

Journal Pre-proof

Recent advances in shrinkage-based high-dimensional inference

Olha Bodnar, Taras Bodnar, Nestor Parolya

PII: S0047-259X(21)00104-4
DOI: <https://doi.org/10.1016/j.jmva.2021.104826>
Reference: YJMVA 104826

To appear in: *Journal of Multivariate Analysis*

Received date : 10 August 2021
Revised date : 25 August 2021
Accepted date : 25 August 2021



Please cite this article as: O. Bodnar, T. Bodnar and N. Parolya, Recent advances in shrinkage-based high-dimensional inference, *Journal of Multivariate Analysis* (2021), doi: <https://doi.org/10.1016/j.jmva.2021.104826>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2021 Elsevier Inc. All rights reserved.

Recent advances in shrinkage-based high-dimensional inference

Olha Bodnar^a, Taras Bodnar^{b,*}, Nestor Parolya^c^aUnit of Statistics, School of Business, Örebro University, Fakultetsgatan 1, SE-70182 Örebro, Sweden^bDepartment of Mathematics, Stockholm University, Roslagsvägen 101, SE-10691 Stockholm, Sweden^cDepartment of Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands**Abstract**

Recently, the shrinkage approach has increased its popularity in theoretical and applied statistics, especially, when point estimators for high-dimensional quantities have to be constructed. A shrinkage estimator is usually obtained by shrinking the sample estimator towards a deterministic target. This allows to reduce the high volatility that is commonly present in the sample estimator by introducing a bias such that the mean-square error of the shrinkage estimator becomes smaller than the one of the corresponding sample estimator. The procedure has shown great advantages especially in the high-dimensional problems where, in general case, the sample estimators are not consistent without imposing structural assumptions on model parameters.

In this paper, we review the mostly used shrinkage estimators for the mean vector, covariance and precision matrices. The application in portfolio theory is provided where the weights of optimal portfolios are usually determined as functions of the mean vector and covariance matrix. Furthermore, a test theory on the mean-variance optimality of a given portfolio based on the shrinkage approach is presented as well.

Keywords: Covariance matrix, High-dimensional asymptotics, High-dimensional optimal portfolio, Mean vector, Precision matrix, Random matrix theory, Shrinkage estimation
2020 MSC: Primary 62H12, Secondary 62F12, 62H15, 62P05

1. Introduction

High-dimensional inference procedures play an important role in many fields of science, like in economics, finance, environmetrics, physics, signal processing, etc., when a statistical model is needed to be fitted to real data. For instance, high-dimensional optimal portfolios are well motivated by the rapid development of technology, which provides investors opportunities to construct a portfolio consisting of a large number of assets traded simultaneously across the world. Moreover, the availability of high-frequent financial data provides a considerable amount of information which can be used in the construction of optimal portfolios.

It is remarkable that the application of the traditional sample estimators is not recommendable in the high-dimensional setting. Although the sample estimators work well when the process dimension is fixed and is significantly smaller than the number of observations, it does not longer hold when the two quantities are comparable. The former case is often used in statistics and it is called the standard asymptotic regime (see, [62]). Under this asymptotic regime the traditional sample estimators, like the maximum likelihood estimator or method of moments estimator, are usually consistent under some regularity conditions. However, it does not longer hold true when the process dimension is comparable to or even larger than the sample size. Here, we are in the situation when both the number of assets and the sample size can tend to infinity. This double asymptotic regime has an interpretation when the ratio between the process dimension and the sample size, also known as the concentration ratio, tends to a finite value as the sample size tends to infinity. This asymptotic regime is known as a high-dimensional asymptotics or "Kolmogorov" asymptotics (see, e.g., [32]). Under the high-dimensional asymptotics the sample estimators behave

*Corresponding author. Email address: taras.bodnar@math.su.se

very unpredictable and they are far from the optimal ones. In general, the greater the concentration ratio the worse are the sample estimators. This well-known problem in statistics is called "the curse of dimensionality".

Recently, new estimators came in play which are biased but can significantly reduce the mean square error in comparison to the traditional estimators. These estimators are known as shrinkage estimators and were introduced in the seminal paper of Stein (see, [80]). A shrinkage estimator is usually defined as a linear combination of the corresponding sample estimator and a known target. The corresponding coefficients in the linear combination are often called shrinkage intensities. It is a very challenging task to find consistent estimators for the shrinkage intensities.

The first shrinkage estimator was developed for the mean vector of a multivariate normal distribution with identity covariance matrix ([57, 80]) and was extended to the case of an arbitrary covariance matrix in [5, 7, 8, 40, 46, 49, 66, 81]. These results were obtained in the standard asymptotic regime, while a high-dimensional version of the James-Stein type estimator was proposed by [37]. Recently, an optimal shrinkage estimator obtained by minimizing the expected quadratic loss function was derived in [83], while a shrinkage estimator of the mean vector shrunk towards an arbitrary target vector was introduced in [22].

The situation becomes more challenging when the covariance matrix is estimated and, especially, when one needs to infer its inverse, the precision matrix. There are some significant improvements when the covariance matrix has a special structure, e.g. sparse, low rank etc. (see, [33, 35, 74]). The results for the covariance matrix that possesses a factor structure were derived in [41–43]. In these cases the covariance matrix can consistently be estimated even for high-dimensional data. However, when no information about a specific structure of the covariance matrix is available, the shrinkage estimator seems to be the most favorable approach in the high-dimensional setting (cf., [20, 64, 65]). The shrinkage estimators for the precision matrix were derived in [21, 61, 82], among others.

In order to handle the curse of dimensionality in the case of the high-dimensional asymptotic regime the results from random matrix theory are usually used. Random matrix theory is a very fast growing branch of probability theory with many applications in statistics and finance. It studies the behavior of the eigenvalues of random matrices under the double asymptotic regime (see, e.g., [2–4, 14, 24, 28, 48, 67, 75, 78]). It is discovered that appropriately transformed random matrix at infinity has a nonrandom behavior and showed how to find the limiting density of its eigenvalues. In particular, Silverstein and Bai [78] proved under very general conditions that the Stieltjes transform of the sample covariance matrix tends almost surely to a nonrandom function which satisfies some equation. This equation was first derived by [67] who showed how the real covariance matrix and its sample estimator are connected at infinity, while a general form of this equation was given in [75]. Finally, using the results of random matrix theory statistical tests on the structure of the covariance matrix were suggested by [15, 36, 45, 51, 85, 87].

Improved estimators of the model parameters constructed by employing random matrix theory, especially shrinkage-based estimators, are widely used in many fields of science, like in signal processing and finance (see, [34, 39, 43, 44, 54, 84, 86]). For instance, an improved calibration of the high-dimensional precision matrix was suggested in [86], while the applications of random matrix theory to signal processing and portfolio theory was discussed in [44]. Furthermore, several authors showed that the shrinkage estimators applied to portfolio weights indeed lead to better results (see, e.g., [23, 26, 47, 50, 64]). In particular, the shrinkage estimator for the covariance matrix was applied to construct an improved estimator of the weights of the global minimum variance portfolio by [64], while the multivariate shrinkage estimator obtained by shrinking the portfolio weights directly was suggested in [50]. The same idea was also used by [47] who constructed a feasible shrinkage estimator for the global minimum variance portfolio which dominates the traditional sample estimator. More recently, the shrinkage estimators based on an arbitrary target vector of portfolio weights were derived by [26] and [23] in the case of the global minimum variance portfolio and mean-variance portfolio, respectively. Finally, statistical test theory on the optimality of portfolio weights was developed in [18, 19] that is based on the shrinkage approach, while sequential procedures derived on the weights of optimal portfolios were established in [9, 10].

The rest of the paper is organized as follows. In Section 2 we present the shrinkage estimator for the high-dimensional mean vector, covariance matrix and precision matrix. Recent results of the application of the shrinkage approach in finance is discussed in Section 3. Discussion of the results is provided in Section 4.

2. Shrinkage estimation of the mean vector and covariance matrix

Let $\mathbf{X}_n = (x_{ij})_{i \in \{1, \dots, p\}, j \in \{1, \dots, n\}}$ with x_{ij} be independent and identically distributed with zero mean and variance equal to one. Throughout the paper it is assumed that the data matrix $\mathbf{Y}_n = [\mathbf{y}_1, \dots, \mathbf{y}_n]$ follows the stochastic model

expressed as

$$\mathbf{Y}_n = \boldsymbol{\mu} \mathbf{1}_n^\top + \boldsymbol{\Sigma}^{1/2} \mathbf{X}_n, \quad (1)$$

where $\mathbf{1}_n$ is the n -dimensional vector of ones and $\boldsymbol{\Sigma}^{1/2}$ is a square root of the positive definite matrix $\boldsymbol{\Sigma}$. We further assume that $\mathbb{E}(|x_{ij}|^{4+\varepsilon}) < \infty$ for any small number $\varepsilon > 0$. No specific distributional assumption is imposed on the element of \mathbf{X}_n .

Under model (1) the observation vectors $\mathbf{y}_1, \dots, \mathbf{y}_n$ are independent and identically distributed with $\mathbb{E}(\mathbf{y}_j) = \boldsymbol{\mu}$ and $\mathbb{Cov}(\mathbf{y}_j) = \boldsymbol{\Sigma}$. The two parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are unknown quantities which have to be inferred by using the observation matrix \mathbf{Y}_n . The most commonly used estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the sample estimators expressed as

$$\bar{\mathbf{y}}_n = \frac{1}{n} \sum_{i=1}^n \mathbf{y}_n = \frac{1}{n} \mathbf{Y}_n \mathbf{1}_n \quad \text{and} \quad \mathbf{S}_n = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{y}_n - \bar{\mathbf{y}}_n)(\mathbf{y}_n - \bar{\mathbf{y}}_n)^\top = \frac{1}{n-1} \mathbf{Y}_n \left(\mathbf{I}_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^\top \right) \mathbf{Y}_n^\top, \quad (2)$$

where \mathbf{I}_n denotes the identity matrix of size n . Under the additional assumption that \mathbf{y}_i are multivariate normally distributed, $\bar{\mathbf{y}}_n$ and $(n-1)\mathbf{S}_n/n$ are also the maximum likelihood estimators for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ and, consequently, they are asymptotically efficient when p is finite and n tends to infinity, i.e., in the classical asymptotic regime. However, both the estimators possess high variability when p becomes comparable to n . As a result, their application to the high-dimensional problems is not desired and new approaches should be employed instead.

In order to reduce the variability which is present in the traditional sample estimators, for example in $\bar{\mathbf{y}}_n$ and \mathbf{S}_n , the shrinkage estimators have been developed in statistical literature, which are usually (slightly) biased but, on other side, they possess considerably smaller variance in comparison to the sample estimators. A shrinkage estimator for a quantity of interest is not uniquely defined and is obtained by minimizing a risk function, which may depend on the application at hand. In Section 2.1 we review the existent shrinkage estimators for the mean vector, while Sections 2.2 and 2.3 present shrinkage estimators for the covariance matrix and the precision matrix. Later, in Section 3 the shrinkage approach is applied to infer the weights of optimal portfolios, which are usually present as functions of the mean vector and covariance (precision) matrix.

2.1. Shrinkage estimation of the mean vector

A shrinkage estimator for the mean vector $\boldsymbol{\mu}$ is usually derived by minimizing the quadratic loss function expressed as

$$L(\hat{\boldsymbol{\mu}}_n, \boldsymbol{\mu}) = (\hat{\boldsymbol{\mu}}_n - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\hat{\boldsymbol{\mu}}_n - \boldsymbol{\mu}) \quad (3)$$

or, its expected value, for an arbitrary estimator $\hat{\boldsymbol{\mu}}_n$,

$$R(\boldsymbol{\mu}) = \mathbb{E} \left((\hat{\boldsymbol{\mu}}_n - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\hat{\boldsymbol{\mu}}_n - \boldsymbol{\mu}) \right). \quad (4)$$

Depending on the imposed condition on $\boldsymbol{\Sigma}$, n , and p several shrinkage estimators exist in the literature which we review below.

The James-Stein shrinkage estimator for the mean vector was derived under the assumption that the covariance matrix $\boldsymbol{\Sigma}$ is the identity matrix and that $n > p > 2$. It is given by

$$\hat{\boldsymbol{\mu}}_{n,JS} = \left(1 - \frac{p-2}{n \bar{\mathbf{y}}_n^\top \bar{\mathbf{y}}_n} \right) \bar{\mathbf{y}}_n. \quad (5)$$

When the concentration ratio $c < 1$ with c defined by $p/n \rightarrow c$ as $n \rightarrow \infty$, a modified version of the James-Stein estimator for $n > p \geq 3$ and an arbitrary covariance matrix is expressed as

$$\hat{\boldsymbol{\mu}}_{n,mJS} = \left(1 - \frac{p-2}{n-p+2} \frac{1}{\bar{\mathbf{y}}_n^\top \mathbf{S}_n^{-1} \bar{\mathbf{y}}_n} \right) \bar{\mathbf{y}}_n, \quad (6)$$

where \mathbf{S}_n is the sample estimator of $\boldsymbol{\Sigma}$ given in (2).

For $p > n \geq 3$ and an arbitrary covariance matrix $\boldsymbol{\Sigma}$, a Baranchik type shrinkage estimator for the mean vector was discussed in [37] and it is given by

$$\hat{\boldsymbol{\mu}}_{n,B} = \left\{ \mathbf{I}_p - \frac{r \mathbf{S}_n \mathbf{S}_n^\top}{\bar{\mathbf{y}}_n^\top \mathbf{S}_n^\top \bar{\mathbf{y}}_n} \right\} \bar{\mathbf{y}}_n,$$

for $\min\{n-1, p\} \geq 3$. The symbol \mathbf{S}_n^+ denotes the Moore-Penrose inverse of \mathbf{S}_n . It was proved in [37], that $\hat{\boldsymbol{\mu}}_{n,B}$ dominates the sample estimator $\bar{\mathbf{y}}_n$ under the quadratic loss when

$$0 \leq r \leq \frac{2(\min\{n-1, p\} - 2)}{n + p - 2 \min\{n-1, p\} + 2}.$$

Chételat and Wells [37] considered a further generalization of the James-Stein estimator in the case of $p > n$ expressed as

$$\hat{\boldsymbol{\mu}}_{n,CW} = (\mathbf{I}_p - \mathbf{S}_n \mathbf{S}_n^+) \bar{\mathbf{y}}_n + \left\{ 1 - \frac{a}{\bar{\mathbf{y}}_n^\top \mathbf{S}_n^+ \bar{\mathbf{y}}_n} \right\}_+ \mathbf{S}_n \mathbf{S}_n^+ \bar{\mathbf{y}}_n,$$

where $b_+ = \max(b, 0)$. They argued that $\hat{\boldsymbol{\mu}}_{n,CW}$ dominates the James-Stein shrinkage estimator when

$$a = \frac{n-3}{p-n+4}.$$

A further shrinkage estimator for the mean vector was derived in [83] who suggested to shrink the sample estimator $\bar{\mathbf{y}}_n$ to the unity target vector. The corresponding shrinkage coefficients are found by minimizing the expected quadratic loss (4). This leads to the following shrinkage estimator (see, [83]):

$$\hat{\boldsymbol{\mu}}_{n,W} = \frac{Z_{1,n} - Z_{4,n}}{Z_{1,n} + Z_{2,n}Z_{4,n}} \bar{\mathbf{y}}_n + \frac{Z_{2,n}Z_{3,n}}{Z_{1,n} + Z_{2,n}Z_{4,n}} \mathbf{1}_n, \quad (7)$$

with

$$\begin{aligned} Z_{1,n} &= \frac{1}{p(n-1)} \sum_{i \neq j} \mathbf{y}_{n,i}^\top \mathbf{S}_n^+ \mathbf{y}_{n,j}, & Z_{2,n} &= \frac{1}{np} \left(\sum_{k=1}^n \mathbf{y}_{n,k}^\top \mathbf{S}_n^+ \mathbf{y}_{n,k} - \frac{1}{n-1} \sum_{i \neq j} \mathbf{y}_{n,i}^\top \mathbf{S}_n^+ \mathbf{y}_{n,j} \right), \\ Z_{3,n} &= \frac{1}{n \mathbf{1}_n^\top \mathbf{S}_n^+ \mathbf{1}_n} \sum_{k=1}^n \mathbf{1}_n^\top \mathbf{S}_n^+ \mathbf{y}_{n,k}, & Z_{4,n} &= \frac{1}{p(n-1) \mathbf{1}_n^\top \mathbf{S}_n^+ \mathbf{1}_n} \sum_{i \neq j} \mathbf{1}_n^\top \mathbf{S}_n^+ \mathbf{y}_{n,i} \mathbf{y}_{n,j}^\top \mathbf{S}_n^+ \mathbf{1}_n, \end{aligned}$$

for $p > n$. The shrinkage estimator (7) is computationally complicated due to the presence of the double sum over p and n in its definition. In order to simplify its computation in practice, the application of its asymptotic counterpart was suggested in [83].

In the case of an arbitrary shrinkage target vector $\boldsymbol{\mu}_0$, a linear shrinkage estimator that minimizes the quadratic loss function (3) was developed by [22]. For $c < 1$, it is given by

$$\hat{\boldsymbol{\mu}}_{n,BOP} = \hat{\alpha}_{mean} \bar{\mathbf{y}}_n + \hat{\beta}_{mean} \boldsymbol{\mu}_0, \quad (8)$$

where

$$\hat{\alpha}_{mean} = \frac{(\bar{\mathbf{y}}_n^\top \mathbf{S}_n^{-1} \bar{\mathbf{y}}_n - \frac{p/n}{1-p/n}) \boldsymbol{\mu}_0^\top \mathbf{S}_n^{-1} \boldsymbol{\mu}_0 - (\bar{\mathbf{y}}_n^\top \mathbf{S}_n^{-1} \boldsymbol{\mu}_0)^2}{\bar{\mathbf{y}}_n^\top \mathbf{S}_n^{-1} \bar{\mathbf{y}}_n \boldsymbol{\mu}_0^\top \mathbf{S}_n^{-1} \boldsymbol{\mu}_0 - (\bar{\mathbf{y}}_n^\top \mathbf{S}_n^{-1} \boldsymbol{\mu}_0)^2} \quad (9)$$

and

$$\hat{\beta}_{mean} = (1 - \hat{\alpha}) \frac{\bar{\mathbf{y}}_n^\top \mathbf{S}_n^{-1} \boldsymbol{\mu}_0}{\boldsymbol{\mu}_0^\top \mathbf{S}_n^{-1} \boldsymbol{\mu}_0}. \quad (10)$$

2.2. Shrinkage estimation of the covariance matrix

In the derivation of the shrinkage estimation of the covariance matrix several loss functions are considered in the literature. Below, we review the approaches which are obtained by minimizing the quadratic loss function which is defined by the Frobenius norm in the matrix case expressed as

$$L(\hat{\boldsymbol{\Sigma}}_n, \boldsymbol{\Sigma}) = \|\hat{\boldsymbol{\Sigma}}_n - \boldsymbol{\Sigma}\|_F^2 \quad (11)$$

for a given estimator $\hat{\boldsymbol{\Sigma}}_n$ of $\boldsymbol{\Sigma}$ with $\|\mathbf{A}\|_F^2 = \text{tr}(\mathbf{A}\mathbf{A}^\top)$ for a square matrix \mathbf{A} .

At first, Ledoit and Wolf ([64]) proposed a linear shrinkage estimator of the covariance matrix Σ which shrinks the sample covariance matrix S_n to the identity matrix and studied its behaviour in the high-dimensional setting. This estimator shrinks the eigenvalues of the sample covariance matrix linearly and is obtained by minimizing the expected quadratic loss function expressed as

$$R(\Sigma) = \mathbb{E}(\|\hat{\Sigma}_n - \Sigma\|_F^2). \quad (12)$$

A generalization of the linear shrinkage estimator of [64] was suggested in [20] where the shrinkage target was chosen to be an arbitrary nonrandom matrix Σ_0 . In contrast to the Ledoit and Wolf linear shrinkage estimator of the covariance matrix, the new shrinkage estimator was derived by minimizing the loss function (11) directly. It is given by

$$\hat{\Sigma}_{n;BGP} = \hat{\alpha}_{cov} S_n + \hat{\beta}_{cov} \Sigma_0, \quad (13)$$

with

$$\hat{\alpha}_{cov} = 1 - \frac{\frac{1}{n} \|S_n\|_{tr}^2 \|\Sigma_0\|_F^2}{\|S_n\|_F^2 \|\Sigma_0\|_F^2 - (\text{tr}(S_n \Sigma_0))^2} \quad (14)$$

and

$$\hat{\beta}_{cov} = \frac{\text{tr}(S_n \Sigma_0)}{\|\Sigma_0\|_F^2} (1 - \hat{\alpha}_{cov}), \quad (15)$$

where $\|A\|_{tr} = \text{tr}[(AA^\top)^{1/2}]$ denotes for the trace norm and Σ_0 is assumed to possess the bounded trace norm. The shrinkage estimator (13) was derived under the assumption $\mathbb{E}(|x_{ij}|^{4+\varepsilon}) < \infty$, while the shrinkage estimator in [64] requires the existent of 8th moments.

The linear shrinkage estimator of [64] also differs from (13) in its structure. First, it is derived for the specific target matrix $\Sigma_0 = 1/p \mathbf{I}$. Second, the expression of $\hat{\alpha}_{cov}$ is different in two approaches. Namely, instead of $\frac{1}{n} \|S_n\|_{tr}^2$ the Ledoit and Wolf shrinkage estimator uses $\frac{1}{n^2} \sum_{i=1}^n \|\mathbf{y}_i \mathbf{y}_i^\top - S_n\|_F^2$ where \mathbf{y}_i are the i -th columns of the observation matrix \mathbf{Y}_n . It is defined by

$$\hat{\alpha}_{cov;LW} = 1 - \frac{\min\{\hat{b}_{cov}^2, \hat{d}_{cov}^2\}}{\hat{d}_{cov}^2}, \quad (16)$$

where

$$\hat{d}_{cov}^2 = \frac{1}{p} \|S_n\|_F^2 - (1/p \text{tr}(S_n))^2, \quad \hat{b}_{cov}^2 = \frac{1}{p} \frac{1}{n^2} \sum_{i=1}^n \|\mathbf{y}_i \mathbf{y}_i^\top - S_n\|_F^2.$$

The shrinkage estimator (13) is unconstrained, while the Ledoit and Wolf estimator is constrained. If $\hat{b}_{cov}^2 > \hat{d}_{cov}^2$ in (16), then $\hat{\alpha}_{cov;LW} = 0$, i.e., the Ledoit and Wolf shrinkage estimator coincides with $\text{tr}(S_n) \frac{1}{p} \mathbf{I}$, independently how large p is with respect to n . In contrast, we always have that $0 < \hat{\alpha}_{cov} \leq 1$ for (13) with $\hat{\alpha}_{cov} = 1$ only if $c = 0$, i.e., the sample covariance matrix possesses the smallest Frobenius loss only if p is much smaller than n . For $c > 0$, the sample covariance matrix is not an optimal estimator for the covariance matrix in terms of minimizing the quadratic loss function (11). Finally, we note that the Ledoit and Wolf estimator is more computationally intensive than ((13) with $\Sigma_0 = \frac{1}{p} \mathbf{I}$, since the quantity \hat{b}_{cov}^2 has to be calculated by a loop.

Further improved estimators of the covariance matrix were suggested in [38, 52, 53, 76] among others. These estimators were derived by minimizing the Stein loss given by (see, [38])

$$L_S(\hat{\Sigma}_n, \Sigma) = \text{tr}(\hat{\Sigma}_n \Sigma^{-1}) - \log(\det(\hat{\Sigma}_n \Sigma^{-1})) - p \quad (17)$$

or the corresponding risk function $R_S(\Sigma) = \mathbb{E}(L_S(\hat{\Sigma}_n, \Sigma))$, and are defined as orthogonal invariant estimators. The class of rotation-equivariant estimators of the covariance matrix coincides with the class of estimators which possess the same eigenvectors as the sample covariance matrix (c.f., [73, Section 5.4]). That is, they are determined as

$$\hat{\Sigma}_{n,OI} = \mathbf{H} \Phi(\mathbf{D}) \mathbf{H}^\top, \quad (18)$$

where $\mathbf{S}_n = \mathbf{H}\mathbf{D}\mathbf{H}^\top$ is the eigenvalue decomposition of the sample covariance matrix \mathbf{S}_n with $\mathbf{D} = \text{diag}(d_1, \dots, d_p)$, $d_1 \geq d_2 \geq \dots \geq d_p$ and $\Phi(\mathbf{D}) = \text{diag}(\phi_1(d_1), \dots, \phi_p(d_p))$ for continuously differentiable function ϕ_i , $i \in \{1, \dots, p\}$. Dey and Srinivasan ([38]) derived the set of functions $\phi_i(\cdot)$, $i \in \{1, \dots, p\}$ for which the orthogonal invariant estimator (18) dominates the sample estimator \mathbf{S}_n .

The orthogonal invariant estimator $\hat{\Sigma}_{n,OI}$ was generalized to the non-linear shrinkage estimator by [65] in the high-dimensional setting, which, for $i \in \{1, \dots, p\}$, is given by

$$\mathbf{S}_{n,LWnonlin} = \mathbf{H}\text{diag}(d_1^{or}, \dots, d_p^{or})\mathbf{H}^\top, \quad d_i^{or} = \begin{cases} \frac{d_i}{|1-c-d_i\check{m}_F(d_i)|^2}, & \text{if } d_i > 0, \\ \frac{1}{(c-1)\check{m}_F(0)}, & \text{if } d_i = 0, \end{cases} \quad (19)$$

where $m_F(z)$ denotes the limiting Stieltjes transform of the sample covariance matrix defined for a distribution function $G : \mathbb{R} \rightarrow \mathbb{R}$ as

$$m_G(z) = \int_{-\infty}^{+\infty} \frac{1}{\lambda - z} dG(\lambda); \quad z \in \mathbb{C}^+ \equiv \{z \in \mathbb{C} : \Im z > 0\}.$$

Moreover, for any $x \in \mathbb{R}$ the quantities $\check{m}_F(x) = \lim_{z \rightarrow x} m_F(z)$ and $\check{m}_F^-(x) = \lim_{z \rightarrow x} m_F^-(z) = \lim_{z \rightarrow x} \frac{c-1}{z} + cm_F(z)$ exist and are finite for $c < 1$ and $c > 1$, respectively. The existence of those limits was proven in [63, 77]. Albeit the oracle shrinkage intensities d_i^{or} depend on the unknown limiting Stieltjes transform, thanks to the recent paper of Ledoit and Wolf [65], they can be fast and efficiently estimated using a simple nonparametric procedure.

2.3. Shrinkage estimation of precision matrix

A linear shrinkage estimator for the precision matrix Σ^{-1} was developed in [21] and it is derived by minimizing the quadratic loss expressed as

$$L(\hat{\Pi}_n, \Sigma^{-1}) = \|\hat{\Pi}_n - \Sigma^{-1}\|_F^2 \quad (20)$$

for an estimator $\hat{\Pi}_n$ of Σ^{-1} . The linear shrinkage estimator for the precision matrix for $c < 1$ is given by

$$\hat{\Pi}_{n,BGP} = \hat{\alpha}_{prec} \mathbf{S}_n^{-1} + \hat{\beta}_{prec} \Pi_0, \quad (21)$$

with

$$\hat{\alpha}_{prec} = 1 - p/n - \frac{\frac{1}{n} \|\mathbf{S}_n^{-1}\|_{tr}^2 \|\Pi_0\|_F^2}{\|\mathbf{S}_n^{-1}\|_F^2 \|\Pi_0\|_F^2 - (\text{tr}(\mathbf{S}_n^{-1} \Pi_0))^2} \quad (22)$$

and

$$\hat{\beta}_{prec} = \frac{\text{tr}(\mathbf{S}_n^{-1} \Pi_0)}{\|\Pi_0\|_F^2} (1 - p/n - \hat{\alpha}_{prec}), \quad (23)$$

where $1/p\Pi_0$ is assumed to have the bounded trace norm.

Another shrinkage estimator of the precision matrix is the scaled standard estimator (SSE) discussed in [61, 71, 79]. It is given by

$$\hat{\Pi}_{SSE} = \frac{n-p-2}{n-1} \mathbf{S}_n^{-1} \delta_{(p < n)} + \frac{p}{n-1} \mathbf{S}_n^+ \delta_{(p \geq n)}, \quad (24)$$

where \mathbf{S}_n^+ is the Moore-Penrose inverse of \mathbf{S}_n and $\delta_{(\cdot)}$ is a Dirac delta function.

The other two shrinkage estimators for the precision were proposed by [40] and by [61] and they are expressed as

$$\hat{\Pi}_{EM} = \frac{n-p-2}{n-1} \mathbf{S}_n^{-1} + \frac{p^2 + p - 2}{(n-1)\text{tr}(\mathbf{S}_n)} \mathbf{I} \quad (25)$$

and

$$\hat{\Pi}_{KS} = p((n-1)\mathbf{S}_n + \text{tr}(\mathbf{S}_n)\mathbf{I})^{-1}, \quad (26)$$

respectively.

3. Application in portfolio theory

In this section we discuss how the theory of shrinkage estimation can be used in portfolio theory where the weights of different optimal portfolio can often be expressed as functions of the mean vector and covariance matrix (see, e.g., [16, 56]). The practical computation of the considered shrinkage estimators of optimal portfolio weights as well as some shrinkage estimators for the mean vector and for the covariance matrix presented in the previous section can be performed in the R-packages, like *HDSHOP* (High-Dimensional Shrinkage Optimal Portfolio, see [17]) and *DOSPortfolio* (Dynamic Optimal Shrinkage Portfolio).

Following the Markowitz theory ([68]), mean-variance optimal portfolios are obtained by minimizing the portfolio variance for a given level of the expected return. The solutions of the Markowitz problem lie on a parabola in the mean-variance space, known as the efficient frontier (see, e.g., [6, 11, 30, 59, 70]). Let $\mathbf{w} = (w_1, \dots, w_p)^\top$ be the p -dimensional vector of portfolio weights. Then the expected return of the portfolio with weights \mathbf{w} is $\mathbf{w}^\top \boldsymbol{\mu}$, while its variance is $\mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w}$. Markowitz optimal portfolios can also be deduced as solutions of other optimization problems (e.g., [25]), like by maximizing the expected utility function (see, [56]) expressed as

$$\mathbf{w}^\top \boldsymbol{\mu} - \frac{\gamma}{2} \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w} \rightarrow \max \quad \text{subject to} \quad \mathbf{w}^\top \mathbf{1}_p = 1, \quad (27)$$

where $\gamma > 0$ is the coefficient of risk aversion that measures the investor's attitude towards risk. The solution of (27) is known as the mean-variance (MV) optimal portfolio and it is given by

$$\mathbf{w}_{MV} = \frac{\boldsymbol{\Sigma}^{-1} \mathbf{1}_p}{\mathbf{1}_p^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}_p} + \gamma^{-1} \mathbf{Q} \boldsymbol{\mu}, \quad (28)$$

where

$$\mathbf{Q} = \boldsymbol{\Sigma}^{-1} - \frac{\boldsymbol{\Sigma}^{-1} \mathbf{1}_p \mathbf{1}_p^\top \boldsymbol{\Sigma}^{-1}}{\mathbf{1}_p^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}_p}.$$

In the case of the fully risk-averse investor, i.e., $\gamma = \infty$, the optimal portfolio is found by minimizing the portfolio variance, i.e.,

$$\mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w} \rightarrow \min \quad \text{subject to} \quad \mathbf{w}^\top \mathbf{1}_p = 1, \quad (29)$$

and its weights are given by

$$\mathbf{w}_{GMV} = \frac{\boldsymbol{\Sigma}^{-1} \mathbf{1}_p}{\mathbf{1}_p^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}_p}. \quad (30)$$

The optimal portfolio with the weights \mathbf{w}_{GMV} is known in financial literature as the global minimum variance (GMV) portfolio. This portfolio lies on the vertex of the efficient frontier, whose equation is given by

$$(R - R_{GMV})^2 = s(V - V_{GMV}),$$

where

$$R_{GMV} = \frac{\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}_p}{\mathbf{1}_p^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}_p}, \quad R_{GMV} = \frac{1}{\mathbf{1}_p^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}_p}, \quad s = \boldsymbol{\mu}^\top \mathbf{Q} \boldsymbol{\mu} \quad (31)$$

are the expected return of the GMV portfolio, the variance of the GMV portfolio, and the slope parameter of the efficient frontier, respectively.

3.1. Traditional sample estimators of portfolio weights

Optimal portfolios cannot be constructed by using (27) and (29), since both formulas depend on the true value of the mean vector and the covariance matrix. Markowitz ([69]) suggested to use historical data of asset returns $\mathbf{y}_1, \dots, \mathbf{y}_n$ to construct the sample estimators $\bar{\mathbf{y}}_n$ and \mathbf{S}_n of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. This leads to the following traditional sample estimators of the MV portfolio weights

$$\hat{\mathbf{w}}_{MV,S} = \frac{\mathbf{S}_n^{-1} \mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_n^{-1} \mathbf{1}_p} + \gamma^{-1} \hat{\mathbf{Q}}_n \bar{\mathbf{y}}_n \quad (32)$$

with

$$\hat{\mathbf{Q}}_n = \mathbf{S}_n^{-1} - \frac{\mathbf{S}_n^{-1} \mathbf{1}_p \mathbf{1}_p^\top \mathbf{S}_n^{-1}}{\mathbf{1}_p^\top \mathbf{S}_n^{-1} \mathbf{1}_p},$$

and of the GMV portfolio weights given by

$$\hat{\mathbf{w}}_{GMV;S} = \frac{\mathbf{S}_n^{-1} \mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_n^{-1} \mathbf{1}_p}, \quad (33)$$

respectively. The distributional properties of the estimators (32) and (33) were studied by [16, 58, 72], among others.

3.2. Naive shrinkage approach

Alternatively, one can replace the unknown $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ (or $\boldsymbol{\Sigma}^{-1}$) in (28) and (30) by any improved estimator as considered in Section 2 as in the case of the GMV portfolio discussed in [60, 64], among others. This leads to

$$\hat{\mathbf{w}}_{MV;nSh} = \frac{\hat{\boldsymbol{\Sigma}}_n^{-1} \mathbf{1}_p}{\mathbf{1}_p^\top \hat{\boldsymbol{\Sigma}}_n^{-1} \mathbf{1}_p} + \gamma^{-1} \hat{\mathbf{Q}}_{n;nSh} \hat{\boldsymbol{\mu}}_n, \quad (34)$$

with

$$\hat{\mathbf{Q}}_{n;nSh} = \hat{\boldsymbol{\Sigma}}_n^{-1} - \frac{\hat{\boldsymbol{\Sigma}}_n^{-1} \mathbf{1}_p \mathbf{1}_p^\top \hat{\boldsymbol{\Sigma}}_n^{-1}}{\mathbf{1}_p^\top \hat{\boldsymbol{\Sigma}}_n^{-1} \mathbf{1}_p}$$

and

$$\hat{\mathbf{w}}_{GMV;nSh} = \frac{\hat{\boldsymbol{\Sigma}}_n^{-1} \mathbf{1}_p}{\mathbf{1}_p^\top \hat{\boldsymbol{\Sigma}}_n^{-1} \mathbf{1}_p}, \quad (35)$$

where $\hat{\boldsymbol{\mu}}_n$ and $\hat{\boldsymbol{\Sigma}}_n$ denote improved estimators of the mean vector and of the covariance matrix, respectively. In the following, we refer to the optimal portfolios constructed by using (34) and (35) as the naive shrinkage estimator of the MV portfolio weights and of the GMV portfolio weights.

3.3. Optimal shrinkage approach

Although the estimators $\hat{\mathbf{w}}_{MV;nSh}$ and $\hat{\mathbf{w}}_{GMV;nSh}$ are constructed by using less volatile estimators of the mean vector and of the covariance matrix, they are not optimal in the sense that they maximize some loss functions. Moreover, it is also questionable, why one has to estimate $p(p+1)/2$ -dimensional and p -dimensional objects, while estimators for $(p-1)$ -dimensional vector of optimal portfolio weights are needed. We deal with this question in the present section by presenting shrinkage estimators for \mathbf{w}_{MV} and \mathbf{w}_{GMV} , which are directly derived for the portfolio weights.

In case of the MV optimal portfolio, the loss function is determined following the optimization problem used in its derivation. Namely, the out-of-sample expected utility function is considered which is given by

$$\hat{\mathbf{w}}_n^\top \boldsymbol{\mu} - \frac{\gamma}{2} \hat{\mathbf{w}}_n^\top \boldsymbol{\Sigma} \hat{\mathbf{w}}_n, \quad (36)$$

where $\hat{\mathbf{w}}_n$ denotes a shrinkage estimator of the MV portfolio weights obtained by a linear combination of its sample estimator $\hat{\mathbf{w}}_{MV;S}$ and the target vector of portfolio weights \mathbf{b} such that $\mathbf{b}^\top \mathbf{1}_p = 1$.

The maximization of (36) leads to the formula of the optimal shrinkage estimator of the MV portfolio expressed as ([23])

$$\hat{\mathbf{w}}_{MV;oSh} = \hat{\alpha}_{n;MV} \hat{\mathbf{w}}_{MV;S} + (1 - \hat{\alpha}_{n;MV}) \mathbf{b} \quad (37)$$

with

$$\hat{\alpha}_{n;MV} = \frac{\gamma^{-1} \left((\hat{R}_{GMV;S} - \hat{R}_{b;S}) \left(1 + \frac{1}{1-c} \right) + \gamma (\hat{V}_{b;S} - \hat{V}_{GMV;c}) + \frac{\gamma^{-1}}{1-c} \hat{s}_c \right)}{\frac{\hat{V}_{GMV;c}}{1-c} - 2 \left(\hat{V}_{GMV;c} + \frac{\gamma^{-1}}{1-c} (\hat{R}_{b;S} - \hat{R}_{GMV;S}) \right) + \gamma^{-2} \left(\frac{\hat{s}_c + c}{(1-c)^3} \right) + \hat{V}_{b;S}}, \quad (38)$$

where $\hat{R}_{GMV:S}$, $\hat{V}_{GMV:c}$, \hat{s}_c , $\hat{R}_{b:S}$, and $\hat{V}_{b:S}$ are consistent estimators of the three parameters of the efficient frontier given in (31), and of the expected return $R_b = \mu^\top \mathbf{b}$ and the variance $V_b = \mathbf{b}^\top \Sigma \mathbf{b}$ of the target portfolio. They are equal to

$$\hat{R}_{GMV:S} = \frac{\bar{\mathbf{y}}_n^\top \mathbf{S}_n^{-1} \mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_n^{-1} \mathbf{1}_p}, \quad \hat{V}_{GMV:c} = \frac{1}{1-c} \hat{V}_{GMV:S}, \quad \hat{s}_c = (1-c)\hat{s} - c, \quad \hat{R}_{b:S} = \hat{\mu}^\top \mathbf{b}, \quad \hat{V}_{b:S} = \mathbf{b}^\top \mathbf{S}_n \mathbf{b}$$

with

$$\hat{V}_{GMV:S} = \frac{1}{\mathbf{1}_p^\top \mathbf{S}_n^{-1} \mathbf{1}_p}, \quad \hat{s} = \bar{\mathbf{y}}_n^\top \hat{\mathbf{Q}}_n \bar{\mathbf{y}}_n.$$

The out-of-sample variance is considered as a loss function when the investor is fully risk averse, i.e., $\gamma = \infty$, (see, e.g., [26, 47]), that is in the case of the GMV portfolio. It is given by

$$\hat{\mathbf{w}}_n^\top \Sigma \hat{\mathbf{w}}_n, \quad (39)$$

where $\hat{\mathbf{w}}_n$ denotes a shrinkage estimator of the GMV portfolio weights defined obtained by a linear combination of its sample estimator $\hat{\mathbf{w}}_{GMV:S}$ and the target vector of portfolio weights \mathbf{b} such that $\mathbf{b}^\top \mathbf{1}_p = 1$.

The solution of (39) is given by (see, [26])

$$\hat{\mathbf{w}}_{GMV;osh} = \hat{\alpha}_{n;GMV} \hat{\mathbf{w}}_{GMV:S} + (1 - \hat{\alpha}_{n;GMV}) \mathbf{b} \quad (40)$$

with

$$\hat{\alpha}_{n;GMV} = \frac{\hat{V}_{b:S} - \hat{V}_{GMV:c}}{\frac{\hat{V}_{GMV:c}}{1-c} - 2\hat{V}_{GMV:c} + \hat{V}_{b:S}}. \quad (41)$$

Another shrinkage estimator of the GMV portfolio weights was developed by [47] and it is given by

$$\hat{\mathbf{w}}_{GMV:FM} = \hat{\alpha}_{n;FM} \hat{\mathbf{w}}_{GMV:S} + (1 - \hat{\alpha}_{n;FM}) \mathbf{b} \quad (42)$$

with

$$\hat{\alpha}_{n;FM} = \frac{p-3}{n-p+2} \frac{\hat{V}_{GMV:S}}{\mathbf{b}^\top \mathbf{S}_n \mathbf{b} - \hat{V}_{GMV:S}}. \quad (43)$$

3.4. Tests on mean-variance optimality of portfolios based on the shrinkage approach

In the previous subsection the point shrinkage estimators for the MV portfolio weights and for the GMV portfolio weights are established. In order to complete this discussion, the interval estimators of the weights of these optimal portfolios are discussed in this subsection. Using the one-to-one correspondence between the interval estimation and the test theory (see [1]), we present first high-dimensional asymptotic tests on the weights of optimal portfolios, and then show how these findings can be used to construct confidence regions for optimal portfolio weights.

Tests for general linear hypotheses imposed on the weights of the MV portfolio and of the GMV portfolio were suggested under the traditional asymptotic setting in [29] and [31], while they were extended to the high-dimensional asymptotic setting by [16]. Unfortunately, these approaches cannot be used to test the whole structure of a portfolio in a single step and allow to make inference only on a finite number of the components of the vector of portfolio weights. In order to deal with the problem, shrinkage-type tests were developed in [19] for the GMV portfolio and in [18] for the MV portfolio under the high-dimensional asymptotic regime.

For the MV portfolio, the goal is to test the hypotheses

$$H_0 : \mathbf{w}_{MV} = \mathbf{w}_0 \quad \text{against} \quad H_1 : \mathbf{w}_{MV} \neq \mathbf{w}_0, \quad (44)$$

i.e., that the portfolio with weights \mathbf{w}_0 is mean-variance efficient under H_0 .

It was shown in [23] that $\hat{\alpha}_{n;MV} \xrightarrow{a.s.} \alpha_{MV}(\mathbf{b})$ for $p/n \rightarrow c \in [0, 1)$ as $n \rightarrow \infty$ where

$$\alpha_{MV}(\mathbf{b}) = \frac{\gamma^{-1} \left((R_{GMV} - R_b) \left(1 + \frac{1}{1-c} \right) + \gamma (V_b - V_{GMV}) + \frac{\gamma^{-1}}{1-c} s \right)}{\frac{V_{GMV}}{1-c} - 2 \left(V_{GMV} + \frac{\gamma^{-1}}{1-c} (R_b - R_{GMV}) \right) + \gamma^{-2} \left(\frac{s+c}{(1-c)^2} \right) + V_b}. \quad (45)$$

Moreover, if \mathbf{w}_0 is mean-variance efficient for the considered risk aversion coefficient γ , then $\alpha_{MV}(\mathbf{w}_0) = 0$, i.e., the shrinkage intensity computed for the target portfolio \mathbf{w}_0 tends almost surely to zero in the high-dimensional asymptotic setting. This observation motivates the consideration of the following hypotheses

$$H_0 : \alpha_{MV}(\mathbf{w}_0) = 0 \quad \text{against} \quad H_1 : \alpha_{MV}(\mathbf{w}_0) \neq 0. \quad (46)$$

Furthermore, it was proved in [18] that the null hypothesis in (44) implies (46), i.e., the rejection of (46) will ensure that \mathbf{w}_0 portfolio is not mean-variance efficient.

For testing (46) the following test statistic was suggested in [18], expressed as

$$T_{MV}(\mathbf{w}_0) = \sqrt{n} \frac{\hat{\alpha}_{n;MV}(\mathbf{w}_0) \hat{B}_n(\mathbf{w}_0)}{\sqrt{\mathbf{d}_0' \hat{\Omega}(\mathbf{w}_0) \mathbf{d}_0}}, \quad (47)$$

where $\hat{\alpha}_{n;MV}(\mathbf{w}_0)$ is the optimal shrinkage intensity as defined in (45) with $\mathbf{b} = \mathbf{w}_0$,

$$\hat{B}_n(\mathbf{w}_0) = \frac{\hat{V}_{GMV;c}}{1-c} - 2 \left(\hat{V}_{GMV;c} + \frac{\gamma^{-1}}{1-c} (\hat{R}_{\mathbf{w}_0;S} - \hat{R}_{GMV;S}) \right) + \gamma^{-2} \left(\frac{\hat{s}_c + c}{(1-c)^3} \right) + \hat{V}_{\mathbf{w}_0;S},$$

$$\mathbf{d}_0 = \begin{pmatrix} \gamma^{-1} + \frac{\gamma^{-1}}{1-c} \\ -1 \\ \frac{\gamma^{-2}}{1-c} \\ -\gamma^{-1} - \frac{\gamma^{-1}}{1-c} \\ 1 \end{pmatrix},$$

and

$$\hat{\Omega}(\mathbf{w}_0) = \begin{pmatrix} \frac{\hat{V}_{GMV;c}(\hat{s}_c+1)}{1-c} & 0 & 0 & \hat{V}_{GMV;c} & -2\hat{V}_{GMV;c}(\hat{R}_{\mathbf{w}_0;S} - \hat{R}_{GMV;S}) \\ 0 & 2\frac{\hat{V}_{GMV;c}^2}{1-c} & 0 & 0 & 2\hat{V}_{GMV;c}^2 \\ 0 & 0 & 2\frac{((\hat{s}_c+1)^2+c-1)}{1-c} & 2(\hat{R}_{\mathbf{w}_0;S} - \hat{R}_{GMV;S}) & -2(\hat{R}_{\mathbf{w}_0;S} - \hat{R}_{GMV;S})^2 \\ \hat{V}_{GMV;c} & 0 & 2(\hat{R}_{\mathbf{w}_0;S} - \hat{R}_{GMV;S}) & \hat{V}_{\mathbf{w}_0;S} & 0 \\ -2\hat{V}_{GMV;c}(\hat{R}_{\mathbf{w}_0;S} - \hat{R}_{GMV;S}) & 2\hat{V}_{GMV;c}^2 & -2(\hat{R}_{\mathbf{w}_0;S} - \hat{R}_{GMV;S})^2 & 0 & 2\hat{V}_{\mathbf{w}_0;S}^2 \end{pmatrix}. \quad (48)$$

Under the null hypothesis in (46) it holds that for $p/n \rightarrow c \in [0, 1)$ as $n \rightarrow \infty$

$$T_{MV}(\mathbf{w}_0) \xrightarrow{d} \mathcal{N}(0, 1)$$

and, hence, the hypothesis that \mathbf{w}_0 are the weights of the MV optimal portfolio is rejected as soon as $|T_{MV}(\mathbf{w}_0)| > z_{1-\delta/2}$, where $z_{1-\delta/2}$ is the $(1 - \delta/2)$ quantile of the standard normal distribution.

Finally, using the correspondence between a statistical test and a confidence region (see, [11]), $(1 - \delta)$ confidence region for mean-variance optimal portfolios corresponding to risk aversion coefficient γ is given by

$$\Omega_{MV;1-\delta}(\mathbf{w}) = \left\{ \mathbf{w} \in \mathbb{R}^p : \mathbf{w}^\top \mathbf{1}_p = 1 \text{ and } |T_{MV}(\mathbf{w})| \leq z_{1-\delta/2} \right\}. \quad (49)$$

Similarly, a test on the weights of the GMV portfolio is constructed. For testing the hypothesis that a portfolio with weights \mathbf{w}_0 coincides with the GMV portfolio, i.e.,

$$H_0 : \mathbf{w}_{GMV} = \mathbf{w}_0 \quad \text{against} \quad H_1 : \mathbf{w}_{GMV} \neq \mathbf{w}_0, \quad (50)$$

we note that $\hat{\alpha}_{n;GMV} \xrightarrow{a.s.} \alpha_{GMV}(\mathbf{b})$ for $p/n \rightarrow c \in [0, 1)$ as $n \rightarrow \infty$ with (see, [26])

$$\alpha_{GMV}(\mathbf{b}) = \frac{V_b - V_{GMV}}{\frac{c}{1-c} V_{GMV} + (V_b - V_{GMV})}, \quad (51)$$

and, consequently, $\alpha_{GMV}(\mathbf{w}_0) = 0$ as soon as \mathbf{w}_0 is the GMV portfolio. Hence, the hypotheses in (50) can be rewritten as

$$H_0 : \alpha_{GMV}(\mathbf{w}_0) = 0 \quad \text{against} \quad H_1 : \alpha_{GMV}(\mathbf{w}_0) \neq 0. \quad (52)$$

For testing (52), the following test statistic was suggested [19]:

$$T_{GMV}(\mathbf{w}_0) = \sqrt{n} \frac{(1-c) \hat{L}_{\mathbf{w}_0}}{c + (1-c) \hat{L}_{\mathbf{w}_0}}, \quad (53)$$

where

$$\hat{L}_{\mathbf{w}_0} = (1-c) \mathbf{w}_0^\top \mathbf{S}_n \mathbf{w}_0 \mathbf{1}_p^\top \mathbf{S}_n^{-1} \mathbf{1}_p - 1.$$

Under the null hypothesis in (52) it holds that for $p/n \rightarrow c \in [0, 1)$ as $n \rightarrow \infty$

$$T_{GMV}(\mathbf{w}_0) \xrightarrow{d} \mathcal{N}\left(0, 2 \frac{1-c}{c}\right).$$

Then, the null hypothesis that \mathbf{w}_0 is the GMV optimal portfolio, is rejected as soon as $|T_{GMV}(\mathbf{w}_0)| > \sqrt{2 \frac{1-c}{c}} z_{1-\delta/2}$.

Finally, $(1 - \delta)$ confidence region for mean-variance optimal portfolios corresponding to risk aversion coefficient γ is given by

$$\Omega_{GMV;1-\delta}(\mathbf{w}) = \left\{ \mathbf{w} \in \mathbb{R}^p : \mathbf{w}^\top \mathbf{1}_p = 1 \text{ and } |T_{GMV}(\mathbf{w}_0)| > \sqrt{2 \frac{1-c}{c}} z_{1-\delta/2} \right\}. \quad (54)$$

3.5. Dynamic shrinkage approach

Recently, two dynamic shrinkage estimators for the weights of the GMV portfolio were developed in [27]. The first dynamic shrinkage estimation strategy corresponds to the case where non-overlapping samples are present, while the second strategy allows overlapping samples. Next, we describe both the approaches.

We consider an investor, who constructs a GMV portfolio at time t_1 by using the shrinkage estimator (40) with the target portfolio \mathbf{b} . The attention of the investor is to continue investing into the GMV portfolio over next T trading periods. Namely, the holding portfolio can be reconstructed at time points t_i for $i \in \{2, \dots, T\}$ as new information arrives on the capital market. This information is presented in this section by the sample of asset returns between t_{i-1} and t_i which is collected into the data matrix \mathbf{Y}_{n_i} . At each time point t_i the investor aims to continue investing in the GMV portfolio and uses the most recent information to update the holding portfolio. Since the transaction costs might be very large, the investor decides to shrink the traditional estimator of the GMV portfolio constructed by using data $\mathbf{Y}_{n_i} = [\mathbf{y}_{n_{i-1}+1}, \mathbf{y}_{n_{i-1}+2}, \dots, \mathbf{y}_{n_i}]$ with $n_0 = 0$ (non-overlapping case) and $\mathbf{Y}_{N_i} = [\mathbf{Y}_{n_1}, \mathbf{Y}_{n_2}, \dots, \mathbf{Y}_{n_i}]$ with $N_i = n_1 + \dots + n_i$ (overlapping case) to the weights of the holding portfolio constructed at time t_{i-1} .

Following model (1), it is assumed that

$$\mathbf{Y}_{n_i} = \mu \mathbf{1}_{n_i}^\top + \Sigma^{1/2} \mathbf{X}_{n_i}, \quad (55)$$

where \mathbf{X}_{n_i} is a $p \times n_i$ matrix which consists of independent and identically distributed random variables with zero mean, unit variance, and finite $4 + \varepsilon$, $\varepsilon > 0$, moments. No specific distributional assumption is imposed on the element of \mathbf{X}_{n_i} , $i \in \{1, \dots, T\}$. Furthermore, \mathbf{Y}_{n_i} , $i \in \{1, \dots, T\}$, are assumed to independent random matrices.

In the non-overlapping case, the sample of asset returns \mathbf{Y}_{n_i} is used to construct the traditional sample estimator of the GMV portfolio at each time t_i given by

$$\hat{\mathbf{w}}_{dS; n_i} = \frac{\mathbf{S}_{n_i}^{-1} \mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_{n_i}^{-1} \mathbf{1}_p}, \quad \mathbf{S}_{n_i} = \frac{1}{n_i - 1} \mathbf{Y}_{n_i} \left(\mathbf{I}_{n_i} - \frac{1}{n_i} \mathbf{1}_{n_i} \mathbf{1}_{n_i}^\top \right) \mathbf{Y}_{n_i}^\top. \quad (56)$$

Then, the shrinkage estimator of the weights of the GMV portfolio at time t_i is obtained by minimizing the out-of-sample variance, namely,

$$\hat{\mathbf{w}}_{n_i}^\top \Sigma \hat{\mathbf{w}}_{n_i}, \quad (57)$$

where $\hat{\mathbf{w}}_{n_i}$ is expressed as linear combination of $\hat{\mathbf{w}}_{S:n_i}$ and the holding portfolio determined at time t_{i-1} , that is $\hat{\mathbf{w}}_{dSh:n_{i-1}}$ for $i \in \{1, \dots, T\}$ with $\hat{\mathbf{w}}_{dSh:n_0} = \mathbf{b}$. The solution to the sequence of optimization problems (57) is given by (see, [27])

$$\hat{\mathbf{w}}_{dSh:n_i} = \hat{\psi}_{dSh:i} \hat{\mathbf{w}}_{S:n_i} + (1 - \hat{\psi}_{dSh:i}) \hat{\mathbf{w}}_{dSh:n_{i-1}} \quad (58)$$

with

$$\hat{\psi}_{dSh:i} = \frac{(n_i - p) \hat{r}_{i-1}}{(n_i - p) \hat{r}_{i-1} + p}, \quad (59)$$

where \hat{r}_i is computed recursively by

$$\hat{r}_i = \hat{\psi}_{dSh:i}^2 \frac{p}{n_i - p} + (1 - \hat{\psi}_{dSh:i})^2 \hat{r}_{i-1} \quad (60)$$

with

$$\hat{r}_0 = \left(1 - \frac{p}{n_1}\right) \mathbf{1}_p^\top \mathbf{S}_{n_1}^{-1} \mathbf{1}_p \mathbf{b}^\top \mathbf{S}_{n_1} \mathbf{b} - 1. \quad (61)$$

Similarly, in the overlapping case, the matrix of asset returns \mathbf{Y}_{N_i} is used to construct the traditional sample estimator of \mathbf{w}_{GMV} given by

$$\hat{\mathbf{w}}_{S:N_i} = \frac{\mathbf{S}_{N_i}^{-1} \mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_{N_i}^{-1} \mathbf{1}_p}, \quad \mathbf{S}_{N_i} = \frac{1}{N_i - 1} \mathbf{Y}_{N_i} \left(\mathbf{I}_{N_i} - \frac{1}{N_i} \mathbf{1}_{N_i} \mathbf{1}_{N_i}^\top \right) \mathbf{Y}_{N_i}^\top, \quad (62)$$

which is shrunk at time t_i to the holding portfolio weights $\hat{\mathbf{w}}_{dSh:N_{i-1}}$. The minimization of the out-of-sample variance at time t_i then leads to

$$\hat{\mathbf{w}}_{dSh:N_i} = \hat{\Psi}_{dSh:i} \hat{\mathbf{w}}_{S:N_i} + (1 - \hat{\Psi}_{dSh:i}) \hat{\mathbf{w}}_{dSh:N_{i-1}}, \quad (63)$$

for $p/N_j \rightarrow C_j \in (0, 1)$ as $N_j \rightarrow \infty$, $j \in \{1, \dots, i\}$ and $i \in \{1, \dots, T\}$, where

$$\hat{\Psi}_{dSh:i} = \frac{(\hat{R}_{i-1} + 1) - \hat{K}_i}{(\hat{R}_{i-1} + 1) + (1 - C_i)^{-1} - 2\hat{K}_i}, \quad (64)$$

with

$$\hat{R}_0 = \hat{r}_0, \quad \hat{R}_i = \hat{\Psi}_{dSh:i}^2 \frac{C_i}{1 - C_i} + (1 - \hat{\Psi}_{dSh:i})^2 \hat{R}_{i-1} + 2\hat{\Psi}_{dSh:i}(1 - \hat{\Psi}_{dSh:i})(\hat{K}_i - 1), \quad (65)$$

and

$$\hat{K}_i = \hat{\beta}_{i-1;0} + \sum_{j=1}^{i-1} \hat{\beta}_{i-1;j} D_{j,i}, \quad (66)$$

where

$$\hat{\beta}_{0;0} = 1, \quad \hat{\beta}_{i-1;i-1} = \hat{\Psi}_{dSh:i-1}, \quad \hat{\beta}_{i-1;j} = (1 - \hat{\Psi}_{dSh:i-1}) \hat{\beta}_{i-2;j}, \quad j \in \{0, \dots, i-2\} \quad (67)$$

and

$$D_{j,i} = 1 - \frac{2(1 - C_j)}{(1 - C_j) + (1 - C_i) \frac{C_j}{C_i} + \sqrt{\left(1 - \frac{C_j}{C_i}\right)^2 + 4(1 - C_i) \frac{C_j}{C_i}}}. \quad (68)$$

Although the dynamic shrinkage estimator of the GMV portfolio weights based on the overlapping sample is more computationally intensive, it possesses a great advantage with respect to non-overlapping samples since it requires that only $n_1 > p$. All other values of n_i can be smaller than p . In contrast, it is needed that all $n_i > p$ in the non-overlapping case. To this end, we note that the practical implementation of both dynamic shrinkage strategies are available in the R package *DOSPortfolio*. The presented dynamic approach for the GMV portfolio weights can further be extended to the MV portfolio with much more involved recursive formulas for the shrinkage intensities $\hat{\Psi}_{dSh:i}$.

4. Discussion and future directions of the research

Estimation of high-dimensional model parameters and functions of high-dimensional model parameters is a challenging task in modern statistical theory. Traditional approaches from frequentist statistics, like the maximum-likelihood estimation or method of moments estimation, do not provide a good answer to the problem by resulting in estimators which are very volatile in the high-dimensional setting.

The shrinkage approach has appeared to be a promising tool to reduce high volatility which is present in the traditional estimators of high-dimensional quantities. This is achieved by shrinking the traditional estimator to deterministic quantities. Although the procedure introduces bias in new estimators, it also considerably reduces the variance such that the mean-square error becomes smaller than the one of the corresponding traditional estimator.

In the present paper we review several shrinkage estimators for the mean vector, covariance matrix, precision matrix and for the weights of optimal portfolios which are functions of the mean-vector and covariance matrix. Although, most of the theoretical results related to the considered shrinkage estimators were in the high-dimensional asymptotic setting, they dealt with the case when the model dimension is smaller than the sample size. On the other hand, in many applications of biostatistics, the sample size is smaller than the model dimension. Under such a setup, the sample covariance matrix is singular and its inverse does not exist. One of the possible solutions is to replace the inverse of the sample covariance matrix by, for example, a generalized inverse or Moore-Penrose inverse (see, e.g., [14, 24, 55]), which will require the derivation of new asymptotic results that might lead to more complicated formulas of the shrinkage intensities. Another line of possible future research might be related to the development of sequential control procedures for monitoring changes in the high-dimensional parameters of stochastic models, like mean vector ([13]) or covariance matrix ([12]), or for sequential surveillance of optimal portfolio weights ([10]). New approaches can be based on the shrinkage approach and can extend the statistical tests on portfolio weights discussed in the present paper.

Acknowledgments

The authors would like to thank Professor Dietrich von Rosen and Professor Tõnu Kollo for their constructive comments that improved the quality of this paper. Olha Bodnar acknowledges valuable support from the internal grant (Rörlig resurs) of the Örebro University. Taras Bodnar was partially supported by the Swedish Research Council (VR) via the project *Bayesian Analysis of Optimal Portfolios and Their Risk Measures*.

References

- [1] J. Aitchison, Confidence-region tests, *Journal of the Royal Statistical Society: Series B (Methodological)* 26 (1964) 462–476.
- [2] G. W. Anderson, A. Guionnet, O. Zeitouni, *An Introduction to Random Matrices*, Cambridge University Press, Cambridge, 2010.
- [3] Z. D. Bai, B. Q. Miao, G. M. Pan, On asymptotics of eigenvectors of large sample covariance matrix, *The Annals of Probability* 35 (2007) 1532–1572.
- [4] Z. D. Bai, J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, Springer, New York, 2010.
- [5] A. J. Baranchik, A family of minimax estimators of the mean of a multivariate normal distribution, *The Annals of Mathematical Statistics* 41 (1970) 642–645.
- [6] D. Bauder, T. Bodnar, N. Parolya, W. Schmid, Bayesian mean–variance analysis: Optimal portfolio selection under parameter uncertainty, *Quantitative Finance* 21 (2021) 221–242.
- [7] J. Berger, M. E. Bock, L. D. Brown, G. Casella, L. Gleser, Minimax estimation of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix, *The Annals of Statistics* 5 (1977) 763–771.
- [8] J. O. Berger, M. E. Bock, Combining independent normal mean estimation problems with unknown variances, *The Annals of Statistics* 4 (1976) 642–648.
- [9] O. Bodnar, Sequential procedures for monitoring covariances of asset returns, in: *Advances in Risk Management*, Springer, 2007, pp. 241–264.
- [10] O. Bodnar, Sequential surveillance of the tangency portfolio weights, *International Journal of Theoretical and Applied Finance* 12 (2009) 797–810.
- [11] O. Bodnar, T. Bodnar, On the unbiased estimator of the efficient frontier, *International Journal of Theoretical and Applied Finance* 13 (2010) 1065–1073.
- [12] O. Bodnar, T. Bodnar, Y. Okhrin, Surveillance of the covariance matrix based on the properties of the singular Wishart distribution, *Computational Statistics & Data Analysis* 53 (2009) 3372–3385.
- [13] O. Bodnar, W. Schmid, CUSUM charts for monitoring the mean of a multivariate Gaussian process, *Journal of Statistical Planning and Inference* 141 (2011) 2055–2070.

- [14] T. Bodnar, H. Dette, N. Parolya, Spectral analysis of the Moore-Penrose inverse of a large dimensional sample covariance matrix, *Journal of Multivariate Analysis* 148 (2016) 160–172.
- [15] T. Bodnar, H. Dette, N. Parolya, Testing for independence of large dimensional vectors, *The Annals of Statistics* 47 (2019) 2977–3008.
- [16] T. Bodnar, H. Dette, N. Parolya, E. Thorsén, Sampling distributions of optimal portfolio weights and characteristics in low and large dimensions, *Random Matrices: Theory and Applications*, to appear (2021).
- [17] T. Bodnar, S. Dmytriv, Y. Okhrin, D. Otryakhin, N. Parolya, HDSHOP: High-Dimensional Shrinkage Optimal Portfolios, 2021. R package version 0.1.0.
- [18] T. Bodnar, S. Dmytriv, Y. Okhrin, N. Parolya, W. Schmid, Statistical inference for the expected utility portfolio in high dimensions, *IEEE Transactions on Signal Processing* 69 (2021) 1–14.
- [19] T. Bodnar, S. Dmytriv, N. Parolya, W. Schmid, Tests for the weights of the global minimum variance portfolio in a high-dimensional setting, *IEEE Transactions on Signal Processing* 67 (2019) 4479–4493.
- [20] T. Bodnar, A. K. Gupta, N. Parolya, On the strong convergence of the optimal linear shrinkage estimator for large dimensional covariance matrix, *Journal of Multivariate Analysis* 132 (2014) 215–228.
- [21] T. Bodnar, A. K. Gupta, N. Parolya, Direct shrinkage estimation of large dimensional precision matrix, *Journal of Multivariate Analysis* 146 (2016) 223–236.
- [22] T. Bodnar, O. Okhrin, N. Parolya, Optimal shrinkage estimator for high-dimensional mean vector, *Journal of Multivariate Analysis* 170 (2019) 63–79.
- [23] T. Bodnar, Y. Okhrin, N. Parolya, Optimal shrinkage-based portfolio selection in high dimensions, *Journal of Business & Economic Statistics*, under revision (2021).
- [24] T. Bodnar, N. Parolya, Spectral analysis of large reflexive generalized inverse and Moore-Penrose inverse matrices, in: T. Holgersson, M. Singull (Eds.), *Recent Developments in Multivariate and Random Matrix Analysis*, Springer, Cham, 2020, pp. 1–16.
- [25] T. Bodnar, N. Parolya, W. Schmid, On the equivalence of quadratic optimization problems commonly used in portfolio theory, *European Journal of Operational Research* 229 (2013) 637 – 644.
- [26] T. Bodnar, N. Parolya, W. Schmid, Estimation of the global minimum variance portfolio in high dimensions, *European Journal of Operational Research* 266 (2018) 371–390.
- [27] T. Bodnar, N. Parolya, E. Thorsén, Dynamic shrinkage estimation of the high-dimensional minimum-variance portfolio, *arXiv preprint arXiv:2106.02131* (2021).
- [28] T. Bodnar, M. Reiss, Exact and asymptotic tests on a factor model in low and large dimensions, *Journal of Multivariate Analysis* 150 (2016) 125–151.
- [29] T. Bodnar, W. Schmid, A test for the weights of the global minimum variance portfolio in an elliptical model, *Metrika* 67 (2008) 127–143.
- [30] T. Bodnar, W. Schmid, Econometrical analysis of the sample efficient frontier, *European Journal of Finance* 15 (2009) 317–335.
- [31] T. Bodnar, W. Schmid, On the exact distribution of the estimated expected utility portfolio weights: Theory and applications, *Statistics & Risk Modeling* 28 (2011) 319–342.
- [32] P. Bühlmann, S. van de Geer, *Statistics for High-Dimensional Data: Methods, Theory and Applications*, Springer Science & Business Media, Berlin, Heidelberg, 2011.
- [33] T. Cai, W. Liu, X. Luo, A constrained l_1 minimization approach to sparse precision matrix estimation, *Journal of the American Statistical Association* 106 (2011) 594–607.
- [34] T. T. Cai, J. Hu, Y. Li, X. Zheng, High-dimensional minimum variance portfolio estimation based on high-frequency data, *Journal of Econometrics* 214 (2020) 482–494.
- [35] T. T. Cai, M. Yuan, Adaptive covariance matrix estimation through block thresholding, *The Annals of Statistics* 40 (2012) 2014–2042.
- [36] S. X. Chen, L.-X. Zhang, P.-S. Zhong, Tests for high-dimensional covariance matrices, *Journal of the American Statistical Association* 105 (2010) 810–819.
- [37] D. Chételat, M. T. Wells, Improved multivariate normal mean estimation with unknown covariance when p is greater than n , *The Annals of Statistics* 40 (2012) 3137–3160.
- [38] D. K. Dey, C. Srinivasan, Estimation of a covariance matrix under Stein’s loss, *The Annals of Statistics* (1985) 1581–1591.
- [39] Y. Ding, Y. Li, X. Zheng, High dimensional minimum variance portfolio estimation under statistical factor models, *Journal of Econometrics* 222 (2021) 502–515.
- [40] B. Efron, C. Morris, Families of minimax estimators of the mean of a multivariate normal distribution, *The Annals of Statistics* (1976) 11–21.
- [41] J. Fan, Y. Fan, J. Lv, High dimensional covariance matrix estimation using a factor model, *Journal of Econometrics* 147 (2008) 186–197.
- [42] J. Fan, Y. Liao, M. Mincheva, Large covariance estimation by thresholding principal orthogonal complements, *Journal of the Royal Statistical Society: Series B* 75 (2013) 603–680.
- [43] J. Fan, J. Zhang, K. Yu, Vast portfolio selection with gross-exposure constraints, *Journal of American Statistical Association* 107 (2012) 592–606.
- [44] Y. Feng, D. P. Palomar, *A Signal Processing Perspective on Financial Engineering*, now Publishers Inc., Boston and Delft, 2016.
- [45] T. J. Fisher, X. Sun, C. M. Gallagher, A new test for sphericity of the covariance matrix for high dimensional data, *Journal of Multivariate Analysis* 101 (2010) 2554–2570.
- [46] D. Fourdrinier, W. E. Strawderman, M. T. Wells, Robust shrinkage estimation for elliptically symmetric distributions with unknown covariance matrix, *Journal of Multivariate Analysis* 85 (2003) 24 – 39.
- [47] G. Frahm, C. Memmel, Dominating estimators for minimum-variance portfolios, *Journal of Econometrics* 159 (2010) 289–302.
- [48] V. L. Girko, A. K. Gupta, Asymptotic behavior of spectral function of empirical covariance matrices, *Random Operators and Stochastic Equations* 2 (1994) 43–60.
- [49] L. J. Gleser, Minimax estimators of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix, *The Annals of Statistics* 14 (1986) 1625–1633.
- [50] V. Golosnoy, Y. Okhrin, Multivariate shrinkage for optimal portfolio weights, *The European Journal of Finance* 13 (2007) 441–458.
- [51] A. K. Gupta, T. Bodnar, An exact test about the covariance matrix, *Journal of Multivariate Analysis* 125 (2014) 176–189.

- [52] L. Haff, An identity for the Wishart distribution with applications, *Journal of Multivariate Analysis* 9 (1979) 531–544.
- [53] L. Haff, Empirical Bayes estimation of the multivariate normal covariance matrix, *The Annals of Statistics* 8 (1980) 586–597.
- [54] T. Holgersson, P. Karlsson, A. Stephan, A risk perspective of estimating portfolio weights of the global minimum-variance portfolio, *AStA Advances in Statistical Analysis* 104 (2020) 59–80.
- [55] S. Imori, D. von Rosen, On the mean and dispersion of the Moore-Penrose generalized inverse of a Wishart matrix, *The Electronic Journal of Linear Algebra* 36 (2020) 124–133.
- [56] J. Ingersoll, *Theory of Financial Decision Making*, Rowman & Littlefield, New Jersey, 1987.
- [57] W. James, C. Stein, Estimation with quadratic loss, *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* 1 (1961) 361–379.
- [58] J. D. Jobson, B. Korkie, Estimation for markowitz efficient portfolios, *Journal of the American Statistical Association* 75 (1980) 544–554.
- [59] R. Kan, D. R. Smith, The distribution of the sample minimum-variance frontier, *Management Science* 54 (2008) 1364–1380.
- [60] A. Kourtis, G. Dotsis, R. N. Markellos, Parameter uncertainty in portfolio selection: Shrinking the inverse covariance matrix, *Journal of Banking & Finance* 36 (2012) 2522–2531.
- [61] T. Kubokawa, M. S. Srivastava, Estimation of the precision matrix of a singular Wishart distribution and its application in high-dimensional data, *Journal of Multivariate Analysis* 99 (2008) 1906–1928.
- [62] L. Le Cam, G. Lo Yang, *Asymptotics in Statistics: Some Basic Concepts*, Springer, New York, 2000.
- [63] O. Ledoit, S. Peche, Eigenvectors of some large sample covariance matrix ensembles, *Probability Theory and Related Fields* 151 (2011) 233–264.
- [64] O. Ledoit, M. Wolf, A well-conditioned estimator for large-dimensional covariance matrices, *Journal of Multivariate Analysis* 88 (2004) 365–411.
- [65] O. Ledoit, M. Wolf, Analytical nonlinear shrinkage of large-dimensional covariance matrices, *The Annals of Statistics* 48 (2020) 3043 – 3065.
- [66] P.-E. Lin, H.-L. Tsai, Generalized Bayes minimax estimators of the multivariate normal mean with unknown covariance matrix, *The Annals of Statistics* 1 (1973) 142–145.
- [67] V. A. Marčenko, L. A. Pastur, Distribution of eigenvalues for some sets of random matrices, *Sbornik: Mathematics* 1 (1967) 457–483.
- [68] H. Markowitz, Portfolio selection, *The Journal of Finance* 7 (1952) 77–91.
- [69] H. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*, John Wiley & Sons, Inc and Chapman & Hall, Ltd., New York, 1959.
- [70] C. R. Merton, On estimating the expected return on the market: An exploratory investigation, *Journal of Financial Economics* 8 (1980) 323–361.
- [71] X. Mestre, M. Á. Lagunas, Finite sample size effect on minimum variance beamformers: Optimum diagonal loading factor for large arrays, *IEEE Transactions on Signal Processing* 54 (2005) 69–82.
- [72] Y. Okhrin, W. Schmid, Distributional properties of portfolio weights, *Journal of Econometrics* 134 (2006) 235–256.
- [73] M. Perlman, *STAT 542: Multivariate Statistical Analysis*, University of Washington (On-Line Class Notes), Seattle, Washington, 2007.
- [74] A. Rohde, A. B. Tsybakov, Estimation of high-dimensional low-rank matrices, *The Annals of Statistics* 39 (2011) 887–930.
- [75] F. Rubio, X. Mestre, Spectral convergence for a general class of random matrices, *Statistics & Probability Letters* 81 (2011) 592–602.
- [76] Y. Sheena, A. Takemura, Inadmissibility of non-order-preserving orthogonally invariant estimators of the covariance matrix in the case of Stein's loss, *Journal of Multivariate Analysis* 41 (1992) 117–131.
- [77] J. Silverstein, S. Choi, Analysis of the limiting spectral distribution of large dimensional random matrices, *Journal of Multivariate Analysis* 54 (1995) 295–309.
- [78] J. W. Silverstein, Z. D. Bai, On the empirical distribution of eigenvalues of a class of large-dimensional random matrices, *Journal of Multivariate Analysis* 54 (1995) 175–192.
- [79] M. Srivastava, Multivariate theory for analyzing high dimensional data, *Journal of The Japan Statistical Society* 37 (2007) 53–86.
- [80] C. Stein, Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, in: *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, University of California Press, Berkeley, Calif., 1956, pp. 197–206.
- [81] C. M. Stein, Estimation of the mean of a multivariate normal distribution, *The Annals of Statistics* (1981) 1135–1151.
- [82] C. Wang, G. Pan, T. Tong, L. Zhu, Shrinkage estimation of large dimensional precision matrix using random matrix theory, *Statistica Sinica* 25 (2015) 993–1008.
- [83] C. Wang, T. Tong, L. Cao, B. Miao, Non-parametric shrinkage mean estimation for quadratic loss functions with unknown covariance matrices, *Journal of Multivariate Analysis* 125 (2014) 222–232.
- [84] L. Wang, Z. Chen, C. D. Wang, R. Li, Ultrahigh dimensional precision matrix estimation via refitted cross validation, *Journal of Econometrics* 215 (2020) 118–130.
- [85] Y. Yang, G. Pan, et al., Independence test for high dimensional data based on regularized canonical correlation coefficients, *The Annals of Statistics* 43 (2015) 467–500.
- [86] M. Zhang, F. Rubio, X. Mestre, D. Palomar, Improved calibration of high-dimensional precision matrices, *IEEE Transactions on Signal Processing* 61 (2013) 1509–1519.
- [87] S. Zheng, Z. Chen, H. Cui, R. Li, et al., Hypothesis testing on linear structures of high-dimensional covariance matrix, *The Annals of Statistics* 47 (2019) 3300–3334.

The contribution of each author of the paper is proportional to the number of coauthors.

Journal Pre-proof