FIR 形ニューラルネットワークのシステム同定への応用

正 員 降 矢 順 治 (城西国際大学)

正 員 西 正 明 (信 州 大 学)

正 員 松 坂 知 行 (八戸工業大学)

Application of FIR-type Neural Network to System Identifications

Junji Furuya, Member (Jyousai International University), Masaaki Nishi, Member (Shinshu University), Tomoyuki Matsuzaka, Member (Hachinohe Institute of Technology)

This paper presents a FIR-type neural network to identify nonlinear systems as a nonlinear identification model. We derive a learning algorithm of this neural network based on the gradient descent method and the conjugate gradient method. In this learning algorithm, we derive the algorithm which updates both weight coefficients contained in FIR-synapses and gains of sigmoid functions in each unit of each layer. Also, we present a generalized sigmoid function as the nonlinear function contained in each unit in order to designate the upper and lower limits of the function. It is shown that the accuracy of the identification model based on FIR-type neural networks is superior to that of the identification model based on ordinary backpropagation neural networks and the conjugate gradient algorithm is superior to the gradient descent algorithm in point of the convergence and the learning cycle.

キーワード：FIR 形ニューラルネットワーク，システム同定，最急降下法，共役勾配法

1. ま え が き

システム同定は自動制御，通信システム，及び地震信号フィルタリング⁽¹⁾のような多くの信号処理の分野において重要な問題である⁽²⁾。システム同定は主としてそのシステムの入力と出力との写像関係を決定することに関心がもたれてきた。過去数十年の間に，未知のパラメータをもつ線形時不変システムの同定には大きな前進が見られ，どのような同定器を選択すべきは，離散時間確率モデル，状態空間モデルなど線形システム理論に基づいて確立されてきた⁽³⁾。

それに対して，非線形システムを同定するために必要とされる非線形同定モデルは，システム（またはプラント）の出力と同定モデルとの間の誤差関数を最小にするに当たって多くの極小値をもち，線形同定モデルのように大域的な収束を得ることが困難になっている。このような問題があるにもかかわらず，非線形ダイナミックシステムの同定にニューラルネットワーク⁽⁴⁾やファジィとニューラルネットワークの融合システム⁽⁵⁾を用いて研究されてきた。これまで，非線形システムの同定に用いられたニューラルネットワークには，多層ニューラルネットワーク，リカレントニューラルネットワーク，RBFニューラルネットワークがある⁽²⁾。K.S.Narendraらは，非線形ダイナミックシステムの同定に多層ニューラルネットワークを取り上げ，最急降下法によ

るバックプロパゲーションを用いて同定を行っている^{(6),(7)}。

本論文では，非線形システムの同定を行うに当たって，FIR 形ニューラルネットワークを用いることを提案している。この FIR 形ニューラルネットワークは，バックプロパゲーションに基づいた通常の階層形ニューラルネットワーク（以下，BP ニューラルネットワークとする）の重み（ニューロンのシナプスに相当する）を FIR 形シナプスに置き換えた階層形ニューラルネットワークである。FIR はフィルタ理論にみられる Finite Impulse Response（有限インパルス応答）を略したものであり，その構造はフィルタの構造と同じである。この FIR 形ニューラルネットワークを取り上げた目的は，非線形システム同定において，FIR 形ニューラルネットワークが文献(6)においてシステム同定に用いられた通常の BP ニューラルネットワークと比較して，同定精度の点でどの程度改善されるかを明らかにすることにある。

FIR 形ニューラルネットワークに関しては，E.A.Wan⁽⁸⁾もこれを取り上げ，学習アルゴリズムをベクトルの導き，化学反応の時系列予測に適用しているが，本論文では各層のユニット（ニューロン）に着目して，誤差逆伝播法に基づいた別個の学習アルゴリズムを導いている。この学習アルゴリズムを最急降下法と共役勾配法との 2 種類の方法によって導き，共役勾配法による学習（トレーニング）が，収

束と学習回数の点で最急降下法による学習よりも改善されるかを検討している。本論文では、各層のユニットに含まれている非線形関数として一般化シグモイド関数を提案し、このシグモイド関数の上下限値を任意の値に指定できるようにしている。また、学習アルゴリズムにおいて、各層のFIR形シナプスに含まれる重みを更新するだけでなく、ユニットのシグモイド関数のゲイン（傾き）も更新できるアルゴリズムを導いている。

2. FIR 形ニューラルネットワークの構成

従来の標準形の階層形ニューラルネットワーク (BP ニューラルネットワーク) の1個の重み (シナプスに相当) の代わりに、図1に示すように、 L 個の遅延素子と $L+1$ 個の重み係数 $w_{kj0}, w_{kj1}, \dots, w_{kjL}$ から成る FIR 形シナプスを用

図1 FIR形シナプス
Fig. 1. FIR-Synapse.

図2 図1の簡略形
Fig. 2. A simplified form.

いる。FIR形シナプスの入力を $z_j(t)$ とすると、出力 $x_k(t)$ は次のようになる。

$$x_k(t) = \sum_{p=0}^L w_{kjp} z_j(t-p) \dots \dots \dots (1)$$

$L+1$ 個の重み係数は、3. で示す学習アルゴリズムにしたがって、学習時に更新されるものである。図1のFIR形シナプスを図2に示すような簡略形を用いて表わすことにする。従来の標準形のBPニューラルネットワークの重みをFIR形シナプスで置換した階層形ニューラルネットワークを考えると、図3に示すようなネットワークの構成となる。このニューラルネットワークをFIR形ニューラルネットワークと呼ぶことにする。図3のニューラルネットワークは、入力層 ($l-3$ 層)、第1中間層 ($l-2$ 層)、第2中間層 ($l-1$ 層) と出力層 (l 層) の4層から成るFIR形ニューラルネットワークである。

入力層のユニット数を N_h 、第1中間層のユニット数を N_i 、第2中間層のユニット数を N_j 、出力層のユニット数を N_k としている。入力層の h 番目のユニット、第1中間層の i 番目のユニット、第2中間層の j 番目のユニット、出力層の k 番目のユニットに着目して、以下に示すように誤差逆伝播法に基づいて学習アルゴリズムを導くことにする。

図3 FIR形4層ニューラルネットワークの構成
Fig. 3. A constitution of four layer FIR-type neural networks.

3. 学習アルゴリズム

〈3・1〉 順方向の計算

①第1中間層の出力 $z_i^{l-2}(t)$ 第1中間層 ($l-2$ 層) の i 番目の内部ポテンシャル $x_i^{l-2}(t)$ と出力 $z_i^{l-2}(t)$ は、入力層 ($l-3$ 層) の h 番目の出力を $z_h^{l-3}(t)$ とすると、

$$x_i^{l-2}(t) = w_{0i}^{l-2} z_{0i}^{l-3}(t) + \sum_{h=1}^{N_h} \sum_{p=0}^L w_{hip}^{l-2} z_h^{l-3}(t-p) \quad \dots\dots\dots (2)$$

$$z_i^{l-2}(t) = f(x_i^{l-2}(t)), \quad i = 1, 2, \dots, N_i \quad \dots\dots\dots (3)$$

となる。

$z_{0i}^{l-3}(t)$ はオフセットであり、通常は1にセットされる。本論文でも全てのオフセットを1にセットしている。 $f(x_i^{l-2}(t))$ は、次の式で表わされる一般化シグモイド関数である。

$$f(x_i^{l-2}(t)) = a + \frac{b-a}{1 + \exp(-g_i^{l-2} x_i^{l-2}(t))} \quad \dots\dots\dots (4)$$

a と b はこの関数の下限値と上限値であり、任意に指定することができる。 g_i^{l-2} はゲインであり、この関数の特性は図4のようになる。以下、このシグモイド関数を用いることにする。

②第2中間層の出力 $z_j^{l-1}(t)$ 第2中間層 ($l-1$ 層) の j 番目の内部ポテンシャル $x_j^{l-1}(t)$ と出力 $z_j^{l-1}(t)$ は、第1中間層の i 番目の出力 $z_i^{l-2}(t)$ を用いると次式となる。 $z_{0j}^{l-2}(t)$ はオフセットである。

$$x_j^{l-1}(t) = w_{0j}^{l-1} z_{0j}^{l-2}(t) + \sum_{i=1}^{N_i} \sum_{p=0}^L w_{ijp}^{l-1} z_i^{l-2}(t-p) \quad \dots\dots\dots (5)$$

$$z_j^{l-1}(t) = f(x_j^{l-1}(t)), \quad j = 1, 2, \dots, N_j \quad \dots\dots\dots (6)$$

③出力層の出力 $z_k^l(t)$ 出力層 (l 層) の k 番目の内部ポテンシャル $x_k^l(t)$ と出力 $z_k^l(t)$ は、第2中間層の j 番目の出力 $z_j^{l-1}(t)$ を用いると、前述と同様に次式となる。

図4 一般化シグモイド関数の特性

Fig. 4. A generalized sigmoid function.

$$x_k^l(t) = w_{0k}^l z_{0k}^{l-1}(t) + \sum_{j=1}^{N_j} \sum_{p=0}^L w_{jkp}^l z_j^{l-1}(t-p) \quad \dots\dots\dots (7)$$

$$z_k^l(t) = f(x_k^l(t)), \quad k = 1, 2, \dots, N_k \quad \dots\dots\dots (8)$$

〈3・2〉 逆方向の計算

出力層からの出力 $z_k^l(t)$ ($k = 1, 2, \dots, N_k$) と望ましい出力 (教師信号) $d_k(t)$ ($k = 1, 2, \dots, N_k$) との自乗誤差 $E(t)$ は、

$$E(t) = \frac{1}{2} \sum_{k=1}^{N_k} (d_k(t) - z_k^l(t))^2 = \frac{1}{2} \sum_{k=1}^{N_k} e_k^2 \quad \dots\dots\dots (9)$$

$$e_k = d_k(t) - z_k^l(t) \quad \dots\dots\dots (10)$$

となる。学習は、この誤差を最小にするように、各層の FIR 形シナプスの重み係数 $w_{hip}^{l-2}, w_{ijp}^{l-1}, w_{jkp}^l$ を更新することである。さらに、各層のシグモイド関数のゲイン $g_k^l, g_j^{l-1}, g_i^{l-2}$ を更新することである。これらの更新を行うに当たって、最急降下法と共役勾配法を用いる。

〈3・2・1〉 最急降下法

①出力層の重み係数 w_{jkp}^l の更新

$$\Delta w_{jkp}^l = -\alpha \frac{\partial E(t)}{\partial w_{jkp}^l} = -\alpha \frac{\partial E(t)}{\partial x_k^l(t)} \frac{\partial x_k^l(t)}{\partial w_{jkp}^l} \quad \dots\dots\dots (11)$$

に基づく最急降下法を用いて、 w_{jkp}^l の更新を行なう。 α は学習定数である。

$$\delta_k^l(t) = -\frac{\partial E(t)}{\partial x_k^l(t)} = -\frac{\partial E(t)}{\partial z_k^l(t)} \frac{\partial z_k^l(t)}{\partial x_k^l(t)} = (d_k(t) - z_k^l(t)) f'(x_k^l(t)) \quad \dots\dots\dots (12)$$

とおくと、

$$\Delta w_{jkp}^l = \alpha \delta_k^l(t) z_j^{l-1}(t-p) \quad \dots\dots\dots (13)$$

$$j = 1, 2, \dots, N_j, \quad k = 1, 2, \dots, N_k$$

$$p = 0, 1, 2, \dots, L$$

となり、学習則は次のようになる。

$$w_{jkp}^l(t+1) = w_{jkp}^l(t) + \Delta w_{jkp}^l \quad \dots\dots\dots (14)$$

$f'(x_k^l(t))$ は、 $f(x_k^l(t))$ の $x_k^l(t)$ に関する微分を表わしている。出力層のオフセット重み係数 w_{0k}^l の更新は、式 (13) の $z_j^{l-1}(t-p)$ を $z_{0k}^{l-1}(t)(=1)$ で置換することによって行われる。

②第2中間層の重み係数 w_{ijp}^{l-1} の更新

$$\begin{aligned}\Delta w_{ijp}^{l-1} &= -\alpha \frac{\partial E(t)}{\partial w_{ijp}^{l-1}} \\ &= -\alpha \frac{\partial E(t)}{\partial x_j^{l-1}(t)} \frac{\partial x_j^{l-1}(t)}{\partial w_{ijp}^{l-1}(t)} \dots\dots\dots (15)\end{aligned}$$

$$\begin{aligned}\delta_j^{l-1}(t) &= -\frac{\partial E(t)}{\partial x_j^{l-1}(t)} \\ &= -\sum_{k=1}^{N_k} \frac{\partial E(t)}{\partial x_k^l(t)} \frac{\partial x_k^l(t)}{\partial z_j^{l-1}(t)} \frac{\partial z_j^{l-1}(t)}{\partial x_j^{l-1}(t)} \dots (16)\end{aligned}$$

とおくと、

$$\delta_j^{l-1}(t) = f'(x_j^{l-1}(t)) \sum_{k=1}^{N_k} \delta_k^l(t) w_{jk0}^l \dots\dots\dots (17)$$

となり、

$$\begin{aligned}\Delta w_{ijp}^{l-1} &= \alpha \delta_j^{l-1}(t) z_i^{l-2}(t-p) \dots\dots\dots (18) \\ i &= 1, 2, \dots, N_i, \quad j = 1, 2, \dots, N_j \\ p &= 0, 1, 2, \dots, L\end{aligned}$$

となる。学習則は次のようになる。

$$w_{ijp}^{l-1}(t+1) = w_{ijp}^{l-1}(t) + \Delta w_{ijp}^{l-1} \dots\dots\dots (19)$$

第2中間層のオフセットの重み係数 w_{0j}^{l-1} の更新は、式(18)の $z_i^{l-2}(t-p)$ を $z_{0j}^{l-2}(t)(=1)$ で置換することによって行われる。

③第1中間層の重み係数 w_{hip}^{l-2} の更新 第2中間層の場合と同様に、最急降下法を用いると、

$$\Delta w_{hip}^{l-2} = \alpha \delta_i^{l-2}(t) z_h^{l-3}(t-p) \dots\dots\dots (20)$$

$$\begin{aligned}\delta_i^{l-2}(t) &= f'(x_i^{l-2}(t)) \sum_{j=1}^{N_j} \delta_j^{l-1}(t) w_{ij0}^{l-1} \dots\dots\dots (21) \\ h &= 1, 2, \dots, N_h, \quad i = 1, 2, \dots, N_i \\ p &= 0, 1, 2, \dots, L\end{aligned}$$

となり、学習則は次のようになる。

$$w_{hip}^{l-2}(t+1) = w_{hip}^{l-2}(t) + \Delta w_{hip}^{l-2} \dots\dots\dots (22)$$

第1中間層のオフセットの重み係数 w_{0i}^{l-2} の更新は、式(20)の $z_h^{l-3}(t-p)$ を $z_{0i}^{l-3}(t)(=1)$ で置換することによって行われる。

④各層のシグモイド関数のゲインの更新 出力層のシグモイド関数(式(4)で表される)のゲイン g_k^l の更新の式は次のようになる。

$$\begin{aligned}\Delta g_k^l &= -\alpha \frac{\partial E(t)}{\partial g_k^l} \\ &= \alpha \delta_k^l(t) x_k^l(t) f'(x_k^l(t)) \dots\dots\dots (23)\end{aligned}$$

$$\delta_k^l(t) = d_k(t) - z_k^l(t) \dots\dots\dots (24)$$

$$\begin{aligned}g_k^l(t+1) &= g_k^l(t) + \Delta g_k^l \dots\dots\dots (25) \\ k &= 1, 2, \dots, N_k\end{aligned}$$

同様に、第2中間層のシグモイド関数のゲイン g_j^{l-1} の更新と第1中間層のシグモイド関数のゲイン g_i^{l-2} の更新の式は、それぞれ

$$\Delta g_j^{l-1} = \alpha \delta_j^{l-1}(t) x_j^{l-1}(t) f'(x_j^{l-1}(t)) \dots\dots\dots (26)$$

$$\begin{aligned}\delta_j^{l-1}(t) &= \sum_{k=1}^{N_k} \delta_k^l(t) f'(x_k^l(t)) w_{jk0}^l \dots\dots\dots (27) \\ j &= 1, 2, \dots, N_j\end{aligned}$$

$$\Delta g_i^{l-2} = \alpha \delta_i^{l-2}(t) x_i^{l-2}(t) f'(x_i^{l-2}(t)) \dots\dots\dots (28)$$

$$\begin{aligned}\delta_i^{l-2}(t) &= \sum_{j=1}^{N_j} \delta_j^{l-1}(t) f'(x_j^{l-1}(t)) w_{ij0}^{l-1} \dots\dots\dots (29) \\ i &= 1, 2, \dots, N_i\end{aligned}$$

となる。上記の FIR 形ニューラルネットワークの学習則を繰り返し(反復法を)用いて、誤差が最小になるように重み係数やゲインの更新を行うに当たっては、次のように収束の判定を定める必要がある。その一つは、学習用入力データとして N 個のサンプルをニューラルネットワークに与えたときの学習を一回の学習と定めた場合に、式(9)から得られる自乗誤差の平均値 $E_a(t)$ が $E_a(t) < \epsilon$ を満たしたときに収束基準値 ϵ (例えば、 $\epsilon = 0.001$) に達したと見なし、学習(更新)を終了することである。もう一つは、学習の制限回数を指定し、反復法により学習回数とその制限回数(例えば、5000回)に達したときに、学習(更新)を終了することである。

〈3・2・2〉 共役勾配法^{(9)~(11)} 式(9)の $E(t)$ を $E(\mathbf{w}, t)$ とし、誤差関数 $\phi(\mathbf{w})$ を次式とする。

$$\begin{aligned}\phi(\mathbf{w}) &= \frac{1}{P} \sum_{n=1}^P E(\mathbf{w}, t_n) \\ &= \frac{1}{2P} \sum_{n=1}^P \sum_{k=1}^{N_k} (d_k(t_n) - z_k^l(t_n))^2 \dots\dots\dots (30)\end{aligned}$$

この式は、 $t = t_n (n = 1, 2, \dots, P)$ における自乗誤差の平均値である。 $\mathbf{w}$ はニューラルネットワークの各層に含まれているすべての重みやシグモイド関数のゲインから成る列ベクトルである。出力層、第2中間層、第1中間層の重み(オフセットの重みを含む)ベクトルを、それぞれ $\mathbf{w}^l, \mathbf{w}^{l-1}, \mathbf{w}^{l-2}$ とし、各層のユニットのシグモイド関数のゲインベクトルを $\mathbf{g}^l, \mathbf{g}^{l-1}, \mathbf{g}^{l-2}$ とすると、 $\mathbf{w}$ は

$$\mathbf{w} = [\mathbf{w}^l, \mathbf{w}^{l-1}, \mathbf{w}^{l-2}, \mathbf{g}^l, \mathbf{g}^{l-1}, \mathbf{g}^{l-2}]^T \dots\dots\dots (31)$$

となる。 T は転置を表している。共役勾配法による学習アルゴリズムを次のような手順で行う。

ステップ1：誤差関数(目的関数) $\phi(\mathbf{w})$ を最小にする点 $\mathbf{w}^*$ の初期値 $\mathbf{w}^{(0)}$ を与える。 $\mathbf{w}^{(0)}$ に含まれる重みベクトル $\mathbf{w}^{l(0)}, \mathbf{w}^{l-1(0)}, \mathbf{w}^{l-2(0)}$ には -0.5 から 0.5 の間のランダム値で初期化を行い、 $\mathbf{w}^{(0)}$ に含まれるゲインベクトル $\mathbf{g}^{l(0)}, \mathbf{g}^{l-1(0)}, \mathbf{g}^{l-2(0)}$ には $0 \sim 1.0$ の間の値を与え、 $\mathbf{w}^{(0)}$ から勾配 (gradient) $\mathbf{D}^{(0)} = -\nabla \phi(\mathbf{w}^{(0)})$ を求める。

ステップ2: $\phi(\mathbf{w}^{(k)} + \alpha_k \mathbf{D}^{(k)})$ を最小にする α_k をライン探索 (Line Search) によって求める。ここで k は反復数を表すものとする。まず, $k=0$ に対しては $\phi(\mathbf{w}^{(0)} + \alpha_0 \mathbf{D}^{(0)})$ を最小にする α_0 を求める。反復法により順次, $\phi(\mathbf{w}^{(1)} + \alpha_1 \mathbf{D}^{(1)})$, $\phi(\mathbf{w}^{(2)} + \alpha_2 \mathbf{D}^{(2)})$, ... を最小にする定数 $\alpha_1, \alpha_2, \dots$ をライン探索によって求め, 誤差関数 $\phi(\mathbf{w})$ が収束値 ϵ に達するまで繰り返す。ライン探索には, $\phi(\mathbf{w}^{(k)} + \alpha_k \mathbf{D}^{(k)})$ の最小値を含む最小の区間を囲い込みアルゴリズムによって求め, 黄金分割法によって最小値を与える α_k を求めることにする⁽¹²⁾。

ステップ3: 重みベクトルの更新を次の式によって求める。

$$\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} + \alpha_k \mathbf{D}^{(k)} \quad \dots\dots\dots (32)$$

$\mathbf{w}^{(k+1)}$ が $\phi(\mathbf{w}^{(k+1)})$ を最小にする値に収束するまで, 順次, $\mathbf{w}^{(2)}, \mathbf{w}^{(3)}, \dots, \mathbf{w}^{(k+1)}$ を求める。

ステップ4: $|\phi(\mathbf{w}^{(k+1)})| \leq \epsilon$ となったときは, 収束値 ϵ に達したとみなす。そうでないときは, 次のステップに移る。

ステップ5: 次の探索方向ベクトル $\mathbf{D}^{(k+1)}$ を求める。 $k+1$ が全体の重みの数とゲインの数の和 S の整数倍になったときは

$$\mathbf{D}^{(k+1)} = -\nabla\phi(\mathbf{w}^{(k+1)}) \quad \dots\dots\dots (33)$$

とし, そうでないときは

$$\mathbf{D}^{(k+1)} = -\nabla\phi(\mathbf{w}^{(k+1)}) + \beta_k \mathbf{D}^{(k)} \quad \dots\dots\dots (34)$$

$$\beta_k = \frac{\nabla\phi(\mathbf{w}^{(k+1)})^T [\nabla\phi(\mathbf{w}^{(k+1)}) - \nabla\phi(\mathbf{w}^{(k)})]}{\|\nabla\phi(\mathbf{w}^{(k)})\|^2} \quad \dots\dots\dots (35)$$

とする。

ステップ6: $k+1 \rightarrow k$ とし, ステップ2に戻る。

式 (30) からわかるように, $\nabla\phi(\mathbf{w})$ を求めることは, $\nabla E(\mathbf{w}, t) = \partial E(\mathbf{w}, t) / \partial \mathbf{w}$ を求めることになる。これは各層の重みとゲインに対する偏微分の列ベクトルから構成される。出力層の重みベクトル $\mathbf{w}^l$, 第2中間層の重みベクトル $\mathbf{w}^{l-1}$ 及び第1中間層の重みベクトル $\mathbf{w}^{l-2}$ は, それぞれ w_{jkp}^l, w_{ijp}^{l-1} 及び w_{hip}^{l-2} を要素とするベクトルであるから, 各層の重みによる $\nabla E(\mathbf{w}, t)$ は, 式 (11), 式 (15), 式 (20) から次のように得られる。

$$\frac{\partial E}{\partial w_{jkp}^l} = -\delta_k^l(t) z_j^{l-1}(t-p) \quad \dots\dots\dots (36)$$

$$\frac{\partial E}{\partial w_{ijp}^{l-1}} = -\delta_j^{l-1}(t) z_i^{l-2}(t-p) \quad \dots\dots\dots (37)$$

$$\frac{\partial E}{\partial w_{hip}^{l-2}} = -\delta_i^{l-2}(t) z_h^{l-3}(t-p) \quad \dots\dots\dots (38)$$

$$h = 1, 2, \dots, N_h, \quad i = 1, 2, \dots, N_i$$

$$j = 1, 2, \dots, N_j, \quad k = 1, 2, \dots, N_k$$

$$p = 0, 1, 2, \dots, L$$

$\delta_k^l(t), \delta_j^{l-1}(t), \delta_i^{l-2}(t)$ は, それぞれ式 (12), 式 (17), 式 (21) と同一である。また, 出力層のジグモイド関数のゲインベ

クトル $\mathbf{g}^l$, 第2中間層のジグモイド関数のゲインベクトル $\mathbf{g}^{l-1}$ 及び第1中間層のジグモイド関数のゲインベクトル $\mathbf{g}^{l-2}$ は, それぞれ g_k^l, g_j^{l-1} 及び g_i^{l-2} を要素とするベクトルであるから, 各層のゲインによる $\nabla E(\mathbf{w}, t)$ は式 (23), 式 (26), 式 (28) から次のように得られる。

$$\frac{\partial E}{\partial g_k^l} = -\delta_k^l(t) x_k^l(t) f'(x_k^l(t)) \quad \dots\dots\dots (39)$$

$$\frac{\partial E}{\partial g_j^{l-1}} = -\delta_j^{l-1}(t) x_j^{l-1}(t) f'(x_j^{l-1}(t)) \quad \dots\dots\dots (40)$$

$$\frac{\partial E}{\partial g_i^{l-2}} = -\delta_i^{l-2}(t) x_i^{l-2}(t) f'(x_i^{l-2}(t)) \quad \dots\dots\dots (41)$$

$$i = 1, 2, \dots, N_i, \quad j = 1, 2, \dots, N_j$$

$$k = 1, 2, \dots, N_k$$

$\delta_k^l(t), \delta_j^{l-1}(t), \delta_i^{l-2}(t)$ は, それぞれ式 (24), 式 (27), 式 (29) と同一である。

4. システム同定への応用

〈4・1〉 システム同定 ニューラルネットワークによるシステム同定の問題は, 図5のような構成において, ニューラルネットワークから成る同定モデルを設定し, 同じ入力を加えたときに, システムの出力 $y_p(t+1)$ と同定モデルの出力 $\hat{y}_p(t+1)$ との間の誤差を最小にするように, 同定モデルのパラメータ (ニューラルネットワークの重み) を調整し更新することである。同定モデルの重みを更新するに当たって, 3. で述べた学習アルゴリズムを用いる。一般的な非線形システムが, 次のように入力 $u(t)$ とその過去の値, 出力 $y_p(t+1)$ の過去の値に依存した差分方程式によって表されるとする。

$$y_p(t+1) = f[y_p(t), y_p(t-1), \dots, y_p(t-n), u(t), u(t-1), \dots, u(t-m)] \quad \dots\dots\dots (42)$$

システムの次数 n, m はわかっているものと仮定し, ニューラルネットワークを用いた同定モデルを次のように設定する。

図5 システム同定の構成

Fig. 5. A constitution of the system identification.

$$\hat{y}_p(t+1) = N[\hat{y}_p(t), \hat{y}_p(t-1), \dots, \hat{y}_p(t-n), \\ u(t), u(t-1), \dots, u(t-m)] \dots (43)$$

ここで、 $\hat{y}_p(t+1)$ は同定モデルの出力である。システムのダイナミクスを同定することは、適切な層をもつ階層形ニューラルネットワークから成る同定モデルに学習を実施し、重みの更新を行うことである。学習期間には、システムと同定モデルと同じ入力を与え、図5のスイッチSをUp側に倒し、ニューラルネットワークはシステムの出力 $y_p(t+1)$ を受け取るようにし、ニューラルネットワークを直並列モデル⁽⁶⁾として扱うことにする。通常のBPニューラルネットワークを用いるときには、種々の時刻 t に対して、入力出力対として $\{[y_p(t), y_p(t-1), \dots, y_p(t-n), u(t), u(t-1), \dots, u(t-m)]^T, y_p(t+1)\}$ を与え、ニューラルネットワークの各層の重みを更新することになる。学習終了後、同定テスト期間に移ってからは、スイッチSをDown側に倒し、式(43)の関係が成り立つようにニューラルネットワークを並列モデルとして扱い、ニューラルネットワークの出力 $\hat{y}_p(t+1)$ を入力側にフィードバックし、システムと同定モデルに同じテスト信号を加えて同定テストを行う。学習が首尾よく行われたならば、同じ入力と同じ初期条件に対して、同定モデルの出力 $\hat{y}_p(t+1)$ はシステムの実際の出力 $y_p(t+1)$ を良好に近似すると期待できる。

〈4・2〉 同定例1 未知のシステム（プラント）として、式(44)に示すように、出力が入力の過去の値と出力の過去の値に非線形的に依存し、2個の付加的結合から成る非線形差分方程式で表されるシステムを取り上げる⁽¹³⁾。

$$y_p(t+1) = f_1(y_p(t), y_p(t-1), y_p(t-2), u(t), u(t-1)) \\ + f_2(y_p(t), y_p(t-1), y_p(t-2), u(t), u(t-1)) \\ = \frac{y_p(t)y_p(t-1)y_p(t-2)u(t-1)(y_p(t-2)-1) + u(t)}{1 + y_p^2(t-1)y_p^2(t-2)} \\ + \frac{y_p(t)y_p(t-1)u(t-1)(y_p(t-2)-1) + u(t)}{1 + y_p^2(t) + y_p^2(t-1)} \dots (44)$$

ここで、 $u(t)$ は入力であり、 $y_p(t+1)$ はシステムの出力である。このシステムを同定するために、式(45)に示すような2個のFIR形ニューラルネットワーク（以下、FIR-NNとする）から成る同定モデルを設定する。

$$\hat{y}_p(t+1) = N_1(y_p(t), y_p(t-1), y_p(t-2), u(t), u(t-1)) \\ + N_2(y_p(t), y_p(t-1), y_p(t-2), u(t), u(t-1)) \dots (45)$$

式(45)は学習期間にこの同定モデルを直並列モデルとして扱った場合の入出力関係である。図6に示すような構成においてこのシステムの同定を行うことは、同定モデルの写像 $N_1(\cdot)$ と $N_2(\cdot)$ の和がシステムの非線形関数 $f_1(\cdot)$ と $f_2(\cdot)$ の和にできるだけ等しくなるように、2個のFIR-NNの重みやシグモイド関数のゲインを調整し、 N_1 と N_2 を決定することである。

$y_p(t+1)$ と $\hat{y}_p(t+1)$ は、それぞれシステムと同定モデルの時刻 t における1個の出力であるので、式(9)の誤差

関数 $E(t)$ は

$$E(t) = \frac{1}{2}(d_k(t) - z_k^l(t))^2 = \frac{1}{2}\epsilon(t+1)^2 \\ = \frac{1}{2}(y_p(t+1) - \hat{y}_p(t+1))^2 \dots \dots \dots (46)$$

となる。この場合、システム出力 $y_p(t+1)$ はFIR-NNの望ましい出力（教師信号）である。この誤差を最小にするように、 N_1 と N_2 の各層のFIRシナプスの重みやゲインを更新するために、前述の最急降下法や共役勾配法による学習アルゴリズムが用いられる。

① ネットワーク構造 ニューラルネットワーク N_1, N_2 いずれにも図3に示した4層から成るFIR-NNを用いることにし、 N_1 と N_2 のネットワーク構造をいずれも2-20(3,1)-10(N_{d1})-1(N_{d2})とする。これは、入力層、第1中間層、第2中間層及び出力層の各ユニット数がそれぞれ2個、20個、10個、1個であることを示している。(3,1)は、 $f_1(\cdot)$ と $f_2(\cdot)$ がともに3個の出力の過去の値と1個の入力の過去の値に依存しているので、第1中間層の各ユニットには1個の入力ユニットからは3個の遅延素子を与え、もう1個の入力ユニットから1個の遅延素子を与えればよいことを示している。 N_{d1} と N_{d2} はそれぞれ第2中間層の各ユニットと出力層の1個のユニットに与えられる遅延素子の数を表している。

同定精度の点で、FIR-NNと比較するために用いられる通常のBPニューラルネットワーク（以下、BP-NNとする）のネットワーク構造を5-20-10-1とする。入力ユニット数を5個としたのは、入力層に $y_p(t), y_p(t-1), y_p(t-2), u(t), u(t-1)$ の5個の信号が加わるからである。

② 同定精度 学習期間には、システムと同定モデルに同じ入力信号を与える。入力信号として、 $[-1, 1]$ の範囲で一様に分布するランダム信号を用いる。ランダム信号を500

図6 非線形システム同定の構成
Fig. 6. A constitution of the nonlinear system identification.

表1 FIR-NN(2-20(3,1)-10(N_{d1})-1(N_{d2})) による同定モデルと BP-NN(5-20-10-1) による同定モデルを用いたときの同定精度

Table 1. A comparison of the identification accuracy of the identification model based on FIR-NN and BP-NN.

FIR-NNの $N_{d1}(=N_{d2})$	最急降下法を用いたときの同定精度	共役勾配法を用いたときの同定精度
3	0.0601	0.0447
4	0.0425	0.0318
5	0.0316	0.0412
6	0.0497	0.0426
7	0.0448	0.0447
8	0.0723	0.0422
9	0.0742	0.0536
10	0.0785	0.0512
BP-NN	同定精度	
5-20-10-1	0.11392	

サンプルを与えたときの学習を1回の学習とみなすことにする。FIR-NNの各層の各ユニットの重みやシグモイド関数のゲインの初期値をランダムに設定し、第1中間層と第2中間層の各ユニットのシグモイド関数の式(4)で表される上下限値 $[b, a]$ を $[1, -1]$ とし、出力ユニットのシグモイド関数の上下限値を $[1.2, -1.2]$ とする。最急降下法を用いたときの学習定数 α を0.25とし、収束基準値を0.001とする。また、共役勾配法を用いたときの収束基準値も0.001とする。テスト期間(学習終了後)には、同定テスト信号 $u(t)$ として、次の信号を加えて、システム出力 $y_p(t)$ と同定モデル出力 $\hat{y}_p(t)$ を求める。

$$u(t) = \begin{cases} \sin(\frac{2\pi t}{250}) & 0 \leq t \leq 500 \\ 0.8 \sin(\frac{2\pi t}{250}) + 0.2 \sin(\frac{2\pi t}{25}) & t > 500 \end{cases} \quad (47)$$

このとき、同定精度 A_p は次の式で求めることにする。

$$A_p = \sqrt{\frac{1}{1000} \sum_{t=0}^{1000} (y_p(t) - \hat{y}_p(t))^2} \quad \dots\dots\dots (48)$$

③テスト結果 N_1 と N_2 のFIR-NNのネットワーク構造を2-20(3,1)-10(N_{d1})-1(N_{d2})としたときに、 $N_{d1} = N_{d2} = 3 \sim 10$ のように遅延素子数を変化させた場合について検討した。最急降下法による学習アルゴリズム(式(2)～式(29))を用いて学習し、収束後に式(47)のテスト信号を加えたときに、式(48)から得られた同定精度と収束基準値に至るまでの学習回数を求めると、表1、表2に示すような結果が得られた。次に、同じネットワーク構造に、共役勾配法による学習アルゴリズム(式(2)～式(10)、式(30)～式(40))を用いて学習させ、収束後、式(47)のテスト信号を与えたときに得られた同定精度と収束基準値に至るまでの学習回数を示すと表1、表2のような結果になった。また、 N_1 と N_2 にBP-NNを用いたときに、ネットワーク構造を5-20-10-1と指定し、最急降下法の学習定数 α を0.25と

表2 FIR-NN(2-20(3,1)-10(N_{d1})-1(N_{d2})) による同定モデルの学習回数

Table 2. The learning cycle of the identification model based on FIR-NN.

FIR-NNの $N_{d1}(=N_{d2})$	最急降下法を用いたときの学習回数	共役勾配法を用いたときの学習回数
3	85	25
4	98	31
5	173	28
6	122	27
7	199	19
8	598	78
9	310	74
10	523	85

図7 最急降下法を用いたときのFIR-NN(2-20(3,1)-10(3)-1(3))の出力 $\hat{y}_p(t)$ とシステム出力 $y_p(t)$

Fig. 7. The output $\hat{y}_p(t)$ of FIR-NN and the output $y_p(t)$ of the system based on the gradient descent method.

図8 最急降下法を用いたときのFIR-NN(2-20(3,1)-10(3)-1(3))の学習回数と誤差の関係

Fig. 8. The relation between the learning cycle of FIR-NN and the error based on the gradient descent method.

し、同じ収束基準値(0.001)で学習した後に、式(47)のテスト信号を加えたときに得られた同定精度は表1に示すような結果となった。

FIR-NNのネットワーク構造を2-20(3,1)-10(3)-1(3)と指定し、最急降下法を用いた場合のシステム出力 $y_p(t)$ 、同定モデル出力 $\hat{y}_p(t)$ と、学習回数と誤差の関係を示すとそれ

図 9 共役勾配法を用いたときの FIR-NN の出力 $\hat{y}_p(t)$ とシステム出力 $y_p(t)$

Fig. 9. The output $\hat{y}_p(t)$ of FIR-NN and the output $y_p(t)$ of the system based on the conjugate gradient method.

図 10 共役勾配法を用いたときの FIR-NN の学習回数と誤差の関係

Fig. 10. The relation between the learning cycle of FIR-NN and the error based on the conjugate gradient method.

図 11 最急降下法を用いたときの BP-NN (5-20-10-1) の出力 $\hat{y}_p(t)$ とシステム出力 $y_p(t)$

Fig. 11. The output $\hat{y}_p(t)$ of BP-NN and the output $y_p(t)$ of the system based on the gradient descent method.

図 12 最急降下法を用いたときの BP-NN (5-20-10-1) の学習回数と誤差の関係

Fig. 12. The relation between the learning cycle of BP-NN and the error.

ぞれ図 7, 図 8 のようになり, 共役勾配法を用いた場合のシステム出力 $y_p(t)$ と同定モデル出力 $\hat{y}_p(t)$ と, 学習回数と誤差の関係を示すとそれぞれ図 9, 図 10 のようになった。また, BP-NN のネットワーク構造を 5-20-10-1 と指定したときの両方の出力と, 学習回数と誤差の関係を示すとそれぞれ図 11, 図 12 のような結果になった。図 7, 図 9, 図 11 と表 1 の結果から, FIR-NN による同定モデルの方が BP-NN による同定モデルよりも同定において改善され, 良好な精度になっていることがわかる。遅延素子数 N_{d1} と N_{d2} を単に増加することは必ずしも精度の向上に寄与するだけでなく, この場合は $N_{d1} = N_{d2} = 4$ または 5 が最適である。このことは N_{d1} と N_{d2} を変化させてシミュレーションを行った結果から得られたものであり, 理論的に最適な N_{d1} と N_{d2} の決定については別途検討する必要があると思われる。共役勾配法を用いた結果は, 精度の点では, 最急降下法を用いた結果よりもとくに良くなっていないが, 表 2 に示されたように, 学習回数を減少させることに寄与しており, また, 図 8 と図 10 からわかるように, 最急降下法に出てくる誤差の振動はみられず, 円滑にかつ急速に収束

している。

学習終了後の同定テスト信号 $u(t)$ として, さらに

$$u(t) = 0.5 \sin\left(\frac{1 \times 2\pi t}{250}\right) + 0.4 \sin\left(\frac{2 \times 2\pi t}{250}\right) + 0.3 \sin\left(\frac{3 \times 2\pi t}{250}\right) + 0.2 \sin\left(\frac{4 \times 2\pi t}{250}\right) + 0.1 \sin\left(\frac{5 \times 2\pi t}{250}\right) \dots (49)$$

及び,

$$u(t) = \begin{cases} \sin\left(\frac{\pi t}{25}\right) & 0 \leq t < 250 \\ 1.0 & 250 \leq t < 500 \\ -1.0 & 500 \leq t < 750 \\ 0.3 \sin\left(\frac{\pi t}{25}\right) + 0.1 \sin\left(\frac{\pi t}{32}\right) + 0.6 \sin\left(\frac{\pi t}{10}\right) & 750 \leq t < 1000 \end{cases} (50)$$

で表される 2 種類の信号を用意し, これらの信号を与えたときの同定モデルの出力とシステム出力を求めてみる。FIR-NN のネットワーク構造を 2-20(3,1)-10(4)-1(4) と指定

図 13 式 (49) のテスト信号を加えたときの FIR-NN の出力 $\hat{y}_p(t)$ とシステム出力 $y_p(t)$

Fig. 13. The output $\hat{y}_p(t)$ of FIR-NN and the output $y_p(t)$ of the system by applying the equation (49).

図 14 式 (49) のテスト信号を加えたときの BP-NN の出力 $\hat{y}_p(t)$ とシステム出力 $y_p(t)$

Fig. 14. The output $\hat{y}_p(t)$ of BP-NN and the output $y_p(t)$ of the system by applying the equation (49).

図 15 式 (50) のテスト信号を加えたときの FIR-NN の出力 $\hat{y}_p(t)$ とシステム出力 $y_p(t)$

Fig. 15. The output $\hat{y}_p(t)$ of FIR-NN and the output $y_p(t)$ of the system by applying the equation (50).

図 16 式 (50) のテスト信号を加えたときの BP-NN の出力 $\hat{y}_p(t)$ とシステム出力 $y_p(t)$

Fig. 16. The output $\hat{y}_p(t)$ of BP-NN and the output $y_p(t)$ of the system by applying the equation (50).

し、最急降下法による学習アルゴリズムを用いて学習し、式 (49) のテスト信号を加えたときのシステム出力 $y_p(t)$ と同定モデル出力 $\hat{y}_p(t)$ を求めると図 13 のようになり、BP-NN のネットワーク構造を 5-20-10-1 と指定し、学習終了後に式 (49) を与えたときの両方の出力を求めると図 14 のようになった。図 13、図 14 から、この場合も、FIR-NN による同定モデルがシステムを良好に同定し、BP-NN による同定モデルよりも同定において改善されていることがわかる。次に、FIR-NN のネットワーク構造を 2-20(3,1)-10(5)-1(5) とし、BP-NN のそれを 5-20-10-1 としたときに、式 (50) のテスト信号を加えたときのシステム出力 $y_p(t)$ と同定モデル出力 $\hat{y}_p(t)$ を求めると、それぞれ図 15、図 16 に示すようになった。図 16 においては、BP-NN による同定モデルの出力とシステム出力との間にかなり (とくに平坦部に) ずれがみられるが、図 15 から FIR-NN による同定モデルは全体的に (平坦部も含めて) システムを良好に同定し、

FIR-NN は非線形システムの同定において BP-NN よりも優れていることがわかる。

以上は、式 (44) が 2 個の付加的結合から成る非線形システムについてであるが、ここで実施した方法は一般に r 個の付加的結合から成る非線形システムの同定にも適用することができる。その場合には r 個の FIR 形ニューラルネットワークが必要となる。

〈4・3〉 同定例 2 未知システムとして、式 (51) に示すように、出力の過去の値に依存する非線形と入力に依存する線形から成るシステムを取りあげる。

$$y_p(t+1) = f(y_p(t), y_p(t-1)) + \alpha_0 u(t) + \alpha_1 u(t-1) \quad (51)$$

$$f(y_p(t), y_p(t-1)) = \frac{y_p(t)y_p(t-1)(y_p(t)+2.5)}{1+y_p^2(t)+y_p^2(t-1)} \quad \dots \quad (52)$$

ここで、 $\alpha_0 = 0.3$ 、 $\alpha_1 = 0.6$ と指定しておくことにする。 $u(t)$ は入力であり、 $y_p(t+1)$ は出力である。このシステム

を同定するモデルとして,

$$\hat{y}_p(t+1) = N(y_p(t), y_p(t-1)) + \hat{\alpha}_0 u(t) + \hat{\alpha}_1 u(t-1) \quad (53)$$

を設ける。写像 N に FIR-NN を用いることにし、線形部分の定数 $\hat{\alpha}_0$ と $\hat{\alpha}_1$ を次の式によって推定する。

$$\begin{aligned} \hat{\alpha}_i^{\text{new}} &= \hat{\alpha}_i^{\text{old}} - \eta(\partial E(t)/\partial t) \\ &= \hat{\alpha}_i^{\text{old}} - \eta(\hat{y}_p(t+1) - y_p(t+1))u(t-1) \cdots \quad (54) \\ i &= 0, 1 \end{aligned}$$

ここで、 $E(t)$ は式 (46) で表わされる誤差関数である。FIR-NN のネットワーク構造を 1-20(2)-10(2)-1(2) とし、第 1 中間層と第 2 中間層の各ユニットのシグモイド関数の上下限値 $[b, a]$ を $[0.5, -0.5]$ とし、出力ユニットのそれを $[4.0, -3.0]$ とする。

学習期間に、 $[-2, 2]$ の範囲で一様に分布するランダム信号を 500 サンプルだけ用意し、500 サンプルによる学習を 1 回の学習とする。最急降下法の学習定数 α と η を 0.05 とし、学習回数の制限回数を 5000 回とする。収束の判定には、同定例 1 のように収束基準値を指定する場合と、学習の制限回数を指定する場合とがあるが、この同定例では学習の制限回数を指定する場合を選ぶことにする。5000 回の学習を行った後に、テスト信号として

$$u(t) = \sin\left(\frac{2\pi t}{25}\right) \cdots \cdots \cdots \quad (55)$$

を与えたとき、システム出力 $y_p(t)$ と同定モデル出力 $\hat{y}_p(t)$ を求めると、図 17 に示すような結果が得られ、精度は 0.05680 となった。推定された $\hat{\alpha}_0$ と $\hat{\alpha}_1$ はそれぞれ 0.30052, 0.59988 となり、線形部分の定数によく接近した値となった。図 17 から、同定モデルの出力がシステムの出力によく一致し、FIR-NN による同定モデルはシステムを良好に同定していることがわかる。

図 17 FIR-NN(1-20(2)-10(2)-1(2)) の出力 $\hat{y}_p(t)$ とシステム出力 $y_p(t)$

Fig. 17. The output $\hat{y}_p(t)$ of FIR-NN and the output $y_p(t)$ of the system.

5. む す び

本論文では、FIR 形ニューラルネットワークを取りあげ、最急降下法と共役勾配法に基づくこのニューラルネットワークの学習アルゴリズムを導き、それを非線形システムの同定問題に適用した。過去の出力と過去の入力に依存して発生する出力をもつ非線形システムの同定において、FIR 形ニューラルネットワークによる同定モデルの同定精度は、通常の BP ニューラルネットワークによる同定モデルの同定精度よりも優れ、良好な同定結果を示した。このことは、汎化能力の点でも、FIR 形ニューラルネットワークが BP ニューラルネットワークよりも勝っていることを示している。システム同定に共役勾配法による学習アルゴリズムを適用した場合と最急降下法による学習アルゴリズムを適用した場合を比較した結果、前者は後者よりも学習回数を減少させることに寄与し、後者は収束値に達するまでに振動を起こすのに対して、前者は円滑にかつ急速に収束するようになり、前者の方が学習アルゴリズムとして優れていることがわかった。FIR 形シナプスに含まれる遅延素子が有効に機能し、同定精度の向上 (または非線形関数の近似の向上) に寄与していることがわかった。

入力とその過去の値 $u(t), u(t-1), \dots, u(t-m)$ 、出力の過去の値 $y_p(t), y_p(t-1), \dots, y_p(t-n)$ とが非線形システムに加わるときに、BP ニューラルネットワークの場合には $m+n+2$ 個の入力ユニットが必要とされるが、FIR 形ニューラルネットワークの場合には 2 個の入力ユニットだけを用意し、2 個の入力ユニットから第 1 中間層の各ユニットに対して、1 個の入力ユニットからは $n+1$ 個の遅延素子を設け、もう 1 個のユニットからは m 個の遅延素子を設ければよいことが明らかになった。本論文では、各層の各ユニットに含まれる非線形関数として一般化シグモイド関数を提案し、この関数の上下限値を任意の値に指定できるようにした。そのため、学習用の入力信号やテスト信号の振幅に応じて、シグモイド関数の上下限値を必要な値に設定できる利点が得られた。

第 2 中間層の各ユニットの遅延素子数 N_{d1} と出力ユニットの遅延素子数 N_{d2} を ($N_{d1} = N_{d2}$ として) 何個にすると同定精度を最も良好にすることができるか (本シミュレーション結果では 4 個または 5 個となったが) についての理論的な最適決定法は重要な課題であり、これについてはさらに検討する必要がある。FIR 形ニューラルネットワークを適応制御系の制御器として使用することについても、今後の検討課題である。

(平成 12 年 6 月 5 日受付, 同 12 年 9 月 27 日再受付)

文 献

- (1) F.U.Dowla, A.J.DeGroot and S.R.Parker, "Back-propagation Neural Networks: Systolic Implementation for Seismic Signal Filtering", Neural Networks, Vol.1, No.3, pp.138-158, 1989-7.
- (2) Fa-Long Luo and Rolf Unbehauen, "Applied Neural Networks

- for Signal processing", Cambridge University Press, 1997.
- (3) 片山徹, "システム同定入門", 朝倉書店, 1993.
 - (4) W.T.Miller, R.S.Sutton and P.J.Werbos, "Neural Networks for Control", MIT Press, 1991.
 - (5) Cheng-Jian Lin, "SISO Nonlinear System Identification Using a Fuzzy-Neural Hybrid Systems", International Journal of Neural Systems, Vol.8, No.3, pp325-337, 1997-6.
 - (6) K.S.Narendra and K.Parthasarathy, "Identification and Control of Dynamic Systems Using Neural Networks", IEEE Trans. on Neural Networks, Vol.1, No.1, pp4-27, 1990-3.
 - (7) K.S.Narendra and K.Parthasarathy, "Gradient Method for the Optimization of Dynamical Systems Containing Neural Networks", IEEE Trans. on Neural Networks, Vol.2, No.2, pp252-262, 1991-3.
 - (8) E.A.Wan, "Time Series Prediction Using Connectionist Network with Internal Delay Lines", Time Series Prediction, pp195-217, Addison Wesley.
 - (9) E.M.Johansson, F.U.Dowla and D.M.Goodman, "Backpropagation Learning for Multilayer Feedforward Neural Networks Using the Conjugate Gradient Method", International Journal of Neural Systems, Vol.2, No.4, pp291-301, 1992.
 - (10) C.Charalambous, "Conjugate gradient algorithm for efficient training of neural networks", IEE Proc.-G, Vol.139, No.3, pp301-310, 1992-6.
 - (11) J.P.Fitch, S.K.Lehman and F.U.Dowla, "Ship Wake-Detection Procedure Using Conjugate Gradient Trained Artificial Neural Networks", IEEE Trans. on Geoscience and Remote Sensing, Vol.29, No.5, pp718-726, 1991-9.
 - (12) 今野浩, "非線形計画法", 日科技連, 1978.
 - (13) 降矢順治, 西正明, "FIR 形ニューラルネットワークのシステム同定への応用", 計測自動制御学会東北支部 35 周年記念学術講演会予稿集, pp.43-44, 1999-10.

降 矢 順 治 (正員) 1936 年生。1960 年 3 月東北大学工学部通信工学科卒業。同年北辰電機製作所入社。1969 年 3 月東京電機大学大学院工学研究科修士課程修了。同年東芝電気(株)入社。1971 年山口大学教育学部講師。1974 年山口大学教育学部助教授。1981 年山口大学教育学部教授。1999 年城西国際大学経営情報学部教授。工学博士。伝送回路網, 神経回路網の研究に従事。電子情報通信学会, 経営情報学会, 日本産業技術教育学会各会員。

西 正 明 (正員) 1956 年生。1978 年 3 月岩手大学工学部電子工学科卒業。1980 年 3 月東北大学大学院工学研究科修士課程修了。同年日立製作所中央研究所入社。1989 年山口大学教育学部助教授。2000 年信州大学教育学部助教授。情報科学博士。高速信号伝送方式, 神経回路網, 教育工学の研究に従事。IEEE, 電子情報通信学会, 日本産業技術教育学会各会員。

松 坂 知 行 (正員) 1937 年生。1960 年 3 月東北大学工学部電気工学科卒業。同年 4 月東芝(株)入社。1972 年 4 月八戸工業大学助教授。1979 年 4 月教授。現在情報システム工学研究所教授。エネルギーシステムの制御, シミュレーションの研究に従事。1986 年 1 月より 1998 年 3 月まで NEDO 大型風力発電技術開発部会委員。工学博士。電子情報通信学会, 計測自動制御学会各会員。