

G-014

複合ニューラルネットワークによる効率的な カオス時系列予測

Efficient Hybrid Neural Network for Chaotic Time Series Prediction

小笠原 亨[†]
Toru Ogasahara

井上 浩孝[‡]
Hirotaka Inoue

1. まえがき

現代における「予測」はコンピュータを利用したシミュレーションが多くなっている。従来の統計的手法は線形近似を利用したものが多く、非線形システムに対しては予測能力に限界があるとされてきた。そこで非線形システムを考慮した予測が可能である。ニューラルネットワークが非線形時系列予測に応用されている [1, 2, 3]。もし精度の高い時系列予測ができれば、株価や為替のレートの予測等多くの分野に役立てることができる。

しかし従来のニューラルネットワークでは精度の高い予測値を得ることができない場合もあった。階層型のようにシンプルなネットワークモデルでは、データの急激な変化に対応しきれないという問題がある。また最近傍決定則を用いたローカルモデルは短期的なカオス時系列予測に対して有効であるが、カオスデータの初期状態に敏感に反応するという特性を持つため、長期予測は不可能とされている。

その精度を改善する為、本研究では、一般的なバックプロパゲーション学習を行う階層型ニューラルネットワークと最近傍決定則を使用する単純なローカルモデルを組み合わせた複合ニューラルネットワークを提案する。この予測器はそれぞれ単独の予測器で出力された予測値を各予測器のもつ出力特性で補正することにより実現されるものである。この複合ニューラルネットワークの性能を評価するにあたり本研究では、Mackey-Glass データと、Santa Fe, K.U.Leuven で開催された時系列予測コンテストで使用されたカオス時系列データに適応させて、それぞれ単独の予測器で予測を行うよりも長期予測が可能であることを示す[§]。

2. 複合ニューラルネットワーク

複合ニューラルネットワークとは、階層型ニューラルネットワークとローカルモデル (Nearest Neighbor 法) それぞれの出力を統合して、新たな予測値としてネットワークの出力とするものである。そして複合により出力された予測値はフィードバックされ、次の時刻における各予測器の入力値として使用される。ある時刻 t の複合ニューラルネットワークの期待出力を x_t とするとネットワークの出力 y_t は次式のように計算される。

$$y_t = \omega_{c_1} y_{c_1 t} + \omega_{c_2} y_{c_2 t} \quad (1)$$

式 (1) における ω_{c_1} , y_{c_1} は階層型ニューラルネットワークの寄与率、及び出力 ω_{c_2} , y_{c_2} はローカルモデルの寄与率、及び出力である。

以下に複合ニューラルネットワークを構成する2つのネットワークについて説明を加える。

2.1 階層型ニューラルネットワーク

図1は階層型ニューラルネットワークの構造である。図1からわかるように、階層型ニューラルネットワークには多数のユニット間結合荷重がある。最初はこれらの結合荷重の適切な値は不明なため、ランダムな値に設定する。この初期状態から各結合荷重を少しずつ修正してゆき、最終的に適切な値に

[†] 呉工業高等専門学校専攻科機械電気工学専攻

[‡] 呉工業高等専門学校電気情報工学科

[§] <http://www.ece.pdx.edu/~mcnames/DataSets/> より入手可能

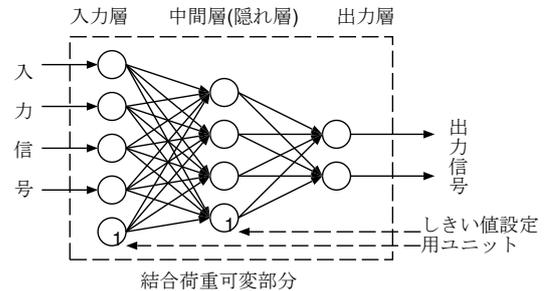


図1 階層型ニューラルネットワークの構造

設定することを階層型ニューラルネットワークの学習という。このような学習法として現在最もよく用いられているのが誤差逆伝播法 (バックプロパゲーション法) である。誤差逆伝播法は、各結合荷重を出力誤差が小さくなるように少しずつ修正するという方法である。

結合荷重 $w_{j_i}^{k-1 k}$ の一回の修正量 $\Delta w_{j_i}^{k-1 k}$ は次式で表される。

$$\Delta w_{j_i}^{k-1 k} = -\varepsilon d_i^k O_j^{k-1} \quad (2)$$

$$d_i^k = \begin{cases} (O_i^m - y_i) O_i^m (1 - O_i^m) & \text{for } k = m \\ (\sum_p w_{i_p}^{k k+1} d_p^{k+1}) O_i^k (1 - O_i^k) & \text{for } k \neq m \end{cases} \quad (3)$$

- $w_{j_i}^{k-1 k}$: $k-1$ 層 j 番ユニットから k 層の i 番ユニットへの結合荷重
- ε : 小さな係数
- m 層 : 出力層
- O_j^{k-1} : $k-1$ 層の j 番目のユニットの出力信号
- y_i : i 番目の出力ユニットの教師信号

誤差逆伝播法において、学習中の時刻 t での結合荷重修正量 $\Delta w(t)$ を、その結合荷重に対する前回の修正量 $\Delta w(t-1)$ を使って次式のようにすることがある。

$$\Delta w(t) = d + m \cdot \Delta w(t-1) \quad (4)$$

ただし、 d は時刻 t での通常の修正量であり、 m はモーメント係数と呼ばれる定数値である。前回の修正量である慣性項 $m \cdot \Delta w(t-1)$ を加えることによって、学習が加速される。ただし、モーメント係数 m を大きくし過ぎると、誤差の極小値付近で振動してしまうなどの弊害が生じるので注意が必要である。

2.2 ローカルモデル

式 (1) における $y_{c_2 t}$ はローカルモデルの出力である。このローカルモデルでの予測は次の2ステップで行うことができる。

まず、最近傍決定則により入力パターンの最近傍点を探し出す。最近傍決定則は一般にパターン認識問題などに用いられる手法で、ある特徴空間内にお

いて最も近い点を同一のクラスとして認識するというものである。本研究では任意数 k 個の近傍点を探す k -NN(k -nearest neighbor) を使用した。

次にこの k 個の近傍点を用いて予測値を求める訳だが、これには平均モデルや線形モデルなど様々な手法がある。本研究では複合ニューラルネットワークの構造をできるだけ単純にするために平均モデルを使用した。平均モデルは次式のように近傍点を単純に平均したものである。

$$y_{c_2t} = \frac{1}{k} \sum_{i=1}^k x_i \quad (5)$$

3. 実験結果

複合ニューラルネットワークの性能を評価するために、白血球の変動に関する非線形モデルである Mackey-Glass データと、Snata Fe で開催された時系列予測コンテストのデータである SantaFe データ、及び 1992 年のニューメキシコで開催された時系列予測コンテストのデータである K.U.Leuven データで予測を行った。

図 2, 図 3, 図 4 は以上に挙げた 3 つのデータそれぞれに iterated 予測の integrated 予測をプログラムに実行させ、何度も試行を重ねた上での最良の予測精度を出力したものである。

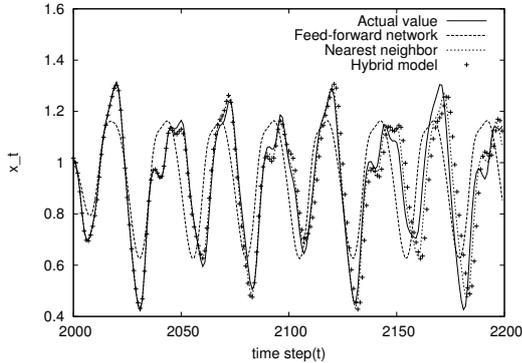


図 2 Mackey-Glass データにおける最良の予測波形

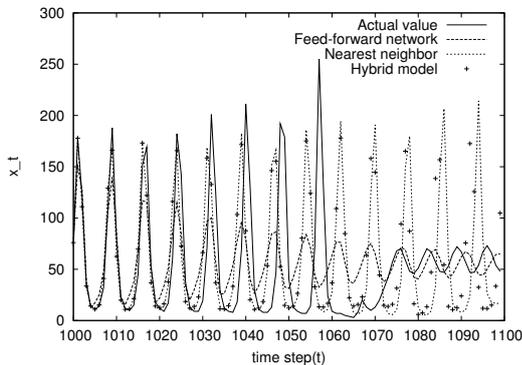


図 3 Santa-Fe データにおける最良の予測波形

評価値の求め方には種々の方法があるが、本研究では NMSE(Normalized Mean Squared Error) を利用した。評価値に NMSE を用いる理由は、NMSE が正規化されている誤差であるため、どの問題に対しても同じ程度の評価値が出力されるからである。NMSE は評価を行う区間を T , 時刻 t における期待出力を a_t , 期待出力の平均値を $E[a] = \frac{1}{T} \sum_{t=1}^T a_t$ とすると (6) 式, (7) 式で表すことができる。

$$e_{c_1} = \frac{\sum_{t=1}^T (y_{c_1t} - a_t)^2}{\sum_{t=1}^T (a_t - E[a])^2} \quad (6)$$

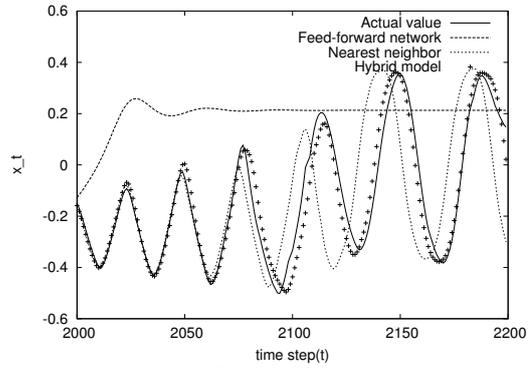


図 4 K.U.Leuven データにおける最良の予測波形

表 1 Mackey-Glass データにおける最良の評価値

モデルの種類	NMSE	処理時間 (秒)
Hybrid Neural Network	0.193852	
Feed-Forward Net	0.451472	16.72
Nearest Neighbor	0.049450	

表 2 Santa-Fe データにおける最良の評価値

モデルの種類	NMSE	処理時間 (秒)
Hybrid Neural Network	1.503876	
Feed-Forward Net	1.601632	11.60
Nearest Neighbor	1.996459	

表 3 K.U.Leuven データにおける最良の評価値

モデルの種類	NMSE	処理時間 (秒)
Hybrid Neural Network	0.063641	
Feed-Forward Net	2.823557	11.92
Nearest Neighbor	0.889379	

$$e_{c_2} = \frac{\sum_{t=1}^T (y_{c_2t} - a_t)^2}{\sum_{t=1}^T (a_t - E[a])^2} \quad (7)$$

4. むすび

得られた結果より、階層型ニューラルネットワークとローカルモデルにより構成された複合ニューラルネットワークは、数値実験の結果それぞれ単独の予測器で予測を行うよりも長期的な予測が可能であることが証明できた。しかし前節、図 2 において、Mackey-Glass データの複合ニューラルネットワークは、そのローカルモデル (Nearest Neighbor 法) の結果よりも精度が劣ってしまっている。この結果の要因として挙げられるのは Mackey-glass データは規則性をもったデータであったため、最近傍決定則を用いたローカルモデル (Nearest Neighbor 法) の方が優秀になってしまったのではないか。それにしてもその他の 2 つの結果は複合ニューラルネットワークの方が勝っている。K.U.Leuven にいたっては約 10 分の 1 の精度が出力されている。そのことから本手法の有効性を明らかにすることができた。

参考文献

- [1] 福永好伸. 複合ニューラルネットワークによる効率的なカオス時系列予測. 修士論文, Feb. 2002.
- [2] 福永好伸, 井上浩孝, 成久洋之. 複合ニューラルネットワークによる効率的なカオス時系列予測. 電子情報通信学会論文誌, Vol. j85-D-2, No. 4, pp. 689-694, April 2002.
- [3] 浜師孝二. ニューラルネットワークを用いた時系列予測に関する研究. 平成 14 年度 卒業研究論文要旨, Feb. 2003.