

資産価格変動パターンの類似性に着目した金融市場予測の評価

Evaluation of financial market forecasting with using similarity of asset price fluctuation patterns

今村 光良 *1*2 中川 慧 *1*3 吉田 健一 *3*
Mitsuyoshi Imamura Kei Nakagawa Kenichi Yoshida

*1日興グローバルラップ株式会社
Nikko Global Wrap Ltd.

*2筑波大学大学院 システム情報工学研究科
University of Tsukuba Graduate School of Systems and Information Engineering

*3筑波大学大学院 ビジネス科学研究科
University of Tsukuba Graduate School of Business Sciences

When predicting stock prices, we may refer to historical markets similarity. In the companion paper [7], we report relatively good results obtained by paying attention to the price fluctuation pattern of TOPIX. In this paper, we apply this method to other stock indexes and foreign exchange rates of the European and US markets. By analyzing the results, we found that the market cycle of target asset is an important factor to represent the price fluctuation pattern.

1. はじめに

株価を予測する際に、現在の相場と似た過去の相場を参考にすることがある。我々は [7] において、日本の株式指数 (TOPIX) について、過去の日次データから形成されるひと月あたりの価格変動パターンと当月を比較し、最も類似するものを予測に活用することで良好な結果が得られることを報告した。当該手法の特徴は価格変動パターンに注目し、かつ、導出の容易な手法で比較的良好な結果が得られるところにある。

本報では、始めに、当該手法を欧米市場の株式指数および為替レートへ適用し、良好な結果が得られる場合と上手くいかない場合を報告する。更に両方の場合を比較・分析する事により、価格変動パターンにおける予測に予測対象に合わせた周期性が重要な変数であることを明らかにする。最終的に、適切な周期を用いることで、欧米市場の株式指数および為替レートについても、良好な結果が得られる。

2. Indexation DTW(IDTW) 法

本報告においても、欧米株式市場指数および為替レートの価格変動パターン抽出に Indexation DTW(IDTW) 法 [7] を用いた。IDTW 法は現在の価格変動に似た過去の価格変動を抽出する際に、月初の価格を 1 と指数化したデータを対象に以下の方法で時系列データ間の類似度 (距離) を計測する。

まず $M \times N$ の大きさの類似度行列を作り、動的計画法により x, y 内の比較する点の組を示すワーピングパスを発見し、距離を計算する。ここで M, N は時系列データ x, y の長さを表す。具体的には以下の手順で計算を行う。

IDTW 距離算出アルゴリズム

1. $M \times N$ の大きさの類似度行列 D を作る
2. $D_{0,0} = 0$, それ以外の類似度行列の値に ∞ を代入し初期化する

連絡先: 今村 光良, 日興グローバルラップ株式会社, 〒103-0016 東京都中央区日本橋小網町 9-2

3. 以下の式により $i = 1, 2, \dots, M, j = 1, 2, \dots, N$ の全ての組み合わせの $D_{i,j}$ を計算する

$$D_{i,j} = d(x_i, y_j) + \min(D_{i,j-1}, D_{i-1,j}, D_{i-1,j-1})$$

d は距離測度であり、マンハッタン距離 $d(x_i, y_j) = |x_i - y_j|$ を使用する

4. 類似度行列 D の $D_{M-1, N-1}$ が時系列データ x, y の IDTW 距離となる

株価や為替レートなどの時系列データは、計測期間 (日次/週次/月次) を変更する事により 1 つの元データから複数の分析用データを作成する事ができる。我々の研究の前にも類似のアルゴリズムでの分析を試みた研究がある (例えば [2, 5, 6]) が、本研究と比して決定すべきパラメータが多い、結果が安定していない、等検討課題が多く残されている。

我々は [7] において、日本の株式指数 (TOPIX) について、過去の日次データから形成されるひと月あたりの価格変動パターンを利用する IDTW 法が有効である事を報告したが、後述するように IDTW 法を為替レートへ単純に適用したのでは、同様に有効な結果を得ることはできない。データの抽出には分析対象に合わせた周期性が重要である。本稿では、この周期性の影響について報告する。

3. 分析結果

3.1 検証方法

2. 章にて示した IDTW を欧米株式市場および為替レートに適用し、その有効性を確認した。分析にあたっては、欧米市場における株価指数として代表的に参照される S&P 500(SPXT), フランス CAC40 指数 (CACR), ドイツ株価指数 (DAX), FTSE100 種総合株価指数 (FTSE100) の 4 種類を対象に、為替レートは、円を基軸として対ドル (USDJPY), 対ユーロ (EURJPY), ドルを基軸とした、対ユーロ (EURUSD) の 3 種類を対象とした。各指数データおよび為替レートは情報端末の Bloomberg より取得した。データ期間は先行研究と同一条件

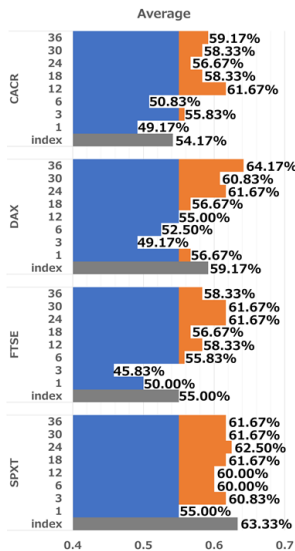


図 1: 欧米の株式指数に IDTW を用いた場合 (平均)

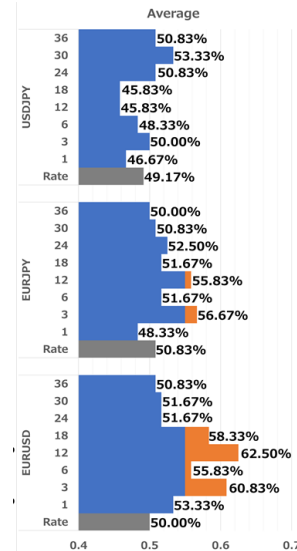


図 3: 為替レートに IDTW を用いた場合 (平均)

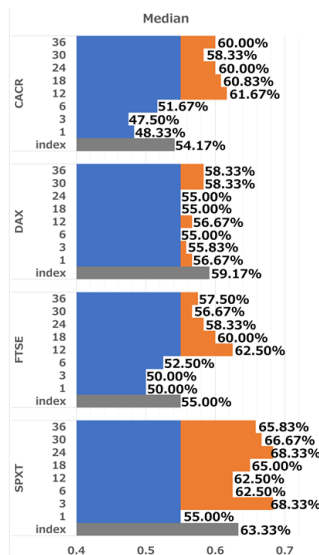


図 2: 欧米の株式指数に IDTW を用いた場合 (中央値)

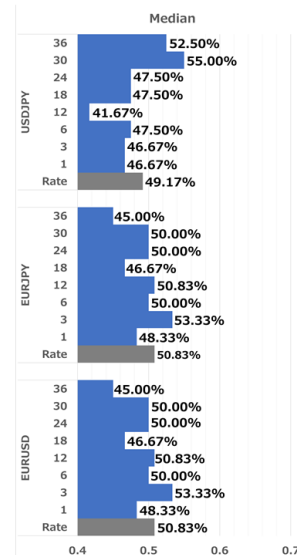


図 4: 為替レートに IDTW を用いた場合 (中央値)

とするため、1989年1月から2017年1月までとし、検証期間は2007年1月から2016年12月までの10年間とした。

分析にあたって検証期間と似た変動パターンを示した月の価格変動を、IDTWを用いて学習期間から探索する。その後、価格変動が似ていると判断された月の翌月の価格を予測値として利用する。これを1か月毎に繰り返し予測を行った。ただし学習期間の始点は1989年1月に固定とした。

3.2 分析 1：予備実験の結果

過去の価格変動パターンに近い月を1, 3, 6, 12, 24, 30, 36月選び、その平均および中央値を使った予測と、次月のリターンの符号が一致していれば正解とした結果を図1から4に示す。正答率のメモリが55%までは青色で表示し、55%以降はオレンジ色で表示している。そのため、オレンジ色の表示がないグラフの箇所は55%以下の正答率となる。また、参考用途に株式指数および為替レートそれぞれに常に買いと判断した場合の正答率を Index および Rate としてグレー色で合

わせて図示している。

図1における各株式指数については、平均値および中央値いずれについてもオレンジ色に図示している箇所が目立つ通り、いずれも55%を超える良好な結果が得られている。対して為替レートについては一部55%を超える箇所も見受けられるが、株式指数と比較した場合、算出方法に依存せず良好な結果が得られているようには見えない。

欧米株式市場の結果がTOPIXと同じく良好だったのに比較して、為替の結果が悪い事は特徴的である。我々は、この差の原因が対象に合わせた周期性にあると考え次節以降の実験を行った。

3.3 分析 2：為替レートの周期長による正答率変化

欧米株式市場の結果と為替の結果の正答率の差異については、時系列データの周期性が影響していることが考えられる。[7]においては、投資家は月単位で価格変動を認識している報告[1]から時系列データの類似性を評価するための比較用デー

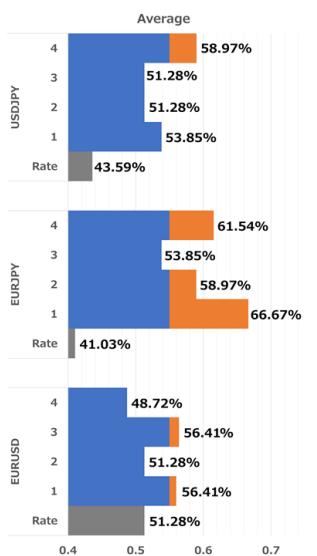


図 5: 為替レートの周期を四半期とした場合 (平均)

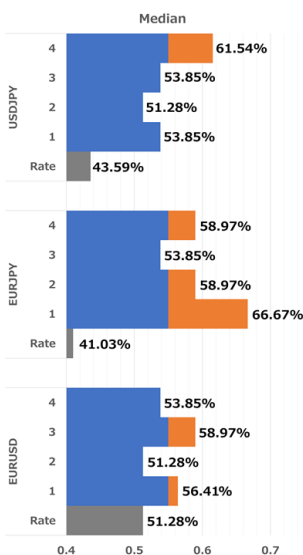


図 6: 為替レートの周期を四半期とした場合 (中央値)

タの切り出しをひと月単位で実施していた。これは株式相場における認識であり、為替相場については周期性に対する異なる認識が存在する可能性がある。

為替相場の認識としては需給による周期性が考えられる。これは、グローバルに展開している企業が海外利益を本国へ送金する際、本国通貨買いが発生することや、外貨資産の本国通貨建て評価確定目的で本国通貨買いが発生するといったことから生じる周期性である。企業が四半期決算を確定させることや、四半期毎に外国債券の償還や利払いが集中するといった、需給の観点から為替相場では、四半期毎の変動が重要であると推察できる。

上記内容を確認するため、まず、為替の周期を四半期毎と仮定し、ひと月の場合との予測精度の違いを調べた。為替レートについては、類似性を比較する周期をひと月から 1-3 月、4-6 月、7-9 月、10-12 月の四半期となるように変更し、再度正答率について確認した。結果は図 5、図 6 に示す通りである。

各為替レートにおいて平均・中央値計算に用いたサンプル数

	最大	最少
CACR	23	18
DAX	23	17
FTSE	23	18
SPXT	23	19

表 1: 各市場の月あたりの日次データ数 (日)

は、母数が三分の一に減少していることを考慮しデータ数を一致させるため、1,2,3,4 としている。図 1 および 2 同様に、正答率のメモリが 55 % までは青色で表示し、55 % 以降はオレンジ色で表示し、常に買いと判断した場合の正答率については Rate にて示している。

周期をひと月としていた場合と比較し、正答率が向上していることが確認できる。また、図 1 および 2 における株式指数と同様に平均値および中央値において同様の結果が得られていることも確認できる。

3.4 分析 3: 株式相場の周期開始日による正答率変化

次に、株式相場のひと月毎の周期性について調査した。現在、各市場のひと月あたり含まれる日次データ数は表 1 のとおりである。17 日から 23 日程度スライドさせると 1 周期程度スライドさせたこととなる。

株式相場に及ぼす「月末」など月単位の周期の影響を調べるため、[7] において価格変動パターンとして用いた月初の日次データから、パターンの開始日を n 日毎スライドさせていき、正答率の変化を調べた。なお、スライドはフレームの開始日と終端日を同時に動かし、スライド前の月に含まれる日次データの個数は同じにした。

図 7 に n 日スライド毎の正答率の変化について図示する。折れ線グラフの値は、図 1 および 2 において、特に正答率の高かったサンプル数 12,18,24,30,36 における n 日スライド毎の正答率の平均値を示している。点線は、日次をスライドしない場合の正答率の平均値となる。

図 7 に示す通り、いずれの株式指数についても、日数スライドによりひと月という周期が損なわれることに比例して正答率が低下していることが確認できる。また、1 周期に相当する 17 から 23 日に近づくと正答率がもとのスライドさせていない水準へ戻っていることが確認できる。

このことから、株式相場がひと月単位で相場を認識している可能性があり、また、当該手法のように価格変動パターンを予測時の特徴量とする場合、周期開始日が予測精度に影響があることが確認できる。

4. まとめ

本研究では、欧米市場の株式指数および為替レートを指数化した価格変動に対して DTW を適用する Indexation DTW (IDTW) を用いて分析し、各金融時系列データにおける価格変動パターンの類似性が将来価格の予測に対して有効か確認した。その結果、当該特徴に着目する場合、周期性を考慮した価格変動パターン抽出が重要であり、対象データの適切な周期を把握できるかが、精度向上への課題となることを確認した。今後の課題としては、手法の改善と前処理の改善が考えられる。

手法の改善については、より類似する価格変動パターンを抽出する方向を考慮する。株式指数の価格変動パターン比較時において、為替や原油といった別資産の価格変動パターンの類似性も考慮するなど、類似性を比較する次元を多次元にする手

法 (MD-DTW)[4] により, 空間軸方向に拡張を施すことが考えられる. 近年, 予測において良好な精度を発揮しているテキストマイニング [8, 9] 等の情報を活用し, 価格変動パターンの類似性とニュース情報等のテキストパターンの類似性を併用するといった拡張も考えられる.

また, 前処理については, 価格変動パターンの周期性を考慮する. 月次予測の場合はひと月が最適解か, 為替同様に四半期や半年, 一年などが適切か, 予測頻度を週次, 日次, へ拡張する場合はどうか考慮が必要である. 価格変動パターンの形成過程では, 月次予測の場合, 終値から成る日次データによる価格変動パターン以外にも, 始値を加えればより細かい価格変動パターンが形成可能である. 価格変動パターン形成時の粒度についても確認する必要がある.

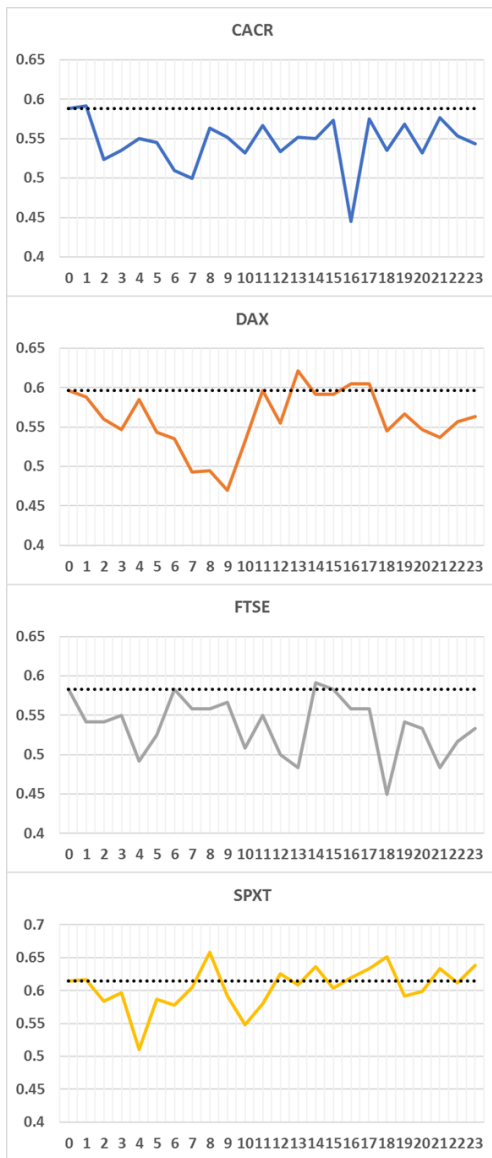


図 7: 日次データスライドによる正答率の変化

参考文献

[1] Bouman, S., and Jacobsen, B.: "The Halloween indicator," Sell in May and go away": Another puzzle. *The American Economic Review*, 92(5), 1618-1635. (2002).

[2] Coelho, M. S.: *Patterns in financial markets: Dynamic time warping*. PhD Thesis NSBE-UNL.(2012).

[3] Keogh, Eamonn J., and Michael J. Pazzani. "Derivative dynamic time warping." *Proceedings of the 2001 SIAM International Conference on Data Mining*. Society for Industrial and Applied Mathematics, 2001.

[4] ten Holt, Gineke A., Marcel JT Reinders, and E. A. Hendriks. "Multi-dimensional dynamic time warping for gesture recognition." *Thirteenth annual conference of the Advanced School for Computing and Imaging*. Vol. 300. 2007.

[5] 小沢育実, 関和広: 関連銘柄同定のための時系列データ類似度尺度の提案. *情報処理学会第 78 回全国大会*.(2016).

[6] 白浜公章: 経済データに対する値と形状に基づく時系列類似尺度の比較. *神戸大学紀要*.(2016).

[7] 中川慧, 今村光良, 吉田健一: 株価変動パターンの類似性を用いた株価予測. *第 31 回人工知能学会全国大会*.(2017).

[8] 藏本貴久, 和泉潔, 吉村忍, 石田智也, 中嶋啓浩, 松井藤五郎, 中川裕志.: 新聞記事のテキストマイニングによる長期市場動向の分析. *人工知能学会論文誌*, 28(3), pp.291-296.(2013).

[9] 五島圭一, 高橋大志, 寺野隆雄.: ニュースのテキスト情報から株価を予測する. *第 29 回人工知能学会全国大会*.(2015).