

時系列変化点の異種時系列への影響度分析 Effects of change points on different kind of time series

杉澤優馬[†]
Yuma Sugisawa

伏見卓恭[‡]
Takayasu Fushimi

斉藤和巳[†]
Kazumi Saito

1. はじめに

我々の生活を取り巻く時系列データとして、株価変動や為替レートなどさまざま存在する。これらの時系列データは、相互に影響しながら変動していることが知られている。一方の時系列が、他方の時系列にどれほど影響を与えるかを分析することは、経済動向の予測などのタスクにおいて重要な研究課題であり、研究が盛んにおこなわれている [2]。変動には、微細な変動から大きな変動まであるが、この変動幅は社会影響の大きさと考えられる。そこで、何らかの社会影響を受け、値が大きく変動した点を検出することで、系列データ同士の関係を調査する。

時系列データの変化点検出の既存研究として、ガウス分布でのモデリングの下で、最小記述長原理を土台として、逐次的に変化点を求める変化点検出法 [3] がある。この研究では、各銘柄ごとと業種ごとにまとめた株価変動の変化点検出結果を比較し、各銘柄個別の要因からなる変動なのか、業種全体を揺るがす社会影響を受けたのかについて分析されている。

本研究では、時系列データの変化点が他の時系列に及ぼす影響について分析する。時系列データ間の関係を分析するには、相関係数や時間のずれを考慮した交差相関係数などが考えられるが、単純な指標では、時系列データ全体で類似の変動をしていなければ、相関係数の値は大きくならず、そのときどきにおける変化の仕方の違いを考慮できない。すなわち、ある期間では正の相関をもち、異なる期間では負の相関をもつような場合では、全体として無相関に近い値となってしまう場合がある。

本研究では、異種時系列への影響度について文献 [3] の変化点検出法を土台に尤度を用いた指標を提案する。変化点検出法により、ある時系列の変化点を抽出し、その変化点を異種時系列に適用した際の尤度の大きさにより、両時系列間の関係性を評価する。そして、株価時系列データ（東証一部）と為替レート時系列データ（アメリカドル）を対象に、いくつかの指標と比較して各銘柄と為替の関係性を分析し、提案指標を評価する。

2. 提案手法

異種時系列への影響分析のための提案手法について説明する。N 個の時系列 $X = \{x_1, x_2, \dots, x_N\}$ を入力データとし、ある時系列 $x_i = (x_{i,1}, \dots, x_{i,T_i}) \in X$ と他の時系列 $x_j = (x_{j,1}, \dots, x_{j,T_j}) \in X$ との影響関係を分析する手法である。ここで、 T_i は時系列 x_i の時刻数を表し、本手法の特色の 1 つとして、 $T_i \neq T_j$ でも適用可能である。提案法の各ステップを以下に示す。

1. 任意の時系列 $x_i \in X$ の変化点を検出;
2. ステップ 1. で検出した変化点を時系列 $x_j \in X \setminus \{x_i\}$ へ適用し、尤度を計算;
3. 尤度で降順にソートし、影響関係の強い時系列を抽出;

時系列データから変化量の多い変化点を検出する方法、ならびに、求められた変化点を適用した際の尤度の計算法について説明する。

2.1. 変化点検出

簡単のため、ある時系列の時刻 t の値を x_t とし、時刻 1 から T までの時系列データは $x = (x_1, \dots, x_T)$ と表す。いま、変化点の個数が K で、それぞれの変化点を古い順に $F(1)$ から $F(K)$ に格納されるとする。また便宜上 $F(0) = 0$ かつ $F(K+1) = T$ と設定し、 $(K+2)$ -次元ベクトル $F = (F(0), \dots, F(K+1))$ を考える。すると、各変化点間を時系列データの平均で近似したときの自乗誤差 $E_K(F)$ は次式で計算することができる。

$$E_K(F) = \sum_{t=1}^T x_t^2 - \sum_{k=1}^{K+1} \frac{1}{\Delta F(k)} (y_{F(k)} - y_{F(k-1)})^2. \quad (1)$$

ここで、 $\Delta F(k) = F(k) - F(k-1)$ を表す。よって、この変化点検出問題は $E_K(F)$ を最小化する F を求める問題として定式化できる。その探索には、文献 [1] と同様に、逐次的に変化点を求める方法を採用する。そして、最小記述長 (MDL) 原理に基づき、最適な変化点数を求める。 i 番目の時系列 x_i の変化点を F_i と MDL による最適変化点数を K_i と表記する。

2.2. 尤度計算

j 番目の時系列に i 番目の時系列の変化点を適用した際の対数尤度 $\mathcal{L}_j(F_i)$ を以下のように計算する。

$$\mathcal{L}_j(F_i) = -\frac{T}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} E_{K_i}(x_j, F_i). \quad (2)$$

ここで、 $E_{K_i}(x_j, F_i)$ は j 番目の時系列に i 番目の時系列の変化点を適用した際の平均自乗誤差を意味する。

$$E_{K_i}(x_j, F_i) = \sum_{t=1}^T x_{jt}^2 - \sum_{K=1}^{K_i+1} \frac{1}{\Delta F_i(K)} (y_{F_i(K)} - y_{F_i(K-1)})^2. \quad (3)$$

3. 評価実験

実験では、東証 1 部 811 社の 2000 年 1 月 1 日から 2009 年 12 月 31 日まで (うち営業日は 2457 日) の終値と同期間 (うち営業日は 2601 日) の為替レートアメリカドル / 日本円の終値を用いる。すなわち、アメリカドルのレートの変動が各株式銘柄に与える影響について分析する。両時系列データの相関係数の絶対値ランキングによる分析結果と比較する。但し相関係数は、株価と為替データの両方に存在する日のみを用いて算出した。

3.1. 実験結果

相関係数の絶対値が大きい銘柄として、アイネス (通信業) やブラザー工業、スタンレー電気などの電気機器業が多く抽出された。これらの銘柄の変動と為替の

[†] 静岡県立大学経営情報学部

[‡] 静岡県立大学経営情報イノベーション研究科

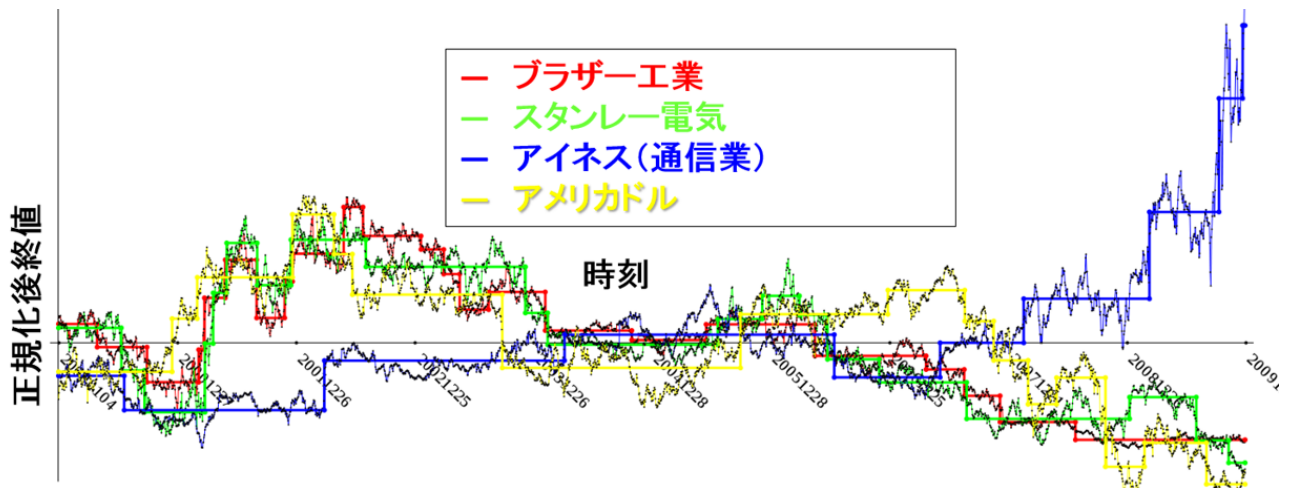


図 1: 相関係数の絶対値によるランキング上位の銘柄

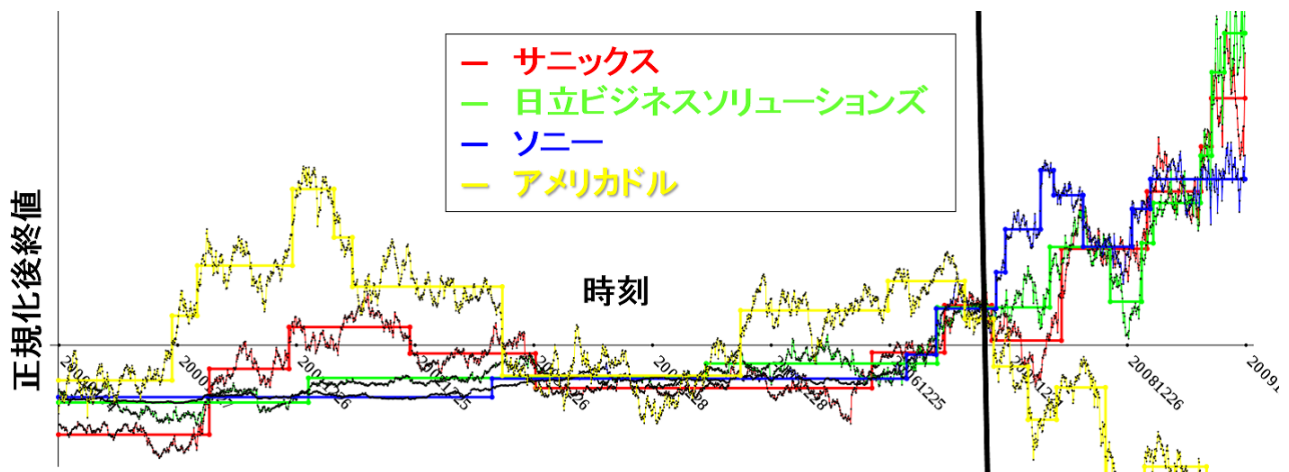


図 2: 提案法による尤度ランキング上位で相関係数の絶対値が低い銘柄

変動を実際にプロットする(図 1)。電気機器業は、海外での売り上げ比率が高いため為替の影響を受けやすいことから為替の変動と高い正の相関関係にある。一方、アイネスなどの国内向けの通信業は円高になると有利になり、円安になると比較的不利になるため、強い負の相関関係にある。さらに順位を下ると、電気機器業、通信業、輸送用機器業など、一般的に知られている、円高や円安に影響を受けやすい業種が上位にランキングされている。実際これらの株は為替敏感株と呼ばれており、為替の影響を受けやすい業種である。

提案手法による尤度が大きい銘柄として、上記の相関係数の絶対値が大きい銘柄も抽出されるが、その他にサニックス、日立ビジネスソリューション、ソニーなども抽出された。これらの銘柄の変動と為替の変動を実際にプロットする。図 2 を見ると、図中黒線の 2007 年中旬までは、正の相関があるといえるが、それ以降は負の相関関係にあるといえる。すなわち、短期的には類似の変動をしているが、長期的にみると類似の変動と逆位相の変動を含んでいるため、単純な相関係数では上位にランキングされないと考えられる。

提案法による尤度のランキングが高いということは、為替変動の変化点が当てはまりの良い銘柄ということになる。変動の仕方がその時々によって異なっていたり、変化点の時刻が多少ずれていたとしても、尤度は比較的高くなると考えられる。また、相関係数のラン

キングの上位にはなかった為替敏感株にあたる企業が尤度のランキング上位では見られるため、相関係数を用いた結果よりも広く、関係の強い銘柄が抽出されたことが示唆された。

4. おわりに

本研究では、株価と為替の影響関係を尤度を用いて抽出した結果、相関係数よりも比較的広く関連する銘柄を抽出できることが示唆された。今後は株価と為替以外の時系列データでも検証を行っていきたい。

謝辞 本研究は科学研究費補助金 (No.23500128) の補助を受けた。

参考文献

- [1] K. Saito, et al., "Burst Detection in a Sequence of Tweets based on Information Diffusion Model," Proc. of DS2012, pp.239–253, 2012.
- [2] R.N. Mantegna, "Hierarchical structure in financial markets," Euro. Phys. J. B 11, pp.193–197, 1999.
- [3] 杉澤優馬, 他, "変化点検出法に基づく業種別社会的影響分析," DEIM2013.