

# BP 学習におけるゆとりニューラルネットワークのもつ 柔軟性について

## Flexibility of an Affordable Neural Network for Back Propagation learning

上手洋子<sup>†</sup>                      西尾芳文<sup>†</sup>  
<sup>†</sup>徳島大学工学部電気電子工学科

Yoko Uwate<sup>†</sup>                      Yoshifumi Nishio<sup>†</sup>  
Dept. of Electrical and Electronic Eng., Tokushima University<sup>†</sup>

### 1 はじめに

日本では、これまでの「詰め込み」教育が批判され、子どもたちの多様な能力をのばすために、いわゆる「ゆとり」教育が導入されている。我々は、ニューラルネットワークの学習においても、「ゆとり」という概念が柔軟性や創造性に対して非常に重要であると考ええる。

ニューラルネットワークはバックプロパゲーション (BP) 学習 [1] が提案されて以来、パターン認識や制御、信号処理など様々な分野で活用され、その有効性が明らかにされている [2]-[4]。実際の脳が何らかの処理をする際には、全てのニューロンが常に一斉に動作しているのではなく、関連するニューロンが必要に応じて動作していると考えられる。我々はこの概念を階層型ニューラルネットワークの中間層に応用したネットワークを提案した [5]。このネットワークは中間層に必要な以上のニューロンを用意し、全ての中間層ニューロンを用いるのではなく、動作する中間層ニューロンが毎更新時に切り替わるというものである。つまり、毎回の更新時に動作していないニューロンが存在することになる。このネットワークのことを“ゆとりニューラルネットワーク”と呼んでいる。文献 [5] では、簡単な二次関数を学習させるシミュレーションを行い、ゆとりニューラルネットワークの有効性を確認している。

本研究では、より難しい問題を対象として様々な計算機シミュレーションを行い、ゆとりニューラルネットワークのもつ様々な性能の調査を行う。具体的な学習例としては、skew tent map によって生成されるカオス時系列を用いる。まず、系列長が 10 の時系列を 100 パターン学習するときの学習精度、収束速度について評価を行う。さらにネットワークの学習が終わったあとで、学習に用いた 100 パターンとは異なる未知のカオス時系列パターン

を入力したときに、入力と同じ時系列パターンを出力できるかの調査を行う。また、中間層ニューロンの切り替わり方法として、カオス、規則、ランダムの場合についての比較を行う。

### 2 ゆとりニューラルネットワーク

実際の脳構造を反映させるために中間層にゆとりニューロンを導入する。あらかじめ必要以上の中間層ニューロンを用意し、全ての中間層ニューロンを用いるのではなく、動作する中間層ニューロンが毎更新時に切り替わるというものである。つまり、毎回の更新時に動作していないニューロンが存在することになる。これがゆとりニューロンである。ゆとりニューロンを導入した階層型ネットワークモデルを図 1 に示す。

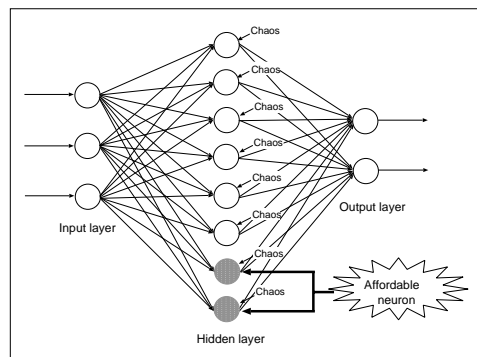


図 1: ゆとりニューラルネットワーク。

## 2.1 カオスの動作方法

我々はこれまで組み合わせ最適化問題を解く Hopfiled NN のニューロンにカオスノイズを注入することで局所解脱出がスムーズになり性能が向上することを確認している。ランダムノイズとの比較も行い、カオスをノイズとして用いた場合で最も良い結果が得られている [6]-[7]。これらのことから、ニューラルネットワークでの局所解脱出にカオスが有効であると推測できる。本研究では、更新時にどのニューロンを動作するかを判定するために、カオス写像である skew tent map を用いる。ここではカオスはあくまでどの中間層ニューロンを選ぶかに用いるのであって、ネットワークやニューロンがカオス的な動作になるのではない。skew tent map はカオスを発生させる非常にシンプルな写像である。skew tent map は式 (1) で表され、図 3 に示すように tent map の頂点座標が  $a$  だけずれたものである。今回は  $a = 0.05$  の skew tent map を用いた。中間層ニューロンに初期値の異なる skew tent map をそれぞれ用意し、更新時に skew tent map の値が大きい順にいくつかの中間層ニューロンを使用する。

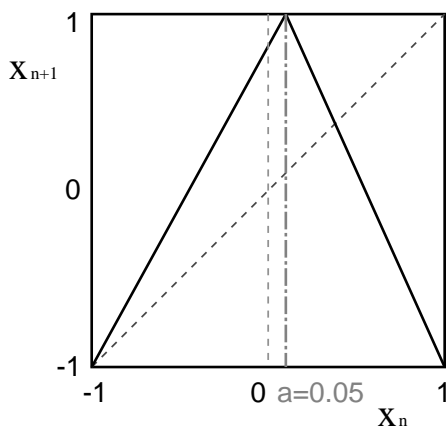


図 2: skew tent 写像.

$$x_{n+1} = \begin{cases} \frac{2x_n + 1 - a}{1 + a} & (-1 \leq x_n \leq a) \\ \frac{-2x_n + 1 + a}{1 - a} & (a < x_n \leq 1) \end{cases} \quad (1)$$

## 2.2 カオス以外の切り替わり方法

文献 [5] では中間層ニューロンの切り替わり方法として、skew tent 写像を用いることによってカオス的な切り替わり方をするネットワークを提案したが、本章では、規則的またはランダムに中間層ニューロンが切り替わった場合についての調査を行う。規則的な切り替わり方とは、中間層の動作するニューロンを順番に選ぶ方法である。ランダムな切り替わり方は、skew tent 写像の時系列の代わりに乱数を用い、乱数の値が大きい中間層ニューロンから動作するニューロンと決定する。

## 3 BP 学習アルゴリズム

BP 学習アルゴリズムとして D.E. Rumelhart らによって提案された標準的なものを用いる [1]。本研究では結合荷重の修正法として一括修正法を使用した。一括修正法は全データに対する誤差を小さくするような結合荷重の組み合わせを探索する。

ネットワークの評価関数として、教師信号と出力値の 2 乗誤差  $E$  を考える。

$$E = \sum_{p=1}^P E_p = \sum_{p=1}^P \left\{ \frac{1}{2} \sum_{i=1}^N (t_{pi} - o_{pi})^2 \right\} \quad (2)$$

ここで、 $P$  は入力データのパターン数、 $N$  は出力層のニューロン数、 $t_{pi}$  は入力データ  $p$  を提示したときの教師信号値、 $o_{pi}$  は入力データ  $p$  を提示したときの出力値を表している。結合荷重は  $E$  を最小に変化すればよいので、結合荷重の変化量は次式のようになる。

$$w_{i,j}^{k-1,k}(m+1) = w_{i,j}^{k-1,k}(m) + \sum_{p=1}^P \Delta_p w_{i,j}^{k-1,k}(m)$$

$$\Delta_p w_{i,j}^{k-1,k}(m) = -\eta \frac{\partial E_p}{\partial w_{i,j}^{k-1,k}} \quad (3)$$

ここで、 $w_{i,j}^{k-1,k}$  は第  $k-1$  層の  $i$  番目のニューロンと第  $k$  層の  $j$  番目のニューロンとの結合荷重、 $m$  は学習回数、 $\eta$  は学習係数である。本研究では、学習の発散を防ぎ、学習の高速化をはかるために次式右辺第 2 項のような慣性項を導入する。

$$\Delta_p w_{i,j}^{k-1,k}(m) = -\eta \frac{\partial E_p}{\partial w_{i,j}^{k-1,k}} + \zeta \Delta_p w_{i,j}^{k-1,k}(m-1) \quad (4)$$

ここで  $\zeta$  は慣性係数である。

## 4 シミュレーション結果

本研究では、学習パターンとして skew tent map によって生成されるカオス時系列を用いる。ここでカオス時系列の系列長は 10 で、パターン数は 100 とした。我々は、これらの学習に対するゆとりニュー

ラルネットワークの性能評価を行う。さらにネットワークの学習後に、学習パターン以外のカオス時系列を入力したときの出力についての調査も行う。本シミュレーションで用いたパラメータは以下のように設定した。ネットワークの学習係数  $\eta = 0.05$ 、慣性係数  $\zeta = 0.02$ 。また、ネットワークの重みの初期値は  $-1.0$  から  $1.0$  の一様乱数で与えた。中間層ニューロンを 20 個用意し、ゆとりニューロンを 0 から 10 個に変化させたときの学習精度について調査を行う。学習精度のシミュレーション結果を図 3 に示す。縦軸は学習パターンに対するネットワーク出力の誤差、横軸は動作する中間層ニューロンの数である。この図から、ゆとりニューロンがあるネットワークのほうが学習の精度がよいことが確認できる。また、中間層ニューロンの切り替わり方法としては、規則的な切り替わりをするよりもカオスまたはランダムに切り替わるほうがネットワークの学習に有効であることがいえる。

次に、学習後に学習パターン以外の未知の入力パターンに対する学習精度の結果を図 4 に示す。縦軸は未知の入力パターンに対するネットワークの誤差、横軸は動作する中間層ニューロンの数である。先ほどと同様に未知の入力パターンに対しても、ゆとりニューラルネットワークのほうがよい性能を得ることが確認できる。また、中間層ニューロンの切り替わり方法としてはカオスまたはランダムに切り替わるネットワークのほうが有効であることがいえる。未知の入力データに対する出力の一例を図 5 に示す。図 5(a) はゆとりニューロンの存在しない従来のネットワーク、(b) はカオスの切り替わるゆとりニューラルネットワーク (ゆとりニューロン数: 2)、(c) は規則的に切り替わるゆとりニューラルネットワーク (ゆとりニューロン数: 2)、(d) はランダムに切り替わるゆとりニューラルネットワーク (ゆとりニューロン数: 2) の場合を表している。これらの図より、ゆとりニューラルネットワークは未知データに対しても誤差が小さいのに対して、従来のネットワークは誤差が大きいことがわかる。

今回の学習例では、入力パターンをそのまま出力するようなネットワークの学習であるので、比較的簡単な問題ではあるが、学習パターン以外のパターンが入力されたときなどには、これまでのネットワーク構造では困難であることがわかった。それに対してゆとりニューロンを用いたネットワークは未知のパターンに対して有効であることが確認できた。ニューラルネットワークの学習には過学習という問題点があるが、ゆとりニューラルネットワークを用いることで、汎化能力の高いネットワーク構造を形成することができるのではないかと考える。

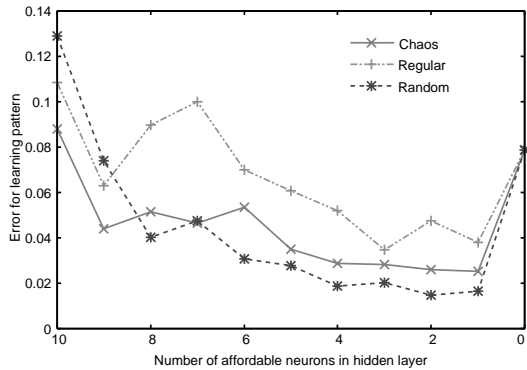


図 3: 学習パターンに対するネットワークの誤差。

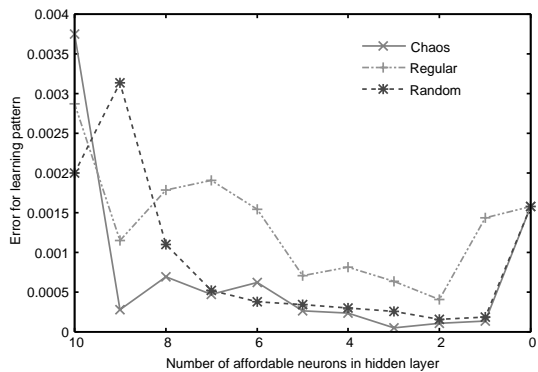


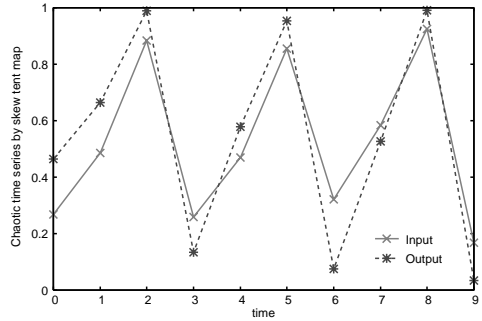
図 4: 未知のパターンに対するネットワークの誤差。

## 5 まとめ

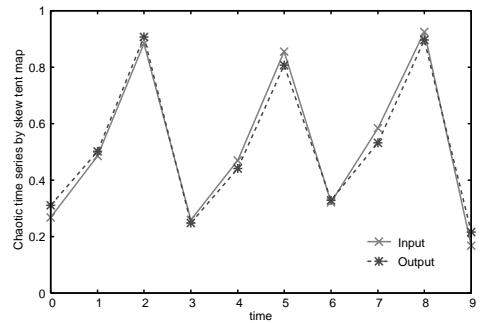
本研究では、BP 学習を行うときに必要以上の中間層ニューロンを用意し、毎更新時に全ての中間層ニューロンを用いるのではなく、動作する中間層ニューロンが切り替わるゆとりニューラルネットワークの性能評価を行った。学習パターンとして、skew tent 写像によって生成されるカオス時系列を用いた。さらに、ネットワークの学習後に未知の入力パターンに対する出力についての調査を行った。シミュレーションの結果、ゆとりニューラルネットワークのほうがネットワークの学習精度、また学習後の未知の入力パターンに対してもよい性能を示すことがわかった。中間層ニューロンの切り替わり方法として、カオス、規則、ランダムの場合においてシミュレーションを行ったが、規則的に切り替わるよりも、カオスやランダムに切り替わるほうがよい学習精度が得られ、未知の入力パターンに対しても有効であることを確認できた。

## References

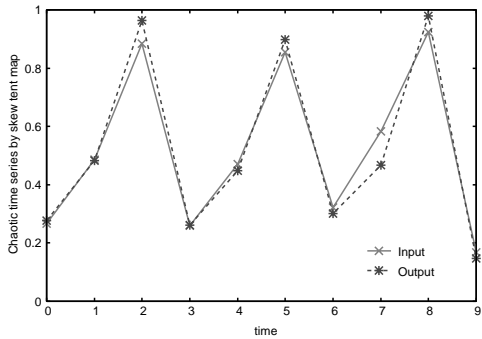
- [1] D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning Internal Representations by Error Propagation," *Parallel Distributed Processing*, vol. 1, MIT Press, MA, pp. 318-362, 1986.
- [2] Y. LeCun, B. Boser, J.S. Denker, et al, "Handwritten Digit Recognition with a Back-Propagation Network," *Advances in Neural Information Processing Systems*, vol. 2, MIT Press, MA, 1990.
- [3] M. Kawano, Y. Uno, M. Isobe and R. Suzuki, "A Hierarchical Neural Network Model for Voluntary Movement with Application to Robotics," *IEEE Control Systems Magazine*, vol. 8, pp. 8-16, 1988.
- [4] T.J. Sejnowski and C.R. Rosenberg, "Parallel Networks that Learn to Pronounce English Text," *Complex Systems*, vol. 1, pp. 145-168, 1987.
- [5] 上手洋子, 西尾芳文 "ゆとりニューロンを導入した階層型ニューラルネットワークのBP学習," 2004年 電子情報通信学会 ソサイエティ大会 講演論文集, no. A-2-9, p. 42, Sep. 2004.
- [6] T. Ueta, Y. Nishio and T. Kawabe, "Comparison between Chaotic Noise and Burst Noise on Solving Ability of Hopfield Neural Networks," *Proc. NOLTA '97*, vol. 1, pp. 409-412, Nov. 1997.
- [7] Y. Uwate, Y. Nishio and A. Ushida, "Markov Chain Modeling of Intermittency Chaos and its Application to Hopfield NN," *IEICE Transactions on Fundamentals*, vol. E87-A, no. 4, pp. 774-779, Apr. 2004.
- [8] Y. Uwate, Y. Nishio, T. Ueta, T. Kawabe and T. Ikeguchi, "Performance of Chaos and Burst Noise Injected to the Hopfield NN for Quadratic Assignment Problems," *IEICE Transactions on Fundamentals*, vol. E87-A, no. 4, pp. 937-943, Apr. 2004.



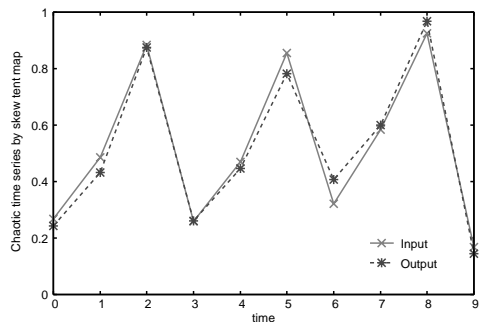
(a) 従来のネットワーク.



(b) ゆとりニューロン2個 (カオス).



(c) ゆとりニューロン2個 (規則).



(d) ゆとりニューロン2個 (ランダム).

図5: 異なる中間層ニューロンの切り替えりによる未知パターンに対する出力結果.