

適者生存型学習則を適用した競合動径基底関数ネットワーク

奥原 浩之[†] 尾崎 俊治[†]

Competitive Radial Basis Function Networks Applying the “Survival of the Fittest Type” Learning Algorithm

Koji OKUHARA[†] and Shunji OSAKI[†]

あらまし 本論文では、シナプス結合荷重間の競合を考慮したシナプス可塑性方程式を導出し、これを学習則として適用した動径基底関数ネットワークを提案する。動径基底関数ネットワークは階層型ニューラルネットワークに比較してニューロンごとの局所的な学習が可能であるなどの優れた点をもつため、関数近似問題やパターン識別に適用され成果を上げている。しかし、動径基底関数ネットワークでは未知の非線形関数を近似するためにあらかじめ必要なニューロン数が不明であるために冗長なニューロンを必要とする。一般に、ニューロンの増加は学習の遅延化や過学習の問題を生じることが知られている。そこで、これらの問題を解決するために、適者生存型学習則に基づいたシナプス可塑性方程式を適用した競合動径基底関数ネットワークを提案する。競合動径基底関数ネットワークではシナプス結合荷重間に競合を生じさせ、学習に必要なニューロンのみが自然に生き残り、学習の効率化を図ることができる。シミュレーションでは、競合動径基底関数ネットワークが通常の動径基底関数ネットワークに比較して高速に学習できることを示す。また、未知の確率密度関数に従い分布する標本点が与えられたとき、冗長なニューロンをもつ競合動径基底関数ネットワークが学習により不必要なニューロンを排除して、最適なニューロン数で元の確率密度関数を推定できることを示す。

キーワード 動径基底関数ネットワーク、シナプス可塑性方程式、適者生存型学習則、学習の高速化、確率密度関数の推定

1. ま え が き

ニューラルネットは素子であるニューロンとそれらをつなぐシナプス結合荷重から構成される。ニューラルネットへの入力を入力ニューロンで受け取られ、ニューロンとシナプス結合荷重による変化を受けて伝達され、出力ニューロンから出力される。このとき、ニューラルネットの学習とは入力に対応する望ましい出力を得るために、ニューロンの入出力特性とシナプス結合荷重を変化させることであると言える。ここで、任意の非線形関数は十分な数の中間ニューロンを用いた3層ニューラルネットワークで、有界閉集合上で一様に近似できるという関数近似定理[1]が示されている。

ところで、入力からそれに対応する望ましい出力を得るために必要とされる中間ニューロンの数は未知で

ある。そのため、学習後に中間ニューロン数をAICを利用して決定する手法[2]や、学習中に中間ニューロンを逐次追加していく手法[3]、逐次削除していく手法[4]などが提案されている。本論文では、ニューラルネットワークの初期状態に冗長なニューロンとシナプス結合荷重が存在する場合において、学習中にシナプス結合荷重間に競合を生じさせて冗長なニューロンとシナプス結合荷重を削除する手法を提案する。

一般に、動径基底関数ネットワーク(Radial Basis Function Networks: 以下, RBFN)[5]においても、冗長なニューロンを考慮することによるシナプス結合荷重の増加は、学習の遅延や過学習の問題を生じる。これらの問題を解決する方法として、望ましい出力と相関が高い入力を伝達しているシナプス結合荷重が生き残り、生き残ったシナプス結合荷重により伝達される入力に変化を受ける入力ニューロンの入出力特性のみを調節することが考えられる。そこで、本論文ではシナプス結合荷重間に入出力の相関に応じた競合が生じるようなシナプス可塑性方程式を導出する。そして、

[†] 広島大学工学部第二類(電気系), 東広島市
Faculty of Engineering, Hiroshima University, Higashi-Hiroshima-shi, 739 Japan

得られたシナプス可塑性方程式を適者生存型学習則として適用した RBFN を競合動径基底関数ネットワーク (Competitive Radial Basis Function Networks: 以下, CRBFN) として提案する。更に, 望ましい出力が動径基底関数の定数倍の足し合せで実現できる特別の場合においては, ターミナルアトラクタ [6] を利用することで望ましい時間で収束するシナプス可塑性方程式が導出できることを示す。シミュレーションでは, RBFN に比較して CRBFN が優れていることを示す。

本論文の構成は次のとおりである。2. では RBFN の概要について述べる。3. ではシナプス結合荷重間の競合を考慮したシナプス可塑性方程式の導出を行う。4. では導出されたシナプス可塑性方程式を適者生存型学習則として適用した CRBFN を提案する。5. では提案した CRBFN の性能をシミュレーションにより確認する。6. ではまとめと今後の課題について述べる。

2. 動径基底関数ネットワークの概要

RBFN は非線形関数 $\eta(\mathbf{x})$ を動径基底関数 $\xi^i(\mathbf{x})$ の足し合せで近似するニューラルネットワークである。動径基底関数としては規格化されたガウス型活性関数などが用いられる。 N 個の入力ニューロンと 1 個の出力ニューロンからなる RBFN は図 1 のような構造をもつ。 d 次元の入力ベクトル $\mathbf{x}_j \in R^d$ ($j = 1, 2, \dots, M$) はすべての入力ニューロンに入力される。第 i 入力ニューロン ($i = 1, 2, \dots, N$) はパラメータ $\phi^i \equiv [\mathbf{m}^i, \Sigma^i]$ をもつ。ここで, $\mathbf{m}^i \equiv [m_1^i, m_2^i, \dots, m_d^i]^T$ であり, $\Sigma^i \in R^{d \times d}$ である。また, Σ^i は正定値対称行列である。第 i 入力ニューロンは入力ベクトル $\mathbf{x}_j$ に対して

$$\xi^i(\mathbf{x}_j) = \exp \left\{ -\frac{1}{2}(\mathbf{x}_j - \mathbf{m}_i)^T \Sigma_i^{-1}(\mathbf{x}_j - \mathbf{m}_i) \right\} \quad (1)$$

を出力する。ここで, 添字の T はベクトルの転置を示す。出力値 $\xi^i(\mathbf{x}_j)$ はシナプス結合荷重 w^i を通して出力ニューロンへ伝達され, 出力ニューロンでこれらは足し合され

$$s(\mathbf{x}_j) = \sum_i^N w^i \xi^i(\mathbf{x}_j) \quad (2)$$

が出力される。RBFN による関数近似は 2 乗誤差評価関数

$$E(\mathbf{w}) = \frac{1}{2} \sum_j^M \{ \eta(\mathbf{x}_j) - s(\mathbf{x}_j) \}^2 \quad (3)$$

を減少させることにより実現される。ここで, $\mathbf{w} \equiv [w^1, w^2, \dots, w^N]^T \in R^N$ である。つまり, RBFN が学習により獲得しなければならないのは, 第 i 入力ニューロンのシナプス結合荷重 w^i , パラメータ $\mathbf{m}^i$ ならびにパラメータ Σ^i である。学習アルゴリズムに Delta ルール [7] を適用することで

$$\begin{aligned} \frac{dw^i}{dt} &= -\Delta \frac{\partial E(\mathbf{w})}{\partial w^i} \\ &= \Delta \sum_j^M \{ \eta(\mathbf{x}_j) - s(\mathbf{x}_j) \} \xi^i(\mathbf{x}_j), \end{aligned} \quad (4)$$

$$\begin{aligned} \frac{dm_1^i}{dt} &= -\Delta \frac{\partial E(\mathbf{w})}{\partial m_1^i} \\ &= \Delta \sum_j^M \{ \eta(\mathbf{x}_j) - s(\mathbf{x}_j) \} w^i \xi^i(\mathbf{x}_j) \\ &\quad \times \frac{(x_j - m_1^i)}{\sigma^{i2}}, \end{aligned} \quad (5)$$

$$\begin{aligned} \frac{d\sigma^i}{dt} &= -\Delta \frac{\partial E(\mathbf{w})}{\partial \sigma^i} \\ &= \Delta \sum_j^M \{ \eta(\mathbf{x}_j) - s(\mathbf{x}_j) \} w^i \xi^i(\mathbf{x}_j) \\ &\quad \times \frac{(x_j - m_1^i)^2}{\sigma^{i3}} \end{aligned} \quad (6)$$

が得られる。但し, 簡単のため $d = 1$ とし $\Sigma^i = \sigma^i$ とした。 Δ は適当な正の定数である。ここで, RBFN についても 3 層ニューラルネットワークと同様な関数近似定理 [8] が示されていることを付記しておく。

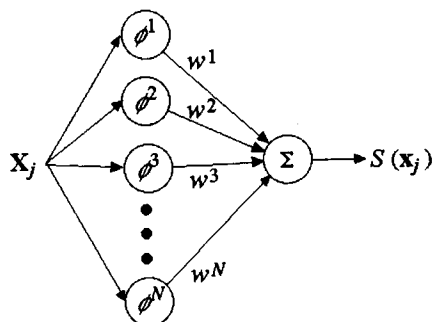


図 1 RBFN の構造
Fig.1 The structure of the RBFN.

ところで、未知の非線形関数を近似するために必要な動径基底関数の個数をあらかじめ知ることができない。そのため、RBFN では初期状態においていくつかの冗長な入力ニューロンを備えている。一般に入力ニューロン数 N は入力ベクトル数 M に比較して少なく設定される ($N \leq M$)。また、入力ニューロンのパラメータ $\mathbf{m}^i$ の決定法には入力ベクトルの部分集合から与える手法と、入力ベクトルには無関係として与える手法がある。パラメータ $\mathbf{m}^i$ を入力ベクトルとは無関係とし非線形関数を近似する手法には、式 (3) の 2 乗誤差評価のこう配に基づき Delta ルールを適用する最急降下手法と、 k -mean 法により入力ベクトル群をクラスタリングして入力ニューロンのパラメータ $\mathbf{m}^i$ を定めると共にパラメータ Σ^i を求め、最小 2 乗法によりシナプス結合荷重を決定するハイブリッド手法 [9] などがある。パラメータ $\mathbf{m}^i$ を入力ベクトルの部分集合から与える手法 [10] では、2 乗誤差評価の降下に大きく寄与するシナプス結合荷重から必要な数だけを取り出すことで RBFN の低次元化を行っている。

しかし、これらの RBFN の低次元化において必要となるニューロン数を決定する基準が確立されていない。そこで、本論文では RBFN の低次元化に適用できる適者生存型学習則を提案する。この手法はシナプス結合荷重間に競合を生じさせることにより、学習しながら自然な低次元化を行うことが可能となる。

3. 適者生存型学習則の導出

3.1 シナプス可塑性方程式の導出

ニューロンは他のニューロンにシナプス間隔を通して神経伝達物質を放出することにより情報を伝達する。第 i ニューロンの第 k 微小領域に付着する第 j ニューロンからのシナプス結合荷重 w_{ik}^j のシナプス可塑性を記述する方程式を

$$\frac{dw_{ik}^j}{dt} = \alpha_{ik}^j w_{ik}^j + g_{ik}^j w_{ik}^j + f_{ik}^j \quad (7)$$

で与える。シナプス結合荷重 w_{ik}^j は Dale 則により、第 j ニューロンが興奮性であるなら正、抑制性であるなら負の値をとる。 α_{ik}^j は内的自然増加率、 g_{ik}^j はシナプス結合荷重 w_{ik}^j における神経成長因子 (Nerve Growth Factor: 以下、NGF) の摂取量である。 f_{ik}^j は揺らぎを表す。内的自然増加率 α_{ik}^j を

$$\alpha_{ik}^j = \int_{x \in B_{ik}} \eta_{ik}(x) \xi_{ik}^j(x) dx \quad (8)$$

で定義する。ここで、 B_{ik} は第 i ニューロンの第 k 微小領域を表す。 ξ_{ik}^j は微小領域 B_{ik} における第 j ニューロンのシナプス前終末発火頻度を表し、これが作用する第 i ニューロンのニューロン膜におけるシナプス後発火頻度を η_{ik} で表す。内的自然増加率 α_{ik}^j は Hebb 則を表すことがわかる。また、NGF の摂取量 g_{ik}^j の時間変化が

$$\frac{dg_{ik}^j}{dt} = \epsilon_{ik}^j (G_{ik} - g_{ik}^j) - \sum_{h=1}^{N_{ik}} \beta_{ik}^{jh} \mu^h w_{ik}^h \quad (9)$$

に従うものとする。ここで、パラメータ ϵ_{ik}^j は正の定数である。 G_{ik} は細胞体から微小領域 B_{ik} への NGF の供給速度であり、ニューロンの特性により決定される定数である。Dale 則を考慮した識別子 μ^j は第 j ニューロンが興奮性であるなら 1、抑制性であるなら -1 となる。 N_{ik} は微小領域 B_{ik} に付着するシナプス前終末の数である。競争係数 $\beta_{ik}^{jh}/\epsilon_{ik}^j$ は、第 j ニューロンと第 h ニューロンが同時に NGF を消費することによる競合の効果であり、

$$\gamma_{ik}^{jh} = \frac{\beta_{ik}^{jh}}{\epsilon_{ik}^j} = \int_{x \in B_{ik}} \xi_{ik}^j(x) \xi_{ik}^h(x) dx \quad (10)$$

で定義する。ここで、シナプス結合荷重 w_{ik}^j の時間変化に対し、NGF の摂取量 g_{ik}^j の時間変化が無視できるものとする

$$g_{ik}^j = G_{ik} - \sum_{h=1}^{N_{ik}} \gamma_{ik}^{jh} \mu^h w_{ik}^h \quad (11)$$

を得る。その結果、Dale 則を考慮したシナプス可塑性方程式として

$$\frac{dw_{ik}^j}{dt} = \left(\alpha_{ik}^j + G_{ik} - \sum_{h=1}^{N_{ik}} \gamma_{ik}^{jh} \mu^h w_{ik}^h \right) w_{ik}^j + f_{ik}^j \quad (12)$$

が導かれる。

簡単のため、以下では微小領域 B_{ik} に着目し、NGF の摂取量が一定 ($G_{ik} = 0$) で揺らぎのない場合 ($f_{ik}^j = 0$) を考える。このとき、正定関数 $V(\mathbf{w}_{ik})$ として

$$V(\mathbf{w}_{ik}) = \frac{1}{2} \int_{x \in B_{ik}} \{\eta_{ik}(x) - s_{ik}(x)\}^2 dx \quad (13)$$

を定義する。ここで、 $\mathbf{w}_{ik} \equiv [w_{ik}^1, w_{ik}^2, \dots, w_{ik}^{N_{ik}}]^T \in R^{N_{ik}}$ であり、

$$s_{ik}(x) = \sum_{j=1}^{N_{ik}} \mu^j w_{ik}^j \xi_{ik}^j(x) \quad (14)$$

である。この式の右辺は微小領域 B_{ik} に付着しているすべてのシナプス前終末のシナプス前発火頻度 $\xi_{ik}^j(x)$ と、Dale 則を考慮したシナプス結合荷重 $\mu^j w_{ik}^j$ との積の総和であるので、 $s_{ik}(x)$ を神経伝達物質放出量と呼ぶこととする。正定関数 $V(\mathbf{w}_{ik})$ はシナプス後発火頻度 $\eta_{ik}(x)$ と神経伝達物質放出量 $s_{ik}(x)$ の差を表す指標である。シナプス後発火頻度 $\eta_{ik}(x)$ が時間に依存しないと仮定すると、その時間変化が

$$\begin{aligned} \frac{dV(\mathbf{w}_{ik})}{dt} &= \sum_{j=1}^{N_{ik}} \frac{\partial V(\mathbf{w}_{ik})}{\partial w_{ik}^j} \frac{dw_{ik}^j}{dt} \\ &= - \sum_{j=1}^{N_{ik}} \mu^j \left[\int_{x \in B_{ik}} \eta_{ik}(x) \xi_{ik}^j(x) dx \right. \\ &\quad \left. - \sum_{h=1}^{N_{ik}} \int_{x \in B_{ik}} \xi_{ik}^j(x) \xi_{ik}^h(x) dx \mu^h w_{ik}^h \right] \frac{dw_{ik}^j}{dt} \\ &= - \sum_{j=1}^{N_{ik}} \mu^j w_{ik}^j \left(\alpha_{ik}^j - \sum_{h=1}^{N_{ik}} \gamma_{ik}^{jh} \mu^h w_{ik}^h \right)^2 \\ &\leq 0 \end{aligned} \quad (15)$$

となるため、正定関数 $V(\mathbf{w}_{ik})$ が Lyapunov 関数となることがわかる。

これから、シナプス後発火頻度 $\eta_{ik}(x)$ を入力 x に対する望ましい出力、シナプス前発火頻度 $\xi_{ik}^j(x)$ を動径基底関数であるとみなすことで、シナプス結合荷重 w_{ik}^j は競合を行いながら RBFN と同様に望ましい出力と動径基底関数の 2 乗誤差評価関数を減少させることが示される。本論文では、式 (7) のシナプス可塑性方程式を適者生存型学習則ということとする。

3.2 特異点を利用したシナプス可塑性方程式

シナプス可塑性方程式に従うシナプス結合荷重では、競合に負けたシナプス結合荷重は平衡状態では 0 になる。ところが、平衡解への漸近は指数関数的に行われるので原理的には有限時間で平衡状態へ到達することはできない。そこで、望ましい出力が動径基底関数を定数倍して足し合わせることで実現できる特別の場合に、あらかじめ与えられた時刻 t^* で平衡解へ収束できるように修正されたシナプス可塑性方程式を導出する。

まず、望ましい時刻 t^* で収束するシナプス結合荷

重の時間変化を Lyapunov 関数を用いて規定する。そこで、Lyapunov 関数の時間変化を

$$\frac{dV(\mathbf{w}_{ik})}{dt} = - \frac{V(\mathbf{w}_{ik})^R V(\mathbf{w}_{ik})^{1/r}}{Rt^*} \quad (16)$$

で定義する [6]。ここで、 r は任意の奇数であり、 $R = (r-1)/r$ である。 $\mathbf{w}_{ik}^0$ は $\mathbf{w}_{ik}$ の初期値である。このような定義が可能となったのは、Dale 則を考慮したシナプス可塑性方程式に対する Lyapunov 関数が導出され、望ましい出力が動径基底関数を定数倍して足し合わせることで実現できる特別の場合を考えているからである。シナプス可塑性方程式は

$$\frac{dw_{ik}^j}{dt} = \Delta_{ik} \left(\alpha_{ik}^j - \sum_{h=1}^{N_{ik}} \gamma_{ik}^{jh} \mu^h w_{ik}^h \right) w_{ik}^j \quad (17)$$

とすることができる。このとき、Lyapunov 関数の時間変化は

$$\begin{aligned} \frac{dV(\mathbf{w}_{ik})}{dt} &= -\Delta_{ik} \sum_{j=1}^{N_{ik}} \mu^j w_{ik}^j \\ &\quad \times \left(\alpha_{ik}^j - \sum_{h=1}^{N_{ik}} \gamma_{ik}^{jh} \mu^h w_{ik}^h \right)^2 \end{aligned} \quad (18)$$

となる。ここで、 Δ_{ik} は式 (16) と式 (18) から決定することができる。以上のことから、望ましい時刻 t^* で平衡解へ収束する修正されたシナプス可塑性方程式を

$$\begin{aligned} \frac{dw_{ik}^j}{dt} &= \frac{(\alpha_{ik}^j - \sum_{h=1}^{N_{ik}} \gamma_{ik}^{jh} \mu^h w_{ik}^h) w_{ik}^j}{\sum_{j=1}^{N_{ik}} \mu^j w_{ik}^j (\alpha_{ik}^j - \sum_{h=1}^{N_{ik}} \gamma_{ik}^{jh} \mu^h w_{ik}^h)^2} \\ &\quad \times \frac{V(\mathbf{w}_{ik})^R V(\mathbf{w}_{ik})^{1/r}}{Rt^*} \end{aligned} \quad (19)$$

で定義することができる。

4. 競合動径基底関数ネットワークの提案

サンプルされた入力ベクトル $\mathbf{x}_j \in R^d$ ($j = 1, 2, \dots, M$) ならびにそれに対する望ましい出力 $\eta(\mathbf{x}_j)$ から、それらを滑らかにつなぐ非線形関数を再構成するための CRBFN の構造は RBFN と同様に図 1 で表せる。この場合、出力ニューロンと微小領域がそれぞれ一つであるので、3. の議論で用いた記号の右下の添字を省略することができる。例えば、シナプス結合荷重 w_{ik}^j は w^j 、内的自然増加率 α_{ik}^j は α^j 、競争係数 γ_{ik}^{jh} は γ^{jh} となる。出力ニューロンにはすべての入力ニューロンからのシナプス結合荷重が付着してい

るので $N_i = N$ となる。また、CRBFN における入力ニューロンの動作規則は式 (1) で与えられ、出力ニューロンの出力は式 (2) で与えられる。

再び、CRBFN における望ましい出力と動径基底関数との 2 乗誤差評価関数を

$$E(\mathbf{w}) = \frac{1}{2} \sum_{j=1}^M \left\{ \eta(\mathbf{x}_j) - \sum_{i=1}^N w^i \xi^i(\mathbf{x}_j) \right\}^2 \quad (20)$$

で与える。ここで、 M はサンプルされた標本点の個数であり、 $\mathbf{w} \equiv [w^1, w^2, \dots, w^N]^T \in R^N$ である。

CRBFN の学習則は 2 乗誤差評価関数を減少させるために第 i 入力ニューロンのシナプス結合荷重 w^i 、パラメータ $\mathbf{m}^i$ の第 k 要素 m_k^i ならびにパラメータ Σ^i の第 k 行 l 列要素 σ_{kl}^i に対し

$$\frac{dw^i}{dt} = \left(\alpha^i - \sum_{h=1}^N \gamma^{ih} w^h \right) w^i \quad (21)$$

$$\frac{dm_k^i}{dt} = -\Delta \frac{\partial E(\mathbf{w})}{\partial m_k^i} \quad (22)$$

$$\frac{d\sigma_{kl}^i}{dt} = -\Delta \frac{\partial E(\mathbf{w})}{\partial \sigma_{kl}^i} \quad (23)$$

で与えられる。ここで、 Δ は適当な正の定数である。内的自然増加率 α^i ならびに競争係数 γ^{ih} は

$$\alpha^i = \sum_{j=1}^M \eta(\mathbf{x}_j) \xi^i(\mathbf{x}_j) \quad (24)$$

$$\gamma^{ih} = \sum_{j=1}^M \xi^i(\mathbf{x}_j) \xi^h(\mathbf{x}_j) \quad (25)$$

で定義される。学習中に $w^i \approx 0$ となった第 i シナプス結合荷重は消滅したものととして、生き残っているシナプス結合荷重、ならびにそれらにより伝達される入力が変化を受ける入力ニューロンのパラメータについてのみ学習を続ける。

ここで、ある確率密度関数に従い発生している M 個の標本点から元の確率密度関数を推定する問題を考える。サンプルされた標本点は CRBFN の入力ベクトル $\mathbf{x}_j$ として用いられる。ところが、入力ベクトルに対する望ましい出力 $\eta(\mathbf{x}_j)$ が不明である。そこで、各標本点に観測誤差を考慮して平均ベクトル $\mathbf{x}_j$ 、共分散行列 $\tilde{\Sigma}$ をもつ正規確率密度関数

$$N(\mathbf{x}_j, \tilde{\Sigma}) \equiv \frac{1}{\sqrt{(2\pi)^d |\tilde{\Sigma}|}}$$

$$\times \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{x}_j)^T \tilde{\Sigma}^{-1} (\mathbf{x} - \mathbf{x}_j) \right\} \quad (26)$$

を考える。ここで、 $|\tilde{\Sigma}|$ は共分散行列 $\tilde{\Sigma}$ の行列式を表す。そして、望ましい出力 $\eta(\mathbf{x})$ として

$$\eta(\mathbf{x}) = \frac{1}{M} \sum_{j=1}^M N(\mathbf{x}_j, \tilde{\Sigma}) \quad (27)$$

を考えることとする。このとき、内的自然増加率 α^i ならびに競争係数 γ^{ih} の定義式にある、サンプルされた標本点 $\mathbf{x}_j$ についての足し合わせは

$$\sum_{j=1}^M f(\mathbf{x}_j) \rightarrow \int f(\mathbf{x}) d\mathbf{x} \quad (28)$$

のように書き換えることができ、積分で与えることが可能となる。

5. シミュレーションならびに考察

5.1 関数近似問題への適用

簡単な 1 次元非線形関数 $f(x)$ の近似に適用することによって、提案された CRBFN が RBFN に比較して高速に関数近似ができることを示す。ここで、1 次元非線形関数 $\eta(x)$ として

$$\eta(x) = 3N(-1.5, 1) + 2N(1, 0.5) \quad (29)$$

を考える。サンプルは $x \in [-5, 5]$ の間から 0.1 ずつ 101 個とり、教師信号として $(x_j, \eta(x_j))$ $j = 1, 2, \dots, 101$ を与える (図 2 参照)。RBFN と CRBFN とともに初期に設定した入力ニューロン数は 6 個であり、パラメータは $[m_1^1, m_1^2, m_1^3, m_1^4, m_1^5, m_1^6]^T =$

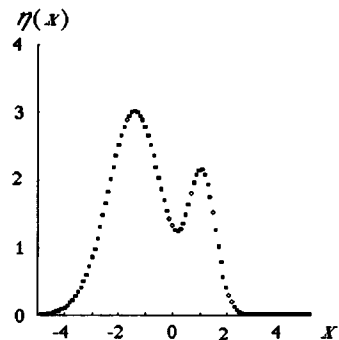


図 2 教師信号の分布図
Fig. 2 The distribution of the teacher signals.

表1 学習前・学習後のシナプス結合荷重

Table 1 The values of synaptic weights before learning and after learning.

	w^1	w^2	w^3	w^4	w^5	w^6
初期値	0.431	0.898	0.324	0.927	0.389	0.161
RBFN	0.290	1.880	2.065	1.054	0.109	0.109
CRBFN	0.000	0.000	2.990	1.986	0.000	0.000

表2 学習が収束するまでの学習回数 (式 (21) の場合)

Table 2 The iteration numbers of learning (the case using Eq. (21)).

	全体ループ	局所ループ
RBFN	219	10723224
CRBFN	11	232710

$[-5, -3, -1, 1, 3, 5]^T$ ならびに $\sigma^i = 1.5$ ($i = 1, 2, \dots, 6$) を与えた。シナプス結合荷重の初期値はともに一様乱数で与えた。学習が収束したときのシナプス結合荷重の値を表1に示す。学習の収束判定は $E(\mathbf{x}) \leq 0.0001$ とした。これより、RBFNでは6個すべての入力ニューロンで関数近似を行うが、CRBFNでは第2, 第3シナプス結合荷重に関係した必要最低限の入力ニューロンで関数近似を行うことがわかる。これは、これらのシナプス結合荷重が望ましい出力と相関が高い入力を伝達していることに起因する。表中の0.000は小数点第5位以下の数値をとっており、消滅したと考えることができるシナプス結合荷重の値である。表2は学習が収束するまでの学習回数である。但し、局所ループはシナプス結合荷重 w^i ならびにパラメータ m_1^i と σ^i を一つ更新するたびに回数を数えている。全体ループの回数はシナプス結合荷重ならびにパラメータをひとつと更新するたびに回数を数えており、途中で学習が収束したときはその回数は数えないものとする。プログラムで異なる点はシナプス結合荷重の更新則が式(4)で与えられるか、あるいは式(21)で与えられるかの違いだけであるので、この学習回数の差はシナプス結合荷重間に競合を考慮したことにより生じていると考えられる。

以上のことから、RBFNではすべての入力ニューロンを用いて関数近似を行うため、シナプス結合荷重 w^i とパラメータ m_1^i と σ^i の調節に多くの学習回数が必要とするが、CRBFNでは必要最低限の入力ニューロンについてのみパラメータの調節をするため、少ない

表3 学習が収束するまでの学習回数 (式 (19) の場合)

Table 3 The iteration numbers of learning (the case using Eq. (19)).

	全体ループ	局所ループ
RBFN	0	912
CRBFN	0	16344

学習回数で学習することができ、過学習を回避できることがわかる。

次に、ターミナルアトラクタの概念を用いて望ましい学習回数 t^* で学習を収束させることができる特別な場合を考える。そこで、1次元非線形関数 $\eta(x)$ として

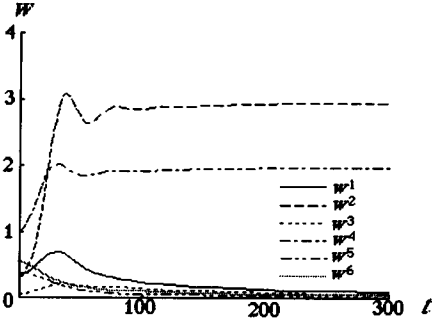
$$\eta(x) = 3N(-3, 1) + 2N(1, 1) \quad (30)$$

を考える。教師信号と初期の設定ならびに学習の収束判定は上で説明したとおりである。表3は学習が収束するまでの学習回数である。ここでも、局所ループはシナプス結合荷重 w^i ならびにパラメータ m_1^i と σ^i を一つ更新するたびに回数を数えており、全体ループの回数はシナプス結合荷重ならびにパラメータをひとつと更新するたびに回数を数えている。全体ループの回数がRBFNとCRBFNともに0となっているのは、学習がシナプス結合荷重の更新のみで収束しているからである。

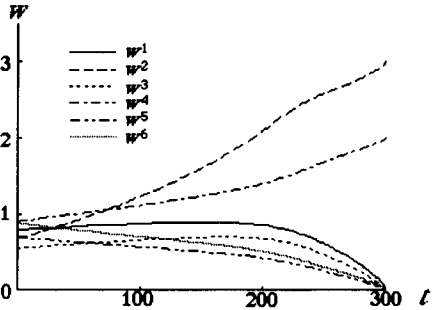
このように、シナプス結合荷重 w^i のみの調節で学習できる特別な関数近似の場合は、CRBFNに比較してRBFNの方が学習回数が少なくてすむことがわかる。しかし、シナプス結合荷重の更新則を式(4)若しくは式(21)で与える限り、学習が終了するまでの学習回数を指定することはできない。図3(a)にシナプス結合荷重の更新則を式(21)で与えたときのシナプス結合荷重の変化を示す。学習は300回以降で終了しているのであるが、平衡解へ徐々に漸近していく様子が図からわかる。図3(b)にシナプス結合荷重の更新則を式(19)で与えたときのシナプス結合荷重の変化を示す。ここでは、望ましい学習回数 t^* を300で与えた。図から望ましい学習回数で学習が終了していることがわかる。

5.2 標本点から確率密度関数の推定

ここで、10個の入力ニューロンをもつRBFNとCRBFNにより、2次元確率密度関数 $p(\mathbf{x})$ に従い発生している M 個の標本点から元の確率密度関数の推定を行う。確率密度関数 $p(\mathbf{x})$ を



(a) Applying Eq. (21) to the synaptic plasticity equation.



(b) Applying Eq. (19) to the synaptic plasticity equation.

図3 シナプス結合荷重の変化

Fig.3 The changes of the synaptic weights.

$$p(\mathbf{x}) = \sum_j^3 p_j N(\mathbf{m}'_j, \Sigma'_j) \quad (31)$$

とする。但し、係数 $\mathbf{p} \equiv [p_1, p_2, p_3]^T$ は

$$\sum_{j=1}^3 p_j = 1 \quad (32)$$

を満たす。ここでは、 $\mathbf{p} = [0.2, 0.3, 0.5]^T$ で与えた。平均ベクトル $\mu'_j \in R^2$ 、共分散行列 $\Sigma'_j \in R^{2 \times 2}$ はそれぞれ表4のように与えた。この確率密度関数の様子を図4に示す。発生させた標本点を2次元平面 $\mathbf{x} \in R^2$ に示したのが図5である。発生させる標本点が多いほどよりよい推定が可能になると考えられるが、ここでは1000点を発生させた。標本点が十分にもとの確率密度関数を反映していれば、この確率密度関数は3個の正規確率密度関数の足し合を規格化することで近似することができる。つまり、3個の入力ニューロンのみで学習が可能であり、残りの7個の入力ニューロンが冗長であると言える。各入力ニューロンのパラメータ $\mathbf{m}^i$ ($i = 1, 2, \dots, 10$) はランダムに発

表4 確率密度関数の平均ベクトルと共分散行列
Table 4 The mean vectors and covariance matrices.

	μ'_1	μ'_2	μ'_3
平均ベクトル	$[2, 2]^T$	$[-1, 1]^T$	$[1, -1]^T$
	Σ'_1	Σ'_2	Σ'_3
共分散行列	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

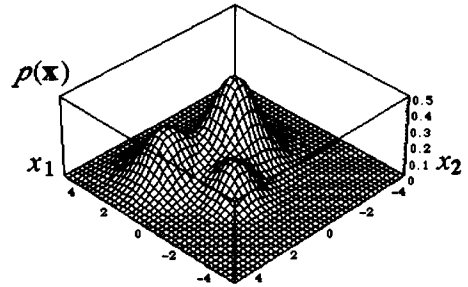


図4 母集団の確率密度関数

Fig.4 The probability density function of the population.

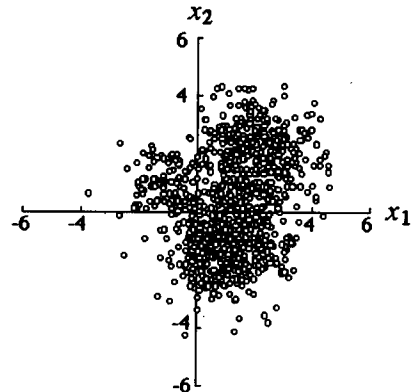


図5 標本点の分布

Fig.5 The distribution of sampling points.

生させる。

図6は各入力ニューロンの受けもつ領域をパラメータ $\mathbf{m}^i$ を中心に半径2の円で描いた概念図である。図中の白抜きの□は各入力ニューロンのパラメータ $\mathbf{m}^i$ を表す。これから、入力ニューロンは初期において、近接した領域を受けもっていることがわかる。このような場合、RBFNでは入力ニューロンのシナプス結合荷重とパラメータが同程度で更新されるため、同じような働きをする冗長な入力ニューロンが複数できるこ

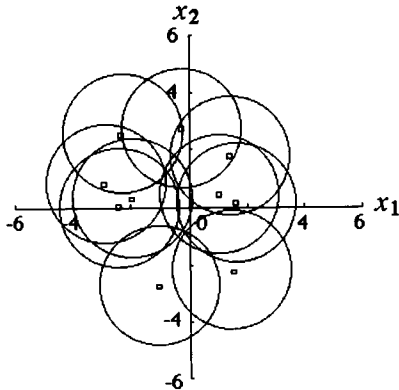


図6 初期の入力ニューロン領域の概念図

Fig. 6 The conceptual figure of the initial distributions of input neurons.

表5 生き残った入力ニューロンのパラメータ Σ^i Table 5 The parameters Σ^i of the surviving input neurons.

共分散行列	Σ^3	Σ^7	Σ^8
	$\begin{bmatrix} 1.30 & 0.02 \\ 0.02 & 1.08 \end{bmatrix}$	$\begin{bmatrix} 0.55 & -0.01 \\ -0.01 & 0.56 \end{bmatrix}$	$\begin{bmatrix} 1.29 & 0.05 \\ 0.05 & 1.53 \end{bmatrix}$

となる。これに対し、CRBFNでは必要最低限の入力ニューロンが生き残り、学習が収束するまでの学習回数を大幅に減らすことができる。CRBFNに与えられる望ましい出力 $\eta(\mathbf{x})$ は、式(27)により与えることができる。各標本点の観測誤差を考慮する共分散行列 $\hat{\Sigma}$ の第 k 行 l 列要素 $\hat{\sigma}_{kl}$ は、 $\hat{\sigma}_{11} = \hat{\sigma}_{22} = 2/3$ ならびに $\hat{\sigma}_{12} = \hat{\sigma}_{21} = 0$ とした。

学習により生き残ったのは第3, 7, 8の3個のシナプス結合荷重である。表5はCRBFNの生き残ったシナプス結合荷重をもつ入力ニューロンのパラメータ Σ^i の学習後の値である。学習前の値はすべて単位行列である。これより、第7入力ニューロンのパラメータ Σ^7 が Σ^3 の値に近づいていることがわかる。図7はCRBFNの生き残ったシナプス結合荷重をもつ入力ニューロンのパラメータ $\mathbf{m}^i$ の学習中の動きである。図中の白抜ききの□は各入力ニューロンの初期のパラメータ $\mathbf{m}^i$ を表す。これらより、第3, 8入力ニューロンのパラメータ ϕ^3, ϕ^8 がそれぞれ $(\mu_1', \Sigma_1'), (\mu_3', \Sigma_3')$ を学習により獲得していることがわかる。CRBFNの学習により獲得されたシナプス結合荷重、パラメータを用いて推定された確率密度関数を図8に示す。標本点を発生させた確率密度関数 $p(\mathbf{x})$ の特徴をうまく推

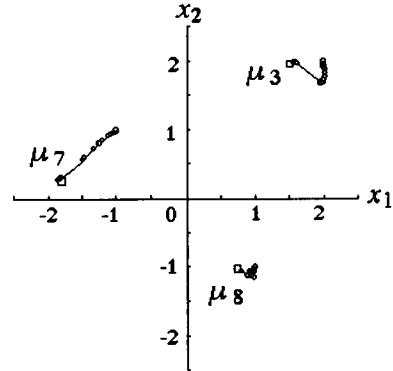
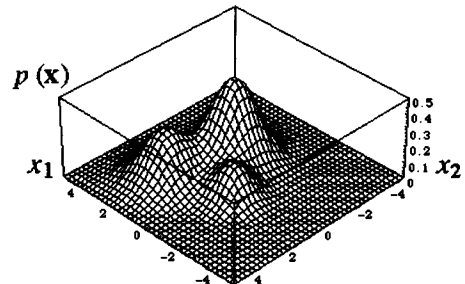
図7 生き残った入力ニューロンのパラメータ $\mathbf{m}_i$ の学習中の軌跡Fig. 7 The trajectories of the parameters $\mathbf{m}_i$ of survival input neurons during learning.

図8 CRBFNにより推定された確率密度関数

Fig. 8 The probability density function estimated by the CRBFN.

定していることがわかる。

6. む す び

本論文では、ニューロンの微小領域において生じている NGF の競合を考慮したシナプス可塑性方程式を導出した。更に、このシナプス可塑性方程式を適者生存型学習則として RBFN に適用した CRBFN を提案した。CRBFN ではシナプス結合荷重間に競合を生じさせ、学習に必要なニューロンのみが自然に生き残り、学習の効率化を図ることができる。シミュレーションでは、シナプス結合荷重の調節のみで学習できる特別な関数近似の場合を除いて、CRBFN は RBFN に比較して学習回数が大幅に少なくてすむことが示された。また、シナプス結合荷重の調節のみで学習できる特別な場合には、望ましい学習回数で学習が終了させることができることが示された。更に、未知の確率密度関数に従い分布する標本点が与えられたとき、冗長

なニューロンをもつ CRBFN が学習により不必要なニューロンを排除して、最適なニューロン数でもとの確率密度関数を推定できることが示された。

今後の課題は、適者生存型学習則となるシナプス可塑性方程式を、その他のニューラルネットワークにも適用できるように拡張していくことである。

謝辞 本研究のシミュレーションを担当して頂いた広島大学大学院学生の佐々木浩二氏に感謝致します。

文 献

- [1] K. Funahashi, "On the approximate realization of continuous mappings by neural networks," Neural Networks, vol.2, pp.148-153, 1989.
- [2] D.B. Fogel, "An information criterion for optimal neural network selection," IEEE Trans. Neural Networks, vol.2, pp.490-497, 1991.
- [3] Z. Wang, C.D. Massimo, M.T. Tham, and A.J. Morris, "A procedure for determining the topology for multi-layer feedforward networks," Neural Networks, vol.7, pp.291-300, 1994.
- [4] A.S. Weigend, D.E. Rumelhart, and B.A. Huberman, "Generalization by weight elimination applied to currency exchange rate prediction," Proc. of the Int. Joint Conf. on Neural Networks, Singapore, vol.3, pp.2374-2379, 1990.
- [5] M.J.D. Powell, "Radial basis functions for multivariable interpolation," A Review, in IMA Conference on Algorithms for the Approximation of Functions and Data, pp.143-167, RMCS, Shrivenham, 1985.
- [6] M. Zak, "Terminal attractors in neural networks," Neural Networks, vol.2, pp.259-274, 1989.
- [7] B. Widrow and R. Winter, "Neural nets for adaptive filtering and adaptive pattern recognition," IEEE Computer, vol.21, pp.25-39, 1988.
- [8] J. Park and I.W. Sandberg, "Universal approximation using radial basis function networks," Neural Computation, vol.3, pp.246-257, 1991.
- [9] J. Moody and C. Darken, "Fast learning in networks of locally tuned processing units," Neural Computation, vol.1, pp.281-294, 1989.
- [10] S. Chen, C.F.N. Cowan, and P.M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," IEEE Trans. Neural Networks, vol.2, no.2, pp.302-309, 1991.
- [11] 奥原浩之, 尾崎俊治, "Dale 則を考慮したシナプス可塑性方程式の解析," システム制御情報学会論文誌, vol.8, no.12, pp.718-720, 1995.

(平成 9 年 2 月 10 日受付, 6 月 16 日再受付)



奥原 浩之 (正員)

平 8 広島大大学院工学研究科システム工学専攻博士課程後期了。同年, 広島大学工学部助手。博士 (工学)。協同現象の数理に興味をもち, 自律分散システムの研究に従事。システム制御情報学会, 日本 OR 学会各会員。



尾崎 俊治 (正員)

昭 45 京大大学院博士課程了。同年広島大学工学部講師, 同年助教授, 昭 61 教授。工博。信頼性工学およびシステム工学の研究に従事。昭 51~52 マンチェスター大学研究員。昭 45~47 南カルフォルニア大学研究員。日本 OR 学会第 4 回大西文献賞受賞。著書「応用確率論」, 「Stochastic System Reliability Modeling」など。情報処理学会, 日本 OR 学会各会員。