

# 移動エントロピーを用いた 銘柄間ネットワークダイナミクスの分析

Analysis of the dynamics of influence networks among stock indexes by using  
Transfer Entropy

小村 和輝<sup>1</sup> 鳥海 不二夫<sup>1</sup> 大橋 弘忠<sup>1</sup>

Kazuki Komura<sup>1</sup>, Fujio Toriumi<sup>1</sup>, and Hirotsada Ohashi<sup>1</sup>

<sup>1</sup> 東京大学工学系研究科

<sup>1</sup>School of Engineering, The University of Tokyo

**Abstract:** Recent research has explored the proper method to analyse the relationships in financial markets for risk management. In this paper, we applied and expanded the method of transfer entropy to construct a stationary network which represents the information propagation between stocks. This network can differ significantly from other static networks, such as correlations network and minimal spanning tree network, because it can include the direction information. We demonstrate that this method reveals meaningful hidden relations of cause and effect between stocks, and show that it is more useful than other methods, such as correlation coefficient network and partial correlation coefficient network.

## 1. はじめに

近年、証券市場における個人投資家の比重が増大しているが、個人投資家と機関投資家には未だ大きな情報格差がある[1]。したがって個人投資家に対して投資判断の支援をおこなう技術の必要性が高まっており、特に突発的な出来事が及ぼす自身の保有銘柄に対する影響の把握が重要となる。そのような影響度を定量化した銘柄間のネットワークを適切に構築することができれば、ある銘柄の価格やボラティリティの変化が個人投資家の保有する他の銘柄にどう波及するのかを事前に知ること役立つ可能性がある。

本研究では、そのような銘柄間の影響度ネットワークを構築する上で、移動エントロピーを拡張した手法が有用であることを示す。さらには個別銘柄間の情報の流れを可視化し、定性的には予測し難い銘柄間の関係性を示すことで、個人投資家のリスク管理に有益な情報を提供することを目的とする。

## 2. 既存研究

銘柄間の関係性を表すネットワーク構築には、こ

れまで様々な手法が用いられてきた。Plerou らはランダム行列理論を用い、ニューヨーク証券取引所の銘柄間についての相互相関行列を分析した[2]。Eom らは最小全域木を用いて構成した銘柄間ネットワークのトポロジー性質から、その関係性を分析した[3]。Kullman らは各銘柄の価格データを用い、時間差をつけて相関係数を計算することで、どの銘柄がどの銘柄に価格の影響を与えているのかの分析を試みた[4]。上山らは3大証券取引所（ニューヨーク市場、ロンドン市場、東京市場）の株価指数を事前データとして学習させたベイジアンネットワークを用いて、次日の株価の動きを予測した[5]。三浦らは単一の予測モデルにより株価を一点で予測するのではなく、複数の予測モデルの学習を行い、各モデルに適切な重みを付けることで予測値の分散を低減する手法を提案した[6]。しかしこれらの手法は因果関係を考慮していないことや定常的な性質に着目していることから、突発的な出来事に対する分析には不向きであり、投資家にとっては有用でない。

移動エントロピーを用いると、上記の手法とは異なり、有向性の定義されていない時系列データから情報伝播の向き、つまりは因果関係を動的に分析できる。また、確率をもとにした計算方法であるため、状態推移の確率変化を見ることも可能となる。

Schreiber は時系列相互相関と比較し、移動エントロピーが因果関係を見る上で優れていることを示した[7]。Greg らは、Twitter の単純な時系列データから

---

連絡先：小村和輝，東京大学 工学系研究科 システム創成学専攻，〒113-8654 東京都文京区本郷7-3-1, komura@crimson.q.t.u-tokyo.ac.jp

ノード間の情報伝播を分析し、移動エントロピーの有用性を示した[8]. Marschinski らは本手法を用い、ダウ平均株価と DAX 指数の間の情報伝播を分析し、ダウ平均株価から DAX 指数への価格情報の流れがあることを示した[9].

Kwon らは対象とするインデックスの種類を 25 に増やし、情報伝播のネットワークを構築しその性質を分析した[10]. さらに Kwon らは、個別銘柄を業界ごとにグループ化したデータを用いることでインデックスから個別銘柄への情報伝播があることを示した[11]. また、鈴木らは、相互相関係数、偏相関係数、移動エントロピーの 3 つの指標を用い、日経株式市場に隠されたスモールワールドネットワークの構造を発見するとともに、移動エントロピーを用いた手法が最も有用であることを示した[12].

ところがこれらの研究はすべて個別銘柄ではなくインデックスや銘柄グループ間の情報伝播を対象としており、且つ定常的なネットワークの分析にとどまっているため、突発的な変化が与える市場に対する影響に関心を持つ投資家に役立つものではない。

### 3. 本研究で用いる手法

#### 3.1 拡張移動エントロピー

移動エントロピーは、離散変数で表された 2 つの定常過程間の情報伝播を定量化するために Schreiber[7]により導入された。ここで、2 つの離散的な定常過程  $I$  と  $J$  を考える。過程  $I, J$  からそれぞれ  $k$  個、 $l$  個のサンプルを用いて計算する場合、過程  $J$  から過程  $I$  の向きへの移動エントロピー  $T_{J \rightarrow I}$  は以下のようになる。

$$T_{J \rightarrow I} = \sum p(i_n, i_{n-1}^{(n-k)}, j_{n-1}^{(n-l)}) \log \frac{p(i_n | i_{n-1}^{(n-k)}, j_{n-1}^{(n-l)})}{p(i_n | i_{n-1}^{(n-k)})} \quad (1)$$

ただし、 $i_n$  と  $j_n$  はそれぞれ過程  $I$  と  $J$  の時間における離散状態を表す、 $i_n^{(n-k)}$  は過程  $I$  の時間  $n$  における状態から  $k$  個分遡った状態までを含めたベクトルであり、 $i_n^{(n-k)} = (i_n, i_{n-1}, \dots, i_{n-k+1})$  と表される。

$j_n^{(l)}$  も同様に、 $j_n^{(n-l)} = (j_n, j_{n-1}, \dots, j_{n-l+1})$  と表される。

$T_{J \rightarrow I}$  は過程  $J$  の情報が、もう一方の過程  $I$  の遷移確率にどれだけ影響を与えるかをあらわす。つまりは  $j_n$  の情報が与えられた際に、過程  $I$  の状態遷移の不確かさをどれだけ減少させるかを定量化している。

よって、以下の式に書き直すことができる。

$$T_{J \rightarrow I} = H\{I_n | I_{n-1}^{(n-k)}\} - H\{I_n | I_{n-1}^{(n-k)}, J_{n-1}^{(n-l)}\} \quad (2)$$

$$H(A|B) = -\sum_{A, B} P(A, B) \log P(A|B) \quad (3)$$

また、この式から分かるように、移動エントロピーは 2 つの離散変数  $i_n$  と  $j_n$  に関して非対称になっており、これにより 2 変数間の情報伝播の方向性がわかる。

しかし、この手法では状態遷移の方向性を区別できず、また遡る状態数に応じて状態の組み合わせが著しく増大してしまう。したがって、今回用いた手法では、 $i_n^{(n-k)} = (i_n, i_{n-k+1}), j_n^{(n-l)} = (j_n, j_{n-l+1})$  と

し、確率計算においても、状態遷移が同一の状態間 ( $ex+1 \rightarrow +1$ ) で起こる場合と、異なる状態間 ( $ex-1 \rightarrow +1$ ) で起こる場合をそれぞれ区別した。前者を正の移動エントロピー、後者を負の移動エントロピーとし、前者から後者を差し引いた値を拡張移動エントロピーまたは拡張 TE とする。

#### 3.2 使用データ

本研究では、2011 年 1 月 1 日から 2011 年 6 月 30 日で、東京証券取引所の取引時間である平日の 9:00~11:00 と 12:30~15:00 の 1 秒ごとの価格情報から得られるリターン情報を用いる。

使用する銘柄は、データの存在する銘柄のうち日経 225 構成銘柄に含まれるもののみを選出している。これは、流動性が小さく注文情報が疎な他の銘柄は、データ量の少なから十分正確に確率が計算できないためである[8].

### 4. 実証分析

#### 4.1 ネットワーク構築

上述したデータをもとに、拡張移動エントロピーを計算し有向重み付きネットワークを構築した。データは 1 秒ごとに離散化し、リターンが正ならば 1、負ならば -1、0 ならば 0 とした。

1 つの拡張移動エントロピーの計算に用いるデータ数は一日の総取引時間を秒数で表した数である 16200 となる。また遡る変数の数を 1 から 10000 まで 1 つずつ変化させ、リターン情報の伝播の期間が 1 秒から 10000 秒までのネットワークをそれぞれ構築した。リターンを計算するティックサイズは 2000

とした。このティックサイズは大きすぎると状態遷移が前の状態に大きく依存し正の移動エントロピーが大きくなってしまい、小さすぎると状態があまり変化しなくなり確率計算に膨大なデータ数が必要になってしまう。今回はいくつかのティックサイズで計算を実行し、後述する手法と比較した際に最も近く、情報伝播の実態を表していると考えられるものを採用した。

また、0, 1, -1 それぞれを等確率で発生させたランダムな離散変数との移動エントロピーをそれぞれ計算し、エッジを張る際の閾値とした。今回は定常的なネットワークを見るために、全期間の拡張移動エントロピーをそれぞれ平均したものをを用いた。

本手法の比較対象としてリターンの相互相関を用いたネットワークと偏相関を用いたネットワークも構築した[12]。交差相関の手法を用い、こちらも上記と同様に、伝播時間が1秒から10000秒までのネットワークをそれぞれ構築した。

## 4.2 結果・考察

先に述べたデータ数の少なさによる確率計算上の問題の影響を小さくするため、構築したネットワークのうち、特に流動性の高い銘柄からの拡張移動エントロピーに着目し情報伝播の様子を分析する。今回は、TOPIX Core30に含まれる銘柄を起点とするエッジのうち、大きさが上位2つのものに着目した。

まず比較対象について考える。銘柄間の相互交差相関の値を図1に表すが、全てほぼ一様な推移をしており、情報伝播の様子は見られない。これは全ての銘柄が自己相関に強く影響を受けるためであると考えられる。したがって本研究では、自己相関の影響を除いた偏相関を比較対象として用いる。

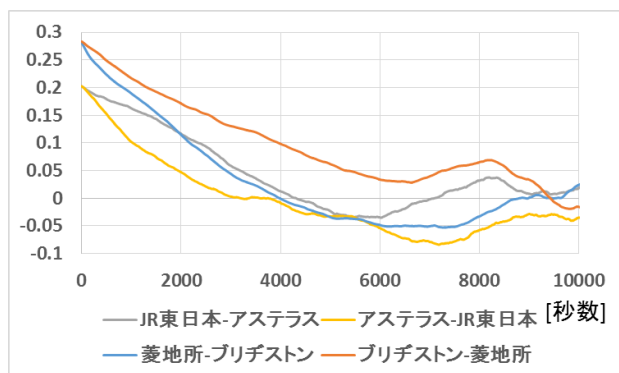


図1：銘柄間の相互交差相関の値

図2に三菱地所-ブリヂストン間の情報伝播の様子を示す。ただし、拡張TEの値は偏相関係数の値域

で正規化しており、逆向き拡張TEとはブリヂストンから三菱地所への向きの拡張TEを示している。これによると、偏相関では一定の情報伝播が継続して観測されるのに対し、拡張TEでは情報伝播を示す値がより細かく変化している。これは、拡張TEが1秒ごとの状態遷移確率を微細に計算しているのに対し、偏相関は一定区間の変数を用い、値が平準化されているためであると考えられる。

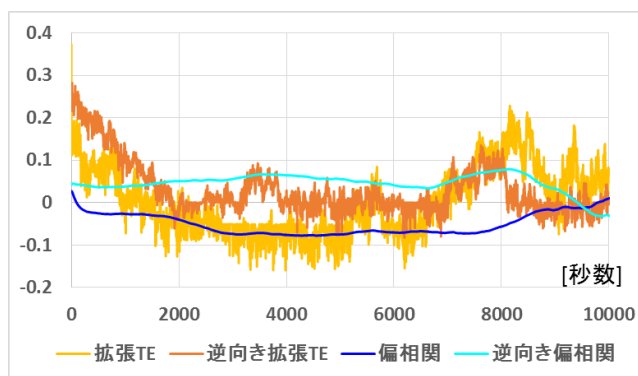


図2：三菱地所-ブリヂストン間の情報伝播

図3に、上記と同様にJR東日本 - アステラス間の情報伝播の様子を示す。図2, 図3共に情報伝播の時間が6000秒を超えた付近から偏相関と拡張TEの値に大きな乖離が見られる。これは、1日毎に計算を行っているために、交差させる変数が大きくなるほど相対的に計算に用いるデータ数が少なくなってしまうことによると考えられる。

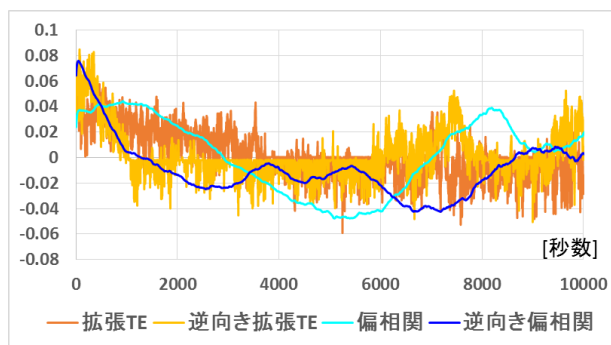


図3：JR東日本 - アステラス間の情報伝播

しかし図2, 図3共に全体としては拡張TEは偏相関と近い値を示しており、本手法が従来手法と同程度には有用であることが分かる。さらに偏相関よりも刻み幅の小さい情報伝播を捉えることができおり、従来は単に正の相関、或いは負の相関と考えられていた区間であっても、細かく見れば正負の確率の大小が数秒ごとに入れ替わっていることが分かる。移動エントロピーのこの性質を利用すれば、従来手法では個人投資家からは見落とされていた数秒単位

でのマージンを得る売買を支援する投資指標を表すことが可能となる。

## 5. おわりに

本研究では、移動エントロピーの定義を拡張し、株式市場における各個別銘柄のリターンの情報から個別銘柄間の定常的な影響度ネットワークを構築した。また拡張移動エントロピーの値と偏相関係数の値を比較することで、銘柄間の情報伝播を捉える上での拡張移動エントロピーの応用可能性を示した。

具体的には、従来手法よりも刻み幅の小さい数秒単位での細かな情報伝播を捉えることができ、ある銘柄がある銘柄に対して従来は常に単なる正の相関を持っていたと考えられていた区間であっても、細かく見れば正負の確率の大小が数秒ごとに入れ替わっていることが分かった。この性質を利用すれば、数秒単位でのマージンを得る投資戦略をとる個人投資家を支援したり、或いはある銘柄の価格が大きく変化した際に、その影響が他の銘柄にどのタイミングでどれほど伝播するのかを数秒単位で予測し知らせること等に役立つ可能性がある。

今後の課題としては、以下の三点が挙げられる。

一つ目に、情報伝播をより正確に表せているかを他の手法と定量的に比較する方法を確立しなければならない。今回は正規化したグラフを定性的に比較することで各手法の相違点を検証したが、提案手法の有用性を示すにはより客観的な指標が必要となる。

二つ目に、リターンの計算をする上で、最適なティックサイズを検証しなければならない。データ数と状態間の依存関係のトレードオフのもと、最適値を決定する手法を確立し、データ数の異なる各銘柄間の組み合わせごとに設定する必要がある。

三つ目に、最終的に個人投資家に対し、どのような情報提供をするか検討しなければならない。今回の研究で、提案手法を用いれば従来よりも小さい時間幅での情報伝播を見ることができたが、具体的にどのような指標に還元し、どういった活用法とともに個人投資家へ提供するかを考える必要がある。またネットワークのマクロな分析も行い、同様に活用できるか検証する必要がある。

## 謝辞

本研究の一部は、科学研究費補助金(40206454)の助成を受けたものである。

## 参考文献

[1] 秦, 劫: 情報の非対称性, 能力の非対称性, と個人投

資家の変, 大阪大学経済学, 57(4) P. 164-P. 176, (2008)

- [2] V, Plerou, P, Gopikrishnan, B, Rosenow, L, A, N, Amaral and H, E, Stanley : Universal and Nonuniversal Properties of Cross Correlations in Financial Time Series, *Physical Review Letters*, 83, 7, pp. 1471-1474, (1999)
- [3] C, Eom, G, Oh, W, -S, Jung, H, Jeong and S, Kim. : Topological properties of stock networks based on minimal spanning tree and random matrix theory in financial time series, *Physica A*, 388, pp. 900-906, (2009)
- [4] L, Kullmann, J, Kertész and K, Kaski. : Time-dependent cross-correlations between different stock returns: A directed network of influence, *Physical Review E*, 66, 026125, (2002)
- [5] 上山薫, 左毅, 上島康孝, 北栄輔, Application of Bayesian Network to Stock Index, SIG-FIN-004-07, (2010)
- [6] 三浦 和起, 日野 英逸, 村田 昇, クロスエントロピー最適化を用いた株価予測値の安定化手法, 情報処理学会研究報告, MPS, 数理モデル化と問題解決研究報告 2010-MPS-81(9), 1-6, (2010)
- [7] T, Schreiber. : Measuring Information Transfer, *Physical Review Letters*, 85, 2, pp. 461-464, (2000)
- [8] Greg Ver Steeg, Aram Galstyan, Information Transfer in Social Media, WWW 2012 - Session: Information Diffusion in Social Networks, (2012)
- [9] R, Marschinski and H, Kantz. : Analysing the information flow between financial time series; An improved estimator for transfer entropy, *The European Physical Journal B*, 30, pp. 275-281, (2002)
- [10] O, Kwon and J, -S, Yang. : Information flow between stock indices, *Europhysics Letters*, 82, 68003, (2008)
- [11] O, Kwon and J, -S, Yang. : Information flow between composite stock index and individual stocks, *Physica A*, 387, pp. 2851-2856, (2008)
- [12] 鈴木智也, 池口徹, 堀尾喜彦, 日経株式市場に隠された複雑ネットワーク構造の推定, THE INSTITUTE OF ELECTRONICS, INFORMATION AND COMMUNICATION, IEICE Technical Report,, NLP2005-122, (2006)
- [13] 鳥海不二夫, 西岡 寛兼, 梅岡 利光, 石井健一郎: 板情報による市場相違性の検出 人工知能学会論文誌 Vol. 27 No. 3 P 143-150(03/2012)
- [14] O, Kwon and J, -S. Yang, Information flow between stock indices, *Europhysics Letters*, 82, 6, 68003, (2008)