

(財)電力中央研究所 正会員 山本広祐

**1. まえがき** 近年、Rumelhartらにより開発された階層型ニューラルネットワーク<sup>1)</sup>が、パターン認識や関数近似問題に用いられている。ネットワークの形態の決定にあたっては、入力層と出力層のユニット数は対象とする問題に応じて一意に決定されるが、中間層のユニット数は試行錯誤的に決定されるのが通例であった。この問題を解決するため、中間層のユニット数を評価する方法が幾つか提案(例えば文献2)~4))されている。著者は、構造工学分野における定量的な関数近似問題に焦点を絞り、中間層のユニット数を自動的に調整できる代表的な2つの方法<sup>3), 4)</sup>を検討した。

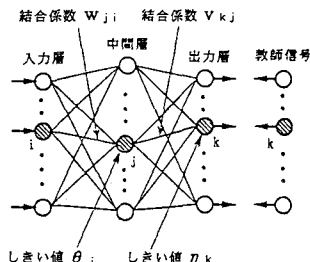
**2. 階層型ニューラルネットワークの原理<sup>1)</sup>** 誤差逆伝搬学習による階層型ニューラルネットワーク(図-1)は、入力信号を受け取ると入力層、中間層、出力層と信号が伝わり、最終的に信号を出力する。入力層の*i*ユニット( $i = 1 \sim n_1$ )から中間層の*j*ユニットへの入力信号を $x_i$ とすると、信号の伝達により*j*ユニットには内部ポテンシャル $r_j$ が蓄積される。この $r_j$ は、シグモイド関数と呼ばれる応答関数の出力として、さらに次層への入力信号 $y_j$ として伝達される。一方、最終的に出力層の*k*ユニットに出力される信号と正しい解(教師信号)との間には誤差が生じるが、この誤差を減らすために各層間の結合係数(結合の重み)と各ユニットのしきい値が修正される。ここで、学習データ*p*に関する誤差関数 $E_p$ を定義する。この $E_p$ を最急降下法を用いて極小化(理想的には最小化)するよう学習が行われ、全学習データに関してこの過程を繰り返すことにより、最小2乗誤差の総和 $E$ の極小化(最小化)が図られることになる。

$$r_j = \sum_{i=1}^{n_1} (w_{ji} \cdot x_i) - \theta_j \quad \dots (1)$$

$$y_j = 1 / \{1 + \exp(-r_j / S)\} \quad \dots (2)$$

$$E_p = \sum_{k=1}^{n_3} (T_{pk} - O_{pk})^2 / 2 \quad \dots (3)$$

$$E = \sum_{p=1}^{n_4} E_p \quad \dots (4)$$



ここに、 $w_{ji}$  : *i*、*j*ユニット間の結合係数

$\theta_j$  : *j*ユニットのしきい値

$T_{pk}$  : 学習データ*p*に関する*k*ユニットの教師信号

$O_{pk}$  : 学習データ*p*に関する*k*ユニットの出力信号

図-1 階層型ニューラルネットワーク

$S$  : シグモイド関数の形状パラメータ

$n_3$  : 出力層ユニット数

$n_4$  : 学習データ総数(総組数)

**3. 中間層ユニット数の自動調整法** 中間層の各ユニットの良否を評価するため、2種類の有効性指数(Badness Factor :  $BAD_j$ <sup>3)</sup>とGoodness Factor :  $Good_j$ <sup>4)</sup>)を導入する。ここで、 $BAD_j$ は中間層の*j*ユニットの全伝搬誤差総和量を表し、 $Good_j$ は*j*ユニットの全伝搬活性化総和量を表す。有効性指数を用いた中間層ユニット数の自動調整アルゴリズムは表-1に示す通りであり、 $BAD_j$ で評価する場合には最大のユニットを、 $Good_j$ で評価する場合には最小のユニットを不良ユニットとして抽出する。

$$BAD_j = \sum_{p=1}^{n_4} \left\{ \sum_{k=1}^{n_3} (v_{kj} \cdot (T_{pk} - O_{pk}))^2 \right\} \quad \dots (5)$$

$$GOOD_j = \sum_{p=1}^{n_4} \left\{ \sum_{k=1}^{n_3} (v_{kj} \cdot O_{pk})^2 \right\} \quad \dots (6)$$

**4. 関数近似問題の設定** Timoshenkoらによって示された周辺単純支持平板が等分布荷重を受ける場合の平板中央線上の応力評価式<sup>5)</sup>の認識問題<sup>6)</sup>を考える。この問題の入力層データは平板の2辺の長さ比と平板中央線上の位置の2ユニットであり、出力層データは単位等荷重に対する2方向の面外曲げモーメントの2ユニットである。入・出力データの総数は70組であり、全て[0,1]に正規化して用いた。なお、 $\Delta$ は、

本問題に関して、入力層-中間層-出力層のユニット数を2-6-2、2-10-2、及び2-15-2の3種類に設定して数値実験を行い、最終的に中間層のユニット数を15に決定して精度の良い解を得ている<sup>9)</sup>。

**5. 検討結果と考察** 一例として、中間層のユニット数を10として計算を開始した時の中間層のユニット数と各学習回における誤差総和Eの推移状況を図-2に示す。ここで、全学習データに対する1通りの学習を1回の学習回と定義した。2500回の学習が終了した時の全学習データに関する正解率は、ほぼ100±3%の範囲内にあった。図-2より、GOOD<sub>2</sub>で評価する方がユニット数の削減傾向が大きい反面、誤差総和Eの消散傾向はBAD<sub>2</sub>の方が良好であることが分かる。このモニタリングでは、BAD<sub>2</sub>で評価する場合には5ユニット、GOOD<sub>2</sub>で評価する場合は4ユニットで誤差総和Eが最小になっている。また、ユニットを1つ削除した後、削除時の誤差総和E<sub>1</sub>の回復を確認するアルゴリズムは、中間層のユニット数を過度に削減しないという機能を有するとともに、ネットワークに適度の"揺らぎ"を与えている。従って、表-1に示したアルゴリズムは局所解からの回避方法としても有効に活用できるものと考えられる。なお、誤差総和Eの消散傾向からみて分かる通り、本問題に関しては有効性指数としてBAD<sub>2</sub>を用いる方が若干有効であるものと考察された。

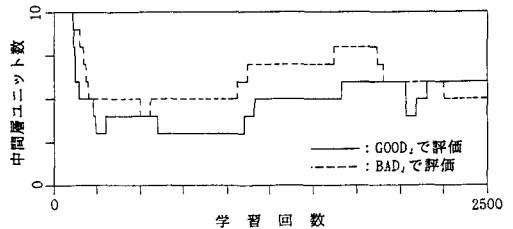
**6. あとがき** 中間層のユニット数は、必要最少限を用いることがネットワークの収束性と計算量の削減の観点から重要であるといわれている。本論で検討した方法は、ネットワークの最適解を求める決定的な方法ではないが、余分な中間層のユニット数を削減し、ネットワーク適度の"揺らぎ"を与える一方法として有効に利用できるであろう。

**参考文献**

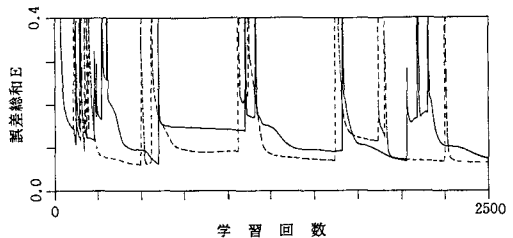
- 1) Rumelhart, D.E., et al. : " Learning Representations by Back-propagating Errors" , Nature, Vol.323, No.9, pp.533-536, 1986
- 2) 栗田多喜夫 : "情報量基準による3層ニューラルネットワークの隠れ層のユニット数の決定法", 電子情報通信学会論文誌 D-II, Vol.J73-D-II, No.11, pp.1872-1878, 1990
- 3) 萩原将文 : "淘汰機能を有するバックプロパゲーション, 学習回数の低減と中間層ユニットの削減法", 電子情報通信学会論文誌 D-I, Vol.J74-D-I, No.6, pp.812-818, 1991
- 4) 松永 豊 他 : "階層型ニューラルネットワークの中間層素子を自動削減する誤差逆伝搬学習アルゴリズム", 電子情報通信学会論文誌 D-II, Vol.J74-D-II, No.8, pp.1118-1121, 1991
- 5) Timoshenko, S.P. and Woinowsky-Krieger, S. : " Theory of Plates and Shells" , pp.113-124, McGraw-Hill
- 6) Alam, J. : " Neural Network Paradigm and its Potential Applications in Structural Analysis" , Extended Abstracts of the 4th ICCBE, pp.91, Tokyo, Japan, 1991

表-1 中間層ユニット数の自動調整アルゴリズム

1) ユニット数を1つ削減する場合
1. 収束基準を設定し、その成立を確認する。
2. 収束基準が成立すれば有効性指数を計算し、不良ユニットを抽出する。
3. 抽出されたユニットを削減する。
2) ユニット数を1つ追加する場合
1. 誤差総和Eをモニタリングする。
2. Eが増加傾向にある場合、所定回数をもって結合係数としきい値を乱数でリセットする。
3. リセット後もEの増加傾向が解消されない場合、ユニットを追加する。
1. 誤差総和Eをモニタリングする。
2. 不良ユニットを削減した時の誤差総和E <sub>1</sub> を記憶する。
3. 削減後、所定回数をもってE <sub>1</sub> を回復できない場合、結合係数としきい値を乱数でリセットする。
4. リセット後もE <sub>1</sub> を回復できない場合、ユニットを追加する。



(a) 中間層ユニット数の推移



(b) 誤差総和Eの推移

図-2 有効性指数に基づく中間層ユニット数の調整