ハイパフォーマンス
コンピューティング 59－3
（1995．12．11）

階層型ニューラルネットワークによる時系列現象の予想

青山智夫，井須芳美[†]，長嶋雲兵[†]
宮崎大学 工学部 電気電子工学科
[†]お茶の水女子大学 理学部 情報科学科

概要

階層型ニューラルネットワークと関数の巡回表現を用いて時系列現象の予想を行った．この方法は，離散化した関数を小区間に分解し，その部分断片を組み合わせて外挿時の関数の形を予想する．断片の中に関数の将来の形と同じものが存在すれば精度良く予想できる．同じ断片がない場合，小区間の学習から元関数の近似関数がネットワーク内に構成できれば（この可能性は低くない）精度良く予想できる．

Extrapolations by using neural-networks and a recurrent representation of functions

Tomoo AOYAMA, Yoshimi ISU[†], and Umpei NAGASHIMA[†]

Miyazaki University
[†]Ochanomizu University

abstract

Extrapolations for time-dependent phenomena are studied on use of multi-layered-neural-networks and recurrent representations of functions. Adopted theories are based on principles which the function can be constructed of many vectorized-fragments, and the fragments can be related to a set, which is a representation for an extrapolated-part of the function.

序

本研究の目的は，時系列現象を階層型ニューラルネットワークを用いて予想する際の精度を定量的に調べることである．階層型ニューラルネットワークにはフィードバック付きと無しの２種類がある．どちらも時系列現象の予想に使用できる．

フィードバック機能つきの階層型ニューラルネットワークを用いた時系列現象予想は幾つか行われている．[1] 試行は株価予想などの経済的な分野に多く，結果は他の経済的指標の補助として使用されるレベルである．予想精度についての議論は少ない．時系列現象を表すデータが予想可能な情報を含んでいるか否かという観点からの研究はなされているが[2]，予想可能な情報が十分に存在するときに精度がどうなるか，情報が少しづつ減っていくと精度がどうなっていくのかという研究が不十分である．

予想とニューロン間の結合数との関連に関する研究もAICを用いた例がある[2]が，もっと多角的な検討が望まれる．

ニューロンの動作関数もシグモイド関数のみが研究の対象になっているが，シグモイド関数を採用しなければならないという根拠はない．

以上の現状認識に立ち階層型ニューラルネットワークの予想精度を再度検討する．予想精度の検討は，単純な動作の良く解っているネットワーク構造から始める必要がある．

本論文では次の構造のニューラルネットワークについて研究した．

①３層構造，

②第２層のニューロンの動作関数はシグモイド関数，

③第３層のニューロンの動作関数は線形関数．[4]

学習時第１層のニューロンに入力　するデータは(0.1,0.9)区間にスケーリングする．

④第１層の１個のニューロンはバイアス・ニューロンとする．

⑤学習時はフィードバック動作を行わない．

時系列現象に関して次の仮定を置く．

①現象は等しい時間間隔で測定できるか，データ処理によって等時間間隔の測定と同等にできる．測定値はスカラーである．

②現象は決定論的部分と非決定論部分からなり，前者が支配的である．

③測定結果を表す連続微分可能な一価関数が存在する．この関数は時間変数に関する合成関数ではない．

予想に関して次の仮定をする．

①現象の非決定論的部分を対象にしない．

1．関数周期

以下「関数」というとき一価実関数である．測定時刻間隔が等しい時系列データ{v}の部分集合 $P_n \equiv \{v_{n-i}, \ldots, v_{n-1}, v_n\}$をパターンという．関数 f(x) の周期を以下のように定義する．[5]

$$G(\theta, \eta) = \int_\Omega \{f(x+\theta) - f(x+\eta)\}^2 dx$$

なる関数$G(\theta, \eta)$を導入する．Ωは積分区間である．f(x)が周期関数の場合，適切な積分区間Ωをとると，G=0を満足する跡は等間隔直線である．η'=定数としたとき，$G(\theta, \eta')$値の変化は次の関係を満足する．

① G=0となるθ値 $\{\theta_1, \theta_2, \ldots, \theta_n\}$ は$\theta_2 - \theta_1 = \theta_4 - \theta_3 = \ldots\ldots$である．

② $\theta_i \leqq \theta \leqq \theta_{i+1}$，$\theta_{i+1} \leqq \theta' \leqq \theta_{i+2}$ である中間のθ，θ'値に対して，$G(\theta, \eta') = G(\theta', \eta')$とすることができる．

周期関数 f(x) を離散化しベクトル{v}を得，Ω区間で定義されるパターン{P}の集合によって，$f(x) \fallingdotseq \{v\} = \{P\}$ とすることができる．

［条件A］{P}の各要素が相互に異なるとき，

［条件B］P_iと任意の実数値ξを一対一に対応させる関数Ξが存在するならば，

$\xi_1 = v_{i+1}, \ldots, \xi_i = v_n$とすることにより，$\{P_1, \ldots, P_{n-1}\}$から$P_n$を計算することができる．$P_n$が$\{P_{n-1}\}$のあるパターン$P_j$に同じならば（この条件は条件Aと{周期＋Ω}分以上のデータがあれば満足される）$P_{n+1} = \{P_n;\ -v_{n-i+1}\ \ +v_j\}$である．ここで{}の中の;の右側の-/+記号は集合の要素を除く/追加することを表す．同様にして$\{P_{n+2}, P_{n+3}, \ldots\}$を計算できる．これを関数の巡回表現と呼ぶ．巡回表現には見掛け上関数の「周期」に関する情報が入っていない．

2．階層型ニューラルネットワークによ周期関数の予想

2.1 原理

［条件A］を満足する関数は，$f(x)=\sin(x)$とおけば，$0<\theta<2\pi, 0<\eta<2\pi$，$\theta \neq \eta$ において，十分な範囲のΩをとれば，

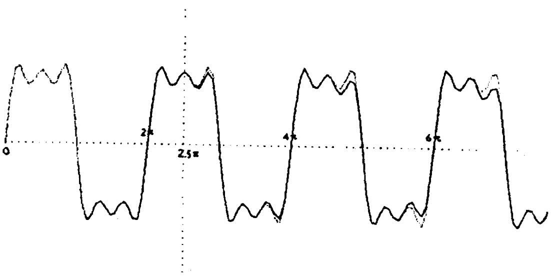
$$G(\theta, \eta)=\int_{\Omega}\{\sin(x+\theta)-\sin(x+\eta)\}^2 dx \neq 0$$

であるから少なくとも１つは存在する．ゆえに，集合$\{\sin(mx), \cos(nx); m=1,2,...,n=1,2,..\}$の各要素の線形結合で表される関数について［条件A］を満足する．

［条件B］を満足する関数Ξの機能は，任意のベクトル{v}を任意のスカラー量{ξ}に対応させることである．これを$\xi=\Xi(P)$と書く．Ξ関数は階層型ニューラルネットワークで近似的に表せる．従って関数の巡回表現とニューラルネットワークを用いて有限フーリエ変換できる周期関数の予想が近似的に可能である．巡回表現を用いたことにより，予想計算時はフィードバック動作が必要である．

2.2 計算結果

下図に$y=\sin(x)+\sin(3x)/3+\sin(5x)/5$（矩形波の近似関数）の結果を示す．学習区間(0, 2.5π)，予想区間(2.5π, 7.5π)である．学習区間は40点で離散化した．Ωは離散化点数で16(=0.4π)，第2, 3層のニューロン数はそれぞれ30, 1である．Ωの値は第１節のG=0が充分に保証されるように決定する（経験上多少大きめが良い）．



予想位置jの精度を$\Delta_j=|y_{j(calc)}-y_j|/\max\{|y_{(learning)}|\}$と定義する．$\Delta_j$がそれぞれ0.1, 0.2, 0.3を越えたj値を学習区間の離散化点数で除算した結果を**Δ**と表示する．この場合**Δ**=(1.88, >3, >3)であった．記号>nはnよりも大きい数ということを示す．

この他，種々の正弦／余弦関数の線形結合，ガウス関数の線形結合で予想精度を計算した．いずれの場合も高精度で長期間予想が可能であった．

2.3 考察と議論

第２層にシグモイド関数，第３層に線形関数を使用した階層型ニューラルネットワークでは，ネットワーク全体では，$\Sigma a_i b_i(x)$ という近似になり，基底関数$\{b_i\}$を第1, 2層で構成し，その係数a_iを第2, 3層間のニューロン間結合荷重で決定することになる．
基底関数はシグモイド関数の線形結合で表されるので，急峻に変化するような時系列現象では数多くの同関数が必要である．緩やかに変化する時系列現象では，第２層のニューロン数を多く取った場合，非決定論的要素の比率が大きい場合「過学習」の危険が生じる．そのときAICによるニューロン数決定が有効である．非決定論的要素がほとんど無い場合，バックプロパゲーションの性質により，第２層のニューロン数が多くても同じような基底関数を生成するだけでニューラルネットワークの機能に問題は生じない．もし測定データ数に較べて多い変数を計算するバックプロパゲーション方程式に不安を感じるならば，第１節のパターン集合{P}を補間によって無限要素集合に変換して学習する．そうすれば現象の時間に関する「連続性」を学習したことにもなる．実際にそのような学習法も試したが，予想能力は連続性を採用しない場合と同じであった．

3．測定データが不十分な周期関数の予想

正弦／余弦／ガウス関数の場合，「１周期＋Ω」期間以上の測定データがあれば高精度の予想が可能である．現実の時系列現象では周期が判然としないことが多く，かりに周期が予想される場合であっても１周期分のデータが得られていないことが多い．このような

場合のニューラルネットワークの予想について議論する．

不十分な学習データから予想する場合のモデルとして，正弦関数を例にとり考察する．数値実験の結果を示す．正弦関数の１周期未満を学習しても，全周期を相当の精度で予想できた．

データ

比率	学習区間	予想区間	Δ		
80%	$0, 2.0\pi$	$2.0\pi, 6.0\pi$	1.2	2.2	>3
64	$0, 1.6\pi$	$1.6\pi, 4.8\pi$	1.9	2.6	>3
48	$0, 1.2\pi$	$1.2\pi, 3.6\pi$	1.6	1.8	2.5
40	$0, 1.0\pi$	$1.0\pi, 3.0\pi$	1.4	1.6	1.7
32	$0, 0.8\pi$	$0.8\pi, 2.4\pi$	2.4	>3	>3
16	$0, 0.4\pi$	$0.4\pi, 1.2\pi$	1.5	1.6	1.8

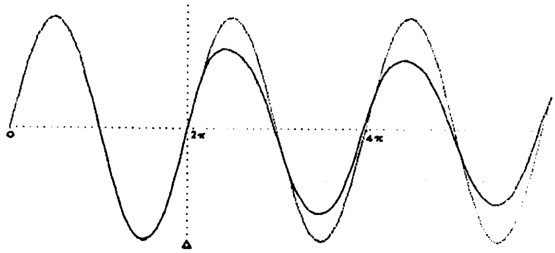
{第2,3層のニューロン数は30,1である}

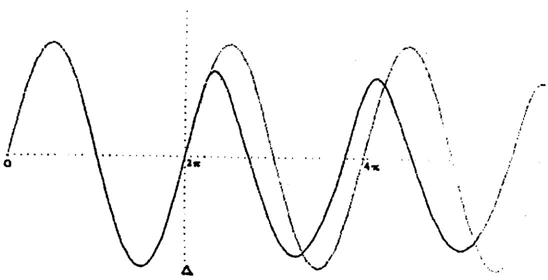
不十分な学習データから予想できる原因は「全関数値を近似する関数をネットワークが見いだした」ためと思われるが，どのような関数を見いだしているのかを調べるために，データ比率80%の場合について，ニューラルネットワーク内部の情報伝搬状況を再構築学習法[3]で調べた．この方法はネットワーク内の有効パスを自動的に強調する．従って見いだされた関数形を推測できるのである．

結果はバイアスニューロンを除き有効なニューロンは第１層で(1,15,16)番目，第２層で1個であった．バイアスニューロン（第１層と第２層に１個づつ必要である）を入れ，ニューロン間6パスのみが有効なネットワークの学習時の出力と教師値の最大2乗誤差は 21.1×10^{-5} であった．十分な精度で学習できている．

正弦関数は2次近似の範囲では，第１層(15,16)番目のニューロン，第２層は2個のニューロンで近似できる．[6]これを確かめるためにバイアスニューロン間パスを含めた8パスのニューラルネットワークで同様の計算を行った．最大2乗誤差は 2.0×10^{-5} となり，さらに良く学習できた．

以上２つの場合において，学習区間では，ネットワークから正しい結果が出力される．しかし関数の巡回表現によって予想を行うと下図のようになった．それぞれ，Δ=(1.15,1.17,1.25)，Δ=(1.17,1.20,1.23)である．この結果は「学習が容易で高精度に可能」なことが必ずしも予想精度を保証しないことを示す．





実際，再構築学習の途中で見いだされた別の8パスからなるニューラルネットワークでは最大2乗誤差は 22.0×10^{-5}，Δ=(1.15,1.20,2.17)であり，全パスを有効にすると最大2乗誤差は 1.0×10^{-5}，Δ=(1.20,2.22,>3)である．学習区間では冗長のように見えるニューロンも予想に関しては何らかの寄与をしていると考えざるを得ない．特にΩ区間の幅の両端のパスが忘却作用の結果生き残ることは，予想誤差を小さくする効果があり，注目すべきである．

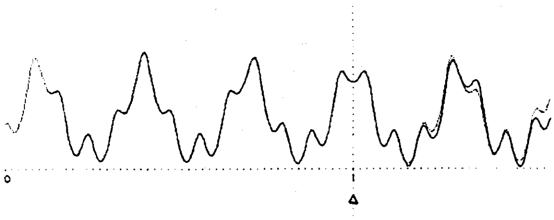
不十分な学習データから相当の精度で予想できる場合は正弦関数のほか，$y=\sin(x)+\sin(8x)/8$，$y=\Sigma\exp\{-\alpha(x-p_i)^2\}+b\Sigma\exp\{-\beta(x-q_i)^2\}$ でも見いだされた．下図に，

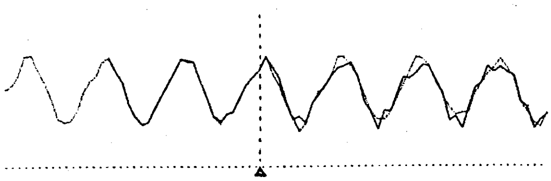
$$y=\sum^{20}\exp\{-\alpha(x-p_i)^2\}+b\sum^{60}\exp\{-\beta(x-q_i)^2\},$$

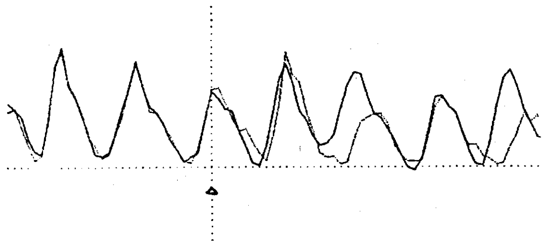
$\alpha=200$, $b=0.4$, $\beta=2000$, $p_i=\{-2+0.3*i;\ i=0,1,2,\dots 19\}$, $q_i=\{-2+0.08*i;\ i=0,1,2,\dots 59\}$，学習区間(0,1)，予想区間(1,1.56)の予想結果を示す．学習区間は160点で離散化した．Δ=(1.53,>1.56,>1.56)であった．

階層型ニューラルネットワークと関数の巡回表現を用いた予想法は，不十分な観測データの中に含まれる

何らかの関数的関連を抽出し，その関連から予想できるように見える．少なくとも，正弦／余弦／ガウス関数の線形結合については可能である．この機能の応用は広範囲である．実例として，月平均最高気温の変動，太陽黒点の相対数変動について予想値の計算を行ったところ，40ヵ月の学習期間から±10%の誤差で6ヵ月，160年の学習期間から±30%の誤差で19年間予想した．







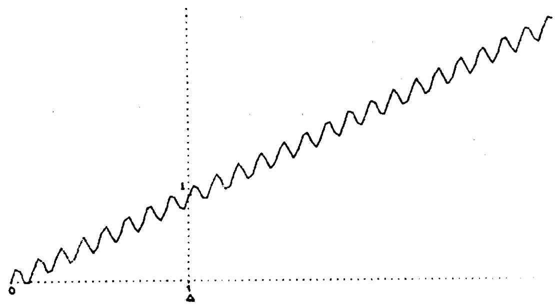
4．非周期関数の予想

非周期関数の場合，一般には，関数の巡回表現と階層型ニューラルネットワークを用いて予想できない．しかし，非周期関数に演算子を作用させて周期関数に変換出来る場合は予想が可能である．もし演算子が見いだせない場合でも，非周期関数の一部に繰り返しパターンがある場合，パータンの類似性を使って予想期間を拡大することはできる．

4.1　過去のデータに見いだされる局所的データの関連が絶対値の異なる状況でも
　　成立するとき

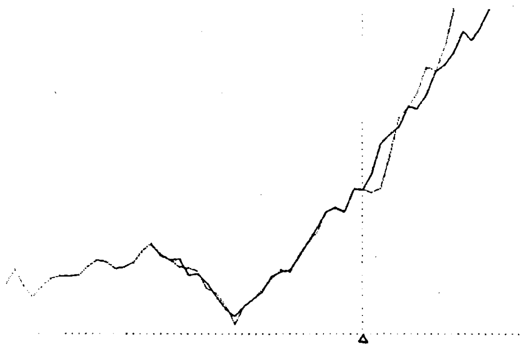
$\{v'\}=\{\ \{v_1-\gamma,\ldots\ldots,v_{i+1}-\gamma\ ;\ \gamma=\min\{v_1,\ldots,v_{i+1}\}\}\ ,\ldots\ldots\}$なるγだけシフトしたパターン集合{v'}とシフトした教師データについてニューラルネットワークで学習する．γ値は入力パターン毎に異なる．このようなシフトしたパターンで学習した場合，ニューラルネットワークで，$\xi=\Xi(P)$ を計算する場合もPの各要素から，その各要素から計算されるγを減じる．ネットワークの出力は$\xi+\gamma$が真のξ値である．

このようにニューラルネットワークを使用するとy=ax+bのような関数でも，関数の巡回表現によって予想が可能である．下図に y=x+sin(8x)/8 の予想を示す．

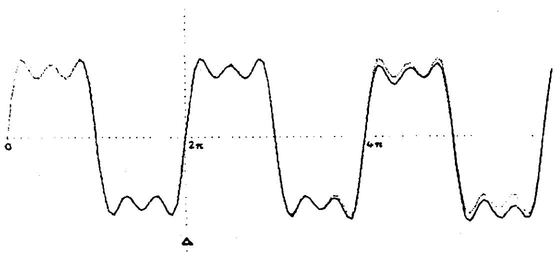


単にニューラルネットワークを用いると，第２層のニューロンの動作関数がシグモイド関数であるため(0,1)区間を越える出力値は精度が保証出来ないが，上記のように入出力データにγ演算子を作用させることによりそれが可能になる．

実例として，特許出願件数の予想を行った．



4.2 時間発展していく系の中に見いだされる局所的な時間に対して対称的な構造を，予想に適用できるとき $\{P'\}=\{\{v_{i+1},\ldots\ldots,v_1\},\ldots\ldots\}$ なるパターン集合 $\{P'\}$ と教師データについて，ニューラルネットワークで学習する．
予想するとき，時間反転しないデータから作成したネットワークと反転したデータから作成したネットワークを切り替えながら使用する．ネットワークを切り替える基準は関数の巡回表現から得られた $P_{n+1}=\{P_n; -v_{n-i+1}\ \ +v_j\}$ が $\{P_1,\ldots,P_n\}$，$\{P'_1,\ldots,P'_n\}$ のどちらの集合の要素に近いかである．２種類のネットワークを使用するのは，１個のネットワークでは $\{P_1,..,P_n; P'_1,..,P'_n\}$ が学習できないからである．
下図に $y=\sin(x)+\sin(3x)/3+\sin(5x)/5$，学習区間 $(0, 2\pi)$，予想区間 $(2\pi, 6\pi)$ の結果を示す．学習区間は40点で離散化した．Ω は16点である．



5．まとめ

(1)本論文で示した階層型ニューラルネットワークと関数の巡回表現を用いた時系列現象の予想法は，関数形を部分断片に分解し，断片を組み合わせて外挿時の関数の形を予想するという原理に基づいている．
従って断片の中に関数の将来の形と同じものが存在すれば精度良く予想可能である．

(2)以上は関数の形から見た予想法の根拠であるが，階層型ニューラルネットワークは学習に用いたデータの中から，ニューロン動作関数の線形結合によって表現できる基底関数を自動的に生成し，この基底関数の線形結合によってデータを再現する関数を自動生成する働きがあるため，この機能が働くような正弦／余弦／ガウス関数の線形結合で表されるような現象では，過去の関数形に将来の関数形を示すような断片が見いだせない場合でも高精度の予想が可能である．

(3)ニューラルネットワークは周期的な時系列現象の予想に有効であるが，無限大に発展していくような現象でも，現象を周期的な関数に投影する演算子が見いだせるならば，上記の原理によって高精度に予想が可能である．

文献

1)上村龍太郎「回帰ニューラルネットワークによる時系列パターンの復元」情報処理学会人工知能研究会報告，No.70，pp.1(1990).
2)松葉郁夫，数理科学8月号，pp.31(1991).
牧秀行，吉原郁夫「ニューラルネットワークを用いた時系列の上下変動予測手法の提案」，情報処理学会第50回全国大会(1995.3).
3)T.Aoyama and H.Ichikawa,"Reconstruction of Weight Matrices in Neural Networks, A Method of Correlating Outputs with Inputs",Chem.Pharm.Bull.,1991,39,pp.1222(1991).
4)上田浩次，山田宗男，堀場勇夫，池谷和夫，鈴木賢治「アナログ出力ニューラルネットワークを用いた駐車率の直接推定方法」情報処理学会論文誌，pp.627(1995).
5)磯部孝編「相関関数およびスペクトル」pp.3-10，東京大学出版会(1968).
6)吉原郁夫，私信(1995.4.19).