The Best Estimation for High-Dimensional Markowitz Mean-Variance Optimization

Bai Zhidong, Li Hua, Wing-Keung Wong
KLASMOE and School of Mathematics and Statistics, NENU
School of Sciences, Chang Chun University
Department of Economics, Hong Kong Baptist University

Abstract

The traditional (plug-in) return for the Markowitz mean-variance (MV) optimization has been demonstrated to seriously overestimate the theoretical optimal return, especially when the dimension to sample size ratio p/n is large. The newly developed bootstrap-corrected estimator corrects the overestimation, but it incurs the "under-prediction problem," it does not do well on the estimation of the corresponding allocation, and it has bigger risk. To circumvent these limitations and to improve the optimal return estimation further, this paper develops the theory of spectral-corrected estimation. We first establish a theorem to explain why the plug-in return greatly overestimates the theoretical optimal return. We prove that under some situations the plug-in return is $\sqrt{\gamma}$ times bigger than the theoretical optimal return, while under other situations, the plug-in return is bigger than but may not be $\sqrt{\gamma}$ times larger than its theoretic counterpart where $\gamma = \frac{1}{1-y}$ with y being the limit of the ratio p/n.

Thereafter, we develop the spectral-corrected estimation for the Markowitz MV model which performs much better than both the plug-in estimation and the bootstrap-corrected estimation not only in terms of the return but also in terms of the allocation and the risk. We further develop properties for our proposed estimation and conduct a simulation to examine the performance of our proposed estimation. Our simulation shows that our proposed estimation not only overcomes the problem of "over-prediction," but also circumvents the "under-prediction," "allocation estimation," and "risk" problems. Our simulation also shows that our proposed spectral-corrected estimation is stable for different values of sample size n, dimension p, and their ratio p/n. In addition, we relax the normality assumption in our proposed estimation so that our proposed spectral-corrected estimators could be obtained when the returns of the assets being studied could follow any distribution under the condition of the existence of the fourth moments.

Keywords: G11; C13

JEL Classification: Markowitz mean-variance optimization, Optimal Return, Optimal Portfolio Allocation, Large Random Matrix, Bootstrap Method.

1 Introduction

This paper aims to develop the best estimation for the problem of the high-dimensional Markowitz mean-variance (MV) portfolio optimization. Our proposed estimation may not be the best estimation, but we believe our approach at least enables academics and practitioners to get closer to obtaining the best estimation for the high-dimensional MV Markowitz optimization problem.

2 Markowitz's Mean-Variance Principle

To distinguish the well-known results in the literature from the ones derived in this paper, all cited results will be called *Propositions* and our derived results will be called *Theorems*. We first discuss Markowitz's MV optimization principle. We suppose that there are p-branch of assets whose returns are denoted by $\mathbf{r} = (\mathbf{r}_1, \cdots, \mathbf{r}_p)^{\mathbf{T}}$ with mean vector $\mu = (\mu_1, \cdots, \mu_p)^{\mathbf{T}}$ and covariance matrix $\Sigma = (\sigma_{ij})$. In addition, we assume that an investor will invest capital C on the p-branch of assets such that she wants to allocate her investable wealth to the assets to attain one of the following:

a. to maximize return subject to a given level of risk, or

b. to minimize risk for a given level of expected return.

Since the above two cases are equivalent, we consider only the first one in this paper. Without loss of generality, we assume $C = 1$ and her investment plan to be $\mathbf{c} = (c_1, \cdots, c_p)^T$. Hence, we have $\Sigma_{i=1}^{p} c_i \leq 1$ in which the strict inequality corresponds to the fact that the investor could invest only part of her wealth. Her anticipated return, R, will then be $\mathbf{c}^T\mu$ with risk $\mathbf{c}^T\Sigma\mathbf{c}$. In this paper, we further assume that short selling is allowed, and hence, any component of **c** could be positive as well as negative. Thus, the above maximization problem can be reformulated as:

$$\max \mathbf{c}^T \mu, \text{ subject to } \mathbf{c}^T \mathbf{1} \leq 1 \text{ and } \mathbf{c}^T \Sigma \mathbf{c} \leq \sigma_0^2 \tag{2.1}$$

where **1** represents the p-dimensional vector of ones and σ_0^2 is a given level of risk. We call $R = \max \mathbf{c}^T \mu$ satisfying (2.1) the **optimal return** and **c** its corresponding **allocation plan**. One could obtain the solution of (2.1) from the following proposition:

Proposition 2.1. *For the optimization problem shown in (2.1), the optimal return, R, and its corresponding investment plan,* **c***, are obtained as follows:*

a. *If*

$$\frac{\boldsymbol{1}^T \Sigma^{-1} \mu \sigma_0}{\sqrt{\mu^T \Sigma^{-1} \mu}} < 1, \tag{2.2}$$

then the optimal return, R, and corresponding investment plan, **c***, will be*

$$R = \sigma_0 \sqrt{\mu^T \Sigma^{-1} \mu} \tag{2.3}$$

and

$$\boldsymbol{c} = \frac{\sigma_0}{\sqrt{\mu^T \Sigma^{-1} \mu}} \Sigma^{-1} \mu \,. \tag{2.4}$$

b. *If*

$$\frac{\boldsymbol{1}^T \Sigma^{-1} \mu \sigma_0}{\sqrt{\mu^T \Sigma^{-1} \mu}} > 1 \,, \tag{2.5}$$

then the optimal return, R, and corresponding investment plan, **c***, will be*

$$R = \frac{\boldsymbol{1}^T \Sigma^{-1} \mu}{\boldsymbol{1}^T \Sigma^{-1} \boldsymbol{1}} + b \left(\mu^T \Sigma^{-1} \mu - \frac{\left(\boldsymbol{1}^T \Sigma^{-1} \mu\right)^2}{\boldsymbol{1}^T \Sigma^{-1} \boldsymbol{1}} \right) \tag{2.6}$$

and

$$\mathbf{c} = \frac{\Sigma^{-1} \boldsymbol{1}}{\boldsymbol{1}^T \Sigma^{-1} \boldsymbol{1}} + b \left(\Sigma^{-1} \mu - \frac{\boldsymbol{1}^T \Sigma^{-1} \mu}{\boldsymbol{1}^T \Sigma^{-1} \boldsymbol{1}} \Sigma^{-1} \boldsymbol{1} \right) \,, \tag{2.7}$$

where

$$b = \sqrt{\frac{\boldsymbol{1}^T \Sigma^{-1} \boldsymbol{1} \sigma_0^2 - 1}{\mu^T \Sigma^{-1} \mu \boldsymbol{1}^T \Sigma^{-1} \boldsymbol{1} - \left(\boldsymbol{1}^T \Sigma^{-1} \mu\right)^2}} \,.$$

3 Large Dimensional Random Matrix Theory

The main advantage of adopting the LDRMT is its ability to investigate the limiting spectrum properties of random matrices when the dimension increases proportionally with the sample size. This turns out to be a powerful tool in dealing with large dimensional data analysis.

4 Markowitz Mean-Variance Optimization Estimation

4.1 Plug-In Estimator

Plugging the sample mean $\overline{\mathbf{x}}$ and the sample covariance matrix S into the formulae of the asset allocation $\mathbf{c}$ in Proposition 2.1:

$$\begin{aligned} \widehat{R}_p &= \hat{\mathbf{c}}_p^T \overline{\mathbf{x}} \,, \\ \hat{\mathbf{c}}_p &= \begin{cases} \frac{S^{-1}\overline{\mathbf{x}}}{\sqrt{\overline{\mathbf{x}}^T S^{-1}\overline{\mathbf{x}}}} & \text{if} \quad \frac{\sigma_0 1^T S^{-1}\overline{\mathbf{x}}}{\sqrt{\overline{\mathbf{x}}^T S^{-1}\overline{\mathbf{x}}}} < 1, \\ \frac{S^{-1}1}{1^T S^{-1}1} + \hat{b}_p\left(S^{-1}\overline{\mathbf{x}} - \frac{1^T S^{-1}\overline{\mathbf{x}}}{1^T S^{-1}1} S^{-1}\mathbf{1}\right) & \text{if} \quad \frac{\sigma_0 1^T S^{-1}\overline{\mathbf{x}}}{\sqrt{\overline{\mathbf{x}}^T S^{-1}\overline{\mathbf{x}}}} > 1, \end{cases} \end{aligned} \tag{4.1}$$

in which

$$\hat{b}_p = \sqrt{\frac{1^T S^{-1} 1 \sigma_0^2 - 1}{\overline{\mathbf{x}}^T S^{-1}\overline{\mathbf{x}} 1^T S^{-1} 1 - (1^T S^{-1}\overline{\mathbf{x}})^2}} \,.$$

4.2 Bootstrap-Corrected Estimation

To circumvent this limitation, Bai, Liu, and Wong (2009, 2009a) propose a bootstrap technique to circumvent the limitation of the "plug-in" estimators. They use the parametric approach of the bootstrap methodology to avoid possible singularity of the covariance matrix estimation in the bootstrap sample.

4.3 Spectral-Corrected Estimators

In this section, we will first discuss how to estimate the eigenvalues of the population covariance matrix, and thereafter, we will develop the theory of the spectral-corrected estimators, which will circumvent all the four defects—the over-prediction phenomenon, the under-prediction problem, the allocation estimation problem, and the problem of big risk. We will discuss the details in the following subsections.

4.3.1 Estimation of the eigenvalues of the population covariance and the population covariance matrix

......

From the estimation of the S.D. H of Σ in the above steps, we obtain the eigenvalue estimators $\hat{a}_1 \geq \hat{a}_2 \geq \cdots \geq \hat{a}_p$. According to the spectral theory, we have

$$S = V\widetilde{\Lambda}V^T, \tag{4.2}$$

where $\widetilde{\Lambda} = \text{diag}(\tilde{\lambda}_1, \cdots, \tilde{\lambda}_p)$ with $\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \cdots \geq \tilde{\lambda}_p$ and the column vectors of V are the orthogonal eigenvectors of S with respect to $\tilde{\lambda}_1, \cdots, \tilde{\lambda}_p$. Suppose that $\widehat{\Lambda} = \text{diag}\{\hat{a}_1, \hat{a}_2, \cdots, \hat{a}_p\}$ in which $\hat{a}_1 \geq \hat{a}_2 \geq \cdots \geq \hat{a}_p$ are the estimations of the eigenvalues for matrix Σ; we put $\widehat{\Lambda}$ in equation (4.2) and obtain the **spectral-corrected covariance**

$$\widehat{\Sigma}_s = V\widehat{\Lambda}V^T. \tag{4.3}$$

The spectral-corrected covariance in (4.3) could be used in the development of the "best" optimal estimation. We will discuss the issue in the following subsections.

4.3.2 Estimation of the optimal return and allocation

After estimating the spectral-corrected covariance $\widehat{\Sigma}_s$ from (4.3) and from the steps discussed in Section 4.3.1, one could plug the sample mean vector $\overline{\mathbf{x}}$ and the spectral-corrected covariance $\widehat{\Sigma}_s$ into the formulae of the asset allocation $\mathbf{c}$ in Proposition 2.1 to obtain

$$\hat{\mathbf{c}}_s = \begin{cases} \dfrac{\sigma_0 \widehat{\Sigma}_s^{-1}\overline{\mathbf{x}}}{\sqrt{\overline{\mathbf{x}}^T\widehat{\Sigma}_s^{-1}\overline{\mathbf{x}}}} & \text{if} \quad \dfrac{\sigma_0 \mathbf{1}^T\widehat{\Sigma}_s^{-1}\overline{\mathbf{x}}}{\sqrt{\overline{\mathbf{x}}^T\widehat{\Sigma}_s^{-1}\overline{\mathbf{x}}}} < 1, \\ \dfrac{\widehat{\Sigma}_s^{-1}\mathbf{1}}{\mathbf{1}^T\widehat{\Sigma}_s^{-1}\mathbf{1}} + \hat{b}_s\left(\widehat{\Sigma}_s^{-1}\overline{\mathbf{x}} - \dfrac{\mathbf{1}^T\widehat{\Sigma}_s^{-1}\overline{\mathbf{x}}}{\mathbf{1}^T\widehat{\Sigma}_s^{-1}\mathbf{1}}\widehat{\Sigma}_s^{-1}\mathbf{1}\right) & \text{if} \quad \dfrac{\sigma_0 \mathbf{1}^T\widehat{\Sigma}_s^{-1}\overline{\mathbf{x}}}{\sqrt{\overline{\mathbf{x}}^T\widehat{\Sigma}_s^{-1}\overline{\mathbf{x}}}} > 1, \end{cases} \tag{4.4}$$

where

$$\hat{b}_s = \sqrt{\frac{\mathbf{1}^T\widehat{\Sigma}_s^{-1}\mathbf{1}\sigma_0^2 - 1}{\overline{\mathbf{x}}^T\widehat{\Sigma}_s^{-1}\overline{\mathbf{x}}\mathbf{1}^T\widehat{\Sigma}_s^{-1}\mathbf{1} - (\mathbf{1}^T\widehat{\Sigma}_s^{-1}\overline{\mathbf{x}})^2}}\,.$$

Since the estimator $\widehat{\Sigma}_s$ is obtained by estimating the eigenvalues of the population covariance, we call $\hat{\mathbf{c}}_s$ the **spectral-corrected allocation**. The corresponding return can be estimated by

$$\hat{R}_s = \hat{\mathbf{c}}_s^T\overline{\mathbf{x}}$$

which we call the **spectral-corrected return**. It can also be expressed as

$$\widehat{\mathbf{R}}_s = \begin{cases} \sigma_0 \sqrt{\overline{\mathbf{x}}^T \widehat{\Sigma}_s^{-1} \overline{\mathbf{x}}} & \text{if} \quad \frac{\sigma_0 \mathbf{1}^T \widehat{\Sigma}_s^{-1} \overline{\mathbf{x}}}{\sqrt{\overline{\mathbf{x}}^T \widehat{\Sigma}_s^{-1} \overline{\mathbf{x}}}} < 1 \,, \\ \frac{\overline{\mathbf{x}}' \widehat{\Sigma}_s^{-1} \mathbf{1}}{\mathbf{1}^T \widehat{\Sigma}_s^{-1} \mathbf{1}} + \hat{b}_s \left(\overline{\mathbf{x}}' \widehat{\Sigma}_s^{-1} \overline{\mathbf{x}} - \frac{\left(\mathbf{1}^T \widehat{\Sigma}_s^{-1} \overline{\mathbf{x}}\right)^2}{\mathbf{1}^T \widehat{\Sigma}_s^{-1} \mathbf{1}} \right) & \text{if} \quad \frac{\sigma_0 \mathbf{1}^T \widehat{\Sigma}_s^{-1} \overline{\mathbf{x}}}{\sqrt{\overline{\mathbf{x}}^T \widehat{\Sigma}_s^{-1} \overline{\mathbf{x}}}} > 1 \,. \end{cases} \tag{4.5}$$

In addition, the risk of the spectral-corrected allocation can be defined as

$$\begin{aligned} Risk_{\mathbf{c}}^s &= \hat{\mathbf{c}}_s^T \Sigma \hat{\mathbf{c}}_s \\ &= \begin{cases} \frac{\sigma_0^2 \overline{\mathbf{x}}' \widehat{\Sigma}_s^{-1} \Sigma \widehat{\Sigma}_s^{-1} \overline{\mathbf{x}}}{\overline{\mathbf{x}}^T \widehat{\Sigma}_s^{-1} \overline{\mathbf{x}}} & \text{if} \quad \frac{\sigma_0 \mathbf{1}^T \widehat{\Sigma}_s^{-1} \overline{\mathbf{x}}}{\sqrt{\overline{\mathbf{x}}^T \widehat{\Sigma}_s^{-1} \overline{\mathbf{x}}}} < 1 \,, \\ \left[\mathbb{A}^T + \hat{b}_s \left(\mathbb{B}^T + \mathbb{C}^T \right) \right] \Sigma \left[\mathbb{A} + \hat{b}_s \left(\mathbb{B} + \mathbb{C} \right) \right] & \text{if} \quad \frac{\sigma_0 \mathbf{1}^T \widehat{\Sigma}_s^{-1} \overline{\mathbf{x}}}{\sqrt{\overline{\mathbf{x}}^T \widehat{\Sigma}_s^{-1} \overline{\mathbf{x}}}} > 1 \,, \end{cases} \end{aligned} \tag{4.6}$$

which we call**spectral-corrected risk**. Here $\mathbb{A} = \frac{\widehat{\Sigma}_s^{-1} \mathbf{1}}{\mathbf{1}^T \widehat{\Sigma}_s^{-1} \mathbf{1}}$, $\mathbb{B} = \widehat{\Sigma}_s^{-1} \overline{\mathbf{x}}$ and $\mathbb{C} = \frac{\mathbf{1}^T \widehat{\Sigma}_s^{-1} \overline{\mathbf{x}}}{\mathbf{1}^T \widehat{\Sigma}_s^{-1} \mathbf{1}} \widehat{\Sigma}_s^{-1} \mathbf{1}$.

4.3.3 The limiting behavior of the spectral-corrected risk

5 Simulation Study

6 Conclusions

References

Bai, Z.D., Chen, J. and Yao, J.-F. 2010. On estimation of the population spectral distribution from a high-dimensional sample covariance matrix. Australian & New Zeland Journal of Statistics , 52, 423-437

Bai, Z.D., Hui, Y.C., Wong, W.K., Zitikis, R., 2012. Evaluating Prospect Performance: Making a Case for a Non-Asymptotic UMPU Test, Journal of Financial Econometrics 10(4), 703-732.

Bai, Z.D., Li, H., Liu, H.X., Wong, W.K., 2011. Test Statistics for Prospect and Markowitz Stochastic Dominances with Applications. Econometrics Journal 122, 1-26.

Bai, Z.D., Liu, H.X., Wong, W.K., 2009. Enhancement of the Applicability of Markowitz's Portfolio Optimization by Utilizing Random Matrix Theory. Mathematical Finance 19(4), 639-667.

Bai, Zhidong, Huixia Liu and Wing-Keung Wong, (2009a), On the Markowitz Mean-Variance Analysis of Self-Financing Portfolios, Risk and Decision Analysis 1(1), 35-42.

......