時系列データの特徴パターンの遷移に着目した 時系列現象の定性的モデリングについての一提案

正員 直江秀憲 (日立製作所)

Proposal of Qualitative Modeling for Time Series Phenomena
based on the Transition of Time Series Data Pattern

Hidenori Naoe, Member (Hitachi Ltd.)

The present paper proposes a modeling method which describes and predicts time series phenomena in a qualitative manner. For the qualitative description, time series data are characterized by the first and second order coefficients which are the outputs of a time dependent polynomial filter. In order to predict by the neural net, the coefficients are transformed and normalized to [0,1] space. From the series of these normalized coefficients, a few teaching patterns for neural net are extracted by clustering algorithm. And by error back propagation algorithm, the future state of the time series phenomena is obtained as symbols. Because the stock price does not necessarily walk randomly, the stock prices are tried to identify and predict as the time series data. Comparing the proposal method with conventional autoregressive model, it is clarified that the former has even high precision concerning with the predicted pattern by the correlation coefficients.

キーワード：時系列現象、多項式、ニューラルネット、クラスター分析、株価、パターン遷移

1. はじめに

本論文の目的は非定常な時系列現象を同定し、予測するための1つの方法について議論することである。従来、このような方法については自己回帰モデル、GMDH (Group Method of Data Handling)などが提案されており、近年では誤差逆伝搬法によるニューラルネットが注目されている^{1) 2)}。このうち、自己回帰モデルは現在の時系列データの値を過去の時系列データの値の線形結合によって説明しようとするものである。GMDHとは要因候補群 $\{X_1, \dots, X_n\}$ と被説明変数 y との関数関係を求めるのに次のような手続きを取る方法である。すなわち、 y を説明するのに要因候補の対 $\{X_i, X_j\}$ から作られる2次式が有意かどうかを調べ、有意ならばあらたに要因候補群に加え、そうでなければ棄却するという手続きを繰り返す。さらに、誤差逆伝搬法によるニューラルネットは、評価関数として出力の誤差の大きさをとり、ニューロン特性を定め、入力を与えたときに教師データに近い出力を得るようにシナプス結合を設定する。

このような時系列データ予測手法に対して、従来の数式による記述方法として特にARモデル(自己回帰モデル)を取り上げてその特徴をみると、以下のことが言える。まず、このモデルの次数はAIC (Akaike Information Criterion, 情報量規準)によって決められて

きた。ARモデルは統計的な線形近似であるため、定常的なデータを取り扱うのには適しているが、非線形性の強い、非定常なデータを取り扱うのには必ずしも適していない。このため、近年では、地震波等の非定常現象をAICの観点から同定しようとする研究、音声認識を状態遷移モデルを導入することによって行なおうとする研究がなされている^{3) 4)}。また、ARモデルでは時系列現象の予測値は数値としてのみ求まり、予測値のパターンをとらえることは容易ではなかった。

一方、株価や経済データの時系列データは、上昇や下降といった特徴パターンの遷移によって、時系列データが構成される傾向にあると思われる。そこで、本論文は、時系列データの特徴パターンの遷移に着目して、時系列データを予測する手法を提案する。具体的には、時間に関する多項式を用いて時系列現象の変動の特徴量(傾き、微分係数)を抽出し、この特徴量の数値としての予測値を誤差逆伝搬法を用いて求め、その値を用いて記号としての予測値を求める。特徴量は記号への変換の容易な量であり、多層ネットにおける学習は特徴量の因果関係付けに相当する。特徴量は「上昇」、「平行」、「下降」を表わす記号と対応付けられている。したがって、時系列現象はこの記号によって表現することができる。すなわち、記号列によって従来は容易ではなかった時系列現

象のパターンとしての予測値をとらえることを試みる。

本論文の構成は次のようになっている。すなわち、第2章において従来の予測技術について述べる。第3章において、株価の変動の傾向を抽出し、変動傾向を記号で表現し、変動を記号の遷移としてとらえる。ついで、株価の変動はランダム・ウォークではないことをふまえ、株価変動の特徴を表わす少数のパターンをクラスター分析の手法を用いて抽出する⁵⁾。そして、この少数のパターンが得られたとき、これを教師データとし、誤差逆伝搬法によって結合定数を定める多層ネットワークを用い、株価の将来の値を予測することを行なう。第4章では時系列現象として強い非線形性を持っている株価の変動を取り上げ、時系列現象を定性的に表現する方法（定性的モデリング）の観点から株価変動の同定、予測を試みる。そして、本論文で提案した方法の有効性を実現値と予測値とを平行移動させて求めた相関係数、予測パターンの記述、の面から検討する。そして、最後に第5章において本論文の結論を述べる。

2. 時系列分析の従来の技術

本章では従来の時系列データ分析の技術とその特徴について述べる。

<2・1> 自己回帰モデル

自己回帰モデルとは時系列データ $\{X(k)\}$ が与えられたとき、

$$X(t) = a_1 X(t-1) + \dots + a_n X(t-n) + \varepsilon_t \quad (1)$$

ε_t : 誤差項

によって将来の時系列の値を求めるといった方法である。次数 n の決定には A I C (Akaike Information Criterion) を用いる。

<2・2> GMDH

GMDHとは要因候補群 $\{X_1, \dots, X_n\}$ と被説明変数 y との関数関係を求めるのに次のような手続きをとる方法である。すなわち、 y を説明するのに要因候補の対 $\{X_i, X_j\}$ から作られる2次式が有意かどうかを調べ、有意ならば新たに要因候補群に加え、そうでなければ棄却するという手続きを繰り返すというものである。

<2・3> 誤差逆伝搬法

誤差逆伝搬法とは次のようなものである。

m 層のネットワークを考え、 k 層の第 i ユニットへの入力総和を I^k_i 、出力総和を O^k_i とし、 $k-1$ 層の第 i ユニットから k 層の第 j ユニットへの結合の重みを ω^{k-1}_{ij} とする。各ユニットの入出力関係を与える関数を f 、

$$O^k_i = f(I^k_i) \quad (2)$$

$$I^k_j = \sum_i \omega^{k-1}_{ij} O^{k-1}_i \quad (3)$$

誤差関数を E とすると、あるパターンの組 (x, y) が与えら

れたとき、

$$E = \sum_j (O^m_j(\omega, x) - y_j)^2 \quad (4)$$

で与えられる。ニューラルネットワークの誤差関数 E を最小化するため、各パターンに関する誤差関数の勾配ベクトルを計算し、 ω を改良するために確率的降下法を用いる。

<2・4> 従来の時系列データ分析手法の特徴

上記に示した代表的な3つの方法は過去の時系列データから将来の時刻 t の一点一点を予測するものであり、時系列データのパターンの遷移に陽に着目し、それを利用するものではない。後に示すように、経済データのような時系列データは上昇、下降をパターンとして繰り返す特徴を持つために、この特徴を利用した時系列データの予測手法を提案することを考えた。本論文では、非定常な時系列現象である株価の変動から特徴量を取り出して、しかる後に誤差逆伝搬法を用い、ARモデルと性能を比較する。ここに、特徴量とパターンとは密接に関連している。

3. 時系列現象の定性的モデリング

<3・1> 変動傾向の記号化

連続事象的なモデル表現では、強度に非線形な現象を表わすには限界があるため、連続事象を記号化（パターン認識）して、記号化した離散空間での因果関係付けを行ない、強度に非線形な現象に対してモデリングをしていく必要があると考えられる。本論文では時間に関する多項式を用いて時系列現象の変動の傾き（微分係数）を求め、変動を「上昇」、「平行」、「下降」と表現して、現象を把握する。これも本論文の特徴の1つである。

いま、時系列データ $\{X_t\}$ ($t = 1, \dots, N$) が与えられたとする。この時系列データに次のような時間の多項式を用いたフィルタをかけ、変動の微係数を求める。ここで、時系列データは連続的には連続関数 $\{f(t)\}$ と表わすことができるものとする。

$$H_m(t) = \sum_{i=0}^m g_{mi} t^{m-i} \quad (m = 0, 1, 2) \quad (5)$$

$$\int_{-\infty}^0 H_m(t) H_n(t) \exp\left(\frac{-t^2}{a}\right) dt = 0 \quad (6)$$

$(m \neq n)$

$$\int_{-\infty}^0 H_m(t)^2 \exp\left(\frac{-t^2}{a}\right) dt = A_m \quad (7)$$

$$f(t) = \sum_{m=0}^2 \frac{1}{\sqrt{A_m}} a_m H_m(t) \quad (8)$$

$$a_m = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{A_m}} H_m(t) \exp\left(-\frac{t^2}{a}\right) dt \quad (9)$$

式(5)は時間の多項式を用いたフィルタを表わしており、式(6) (7)はこのフィルタが直交関数であることを示している。式(8)は時系列関数がこのフィルタを用いてどのように表わされるかを示しており、式(9)は式(8)中の未定係数がどのように求められるかを示している。

ここに、取り扱おうとする現象の時間スケールパラメータに相当する a の値の取り方に任意性がある。本論文では時系列現象を時間の m 次の多項式によって近似することを考える。このとき、上記のようなフィルタをかけると時系列データの時間についての m 次の微係数が求まる。すなわち、0次の項 a_0 は時系列データの平均的傾向を、1次、2次の項はそれぞれ時系列データの1次、2次の変動傾向を表わしている。本論文では主として1次、2次の項 a_1, a_2 を考える。係数は時刻ごとに求まる。そこで、時刻 t における1次、2次の係数を a_{1t}, a_{2t} とする。

次に、この係数を正規化フィルタによってそれぞれ $[0, 1]$ 空間に変換する。これは後にこの変換した値を $[0, 1]$ 空間の値を扱う多層ネットへ入力して変換値の将来の値を予測することを目的としているためである。この変換には次のような sigmoid 関数を用いる。

$$q_{it} = \frac{1}{2} \left[1 + \tanh\left(\frac{a_{it} - \bar{a}_i}{\varepsilon \sigma_i}\right) \right] \quad (10)$$

$$\bar{a}_i = \frac{1}{N} \sum_{i=1}^N a_{it} \quad (11)$$

$$\sigma_i^2 = \frac{1}{N} \sum_{i=1}^N (a_{it} - \bar{a}_i)^2 \quad (12)$$

ε : 定数
 $i = 0, 1, 2$

この変数は、多層ネットによる予測値を復元(defuzzy)し、もとの係数の値を求めるのに適している。すなわち、予測値 q_i が得られたとき、もとの係数の値を求めるのに適している。すなわち、予測値 $\hat{q}_i$ がえられたとき、次式(13)によって係数 a_i の予測値 $\hat{a}_i$ が得られる。

$$\hat{a}_i = \bar{a}_i + \frac{\varepsilon \sigma_i}{2} \log\left(\frac{\hat{q}_i}{1 - \hat{q}_i}\right) \quad (13)$$

$i = 0, 1, 2$

時系列データ $\{a_{1t}, a_{2t}\}$ ($t = 1, \dots, N$)が与えられたとき、これらの値によって、時系列データが「上昇中」なのか、「平行」なのか、「下降中」なのかが分かる。すなわち、 $a_i > 0.0$ ならば「上昇中」であり、 $a_i = 0.0$ ならば「平行」であり、 $a_i < 0.0$ ならば「下降中」である。表1には a_1, a_2 の値と状態名(記号)との対応を示した。すなわち、 a_{1t} の状態は $a_{1t} > 0.0$, $a_{1t} = 0.0$, $a_{1t} < 0.0$ によって分類する。 $a_{1t} > 0.0$ のときは「上昇中」、

$a_{1t} = 0.0$ のときは「平行」、 $a_{1t} < 0.0$ のときは「下降中」と表現できる。1次微分係数および2次微分係数を考えると9通りの状態が抽出され、それぞれの状態に名前をつけることができる。すなわち、時系列現象はこの9つの状態の遷移として表現することができる。このように時系列データにフィルタをかけ、特徴量を抽出し、この特徴量の予測値を数値として求め、この値を記号に変換するという方法により、時系列現象のパターンとしての遷移をとらえることができる。従来は時系列現象の予測値は数値としてのみ求められていた。このため、将来の状態のパターンを把握することは容易ではなかった。提案方法においては、時系列現象の予測値を記号として求めることによりパターンを容易に把握することができる。

表1 時系列パターンと記号

Table.1 Time Series Patterns and Symbols

Symbol	a_1	a_2	Name of Pattern
1	$a_1 > 0.0$	$a_2 > 0.0$	up, rapidly
2	$a_1 > 0.0$	$a_2 = 0.0$	up
3	$a_1 > 0.0$	$a_2 < 0.0$	up, gradually
4	$a_1 = 0.0$	$a_2 > 0.0$	up, flat
5	$a_1 = 0.0$	$a_2 = 0.0$	flat
6	$a_1 = 0.0$	$a_2 < 0.0$	down, flat
7	$a_1 < 0.0$	$a_2 > 0.0$	down, gradually
8	$a_1 < 0.0$	$a_2 = 0.0$	down
9	$a_1 < 0.0$	$a_2 < 0.0$	down, rapidly

<3・2> クラスタ分析による時系列パターンの分類と特徴量の抽出

いま、前節の如く、時系列現象 $(X_1, X_2, \dots, X_N)$ に時間の多項式を用いたフィルタを用いて取り出した係数の列を変換して得られた時系列データの列

$(q_{11}, q_{21}), \dots, (q_{1N}, q_{2N})$

が得られたとき、この時系列データの変動の特徴を表わす少数のパターンを抽出することを試みる。これは、多層ネットによる学習には多くの時間がかかるため、少数の特徴量を抽出して、それを教師データとして学習することが必要となるからである。このため、得られた時系列データ $\{q_{11}, q_{21}\}$ ($t=1, \dots, N$) をクラスター分析により変動傾向ごとに分類する。

本論文では、クラスター分析の手法として非階層的手法の1つである k -means 法を用いて時系列データを分類する。アルゴリズムは次のとおりである。すなわち、

(1) 初期点 (初期重心) を n 個定める。ただし、次数を l としたとき、

$$q_{1k_i}, q_{2k_i}, \dots, q_{1k_{i+1}-1}, q_{2k_{i+1}-1} \\ i=1, \dots, n$$

の如くおく。

(2) 各点 (ただし、1次元の連続した時系列データのベクトル) から最短の初期点を求め、初期グループを決める。

(3) 各グループの重心を求める。

(4) 各点から最短の重心を求め、新たなグループを決める。

(5) (3)、(4) のステップを M 回繰り返す。繰り返し回数 M は、これ以上繰り返しても値に大きな変化が見られなくなるまでとする。

これによって、特徴量である少数のパターン

$$r^i = (q_{11}^i, q_{21}^i, \dots, q_{1l}^i, q_{2l}^i) \\ i=1, \dots, n$$

が得られる。

クラスター分析によって分類すると、時系列データ $(X_1, X_2, \dots, X_N)$ の代表パターンがわかる。代表パターンの遷移の関係を学習することによりあるパターンから次のパターンを予測することができる。

<3・3> 誤差逆伝搬法を用いた時系列パターンの遷移の学習

時系列現象の少数の特徴量が抽出されたならば、3層構造の多層ネットを用いた誤差逆伝搬法によって将来の値を予測する。この際の入力データは k -means 法で求めた特徴量である少数のパターン r^i ($i=1, \dots, n$) であり、出力データは予測値である特徴量である。ここでは75日分を学習させ、75日以後の10日間を予測した。

3層のネットワークを考え、 k 層の第 i ユニットへの入力の総和を I_k^i 、出力の総和を O_k^i とし、 $k-1$ 層の第 i ユニットから k 層の第 j ユニットへの結合の重みを $\omega^{k-1, k}_{ij}$ とする。各ユニットの入出力関係を与える関数を f とすると、

$$O_k^i = f(I_k^i) \quad (14)$$

$$I_k^j = \sum \omega^{k-1, k}_{ij} O_k^i \quad (15)$$

を得る。誤差関数 E を最小化するため、各パターンに関する誤差関数の勾配ベクトルを計算し、 ω を改良するために確率的降下法を用いた。

図1 予測パターン

Fig.1 Predicted Pattern

4. 数値実験による提案方法の評価

<4・1> 対象とする時系列データと提案方法

本節では、提案方法である非定常な時系列現象である株価の変動から特徴量を取り出し、しかる後に誤差逆伝搬法を用いる提案方法と従来方法のARモデルとで性能を比較する。

対象としたデータは1984年1月から1988年12月までの6か月ごと、75日間の9個の電機銘柄の株価である。 n 個 ($n=20$) の長さ (次数) l ($l=4$) のパターンを教師データとし、多層ネットを学習させ、75日以後の10日間の値を予測した。そして、次のような実現値と予測値とを平行移動させて求める相関係数で評価した。これは実現値と予測値そのものが問題なのではなく、多少ラグがあっても実現値のパターンとの類似性を問題とするためである。

$$\rho = \frac{\sigma_{12}}{\sigma_1 \sigma_2} \times 100 \quad (16)$$

ρ : 相関係数
 σ_{12} : 実現値と予測値との共分散
 σ_1 : 実現値の標準偏差
 σ_2 : 予測値の標準偏差

<4・2> 数値実験の結果と評価

予測値と実現値とをプロットしてみると、従来方法では単調増加、または、単調減少の予測傾向しか得られないのに対し、提案方法ではある程度のパターンをとらえた予測傾向が得られていることが分かる。目視によると、提案方法の場合、実現値と予測値とはラグを考慮すると極めてよく一致している（図1）。すなわち、目視によると提案方法は従来方法よりもパターンの予測という点で優れている。表2は、1984年から1988年までのラグを考慮した相関係数の値を示したものである。相関係数が大きく、かつ、100に近い方が予測精度は高い。＊は相関係数の大きい方を示したものであり、■は相関係数が大きく、かつ、その値が50以上であるものを示したものである。期間Ⅰは1月4日からの75日間を表わしており、期間Ⅱは7月1日からの75日間を表わしている。上段は提案方法、下段は従来方法の結果を表わしている。各手法とも実現値との平行移動は3日までであり、その間で値が最大となるものをとった。この表より、実現値と予測値とを平行移動させて相関をとっ

てみると、相関係数の大小だけでは提案方法は従来方法より52対36で優位であることが分かる。相関係数が大きく、かつ、その値が50以上であるものの数を比較してみると、提案方法は従来方法よりもやや優れた性能をもつことが結論される。

<4・3> 予測パターンの記述

本論文では、時系列データに時間についての多項式を用いたフィルタをかけ、記号に変換することの容易な特徴量を抽出し、この特徴量の予測値を数値として求め、その値を用いて記号としての予測値を求めるという方法を提案した。従来のARモデルのみによる予測では数値としての予測値しか得られなかった。従って、時系列現象の変動傾向のパターンを把握するのは難しかった。しかしながら、提案方法では数値としてのみならず、表1に示したように「上昇」、「平行」、「下降」を表わす記号として予測値を得ることができる。予測値を記号として定性的に表わした方が時系列現象の変動の把握は容易である。表3は図1に対応して提案方法による数値としての予測値と記号としての予測値を示したものである。1985年1月4日からの76日目から85日目までを予測したもので、銘柄番号は6501である。図1をみると提案方法はパターンの予測という面で実現値をよくとらえている。しかしながら、これを数値で表現したのではパターンの把握は難しい。記号として表わせば予測傾向は比較的容易につかめる。この例では記号列は上昇から下降に転じ、再び上昇するという傾向をよく表わしている。このように、時系列現象の状態を記号によって定性的に表現することに提案方法の特徴の1つがある。

表2 評価結果

Table 2 Result of Estimation

Stock No.	84		85		86		87		88	
	I	II	I	II	I	II	I	II	I	II
6501	64.8*	66.6*	88.0*	33.9	49.9*	49.3*	61.1*	70.6	17.5	-19.7*
	-74.7	16.0	-90.2	56.8*	-84.1	32.5	54.2	77.1*	72.3*	-56.7
6502	57.4*	12.5*	-5.2*	33.5*	35.8*	87.2*	69.5*	57.9*	-28.4	18.1*
	-65.9	-27.7	-87.0	-21.4	-71.3	86.5	-72.7	-42.0	23.1*	-45.4
6503	66.2*	29.3*	85.3	-32.3	8.0*	65.5	59.5*	57.8*	67.8	-7.2
	-64.2	-17.7	94.6*	65.2*	5.6	92.9*	-52.6	55.7	76.0*	13.2*
6504	35.0	0.9*	78.2*	-5.3*	38.7	8.5*	-14.8	59.6*	28.4	53.0*
	83.8*	-2.4	76.8	-18.7	-69.0*	-25.6	-1.8*	-49.1	77.7*	-64.4
6701	57.3*	21.4*	-40.3	25.8*	52.3*	87.7*	29.2*	63.0	-44.3	5.5
	-25.8	-10.3	17.8*	-91.3	-85.3	-87.2	-43.9	84.9*	82.4*	36.6*
6702	65.4*	-15.9*	27.6*	69.7*	-9.3*	65.7*	-7.0	62.2	-27.4	-31.9
	13.6	-45.0	3.3	-92.2	-50.6	54.9	50.1*	83.2*	78.8*	-22.4*
6752	53.3*	47.8	80.2*	46.5	50.3*	66.5*	31.3	37.5	-2.1	-37.6
	-73.8	64.9*	-22.8	72.6*	-91.3	39.8	65.1*	84.8*	84.1*	-15.2*
6758	81.4*	-20.6	78.9	35.9	52.1	36.7	82.2*	70.1*	7.9	36.7*
	-75.4	25.3*	78.9	56.5*	77.1*	69.7*	57.3	0.6	69.5*	0.3
6773	82.6*	-30.8*	-33.2	-23.2	66.5*	88.5*	54.5*	-3.6	2.0	46.8*
	-92.4	-61.6	-3.9*	71.1*	-66.0	-40.2	42.0	92.4*	67.7*	-6.6

表 3 予測値の列
Fig.3 Series of Predicted Value

Day	Predicted Value	Predicted Pattern
1	815	9
2	817	9
3	820	5
4	823	1
5	824	1
6	824	1
7	822	2
8	820	6
9	817	9
10	813	9

5. おわりに

本論文では、時系列現象を定性的に表現し、予測する定性的モデリングについて考察した。時系列データが与えられたとき、この時系列データに時間についての多項式を用いてフィルタをかけ、時系列データの変動傾向を時間による微分係数という形で求める。時間による1次微分、2次微分を表わす1次、2次の係数を取りあげ、これを後に多層ネットを用いて予測するために、

[0, 1] 空間に変換し、正規化した。そして、時系列現象の特徴量の列から変動傾向の特徴を表わす少数のパターンを非階層型のクラスター分析の手法である

k-means 法を用いて抽出した。そして、この少数のパターンを教師データとし、多層ネットを用いて、時系列現象の将来の状態を求めた。さらに、微分係数の正、0、負によって時系列現象を変動ごとに分類し、変動傾向を「上昇」、「平行」、「下降」といった状態名(記号)として求め、時系列現象の将来の状態をこの状態名(記号)として表わした。本論文では、株価の変動は必ずしもランダム・ウォークではなく、同定、予測可能性があることをふまえて、時系列現象として株価変動を取りあげた。従来方法としてARモデルを取りあげて、提案モデルの性能を評価することを試み、実現値と予測値とを平行移動させて求めた相関係数を用いると提案方法は従来方法よりもパターンの予測という点でやや高い精度を有していることを確認した。

なお、本報告における解析では、(株)日本経済新聞社のNEEDSデータベース株価データを用いた。

(平成10年7月3日受付、平成11年1月13日再受付)

参考文献

- (1) H.Akaike : A New Look at the Statistical Model Identification, IEEE transactions on Automatic Control, AC-19, Vol. 6, 716-723 (1974)
- (2) D.E.Rumelhart, J.L.McClelland et al.: Pararell Distributed Processing, MIT Press (1986)
- (3) Kitagawa, G. : Non-Gaussian State-Space Modeling of nonstationary Time Series, Journal of the American Statistical Association, Vol.82, No.400, 1032-1063 (1987)
- (4) Rabiner, L.R. and Juang, B.H.: An Introduction to Hidden Markov Models, IEEE ASSP Magazin, Jan. 4-16, (1986)
- (5) 大橋靖雄：分類手法概論、計測と制御、第24巻、第11号、999-1006 (1985)

直江 秀憲

(正員) 1958年生。

83年東京大学理学部情報科学科卒業。86年筑波大学大学院修了。同年(株)日立製作所入社。システム開発研究所においてニューラルネットワーク、時系列分析の研究に従事。