



株式市場に埋め込まれたグループ相関構造 日米比較

吉川 丈夫*1 · 家富 洋*2

Group Correlation Structures Embedded in Financial Markets – Comparative Study Between Japan and U.S.

Takeo YOSHIKAWA*1 and Hiroshi IYETOMI*2

Abstract– We study group correlation structures of financial markets in Japan and U.S. from a network-theoretic point of view. The correlation matrix of stock price changes, purified by the random matrix theory, is regarded as an adjacency matrix for a network. The weighted links in the networks thus constructed can have negative sign corresponding to anticorrelation between stocks. To identify groups in such a network, we search for the optimum decomposition of nodes which maximizes the total sum of weights of links inside groups. We then find that the network of Tokyo Stock Exchange is decomposed into four groups. The stock prices comove almost perfectly inside the groups and move oppositely across the groups. Also we apply the same analysis to the S&P 500 stocks. The U.S. stock market shows frustrated behavior similar to that embedded in the Japanese market.

Keywords– principal component analysis, financial market, RMT: Random Matrix Theory, group correlation, network, frustration

1. はじめに

時間方向に T 個の観測点がある N 変数の時系列で記述された複雑系を考える。各時刻における系の状態は N 次元空間における 1 点で記述され、系の相関構造は N 次元空間における T 個の状態点の分布形状に反映される。そのような多変量相関を計量する有力な手法の 1 つとして主成分分析がある。具体的には、相関係数からなる行列（以後、相関行列と呼ぶ）の固有値問題を解けばよい。得られる固有値を大きいものから順に並べると、第 1 固有値が分散の最大値を与え、付随する固有ベクトルが第 1 主成分軸を決め、最も重要な相関構造を表す。第 2 固有値とその固有ベクトルが、第 1 主成分に対する直交補空間で最も重要な相関情報（分散値、主成分軸）を提供する。以下同様である。このような主成分分析は、力学で例えば、質点系に対する慣性主軸の決定と等価である¹。

しかし、すべての主成分が統計的に有意というわけではない。まったくランダムな時系列データでも、ただ 1 度の有限の観測では擬似的な相関がどうしても含まれてしまう。主成分分析において、いったいどれほどの大きさの固有値まで統計的に意味あるものとして選択するかが大問題である。それに解答を与えるのがランダム行列理論（RMT）[1] である。ランダム時系列の相関行列は、特別な無限次元極限（データ数 N と時間方向の観測点数 T との比 $Q = T/N$ を固定）で、普遍的な固有値分布や固有ベクトル成分分布をもつことが知られ、応用上有用な解析的表式が導かれている。もちろん、時間方向の観測点数を増やせば増やすほど、固有値の広がり方は 1 に収束する。もし RMT によって予測される最大固有値より大きい固有値が存在すれば、それは真の相関の現れと判断できる。

1999 年に経済物理学にとって記念碑的論文 2 編が同時に出版された [2, 3]。それらの論文は、米国株式市場における株価相関を明らかにするために、S&P 500 構成銘柄がつくる相関行列の固有値分布と RMT の解析式とを比較し、市場全体の“重心運動”（市場モード）を含めて複数の統計的に有意な主成分があることを示した。これらの論文を契機に多変量時系列に対する主成分分析と RMT の複合解析手法が確立され [4–6]、多くの研究

*1 新潟大学大学院自然科学研究科 新潟市西区五十嵐二の町 8050 番地

*2 新潟大学理学部 新潟市西区五十嵐二の町 8050 番地

*1 Graduate School of Science and Technology, Niigata University, 8050, Ikarashi 2-no-cho, Nishi-ku, Niigata

*2 Faculty of Science, Niigata University, 8050, Ikarashi 2-no-cho, Nishi-ku, Niigata

Received: 28 July 2013, 6 September 2013

1. 相関行列が慣性テンソル、固有値が主慣性モーメント、固有ベクトル（主成分軸）が慣性主軸にそれぞれ対応する。

者が株価相関の研究に RMT を応用した論文を発表している。これまでに我が国や米国などの成熟した株式市場において市場全体の振る舞いの背後に株式銘柄グループ（同じ値動きをする銘柄集団）が存在すること [4, 7, 8]、比べてインドのような新興市場ではこのようなグループ構造があまり目立たないことが明らかにされた [9]。さらに、その解析手法の応用範囲は金融分野に留まらず、気象科学 [10]、情報通信 [11]、脳科学 [12] などへ拡大している。

株式市場における銘柄のグループ相関（多粒子系における重心から観測した各粒子の運動間の相関に対応）を解析するにあたっては、これまで銘柄同士の相関関係のみに焦点が当てられてきた。しかし、銘柄同士は相関関係ばかりではなく反相関関係にある場合もある。我々は、銘柄間の正負の相関関係をネットワーク的な視点から同等に扱い [13–15]、株式市場における隠れた相関構造を明らかにすることを目指している。すでに東京市場については先行して研究を進めた [16]。本稿では、米国市場についても同様の解析を行い、日米市場における相関構造について比較した結果について紹介する。

2. 相関行列

ここでの具体的な解析対象は、東証一部上場企業（TSE）の 33 業種 557 銘柄、2707 営業日（1996 年 1 月 4 日から 2006 年 12 月 29 日まで）の株価日次データ、ならびに S&P 500 の 24 業種 483 銘柄、1009 営業日（2008 年 1 月 2 日から 2011 年 12 月 30 日まで）の株価日次データである²。

銘柄 n の時刻 t における価格を $S_{n,t}$ ($n = 1, 2, \dots, N; t = 1, 2, \dots, T + 1$) とする。株価時系列 $\{S_{n,t}\}$ を定常化するため、対数収益率 $R_{n,t} \equiv \ln S_{n,t+1} - \ln S_{n,t}$ を計算し、さらに次のように標準化（平均 0、分散 1）する：

$$G_{n,t} \equiv \frac{R_{n,t} - \langle R_n \rangle}{\sigma_n} \quad (1)$$

ここで $\langle \dots \rangle$ は時系列 $\{R_{n,t}\}$ の時間平均、 σ_n はその標準偏差である：

$$\langle R_n \rangle \equiv \frac{1}{T} \sum_{t=1}^T R_{n,t}, \quad (2)$$

$$\sigma_n \equiv \frac{1}{T} \sum_{t=1}^T (R_{n,t} - \langle R_n \rangle)^2. \quad (3)$$

このような $G_{n,t}$ を (n, t) 要素とする $N \times T$ 行列を G とすると (G の各行は各銘柄の標準化した対数収益率の時系

列である)、相関行列 C は、

$$C = \frac{1}{T} G G^T \quad (4)$$

と与えられる。ここで G^T は行列 G の転置を表す。

3. RMT を用いた主成分の判定

相関行列は、その固有値と対応する固有ベクトルを用いて次のように書くことができる（スペクトル分解）：

$$C = \sum_{\ell=1}^N \lambda_{\ell} u_{\ell} u_{\ell}^T \quad (5)$$

ここでは $\lambda_1 > \lambda_2 > \dots > \lambda_N$ とする。式 (5) からわかるように、より大きな固有値と付随する固有ベクトル（主成分）に相関に関するより重要な情報が含まれている。では、どの固有ベクトルまでが統計的に意味のある情報を持っているのか。従来、対応する固有値が 1 以上のもの、あるいは累積寄与率が 80% を超えるまで、などという主成分の判定基準が用いられていたが、実際にはそれぞれの分析においては“適切”に判断する必要があった。

ここで、各行列要素が平均 0、分散 1 の独立同分布の確率変数からなる $N \times T$ 行列 H を用い、 $\frac{1}{T} H H^T$ で計算される行列を考える。これは H の各行をひとつの時系列とみなすと、 $\frac{1}{T} H H^T$ は N 個のランダムな時系列の相関行列（ランダム相関行列）と考えられる。RMT によれば、このランダム相関行列において、 $Q = T/N$ を一定のまま $N, T \rightarrow \infty$ の極限をとると、固有値分布、固有ベクトル成分分布、最近接固有値間隔分布などに対して様々な普遍的な性質が現れる。例えば、ランダム相関行列の固有値分布はそのような極限操作下で次の解析式に従う [17, 18]：

$$\rho(\lambda) = \frac{Q}{2\pi} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda}. \quad (6)$$

ここで $\lambda_{\pm} = (1 \pm \sqrt{1/Q})^2$ であり、固有値は区間 $[\lambda_-, \lambda_+]$ の中に分布することがわかる。ランダム相関行列に関するこれらの性質と実際の相関行列がもつ性質を比較することにより、統計的に有意な相関（主成分）を曖昧性無く抽出することができる。

東京市場および S&P 500 市場の固有値分布と対応するランダム相関行列（実際のデータ行列と同じアスペクト比 $Q = T/N$ をもつ）の固有値分布を Fig. 1 および Fig. 2 でそれぞれ比較する。ここで、実際の固有値分布のうち、式 (6) の分布から外れる固有値、特に λ_+ より大きな固有値に対応する固有ベクトルに有意な相関が含まれていると考えるのが RMT による主成分の判定規準である。東京市場においては、 $\lambda_{13} (\cong 2.17) > \lambda_+ (\cong 2.11) > \lambda_{14} (\cong 2.07)$ であるので 13 番目までを主成分として採用し、同様に S&P 500 市場

2. ここでの「業種」とは東京市場については証券コード協議会（SICC）による 33 業種（中分類）であり、S&P 500 については Global Industry Classification Standard (GICS) による 24 の産業グループを指す。それぞれに一つ上の分類があり、東京市場では 10 の大分類、GICS では 10 のセクターとなっている。

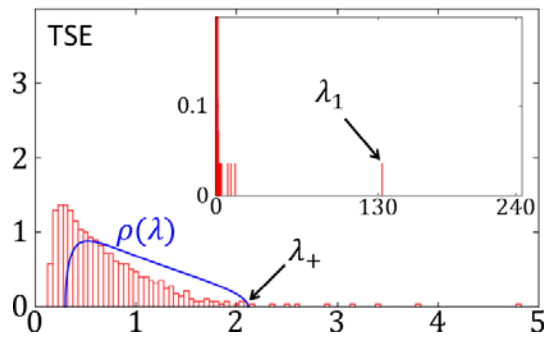


Fig. 1: Probability density distribution of the eigenvalues of the correlation matrix in the Tokyo Stock Exchange and the corresponding distribution $\rho(\lambda)$, Eq. (6), predicted by the RMT.

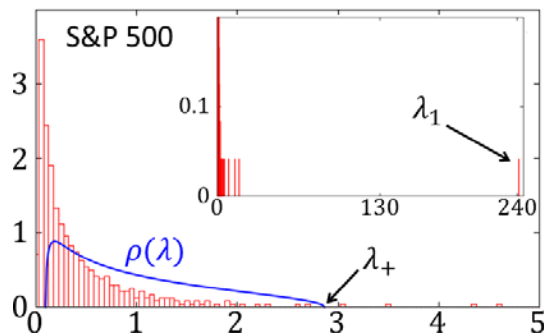


Fig. 2: Same as Fig. 1, but in the S&P 500.

においては、 $\lambda_9 (\cong 3.07) > \lambda_+ (\cong 2.86) > \lambda_{10} (\cong 2.71)$ であるので 9 番目までを採用する。

Fig. 3 および Fig. 4 は東京市場および S&P 500 市場それぞれの固有ベクトル u_ℓ ($\lambda_\ell > \lambda_+$ に対応) における業種の偏極を表したものである (u_1 は除く)。業種の並びは、東京市場については SICC の、S&P 500 市場については GICS の並びに従った。業種 S の偏極率は固有ベクトル成分の和と絶対値の和の割合を業種ごとに計算する:

$$P_S \equiv \frac{\sum_{j \in S} u_{\ell,j}}{\sum_{j \in S} |u_{\ell,j}|}. \quad (7)$$

後述するように第一固有ベクトル成分はすべて正なので全業種に対して偏極率は 1 である。両市場とも多くの業種で高いまとまりを示しており、これは先行研究の結果 [4, 7, 9] と一致する。東京市場と比べて S&P 500 市場で顕著なのが並び順の近い (すなわち上位の分類で同じカテゴリに分類される) 業種間の連動である。東京市場でも u_2, u_6, u_7 などにはその傾向はみられるが、S&P 500 市場ではすべての固有ベクトルにその傾向があり、これは S&P 500 市場の構成銘柄は、GICS の産業グループの上位の分類、セクターとしても連動していることを示している。

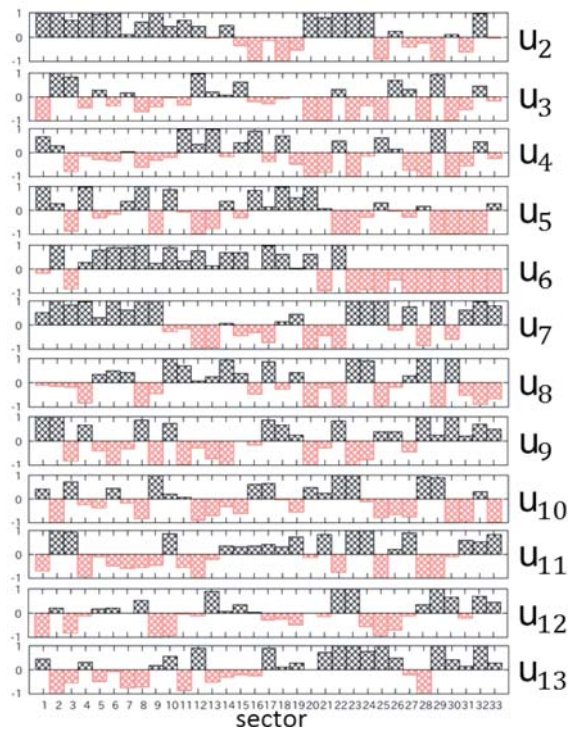


Fig. 3: Polarization of each sector defined by Eq. (7) in the statistically meaningful eigenvectors for the Tokyo Stock Exchange.

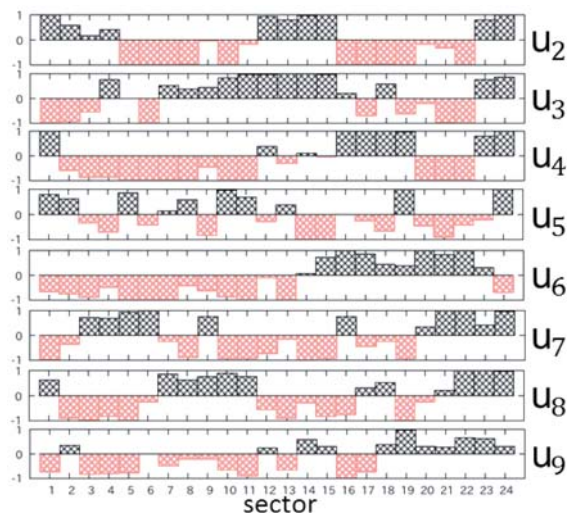


Fig. 4: Same as Fig. 3, but for the S&P 500.

4. 株価相関ネットワーク

ここで市場の相関構造を理解するために、ネットワークの視点を導入する。相関行列をネットワークの隣接行列と読み替えることで、銘柄をノード、銘柄間の相関係数をリンクとする株価相関ネットワークを構築することができる。株価の対数リターンから計算される相関行列

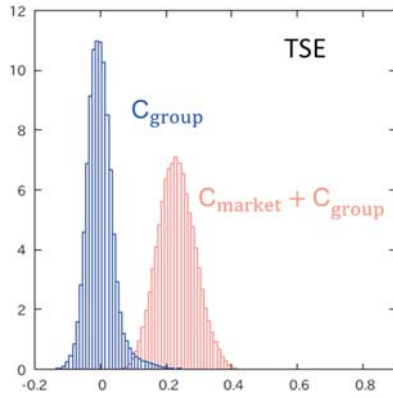


Fig. 5: Distribution of the matrix elements of C_{group} and $C_{\text{market}} + C_{\text{group}}$ in the Tokyo Stock Exchange.

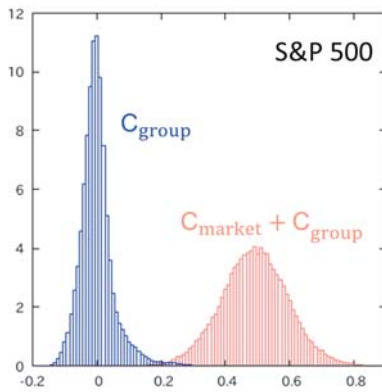


Fig. 6: Same as Fig. 5, but in the S&P 500.

C を次のように分解する [3, 4, 8] .

$$C \equiv C_{\text{market}} + C_{\text{group}} + C_{\text{random}} \quad (8)$$

$$= \lambda_1 u_1 u_1^T + \sum_{\ell=2}^{N_c} \lambda_{\ell} u_{\ell} u_{\ell}^T + \sum_{\ell=N_c+1}^N \lambda_{\ell} u_{\ell} u_{\ell}^T. \quad (9)$$

ここで N は全銘柄数, N_c は固有値が RMT の最大固有値を上回る成分数 (TSE については $N_c = 13$, S&P 500 市場については $N_c = 9$) である. C_{random} については 3 節でみたように, その寄与を積極的にノイズと区別する理由はない. 第 1 固有値 λ_1 は第 2 固有値に比べ非常に大きく (TSE については $\lambda_1 \cong 133$, $\lambda_2 \cong 15$, S&P 500 については $\lambda_1 \cong 241$, $\lambda_2 \cong 17$), しかも対応する固有ベクトル u_1 の成分はすべて符号が同じであるため, C_{market} は市場全体の動きを表す. これは, Fig. 5 および Fig. 6 に示す東京市場および S&P 500 市場における C_{group} および $C_{\text{market}} + C_{\text{group}}$ の行列要素の分布からもわかる. どちらの市場においても C_{group} は正負両方の領域に分布しているが, $C_{\text{market}} + C_{\text{group}}$ は正の領域のみに分布している. 以上から, C_{group} が市場の特徴的な相関構造を表すと考えられるので, これを隣接行列とする株価相関ネッ

トワークを構築する. さらに, 各主成分の相関構造への寄与を調べるため,

$$C_{\text{group}}^{(m)} = \sum_{\ell=2}^m \lambda_{\ell} u_{\ell} u_{\ell}^T \quad (m = 2, \dots, N_c) \quad (10)$$

を隣接行列とするネットワークも合わせて考える ($m = N_c$ が C_{group} を与える).

5. ネットワークとコミュニティ抽出

一般にネットワークは, ノード間のリンクの重みをその要素とする行列 (リンクがない場合の重みは 0), いわゆる隣接行列を使って記述することができる. 最も単純なネットワークは, リンクの重みを 0 か 1 の 2 値で表すものである. 例えば, 人をノードとし, 人と人を友人関係の有無にリンクの有無を対応させれば友人関係ネットワークを作ることができる. このようなネットワークでは, リンクの分布はランダムでもなく, また一様でもない. ネットワーク中のリンクの密な部分は, ノード同士の関係の強い部分である. 友人関係ネットワークでは, 仲の良い集団であり, 企業の取引ネットワークでは, 取引が活発な集団である. ネットワーク中のリンクの密な部分をコミュニティと呼び, その情報はネットワーク構造を理解する上で重要である [19].

ノード i が属するコミュニティを σ_i とし, すべてのノードの属するコミュニティを順に並べた $\{\sigma\} = \{\sigma_1, \sigma_2, \dots, \sigma_N\}$ をネットワークの分割と呼ぶ. コミュニティを決定することは, 分割 $\{\sigma\}$ を決めることである. 複雑ネットワーク中のコミュニティを抽出する方法として, モジュラリティ [19, 20] を最大化する手法がよく用いられる. ネットワーク分割の評価関数であるモジュラリティ $Q(\{\sigma\})$ は次の様に定義される:

$$Q(\{\sigma\}) \equiv \frac{1}{m} \sum_{i,j} [A_{i,j} - p_{i,j}] \delta(\sigma_i, \sigma_j), \quad p_{i,j} = \frac{k_i k_j}{m}. \quad (11)$$

ここで $A_{i,j}$ は隣接行列の (i, j) 要素, $k_i = \sum_j A_{i,j}$ はノード i の次数 (そのノードにつながるリンク数, あるいはリンクの重みの合計), $m = \sum_i \sum_j A_{i,j}$ はネットワーク全体の次数の合計であり, クロネッカーのデルタ $\delta()$ はノード i とノード j が同一のコミュニティに所属していることを保証する. また, $p_{i,j}$ はネットワークをランダム化した場合にノード i, j 間にリンクが張られる確率 (ノード i, j 間の重みの期待値) である. つまり, もし対象のネットワークが対応するランダム・ネットワークに比べて有意のコミュニティ構造をもつならば, 最大化されたモジュラリティは正の値をとる.

コミュニティ抽出の有用性を示す代表例は Zachary の空手クラブ・ネットワーク [21] である. Zachary は, 米国のある大学における空手クラブが運営方針の対立から

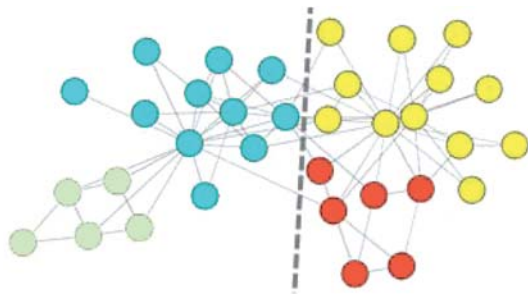


Fig. 7: Four communities obtained through the modularity maximization for the Karate Club network[21], and the dashed line depicts its actual split into two groups.

2 派に分裂する様子を観察した。Fig. 7 は、この空手クラブの交友関係のネットワークからモジュラリティの最大化により抽出したコミュニティであり、コミュニティはノードの濃淡で区別されている。モジュラリティによるコミュニティ抽出により、実際に起こった内部分裂の境界（図中の破線）を正確に検出できている。

6. 株価相関ネットワークのコミュニティ

Kim and Jeong[8] や Pan and Sinha[9] はグループ相関行列 C_{group} からネットワークを作って相関の強い銘柄グループを抽出している。彼らはある閾値を設定し、その閾値を超える相関係数をもつ銘柄をリンクすることによりネットワークとした。我々は相関の強い銘柄グループを抽出するとともに、抽出されたグループ間の関係性を知るために負の値の相関係数もそのままに C_{group} を隣接行列として分析する。つまり、値動きに正の相関のある銘柄同士は正の重みを持つリンクでつながっており、逆に、負の相関のある銘柄同士は負の重みのリンクでつながっているネットワークを考える。値動きに正の相関のある銘柄を同じグループに集め、負の相関のある銘柄を別のグループへと分けるためにフラストレーション³と呼ばれる次の目的関数を最小化する [14]:

$$F(\{\sigma\}) = -\sum_{i,j} A_{i,j} \delta(\sigma_i, \sigma_j). \quad (13)$$

すなわち、グループ内の重みの総和を最大化するようにネットワークを分割することで、グループ内には正のり

3. 第 5 節で紹介したモジュラリティは負の重みを持つリンクに対応していない。そこで負のリンクに対応するようにモジュラリティを拡張した新たな量、ハミルトニアンが提案された [13, 14]。それは次のように定義される:

$$\mathcal{H}(\{\sigma\}) \equiv -\sum_{i,j} \left[A_{i,j} - (\gamma^+ p_{i,j}^+ - \gamma^- p_{i,j}^-) \right] \delta(\sigma_i, \sigma_j), \quad (12)$$

ここで、 $p_{i,j}^\pm$ はノード i, j 間の正（負）のリンクの次数の期待値であり、 γ^\pm はノードのまとまりやすさに関するパラメータである。フラストレーション $F(\{\sigma\})$ はハミルトニアンの特な場合 ($\gamma^\pm = 0$) である。

Table 1: Decomposition of stocks into communities at varied m in the Tokyo Stock Exchange. C_{group} and $C_{\text{group}} + C_{\text{random}}$ have almost the same structures.

| m | Community | | | |
|--|-----------|-----|-----|-----|
| | I | II | III | IV |
| 2 | 279 | 278 | – | – |
| 3 | 199 | 192 | 166 | – |
| 4 | 175 | 149 | 119 | 114 |
| 5 | 175 | 149 | 119 | 114 |
| 6 | 177 | 149 | 117 | 114 |
| 7 | 179 | 148 | 117 | 113 |
| 8 | 179 | 150 | 117 | 111 |
| 9 | 178 | 146 | 118 | 115 |
| 10 | 178 | 145 | 119 | 115 |
| 11 | 173 | 146 | 123 | 115 |
| 12 | 174 | 146 | 122 | 115 |
| 13 (C_{group}) | 175 | 148 | 118 | 116 |
| $C_{\text{group}} + C_{\text{random}}$ | 177 | 148 | 117 | 115 |

Table 2: Decomposition of stocks into communities at varied m in the S&P 500. C_{group} and $C_{\text{group}} + C_{\text{random}}$ have the same structures.

| m | Community | | | |
|--|-----------|-----|-----|----|
| | I | II | III | IV |
| 2 | 247 | 236 | – | – |
| 3 | 197 | 161 | 125 | – |
| 4 | 184 | 159 | 140 | – |
| 5 | 183 | 162 | 138 | – |
| 6 | 171 | 161 | 151 | – |
| 7 | 170 | 161 | 152 | – |
| 8 | 168 | 163 | 152 | – |
| 9 (C_{group}) | 166 | 157 | 93 | 67 |
| $C_{\text{group}} + C_{\text{random}}$ | 166 | 157 | 93 | 67 |

ンクが集まり、グループ間には負のリンクが集積する。つまり、協調して値動きする銘柄がグループを形成し、銘柄グループ同士は反相関の関係となる。

フラストレーションの最小化により C_{group} , $C_{\text{group}}^{(m)}$, および $C_{\text{group}} + C_{\text{random}}$ から抽出されたコミュニティの数とサイズを Table 1 および Table 2 に示した。フラストレーションの最小化には焼き鈍し (simulated annealing) 法 [22] を利用した。東京市場ならびに S&P 500 市場において、どちらについても最終的に 4 つのグループが得られた。東京市場については上位 3 つの固有成分 ($\ell = 2, 3, 4$) だけでグループ分割が完全に決定されることを見

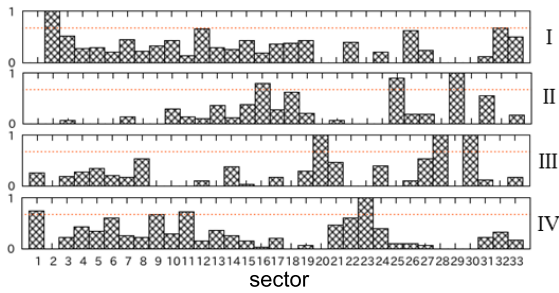


Fig. 8: Fragmentation of the business sectors into communities in the Tokyo Stock Exchange. The dotted line represents the fraction level of 2/3.

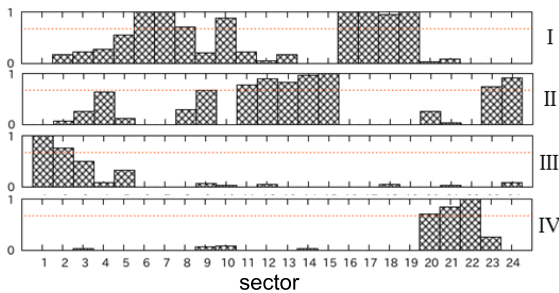


Fig. 9: Same as Fig. 8, but in the S&P 500.

いだした．他方，S&P 500 市場については上位 2 つの固有成分 ($\ell = 2, 3$) および $\ell = 9$ の 3 成分だけでグループ分割が決定される．いずれにしても C_{random} (東京市場では 64% の寄与，S&P 500 市場では 38% の寄与をもつ) がグループ分割に何も影響を与えず，このグループ分割の結果は RMT による主成分判定の結果と矛盾しない．

Fig. 8 および Fig. 9 はそれぞれのコミュニティに各業種が属する割合を示したものである．東京市場で特にまとまりのよい業種は，第 II コミュニティの証券・先物取引業 (29 番)，第 III コミュニティの銀行業 (28)，電力・ガス業 (20)，保険業 (30) などであるが，それほどまとまりがよくない業種も多く，2 つあるいは 3 つのコミュニティに分かれている．

一方，S&P 500 市場はそれぞれの業種 (産業グループ) ごとのまとまりもよく，さらに上位のカテゴリ (セクター) としても非常によいまとまりを示している．そのため 4 つのコミュニティを，第 I コミュニティは Consumer Discretionary と Financials，第 II コミュニティは Industrials，Consumer Staples，Telecommunication Services と Utilities，第 III コミュニティは Energy と Materials，第 IV コミュニティは Information Technology のように，セクターで特徴づけることもできる．

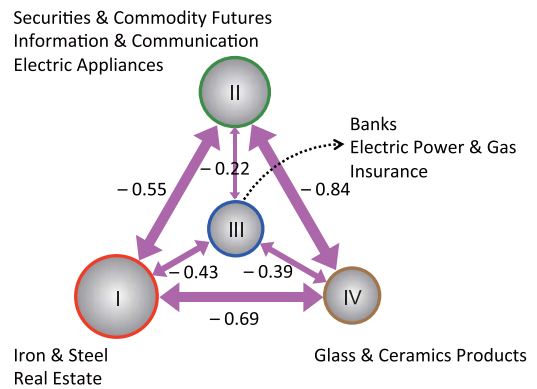


Fig. 10: Frustrated stock group structure in the Tokyo Stock Exchange. The arrows designate the anticorrelation relationship between communities with the corresponding values of P_{corr} , Eq. (14), represented by their width.

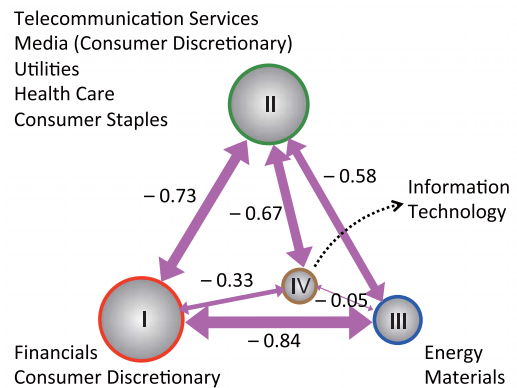


Fig. 11: Same as Fig. 10, but in the S&P 500.

7. フラストレーション構造

グループ内ならびにグループ間の相関関係を定量的に計測するために，各場合に対してリンクの重みの偏極率を導入した．それは次のように計算される：

$$P_{\text{corr}} \equiv \frac{\sum_{\text{within/between}} [C_{\text{group}}]_{i,j}}{\sum_{\text{within/between}} [C_{\text{group}}]_{i,j}}. \quad (14)$$

ここで， $\sum_{\text{within/between}}$ はグループ内ならびにグループ間でのリンクの重みの総和を表す．Fig. 10 および Fig. 11 は，東京市場，S&P 500 市場それぞれについて得られた 4 つのグループの業種の特徴と相互相関の様子を示す．グループ内での偏極率は 80% を優に超え，各グループの銘柄が高い協調関係にあることを確認できる．グループ間の偏極率については，図中のリンクの矢印に添えて表示している．どの相互リンクも負の偏極率をもち，グループ間が競合関係にある．このように相反する 4 つのグループには，すべてを満足させる状態は決して実現し

ない。つまり、日米の株式市場の背後にはフラストレーション状態⁴にある4極構造が共通して存在する。東京市場における4つの極の内、3つは互いに強い反相関関係にあり、第3極（金融、電気・ガスなど）は、他の相反する3つの極と比較的独立した地位を占めている。S&P 500市場では東京市場に比べて業種ごとのかたまりがより鮮明であり、金融業も相反する1極をなしている。また、第4の極であるIT業は第3極を構成しているエネルギー、材料業とほとんど無相関である。

このような多極のフラストレーション構造を直感的に理解するために、銘柄 j に対して次の「相関状態ベクトル」 d_j を導入する:

$$d_j^T = \left[\sqrt{\lambda_2} u_{2,j}, \dots, \sqrt{\lambda_{N_c}} u_{N_c,j} \right]. \quad (15)$$

このように定義された相関状態空間は東京市場においては13次元、S&P 500市場では9次元であり、銘柄 i, j 間のグループ相関係数は対応する相関状態ベクトルの内積で与えられる:

$$[C_{\text{group}}]_{i,j} = d_i^T d_j \quad (16)$$

4つの極の重心を通る3次元空間を切り出し、それらの重心ベクトルを含めて日米の市場における銘柄間の相関構造を可視化したものが Fig. 12 と Fig. 13 である。それぞれの市場について4極フラストレーション構造を容易に見て取ることができる。

8. おわりに

多変量時系列の主成分分析においてランダム行列理論 (RMT) が与える固有値分布、特に最大固有値を主成分判定規準として有意な相関を抽出する手法を用いて、日米の株式市場の中に埋め込まれた相関構造を明らかにすることを試みた。その際に、コミュニティ抽出というネットワーク科学の手法を取り入れた。

解析の結果、東京市場で先に発見した株式銘柄間のフラストレーション構造が米国市場においても存在することを確認し、成熟した市場で普遍的な性質であろうとの強い示唆を得た。また、そのようなフラストレーション構造が、株式市場の複雑な振る舞いを生む要因となっているかもしれない。現在、東証一部、S&P 500の両市場について、2001年 - 2011年の新しいデータについて解析中である。全期間を前期、中期、後期に分け、それぞれの時期でコミュニティを抽出してみると、いずれの市場においてもフラストレーション構造が継続して存在していることを確認できる。そのような特異の相関構造が時間的にどのように変化しているのか、市場モードの

4. 社会ネットワーク（例えば、友好関係にあるときは正の重み、敵対関係にあるときは負の重みを仮定）においては“unbalanced”との表現が用いられる。

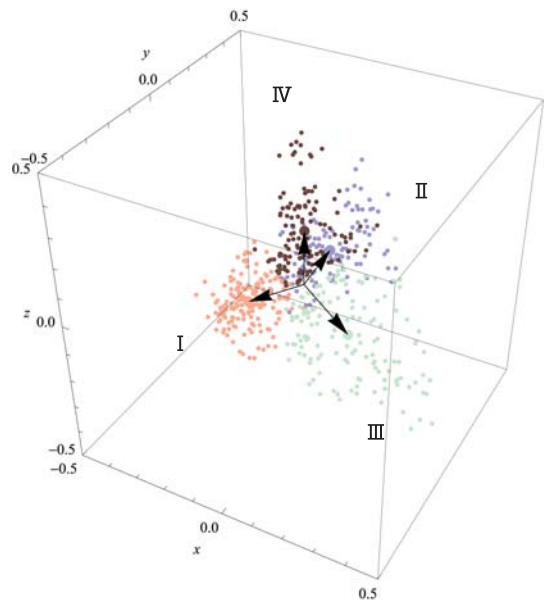


Fig. 12: Three-dimensional correlation state vectors of the stocks in the Tokyo Stock Exchange. The stocks belonging to different communities are distinguished by varied degrees of shading. The arrows point to the positions of the center of gravity in individual communities with their designations.

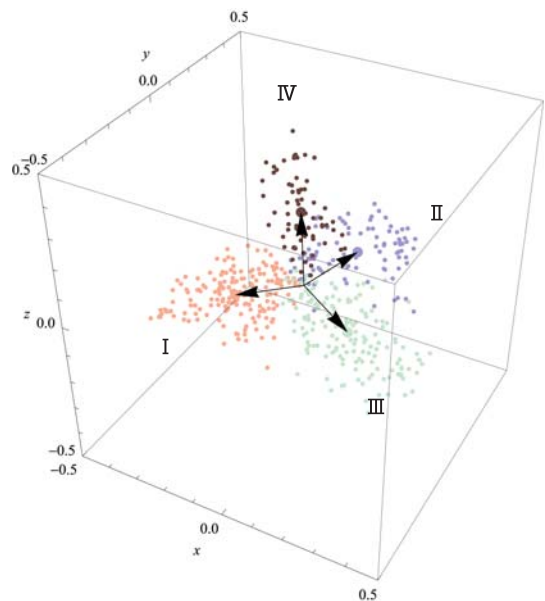


Fig. 13: Same as Fig. 12, but in the S&P 500.

振る舞いと関係があるのか無いのか、もしあるならばどのような関係があるのか等々、今後さらに解析を進めていく予定である。加えて、日米市場の枠を超えた株価連関、新興市場における株価相関の様子など研究の興味は尽きない。

相関行列は複雑系の研究においてしばしば登場する。

本稿では株式市場における相関構造の解析を題材に最新
の多変量解析の研究手法について解説した。具体例の中
に手法の要点が溶け込んでいるが、慧眼なる読者諸氏は
その手法の汎用性についてすでにお気づきのことと思う。

謝辞: 本研究は JSPS 科研費 22300080, 25400393 の助成を
受けたものである。

参考文献

- [1] M. L. Mehta: “Random Matrices,” 3rd (Ed.), Academic Press, 2004.
- [2] L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters: “Noise dressing of financial correlation matrices,” Phys. Rev. Lett., Vol.83, pp. 1467-1471, 1999.
- [3] 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,” Phys. Rev. Lett., Vol.83, pp. 1471-1474, 1999.
- [4] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, T. Guhr, and H. E. Stanley: “Random matrix approach to cross correlations in financial data,” Phys. Rev. E, Vol.65, p. 066126, 2002.
- [5] 青山秀明, 家富洋, 池田裕一, 相馬亘, 藤原義久: 経済物理学, 共立出版, 2008.
- [6] J.-P. Bouchaud and M. Potters: “Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management,” 2nd (Ed.), Cambridge University Press, 2009.
- [7] A. Utsugi, K. Ino, and M. Oshikawa: “Random matrix theory analysis of cross correlations in financial markets,” Phys. Rev. E, Vol.70, p. 026110, 2004.
- [8] D. H. Kim and H. Jeong: “Systematic analysis of group identification in stock markets,” Phys. Rev. E, Vol.72, p. 046133, 2005.
- [9] R. K. Pan and S. Sinha: “Collective behavior of stock price movements in an emerging market,” Phys. Rev. E, Vol.76, p. 046116, 2007.
- [10] M. S. Santhanam and P. K. Patra: “Statistics of atmospheric correlations,” Phys. Rev. E, Vol.64, p. 016102, 2001.
- [11] M. Barthélemy, B. Gondran, and E. Guichard: “Large scale cross-correlations in Internet traffic,” Phys. Rev. E, Vol.66, p. 056110, 2002.
- [12] P. Šeba: “Random matrix analysis of human EEG data,” Phys. Rev. Lett, Vol.91, p. 198104, 2003.
- [13] S. Gómez, P. Jensen, and A. Arenas: “Analysis of community structure in networks of correlated data,” Phys. Rev. E, Vol.80, p. 016114, 2009.
- [14] V. A. Traag and J. Bruggeman: “Community detection in networks with positive and negative links,” Phys. Rev. E, Vol.80, p. 036115, 2009.
- [15] D. Easley and J. Kleinberg: “Networks, Crowds, and Markets: Reasoning About a Highly Connected World,” Cambridge University Press, 2010.
- [16] T. Yoshikawa, T. Iino and H. Iyetomi: “Observation of frustrated correlation structure in a well-developed financial market,” Prog. Theor. Phys. Suppl., Vol.194, pp. 55-63, 2012.
- [17] V. A. Marčenko and L. A. Pastur: “Distributions of eigenvalues for some sets of random matrices,” Math. USSR-Sb., Vol.1, pp. 457-483, 1967.
- [18] A. M. Sengupta and P. P. Mitra: “Distributions of singular values for some random matrices,” Phys. Rev. E, Vol.60, pp. 3389-3392, 1999.
- [19] S. Fortunato: “Community detection in graphs,” Phys. Rep., Vol.486, pp. 75-174, 2010.
- [20] M. E. J. Newman: “Modularity and community structure in networks,” PNAS, Vol.103, pp. 8577-8582, 2006.
- [21] W. W. Zachary: “An information flow model for conflict and fission in small groups,” J. Anthropological Res., Vol.33, pp. 452-473, 1977.
- [22] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi: “Optimization by simulated annealing,” Science, Vol.220, pp. 671-680, 1983.

吉川 丈夫



2012 年新潟大学大学院自然科学研究科数理物質科学専攻博士前期課程修了。現在、同後期課程在籍。

家富 洋



1984 年 3 月東京大学大学院理学系研究科物理学専攻博士課程修了。理学博士。同年 4 月東京大学理学部助手, 1993 年 4 月株式会社日立製作所エネルギー研究所研究員, 1995 年 1 月新潟大学理学部助教授, 2002 年 1 月同教授, 2012 年 4 月東京大学大学院経済学研究科特任教授, 2013 年 5 月新潟大学理学部教授, 現在に至る。
