

# Theoretical Analysis and Application on Factor Analysis-Shrinkage Method to Portfolio

Huina Zhang, Yumei Li\*

Department of Mathematics, School of Science, Beijing Technology and Business University, Beijing  
Email: [zh\\_huina@126.com](mailto:zh_huina@126.com), [liwjyumei@163.com](mailto:liwjyumei@163.com)

Received: Jul. 3<sup>rd</sup>, 2015; accepted: Jul. 23<sup>rd</sup>, 2015; published: Jul. 30<sup>th</sup>, 2015

Copyright © 2015 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

In the portfolio with no short sales, investors usually want to get the optimal portfolio model in order to achieve maximum benefits and minimum risk. In this paper, we research the Markowitz model and the combination model of resampling and Markowitz model. However, due to the distribution of investment targets is not clear, the resampling method may overestimate the distribution of population. Then the application of shrinkage method has improved the variance of estimation. Based on shrinkage method and factor analysis, factor analysis-shrinkage method is proposed to further optimize the estimation. Experimental results using the real data of stocks and R software verify the analysis factor-shrinkage method.

## Keywords

Portfolio, Markowitz Model, Shrinkage Method, Factor Analysis-Shrinkage Method, R Software

---

# 因子分析-Shrinkage法在投资组合中的理论分析及应用

张慧娜, 李裕梅\*

北京工商大学理学院数学系, 北京  
Email: [zh\\_huina@126.com](mailto:zh_huina@126.com), [liwjyumei@163.com](mailto:liwjyumei@163.com)

\*通讯作者。

收稿日期：2015年7月3日；录用日期：2015年7月23日；发布日期：2015年7月30日

## 摘要

在不允许卖空的投资组合中，投资者往往希望通过投资组合模型得到最优投资组合以达到收益最大且风险最小的目标。本文对马科维茨模型以及重抽样和马科维茨模型相结合的模型进行了研究，但由于投资对象的总体分布不明确，用重抽样方法来估计总体分布可能会高估。Shrinkage法的应用缩减了估计量的方差。基于Shrinkage法及因子分析的思想，本文提出因子分析-Shrinkage法进一步优化估计量。用R软件对股票数据进行了实例分析与验证。

## 关键词

投资组合，马科维茨模型，Shrinkage法，因子分析-Shrinkage法，R软件

## 1. 引言

马科维茨投资组合理论[1]是现代投资组合理论的奠基石，也是金融理论的基础。投资组合理论主要研究怎样在未来结果不确定的情况下对有限资产进行投资，以使预期收益最大化和风险最小化得到最好的权衡。该问题一直是金融界研究的热点，吸引了许多优秀学者在这方面开展工作[2] [3]。在马科维茨模型的基础上，重抽样法、贝叶斯估计法、Shrinkage 法等[4]在投资组合方面得到了比较好的应用。在投资过程中，投资者往往希望尽可能规避风险而利益最大化，或是利益一定时最小化风险，利益和风险的最佳权衡就是有效的最优投资组合[5]。本文基于 Shrinkage 法的思想对其改进，得到一种更优的减小估计量的方法，即因子分析-Shrinkage 法，并用实例进行了验证。本文实验都是以 R 软件[6]来实现的。

## 2. 马科维茨模型介绍及实例分析

### 2.1. 马科维茨模型介绍

马科维茨投资组合理论认为，投资者是规避风险的，总是追求高的预期收益，他们不愿意承担没有相应的预期收益加以补偿的额外风险。马科维茨模型根据风险分散的原理[7]，应用二次规划的数学方法，揭示了投资组合的有效边界，及如何找出满足风险水平一定下获得最大的收益，或是预期收益一定下如何最小化投资风险的组合。

假设对投资者来说，可以选择  $N$  种带风险的资产，用  $R_i$  来表示第  $i$  个风险投资的收益率，其预期收

益率用  $\mu_i$  表示。把  $N$  种带风险的资产的收益率用向量形式表示： $R = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_N \end{pmatrix}$ ，预期收益的向量形式为：

$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{pmatrix}$ ，矩阵  $R$  的协方差矩阵用  $\Sigma$  表示。一般要求，各资产的权重用  $W_i$  表示，且满足  $W_1 + W_2 + \cdots + W_N = 1$ 。

于是，投资组合的预期收益是  $\sum_{i=1}^N W_i \mu_i = W^T \mu$ 。因此，投资组合的方差可以表示为： $W^T \Sigma W$ 。

假设投资组合的有效收益是获得一定收益时，投资的风险最小。投资组合的风险用其方差来刻画。给定一个预期收益  $\mu_T$ ，则二次规划模型的目标函数为  $\min W^T \Sigma W$ ，为了适用 R 软件中回归模型函数的求解，将目标函数变形为  $\min \frac{1}{2} W^T (2\Sigma) W$ 。在不允许卖空条件下，该模型所表示的最优投资组合[8]应满足：

$$\begin{aligned} \text{目标函数: } & \min \frac{1}{2} W^T (2\Sigma) W \\ \text{约束条件: } & \begin{cases} W^T \mu = \mu_T \\ W_1 + W_2 + \dots + W_N = 1 \\ W_i > 0 \end{cases} \end{aligned}$$

一般情况下，为了分散风险，投资者在进行投资选择时，不仅会选择高风险高利率的资产，还会选择一些无风险的资产进行投资来规避不必要的风险，像市场上所发行的基金、国债等都是无风险投资项目的最佳代表。马科维茨模型从理论上给出了最优投资组合的合理解释，但是求解过程缺乏可操作性，于是斯坦福大学威廉夏普教授提出了资本市场线模型理论，能更快的找到风险资产和无风险资产之间的线性关系，通过夏普比率揭示了有无风险投资的最佳组合[9]。夏普比率越大则说明该投资组合越好，反之越差。用数学模型表示为：

$$E(R_T) = \mu_f + \left( \frac{E(R_T) - \mu_f}{\sigma_{R_T}} \right) \sigma_{R_T}$$

其中， $\mu_f$  表示无风险资产的收益， $E(R_T)$  表示有风险投资组合的期望， $\sigma_{R_T}$  表示有风险投资组合的方差， $\frac{E(R_T) - \mu_f}{\sigma_{R_T}}$  表示夏普比率。

## 2.2. 实例分析

假设某位投资者选择了如下 5 只带有风险的股票，分别是 BGGF (包钢股份)、ZGSH (中国石化)、TYZG (太原重工)、ZGJZ (中国建筑)、ZJKY (紫金矿业)，并选取了这 5 只股票在 2014 年全年的日收益率情况分析。同时，为了规避一定风险，还选择了一只年收益率为 1.3% 的基金。利用马科维茨模型得到最佳投资组合(见图 1)。图 1 表示不允许卖空条件下，五只风险股票和一只无风险股票的最佳投资组合，实线表示有效边界，虚线表示有效投资组合，星号表示无风险资产和切点投资组合，加号表示最小方差投资组合。

由图 1 可知，F 点是无风险投资点，P 点以上的实线表示有效投资组合，T 点是有风险投资和无风险投资组合的切点，该点的夏普比率为 0.222，即夏普比率最大的点，也就是 5 只风险股票和一只无风险股票组合的最优点。此时，该五只有风险股票的投资组合系数见表 1。

由表 1 结果可知，在仅有 5 只带风险股票投资组合下，可以将资金的 5.38% 分配给包钢股份，20.41% 分配给中国石化，25.93% 分配给太原重工，32.38% 分配给中国建筑，15.89% 投资给紫金矿业。于是，由 5 只股票的预期期望和对应股票的投资系数可以得到对整个风险股票的组合收益为 0.395。由该结果和无风险的基金进行组合。若想要获得 20% 的收益，则需要 5 只有风险股票上的投资总资产的 49%，剩余 51% 投资到无风险资产上。

## 3. 重抽样对马科维茨模型的优化

马科维茨模型虽然得到了广泛的认可和应用，但是它对模型输入参数的变动比较敏感，微小的波动

就会对结果优化产生较大偏差。投资组合的重抽样能很好的解决这个问题，通过重复抽样的方法增加样本的信息量，从而大大降低马科维茨模型对输入参数的敏感性[10]。

重抽样的主要思想是，把原始数据当作总体，把重抽样得到的数据当作样本，通过重抽样得到的样本来估计总体的分布。一般来说，重抽样的次数越多，得到的样本信息量越大，拟合总体的分布越好。由上例得到其最大夏普比率为 0.222，这个比率比理论夏普比率高。通过重抽样的方法，利用箱线图、实际夏普比率和估计夏普比率，可以看到高估的程度(见图 2)。图 2 表示在不允许卖空条件下，利用重抽样方法得到的估计夏普比高估理论夏普比的程度，红色虚线代表理论夏普比。左边箱线图表示经 500 次重抽样得到的切线投资组合实际夏普比与理论夏普比之间相差程度，右边箱线图表示经 500 次重抽样的得

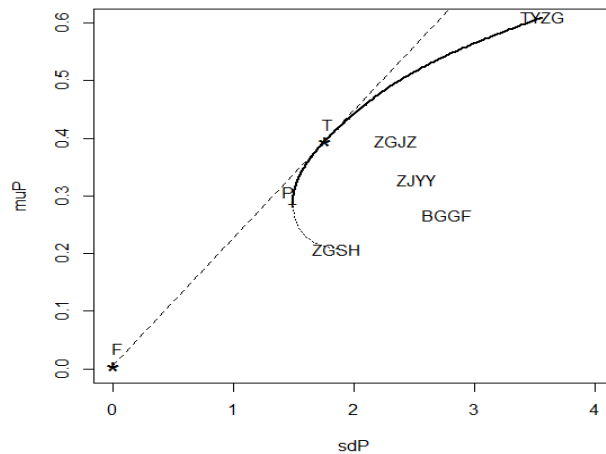


Figure 1. The optimal portfolio with short sales prohibited  
图 1. 不允许卖空条件下的最佳投资组合

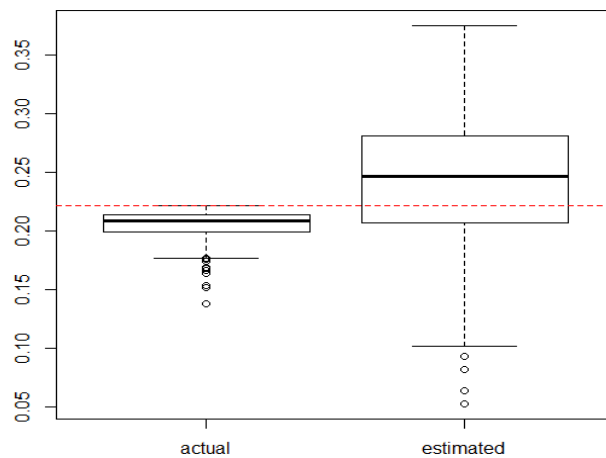


Figure 2. Versus theoretical Sharpe's ratio, the overestimated extent of estimated Sharpe's ratio by using resampling method  
图 2. 用重抽样方法得到的估计夏普比率比理论夏普比率的高估程度

Table 1. The portfolio's weights of the five risk rocks  
表 1. 五只风险股票的投资组合系数

| 包钢股份   | 中国石化   | 太原重工   | 中国建筑   | 紫金矿业   |
|--------|--------|--------|--------|--------|
| 0.0538 | 0.2042 | 0.2593 | 0.3238 | 0.1589 |

到的切线投资组合估计夏普比与理论夏普比之间的高估程度。

由图 2 可知, 估计夏普比率确实比理论夏普比率高, 其中红色虚线为理论上的夏普比率 0.222。这是由于估计的夏普比率是通过重抽样得到的样本数据进行拟合得到的, 它可以很好的拟合该样本数据的特征, 所以比估计实际分布时的夏普比率值高。

## 4. 因子分析-Shrinkage 法

### 4.1. Shrinkage 法

为了减少实际值与估计值之间的误差, 有许多方法已经广泛应用, 并用来缓解这个问题, 例如贝叶斯估计法、因子模型、Shrinkage 法等。我们这里先介绍 Shrinkage 法, 然后基于 Shrinkage 法的思想对其改进, 得到一种新的减小估计量的方法, 即因子分析-Shrinkage 法。

Shrinkage 法的主要目的是减小估计量的方差。假设第  $i$  只股票的均值是  $\bar{Y}_i$ , 则  $\bar{Y} = \frac{\sum_{i=1}^5 \bar{Y}_i}{5}$  为 5 只股票均值的均值。则第  $i$  只股票期望收益的估计为:

$$\hat{\mu}_i = w\mu_i + (1-w)\bar{Y}$$

其中,  $w \in [0,1]$  是权重系数。因此  $\hat{\mu}_i$  的方差会随着  $w$  的减小而趋于 0。通过 R 软件可以得到最大夏普比与真实夏普比之间、最大夏普比与估计夏普比之间绝对误差的波动情况(见图 3)、最大夏普比和权重  $w$  的关系以及真实夏普比与估计夏普比之间的绝对误差随权重的波动情况(见图 4)。在不允许卖空条件下, 图 3 表示 Shrinkage 法得到的真实夏普比、估计夏普比分别与最大夏普比之间绝对误差的波动情况, 分别用蓝色实线、红色虚线表示; 图 4 左图表示最大夏普比与权重的变化关系, 右图表示真实夏普比与估计夏普比之间绝对误差随权重的波动情况。

由图 3 可以看到, 最大夏普比与估计夏普比之间的绝对误差比与实际夏普比之间的大, 且其波动大。

由图 4 左图可知, 最大夏普比随权重  $w$  的增大而增大, 开始增长速度较快, 然后变得越来越缓慢, 直到趋于一个常数, 保持不变, 达到夏普比的最大值。由图 4 右图可知, 真实夏普比与估计夏普比之间绝对误差的波动是由剧烈变得缓慢, 然后又变得剧烈。在其波动中可以得到最优的权重  $w = 0.289$ 。

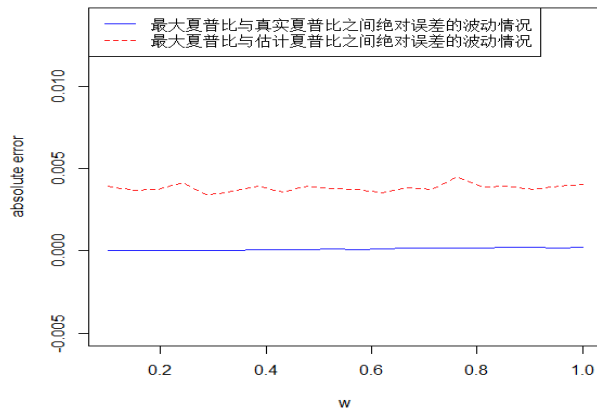
利用 Shrinkage 法可得, 实际和估计值之间夏普比的箱线图如下(见图 5)。图 5 表示在不允许卖空条件下, 利用 Shrinkage 法估计的夏普比高估理论夏普比的程度, 红色虚线代表理论夏普比。左边箱线图表示经 500 次重抽样得到的切线投资组 and 实际夏普比与理论夏普比之间相差程度, 右边箱线图表示经 500 次重抽样的得到的切线投资组合估计夏普比与理论夏普比之间的高估程度。

和图 2 相比, 由图 5 发现实际的夏普比率和估计的夏普率比更接近理论夏普比值, 而且实际夏普比与估计夏普比之间的偏差也在不断的缩减。这就是 Shrinkage 法的真正意义所在。

### 4.2. 因子分析-Shrinkage 法

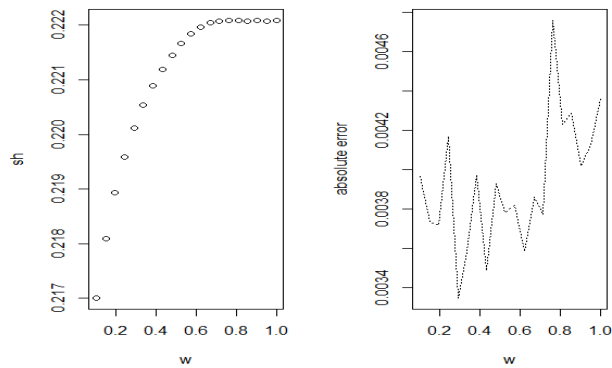
因子分析[11]是指从变量群中提取出公共因子的统计技术。它可以在许多变量中找出隐藏的具有代表性的因子。为了更进一步减小估计量的波动程度, 使得实际夏普比与估计夏普比之间的偏差缩减的更小, 更能充分解释用样本数据拟合总体分布的合理性, 我们提出因子分析-Shrinkage 法。其主要思想是基于 Shrinkage 法, 将  $\hat{\mu}_i = w\mu_i + (1-w)\bar{Y}$  式中的  $\bar{Y}$  用重抽样样本的公共因子的均值代替, 从而对 Shrinkage 法得到了进一步优化。

利用因子分析-Shrinkage 法同样得到最大夏普比与真实的夏普比之间、最大夏普比与估计的夏普比之间的绝对误差的波动情况(见图 6)。在不允许卖空条件下, 图 6 表示因子分析-Shrinkage 法得到的真实夏



**Figure 3.** The variances of the absolute errors of maximum Sharpe's ratio and actual Sharpe's ratio; and the variances of maximum Sharpe's ratio and estimated Sharpe's ratio based on Shrinkage method

**图 3.** 用 Shrinkage 法得到的最大夏普比分别与真实夏普比和估计夏普比之间绝对误差的波动情况



**Figure 4.** With Shrinkage method, the changes between the weight and the maximum Sharpe's ratio; and the changes between the weight and the absolute error of actual Sharpe's ratio and estimated Sharpe's ratio

**图 4.** 用 Shrinkage 法得到的最大夏普比与权重以及真实夏普比与估计夏普比之间的绝对误差与权重的变化情况

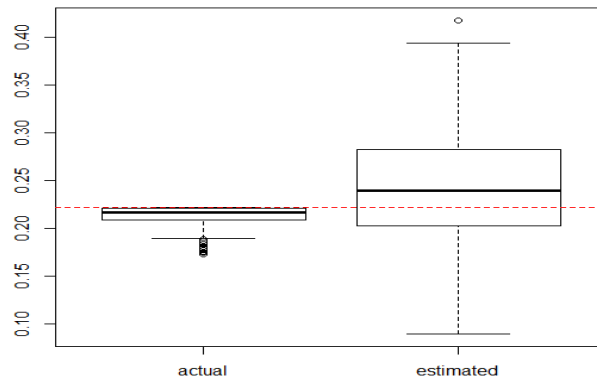
普比、估计夏普比与最大夏普比之间绝对误差的波动情况，分别用蓝色实线、红色虚线表示。

由图 6 可以发现，最大夏普比与估计夏普比之间的绝对误差比与实际夏普比之间的大，且其波动大，这与图 2 所表达的结果一致。但不同的是，两者绝对误差的波动情况均随着权重系数  $w$  的增大而减小。两条线的波动情况整体处于下降趋势，在  $w = 0.85$  的位置趋于平缓。

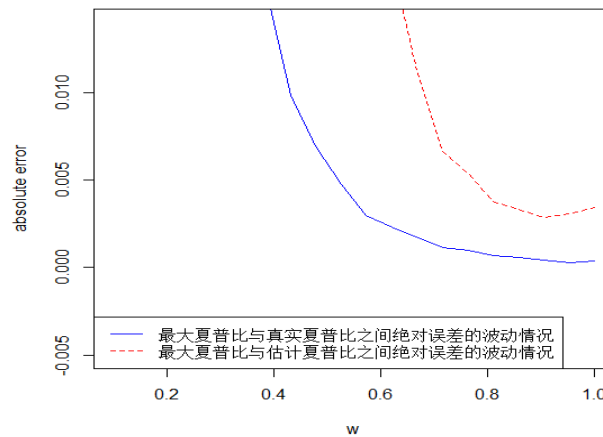
通过运算，得到最大夏普比和权重  $w$  的关系及真实的夏普比与估计的夏普比之间绝对误差的波动情况(见图 7)。图 7 左图表示最大夏普比与权重的变化关系，右图表示真实夏普比与估计夏普比之间绝对误差随权重的波动情况。

由图 7 左图可知，最大夏普比随着权重  $w$  的增大而减小。由右图发现，估计夏普比与实际夏普比之间绝对误差的波动也随着权重  $w$  的增大而减小，逐渐趋于平缓。

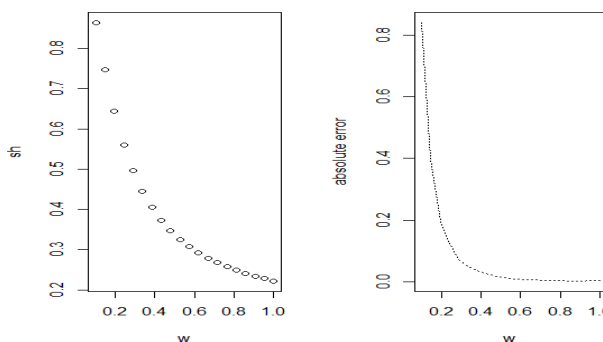
最后，用实际夏普比与估计夏普比的箱线图来展示因子分析-Shrinkage 法的优化成果(见图 8)。图 8 表示在不允许卖空条件下，利用因子分析-Shrinkage 法估计的夏普比高估理论夏普比的程度，红色虚线代表理论夏普比。



**Figure 5.** Versus theoretical Sharpe's ratio, the overestimated extent of estimated Sharpe's ratio by using Shrinkage method  
**图 5.** 用 Shrinkage 法得到的估计夏普比率比理论夏普比率的高估程度

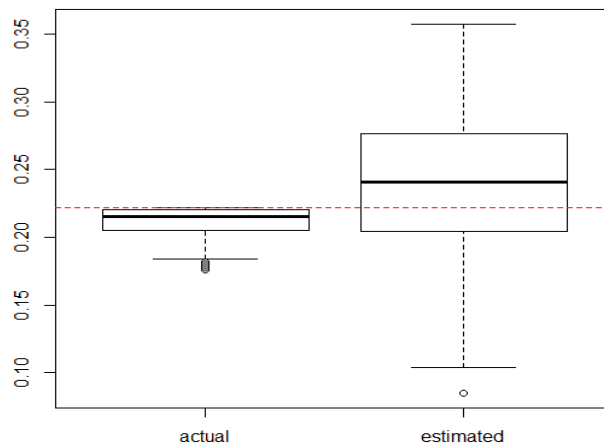


**Figure 6.** The variances of the absolute errors of maximum Sharpe's ratio and actual Sharpe's ratio; and the variances of maximum Sharpe's ratio and estimated Sharpe's ratio based on factor analysis-Shrinkage method  
**图 6.** 用因子分析-Shrinkage 法得到的最大夏普比分别与真实夏普比和估计夏普比之间绝对误差的波动情况



**Figure 7.** With factor analysis-Shrinkage method, the changes between the weight and the maximum Sharpe's ratio; and the changes between the weight and the absolute error of actual Sharpe's ratio and estimated Sharpe's ratio  
**图 7.** 用因子分析-Shrinkage 法得到的最大夏普比与权重以及真实夏普比与估计夏普比之间的绝对误差与权重的变化情况





**Figure 8.** Versus theoretical Sharpe's ratio, the overestimated extent of estimated Sharpe's ratio by using factor analysis-Shrinkage method

**图 8.** 用因子分析-Shrinkage 法得到的估计夏普比率比理论夏普比率的高估程度

**Table 2.** The comparison of the variance based on Shrinkage and factor analysis-Shrinkage methods  
**表 2.** 基于 Shrinkage 和因子分析-Shrinkage 法的两种波动的比较

|           | Shrinkage 法 | 因子分析-Shrinkage 法 |
|-----------|-------------|------------------|
| actual    | -0.007      | -0.0064          |
| estimated | 0.022       | 0.0175           |

图 8 中, 红色虚线仍代表理论夏普比 0.222, 图中 actual 为重抽样得到的实际夏普比的箱线图, estimated 为重抽样得到的估计夏普比的箱线图。和图 5 相比, 图 8 中显示的实际夏普比、估计夏普比都与理论夏普比更接近了, 基于 Shrinkage 法和因子分析-Shrinkage 法波动的比较见表 2。在图 5 中用 Shrinkage 法得到的实际夏普比与理论夏普比之间的差距为-0.007, 在图 8 中用因子分析-Shrinkage 法得到的实际夏普比与理论夏普比之间的差距为-0.0064, 可见后者得到的差距更小。在图 5 中用 Shrinkage 法得到的估计夏普比与理论夏普比之间的差距为 0.0222, 在图 8 中用因子分析-Shrinkage 法得到的估计夏普比与理论夏普比之间的差距为 0.0175, 可见仍是因子分析-Shrinkage 法达到进一步优化。这说明, 优化的因子分析-Shrinkage 法比传统的 Shrinkage 法效果好, 真实夏普比和估计夏普比均与理论的夏普比更接近, 即使得重抽样的数据更加符合真实分布。

## 5. 结论

本文围绕马科维茨投资组合理论, 结合具体事例开展了分析。先由马科维茨模型分析了 5 只风险股票和无风险基金的最佳投资组合, 得到最大夏普比, 利用重抽样方法降低马科维茨模型对输入参数的敏感性, 分析了估计夏普比的高估性, 然后用 Shrinkage 法减小估计量的方差, 缩减了实际夏普比、估计夏普比与理论夏普比的绝对误差, 进一步提出因子分析-Shrinkage 法, 实例证明因子分析-Shrinkage 法的确比 Shrinkage 法效果好, 使模型更加符合实际情况, 进一步提高了模型的实用效率。

## 基金项目

国家自然科学基金(11101012); 北京工商大学学科建设与研究生教育专项(2015XWYJS018)。



## 参考文献 (References)

- [1] 顾岚, 薛继锐, 罗立禹, 徐悦 (2001) 中国股市的投资组合分析. *数理统计与管理*, **5**, 56-60.
- [2] 杨焕云 (2013) 投资组合中一种极大极小化算法及其应用. *统计与决策*, **2**, 62-64.
- [3] 丰雪, 张闾(2005) 投资组合理论的脉络及发展. *辽宁大学学报*, **2**, 142-144.
- [4] Ruppert, D. (2011) *Statistics and data analysis for financial engineering*. Springer, New York, 131-147, 285-305, 531-577. <http://dx.doi.org/10.1007/978-1-4419-7787-8>
- [5] 徐少君 (2004) 投资组合理论发展综述. *浙江社会科学*, **4**, 198-202.
- [6] 薛毅, 陈立萍 (2007) *统计建模与 R 软件*. 清华大学出版社, 北京.
- [7] 王辉, 陈立文, 杨艳芳 (2004) 投资组合风险的分散化研究. *数理统计与管理*, **1**, 53-57.
- [8] 张鹏 (2008) 不允许卖空情况下均值 - 方差和均值 - VaR 投资组合比较研究. *中国管理科学*, **4**, 30-35.
- [9] 张健, 刘嘉惠 (2008) 夏普比率在投资组合管理中的应用. *云南农业大学学报*, **4**, 4-9.
- [10] 张炜, 曾勇(2006) 投资组合再抽样方法及其在沪市 A 股的实证研究. *金融管理*, **18**, 3-8.
- [11] 吕盛鸽 (2005) 因子分析在证券组合投资中的应用. *统计与决策*, **5**, 121-122.