

# ARモデルを用いたカオス時系列予測法の再検討

長瀬 隆 久†

本研究では, ARモデルを用いたカオス時系列予測法を再考するが, そのような手法としては SETAR やダミー変数を加えた AR (DAR) モデルなどがある. しかし, DAR に関してはその限界も指摘されている. そこで, 本研究では, まず DAR の改善法を検討し, さらに後半では SETAR との特徴や精度の違いを示した. 次に, 連続系のカオスモデルを AR 系のモデルで予測することも検討し, テーラー展開を用いたパラメータ与の AR モデルで良好な結果が得られた. また, 両検討において「ローレンツの類推法」や「Farmer と Sidorowich の方法」を簡略化した AR タイプのモデルとの比較も行った. そして最後にカオスが発生するモデルであっても, 観測する時間間隔が長いと確率過程として同定される場合があることを指摘し, カオスと確率過程について考察した.

## Reconsideration of Chaos Time Series Forecasting Method Using AR Model

TAKAHISA NAGASE†

In this paper we reconsider chaos time series forecasting method using the AR model. There are the SETAR or the AR model adding dummy variables (DAR) about this kind of methods. But it has been pointed the limitations about the DAR. In our study, at first we improved the DAR and compared the feature and accuracy with the SETAR. Secondly we investigated forecasting continuous chaos model with a series of the AR models. According to this analysis, the AR model using Taylor series expansion has good property. Moreover, in both examination comparison with the other models of the AR type which is simplified the Lorenz's method of Analogues and the Farmer and Sidorowich's method was also performed. Finally we pointed that there are some cases which chaos is identified stochastic process when time interval is long and thought the relation between chaos and stochastic process.

### 1. はじめに

近年, 統計的なモデルである「ARモデル」を用いたカオス時系列予測法において, Tong による SETAR (TAR)<sup>1)</sup> や長瀬によるダミー変数を加えた AR モデル<sup>2)</sup> (以下, DAR と略す) などが提案されている. 初期値への鋭敏性を有し, 複雑な挙動をするカオス時系列であるが, 区分線形近似によるこのような「簡便法」でも, ある程度予測することが可能であることが示されている. しかし, SETAR に関してはすでに研究がなされているが, DAR に関してはまだ提案されて間もないため, ほとんど研究されていない.

そこで, 本研究ではこの一連の AR 系のモデルによるカオス時系列の予測について考察するが, 特に DAR に焦点を絞る. そして, 以前から指摘されている<sup>2)</sup>

その欠陥の改善や連続系のカオスへの応用についてもふれ, また SETAR との比較も行った. その際「局的」な近似モデルとして用いられる「ローレンツの類推法」や「Farmer と Sidorowich の手法」を簡略化した手法との精度の検証も行った.

### 2. DAR の改善

#### 2.1 相関係数行列分析

DAR は, 離散型カオス時系列の予測を行う際に, 原系列の差分 (階差) をとることにより部分的に線形性を見出し, その線形部分に AR モデルをあてはめ, それ以外の部分 (残差の大きい部分) をダミー変数で処理する手法である. たとえば, ロジスティック写像

$$X_t = \alpha X_{t-1}(1 - X_{t-1}) \quad X_0 = 0.1 \quad (1)$$

において, パラメータが  $\alpha = 3.7$  の場合を考える. この時系列の 1 次階差をとった系列を  $X'_t$  として,  $X'_t$  と  $X'_{t-1}$  をプロットしたものが図 1 であり, このように階差をとると原点付近に部分的な線形性が現れることが知られており, この階差系列に DAR を適用す

† 中央大学大学院理工学研究科情報工学専攻  
Department of Information and System Engineering,  
Graduate School of Science and Engineering, Chuo University

ることが可能であることがすでに示されている<sup>2)</sup>。

しかし、ロジスティック写像でもパラメータが 4.0 などの場合、差分化しても図 2 のように非線形性がまだ強く残り、これは 2 次階差などをとっても除去されない。それゆえ、線形 AR モデルにダミー変数を加えても、あまり精度が上がらないという欠陥が同時に指摘されていた<sup>2)</sup>。そこで、本研究ではまずこの DAR の欠陥の改善を試みる。

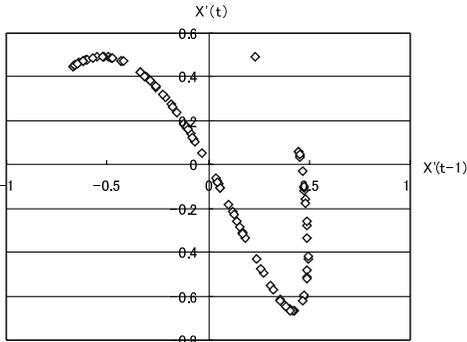


図 1 ロジスティック写像の 1 次階差 (α = 3.7)  
Fig. 1 The first-order difference of Logistic map (α = 3.7).

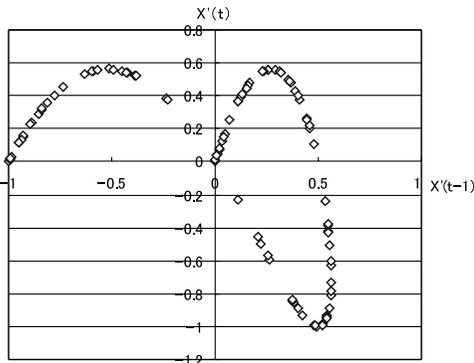


図 2 ロジスティック写像の 1 次階差 (α = 4.0)  
Fig. 2 The first-order difference of Logistic map (α = 4.0).

具体的には、1 次階差の自己回帰は精度があまりあがっていないことをふまえ、自己回帰型に限定せずに原系列や 2 次階差系列など、どのような系列を従属変数ならびに説明変数に設定すればよいのかをまず検討する。

今、ロジスティック写像 (α = 4.0) の原系列から 2 次階差系列までを考え、タイムラグは 2 まで考えとする。それらの系列すべての「相関係数行列」を求めると表 1 のようになる。この表の中で、たとえば「原(-1)」は原系列  $X_t$  に対してタイムラグ 1 の  $X_{t-1}$  の系列を示している。階差系列も同様である。ただし、注意しなければならないのは、説明変数にどの系列を持ってくるにせよ、それは従属変数に比べてかならずタイムラグが 1 以上遅れていなければ予測モデルとしては成立しないということである。つまり、この「相関係数行列」の中で実際上の分析で意味があるのは、「重複部分」を除いた「斜字体の部分」だけということである(観測時間が同じものどうしと、同じ系列どうしでタイムラグを平行移動させただけのものを除いたのが斜字体の部分)。

斜字体でさらに太字のところが関連の強いところであるが、この表 1 を眺めると、ロジスティック写像 (α = 4.0) においては、1 次階差にラグ 1 の原系列を回帰させてもよいが、2 次階差系にラグ 1 の原系列と同じくラグ 1 の 1 次階差および 2 次階差を回帰させるのが適当である。そのような設定で回帰分析を行うと、自由度調整済み決定係数(以下  $\bar{R}^2$  とする)は  $\bar{R}^2 = 0.842282$  と良好な値を示した。このような手法をここでは仮にカオス時系列の「相関係数行列分析」と呼ぶことにする。

通常、回帰分析においては  $R^2$  (決定係数) を用い、これは  $R^2 = 1 - \frac{\text{誤差変動}}{\text{全変動}} = 1 - \frac{\sum e_i^2}{\sum (Y_t - \bar{Y})^2}$  であるが、自由度調整済み決定係数は、 $\bar{R}^2 = 1 - \frac{(N-1) \sum e_i^2}{(N-m-1) \sum (Y_t - \bar{Y})^2}$  (N : 全データ数, m : 説明変数の数) である。これは

表 1 ロジスティック写像の相関係数行列

Table 1 Correlation coefficient matrix of logistic map.

	原系列	1 次階差	2 次階差	原(-1)	1 次(-1)	2 次(-1)	原(-2)	1 次(-2)	2 次(-2)
原系列	1								
1 次階差	0.729767	1							
2 次階差	0.412007	0.863091	1						
原(-1)	-0.06127	<b>-0.7271</b>	<b>-0.8463</b>	1					
1 次(-1)	0.015033	-0.49348	<b>-0.8652</b>	0.735526	1				
2 次(-1)	0.030401	-0.2652	-0.65532	0.417681	0.865232	1			
原(-2)	-0.08342	-0.00152	0.426051	-0.08155	-0.73522	-0.85496	1		
1 次(-2)	-0.03756	-0.03257	0.271679	0.009835	-0.5002	-0.86693	0.745691	1	
2 次(-2)	-0.01165	-0.03594	0.136522	0.040777	-0.27079	-0.65864	0.439165	0.868323	1

表 2 相関係数行列分析の結果  
Table 2 Results of the analysis of correlation coefficient matrix.

モデル	従属変数	説明変数	$\bar{R}^2$
ロジスティック (3.7)	2 次階差	原 (-1), 1 次 (-1)	0.963211
ロジスティック (4.0)	2 次階差	原 (-1), 1 次 (-1), 2 次 (-1)	0.841919
エノン写像	2 次階差	原 (-1), 1 次 (-1), 2 次 (-1)	0.917951
3 次の非線形写像	2 次階差	1 次 (-1), 2 次 (-1)	0.761811
池田写像	2 次階差	原 (-1), 1 次 (-1), 2 次 (-1)	0.973838
指数型 AR モデル	2 次階差	原 (-1)	0.950746

モデルの自由度を考慮して説明変数の数と推定精度のバランスをとった指標である (AIC が使われる場合も多い) が, 本研究では以下, 回帰モデルの評価を  $\bar{R}^2$  を用いて行う.

そこで, この「相関係数行列分析」を用いて, 様々なカオス時系列の分析を試みる. 対象としては,

#### ロジスティック写像 1

$$X_t = \alpha X_{t-1}(1 - X_{t-1}) \quad (\alpha = 3.7, X_0 = 0.1) \quad (2)$$

#### ロジスティック写像 2

$$X_t = \alpha X_{t-1}(1 - X_{t-1}) \quad (\alpha = 4.0, X_0 = 0.1) \quad (3)$$

#### エノン写像

$$\begin{aligned} X_t &= 1 + Y_{t-1} - 1.4X_{t-1}^2 \\ Y_t &= 0.3X_{t-1} \quad (X_0 = 0.1, Y_0 = 0.1) \end{aligned} \quad (4)$$

#### 3 次の非線形写像<sup>3)</sup>

$$\begin{aligned} X_t &= 1.9X_{t-1} - X_{t-1}^3 + Y_{t-1} \\ Y_t &= 0.5X_{t-1} \quad (X_0 = 0.1, Y_0 = 0.1) \end{aligned} \quad (5)$$

#### 池田写像<sup>4)</sup>

$$\begin{aligned} X_t &= q + b(X_t \cos \theta - Y_t \sin \theta) \\ Y_t &= b(X_t \cos \theta - Y_t \sin \theta) \\ \theta &= \gamma - a/(1 + X_t^2 + Y_t^2) \end{aligned} \quad (6)$$

$$(X_0 = 0.5, Y_0 = 0.7, q = 1.0, \gamma = 0.4, a = 6.0, b = 0.7)$$

#### 指数型 AR モデル<sup>5)</sup>

$$\begin{aligned} X_t &= \{0.99 - 0.36 \exp(-X_{t-1}^2)\} X_{t-1} \\ &\quad - \{0.36 - 40 \exp(-X_{t-1}^2)\} X_{t-2} \end{aligned} \quad (7)$$

を考える. 最後にあげた「指数型 AR モデル」は, 上記のケースを含む特定のパラメータのときにカオスが発生することが知られている<sup>5)</sup>.

これらの時系列の各々データ数: 100 に対して「相関係数行列分析」を適用した分析結果が表 2 である (従属変数がすべて 2 次階差となったが, これが偶然か

理論的な背景があるのかという考察は本研究では行っていない). また, 以下, 離散型のモデルの分析の際はデータ数を 100 とする.

$\bar{R}^2$  が 0.85 程度でよしとするならば, 3 次の非線形写像以外は以上の分析でさしつかえないことになる. カオス時系列は非線形の極みのようなイメージがあるが, このように原系列を階差系列に分解して, それらの「相関係数行列」を眺めると, 実はその内部にこれだけの線形性が潜んでいる. おそらくこのことはこれまでに指摘されていなかった事実であろう.

前回の研究<sup>2)</sup> でロジスティック写像 2 ( $\alpha = 4.0$ ) に階差系列の自己回帰を用いた「ダミー変数なし」の分析で  $\bar{R}^2 = 0.3426$  であつたものが, 同じく「ダミー変数なし」で  $\bar{R}^2 = 0.842282$  になったことを考えると, 様々な系列を用いることで飛躍的に精度があがる可能性が理解できよう.

この  $\bar{R}^2$  の違いや表 2 の全体的な  $\bar{R}^2$  の水準を考えれば, DAR の精度の改善は「ダミー変数の設定」よりも, まずどのような系列を説明変数に設定するか, すなわちいかに「線形性の強い部分を見つけ出すか」ということがポイントとなることが分かる.

#### 2.2 ダミー変数を用いた精度の向上

「相関係数行列分析」においては, 良好な結果が得られた. しかし, これ以上の精度を求めらるるのであれば, まだ分析を続けなければならない. そこで,  $\bar{R}^2$  が 0.95 に満たないロジスティック写像 2 ( $\alpha = 4.0$ ), エノン写像, 3 次の非線形写像に対して, ダミー変数を用いた改善を検討する.

相関係数行列表で相関が高いところを太字としたが, 相関図はどのようになっているのであるうか. ロジスティック写像 2 ( $\alpha = 4.0$ ) について, 相関係数が 0.8 を超える 2 つの相関図が図 3, 図 4 である. 図 3 においては, ある程度線形性が存在するのは確かだが, 説明変数の横軸の領域で分割できるような形状ではなく, 「ダミー変数」で処理できる部分がない. 図 4 の場合, 横軸の両側で線形性が崩れているが, 左側の  $-0.7$  以下はダミー変数として処理できそうである. 前回の研究<sup>2)</sup> では, このような相関図の場合でもまず回帰分析

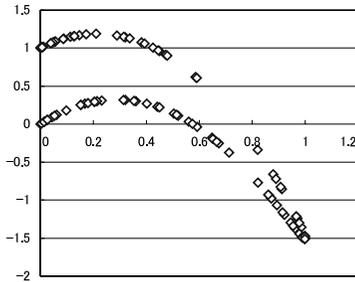


図 3 原系列から 2 次階差への回帰 ( $\alpha = 4.0$ )

Fig. 3 Regression from the original series to the second-order difference one ( $\alpha = 4.0$ ).

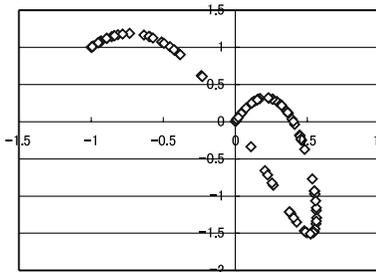


図 4 1 次階差から 2 次階差への回帰 ( $\alpha = 4.0$ )

Fig. 4 Regression from the first-order difference series to the second-order difference one ( $\alpha = 4.0$ ).

を行って、その残差を分析してダミー変数を設定したが、このように相関図そのものを見た段階でダミー変数を設定することも可能である。そこで、本研究では相関図で線形性の崩れた部分にダミー変数を設定し、その後、残差分析を行い、必要があればさらにダミー変数を追加することを考える。

そこでロジスティック写像 ( $\alpha = 4.0$ ) に関して、2 次階差系列にラグ 1 の原系列、1 次階差系列、2 次階差系列を回帰させ、図 4 をもとに 1 次階差系列の値が  $-0.7$  以下で「ある：1」、「ない：0」というダミー変数を用いた結果、 $\bar{R}^2 = 0.869903$  であり、2 次階差系列の観測値と予測値は図 5 であり、その残差が図 6 である。この図 5、6 を見ると、おおむね全体像は合致しているが、2 次階差系列の原点付近での振舞いとらえきれていないのが如実に分かる。

そこで、新たにラグ 1 の 2 次階差系列の絶対値が  $0.1$  以下で「ある：1」、「ない：0」というダミー変数を加えると、ここでは図は示さないが  $\bar{R}^2 = 0.900317$  と良好な値を示した。

次にエノン写像であるが、「相関係数行列分析」のみでも  $\bar{R}^2$  が  $0.9$  を超えている。分析結果では 2 次階差の AR を含む形となっているが、エノン写像の 2 次階

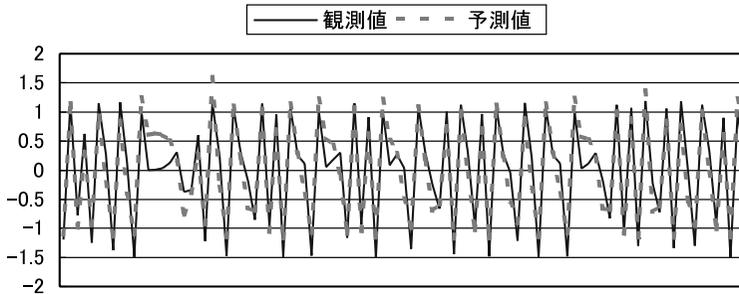


図 5 ロジスティック写像 ( $\alpha = 4.0$ ) の 2 次階差系列の予測

Fig. 5 Forecasting of the second-order series of logistic map ( $\alpha = 4.0$ ).

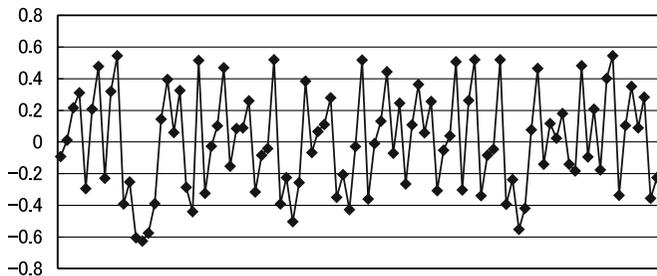


図 6 ロジスティック写像 ( $\alpha = 4.0$ ) の 2 次階差系列の予測誤差

Fig. 6 Forecasting error of the second-order series of logistic map ( $\alpha = 4.0$ ).

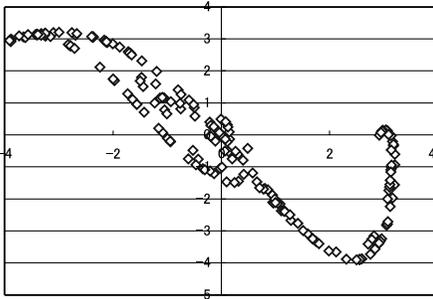


図 7 エノン写像の 2 次階差の相関図

Fig. 7 Correlation diagram of the second-order difference of henon map.

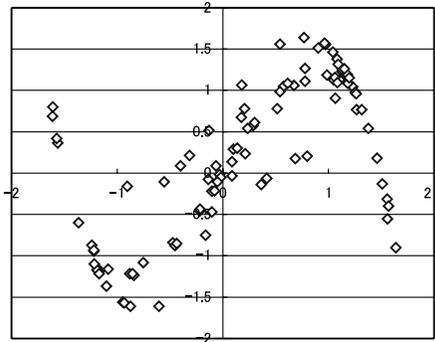


図 9 3 次の非線形写像のアトラクタ

Fig. 9 Attractor of the third-order nonlinear map.

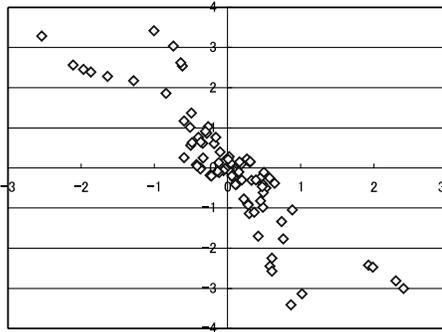


図 8 3 次の非線形写像 (1 次階差から 2 次階差への回帰)

Fig. 8 The third-order nonlinear map (Regression from the first-order difference to the second-order difference).

差系列の相関図は図 7 のような形状をしている。これもロジスティック写像と同様、この相関図の段階で横軸 ( $X_{t-1}$ ) が  $-2.5 \sim +2.5$  の範囲ではほぼ線形であり、それ以外をダミー変数で処理することが考えられる。

そこで、「相関係数行列分析」で設定された説明変数以外に、ラグ 1 の 2 次階差系列が、 $-2.5$  以下で「ある: 1」、「ない: 0」、同じく  $+2.5$  以上で「ある: 1」、「ない: 0」という 2 つのダミー変数を用いると、 $\bar{R}^2 = 0.932053$  との改善が見られた。

さらに、3 次の非線形写像であるが、表 1 の従属変数と各々の説明変数の相関図は 2 つとも図 8 のようになり、ロジスティック写像 2 ( $\alpha = 4.0$ ) の場合と形状こそ違えど、やはり同様に相関が高い系列だが、ダミー化できる部分が見出せない。そうすると、この系列はそもそも前回の研究<sup>2)</sup> で扱ったときと同様、図 9 のように原系列に線形性がある系列であるので (相関係数 0.66)、通常の前系列に対する自己回帰モデルにダミー変数を加えた最も基本的な DAR モデルをあてはめた方が精度が良い可能性がある。

図 9 を直観的にとらえれば、 $-1 \sim +1$  まだが線形

部分であり、それ以上とそれ以下をダミー変数で処理することが考えられる。SETAR ならば、3 つに区別化して AR(1) をあてはめる形状であるが、両側部分も各々ダミー変数 1 つずつで扱うにはその変位が大きい。よって、さらに細かく  $\pm 1.5$  でも分割し (この設定は前回の研究でも用いた) 計 5 つの区分に分けてダミー変数を 4 つ用いて推定すると、 $\bar{R}^2 = 0.921294$  となる。また、3 つの写像ともこれ以上の残差分析においては特徴的な部分は見られなかった。

このように相関係数行列表からの分析結果よりも精度を向上させようと考えた場合は、相関の強い系列というよりは、相関図を描き「ダミー変数変化」の容易な系列を考え、その後、前回の研究のような残差分析を行うのが有効であるが、このような判断の際にはまだ幾分かの恣意的な要素が残ることとなる。ロジスティック写像 ( $\alpha = 4.0$ ) の場合では、別の観点からダミー変数を組み込み  $\bar{R}^2 = 0.93$  程度まで精度を改善することも可能であったが、ダミー変数の設定に恣意性が強く、あまり他の写像への汎用性もない設定であったので今回は割愛することとした。

### 2.3 DAR 改善のまとめ

DAR において、ロジスティック写像 2 ( $\alpha = 4.0$ ) に適用した際の欠陥が指摘されていた<sup>2)</sup> が、DAR における精度は本研究の「相関係数行列分析」に基づけば、ダミー変数の設定よりも、むしろどのような系列を従属変数・説明変数に設定するかという問題の方が大きい。前述のような当該階差系列どうしの自己回帰で精度が思わしくないケースでも適切な系列を選べば改善されることが分かった。

さらに「相関係数行列分析」よりも精度の向上を望む場合、ダミー変数を設定するため、相関の比較的高い系列どうしの相関図を図示することが重要である。その際、相関が最も高い系列どうしの相関図にダミー

変数が設定しにくくとも、それ以外の相関図でダミー化できる場合がある。また、ダミー変数のさらなる設定には、前回の分析で行ったような従属変数(観測値)に対する残差分析に加え、本研究のように拡張した場合、説明変数に用いた系列に対する残差分析でも精度が向上する場合がある。この残差分析によるダミー変数の追加はいまだに恣意的な部分が排除しきれないので、さらなる体系化が今後の課題であるが、本研究では代表的な6つの離散型カオスのモデルにおいて「相関係数行列」ならびに「ダミー変数」を用いることによって最終的にはすべてのケースで  $\bar{R}^2$  は0.9以上とすることが可能であった。

また前述の「相関係数行列分析」のように当該系列以外の系列を回帰させた場合に、もはやそれは自己回帰とは呼ばれないものである。しかし、たとえば2次階差系列にラグ1の原系列を回帰させた場合などを考えると、それ自体は自己回帰ではなく、 $(X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) = \beta X_{t-1} + \varepsilon_t$  のような式になる。これを原系列の予測に用いる際には、左辺を  $X_t = (\beta + 2)X_{t-1} - X_{t-2} + \varepsilon_t$  などとして移項して用いるので、結局、自己回帰型となり階差系列などを多用したとしてもこれは最終的に「ARモデル」の1種となる。

### 3. 連続系のカオスとARモデルについて

#### 3.1 テーラー展開を用いた予測法

以前に指摘されていたDARの欠陥は前述のように改善されたが、そのほかにもローレンツ方程式、ダフィング方程式、レスラー方程式などに代表される連続系のカオス時系列に対しては有効であるのかどうかということも課題とされていた<sup>2)</sup>。

ところが、これはあまり知られていない事実であるが、実は連続系のカオス・モデルをルンゲ・クッタ法で近似した離散データに対しては、DARやSETARよりもさらに単純なモデルで1期先の予測が可能である。

具体的には、テーラー展開を用いる。連続関数はテーラー展開で近似できることが知られているが、このテーラー展開の微分を後進差分で近似するだけで、ローレンツ方程式などに4次のルンゲ・クッタ法(以下、本研究では微分方程式を扱う際、すべてルンゲクッタ法の4次で、きざみ幅： $h = 0.01$ のものを用いる)を適用した離散データの1期先予測が行える。

テーラー展開は、連続関数の一種の時系列予測と考えることもできる<sup>6)</sup>。これに着目し、離散型時系列の予測に適用するため、テイラー展開の差分化を考える。

ある時間に関する関数  $f(t)$  の、時点  $a$  におけるテイラー展開の第0次近似、第1次近似および第2次近似を示せば以下ようになる(ちなみに、第1次近似はいわゆる接線の方程式である)。

$$f(t) = f(a) \quad (8)$$

$$f(t) = f(a) + (t-a) \cdot f'(a) \quad (9)$$

$$f(t) = f(a) + (t-a) \cdot f'(a) + (t-a)^2 / 2 \cdot f''(a) \quad (10)$$

これらは連続時間におけるものであるから、離散時間に応用するために、 $a = t-1$ 、1階微分を1階差分(後進差分)、2階微分を2階差分で近似することを考えると、以下のような非常に簡潔な式が得られる。

$$f(t) = f(t-1) \quad (11)$$

$$f(t) = 2f(t-1) - f(t-2) \quad (12)$$

$$f(t) = 2.5f(t-1) - 2f(t-2) + 0.5f(t-3) \quad (13)$$

離散型に拡張した式(11)、(12)、(13)において、重要なポイントは、連続時間の場合には具体的な関数が分かっていなければ導関数が導き出せず、近似・予測ができないのに対し、微分を後進差分(離散型時系列の予測に応用する際には、前進差分、中心差分は使えない)で近似したことにより、離散型時系列がある関数  $f(t)$  に従っていると仮定し、それが具体的にどういった関数なのかということが分からなくても、時系列データさえあればその値から予測が行えるようになることである。もちろんこれは、ある時間に関する離散型時系列  $f(t)$  がカオスである場合にも、連続関数に近い十分に滑らかな場合は成り立つ。

そこで、この(11)~(13)の予測式で、先に述べた3つの連続系カオス時系列を4次のルンゲクッタ法で離散化したデータに適用してみる。

#### ローレンツ方程式

$$\begin{aligned} \dot{x} &= -\sigma x + \sigma y \\ \dot{y} &= -xz + rx - y \\ \dot{z} &= xy - bz \end{aligned} \quad (14)$$

$$(\sigma = 10, r = 28, b = 8/3)$$

$$x(0) = 20, y(0) = 30, z(0) = 10$$

#### ダフィング方程式

$$\begin{aligned} m\ddot{x} + c\dot{x} + kx^3 &= F \cos \omega t \\ (m = 1, c = 0.05, k = 1, F = 7.5, \\ \omega = 1, x(0) = 3.5) \end{aligned} \quad (15)$$

#### レスラー方程式

$$\dot{v} = -du - ez$$

表 3 テーラー展開を用いた予測  
Table 3 Forecasting using Taylor series expansion.

MSE	ローレンツ	ダフィンゲ	レスラー
0 次近似	0.187263	0.000656	0.002421
1 次近似	0.002725	4.64E-07	5.93E-07
2 次近似	7.30E-04	1.16E-07	1.49E-07
RAR(1000)	0.026554	0.000586	0.000914
FAR(100)	0.000216	2.54E-09	4.94E-09
FAR(500)	1.64E-06	5.53E-09	5.49E-09
FAR(1000)	2.71E-07	3.77E-10	5.52E-09

$$\dot{u} = hv + fu \quad (16)$$

$$\dot{z} = b + gz(v - c)$$

$$(d=1, e=1, h=1, f=0.2, b=0.2, g=1, c=0.4, v(0)=3.5, u(0)=4, z(0)=0)$$

これらの連立常微分方程式の第 1 方程式の変数 (それぞれ  $x, x, v$ ) の 5,000 個のデータを各予測式で予測した場合の MSE (平均 2 乗誤差) が表 3 の上段に示されているが、カオス時系列を扱ったとは思えないほど MSE が小さく、原系列と予測系列は、図示してもほぼ完全に重なって判別しにくいので図は割愛する。この 3 種類の連続系カオスにおいて、すべて MSE は、0 次近似 > 1 次近似 > 2 次近似、であり 2 次の近似から導出した予測式 (13) が最も精度が良かった。

以上のようなモデルで予測が行えるということは、確かに (11)~(13) の予測式は AR タイプのものであるが、時系列が滑らかである場合、1 期先予測は、パラメータ所与のモデルの利用が可能であり、最小 2 乗法すら必要としないということである。つまり、DAR や SETAR を用いる必要性がない。

また、この 0 次近似、1 次近似のモデルは経済時系列の予測の際に、ごくまれにナイーブ・モデルという名でベンチ・マーク基準のような形で使われることがあり<sup>7),8)</sup>、このような単純なモデルでも場合によっては、他の手法よりも精度の良い予測をはじき出すことから、あたらない経済予測をそういった観点から揶揄するような啓蒙書なども存在する<sup>9)</sup>。ところが、ナイーブ・モデルは単に適応的に予測を行う期待のモデルであり<sup>10)</sup>、根拠があまりないモデルとされている。しかし、本研究のように示せば、それはテーラー展開から導出されることが明らかである (0 次近似の場合、原系列にランダム・ウォークを仮定しても同じ式であり、1 次近似の場合は、1 次階差にランダム・ウォークを仮定しても導ける)。

ただし、テーラー展開から導出したということは、あくまでも滑らかな時系列の 1 期先予測に有効ということであって、モデルの非線形ダイナミクスなどを把

握するには、このような単純なモデルでは不可能である。また、振動の激しいデータ (たとえば、確率過程でいえば、自己相関が負になるようなデータ) には不向きな予測法である。つまり、結果はここに示さないが、カオスでも離散型のロジスティック写像などでは、テーラー展開を用いて精度の良い予測は行えない。

式 (11)~(13) はパラメータが所与の AR モデルでもあり、連続系カオスと AR モデルにはこのような関係がある。ちなみに、これら連続系のカオスにおいても「相関係数行列分析」はもちろん可能だが、結果として得られるモデルは式 (11)~(13) とほぼ同じであるということになる。

### 3.2 他の予測法との比較

カオス時系列予測法の中でも「局所線形近似」の代表的な手法としては、「ローレンツの類推法」<sup>11)</sup> や、「Farmer と Sidorowich の手法」<sup>12)</sup> がある。ここでは前者を簡略化したものを「RAR (ローレンツの AR)」とし、後者の手法を簡略化したものを「FAR (ファーマーの AR)」としてテーラー展開を用いた予測法の精度を比較する。

今、状態空間内におけるアトラクタの軌道を考え、この軌道上の 1 点を  $x(T)$  とする。この  $x(T)$  の近傍を状態空間内における過去の値のデータベース中より探索し近傍群  $x(k_n)$  を作成し、最近傍を  $x(k_1)$  とする。

Farmer らの手法<sup>12)</sup> においていわゆる「0 次近似」と呼ばれているものは、予測対象を  $x(T+1)$  とすると

$$x(T+1) \doteq x(k_1 + 1) \quad (17)$$

として予測する手法であり、これはローレンツの類推法と等価である。

また、Farmer らの手法において「1 次近似」と呼ばれているものは松葉らの研究で比較的平易に解説されているが<sup>13),14)</sup>、簡単に説明すれば、近傍群  $x(k_n)$  から状態空間内のその先の挙動を

$$x(k_n + 1) = A \cdot x(k_n) + B \quad (18)$$

といった具合に線形近似により推定し、 $x(T)$  と予測対象  $x(T+1)$  との関係を

$$x(T+1) \doteq A \cdot x(T) + B \quad (19)$$

のように予測する方法である。通常「Farmer と Sidorowich の手法」といった場合、この 1 次近似を指す場合が多い。

また、「ローレンツの類推法」や「Farmer と Sidorowich の手法」においてモデル・ビルディングの際に必要なパラメータは「状態空間の埋め込み次元:  $m$ 」と「遅れ時間:  $\tau$ 」であり、1 次近似の場合は近傍群の「点の数:  $n$ 」かあるいは「半径:  $r$ 」が必要

となり、予測式の推定には最小 2 乗推定が行われる。これらのパラメータの設定のうち「埋め込み次元」については、ターケンスの埋め込み定理によりカオス時系列のアトラクタ次元を  $d$  とすれば、 $m$  (整数)  $\geq 2d+1$  であれば埋め込み可能であることが知られている。また「遅れ時間:  $\tau$ 」について多くの設定基準がある<sup>11)</sup> が、特にアトラクタ次元  $d$  の推定については、GP 法 (Grassberger-Procaccia algorithm) が最も一般的である。しかし、この GP 法は非常に膨大な演算量がかかり多くのデータ量が求められる。このように「ローレンツの類推法」や「Farmer と Sidorowich の手法」は、予測の前処理で演算量ばかりか「適切な」次元の推定にかなりのデータ数が要求されるという欠点がある。「簡便法」を目的とするタイプ<sup>2)</sup> である DAR やテラー展開を用いた予測法との比較対象としては有用ではないし、また、現実の経済時系列データの予測などを念頭におけば、次元を正確に推定することも簡単ではない<sup>20),24)</sup>。そこで、ここでは Farmer らの手法を簡略化した「FAR」を考えることにする。

$\tau = 1$  と考えると、Farmer らの手法は領域の数が非常に多い SETAR (しきい値モデル) と等価であることが知られている (たとえば文献 13) の p.159)。本研究では、 $m = 2$ ,  $\tau = 1$  と設定した局所的な「AR モデル」を「FAR」とする。 $m$  が小さいとターケンスの埋め込み定理からすれば埋め込める保障はないが、「局所 AR モデル」と考えれば「簡便法」としては十分に通用するはずである。また、本研究の目的からも  $\tau = 1$  とする「AR」タイプのモデルとの比較の方が望ましい！ローレンツの類推法は同様に、 $m = 2$ ,  $\tau = 1$  としたものを「RAR」とする。

表 3 の下段に「RAR」と「FAR」による 3 タイプの連続系カオスに対する精度を示している。ここで FAR(1000) などは、最初のデータ 1,000 から状態空間を構成したものをデータベースとして用意し、それにより残りの 4,000 の予測を試みたという意味であり、また近傍点の数は  $n = 30$  と設定した。

表 3 におけるテラー展開の 1 次近似、2 次近似を用いたものはパラメータ所与のモデルであり、データ数もほとんど必要としないが、簡略化した「RAR」に比べ勝っており「FAR」と比べると厳密に言えば劣るが実用上はそれほど遜色のないレベルであると考えられる。

#### 4. DAR と SETAR との違い

本研究では AR 系モデルによるカオス時系列の予測を考察しており、特に DAR に焦点を絞っている。前

章までで、前回の研究<sup>2)</sup> における課題に対処する形で論じたが、この章と次の章では SETAR との比較を行う。

ここまでで改善が見られた DAR は SETAR と類似した統計モデルであるが、ここではその違いを認識しておきたいと思う。DAR と SETAR はたとえば以下のように表される。

DAR: 以下は、ダミー変数が 2 つの場合の例である。また、 $r_n$  の選び方は任意である。

$$X_t = \alpha_0 + \alpha_1 D_1 + \alpha_2 D_2 + \sum_{i=1}^j \beta_i X_{t-i} + \varepsilon_t$$

$$\left( D_1 = \begin{cases} 1 & \text{if } X_{t-d} \leq r_1 \\ 0 & \text{if } X_{t-d} > r_1 \end{cases} \right. \\ \left. D_2 = \begin{cases} 1 & \text{if } X_{t-d} > r_2 \\ 0 & \text{if } X_{t-d} \leq r_2 \end{cases} \right) \quad (20)$$

SETAR: 以下は、分割する領域が 2 つの場合の例である。ここでも、 $r$  は任意である。

$$X_t = \begin{cases} a_0 + \sum_{n=1}^{m_1} a_n X_{t-n} + \varepsilon_t & \text{if } X_{t-d} \leq r \\ b_0 + \sum_{n=1}^{m_2} b_n X_{t-n} + \sigma_t & \text{if } X_{t-d} > r \end{cases} \quad (21)$$

DAR と SETAR は非線形時系列を区分線形 AR モデルで推定するという観点は同じだが、その AR モデルの構造が異なる。DAR は、区分化された各領域において AR 係数は一定であるが、その代わりにダミー変数による定数項の調整を行っている。ところが、SETAR は、各領域において異なる AR 係数および定数項を推定する。

SETAR の利点は、やはりそういったモデルの表現力であるが、領域の数を増やすとデータの分割の作業や、領域の数だけ最小 2 乗推定をするなど解析に少し手間がかかり、1 領域あたりのデータ数も少なくなる。Tong による分析<sup>1)</sup> においても、そのほとんどが、領域の数は 2 つであり、多くても 3 つである。

ところが、DAR の場合、AR 係数が一定であり、一見表現力が乏しいように感じられるが、ダミー変数 (領域の数) をいくつ増やしても最小 2 乗法自体は 1 回ですむ。このように定数項の変化だけに限定しているため、分割する領域の数が多くても容易に解析できるという意味において、モデルの表現力は豊かである。本研究でもダミー変数は最大で 4 つ用いている。

また、式 (4)、(5) から考えれば、DAR は以下のべ

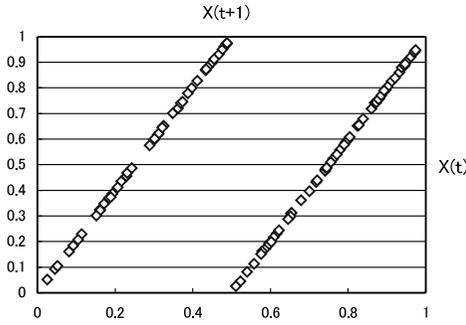


図 10 ベルヌーイシフト写像  
Fig. 10 Bernoulli-shift map.

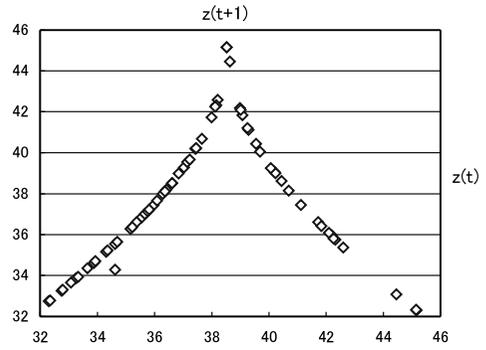


図 12 ローレンツ・プロット  
Fig. 12 Lorenz-plot.

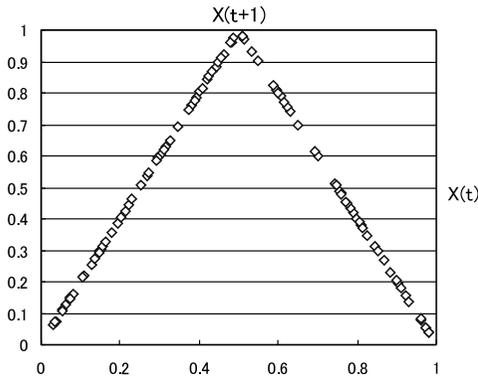


図 11 テント写像  
Fig. 11 Tent map.

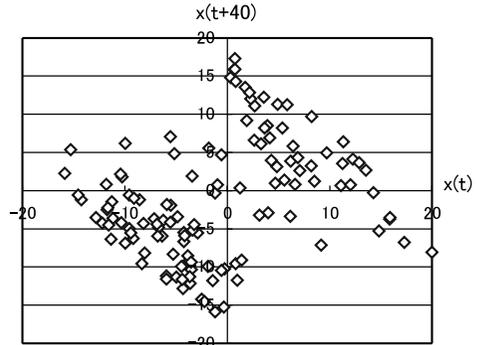


図 13 ローレンツ方程式 (40 個とび)  
Fig. 13 Lorenz equation (at forty intervals).

ルヌーイ・シフト写像 ( 図 10 )

$$\begin{aligned}
 X_{t+1} &= 2X_t \pmod{1} \\
 &= \begin{cases} 2X_t & \text{if } 0 \leq X_t < 0.5 \\ 2X_t - 1 & \text{if } 0.5 \leq X_t < 1 \end{cases} \quad (22)
 \end{aligned}$$

SETAR はテント写像 ( 図 11 )

$$X_{t+1} = \begin{cases} 2X_t & \text{if } 0 \leq X_t < 0.5 \\ 2 - 2X_t & \text{if } 0.5 \leq X_t < 1 \end{cases} \quad (23)$$

のような、写像から発生される時系列データの推定に最も適しているということが、容易に理解できる。前者は AR 係数一定で定数項が異なり、後者は、AR 係数も定数項も異なる。

また、先に連続系カオスの予測を考えた場合、テラー展開から容易に予測が可能であったが、あくまでも 1 期先の予測の場合である。そこで、連続系の長期予測の際の DAR と SETAR について考えてみる。

ローレンツは、時系列の極大値  $z_t$  ( 横軸 ) と次の極大値  $z_{t+1}$  ( 縦軸 ) とをプロットすることを考えた。これをローレンツ・プロットという<sup>15)</sup> が、たとえば

ローレンツ方程式

$$\begin{aligned}
 \dot{x} &= -\sigma x + \sigma y \\
 \dot{y} &= -xz + rx - y \\
 \dot{z} &= xy - bz \quad (24)
 \end{aligned}$$

( パラメータ, 初期値は前述と同様 )

の変数  $z$  のローレンツ・プロットは、図 12 のようになる。これは、テント写像と類似の形状であるので、SETAR でローレンツ方程式の次の極大値はいくつになるかという予測が首尾よく可能であることは明らかである ( DAR の場合は、階差をとれば可能である )。

さらに今度は、ローレンツ方程式の、ローレンツ・プロットではなく単純に 40 個とびのデータ ( 30 でも可 ) を用いて、 $X_t$  ( 横軸 ) と  $X_{t+40}$  ( 縦軸 ) をプロットすると、図 13 のように、ベルヌーイ・シフト写像を反転したようなアトラクタになる ( 初期値からの 40 個とびだけでなく、様々な値から同数ずつとばしてもほぼ同様のアトラクタである )。このようなタイプはやはり DAR が有効である。

以上のように、DAR と SETAR は類似のモデルであるが、最もストレートにあてはまるタイプの時系列には違いがある。

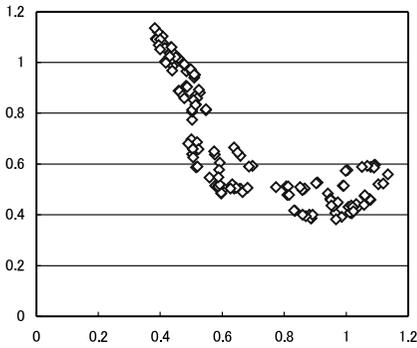


図 14 池田写像のアトラクタ  
Fig. 14 Attractor of Ikeda map.

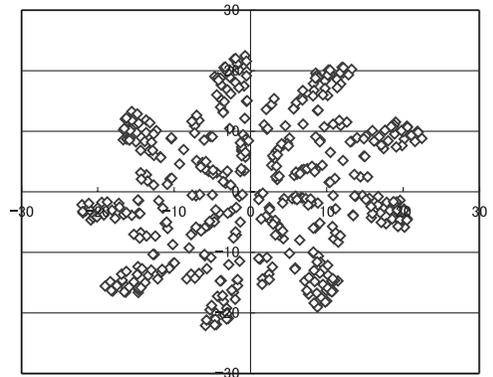


図 15 指数型 AR モデルのアトラクタ  
Fig. 15 Attractor of Exponential AR model.

## 5. DAR と SETAR の比較検討

### 5.1 DAR と SETAR の精度の比較

冒頭で DAR の欠陥は改善され、前章では DAR と同様の統計モデルである SETAR の各々の利点などを述べた。また、最もよくあてはまるケース(テント写像, ペルヌーイ・シフト写像)を紹介したが、ここでは完全には両者の枠組みに属さないタイプのデータに対しての両者の精度の比較を試みる。データとしては、冒頭での検証にも用いた 6 つのモデルから発生させたものを用いる。6 つのモデルとは、ロジスティック写像 ( $\alpha = 3.7$ ), ロジスティック写像 ( $\alpha = 4.0$ ), エノン写像, 3 次の非線形写像, 池田写像, 指数型 AR モデルである。

DAR に関しては、2 章で推定したものをを用いた。つまり、ロジスティック写像 1, 池田写像, 指数型 AR モデルは表 2 のもの、それ以外はダミー変数を加えたものである。

SETAR の推定方法であるが、6 つのモデルのうち、ロジスティック写像とエノン写像は、直観的に最も自然である、2 次関数の形状を真ん中で区分化し各々 AR(1) で推定する手法<sup>16)</sup>をとった。また、池田写像のアトラクタは図 14 であるが、これは横軸における 0.7 の値で分けて各々 AR(1) を用いた。また、3 次の非線形写像は図 9 であるが、 $X_{t-1}$  が  $\pm 1$  の時点で区切り 3 つの領域に分けて各々 AR(1) を用いた。また、推定に用いたデータ数は 100 である。

注意点であるが、例外として指数型 AR モデルでは SETAR の推定は行わないものとした。理由は次のとおりである。図 15 が指数型 AR モデルのアトラクタであるが、これはどう考えても従来の SETAR の枠組み<sup>1)</sup>では区分線形化が不可能である。また「相関係数行列分析」においても、特にダミー変数を用いなくても良好な精度であったので、階差系列などをとったと

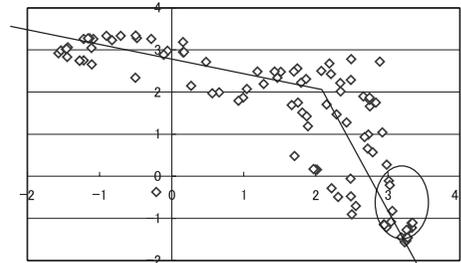


図 16 ダフィング方程式のローレンツ・プロット  
Fig. 16 Lorenz-plot of duffing equation.

しても結局、区分線形化する箇所がなく DAR と同じ推定にならざるをえないからである。以上を踏まえて推定を行った結果の MSE が表 4 の上段にある。ロジスティック写像, エノン写像に関しては SETAR の方がわずかに優れているが、池田写像や 3 次の非線形写像においては同程度であるという結論が得られた。

### 5.2 特徴的なケースにおける実際上の DAR と SETAR

前章では DAR と SETAR の理論的な関係を論じたが、ここではその実際上の適用の際における特徴的なケースを解説する。そのために今、ダフィング方程式

$$m\ddot{x} + c\dot{x} + kx^3 = F \cos \omega t$$

$$(m = 1, c = 0.04, k = 1,$$

$$F = 7.5, \omega = 1, x(0) = 3.5) \quad (25)$$

のローレンツ・プロットを考える。ルンゲクッタ法で得たデータ 5,000 個からローレンツ・プロットを抽出するとデータ数は約 100 となる。それは図 16 のようになるが、直観的には明らかに図中の直線で示した SETAR の区分線形近似が妥当と思われる。ところが、実際上は右下の丸枠で囲んだ、横軸で 3 以上のデータをダミー変数で処理した DAR の方が精度が良い。

今、DAR, SETAR とともに図 16 の写像から単純に

表 4 DAR, SETAR, RAR, FAR の精度  
Table 4 Accuracy of DAR, SETAR, RAR and FAR.

MSE	ロジスティック ( $\alpha = 3.7$ )	ロジスティック ( $\alpha = 4.0$ )	エノン 写像	池田 写像	指数型 AR モデル	3 次 非線形写像	ダフィングの ローレンツ・プロット
DAR	0.019542	0.016880	0.176553	0.019459	14.795318	0.070034	0.752919
SETAR	0.004181	0.005795	0.068617	0.019553	—	0.183303	0.785296
RAR	0.000359	0.009367	0.013293	0.003612	14.999018	0.044630	0.415687
FAR[30]	0.001363	0.044755	0.143010	0.005085	12.422472	0.255376	0.650372
FAR[20]	0.000420	0.014212	0.042150	0.004658	12.133509	0.134252	0.583675
FAR[10]	0.000003	0.003752	0.023672	0.003178	9.580874	0.094618	0.405897

AR(1) を考える．縦軸を  $Z_t$ ，横軸を  $Z_{t-1}$  として  
DAR :

$$Z_t = \alpha_0 + \alpha_1 D_1 + \alpha_2 Z_{t-1} + \varepsilon_t$$

$$\left( D_1 = \begin{cases} 1 & \text{if } Z_{t-1} \geq 3 \\ 0 & \text{if } Z_{t-1} < 3 \end{cases} \right) \quad (26)$$

SETAR :

$$Z_t = \begin{cases} a_0 + a_1 Z_{t-1} + \varepsilon_t & \text{if } Z_{t-1} > 2 \\ b_0 + b_1 Z_{t-1} + \sigma_t & \text{if } Z_{t-1} \leq 2 \end{cases} \quad (27)$$

を最小 2 乗推定した結果の MSE は，表 4 右端にあるが，若干 DAR が勝る（ただし，ダフィング方程式のパラメータや初期値が上記と異なる場合などは  $(Z_t, Z_{t-1}) = (3, 3)$  付近にデータが集中する傾向がある影響でダミー変数が有効ではなくなる）。

以上のように「分割する領域の数を同じにする」と（ダミー変数を 1 つ用いると分割する領域は 2 つ），理論的には SETAR が完全に DAR を包含するはずだが，区分化する「しきい値」を理路整然と求めることが困難であり，最適な SETAR を求めることが難しい。そのような実際上の状況においては，このダフィング方程式のローレンツ・プロットの場合のように，同じ数の領域に分割しても，各々直観的に妥当と思われる「しきい値」を選んだ場合，DAR の方が有効である場合がある。しかしながら，表 4 のその他の多くのケースを考えると DAR も SETAR もおおむね同程度の精度であると考えられる。

### 5.3 他の予測法との比較

連続系と同様に「RAR」と「FAR」との比較を行ったのが表 4 の下段である。離散型では 100 という少ないデータベースから状態空間 ( $m = 2, \tau = 1$ ) を構成し，その後の 100 を予測した。またこのような現実の予測問題に近い少ないデータ数での推定であるので，GP 法により次元を推定するのももちろん困難であるが，近傍の範囲が広いと大きな誤差を生む可能性が高いので，3 通りの近傍点数 (10, 20, 30) により

推定を行った。FAR[30]などは，局所近傍の 30 点から線形近似を推定したことを意味する。表 4 を見れば，「RAR」は「DAR」，「SETAR」と優劣をつけたいが「FAR」は特に近傍の数が 20 以下では優れている。「RAR」や「FAR」はこのようにデータ量が少なくても十分通用するが，その際，近傍郡の点の数には注意が必要である。このケースでは，データベースが小さいので，この場合の近傍点の数  $n = 30$  というのは，精度を見る限りもはや近傍とはなりえないようである。

「Farmer と Sidorowich の手法」は，このようにデータ数が少なく次元を判断できない場合でも「局所 AR モデル」であると考えることにより，ある程度適応可能であるが，データ数に比して近傍をどのくらいにとるかなどのチューニングはそれほど容易ではなく，いくつかのケースで思考錯誤するしかないと考えられる（文献 14）などでも同様である。

## 6. カオスと確率過程

これまで，連続系カオスのローレンツ・プロットや 30, 40 個とびのデータが，DAR や SETAR などの区線形 AR モデルである程度推定できることを見てきたが，今度は前述のローレンツ方程式の変数  $z$  の 50 個とびデータを考える（パラメータなどはそのまま）。図 17 が原系列 (5,000 個)，図 18 が 50 個とびの時系列 (100 個) であるが，図 17 の原系列の場合，直観的に確率過程からのデータとは考えにくい，図 18 は外見上，カオスというよりは，確率過程から発生されたデータに見える（これも前述と同様に初期値からだけでなく，様々な値から同数ずつとばしてもほぼ同様であるが，以下の分析結果は初期値から 50 個とびのものであり，初期値以外からのデータでは分析結果が少しずつ異なる）。

この 50 個とびの時系列に対して，通常のボックス・ジェンキンス法における ARIMA モデルの同定を行うことを考える。図 19，図 20 がこのデータに対する ACF (AutoCorrelation Function) と PACF

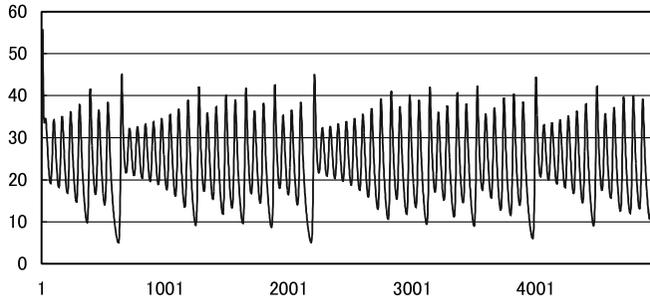


図 17 ローレンツ方程式 ( $z$ )  
Fig.17 Lorenz equation ( $z$ ).

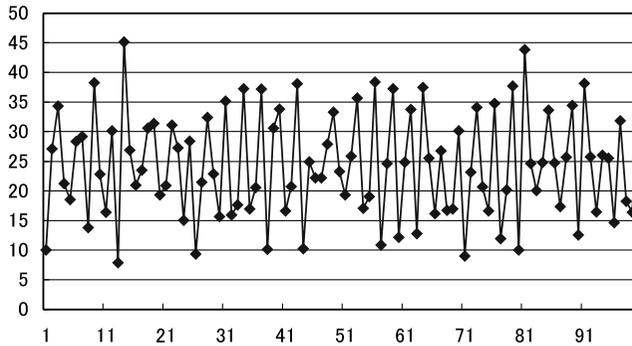


図 18 ローレンツ方程式の 50 個とび  
Fig.18 Lorenz equation (at fifty intervals).

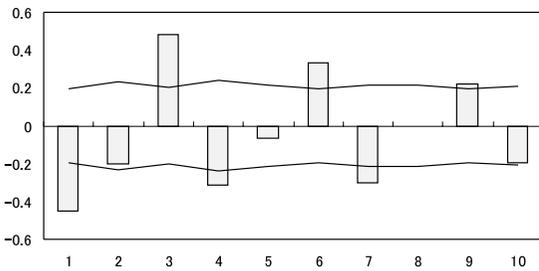


図 19 ACF (ローレンツ方程式の 50 個とび)  
Fig.19 ACF (Lorenz equation at fifty intervals).

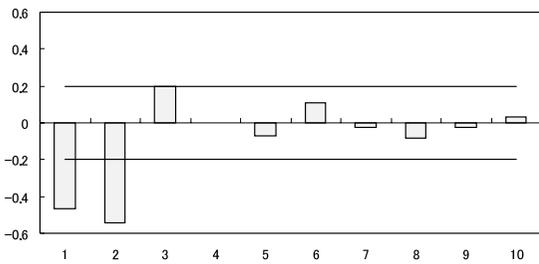


図 20 PACF (ローレンツ方程式の 50 個とび)  
Fig.20 PACF (Lorenz equation at fifty intervals).

( Partial AutoCorrelation Function ) であるが、これは AR(2) と診断される<sup>17)</sup>。この場合、時系列の発生メカニズムはカオスであるが、観測する時間間隔が長

くなると、ある種の確率過程であると同定され、時系列の外見上からもその区別が難しくなる。

このことはアナロジーでいえば合原がすでに「決定論的非線形予測法とその限界」<sup>18)</sup>として、時間が離れるにつれ、「線形予測」→「決定論的非線形予測」→「統計量の推定」といった具合に、適用できる予測手法が移り変わることを示しているが、本研究ではそれをモデルから具体的に示すこともできる。

たとえば、ローレンツ方程式の変数  $z$  に着目すれば、もちろんこの  $z$  も 1 期先予測には「テラー展開」を用いた「線形予測」が利用できる。また、 $z$  のローレンツ・プロットは、間隔の平均を求めるとおおむね 35 個程度であるが、「決定論的非線形予測」の代わりに前述のように「区分線形タイプの AR」を用いることができた。最後にモデルがカオスであっても、観測時間を 50 個とびにすると、純粋な確率過程として同定され AR(2) で予測することになる。

これは何を意味するかというと、たとえばある銘柄の株価の「秒単位」の動きが、実際は「カオス」や「カオス+ノイズ」であったとしても、「時間単位」、「日次単位」の動きは分析上、確率過程として同定されている可能性があるということである。そして、それはもはや非線形予測法での予測が難しく、本来は決定論に

従う部分までもノイズとして扱わなければならない可能性を示唆する。しかし、確率過程と同等されても、厳密に言えば図 18 などはカオスから生成されているので、そういった観点からは、株価の予測などをカオスを前提として、分析・予測することもやはり否定はできない<sup>19)</sup>。

また、この種の関係は、通常フラクタルの観点から、実際の株価の秒単位、時間単位、日次のデータなどを図示して、自己相似性を指摘するものが多いが、実際の経済時系列データではなく、上記のようにカオスが発生するモデルから眺めてみるということもできる。

## 7. まとめと今後の課題

本研究では、AR モデルとカオス時系列予測との関係を、多岐にわたって検討したため、いくぶん焦点がぼやけたかもしれない。各章はある程度、独立的な話でもあるが、最後に少しポイントをまとめた。本研究のポイントは、

- ① DAR をロジスティック写像に適用した際の欠陥を改善した(その際「相関係数行列分析」+「ダミー変数の設定」でその他の多くの離散型カオスモデルにおいても対応が可能であった)。
- ② 連続系カオスはテラー展開を利用して 1 期先予測が容易に可能(精度は、簡略化した「ローレンツの類推法」や「Farmer らの手法」にほぼ等しい)。
- ③ DAR と SETAR の特徴と精度の比較(精度は両手法ともほぼ同じだが「指数型 AR モデル」などは SETAR では分析不可能である。また、簡略化した「ローレンツの類推法」や「Farmer らの手法」の方が精度を期待できる)。
- ④ カオスと確率過程の関係、である。

① に関しては、「相関係数行列分析」は比較的すっきりした話であるのに対し、ダミー変数の設定はまだまだ恣意的な部分を含んでいるので、それらのさらなる体系化は今後の課題としたい。② や ③ では、「ローレンツの類推法」や「Farmer らの手法」を用いたが、現実の経済時系列データなどではアトラクタの次元を正確に推定するのが困難であることと(たとえば、文献 20) や 24)), 簡便な手法としては演算量が大きくなるということにより、各々簡略化した「RAR」や「FAR」を比較の対象とした。② のテラー展開を用いたものは、これらと遜色はなかったが、③ の離散型においては「DAR」よりも「FAR」の方が精度が良かった。ただし「FAR」は少ないデータの下では、正

確な次元を求められないほかに、近傍のチューニングなども難しいという点があり「DAR」なども簡便法としては十分有効であると考えられる。また、同じく離散型において「DAR」は「指数型 AR モデル」の推定などにおいて「SETAR」と比べての利点もあった。今後、③ に関してはデータ数を多くとれば、おそらく「FAR」が有効であろうから、そのほかには  $n$  期先予測での比較検証が課題となろう。

また、④ に関しては、従来は経済時系列の分析・予測などでは伝統的にボックス・ジェンキンス法による ARIMA モデルや、計量経済モデルなどが用いられていた。最近の ARCH, GARCH や VAR などそれぞれの延長線上にある。

ところが、一方では工学的な立場から、経済時系列に対するカオス理論によるアプローチもここ数十年でかなりの数の論文が出されているようである<sup>20),21)</sup>。そういった状況を念頭におくと、経済時系列はカオスなのかそれとも確率過程なのかという問題が出てくる。

また、統計的時系列解析の分野では、時間差が増大しても、なお相関が残るような問題に対して、カオスよりは ARFIMA (Fractional ARIMA) モデルが議論されている。この ARFIMA はいわゆる Long-Memory モデルとして知られ、通常の ARIMA モデルにおいて和分の次数を「整数」から「実数」に拡張したものである<sup>22)</sup>。さらにこのモデルは、経済時系列にとどまらず「トラヒック・データ解析」などでも議論されている<sup>23)</sup>が「ハースト指数」と直接関係するモデルであり、またこの「ハースト指数」は、カオス・フラクタルと密接な関係があることはすでに以前から指摘されている<sup>24)</sup>。

そういった背景を考えると、④ のローレンツ方程式の 50 個とびのデータが確率過程の構造を有するという事は、ARFIMA モデル、フラクタルなどその構造上において関連が深いのではないかと考えることもできる。そこで、これらを念頭におきながら「カオス」と「確率過程」の関係を明らかにし、また種々の経済時系列やトラヒック時系列の生成メカニズムはどのようなものかと考えるのが妥当なのかを検討するのが今後の課題である。

## 参 考 文 献

- 1) Tong, H.: *Non-linear Time Series—A Dynamical System Approach*, Oxford University Press (1990).
- 2) 長瀬隆久: AR モデルにダミー変数を加えたカオス時系列予測法, 情報処理学会論文誌, Vol.43,

- No.10, pp.3247–3250 (2002).
- 3) 潮 俊光: カオス制御, p.75, 朝倉書店 (1996).
  - 4) Ikeda, K.: Multiple-Valued Stationary State and Its Instability of The Transmitted Light by a Ring Cavity System, *Optics Communications*, Vol.30, No.2, pp.257–261 (1979).
  - 5) 合原一幸, 相澤洋二 (編著): カオス研究の最前線—非線形科学の正規へ向けて, p.101, サイエンス社 (1999).
  - 6) 大村 平: 微積分のはなし (下), pp.184–204, 221–229, 日科技連 (1972).
  - 7) Granger, C.W.J. and Newbold, P.: *Forecasting Economic time series*, p.291, Academic Press (1977).
  - 8) Makridakis, S. and Andersen, A., et al.: The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition, *Journal of Forecasting*, Vol.1, pp.111–153 (p.143) (1982).
  - 9) ウィリアム・シャーデン (著), 森 孝恵 (訳): 予測ビジネスで儲ける人々, ダイヤモンド社 (1999).
  - 10) マダラ -, G.S. (著), 和合 肇 (訳): 計量経済分析の方法, p.301, シーエーピー出版 (1996).
  - 11) 合原一幸 (編): カオス時系列解析の基礎と応用, 産業図書 (2000).
  - 12) Farmer, J.D. and Sidorowich, J.J.: Predicting Chaotic Time Series, *Physical Review Letters*, Vol.59, No.8, pp.845–848 (1987).
  - 13) 松葉育雄: 非線形時系列解析, 朝倉書店 (2000).
  - 14) 田中 賢, 王 麟元, 松葉育雄: 経済時系列データの構造変化の抽出と予測, 電子情報通信学会論文誌 (A), Vol.J81-A, No.4, pp.658–663 (1998).
  - 15) 合原一幸: カオス学入門, 放送大学教育振興会, p.45 (2001).
  - 16) 松葉育雄: ラグ回帰, しきい値モデル, カオスの臨界特性, 電子情報通信学会論文誌 (A), Vol.J81-A, No.3, pp.389–396 (1998).
  - 17) ヴァンデル, W. (著), 蓑谷千風彦, 廣松 毅 (訳): 時系列入門, p.99, 多賀出版 (1988).
  - 18) 合原一幸: 身近に存在する“決定論的カオス”, カオス学=最前線, 日経サイエンス'96 1月号 (1996).
  - 19) 上田太一郎: データマイニング実践集, pp.126–136, 共立出版 (1999).
  - 20) 寺崎 健, 池口 徹, 合原一幸, 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, 電子情報通信学会論文誌 (A), Vol.J78-A, No.12, pp.1601–1617 (1995).
  - 21) 亀川博史, 松丸正延: 外国為替レートにおける動的カオス解析の研究, 日本経営工学会論文誌, Vol.52, No.4, pp.239–255 (2001).
  - 22) 矢島美寛: Long-Memory モデルとその統計的性質, 日本統計学会誌, Vol.19, No.2, pp.219–236 (1989).
  - 23) 小沢利久, 町原文明, 石橋圭介: マルチメディアトラヒック理論の最新動向, 電子情報通信学会誌, Vol.81, No.5, pp.506–515 (1998).
  - 24) エドガー・ピーターズ (著), 新田 功 (訳): カオスと資本市場, 白桃書房 (1994).
- (平成 15 年 2 月 21 日受付)  
(平成 16 年 1 月 6 日採録)



長瀬 隆久 (正会員)

1998 年東京理科大学経営学部卒業。2000 年同大学院修士課程入学。2002 年同大学院修了。同年総務省郵政研究所, 2003 年 (財) 土地総合研究所。また, 同年中央大学大学院理工学研究科博士後期課程 (情報工学) 入学。現在も在学中。研究の中心は非線形デジタル信号処理および非線形時系列解析。