小山工業高等専門学校研究紀要
第 36 号 (2004)， 137 － 142

カオス時系列局所線形近似予測に対する近傍点依存性

The Near Point Dependability of the Prediction of Local Linear Approximation to Chaotic Time Series.

渡 辺 達 男

Tatsuo Watanabe

１．緒言

カオス現象は現在まで様々な研究が行われている。数値計算を用いた実験、数学的構造の研究、自然の中に含まれるカオス現象の研究等様々な分野に渡っている。これはカオスが自然の中の基礎的な性質の一側面を表しているからであろう。

この論文では以前から興味が持たれている決定論的カオスの線形予測に於ける幾つかの性質を議論する。

決定論的カオスではその力学系を把握することにより、本来予測が難しいカオスに対して近未来予測が可能であることが指摘されている[1)]。カオスの本来の性質の初期値敏感性から、長時間に渡る予測は原理的に不可能であるが、近未来の予測はある程度実用になる。

カオスの時系列予測問題というのは、１次元のカオス列の予測問題であり、様々な方法が考え出されているが、大きく分けて局所線形近似予測と非線形予測に分けられる[1)]。局所線形近似にはLorenzの類推法[2)]、ヤコビ行列推定法[3)] 等があり、それぞれに特徴がある。また非線形予測には動径基底関数ネットワーク法（RBF）[4)] 等がある。これはニューロネットワークとしても知られている。

この論文ではいろいろな予測法の中で比較的簡単なLorenz類推法を取り上げ、その性質を若干検証する。Lorenz類推法は簡単であり、計算が容易であるが、データベースの中に雑音があると予測精度が落ちるという欠点がある。しかし、十分な量のデータベースを与えると予測が良くなる可能性がある。さらにはLorenz類推法では近傍点を最近傍点１つを用い予測するが、ここでは近傍点の数を増やすことにより予測精度が良くなるのではないかと考える。そして近傍点数と予測精度の関係を考察することにする。

Logistic写像を決定論的カオスの例に取り、Lorenz類推法により予測精度を考察する。予測のためのデータベースの量、及び近傍点数の依存性を調べる。

さらには、現実の系として年間降水量の変化を例にとり予測を行い仮説の検証を行う。

２章は局所線形予測の解説に、３章はLogistic写像及び年間降水量の予測実験に、４章は考察に、５章はまとめにあてられる。

２．局所線形予測の方法

１９６９年にLorenzは類推法を考案した[1)]。最初に１次元カオス時系列上の再構成された状態空間内に於けるアトラクタの軌道を考える。この軌道上のベクトルａ（Ｔ）を考える。このａ（Ｔ）が与えられた時、ａ（Ｔ＋１）を予測することを問題とする。ベクトルａ（Ｔ）の近傍を状態空間内で検索し、その近傍ベクトルａ（t_i）（ｉ＝１，２・・・Ｍ）とする。これらを距離が近い順番に並べる。その時、ａ（Ｔ＋１）はａ（t_i＋１）と予測する。すなわち、状態空間内の最近傍点ベクトルを予測ベクトルとする。

この方法では最近傍点のみを用いるが、最近傍点の近い点を数点とりあげ、それらの点の時系列の次の点から、予測点を予想することが出来る。

$$a(T+1) = f(a(t_i+1), i = 1, M) \qquad (1)$$

ただし、関数ｆは適当な重み関数である。

この関数ｆは算術平均、重み付け平均等で予想精度が変化すると予想される。

距離の計算にはユークリッド計量を使うが、その他の方法もある。

ここで再構成された状態空間とは、１次元カオ

ス時系列適当な次元数だけ区切り、そのベクトルの張る空間のことを言う。予測にはこの空間における点を取り扱うことで行われる。

3．予測結果

（1）Logistic 写像の1ステップ先予測結果

Logistic写像に対して数値計算を用い、実際に予測を行った。

Logistic写像は、

$$x_{i+1} = ax_i(1-x_i) \quad (i = 0,1,\cdots) \qquad (2)$$

で与えられ、適当なaに対して、xの数列はカオスであることが広く知られている。ここでは初期値x_0＝0．5とa＝3．9として計算を行った。

埋め込み次元は3次元として、xの数列を3個づつ区切り状態空間のベクトルとし、このベクトルをもとに予測を行う。図1にこの状態空間上のベクトルの点を図示する。

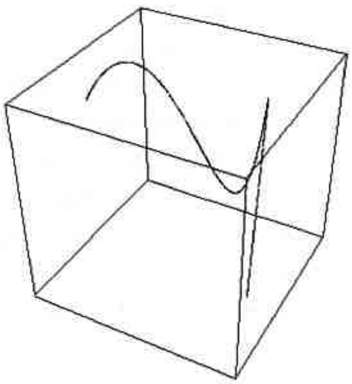

図1　状態空間上の点

予測を行うための予測値以前のxをデータベースと名付けた。

はじめに、ある点から1ステップ先の予測を行った。その時、データベース数を100から2000まで変化させた。また、近傍点数は2点から8点まで変化させ、近傍点数で予測がどう変化するかを調べた。関数fは算術平均を用いた。図2にデータベース数を100から2000まで変化させた時の近傍2点近似を用いた予測誤差の図を示す。

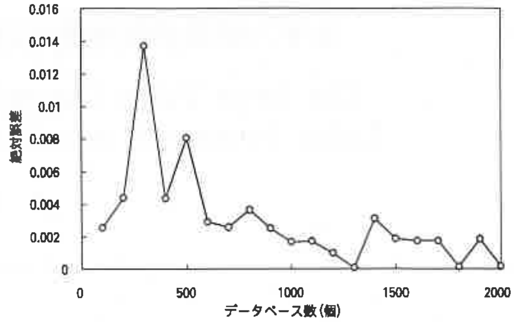

図2　近傍2点のデータベース数と誤差

この図を見るとデータベース数が少ないと予測誤差が大きいが、データベース数が多くなると予測誤差が少なくなる。ただし、データベース数が100程度の場合はたまたま近傍点が近くにあったのか、誤差が少なくなっている。

近傍点を2点から8点まで増やして同じ図にまとめたものを図3に示す。

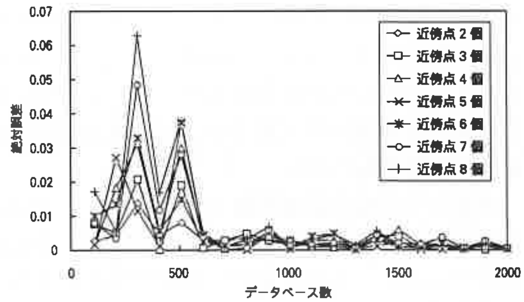

図3　データベース数と誤差

この図から、データベース数を増やすと近傍点数に関わらず誤差が減少することがわかる。

次に、近傍点数と予測誤差の関係を見るために、データベース数をパラメーターに取り、近傍点数

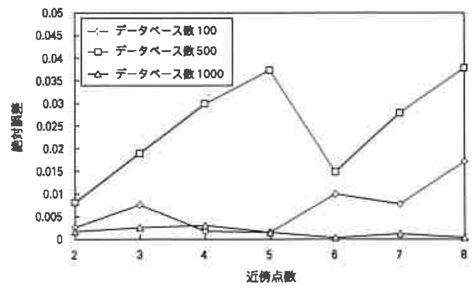

図4　近傍点数と誤差1（1ステップ予測）

と誤差の関係を図４に示す。図４はデータベース数が少ない時の図である。

図を見ると、どのグラフにも誤差最小点があり、その時の近傍点数は必ずしも８個ではなく、データベース数１００点では近傍点５個、データベース数５００点では近傍点６個、１０００点では６個になっている。これを見ると近傍点数が増えることにより誤差が少なくなるのではなく、データベース数により、ある近傍点数で誤差が最小値を取ることがわかる。

同じ図でデータベース数を多くしたものを図５に示す。

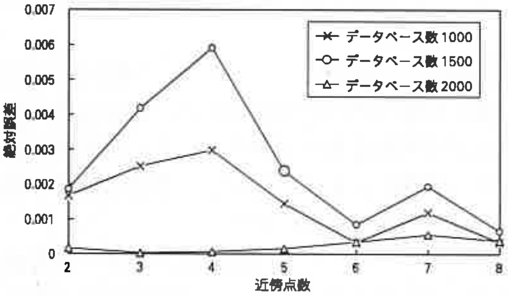

図５　近傍点数と誤差２（１ステップ予測）

図からやはり誤差が最小に近傍点数が存在することがわかる。データベース数１５００点では近傍点数６個、２０００点では３個になる。さらにデータベースを増やした図を図６に示す。

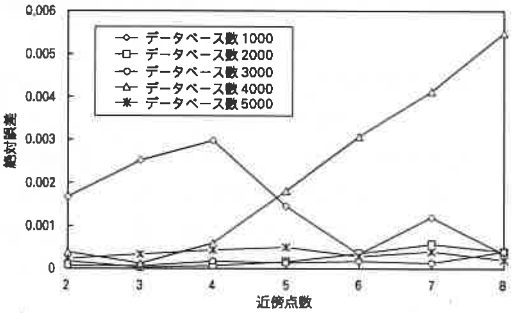

図６　近傍点数と誤差３（１ステップ予測）

図を見ると、データベース４０００点で変わった変化を示すが、データベース数が多いと、予測誤差が近傍点により極端に変化することが少なくなることがわかる。

（２）Logistic写像の２ステップ以上の予測

（１）では１ステップ先の予測の誤差評価を行った。ここでは１０ステップ先まで予測を行い、誤差がどのようになるのか、近傍点数との関係を調べた。

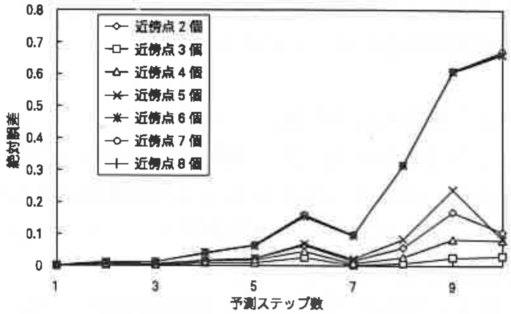

図７　予測ステップ数と誤差

図７は予測ステップ数と誤差の関係を近傍点数をパラメーターにして描いたものである。データベース数は２０００点とした。

この図を見ると、より先を予測すると誤差が増加することがわかる。また近傍点３個では１０ステップまで誤差が少ないのに比べて、６個以上の近傍点を用いた予測では４ステップ目から誤差が増加しはじめる。このことから、近傍点数はいたずらに多くすると誤差が増えてしまうことがわかる。この実験では３個が誤差が少なかった。また３ステップまでは非常に誤差が少ないことがわかる。

次の上記の関係をより詳しく見るために、近傍点数と誤差を予測ステップ数をパラメーターにして描いた図を図８に示す。

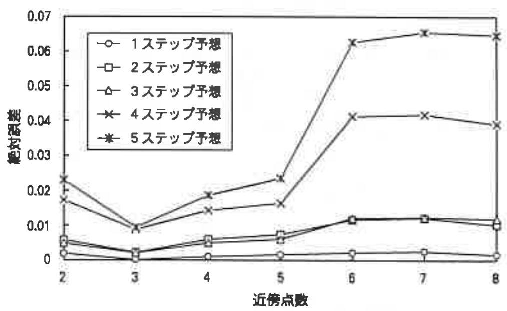

図８　近傍点数と誤差

図を見ると、近傍点３個の時どのステップの予測も誤差が少なくなる。１ステップ先の予測では、近傍点数に依存がほとんど見られないが、４、５

ステップ先では明らかに近傍点数で予測誤差が多くなる。近傍点３個は安定して誤差が少ないが、近傍点６個以上では、少し誤差が多くなる。

このことからも、近傍点数が多いことが必ずしも予測を良くすることではなく、誤差が少ない近傍点数が存在することがわかる。

（３）自然現象の予測

この方法を利用して、実際の自然現象を予測してみた。例として栃木県小山市の年間降水量の変化を予測した。データは気象庁のホームページから転記させていただいた[5]。

図９に近傍点２個の真値と予測値の図を、図１０に近傍点３個の真値と予測値の図を示す。

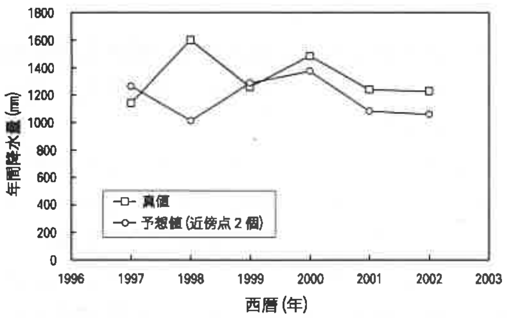

図９　年間降水量の予測（近傍点２個）

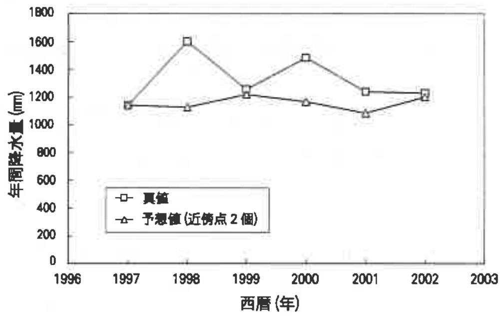

図１０　年間降水量の予測（近傍点３個）

図９、１０でははは１９７９年からのデータを用いて近傍点２個、近傍点３個を用いた予測を行った。

図９から近傍点２個が良い予測を行っている。１９９８年こそ誤差が大きいが、それ以外の年は非常に良く予測されている。図３の近傍点３個の場合は予測のずれが少ないが、データベース数が少ないためか年による変化をそれほど予測していないように思える。全ての予測値がだいたい平均値を示しているのではないかと思われる。ただ、２００２年は良く一致している。これにより、データベース数の少ないこのような自然現象は近傍点２個（もしくは出来るだけ少ない近傍点数）が予測に適していると思われる。

この現象はカオスかどうかわからない。データベース数が少ないためどのようなアトラクターがあるのかわからない。

４．考察

３章ではカオス時系列をLogistic写像を例に取り、実際の予測を行った。また自然現象の予測の例として、栃木県小山市の年間降水量の変化の予測を行った。

Logistic写像の場合、データベース数が多い場合、再構成空間内の点が稠密なので、近傍点数を多く取れば予測精度が上がるのではないかと予想されるが、実際にはそうではなく、近傍点数が３個の時、最も誤差が少なかった。

これは、データベース数が稠密とはいえ再構成空間の点間の距離があり、近傍点数を多く取るとその分より遠くの点の情報を取り入れることになり、より遠くの点の情報は近似の精度が悪く、その結果誤差を増加させる原因になっているのではないかと思われる。

上記の結果から最も誤差が少なくなるデータベース数、近傍点数、予測ステップ数の関数があると思われる。近傍点は今回の実験では３個と６個で誤差少なくなった。何か関係がありそうである。

年間降水量の変化の予測では、データベース数が少ないので、近傍点２個がよい予測を示した。

３個の近傍点を取ると、非常に遠い点の情報を取り入れてしまうので予想が良くないと思われたが、あまり大きな変化をする量ではないからか、誤差はそれほど多くはならなかったのではないかと思われる。

今回の報告では、各パラメータと誤差の最小に関する関数の特定をまだ行っていないが、今後関数の特定を行いたい。

５．まとめ

Logistic写像に関して、Lorenzの類推法を用い近傍点数と予測誤差の関係を調べた。

その結果、データベース数と近傍点数、予測ス

テップ数と誤差の間には密接な関係があることがわかった。特に近傍点数はいたずらに増やしても誤差は少なくならず、実験では近傍点数３個、６個が最も予測に適していることがわかった。

年間降水量の予測では良い予測が得られることがわかった。

今後はデータベース数、近傍点数、予測ステップ数、誤差間の関連性をより詳しく調べる必要があると思われる。

また、他の予測法の研究も必要と思われる。

参考文献

1) 合原一幸編：カオス時系列解析の基礎と応用、産業図書（２０００）.
2) E.N.Lorenz:J.Atmos.Sci.26(1969)636.
3) M.Sano,Y.Sawada:Phys.Lev.Lett.55(1985)1082.
4) M.Casdagli:Physica,35(1989)335.
5) http://www.jma.go.jp/JMA_HP/jma/index.html

「受理年月日　2003年９月25日」