

二次遅れを考慮したニューラルネットワークによる 時系列の学習

7H-7

本間 宏樹

芳沢 聡

外山 淳

新保 勝

北海道大学工学部

1 はじめに

脳の記憶機構を考える上で、ある時系列の一部分から全体を連想することは特に記憶の連想をモデル化する重要な問題の一つである。時系列を学習する神経回路モデルとしては Doya ら [1] によるニューラルネットワークに非対称結合を持たせ、状態遷移を起こさせることにより周期パターンを発生させるものがある。また、エルマン [2] は一次遅れを持つリカレントネットワークを提案している。これらのモデルを用いることにより、時系列を学習し、その一部分からそれに続く時系列を予測することができる。しかし、この方法では繰り返しを含む時系列をうまく学習しない場合がある。

本研究では、このリカレントネットワークに二次遅れの層を持たせ、時系列の学習を強化するモデルを提案する。

2 エルマンのモデルと学習方法

エルマンのモデルは入力層と中間層、出力層、CONTEXT 層のそれぞれにユニットを持つモデルである（図1）。入力層に入った信号は中間層を経て出力層と CONTEXT 層に同時に伝搬する。ここで、中間層から CONTEXT 層へは一对一の対応であり、結合係数は全て1に固定している。このため中間層と CONTEXT 層のユニットの数は同じである。また、入力層と出力層のユニットの数も同じである。次の信号が入力層に入り中間層へ伝わる時刻には、一時刻前に CONTEXT 層に蓄えられていた

信号も同時に中間層へ伝搬する。ただし、中間層から CONTEXT 層への結合以外は全て相互結合している。

教師信号は入力層へ入る信号の次の時刻の信号である。このモデルに二見ら [3] は誤差逆伝搬法を適用し、図1に示す結合係数 w, v をそれぞれ修正することを試みている。

3 提案法

エルマンのモデルでは、繰り返しを含む時系列の含むような入力に対して、うまく時系列を学習できない場合がある。中間層へ入力される CONTEXT 層からの出力は一時刻前の影響が最も強いため、時系列が繰り返しを含むような場合は学習がうまく出来ないと考えられる。そこで、図2に示すように二次遅れの層を加えることにより時系列学習を、より強化することができる。

中間層への入力は入力層の出力と一時刻前に中間層が出力した信号、二時刻前に中間層が出力した信号がそれぞれの結合係数を介して同時に送られる。学習はそれらの結合係数と中間層-出力層間の結合係数を全て修正させるものであり、エルマンのモデルに誤差逆伝搬法を適用するという点では二見らと同じである。モデルのダイナミクスは以下の通りである。

中間層の入力、出力をそれぞれ $u_i(t), v_i(t)$ とし、 θ_j を閾値、 $p_i(t), q_i(t)$ をそれぞれ一次遅れ、二次遅れのユニットの出力、 $x_j(t)$ を入力層のユニットの出力、 w_{ji} を入力層-出力層間の結合係数、 $w_{ji}^{(k)}$ を k 次遅れの層-中間層間の結合係数とすると、

$$u_i(t) = \sum_j w_{ji} x_j(t) + \sum_j w_{ji}^{(1)} p_j(t) + \sum_j w_{ji}^{(2)} q_j(t) + \theta_j$$

Learning of Time Series Patterns Using Neural Network with Second-Order Time Lag Components.

Hiroki Homma, Satoshi Yoshizawa, Jun Toyama and Masaru Shimbo.

Faculty of Eng., Hokkaido Univ.

Kita 13, Nishi 8, Sapporo 060, Japan.

$$v_i(t) = s(u_i(t)) \quad s: \text{シグモイド関数}$$

となる。また、教師信号を $T_j(t)$ としたときの全体の誤差 $E(t)$ は

$$E(t) = \frac{1}{2} \sum_j (T_j(t) - v_j(t))^2$$

また、学習則には慣性項 α も加えて一括修正し、次の量だけ結合係数を変化させる。

$$\Delta w_{ji} = -\eta \left[\frac{\partial E(t)}{\partial w_{ji}} + \left(\frac{\partial E(t)}{\partial w_{ji}^{(1)}} \right)^{-1} + \left(\frac{\partial E(t)}{\partial w_{ji}^{(2)}} \right)^{-2} \right] + \Delta^{-1} w_{ji} \cdot \alpha$$

$$\Delta w_{ji}^{(1)} = -\eta \frac{\partial E(t)}{\partial w_{ji}^{(1)}} + \Delta^{-1} w_{ji}^{(1)} \cdot \alpha$$

$$\Delta w_{ji}^{(2)} = -\eta \frac{\partial E(t)}{\partial w_{ji}^{(2)}} + \Delta^{-1} w_{ji}^{(2)} \cdot \alpha$$

ただし、 η は学習定数であり、 $\Delta^{-1}y$ は一時刻前の y の修正量を表す。

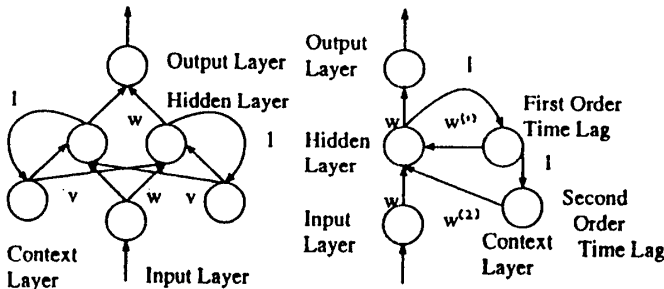


図1：一次遅れを持つネットワーク 図2：二次遅れを持つネットワーク

4 実験

実験では繰り返しを含む時系列と、含むものを用いて学習の度合を比較する。まず、入力信号として表1のようなものを考える。入力層のユニット数を4、中間層のユニット数を6とする。

表1：実験に用いた時系列

ド	0001	ソ	0101	#レ	1010
レ	0010	ラ	0110	#ファ	1011
ミ	0011	シ	0111	#ソ	1100
ファ	0100	#ド	1001	#ラ	1101

まず、繰り返しを含まない例として、 $I = [\text{ミ}\#\text{レ}\text{ソ}\text{シ}\text{レ}\text{ド}\text{ラ}]$ を入力信号とする学習を行なう。結果は、 $E(t)$ の平均を学習誤差とすると、一次遅れだけを持つモデルでは、学習回数511回で学習誤差が0.1以下となった。また、二次遅れを考慮したものは、

学習回数434回で学習誤差が0.1以下となった。

次に、繰り返しを含む例として $I = [\text{ミ}\#\text{レ}\text{ミ}\#\text{レ}\text{シ}\text{レ}\text{ド}\text{ラ}]$ を入力信号とする学習を行なう。結果は表2に示す。

表2：学習結果

教師信号	出力信号 (1次遅れ)	(2次遅れ)
ミ 0011	.00,.00,.89,.97	.00,.00,.94,.94
#レ 1010	.54,.15,.89,.18	.66,.33,.90,.26
ミ 0011	.03,.00,.67,.98	.00,.00,.95,.97
#レ 1010	.76,.20,.95,.17	.79,.23,.91,.20
ミ 0011	.01,.00,.94,.98	.00,.00,.95,.95
シ 0111	.89,.13,.96,.14	.85,.15,.91,.15
レ 0010	.09,.01,.99,.17	.02,.12,.99,.03
ド 0001	.13,.00,.82,.98	.07,.01,.13,.96
ラ 0110	.15,.96,.99,.02	.04,.93,.99,.01

表2より一次遅れだけのモデルでは'シ'と'ド'の学習が失敗したほか、2番目の'#レ'についても学習が思わしくなかった。二次遅れを考慮した場合は'シ'が学習できなかったが、その他の信号についてはうまく学習できた。

5 おわりに

二次遅れを考慮したニューラルネットワークを用いた時系列の学習について検討した。二次遅れを考慮したモデルの方が一次遅れの層だけを持つモデルより時系列の学習を強化することができた。このことから、中間層へ入る過去の信号が一時刻前の影響ばかりにとらわれないため、学習が強化できたと考えることができる。しかし、繰り返し頻度が多くなる場合には、二次遅れを考慮しても学習できない部分があった。今後の課題としては、過去の時系列の影響をどの程度まで考慮すればよいのか検討する必要がある。

文献

[1]K.Doya and S.Yoshizawa: Adaptive Neural Oscillator Using Continuous-Time Back-propagation learning, Neural Networks, 2(1989), 375-385.
 [2]Jeffrey L.Elmán: Finding Structure in Time Cognitive Science, 14(1990), 179-211.
 [3]二見、星宮:階層構造の神経回路モデルによる時系列パターンの発生と認識. 信学技報, NC90-6(1990), 39-46.