

Title	高次元データに関するポートフォリオ最適化問題
Sub Title	Portfolio optimization problem for high dimensional data
Author	白石, 博(Shiraishi, Hiroshi)
Publisher	
Publication year	2017
Jtitle	科学研究費補助金研究成果報告書 (2016.)
JaLC DOI	
Abstract	<p>本研究では, 投資資産が膨大となるような高次元設定の下で (1) 『分散共分散行列の逆行列に対する漸近不偏な推定量を用いた最適ポートフォリオ推定量』および (2) 『ポートフォリオ比率に対する縮小推定量を用いた最適ポートフォリオ推定量』の2つの推定量を提案した。まず, これらの推定量の理論的, 数値的正当性を確認すると同時に, 日本の株価データ(200銘柄)を用いて既存のポートフォリオとのパフォーマンスを(シャープ比や効用関数を用いた指標により)比較し, 提案手法の優位性があることを確認した。</p> <p>In this research, we propose two types of optimal portfolio estimators when an investor aims to invest a huge number of assets. Suppose that high dimensional data with respect to the past asset return is available. Then we propose (1) Optimal portfolio estimator by using an unbiased estimator for the inverse of the covariance matrix, (2) Shrinkage type estimator for the plug-in portfolio weights estimator. First, we derive the theoretical properties and confirm numerical validity of these estimators. Then, we compare the performance (in terms of the sharp ration and the utility function) with existing portfolios based on the Japanese stock price data (200 issues) and realized the superiority of the proposed method.</p>
Notes	研究種目: 若手研究(B) 研究期間: 2012~2016 課題番号: 24730193 研究分野: 統計科学
Genre	Research Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KAKEN_24730193seika

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the Keio Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 1 日現在

機関番号：32612

研究種目：若手研究(B)

研究期間：2012～2016

課題番号：24730193

研究課題名(和文)高次元データに関するポートフォリオ最適化問題

研究課題名(英文)Portfolio optimization problem for high dimensional data

研究代表者

白石 博(Shiraishi, Hiroshi)

慶應義塾大学・理工学部(矢上)・准教授

研究者番号：90454024

交付決定額(研究期間全体)：(直接経費) 2,000,000円

研究成果の概要(和文)：本研究では、投資資産が膨大となるような高次元設定の下で(1)『分散共分散行列の逆行列に対する漸近不偏な推定量を用いた最適ポートフォリオ推定量』および(2)『ポートフォリオ比率に対する縮小推定量を用いた最適ポートフォリオ推定量』の2つの推定量を提案した。まず、これらの推定量の理論的、数値的正当性を確認すると同時に、日本の株価データ(200銘柄)を用いて既存のポートフォリオとのパフォーマンスを(シャープ比や効用関数を用いた指標により)比較し、提案手法の優位性があることを確認した。

研究成果の概要(英文)：In this research, we propose two types of optimal portfolio estimators when an investor aims to invest a huge number of assets. Suppose that high dimensional data with respect to the past asset return is available. Then we propose (1)Optimal portfolio estimator by using an unbiased estimator for the inverse of the covariance matrix, (2)Shrinkage type estimator for the plug-in portfolio weights estimator. First, we derive the theoretical properties and confirm numerical validity of these estimators. Then, we compare the performance (in terms of the sharp ratio and the utility function) with existing portfolios based on the Japanese stock price data (200 issues) and realized the superiority of the proposed method.

研究分野：統計科学

キーワード：最適ポートフォリオ 統計的推定 漸近理論 高次元データ 縮小推定量

1. 研究開始当初の背景

近年、情報化の進展に伴い、得られるデータの次元 p が標本サイズ n に比べて同等またはより大きい場合が見受けられる。本研究では、このような高次元データに対するポートフォリオ選択問題を扱う。

Markowitz によって提案された平均分散最適ポートフォリオ理論では、資産から得られる期待収益率と期待通りに収益が得られるかどうか投資の危険度を示すリスクのトレードオフ関係が分析される。この理論では、投資資産の収益率を確率ベクトルとして考え、その平均ベクトルと分散共分散行列の逆行列の関数で目的関数を定義し、その目的関数を最大にするような分配比率から最適ポートフォリオが導出される。

実際には、上記の平均ベクトルと分散共分散行列の逆行列は株価データから推定する必要があり、上述した高次元データを扱う場合には具体的に次のような問題点が指摘されている。

(1) 平均ベクトルおよび分散共分散行列に対する最も自然かつ古典的な推定量として、標本平均ベクトルおよび標本分散共分散行列がある。高次元データから標本分散共分散行列を構成した場合、その逆行列が漸近的に不偏性を持たなかったり、 $p > n$ の場合では逆行列が存在しないことがある。

(2) もう一つの手法として、収益率の確率ベクトルに分散に関する構造を導入することで次元圧縮する方法が提案されている。CAPM(Capital Asset Pricing Model) や APT(Asset Pricing Theory) を背景としたファクターモデルは、ポートフォリオを構成する資産をファクターと呼ばれる種々の経済指標(日経平均や GDP 成長率など)で回帰するモデルであり、資産数 (p) をより少ないファクター数 (k) で説明することにより逆行列の問題を回避できることが知られている。しかし、モデルの妥当性、ファクターの選択方法など実用に関しては幾つかの問題がある。

2. 研究の目的

まず、本研究において、サンプルサイズを n 、資産数を p としたとき

$$n \rightarrow \infty, p \rightarrow \infty, p/n \rightarrow c_0 \in (0,1)$$

が成り立つと仮定する(以下、この仮定が成立する状況を「高次元漸近理論の枠組み」と書くこととする。)。この仮定は本研究課題である高次元データを反映したものである。藤越(2011)によると、 $c_0 > 1$ であったり、 $p \gg n$ のような高次元小標本と呼ばれるデータの解析についての研究も近年ではされているが、本研究では次元およびサンプルサイズが発散するが、 $p < n$ という状況は保っていると仮定する。その下で上記の2つの問題への対応を考える。

(1)の問題点に着目し、推定量の改良を考える。Fujikoshi et al. (2010) などによ

ると、高次元漸近理論の枠組みの下では、従来の標本分散共分散行列の逆行列が漸近不偏となることが知られている。これは、 p 固定の下で $n \rightarrow \infty$ とする大標本漸近理論では漸近的に無視されていた p/n の項が無視できなくなるためであり、この点を考慮した新たな推定量を提案し、理論的正当性を確認する。

また、(2)の問題点に対する対応も同時に考える。高次元ファクターモデルについての研究は、比較的古くから行われており、先行文献の調査および実務的に活用可能かどうかを調べる。

最終的には、本研究課題で提案する最適ポートフォリオの理論的結果をシミュレーションで検証すると共に、実データを用いて既存のポートフォリオとのパフォーマンスの比較を行う。これにより、投資資産が膨大となるような高次元設定の下で、提案手法が一定の優位性を持つことを理論的・実務的に実証することが出来ると期待される。上記の研究を通じて、近年におけるポートフォリオ選択問題による理論と実務のギャップを解消することが本研究の目的である。

3. 研究の方法

具体的には次の手順で研究を進めた。

(1) 先行研究の調査：高次元漸近理論の枠組みにおけるポートフォリオ選択問題に関する先行文献の調査を行った。特に、

ランダム行列理論を用いた研究(Bai et al., 2009 など)

縮小推定量を用いた研究(Ledoit and Wolf, 2003 など)

ファクターモデルを仮定した研究(Ao et al., 2015 など)

について、現時点で判明している結果を調査し、現実とのギャップは存在しないか? 新たな研究の余地は無いか? を調査した。

(2) 推定量の提案：次に(1)を受けて、新たな推定量の提案を考えたい。具体的には、漸近不偏性を持つ分散共分散の逆行列の推定量を用いたポートフォリオ推定量の提案

ポートフォリオ比率の推定量に関する縮小タイプの推定量

を提案した。

(3) 理論の構築：(2)で提案した推定量についての漸近的性質として、漸近不偏性、一致性、漸近正規性などの導出を試みた。

(4) 数値実験：(2)で提案した推定量について、(3)で試みた結果を確認するためのシミュレーションを行った。また、(2)で提案した推定量が実務的に利用可能であるか? また、(1)で調査する過去提案された推定量に対して提案推定量が優位性を持っているか? を調べるために、実データを用いた解析を行った。

4. 研究成果

(1) 先行研究の調査について：

Bai et al. (2009)は、確率論で発展してきたランダム行列理論を用いて、高次元漸近理論の枠組みでの最適ポートフォリオの推定問題を議論している。また、Karoui(2010)は、資産の収益率が楕円型分布族に従う場合の漸近的性質の導出を、ランダム行列理論を用いないで行っている。Bodnar et al.(2014)は、 $p/n < c_0$ ($0,1$)の場合だけでなく、 $c_0 > 1$ の場合であっても、一般化逆行列を用いれば、一致推定量が導出できることを示した。

James and Stein によって提案された縮小推定量は、平均二乗誤差(MSE)を減少させる推定量として広く知られている。Jorion(1986)は、ポートフォリオ比率の推定量における標本平均の部分を縮小推定量に置き換えた推定量を提案している。また、Ledoit and Wolf(2003)などは、ポートフォリオ比率の推定量における標本分散共分散の部分を縮小推定量に置き換えた推定量を提案している。

Camberlain and Rothschild(1983 a,b)は、Ross が提案した古典的なファクターモデルの正当性を高次元漸近理論の枠組みでも示している。さらに、Forni and Lippi(2001)は、この議論を時系列モデルでの設定に拡張し、generalized dynamic factor model (GDFM)というモデルを提案した。また、古典的なファクターモデルを仮定したとき、最適ポートフォリオ比率を推定するための統計的手法は、状況によって2通りの方法が考えられることが判明した。

i) ファクターが観測されている場合：この場合は回帰モデルに帰着され、そのパラメータ推定法は従来の最小二乗法よりも、L1 罰則項を付与したLASSO などを使用した方が結果が良くなる場合が多々ある事が知られている。特に、Ao et al(2015)は、真の最適ポートフォリオ比率にスパース性（比率が0の資産が多くある）があるという仮定をし、LASSO を用いて最適ポートフォリオ比率を導出している。

ii) ファクターが観測されていない場合：この場合は static approach と呼ばれる PCA(principal component analysis: 主成分分析)を使った手法が Stock and Watson(2002)らによって提案されている。さらに、GDFM に対しては、dynamic approach と呼ばれる手法が Forni et al.(2000,2005)によって提案されている。

(2) 推定量の提案について：

ポートフォリオ比率の推定量における標本分散共分散の部分を、逆行列が漸近不偏となるように p/n という項で補正した推定量に置き換えたポートフォリオ推定量を提案した。

平均分散基準で考えると、最適ポートフォリオ比率は、資産の平均ベクトル(μ)と分散共分散行列(Σ)の関数 $g(\mu, \Sigma)$ として定義される。先行文献で提案された縮小ポートフォリオ推定量は、従来の μ や Σ の古典的な推定量($\hat{\mu}, \hat{\Sigma}$)を縮小推定量に置き換えた議論をしていたが、新たに $g(\hat{\mu}, \hat{\Sigma})$ を縮小推定した推定量を提案した。

(3) 理論の構築について：

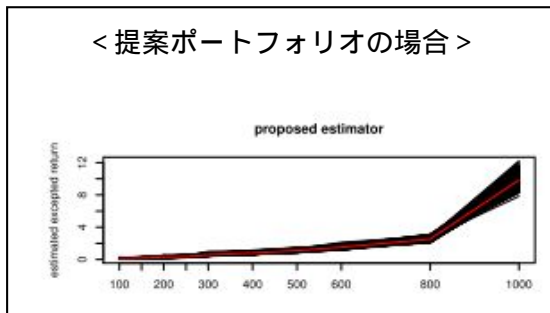
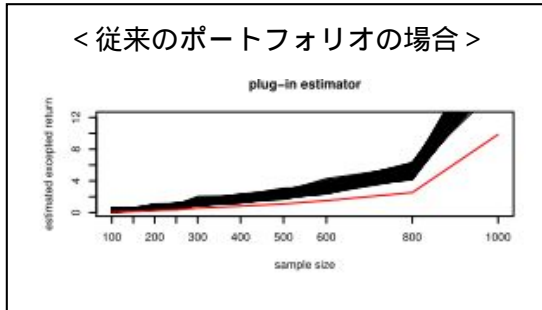
(2) で提案した推定量の一致性、漸近正規性を調べた。高次元漸近理論の枠組みでは、従来のポートフォリオ推定量が一致性を持たないことを示し、それを補正した提案推定量が一致性を持つことを示した。さらに、提案推定量の漸近正規性を導出した。漸近分散を明示的に記述し、その一致推定量を構成できたことから、最適ポートフォリオで構成される有効フロンティアの信頼領域を明示的に構成できた。

James and Stein が提案した縮小推定量は、次元 p が3以上であるときにそのMSEが標本平均と比べて減少することが知られている。その拡張として、(2) で提案した推定量が従来の推定量よりもMSEの意味で改良される条件を導くことに成功した。

(4) 数値実験について：

最後に、本研究課題で提案した最適ポートフォリオ推定量の理論的結果の検証をシミュレーションにより行い、さらに実データを用いて既存の推定量とのパフォーマンスの比較を行った。

まず、(2) で提案したポートフォリオ推定量に関してシミュレーションを行い、高次元漸近理論の枠組みで推定量が漸近不偏であることを確認した。下図は、資産数 p とサンプルサイズ n との比（つまり、 p/n ）を0.5で固定した下で、 n （横軸）を大きくした場合のポートフォリオの期待リターン推定値を1000回繰り返しプロットしたものであり、赤線は真の期待リターンである。従来のポートフォリオの場合、 n が大きくなると明らかに偏り（バイアス）が生じていることが分かる。一方、提案ポートフォリオの場合、偏りが補正されていることが分かる。



次に、実データ解析を行った。既存のポートフォリオとして

- a) 配分比率が等しいポートフォリオ (NP; 頑健性があることが知られている。)
- b) 標本平均ベクトル、標本共分散行列を用いた従来のポートフォリオ (ML)
- c) 標本平均を縮小推定したポートフォリオ (BS)
- d) 分散共分散行列を縮小推定したポートフォリオ (CS)

を採用した。まず、日経 225 の 209 銘柄の株価データを用いて、(2) で提案したポートフォリオと既存のポートフォリオ (NP, ML, CS) を効用関数を用いた指標を用いて比較した結果、リスク回避的な状況 (つまり、リターンよりも分散を比較的重視する効用関数) においては、提案手法が最も優位であることを確認した。

また、20 銘柄を用いて、(2) で提案したポートフォリオで、特に 3 つの縮小関数を用いたものと既存のポートフォリオ (NP, ML, BS, CS) を比較した。シャープ比や効用関数を用いた指標により、縮小関数を適切に選べば提案手法が優位性を持つことを確認した。

(5) まとめと今後の課題について :

本研究では、サンプルサイズ n と次元 p が、 $n > p$ という条件を満たしながら共に大きくなる“高次元漸近理論の枠組み”での最適ポートフォリオの推定法を調査し、新たな推定量を提案した。2 つの推定法を提案し、理論的結果の導出およびシミュレーションによる理論的結果の検証を行った。また、実データ解析により、提案手法の優位性を

確認した。しかしながら、これらの結果は幾つか条件を必要としており、現実とのギャップは依然として残っていると思われる。特に、理論的結果の導出の際に正規性を仮定したり、時間的従属性を無視しており、これらの仮定を緩和した理論の導出は今後の課題として残っている。その他にも、GDFM を仮定した場合の動的最適ポートフォリオ推定量の提案、ファクターが観測されていない場合の推定量の提案などが今後の課題と言える。しかし、Fan et al.(2008)は、ファクターモデルを仮定した場合の最適ポートフォリオ推定量の収束レートに関する議論をしており、この結果を拡張することなど、上記の問題解決に繋がる新たな着想も得ることができた。

<引用文献>

- Fujikoshi, Y., Ulyanov, V.V., Shimizu, R., Multivariate statistics. High-dimensional and large-sample approximations., Wiley Series in Probability and Statistics., 2010.
- 藤越康祝、21 世紀の統計学への挑戦 滝課題と展望、「21 世紀の統計科学」第 3 巻、第 3 章、2011
- Bai, Z., ; Liu, H., Wong, W.K., Enhancement of the applicability of markowitz's portfolio optimization by utilizing random matrix theory., Mathematical Finance, 19(4), 639-667, 2009
- Ledoit, O., and Wolf, M., Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection., Journal of Empirical Finance, 10, 603-621, 2003
- Ao, M., Li, Y., and Zheng, X., Solving the high-dimensional Markowitz optimization problem: When sparse regression meets random matrix theory., Available at SSRN, 2015
- EL Karoui, N., High-Dimensionality Effects in the Markowitz Problem and Other Quadratic Problems with Linear Constants: Risk Underestimation. The Annals of Statistics, 38(6), 3487-3566, 2010
- Bodnar, T., Parolya, N., and Schmid, W., Estimation of the global minimum variance portfolio in high dimensions., arXiv:1406.0437, 2014
- Jorion, P., Bayes-Stein Estimation for Portfolio Analysis., Journal of Financial and Quantitative Analysis., 21(3), 279-292, 1986
- Camberlain, G., and Rothschild, M., Arbitrage, Factor structure, and Mean-Variance Analysis on Large Asset Markets., Econometrica, 51(5),

1281-1304, 1983
Camberlain, G., Funds, Factors and Diversification in Arbitrage Pricing Models., *Econometrica*, 51(5), 1305-1323, 1983
Forni, M., and Lippi, M., The generalized dynamic factor model: representation theory., *Econometric Theory*, 17, 1113-1141, 2001
Stock, J. and Watson, M., Forecasting Using Principal Components From a Large Number of Predictors, *Journal of the American Statistical Association*, 97(460), 1167-1179, 2002
Forni, M., Hallin, M., Lippi, M., and Reichlin, L., The generalized factor model: identification and estimation., *The Review of Economics and Statistics*, 82, 540-554.
Forni, M., Hallin, M., Lippi, M., and Reichlin, L., The generalized factor model consistency and rates., *Journal of Econometrics*, 119, 231-255, 2004.
Fan, J., Fan, Y., and Lv, J., High dimensional covariance matrix estimation using a factor model., *Journal of Econometrics*, 147, 186-197, 2008

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計2件)

岡紘之、白石博、高次元における有効フロンティアの統計的推定、早稲田大学理工研報告特集号、査読無、第14号、2017、15-31

白石博、GPIGのポートフォリオ解析について、早稲田大学理工研報告特集号、査読有、第12号、2015、3-17

〔学会発表〕(計3件)

岡紘之、白石博、高次元の下での有効フロンティアの統計的推定、大規模複雑データの理論と方法論:最前線の動向、2015年11月17日、筑波大学(茨城県つくば市)

Hiroshi Shiraishi、Universal Portfolio with Optimal Categorized Side Information、Workshop on Statistics for High-dimensional and Dependent Data、2014年3月21日、Taipei(Taiwan)

Hiroshi Shiraishi、Review of Statistical Portfolio Theory、Waseda International Symposium on "Stable Process, Semimartingale, Finance & Pension Mathematics"、2014年3月5日、早稲田大学(東京都新宿区)

〔図書〕(計0件)

〔産業財産権〕

出願状況(計0件)

取得状況(計0件)

〔その他〕

6. 研究組織

(1) 研究代表者

白石 博 (SHIRAISHI, Hiroshi)

慶應義塾大学・理工学部・准教授

研究者番号: 90454024