

# 株式市場における市場心理の ニューラルネットワークによる可視化

岩田 英朗

## 1. はじめに

公開株式市場の上場株式価格を基に特定の計算式より算出される株価指数は、地理的・社会的・政治的制約下での重要な経済指標の一つと認識され、時系列データの入手容易性及び詳細性も相まって、しばしば分析対象となっている<sup>1)</sup>。又、株式指標を含む様々な経済時系列データの分析に際しては、時系列における単位根問題の存在が広く知られている<sup>2)</sup>。

株価指数時系列の分析手法はファンダメンタルズ分析<sup>3)4)</sup>とテクニカル分析に大別できる。人間の社会活動全体が決定論的カオスだと仮定すれば、その一形態として発現する株価指数時系列はカオス理論を用いたモデル化の対象となり、株式市場の状態理解を促進できる<sup>5)</sup>。しかしモデル化に際しては、株式市場に影響を与えるファンダメンタルズ候補が無数に存在する為、共分散構造分析や探索的多変量解析を用いた現状分析及び解析は容易ではない<sup>6)</sup>。天変地異や戦争といった突発的・偶発的要因が政治・経済・社会状況を大きく変化させる可能性も考慮しなければならない。つまり、変動との相関の高いファンダメンタルズが時間と共に動的変化し、状況に対応可能なモデルとしなければならない。

テクニカル分析の分野では、市場モデル構築の際に効率的市場仮説 (efficient-market hypothesis)<sup>7)8)</sup>と限定合理性に基づく満足化仮説 (satisficing hypothesis)<sup>9)10)</sup>のどちらを採用するのが妥当であるか、に関する議論が続いている。単位時間当たりの変化量がブラウン運動であるとする確率微分方程式に基づくモデルは前者に立脚し、金融工学の現場で広く用いられ一定の成果を挙げている。ただし、同モデルが真である場合は正確な株価予測が事実上不可能となり、確率論に立脚した議論のみ存続可能となる。一方で後者が真であった場合、外的環境の変動が時系列変動に何らかの周期性を生じさせている可能性が生まれる。

Mandelbrot<sup>11)</sup>は株価時系列の振舞いがフラクタル性を有する事実を発見したが、筆者らはシミュレーションにより、確率微分方程式によって表記された時系列の振舞いがフラクタル構造 (自己相似形) を有することを確認している<sup>12)</sup>。自己相似形の存在をもって両仮説の優劣を判断することはできないが、株価指数時系列が示すある種の周期性を実証的に解明する行為は、現実に即したモデル構築に有効である。又、モデル設定による振舞い予測と現実の差異を検証し、モデルの優位性を議論できる。

これまでに、ニューラルネットワーク (NN) やエージェントモデル、GA、フラクタル解析

等のヒューリスティックな手法や、AR モデル及び ARCH 等の分散不均一モデルを含む AR 拡張モデル、NHA (Non-Harmonic Analysis) 等を用いて、経済時系列の周期性解析及び同予測が試みられている<sup>13)~22)</sup>。株価指数の決定権を有するのは市場参加者である。個々の市場参加者は離散時間  $t$  において、株式の「購入」「売却」「保持」及び「無反応 (無関心)」という 4 形態から一つを選択し、その総和が離散市場価格  $p$  に反映される。株価指数の時系列変動は  $p$  の連続であり、判断のポイントは  $t$  における市場参加者の売買意欲 (市場心理: market sentiment) となる。

ところが、時系列の過去の振舞いが市場心理に影響を与えている可能性は高く<sup>23) 24)</sup>、それが時系列の振舞いにフラクタル性を生み出す原因の一つと考えることができる。実際、市場参加者の意思決定支援ツールとして、日足の各値を用いた回帰直線である移動平均線が古くから用いられており、その代表的タイムスパンは 5・25・75・200 日 (いずれも開場日) である。

そこで本研究では、学習可能なエージェントを複数用意し、期間の異なる時系列から抽出した回帰直線の傾き及び同直線に対する分散で構成する学習セット群を用いた学習を行った。学習によって過去の一定期間における指標の振舞いのある程度記憶するエージェントが生成されたと仮定し、個々のエージェントに近未来の株価指数を予測させた。エージェント個々の予測値と実際値を比較することにより、 $t$  における市場心理の状態を解明しようというのが、本研究の目的である。

## 2. 学習機能を有するエージェントの構築

本研究で用いた、NN に基づく学習が可能なエージェントのモデル構成及び実験に用いたデータの詳細を以下に示す。

### 2.1. 学習機能の詳細

エージェントは、入力層 160、中間層 250、出力層 80 のユニットで構成されたバックプロパゲーション (BP) 法による 3 階層 NN を有する。各ユニットは 0 から 1 の値を入出力し、学習係数  $\alpha$  と  $\beta$  はいずれも 0.8 と定めた。同時に、NN の学習終了条件を「15 万回の学習試行」あるいは「出力層ユニット 1 個当たりの教師信号との誤差平均 (Unit Error Average: 以下 UEA) が 0.05 以下」と定めた。

### 2.2. 実験に用いるデータ

東京証券取引所一部上場株式銘柄全てを対象とした指標である TOPIX と、米国を代表する大型株 500 銘柄から算出される S&P500 の、2005 年 1 月 1 日から 2014 年 12 月 31 日の日足を実験対象とした。なお S&P500 は、ニューヨーク証券取引所等の米国内 3 取引所に上場する銘

柄より信用格付会社である S&P Dow Jones Indices LLC が選出し、算出している。

拙論<sup>12)</sup>で明らかにした通り、時間  $t$  に市場休場日を含めない方が振舞いにフラクタル構造が現れ易い。そこで実験に際しては休場日を除外し開場日のみを時間  $t$  に配し、各指標の日足における高値と安値を用いた。なお Box-counting Method 法<sup>25) 26)</sup>による推定フラクタル次元は、TOPIX 高値 1.44, 同安値 1.46, S&P500 高値 1.38, 同安値 1.40 であった。

### 2.3. 学習データの作成手法

開場日  $t$  における実際の日足・高値を  $Rh_t$  とする時、 $t-(w-1)$  から  $t$  の  $w$  開場日間の高値集合  $Fh_R = \{Rh_{(t-k)} \mid 0 \leq k \leq w-1\}$  を用いて以下の作業を行い、傾き  $ah_w$  及び分散  $sh_w$  を算出する。

1. 集合  $Fh_R$  より、最小二乗法により式 (1) の回帰直線における傾き  $ah_w$  を算出

$$Th_{(u)} = ah_w u + bh_w \quad \dots (1)$$

ただし、 $t-(w-1) \leq u \leq t$

2. 式 (1) を用いて集合  $Fh_T = \{Th_{(t-k)} \mid 0 \leq k \leq w-1\}$  を形成し、式 (2) を用いて分散  $sh_w$  を算出

$$sh_w = \frac{1}{w} \sum_{k=0}^{w-1} (Dh_{(t-k)} - \overline{Dh_w})^2 \quad \dots (2)$$

$$\text{ただし、} Dh_{(u)} = |Rh_{(u)} - Th_{(u)}|, \quad \overline{Dh_w} = \frac{1}{w} \sum_{k=0}^{w-1} Dh_{(t-k)}$$

続いて、開場日  $t$  の日足・安値を  $Rl_t$  とした時の安値集合  $Fl_R = \{Rl_{(t-k)} \mid 0 \leq k \leq w-1\}$  を用いて、上記と同様に  $al_w$  及び  $sl_w$  を算出する。いずれの値も、NN への入力データとするために 256 段階に離散化した後、8bit に符号化している。

### 2.4. 学習セット群とエージェント

NN による学習済みエージェントは、当日  $t$  の高値と安値を開場前に予測する。学習に必要なデータは前日  $t-1$  までの市況から作成する必要がある、その最も大きな集合を学習セット群 (learning data sets) と名付けた。過去の市況において、学習セット群として提供可能な教師信号の時間的限界は  $t-1$  である。そこで、 $t-(a+w)$  から  $t-a$  までの市況 (ただし  $a \geq 2$ ) を NN に入力した際の出力を  $t-(a-1)$  の予測市況とし、 $t-(a-1)$  の実際の市況を教師信号 (teacher signal) と定めた。 $t-(a-1)$  における予測値と実際値を比較し、両者の誤差を逆伝播させることで NN の構成変化を促す行為が、エージェントによる学習の実態となる。また本システムによる予測は、学習が終了したエージェントに  $t-(1+w)$  から  $t-1$  までの市況をテストデータ (test data) とし

て入力し、 $t$ の予測市況を導出する作業となる。

学習セット群作として、学習セット作成の為に遡る日数の大きさに応じ、短期 (short) 潜在記憶  $W_{D10}$ 、中期 (middle) 潜在記憶  $W_{D20}$ 、長期 (long) 潜在記憶  $W_{D30}$  の3種を用意した。エージェントも3個用意し、個々のエージェントは自身が受け持つ学習セット群1種のみで学習を行い、予測も独立して行う。

## 2.5. 潜在記憶に対応した学習セット群の構成

例として、短期潜在記憶  $W_{D10}$  に基づいてエージェントが TOPIX・2014年12月30日の予測を行う際に用いた学習セット群を表1に表す。縦に  $t-2$  から  $t-11$  まで10の学習セット (learning set) が並び、横には学習セットを構成する5組の学習パターン (learning pattern) と1組の教師信号が並ぶ。

12月30日市況予測に利用可能なデータは、12月29日までの市況に限られる。そこで、直近10開場日 (12月15日から29日) 個々において、当該日の市況を教師信号とする学習セットを規則に基づき用意する。つまりエージェントは、教師信号の日付が異なる10個の学習セットを一つの群として扱い、各学習セットを順不同かつ繰り返し入力することで適切な予測手法の学習に励み、12月30日の市況予測に備える。本実験では学習必要時間や学習性能を考慮し、群を形成する学習セット数を10と定めている。

学習セットの構成要素である学習パターンは、高値と安値各々の傾きと分散  $ah$ ,  $sh$ ,  $al$ ,  $sl$  の計4個の学習データ ( $8\text{bit} \times 4 = 32\text{bit}$ ) で構成されるが、算出には前述の潜在記憶が大きく影響する。例えば  $W_{D10}$  では、 $w \in \{10, 20, 30, 40, 50\}$  における高値集合  $Fh_R$  及び安値集合  $Fl_R$  を用い、5組の学習パターンが用意される。従って、 $W_{D10}$  における各学習パターンは、2.3. に従い算出した入力データ集合  $IN = \{ah_{10}, sh_{10}, al_{10}, sl_{10}, ah_{20}, sh_{20}, al_{20}, sl_{20}, ah_{30}, sh_{30}, al_{30}, sl_{30}, ah_{40}, sh_{40}, al_{40}, sl_{40}, ah_{50}, sh_{50}, al_{50}, sl_{50}\}$  を構成する。そのデータ長は合計で160bit ( $32\text{bit} \times 5$ ) となる。具体的には、表1の学習セット  $t-2$  における  $w=10$  の「12月12日から26日まで」は、この期間内に存在する10開場日に記録した高値と安値を用いた、1組の学習パターン生成を表す。同様に  $w=50$  では、10月15日から12月26日までの50開場日から1組の学習パターンを生成する。中期潜在記憶  $W_{D20}$  では  $w \in \{20, 40, 60, 80, 100\}$  となり、 $w=20$  から  $w=100$  の最大100開場日前、長期潜在記憶  $W_{D30}$  では  $w \in \{30, 60, 90, 120, 150\}$  で最大150開場日前までの潜在記憶を内包した学習パターンをもって  $IN$  集合を形成する。

学習セットのもう一方の構成要素である教師信号は、1組の学習パターンが表す回帰直線と実際値の符号付距離を256段階に離散化した後、8bitに符号化する。高値と安値それぞれに5組の学習パターンが存在する為、学習セット1個毎に10の教師信号 ( $8\text{bit} \times 10 = 80\text{bit}$ ) が用意される。

## 2.6. エージェントによる予測値の導出

表1に従うと、 $t=12$ 月30日の高値と安値を予測する短期潜在記憶担当エージェントは、前述の通り  $t-10$  の12月15日から  $t-1$  の12月29日を教師信号とする10個の学習セットを順不同で学習する。学習が終了した時点で当該エージェントは、前日  $t-1$  までの市況から当日  $t$  の市況を予測する能力を身に付けていると期待する。

表1の学習セット群を用いた学習が終了したNNに、同じく表1のTest Dataに記された期間の市況から2.3.に従って算出した傾きと分散を入力すれば、エージェントは12月30日の高値と安値それぞれの予測値となる8bitデータを、学習パターン毎に5組・計10個出力する。そこで、教師信号作成手順を逆に辿ることで学習パターンに対応する回帰直線との符号付距離を求め、当該入力データによる予測値としている。ただし、この手順により算出される予測値は、離散化の影響によって常に一定の幅を持つ為、本実験ではその中央値を予測値と定めた。更に高値と安値それぞれで、出力された5組を相加平均した値を当該エージェントの予測値（高値及び安値）と定める。

表1  $t=2014/12/30$ における、短期潜在記憶  $W_{D10}$  の学習セット群及びテストデータ

Learning Set	Learning Patterns						Teacher Signal
	$w=10$	$w=20$	$w=30$	$w=40$	$w=50$	to	
	from	from	from	from	from		
$t-2$	12-Dec	28-Nov	13-Nov	29-Oct	15-Oct	26-Dec	29-Dec
$t-3$	11-Dec	27-Nov	12-Nov	28-Oct	14-Oct	25-Dec	26-Dec
$t-4$	10-Dec	26-Nov	11-Nov	27-Oct	10-Oct	24-Dec	25-Dec
$t-5$	9-Dec	25-Nov	10-Nov	24-Oct	9-Oct	22-Dec	24-Dec
$t-6$	8-Dec	21-Nov	6-Nov	23-Oct	8-Oct	19-Dec	22-Dec
$t-7$	5-Dec	20-Nov	7-Nov	22-Oct	7-Oct	18-Dec	19-Dec
$t-8$	4-Dec	19-Nov	5-Nov	21-Oct	6-Oct	17-Dec	18-Dec
$t-9$	3-Dec	18-Nov	4-Nov	20-Oct	3-Oct	16-Dec	17-Dec
$t-10$	2-Dec	17-Nov	31-Oct	17-Oct	2-Oct	15-Dec	16-Dec
$t-11$	1-Dec	14-Nov	30-Oct	16-Oct	1-Oct	12-Dec	15-Dec
Test Data	15-Dec	1-Dec	14-Nov	30-Oct	16-Oct	29-Dec	

## 3. 市況予測実験の結果

本研究では、2005年1月1日から2014年12月31日の10年間において予測値を導出した。同期間におけるTOPIXとS&P500の終値・時系列変動を図1に示す。両指数の2005年最初の算出値は、TOPIX=1153.38、S&P500=1208.08と似通っている。そこで作図に伴う数値変換は行わず、指数をそのまま縦軸に用いている。また、各指数の算出（開場）状況を表2に示す。

市場の開場に伴い指標算出された日数を Open に、閉場により指標算出がない日数を Close に表す。

エージェントによる予測値と実際値の絶対誤差を実際値に対する百分率で表したものを予測誤差（以下 Error）と定義する。表3は TOPIX・高値の、表4は S&P500・高値の Error 状況一覧である。縦に、Error の最大・最小・平均（Average of Errors：以下 AoE）・標準偏差（以下 SD）及び Error 分布状況を示す。なお、SD は百分率を単位として算出している。横には、潜在記憶の長さが異なる3つのエージェント（短期、中期、長期それぞれの潜在記憶を学習セット群として学習を行った NN）による予測成績を配している。加えて、3 エージェントの出力を相加平均する Equalize 戦略を採用した場合と、場が閉じた後に初めて判明する、3 個のエージェントの中で実際値に最も近い値を予測していたエージェントを開場日単位で選出した最善選択（Best）の成績を記す。Best は市場結果が出そろった後に最良な予測を行ったエージェントを選出しているため、本システムによる予測の理論上限となる。

表5は、TOPIX の予測導出における各エージェントの NN 学習状況と AoE の関係を、表6は S&P500・高値における両者の関係を示している。興味深いのは、いずれにおいても、UEA が0.05を下回った場合と比較し学習試行回数が15万回に達して学習を打ち切った場合の方が、統計的に AoE が低くなっている点である。前日までのデータによる学習への適応度が低いエージェントの方が、高い場合と比べ実際値に近い予測を行う傾向が示されている。この傾向は、TOPIX 及び S&P500 の安値でも確認された。

表7に、本研究で実施した予測全ての成績を示す。これら数値が示す意味を考える目的で、同一期間内の TOPIX 日足・値振動率 = (高値 - 安値) / 始値を百分率で求めると、最大 12.71%・最小 0.21%・平均 1.29%・標準偏差 0.90 であった。また、S&P500 のそれは最大 11.22%・最小 0.09%・平均 1.34%・標準偏差 1.11 であった。最も成績の良い Equalize 戦略を採用しても、予測誤差の規模は1日の値振幅に近いため、本システムの予測精度は低く、予測システムとして利用することは困難である。

一方で、学習セット群における潜在記憶の長短は予測成績に大きな差異を与えていない。更に Equalize 戦略はいずれのエージェントよりも成績が優れており、Best に最も近い。この傾向はいずれの指標の高値と安値で等しく認められる。又、高値・安値共に S&P500 の方が TOPIX より成績が良く、一方で指標に関係なく安値より高値の予測が得意という本システムの特徴も明らかとなった。

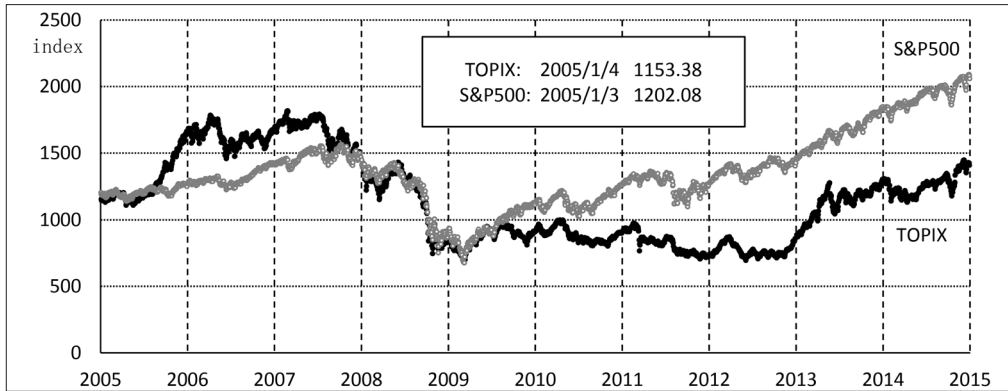


図1 2005年1月1日から2014年12月31日における、TOPIXとS&P500の指標・終値の時系列変動

表2 2005年1月1日から2014年12月31日における、TOPIXとS&P500の指標算出日の状況

from 2005/1/1 to 2014/12/31		S&P500		
		Open	Close	Total
TOPIX	Open	2372 (65.0%)	81 (2.2%)	2453 (67.2%)
	Close	145 (4.0%)	1054 (28.9%)	1199 (32.8%)
	Total	2517 (68.9%)	1135 (31.1%)	3652 (100%)

表3 2005年1月1日から2014年12月31日における TOPIX・高値の予測成績

予測成績	潜在記憶			Equalize 戦略	Best
	短期	中期	長期		
Error 最大	12.29%	12.76%	12.70%	10.95%	8.99%
Error 最小	0.00%	0.00%	0.00%	0.00%	0.00%
AoE	1.53%	1.56%	1.53%	1.34%	0.81%
Error の SD	1.49	1.57	1.53	1.29	0.97
Number of Days					
Error < 1%	1104 (45.0%)	1139 (46.4%)	1184 (48.3%)	1239 (50.5%)	1800 (73.4%)
1% ≤ Error < 2%	708 (28.9%)	655 (26.7%)	644 (26.3%)	681 (27.8%)	436 (17.8%)
2% ≤ Error < 3%	336 (13.7%)	313 (12.8%)	310 (12.6%)	298 (12.1%)	139 (5.7%)
3% ≤ Error < 4%	142 (5.8%)	159 (6.5%)	135 (5.5%)	130 (5.3%)	39 (1.6%)
4% ≤ Error < 5%	81 (3.3%)	93 (3.8%)	83 (3.4%)	55 (2.2%)	17 (0.7%)
5% ≤ Error < 6%	36 (1.5%)	41 (1.7%)	44 (1.8%)	27 (1.1%)	10 (0.4%)
6% ≤ Error < 7%	22 (0.9%)	22 (0.9%)	24 (1.0%)	8 (0.3%)	6 (0.2%)
7% ≤ Error < 8%	8 (0.3%)	12 (0.5%)	15 (0.6%)	7 (0.3%)	2 (0.1%)
8% ≤ Error < 9%	7 (0.3%)	5 (0.2%)	6 (0.2%)	4 (0.2%)	4 (0.2%)
9% ≤ Error < 10%	3 (0.1%)	9 (0.4%)	2 (0.1%)	0	0
10% ≤ Error	6 (0.2%)	5 (0.2%)	6 (0.2%)	4 (0.2%)	0
Total	2453				

表4 2005年1月1日から2014年12月31日におけるS&amp;P500・高値の予測成績

予測成績	潜在記憶			Equalize 戦略	Best
	短期	中期	長期		
Error 最大	12.69%	12.45%	10.87%	8.69%	6.22%
Error 最小	0.00%	0.00%	0.00%	0.00%	0.00%
AoE	1.10%	1.10%	1.15%	0.99%	0.60%
Error の SD	1.23	1.24	1.26	1.07	0.75
Number of Days					
Error < 1%	1570 (62.4%)	1576 (62.6%)	1532 (60.9%)	1659 (65.9%)	2053 (81.6%)
1% ≤ Error < 2%	553 (22.0%)	546 (21.7%)	541 (21.5%)	540 (21.5%)	324 (12.9%)
2% ≤ Error < 3%	225 (8.9%)	217 (8.6%)	256 (10.2%)	183 (7.3%)	94 (3.7%)
3% ≤ Error < 4%	89 (3.5%)	97 (3.9%)	88 (3.5%)	77 (3.1%)	27 (1.1%)
4% ≤ Error < 5%	36 (1.4%)	40 (1.6%)	48 (1.9%)	26 (1.0%)	10 (0.4%)
5% ≤ Error < 6%	20 (0.8%)	15 (0.6%)	27 (1.1%)	14 (0.6%)	7 (0.3%)
6% ≤ Error < 7%	9 (0.4%)	9 (0.4%)	14 (0.6%)	11 (0.4%)	2 (0.1%)
7% ≤ Error < 8%	6 (0.2%)	10 (0.4%)	2 (0.1%)	3 (0.1%)	0
8% ≤ Error < 9%	2 (0.1%)	3 (0.1%)	4 (0.2%)	4 (0.2%)	0
9% ≤ Error < 10%	4 (0.2%)	0	3 (0.1%)	0	0
10% ≤ Error	3 (0.1%)	4 (0.2%)	2 (0.1%)	0	0
Total	2517				

表5 TOPIX・高値の予測成績とNNの学習状況

TOPIX・高値		Total	Unit Error Average < 0.05	Stop at 150000 times
Number of Days	短期潜在記憶	2453	2051 (83.6%)	402 (16.4%)
	中期潜在記憶		1467 (59.8%)	986 (40.2%)
	長期潜在記憶		987 (40.2%)	1466 (59.8%)
AoE	短期潜在記憶	1.53%	1.58%	1.29%
	中期潜在記憶	1.56%	1.57%	1.55%
	長期潜在記憶	1.52%	1.57%	1.48%

表6 S&amp;P500・高値の予測成績とNNの学習状況

S&P500・高値		Total	Unit Error Average < 0.05	Stop at 150000 times
Number of Days	短期潜在記憶	2517	1955 (77.7%)	562 (22.3%)
	中期潜在記憶		991 (39.4%)	1526 (60.6%)
	長期潜在記憶		603 (24.0%)	1914 (76.0%)
AoE	短期潜在記憶	1.10%	1.21%	0.71%
	中期潜在記憶	1.10%	1.37%	0.92%
	長期潜在記憶	1.15%	1.61%	1.00%



表7 TOPIX と S&amp;P500 の予測成績

			潜在記憶			Equalize 戦略	Best
			短期	中期	長期		
TOPIX	高値	AoE	1.53%	1.56%	1.52%	1.34%	0.81%
		SD	1.49	1.57	1.53	1.29	0.97
	安値	AoE	1.72%	1.63%	1.76%	1.51%	0.92%
		SD	1.82	1.81	1.87	1.57	1.21
S&P500	高値	AoE	1.10%	1.10%	1.15%	0.99%	0.60%
		SD	1.23	1.24	1.26	1.07	0.75
	安値	AoE	1.32%	1.29%	1.36%	1.15%	0.70%
		SD	1.50	1.51	1.54	1.31	0.95

## 4. 結果の検証及び有効性の考察

### 4.1. 結果の検証

2005年から2014年のTOPIX及びS&P500における、Equalizeの月単位AoE変動を図2と図3に示す。2008年10月には、TOPIX、S&P500共に本システムによる予測精度が大きく低下している。これは米国における住宅バブルが2006年にピークを迎え、2007年のサブプライム住宅ローン危機顕在化、2008年9月15日のLehman Brothers Holdings Inc. 破綻という一連の世界的金融危機の状況と符合する。世界規模での経済的混乱によって米国・東京の両市場で起った予測精度の低下は、時間の経過と共に落ち着きを取戻し、2010年末には両市場とも予測精度が最高水準まで向上する。しかし2011年に入って世界規模での資源価格高騰が問題視され、例えば原油価格は1バレル100ドル以上となった。このような経済情勢が市場参加者の先行き不透明感を誘い、同年後半における両者の予測精度悪化に繋がったと考えられる。

又、TOPIXとS&P500の2005年から2011年のAoE変化は、東日本大震災が発生した2011年3月を除き似通っているが、2012年12月以降は様相が異なる。その原因として、同年12月26日に発足した第2次安倍内閣が打ち出した日本独自の経済政策が想定される。政府による新たな経済政策の導入が、東京市場に特に大きな影響を与えたと推測できる。

市場の形成メカニズム解明においては、市場心理という定量化の難しい存在を常に考慮する必要がある。市場心理によって市場は時に実体経済を超越し、バブルや恐慌を生み出すと考えられている。しかし、市場心理をマクロ的かつ定量的に測量する枠組みの提示はほとんど成されていない。そこで、本システムを活用した市場心理分析手法について提案する。

ただし、ここでも時系列分析における単位根問題に直面する。本研究では対象を日本及び米国株式市場参加者と定めている。従って、人為的単位根ではあるがこれら地域で一般的かつ生活習慣に根付いた時間区分であるグレゴリオ暦を基軸として、本システムの活用手法を以下に提示する。

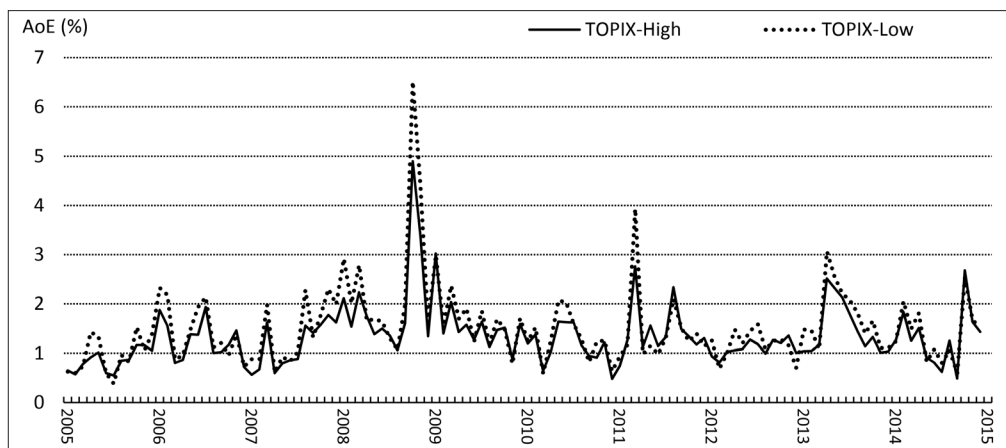


図2 TOPIX・Equalizeの月単位Average of Errors変動

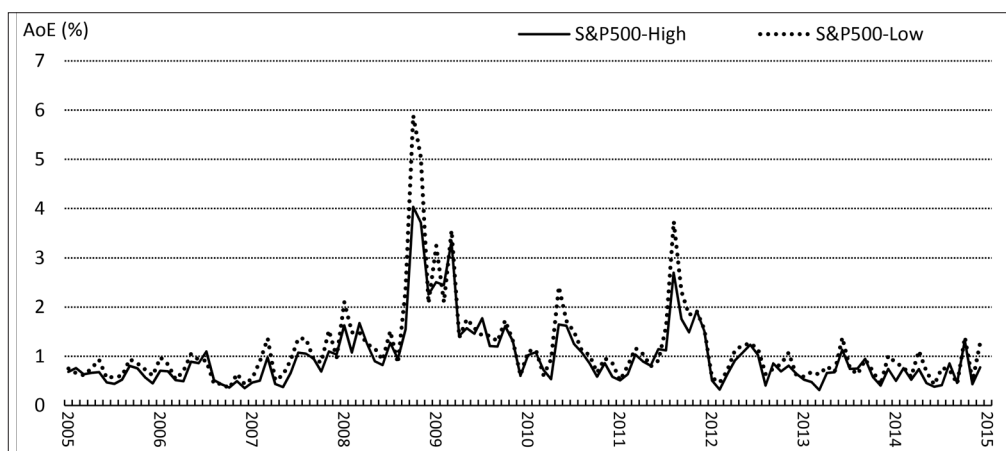


図3 S&P500・Equalizeの月単位Average of Errors変動

#### 4.2. システム活用手法の提案

3. に示した通り、統計的に予測精度が最も高いEqualize戦略のAoEは1日の値振幅に近い規模である為、本システムを予測システムとして利用するのは困難である。しかし、事後に確定するBestのAoEは1日の値振動を十分下回っている。そこで、短期・中期・長期潜在記憶の予測値と実際値を比較し、一定期間内でのBest選出状況から市場心理を定量的に議論する手法を提案する。

表8に、2005年から2014年のTOPIX及びS&P500・高値と安値予測においてBestに選出されたエージェントの割合を年単位で示す。TOPIXとS&P500を比較した場合、高値では選出比率に明確な違いが認められる。TOPIXでは、長期潜在記憶による学習結果がBestに選出される年は2005年から2014年の10年間の6割を占めるが、S&P500では短期潜在記憶の選出年

が6割となっている。高値は、市場参加者にとって一つの心理的障壁と考えることができる。その時点での経済的・社会的情勢を鑑みた市場参加者は、心理的上限としての高値（心理高値上限）を設定するのが一般的である。指標が高値設定値に接近又は超過すると当該市場参加者は高値警戒感を抱き、従来とは異なる行動規範の適用について検討を始める。本結果は「東京証券取引所の市場参加者は、長い期間の移動平均線を意識して心理高値上限を設定する傾向にあるが、ニューヨーク証券取引所等の市場参加者は、短い期間の移動平均線を意識して設定する傾向にある」との可能性を示す。

一方の安値では、両指標とも中期潜在記憶選出年が多数となっている（7割と6割）点で似た傾向を示している。高値と同様に心理的障壁としての安値が存在すると想定できる。よって、「市場参加者は参加する市場に関係なく、心理安値下限の設定に際し中程度の期間に設定した移動平均線を意識する傾向にある」可能性が生まれる。

図2及び図3からも明らかな通り、本システムによる予測精度が低くなった原因の一つは、2007年から2009年に顕著な「高いAoEの持続」である。そこで、2007年から2009年のTOPIX・高値及び安値におけるBest選出比率を月単位で示したグラフを図4及び図5に、S&P500のそれを図6及び図7に示す。

図4から図7を比べると、同一指標の高値と安値間にBest選出比率の連動は認められない。異なる指標の高値同士あるいは安値同士にも、連動は見出せない。しかし個々のグラフを見ると、心理高値上限及び心理安値下限設定時に意識する移動平均線の期間には周期性が見て取れる。例えば、図6のS&P500・高値では、2007年8月以降2009年7月までは短期が主たる選出元となる。しかし2008年9月から11月は短期と長期の選出が拮抗し、12月には約6割を短期が占める。2009年1月は長期が4割、2月は短期が4割強と揺れ動いた後、3月は中期6割、4月は中期4割となり、以降は徐々に落ち着きを取戻し緩やかな循環へと市場心理が移行する様が見て取れる。図7の安値では2008年に入ると長期の選出が目立つようになるが、9月には3者が拮抗する状態となり、10・11月と短期が優勢となる。その後は長期から中期・短期・中期という緩やかな循環に移行し、市場心理は落ち着きを取戻す。

このように、期間の異なる3種の潜在記憶を各々学習したエージェント群の予測値を実際値と比較することにより、設定した単位根における市場心理の定量的議論が可能となる。同時に、市場心理の変化に周期性又は規則性が存在する可能性も認められる。本システムを活用すれば、市場心理の時系列変化を数値化できる。これは、市場心理というファンダメンタルを組み込んだ、株価指数時系列分析に有用なモデル構築に道を開く。

表 8 2005 年から 2014 年の TOPIX 及び S&P500・Best に  
おける年単位での潜在記憶選択比率 (%)

指標	西暦	高値			安値		
		短期	中期	長期	短期	中期	長期
TOPIX	2005	28.6	<b>38.8</b>	32.7	31.8	31.0	<b>37.1</b>
	2006	<b>36.7</b>	33.1	30.2	31.9	<b>34.3</b>	33.9
	2007	29.8	30.2	<b>40.0</b>	32.2	<b>38.0</b>	29.8
	2008	32.7	29.4	<b>38.0</b>	33.1	32.2	<b>34.7</b>
	2009	<b>36.2</b>	29.6	34.2	33.3	<b>35.8</b>	30.9
	2010	36.3	26.1	<b>37.6</b>	<b>36.3</b>	32.2	31.4
	2011	31.8	31.0	<b>37.1</b>	28.2	<b>38.4</b>	33.5
	2012	34.3	<b>35.1</b>	30.6	33.5	<b>34.7</b>	31.9
	2013	33.5	32.2	<b>34.3</b>	31.8	<b>38.4</b>	29.8
	2014	32.4	31.6	<b>36.1</b>	34.4	<b>35.7</b>	29.9
S&P500	2005	33.3	33.3	33.3	<b>36.1</b>	34.1	29.8
	2006	<b>33.9</b>	<b>33.9</b>	32.3	33.5	31.9	<b>34.7</b>
	2007	34.3	<b>34.7</b>	31.1	31.5	<b>34.7</b>	33.9
	2008	<b>39.9</b>	28.5	31.6	32.0	28.5	<b>39.5</b>
	2009	32.1	<b>35.7</b>	32.1	29.0	<b>39.3</b>	31.7
	2010	<b>36.5</b>	31.3	32.1	33.3	31.3	<b>35.3</b>
	2011	<b>34.5</b>	31.3	34.1	31.0	<b>34.5</b>	<b>34.5</b>
	2012	<b>35.6</b>	33.2	31.2	31.6	<b>36.0</b>	32.4
	2013	33.7	<b>34.1</b>	32.1	34.5	<b>35.7</b>	29.8
	2014	<b>37.3</b>	29.8	32.9	32.5	<b>36.1</b>	31.3

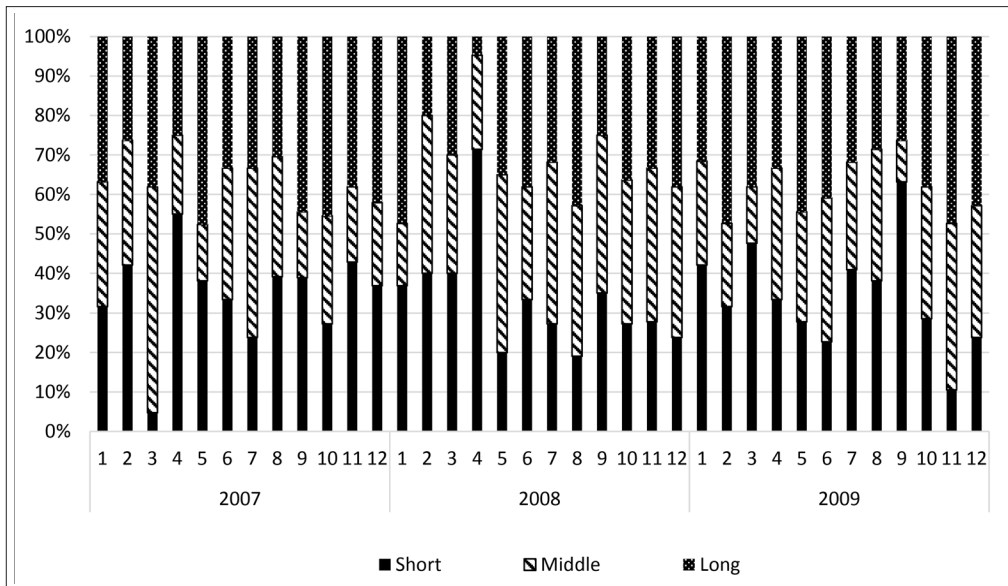


図4 2007年から2009年のTOPIX高値・Bestにおける、月単位・潜在記憶選択比率

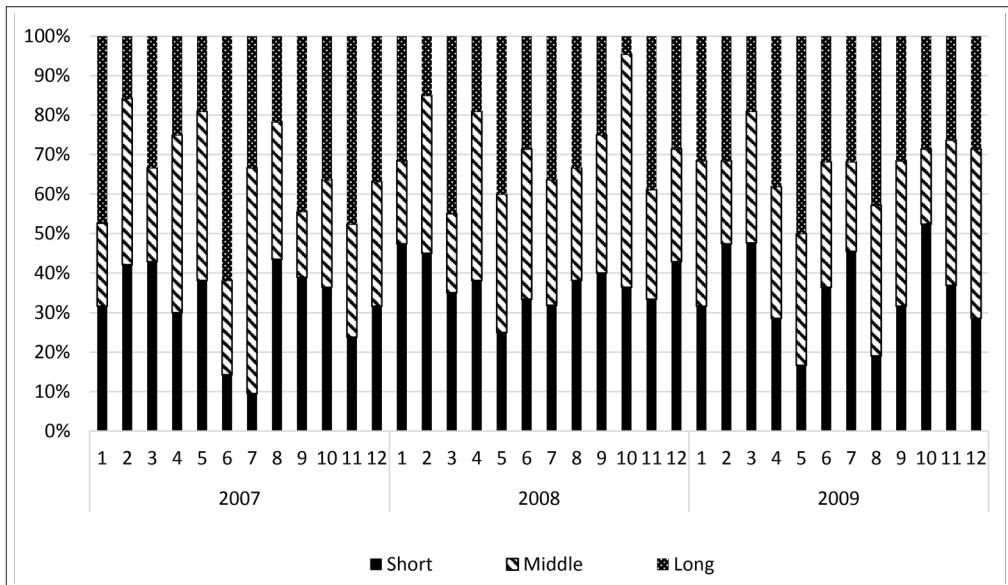


図5 2007年から2009年のTOPIX安値・Bestにおける、月単位・潜在記憶選択比率

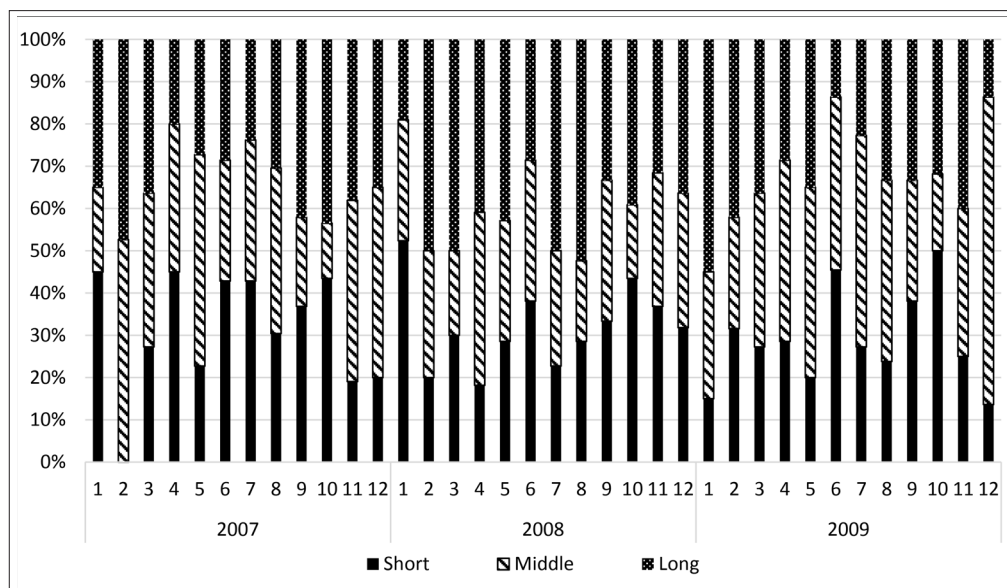


図6 2007年から2009年のS&P500高値・Bestにおける、月単位・潜在記憶選択比率

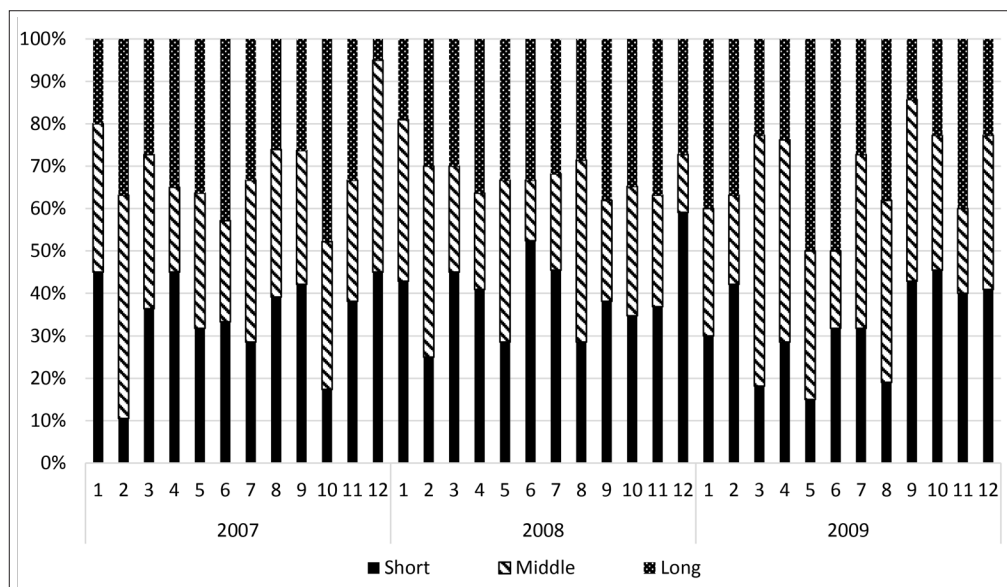


図7 2007年から2009年のS&P500安値・Bestにおける、月単位・潜在記憶選択比率

## 5. むすび

本研究ではBP法を用いた3階層NNをフィルタとして活用し、株価指数時系列に潜み自己複製をもたらす要素の可視化を試みた。具体的には、NNを備えたエージェントを対象に、一

定区間の株価指数より算出したパラメータで構成する学習セット群による学習を行い、振舞いを記憶したエージェントを複数出現させた。個々のエージェントは、起点は同じであるが区間の長さがそれぞれ異なるデータより算出した学習セット群による学習を独立して行う。1個のエージェントが行う予測の精度は実用的で無いが、各予測を相互比較し学習セット群と精度の相関に着目することで、その時点の市場心理を数値表現する手法を提案した。

テクニカル分析によって経済時系列の階層的クラスタ構造を自律的に導出する本手法は自己符号化器そのものであり、深層学習に向けた有効な事前学習と見做すことができる<sup>27)</sup>。そしてこのような自己符号化器を積層することにより、精度の高い経済時系列予測が可能となる NN 形成の可能性が認められる<sup>28)</sup>。時系列の区間を動的に区切って回帰分析を繰り返す本手法は、ARCH モデルによる分析又はサポートベクトル回帰分析に近似する。従って、現在の株式市場の状態を市場心理の側面から可視化し、市場参加者の株式売買に関する意思決定を支援するサポートベクトルマシンの構築も想定できる<sup>29)</sup>。株式市場をマルコフ連鎖と捉えるファンダメンタルズ分析では、株価指標時系列は本手法により可視化される市場心理を反応物質の一つとする、反応拡散方程式系だと仮定できる。その場合、市場心理の濃度が遷移ダイナミクスを左右する<sup>30)</sup>。そして、自己複製と自己崩壊を繰り返す時空カオス<sup>31)</sup>が生み出されていると考えられる。

本システムの優位性は、心理という定式化が困難な情報を定量的に扱うことが可能となった点にある。出力値を統計処理することで従来手法では解明が困難であった、隠れマルコフ連鎖を発見できる可能性が生まれた。本研究が明らかにした潜在記憶というパラメータの妥当性を引き続き検証すると同時に、本システムを改良・発展させることで、株式市場を適切に表現可能なモデルを構築するのが今後の課題である。

### 【参考文献】

- 1) 寺崎 健, 池口 徹, 合原一幸 [他], 田中 智, “経済時系列データの決定論的非線形ダイナミカル特性に関する解析”, 信学論 (A), vol.J78-A, no.12, pp.1601-1617, Dec. 1995.
- 2) 黒住英司, “経済時系列分析と単位根検定: これまでの発展と今後の展望”, 日本統計学会誌, vol.38, no.1, pp.39-57, Sep. 2008.
- 3) 浅野敬志, 大坪史尚, 天白隼也, “株価水準を評価するうえで有用な利益情報: 公正価値情報や減損損失はノイズなのか”, 金融研究, vol.35, no.1, pp.31-69, Jan. 2016.
- 4) 佐野一雄, “真のファンダメンタルズ仮説—株式市場におけるバイアスの検出—”, 行動経済学, vol.8, pp.66-72, Nov. 2015.
- 5) 長島知正, 永井喜則, 荻原利彦, 土屋 尚, “時系列データ解析とカオス”, 計測と制御, vol.29, no.9, pp.839-846, Sep. 1990.
- 6) 回測純治, 宮崎浩一, 岡本雅生, “5パラメータ・ローカルボラティリティ・モデルのオプション市場価格再現力と株価予測力”, 情処学論: 数理モデル化と応用, vol.2, no.2, pp.58-69, Apr. 2009.
- 7) E.F. Fama, “Efficient Capital Markets: A Review of Theory and Empirical Work”, Journal of Finance, vol.25, no.2, pp.383-417, May 1970.

- 8) 前田廉孝, “経済史研究における計量分析の方法と課題—効率的市場仮説をめぐる分析を中心に—”, 西南学院大学経済学論集, vol.49, no.2・3, pp.169-190, Dec. 2014.
- 9) H.A. Simon, “From substantive to procedural rationality”, in Latsis, Method and Appraisal in Economics, Cambridge University Press, pp.129-148, Cambridge, 1976.
- 10) 福田収一, “満足化設計: 限定合理性, 人工知能, 感情”, 設計工学・システム部門講演会講演論文集, vol.2012, no.22, pp.“2316-1”-“2316-5”, Sep. 2012.
- 11) B. Mandelbrot, “The Variation of Certain Speculative Prices”, The Journal of Business, vol.36, no.4, pp.394-419, Oct. 1963.
- 12) 岩田英朗, 芦田昌也, 竹内昭浩, “株価変動に見るフラクタル性の研究”, 経済理論, vol.343, pp.25-51, Mar. 2008.
- 13) 時永神三, 森保 洋, 宮崎明雄 [他], 島津宣之, “フラクタル的な性質をもつ時系列の一予測手法とその応用”, 信学論 (A), vol.J79-A, no.11, pp.1793-1800, Nov. 1996.
- 14) 時永祥三, 森保 洋, 宮崎明雄 [他], 島津宣之, “スケール伸長変換およびウェーブレット変換によるパラメータ推定を用いたフラクタル時系列の予測”, 信学論 (A), vol.J79-A, no.12, pp.2054-2062, Dec. 1996.
- 15) 松本義之, 和多田淳三, “関連データを同時に埋め込んだカオスによる短期予測に関する研究”, 日本経営工学会論文誌, vol.49, no.4, pp.209-217, Oct. 1998.
- 16) 肥後雅博, 中田 (黒田) 祥子, “経済変数から基調的変動を抽出する時系列的手法について”, 金融研究, vol.17, no.6, pp.39-97, Dec. 1998.
- 17) 池田欽一, 時永祥三, “フラクタル時系列の予測手法を用いた株価予測とその応用”, Journal of the Operations Research Society of Japan, vol.42, no.1, pp.18-31, Mar. 1999.
- 18) 渡部敏明, 佐々木浩二, “ARCH 型モデルと “Realized Volatility” によるボラティリティ予測とバリュエーション・リスク”, 金融研究, vol.25, pp.39-74, Mar. 2006.
- 19) 下川哲矢, 参沢匡将, 渡邊恭子, “エージェントモデルを用いた情報伝達のモデル化と株価の予測可能性との関係”, 人工知能誌, vol.21, no.4, pp.340-349, Nov. 2006.
- 20) 広林茂樹, “1990年以降の日経平均株価における Non-Harmonic Analysis を用いた予測法の検証”, 信学論 (A), vol.J91-A, no.1, pp.153-161, Jan. 2008.
- 21) 八木 勲, 水田孝信, 和泉 潔, “人工市場を用いた市場暴落後における反発メカニズムの分析”, 情処学誌, vol.53, no.11, pp.2388-2398, Nov. 2012.
- 22) 鳥海不二夫, 石井健一郎, “人工市場を用いた予測市場の予測メカニズムの分析”, 人工知能誌, vol.27, no.6, pp.346-354, Nov. 2012.
- 23) 高橋弘志, 大町美歩, 松葉育雄, “経済時系列の時空間分布と一般化エントロピー”, 信学技報 NLP, vol.101, no.614, pp.1-6, Jan. 2002.
- 24) 関本信太郎, 森 康久仁, 松葉育雄, “Granger 理論による経済データの長期記憶性の分析”, 信学技報 NLP, vol.110, no.122, pp.81-84, Jul. 2010.
- 25) L.F. Burlaga and L.W. Klein, “Fractal structure of the interplanetary magnetic field”, Journal of Geophysical Research, vol.91, Iss.A11, pp.347-350, Jan. 1986.
- 26) L.F. Burlaga and L.W. Klein, “Large-scale fluctuations in the interplanetary medium”, Journal of Geophysical Research, vol.92, Iss.A21, pp.1261-1266, Feb. 1987.
- 27) 麻生英樹, “多層ニューラルネットワークによる深層表現の学習”, 人工知能誌, vol.28, no.4, pp.649-659, Jul. 2013.
- 28) Y. Bengio, P. Lamblin and D. Popovici, “Greedy Layer-Wise Training of Deep Networks”, Advances in Neural Information Processing Systems 19, pp.153-160, Dec. 2006.
- 29) 山口暢彦, “教師あり学習を用いたGTMによるデータの可視化”, 信学技報 NC, vol.113, no.500, pp.53-58, Mar. 2014.
- 30) 三村昌泰, 上山大信, 西浦廉政, 永山雅晴, 柴 伸一郎, パターン形成とダイナミクス, 三村昌泰 (編), (財) 東京大学出版会, 東京, 2006.



- 31) Y. Nishiura, D. Ueyama, "Spatio-temporal chaos for the Gray-Scott model", *Physica D: Nonlinear Phenomena*, vol.150, Iss. 3-4, pp.137-162, Apr. 2001.

## Visualization of the Market Sentiment in Stock Markets based on Neural Network

Hideaki IWATA

### Abstract

In this study, we tried to visualize the market sentiment that affects the judgment and behavior of stock market participants. Its purpose is to clarify the periodicity contained in the time series data of TOPIX and S&P500. The trend and volatility for a given period of time were calculated for price fluctuations from 2005 to 2014 and we used as training data for the neural network. We prepared three types of learning data: short, medium, and long term, according to the length of the period over which we calculate the trend and volatility. An agent with a neural network learns using one of these three types of training data, and each predicted a near-future price after the training was completed. Therefore, we used three agents in this study. By comparing the predictive results of these three agents, we attempted to visualize market sentiment.