

多段ファジイ推論を用いたカオス時系列予測とその性能 評価

*九州大学経済学部 岸川 善紀 KISHIKAWA Yoshinori

1 まえがき

本論文では多段ファジイ推論を用いたカオス時系列予測と、その応用について述べる。入力を数段階に分散して用いる多段ファジイ推論は、1段のシステムと比較してルール数を大幅に削減でき、しかも性能が同程度であるので、入力変数を多く用いる決定問題に適している [1][3]。一方、カオスは決定論的にシステムが記述されるので、ファジイ推論で関数近似を変換することにより、予測を行なう問題として取り扱える。すなわち、経済時系列解析の分野においては、線形あるいは非線形の確率過程モデルを用いて時系列の特徴づけや予測がなされるが、カオス理論では決定論的なモデルを用いているので、方程式が推定されると、将来の値の予測は簡単な計算により実行できることになる。

多段ファジイ推論を時系列予測へ応用する場合、カオスとしてモデル化できる方程式系を推定したのち、これを用いて時系列の予測を行う。予測誤差などを用いてカオス理論の性能を検定することも重要であるが、予測値と観測値との相関を用いて時系列へのモデルあてはめの適切さを分類することも行われている。

応用としてレスラー方程式で生成される時系列の予測をとりあげる。更に、いくつかの実際的な時系列予測へ適用する。2では多段ファジイ推論について述べる。3ではGAによるメンバーシップ関数の最適化および、逆伝搬法 (back propagation) を用いたルールのウェイトの最適化を合わせて述べる。4

ではカオス時系列とその予測について述べる。5では応用としてレスラー方程式で生成された時系列やその他の時系列について実際に予測を行なってみる。6ではアトラクタ再構成による時系列予測と比較することにより多段ファジイ推論システムの予測の性能を評価する。

2 多段ファジイ推論

最初に、多段ファジイ推論のルールとその出力について整理しておく。いま N 段階の多段ファジイ推論を考え $i (i = 1, 2, \dots, N)$ 段目のルール集合には $n_i (i = 1, 2, \dots, N)$ 個の if-then ルールがある。

If x_1 is A_{11}^1 and ... and x_M is A_{1M}^1
then y_1 is w_1^1

If x_1 is A_{11}^2 and ... and x_M is A_{1M}^2
then y_1 is w_1^2

.....

If x_1 is $A_{11}^{n_1}$ and ... and x_M is $A_{1M}^{n_1}$
then y_1 is $w_1^{n_1}$

.....

If x_1 is $A_{N1}^{n_N}$ and ... and x_M is $A_{NM}^{n_N}$ and y_{N-1} is $B_N^{n_N}$

then y_N is $w_N^{n_N}$ (1)

ここで $x_j (j = 1, 2, \dots, M)$ はそれぞれの段階における入力であり、通常は各段において x_j の一部のみが入力変数として用いられるとする。 A_{ij}^k はファジイ集合である。数値 w_i^k はそれぞれのルールのウェイトである。 i 段における出力 $y_i (i = 1, 2, \dots, N-1)$

は中間的な変数であり、これは次の段における入力として用いられる。最終的な N 段における出力 y_N がシステムの出力である。このとき、簡易推論では i 段の出力は次のようにして計算される (defuzzification)[2]。

$$\mu_i^k = \prod_{j=1}^{M^*} \mu A_{ij}^k(x_j) \quad (2)$$

$$y_i = \frac{\sum_{k=1}^{N_i} \mu_i^k w_i^k}{\sum_{k=1}^{N_i} \mu_i^k} \quad (3)$$

ここで μA_{ij}^k はファジイ集合 A_{ij}^k のメンバーシップ関数であり、 μ_i^k は、 i 段目、 k 番目のルールの適合度である。 M^* は i 段階への入力数であり、前段からの出力を入力として用いているため、その最大値は $M+1$ である。しかし、通常は入力の一部しか利用しないので M^* は $M+1$ より十分に小さい。

μ_i^k をウェイト w_i^k で掛けながら合計すると出力がえられる。メンバーシップ関数 $\mu B_i^k(y_{i-1})$ は出力変数 y_N のため $\mu A_{ij}^k(x_j)$ の代わりに用いられる。

以下の議論ではウェイト w_i^k とメンバーシップ関数の形状がシステムの性能を左右する。 w_i^k はニューラルネットワーク設計における逆伝搬法を用いて最適化できる。メンバーシップ関数の形状は GA を用いて最適化する。これらの最適化手法は交互に使用される。

最初、 w_i^k を固定しメンバーシップ関数を最適化し、この後メンバーシップ関数の形状を固定し、 w_i^k を最適化する、このステップを繰り返しおこなうのである。以下の式で、 y は最終端の出力であり、 y^B は学習に用いる B 番目の入力変数の組に対し、望ましい出力の値として表される。

$$J = (y - y^B)^2 / 2 \quad (4)$$

を最小化する w_i^k は次のように与えられる。

$$\delta_i = y - y^B \quad (5)$$

$$\Delta w_{i-2}^k(t) = -\alpha \delta_{i-1} \frac{\mu_{i-2}^k}{\sum_k \mu_{i-2}^k} + \eta \Delta w_{i-2}^k(t-1) \quad (6)$$

$$\delta_{i-1} = \sum_k \delta_i \frac{\sum_k w_{i-1}^k - \sum_k w_{i-1}^k \mu_{i-1}^k}{(\sum_k \mu_{i-1}^k)^2}$$

$$\frac{\partial \prod_{j=1}^{M^*} \mu A_{i-1j}^k(q_{i-2}^j)}{\partial q_{i-2}^k} \quad (7)$$

q_i^k は i 段、 k 番目の入力変数、 α, η は収束のためのパラメータである。この式の導出は文献 [10] に示されている。

3 GA によるメンバーシップ関数最適化

メンバーシップ関数の形状として三角形を仮定し、その特徴点は底辺の 2 つの座標 b_1, b_2 と、数値 1 をとる座標 b_c の 3 つをとり、これを次のようなストリングで表現したものを個体とする [4]。

$$b_1^1, b_c^1, b_2^1, \dots, b_1^K, b_c^K, b_2^K$$

ただし、これらの変数の上の添字はすべてのメンバーシップ関数に対応し、第 1 段の第 1 番号ルールの第 1 変数から始まり、最終的に第 N 段の第 n_N 番号ルールの第 M^* 変数まで変化する。 K はこれらの合計である。簡単のため $b_1^1, b_c^1, b_2^1, \dots$ を順序通り並べて、 s_1, s_2, \dots として記述することにする。個体の交差処理、突然変異処理は通常の GA と同じである。

ウェイト w_i^k の最適化を含めてアルゴリズムとして示すと次のようになる。

(Step1): ウェイトとメンバーシップ関数の形状の初期値を与える

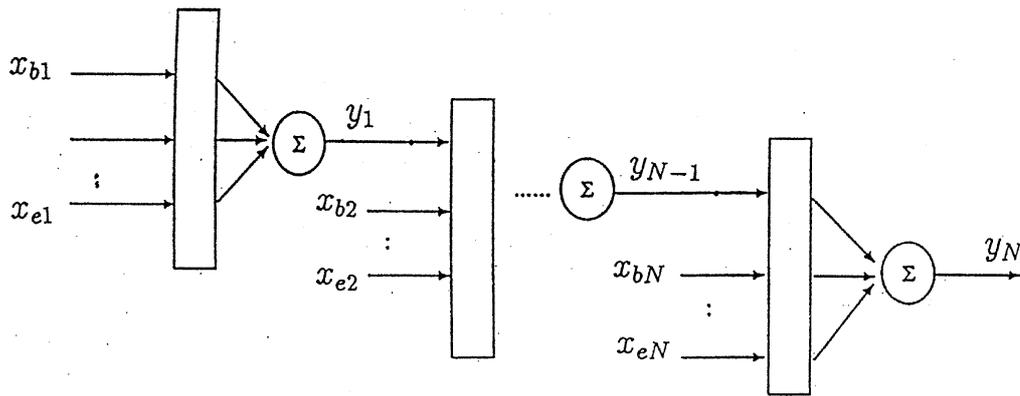
初期値としてルールのウェイト w_i^k とメンバーシップ関数の形状が与えられる。すなわち底辺と数値 1 をとる頂点の横座標 b_1, b_2 と、 b_c が与えられ、それを並べたストリングが GA における個体となる。

(Step2): ウェイトを最適化

ルールのウェイトを逆伝搬アルゴリズムを用いることにより最適化する。この操作の段階ではメンバーシップ関数の形状は固定されている。

(Step3): 個体ごとの適応度を計算

(2) の操作を必要回数行なった後 (その数は後に記す)、最適なウェイトの値が計算される。ここから



入力変数 x_j は x_1, x_2, \dots, x_M から選択

各 i 段について x_{bi}, \dots, x_{ei} を入力

図 1: 多段ファジイ推論システム

GA の操作が適用されるのであるが、まずは各個体の適応度を計算し、適応度の高いものの順に並べ替える。

(Step4): 適応度順に個体ペアを交叉処理, 新個体生成

より高い適応度をもつ個体の中から一对の組を選択し, 交叉を適用する。また, 個体群の中から全く新しい個体を発生させるために突然変異操作をおこなう。突然変異の発生確率は 0.15 とする。

(Step5): 適応度の低い個体の置き換え

GA 操作のために新たに発生した個体と, 現在ある個体で適応度の低い個体とを入れ換える。個体全体の 1/4 を入れ換える。新たに適応度順に並べ替える。

(Step6): GA を必要回数だけを適用 ((2) から (5) まで)

GA を適用する条件は以下のとおりである。

- i) GA の個体数は 50
- ii) 突然変異の確率は 0.15
- iii) 個体の 1/4 を入れかえる
- iv) BP (back propagation) は 100 回繰り返し
- v) GA 操作は 30 回で終了

BP による学習は GA 操作毎に 100 回行なわれることとなる。GA の適用には適応度に比例して個体を選択されるルーレット戦略を用いる。

4 カオス時系列

カオスは, 制御理論の分野などにおいて, もともと古くからその存在が知られていたが, 現実に数理モデルとして利用できるものとして認識されてはならず, 取り扱いに困るノイズとして考えられていた。しかし, ローレンツ (Lorenz) により気象現象を記述する数理モデルとして注目されるようになってから, 不確実な現象を説明する理論として研究されている [5]- [7]。

経済学分野では, 経済時系列における不規則性を説明するモデルとして応用されたり, 不均衡経済マクロモデルにおける変動のモデルなどの分野へ応用されている。

カオス現象は, 1 つの不確実な現象であるが, ランダムな変化ではなく規則的な変動とも大きく離れてはいない。従って, 確率過程により生成される時系列のような不規則な変動ではなく, 時系列は周期的な波に収束もしないかわりに, 明確な不規則振動を示さず, これらの中間的な動きをする。従って, カオス理論が展開される以前には, やっかいな雑音 (ノイズ) と見なされたり, ARMA モデルなどの確率過程のモデルの変形として理解された。

しかし, このように観測されるデータの違いの基

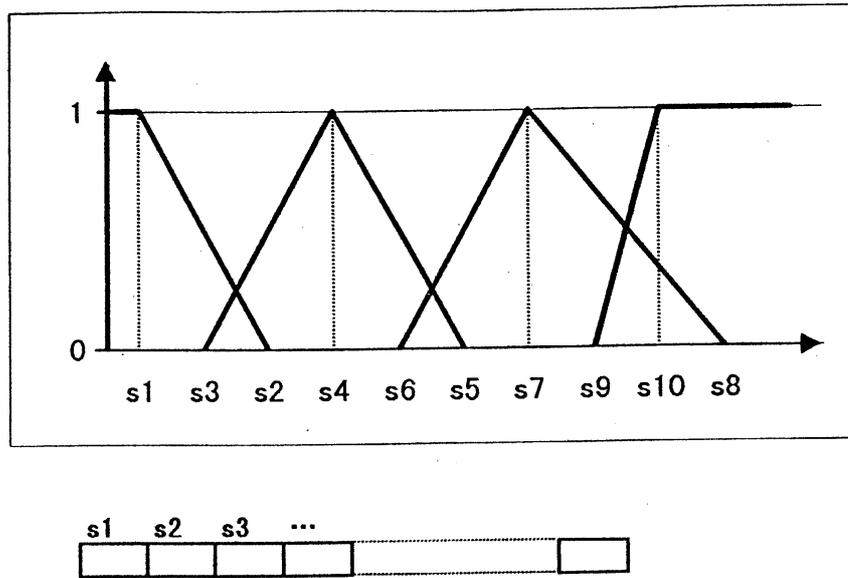


図 2: メンバシップ関数とストリング

礎となる理論は大きく異なっており、カオスは非線形方程式により記述されるシステムにより生成されるのに対して、確率過程はシステムを駆動する入力やノイズが確率的な変化をすることを前提としていることである。カオスは、いわゆる決定論的な方程式により生成され、その中には確率過程のノイズの不確実な要素が含まれない。

このようなことから、カオスを経済分野に応用する場合、方程式から出発するのか、観測データから出発するのかの、2つのアプローチがあると考えられる。その1つは観測されたデータ（例えば時系列データ）を処理する方法として用いることであり、もう1つはあらかじめカオス発生モデルを仮定して、現実に観測された現象をこのモデルにより説明できるかを調べることである。今回は方程式のモデルを仮定し、それにより生成される時系列を予測するものである。

カオスは方程式により決定論的に生成されているので、時刻ごとに観測データが一意に決まる。このことにより、カオス時系列の2つの時刻を同時に図として描くことにより明確な形状が観測される。例えば、logistic map とよばれるカオスでは、時刻 t の値 $x(t)$ を用いて時刻 $t+1$ の値 $x(t+1)$ が生成されるので、 $x(t)$ 、 $x(t+1)$ は方程式により結ばれている [8][9]。従って、2次元平面に $x(t)$ を横軸、

$x(t+1)$ を縦軸に描くと、この方程式に対応する1つの閉じた曲線になる。これをアトラクタとよぶ。これに対して、確率過程のモデルにより発生された時系列では、 $x(t)$ 、 $x(t+1)$ は一般には相関がないので、2次元平面には明確な図形は現れない。

更に、変数の数が多くなった場合でも3次元平面、4次元平面に、生成されるメカニズム（正確にいえば n 次元多様体）に対応するアトラクタが描かれることになる。良く知られている例として、経済モデルでも良く引用されるレスラー（Rossler）方程式系では3次元の方程式で記述がなされるので、3次元のアトラクタとなる。

5 応用例

5.1 レスラー系で生成された時系列の予測

確率的な手法の特徴は、モデルを駆動する外力としてノイズを仮定することであり、常に不確実性を含んでいる。従って、モデルを同定したあとのシミュレーションにおいても、モデルとは無関係のノイズを導入する必要がある。これに対して、カオスによる解析では、モデルが確定的に得られるので、ノイ

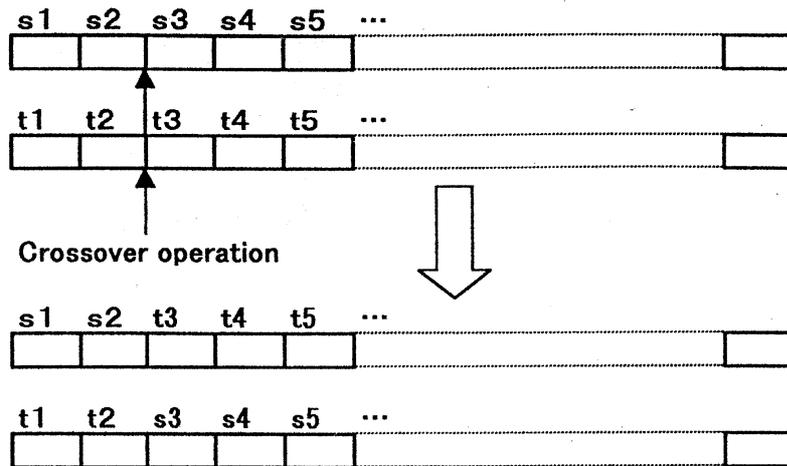


図 3: GA 操作 (交叉)

ズを含ませる必要はなく、計算も簡単であり、モデルの意味も分かりやすい。更に、線形や非線形のモデルにより時系列モデルを推定する場合に特有の問題として生じる位相のずれに起因する問題が発生しない。

これは、例えば ARMA モデル、線形予測などにおいて、現在までの観測データを用いて将来の値を予測する場合に、情報が過去にかたよっているために、どうしても予測に時間的なずれ(位相のずれ)が発生してしまうことが理論的に証明されていることである。これに対してカオス理論では、モデルが決定論的なモデルであるので、このような位相のずれは発生しない。

もともと、時系列がカオス的な性質をもっている場合には、初期値に対してシステム応答が敏感であるなどの理由により、長期的な時系列の予測は困難である。しかし、システムの方程式は決定論的に求められるので、短期的にはこの特性を生かして精度のよい予測が可能となる。また、短期的予測だけでなく、長期的予測に関しても、ダイナミックスが決定論的に求まるので、条件が整理されていれば、精度の良い予測が得られることが期待される。

以下では、レスラー (Rossler) 方程式により生成

される時系列の予測問題を考察する。

$$\frac{dx}{dt} = -y - z \quad (8)$$

$$\frac{dy}{dt} = x + \alpha y \quad (9)$$

$$\frac{dz}{dt} = \beta - \gamma z + xz \quad (10)$$

これに対し多段ファジイを適用した結果をまとめる。

時系列のサンプル数を 1000 とし、次の 2 つの学習条件を与える。予測する時系列を $x(t)$ としてみると、入力には現在までの $x(t)$ の値 9 個を用いる。多段ファジイ推論システムの段数を 3 とし、それぞれの入力個数を 3, 3, 3, ..., とする。

ケース 1: 最初の 700 個を用いて学習残りでの予測を検証

ケース 2: 最初の 200 個を用いて学習残りでの予測を検証

図 4.5 にはそれぞれケースの場合の予測の状況を示している。これより分るように、ケース 1 では良好な結果を与えている。ケース 2 でも満足できる結果である。

また、残りの $y(t)$, $z(t)$ についても同様に予測を行なう。予測の精度の評価は観測値との二乗誤差によって行なう。表 1 にそれを示す。

表 1. レスラー方程式で生成される時系列の 1 期先の予測誤差

	$x(t)$	$y(t)$	$z(t)$
700 個学習	0.035	0.048	0.003
200 個学習	0.142	0.124	0.335

ここで、予測誤差は次の式で与えられる。

$$error = \left[\sum_{t=1}^T (x(t) - \overline{x(t)})^2 \right]^{1/2} / \sigma \quad (11)$$

5.2 観測された時系列の予測

また、実際に観測されたいくつかの時系列に関して多段ファジイを適用してみる。レスラー方程式系の時と同じくサンプルが 1000 個以上あるものに関しては上の 2 つのケースと同様な条件で予測を検証し、サンプル数が足りない場合についてはケース 1 としてサンプルの最初の 7 割を学習に用いて残り 3 割で予測を検証した場合、ケース 2 としてサンプルの最初の 2 割を学習に用いて残り 8 割で予測を検証した場合とする。ここでは、1 期前までの観測値までを予測に用いる 1 期先予測のみではなく、予測値をその次の期の予測に用いて、逐次的に予測値を求める n 期先予測も行なっている。以下が、多段ファジイ推論を用いた場合の予測誤差である。

表 2. ファジイ推論による 1 期先の予測誤差

学習	ushiki	henon	att	sales	sunspots
7 割	0.14	0.22	0.05	0.12	0.19
2 割	0.31	0.42	0.05	0.23	0.55

表 3. ファジイ推論による n 期先の予測誤差

学習	ushiki	henon	att	sales	sunspots
7 割	0.15	0.24	0.05	0.16	0.20
2 割	0.33	0.46	0.05	0.29	0.58

カオス時系列の場合、初期値鋭敏性があるため長期予測になる程予測精度が急激に落ちることが考えられているが、ここでは n 期先予測の場合も 1 期

先予測の場合と比較してもそれほど悪くなっていない。詳しい検証はまだ行っていないが、ファジイ推論を用いた n 期先予測の応用の可能性も考えられる。

なお、GA を用いないでウェイトの最適化だけを逆伝搬法を用いて最適化する場合、あるいは通常モデルを用いた予測の場合との比較は省略しているが、ここにあげた結果より悪くなっている。

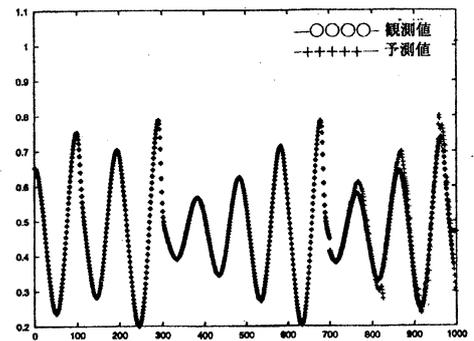


図 4. $x(t)$ の予測結果 (ケース 1)

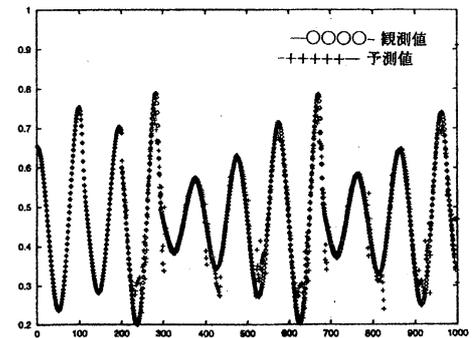


図 5. $x(t)$ の予測結果 (ケース 2)

6 アトラクタ再構成による時系列予測との比較

アトラクタの将来の位置を予測する方法には、ヤコビ行列を用いた局所近似法, radial basic function 法, ニューラルネットワーク法などがあるが、いずれもアトラクタ上での推定を用いることは共通しており、予測精度を向上させるために従来方法との結合を行うかの違いであると言える。従って、以

下では局所近似法を中心として述べていく。

いま、 n 次元のアトラクタ空間(再構成状態空間)における点を vt とし、ここにおけるヤコビ行列(微分値行列)が計算されているとする。すなわち、もとのカオスのダイナミックスを

$$x(t+1) = F(x(t))$$

としたとき、変数 $x(t)$ の微小変化(摂動項)の最初の $D_f(x)$ がヤコビ行列である。

予測するまでの時刻の差 p が十分に小さいと仮定すると、 $x(t)$ の変化も微小であると言えるので、上にあげたヤコビ行列を用いて予測することが可能となる。ここで、ヤコビ行列が過去のデータから推定されていれば、微小変化の将来の値である左辺が得られるので、問題はヤコビ行列をどのように推定するかの方法となる。

これには、アトラクタ上で、現在の時刻における状態を求め、この近傍の点を利用してヤコビ行列を推定する方法を用いる。

以下では、この計算手順を分かりやすく示す。まず、観測された時系列 $x(t)$ が得られている場合に、Takens の定理を用いてアトラクタを構成するための系列を生成する。

$$X(t) = (x(t), x(t-L), x(t-2L), \dots, x(t-(m-1)L)) \quad (12)$$

このとき、 $m \geq 2 \nu + 1$ であれば、この生成された m 次元空間にアトラクタの構造が保存される(Takens の定理)。

いま、時刻 t までの時系列の観測データ $x(t)$ が得られているとき、 p 時刻先の値 $x(t+p)$ を予測する問題を考える。アトラクタ上での $x(t)$ に対応する点 $X(t)$ が与えられているときに、 p 時刻後のアトラクタ上での点 $X(t+p)$ を $X(t)$ を用いて予測する問題となる。

$$X(t+p) = F(X(t)) \quad (13)$$

ここで、関数 F は予測を行うための近似関数である。いま、時間間隔 p が十分に小さい場合には、アトラクタの時間変化は、ほぼ $X(t)$ の近傍点の時間変化により近似できると考えられるので、近傍の点

を複数個適当に選んで、これらの平均的な推移を最小2乗法などにより求めて、動きを推定する。

なお、再構成された空間では時間遅れ要素を用いて m 次元空間を構成するが、ここでは、時刻 $t+p$ の時系列の値だけ、すなわち、 m 次元空間の第1番目の要素 $x(t+p)$ だけを予測すればよく、その他の要素は予測する必要がないので、近似的な予測式は次のようになる。

$$x(t+p) = F(X(t)) \quad (14)$$

最も簡単な予測式として次のものを用いる。

$$x(t+p) = a_0 + \sum_{k=1}^m a_k x(t - (k-1)L) \quad (15)$$

この式における係数は次のように決められる。まず、 $X(t)$ を除くすべての $X(i)$ についての $X(t)$ とのユークリッド距離を計算する。

$$R = \text{dist}(X(t), X(i)), i \neq t \quad (16)$$

この R の小さい順に n 個の点を $X(t)$ の近傍点として採用し、

$$X_{\tau_1}, X_{\tau_2}, \dots, X_{\tau_n} \quad (17)$$

としておく。このもとで、次の式を最小にする係数を求める。

$$\sum_{i=1}^n (X_{\tau_i+p} - F(X_{\tau_i}))^2 \rightarrow \min \quad (18)$$

この計算は、具体的には、 X_{τ_i+p} の1番目の要素 x_{τ_i+p} を n 個集めたベクトル x と、 $n \times (m+1)$ 列のヤコビ行列と、 $m+1$ 個の要素からなる係数ベクトルを用いて次の観測方程式をたてることになる。

$$x = GY \quad (19)$$

ここで

$$x = (x_{\tau_1+p}, x_{\tau_2+p}, \dots, x_{\tau_n+p}) \quad (20)$$

$$G = \begin{pmatrix} 1 & x_{\tau_1} & x_{\tau_1-L} & \dots & x_{\tau_1-(m-1)L} \\ 1 & x_{\tau_2} & x_{\tau_2-L} & \dots & x_{\tau_2-(m-1)L} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{\tau_n} & x_{\tau_n-L} & \dots & x_{\tau_n-(m-1)L} \end{pmatrix}$$

(21)

$$a = (a_0, a_1, \dots, a_n) \quad (22)$$

係数ベクトルは、式(18)に示される左辺と右辺の差異が最小2乗の意味で最小となるように求められる。従って、結論的にこの定式化を振りかえると、時系列の線形予測を複数地点で行うことにより、線形予測の係数を求めることになっており、その意味では通常の予測手法を大きくかわるものではない。しかし、アトラクタで近傍に位置する点を用いてパラメータを推定していること、また、このことが時系列がカオスであるのと最大限生かしていることなど、大きな相違点となっている。

式(19)におけるヤコビ行列 G を Householder 法などを用いて QR 分解することにより、係数 a を計算できる。例えば、Fortran プログラムでこの問題を解くには、サブルーチンライブラリなどに登録されている Householder 法を用いることが有効である。この係数 a を式(15)に代入することにより、時系列に値 $x(t+p)$ を予測することができる。ここで、式(19)が解を持つための条件として次がある。

$$n \geq m + 1$$

n の選び方のより予測精度が大きな影響を受けることが知られている。すなわち、 n が小さいと最小2乗の計算過程で大きな誤差を生じることになるが、一方、大きくしすぎると、 $X(t)$ からかなり離れた点までも用いて予測することになるため、予測誤差は拡大する。現実には試行錯誤的に求める場合が多いが、いくつかの研究では自動的に設定できるように工夫しているケースもある。

ファジイ推論の場合とアトラクタ再構成の手法を用いて予測した場合の結果との比較は以下の通りとなる。

表4. ファジイ推論, アトラクタ法による予測誤差

	ushiki	sunspots
ファジイ推論	0.14	0.19
アトラクタ法	0.23	0.44

どちらの場合もファジイ推論の方が結果が良好であることがわかる。

7 むすび

今回は、多段ファジイ推論を用いたカオス時系列の予測をレスラー方程式によって生成される時系列や実際に観測された時系列に対して適用し、その性能の評価をおこなった。また、アトラクタ再構成によるカオス時系列予測の手法との比較もおこなった。

今後の課題として n 期先予測への応用や、非カオス時系列に対する多段ファジイの適用可能性などが考えられ、検証を進める予定である。

参考文献

- [1]L.A.Zadeh:"Fuzzy sets", Information and Control,8, pp.338-353 (1996).
- [2]S.Murakami and M.Maeda:"Automobile speed control system using a fuzzy logic controller", in Industrial Applications of Fuzzy Control, ed. M.Sugeno, pp.105-123, Elsevier Science Publishers (1985).
- [3]L.Wang and J.M.Mendel:"Generating fuzzy rules by learning from examples",IEEE Trans.,System,Man, Cybernetics, Vol.22,No.6,pp.1414-1427 (1992).
- [4]K.Tan and S.Tokinaga:"Optimization of fuzzy inference rules by using the Genetic Algorithm and its application to the bond rating", Proc. JORSJ,Vol.42,No.2 to appear(1999).
- [5]合原一幸編:「応用カオス」,サイエンス社(1994)
- [6]合原一幸編:「カオス」,サイエンス社(1990)
- [7]E.E.Peters: Chaos and Order in the Capital Markets, John Wiley nad Sons, Inc.(1991) (邦訳: 新田功訳「カオスと資本市場」,白桃書房)
- [8]E.N.Lorenz:"Deterministic non-periodic flows", Journal of Atmos. Science, 20, pp.130-141 (1963).
- [9]O.E.Rossler:"An equation of continuous chaos", Physical Letters, 57A, p.397 (1976).
- [10]Y.Kishikawa and S.tokinaga:"Prediction of Chaotic Time Series by using Multi-stage Fuzzy Inference System optimized by the GA", Trans.IEICE Section E to

appear(1999).