

移動エントロピーを用いた 銘柄間の影響度ネットワークの分析

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

小村 和輝¹ 鳥海 不二夫¹ 大橋 弘忠¹

Kazuki Komura¹, Fujio Toriumi¹, and Hirotada Ohashi¹

¹ 東京大学工学系研究科

¹School 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 apply 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.

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用いる手法が有用であることを示す。延いては、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、Shiらはアジア通貨危機の前後で中国とアメリカの市場間の情報伝播の仕方が変化したことを示した[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

2.1 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

る離散状態を表す、 $i_n^{(k)}$ 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、 $i_n^{(k)} = (i_n, i_{n-1}, \dots, i_{n-k+1})$ と表される. $j_n^{(l)}$ も

同様に、 $j_n^{(l)} = (j_n, j_{n-1}, \dots, j_{n-k+1})$ 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.

3.2 使用データ

単純な価格変化等に比べ、約定前の注文情報が含まれる板情報には投資家心理がよりよく反映される[14]. したがって本研究では、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

表1: 使用銘柄一覧

日経225先物	双日	昭電工	塩野義	東海力	オークマ	シャープ	トヨタ	りそなHD
日水	キッコマン	住友化	中外薬	TOTO	コマツ	ソニー	三菱自	三井住友トラ
国際石開帝石	味の素	日産化	エーザイ	ガイシ	住友重	TDK	マツダ	三井住友FG
コムシスHD	ニチレイ	日曹達	テルモ	新日鉄住金	クボタ	ミツミ	ホンダ	横浜銀
大成建	JT	東ソー	第一三共	神戸鋼	荏原	アルプス	スズキ	みずほFG
大林組	東洋紡	電化	ヤフー	JFE	千代建	パイオニア	キャノン	大和
清水建	ユニチカ	信越化	トレンド	日製鋼	ダイキン	横河電	シチズンHD	野村
鹿島	日清紡HD	協和キリン	富士フィルム	三井金	日精工	アドテスト	伊藤忠	T&D
ハウス	日東紡	三井化学	ユニカミノル	東邦鉛	NTN	デンソー	丸紅	菱地所
積ハウス	セブン&アイ	三菱ケミHD	資生堂	三菱マ	ジェイテクト	カシオ	三井物	
日揮	帝人	宇部興	昭和シェル	住友鉱	ミネベア	ファナック	東エレクト	
日清粉G	東レ	日化薬	浜ゴム	DOWA	日立	京セラ	住友商	
日ハム	クラレ	電通	ブリヂストン	古河機金	東芝	三菱重	三菱商	
サッポロHD	旭化成	花王	旭硝子	古河電	三菱電	川重	丸井G	
アサヒ	王子HD	武田	板硝子	住友電	NEC	IHI	イオン	
キリンHD	北越紀州紙	アステラス	住友大阪	フジクラ	富士通	日産自	新生銀	
宝HLD	日本製紙グループ本社	大日本住友	太平洋セメ	洋缶HD	パナソニック	いすゞ	三菱UFJ	

ある平日の 9:00~11:00 と 12:30~15:00 の 1 秒ごとの注文情報を用いる。使用する 145 銘柄の一覧を表 1 に示す。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、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、買い注文または売り注文が入っていれば 1、入っていなければ 0 とした。1 つの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。また先行研究を参考に遡る変数の数は 1、つまり式 (1) で $k=l=1$ とした[7][8]。また、0, 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、全期間の移動エントロピーをそれぞれ平均したものをを用いた。構築したネットワークの例を図 1 に示す

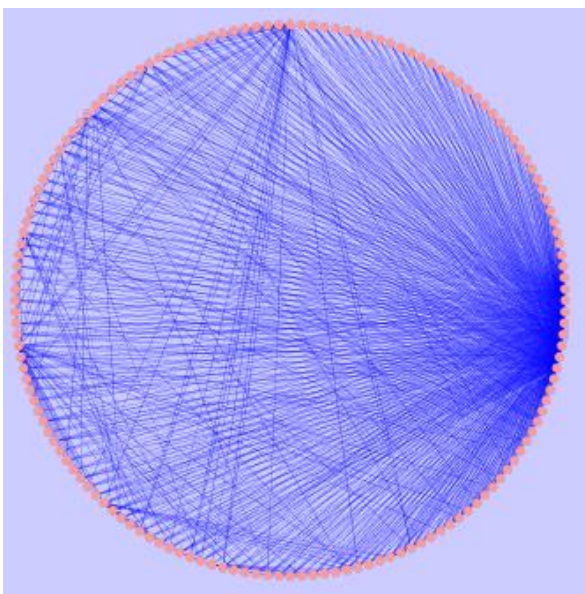


図 1: 移動エントロピーによる影響度ネットワーク

4.2 結果・考察

各個別銘柄間の移動エントロピーのうち、大きさ上位 10 組を表 2 に、下位 10 組を表 3 に示す。この結果から、移動エントロピーの低い組み合わせには流動性の低い銘柄及び高い銘柄いずれも含まれるのに対し、移動エントロピーの高い組み合わせには流動性の低い銘柄しか含まれていないことが分かる。これは、先に述べた注文数の少なさによる確率計算上の問題であると考えられる。そこで、特に流動性の高い銘柄同士の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移動エントロピーのみを用いた。そのうち上位 3 つの銘柄について、最も大きな値、最も小さな値それぞれを表 4 に示す。

表 2: 移動エントロピーの大きさ上位 10 組

	エッジの向き	移動エントロピー
1	東海カ→信越化	0.005303783
2	浜ゴム→日産化	0.005074696
3	東海カ→京セラ	0.005063648
4	浜ゴム→信越化	0.005057146
5	東海カ→コニカミノル	0.004973537
6	東海カ→日産化	0.004948509
7	日化薬→信越化	0.004933296
8	浜ゴム→クラレ	0.004913129
9	東海カ→ハウス	0.004841444
10	東海カ→テルモ	0.004835905

表 3: 移動エントロピーの大きさ下位 10 組

	エッジの向き	移動エントロピー
1	三井住友FG→日化薬	0.000110865
2	三井住友FG→味の素	0.000118652
3	三井住友FG→大和	0.000122768
4	ソフトバンク→古河電	0.000127951
5	三井住友FG→ホンダ	0.000128808
6	三井住友FG→三菱重	0.000146948
7	三菱商→北越紀州紙	0.000152063
8	ソフトバンク→双日	0.000152469
9	ソフトバンク→住友大阪	0.000152469
10	トヨタ→マツダ	0.000153401

表 4: 流動性の高い銘柄からの移動エントロピー

エッジの向き	移動エントロピー
JR東日本→アステラス	0.002749892
JR東日本→ユニチカ	0.000473615
信越化→伊藤忠	0.001721551
信越化→日東紡	0.000308674
菱地所→ブリヂストン	0.001660022
菱地所→古河機金	0.000449489

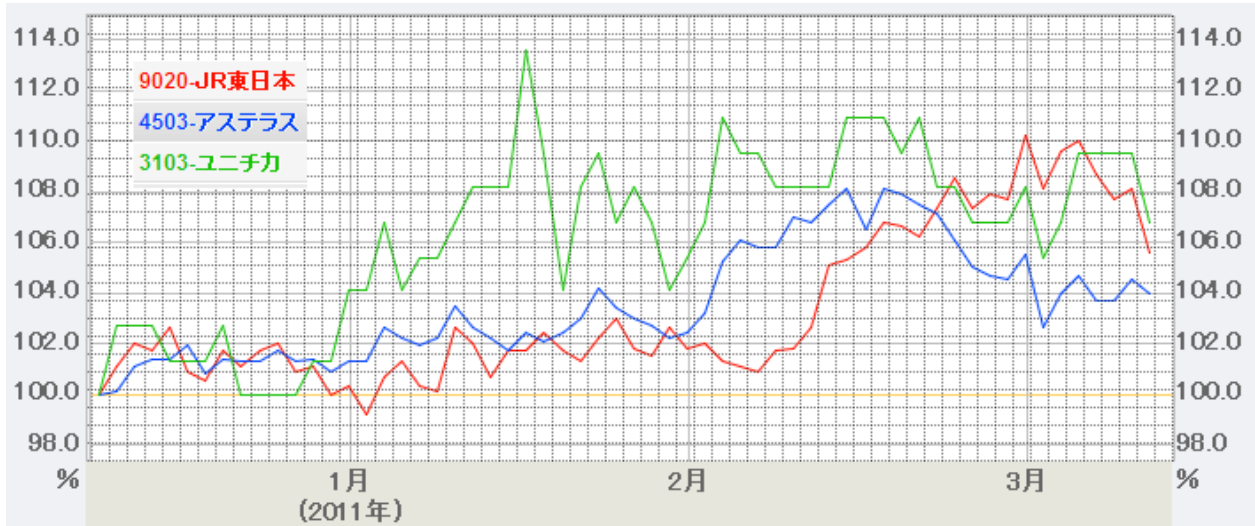


図2：JR 東海・アステラス・ユニチカの価格変化

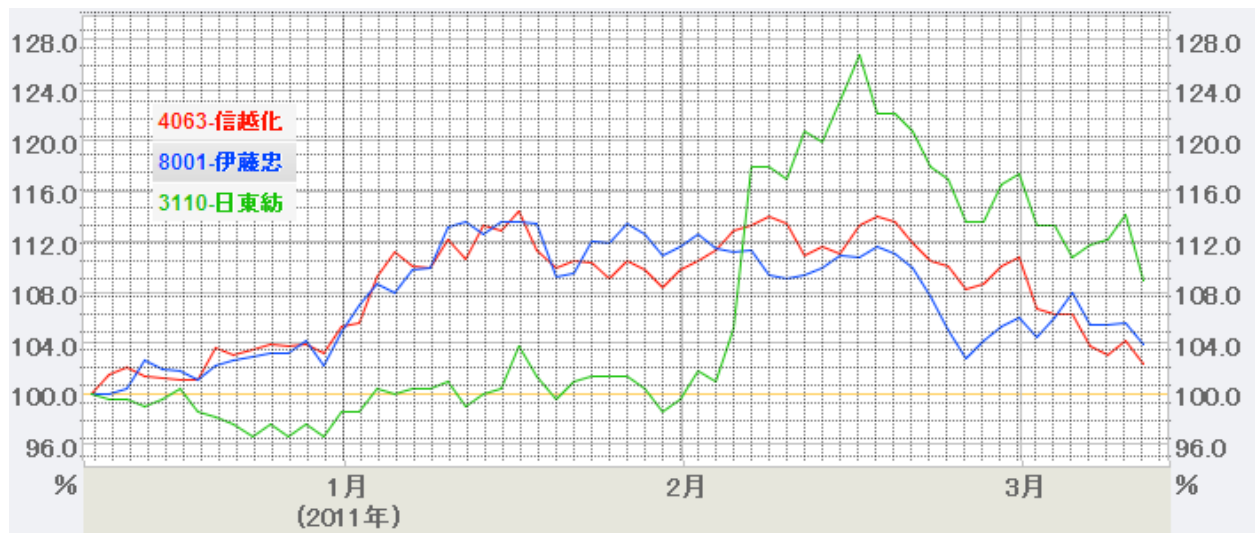


図3： 信越化・伊藤忠・日東紡の価格変化

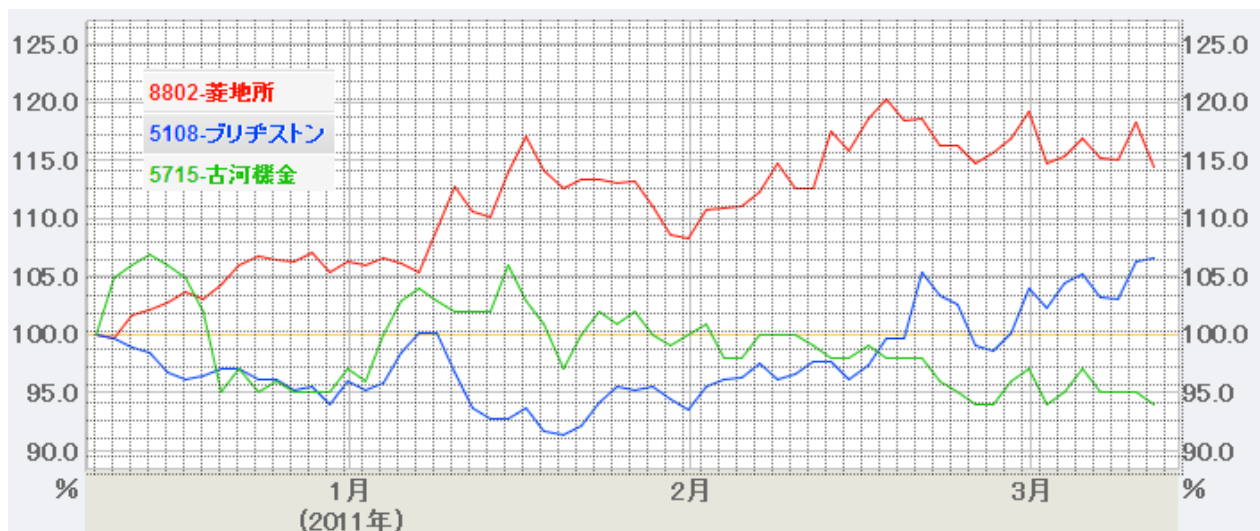


図4： 菱地所・ブリヂストン・古河機金の価格変化

この値が情報の伝播を表していることを確かめるため、各組み合わせについて、価格変化の様子を図2、図3、図4に示した。ただし、震災後は銘柄間の関係が定常状態でなくなるため、期間は震災直前までの3ヶ月間である2010年12月10日~2011年3月10日とした。また、価格変化は期間初日を100%とした割合で表した。

図1では、およそJR東日本が主導する形でアステラスの株価が推移しているのに対し、ユニチカにはJR東日本の価格変化に対して特定の傾向は見られない。図3も図2と同様に、伊藤忠に対しては信越化が主導する値動きが見られるが、日東紡は信越化とほぼ無関係に価格が変化しているように見える。図4では、菱地所が主導する形で、ブリヂストンにそれとは逆の値動きが起こっているが、古河機金には菱地所との関係性はあまり見られない。これらを定量的に示すために、各元ノードの銘柄の値動きを1日~10日間過去方向へ平行移動し、それぞれ前日からの価格変動の割合(例:1000円→1100円ならば0.1)を変数に相関係数を計算した。そのうち、絶対値が最大となった場合を表5に示す。この結果から、先に述べた移動エントロピーの高い先ノードの銘柄の方が元ノードの銘柄主導で値動きが起こる傾向があることが分かる。

また、逆の値動きの場合にも移動エントロピーが高い値を示した理由は、売り注文と買い注文を区別していなかった点にあると考えられる。売り注文の後に売り注文、買い注文の後に買い注文ではなく、売り注文の後に買い注文、買い注文の後に売り注文が行われる確率が高い場合でも移動エントロピーは高い値を示す。さらに、単に注文のタイミングの情報のみを用いて移動エントロピーを計算しているにも関わらず、高い値を示す先ノードの銘柄の価格変化は元ノードの銘柄と同じ動きまたは逆の動きで一貫していた。このことから、銘柄間を伝播する情報について、向きと大きさだけでなく、その正負に関する定常的な傾向を見ることができたといえる。

表5: 最大値をとる時の平行移動日数と相関係数

	平行移動した日数	相関係数
JR東海-アステラス	4	0.352085085
JR東海-ユニチカ	8	0.121122289
信越化-伊藤忠	1	0.332531042
信越化-日東紡	4	-0.26383753
菱地所-ブリヂストン	2	-0.37802548
菱地所-古河機金	2	-0.225190644

5. おわりに

本研究では、株式市場における各個別銘柄の注文のタイミングのみから移動エントロピーを計算し、個別銘柄間の定常的な影響度ネットワークを構築した。また移動エントロピーの値と価格変化の関係を見ることで、移動エントロピーが銘柄間の情報伝播を捉える上で有用であることを示した。伝播する情報については、向きと大きさだけでなく、その正負の性質を見ることができた。

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どうかを他の手法と比較すべきであると考える。単純な時系列データから有向性を検証できる点や任意の区切り幅で分析できる点で本手法は他の手法よりも優れているといえるが、データの連続性を考慮しているという特長も検証する必要がある。

二つ目に、銘柄ごとの最適な閾値を設定する手法を検討すべきであると考える。先に述べた疎なデータによる確率計算などの要因から、銘柄ごとに移動エントロピーの分布が異なるため、それらを個別に考慮する必要がある。

三つ目に、最終的に個人投資家にどのような形で情報提供をするか検討すべきであると考える。今回の研究で、定常的なネットワークが価格の情報伝播の様子を表すことは示されたが、移動エントロピーの値の大きさと情報伝播の大きさの関係性等は今後定量的に検証する必要がある。最終的に個人投資家に向けてどのようなリコメンデーションを施すのかを明確にしなければならない。

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)
- [1 0] O, Kwon and J, -S, Yang. : Information flow between stock indices, *Europhysics Letters*, 82, 68003, (2008)
- [1 1] O, Kwon and J, -S, Yang. : Information flow between composite stock index and individual stocks, *Physica A*, 387, pp, 2851-2856, (2008)
- [1 2] W, Shi and P, Shang. : Cross-sample entropy statistic as a measure of synchronism and cross-correlation of stock markets, *Nonlinear Dynamics*, 71, pp, 539-554, (2013)
- [1 3] 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)
- [1 4] O, Kwon and J, -S. Yang, Information flow between stock indices, *Europhysics Letters*, 82, 6, 68003, (2008)