カオス時系列予測に於ける、予測誤差の性質を考慮した誤差補正法について

渡邉　達男*1

About the Error Correction Method in Consideration of the Character of the Prediction Error in Chaotic Time Series Prediction

Tatsuo WATANABE

In this paper, the new correction method of the prediction error Lorentz's method of guessing in chaotic time series prediction was proposed. By understanding the pattern of the error of a time series beforehand showed that predictive accuracy improved. Moreover, the improvement in predictive accuracy was also understood are restricted in early stages of prediction. Furthermore, the theoretic limit of improvement in predictive accuracy was considered. The proposed method may be useful also for prediction of natural phenomena, such as an earthquake.

KEYWORDS： chaos, chaos time series, forecast, new correction method

1. はじめに

自然界の様々な災害は、現在まで予測することが困難であり、また一度発生すると甚大な被害が発生することが多い。

平成２３年３月１１日に東北地方で発生した大震災は如実にそのことを物語っている。地震と同時に津波が発生し、多くの人命を奪った。また、原子力発電所の機能が停止し、炉心溶融というかつて無い危機的な状況が発生した。誰がこのような事態を予測したであろうか。多くの地震予知の重要さが指摘されているなかで、誰もが予測し得なかった地震が発生した。炉心溶融に関する被害は今でも発電所近辺での重大な放射能汚染という、かつて無い事態を招いている。

さらには、平成２３年９月に入り、台風１２号の西日本での被害は甚大であった。速度の極めて遅い台風の大雨により、崖崩れ、洪水等が多発した。そのために多くの人命がまた奪われた。

筆者はかねてから自然界の現象の予測の重要さを指摘し、またその方法を模索し続けてきた。

自然界の予測を行うことは人類の歴史とともに始まった大きな問題である。太古の時代には、占いや祈祷などで予測等を行っていた。最近では様々な方法が考えられている[1)]。

自然界の多くの現象は１変数モデルに置き換えられ、カオス時系列と近似的に考えられる。カオス時系列は簡単な規則に従うのにも関わらず、予測が困難な時系列である。カオスの性質上、長期予測が困難で、誤差が指数関数的に大きくなっていく性質がある。

しかし、様々な方法での予測が試みられてきた。その結果、近未来ではある程度の精度で予測が行われるようになってきた。

古くはLorenzによって、気象の予測に対する、Lorenz の類推法[1,2)]が考えられた。それによると、時系列を任意の次元に区切り、空間に埋め込む。

*1 電子制御工学科(Dept. of Electronic Control Engineering), E-mail: watanabe@oyama-ct.ac.jp

その埋め込まれた空間上の点として、時系列が扱われる。埋め込まれた空間の点はその点一つが時系列の変化のパターンを示しており、点を扱うことで、時系列パターンの類似性をもって予測を行おうとした。しかしLorenzの方法は単純すぎて、それだけでは予測精度があまり良いと言えない。

筆者はこれまで Lorenz の類推法に関していくつかの改良を加え、予測精度の向上を考察してきた[3-8]。特に、空間次元や近傍点数に依存して予測精度が変化することを示し、最適予測のための次元数や近傍点数に関する提案をしてきた。しかし、まだまだパラーメタの任意性の改善や、方法の改良が考えられ、単純であるが、予測精度向上のための様々な方法が残されていると思われる。

この論文では、Lorenz の類推法を改良する方法として、“予測誤差時系列を用いた補正法” を提案する。これは Lorenz の類推法の改良法で予測した時系列を真値と比較し、あらかじめ誤差の時系列ベクトルを求める。そして、これから求めようとする時系列に最も適する誤差時系列を用い、誤差の予測をする。その予測された誤差時系列と予測時系列の差分を取ることにより、予測精度の向上がはかられる方法である。

数値実験の結果、予測精度は向上した。

２章は予測機である、Lorenz の類推法の解説と、“予測誤差時系列を用いた補正法” を説明する。３章は数値実験での予測向上について述べる。４章は考察に、５章はまとめに当てられる。

２． Lorenz の類推法と誤差について

E.N.Lorenz は、1969 年、類推法 (the method of analogues)と呼ばれる手法を提案した[1,2]。

今、１次元時系列 x(0),x(1),…,x(t) を要素に持った m 次元ベクトルを考えると、v(t)={x(t-m),x(t-m+1),…,x(t)}に対して、そのベクトル空間上で、今わかっている最後の点のベクトルの最近傍点の次の時系列ベクトルをもって、Lorenz は予測点(まだ知られていない次のベクトル)とした。これは今の時系列パターンと同じパターンを過去に見出し、過去のパターンが繰り返されるという仮定をもとに、その情報から予測を行うものである。

データにノイズが無い場合には非常に有効な予測手段であるが、ノイズの量が多いと、近傍点が果たして最も良い点であるかが問題となり、予測精度が悪くなる事が考えられる。しかし、最近傍点数を増やす事で、ノイズを低減させ、予測精度が向上する。複数の近傍点からどのように予測ベクトルを一つ決定するかにより、予測精度も変化する。もっとも一般的なのは複数のベクトルの加重平均を取ることである。すなわち、近傍点ベクトル v(k_i)、近傍点数 N、近傍点と最後の点のベクトルとの距離 d_i とすると、予測点ベクトル v(T)は、

$$v(T) = \frac{\sum_{i=1}^{N} d_i\, v(k_i+1)}{\sum_{i=1}^{M} d_i} \qquad (1)$$

で与えられる。

次に、Lorenz の類推法の改良方法をさらに改良する方法として“予測誤差時系列を用いた補正法” を提案する。

“予測誤差時系列を用いた補正法” は、あらかじめ１次元時系列 x(0),x(1),…,x(T-1)に対して、任意の時間、例えば s (1 ≤ s ≤T-1)から p ステップ予測を行う。その結果 x’(s+1),x’(s+2),…,x’(s+p)の予測値が得られる。

得られた予測値と真値 x (s+1),x(s+2),…,x (s+p) の差分で作られる誤差ベクトル e(s)は、

$$e(s) = \{x'(s+1) - x(s+1), \cdots, x'(s+p) - x(s+p)\} \qquad (2)$$

この p 次元誤差ベクトル e(s)が張る p 次元空間にすべての誤差ベクトルを埋め込む。

次に、これから予測しようとする既知の最後の点 x(T-1)に関して p ステップ前からの予測を行い、真値との誤差ベクトル e’(t)を求める。

$$e'(t) = \{x'(t-p) - x(t-p), \cdots, x'(t) - x(t)\} \qquad (3)$$

この誤差ベクトル e’(t)に対して、誤差ベクトルが埋め込まれた p 次元空間においての最近傍ベクトル e”(r) (1 ≤r ≤ T-1)を求める。

誤差ベクトル e”(r+1)が T からの予測ベクトルの誤差ベクトルを表すと考え、これを予測されたベクトル v (T) に対して、

$$v'(T) = v(T) - e''(r+1) \qquad (4)$$

なる補正を加えて、補正された予測ベクトルと考える。

この方法は、誤差の時系列があるパターンを持っていると仮定される時、大きな効果がある。

３．数値実験結果

予測は LogisticMap に対する予測、すなわち、

$$x_{n+1} = ax_n(1-x_n) \qquad (5)$$

なる１次元時系列に対する予測を行った。LogisticMap は当初は生物の個体数の変動を表す式として提案された。なお、$a=3.8, x_0=0.5$ とした。

予測に当たっては Lorenz の類推法の改良法を用いた。すなわち、埋め込み次元は３次元、近傍点数は２点、そして近傍点の１ステップ先の点の内分点を予測点とした。また予測は 11000 ステップ目から 20 ステップ予測を行った。

図１は従来の方法[3-6]で予測した結果を示してある。

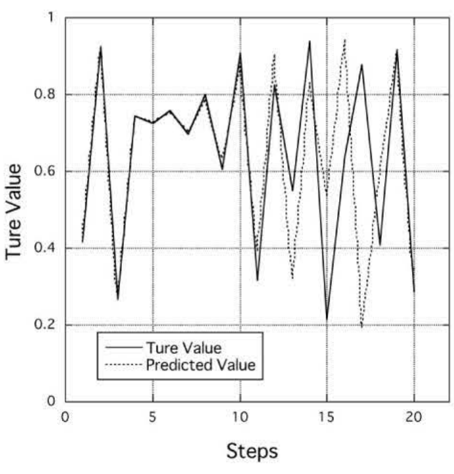

図１　従来法の予測結果

真値と予測値は予測後 10 ステップあたりで誤差が大きくなり、当たらないのがわかる。

そこで、誤差がどのように変化しているのかを調べてみた。

図２は図１の絶対誤差（実線）を示したのもである。

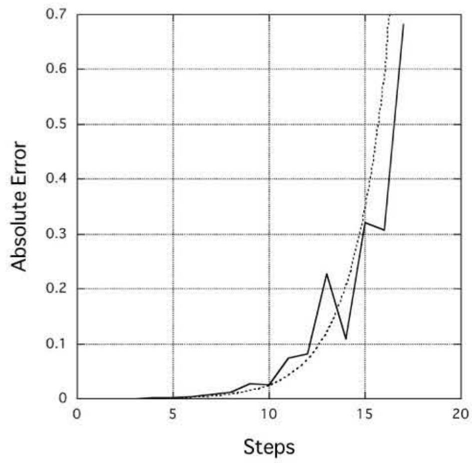

図２　誤差と指数関数の関係

図で示されるように、10 ステップあたりから急激に誤差が大きくなる。図で点線は指数関数をフィットさせたグラフである。図で見ると誤差は指数関数的に増加していることがわかる。この結果からも時系列がカオティックであることがわかる。

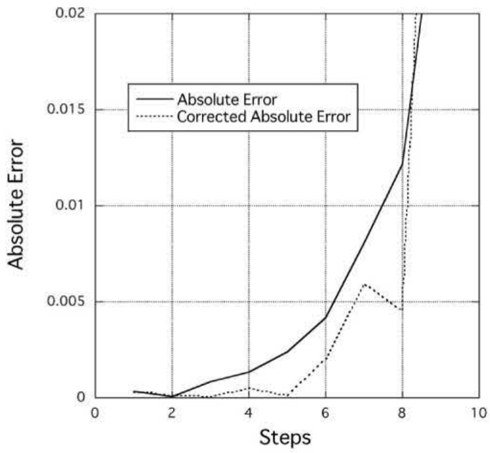

図３　補正法を用いた予測の誤差

次に、実際に“予測誤差時系列を用いた補正法”を用いた、予測した時の予測誤差を図３に示す。この予測では元となるデータの 10000 ステップから 11000 ステップまでのデータに関して誤差ベクトルを計算し、予測を行った。

図を見ると、8 ステップあたりまでは、補正法を用いた予測が誤差が少ないことがわかる。しか

し、9ステップからはこの場合は誤差が逆転していることがわかる。さらに先のステップまで見てみたものが図4である。

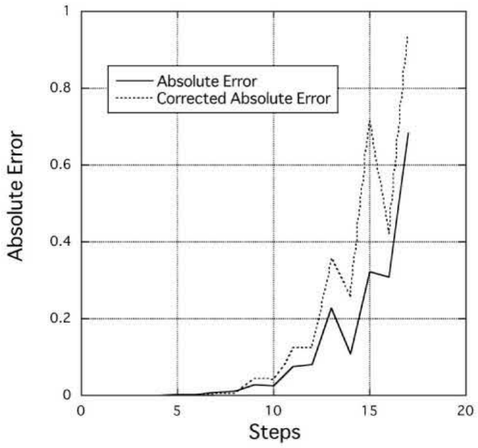

図4 補正法を用いた予測の誤差2

10 ステップあたりから誤算は補正を加える前に比べて増加している。従って、この補正法は比較的ステップ数が少ないときには誤差を減少させることがわかる。

4．考察

数値実験から、提案した補正法は予測し始めの頃は誤差を少なくすることがわかった。しかしこれはどうしてなのか。

図5に誤差ベクトル e(s)の3次元空間上の分布を示す。

図を見ると、カオスなので明確なアトラクターが無いのがわかる。そして、分布は原点 (0, 0, 0) に集中していることがわかる。

予測の最初のステップでは、誤差が少ないので、誤差ベクトルが密集した部分に誤差ベクトルが存在し、近傍点ベクトルとの距離が短いため誤差のパターンを比較的正確に表していると考えられる。

しかし予測ステップが経過し、誤差が大きくなると、まばらな空間に移動し、近傍点の距離も大きくなり、その結果、誤差パターンを良く表さない近傍点を選択せざるを得なくなるからではないかと考えられる。

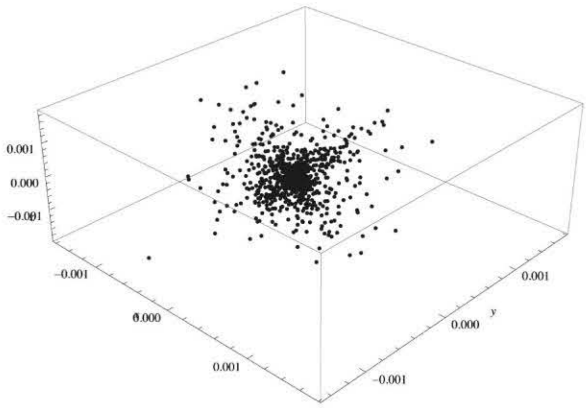

図5 e(s)の分布

このことは誤差ベクトルの数を増やしても原理的には変わらないので、この補正法の原理的な限界であると思われる。しかし、誤差ベクトルを増やせば(データ数が増えるということでもある)、誤差は少なくはなること考えられる。

カオス時系列では無い場合はどうであろうか。図5に相当する図は何らかのアトラクターを持つことが可能であると思われる。従って、この方法は予測精度を上げることが期待される。なぜなら、アトラクターがあれば、その付近の空間では、密に誤差ベクトルが存在し、近傍点との距離が近いからである。

5．まとめ

Lorenz の類推法の近傍点から予測点を求める方法には様々な方法があるが、本論文では“予測誤差時系列を用いた補正法”を提案した。その結果、予測し始めは誤差を少なくすることがわかった。従ってこの方法を用いて、例えば、数ステップ予測してから、真値を用いて修正し、さらに数ステップ予測してというふうに、真値は数ステップ毎に利用できる場合、高精度の予測ができる可能性がある。

この方法が実際にどの程度自然界の予測に使用できるかを検証する必要がある。

カオス時系列予測には様々なバリエーションが考えられるが、まだまだ改良の余地があり、精度が改善されると思われる。

自然界の予測がさらに高精度に行われるようになり、地震等の災害から少しでも人々が守られることを祈りたい。

参考文献

1) 合原一幸編:カオス時系列の基礎と応用, 産業図書(2000).
2)E.N.Lorenz:Jornal of the atomospheric science, Vol.20,p.130 (1963).
3)渡辺他:小山高専紀要,Vol38,pp.101-106(2006).
4)渡辺:小山高専紀要,Vol39,pp.107-112(2007).
5)渡辺:小山高専紀要,Vol40,pp.91-94(2008).
6)渡辺:小山高専紀要,Vol41,pp.117-122(2009).
7)渡辺:小山高専紀要,Vol42,pp.103-108(2010).
8) 渡辺:小山高専紀要,Vol43,pp.105-110(2011).

【受理年月日　2011年 9月30日】