冗長な隠れ層・隠れユニットの削除による 階層型ニューラルネットのコンパクト構造化

正員 増田 達也 (大阪工大)
正員 佐藤 幸史 (シャープ)
正員 池谷 浩彦 (日本光電工業)
正員 藤井 善行 (三菱電機)

A Compact Structuring Method for Hierarchical Neural Networks by Eliminating Extra Hidden Layers and Units

Tatsuya Masuda, Member (Osaka Institute of Technology)

Koji Sato, Member (Sharp Corporation)

Hirohiko Ikeya, Member (Nihon Kohden Corporation)

Yoshiyuki Fujii, Member (Mitsubishi Electric Corporation)

When we apply a hierarchical neural network based on the back-propagation algorithm to a particular problem, we must determine beforehand the suitable size of network for the problem. But it is a very difficult problem. Too small a network will not learn at all, while too large a network will be inefficient and worsen its generalization ability due to overfitting.

In order to solve this problem, in this paper we propose a compact structuring method based on learning with a large size network and then compacting gradually the network by eliminating extra hidden layers and units. The result is a small and efficient network that performs better than the original. Also we demonstrate the effectiveness of this method by applying it to an identification problem of logic function.

キーワード：ニューラルネット，コンパクト構造化，バックプロパゲーション，隠れ層，隠れユニット

1. まえがき

D. E. Rumelhart らによって提案されたバックプロパゲーション学習（以後、BP学習と略す）に基づく階層型ニューラルネット⁽¹⁾は、様々な処理に適用できる汎用性の高いネットワークであるため、認識問題、同定問題、予測問題、制御問題などに数多く応用されている^{(2)・(3)・(4)}。この階層型ニューラルネットを実際問題に適用する際に問題となることは、対象の問題に適したサイズのネットワークをいかに用意するかということである。使用するネットワークのサイズが小さすぎると、隠れユニット数不足のために、十分な学習ができない。逆に大きすぎると、過剰自由度のためオーバーフィッティングを起こして個別パターンやノイズまで学習してしまい、ネットワークの汎化能力の低下を招く。したがって、良好な結果を得るためには、個々の学習問題に対して適切なサイズのネットワークを用いることが非常に重要となる。

最近、この問題を克服するために、学習段階でネットワークの構造を逐次変化させながら最終的に最適な構造のネットワークを得ようという試みがいくつかなされるように

なってきた⁽⁵⁾。筆者らもその1つの方法として、学習途中のネットワークの各隠れユニットの出力系列に着目して、それらの相関係数を指標に、同一層内の相関の高い隠れユニットを逐次合成していくコンパクト構造化手法を先に提案した⁽⁶⁾。そして、この方法によって同一層内での冗長な隠れユニット数の削減とネットワークの汎化能力の向上を達成できることを明らかにした。しかし、問題に適したサイズとは当然隠れ層の数までも含めた意味を持ち、先に提案した手法では冗長な隠れ層の削減は行うことができなかった。ネットワーク内に冗長な隠れ層が存在すると、冗長な隠れユニットと同様にネットワーク構造を不鮮明にし、さらには汎化能力の低下をも招く恐れがある。

この隠れ層の数に関しては、十分多くの隠れユニットを用意すれば1層の隠れ層で任意の連続関数（写像）が近似可能であるという報告がいくつかなされている^{(7)・(8)・(9)}。ただ、これらの報告は一般に写像を実現する3層ネットワークの存在を主張するものであり、3層ネットワークが学習的に、あるいは汎化能力的に最も優れていると主張しているわけではない。実際、問題の構造によっては、4層以上のネットワークを用いることにより全体での必要な隠れ

ユニット数がずっと少なくなることも十分あり得る。すなわち、対象とする問題によって隠れユニット数同様、最適な隠れ層数も異なってくる。

そこで、本論文では先に提案した手法をさらに拡張して、隠れユニットの出力系列の相関係数および分散を指標に冗長な隠れユニットだけでなく、冗長な隠れ層までも削除する新たなコンパクト構造化手法を提案する。この手法によって、ネットワーク内には最終的に必要な数の隠れユニットおよび隠れ層だけが残る、真の意味でネットワークのコンパクト化が実現できる。さらに、本論文ではこの手法を論理関数の同定問題に適用することにより、その有効性を明らかにする。

2. 冗長な隠れ層の削除

〈2・1〉 判定条件 階層型ニューラルネットの情報伝達機構は、各層ごとで行われる情報変換の繰り返しであり、その変換に非線形性を持たせることによって様々な入出力関係が実現できる。ただ、その変換形態がBP学習によって獲得されるため、必要以上に隠れ層を用意した場合、情報伝達的に同じ働きをする冗長な隠れ層がネットワーク内に出現することになる。たとえば、図1において、矢印で対応付けられたユニット対が各入力パターンに対して情報伝達的に同じ働きをしている場合、J層からI層にかけての変換は無意味であり、J層あるいはI層のいずれか一方は冗長な隠れ層といえる。

我々は先に提案した手法⁽⁶⁾の中で、情報伝達的に同じ働きをしている同一層内の2つの隠れユニット*i*、*j*を、*n*個の訓練用データに対するこれらの隠れユニットの出力系列 $\{V_{ip} | p=1, \dots, n\}$ 、 $\{V_{jp} | p=1, \dots, n\}$ の相関係数、

$$r_{ij} = \frac{\frac{1}{n} \sum_{p=1}^n V_{ip} V_{jp} - \overline{V_i} \overline{V_j}}{\sqrt{\frac{1}{n} \sum_{p=1}^n V_{ip}^2 - \overline{V_i}^2} \sqrt{\frac{1}{n} \sum_{p=1}^n V_{jp}^2 - \overline{V_j}^2}} \quad (-1 \leq r_{ij} \leq 1) \quad \dots\dots\dots(1)$$

図1 冗長な隠れ層

Fig.1. An example of extra hidden layer

および、それらの分散 s_i^2 、 s_j^2 を用いて、つぎのように判別した。

$$|r_{ij}| \geq \theta_1 \quad \text{かつ} \quad s_i^2, s_j^2 \geq \theta_2 \quad \dots\dots(2)$$

すなわち、(2)式を満たすとき、隠れユニット*i*、*j*は同じ働きをしているものとみなして1つに合成した。この相関係数および分散を用いると、上述の冗長な隠れ層の判定も容易に行うことができる。以下にその判定条件を示す。

条件1：互いに隣接し合う隠れ層I層（I層は出力層でもよい）とJ層に属するユニット数が同数である。

条件2：各訓練パターンに対してI層のあるユニット*i*と情報伝達的に同じ働きをするユニット*j*が必ずJ層に1個存在する。すなわち、(2)式を満たす相関の高いユニット対(*i*, *j*)が存在する。そして、このようなユニット対がI層およびJ層のすべてのユニットに対して、図1に示すように1対1の対応関係で存在する。

条件1は、同じ働きをしている隠れ層とみなすための必要不可欠な条件である。つまり、ユニット数が異なる2つの隠れ層は、情報伝達的に同じ働きをしているとはいえない。一方、条件2は相関の高いユニットが重複することなく1対1の対応関係で存在していることを示している。

本論文では、上記2つの条件を満たしているときに限り、I層とJ層は情報伝達的に同じ働きをしているものとみなす。そして、これら2つの隠れ層は出力層側のI層で代表させ、J層は冗長な隠れ層とみなして削除する。これによって、最終的に残る隠れ層は情報伝達的に互いに働きの異なる隠れ層となる。

〈2・2〉 削除方法 具体的な隠れ層の削除方法を以下に述べる。冗長な隠れ層の削除は、その層に属する各ユニットの働きを隣接する上層のユニットに代表させることによって行う。図2(a)は、階層型ニューラルネットの隣接する任意の3層（I、J、K層）を表したもので、K層が入力層側、I層は出力層側である。もちろんI層を出力層とみなしてもよい。いま、同図において、I層とJ層が〈2・1〉の条件を満たし、J層が冗長な隠れ層と判定されたとする。このとき、出力系列の相関が高い任意のユニット対(*i*, *j*)に関しては、ユニット*i*、*j*の入出力関係がシグモイド関数等の単調増加関数で表現されているため、当然これらのユニットへの入力系列 $\{U_{ip} | p=1, \dots, n\}$ 、 $\{U_{jp} | p=1, \dots, n\}$ の相関も高いと考えられる。つまり、 U_i 、 U_j の関係を、次の線形回帰式で近似的に表現することができる⁽¹⁰⁾。

$$U_i = a U_j + b \quad \dots\dots\dots(3)$$

ただし、

$$a = \frac{\frac{1}{n} \sum_{p=1}^n U_{ip} U_{jp} - \overline{U_i} \overline{U_j}}{\frac{1}{n} \sum_{p=1}^n U_{jp}^2 - \overline{U_j}^2}$$

$$b = \overline{U_i} - a \overline{U_j}$$

図 2 隠れ層の削除
Fig.2. Eliminating method of extra hidden layer

ここで、J 層のユニット j への入力 U_j は

$$U_j = \sum_{k=1}^m w_{jk} V_k + w_{jb} \quad \dots\dots\dots (4)$$

であるから、この関係を(3)式に代入すると、

$$\begin{aligned} U_i &= a U_j + b \\ &= a \left(\sum_{k=1}^m w_{jk} V_k + w_{jb} \right) + b \\ &= \sum_{k=1}^m a w_{jk} V_k + (a w_{jb} + b) \cdot 1 \quad \dots\dots\dots (5) \end{aligned}$$

となる。すなわち、図 2(b)に示すように K 層のすべてのユニットから I 層のユニット i へ新たにリンクを接続して、そのリンク荷重 w_{ik} 、 $k=1, \dots, m$ およびバイアスユニットからのリンク荷重 w_{ib} をそれぞれ

$$\begin{aligned} w_{ik} &= a w_{jk} \quad \dots\dots\dots (6) \\ w_{ib} &= a w_{jb} + b \end{aligned}$$

とすることによって、J 層のユニット j を除去することができる。そして、この操作を相関の高いすべてのユニットに対して施すことにより、結果的に J 層全体が削除されることになる。

3. アルゴリズム

2 章で述べた冗長な隠れ層の削除方法を、先に提案した

隠れユニットの合成手法⁽⁶⁾に組み込むと、そのアルゴリズムはつぎようになる。

- ① 初期化：階層型ニューラルネットの初期構造としては、隠れ層数、隠れユニット数とも大きめのサイズのものを用意する。一様乱数によるネットワークのリンク荷重の初期化および学習係数 η 、モーメント係数 α 、隠れユニットの合成および隠れ層の削除時期のしきい値 ε_1 、学習終了のしきい値 ε_2 、合成のしきい値 θ_1 、 θ_2 の設定を行う。学習回数を $k = 1$ とする。
- ② BP 学習： n 個の訓練用データに対する隠れユニットの出力系列 $\{V_{ip}\}$ およびネットワークの出力二乗誤差 $E = \frac{1}{2} \sum_p \sum_k (O_{kp} - T_{kp})^2$ を計算する。続いて、一般化デルタルールにより、ネットワークのすべてのリンク荷重の修正を行う。
- ③ 隠れユニットの合成、隠れ層の削除および学習終了時期の判定：ネットワークの出力二乗誤差 E が隠れユニットの合成および隠れ層の削除時期の条件 $E \leq \varepsilon_1$ を満たしていなければ、 $k = k + 1$ として②に戻る。条件を満たして、 $\varepsilon_1 \geq E > \varepsilon_2$ ならば、④へ進む。学習終了条件 $E \leq \varepsilon_2$ を満たしていれば、学習を終了する。
- ④ 分散、相関係数の計算：各隠れ層の隠れユニットの出力系列の分散 s_i^2 および相関係数 r_{ij} を計算して、⑤へすすむ。
- ⑤ 同一層内の隠れユニットの合成：(2)式の合成条件を満たす、同一隠れ層内のすべての隠れユニット対 (i, j) を文献(6)の合成方法を用いて合成し、 $k = k + 1$ として②にもどる。合成条件を満たす隠れユニット対が同一隠れ層内に存在しない場合は、そのまま⑥へ進む。
- ⑥ 冗長な隠れ層の削除：2・1で示した判定条件を満たす冗長な隠れ層が存在すれば、2・2の方法で削除して、 $k = k + 1$ として②にもどる。冗長な隠れ層が存在しない場合は、そのまま $k = k + 1$ として②にもどる。

4. 適用例

本手法を簡単な論理関数の同定問題に適用してみる。学習させる論理関数は次式の 6 入力 2 出力の論理関数である。

$$\begin{aligned} g(a, b, c, d) &= (a \vee b) \wedge (c \vee d) \\ h(c, d, e, f) &= (c \vee d) \wedge (e \vee f) \quad \dots\dots\dots (7) \end{aligned}$$

ここで $\vee$ 、 $\wedge$ はそれぞれ論理和および論理積である。(7)式は 6 入力の関数であるから、入出力パターンは 2^6 通りある。本実験ではこのうちランダムに選びだした 48 個を訓練用に、残りの 16 個を判別用に用いた。また、ネットワークの初期構造としては、図 3 に示すように 1 つの入力層 (6 ユニット)、3 つの隠れ層 (各 6 ユニット)、1 つの出力層 (2 ユニット) をもつ 6-6-6-2 構造のものを用意した。実験条件は以下の通りである。

図3 あらかじめ用意するネットワーク構造
Fig.3. Initial structure of network for identifying the logic function given by Eq.(7)

図4 本手法を用いた場合の学習曲線
Fig.4. Learning curve of network

学習定数: $\eta = 1.0$
モーメント係数: $\alpha = 0.9$
隠れユニットの合成および隠れ層の削除時期のしきい値: $\varepsilon_1 = 0.001$
学習終了のしきい値: $\varepsilon_2 = 0.0001$
相関係数のしきい値: $\theta_1 = 0.8$
分散のしきい値: $\theta_2 = 0.01$

なお、上記 θ_1 , θ_2 の設定値の妥当性については文献(6)で詳しく検討されている。
また、ネットワークのリンク荷重の初期値は区間 $[-0.1, 0.1]$ の一様乱数で割り当てた。
まず、本手法を適用した場合の学習状況について述べる。

表1 隠れユニット合成時の相関係数と分散
Table 1. Correlation coefficients r_{ij} and variance s_i^2 of hidden units

	$i \backslash j$	相関係数 r_{ij}						分散 s_i^2
		①	②	③	④	⑤	⑥	
第1隠れ層	①	1.00	-0.09	-0.02	-0.15	0.84	0.13	0.135
	②	-0.09	1.00	0.97	0.22	-0.21	0.75	0.124
	③	-0.02	0.97	1.00	0.09	-0.17	0.68	0.093
	④	-0.15	0.22	0.09	1.00	-0.03	0.46	0.148
	⑤	0.84	-0.21	-0.17	-0.03	1.00	0.07	0.081
	⑥	0.13	0.75	0.68	0.46	0.07	1.00	0.131
第2隠れ層	①	1.00	0.29	0.67	0.50	0.69	0.22	0.144
	②	0.29	1.00	0.84	0.91	0.83	0.99	0.116
	③	0.67	0.84	1.00	0.98	1.00	0.77	0.081
	④	0.50	0.91	0.98	1.00	0.97	0.86	0.084
	⑤	0.69	0.83	1.00	0.97	1.00	0.76	0.082
	⑥	0.22	0.99	0.77	0.86	0.76	1.00	0.131
第3隠れ層	①	1.00	1.00	0.79	1.00	0.54	0.99	0.126
	②	1.00	1.00	0.77	1.00	0.52	0.99	0.127
	③	0.79	0.77	1.00	0.76	0.94	0.48	0.072
	④	1.00	1.00	0.76	1.00	0.51	0.98	0.136
	⑤	0.54	0.52	0.94	0.51	1.00	0.61	0.177
	⑥	0.99	0.99	0.48	0.98	0.61	1.00	0.103

図4は本手法を適用したときの学習曲線である。本実験の場合、学習開始後445回目でネットワークの出力二乗誤差 E が合成時期の条件 $E \leq \varepsilon_1$ を満たし、各層の隠れユニット同士の合成が行われた。表1に合成時の第1、第2、第3隠れ層に属する隠れユニットの出力系列の相関係数、分散を示す。同表より、第1隠れ層においてはユニット①と⑤およびユニット②と③の相関が高く、これらがそれぞれ1つに合成されることがわかる。同様に第2隠れ層においてはユニット②、③、④、⑤がユニット⑥に、第3隠れ層ではユニット①、②、⑥がユニット④に、また、ユニット③がユニット⑤に合成される。この結果、ネットワーク構造は初期構造である6-6-6-6-2構造から図5(a)に示す6-4-2-2-2構造までコンパクト化された。

この合成によって、ネットワークの出力二乗誤差 E は一旦上昇するが、その後速やかに減少して学習回数447回目に第3隠れ層が冗長な隠れ層と判定されて削除された。このときの状況を表2と図5(b)に示す。先の合成によって、第3隠れ層のユニット数はすでに2つになっており、出力層のユニット数と同数となるため、 $\langle 2 \cdot 1 \rangle$ の条件1は満たされている。表2は削除される直前の第3隠れ層のユニットと出力ユニットの相関係数を示したものである。同表より、出力層のユニット①と第3隠れ層のユニット⑤、出力層のユニット②と第3隠れ層のユニット④の相関がそれぞれ高くなっており、1対1の対応関係を持つ相関の高いユニット対がユニット数と同数だけあることがわかる。すなわち、 $\langle 2 \cdot 1 \rangle$ の条件2も満たされる。したがって、第3隠れ層は出力層と同じ働きをする冗長な隠れ層と判定され、前節の削除方法によって図5(b)のように削除される。

図5 ネットワークのコンパクト構造化状況
Fig.5. Change process of network structure in the identification problem of logic function

表2 出力層と第3隠れ層間の相関係数
Table 2. Correlation coefficients just before the elimination of hidden layer 3

		第3隠れ層	
出力層	$i \setminus j$	④	⑤
	①	0.49	1.00
	②	0.98	0.39

表3 出力層と第2隠れ層間の相関係数
Table 3. Correlation coefficients just before the elimination of hidden layer 2

		第2隠れ層	
出力層	$i \setminus j$	①	⑥
	①	-0.25	-0.97
	②	-0.96	-0.61

Fig.6. Final structure of network obtained by the proposed method

表4 各ユニットの出力状況
Table 4. Output value of each unit for 12 input-output patterns

入力層						第1隠れ層			出力層		教師信号	
a	b	c	d	e	f	①	②	④	①	②	①	②
1	1	0	0	1	1	0.00	0.97	0.00	0.02	0.02	0	0
1	0	0	1	0	1	0.05	0.02	0.04	0.99	0.99	1	1
0	1	0	0	1	0	0.01	0.94	0.07	0.01	0.00	0	0
0	0	0	0	1	0	0.03	0.91	0.98	0.00	0.02	0	0
0	0	1	1	0	1	0.04	0.00	0.94	0.02	0.99	0	1
1	0	1	0	0	0	0.88	0.01	0.03	0.99	0.01	1	0
0	0	0	1	0	0	0.83	0.01	0.96	0.01	0.01	0	0
1	0	0	0	0	0	0.86	0.89	0.05	0.01	0.00	0	0
1	0	1	1	1	1	0.00	0.00	0.03	0.99	1.00	1	1
1	1	0	0	1	0	0.01	0.95	0.00	0.02	0.02	0	0
1	1	0	0	0	1	0.02	0.95	0.00	0.02	0.02	0	0
1	0	0	1	0	0	0.87	0.01	0.00	0.99	0.01	1	0

その後、学習回数453回目に第2隠れ層も冗長な隠れ層と判断され、2度目の隠れ層の削除が行われた。この状況を表3と図5(c)に示す。さらに、472回目に第1隠れ層内で2度目の隠れユニット同士の合成が行われ、図5(d)に示すように隠れユニット⑥が隠れユニット②に合成された。そして、学習回数775回目に学習終了条件 $E \leq \varepsilon_2$ が満たされて、ネットワークの学習が終了した。図6は本アルゴリズムによって最終的に得られたネットワーク構造を図示したものである。同図において、リンク上の四角で囲んだ数値は、上下のユニット間のリンク荷重値を、ユニットの横の鍵かっこ内の数値はバイアスユニットからのリンク荷重値をそれぞれ表している。図3のネットワークと比較してかなりコンパクトな6-3-2構造のネットワークが得られていることが同図よりわかる。

つぎに、図6のネットワークの各ユニットが情報伝達的にそれぞれどのような役割を演じているかについて検討してみる。表4は、学習に用いた入出力データのうち、典型的な12個のパターンに対する隠れ層および出力層に属する各ユニットの出力値を示したものである。ここで、上段の4つの入出力データは、入力 a, b の組み合わせにだけ注目して無作為に抽出したものであり、中段の4つは入力 c, d の組み合わせに、下段の4つは e, f の組み合わせに注目したもので、それぞれを破線で区切っている。まず、図6のリンク荷重値から、隠れ層のユニット④は、入力ユニット a, b から大きな影響を受けていることがわかる。そこで、表4の上段の4つの入力パターンに対する隠れ層のユニット④の出力値を見ると、 a, b が共に0のときにのみ発火していることがわかる。ゆえに、隠れ層のユニット④は入力 a, b から負のリンク荷重を媒介にしていることを考慮すると、入力 a, b の否定の論理積 $\overline{a} \wedge \overline{b}$ を計算していることになる。同様に考えると、隠れ層のユニット②は入力 c, d の否定の論理積 $\overline{c} \wedge \overline{d}$ を、ユニット①は入力 e, f の否定の論理積 $\overline{e} \wedge \overline{f}$ を計算していることになる。一方、出力層のユニット①は、図6より隠れ層のユニット②と④から大きい影響を受けていることがわかる。表4において、隠れ層のユニット②と④の出力値と出力層のユニット①の出力値を見ると、隠れ層のユニット②と④が共に0のときにのみ出力層のユニット①が発火している。したがって、出力層のユニット①は、負のリンク荷重を媒介にして隠れ層のユニット②と④の否定の論理積、

$$(\overline{a} \wedge \overline{b}) \wedge (\overline{c} \wedge \overline{d}) = (a \vee b) \wedge (c \vee d) \quad \dots\dots\dots(8)$$

すなわち、(7)式の出力 g を計算していることがわかる。同様に、出力層のユニット②は、隠れ層のユニット①と②からの負のリンク荷重を媒介にして、

$$(\overline{c} \wedge \overline{d}) \wedge (\overline{e} \wedge \overline{f}) = (c \vee d) \wedge (e \vee f) \quad \dots\dots\dots(9)$$

すなわち、(7)式の出力 h を計算している。このように、

本アルゴリズムで得られたネットワークは、ド・モルガンの法則を暗に成立させながら正しい出力を出す、非常に興味深い構造になっていることがわかる。

最後に、本アルゴリズムを用いて得られた図6のネットワーク(6-3-2構造)と、図3の初期構造から始めて隠れユニットの合成のみを施して得られたネットワーク(6-3-2-2-2構造)、ならびに図3の構造のままB P学習のみを行って得られたネットワーク(6-6-6-6-2構造)の判別能力を比較してみる。このため、64個の入出力パターンのうち、学習に使用しなかった16個の判別用パターンに対する各ネットワークの出力二乗誤差 E_1, E_2, E_3 を求めてみた。その結果は、それぞれ

$$E_1 = 0.804 \times 10^{-4}$$

$$E_2 = 0.823 \times 10^{-4}$$

$$E_3 = 1.984 \times 10^{-4}$$

となり、何も施さないネットワークの出力二乗誤差 E_3 より隠れユニットの合成のみを施したネットワークの出力二乗誤差 E_2 の方が、また E_2 より本手法によって隠れユニットの合成および冗長な隠れ層の削除を行ったネットワークの出力二乗誤差 E_1 の方がより小さい値になっている。すなわち、本手法で得られた図6のネットワークが最も優れた判別能力を有していると言える。このことから、冗長な隠れユニットや冗長な隠れ層の存在がネットワークの汎化能力に悪影響を与えることがわかる。

5. あとがき

本論文では、先に提案した隠れユニットの合成による階層型ニューラルネットのコンパクト構造化手法をさらに拡張して、冗長な隠れ層の削除も同時に行うことができる新たな構造化手法を提案した。さらに本手法を論理関数の同定問題に適用して、隠れユニットの数だけでなく、隠れ層の層数的にも冗長性の少ないコンパクトな構造のネットワークが容易に得られること、また本手法によって得られたネットワークが汎化能力的にも優れていることを明らかにした。

今後は本手法をより現実的な問題に適用して、その有効性を実証していく予定である。

(平成6年3月2日受付)

文 献

- (1) D. E. Rumelhart・J. L. McClelland and the PDP Research Group: "Parallel Distributed Processing", Vol. 1~2, The MIT Press (1986)
- (2) 中野 馨・飯沼 一元・ニューロンネットグループ・桐谷 滋: 「ニューロコンピュータ」, 技術評論社 (1989)
- (3) 杉江 昇: 「ニューロコンピュータ part II」, コンピュータローラ, 29, 1~128 (1990)
- (4) システム制御情報学会: 「ニューラルネットワークの応用総合特集号」, システム/制御/情報, 35-1, 2~58

(1991)

- (5)石川 眞澄:「ネットワーク学習アルゴリズムの最近の話題」, 計測と制御, 30-4, 285~290 (1991)
- (6)増田 達也・藤井 善行:「隠れユニットの合成による階層型ニューラルネットワークのコンパクト構造化」, 計測自動制御学会論文集, 28-4, 519~527 (1992)
- (7)G. Cybenko: "Continuous Valued Neural Networks with Two-layers are Sufficient", Tech. Report, Tuft Univ. (1988)
- (8)B. Irie and S. Miyake: "Capabilities of Three-layered Perceptrons", Proc. of IEEE International Conference on Neural Networks, Vol. I, 641~648, San Diego (1988)
- (9)R. Hecht-Nielsen: "Theory of the Backpropagation Neural Network", Proc. of International Joint Conference on Neural Networks, Vol. I, 593~605, Washington D.C. (1989)
- (10)奥野 忠一・久米 均・芳賀 敏郎・吉澤 正:「多変量解析法」, 日科技連 (1971)

増田 達也 (正員) 昭和25年8月2日生。昭和56年3月大阪大学大学院博士課程単位修得退学

(電気工学専攻)。日本学術振興会奨励研究員を経て, 昭和60年4月大阪大学工学部電気工学科助手。昭和62年4月大阪工業大学工学部電気工学科講師, 現在同教授。工学博士。ニューラルネット, ファジィ制御, 遺伝的アルゴリズム, 多目的意志決定, シ

ステム構造分析などの研究に従事。電子情報通信学会, 計測自動制御学会, システム制御情報学会, 日本ファジィ学会, 日本オペレーションズ・リサーチ学会, IEEE会員。

佐藤 幸史 (正員) 昭和44年4月24日生。平成6年3月大阪工業大学大学院修士課程(電気工学専攻)修了。同年4月シャープ(株)に入社, 現在に至る。在学中, 遺伝的アルゴリズムの理論および応用に関する研究に従事。電子情報通信学会, 計測自動制御学会会員。

池谷 浩彦 (正員) 昭和44年1月11日生。平成5年3月大阪工業大学大学院修士課程(電気工学専攻)修了。同年4月日本光電工業(株)に入社, 現在に至る。在学中, ニューラルネットの理論および応用に関する研究に従事。計測自動制御学会, システム制御情報学会会員。

藤井 善行 (正員) 昭和42年2月27日生。平成4年3月大阪工業大学大学院修士課程(電気工学専攻)修了。同年4月三菱電機(株)に入社, 現在に至る。在学中, ニューラルネットの理論および応用に関する研究に従事。計測自動制御学会, システム制御情報学会会員。