

ニューラルネットワークを活用した、 経済指標の時系列変動に現れるフラクタル性の解明と理解

岩 田 英 朗

1. はじめに

公開市場における株券売買価格は、当該企業が有する多様な社会的・経済的価値を反映する。従って、ある市場に上場する複数の企業株価より算出される株価指標は、市場が属する国や地域がその時点において保持する、包括的な経済および社会環境を表していると考えられている。株価指標の時系列変動は一見しただけでは規則性を認めることが困難であるが、その変動チャートにフラクタル性(自己相似性)が認められる事実を B. B. Mandelbrot [1] は発見した。時系列変動チャートにおいてフラクタル性が常に認められるのであれば、フラクタル性を生み出している要因を探究することにより、統計的分析に基づく変動規則の導出可能性が高まるだけでなく、指標変動予測への応用も期待できる [2]。

そこで拙著 [3] では、代表的な株価指標の時系列変動が有するフラクタル構造(自己相似構造)に注目し、構造の解明およびフラクタル性をもたらす要因の考察に取り組んだ。その概要を以下に示す。

1. Burlaga と Klein が考案した、時系列データを要素とする集合におけるフラクタル次元推定手法である BK 法 [4] [5] [6] に基づき、株価指標変動チャートのフラクタル次元を推定し、フラクタル性の存在を確認。

【BK 法に基づく D_{BK} 推定方法】

- A) 時間 t に依存する n 個のデータで構成される集合 $F = \{x(t) \mid 1 \leq t \leq n\}$ において
- B) 時間 t に沿って順番かつ重複しないように要素を個単位のグループに区分
- C) 形成された $[n/k]$ 組における i 番目のグループを $\bar{x}_i(k)$ と表記 ($[]$ はガウス記号)
- D) i 番目のグループ $\bar{x}_i(k)$ の平均を \bar{x}_i と表記
- E) 式 1 を用いて時系列平均変動 $L_{BK}(k)$ を算出

$$L_{BK}(k) = \frac{1}{k} \sum_{i=1}^{[n/k]} | \overline{x_{i+1}(k)} - \overline{x_i(k)} | \quad (\text{式1})$$

F) k の値を様々に変化させた時, 式 2 および式 3 の関係を満たすならば, 集合 F はフラクタル構造を持つと判断し, そのフラクタル次元を D_{BK} と推定

$$k \propto L_{BK}(k)^{-D_{BK}} \quad (\text{式 2})$$

$$D_{BK} = -\frac{\log L_{BK}(k)}{\log k} \quad (\text{式 3})$$

2. 時間 t における株価指標を $x(t)$ とした場合, 差分 $dx(t) = x(t+dt) - x(t)$ は確率微分方程式 (式 4) により線形表現可能 [7] [8] [9] との認識に基づき, チャートが示すフラクタル性と確率微分方程式を構成する各要素の関連性をシミュレーションにより探究。

$$dx(t) = a(t)dt + s(t)dw(t) \quad (\text{式 4})$$

ただし,

$$dx(t) \equiv x(t+dt) - x(t)$$

$$dw(t) \equiv x(t+dt) - w(t)$$

$a(t)$: ドリフト項 (トレンド)

$s(t)$: 標準偏差項 (ボラティリティ)

$w(t)$: 幾何ブラウン運動 (ウィーナー過程)

3. ブラウン運動の確率密度関数を決定する標準偏差 σ の増減はフラクタル次元の推定に大きな影響を与えない反面, トレンド項やボラティリティ項はフラクタル次元の推定やフラクタル性の発現形態に強い影響を与える事実を確認。

【ブラウン運動の定義】

$B(t)$ が次の条件を満たす時, $B(t)$ は標準偏差 σ の 1 次元ブラウン運動である。

(1) $0 = t_0 < t_1 < \dots < t_n$ であれば, $B(t_1) - B(t_0), \dots, B(t_n) - B(t_{n-1})$ は独立 (独立増分)

(2) $B_t - B_s$ の分布は $t-s$ にのみ依存 (定常増分)

(3) $B(t)$ は正規分布 $N(\mu, \sigma^2 t)$ に従う

$$\text{ただし, } N(\mu, \sigma^2 t) \text{ の密度関数 } f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

(4) $t \rightarrow B^t$ は連続である

4. 以上より, 確率微分方程式を応用したフラクタル次元推定法である Slope 法 (式 5 および式 6) を提案すると共に, Sloop 法により推定したフラクタル次元 D_{SL} の有効性を D_{BK} との比較により証明。

$$k \propto L_{SL}(k)^{-D_{SL}} \quad (\text{式 5})$$

$$D_{SL} = -\frac{\log L_{SL}(k)}{\log k} \quad (\text{式 6})$$

ただし

$$\text{集合 } F = \{x(t) \mid 1 \leq t \leq n\}$$

$$i \text{ 番目のグループ } \tilde{x}_i(k) = \{x(t) \mid x((i-1)k+1) \leq t \leq x(i \cdot k)\}$$

$$\widehat{a}_1(k) = \frac{\sum_{j=1}^k j \cdot x((i-1)k+j) - \sum_{j=1}^k j \cdot \sum_{j=1}^k x((i-1)k+j)}{\sum_{j=1}^k j^2 - \left(\sum_{j=1}^k j\right)^2}$$

$$L_{SL}(k) = \sum_{i=1}^{\lceil n-1/k \rceil - 1} |\widehat{a}_1(k)|$$

$\widehat{a}_1(k)$ は i 番目のグループに属する k 個の要素を基に、最小二乗法を用いて求めた回帰直線の傾き

筆者らが提示した Sloop 法によるフラクタル次元推定の有効性は、一定範囲毎に離散化した時間軸内の $a(t)$ および $s(t)$ 変化を追跡することにより、時系列変動チャートが示すフラクタル性の理解可能性を示している。他方では、短期的な指標変動に対しては $w(t)$ の影響が大きいため、線形分析のみでは予測に実用性が認められない事実も明らかにした。

そこで本研究では、確率微分方程式によって予見される無数の未来に対し、過去の $a(t)$ および $s(t)$ 変化を用いて自己学習したニューラルネットワーク（以下、NN）[10] [11] [12] [13] を用いることにより、指標変動の近未来予見手法について考察する。

2. ニューロンモデル

NN は、神経細胞が有する生物学的特性を模した複数のユニットから構成される。1 個のユニットは図 1 に示す通り多入力 1 出力の構造を持ち、軸索突起に相当するラインにより他のユニットと結合する。個々のラインはユニット間の結合の強さを表す固有の重み（結合係数 W_1, W_2, \dots, W_n ）を有し、各々の入力ラインより刺激 X_1, X_2, \dots, X_n が伝達された際のユニット受容信号の強さ X は式 7 で表される。ユニットは共通する応答関数 f 以外にユニットの自己特性に相当するユニークな閾値 (offset) θ を有しており、式 8 で表される出力値 Y を出力する。なお、本研究では応答関数 f として式 9 に示すシグモイド関数を採用している。

$$X = \sum_{i=1}^n X_i \cdot W_i \quad (\text{式 7})$$

$$Y = f(X + \theta) \quad (\text{式 8})$$

$$f(x) = \frac{1}{1 + e^{-x}} \quad (\text{式 9})$$

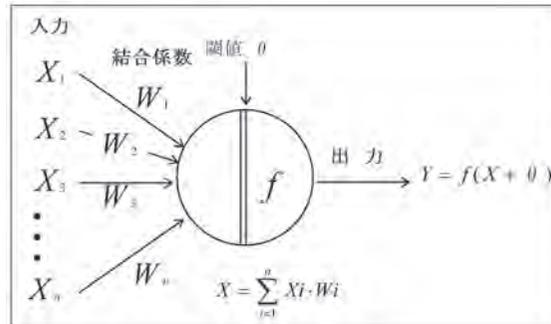


図1 ユニットの模式図

3. 誤差逆伝播法と学習プロセス

本研究では、誤差逆伝播法（バックプロパゲーション法：以下、BP）[14] [15] [16] [17]を用いた学習プロセスを有する3階層型NNモデル（図2）を活用している。本モデルでは、入力信号は入力層・中間層・出力層の順に一方向のみ伝達されるが、入力信号パターンに対応する教師信号（模範解答） T を別に用意することで、式10に示す出力層ユニットの出力値と教師信号値の最小二乗誤差総計 E を学習に利用する。つまり本モデルにおける学習とは、ラインおよびユニットの特性（ W および θ ）を変更することによって E を最小化しようとする試みである。

$$E = \frac{\sum (T_i - X_i(o))^2}{2} \quad (\text{式10})$$

$X_i(o)$ ：出力層に位置するユニット i の出力

T_i ：ユニット i の教師信号

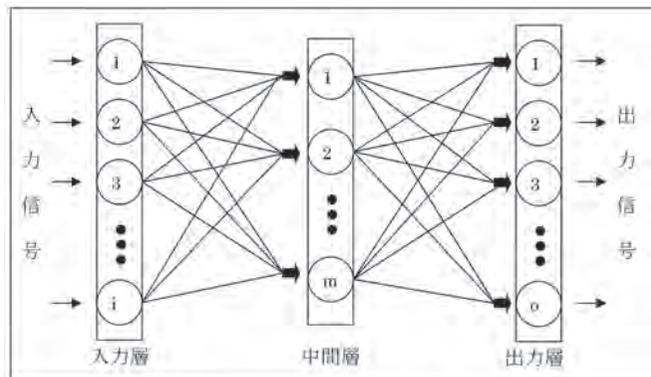


図2 本研究で用いた3階層型NNモデル

Eに起因する θ および W の修正情報は入力信号とは逆に中間層・入力層へと伝播され、ユニットおよびラインの特性変更を促す。これがBPによるNNの学習であり、生物を構成する神経細胞が有する学習メカニズムを模倣した自己組織型学習モデルである。学習のプロセスを図3に示す。

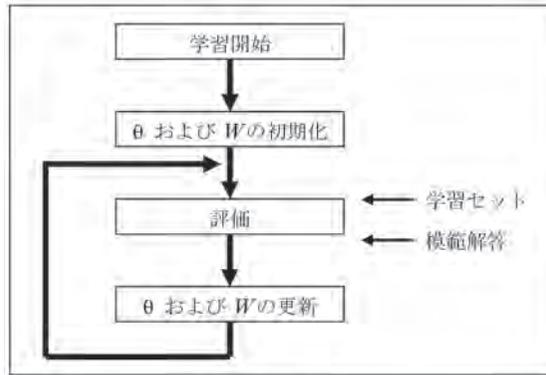


図3 学習プロセス

本研究では W および θ の更新に際し、BPの中でも最も一般的な形式を採用している。BPの数理モデルを以下に記す。なお、 α および β は学習係数と呼ばれ、入力信号パターンに則した学習係数を選択すれば効率的な学習が実現する。

中間層ユニット m と出力層ユニット o の結合係数 W_{mo} の更新値 $\Delta W_{mo} = \alpha \cdot \sigma_o \cdot Y_m$

出力層ユニット o の閾値 θ_o の更新値 $\Delta \theta_o = \beta \cdot \sigma_o$

入力層ユニット i と中間層ユニット m の結合係数 W_{im} の更新値 $\Delta W_{im} = \alpha \cdot \rho_m \cdot Y_i$

中間層ユニット m の閾値 θ_m の更新値 $\Delta \theta_m = \beta \cdot \rho_m$

ただし $\sigma_o = -(T_o - Y_o) \cdot Y_o \cdot (1 - Y_o)$

$$\rho_m = \sum_o (\sigma_o \cdot W_{mo}) \cdot Y_m \cdot (1 - Y_m)$$

4. 学習パターンの作成

実験に用いたデータは、東京証券取引所第一部上場株式銘柄全てを対象とした指標であるTOPIX (Tokyo Stock Price Index: 東証株価指数) の、2004年1月5日から2013年12月30日までの値動きである。同期間内の開場日は計2455日であった。また、拙著 [3] で明らかにした通り、市場休場時間帯に発生した様々な社会的事象は変動チャートのフラクタル性に大きな影響を与えない。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軸に配することとした。

市場における一日の価格変動状況を把握する際に用いられる代表的な値が日足である。日足は、1日の取引における始値・最高値を表す高値・最低値を表す安値・終値の4値から成る。そこで本研究では、前日までの価格変動から開場前に対象日の高値および安値予測を試みることで、売買タイミングの判断を支援するシステムの開発に目標を絞った。

具体的には、開場日 t における実際の高値を $H_r(t)$ と定義した時、 $t-1$ から $t-w$ の w 開場日間の経験を要素とする集合 $F = \{H_r(t-k) \mid 1 \leq k \leq w\}$ より、当日の予測高値 $H_e(t)$ を開場前に予測するシステムの開発を目指した。従って集合 F は、予測システムの根幹を成す NN 学習プロセスの入力信号パターンとなり、同パターンに対応する教師信号は $H_r(t)$ となる。しかしながら、実際の値をそのままの形で NN に学習させた場合、予測精度が極端に低くなるのが既に知られている [18] [19]。そこで本研究では、先に述べたチャートに現れるフラクタル性を利用した学習パターン（入力信号パターンおよび対応する教師信号）で構成する複数の学習セットを用意した。学習パターンおよび学習セットの作成手順を以下に示す。

【1 個の学習パターン作成手順】

1. 式 11 および式 12 を用いて集合 F より回帰直線 $Y = a_{hw}(t) \cdot X + b_{hw}(t)$ を算出

$$a_{hw}(t) = - \frac{w \sum_{k=1}^w ((t-k) \cdot H_r(t-k)) - \sum_{k=1}^w (t-k) \cdot \sum_{k=1}^w H_r(t-k)}{w \sum_{k=1}^w (t-k)^2 - (\sum_{k=1}^w (t-k))^2} \quad \text{式(11)}$$

$$b_{hw}(t) = - \frac{\sum_{k=1}^w (t-k)^2 \cdot \sum_{k=1}^w H_r(t-k) - \sum_{k=1}^w ((t-k) \cdot H_r(t-k)) \sum_{k=1}^w H_r(t-k)}{w \sum_{k=1}^w (t-k)^2 - (\sum_{k=1}^w (t-k))^2} \quad \text{式(12)}$$

2. 回帰直線より算出可能な開場日 t の理論高値を $H_r(t) = a_{hw}(t) \cdot t + b_{hw}(t)$ と定めた時、 $H_r(t)$ と $H_e(t)$ の距離を $H_d(t) = H_e(t) - H_r(t)$ と表記

3. 距離 H_{dw} を要素とする集合 $G = \{|H_{dw}(t-k)| \mid 1 \leq k \leq w\}$ を形成

4. 集合 G より式 13 に従い標準偏差 S_{hw} を算出

$$S_{hw} = \sqrt{\frac{1}{w} \sum_{k=1}^w (|H_{dw}(t-k)| - \overline{H_{dw}})^2} \quad \text{(式 13)}$$

ただし

$$\overline{H_{dw}} = \frac{1}{w} \sum_{k=1}^w |H_{dw}(t-k)|$$

前日までの変動データを基として、上記における a_{hw} は式 4 に示した確率微分方程式におけるドリフト項（トレンド）に、 S_{hw} は標準偏差項（ボラティリティ）に相当する。この2値より対象日 t の $H_d(t)$ を開場前に算出できれば、対象日 t の高値予測に成功したことになる。そこで、 $H_d(t)$ の予測に供する入力信号パターンを $I_{hw}(t) = \{a_{hw}(t), S_{hw}(t) \mid w = q_1, q_2, \dots, q_n\}$ 、 $I_{hw}(t)$ に対

応する教師信号を $T_{hw}(t) = \{H_{dw}(t) \mid w = q_1, q_2, \dots, q_n\}$ と定めた上で、開場日 t における複数の w 値からなる学習セット $P_{hw}(t) = \{I_{hw}(t), T_{hw}(t) \mid w = q_1, q_2, \dots, q_n\}$ を準備した。

また本研究で用いた NN の場合、入力信号・出力信号いずれにおいても強さは 0 又は 1 に制限されている。従って NN を用いた学習を実現するためには、学習パターンの実装に更なる工夫が求められる。そこで、各種データを指標化・離散化・量子化することにより、学習セット $P_{hw}(t)$ を 2 値 (bit) 表現することとした。

具体的な作業手順を以下に示すとともに、2013 年 1 月 4 日の TOPIX 日足を例に、各作業の結果を表 1 および表 2 に記す。

【指標化・離散化・量子化の手順】

1. 学習セット内に TOPIX 以外の経済指標を組み込むという発展性を考慮し、2004 年 1 月 5 日の始値を 100 として期間内の TOPIX を指標化 (表 1)
2. 1 の指標化済みデータを用いて $a_{hw}(t)$, $S_{hw}(t)$, $H_{dw}(t)$ を算出
3. $a_{hw}(t)$, $S_{hw}(t)$, $H_{dw}(t)$ 各々に対し、期間内における上限・下限を基に各値を 256 段階に離散化 (表 2)
4. 離散化データをビット数 8 で量子化 (表 2)

表 1 2013 年 1 月 4 日の TOPIX 日足データ

| 2013/01/04 | 始値 | 高値 | 安値 | 終値 |
|--------------------------|--------|--------|--------|--------|
| TOPIX | 876.97 | 888.53 | 876.83 | 888.51 |
| 2004/01/05 始値を 100 とした指標 | 83.119 | 84.215 | 83.106 | 84.213 |

表 2 $t=2013$ 年 1 月 4 日, $w=10$ の $a_{hw}(t)$, $S_{hw}(t)$, $H_{dw}(t)$: 離散化および量子化

| | $a_{hw}(t)$ | $S_{hw}(t)$ | $H_t(t)$ | $H_{dw}(t)$ |
|----------------|-------------------------|-----------------------|----------|----------------------------|
| $t=2013/01/04$ | 0.5819 | 0.3172 | 82.685 | 1.5296 |
| 期間内上限 | 1.7019 | 3.2754 | / | 10.3509 |
| 期間内下限 | -2.7330 | 0.0834 | / | -9.3558 |
| 離散化データ | $191 \leq a_w(t) < 192$ | $18 \leq S_w(t) < 19$ | / | $141 \leq H_{dw}(t) < 142$ |
| 量子化データ | 10111111 | 00010010 | / | 10001101 |

開場日 t における日足・安値を $L_r(t)$ と定義し、高値と同じ作業を実施することにより $P_{hw}(t) = \{I_{hw}(t), T_{hw}(t) \mid w = q_1, q_2, \dots, q_n\}$ (ただし $I_{hw}(t) = \{a_{hw}(t), S_{hw}(t)\}$ と $T_{hw}(t) = \{L_{dw}(t)\}$) を算出し、 $P_{hw}(t)$ と $P_{hw}(t)$ を併せ持つ最終学習セット $P_w(t)$ を形成した。

【学習セット1個の内訳】

$$P_w(t) = \{I_w(t), T_w(t) \mid w = q_1, q_2, \dots, q_n\}$$

$$\text{ただし, } I_w(t) = \{a_{hw}(t), S_{hw}(t), a_{lw}(t), S_{lw}(t)\}$$

$$T_w(t) = \{H_{dw}(t), L_{dw}(t)\}$$

5. NNによる学習実験の実行

学習実験を行うに当たり、学習を成立ならしめるために最適な各種パラメータを確定する必要が認められる。例えばBPの特性上、入力層・中間層・出力層それぞれに配置するユニット数の増加は計算量の増加とそれに伴う計算時間の増加を生むだけでなく、学習不成功のリスク高騰を覚悟しなければならない。そこで、複数の予備実験を行った上で表3の実験環境を構築した。

表3 実験環境を形成する各種パラメータ

| 項目 | パラメータ名 | 値 |
|----|--------------------------|---|
| 1 | 学習パターン数を決定するブロック化の深さ n | 5 |
| 2 | 1度の学習に供する学習セット数 u | 10 |
| 3 | NNの入力層ユニット数 | 160 |
| 4 | NNの中間層ユニット数 | 250 |
| 5 | NNの出力層ユニット数 | 80 |
| 6 | BPにおける学習係数 α | 0.8 |
| 7 | BPにおける学習係数 β | 0.8 |
| 8 | 学習終了条件 | 15万回の学習試行 または 出力層ユニット1個当たりの 誤差平均が0.001以下 |

項目1および2は、本NNに学習を期待する情報に直結するパラメータである。指標チャートに現れるフラクタル性を学習により自律認識したNNを用意することで、未来の指標チャート状態を予測しようというのが本研究の着眼点である。従って、多様な状況下での自律学習により複数の変動環境に対応可能なNNを求めるのであれば、1個の学習セットに含まれる学習パターン数を規定するブロック化の深さ n 、および1度の学習に供する学習セット数 u は大きいほど良い。しかし n および u の増加による学習内容の複雑化は、学習の不成功確率を高める。本研究で使用した学習パターンおよび学習セットを明確化する目的で、2013年1月4日の高値または安値を予測するNNを構築するために必要な $n=2$ 、 $w=\{3, 6\}$ 、 $u=3$ における学習セッ

ト群を表4に例示する。

表4, $n = 2$, $w = \{3, 6\}$, $u = 3$ における予測日 = 2013年1月4日の学習セット群

| u | n | w | w = { 3, 6 } (2012年) | | | | | |
|---|---|----------------------------|----------------------|-------|-------|---|---|---|
| | | | 1 | 2 | ③ | 4 | 5 | ⑥ |
| 1 | 1 | $I_3(u_1)$ の作成範囲 | 12/27 | 12/26 | 12/25 | | | |
| | 2 | $I_6(u_1)$ の作成範囲 | 12/27 | 12/26 | 12/25 | | | |
| | | $T_3(u_1), T_6(u_1)$ の算出対象 | 12月28日 | | | | | |
| 2 | 1 | $I_3(u_2)$ の作成範囲 | 12/26 | 12/25 | 12/21 | | | |
| | 2 | $I_6(u_2)$ の作成範囲 | 12/26 | 12/25 | 12/21 | | | |
| | | $T_3(u_2), T_6(u_2)$ の算出対象 | 12月27日 | | | | | |
| 3 | 1 | $I_3(u_3)$ の作成範囲 | 12/25 | 12/21 | 12/20 | | | |
| | 2 | $I_6(u_3)$ の作成範囲 | 12/25 | 12/21 | 12/20 | | | |
| | | $T_3(u_3), T_6(u_3)$ の算出対象 | 12月26日 | | | | | |

表4に示した学習セット群は全て2013年1月4日の日足・高値および安値を予測するために供されるデータであるため、2012年12月28日以前のデータに限定される。 $u=3$ より12月28日・27日そして26日を学習の起点とする直前3開場日の経験(3個の学習セット)が学習対象となる。また $n=2$, $w = \{3, 6\}$ より、当該学習セット群の構築には予測日から遡ること9開場日前までの日足情報を必要とする。1個の学習セットには入力信号8種と教師信号4種が含まれるため、本ケースにおける学習セット群は24種の入力信号と12種の教師信号で形成される。

u と n の設定を変化した複数の予備実験の結果を、トレードオフの関係にある学習不成功率と現実的な計算時間という二側面から検討した結果、本実験では $n=5$ および $u=10$ と設定した。 $n=5$ により、1つの学習セットが持つ入力信号は高値および安値のトレンドとボラティリティ5パターン計20個、出力信号は高値および安値での距離 $H_d(t)$ と $L_d(t)$ の5パターン計10個となる。各信号は8bit量子化されるため、NNの入力層ユニット数は160、出力層ユニット数は80となる。

中間層ユニット数の設定においては、その値を増加させることにより学習の分解能が高まり、より多様な状況を適切に学習できる可能性が高まる反面、計算量の爆発による計算時間の劇的増加を生む。またBPでは、中間層ユニット数の増加が学習能力の向上に寄与する保証は存在せず、逆に学習不成功率を高める危険性が存在する。そこで本パラメータについても幾つかの

予備実験結果を考慮し、250 と定めた。また BP における学習係数 α および β については、一般的に広く用いられる値を採用した。

項目 1 から 7 のパラメータを確定した上でなお、実験の成否を大きく作用するのが項目 8 の学習終了条件である。最急降下法による解の探索を行う BP では、最適解を常に発見できる保証は存在しない。BP は見つかった解が優良解か最適解かを判断する機能を有しておらず、NN の出力信号と教師信号との誤差または学習回数に制限を設けることで学習を強制終了させる以外に学習を終わらせることは出来ない。

実用的な学習時間確保を理由に、本実験においては 1 度の学習回数上限を 15 万回と定めた。もう一方の学習打ち切り誤差の設定に際しては、出力層ユニット 1 個当たりの誤差平均 (Error Average) と学習状況の関連性を事前に確認する必要が認められた。そこで、 $W_1 = \{10, 20, 30, 40, 50\}$ において 2013 年 1 月 4 日の TOPIX 高値と安値を予測するために用いる NN の学習状況を表したのが図 4・図 5 および表 5 である。

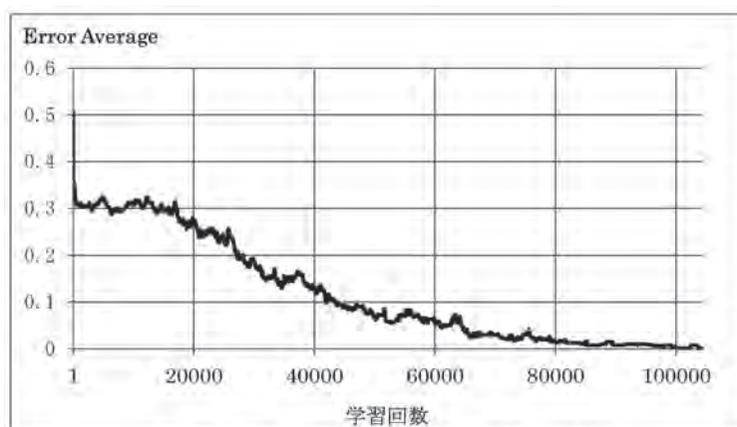


図 4 $W_1 = \{10, 20, 30, 40, 50\}$ における誤差平均推移

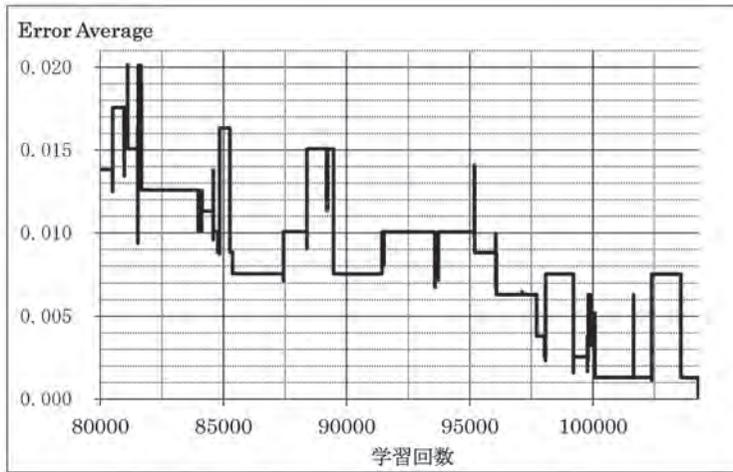


図5 図4における，学習回数80000以降の誤差平均推移詳細

表5 学習終了条件の設定と学習状況

| 学習終了条件 Error Average ≤ | 学習終了回数 | 誤差平均 | 学習失敗 | |
|---------------------------|--------|---------|------|-----|
| | | | 実数 | 発生率 |
| 0.010 | 81523 | 0.00939 | 10 | 10% |
| 0.009 | 84767 | 0.00895 | 7 | 7% |
| 0.008 | 85370 | 0.00766 | 6 | 6% |
| 0.007 | 93581 | 0.00677 | 8 | 8% |
| 0.006 | 97717 | 0.00388 | 3 | 3% |
| 0.005 | 97717 | 0.00388 | 3 | 3% |
| 0.004 | 97717 | 0.00388 | 3 | 3% |
| 0.003 | 98038 | 0.00256 | 2 | 2% |
| 0.002 | 99207 | 0.00160 | 2 | 2% |
| 0.001 | 104272 | 0.00015 | 0 | 0% |

本学習セット群には計100個の教師信号が含まれるが，セット群はt-1からt-10までの10個の学習セットで形成されるため，1個の学習セットは10個の教師信号を内包している。8bit量子化された教師信号1つは8個の出力層ユニットで担当するため，出力層ユニットから同時出力される信号と教師信号の誤差合計は最大で80となる。そこで図4では出力層ユニット1個当たりの誤差平均に着目し，BPによる学習成果推移を時系列で表している。また，特に80000回以降の時系列変化について詳しく記述したグラフが図5となる。

表5より，104272回の学習を繰り返すことで誤差平均が0.001を下回った事実が判る。また，学習終了条件をError Average ≤ 0.010から0.001まで0.001刻みで設定した際の学習済みNNを

対象に、学習の適切さを検証した結果が「学習失敗」項となる。具体的には、学習に供した学習セット群を構成する入力信号パターン I_w を学習済み NN に改めて入力する。それによって出力層ユニットの出力信号 $E_w(t) = \{E_{hw}(t), E_{lw}(t) \mid w = q_1, q_2, \dots, q_n\}$ を求め、教師信号 $T_w(t)$ との比較により適正に学習が行われたか否かを検証した。例えば Error Average ≤ 0.010 と設定した場合、適切な学習が行われなかった結果、模範解答による反復学習を既に行った 100 の出力信号のうち 1 割の 10 で依然として間違ってしまう NN が形成される事実が認められる。また本学習セット群においては、Error Average ≤ 0.001 と設定することで誤学習をゼロとすることが可能であった。その他の幾つかの予備実験結果も勘案し、本実験では Error Average ≤ 0.001 と定めた。

5. 先行実験の実施と問題点の把握

これまでに紹介した手法に従い、2011年12月21日から2013年12月30日までの500開場日を対象に、TOPIXにおける日足高値および安値の予測を試みた。具体的には、予測前日を起点に w 日間に遡った日足値から算出されるトレンド $a_w(t-1)$ とボラティリ $S_w(t-1)$ を学習済みネットワークに入力することにより、予測日 t における回帰直線 $y = a_{hw}(t-1) \cdot x + b_{hw}(t-1)$ との距離 $H_{dw}(t)$ および $y = a_{lw}(t-1) \cdot x + b_{lw}(t-1)$ との距離 $L_{dw}(t)$ を応答するよう NN に求めた。

NN が応答する $H_{dw}(t)$ および $L_{dw}(t)$ は 8bit データであるため、4. で示した【指標化・離散化・量子化の手順】を逆に辿ることで TOPIX データに変換する。ただし、離散化によって欠落した情報を補完する際には、階級の上限と下限による単純平均を用いた。

予測日 t の予想高値 $H_e(t)$ および予想安値 $L_e(t)$ は、 $H_e(t) = H_t(t) + H_a(t)$ 、 $L_e(t) = L_t(t) + L_a(t)$ で算出されるが、本実験では $n = 5$ と設定しているため高値については $H_{d10}(t)$ から $H_{d50}(t)$ 5種、安値についても同じく $L_{d10}(t)$ から $L_{d50}(t)$ の5種が予測される。そこで本実験では式14および式15を用いて $H_d(t)$ と $L_d(t)$ を算出することとした。

$$H_d(t) = \frac{1}{5}(H_{d10}(t) + H_{d20}(t) + \dots + H_{d50}(t)) \quad (\text{式14})$$

$$L_d(t) = \frac{1}{5}(L_{d10}(t) + L_{d20}(t) + \dots + L_{d50}(t)) \quad (\text{式15})$$

表6および図6は、当日の立会時間終了後に判明する $H_r(t)$ および $L_r(t)$ を基準に、式16および式17で求まる予測値と実際値の差異状況を表している。なお参考として、式18で求まるの同一期間内における状況も表6に記載する。

$$E_h(t) = \frac{|H_e(t) - H_r(t)|}{H_r(t)} \quad (\text{式16})$$

$$E_l(t) = \frac{|L_e(t) - L_r(t)|}{L_r(t)} \quad (\text{式17})$$

$$HL(t) = \frac{H_r(t) - L_r(t)}{(H_r(t) + L_r(t))/2} \quad (\text{式18})$$

表6 予測値と実際値の差異状況

| 差異の範囲 | 発生日数 | | |
|-------------|----------|----------|------------|
| | $E_h(t)$ | $E_l(t)$ | HL(t) : 参考 |
| 0%以上 1%未満 | 221 | 189 | 236 |
| 1%以上 2%未満 | 143 | 143 | 214 |
| 2%以上 3%未満 | 74 | 71 | 38 |
| 3%以上 4%未満 | 27 | 44 | 5 |
| 4%以上 5%未満 | 19 | 16 | 5 |
| 5%以上 6%未満 | 7 | 12 | 0 |
| 6%以上 7%未満 | 4 | 15 | 1 |
| 7%以上 8%未満 | 3 | 5 | 0 |
| 8%以上 9%未満 | 1 | 1 | 1 |
| 9%以上 10%未満 | 0 | 3 | 0 |
| 10%以上 11%未満 | 0 | 0 | 0 |
| 11%以上 12%未満 | 1 | 0 | 0 |
| 12%以上 13%未満 | 0 | 1 | |
| 合計 | 500 | | |
| 最大(単位:%) | 11.02% | 12.72% | 8.19% |
| 最小(単位:%) | 0.00% | 0.00% | 0.31% |
| 平均(単位:%) | 1.56% | 1.88% | 1.21% |
| 標準偏差(単位:%) | 1.4534 | 1.8022 | 0.7577 |

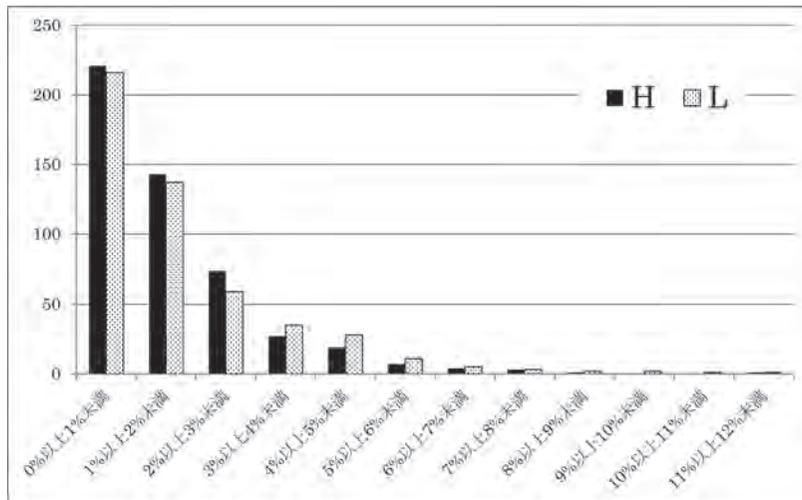


図6 予測値と実際値の誤差状況

表6からも明らか通り、本実験によって得られた予測の精度は低い。予測値と実際値との差異発生状況は、高値と安値の差分が両者の平均に占める割合 $HL(t)$ の発生状況より全体として大きく、実用に耐える状況にはないと判断できる。このような状況が生じた原因として、学習打ち切り条件の設定が不適切であった可能性が考えられる。そこで、適用された学習終了条件と $E_h(t)$ および $E_l(t)$ の関連について表7に示す。また表8では、各々のケースにおける誤差平均と $E_h(t)$ および $E_l(t)$ の相関係数を示す。

表7 適用された学習終了条件と $E_h(t)$ および $E_l(t)$ の関連

| 適用された終了条件 対象 | 誤差平均 0.001 以下 | | 15万回の学習試行 | |
|-----------------------------------|---------------|----------|-----------|----------|
| | $E_h(t)$ | $E_l(t)$ | $E_h(t)$ | $E_l(t)$ |
| 発生回数 (500 回中) | 212 回 | | 288 回 | |
| 学習回数平均 | 101605.6 | | 150000 | |
| 誤差平均の平均 | 0.0004 | | 0.07404 | |
| $E_h(t)$, $E_l(t)$ の平均 (単位: %) | 1.65% | 2.00% | 1.36% | 1.46% |
| $E_h(t)$, $E_l(t)$ の標準偏差 (単位: %) | 1.5402 | 1.8951 | 1.1166 | 1.2665 |

表 8 誤差平均の平均と $E_h(t)$ および $E_l(t)$ の相関係数

| | | 全体 | 適用された終了条件 | |
|----------------|----------|---------|---------------|------------|
| | | | 誤差平均 0.001 以下 | 15 万回の学習試行 |
| 発生回数 | | 500 回 | 212 回 | 288 回 |
| 誤差平均との 相関係数 | $E_h(t)$ | 0.1948 | -0.0933 | 0.1683 |
| | $E_l(t)$ | -0.0279 | -0.0453 | 0.0900 |

表 8 より、各々の学習における誤差平均と $E_h(t)$ および $E_l(t)$ には相関が認められない事実が判る。また表 7 に示した通り、誤差平均が 0.001 以下となって学習を終了した場合と比較し、誤差が十分に小さくならず学習を打ち切った方が $E_h(t)$ および $E_l(t)$ の平均・標準偏差共に小さくなっている。つまり、過去の経験をより厳密に学習する行為は未来予測に必ずしも資するとは言えず、むしろある程度の曖昧さを含んだ学習の方が未来予測には適している可能性が示されている。

他方では、例えば図 7 に示す $E_h(t)$ の時系列推移（2012 年 10 月 11 日から 2013 年 3 月 11 日までの 100 開場日）には一定の周期性を認めることが可能であり、予測値と実際値の誤差変化からの法則抽出による予測精度の向上が予見できる。

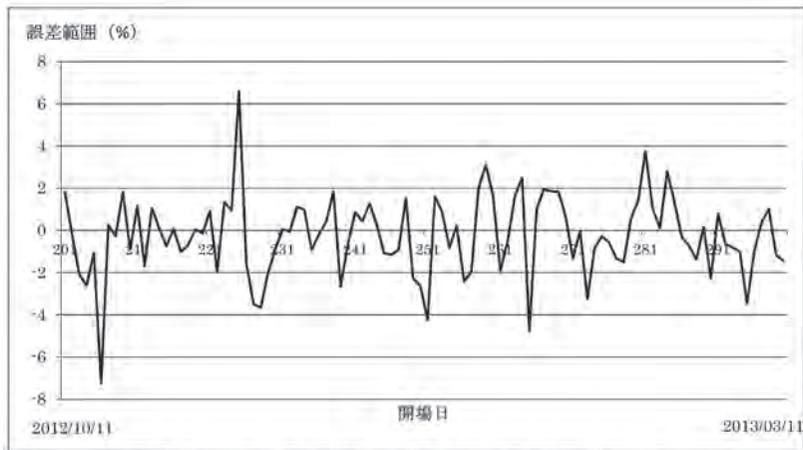


図 7 2012 年 10 月 11 日から 2013 年 3 月 11 日まで $E_h(t)$ の時系列推移

6. 実験および考察

先行実験の結果を踏まえ、同じく 2011 年 12 月 21 日から 2013 年 12 月 30 日までの 500 開場日を対象に、集合 w の値が異なる複数の学習セット群による学習および未来予測実験を実施した。具体的には、従来の $W_1 = \{10, 20, 30, 40, 50\}$ 以外に $W_2 = \{15, 30, 45, 60, 75\}$ および

$W_3 = \{20, 40, 60, 80, 100\}$ を対象とした実験を行うことで、学習セット群の違いによる予測精度を検証した。

そこで以降では、例えば W_1 によって求めた予測値を基とした $E_h(t)$ を $W_1 \rightarrow E_h(t)$ と記す。また、 W_1, W_2, W_3 それぞれでの予測値を単純平均化した値を基とした $E_h(t)$ を $Ave \rightarrow E_h(t)$ と表すこととする。表9に実験結果を示す。

表9 W_1, W_2, W_3 における予測値と実際値の差異状況

| 差異の範囲 | 発生日数 | | | | | | | |
|-------------|-------------------|----------|-------------------|----------|-------------------|----------|-------------------|----------|
| | $W_1 \rightarrow$ | | $W_2 \rightarrow$ | | $W_3 \rightarrow$ | | Ave \rightarrow | |
| | $E_h(t)$ | $E_l(t)$ | $E_h(t)$ | $E_l(t)$ | $E_h(t)$ | $E_l(t)$ | $E_h(t)$ | $E_l(t)$ |
| 0%以上 1%未満 | 221 | 189 | 228 | 193 | 232 | 187 | 238 | 211 |
| 1%以上 2%未満 | 143 | 143 | 149 | 134 | 135 | 137 | 159 | 135 |
| 2%以上 3%未満 | 74 | 71 | 62 | 73 | 66 | 79 | 65 | 74 |
| 3%以上 4%未満 | 27 | 44 | 26 | 47 | 32 | 50 | 22 | 42 |
| 4%以上 5%未満 | 19 | 16 | 16 | 24 | 13 | 17 | 6 | 15 |
| 5%以上 6%未満 | 7 | 12 | 6 | 12 | 9 | 13 | 3 | 11 |
| 6%以上 7%未満 | 4 | 15 | 5 | 9 | 7 | 3 | 2 | 8 |
| 7%以上 8%未満 | 3 | 5 | 4 | 3 | 3 | 4 | 2 | 0 |
| 8%以上 9%未満 | 1 | 1 | 0 | 3 | 0 | 4 | 1 | 3 |
| 9%以上 10%未満 | 0 | 3 | 1 | 1 | 1 | 4 | 1 | 0 |
| 10%以上 11%未満 | 0 | 0 | 2 | 0 | 1 | 1 | 1 | 1 |
| 11%以上 12%未満 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 12%以上 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 |
| 合計 | 500 | | | | | | | |
| 最大(単位:%) | 11.02% | 12.72% | 12.82% | 11.26% | 12.92% | 13.47% | 10.38% | 10.57% |
| 最小(単位:%) | 0.00% | 0.00% | 0.00% | 0.02% | 0.01% | 0.00% | 0.01% | 0.02% |
| 平均(単位:%) | 1.56% | 1.88% | 1.53% | 1.87% | 1.56% | 1.89% | 1.33% | 1.68% |
| 標準偏差(単位:%) | 1.4534 | 1.8022 | 1.6014 | 1.6974 | 1.6005 | 1.8011 | 1.2837 | 1.5205 |

いずれの学習セット群においても予測値と実際値の差異発生状況は似通っており、 W の値による学習セット群の相違が予測のための学習に与える影響は少ないと考えられる。しかし、Ave \rightarrow においては $E_h(t)$ と $E_l(t)$ の双方で平均・標準偏差共に減少しており、個々の学習セット群のみの場合との比較において優位性が認められる。そこで Ave \rightarrow が有する優位性の原資を探る目的で、開場日 t において W_1, W_2, W_3 それぞれの予測値と実際値を比較し、最も差異の小さな予測値を選出した Opt $\rightarrow E_h(t)$ および Opt $\rightarrow E_l(t)$ を導出した。

Opt \rightarrow は、 W_1, W_2, W_3 という異なった学習セット群を学習した3個のエージェントが各々下した未来予測を目前に並べ、最も賢明な選択を行ったとの仮定に基づく理想状態である。Opt \rightarrow における各エージェントの選出状況および差異状況を表10に示すが、 W_1, W_2, W_3 の構

成比率は概ね 1/3 となっている。また、結果に基づく理想的選択であるが故に当然であるが、Ave → と比較し、平均・標準偏差共に大きく減少している。

表 10 Opt → $E_h(t)$ および Opt → $E_l(t)$ の構成状況

| | Opt → | | $W_1 \rightarrow$ からの選出 | | $W_2 \rightarrow$ からの選出 | | $W_3 \rightarrow$ からの選出 | |
|------------|----------|----------|-------------------------|----------|-------------------------|----------|-------------------------|----------|
| | $E_h(t)$ | $E_l(t)$ | $E_h(t)$ | $E_l(t)$ | $E_h(t)$ | $E_l(t)$ | $E_h(t)$ | $E_l(t)$ |
| 選出日数 | 500 | | 168 | 157 | 167 | 194 | 165 | 149 |
| 割合 | 100.0% | | 33.6% | 31.4% | 33.4% | 38.8% | 33.0% | 29.8% |
| 最大(単位:%) | 10.91% | 12.25% | 7.35% | 6.42% | 7.37% | 12.25% | 10.91% | 6.43% |
| 最小(単位:%) | 0.00% | 0.00% | 0.00% | 0.00% | 0.01% | 0.00% | 0.00% | 0.00% |
| 平均(単位:%) | 0.71% | 0.81% | 0.84% | 0.71% | 0.65% | 0.88% | 0.64% | 0.80% |
| 標準偏差(単位:%) | 0.9999 | 1.0775 | 1.0599 | 1.6974 | 0.8231 | 1.3484 | 1.0895 | 0.8808 |

表 10 より明らかな通り、過去の変動に対し異なる角度での学習を行った 3 個の独立エージェントによる未来予測を参照するという Ave → の戦略は、理に適っている。よって、1 個のエージェントが下した未来予測のみを信用するのではなく、複数のエージェントが行う予測を何らかの手段をもって平準化することで更なる予測を行う戦略は有用だと判断する。しかし、答えが既知な状況の理想戦略 Opt → と比べるとその精度はまだ不十分であり、実用性を欠いている。そこで、Opt → を参考に理想戦略に近づく方策を以下で考察する。

表 11 は、Opt → における選択の連続性を示している。エージェントの選択が連続しない（単独であった）日数は $E_h(t)$ で 208 日・ $E_l(t)$ で 215 日であったが、3 日以上連続した日数の合計はそれぞれ 172 日と 147 日であった。以上からも明らかな通り、理想戦略 Opt → におけるエージェントの選択には一定の規則性が認められる。これは、TOPIX 変動が有するフラクタル性に起因するものと考えられる。

表 11 Opt → $E_h(t)$ および Opt → $E_l(t)$ に認められる連続性

| | | 連続日数 | | | | | | 日数 |
|-------------------|----------|------|------|------|-----|-----|-----|-----|
| | | 単独 | 2 日 | 3 日 | 4 日 | 5 日 | 6 日 | |
| $W_1 \rightarrow$ | $E_h(t)$ | 64 回 | 21 回 | 12 回 | 5 回 | 0 回 | 1 回 | 168 |
| | $E_l(t)$ | 75 回 | 24 回 | 10 回 | 1 回 | 0 回 | 0 回 | 157 |
| $W_2 \rightarrow$ | $E_h(t)$ | 76 回 | 23 回 | 5 回 | 2 回 | 2 回 | 2 回 | 167 |
| | $E_l(t)$ | 76 回 | 24 回 | 11 回 | 8 回 | 1 回 | 0 回 | 194 |
| $W_3 \rightarrow$ | $E_h(t)$ | 68 回 | 16 回 | 13 回 | 5 回 | 0 回 | 1 回 | 165 |
| | $E_l(t)$ | 64 回 | 21 回 | 10 回 | 2 回 | 1 回 | 0 回 | 149 |

そこで Opt → において、一定期間における各エージェントの選択状況を分析した結果、例えば表 12 に示す周期性の存在が確認できた。表 12 では、対象期間を連続する 50 開場日単位で区切り、Opt → として選ばれたエージェントが占める日数の割合を表している。各々で 40% を超過するケースを強調して表示しているが、 $E_h(t)$ と $E_l(t)$ いずれにおいても明確な周期性が確認できる。これら周期性を感知する機能を組み込み、エージェントの選択に何らかの優先順位を付与する機能を設けることができれば、より高い精度での未来予測が期待できる。

表 12 Opt → $E_h(t)$ および Opt → $E_l(t)$ に認められる周期性 (50 開場日単位)

| 区間 | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|----------|-------------------|---------------|---------------|---------------|----------------|----------------|---------------|---------------|---------------|---------------|
| 始まり開場日 | | 2012/ 2/13 | 2012/ 4/24 | 2012/ 7/6 | 2012/ 9/18 | 2012/ 11/29 | 2013/ 2/18 | 2013/ 5/1 | 2013/ 7/12 | 2013/ 9/25 |
| 終り開場日 | | 2012/ 4/23 | 2012/ 7/5 | 2012/ 9/14 | 2012/ 11/28 | 2013/ 2/15 | 2013/ 4/30 | 2013/ 7/11 | 2013/ 9/24 | 2013/ 12/5 |
| $E_h(t)$ | $W_1 \rightarrow$ | 36% | 44% | 42% | 46% | 24% | 28% | 28% | 34% | 28% |
| | $W_2 \rightarrow$ | 26% | 22% | 26% | 32% | 44% | 30% | 44% | 44% | 40% |
| | $W_3 \rightarrow$ | 38% | 34% | 32% | 22% | 32% | 42% | 28% | 22% | 32% |
| $E_l(t)$ | $W_1 \rightarrow$ | 32% | 32% | 20% | 44% | 30% | 28% | 22% | 40% | 28% |
| | $W_2 \rightarrow$ | 44% | 24% | 54% | 28% | 40% | 36% | 38% | 36% | 54% |
| | $W_3 \rightarrow$ | 24% | 44% | 26% | 28% | 30% | 36% | 40% | 24% | 18% |

7. まとめと今後の課題

本研究では、TOPIX 推移を対象に確率微分方程式によって予見される無数の近未来からの選択に際し、自己学習機能を有する NN をエージェントとして活用する手法について探究した。

はじめに、予測対象日から遡る期間が異なる複数の区間を用意し、それぞれの区間内におけるトレンドとボラティリティで構成されるデータを NN の学習セット群と定めた。実験に用いた NN の特性上、学習に供するデータ数を増加させると学習に要する時間が激増するだけでなく、学習成果も芳しくなくなる。そこで本研究では、1 種類の学習セット群が有する情報を制限する代わりに、異なる視点からの分析に基づく複数の学習セット群を用意した。同時に、NN を内包するエージェントを学習セット群の数だけ用意し、各エージェントは対応する学習セット群のみを用いて独立して学習する環境とした。

結果、学習の成果としてもたらされる未来予測はエージェントの数だけ出現するが、それらの単純平均によって求まる予測値は個々のエージェントが下した予測よりも高い精度を示すことが明らかとなった。つまり、TOPIX 変動が有するフラクタル性が有する特性の一部を学習セット群が抽出しただけでなく、NN による学習がそれを一層先鋭化させていると考えられる。また、フラクタル性に含まれる異なる特性を学習した 3 個の独立エージェントの出現によって、

複雑系に潜む、直感では認識が困難な周期性を可視化できる事実も明らかにした。

TOPIX 変動に内在するフラクタル性に起因する状況変化をいち早く察知できる機能をエージェントに付与することで、より優れたエージェント選別が可能な予測システムを構築することによって、予測精度を高めるのが今後の課題である。

【謝辞】

本研究は、和歌山大学経済学部・平成 23 年度研修専念制度の活用によって得られた成果の一つである。貴重な研究機会を提供下さった関係者各位に深謝申し上げます。

【参考文献】

- [1] B.B. Mandelbrot, 広中平祐 監訳, フラクタル幾何学, 日経サイエンス社, 1984
- [2] 田中航二, TOPIX (1990-1970) にみられる変動の非線形特性について, 大阪商業大学論集, 第 113 巻, 871-885, 1999
- [3] 岩田英朗, 株価変動に見るフラクタル性の研究, 経済理論 343 号, 和歌山大学経済学会, 2008
- [4] 高橋朋一・長坂健二, 時系列データのフラクタル次元の推定法, 計測自動制御学会論文集, Vol.33, No.9, 987-980, 1997
- [5] L.F. Burlaga and L.W. Klein, Fractal structure of the interplanetary magnetic field, Journal of Geophysical Research, 91, 347-350, 1986
- [6] L.F. Burlaga and L.W. Klein, Large-scale fluctuations in the interplanetary medium, Journal of Geophysical Research, 92, 1261-1266, 1987
- [7] 山本 拓, 経済の時系列分析, 創文社, 1998
- [8] B. Oksendal, 谷口説男 訳, 確率微分方程式, シュプリンガー・ジャパン, 1999
- [9] 北坂真一, 統計学から始める計量経済学, 有斐閣, 2005
- [10] Michael A. Arbib : *The Metaphorical Brain 2*, John Wiley & Sons. Inc., 1989
(邦訳「ニューラルネットと脳理論」, 訳 = 金子隆芳, サイエンス社)
- [11] F. Rosenblatt : *Principle of Neurodynamics*, Spartan, 1961.
- [12] J.K. Hopfield : *Neural network and physical systems with emergent collective computational abilities*, Proc. Natl. Acad. Sci. USA, Vol.79, 1982.
- [13] D.H. Ackley, G.E. Hinton, and T.J. Sejnowski : *A Learning Algorithm for Boltzmann Machines*, Cognitive Sci., 9, 1985.
- [14] D.E. Rumelhart, J.L. McClelland, and PDP Research Group : *Parallel Distributed Processing*, Vol. I & II , MIT press, 1986.
- [15] 甘利俊一 : 「神経回路網の数理」, 産業図書, 1978.
- [16] D.E. Rumelhart, J.L. McClelland, and PDP Research Group : *Parallel Distributed Processing*, Vol. I & II , MIT press, 1986
- [17] E.F. Scott : *An Empirical Study of Learning Speed in BackPropagation Networks*, CMU-CS-88-162
- [18] 甘利俊一 : 「神経回路網の数理」, 産業図書, 1978
- [19] 上坂吉則 : 「ニューロコンピューティングの数学的基礎」, 近代科学社, 1993

Investigating and Understanding the Fractal Theory of Changes in Economic Indicators Over Time Using Neural Networks

Hideaki IWATA

Abstract

In this paper, I focused on the fractal structure of change charts for economic indicators, and attempted to calculate realization probabilities for the innumerable futures that can be forecast. Specifically, at first, I quantified changes in trends and volatility using TOPIX change charts from January 5, 2004, to December 30, 2013, and produced learning patterns for neural networks including input and instruction signals based on these values. After that, I stimulated self-organization of neural networks using these patterns. I predicted near-future TOPIX values using these newly-created networks, and measured errors between the predicted and actual values.

This research is based on the supposition that past changes in economic conditions affect future alternative possibilities. So, when producing learning patterns, in the technique for estimating the fractal dimension in changes over time I made reference to my previously proposed slope method. In this way, using the linear technique of stochastic differential equations, past changes are stored as patterns. In my research, I try to increase prediction accuracy by using three-layered neural networks based on the non-linear forecasting technique of backpropagation.