

## 論文

## 階層型ニューラルネットワークの構成と耐故障性の関係

高瀬 治彦 井上 智紀<sup>†</sup> 林 照峯

Haruhiko TAKASE, Tomonori INOUE, Terumine HAYASHI

(電気電子工学科 Electrical and Electronic Engineering)

(Received September 18, 2000)

## Abstract

We propose a new learning algorithm to enhance fault tolerance of multi-layer neural networks (MLN). This method is based on the idea that strong weights make MLN sensitive to faults. The purpose of new learning algorithm is to remove obstructions of fault tolerance from MLNs. We discuss about obstructions (strong connection and bias of each unit). As a result, we proposed new learning algorithm which is restricting the absolute value of weight and constructing MLNs without bias. We apply this algorithm to pattern recognition problems. It is shown that the fault tolerance of MLNs is improved.

Key words: Fault tolerance, Neural networks, Back Propagation algorithm

## 1 はじめに

従来、ニューラルネットワークは、その並列分散処理という形態のため、冗長な構造を形成することが容易であり、その結果、耐故障性を容易に獲得できるシステムであると考えられてきた。しかし、現在よく用いられる階層型ニューラルネットワーク (MLN: Multi Layer Neural networks) と、BP (Back Propagation) 学習法 [1] の組み合わせでは、高い耐故障性を示す MLN を獲得することが困難であり、高い耐故障性を示す MLN を獲得するためには、適切な学習法 (耐故障化学習法) を用いなければならないことが明らかになってきた。

そのため、これまでさまざまな耐故障化学習法が提案されてきた。第一には、学習時にノイズを付加する手法が挙げられる。これは MLN の学習時に、その学習パターン、MLN 内 (結合荷重など) にノイズを加え学習する手法 [2] である。他には、学習時の評価関数に耐故障性を表す項を付加する手法も挙げることができる。これは、学習目標である評価関数に、無故障時の MLN の出力誤差のみではなく、故障時の MLN の出力誤差も加え、これを最適化する手法 [3, 4] である。また、MLN の構造を適切に保ちながら学習する手法も挙げることができる。これは学習中に、MLN の耐故障化を妨げる大きい結合荷重を強制的に小さくすることで、耐故障性を高める手法 [5, 6] である。

しかし、これらの方法は一般的に、いずれの場合も付加的な計算が増えがちであり、もともと時間がかかっていた MLN の学習時間の増大につながりやすい。

<sup>†</sup>現、株式会社 DENSO

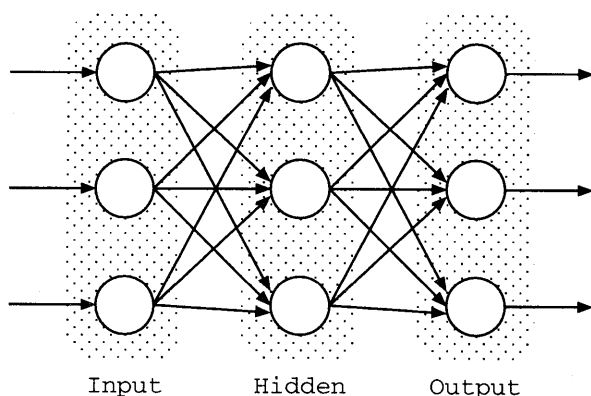


図 1: 階層型ニューラルネットワーク

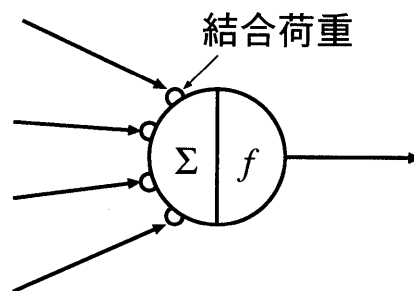


図 2: ユニット

そこで、本研究では、簡単な手法により MLN を耐故障化するために、特に結合荷重の値に着目することで、MLN の構造を適切に保ちながら学習する手法を提案する。[5, 6] でも、同様に結合荷重の値に着目しているが、この手法では、1 回の結合荷重更新のたびに、全結合荷重の分散等を計算するため付加的な計算量が増大していた。本論文では、そのような計算が不要であるような手法の提案を行う。

この手法による学習の結果、耐故障性を悪化させる要因が排除され、学習後の MLN の耐故障性の向上が期待できる。

## 2 階層型ニューラルネットワーク

### 2.1 構成

図 1, 2 に、それぞれ本論文で対象とする MLN および MLN を構成するユニットの構造をそれぞれ示す。MLN は、その基本構成単位であるユニットが入力層、中間層、出力層の 3 層に分かれ並ぶ構造を持つ。入力層、中間層の各ユニットの間と、中間層、出力層の各ユニットの間は、有向（左から右）の辺で結合しており、接続先のユニットとの接点に結合荷重と呼ばれる重みを持つ。MLN への入力信号は、入力層ユニットの入力として与え、出力層ユニットの出力を MLN の出力信号として得る。

入力層のユニットは、入力信号を中間層の各ユニットに分配するのみで値の変換は行わない。中間層、出力層の各ユニット  $i$  は、式 (1) に従い多入力 1 出力の素子として動作する。すなわち、各ユニットは、入力の重み付き和に対して非線形関数による変換を行う。

$$\begin{aligned} x_i &= \sum_j w_{ij} y_j + \theta_i \\ y_i &= f(x_i) = \frac{1}{1 + \exp(-x_i)} \end{aligned} \quad (1)$$

ただし、 $x_i$ ,  $y_i$ ,  $\theta_i$ ,  $w_{ij}$ ,  $f(x_i)$  を、それぞれユニット  $i$  の内部状態、ユニット  $i$  の出力、ユニット  $i$  のバイアス、ユニット  $j$  からユニット  $i$  への結合強さ（結合荷重）、入出力変換関数とする。

MLN は、十分な数の中間層ユニットを用意することで、任意の精度で任意の入出力パターン集合を学習できることが知られている [7]。しかし、冗長な数の中間層ユニットを用意しても、必ずしも耐故障性が高くないことも示されている [8]。

### 2.2 BP 学習法

MLN に、所望の入出力関係を実現させるための手法として BP（誤差逆伝搬）学習法が知られている。これは、与えられた入力パターン（教師入力パターン）と、それに対して所望する出力パターン（教師出力パ

ターン) に関して, 式 (2) で定義される誤差  $E$  を学習評価関数とし, これを減少させるように結合荷重, バイアスを調整することで行う.

$$E = \sum_p \sum_k \frac{1}{2} (y_k^p - Y_k^p)^2 \quad (2)$$

ただし,  $\sum_p$  は全教師パターンに対する総和,  $\sum_k$  は全出力層ユニットに対する総和であるとし,  $y_k^p, Y_k^p$  はそれぞれ, パターン  $p$  の教師入力パターンに対する MLN の出力層ユニット  $k$  の出力, および教師出力信号であるとする.

上記目的を達成するために, BP 学習法では, 式 (3) に従い, 所望の入出力関係を実現できるまで, すべての結合荷重, バイアスの更新を繰り返す.

$$\Delta w_{ij}(n) = -\eta \frac{\partial E}{\partial w_{ij}} + \alpha \Delta w_{ij}(n-1) \quad (3)$$

$$\frac{\partial E}{\partial w_{ij}} = \sum_p \delta_i^p y_j^p \quad (4)$$

$$\delta_i = \begin{cases} f'(x_i)(y_i^p - Y_i^p) & (\text{ユニット } i \text{ が出力層ユニット}) \\ f'(x_i) \sum_k \delta_k w_{ki} & (\text{上記以外}) \end{cases}$$

ただし,  $\Delta w_{ij}(n)$  は  $n$  回目の学習における結合荷重  $w_{ij}$  の更新量であり,  $f'(x_i)$  は  $x_i$  における  $f()$  の微係数であるとする. また,  $\alpha, \eta$  はパラメータでありそれぞれ, 慣性係数, 学習率と呼び,  $0 \leq \alpha < 1, 0 < \eta$  である.

ユニット  $i$  のバイアスの更新量  $\Delta \theta_i$  は, 定数 1 を出力するユニット  $b$  からユニット  $i$  への荷重更新量  $\Delta w_{ib}$  と等しい.

MLN においては, その学習と実際の利用は別に行われる. すなわち, 上記の手順に基づいて十分な回数の荷重の更新を行ったのち, MLN を実際に利用する.

### 3 階層型ニューラルネットワークの耐故障化

#### 3.1 故障モデル

MLN には, さまざまな故障可能箇所が存在するが, ここでは, 任意のユニット間の結合荷重の断線故障を対象とする. この故障モデルは, 断線した結合からの入力が入力が 0 になる場合を考えたとき, 故障した結合荷重が 0 となる故障と等価となる. 本論文でもそのように考える. これは, 他の MLN の耐故障性を扱った論文でも多く取り上げられているモデルである.

また, ユニット  $i$  のバイアスは 2.2 節でも述べたように, 定数 1 を出力するユニット  $b$  と, ユニット  $b$  からユニット  $i$  への結合 (結合荷重は  $w_{ib} = \theta_i$ ) により実現できる. このため, バイアスの故障は, バイアスに対応する結合の断線故障と等価と見なす. すなわち, バイアスは, その故障によりその値が 0 となるものとした.

本論文では, 学習終了後に, この結合の断線故障が発生した場合の MLN の耐故障性を扱う.

#### 3.2 MLN の故障とその影響

式 (2) から分かる通り, BP 学習法では, 教師信号と MLN の出力の差を小さくすることのみを目標としているため, 学習後の MLN が必ずしも高い耐故障性を示さない. そこで, MLN の耐故障性向上させる手法の提案の前に, 本章では MLN の耐故障性を悪化させる要因について考察する.

[5] でも述べられているように,  $|w_{ij}|$  の大きさが小さいほど, その MLN の耐故障性は高くなる. 実際, 結合荷重  $w_{ij}$  である結合が故障した場合のユニット  $i$  の出力変化  $\Delta y_i$  は, 式 (5) のようになる.

$$\Delta y_i = |f(x_i) - f(x_i - w_{ij} y_j)| \quad (5)$$

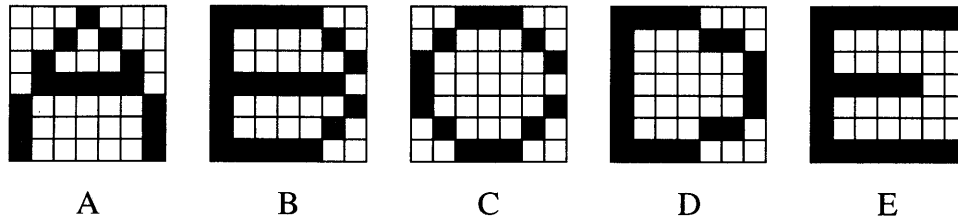


表 1: 教師入力パターンの例

式 (5) より,  $|w_{ij}y_j|$  の大きさが, 正常時と故障時の応答の差に大きな影響を与えることがわかる.

ここで,  $y_j$  がユニット  $i$  への入力信号であり MLN への入力パターンにより変化することを考えると,  $|w_{ij}|$  が小さいとき, どのような入力パターンに対しても  $\Delta y_i$  が小さくなる. すなわち,  $|w_{ij}|$  が大きな結合の故障は, それが小さな結合の故障と比べ, 故障によるユニット出力への影響が大きいといえる.

また,  $|w_{ij}y_j|$  の大きさを, 結合の故障, バイアスの故障がそれぞれ発生した場合について比べる. ユニット  $i$  への入力  $y_j$  は, 他ユニットからの入力の場合は  $0 < y_j < 1$  であり, バイアスの場合は  $y_j = 1$  と見なすことができる. ここで, 結合荷重, バイアスの値が同程度であれば,  $y_j$  の値より, バイアスの故障の方が結合の故障よりもユニット出力に大きな影響を与えるといえる.

すなわち, これらの要因により, 故障により MLN の出力へ大きな影響が生じ, MLN の耐故障性が低下するといえる.

### 3.3 耐故障化学習法

[5] では, 大きな結合荷重を持つ結合の故障が, ユニットの出力に大きな影響を与えていることに着目して, 学習中に結合荷重の平均, 標準偏差を求め, これらに基づいて大きいと判断された結合荷重の値を強制的に小さく更新している. この手法では結合荷重の更新を行うたびに, 全結合荷重の平均, 標準偏差を求めため, 計算量が増えがちになる.

そこで本論文では, 結合荷重の大小を判断する基準を結合荷重の更新を行うたびに, 動的に変化させるのではなく, BP 法による学習中に結合荷重の値の絶対値が一定値以下となるように制限する手法を提案する.

加えて, 結合の故障とバイアスの故障では, 後者の方がユニット出力に与える影響が大きいことから, 学習対象の MLN の構成ユニットよりバイアスを除くことも提案する.

以上の 2 手法により, 3.2 節で考察した, MLN の耐故障性を低下させる要因が学習後の MLN より排除され, 耐故障性が向上することが期待される.

## 4 数値実験

提案した学習法の有効性を確かめるため, 以下に示す文字認識の学習を行いその結果得られた MLN を評価する. この際, バイアスを持つユニットにより構成された MLN をオリジナルの BP 法により学習したもの (従来法), バイアスを持つユニットにより構成された MLN の結合荷重の絶対値を制限しながら学習したもの (バイアスあり提案法), バイアスを持たないユニットにより構成された MLN を結合荷重の絶対値を制限しながら学習したもの (バイアスなし提案法) の 3 種類の比較を行う.

評価は, (1) 各結合が故障した時の MLN 出力の変化, (2) MLN 内の故障が増加していった時の認識率の変化により行う.

対象とするのは, 図 1 のように,  $7 \times 7$  の格子上に白黒で描かれたアルファベット A~Z の認識である.

これを実現するため, 49 入力 26 出力の MLN を用意する. また, 中間層ユニット数は, 学習対象を十分に余裕をもって学習できるように 30 個用意した.

MLN への入力パターンは, 格子上の黒点を 1, 白点を 0 として各点を入力する. 出力パターンは, アル

ファベット A~Z にそれぞれ対応したユニットを用意し、最も出力の大きかったもユニットに対応するアルファベットを認識結果とする。

また、特に記さない限り、BP 学習法のパラメータは  $\eta = 0.05$ ,  $\alpha = 0.9$ , 初期荷重は  $[-1, 1]$  の一様乱数とした。提案法では、結合荷重の絶対値を 1 以下に制限し学習を行った。学習の終了条件は、いずれの文字に対する MLN の出力 2 乗誤差も 0.01 を下回ったときとした。

#### 4.1 各故障の MLN 出力への影響

従来法、バイアスあり提案法、バイアスなし提案法により学習した MLN について、各結合が故障した時の MLN 出力の変化を調べる。これにより、ひとつの故障が MLN の出力へ及ぼす影響を知ることができる。

図 3, 4, 5 に、それぞれ従来法、バイアスあり提案法、バイアスなし提案法で学習した MLN に対する結果を示す。横軸は故障した結合の結合荷重、バイアスの値を、縦軸は故障が発生する前と後での出力層の各ユニット出力の変化の絶対値の和（出力変動）を示す。また図中の各点は、ひとつの結合またはバイアスを表す。このグラフで、上方にある点ほど、その結合、バイアスの故障による MLN 出力への影響が大きいといえる。

これらの結果より、以下のことがわかる。

まず、3.2 節で考察したように、MLN の出力に大きな影響を与えるのは、絶対値の大きな結合荷重またはバイアスの故障であることがわかる。これは、いずれの図でも、上方 ( $\Delta y$  が大きい部分) に存在する点は、荷重の絶対値が比較的大きい部分に存在することにより示されている。

また、結合荷重の絶対値に制限を与えることで、MLN の出力に大きな影響を与える要因が除かれていることも確認できる。これは、図 3 における荷重の絶対値が 1 を超える部分が、図 4 で除かれた結果、従来法で 5 以上の出力変動をもたらしていた結合が除かれていることにより示されている。

さらに、図 4 において、結合、バイアスの各値でそれらの故障による出力変動の最大値を比べると、バイアスの故障によりより大きな出力変動が発生していることが確認できる。バイアスなし提案法により、バイアスを除き学習を行うことで、図 5 に示されているように、大きな出力変動をもたらしていたバイアスがなくなることで、出力変動の最大値がさらに小さくなった。

#### 4.2 故障率の増加に伴う認識率の変化

さまざまな故障数において、故障した結合荷重数の全結合荷重数に対する割合と、各故障数における MLN の認識率の関係をグラフに描き評価する。これは、耐故障性を扱った論文でよく用いられる評価法であり、グラフが上であるほど、また、グラフの傾きが小さいほど、故障箇所を増大に対する性能低下が少なく、耐故障性が高いといえる。

本実験では、教師パターン 1 個あたり、その入力パターンの 3 マスの白黒を反転したものを 20 種作成し、それらに対してそれぞれ、誤認識したパターンの割合を計算した。これを 20 種の初期荷重から同じ教師パターンを学習した MLN に対して計算し、図 6 に故障率と認識率の関係として示す。

この結果より、例えば認識率が 95% に低下した時点の故障率を比較すると、従来法は 1.2%、バイアスあり提案法は 3.6%、バイアスなし提案法は 4.8% となっている。また、グラフの傾きも、従来法、バイアスあり提案法、バイアスなし提案法の順に緩やかになっている。いずれも、提案法による耐故障性の向上を示している。

## 5 むすび

高い耐故障性を有する MLN を学習するために、耐故障性を悪化させる要因について考察した。その結果、絶対値の大きな結合荷重、バイアスの存在が MLN の耐故障性を悪化させる要因となることを明らかにした。

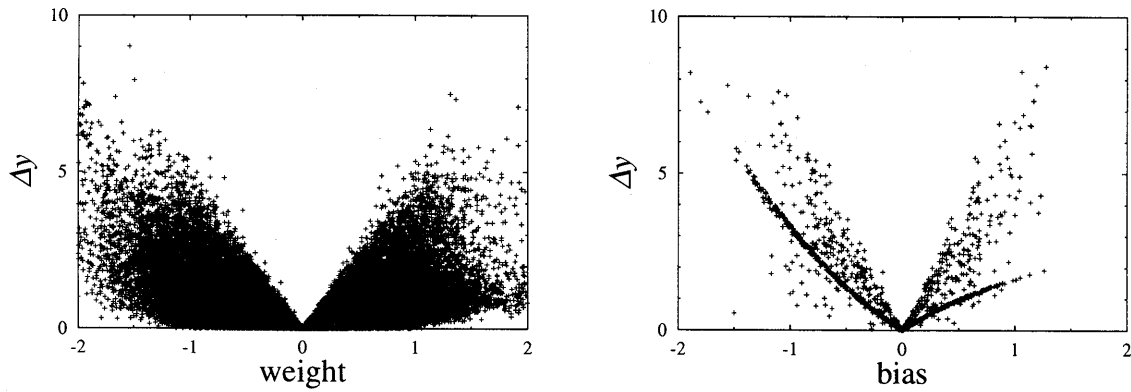


図 3: 故障による出力変動 (従来法)

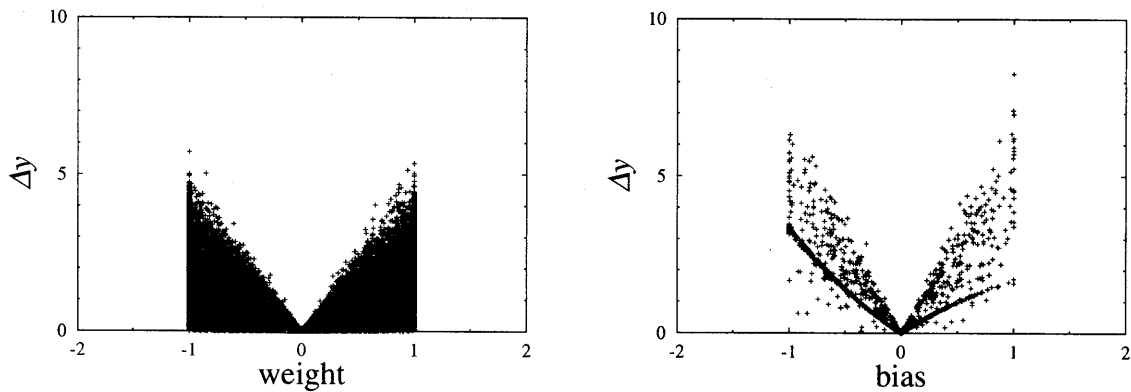


図 4: 故障による出力変動 (バイアスあり提案法)

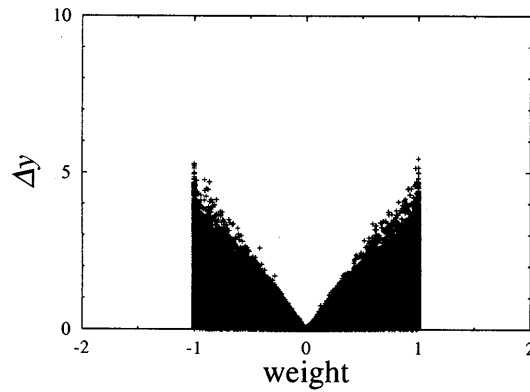


図 5: 故障による出力変動 (バイアスなし提案法)

このことをふまえて、これらの要因を排除してMLNの学習を行うために、BP法による学習時に結合荷重の絶対値を一定値以下となるように制限をすること、および、MLNの構成ユニットとしてバイアスを持たないユニットを用いることを提案した。これらにより、耐故障性を悪化させる要因が学習後のMLNより排除され、MLNの耐故障性が向上する。

以上を、文字認識を扱った数値実験により確認した。

## 参考文献

- [1] D.E.Rumelhart and J.L.McClelland "Parallel Distributed Processing", MIT Press (1986).

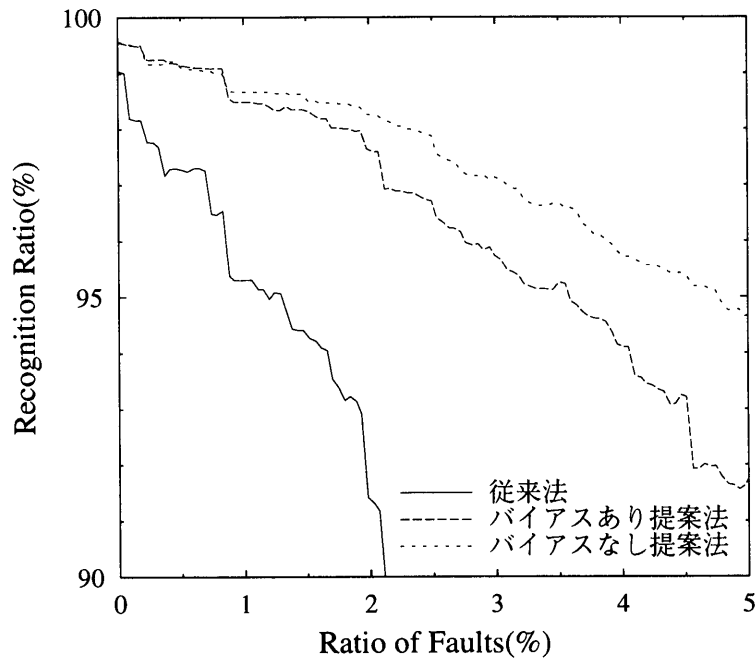


図 6: 故障率の増加に伴う誤認識率の変化

- [2] Alan F. Murray and Peter J. Edwards “Enhanced MLP Performance and Fault Tolerance Resulting from Synaptic Weight Noise During Training”, IEEE Transactions on neural networks, Vol.5, No.5, pp. 792–802 (1994).
- [3] 丹 康雄, 南谷 崇 “フォールトトレランスを有する階層型ニューラルネットワークとその性質”, 電子情報通信学会論文誌 D-I, Vol. J76-D-I, No.7, pp.380–389 (1993).
- [4] 高浪 五男, 楊 雲平 “故障注入による階層型ニューラルネットワークの重み故障に対する耐故障化”, 第 40 回 FTC 研究会 (1999).
- [5] 上浦 尚武, 畑 豊 “階層型ニューラルネットワークのフォールトトレラント学習アルゴリズムに関する一考察”, 第 39 回 FTC 研究会 (1998).
- [6] Nait Charif Hammadi and Hideo Ito “A Learning Algorithm for Fault Tolerant Feedforward Neural Networks”, IEICE Transaction Information and Systems, Vol. E80-D, No.1, pp.21–27 (1997).
- [7] 舟橋 賢一 “ニューラルネットワークの capability について”, 電子情報通信学会技術研究報告 MBE88-52 (1988).
- [8] Martin D. Emmerson and Robert I. Damper “Determining and Improving the Fault Tolerance of Multilayer Perceptrons in a Pattern Recognition Application”, IEEE Transactions on Neural Networks, Vol.4, No.5, pp. 788–793 (1993).