

A-010

ランダム行列の固有値分布との比較による米国株価日次変動のトレンド抽出 Trend-extraction of Daily Stock Prices in the American Market by Means of RMT-PCM

木戸 丈剛†
Takemasa Kido

田中 美栄子†
Mieko Tanaka-Yamawaki

1. はじめに

金融価格時系列の持つ強いランダム性に埋もれた有意な情報を抽出することは興味深い課題である。中でも、証券市場で取引される数百以上の株式価格間の関連性を計算機を利用して扱う手法の開発は、証券市場のシステム化に伴って近年急激に重要性を帯びてきた。

一方、ランダム行列理論[1]が様々な分野での応用を視野に入れて大きな期待を寄せられる今日、この理論より導かれる公式[2]を株価時系列の相関行列の固有値分布と比較することで有意成分を分離する方法が注目を集めている[3-6]。この方法を、ランダム行列理論(Random Matrix Theory:RMT と略)を応用した主成分分析手法(RMT oriented Principal Component Analysis:RMT-PCA と略)と名付けておく。本稿はその有効性を、株式市場における個々の株価の価格変動データ解析に対し検証しようとするものである。

通常の相関行列の固有値・固有ベクトルを用いた主成分分析手法では、第 n 固有ベクトルを第 n 主成分とし、各データとの内積をとって作成した n 成分得点からなる n 次元空間の中でクラスター解析を行う方法が取られるが、これはデータ数が比較的少数の場合にしか適用できないため、株式市場のように数百以上の株価が連動して複雑な動きを形成するような多次元に対しては本稿で考察する RMT を取り入れた手法が有効となる[7, 8]。

この方法を我々は、米国株式市場の16年分の日次データに適用し、2年分を1データとして解析した結果、顕著な動きを示す業種の時間変化を抽出することができた。

本論文の2章以降の構成は以下のとおりである。2章から4章では研究手法とその具体例を示し、5章に4章までに示した手法を全ての期間に適用した結果とその結果に対する考察を述べる。

2. 株価データの取得と編集

我々は2009年度版のS&P500リストにある500社の毎日の終値を対象に分析を行った。アメリカの株価のうちStandard&Poors社が代表的な500銘柄を選出したものがS&P500である。これはYahoo FINANCE USAのwebサイトから会社の上場日から前日までの日次データをExcelファイル形式で取得できる。

分析対象は1994年1月から2009年9月までとした。日次データを扱うため1社に対する1年間のデータ量は証券市場の営業日である約250となる。1994年から1995年、1996年から1997年の様に2年毎を1期間とすることにより、銘柄数500を超える時系列長を確保した。

同時刻時系列間の内積を計算するには、全ての株の全時刻の株価が揃わなければならないが、市場で取引される全ての株がこの条件を満たすわけではない。そこで、毎日取引のあった株のみを選んで解析が行うことになる。しかし、本研究では2年間の営業日に対し数日分足りないだけの銘柄のデータを捨てるよりも、できるだけ多くの銘柄につい

て解析を行うため、取引が行われなかった日付の終値には変動なしとして前日の終値を補うことで全時刻を揃えた。この操作によって対象の銘柄数を各期間で10程度増やすことができた。

使用データの各期間の銘柄数、時系列長、時系列長/銘柄数を表1に表記する。銘柄数が期間によって変動するのは2009年度版のリストの銘柄を対象とするため期間によってまだ上場されていない銘柄が存在するためである。

表1 期間毎のデータ数

期間	N(銘柄数)	L(時系列長)	Q(L/N)
94~95	373	504	1.35
96~97	397	507	1.28
98~99	419	504	1.20
00~01	447	500	1.12
02~03	464	504	1.09
04~05	472	504	1.07
06~07	486	502	1.03
08~09	494	436	0.88

3. RMT-PCA手法に基づく有意な固有値の分離

3.1 時系列データ行列Gの作成

銘柄数(N)と時系列長(L=取引日-1)のデータに対し、時系列の変動を見るために i 番目の株価の k 番目の時系列($S_{i,k}$)とその1つ前の時系列との差をとる($i=1, \dots, N$; $k=1, \dots, L$)。変動の値ではなく、変動の割合を見るために、各値を対数化し、その対数収益率を計算する。

$$X_{i,k} = \ln(S_{i,k} / S_{i,k-1}) \quad (1)$$

式(1)から、 i 番目の株価の長さLについての平均値($\langle X_i \rangle$)を差し引き標準偏差(σ)で割って正規化する。

$$G_{i,k} = \frac{X_{i,k} - \langle X_i \rangle}{\sigma} \quad (2)$$

こうして得た i 番目の株価時系列を i 行目とした $N \times L$ の行列 G を作成する。

$$G = \begin{pmatrix} G_{1,1} & G_{1,2} & \cdots & G_{1,L} \\ G_{2,1} & G_{2,2} & \cdots & G_{2,L} \\ \vdots & \vdots & \ddots & \vdots \\ G_{N,1} & G_{N,2} & \cdots & G_{N,L} \end{pmatrix} \quad (3)$$

† 鳥取大学大学院工学研究科情報エレクトロニクス専攻,
Tottori University, Graduate School of Engineering,
Department of Information and Electronics.

3.2 同時刻自己相関行列 C の作成

同時刻自己相関行列 C を行列 G とその転置行列 G^T の積により作成する。これは株価 i と株価 j の時系列の内積を時系列長で割ったものを成分とする N×N の行列となる。

$$C_{i,j} = \frac{1}{L} \sum_{k=1}^L G_{i,k} G_{j,k} \quad (4)$$

行列形式にすると (5) のようになる。

$$C = \begin{pmatrix} C_{1,1} & C_{1,2} & \dots & C_{1,N} \\ C_{2,1} & C_{2,2} & \dots & C_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ C_{N,1} & C_{N,2} & \dots & C_{N,N} \end{pmatrix} \quad (5)$$

定義により行列 C は対称行列であり、対角成分は全て 1 となるように規格化されている。非対角成分は他との相関であるから絶対値が 1 以下の値をとる。C の具体的な形を見るため、2006 年から 2007 年の実データ (以下 06-07 データと表記) を例にその数値を示すと次のようになる。

$$C = \begin{pmatrix} 1.0000 & 0.2403 & \dots & 0.2949 \\ 0.2403 & 1.0000 & \dots & 0.2618 \\ \vdots & \vdots & \ddots & \vdots \\ 0.2949 & 0.2618 & \dots & 1.0000 \end{pmatrix} \quad (6)$$

3.3 有意な固有値の判定

相関行列 C の固有値を求め、ランダム時系列の固有値分布に対する結果と比較することにより有意な固有値を取り出す。主成分分析では値の大きなものから加算していきその累積が全体の 8 割に達する時点までの成分を有意な成分とし、分析するのが一般的である。しかしランダム行列理論ではランダム時系列の行列の固有値と比較することで有意な固有値を抽出する。株価を扱う場合の様な多変量解析で累積加算が 8 割となると、主成分とされる成分が多くなり雑音が多すぎるためである。

ランダム時系列に対しては

$$N \rightarrow \infty, L \rightarrow \infty, 1 < Q \equiv L/N \quad (7)$$

の条件で固有値の上限(λ₊)と下限(λ₋)は

$$\lambda_{\pm} = 1 \pm \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}, \lambda_- < \lambda < \lambda_+ \quad (8)$$

として理論式がわかっているのので、この範囲内にある固有値はランダム成分、範囲外にある固有値を有意成分として分離することができる。

また、ランダム時系列の相関行列に対する固有値分布はランダム行列理論から導かれる以下の公式

$$P_{rm}(\lambda) = \frac{Q}{2\pi} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda} \quad (9)$$

が知られているので[2]、これと行列 C の固有値分布のヒストグラムを比較することにより、有意成分とランダム成分とを分離できる。

しかし式(8)の条件を満たすには、銘柄数 N より長い時系列長 L が必要である。従って株式市場で重要な要素となる 400 以上の銘柄の終値を対象とする場合、1 年分のデータでは足りないことになり、最低でも 2 年分を連結する必要がある。この理由により Plerou の論文[3]では 7 年分を連結して用いた結果が書かれていた。本研究では年次変化を追うため、2 年分ずつを 1 データとした。

固有値分布の有意部分を図 1 に、ランダム部分を図 2 に図示する。

図 1 は式(9)の破線と、その範囲に収まらなかった 06-07 データの固有値分布である実線を、比較するために縦軸の表示範囲を 0 から 0.10 に縮小した図である。ランダム行列の固有値分布から逸脱した固有値が 06-07 データに表れていることがわかる。

図 2 では式(8)から計算したランダム行列の固有値の上限と下限の範囲に注目した図である。ランダム行列の固有値分布の式(9)である破線の固有値 λ が 0 付近で急激に増えてしまっているが、形はほぼ一致している。

このことから実データから作成した行列の固有値で理論式(9)の λ₋ から λ₊ の範囲にある固有値はランダムな行列からでも出現する固有値といえる。逆に図 1 に出てきているような λ₊ 以上の固有値はランダムな行列ではありえない有意な固有値といえることができる。

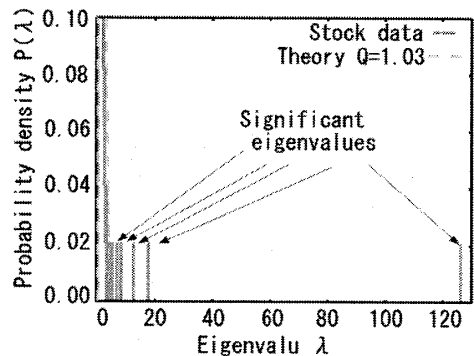


図 1 理論式(9)と 06-07 データの固有値分布の比較

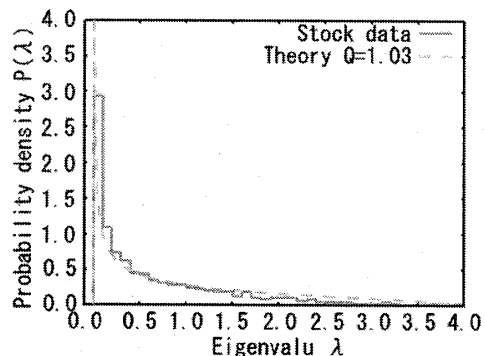


図 2 理論式(9)と 06-07 データの固有値分布の比較 (固有値 λ の範囲 0-4)

対象とした1994年から2009年の米国株の日次変動データでは有意な固有値の基準値 λ_+ 以上の固有値は5~10個となった。例えば06-07データは $N=486, L=501, Q=1.03, \lambda_+=3.94$, となり8個の有意な固有値を分離できた。

これにより図3に示すように、通常の主成分分析では固有値の累積加算が固有値全体の8割になるまでの固有値を主成分とする手法を本研究の様なデータに適用すると、主成分の数は数十~数百になってしまう、しかし図4のようにランダム行列理論を利用し実データの固有値をランダム行列の固有値と比較することで主成分を10以下に絞ることができた。

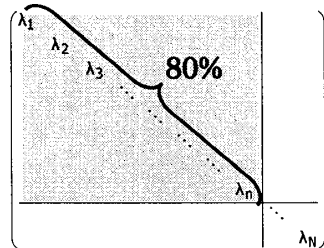


図3 主成分の基準を累積固有値8割とした場合の主成分

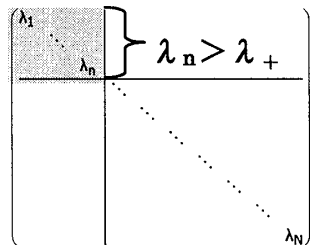


図4 主成分の基準にRMTを用いた場合の主成分

4. 固有ベクトルの分析

4.1 固有ベクトルの成分の相関

これまでの議論で主成分を絞って有意な固有値を抽出する方法ができたので、次に固有ベクトルを分析する。最大固有値 λ_1 の固有ベクトル u_1 の成分には目立った特徴が見えないので、2番目に大きな固有値 λ_2 以降の有意な固有値に対する固有ベクトル u_2 以降に注目する。

固有ベクトル u_2 の成分のなかで顕著に大きなものから20成分に注目する。本来固有値 λ_2 への影響を考慮し、どの成分に着目するかを定めるべきだが、着目すべき成分を決定する手法は今後の課題とし本稿では一意に20成分とした。

固有ベクトルの各成分がどのような値をとっているかを確認するために06-07データの u_2 の各成分の値を棒グラフにしたものを図5に図示する。

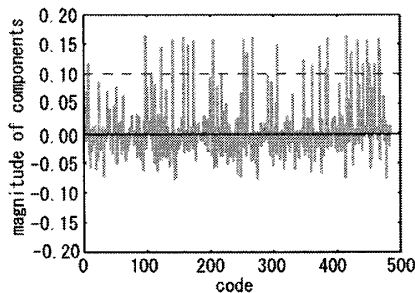


図5 固有ベクトル u_2 の成分内訳(06-07)

注目する20成分をその固有ベクトルの主要成分とし、主要成分同士の相関を確認するために、06-07データの行列Cから20成分同士に対応する要素を抽出する。表2は主要20成分から4成分を、表3は行列Cからランダムに4成分をそれぞれ抽出し表に示したものである。また、本稿での対象データは株価であり1成分がそれぞれの株に対応することを記しておく。

表2 固有ベクトル u_2 の主要成分の相関(06-07)

u_2 の成分の値	symbol	DVN	APA	RRC	SWN
0.164	DVN	1.00	0.77	0.75	0.71
0.164	APA	0.77	1.00	0.75	0.71
0.161	RRC	0.75	0.75	1.00	0.75
0.161	SWN	0.71	0.71	0.75	1.00

表3 固有ベクトル u_2 の主要成分以外の相関(06-07)

u_2 の成分の値	symbol	CVX	BTU	COG	OXY
0.097	CVX	1.00	0.30	0.23	0.16
0.094	BTU	0.30	1.00	0.22	0.15
0.085	COG	0.23	0.22	1.00	0.22
0.085	OXY	0.16	0.15	0.22	1.00

表2と表3を比較すると表3のランダムに抽出した成分同士の相関は大きくても3割程度だった、それに対して表2値の4成分同士では相関の値が7割程度あり、比較的大きい。表2、表3のそれぞれ4成分の株価の変動をそれぞれ図6、図7に図示する。

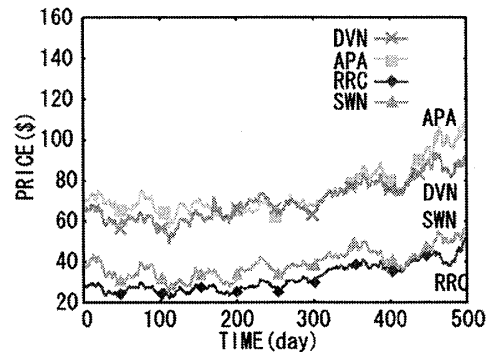


図6 固有ベクトル u_2 の主要成分の変動(06-07)

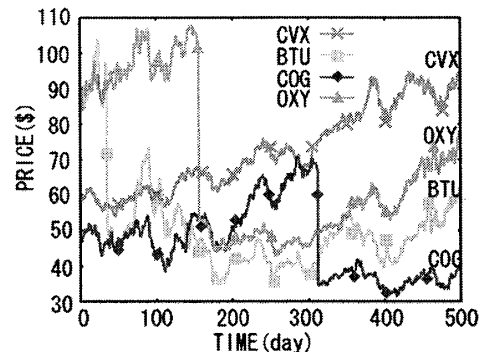


図7 固有ベクトル u_2 の主要成分以外の変動(06-07)

図6と図7を比べるとやはり相関の値が大きいと変動が似ていることが見て取れる。

まとめるとランダム行列理論を利用し主成分を絞り込む。主成分とされた固有値に対応する固有ベクトルの成分で値の大きな成分に注目すると相関関係にあることがわかった。相関関係があるということは注目された成分はランダムな変動ではなく有意な変動をしているといえる。

4.2 各銘柄の業種分類

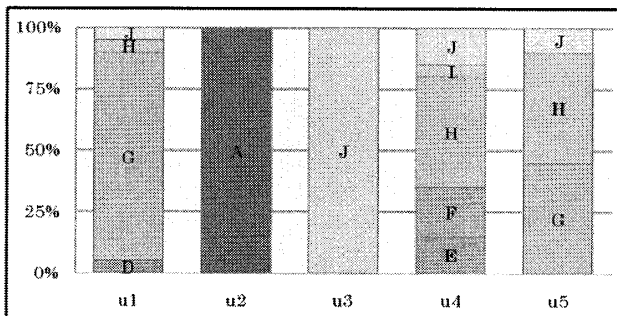
前節までに有意な変動をしている株を抽出することができた。これらの株に共通する特徴を見いだすことができれば、その期間のトレンドを抽出できたことになる。特徴の1つとして抽出された株の業種に注目した。

本研究対象のS&P500に含まれる全ての銘柄には業種コードが割り振られている。各銘柄の業種分類はGICS(ギックス)コードによって行う。GICSコードはS&P社とMorgan Stanley Capital International(MSCI)が共同で作成した(Global Industry Classification Standard=GICS)であり、世界産業分類基準である。世界産業分類基準は、10のセクターと24の産業グループ、67の産業、および147 S&P500の内訳の産業サブグループからなり、1つの企業は各階層で1つのグループにしか入ることができない。

3つ以上の分野にまたがって事業展開している多角化企業で、売上高または利益のどちらかが全体の半分以上を占める事業部門がない場合は、コングロマリット産業サブグループ(一般事業会社セクター)、またはマルチセクター持株会社産業サブグループ(金融セクター)に分類される。

4.3 固有ベクトルの主成分の業種構成

GICSコードによる業種分類を06-07データで抽出された成分に対して行った結果、本研究でも企業の業種に偏りがあることがわかった。その結果を図9に図示する。



※A:エネルギー B:素材 C:資本財 D:サービス
E:生活必需品 F:ヘルスケア G:金融 H:情報技術
I:電気通信 J:公益事業

図9 06-07データの主成分の固有ベクトルの上位20成分の業種内訳

06-07データでは u_1, u_2, u_3, u_5 業種が偏り、特に u_1, u_2, u_3 でそれぞれ金融、エネルギー、公益事業に偏り、顕著な結果が表れた。有意な変動をする株を分析した結果06-07データでは金融、エネルギー、公益事業関係の株が集まった、つまりこの3種の業種がこの期間のトレンドとすることができる。

5. 結果

5.1 各期間の有意な固有値とRMTによる基準値 λ_+

06-07データで行った分析と同様の分析を表1に示した8データすべてに対して行った。その結果全ての期間で固有値 $\lambda_1 \sim \lambda_5$ が主成分となったため、その比較を行う。

表4に8データの固有値 $\lambda_1 \sim \lambda_5$ の値、各データでのランダム行列の固有値の上限 λ_+ と各データの主成分となった固有値の数を示す。

表4 各データの固有値と理論との比較

	94-95	96-97	98-99	00-01
λ_1	33.40	52.62	54.29	64.16
λ_2	9.75	10.50	15.42	27.93
λ_3	6.73	6.71	12.12	17.26
λ_4	5.02	5.37	8.14	10.84
λ_5	4.59	4.39	5.99	8.51
λ_+	3.46	3.54	3.61	3.78

有意な固有値の数

	6	7	6	11
--	---	---	---	----

	02-03	04-05	06-07	08-09
λ_1	139.65	95.56	126.28	240.58
λ_2	15.19	18.70	17.64	20.72
λ_3	13.72	13.36	12.48	15.50
λ_4	10.50	8.08	8.14	9.11
λ_5	7.67	5.78	6.76	5.95
λ_+	3.83	3.86	3.94	4.27

有意な固有値の数

	10	9	8	7
--	----	---	---	---

有意な固有値の閾値に式(8)の λ_+ を適用しているが、この λ_+ はランダム行列の固有値の上限の理論値である。実際に乱数を使い実データと同様の方法で固有値を求めた場合多少の誤差が発生する可能性も考慮する。例えば06-07データでは表4に示されていないが、 λ_5 以降の固有値は $\lambda_6=5.07, \lambda_7=4.82, \lambda_8=4.15, \lambda_9=3.95$ となり $\lambda_+=3.94$ となっている。有意成分の数が十分に確保できている場合は誤差を考慮し、閾値と近い固有値は主成分としては取り除くことも考える必要がある。

5.2 固有ベクトルの主要成分 20 社の業種内訳

表4で示された有意な固有値に対する固有ベクトルの主要成分を4.2節で述べたGICSコードによって業種を分類する。各固有ベクトルの主要成分の業種の内訳を表5に示す。

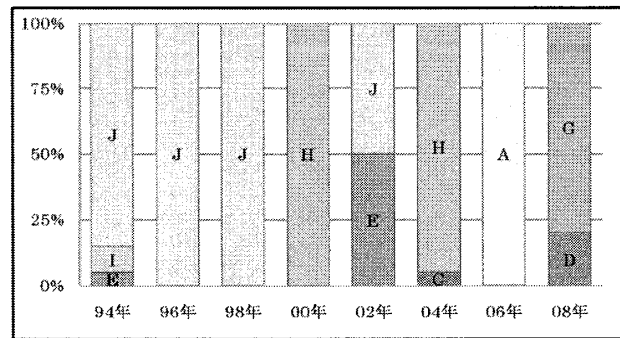
表5 各固有ベクトルの上位20成分の主な業種

	94-95	96-97	98-99	00-01
u_1	G:6 J:8	G:8	G:14	C:6 G:10
u_2	J:17	J:20	J:20	H:20
u_3	A:16	A:19	H:11	E:8 G:12
u_4	偏りなし	偏りなし	B:12 C:8	J:20
u_5	E:8 F:8	G:8	F:7 G:7	F:6 G:9

	02-03	04-05	06-07	08-09
u_1	G:16	G:12	G:17	C:9
u_2	E:10 J:10	H:19	A:20	G:16
u_3	A:19	G:13 J:7	J:20	E:8 F:6
u_4	J:20	偏りなし	H:9	D:10 H:10
u_5	F:20	D:13	G:9 H:9	D:8 F:9

※A:エネルギー B:素材 C:資本財 D:サービス
E:生活必需品 F:ヘルスケア G:金融 H:情報技術
I:電気通信 J:公益事業

表5をデータの比較のために各データの u_2 についてまとめたものを図10に図示する。



※A:エネルギー B:素材 C:資本財 D:サービス
E:生活必需品 F:ヘルスケア G:金融 H:情報技術
I:電気通信 J:公益事業

図10 固有ベクトル u_2 の上位20成分の業種内訳

各期間ともに u_2 の主要成分が特定の業種に集中し、その業種が期間によって変化していることがわかる。これをその期間のトレンドと考えることができる。そこで表5から読み取れる各期間の特徴を以下に述べる。

94-95データでは u_2, u_3, u_5 には業種の偏りがみられるが、 u_4 では業種の偏りがみられない。94-95データでは偏って大きい成分が u_4 以外の固有ベクトルでは10成分以上有るのに対し、 u_4 は4成分しかなかったため、20成分までみると相関関係がなくなり、業種の偏りがみられないと考えられる。このことから94-95データの特徴は u_2, u_3, u_5 から読み取ると公益事業、エネルギー、ヘルスケア、生活必需品関連の銘柄の株がランダムでない動きをしていると考えられる。

同様に96-97データでは u_2, u_3 には業種の偏りがみられるが、 u_4, u_5 では顕著に大きな成分が少なかったため業種の偏りがみられない。また u_5 の固有値 λ_5 が4.39と低いため、ランダムな固有値とも考えられる。このことから96-97データの特徴は u_2, u_3 からのみ読み取ると公益事業、エネルギー関連の銘柄に絞られる。

98-99データでは u_2 の公益事業と u_4 の素材関連の銘柄がランダムでない動きをしていると考えられる。

00-01データでは情報技術、金融、生活必需品、公益事業関連の銘柄が、02-03データでは $u_2 \sim u_5$ 全ての固有ベクトルで業種の偏りがみられ、公益事業、エネルギー、ヘルスケア、生活必需品がランダムでない変動が起こり特に公益事業について特徴的な変動あったと推測される。04-05データでは、情報技術、金融、サービス関連の銘柄でランダムでない変動があったと推測される。06-07データでは、エネルギー、公益事業、情報技術、金融関連の銘柄で、08-09データでは、金融が目立ち次いでサービス、ヘルスケア、情報技術の銘柄でランダムでない変動があったと推測される。

次に業種の偏り以外にも上位20成分の相関関係を見る。00-01データの業種の偏らなかつた u_5 の主要な成分は、ランダムな変動をする銘柄の集まりかといえそうではなかつた。 λ_5 の値も大きく u_5 の値の大きな成分も多く、相関の値も大きかつたため主要成分同士は相関関係にあると思われる。つまり、00-01データの u_5 の主要成分は業種によらない相関関係にある銘柄だと考えられる。他にも04-05データの u_2 の成分では大きな要素はなかつたが相関関係にある要素の集まりであった。

最後に第1固有ベクトル u_1 について述べる。 u_1 は市場全体を牽引する優良企業が集まるとされている[3, 4]。本研究では各期間の u_1 は金融が目立つ。このことから金融関係の株の影響が市場にも大きな影響があると推測できる。

1994年から2009年まで通して公益事業が登場するが、2000年以前は u_2 で出現し有力な業種だったが、2000年以降 u_2 は情報技術などに変わり、公益事業は u_3, u_4 で出現するようになった。これは2000年以降情報技術の進歩と業種としての情報技術の進歩によって、株の売買が盛んに行われるようになることで、1日ではあまり株価が変動しない公益事業よりも、注目されている業種の変動がみられるようになったなどと考えることができる。わかりやすい例として08-09で金融が目立ち、 u_1 に資本財(住宅関係)が現れたのは金融危機による変動などである。

6. まとめ

本研究では2009年のS&P500のリストにある米国株価の500銘柄の日次変動データを用いて解析を行い、1994年から2009年の16年間の2年毎の特徴を、有意な変動をする銘柄の業種の偏りとして抽出することを目的とした。有意な変動をする銘柄を抽出する手法として、ランダム行列理論を用いて行った。

その過程でNが約500あれば、全ての有意成分を漏れなく抽出することはできないが、という条件はつくが、 $Q < 1$ の場合でも有意成分とランダム成分を区別する目的に限っては分離の閾値としてランダム行列の理論値を使用可能だということがわかつた。実験結果からランダムでない変動をする銘柄とその業種の構成がわかり、そして期間毎に比較することで、各期間の有望な業種を知ることができ

た. このことで長期的な投資戦略の指標として利用できる可能性が見えたといえる.

だが本研究では問題点として, 有意な固有値を分離したあとの固有ベクトルの主要成分を, 値の大きなものから20成分と一意的に決めてしまった. また, 今回の固有ベクトルの分析は正方向の成分だけに行った結果である. このことを踏まえて

- ・固有ベクトルの主要成分の正確な閾値を決定する方法の確立
- ・寄与率の大きな正負両方向の固有ベクトル成分の分析の2点が今後の課題となる.

参考文献

- [1] M. L. Mehta, "Random Matrices", Academic Press 3rd edition, 2004.
- [2] A. M. Sengupta and P. P. Mitra, "Distribution of singular values for some random matrices", Physical Review E vol. 60, pp. 3389, 1999.
- [3] V. Plerou, et. al, "Random matrix approach to cross correlation in financial data", Physical Review E 65, 066126, 2002.
- [4] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, and H. E. Stanley, Physical Review Letters, 83, 1471, 1999.
- [5] L. Laloux, P. Cizeaux, J. -P. Bouchaud, and M. Potters, 83, 1467, 1999.
- [6] J. -P. Bouchaud and M. Potters, "Theory of Financial Risks", Cambridge University Press, 2000: "金融リスクの理論"(森平監訳)朝倉書店, 2003.
- [7] 永尾太郎: ランダム行列の基礎, 東京大学出版会, 2005.
- [8] 青山秀明, 他: 経済物理学, 共立出版, 2008.
- [9] 田中美栄子, 田中瑶子, 伊藤大哲, 中村元紀, 木戸丈剛, 川村綾, 佐藤彰洋: ランダム行列との比較による NYSE 株価 1 時間変動の相関行列分析(1), 素粒子論研究(京都大学基礎物理学研究所)117 巻 5 号, E85-E86, 2009 年 12 月.
- [10] 田中美栄子, 伊藤大哲, 田中瑶子, 木戸丈剛: ランダム行列理論との比較による NYSE 株価 1 時間変動の解析(2), 素粒子論研究(京都大学基礎物理学研究所)117 巻 5 号, E87-E88, 2009 年 12 月.
- [11] 田中美栄子, 田中瑶子, 伊藤大哲: ランダム行列との比較による NYSE 株価 1 時間変動の相関行列解析, 統計数理研究所共同研究レポート第 241 巻, 「経済物理とその周辺(6)」(統計数理研究所), 27-31, 2010 年 3 月.
- [12] 伊藤大哲: ランダム行列理論の固有地分布に基づく主成分分析手法の適用条件, 鳥取大学工学部卒業論文, 2010 年 3 月