A Note on the Mean-Variance Analysis of Self-Financing Portfolios

Zhidong Bai
Department of Statistics
Northeast Normal University, China

Huixia Liu
Department of Statistics
National University of Singapore, Singapore

Wing-Keung Wong
Department of Economics
National University of Singapore, Singapore

September 15, 2006
A Note on the Mean-Variance Analysis of Self-Financing Portfolios

Abstract

This paper extends the work of Korkie and Turtle (2002) by first proving that the traditional estimate for the optimal return of self-financing portfolios always over-estimates from its theoretic value. To circumvent the problem, we develop a Bootstrap estimate for the optimal return of self-financing portfolios and prove that this estimate is consistent with its counterpart parameter. We further demonstrate the superiority of our proposed estimate over the traditional estimate by simulation.
I. Introduction

Markowitz (1952, 1959) introduce a mean-variance (MV) portfolio optimization procedure in which investors incorporate their preferences on risk and expectation of return to seek the best allocation of wealth by selecting the portfolios that maximize anticipated profit subject to achieving a specified level of risk, or equivalently minimize variance subject to achieving a specified level of expected gain. This theoretical MV optimization procedure is expected to be a powerful tool for portfolio optimizers to efficiently allocate their wealth to different investment alternatives to achieve the maximum expected profit. However, many studies (see, for example, Michaud (1989), Canner, Mankiw, and Weil (1997), Simaan (1997)) have found the MV-optimized portfolios to be unintuitive; thereby making their estimates to do more harm than good. Michaud (1989) refers to MV optimization process as one of the outstanding puzzles in modern finance, arguing that it has yet to meet with widespread acceptance by the investment community, particularly as a practical tool for active equity investment management. He terms this puzzle the “Markowitz optimization enigma” and calls the MV optimizers to be “estimation-error maximizers”.

This phenomenon also holds when we conduct the mean-variance analysis to obtain the optimal return from self-financing portfolios studied by Korkie and Turtle (2002) and others. In this paper, we complement their theoretical work by starting off to first prove that when the number of assets is large, the traditional return estimate for the optimal self-financing portfolio obtained by plugging the sample mean and covariance matrix into its theoretic value is always over-estimated and, in return, makes the self-financing MV optimization procedure impractical. We call this return estimate “plug-in” return and call this phenomenon “over-prediction.” This inaccuracy makes the self-financing MV optimization procedure to be impractical. To circumvent this over-prediction problem, we invoke the bootstrap technique to develop a new estimate that analytically corrects the over-prediction and reduces the error drastically. Furthermore, we theoretically prove that this bootstrap estimate is consistent to its counterpart parameter and thus our ap-
proach makes it a possibility to implement the optimization procedure; thereby making it practically useful. Our simulation further confirms the consistency of our proposed estimates; implying that the essence of the portfolio analysis problem could be adequately captured by our proposed estimates. Our simulation also shows that our proposed method improves the estimation accuracy so substantially that its relative efficiency could be as high as 205 times when compared with the traditional “plug-in” estimate for 350 number of assets with sample size of 500. Naturally, the relative efficiency will be much higher for bigger sample size and larger number of assets. The improvement of our proposed estimates are so big that there is a sound basis for believing our proposed estimate to be the best estimate to date for seeking the optimal return, yielding substantial benefits in terms of both profit maximization and risk reduction.

II. Theory

Suppose that there are \( p \)-branch of assets, \( S = (s_1, \ldots, s_p)^T \), whose returns are denoted by \( r = (r_1, \ldots, r_p)^T \) with mean \( \mu = (\mu_1, \ldots, \mu_p)^T \) and covariance matrix \( \Sigma = (\sigma_{ij}) \). In addition, we suppose that an investor will invest capital \( C \) on the \( p \)-branch of assets \( S \) such that s/he wants to allocate her/his investable wealth on the assets to obtain any of the following:

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

2. to minimize her/his risk for a given level of expected return.

Since the above two problems are equivalent, we only look for solution to the first problem in this paper. To obtain self-financing portfolio, we have \( C = 0 \). As her/his investment plan (or assets allocation) to be \( c = (c_1, \ldots, c_p)^T \), we have \( \sum_{i=1}^p c_i = C = 0 \). Also, her/his anticipated return, \( R \), will then be \( c^T \mu \) with risk \( c^T \Sigma c \). In this paper, we further assume that short selling is allowed and hence any component of \( c \) could be negative. Thus, the above maximization problem can be re-formulated to the following
optimization problem:

\[
\max R = c^T \mu, \text{ subject to } c^T 1 = 0 \text{ and } c^T \Sigma c \leq \sigma_0^2
\]  

where \( I \) represents a vector of ones with length conforming to the rules of matrix algebra and \( \sigma_0^2 \) is a given risk level. We call \( R \) satisfying (1) to be optimal return for the corresponding self-financing portfolio. The solution of (1) can be obtained in the following theorem:

**Theorem 1** For the optimization problem shown in (1), the optimal return, \( R \), and its corresponding investment plan, \( c \), are, respectively,

\[
R = \sigma_0 \sqrt{\mu^T \Sigma^{-1} \mu - \frac{(I^T \Sigma^{-1} \mu)}{I^T \Sigma^{-1} I}^2}
\]  

and

\[
c = \frac{\sigma_0 \sqrt{\mu^T \Sigma^{-1} \mu - \frac{(I^T \Sigma^{-1} \mu)}{I^T \Sigma^{-1} I}^2}}{(\Sigma^{-1} \mu - \frac{I^T \Sigma^{-1} \mu}{I^T \Sigma^{-1} I})}.
\]

The proof of Theorem 1 is straightforward. The set of efficient feasible portfolios for all possible levels of portfolio risk forms the self-financing MV efficient frontier. For any given level of risk, Theorem 1 seems to provide us a unique optimal return with its corresponding self-financing MV-optimal investment plan or asset allocation to represent the best investment alternative given the selected assets, and thus it seems to provide a solution to the self-financing MV optimization procedure. Nonetheless, it is easy to expect the problem to be straightforward; however, this is not so as the estimation of the optimal self-financing return is a difficult task.

Suppose that \( \{x_{jk}\} \) for \( j = 1, \cdots, p \) and \( k = 1, \cdots, n \) is a set of double array of independent and identically distributed (iid) random variables with mean zero and variance
\[ \sigma^2 \]. Let \( x_k = (x_{1k}, \cdots, x_{pk})^T \) and \( X = (x_1, \cdots, x_n) \). Then, the traditional return estimate (We call it plug-in return in this paper) for the optimal self-financing portfolio is obtained by “plugging-in” the sample mean vector \( \overline{X} \) and the sample covariance matrix \( S \) into the formulae in Theorem 1 such that

\[ \hat{R}_p = \hat{c}_p^T \overline{X} \]  

(3)

where

\[ \overline{X} = \frac{1}{n} \sum_{k=1}^{n} x_k / n , \quad S = \frac{1}{n-1} \sum_{k=1}^{n} (x_k - \overline{X})(x_k - \overline{X})^T \]  

(4)

and

\[ \hat{c}_p = \frac{\sigma_0}{\sqrt{X^T S^{-1} X - (1^T S^{-1} X)^2}} (S^{-1} X - 1^T S^{-1} \frac{1}{1^T S^{-1} 1} 1) . \]

Our simulation results shown in Figure 1 show that this plug-in return is far from satisfaction, especially when sample size become large. From the figure, we find that the plug-in return is always larger than its theoretical value. In this paper, we prove this over-prediction phenomenon by the following theorem:

**Theorem 2** Assume that \( X_1, \cdots, X_n \) are normally distributed with \( p \times 1 \) mean vector \( \mu \) and \( p \times p \) covariance matrix \( \Sigma \). If the dimension-to-sample-size ratio index \( p/n \to y \in (0, 1) \), we have

\[ \frac{\mu^T \Sigma^{-1} \mu}{n} \to a_1 , \quad \frac{1^T \Sigma^{-1} 1}{n} \to a_2 , \quad \frac{1^T \Sigma^{-1} \mu}{n} \to a_3 , \]

and if \( a_1 a_2 - a_3^2 > 0 \), then, we have

\[ \lim_{n \to \infty} \frac{\hat{R}_p}{\sqrt{n}} \to \sigma_0 \sqrt{\gamma (a_1 a_2 - a_3^2) / a_2} > \lim_{n \to \infty} \frac{R}{\sqrt{n}} = \sigma_0 \sqrt{a_1 a_2 - a_3^2 / a_2} , \]

where \( \gamma = \int_a^b \frac{1}{x} dF_y(x) = \frac{1}{1-y} > 1 \), \( a = (1 - \sqrt{y})^2 \) and \( b = (1 + \sqrt{y})^2 \).
The proof of Theorem 2 can be obtained by applying some fundamental limit theorems in the theory of large dimensional random matrix theory (see, for example, Jonsson (1982), Bai and Yin (1993) and Bai (1999)). From this theorem, we know that the plug-in return, $\hat{R}_p$, is always bigger than the theoretical optimal return, $R = c^T \mu$, defined in (1).

Employing the bootstrap technique, in this paper we further develop an efficient estimator for the optimal self-financing return to circumvent this over-prediction problem. We now describe the procedure to construct a parametric bootstrap estimate from the estimate of plug-in return, $\hat{R}_p$, defined in (3) as follows: First, we draw a resample $\chi^* = \{X^*_1, \ldots, X^*_n\}$ from the $p$-variate 'normal distribution' with mean vector $\bar{X}$ and covariance matrix $S$. Then, by invoking the self-financing optimization procedure again on the resample $\chi^*$, we obtain the bootstrapped “plug-in” allocation, $\hat{c}_p^*$, and the bootstrapped “plug-in” return, $\hat{R}_p^*$, such that

$$\hat{R}_p^* = \hat{c}_p^T \bar{X}^*$$

(5)

where $\bar{X}^* = \frac{1}{n} \sum_1^n X^*_i$. Thereafter, we obtain the bootstrap corrected return estimate $\hat{R}_b$ as shown in the following theorem:

**Theorem 3** Under the conditions in Theorem 2 and using the bootstrap correction procedure described above, the bootstrap corrected return estimate, $\hat{R}_b$, given by:

$$\hat{R}_b = \hat{R}_p + \frac{1}{\sqrt{\gamma}}(\hat{R}_p - \hat{R}_p^*).$$

(6)

possesses the following property:

$$\sqrt{\gamma}(R - \hat{R}_p) \approx \hat{R}_p - \hat{R}_p^*$$

(7)

where $\gamma$ is defined in Theorem 2, $R$ is the theoretic optimal self-financing return obtained
from Theorem 1, $\hat{R}_p$ is the plug-in return estimate defined in (3) and obtained by using the original sample $\chi$, and $\hat{R}^*_p$ is the bootstrapped plug-in return estimate defined in (5) and obtained by using the bootstrapped sample $\chi^*$ respectively.

The proof of Theorem 3 can be easily obtained by applying Theorem 2 and some fundamental limit theorems in the theory of large dimensional random matrix theory (see, for example, Jonsson (1982), Bai and Yin (1993) and Bai (1999)). We note that the relation $A \simeq B$ means that $A/B \to 1$ in the limiting procedure. The above theorem has shown that the bootstrap corrected return estimate, $\hat{R}_b$, is consistent to the theoretic optimal self-financing return, $R$.

III. Simulation Study

To illustrate the over-prediction problem, we, for simplicity sake, generate a $p$-branch of standardized asset returns from a multivariate normal distribution with mean $\mu = (\mu_1, \cdots, \mu_p)^T$ and identity covariance matrix $\Sigma = (I_{jk})$ in which $I_{jk} = 1$ when $j = k$ and $I_{jk} = 0$ otherwise. Given the level of risk, the known population mean vector, $\mu$, and the known population covariance matrix, $\Sigma$, we can compute the theoretic optimal allocation, $c$, and, thereafter, compute the theoretic optimal return, $R$, for the self-financing portfolios. These values will then be used to compare the performance of all the estimators being studied in our paper. Using this dataset, we apply the formula in (4) to compute the sample mean, $\overline{X}$, and sample covariance, $S$, which, in turn, enable us to obtain the plug-in return, $\hat{R}_p$, and its corresponding plug-in allocation, $\hat{c}_p$, by substituting $\overline{X}$ and $S$ into $\mu$ and $\Sigma$ respectively in the formula of $\hat{R}_p$ and $\hat{c}_p$ as shown in (3). To illustrate the over-prediction problem, we first plot the theoretic optimal self-financing returns, $R$, and the plug-in returns, $\hat{R}_p$, for different values of $p$ with the same sample size $n = 500$ in Figure 1.

From Figure 1, we find the following: (1) the plug-in return $\hat{R}_p$ defined in (3) is a good
estimate to the theoretic optimal return $R$ when $p$ is small ($\leq 30$); (2) when $p$ is large ($\geq 60$), the difference between the theoretic optimal return $R$ and the plug-in return $\hat{R}_p$ becomes dramatically large; (3) the larger the $p$, the greater the difference; and (4) when $p$ is large, both the plug-in return $\hat{R}_p$ is always larger than the theoretic optimal return, $R$, computed by using the true mean and covariance matrix. These confirm the “Markowitz optimization enigma” that the plug-in returns $\hat{R}_p$ should not be used in practice.

In order to show the superiority of the performance of $\hat{R}_b$ over that of plug-in return $\hat{R}_p$, we define the bootstrap corrected difference, $d^R_b (= \hat{R}_b - R)$, for the return estimate to be the difference between the bootstrap corrected optimal return estimate $\hat{R}_b$ and the theoretic optimal return $R$ and the plug-in difference, $d^R_p (= \hat{R}_p - R)$, for the return estimates $\hat{R}_p$ and $\hat{R}_b$ respectively. We then simulate 1000 times to compute $d_x^R$ for $x = p$ and $b$, $n = 500$ and $p = 100, 200$ and $300$. The results are depicted in Figure 2.

From Figure 2, we find the desired property for our proposed estimate that $d^R_b$ is much smaller than $d^R_p$ in absolute values for all the cases. This infers that the values of bootstrap corrected method are much more accurate in estimating the theoretic value than those obtained by using the plug-in procedure. Furthermore, as $p$ increases, the two lines on each level as shown in Figure 2 further separated from each other; implying that the magnitude of improvement from $d^R_p$ to $d^R_b$ are remarkable.

To further illustrate the superiority of our estimate over the traditional plug-in estimate, we present in Table 1 the mean square errors (MSEs) of the different estimates for different $p$ and present their relative efficiencies (RE) for returns such that

$$RE_{p,b}^R = \frac{MSE(d^R_p)}{MSE(d^R_b)}.$$  \hfill (8)

Comparing the MSE of $d^R_b$ with that of $d^R_p$ in Figure 2 and Table 1, $d^R_b$ has been reduced dramatically from those of $d^R_p$, indicating that our proposed estimates are superior. We find that the MSE of $d^R_b$ is only 0.05, improving 3.4 times from that of $d^R_p$ when $p = 50$. 

8
When the number of assets increases, the improvement become much more substantial. For example, when \( p = 350 \), the MSE of \( d_R^b \) is only 1.23 but the MSE of \( d_R^p \) is 252.89; improving 205.60 times from that of \( d_R^p \). This is an unbelievable improvement. We note that when both \( n \) and \( p \) are bigger, the relative efficiency of our proposed estimate over the traditional plug-in estimate could be much larger.

IV. Conclusions

Holding both theoretical and practical interest, the basic problem for MV analysis is to identify those combinations of assets that constitute attainable efficient portfolios. The purpose of this paper is to solve this problem by developing a new optimal return estimate to capture the essence of self-financing portfolio selection. With this in mind, we go on to first theoretically prove that the estimated plug-in return obtained by plugging in the sample mean and sample covariance into the formulae of the self-financing optimal return is inadequate as it is always larger than its theoretical value when the number of assets is large. To recall, we call this problem “over-prediction.” To circumvent this problem, we develop the new estimator, the bootstrap corrected return, for the theoretic self-financing optimal return by employing the parametric bootstrap method.

Our simulation results confirm that the essence of the self-financing portfolio analysis problem could be adequately captured by our proposed bootstrap correction method which improves the accuracy of the estimation dramatically. As our approach is easy to operate and implement in practice, the whole efficient frontier of our estimates can be constructed analytically. Thus, our proposed estimator facilitates an unquestionable ease of implementing the self-financing MV optimization procedure, and ensuring it is practically useful.

We note that our model includes the situation in which one of the assets is a riskless asset so that investors can risklessly lend and borrow at the same rate. In this situa-
tion, the separation theorem holds and thus our proposed return estimate is the optimal combination of the riskless asset and the optimal self-financing risky portfolio. We also note that the other assets listed in our model could be common stocks, preferred shares, bonds and other types of assets so that the optimal return estimate proposed in our paper actually represents the optimal self-financing return for the best combination of riskless rate, bonds, stocks and other assets.
References


Figure 1: Empirical and theoretical optimal returns for different numbers of assets

![Graph showing empirical and theoretical optimal returns.]

Solid line — the theoretic optimal return ($R$); Dotted line—the plug in return ($\hat{R}_p$).
Figure 2: Comparison between the Empirical and Corrected Portfolio Allocations and Returns

Table 1: MSE and Relative Efficiency Comparison

<table>
<thead>
<tr>
<th>p</th>
<th>$\text{MSE}(d^R_p)$</th>
<th>$\text{MSE}(d^R_b)$</th>
<th>$RE^R_{p,b}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.17</td>
<td>0.05</td>
<td>3.40</td>
</tr>
<tr>
<td>100</td>
<td>1.64</td>
<td>0.11</td>
<td>14.91</td>
</tr>
<tr>
<td>150</td>
<td>5.22</td>
<td>0.15</td>
<td>34.80</td>
</tr>
<tr>
<td>200</td>
<td>17.17</td>
<td>0.28</td>
<td>61.32</td>
</tr>
<tr>
<td>250</td>
<td>45.81</td>
<td>0.50</td>
<td>91.62</td>
</tr>
<tr>
<td>300</td>
<td>94.42</td>
<td>0.59</td>
<td>160.03</td>
</tr>
<tr>
<td>350</td>
<td>252.89</td>
<td>1.23</td>
<td>205.60</td>
</tr>
</tbody>
</table>