

現代的な投信環境における最適な共分散行列の推定 について

國廣 亮生

February 13, 2026

要約

現代ポートフォリオ理論において、高次元・小標本の設定では、推定誤差の増大により最適化ポートフォリオが不安定化するという問題がある。この問題に対し、従来の実証研究では、パラメータ推定を完全に放棄する「等配分戦略 (1/N)」が、多くの洗練されたモデルを上回る頑健性を示すことが報告されてきた。本研究では、情報を「完全に捨てる (等配分)」のではなく、統計的な縮小推定を用いて情報を「適切に正則化して活用する」アプローチの有効性を検証する。米国の株式市場データを用いた包括的な実証実験の結果、縮小推定は高次元データに含まれるノイズを抑制しつつ、有用な相関構造を抽出することに成功し、等配分戦略や代表的なファクターモデルを上回るパフォーマンスを実現することを明らかにした。

1 導入

1.1 背景と課題

現代のポートフォリオ理論は、Markowitz [1] によって確立された平均・分散アプローチ (Mean-Variance Optimization) に基づいている。この枠組みにおいて、投資家は所与の期待リターンに対してリスク (分散) を最小化するポートフォリオを構築することを目的とする。この最適化計算には、各資産の期待リターンベクトルと、資産間の相関構造を表す分散共分散行列 (Covariance Matrix) の正確な推定が必要不可欠である。

理論上、真の共分散行列が既知であれば最適なポートフォリオを一意に決定できる。しかし、実務上において真のパラメータは観測不可能であり、過去の市場データから推定する必要がある。最も一般的に用いられるのは**標本共分散行列 (Sample Covariance Matrix)**であるが、これは「次元の呪い」に対して極めて脆弱であるという致命的な欠点を持つ。特に、運用する銘柄数 N が観測期間 T に対して大きい場合（高次元・小標本設定）、推定誤差が増大し、算出されるポートフォリオのウェイトは極端かつ不安定なものとなる。これは、標本共分散行列がデータに含まれるノイズを過学習してしまうことに起因する。

1.2 既存のアプローチと本研究の目的

この推定誤差の問題に対処するため、本研究では主に3つのアプローチに着目する。

第一に、相関構造の推定そのものを放棄する**等配分戦略 (Equal Weight Strategy)**である。これは、推定誤差が極めて大きい環境下では、バイアスのあるモデルやノイズまみれの最適化を行うよりも、単純な均等配分の方がロバストな結果をもたらすという考えに基づいている。一方で、データに含まれる有用な情報までも全て捨ててしまうという欠点を持つ。

第二に、経済的あるいは統計的な事前知識に基づいてデータに構造的制約を課す**ファクターモデル**を用いたアプローチである。CAPM などの市場ファクターモデルや主成分分析 (PCA)、POET などがこれに該当する。これらは推定すべきパラメータ数を大幅に削減できる利点があるが、仮定したモデル構造が現実の市場と乖離していた場合（特定化誤差）、大きなバイアスを生むリスクがある。

第三に、統計的な手法によりノイズを除去する**縮小推定 (Shrinkage Estimation)**である。Ledoit and Wolf [2] による線形縮小推定は、バイアスは小さいがノイズは大きい標本共分散行列と、バイアスは大きいがノイズは小さい構造化されたターゲット行列の加重平均をとることで、バイアスと分散のトレードオフを最適化する。近年では、De Nard [3] らにより、ランダム行列理論を用いて固有値分布そのものを補正する非線形縮小推定へと発展している。

現代ポートフォリオ理論の実証研究において、高次元・小標本の設定では、推定誤差の影響を回避するためにパラメータ推定を放棄する等配分戦略が、しばしば最適化モデルを上回ることが古くから指摘されている [4, 5, 6]。しかし本研究では、情報を「完全に捨てる（等配分）」のではなく、縮小推定を用いて情報を「適切に平均化・正則化して活用する」アプローチの方が、リスク低減とリターン向上の両面でより優れたパフォーマンスを発揮できるという仮説を立てる。

この仮説を検証するため、米国の株式市場データを用い、等配分戦略や、近年実務で広

く利用されるファクターモデル（Market Factor, PCA, POET）との包括的な比較実験を行う。そして、現代の複雑な投資環境において、縮小推定が最も有効なアプローチであることを実証する。

2 ポートフォリオ最適化問題

現代のポートフォリオ理論は、Markowitz [1] によって確立された平均・分散アプローチ（Mean-Variance Optimization）に基づいている。この理論的枠組みにおいて、投資家は以下の 2 つの目的の間で最適なトレードオフを決定する。すなわち、(1) 特定のリスク水準の下で期待リターンを最大化すること、あるいは (2) 特定の期待リターン水準を達成しつつリスク（分散）を最小化することである。

ポートフォリオの構築において、各資産の期待リターン μ の推定は極めて困難であり、その推定誤差が最適化結果に与える悪影響は大きいことが知られている [7]。そのため、共分散行列の推定精度の評価を主目的とする多くの研究（2; 3 等）では、期待リターンの推定誤差の影響を排除できる**大域的最小分散ポートフォリオ (Global Minimum Variance Portfolio: GMVP)** に焦点を当てている。

本研究でもこれに倣い、期待リターンの予測を行わず、リスクの最小化のみを目的とする最小分散ポートフォリオ問題を扱う。以下では、その定式化と、実務上の課題について述べる。

2.1 最小分散ポートフォリオ

現代ポートフォリオ理論は、Markowitz [1] が 1952 年に発表した論文で提唱した最適化理論を元に構築されており、投資家が直面するリターンとリスクのトレードオフを定量的に解決する方法に従う。この理論の目的は特定のリスク水準の下で期待リターンを最大化し、特定の期待リターン水準を達成する中でリスクを最小化することにある。

ポートフォリオの期待リターン $E(r_p)$ は、個々の資産の期待リターン $E(r_i)$ を、その資産の投資ウェイト w_i で加重平均した値で求める。

$$E(r_p) = \sum_{i=1}^N w_i E(r_i) = \mathbf{w}^T \boldsymbol{\mu} \quad (1)$$

\mathbf{w} は各資産に対する投資比率を縦に並べたウェイトベクトルであり、 $\boldsymbol{\mu}$ は各資産の期待リターンを並べた列ベクトルである。また、ポートフォリオの分散 σ_p^2 は、ポートフォリオのリスクの大きさを表す。単に各資産の分散を合計するのではなく、資産間の共分散を計算することで求める。

$$\sigma_p^2 = \sum_{i=1}^N \sum_{j=1}^N w_i w_j \text{Cov}(r_i, r_j) = \mathbf{w}^\top \mathbf{\Sigma} \mathbf{w} \quad (2)$$

$\text{Cov}(r_i, r_j)$ は資産 i のリターン r_i と資産 j のリターン r_j の共分散でリターンの相関を示し、 $\mathbf{\Sigma}$ はすべての資産ペアの共分散と分散をまとめた分散共分散行列である。この分散の計算において、資産間の共分散が小さい（または負の相関がある）ほど、ポートフォリオ全体のリスクは小さくなる。そのため、ポートフォリオを最適化するには、分散共分散行列である $\mathbf{\Sigma}$ を最小化する必要がある。

したがって、最小分散ポートフォリオを求める最小化問題は以下のように定式化できる。

$$\min_{\mathbf{w}} \quad \mathbf{w}^\top \mathbf{\Sigma} \mathbf{w} \quad (3)$$

$$\text{s.t.} \quad \mathbf{C}^\top \mathbf{w} = \boldsymbol{\gamma} \quad (4)$$

\mathbf{C} は、 N を資産数、 M を制約数としたときにおける $N \times M$ の行列であり、 $\boldsymbol{\gamma}$ は各制約条件の目標値を格納した $M \times 1$ の列ベクトルである。この最小化問題を解くことによって、リスクを最小限に抑える最適なウェイトベクトル \mathbf{w} が以下のように求められることが知られている。

$$\begin{aligned} \mathbf{w} &= \mathbf{\Sigma}^{-1} \mathbf{C} (\mathbf{C}^\top \mathbf{\Sigma}^{-1} \mathbf{C})^{-1} \boldsymbol{\gamma} \\ &= \sum_{i=1}^N \frac{\mathbf{C}^\top \cdot \mathbf{u}_i}{\lambda_i} \mathbf{u}_i \end{aligned} \quad (5)$$

ここで、 λ_i は分散共分散行列の各固有値を表し、 \mathbf{u}_i は分散共分散行列の各固有ベクトルを表す。

2.2 標本共分散行列

理論的なポートフォリオ最小化問題では、真の分散共分散行列 $\mathbf{\Sigma}$ を用いて推定されるが、実務上では $\mathbf{\Sigma}$ を観測することは不可能であるため、過去の収益率 \mathbf{R} （中心化済みとする）から推定した標本共分散行列 $\hat{\mathbf{\Sigma}} = \frac{1}{T-1} \mathbf{R}^\top \mathbf{R}$ が用いられる。

実務上では、ポートフォリオの最適化は (3) 式を用いて以下のように行われる。

$$\min_{\mathbf{w}} \quad \mathbf{w}^\top \hat{\mathbf{\Sigma}} \mathbf{w} \quad (6)$$

$$\text{s.t.} \quad \mathbf{C}^\top \mathbf{w} = \boldsymbol{\gamma} \quad (7)$$

また、(5) 式も標本共分散行列を用いて以下のように表される。

$$\begin{aligned}\mathbf{w} &= \hat{\Sigma}^{-1} \mathbf{C} (\mathbf{C}^T \hat{\Sigma}^{-1} \mathbf{C})^{-1} \boldsymbol{\gamma} \\ &= \sum_{i=1}^N \frac{\mathbf{C}^T \cdot \mathbf{u}_i}{\lambda_i} \mathbf{u}_i\end{aligned}\tag{8}$$

しかし、特に銘柄数 N が観測期間 T に対して大きい場合、標本共分散行列は以下のような問題を引き起こすことが知られている。

第一に、次元性の呪いと推定誤差である。推定すべきパラメータ数 $N(N+1)/2$ がデータ量 $N \times T$ に対して過大である場合、標本共分散行列は真の行列に対する一致推定量とならず、大きな推定誤差を含むことになる。

第二に、固有値のバイアスである。Ledoit [8] によれば、標本共分散行列の固有値は、真の固有値よりも分散して推定される傾向がある。具体的には、大きな固有値はより大きく、小さな固有値はより小さく（ゼロに近く）ゆがんで推定される。

第三に、このバイアスが引き起こす誤差の最大化である [5]。式 (5) に示されるように、最適ウェイトの算出には共分散行列の逆行列が用いられるため、各固有ベクトルへの配分比率はその固有値の逆数に比例する。前述のバイアスにより、本来のリスクよりも過小に推定された（極端に小さな）固有値に対して、計算上は莫大な投資ウェイトが割り当てられてしまうことになる。

結果として、標本共分散行列を用いた最適化は、真のリスクが低い銘柄ではなく「推定誤差が大きくリスクが過小評価された銘柄」に集中投資することになり、極端で不安定なポートフォリオが出力される [3]。本研究で扱う $N \gg T$ のケースでは、行列がランク落ちして逆行列が存在しない（特異行列となる）ため、これらの問題はさらに決定的となる。

3 共分散行列の推定手法

本研究では、標本共分散行列の問題点を克服するために提案されたいくつかの手法を比較検討する。

3.1 等配分戦略 (Equal Weight Strategy)

等配分戦略（ $1/N$ ポートフォリオ）とは、最適化やパラメータ推定を一切行わず、すべての資産に対して均等なウェイトを配分する単純な戦略である。ポートフォリオの各資産 i への投資比率 w_i は以下のように定義される。

$$w_i = \frac{1}{N}, \quad i = 1, \dots, N \quad (9)$$

この戦略は、一見すると単純すぎるように思えるが、高次元データやモデルの不確実性が高い環境においては、非常に強力なベンチマークとして機能することが知られている。DeMiguel et al. [6] は、推定誤差のために最適化ポートフォリオが不安定になる場合、パラメータ推定を必要としない等配分戦略が、標本共分散行列を用いた平均分散最適化をはじめとする多くの洗練されたモデルよりも優れたパフォーマンス（高いシャープ・レシオ）を示すことを実証した。

本研究においては、この等配分戦略を一切の情報を利用しない（正則化を無限大にした）極端なケースとしてのベンチマークに位置づけ、縮小推定の有効性を測る基準とする。

3.2 市場ファクターモデル (Market Factor Model)

市場全体の動きを表す単一のファクターを用いて個別の資産リターンを説明するモデルである。資産 i のリターン r_i は、市場ファクター f とその感応度 β_i 、および固有リスク ϵ_i を用いて以下のようにモデル化される。

$$r_i = \beta_i f + \epsilon_i \quad (10)$$

ここで、 ϵ_i は市場の動きとは無相関な個別資産固有の変動であり、資産間でも互いに無相関であると仮定する ($\text{Cov}(\epsilon_i, \epsilon_j) = 0, i \neq j$)。この仮定により、市場ファクターモデルに基づく共分散行列 Σ_{factor} は以下のような構造を持つ。

$$\Sigma_{\text{factor}} = \sigma_f^2 \boldsymbol{\beta} \boldsymbol{\beta}^\top + \mathbf{D} \quad (11)$$

ここで、 σ_f^2 は市場ファクターの分散、 $\boldsymbol{\beta}$ は各資産の感応度を並べたベクトル ($\boldsymbol{\beta} = [\beta_1, \dots, \beta_N]^\top$)、 \mathbf{D} は各資産の固有リスクの分散 ($\sigma_{\epsilon_i}^2$) を対角成分に持つ対角行列である。

通常の標本共分散行列では $N(N + 1)/2$ 個のパラメータを推定する必要があるが、このモデルを用いることで、推定すべきパラメータ数は $2N + 1$ 個 (N 個の β 、 N 個の固有分散、1 個のファクター分散) にまで大幅に削減され、推定誤差を抑制することができる。

なお、本研究における市場ファクター f としては、米国株式市場の包括的なベンチマークである **S&P Composite 1500** 指数の対数収益率を採用する。S&P Composite 1500 は、

大型株 (S&P 500), 中型株 (S&P MidCap 400), 小型株 (S&P SmallCap 600) を統合して構築されており, 米国株式市場の時価総額の約 90% をカバーしている. これにより, 単なる大型株指数である S&P 500 よりも, 市場全体の動向をより正確かつ包括的に捉えることが可能となる.

3.3 統計的ファクターモデル (PCA と POET)

市場インデックスなどの外部変数を用いず, データそのものから主要な変動要因 (ファクター) を統計的に抽出する手法について述べる. 代表的な手法として主成分分析 (PCA) と, その拡張である POET が挙げられる.

PCA は, 資産リターンの標本共分散行列 $\hat{\Sigma}$ をスペクトル分解し, 固有値の大きい上位 K 個の主成分を用いて共分散行列を再構成する手法である. 標本共分散行列の固有値分解を $\hat{\Sigma} = \sum_{j=1}^N \lambda_j v_j v_j^\top$ (ただし $\lambda_1 \geq \dots \geq \lambda_N$ は固有値, v_j は対応する固有ベクトル) とすると, PCA に基づく推定量 $\hat{\Sigma}_{PCA}$ は以下のように表される.

$$\hat{\Sigma}_{PCA} = \sum_{j=1}^K \lambda_j v_j v_j^\top + \hat{\mathbf{D}}_K \quad (12)$$

ここで, 第 1 項は K 個の共通ファクターによる構造を表し, $\hat{\mathbf{D}}_K$ は残差部分の対角成分のみを取り出した対角行列である. このモデルは次元圧縮に有効であるが, ファクターで説明できない残差部分は「互いに無相関」とであると仮定するため, ファクター以外の局所的な相関構造 (セクター内相関など) は無視されてしまうという課題がある.

この課題に対処するために Fan et al. [9] によって提案された手法が POET (Principal Orthogonal complement Thresholding) である. POET では, PCA によって抽出された共通ファクター部分を除いた残差行列 $\hat{\mathbf{R}}_K = \hat{\Sigma} - \sum_{j=1}^K \lambda_j v_j v_j^\top$ に対して, 閾値処理 (Thresholding) を適用する. 残差行列の各要素に対し, 微小な値をゼロにする閾値関数 $s(\cdot)$ を用いることで, POET の推定量 $\hat{\Sigma}_{POET}$ は以下のように定義される.

$$\hat{\Sigma}_{POET} = \sum_{j=1}^K \lambda_j v_j v_j^\top + s(\hat{\mathbf{R}}_K) \quad (13)$$

このように, POET は共通ファクター構造を維持しつつ, 残差部分に含まれるスパースな (まばらな) 相関構造も柔軟に取り込むことが可能である.

3.4 縮小推定 (Shrinkage Estimation)

標本共分散行列は、次元数 N が観測期間 T に対して大きい場合、推定誤差が大きくなりポートフォリオのパフォーマンスを悪化させる。この問題に対処するため、Ledoit and Wolf [2] は「縮小推定」と呼ばれるアプローチを確立した。本研究では、その代表的な手法である線形縮小推定と、近年発展した非線形縮小推定を採用する。

3.4.1 線形縮小推定 (Linear Shrinkage)

Ledoit and Wolf (2004) によって提案された線形縮小推定は、バイアスのない標本共分散行列 $\hat{\Sigma}$ と、バイアスの大きいターゲット行列 \mathbf{F} の凸結合をとることで、バイアスと分散のトレードオフを最適化する手法である。推定される共分散行列 $\hat{\Sigma}_{Lin}$ は以下の式で定義される。

$$\hat{\Sigma}_{Lin} = \delta \mathbf{F} + (1 - \delta) \hat{\Sigma} \quad (14)$$

ここで $\delta \in [0, 1]$ は縮小強度 (Shrinkage Intensity) である。本手法の核心は、真の共分散行列 Σ との推定誤差を表す以下の二次損失関数を最小化する最適な δ を決定することにある。

$$\min_{\delta} \mathbb{E} [\|\hat{\Sigma}_{Lin} - \Sigma\|_F^2] \quad (15)$$

ここで $\|\cdot\|_F$ はフロベニウスノルムである。

縮小推定量の二乗平均誤差は、以下のようにターゲット行列のバイアスの二乗と標本の分散の重み付き和に分解される。

$$\text{MSE}(\delta) = \mathbb{E} [\|\hat{\Sigma}_{Lin} - \Sigma\|_F^2] \quad (16)$$

$$= \delta^2 \|\mathbf{F} - \Sigma\|_F^2 + (1 - \delta)^2 \|\hat{\Sigma} - \Sigma\|_F^2 \quad (17)$$

証明は以下の通りである。

まず、二乗平均誤差の定義式に $\hat{\Sigma}_{Lin}$ の式を代入する。

$$\text{MSE}(\delta) = \mathbb{E} [\|\delta \mathbf{F} + (1 - \delta) \hat{\Sigma} - \Sigma\|_F^2] \quad (18)$$

$$= \mathbb{E} [\|\delta \mathbf{F} + (1 - \delta) \hat{\Sigma} - \Sigma + \delta \Sigma - \delta \Sigma\|_F^2] \quad (19)$$

$$= \mathbb{E} [\|\delta(\mathbf{F} - \Sigma) + (1 - \delta)(\hat{\Sigma} - \Sigma)\|_F^2] \quad (20)$$

ここで、ノルムの性質 $\|\mathbf{A} + \mathbf{B}\|_F^2 = \|\mathbf{A}\|_F^2 + \|\mathbf{B}\|_F^2 + 2\langle \mathbf{A}, \mathbf{B} \rangle$ を利用して展開する。 \mathbf{F} と Σ は定数行列であることに注意すると、

$$\text{MSE}(\delta) = \delta^2 \|\mathbf{F} - \Sigma\|_F^2 + (1 - \delta)^2 \mathbb{E} [\|\hat{\Sigma} - \Sigma\|_F^2] + 2\delta(1 - \delta) \mathbb{E} [\langle \mathbf{F} - \Sigma, \hat{\Sigma} - \Sigma \rangle] \quad (21)$$

第三項について考える。内積の線形性と期待値の線形性を用いると、

$$\mathbb{E} [\langle \mathbf{F} - \Sigma, \hat{\Sigma} - \Sigma \rangle] = \langle \mathbf{F} - \Sigma, \mathbb{E}[\hat{\Sigma} - \Sigma] \rangle \quad (22)$$

$$= \langle \mathbf{F} - \Sigma, \Sigma - \Sigma \rangle \quad (\because \mathbb{E}[\hat{\Sigma}] = \Sigma : \text{不偏推定量}) \quad (23)$$

$$= 0 \quad (24)$$

となり、第三項は消失する。

したがって、二乗平均誤差は以下のように簡略化される。

$$\text{MSE}(\delta) = \delta^2 \|\mathbf{F} - \Sigma\|_F^2 + (1 - \delta)^2 \mathbb{E} [\|\hat{\Sigma} - \Sigma\|_F^2] \quad (25)$$

この式を δ について最小化することで、最適な縮小強度は以下のように導出される。

$$\frac{\partial \text{MSE}(\delta)}{\partial \delta} = 2\delta \|\mathbf{F} - \Sigma\|_F^2 - 2(1 - \delta) \mathbb{E} [\|\hat{\Sigma} - \Sigma\|_F^2] = 0 \quad (26)$$

$$\delta^* = \frac{\mathbb{E} [\|\hat{\Sigma} - \Sigma\|_F^2]}{\mathbb{E} [\|\hat{\Sigma} - \Sigma\|_F^2] + \|\mathbf{F} - \Sigma\|_F^2} \quad (27)$$

ここで、以下のように変数を定義する。

$$\tau^2 := \mathbb{E} [\|\hat{\Sigma} - \Sigma\|_F^2] \quad (\text{標本共分散行列の分散}) \quad (28)$$

$$l^2 := \|\Sigma - \mathbf{F}\|_F^2 \quad (\text{ターゲット行列のバイアス}) \quad (29)$$

$$\zeta^2 := \mathbb{E} [\|\hat{\Sigma} - \mathbf{F}\|_F^2] \quad (\text{標本共分散行列とターゲット行列の平均二乗誤差}) \quad (30)$$

先ほどの内積の期待値が 0 になる性質 $\mathbb{E} [\langle \mathbf{F} - \Sigma, \hat{\Sigma} - \Sigma \rangle] = 0$ を用いて ζ^2 を展開すると、以下の関係式が得られる。

$$\zeta^2 = \mathbb{E} [\|(\hat{\Sigma} - \Sigma) + (\Sigma - \mathbf{F})\|_F^2] \quad (31)$$

$$= \mathbb{E} [\|\hat{\Sigma} - \Sigma\|_F^2] + \|\Sigma - \mathbf{F}\|_F^2 + 2\mathbb{E} [\langle \hat{\Sigma} - \Sigma, \Sigma - \mathbf{F} \rangle] \quad (32)$$

$$= \tau^2 + l^2 \quad (33)$$

したがって、最適な縮小強度 δ^* は以下のように極めてシンプルな形で表現できる。

$$\delta^* = \frac{\tau^2}{\tau^2 + l^2} \quad (34)$$

$$= \frac{\tau^2}{\zeta^2} \quad (35)$$

この結果は、最適な縮小強度が「標本共分散行列の推定誤差（分散 τ^2 ）」を「全体の推定誤差（MSE ζ^2 ）」で割った比率として決定されることを示している。すなわち、推定誤差 τ^2 が大きい（ N が大きい）場合や、ターゲットのバイアス l^2 が小さい場合ほど、 δ^* は 1 に近づき、ターゲット行列への配分が増加する。 ζ^2 はデータから直接計算可能な不偏推定量が存在するが、 τ^2 は未知の真の共分散行列 Σ に依存するため、実務的には Ledoit and Wolf [2] の手法を用いて τ^2 の一致推定量を計算し、 $\delta^* = \hat{\tau}^2 / \hat{\zeta}^2$ として適用する。

なお、本研究ではターゲット行列 \mathbf{F} として、スケーリングされた単位行列（Scaled Identity Matrix）を採用する。これは相関構造を一切仮定しない最もシンプルなモデルであり、高次元・小標本下における推定誤差の混入を最小限に抑えることが期待できる。また、ファクター構造を積極的に取り入れる POET などの手法と対比する上でも、最も構造的仮定の少ないベースラインとして適している。近年では、資産を産業セクター等にクラスタリングしてブロックごとの相関構造を反映させる Blockbuster 手法 [3] など提案されているが、本研究ではこのように最も基礎的なターゲットにおける有効性を検証する。

3.4.2 非線形縮小推定 (Nonlinear Shrinkage)

線形縮小推定はすべての固有値を一律に補正するが、Ledoit and Wolf [10, 11] および De Nard [3] は、ランダム行列理論（Random Matrix Theory: RMT）の知見に基づき、個々の固有値に対して最適な非線形変換を施す手法を提案した。

高次元設定（ $N/T \rightarrow c > 0$ ）において、標本共分散行列の固有値分布は、Marchenko-Pastur 則に従って真の固有値分布よりも広く拡散することが知られている。具体的には、小さな固有値はより小さく、大きな固有値はより大きく推定されるバイアスを持つ。非線

形縮小推定では、標本共分散行列のスペクトル分解を $\hat{\Sigma} = \sum_{i=1}^N \lambda_i u_i u_i^\top$ としたとき、固有ベクトル u_i はそのまま保持し、固有値 λ_i のみを最適化された固有値 \tilde{d}_i に置き換える。

$$\hat{\Sigma}_{NonLin} = \sum_{i=1}^N \tilde{d}_i u_i u_i^\top \quad (36)$$

ここで \tilde{d}_i は、観測された固有値分布から真の固有値分布を復元する QuEST (Quantized Eigenvalues Sampling Transform) 関数を用いて推定される。これにより、線形縮小よりも柔軟かつ高精度に、大次元共分散行列特有のバイアスを除去することが可能となる。

4 実験内容

提案手法の有効性を検証するため、米国の株式市場の個別銘柄のデータを用いたバックテストを行う。

4.1 データセット

- **対象:** 米国株式市場における個別銘柄
- **期間:** 1995 年 3 月 1 日から 2023 年 12 月 29 日
- **データ処理:** 日次リターンを使用し、欠損値は 0 で補完する。

なお、本研究のソースコードは GitHub にて公開している (https://github.com/kk1341/Hirodai_Research)。

4.2 実験設定

本研究では、目的の異なる以下の 2 つの実験を行う。

4.2.1 実験 1：統計的性能評価 (Statistical Performance Evaluation)

中規模から高次元の設定において、各手法の平均的なパフォーマンスと安定性を統計的に評価することを目的とする。

- **銘柄数 (N):** 30, 50, 100, 200, 500 銘柄
- **試行回数:** 各設定につき 30 回のモンテカルロ・シミュレーション
- **訓練期間 (T_{train}):** 21 日 (約 1 ヶ月)
- **バックテスト期間:** 訓練期間直後の 1 日 (ローリングウィンドウ方式)

4.2.2 実験 2：大規模実証実験 (Large-Scale Demonstration)

TOPIX などの主要指数全体を運用する場合を想定し，超高次元環境における縮小推定の実用性を検証することを目的とする．

- 銘柄数 (N): 1000 銘柄
- 試行回数: 1 回 (実運用を想定したケーススタディ)
- その他の設定: 実験 1 と同様

4.3 評価指標

両実験において，以下の指標を用いて評価を行う．

- シャープ・レシオ (Sharpe Ratio): 年率換算したリターンをリスク（標準偏差）で除したもの．

5 実験結果

本節では，前節で設定した 2 つの実験環境における各手法のパフォーマンスを報告する．

5.1 実験 1：統計的性能評価 ($N = 30 \sim 500$)

中規模から高次元の設定における最小分散ポートフォリオの平均シャープ・レシオを表 1 に示す．

Table1 各手法における平均シャープ・レシオ (Mean Sharpe Ratio)

N	Equal Weight	Linear Shrinkage	Market Factor	Nonlinear Shrinkage	PCA	POET	Sample
30	0.6980	0.6762	0.6372	0.5778	0.5432	0.5256	0.0082
50	0.7387	0.7773	0.7180	0.7208	0.6674	0.6744	0.0408
100	0.7638	0.8320	0.7581	0.8231	0.7464	0.7456	0.0186
200	0.7735	0.9405	0.7694	0.9781	0.7648	0.7652	0.1324
500	0.7857	1.0636	0.7713	1.1039	0.7696	0.7687	0.0950

表 1 の結果から，以下の傾向が確認された．

まず，ポートフォリオ構築における標本共分散行列の限界が明らかとなった．標本共分散行列は，いずれの N においてもシャープ・レシオが 0.008 \sim 0.132 という極めて低い水準に留まっている．Figure 2 の箱ひげ図を参照すると，この手法は N の増加とともに分布

が極端に拡大しており、高次元下での推定誤差の増大が運用パフォーマンスを著しく不安定化させていることが確認できる。

次に、低次元領域における単純な戦略の堅牢性が挙げられる。 $N = 30$ のように銘柄数が訓練期間 ($T = 21$) に近い設定では、マーケットファクターモデルが 0.637 と比較的安定した結果を示した。しかし、これ以上に高いパフォーマンスを記録したのが等配分戦略 (0.698) や線形縮小推定 (0.676) である。すなわち、次元が低い段階では、強力な構造を仮定するモデルよりも、単純な等配分や縮小推定の方が有効に機能している。

一方で、 $N = 50$ を超えると傾向は一変し、縮小推定の優位性が明確となる。次元が増加するにつれて線形縮小推定がマーケットファクターを逆転し、より高いパフォーマンスを発揮するようになった。特に $N = 100, 200, 500$ と次元が増すほど、線形および非線形縮小推定の優越性は顕著となり、Figure 2 においても、高い平均値を維持しつつ四分位範囲が狭いという、安定した特性を示している。

5.1.1 考察

実験 1 の結果から、現代の株式市場における各手法のパフォーマンスについて以下のような考察が得られる。

第一の論点は、推定誤差と最適化のトレードオフである。 $N = 30$ という低次元設定において等配分戦略が高いパフォーマンスを維持したのは、限られたデータからパラメータを推定する「コスト（誤差）」が、最適化を行うことで得られる「恩恵（リターン改善）」を上回ったためと解釈できる。情報量が不十分な状況下では、ノイズを含むデータで無理に最適化を行うよりも、パラメータ推定を一切行わない等配分戦略の方が、結果としてロバストな解となる場合があることを示唆している。

第二に、高次元・小標本という条件下において、縮小推定が等配分戦略を上回るパフォーマンスを示した点である。通常、高次元データでは標本共分散行列の推定誤差が支配的となり、パラメータ推定を放棄する等配分戦略が最適な解となることが多い。しかし、本実験では縮小推定が等配分戦略を上回るパフォーマンスを示した。これは、縮小推定が単にデータと単位行列を混ぜ合わせているだけでなく、標本共分散行列に含まれる固有値の拡散といったノイズを抑制しつつ、データの中に存在する相関構造の抽出に成功していることを示唆している。つまり、高次元データにおいては情報をすべて捨てる等配分戦略よりも、情報を正則化して活用する縮小推定の方が、リスク低減において有意であることが確認された。

第三に、ファクターモデルのパフォーマンスが中程度に留まった要因として、モデルの特定化誤差が挙げられる。マーケットファクターや POET といった手法は、少数の共通要因でリターンを説明しようとするが、現実の高次元データにはファクターでは捉えきれな

い複雑な局所的相関が残差として存在する．特に本実験のように銘柄数に対してサンプル数が少ない場合，POET 等の手法では残差に含まれるシグナルとノイズの分離が困難であり，その結果，パフォーマンスが縮小推定ほど伸びなかったと推察される．

最後に，縮小推定が等配分戦略を大きく上回った事実は，市場における情報の存在を証明している．等配分戦略は相関情報をすべて破棄するアプローチであるが，縮小推定は適切な正則化を通じてその情報を抽出することに成功している．現代のような高次元の投資環境においては，情報を捨てるのではなく，適切に制御して使い切るアプローチが不可欠であることが示された．

5.2 実験 2：大規模な銘柄数における実証実験 ($N = 1000$)

次に，TOPIX などの主要指数全体を運用する場合を想定した，超高次元設定 ($N = 1000$) におけるシミュレーション結果を表 2 に示す．

Table2 大規模な銘柄数 ($N = 1000$) におけるシャープ・レシオ

Method	Sharpe Ratio
Equal Weight	0.7884
Linear Shrinkage	1.1159
Market Factor	0.7717
Nonlinear Shrinkage	1.1484
PCA	0.7709
POET	0.7709
Sample Covariance	0.1035

この結果は，実験 1 で得られた知見が，実務的な大規模ポートフォリオにおいても成立することを強く支持している．

第一に，縮小推定の堅牢性である． $N = 1000$ という超高次元環境であっても，線形縮小推定 (1.116) および非線形縮小推定 (1.148) は 1.1 を超える高いシャープ・レシオを達成した．実験 1 の考察で述べたように，これは縮小推定が標本共分散行列のノイズを抑制しつつ，有用な相関構造を抽出するという効果が，規模が拡大しても破綻することなく機能し続けていることを証明している．

第二に，ファクターモデルおよび等配分戦略の限界である．マーケットファクター，PCA，POET などのファクターモデルは 0.77 ～ 0.79 程度に留まり，等配分戦略 (0.788) と同等の水準であった．これは，実験 1 で確認された特定化誤差の影響が大規模化によって解消されるわけではないことを示唆している．むしろ，銘柄数が増えるほど局所的な相関

構造は複雑化するため、少数の共通ファクターで全体を説明するアプローチの限界がより鮮明になったと言える。

第三に、縮小推定の中でも非線形縮小推定が線形縮小推定を上回るパフォーマンスを示した要因である。これは、両手法における固有値の扱いの違いに起因する。線形縮小推定は、すべての固有値を一律の強度 δ でターゲット方向へ縮小させるため、ノイズの除去とシグナルの維持の間に強いトレードオフが存在する。対照的に、非線形縮小推定はランダム行列理論に基づき、個々の固有値に対して最適な非線形変換を適用する。 $N = 1000$ のような超高次元下では、標本固有値の分布が極端に拡散するため、一律の縮小を行う線形手法よりも、固有値ごとにきめ細かな補正を行う非線形手法の方が、ノイズを効果的に抑制しつつ有用な情報を保持できたと解釈できる。

結論として、TOPIX 全体のような大規模ユニバースを運用する場合においても、情報を適切に制御して活用する縮小推定アプローチが、ファクターモデルや等配分戦略よりも優れたパフォーマンスを発揮することが実証された。

6 まとめ

本研究では、高次元・小標本環境におけるポートフォリオ最適化問題に対し、Ledoit and Wolf [2] の縮小推定法をはじめとする共分散行列推定手法の有効性を、米国の株式市場データを用いて検証した。

実証実験の結果、以下の3つの知見が得られた。第一に、情報量が限られる低次元設定では、推定誤差を伴う最適化よりも、パラメータ推定を行わない等配分戦略がロバストな解となる場合があることである。第二に、従来の手法が破綻する高次元設定において、縮小推定アプローチは標本共分散行列のノイズを抑制しつつ、有用な相関構造を抽出することで、等配分戦略をも上回る優れたパフォーマンスを発揮することである。第三に、ファクターモデルは共通要因の抽出には有効であるが、高次元データ特有の局所的な相関構造を捉えきれず、パフォーマンスの改善には限界があることである。

結論として、現代の複雑な投資環境においては、情報を捨てるのではなく、縮小推定を通じて適切に正則化して活用するアプローチが、リスク低減とリターン向上の両面において最も有効であることが示された。

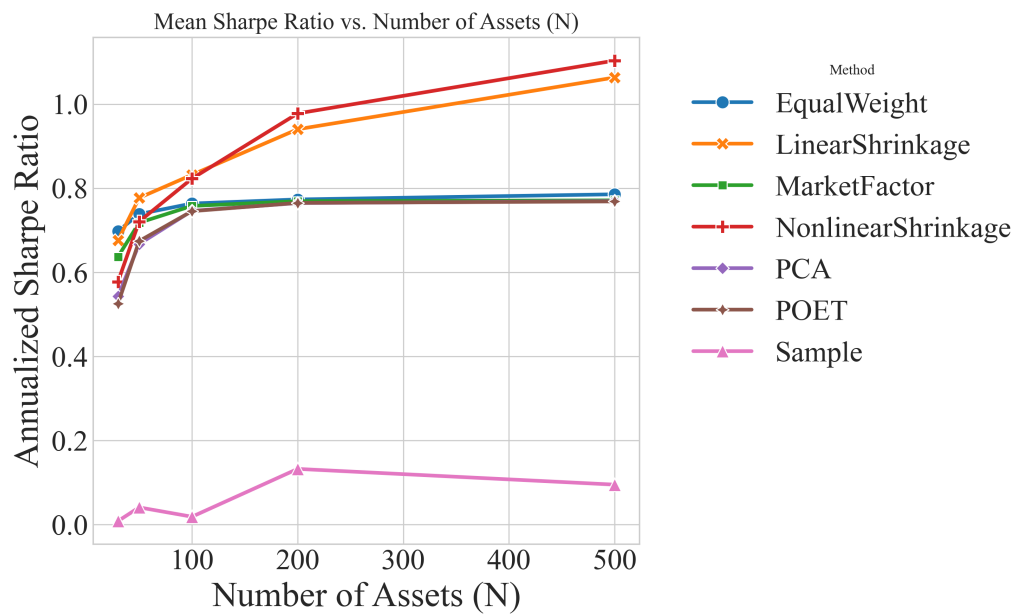


Figure 1 銘柄数 (N) と平均シャープ・レシオの関係

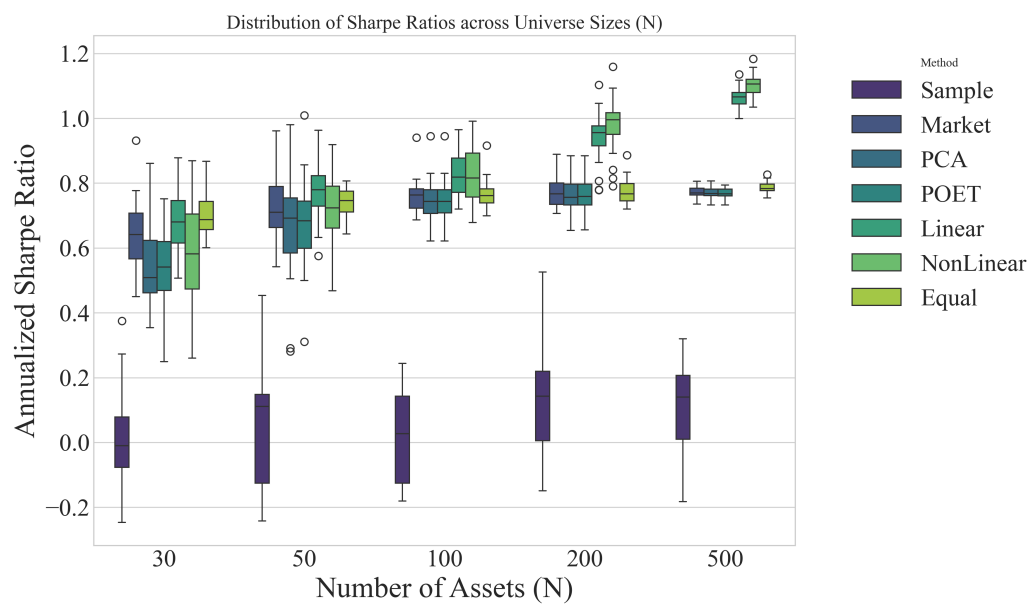


Figure 2 各手法におけるシャープ・レシオの分布比較 ($N = 30 \sim 500$)

References

- [1] Harry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77--91, 1952.
- [2] Olivier Ledoit and Michael Wolf. Honey, I shrunk the sample covariance matrix. *The Journal of Portfolio Management*, 30(4):110--119, 2004.
- [3] Gianluca De Nard. Oops! I shrunk the sample covariance matrix again: Blockbuster meets shrinkage. *Journal of Financial Econometrics*, 20(4):569--611, 2022.
- [4] J. D. Jobson and Bob Korkie. Estimation for markowitz efficient portfolios. *Journal of the American Statistical Association*, 75(371):544--554, 1980.
- [5] Richard O Michaud. The markowitz optimization enigma: Is 'optimized' optimal? *Financial Analysts Journal*, 45(1):31--42, 1989.
- [6] Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The Review of Financial Studies*, 22(5):1915--1953, 2009.
- [7] Robert C Merton. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8(4):323--361, 1980.
- [8] Olivier Ledoit. *Essays on risk and return in the stock market*. PhD thesis, Massachusetts Institute of Technology, 1995.
- [9] Jianqing Fan, Yingying Liao, and Martina Mincheva. Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75(4):603--680, 2013.
- [10] Olivier Ledoit and Michael Wolf. Nonlinear shrinkage estimation of large-dimensional covariance matrices. *The Annals of Statistics*, 40(2):1024--1060, 2012.
- [11] Olivier Ledoit and Michael Wolf. Spectrum estimation: A unified framework for covariance matrix estimation and pca in large dimensions. *Journal of Multivariate Analysis*, 139:360--384, 2015.