

反応度と誤差を利用して隠れユニット数を動的に変更する階層型ニューラルネットワーク

中尾 智子 鈴木 豪 乾 伸雄 小谷 善行

東京農工大学工学部電子情報工学科

1 はじめに

階層型ニューラルネットワークでは隠れ層の構成、特に隠れユニット数をどのように決定するかが問題となっている。一般に、隠れユニット数が少ないと解けない問題が存在し、また、隠れユニット数が必要以上に多くあってもネットワークの機能は大きくは変わらないことが知られている[3]。そこで、与えられた問題を解決するために必要とされる精度を達成できる、より少ない数の隠れユニット数を求める手法が必要とされている。

本研究では、隣り合う層のユニットが全て結合した3層の階層型ニューラルネットワークにおいて、ネットワークの反応度と誤差を利用して隠れユニット数を追加していく逐次的生成方法の一手法を提案する。

2 従来手法

動的に隠れユニット数を求める方法として、小さいネットワークから学習を始めて、必要な隠れユニットを追加していく逐次的生成方法と、大きいネットワークから学習を始めて、不要な隠れユニットを削除していく逐次的削減方法がある。逐次的削減方法には、G-Factor や E-Factor といった評価関数を使って隠れユニットを削除する方法などが提案されている[2]。逐次的生成方法には、隠れユニットをひとつ追加して全ての入・出力ユニットと結合させ、全てのネットワークを学習する DNC、追加した隠れユニットのみを学習する PPR、また、追加した隠れユニットを任意のユニットに結合させる GMDH などが提案されている[1]。しかし、実際には隠れユニット数を求める方法として、試行錯誤的に隠れユニット数を与えて学習を試みる方法が使用されている。

3 反応度と誤差を利用した動的変更方法

競合学習における逐次的生成方法に、正しく学習できていない入力データの値を重みベクトルに持つ隠れユニットを生成して、教師値に合う出力ユニットにのみ結合させる RCE モデルが提案されている[3]。本稿では、この

RCE モデルを全結合した3層の階層型ニューラルネットワークに応用し、学習を行う方法を提案する。

ユニット i への入力値を x_i 、ユニット i とユニット j 間の重みを w_{ij} 、ユニット j の閾値を θ_j とすると、隠れ層・出力層のユニット j からの出力値 O_j は、次式で計算される。伝達関数 $f(x)$ にはシグモイド関数を利用する。

$$O_j = f\left(\sum_i w_{ij}x_i - \theta_j\right)$$

ネットワークの学習には、バックプロパゲーション法を用いる。ここでは、学習データの入力値と教師値は0,1の2値で与えると仮定する。

学習データ k に対するネットワークの反応度 $r(k)$ を次式で定義する。

$$r(k) = \frac{1}{1 + E(k)}$$

ここで、 $E(k) (\geq 0)$ は学習データ k に対する誤差である。ネットワークの反応度が小さいほど学習データに対して正しい反応ができていないことを意味する。

隠れユニット数の動的変更には、次の(1)から(4)の条件にしたがって隠れユニットをひとつずつ追加する方法をとる。隠れユニットを追加した後は、全てのネットワークについて学習を行う。

(1) 学習の終了条件

誤差が既定値よりも小さくなったときに終了する。

(2) 初期化

ネットワークの初期化を次のように行う。

- ① 隠れユニット数は1個とする。
- ② 重みは乱数で初期化する。
- ③ 閾値は0とする。

(3) 隠れユニットを追加するタイミング

隠れユニットを追加するタイミングは、与えられた M に対して式(1)を満たしたときとする：

$$\max_{t-M < i \leq t} |E^i - E^{i-1}| < \varepsilon \quad \dots(1)$$

ε : 定数

ここで、 E^i は $i \times N$ 回目 (ただし、 N は学習データ数) の学習を行った後での全学習データに対する誤差を表わす。式(1)は、隣接する誤差の差が定数 ε よりも小さ

い場合が M 回続いた場合に隠れユニットをひとつ追加することを意味する。

(4) 追加した隠れユニットに関する重み・閾値の初期値設定方法

追加した隠れユニットに関する重みと閾値は、初期値を次のように設定する。

- ① ネットワークの反応度が最小である学習データ p を求める。ここで、 LP は学習データの集合を表わす。

$$p = \arg \min_{k \in LP} r(k)$$

(ただし、ネットワークの反応度が最小である学習データが複数個ある場合には、与えられた学習データの順番が先にある方とする)

- ② 入力ユニット i と追加した隠れユニット h 間の重み w_{ih} は、学習データ p の入力値 x_{pi} と同じ値にする。

$$w_{ih} = x_{pi}$$

- ③ 追加した隠れユニット h と出力ユニット j 間の重み w_{jh} は、学習データ p の教師値 t_{pj} と同じ値にする。

$$w_{jh} = t_{pj}$$

- ④ 追加した隠れユニット h の閾値 θ_h は 0 とする。

$$\theta_h = 0$$

これは、隠れユニットを追加する前のネットワークでは正しく反応できなかった学習データに対して、正しい反応をするような隠れユニットを追加する手続きになっている。

4 学習実験

本手法を用いて、3OX ゲームの評価関数の学習実験を行った。入力データには局面の生情報を与え、教師値には探索で求めた局面の完全解を与えた。入力ユニット数は 64 個、出力ユニット数は 3 個、学習データ数は 128170 個である。誤差を求めるのには平均二乗誤差を使用した。ネットワークの反応度を求めるための誤差には二乗誤差を使用した。

学習率は 0.75、ネットワークにおける重みの初期化は範囲 $[-0.3, 0.3]$ の乱数で行い、初期の隠れユニット数は 1 個、隠れユニットを追加するタイミング用のパラメータは $\epsilon = 0.02$ 、 $M = 50$ として実験を行った。実験で得られた誤差曲線の例を図 1 に示す。学習の結果、最終的な隠れユニット数は 76 個となり、学習回数は 5206×128170 回、誤差

は 0.15、正解率は 80.0% であった。

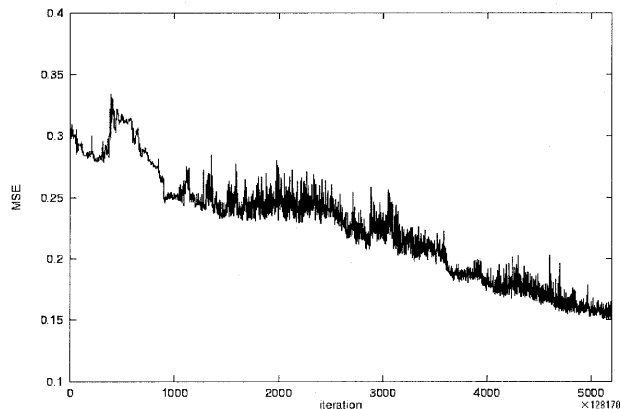


図 1 誤差曲線

図 1 から、学習の終了条件が満たされるまで、誤差が順調に減少していることが分かる。天下りに隠れユニット数を 35, 60, 80, 100 個と与えて学習した場合でも、正解率が 80% を超えたものは 80 個と 100 個のときであった。学習時間の多くが、大きなネットワークに費やされることを考えると、本手法による隠れユニット数の決定方法は、試行錯誤的に隠れユニット数を与えて学習するよりも妥当であると考えられる。

5 おわりに

3 層の階層型ニューラルネットワークの隠れユニット数を求める方法として、ネットワークの反応度と誤差を利用して隠れユニットを追加していく方法を提案した。

本手法は、3OX ゲームの評価関数の学習実験において終了条件を満たす小さいネットワークを見つけることができた。

参考文献

- [1] Tin-Yau Kwok and Dit-Yan Yeung, Constructive Algorithms for Structure Learning in Feedforward Neural Networks for Regression Problems, IEEE Transactions on Neural Networks Vol. XX No. Y, 1999.
- [2] 真島 他, 中間層素子削減のための評価関数 E-factor とその有効性, 日本神経回路学会誌 Vol. 1 No2, 1994.
- [3] 臼井 他, 基礎と実践 ニューラルネットワーク, コロナ社, 1995.

付録 3OX ゲーム

3 目並べの一種。コマを置けるのは最大 3 つまでで、4 つ目に置くとときには 1 つ目に置いたコマはなくなる。

