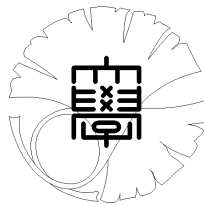


数理科学実践研究レター 2026-4 July 14, 2026

**Ridge-regularized mosaic permutation test**

by

澤谷 一磨



**UNIVERSITY OF TOKYO**

GRADUATE SCHOOL OF MATHEMATICAL SCIENCES

KOMABA, TOKYO, JAPAN

## Ridge-regularized mosaic permutation test

澤谷一磨<sup>1</sup>（東京大学大学院経済学研究科）

Kazuma Sawaya (Graduate School of Economics, The University of Tokyo)

## 概要

ファンダメンタル因子モデルは、金融分野の実務で頻繁に用いられるモデルだが、そのモデル自体が実データに即しているかを検証することは重要な問題となる。本研究では、ファンダメンタル因子モデルに対する適合度検定の一種であるモザイク並べ替え検定を拡張し、その安定性と検出力を向上するためにリッジ正則化モザイク並べ替え検定を提案する。提案法の数値的性能を人工データと S&P500 対象銘柄を用いた実データから検証し、その有用性を明らかにした。

## 1 はじめに

ファンダメンタル因子モデルは、経済学および金融におけるリスク管理のためのもっとも一般的な道具の一つである (Rosenberg, 1974; Grinold and Kahn, 1994; Giglio *et al.*, 2022)。このモデルを通じて資産リターン間の相関関係をモデリングすることで、ポートフォリオ内の多くの資産が同時に価値を失う確率を推定することが可能になる。一方で、コロナショックや金融危機など、状況の変化に応じてモデル自体が依然として適切に現実を説明しているか否かを評価する必要がある。実際、誤ったモデルに基づくリスク推定はバイアスを生み、誤った意思決定を招く恐れがある。そこで本研究では、ファンダメンタル因子モデルの実データに対する妥当性の検定手法を提案する。以下では、まず 2 節でファンダメンタル因子モデルを定義し、3 節でモザイク並べ替え検定という先行手法をレビューする。4 節において提案手法を提示し、5 節においてその数値的な検証を行う。

## 2 ファンダメンタル因子モデル

各時刻  $t = 1, \dots, T$  に対して、 $p$  個の資産に対するリターン  $Y_t \in \mathbb{R}^p$  が次のファンダメンタル因子モデルに従うとする：

$$Y_t = L_t X_t + \epsilon_t, \quad (1)$$

ここで、 $X_t, L_t, \epsilon_t$  は以下のように定義される。

- $X_t \in \mathbb{R}^k$  は資産間の相関構造を駆動する  $k$  ( $\ll p$ ) 個の因子のリターンを表すベクトル。
- $L_t \in \mathbb{R}^{p \times k}$  は因子負荷行列と呼ばれる行列で、 $(L_t)_{j\ell}$  は  $\ell$  番目の因子の  $j$  番目の資産に対する重みを与える（以下、フルランクとする）。
- $\epsilon_t \in \mathbb{R}^p$  は因子で説明できない個別的なノイズで、以下では残差と呼ぶ。

このモデルは capital asset pricing model (CAPM) (Sharpe, 1964), arbitrage pricing theory (APT) (Ross, 1976), Fama-French (Fama and French, 1993, 2015), BARRA (Grinold and Kahn, 2000) など多くの実務的なモデルを含む。 $L_t$  はたとえば財務会計データから計算されるほか、バイナリ値の場合であれば各要素は対応銘柄がどの業界に属しているかを表現できる。本稿では、(Spector *et al.*, 2024) に従い、実務的に広く普及しているアプローチとして  $X_t$  を未知、 $L_t$  を既知として扱う。

## 3 モザイク並べ替え検定

本節では、本論文で考慮する適合度検定の帰無仮説を提示するほか、先述のモデル (1) が正確に資産間の相関構造をモデリングできているか否かを統計的に検定する手法であるモザイク並べ替え検定

<sup>1</sup>sawaya@g.ecc.u-tokyo.ac.jp

(Spector *et al.*, 2024) を紹介する. まず,  $\epsilon_{\cdot,j} := (\epsilon_{1,j}, \dots, \epsilon_{T,j}) \in \mathbb{R}^T$  としたときに, 以下の帰無仮説を考える:

$$\mathcal{H}_0 : \epsilon_{\cdot,1}, \dots, \epsilon_{\cdot,p} \in \mathbb{R}^T \text{ are jointly independent.} \quad (2)$$

これは各資産に対する残差が互いに独立であるという帰無仮説であり, 因子で説明できない資産間依存性が存在しないという意味でファンダメンタル因子モデルのデータへの適合性をあらわしている. この仮説は残差の自己相関・非定常性・不均一性を許容している点で金融データの特徴を抑えているため, 対応する検定手法が実応用上重要となる.

### 3.1 並べ替え検定

まず, ノンパラメトリックな独立性検定の古典的な方法として並べ替え検定 (permutation test) (van der Vaart, 1998) を紹介する. 独立同分布要素をもつ 2 つの確率ベクトル  $(X_1, \dots, X_n)$  と  $(Y_1, \dots, Y_n)$  の独立性検定を考える. 非独立性の尺度を  $\mathcal{T} := \mathcal{T}((X_1, Y_1), \dots, (X_n, Y_n))$  とする. ここで重要な性質として,  $X$  と  $Y$  が独立であれば, 任意の置換  $\sigma : [n] \rightarrow [n]$  について, 帰無仮説 (2) のもとでの置換不変性が知られている:

$$\mathcal{T}((X_1, Y_1), \dots, (X_n, Y_n)) \stackrel{d}{=} \mathcal{T}((X_1, Y_{\sigma(1)}), \dots, (X_n, Y_{\sigma(n)})). \quad (3)$$

したがって, 並べ替えによって生じる  $n!$  通りの  $\mathcal{T}(\cdot)$  のなす経験分布が帰無分布となり, それらのうち最も極端な上位 100 $\alpha$ % に実現値が入っていれば棄却するという有意水準  $\alpha$  の検定を構築できる. なお, 全通りの計算は困難なので通常はモンテカルロ法で近似をおこなう. よって具体的な手順としては, 一様ランダムに  $B$  個の置換に対応する  $\mathcal{T}_1, \dots, \mathcal{T}_B$  を得て,

$$\frac{1 + \sum_{k=1}^B 1\{\mathcal{T} \leq \mathcal{T}_k\}}{B + 1} \leq \alpha. \quad (4)$$

である場合に独立性の仮説を棄却する.

では, この一般的な方法を今回のファンダメンタル因子モデルに対する検定問題 (2) に適用することを考える. 説明の簡単のために  $\epsilon_1, \dots, \epsilon_T$  が独立同分布に従うと仮定する. このときファンダメンタル因子モデルの場合, 残差が既知であれば, 独立性に関する帰無仮説のもと, 各資産ごとの任意の置換  $\pi_1, \dots, \pi_p$  について以下が成立する:

$$\epsilon := \begin{bmatrix} \epsilon_{1,1} & \epsilon_{1,2} & \cdots & \epsilon_{1,p} \\ \epsilon_{2,1} & \epsilon_{2,2} & \cdots & \epsilon_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{T,1} & \epsilon_{T,2} & \cdots & \epsilon_{T,p} \end{bmatrix} \stackrel{d}{=} \begin{bmatrix} \epsilon_{\pi_1(1),1} & \epsilon_{\pi_2(1),2} & \cdots & \epsilon_{\pi_p(1),p} \\ \epsilon_{\pi_1(2),1} & \epsilon_{\pi_2(2),2} & \cdots & \epsilon_{\pi_p(2),p} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{\pi_1(T),1} & \epsilon_{\pi_2(T),2} & \cdots & \epsilon_{\pi_p(T),p} \end{bmatrix}. \quad (5)$$

したがって, 先述の手順に従って並べ替え検定が構成できる. しかし, 残差  $\epsilon$  は直接観測できないため, 上記の構造を保持した残差推定量を構成する必要がある. もっとも代表的な残差の推定量は, 因子リターンの最小二乗推定量

$$\hat{X}_t = \arg \min_{X \in \mathbb{R}^k} \|Y_t - L_t X\|_2^2 = (L_t^\top L_t)^{-1} L_t^\top Y_t \quad (6)$$

をモデルにプラグインした,

$$\hat{\epsilon}_t^{\text{ols}} = H_t Y_t, \quad H_t = I - L_t (L_t^\top L_t)^{-1} L_t^\top \quad (7)$$

である. なお,  $L_t$  のフルランク性と  $k < p$  を仮定しているため逆行列は常に well-defined である. しかし, この残差推定量  $\hat{\epsilon}_t^{\text{ols}}$  は帰無仮説 (2) のもとで本来の残差がもつ独立性を保存しないため, 直接的に妥当な検定手順を構築することができない. このことは残差推定量の共分散行列  $\text{Cov}(\hat{\epsilon}_t^{\text{ols}}) = H_t \text{Cov}(\epsilon_t) H_t$  が一般に対角行列とならないことから確かめられる. この問題を解決したのが次に述べるモザイク並べ替え検定である.

### 3.2 モザイク並べ替え検定

モザイク並べ替え検定 (Spector *et al.*, 2024) では、まず資産リターンを時間方向に並べた行列  $\mathbf{Y} \in \mathbb{R}^{T \times p}$  を時間・資産の両軸に沿って  $M$  個の排反な長方形へと分割する。この各分割をタイルと呼ぶ。具体的な手順は Algorithm 1 に示した通り。

---

**Algorithm 1** モザイク並べ替え検定の擬似コード

---

**Require:** 資産リターン  $\mathbf{Y} \in \mathbb{R}^{T \times p}$ , 因子負荷行列  $L_t \in \mathbb{R}^{p \times k}$ ,  $t = 1, \dots, T$ , タイルの分割, 検定統計量  $S: \mathbb{R}^{T \times p} \rightarrow \mathbb{R}$ , モンテカルロ計算の繰り返し数  $R$ .

- 1: 各タイル  $m = 1, \dots, M$  における収益と残差推定量の小行列を  $\mathbf{Y}_{(m)}$ ,  $\hat{\epsilon}_{(m)}$  として、後者をタイル内の最小二乗推定量から計算する。  $\{\hat{\epsilon}_{(m)}\}_{m=1, \dots, M}$  からなる行列全体を  $\hat{\epsilon} \in \mathbb{R}^{T \times p}$  とおく。
- 2: 各タイルでの残差推定量に対して、ランダムな行の置換をしたときの全体の行列を  $\tilde{\epsilon}^{(r)} \in \mathbb{R}^{T \times p}$ ,  $r = 1, \dots, R$  とする。
- 3: 任意の検定統計量  $S$  に対して次の  $p$  値を計算する:

$$p_{\text{val}} := \frac{1 + \sum_{r=1}^R \mathbf{1}\{S(\hat{\epsilon}) \leq S(\tilde{\epsilon}^{(r)})\}}{R + 1}. \quad (8)$$


---

タイル内での因子負荷行列の不変性及びタイル内における残差の交換可能性を仮定する。

**仮定 1**  $L_t$  は同一タイル上で定数。

**仮定 2 (局所交換可能性)** 各資産  $j = 1, \dots, p$  について、 $\mathbf{Y}_{t,j}$  と  $\mathbf{Y}_{\pi(t),j}$  が常に同じタイルに入るような任意の置換を  $\pi: [T] \rightarrow [T]$  とする。このとき、

$$(\epsilon_{1,j}, \dots, \epsilon_{T,j}) \stackrel{d}{=} (\epsilon_{\pi(1),j}, \dots, \epsilon_{\pi(T),j}). \quad (9)$$

これらの仮定は (i) 残差と因子の分布の任意の裾重性, (ii) 因子の任意の自己相関と非定常性, (iii) タイル間残差の非定常性, (iv) 任意の検定統計量の使用を許容している点で一般的である。このもとで以下が言える。

**定理 3 (Spector *et al.* (2024))** 仮定 1–2 が成立するとき、 $\mathcal{H}_0$  のもとで  $p_{\text{val}}$  は有効である。すなわち、任意の  $\alpha \in (0, 1)$  に対して  $\Pr(p_{\text{val}} \leq \alpha) \leq \alpha$ 。

証明略。

## 4 リッジ正則化モザイク並べ替え検定

本節では、残差推定量として最小二乗法ではなく、リッジ正則化推定量を用いた方法を提案する。具体的には、各タイル  $m = 1, \dots, M$  上で因子リターンの最小二乗推定に対してユークリッドノルムの二乗で罰則を付けたもの

$$\tilde{X}_t = \arg \min_{X \in \mathbb{R}^k} \|\mathbf{Y}_{(m)} - L_{(m)}X\|_2^2 + \lambda \|X\|_2^2 = (L_{(m)}^\top L_{(m)} + \lambda I)^{-1} L_{(m)}^\top \mathbf{Y}_{(m)}, \quad \lambda > 0 \quad (10)$$

をモデルにプラグインして得られる残差推定量

$$\tilde{\epsilon}_{(m)} = \tilde{H}_{(m)} \mathbf{Y}_{(m)}, \quad \tilde{H}_{(m)} = I - L_{(m)}(L_{(m)}^\top L_{(m)} + \lambda I)^{-1} L_{(m)}^\top, \quad (11)$$

を  $\hat{\epsilon}_m$  の代わりに Algorithm 1 で用いる。このようにして残差推定量 (11) からつくられる検定手法をリッジ正則化モザイク並べ替え検定と呼ぶことにする。現代の数理統計学においては、リッジ正則化推定量は推定の安定性を高め、それに基づく仮説検定の検出力向上をもたらすことが知られているため、より信頼性の高い適合度検定となることが期待される。

## 5 数値実験

### 5.1 人工データ

まず人工的に生成したデータに対するモザイク並べ替え検定とリッジ正則化モザイク並べ替え検定の結果を比較する。今回、 $(T, p, k) = (1000, 100, 90)$  とし、 $t = 1, \dots, 500$  と  $t = 501, \dots, T$  で 2 つのレジームに分けた。具体的には、Toeplitz 行列で構成した因子負荷行列  $(L_t)_{ij} = \rho^{|i-j|}$  が前半のレジームでは  $\rho = 0.9$ 、後半では  $\rho = 0.4$  で固定した。また、AR(1) 過程から因子リターンを生成した： $t = 1, \dots, T-1$  に対し、

$$X_{t+1} = 0.8X_t + \mathcal{N}(0, 1), \quad X_1 \sim \mathcal{N}_k(0, I). \quad (12)$$

これに対して  $(\epsilon_t)_j \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 0.1)$  を加えることでリターン  $Y_t$  をそれぞれ生成する。なお、以下全ての実験ではモンテカルロ繰り返し回数  $R = 500$  を採用し、Spector *et al.* (2024) の Section 3.3 で推奨されているデフォルトタイル選択方式を用いたほか、 $\lambda = 0.5$  を固定する。

**null ケース.** まず、帰無仮説を棄却すべきでない、データ生成に用いた真の  $L_t$  を既知として  $t = 1, \dots, 500$  の観測のみを用いたケースを扱う。図 1–2 の結果は実際の並べ替えによって生じた統計量の経験分布をヒストグラムで、観測に対応する統計量を縦赤線でプロットしている。null ケースを扱っているため赤線がヒストグラムの右側に寄り過ぎないことが期待されるが、既存法はそれに失敗し、提案法は成功している。

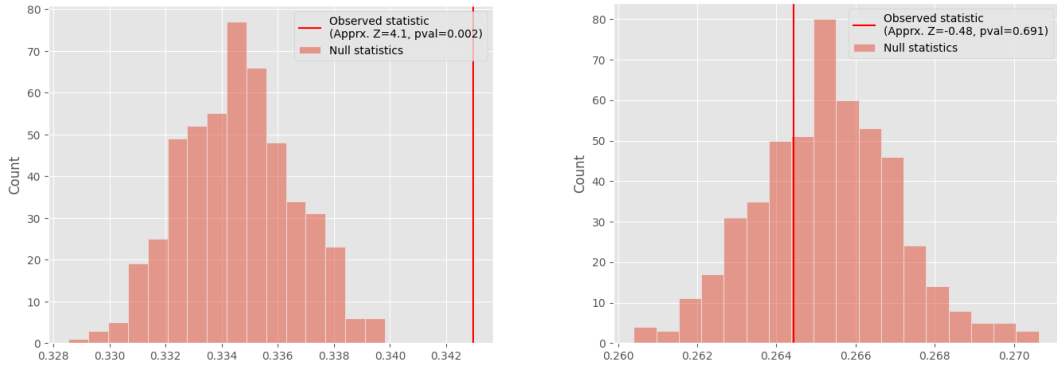


図 1: 帰無仮説が正しいもとでの先行手法. 誤って棄却している. 図 2: 帰無仮説が正しいもとでの提案手法. 正しく棄却せずにいる.

**non-null ケース.** 続いて、 $t = 1, \dots, T$  のフルサンプルを用いて、しかし前半の因子負荷行列のみを全サンプルで使い続ける場合を考える。この場合、因子負荷行列に関するモデル誤特定があるため、帰無仮説の棄却を強く支持することが望ましい。結果が図 3–4 で、既存法と提案法のいずれも正しく棄却を支持している（赤線がヒストグラムの右側にある）一方で、提案手法のほうがより強く棄却を支持しており、リッジ正則化推定量のもたらす検出力の高さが観察できる。

### 5.2 実データ

2025 年 9 月 30 日時点における S&P500 の対象銘柄のうち、直近 90 日間の平均売買代金が多かった 100 銘柄を 2016 年 1 月 1 日から 2025 年 9 月 30 日まで取得し、サイズ・バリュー・モメンタム・ボラティリティ・セクターダミーを因子として計算した。具体的には、Yahoo!ファイナンスの API から各銘柄の情報を取得し、以下の各要素を並べた行列を因子負荷行列として構成した（合計ファクター数は 14）：

- サイズ因子：月末価格と発行株数の積の対数により各月ごと定義。

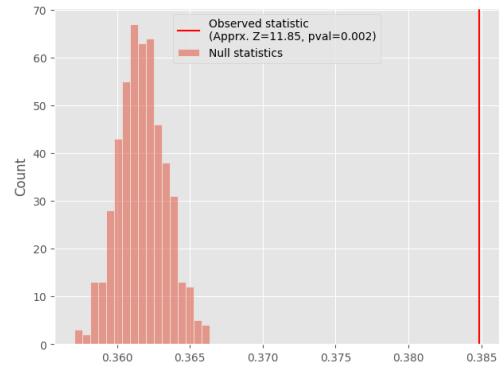
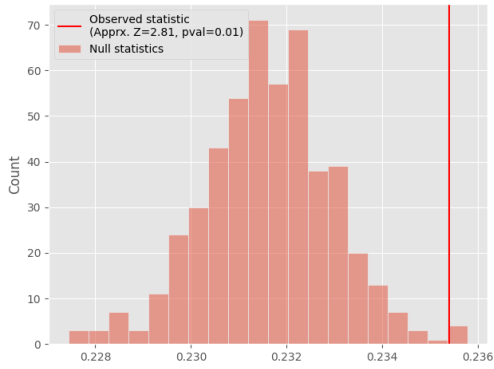


図 3: 帰無仮説が誤りのもとでの先行手法. 検定 図 4: 帰無仮説が誤りのもとでの提案手法. より強く正しい棄却を支持している.

- バリュースコア因子：価格対簿価比率（P/B）の逆数を使用。
- モメンタム因子：直近 1 か月を除いた過去 11 か月の累積リターンを計算。
- ボラティリティ因子：過去 252 日間の日次リターンから計算される標準偏差を使用。
- セクターダミー：Yahoo!ファイナンスから取得可能である全 10 セクター（Basic Materials, Communication Services, Consumer Cyclical, Consumer Defensive, Energy, Financial Services, Healthcare, Industrials, Technology, Utilities）のうち各銘柄の属するセクターをあらわす静的な one-hot ベクトルを使用<sup>2</sup>。

そして、各銘柄の月次リターンを  $Y_t$  とした。このデータに対して既存法と提案法を適用した結果が図 5-6 である。いずれも同程度に帰無仮説の誤りを強く支持しており、今回構築したモデル（とりわけ因子負荷行列の設計）が現実を上手く説明しきれていないことを根拠づけた。

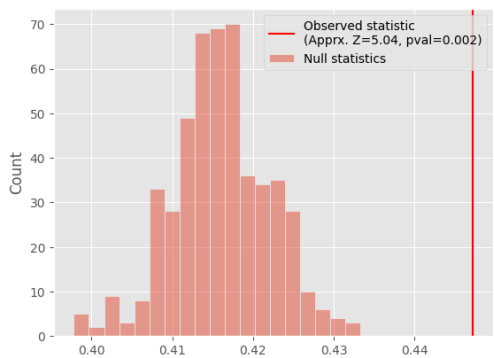


図 5: 帰無仮説の誤りを強く支持。

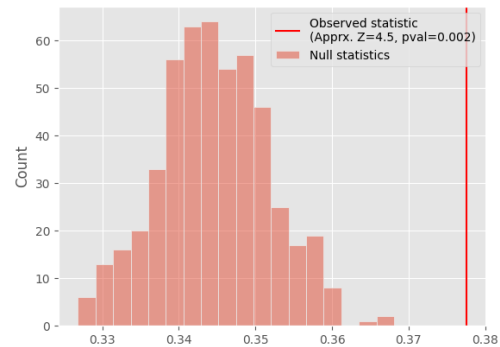


図 6: 先行手法と同程度に帰無仮説の誤りを強く支持。

## 6 終わりに

本研究で提案したリッジ正則化モザイク並べ替え検定は、経済学や金融の実証で頻繁に活用されるファンダメンタル因子モデルが現実のデータと乖離しているかどうかを統計学的に判断するうえで有用であることを数値的に確認した。

一方で、さらに実用を見据えるのであれば、因子負荷行列の設計に関するガイドラインや結果のロバスト性の向上、異なる罰則付き推定等を検討することは有用な方針となる。

<sup>2</sup>S&P 社が用いている Global Industry Classification Standard (GICS) におけるセクターと比較して、Real Estate のセクターを含まない形となる。

## 謝辞

本研究に際し課題の提供をいただいた株式会社三菱 UFJ 銀行に感謝いたします。また、株式会社三菱 UFJ 銀行の上田慎一郎様、荻野卓啓様、島田耕平様、中島卓也様、及び東京大学大学院数理科学研究科の浅井聡太先生、齊藤宣一先生、坪内俊太郎先生、そして同じ利益班として課題に取り組んだ渡部匠様には多くの助言やコメントをいただきました。ここに感謝申し上げます。加えて、本稿の正確性と明快さを改善するために大きく寄与してくださった匿名の査読者の先生に感謝いたします。

## 参考文献

- Fama, E. F. and French, K. R. (1993) Common risk factors in the returns on stocks and bonds, *Journal of financial economics*, **33**, 3–56.
- Fama, E. F. and French, K. R. (2015) A five-factor asset pricing model, *Journal of financial economics*, **116**, 1–22.
- Giglio, S., Kelly, B. and Xiu, D. (2022) Factor models, machine learning, and asset pricing, *Annual Review of Financial Economics*, **14**, 337–368.
- Grinold, R. and Kahn, R. N. (1994) Multiple-factor models for portfolio risk, *A practitioner's guide to factor models*, pp. 59–85.
- Grinold, R. C. and Kahn, R. N. (2000) *Active portfolio management*, McGraw Hill New York.
- Rosenberg, B. (1974) Extra-market components of covariance in security returns, *Journal of Financial and quantitative analysis*, **9**, 263–274.
- Ross, S. A. (1976) The arbitrage theory of capital asset pricing, *Journal of Economic Theory*, **13**, 341–360.
- Sharpe, W. F. (1964) Capital asset prices: A theory of market equilibrium under conditions of risk, *The journal of finance*, **19**, 425–442.
- Spector, A., Barber, R. F., Hastie, T., Kahn, R. N. and Candès, E. (2024) The mosaic permutation test: an exact and nonparametric goodness-of-fit test for factor models, *arXiv preprint arXiv:2404.15017*.
- van der Vaart, A. W. (1998) *Asymptotic statistics*, Cambridge University Press.