [[CrossRef](#)]