Financial Statistical Journal 2024, 7(2), 9912.
https://doi.org/10.24294/fsj9912

Article

Multi-target linear shrinkage estimation of large precision matrix

Yuang Xue, **Lihong Wang***

School of Mathematics, Nanjing University, Nanjing 210093, China
* **Corresponding author:** Lihong Wang, lhwang@nju.edu.cn

CITATION

Xue Y, Wang L. Multi-target linear shrinkage estimation of large precision matrix. Financial Statistical Journal. 2024; 7(2): 9912. https://doi.org/10.24294/fsj9912

ARTICLE INFO

Received: 28 October 2024
Accepted: 22 November 2024
Available online: 5 December 2024

Abstract: In this paper we propose a multi-target linear shrinkage estimator of the precision matrix by shrinking the inverse of the sample covariance matrix directly, which is a generalization of the single-target linear shrinkage estimator. The explicit expression of the weights of multi-target linear shrinkage estimator is derived when the ratio of the dimension p and the sample size n tends to a positive constant $c \in (0, 1)$. The numerical simulation and an empirical analysis of financial market data are provided to compare the multi-target linear shrinkage estimator with some other estimators of the precision matrix proposed in the literature. The computation results show the improvement of the multi-target linear shrinkage estimator.

Keywords: frobenius norm; linear shrinkage estimator; multiple target matrices; precision matrix

1. Introduction

The estimation of a large covariance matrix and its inverse matrix, known as the precision matrix, is central to statistical learning theory and econometrics and has been receiving growing attention from both researchers and practitioners. For example, principal component analysis and factor analysis involve covariance matrix estimation, while Fisher linear discriminant analysis needs precision matrix estimation. Moreover, to implement the mean-variance portfolio in practice, the accuracy in estimating the covariance structure of returns and its precision matrix is vital. As pointed out in Elton et al. [1] and Markowitz [2], a suitable precision matrix estimator leads to a good estimation for different types of optimal portfolios.

Various methods have been proposed to estimate the covariance matrix and the precision matrix in high-dimensional setting from different directions. Bodnar et al. [3] reviewed the recent advances in the shrinkage-based estimation for high-dimensional covariance and precision matrices. See also, e.g., Bodnar et al. [4]; Fan et al. [5]; Fan et al. [6]; Friedman et al. [7]; Ikeda et al. [8]; Kuismin et al. [9]; Ledoit and Wolf [10, 11]; Liu and Tang [12]; van Wieringen and Peeters [13]; Wang et al. [14]; and Wang et al. [15].

For the estimation of precision matrix, one direction is to construct linear shrinkage estimate of the covariance matrix, and then use its inverse as the estimator of the precision matrix. The shrinkage estimator shrinks the eigenvalues of the sample covariance matrix and forms a linear combination of the sample covariance matrix and a pre-chosen target matrix, see, e.g., Dey and Srinivasan [16], Ledoit and Wolf [17, 18]. Another approach is to shrink directly the inverse of the sample covariance matrix itself, instead of shrinking the sample covariance matrix and then inverting it. When both the sample size n and the dimension p tend to infinity but their ratio tends to a positive constant, Bodnar et al. [19] proposed a direct single target shrinkage estimator for the precision matrix.

Most of the linear shrinkage estimates in the literature use single target matrix. However, the performance of the estimator strongly depends on the choice of the target matrix (cf. Engel et al. [20] and Gray et al. [21]). In order to incorporate uncertainty about the target choice, Gray et al. [21] proposed a multi-target linear shrinkage estimator for the covariance matrix that allows for shrinkage of the sample covariance matrix towards multiple targets simultaneously. The multi-target estimator is less sensitive to target misspecification and leads to equal or improved estimates compared to single-target linear shrinkage estimators, see also Bartz et al. [22] and Lancewicki and Aladjem [23]. However, to the best of our knowledge, we have not found any multi-target shrinkage estimation methods for the precision matrix in the literature. The explicit expression of the shrinkage weights and the theoretical properties of the multi-target linear shrinkage estimation of precision matrix, when both the sample size n and the dimension p tend to infinity, are still unavailable.

Inspired by the works of Bodnar et al. [19] and Gray et al. [21], this paper considers the linear shrinkage estimation of the precision matrix based on multiple shrinkage target matrices. This multi-target shrinkage estimator forms a linear combination of the inverse of the sample covariance matrix and multiple general shrinkage target matrices. In Section 2, we derive the explicit expression of the weights of the multi-target linear shrinkage estimator in the case of $p/n \to c \in (0,1)$. In Section 3, numerical simulation and an empirical analysis of portfolio optimization are provided to compare the multi-target linear shrinkage estimator with some other estimators for the precision matrix proposed in the literature. The simulation results show the improvement of the multi-target linear shrinkage estimator. The proofs of the theorems are placed in the appendix.

2. Multi-target linear shrinkage estimator

The following notations are used throughout the paper. Let n be the sample size, $p = p(n)$ is the number of the variables. Σ_n^{-1} stands for the true precision matrix, and $\widehat{\Sigma}_n^{-1}$ denotes the estimator of Σ_n^{-1}. Since the dimension p is a function of the sample size n, we use the subscript n for the covariance matrix Σ_n which depends on n via $p(n)$. Let $H_n(t)$ denote the empirical distribution function of the eigenvalues of Σ_n.

Let Y_n be a $p \times n$ matrix which consists of independent and identically distributed (i.i.d.) random variables with zero mean and unit variance. The observation matrix is defined as $X_n = \Sigma_n^{\frac{1}{2}} Y_n$. S_n denotes the sample covariance matrix, i.e. $S_n = \frac{1}{n} X_n X_n' = \frac{1}{n} \Sigma_n^{\frac{1}{2}} Y_n Y_n' \Sigma_n^{\frac{1}{2}}$.

$\|A\|_F^2 = tr(AA')$ denotes the Frobenius norm of a matrix A, and $\frac{1}{p}\|A\|_F^2$ denotes the normalized Frobenius norm of $p \times p$ matrix A, while $\|A\|_{tr} = tr\left[(AA')^{\frac{1}{2}}\right]$ stands for the trace norm, and $\frac{1}{p}\|A\|_{tr}$ is the normalized trace norm. Define the Frobenius norm loss function as $L_F^2(\widehat{\Sigma}_n^{-1}, \Sigma_n^{-1}) = \|\widehat{\Sigma}_n^{-1} - \Sigma_n^{-1}\|_F^2$.

Bodnar et al. [19] proposed a single-target linear shrinkage estimator for the precision matrix which is a convex combination of the inverse of the sample covariance matrix and a target matrix and minimizes the Frobenius norm loss. The linear shrinkage estimators are motivated from Bayes methods which are developed to nonparametric linear estimation by estimating the regularization parameter ω nonparametrically through the optimal weight. For more details on the justification, motivation and application of such linear shrinkage estimation, we refer to Bodnar et al. [19]; Ikeda et al. [8]; Ledoit and Wolf [17]; and Robbins [24].

One of the disadvantage of the single-target linear shrinkage estimation is that the performance of the shrinkage estimator depends to a great extent on the selection of target matrix. The choice of target should be guided by the presumed structure of the population covariance matrix. There is often no single ideal target, and it is difficult to identify a sensible target. The single-target estimator may be misspecified, because every target has a different bias-variance trade-off with respect to the unknown population covariance matrix. In order to reduce the estimation error caused by the improper selection of the target matrix, we consider using a set of target matrices.

Let $\Pi_1, \cdots, \Pi_k$ be the set of target matrices, where Π_i is symmetric, positive definite and has uniformly bounded normalized Frobenius norm for $i = 1, \cdots, k$. We propose the multi-target linear shrinkage estimator $\widehat{\Sigma}_n^{-1}$ of the precision matrix

$$\widehat{\Sigma}_n^{-1} = (1 - \omega_1 - \cdots - \omega_k) S_n^{-1} + \omega_1 \Pi_1 + \cdots + \omega_k \Pi_k \tag{1}$$

which minimizes the loss function

$$L_F^2(\widehat{\Sigma}_n^{-1}, \Sigma_n^{-1}) = \|(S_n^{-1} - \Sigma_n^{-1}) - \omega_1(S_n^{-1} - \Pi_1) - \cdots - \omega_k(S_n^{-1} - \Pi_k)\|_F^2$$

Here $\Omega = (\omega_1, \cdots, \omega_k)'$ is restricted by $\omega_i \geq 0$ for $i = 1, \cdots, k$ and $\omega_1 + \cdots + \omega_k \leq 1$.

Taking the derivatives of L_F^2 with respect to ω_i, $i = 1, \cdots, k$, and setting them equal to zero,

$$\begin{aligned} &\frac{\partial L_F^2(\widehat{\Sigma}_n^{-1}, \Sigma_n^{-1})}{\partial \omega_i} \\ &= -2\, tr\Big[\big((S_n^{-1} - \Sigma_n^{-1}) - \omega_1(S_n^{-1} - \Pi_1) - \cdots - \omega_k(S_n^{-1} - \Pi_k) \big) \big(S_n^{-1} - \Pi_i \big) \Big] \\ &= 0 \end{aligned}$$

we obtain the optimal weight Ω^* satisfying

$$A_n \Omega^* = B_n \tag{2}$$

where

$$A_n = \frac{1}{p} \begin{pmatrix} tr[(S_n^{-1} - \Pi_1)(S_n^{-1} - \Pi_1)] & \cdots & tr[(S_n^{-1} - \Pi_k)(S_n^{-1} - \Pi_1)] \\ \vdots & \ddots & \vdots \\ tr[(S_n^{-1} - \Pi_1)(S_n^{-1} - \Pi_k)] & \cdots & tr[(S_n^{-1} - \Pi_k)(S_n^{-1} - \Pi_k)] \end{pmatrix}$$

and

$$B_n = \frac{1}{p} \begin{pmatrix} tr[(S_n^{-1} - \Sigma_n^{-1})(S_n^{-1} - \Pi_1)] \\ \vdots \\ tr[(S_n^{-1} - \Sigma_n^{-1})(S_n^{-1} - \Pi_k)] \end{pmatrix}$$

Since S_n^{-1} is a biased estimator of Σ_n^{-1} when $p/n \to c \in (0, 1)$, the shrinkage estimator (1) actually balances the trade-off between bias and variance in an effective way. Thus $\widehat{\Sigma}_n^{-1}$ can be regarded as a maximum likelihood estimator with special ridge penalty terms, where ω_i, $i = 1, \cdots, k$, are the regularization parameters.

To estimate the optimal weight Ω^*, we suggest a method for estimating A_n and B_n consistently. The following lemma provides the limits of $\|S_n^{-1}\|_F^2$ and $tr(S_n^{-1}\Theta)$ for some symmetric positive definite matrix Θ. The proof can be found in Theorem 3.2 of Bodnar et al. [19]. Using this result one can easily obtain the consistent estimator of the weight Ω^*.

Lemma 1. *Let $p/n \to c \in (0, 1)$. Assume that the elements of the matrix Y_n have uniformly bounded $4+\epsilon$ moments, where $\epsilon > 0$, $H_n(t)$ converges to a limit $H(t)$ at all points of continuity of H, and for n large enough there exists the compact interval $[h_0, h_1] \subset (0, +\infty)$ which contains the support of H_n. Then as $n \to \infty$,*

$$\frac{1}{p} \left| \|S_n^{-1}\|_F^2 - \Big(\frac{1}{(1-c)^2} \|\Sigma_n^{-1}\|_F^2 + \frac{c}{p(1-c)^3} \|\Sigma_n^{-1}\|_{tr}^2 \Big) \right| \xrightarrow{a.s.} 0 \tag{3}$$

and for symmetric positive definite matrix Θ which has uniformly bounded trace norm,

$$\left| tr(S_n^{-1}\Theta) - \frac{1}{1-c} tr(\Sigma_n^{-1}\Theta) \right| \overset{a.s.}{\to} 0 \tag{4}$$

Throughout the paper we assume that the conditions of Lemma 1 are satisfied. These conditions also ensure that Σ_n^{-1} has uniformly bounded normalized Frobenius norm and normalized trace norm. Note that we need the condition $c < 1$ to keep the sample covariance matrix S_n invertible. The case of $c > 1$ is very difficult to handle because of the loss of information as the dimension p is greater than the sample size n (cf. Bodnar et al. [19]). Although the estimator S_n^{-1} can be replaced by the generalized inverse matrix of the sample covariance matrix, it is not clear how to estimate the optimal weight consistently. Since the theory is yet to be developed fully for the $c > 1$ case, we leave it for future research.

Let $A_0 = (a_{0ij})_{k\times k}$ and $B_0 = (b_{0i})_k$, where

$$\begin{aligned} a_{0ij} = {} & \frac{1}{p(1-c)^2}\|\Sigma_n^{-1}\|_F^2 + \frac{c}{p^2(1-c)^3}\|\Sigma_n^{-1}\|_{tr}^2 - \frac{1}{p(1-c)} tr[\Sigma_n^{-1}(\Pi_i + \Pi_j)] \\ & + \frac{1}{p} tr(\Pi_i \Pi_j) \end{aligned}$$

and

$$b_{0i} = \frac{c}{p(1-c)^2}\|\Sigma_n^{-1}\|_F^2 + \frac{c}{p^2(1-c)^3}\|\Sigma_n^{-1}\|_{tr}^2 - \frac{c}{p(1-c)} tr(\Sigma_n^{-1}\Pi_i)$$

Theorem 1 shows the non-random limit of A_n and B_n.

Theorem 1. *Under the conditions of Lemma 1, we have as $n \to \infty$,*

$$A_n - A_0 \overset{a.s.}{\to} 0, \ \ B_n - B_0 \overset{a.s.}{\to} 0$$

in the sense that each element converges almost surely.

Theorem 1 implies that the asymptotic optimal weight vector $\Omega_0 = (\omega_{01}, \cdots, \omega_{0k})'$ satisfies

$$A_0 \Omega_0 = B_0 \tag{5}$$

From Equation (5), we have $\Omega_0 = 0$ and then Ω^* tends to 0 in the case of $c = 0$, which means that the inverse of the sample covariance matrix is an asymptotically best estimator for the precision matrix in terms of minimizing the Frobenius norm loss. In contrast if p increases, the sample covariance matrix becomes ill-conditioned and hence the linear shrinkage estimator (1) improves the performance of the sample estimator. The impact of this improvement becomes larger as p approaches n, i.e. as c approaches 1. In this case, each element of Ω_0 tends to $1/n$ and hence Ω^* approaches $(n^{-1}, \cdots, n^{-1})'$.

By Equation (5), to derive the a.s. consistent estimator of Ω_0, it suffices to construct $\widehat{A}_0$ and $\widehat{B}_0$, the a.s. consistent estimators of A_0 and B_0, which are provided in the following Theorem 2.

Theorem 2. *Under the conditions of Lemma 1, the a.s. consistent estimators of A_0 and B_0 are given by*

$$\widehat{A}_0 = A_n, \ \ \widehat{B}_0 = \frac{1}{p}\begin{pmatrix} \frac{p}{n}\|S_n^{-1}\|_F^2 + \frac{1}{n}\|S_n^{-1}\|_{tr}^2 - \frac{p}{n}\, tr(S_n^{-1}\Pi_1) \\ \vdots \\ \frac{p}{n}\|S_n^{-1}\|_F^2 + \frac{1}{n}\|S_n^{-1}\|_{tr}^2 - \frac{p}{n}\, tr(S_n^{-1}\Pi_k) \end{pmatrix}$$

Assume that $\widehat{A}_0$ is invertible. Then we get the estimator $\widetilde{\Omega}$ of the asymptotic optimal weight Ω_0

$$\widetilde{\Omega} = (\tilde{\omega}_1, \cdots, \tilde{\omega}_k)' = \widehat{A}_0^{-1}\widehat{B}_0 \quad (6)$$

Since Ω is restricted on $\omega_i \geq 0$ for $i = 1, \cdots, k$ and $\omega_1 + \cdots + \omega_k \leq 1$, it is reasonable to define the a.s. consistent estimators of the shrinkage intensities as follows.

Rewrite the multi-target linear shrinkage estimator (1) as

$$\widehat{\Sigma}_n^{-1} = (1-\beta)S_n^{-1} + \beta\big\{(1-\alpha_1-\cdots-\alpha_{k-1})\Pi_1 + \alpha_1\Pi_2 + \cdots + \alpha_{k-1}\Pi_k\big\}$$

where β and α_i, $i = 1, \cdots, k-1$ satisfy that $0 \leq \alpha_i,\ \beta \leq 1$, and

$$\omega_1 + \cdots + \omega_k = \beta,\ \ \omega_1 = (1-\alpha_1-\cdots-\alpha_{k-1})\beta,\ \ \omega_i = \alpha_{i-1}\beta \quad (7)$$

for $i = 2, \cdots, k$. Then we construct the estimators of $(\alpha_1, \cdots, \alpha_{k-1}, \beta)$ using Equation (6) and the relation Equation (7), namely

$$\tilde{\beta} = \tilde{\omega}_1 + \cdots + \tilde{\omega}_k,\ \ \tilde{\alpha}_i = \tilde{\omega}_{i+1}/\tilde{\beta},\ \ i = 1, \cdots, k-1$$

Note that $0 \leq \alpha_i,\ \beta \leq 1$, we estimate β and α_i, $i = 1, \cdots, k-1$, by $\hat{\beta} =$ max $\big(0, \min(1, \tilde{\beta})\big)$ and $\hat{\alpha}_i =$ max $\big(0, \min(1, \tilde{\alpha}_i)\big)$, $i = 1, \cdots, k-1$. Using Equation (7) again, we obtain the estimators

$$\hat{\omega}_1 = (1-\hat{\alpha}_1-\cdots-\hat{\alpha}_{k-1})\hat{\beta},\ \ \hat{\omega}_i = \hat{\alpha}_{i-1}\hat{\beta},\ \ i = 2, \cdots, k$$

Now the genuine multi-target linear shrinkage estimator of the precision matrix is given by

$$\widehat{\Sigma}_n^{-1} = (1-\hat{\omega}_1-\cdots-\hat{\omega}_k)S_n^{-1} + \hat{\omega}_1\Pi_1 + \cdots + \hat{\omega}_k\Pi_k \quad (8)$$

Theorems 1 and 2 immediately imply the estimator $\widehat{\Omega} = (\hat{\omega}_1, \cdots, \hat{\omega}_k)'$ converges almost surely to the asymptotic weight vector Ω_0 in Equation (5), and thus converges almost surely to its oracle optimal intensity Ω^* in Equation (2) as $n \to \infty$. This result is presented in the following Theorem 3, which implies that the multi-target linear shrinkage estimator in Equation (8) performs as well as its oracle one.

Theorem 3. *Assume that $\hat{A}_0$ is invertible. Then, under the conditions of Lemma 1, as $n \to \infty$,*

$$\widehat{\Omega} - \Omega^* \xrightarrow{a.s.} 0$$

Let

$$A_n^0 = \frac{1}{p}\begin{pmatrix} tr[(\Sigma_n^{-1}-\Pi_1)(\Sigma_n^{-1}-\Pi_1)] & \cdots & tr[(\Sigma_n^{-1}-\Pi_k)(\Sigma_n^{-1}-\Pi_1)] \\ \vdots & \ddots & \vdots \\ tr[(\Sigma_n^{-1}-\Pi_1)(\Sigma_n^{-1}-\Pi_k)] & \cdots & tr[(\Sigma_n^{-1}-\Pi_k)(\Sigma_n^{-1}-\Pi_k)] \end{pmatrix}$$

The following theorem shows that the shrinkage intensities tend almost surely to zero when $p/n \to c = 0$ as $n \to \infty$. This implies that the multi-target linear shrinkage estimator is asymptotically equivalent to the sample estimator. Bai and Shi [25] showed that the sample precision matrix is consistent in the case when $p/n \to 0$. This implies that the multi-target linear shrinkage estimator $\widehat{\Sigma}_n^{-1}$ is a consistent estimator for the precision matrix in this case.

Theorem 4. *Assume that A_n^0 is invertible and $p = o(n)$. Then, as $n \to \infty$,*

$$\widehat{\Omega} \overset{a.s.}{\to} 0$$

Regarding the choice of the target matrices, selecting suitable targets require some diligence. It is worth mentioning that the theoretical results derived in Theorems 1–4 are based on the assumption that the target matrices are non-random. However, the data driven target matrices may result in more accurate estimators, as these target matrices provide more information about the structure of the precision matrix. Therefore, in practice, one can use the estimators $\hat{\Pi}_1, \cdots, \hat{\Pi}_k$ of the targets to construct the multi-target linear shrinkage estimator of the precision matrix. The theoretical investigation on the properties of the estimator when using the estimated targets is not yet available. This topic will be pursued in the future research.

In the absence of prior information, the nine target matrices described in **Table 1** in next section can be included due to their popularity in the literature (e.g. Gray et al. [21]). This is illustrated in Section 3 using simulations and a real data example. It is also possible to further enrich the target set with any covariance structures not listed in **Table 1**. Examples include Toeplitz, higher-order autoregressive, or latent factor structures, see, e.g., Chen [26] and Ledoit and Wolf [27].

An overview was given in Schäfer and Strimmer [28], where six types of commonly used targets were proposed. These targets are included in the following **Table 1**, denoted as $\hat{\Pi}_1$ (diagonal, unit variance), $\hat{\Pi}_2$ (diagonal, common variance), $\hat{\Pi}_3$ (diagonal, unequal variance), $\hat{\Pi}_4$ and $\hat{\Pi}_5$ (common (co)variance), $\hat{\Pi}_6$ (unequal variance, constant correlation), and $\hat{\Pi}_7$, $\hat{\Pi}_8$ and $\hat{\Pi}_9$ (decaying correlation). In the literature it is easy to find examples where one of the above targets is employed, see, e.g., Dobra et al. [29]; Friedman [30]; Hastie et al. [31]; and Ledoit and Wolf [27,32].

3. Numerical and empirical studies

The purpose of this section is to compare the performance of the proposed approach with existing ones. We also apply the proposed multi-target linear shrinkage estimator for an empirical analysis of financial market data.

3.1. Simulation study

In this subsection we investigate the numerical performance of the proposed estimator through simulation. We generate data as follows. Let $d_i = 0.1 + 10 \times U_i,\ U_i$ is generated from the uniform distribution on the interval (0, 1). Let σ_{ij} be the (i, j)th element of the true covariance matrix Σ. We consider six types of covariance structures:

(Model 1) $\sigma_{ij} = d_i d_j \rho^{|i-j|}$,

(Model 2) $\sigma_{ij} = \rho^{|i-j|}$,

(Model 3) $\sigma_{ij} = d_i d_j \{|i-j+1|^{2h} - 2|i-j|^{2h} + |i-j-1|^{2h}\}/2$,

(Model 4) $\sigma_{ij} = \{|i-j+1|^{2h} - 2|i-j|^{2h} + |i-j-1|^{2h}\}/2$,

(Model 5) $\Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix}$ is a block diagonal matrix, where the elements of Σ_1 and Σ_2 are from (Model 1) and (Model 3), respectively,

(Model 6) Σ is a block diagonal matrix similar to (Model 5), where the elements of the two blocks are from (Model 2) and (Model 4), respectively.

(Model 2) and (Model 4) correspond to the covariance structures of an autoregressive process and a fractional Brownian Motion, respectively. (Model 5) and (Model 6) represent more complex covariance structures. The data $x_1, ..., x_n$ are generated by $x_i = \Sigma^{\frac{1}{2}} y_i$, where $y_i = (y_{1i}, ..., y_{pi})'$. $y_{1i}, ..., y_{pi}$ are mutually independently distributed as

(Case 1) $y_{ji} \sim N(0,1)$,
(Case 2) $y_{ji} \sim \sqrt{(m-2)/m} z_{ji},\ z_{ji} \sim t_m$,
(Case 3) $y_{ji} = (z_{ji} - \tau)/\sqrt{2\tau},\ z_{ji} \sim \chi^2_\tau$,
where t_m is a t-distribution with m degrees of freedom and χ^2_τ is a chi-square distribution with τ degrees of freedom. (Case 2) and (Case 3) include heavy-tailed and skewed distributions, respectively.

Gray et al. [21] listed nine popular target matrices in the literature. We choose them as the target matrix set of our multi-target linear shrinkage estimator for the precision matrix, see **Table 1** for the set of target matrices.

Table 1. Target matrices.

correlation	**zero($r_{ij} = 0$)**	**constant ($r_{ij} = \bar{r}$)**	**decaying($r_{ij} = \bar{r}^{\lvert i-j \rvert}$)**
unit variance ($\nu_i = 1$)	$\hat{\Pi}_1$	$\hat{\Pi}_4$	$\hat{\Pi}_7$
common variance ($\nu_i = \bar{s}$)	$\hat{\Pi}_2$	$\hat{\Pi}_5$	$\hat{\Pi}_8$
unequal variances ($\nu_i = s_{ii}$)	$\hat{\Pi}_3$	$\hat{\Pi}_6$	$\hat{\Pi}_9$

Here the target matrix $\hat{\Pi} = V^{\frac{1}{2}} R V^{\frac{1}{2}}$, with $V = diag(\nu_1, ..., \nu_p)$ a diagonal matrix, $R = (r_{ij})_{p\times p}$ a correlation matrix. s_{ij} denotes the (i,j)th element of the sample covariance matrix S_n, while $\bar{s},\ \bar{r}$ are the averages of the sample variances and correlations, respectively.

To assess the performance of the estimators, we calculate the PRIAL (Percentage Relative Improvement in Average Loss) presented in Ledoit and Wolf [18]. The PRIAL indicates the extent to which the MSE of the estimator $\widehat{\Sigma}_n^{-1}$ outperforms in percentage terms the MSE of the sample precision matrix. Let $\widehat{\Sigma}_n^{-1}$ be an estimator of the precision matrix, the PRIAL is defined by

$$\text{PRIAL}(\widehat{\Sigma}_n^{-1}) = \left(1 - \frac{\mathrm{E}\|\widehat{\Sigma}_n^{-1} - \Sigma_n^{-1}\|_F^2}{\mathrm{E}\|S_n^{-1} - \Sigma_n^{-1}\|_F^2}\right) \times 100\%$$

which actually measures the relative improvement of an estimator over the sample precision matrix. Thus, the improvement of the shrinkage estimator over the sample precision matrix will be measured by how closely this estimator approximates Σ_n^{-1} relative to the sample precision matrix. The PRIAL being closer to 100% indicates the stronger improvement of an estimator.

In order to investigate the performance of the suggested multi-target linear shrinkage (MULTI) estimator for the precision matrix, we introduce three benchmark estimators. Since there exist extensive comparisons of the single-target linear shrinkage estimators and other estimators given in the statistical literature, see, e.g. Bodnar et al. [19] and Ikeda et al. [8], here we only compare the multi-target estimator with the single-target estimators.

The first estimator considered is the nonlinear shrinkage estimator "EV" proposed and studied in Ledoit and Wolf [10,18,33], which is defined as

$$\widehat{\Sigma}_{EV}^{-1} = UA^*U'$$

where U is an orthogonal matrix whose columns are the eigenvectors of the sample covariance matrix S_n, A^* is a diagonal matrix whose elements are real univariate functions which depend on S_n. The exact formula of A^* can be found in (4.3) of Ledoit and Wolf [18].

The second one is based on a single-target linear shrinkage estimator (OLSE) of the covariance matrix provided by Bodnar et al. [34], then the OLSE estimator of the precision matrix is given by

$$\widehat{\Sigma}_{OLSE}^{-1} = \left(\widehat{\Sigma}_{OLSE}\right)^{-1},\ \ \widehat{\Sigma}_{OLSE} = \alpha S_n + \beta \Pi$$

where $\alpha = 1 - \frac{\frac{1}{n}\|S_n\|_{tr}^2 \|\Pi\|_F^2}{\|S_n\|_F^2\|\Pi\|_F^2 - (tr(S_n\Pi))^2}$, $\beta = \frac{tr(S_n\Pi)}{\|\Pi\|_F^2}(1-\alpha)$. The third estimator is the single-target linear shrinkage (ONE) estimator with $k = 1$ in Equation (1).

For both linear shrinkage estimators, we choose the target matrix from **Table 1**. In simulations we tried each of the shrinkage targets in **Table 1**. Unsurprisingly, the performance of both estimators varies across the different scenarios and depends on the choice of shrinkage target. Overall, the single-target estimators with the best choice of target matrix does not outperform the suggested multi-target estimator which does not need to choose the suitable shrinkage target. It appears that the proposed multi-target estimator is less sensitive to misspecification of the targets.

Now we report the comparison of the performance of the estimators EV, OLSE, ONE and MULTI. Since the comparisons show no significant difference among the nine shrinkage targets in **Table 1**, we only present the results using the target matrix $\hat{\Pi}_1$ for OLSE and ONE estimators.

Table 2 summarises the simulation results with $m = 5$, $\tau = 3$, $p = 100$, $\rho = 0.2, 0.8$, $h = 0.8$, and $n = 120, 200, 300, 500$, respectively. The PRIALs are approximated by an average over 1000 simulation runs for each scenario.

It is observed that the multi-target shrinkage estimator performs better than the other three methods for all six models in almost all scenarios, except for the case of $n = 120$, where MULTI and EV have similar performance. Moreover, we conclude that, for all estimators in each scenario, the relative improvement measure PRIAL decreases when the sample size n increases. This shows that, when the dimension p is fixed, the performance of the sample estimator improves as the sample size increases. It is also seen that even if the sample size n becomes larger, the PRIAL of MULTI estimator is still greater than 95.5%, better than EV estimator and much better than the other two estimators. In general, EV estimator outperforms OLSE and ONE estimators, while the performance of MULTI estimator is the best and most stable one.

On the other hand, the computation results show that the performance of OLSE and ONE estimators differs greatly for different covariance matrix structures. This reveals that the correct selection of the shrinkage target is crucial to OLSE and ONE methods. In contrast, the MULTI estimator achieves a similar performance for all models without having to choose the correct shrinkage target, especially for (Model 5) and (Model 6) with complex covariance structures. This demonstrates the robustness of the proposed multi-target method in estimating the complicated precision matrices compared to single-target shrinkage.

Next, we consider the case of fixing $c = p/n = 0.2, 0.4, 0.6, 0.8$, where n takes values in $\{50 + 20m, m \in \mathbb{N}\} \cap [50, 500]$. **Figures 1** and **2** depict the PRIALs of EV, OLSE, ONE and MULTI estimators for different c values under two covariance matrix structures, respectively.

It can be seen from **Figure 1** that the multi-target linear shrinkage estimator has the best performance for sample size $n \geq 90$. When the sample size is less than 90, MULTI still outperforms other estimators except for the case of $c = 0.8$. This highlights the multi-target estimator's relative benefit in small sample scenarios and provides insights for its applications with severe data constraints. Moreover, with the increase of sample size n, the PRIAL of MULTI estimator increases continuously. The PRIAL of OLSE estimator is unstable and having sudden drops. For (Model 2), the performance of ONE estimator is the worst, which may be caused by improper selection of shrinkage target. This again shows the importance of using multiple target matrices in the estimation procedure. Compared with the EV, OLSE and ONE estimators, the performance of MULTI estimator is the best and the PRIALs are the most stable. The PRIALs of the EV, OLSE and ONE estimators have relatively large volatility when $c = 0.2$.

Figure 2 further shows the superiority of multi-target linear shrinkage estimator for sample size $n \geq 90$. The simulation results indicate that three direct shrinkage procedures outperform the OLSE method for (Model 4). It is noted that the PRIAL of the OLSE estimator becomes smaller as the sample size n increases. The reason may be that the OLSE estimator needs to first estimate the covariance matrix and then invert the estimator of the covariance matrix to obtain the estimator of the precision matrix, while the MULTI, EV and ONE estimators directly shrink the sample precision matrix.

Table 2. PRIALs of EV, OLSE, ONE and MULTI estimators for six models based on 1000 replications.

		Model 1				**Model 2**			
	$\boldsymbol{n}$	**EV**	**OLSE**	**ONE**	**MULTI**	**EV**	**OLSE**	**ONE**	**MULTI**
$N(0,1)$ $\rho=0.2$	120	100	100	99.7	99.9	100	100	100	100
	200	98.9	98.7	98.7	99.8	98.9	98.8	98.7	99.7
	300	95.8	95.4	95.4	99.7	95.7	95.4	95.2	99.6
	500	89	87.9	87.8	99.2	88	87.1	86.7	99
t_5 $\rho=0.2$	120	100	100	100	100	100	100	99.9	100
	200	99.1	99.1	99.1	99.1	98.9	98.7	98.7	99.3
	300	95.9	95.5	95.4	98.8	95.9	95.6	95.5	98.6
	500	89.3	88.1	87.7	97.1	89.4	88.6	88.1	97.8
$N(0,1)$ $\rho=0.8$	120	99.9	99.8	99.5	99.5	99.9	99.7	99.2	99.9
	200	96.9	94.8	91.1	99.7	96.8	94.7	90.5	99.9
	300	92.3	88.1	79.8	99.5	91.3	86.8	78.2	99.7
	500	81.8	73.6	61.3	99.3	80.9	72.3	60.1	99.2
t_5 $\rho=0.8$	120	99.9	99.9	99.6	100	99.9	99.8	99.6	100
	200	97.3	95.6	92	99.6	97.7	96.5	92.7	99.7
	300	92.5	88.7	79.6	98.7	93.3	89.7	81	98.9
	500	81.3	72.5	59.5	98.6	82.4	75.1	63.7	99
		Model 3				Model 4			
	n	EV	OLSE	ONE	MULTI	EV	OLSE	ONE	MULTI
$N(0,1)$	120	100	99.8	99.9	100	100	99.9	99.7	99.8
	200	98.1	92.5	95.7	99.8	97.8	92.3	95.6	99.9
	300	94	77.9	89.2	99.4	93.3	77.6	87.8	99.5
	500	83.7	48.8	71.5	98.2	83.8	50.9	74.1	98.6
χ^2_3	120	99.9	99.7	99.8	100	100	99.8	99.9	100
	200	98.1	90.7	96	99.5	98.1	92.1	95.9	99.5
	300	93.7	73.5	88.4	98.7	94.2	76.1	88.6	98.7
	500	84.7	53.9	73.2	97.8	83.7	53.8	74.1	98
		Model 5				Model 6			
	n	EV	OLSE	ONE	MULTI	EV	OLSE	ONE	MULTI
$N(0,1)$ $\rho=0.2$	120	97.5	89.8	89.2	97.8	97.2	87.3	88.9	97.3
	200	95.1	83.5	84.6	97.5	95.8	83.4	85.2	97.1
	300	90.1	69.7	78.2	97	92.6	70.5	75.9	96.7
	500	81.7	44.5	60.7	96.9	80.8	46	63.1	96.2
t_5 $\rho=0.2$	120	97.2	88.8	89	96.8	97	86.8	87.6	97
	200	94.9	81.7	84.9	96.5	94.6	82.5	82.4	96.6
	300	90.8	65.1	73.9	96.2	91.9	70.1	72.5	96.5
	500	81.2	50.1	55.3	95.5	77.3	44.9	55.4	95.8

Generally speaking, the above simulation evidence reveals that the proposed multi-target linear shrinkage estimator has better performance than the single-target linear shrinkage estimators. It is a great alternative to the existing estimation methods.

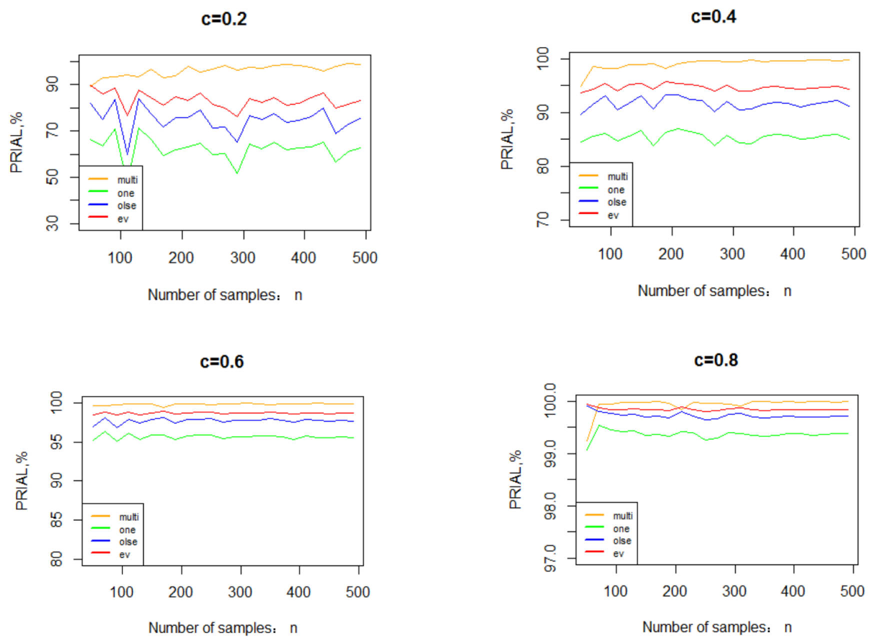


Figure 1. The PRIALs for (Model 2) with ρ=0.8 and $y \sim t_5$.

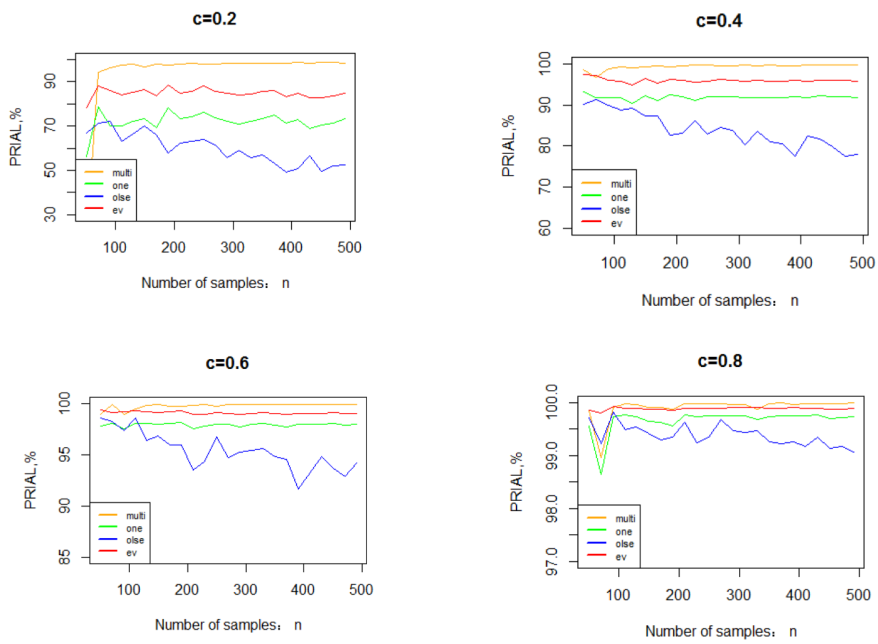


Figure 2. The PRIALs for (Model 4) with $h = 0.8$ and $y \sim N(0, 1)$.

3.2. Empirical study

The ground breaking work of Markowitz [2]—the mean-variance efficient portfolio theory is one of the key tools for portfolio management. However, one needs to know the unobservable covariance matrix and precision matrix to implement this framework. Therefore, it is of vital importance to construct a high-performance precision matrix estimator.

In this subsection we apply the estimator of precision matrix to portfolio optimization problems. Namely, we use the proposed multi-target linear shrinkage method to estimate the weights of a portfolio. We consider a portfolio with p stocks. Denote the expected return of p stocks as $\mu = (\mu_1, ..., \mu_p)'$, the covariance matrix as Σ. Let $\gamma = (\gamma_1, ..., \gamma_p)'$ be the weight of the portfolio, then the return of the portfolio is $\mathrm{E}(R_p) = \gamma'\mu$, the risk (variance) is $\mathrm{Var}(R_p) = \gamma'\Sigma\gamma$. Let $\mathbf{1} = (1, \cdots, 1)'_p$.

The following two popular models are applied to find the efficient frontier of the portfolio (see, e.g., Amenc and Sourd [35]; Cai et al. [36]; Ding et al. [37]; and Joo and Park [38]):

(I). Global minimum variance portfolio

$$\min_{\gamma} \mathrm{Var}(R_p), \text{ subject to } \gamma'\mathbf{1} = 1$$

(II). Maximum expected return portfolio with fixed risk σ_p^2

$$\max_{\gamma} \mathrm{E}(R_p), \text{ subject to } \mathrm{Var}(R_p) = \sigma_p^2, \;\; \gamma'\mathbf{1} = 1$$

The solution to Model (I) is

$$\gamma_1^* = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}$$

From Amenc and Sourd [35], the solution of Model (II) is $\gamma_2^* = (\gamma_{21}^*, ..., \gamma_{2p}^*)'$, where

$$\gamma_{2i}^* = \frac{\mathrm{E}(R_p)\sum_{j=1}^{p} \nu_{ij}(C\mu_j - A) + \sum_{j=1}^{p} \nu_{ij}(B - A\mu_j)}{BC - A^2}, \quad i = 1, ..., p \tag{9}$$

ν_{ij} is the (i, j)th element of Σ^{-1}, and

$$A = \sum_{i=1}^{p}\sum_{j=1}^{p} \nu_{ij}\mu_j, \;\; B = \sum_{i=1}^{p}\sum_{j=1}^{p} \nu_{ij}\mu_i\mu_j, \;\; C = \sum_{i=1}^{p}\sum_{j=1}^{p} \nu_{ij},$$

$$\mathrm{E}(R_p) = \frac{A}{C} + \frac{1}{C}\sqrt{(BC - A^2)C\big(\sigma_p^2 - \frac{1}{C}\big)}$$

We now consider the portfolio management for the stocks in the China CSI Smallcap 500 index which is a prominent benchmark that measures the performance of 500 mid and small-cap A-share stocks listed on the Shanghai and Shenzhen Stock Exchanges. In order to build an efficient portfolio, we use $p = 120$ randomly selected stocks from the components of CSI 500 index with a relatively large total market value and a high turnover ratio. Specifically, the data set contains the daily closing price of the selected 120 stocks from 21 June 2017 to 20 December 2019 with $n = 613$ observations.

The data are divided into three sets: the training set (the first 240 observations), the validation set (the next 240 observations) and the test set (the last 133 observations). The training set is used to estimate the precision matrix, the validation set is utilized to determine the optimal portfolio weight γ^*, while the test set is applied to evaluate the return and risk of the portfolios under different methods.

The parameter μ is calculated from the average return of each stock, Σ^{-1} is estimated by five methods: the sample precision matrix (S^{-1}), EV estimator, OLSE estimator, single-target linear shrinkage (ONE) estimator and the multi-target linear shrinkage (MULTI) estimator, respectively. Among the 120 stocks, the unequal variances and correlations are dominant structures. Based on the data characteristics, we compute two MULTI estimators, one with the total nine targets and the other one (MULTI_6) obtained when using only six targets $\hat{\Pi}_1$, $\hat{\Pi}_3$, $\hat{\Pi}_6$, $\hat{\Pi}_7 - \hat{\Pi}_9$.

Using the estimators $\hat{\mu}$ and $\hat{\Sigma}^{-1}$, the portfolio weight vectors for both models can be calculated by $\hat{\gamma}_1^* = \frac{\hat{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\hat{\Sigma}^{-1}\mathbf{1}}$, and $\hat{\gamma}_2^* = (\hat{\gamma}_{21}^*, ..., \hat{\gamma}_{2p}^*)'$, where $\hat{\gamma}_{2i}^*$, $i = 1, ..., p$, are obtained by plugging in these estimators in Equation (9). Now one can compute the expected return and the risk of a portfolio as well as the CV, where CV is the coefficient of variation, which reflects the ratio of the square root of the risk to the expected return. Obviously, a smaller value of CV means the better risk-return trade-off of the portfolio.

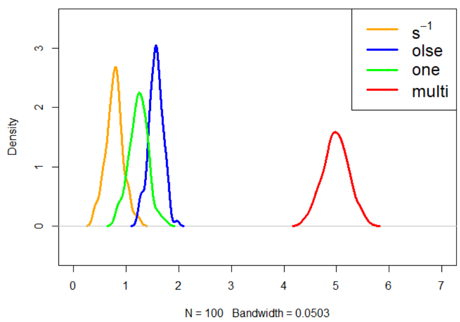
The results of the portfolio obtained by Model (I) and Model (II) are shown in **Table 3**. **Table 3** indicates that in Model (I), the performance of MULTI estimator is the best with the maximum return, minimum risk and lowest CV. MULTI_6 achieves a similar result. EV and ONE estimators have similar performance. The return obtained by S^{-1} is the smallest, while the risk and CV of OLSE estimator are larger than S^{-1}, EV and ONE methods. Meanwhile, in Model (II), when the investment risk is fixed to be equal to 0.1, the MULTI estimators still result in the highest expected return and smallest CV. Moreover, the S^{-1}, EV and ONE estimators have similar results, but the OLSE estimator is the worst.

In both cases, MULTI and MULTI_6 have very similar performance. This highlights the key strength of the proposed multi-target estimator that it is less sensitive to misspecification of the targets.

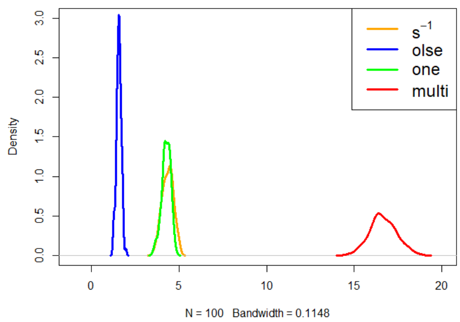
Table 3. The results of the portfolio obtained by Models (I) and (II).

		S^{-1}	EV	OLSE	ONE	MULTI_6	MULTI
Model (I)	return	0.9	1.4	1.8	1.3	4.9	5
	risk	2.3×10^{-4}	5.5×10^{-4}	7.4×10^{-3}	5.2×10^{-4}	1.7×10^{-4}	1.7×10^{-4}
	CV	1.7×10^{-2}	1.7×10^{-2}	4.8×10^{-2}	1.8×10^{-2}	2.7×10^{-3}	2.6×10^{-3}
Model (II)	return	42.4	55	17.5	48.2	133.9	135.1
(with risk = 0.1)	CV	7.5×10^{-3}	5.8×10^{-3}	1.8×10^{-2}	6.6×10^{-3}	2.4×10^{-3}	2.3×10^{-3}

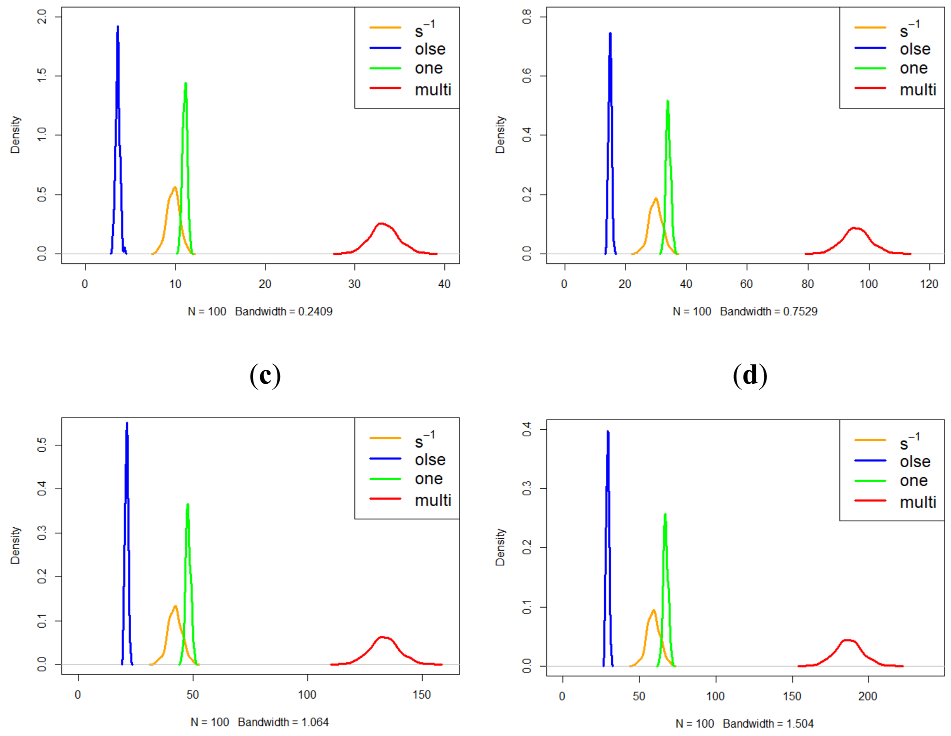
In order to further evaluate the performance of the estimators, we randomly divide 613 observations into the training set, validation set and test set with sample size of 240, 240 and 133, respectively. Since the performance of EV and ONE are quite similar, and we are more interested in comparing the single- and multi-target linear shrinkage methods, we use only OLSE, ONE and MULTI estimators together with the sample precision matrix to calculate the return and risk of the portfolios. The procedure is repeated 100 times and the simulation results are recorded by **Figure 3** which shows the density functions of portfolio returns in 100 replications using S^{-1}, OLSE, ONE and MULTI estimators under Model (I) and Model (II) with the investment risk changing from 0.001 to 0.2, respectively. In **Figure 3a** represents the density for Model (I), **Figure 3b–f** represent the densities for Model (II) with risk = 0.001, 0.005, 0.05, 0.1 and 0.2, respectively.



(a)



(b)



(c) (d) (e) (f)

Figure 3. The density functions of portfolio returns with 100 realizations.

The simulation evidence in **Figure 3** illustrates that the greater the investment risk, the larger the return, and for both models, the return of the portfolio obtained by using MULTI estimator is significantly higher than the other estimators. For model (I), the OLSE estimator is superior to S^{-1} and ONE estimator. For model (II), the performance of the OLSE estimator is poor. As the risk increases, the ONE estimator performs better than the sample precision matrix. Overall, the proposed multi-target linear shrinkage estimator has superior performance in all cases. Our findings show that the multi-target shrinkage approach is quite useful for reducing the estimation errors of the precision matrix and increasing the performances of the portfolios. The proposed method yields more accurate portfolio weights than those of other methods, resulting in higher returns and lower risks.

4. Conclusions

A new estimation method for the precision matrix is considered in this paper. The multi-target linear shrinkage estimator by shrinking the inverse of the sample covariance matrix directly is proposed. This approach generalizes single-target shrinkage methods by allowing the estimator to incorporate multiple targets. The estimator is applied to the simulated data and a financial market dataset, and compared with several existing estimators. The computation results show the improvement of the multi-target linear shrinkage estimator particularly for high-dimensional problems where choosing a single target matrix might limit performance.

It is clear that a careful analysis on determining the optimal number and the optimal choice of the target matrices can greatly help in improving the performance of the estimator. Designing the adaptive methods that automatically select optimal targets based on empirical data characteristics would be an interesting research topic to explore in the future. Another open question worth pursuing further is to investigate the multi-target OLSE method which should also have some nice properties. The multi-target OLSE procedure first creates the multi-target linear shrinkage for the sample covariance matrix and thereafter inverts it to obtain the estimator of the precision matrix. The research of this topic is ongoing.

Author contributions: Conceptualization, YX and LW; methodology, YX and LW ; software, YX; validation, YX and LW; formal analysis, YX and LW; investigation, YX and LW; resources, YX and LW; data curation, YX; writing—original draft preparation, YX and LW; writing—review and editing, LW; visualization, YX; supervision, LW; project administration, LW; funding acquisition, LW. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by National Natural Science Foundation of China (NSFC) under Grant 11671194.

Conflict of interest: The authors declare no conflict of interest.

References

1. Elton EJ, Gruber MJ, Brown SJ, Goetzmann WN. Modern Portfolio Theory and Investment Analysis. Wiley Publishing; 2009.
2. Markowitz H. Portfolio selection. J. Finance. 1952; 7: 77–91.
3. Bodnar O, Bodnar T, Parolya N. Recent advances in shrinkage-based high-dimensional inference. J. Multivar. Ana. 2022; 188: 104826.
4. Bodnar T, Okhrin Y, Parolya N. Optimal shrinkage based portfolio selection in high dimensions. J. Bus. Econ. Stat. 2023; 41(1): 140–156.
5. Fan J, Liao Y, Liu H. An overview of the estimation of large covariance and precision matrices. Econom. J. 2016; 19: C1–C32.
6. Fan J, Liao Y, Mincheva M. High-dimensional covariance matrix estimation in approximate factor models. Ann. Stat. 2017; 39: 3320–3356.
7. Friedman JH, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. Biostatistics. 2008; 9: 432–441.
8. Ikeda Y, Kubokawa T, Srivastava MS. Comparison of linear shrinkage estimators of a large covariance matrix in normal and non-normal distributions. Comput. Statist. Data Anal. 2016; 95: 95–108.
9. Kuismin MO, Kemppainen JT, Sillanpää MJ. Precision matrix estimation with ROPE. J. Comput. Graph. Statist. 2017; 26: 682–694.
10. Ledoit O, Wolf M. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. Rev. Financ. Stud. 2017; 30: 4349–4388.
11. Ledoit O, Wolf M. Shrinkage estimation of large covariance matrices: Keep it simple, statistician? J. Multivar. Anal. 2021; 186: 104796.
12. Liu C, Tang CY. A quasi-maximum likelihood approach for integrated covariance matrix estimation with high frequency data. J. Econom. 2014; 180: 217–232.
13. van Wieringen WN, Peeters CFW. Ridge estimation of inverse covariance matrices from high-dimensional data. Comput. Statist. Data Anal. 2016; 103: 284–303.
14. Wang L, Chen Z, Wang CD, Li R. Ultrahigh dimensional precision matrix estimation via refitted cross validation. J. Econom. 2020; 215: 118–130.
15. Wang X, Zhang B. Target selection in shrinkage estimation of covariance matrix: A structural similarity approach. Stat. Prob. Lett. 2024; 208: 110048.
16. Dey DK, Srinivasan C. Estimation of a covariance matrix under Stein's loss. Ann. Stat. 1985; 13: 1581–1591.
17. Ledoit O, Wolf M. A well-conditioned estimator for large-dimensional covariance matrices. J. Multivar. Anal. 2004a; 88: 365–411.
18. Ledoit O, Wolf M. Nonlinear shrinkage estimation of large-dimensional covariance matrices. Ann. Statist. 2012b; 40: 1024–1060.
19. Bodnar T, Gupta AK, Parolya N. Direct shrinkage estimation of large dimensional precision matrix. J. Multivar Anal,

2016; 146: 223–236.
20. Engel J, Buydens L, Blanchet L. An overview of large-dimensional covariance and precision matrix estimators with applications in chemometrics. J. Chemometrics. 2017; 31: e2880.
21. Gray H, Leday GGR, Vallejos CA, Richardson S. Shrinkage estimation of large covariance matrices using multiple shrinkage targets. arXiv preprint. 2018; 1809: 08024.
22. Bartz D, Höhne J, Müller KR. Multi-target shrinkage. arXiv preprint. 2014; 1412: 2041.
23. Lancewicki T, Aladjem M. Multi-target shrinkage estimation for covariance matrices. IEEE Trans. Signal Process. 2014; 62: 6380–6390.
24. Robbins H. Some thoughts on empirical Bayes estimation. Ann. Stat. 1983; 11: 713–723.
25. Bai J, Shi S. Estimating high dimensional covariance matrices and its applications. Ann. Econom. Finance. 2011; 12: 199–215.
26. Chen CF. Bayesian inference for a normal dispersion matrix and its application to stochastic multiple regression analysis. J. Royal Stat. Soc. Ser. B Stat. Methodol. 1979; 41: 235–248.
27. Ledoit O, Wolf M. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. J. Empir. Finance. 2003; 10: 603–621.
28. Schäfer J, Strimmer K. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. Stat. Appl. Genet. Mol. Biol. 2005; 4: 32.
29. Dobra A, Hans C, Jones B, et al. Sparse graphical models for exploring gene expression data. J. Multivar. Anal. 2014; 90: 196–212.
30. Friedman JH. Regularized discriminant analysis. J. Amer. Statist. Assoc. 1989; 84: 165–175.
31. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. Springer Publishing; 2001.
32. Ledoit O, Wolf M. Honey, I shrunk the sample covariance matrix. J. Portfolio. Management. 2024c; 30: 110–119.
33. Ledoit O, Wolf M. Analytical nonlinear shrinkage of large-dimensional covariance matrices. Ann. Stat. 2020d; 48: 3043–3065.
34. Bodnar T, Gupta AK, Parolya N. On the strong convergence of the optimal linear shrinkage estimator for large dimensional covariance matrix. J. Multivar. Anal. 2014; 132: 215–228.
35. Amenc N, Sourd VL. Portfolio Theory and Performance Analysis. John Wiley & Sons Publishing; 2005.
36. Cai T, Hu J, Li Y, Zheng X. High-dimensional minimum variance portfolio estimation based on high-frequency data. J. Econom. 2019; 214: 482–494.
37. Ding Y, Li Y, Zheng X. High dimensional minimum variance portfolio estimation under statistical factor models. J. Econom. 2021; 222: 502–515.
38. Joo YC, Park SY. Optimal portfolio selection using a simple double-shrinkage selection rule. Finance Res. Lett. 2021; 43: 102019.

Appendix

Proof of Theorem 1. Let $a_n \overset{a.s.}{\sim} b_n$ denotes $a_n - b_n \overset{a.s.}{\to} 0$. By Lemma 1, we obtain

$$
\begin{aligned}
&\frac{1}{p}\|S_n^{-1}\|_F^2 \overset{a.s.}{\sim} \frac{1}{p(1-c)^2}\|\Sigma_n^{-1}\|_F^2 + \frac{c}{p^2(1-c)^3}\|\Sigma_n^{-1}\|_{tr}^2,\\
&\frac{1}{p}tr(S_n^{-1}\Pi) \overset{a.s.}{\sim} \frac{1}{p(1-c)}tr(\Sigma_n^{-1}\Pi),\\
&\frac{1}{p}tr(S_n^{-1}\Sigma_n^{-1}) \overset{a.s.}{\sim} \frac{1}{p(1-c)}\|\Sigma_n^{-1}\|_F^2
\end{aligned}
$$

These yield that, as $n \to \infty$,

$$
A_n - A_0 \overset{a.s.}{\to} 0, \;\; B_n - B_0 \overset{a.s.}{\to} 0
$$

in the sense that each element converges almost surely.□

Proof of Theorem 2. By Lemma 1, the a.s. consistent estimators of $\frac{1}{p}\|\Sigma_n^{-1}\|_{tr}$, $\frac{1}{p}tr(\Sigma_n^{-1}\Pi)$, and $\frac{1}{p}\|\Sigma_n^{-1}\|_F^2$ can be given by

$$
\frac{1}{p}\widehat{\|\Sigma_n^{-1}\|}_{tr} = \frac{(1-p/n)}{p}\|S_n^{-1}\|_{tr} \tag{A1}
$$

$$
\frac{1}{p}tr\widehat{(\Sigma_n^{-1}\Pi)} = \frac{(1-p/n)}{p}tr(S_n^{-1}\Pi) \tag{A2}
$$

and

$$
\begin{aligned}
\tfrac{1}{p}\widehat{\|\Sigma_n^{-1}\|}_F^2 &= \tfrac{(1-p/n)^2}{p}\|S_n^{-1}\|_F^2 - \tfrac{1}{pn(1-p/n)}\widehat{\|\Sigma_n^{-1}\|}_{tr}^2\\
&= \tfrac{(1-p/n)^2}{p}\|S_n^{-1}\|_F^2 - \tfrac{(1-p/n)}{pn}\|S_n^{-1}\|_{tr}^2
\end{aligned} \tag{A3}
$$

Hence the a.s. consistent estimator of a_{0ij} is given by

$$
\begin{aligned}
\hat{a}_{0ij} &= \frac{1}{p}\|S_n^{-1}\|_F^2 - \frac{1}{pn(1-p/n)}\|S_n^{-1}\|_{tr}^2 + \frac{1}{pn(1-p/n)^3}(1-p/n)^2\|S_n^{-1}\|_{tr}^2\\
&\quad - \frac{1}{p(1-p/n)}(1-p/n)\; tr[S_n^{-1}(\Pi_i+\Pi_j)] + \frac{1}{p}tr(\Pi_i\Pi_j)\\
&= \frac{1}{p}\Big(\|S_n^{-1}\|_F^2 - tr(S_n^{-1}\Pi_i) - tr(S_n^{-1}\Pi_j) + tr(\Pi_i\Pi_j)\Big)
\end{aligned}
$$

Therefore $\widehat{A}_0 = A_n$.

Similarly, by Equations (A1)–(A3), the a.s. consistent estimator of b_{0i} is given by

$$
\begin{aligned}
\hat{b}_{0i} &= \frac{1}{n(1-p/n)^2}\left((1-p/n)^2\|S_n^{-1}\|_F^2 - \frac{(1-p/n)}{n}\|S_n^{-1}\|_{tr}^2\right)\\
&\quad + \frac{1}{pn(1-p/n)^3}(1-p/n)^2\|S_n^{-1}\|_{tr}^2 - \frac{1}{n(1-p/n)}(1-p/n)tr(S_n^{-1}\Pi_i)\\
&= \frac{1}{n}\|S_n^{-1}\|_F^2 + \frac{1}{pn}\|S_n^{-1}\|_{tr}^2 - \frac{1}{n}\; tr(S_n^{-1}\Pi_i)
\end{aligned}
$$

Thus

$$
\widehat{B}_0 = \frac{1}{p}\begin{pmatrix} \frac{p}{n}\|S_n^{-1}\|_F^2 + \frac{1}{n}\|S_n^{-1}\|_{tr}^2 - \frac{p}{n}\; tr(S_n^{-1}\Pi_1)\\ \vdots \\ \frac{p}{n}\|S_n^{-1}\|_F^2 + \frac{1}{n}\|S_n^{-1}\|_{tr}^2 - \frac{p}{n}\; tr(S_n^{-1}\Pi_k)\end{pmatrix}
$$

□

Proof of Theorem 4. Note that, in this case, $p/n \to c = 0$ as $n \to \infty$. Then, for $i = 1, \cdots, k$, $\hat{b}_{0i} \overset{a.s.}{\sim} b_{0i} = o(1)$. Hence, $\widehat{B}_0 \overset{a.s.}{\to} 0$ as $n \to \infty$.

On the other hand, for $i, j = 1, \cdots, k$, the (i, j)th element of A_n,

$$
\begin{aligned}
\frac{1}{p} tr[(S_n^{-1} - \Pi_i)(S_n^{-1} - \Pi_j)] &\overset{a.s.}{\sim} a_{0ij} \\
&\overset{a.s.}{\sim} \frac{1}{p}\Big(\|\Sigma_n^{-1}\|_F^2 - tr[\Sigma_n^{-1}(\Pi_i + \Pi_j)] + tr(\Pi_i \Pi_j)\Big) \\
&= \frac{1}{p} tr[(\Sigma_n^{-1} - \Pi_i)(\Sigma_n^{-1} - \Pi_j)]
\end{aligned}
$$

Thus $\widehat{A}_0 \overset{a.s.}{\sim} A_n^0$.

Notice that

$$
\begin{aligned}
&\frac{1}{p} tr[(\Sigma_n^{-1} - \Pi_i)(\Sigma_n^{-1} - \Pi_j)] \\
&\leq \frac{1}{p}\|\Sigma_n^{-1}\|_F^2 + \Big(\frac{1}{p}\|\Sigma_n^{-1}\|_F^2\Big)^{1/2}\Big\{\Big(\frac{1}{p}\|\Pi_i\|_F^2\Big)^{1/2} + \Big(\frac{1}{p}\|\Pi_j\|_F^2\Big)^{1/2}\Big\} \\
&\quad + \Big(\frac{1}{p}\|\Pi_i\|_F^2 \frac{1}{p}\|\Pi_j\|_F^2\Big)^{1/2} \\
&= O(1)
\end{aligned}
$$

Then we obtain $\widehat{A}_0^{-1} = O(1)$ almost surely. That is, $\widetilde{\Omega} = \widehat{A}_0^{-1}\widehat{B}_0 \overset{a.s.}{\to} 0$ as $n \to \infty$, This yields $\widehat{\Omega} \to 0$ almost surely.□