

変化点検知とスパース推定に基づくポートフォリオ選択

Markowitz Portfolio Selection with Change-point Detection and Sparse Estimation

岡 滉 河原 吉伸 鷲尾 隆
Akira Oka Yoshinobu Kawahara Takashi Washio

大阪大学 産業科学研究所
The Institute of Scientific and Industrial Research, Osaka University

In recent years, a portfolio selection based on sparse markowitz model has been studied extensively, and generally it is assumed that the length of the learning period and testing period are fixed. However, dynamics of a market is often changed in learning process or test process, and we can't get supposed results in this case. In this paper, we propose a sparse markowitz portfolio selection with change-point detection to consider changing of dynamics of a market. Additionally, we also use clusters of the past data to overcome the weakness of the method we propose. We show some empirical examples, where the proposed portfolio selections are compared with the conventional one by using TOPIX500 data.

1. はじめに

近年、投資におけるポートフォリオ選択において、1952年にH.Markowitz氏が提案したリスク最小化に基づくポートフォリオ選択モデルを基礎として、様々な研究がなされている[1]. Markowitzモデルとは、ある程度のデータが与えられた時、そこから期待されるリターンの分散を最小化することでより良い資産分配の割合を決めるものである。

最近では、このMarkowitzモデルに様々な正則化項や制約を課したものが提案されている。J.Brodie氏は2009年にMarkowitzモデルにL1正則化項を組み込むことで、より実用的なポートフォリオ選択となることを発表している[2]. また、同じく2009年にDeMiguel氏らは、投資割合にノルム制約を課した上で分散を最小化することで、ポートフォリオの推定誤差を減少できることを指摘した[3].

これらの手法では、図1上側のタイムチャートに示すように、基本的に等期間の訓練と等期間の投資を仮定しており、途中でダイナミクスが変動する可能性を考慮していない。しかし、実際の株式市場では途中で分布が大きく変動することは多々あり、その際に想定した結果が得られないという問題がある。

そこで本研究では、図1下側のタイムチャートに示すように、変化点検知を用いて市場の変化を検知することで、前述の問題を緩和させる手法について提案する。変化の検知は、密度比推定に基づく変化点検知を用いる[4]. 変化を検知したところで投資を中断し、新たに訓練を開始することで、分布が大きく異なる期間で同じポートフォリオで投資を続けることを防ぐ。

しかしながら、この提案手法のままでは訓練期間が多くなる傾向があり、実際に投資を行わない空白期間が増えてしまう問題を有する。そこで、過去のデータを密度比に応じてクラスタリングしておき、訓練期間の一部として用いることで、空白期間を減らし、投資する期間を増やすことを試みる。

最後に、これらの手法を2003~2013年のTOPIX500の銘柄を用いて検証を行った結果を示す。

本稿の構成は以下のようである。まず、2.では先行研究として、ポートフォリオ選択モデルと、本研究に用いる密度比推定に基づく変化点検知について述べる。3.では提案手法につい

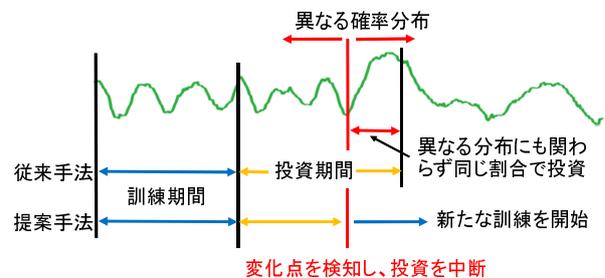


図1: 従来手法と提案手法のタイムチャート

て説明する。3.1.で変化点検知を用いた手法について提案し、3.2.では過去のデータのクラスタリングを用いた手法について説明する。4.ではTOPIX500のデータを用いて、提案手法の検証を行った結果を示し、最後に5.では結論を述べる。

2. 先行研究

2.1 ポートフォリオ選択モデル

Markowitzモデルは、リターンの分散をリスクとして、必要とするリターンを制約式として組み込んだ上で、リスクを最小化することで、目的とする低いリスクと高いリターンを実現しようとするものである。銘柄数 D の株式の時間 t におけるリターンベクトルを \mathbf{r}_t として $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_T)^\top$ とし、 \mathbf{r}_t の分散共分散行列を $\hat{\Sigma}$ 、 \mathbf{w} を投資する割合のベクトル、 $\boldsymbol{\mu}$ を各銘柄資産の期待収益率を成分に持つベクトル、 ρ を必要とする収益率として、Markowitzモデルは以下のように表される。

$$\min_{\mathbf{w} \in \mathbb{R}^D} \mathbf{w}^\top \hat{\Sigma} \mathbf{w} \quad \text{s.t.} \quad \mathbf{w}^\top \mathbf{1}_D = 1, \mathbf{w}^\top \boldsymbol{\mu} = \rho$$

ここで、データから得られた期待値、分散を代入して、簡単な式変形を行なうことで、以下のような線形回帰の式に変形できる。

$$\min_{\mathbf{w} \in \mathbb{R}^D} \frac{1}{T} \|\rho - \mathbf{R}\mathbf{w}\|_2^2 \quad \text{s.t.} \quad \mathbf{w}^\top \mathbf{1}_D = 1, \mathbf{w}^\top \boldsymbol{\mu} = \rho \quad (1)$$

連絡先: 岡 滉, 大阪大学 産業科学研究所, 567-0047 大阪府 茨木市美穂ヶ丘 8-1, oka@ar.sanken.osaka-u.ac.jp

Brodie モデルは, (1) 式に L1 正則化を施したもので, τ を正則化パラメータとして以下のように表される.

$$\min_{\mathbf{w} \in \mathbb{R}^D} \|\rho - \mathbf{R}\mathbf{w}\|_2^2 + \tau \|\mathbf{w}\|_1 \quad \text{s.t.} \quad \mathbf{w}^\top \mathbf{1}_d = 1, \mathbf{w}^\top \boldsymbol{\mu} = \rho$$

L1 正則化を施すことによって, \mathbf{w} がよりスパースとなり, 売買に関連する銘柄を減らすことができることや, $\mathbf{w}^\top \mathbf{1}_d = 1$ という制約式とあわせることで, \mathbf{w} でマイナスとなる重みを減して, 空売りのポジションを減らせるなど, その実用性が報告されており, 本実験でもこのモデルを用いてポートフォリオ選択を行う [2].

2.2 密度比推定に基づく変化点検知

今回用いる変化点検知では, 変化を検知するために参照する期間の最初の時点 t_{rf} , 変化点かどうかを調べたい時点 t_{te} として, 時系列データ $\mathbf{r}_t(t = t_{rf}, \dots, T)$ に対して, $\mathbf{r}_t(t = t_{rf}, \dots, t_{te})$ と $\mathbf{r}_t(t = t_{te}+1, \dots, T)$ が一つの確率密度分布 $q_{rf}(\mathbf{r})$ から生成されているという仮説 H_0 と, それぞれが別の確率密度分布 $q_{rf}(\mathbf{r})$, $q_{te}(\mathbf{r})$ から生成されているという仮説 H_1 の尤度比から変化点の検知を行う. ここで簡単のため, 以下のように定義する.

$$\begin{aligned} \mathbf{r}_{rf}(i) &= \mathbf{r}_{t_{rf}+i-1}, & \mathbf{r}_{te}(i) &= \mathbf{r}_{t_{te}+i-1} \\ n_{rf} &= t_{te} - 1, & n_{te} &= T - t_{te} \end{aligned}$$

仮説 H_0 と H_1 の尤度比 Λ は, 以下のように表される.

$$\begin{aligned} \Lambda &= \frac{\prod_{i=1}^{n_{rf}} q_{rf}(\mathbf{r}_{rf}(i)) \prod_{i=1}^{n_{te}} q_{te}(\mathbf{r}_{te}(i))}{\prod_{i=1}^{n_{rf}} q_{rf}(\mathbf{r}_{rf}(i)) \prod_{i=1}^{n_{te}} q_{rf}(\mathbf{r}_{te}(i))} \\ &= \frac{\prod_{i=1}^{n_{te}} q_{te}(\mathbf{r}_{te}(i))}{\prod_{i=1}^{n_{te}} q_{rf}(\mathbf{r}_{te}(i))} \end{aligned} \quad (2)$$

この尤度比 Λ に対数をとった指標 S が設定した閾値を超えたか否かによって市場の変化を検知する.

$$S = \sum_{i=1}^{n_{te}} \log \frac{q_{te}(\mathbf{r}_{te}(i))}{q_{rf}(\mathbf{r}_{te}(i))} \quad (3)$$

それぞれの確率密度の推定は困難な問題とされており, それを回避するため, 本研究では指標 S を分母分子の確率分布を推定することなく直接確率密度比を推定する手法 KLIEP を用いる [4]. KLIEP では, 確率密度比をパラメータとガウスカーネルの線形和としてモデル化し, データからパラメータを推定することで確率密度比の推定を行う. 推定した確率密度比を (3) 式に代入することで, 尤度比 Λ を算出する. 算出した尤度比 Λ が設定する閾値 μ を超えた時, 変化点とする.

今回の変化点検知では, バッチ式の計算を必要とする KLIEP の応用であり, データが入力されるたびにパラメータを更新していくオンライン式で密度比推定を行う手法を用いる [5]. また, 過去のデータのクラスタリングはバッチ式の KLIEP を用いて行なう.

3. 提案手法

従来の手法では, 基本的に等期間の訓練または投資を仮定しており, 途中で株式市場のダイナミクスが変化したとき, 訓練を行いポートフォリオを作成しているときの分布と実際に投資をする際の分布が変わるため, 想定したリターンが得られないという問題を抱えている. そこで変化点検知を用いて, 分布

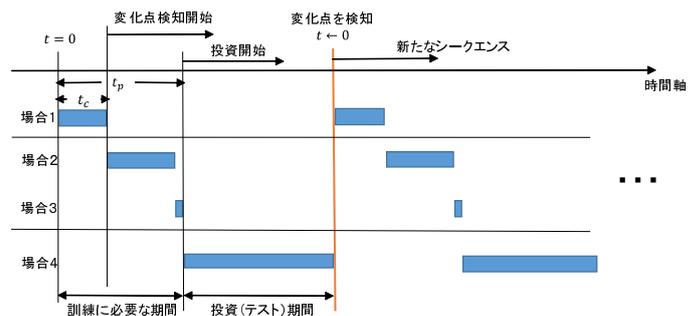


図 2: 変化点検知を用いた手法図

が大きく変動した際に投資を中断し, 訓練を再度やり直すことで前述の問題を緩和させることを提案する.

しかし, この手法では, 分布の変動を検知するたびに訓練を行うため, 全体として訓練期間が長くなり, 投資しない期間が増えるという問題がある. そこで, 過去のデータを予めクラスタリングしておき, 現在のデータと似たクラスタを訓練期間の一部として利用することで, 投資する期間を増やし, 最終利益を向上させることを提案する.

3.1 変化点検知を用いた手法

リターンのデータ $\mathbf{r}_t(t = 0 \dots T)$ を保持しているときに, 新たなリターンのデータ \mathbf{r}_{T+1} が与えられたとき, Brodie モデルの学習のための訓練期間を $1, \dots, t_p$, 変化点を検知するための参照期間を $1, \dots, t_c (\leq t_p)$, 各銘柄に投資する割合 (重み) を \mathbf{w} として以下のような処理を行う.

(場合 1). $T+1 < t_p$ かつ $T+1 \leq t_c$ のとき

$T+1 > t_p$ となるまで投資を行わず, $\mathbf{w} = \mathbf{0}$ とする. 変化点検知は参照期間が足りないため実施しない.

(場合 2). $T+1 < t_p$ かつ $T+1 > t_c$ のとき

投資は行わないが変化点検知は行い, 変化点を検知した場合はこれまでのデータを捨て, 次に与えられたデータを \mathbf{r}_1 として新たにポートフォリオの作成を始める.

(場合 3). $T+1 = t_p$ のとき

データ $\mathbf{r}_t(t = 1, 2 \dots T+1)$ を用いてポートフォリオの学習を行い, 投資する割合 \mathbf{w} を決める. 変化点を検知した場合は, 場合 2 と同様である.

(場合 4). $T+1 > t_p$ のとき

すでに学習してある重み \mathbf{w} を用いて投資を行う. このとき新たにポートフォリオ選択のための学習は行わない. 変化点を検知した場合は投資をそこでやめ, 場合 2 と同様新たにポートフォリオ作成を始める.

以上のような場合分けで, オンラインで変化点を検知しつつポートフォリオ学習を行い, 図 2 に示す流れのポートフォリオ訓練及び投資を実現する.

3.2 変化点検知と過去のクラスタリングを用いた手法

過去のリターンのデータ $\mathbf{r}'_t(t = 1 \dots T')$ があるとき, 以下のようにクラスタリングを行う. a は訓練期間の一部として, 過去のデータを利用する日数を決めるパラメータである. 本手法ではすべてのクラスタが日数 a を下回らないように分割を行った.

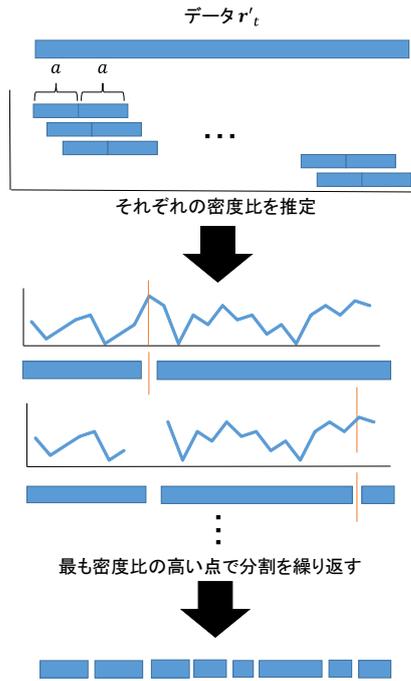


図 3: クラスタリング手法図

- (i) $t'(t' = a + 1 \dots T' - a)$ に対して, $r'_t(t = t' - a \dots t')$ を参照データ, $r'_t(t = t' + 1 \dots t' + a)$ をテストデータとして, それぞれの確率密度分布の密度比を KLIEP を用いて推定する.
- (ii) (i) において推定した密度比の中で最も高い点でデータを分割する.
- (iii) これまでに分割された点の前後 a の期間を除いて, もっとも密度比が高い点で新たにデータを分割する.
- (iv) (iii) を繰り返し, 全てのクラスタの長さが $2a$ 未満となったとき工程を終了する.

この操作を図 3 に示す. 上記の分割で最終的に得た区間をそれぞれクラスタとし, クラスタからはじめの a 期間を取り出して訓練期間の一部として用いる. 過去のクラスタそれぞれに番号 $n(n = 1, \dots, N)$ を振り, クラスタそれぞれが保有するデータを $r'_t[n](t = 1 \dots a, n = 1 \dots N)$, とする. 現時刻 T までの, リターンのデータ $r_t(t = 0 \dots T)$ に対し, 新たなリターンのデータ r_{T+1} が与えられた場合, Brodie モデルの学習のための訓練期間を t_p , 変化点を検知するための参照期間を $t_c(\leq t_p)$, 各銘柄に投資する割合を w として以下の処理を行う.

- (場合 1). $T + 1 + a < t_p$ かつ $t \leq t_c$ のとき
 $T + 1 + a > t_p$ となるまで投資を行わず, $w = 0$ とする.
 変化点検知は参照期間が足りないため実施しない.
- (場合 2). $T + 1 + a < t_p$ かつ $t > t_c$ のとき
 投資は行わないが変化点検知は行い, 変化点を検知した場合はこれまでのデータを捨て, 次に与えられたデータを r_1 として新たにポートフォリオの作成を始める.
- (場合 3). $T + 1 + a = t_p$ のとき
 データ $r_t(T = 1, 2 \dots t - a)$ とそれぞれのクラスタが保持

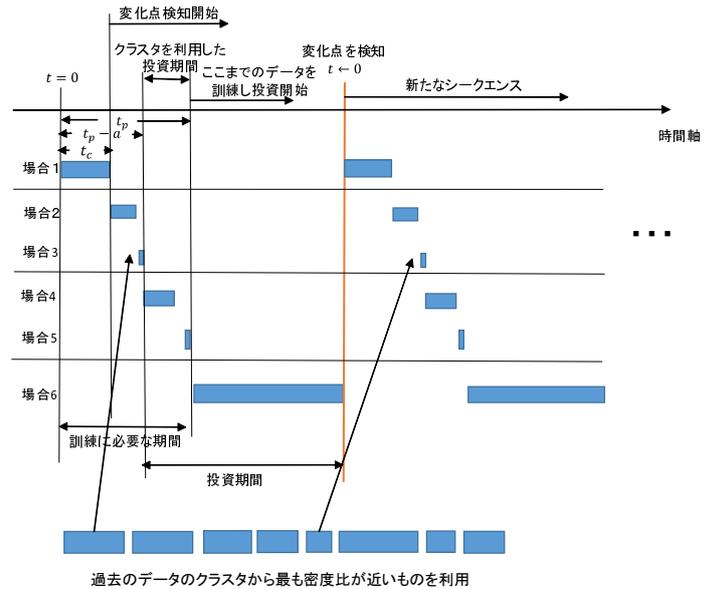


図 4: 変化点検知に加えクラスタを利用した手法図

するデータ $\{r'_t[n]|t = 1 \dots a\} (n = 1 \dots N)$ との間の密度比を KLIEP を用いて推定し, 最も 1 に近いクラスタの番号を n' として, $r_t(T = 1, 2 \dots t - a)$ 及び $\{r'_t[n']|t = 1, 2 \dots a\}$ を用いてポートフォリオの学習を行い, $T + 1 > t_p$ となるまでの間, 投資する割合 w' を決める変化点を検知した場合は場合 2 と同様である.

- (場合 4). $T + 1 + a > t_p$ かつ $T + 1 < t_p$ のとき
 新たに学習は行わず, クラスタを利用して学習した重み w' を用いて投資を行う. 変化点を検知した場合は場合 2 と同様である.
- (場合 5). $T + 1 = t_p$ のとき
 データ $r_t(t = 1, 2 \dots T + 1)$ を用いてポートフォリオの学習を行い, 投資する割合 w を決める. 変化点を検知した場合は場合 2 と同様である.
- (場合 6). $T + 1 > t_p$ のとき
 すでに学習してある重み w を用いて投資を行う. このとき新たにポートフォリオ選択のための学習は行わない. 変化点を検知した場合は投資をそこでやめ, 場合 2 と同様新たにポートフォリオ作成を始める.

以上のように過去のデータのクラスタを利用することで投資しない期間を減らすことが可能になる. この流れを図 4 に示す.

4. 検証実験

本稿では 3. で提案した手法の検証を行う. 2003~2013 年の TOPIX500 の銘柄を対象に, 変化点検知のみを用いた場合 (手法 1), 変化点検知と過去のクラスタリングの両方を用いた場合 (手法 2), 従来手法として, 幅 1 年の移動窓で半年ずつずらしながらはじめの半年を訓練データ, 残りの半年をテストデータとして投資を行った場合 (手法 3), これらの比較用として, 全銘柄に等しく投資を行った場合 (手法 4) の以上の 4 つの手

法の比較実験を行った。変化点検知に用いるパラメータは事前実験により、ガウスカネル幅である $\sigma = 30$, $n_{rf} = 45$ とし、変化点の検知のための閾値 μ は 0.5~4 の間で 0.25 ずつ変化させて実験を行った。ポートフォリオ学習における正則化パラメータは解がなくなる限界まで \mathbf{w} がスパースになるように設定し、設定するリターンは $\rho = 1$ 及び $\rho = 3$ とした。評価方法として、一般に広く用いられている Sharpe Ratio(SR) を用いる。SR は一般には次のように定義される。

$$SR = \frac{\text{平均収益率} - \text{リスクフリーレート}}{\text{収益率の標準偏差}} \quad (4)$$

リスクフリーレートは現在の日本では限りなく 0 に近いため今回は無視する。今回は投資した日のリターンの平均を標準偏差で割った値である日率 Sharpe Ratio(SR) と、投資していない日も考慮して、1 年毎のリターンの平均を標準偏差で割った値である年率 SR を用いた。ポートフォリオ訓練で設定するリターン ρ を 1 と 3 とし、変化点検知の閾値を 0.5 から 4 まで 0.25 刻みで変化させて横軸として、日率、年率における SR を縦軸にして描いたグラフが図 5~8 である。変化点検知の閾値が低いほど変化を敏感に検知し、変化点が多くなる。青が提案手法 1、赤が提案手法 2、緑が手法 3、黒が手法 4 である。手法 3、手法 4 では変化点検知を用いていないため、横軸に対して一定の値をとっている。図 5、図 6 から、変化点検知を用いることで用いないものに比べ高い日率 SR を得ることが分かる。また図 7、図 8 から、変化点検知及びクラスタリングを利用することで高い年率 SR を得ることが分かる。このことから市場の変動分布には変化点が存在し、それを捉えることが有用であると言える。

5. まとめ

本稿では変化点検知や過去データのクラスタリングなどを用いて、市場の変化を考慮してポートフォリオを選択する手法を提案した。また、TOPIX500 のデータを用いた検証実験から、市場には変化点が存在し、提案手法を用いることでより高い SR が得られることが確認された。

参考文献

- [1] H. Markowitz, *Portfolio Selection*, The Journal of Finance, **7**, 1952, 77-91.
- [2] J. Brodie, I. Daubechies, C. De Mol, and D.Giannone, *Sparse and stable Markowitz portfolios*, Proceedings of the National Academy of Sciences of the United States of America, **106**, 2009, 12267-12272.
- [3] V. Demiguel, L. Garlappi, F.J. Nogales, and R. Uppal, *A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms*, Management Science, **55**(5), 798-812, 2009.
- [4] M. Sugiyama, T. Suzuki, S. Nakajima, H. Kashima, P. von Bunau, and M. Kawanabe, *Direct importance estimation for covariance shift adaptation*, Ann Inst Stat Math, **60**(4), 2008, 699-746.
- [5] Y. Kawahara and M. Sugiyama, *Sequential Change-Point Detection Based on Direct Density-Ratio Esti-*

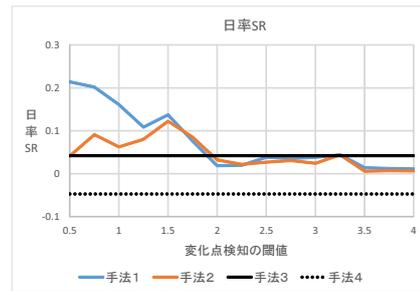


図 5: $\rho = 1$ の日率 SR

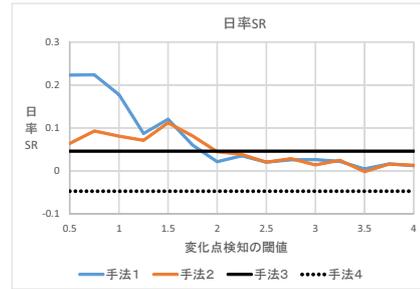


図 6: $\rho = 3$ の日率 SR

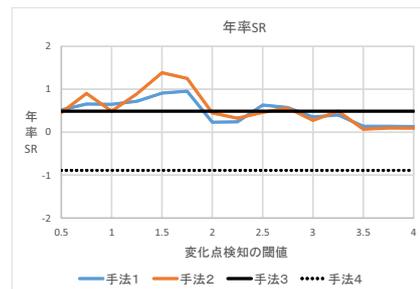


図 7: $\rho = 1$ の年率 SR

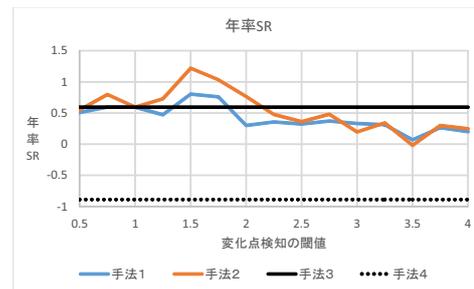


図 8: $\rho = 3$ の年率 SR

tion, Statistical Analysis and Data Mining, **5**(2), 2012, 114-127.