

Original

A Study on a Method Improving Efficiency of Search Based on Conjugate Gradient Method

Takahiro YOSHIDA,* Masayuki GOTOH** and Nobuhiko TAWARA*

Abstract

BP (Back Propagation) learning algorithm based on Conjugate Gradient Method is an effective method for high-speed learning. This method has originally been introduced for the problem to minimize the quadratic function and is guaranteed to converge to it in the number of times equal to the search surface dimensions. It is also applied as it is to a generalized function like BP learning algorithm, because in this case too, it can approximate locally to a quadratic function. However, it fails to converge in the number of times equal to the search dimensions in this case, requiring the step to take the second approximation process after the search dimensions reached a certain number. This step is called "Restart". However, this method occasionally falls into local minimal depending on search surface, because it locally approximates a generalized function to a quadratic function. The cause is considered that it continues searching until its number of times reaches the conventionally determined number of times before the restart, even if the approximation to a quadratic function is low in precision; therefore this restart method is considered to improve the search efficiency when the quadratic function approximation is low in precision. We therefore propose an efficient learning algorithm based on Conjugate Gradient Method. The method proposed here can decide whether the restart is needed according to new parameters which evaluate the quadratic approximation accuracy at every search point. We call the proposed method "An improved Conjugate Gradient Method".

Key words : Conjugate Gradient Method, nonlinear programming, Back Propagation, neural networks

* Musashi Institute of Technology
** Waseda University

論文

共役勾配法における探索効率向上法に関する一考察

吉田 隆 弘*, 後 藤 正 幸**, 俵 信 彦*

ニューラルネットワークのバックプロパゲーション学習（以下BP学習）に共役勾配法を適用することにより、ある程度的高速化が達成できると報告されている。しかし、共役勾配法は高速である反面、初期値に依存するという不安定さを合わせ持つ。また、従来の共役勾配法を適用したBP学習は再出発までの学習回数を一定に設定しているため、二次近似の精度が良くない場合でも再出発までの学習回数に達するまで探索をそのまま続けているので、学習の効率が悪くなるという問題もある。そこで、本研究では、二次近似の精度を各探索点で評価することにより再出発の必要性の有無の判定を行い、適応的に近似範囲を調整する改良型共役勾配法を提案する。

キーワード：共役勾配法，非線形計画法，バックプロパゲーション，ニューラルネットワーク

1. はじめに

階層型ニューラルネットワークのBP学習は、学習速度、学習の不安定性などの学習の問題が生じている〔1〕～〔10〕。学習速度の問題に対しては、最急降下法を適用した従来のBP学習にかえて、代表的な非線形最適化手法である共役勾配法を適用することにより、ある程度的高速化が達成できることが明らかとなっている〔1〕～〔3〕。

共役勾配法は、もともと二次関数の最小化問題に対して導出され、二次関数に対しては探索次元数と同じ回数で収束することが保証されている。また、この方法はBP学習のような一般の関数に対しても局所的に二次近似できるという考えからそのまま適用されているが、この場合は探索次元数と同じ回数では収束しないので、一定の回数に達したら二次近似しなおして再び探索を開始するという手続きをとることが必要になる。この手続きのことを再出発と呼んでいる。したがって、共役勾配法は再出発から再出発までの間に探索した範囲を二次近似していることになる〔6〕、〔7〕。

しかし、共役勾配法による探索は高速である反面、一般の関数を局所的に二次近似しているため、探索局面の性質によっては二次近似がうまくいかずに結果がばらつき、場合によっては最急降下法では陥らないような誤差の大きな局所解に陥ってしまうという問題がある。これは、従来は再出発までの学習回数を一定に設定しているため、二次近似の精度が良くない場合で

も再出発までの学習回数に達するまでは探索をそのまま続けていることが原因と考えられる。したがって、共役勾配法による探索では、このように二次近似の精度が原因と考えられる。したがって、共役勾配法による探索では、このように二次近似の精度が悪い場合には、その時点で再出発を行い新たに二次近似しなおしたほうが全体としての学習の効率が良くなると思われる。

そこで本研究では、二次近似の精度を各探索点で評価することにより再出発の必要性の有無を判定して、適応的に近似範囲を調整する改良した共役勾配法をBP学習に適用し、シミュレーションによってその有効性を示す。

2. 共役勾配法

まず、二次関数の最小化問題に対する共役勾配法の定義を以下に示す。

【定義1】 正定対称行列 $Q \in R^{n \times n}$ に対して、 l 個のベクトル $d^1, d^2, \dots, d^l \in R^n$ が

$$(d^i)^T Q d^j = 0, i \neq j \text{ かつ } i, j = 1, 2, \dots, l \quad (1)$$

を満たすとき、 d^1, d^2, \dots, d^l は Q に対して互いに共役であるという。ここで、 $l \leq n$ である。 □

【定義2】 $c \in R^1, b \in R^n$ で $Q \in R^{n \times n}$ を対称な正定値行列として、二次関数

$$q(x) = c + b^T x + \frac{1}{2} x^T Q x \quad (2)$$

を最小化するために、探索方向が Q に対して互いに共役になる方向に直線探索を行うアルゴリズムを共役方向法といい、その中でも目的関数の微係数を利用して共役方向を生成する方法を共役勾配法という。

共役勾配法では、適当な直線探索を行えば n 回の

* 武蔵工業大学

** 早稲田大学

受付：1996年12月19日，再受付（1回）

受理：1997年9月22日

探索で二次関数の最適解を得ることができる。

l 回目の探索点を \mathbf{x}^l とすると、次の探索点 \mathbf{x}^{l+1} を以下のように計算する。目的関数を q 、探索方向ベクトルを \mathbf{d}^l として、

$$\mathbf{x}^{l+1} = \mathbf{x}^l + \alpha^l \mathbf{d}^l \quad (l=1, 2, 3, \dots) \quad (3)$$

$$\alpha^l = \arg \min_{\alpha > 0} \{q(\mathbf{x}^l + \alpha \mathbf{d}^l)\} \quad (4)$$

$$\mathbf{d}^l = -\nabla^T q(\mathbf{x}^l) \quad (5)$$

とすれば、 l 回目の共役方向 \mathbf{d}^l は、以下の式で生成できる [11], [12]。

$$\mathbf{d}^l = -\nabla q(\mathbf{x}^l) + \beta^l \mathbf{d}^{l-1} \quad (6)$$

式(4)の計算を直線探索といい、 $\alpha^l \in \mathbf{R}^1$ をステップ幅という。 $\beta^l (l=2, \dots, n)$ の計算法としては、次のような計算法が知られている。

1) Fletcher-Reeves の公式

$$\beta^l_{FR} = \frac{\|\nabla q(\mathbf{x}^l)\|^2}{\|\nabla q(\mathbf{x}^{l-1})\|^2} \quad (7)$$

2) Polak-Ribiere-Polyak の公式

$$\beta^l_{PRP} = \frac{\nabla^T q(\mathbf{x}^l)(\nabla q(\mathbf{x}^l) - \nabla q(\mathbf{x}^{l-1}))}{\|\nabla q(\mathbf{x}^{l-1})\|^2} \quad (8)$$

3) Sorenson-Wolfe の公式

$$\beta^l_{SW} = \frac{\nabla^T q(\mathbf{x}^l)(\nabla q(\mathbf{x}^l) - \nabla q(\mathbf{x}^{l-1}))}{(\mathbf{d}^{l-1})^T (\nabla q(\mathbf{x}^l) - \nabla q(\mathbf{x}^{l-1}))} \quad (9)$$

二次形式の目的関数に対して得られた共役勾配法のアルゴリズムは、そのまま一般の関数に適用することができる。これは一般の関数でも局所的には二次近似できるという理由からである。

この場合、繰り返し計算は n 回 (探索する空間の次元数) では収束しない。そこで、一定回数繰り返しの後に探索方向ベクトルを最急降下方向にセットしなおして再び探索をするという手続きがとられ、これを再出発と呼ぶ。再出発をしないと収束が遅くなることが実験的に示されており、従来の共役勾配法ではあらかじめ設定された回数ごとに再出発を行うという方法がとられている。しかし、再出発までの回数をどのように設定すればよいかについては明らかになっていない。

3. 二次近似範囲を調整する改良型共役勾配法

従来の共役勾配法では、再出発までの学習回数を一定回数に設定しており、二次近似の精度を考慮していないために、探索の効率が悪くなる場合がある。そこで、共役勾配法の探索効率を向上させる有効な方法として、再出発の必要性を探索点ごとに逐次的に判定を行い、二次近似の精度を維持するというような方法を考えることにする。

そのためには、目的関数が二次近似できているかを評価することが必要である。そこで、目的関数を各探

索点の周りでテーラー展開することを考える。

$$q(\mathbf{x} + \mathbf{y}) = q(\mathbf{x}) + q'(\mathbf{x})\mathbf{y} + \frac{1}{2}\mathbf{y}^T q''(\mathbf{x})\mathbf{y} + O(\mathbf{y}^3) \\ \stackrel{\text{def}}{=} q_{qw}(\mathbf{y}) + O(\mathbf{y}^3) \quad (10)$$

ただし、 $O(\mathbf{y}^3)$ は三次以下の項である。

$\mathbf{y} = \alpha \mathbf{d}$ とおき、 $q_{qw}(\alpha \mathbf{d})$ を α で偏微分すると、

$$\frac{\partial q_{qw}}{\partial \alpha} = q'(\mathbf{x})\mathbf{d} + \alpha \mathbf{d}^T q''(\mathbf{x})\mathbf{d} \stackrel{\text{def}}{=} q'_{qw}(\alpha) \quad (11)$$

$$\frac{\partial^2 q_{qw}}{\partial \alpha^2} = \mathbf{d}^T q''(\mathbf{x})\mathbf{d} \stackrel{\text{def}}{=} q''_{qw}(\alpha) \quad (12)$$

となる。ここで、(11)式を $\alpha=0$ とおき、三次の項を無視することにより、最適ステップ幅の二次近似推定値

$$\hat{\alpha} = \frac{-q'(\mathbf{x})\mathbf{d}}{\mathbf{d}^T q''(\mathbf{x})\mathbf{d}} = \frac{-q'_{qw}(0)}{q''_{qw}(0)} \quad (13)$$

が得られる。齊藤、中野 [4] は、共役勾配法自体二次近似する方法であることから、正確な直線探索の代わりにこの(13)式を用いた方法を提案している。

最適ステップ幅が決定されれば、二次近似の精度をみるための評価値として、

$$\Delta_l = \frac{q(\mathbf{x}^l) - q(\mathbf{x}^l + \hat{\alpha}^l \mathbf{d}^l)}{q(\mathbf{x}^l) - q_{qw}(\hat{\alpha}^l)} \quad (14)$$

を定義することができる。これは、目的関数の減少量と目的関数を二次近似した関数の減少量の比であり、二次近似の精度が良いほど 1 に近づく。また、この値が 0 に近くなるほど二次近似で予想される目的関数減少量を達成できないことを示している。

ここで、(14)式を計算しやすいように展開を行う。二次近似の精度が良ければ、最適ステップ幅は(13)式を満たすことから、

$$q(\mathbf{x}^l) - q_{qw}(\hat{\alpha}^l \mathbf{d}^l) = -q'(\mathbf{x}^l) \hat{\alpha}^l \mathbf{d}^l \\ - \frac{1}{2} \hat{\alpha}^{l2} \mathbf{d}^{lT} q''(\mathbf{x}^l) \mathbf{d}^l \\ = \frac{\{q'_{qw}(0)\}^2}{q''_{qw}(0)} - \frac{1}{2} \frac{\{q'_{qw}(0)\}^2}{q''_{qw}(0)} \\ = \frac{1}{2} \frac{\{q'_{qw}(0)\}^2}{q''_{qw}(0)} \quad (15)$$

であるから、

$$\Delta^l = \frac{2q''_{qw}(0)\{q(\mathbf{x}^l) - q(\mathbf{x}^l + \hat{\alpha}^l \mathbf{d}^l)\}}{\{q'_{qw}(0)\}} \quad (16)$$

となる。この(16)式は(13)式の計算の際に既に求めている $q'_{qw}(0)$ と $q''_{qw}(0)$ から計算できるので、比較的効率よく計算が可能な形となっている [4]。

以上より、(16)式に従って各探索点で Δ^l の計算を行っていけば、二次近似の精度をみながらの探索が可能となる。そこで、この二次近似の精度をみるための評価値 Δ^l がある一定の値を下回れば二次近似の精度が悪く、探索の効率が悪化しているとみなすことができるので、この時点で再出発を行うことにする。

4. 階層型ニューラルネットワークモデル

$\{(x_1, y_1), \dots, (x_m, y_m)\}$ を事例集合, $x_i = (x_{i0}, \dots, x_{in})$ をニューラルネットへの n 次元入力ベクトル, y_i をニューラルネットからの目標出力値とする. また, 三層ニューラルネットにおいて, h を中間ユニット数, $w_j (j \geq 1)$ を入力ユニットから中間ユニットへのウエイト, $w_0 = (w_{00}, \dots, w_{0h})^T$ を中間ユニットから出力ユニットへのウエイトとする. ここで, w_{j0} はバイアスを表し, $x_{i0} = 1$ とする. 以下では, (w_0, \dots, w_h) を w と略記する. ニューラルネットの学習とは,

$$E_i(w) = y_i - \left\{ w_{00} + \sum_{j=1}^h w_{0j} \sigma(w_j^T x_i) \right\} \quad (17)$$

とおけば,

$$E(w) = \frac{1}{2} \sum_{i=1}^m E_i(w)^2 \quad (18)$$

を最小化する問題として定義される. ここで, $\sigma(z)$ は適当な非線形関数であり, 通常以下のように sigmoid 関数を使用する.

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (19)$$

本研究では, (1), (2) 式で定義されるニューラルネットの学習について考える.

5. 共役勾配法による BP 学習

5.1 BP 学習

(17), (18) 式で定義されるニューラルネットにおいて, 各ウエイトに対する勾配ベクトルは,

$$\frac{\partial E(w)}{\partial w_{0j}} = - \sum_{i=1}^m E_i(w) \sigma_{ij} \quad (20)$$

$$\frac{\partial E(w)}{\partial w_{jp}} = - \sum_{i=1}^m E_i(w) w_{0j} \sigma_{ij}' x_{ip} \quad (21)$$

となる. ここで, $\sigma_{ij} (= \sigma(w_j^T x_i))$ は, 各事例 i に対する中間ユニット j の出力値であり, σ_{ij}' はその一回微分値を表し, $\sigma_{i0} = 1$ と定義する. なお, sigmoid 関数の場合には,

$$\sigma_{ij}' = \sigma_{ij}(1 - \sigma_{ij}) \quad (22)$$

という関係が成り立つ.

BP においてウエイトの更新は以下のように行う.

$$w(t+1) = w(t) + \Delta w(t) \quad (23)$$

t は更新回数で $\Delta w(t)$ はウエイトの更新量である. $\Delta w(t)$ は,

$$\Delta w(t) = \alpha(t) d(t) \quad (24)$$

に従って求める. ただし, $\alpha(t)$ はステップ幅, $d(t)$ は更新方向ベクトルである.

5.2 従来の共役勾配法による BP 学習

共役勾配法による BP 学習では, ウエイトの更新は

以下のように行う.

$$\begin{aligned} \Delta w(t) &= \alpha(t) \left\{ -\nabla E(t) + \beta(t) \frac{\Delta w(t-1)}{\alpha(t-1)} \right\} \\ &= \alpha(t) d(t) \end{aligned} \quad (25)$$

$-\nabla E(t)$ は各学習パターンへの二乗誤差に対する最急降下方向の和とし, BP アルゴリズムで生成する. $\beta(t)$ は以下の Polak-Ribiere-Polyak の公式で計算する.

$$\beta(t)_{PRP} = \frac{\nabla^T E(t) (\nabla E(t) - \nabla E(t-1))}{\|\nabla E(t-1)\|^2} \quad (26)$$

また, 本研究では, 共役方向へのステップ幅を決める計算方法として, (13) 式をもとにして $E(w + \alpha d)$, $E_i(w + \alpha d)$ をそれぞれ $e(\alpha)$, $e_i(\alpha)$ で表せば, 最適ステップ幅の推定値は,

$$\hat{\alpha} = \frac{-\nabla^T E(w) d}{d^T \nabla^2 E(w) d} = \frac{-e'(0)}{e''(0)} \quad (27)$$

となる. ここで, (17), (18) 式で定義されるニューラルネットでは, (27) 式を効率よく計算できる. まず, 微分した後 $e(\alpha)$, $e_i(\alpha)$ の関係は,

$$e'(\alpha) = \sum_{i=1}^m e_i(\alpha) e_i'(\alpha) \quad (28)$$

$$e''(\alpha) = \sum_{i=1}^m \{ (e_i'(\alpha))^2 + e_i(\alpha) e_i''(\alpha) \} \quad (29)$$

となる. 一方, 更新ベクトル d を各ウエイトに対応するベクトル (d_0, \dots, d_h) に分解すれば, $e_i'(0)$, $e_i''(0)$ の値は,

$$e_i'(0) = -d_{00} - \sum_{j=1}^h \{ d_{0j} \sigma_{ij} + w_{0j} \sigma_{ij}' d_j^T x_i \} \quad (30)$$

$$e_i''(0) = - \sum_{j=1}^h \{ 2 d_{0j} \sigma_{ij}' (d_j^T x_i) + w_{0j} \sigma_{ij}'' (d_j^T x_i)^2 \} \quad (31)$$

となる. ここで, $\sigma_{ij}'' (= \sigma''(w_j^T x_i))$ は, 各事例 i に対する中間ユニット j の二回微分値を表す. なお, sigmoid 関数の場合には,

$$\sigma_{ij}'' = \sigma_{ij}(1 - \sigma_{ij})(1 - 2\sigma_{ij}) \quad (32)$$

という関係が成り立つ.

以上より, 実際にニューラルネットワークに学習させる際の共役勾配法 BP 学習アルゴリズムは以下のようになる.

【共役勾配法 BP 学習アルゴリズム】

- [step 1] 終了条件, 再出発までの学習回数 T の設定を行う.
- [step 2] 更新回数を $t=1$ とおき, $\beta(1)=0$ とする.
- [step 3] (25), (26), (27) 式に従って, ウエイトの更新ベクトル, ステップ幅を計算し, ウエイトを更新する.

[step 4] 終了条件を満たしていれば学習を終了する。

[step 5] $t < T$ ならば $t = t + 1$ として [step 3] に戻り $t = T$ ならば [step 2] に戻る。(再出発)

5.3 改良共役勾配法による BP 学習

本研究で提案する近似範囲を調整する改良共役勾配法の BP 学習アルゴリズムにおいて、再出発の判定の評価値は以下のように計算する。

$$\Delta(t) = \frac{2e''(0)\{\nabla E(t)\}}{\{e'(0)\}^2} \quad (33)$$

(33)式の再出発判定の評価値の計算は(27)~(31)式の最適ステップ幅の推定値の計算過程を利用することにより BP においては計算が容易となるので、(33)式の計算量というのは全体の計算量からすれば非常に小さくなっているといえる。したがって、共役勾配法 BP 学習アルゴリズムと改良型共役勾配法 BP 学習アルゴリズムは同程度の計算量になることがいえる。

【改良型共役勾配法 BP 学習アルゴリズム】

[step 1] 終了条件、再出発までの学習回数 T の設定を行う。

[step 2] 更新回数を $t=1$ とおき、 $\beta(1)=0$ とする。

[step 3] (25), (26), (27)式に従って、ウェイトの更新ベクトル、ステップ幅を計算し、ウェイトを更新する。

[step 4] 終了条件を満たしていれば学習を終了する。

[step 5] (33)式に従って $\Delta(t)$ を計算し、 $\Delta(t) \geq \gamma$ ならば、 $t = t + 1$ として [step 3] に戻り、 $\Delta(t) < \gamma$ ならば、[step 2] に戻る。(再出発)

6. シミュレーション

本節では、まず乱数で発生させた複数のデータに対して、共役勾配法によって、再出発までの学習回数をいくつかの回数ごとにおいて学習させ、再出発までの学習回数が共役勾配法の効率に対しどのように影響するのかを考察する。

さらに、共役勾配法と改良型共役勾配法により同じデータに対して学習させて結果の比較を行い提案法の有効性を検証する。

6.1 シミュレーション条件

使用したニューラルネットワークモデルは、入力層素子数 3、中間層素子数 5、出力層素子数 1 の 3 層ネ

ットワークである。ウェイトの初期値は $-1.0 \sim 1.0$ の範囲で 10 通り乱数発生させた。今回のシミュレーションで使用する学習データは、入力ベクトル x_i を $0 \sim 1.0$ の範囲で各 20 通り乱数発生させ、目標出力値 y_i を、

$$y_i = x_{i1} + x_{i2} + x_{i3} \quad (34)$$

に従って発生させた線形データ (data 1) と、入力ベクトル x_i 、目標出力値 y_i とともに $0 \sim 1.0$ の範囲で各 20 通り乱数発生させた非線形データ (data 2) の二種類のデータである。学習回数は data 1 については 300 回、data 2 については 2000 回とし、それぞれの学習終了時の二乗誤差の初期値による平均値、標準偏差を行う [5], [6]。

また、共役勾配法については、再出発までの学習回数を data 1 では 20 回から 100 回まで 20 回刻みで計 5 パターン data 2 では 20 回から 200 回まで 20 回刻みで計 10 パターンとし、改良型共役勾配法については $\gamma=0.75$ とした。

6.2 シミュレーション結果および考察

図 1、図 2 に従来の共役勾配法 BP 学習による data 2 の学習 2000 回終了時に得られた各々の二乗誤差の初期値による平均値と標準偏差を示し、図 3、図 4 に再出発の回数に関するそれらの最大値、平均値、

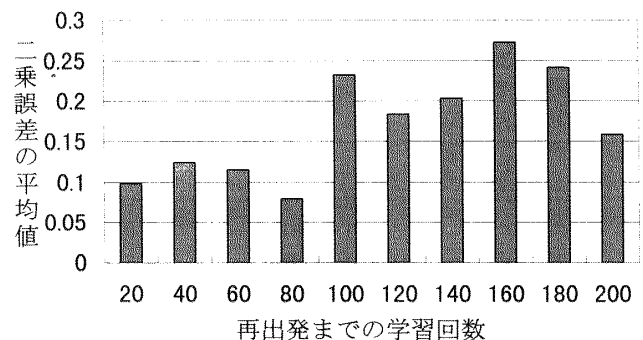


図 1 従来の共役勾配法の二乗誤差の平均値 (data 1)

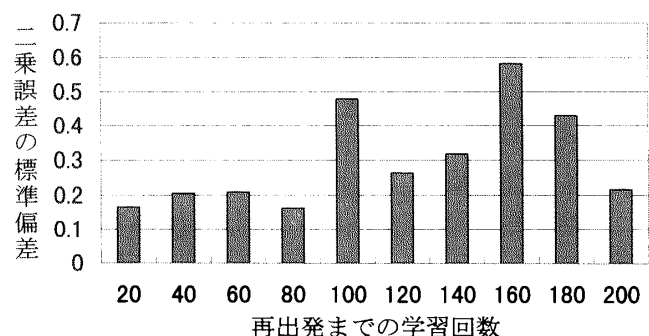


図 2 従来の共役勾配法の二乗誤差の標準偏差 (data 1)

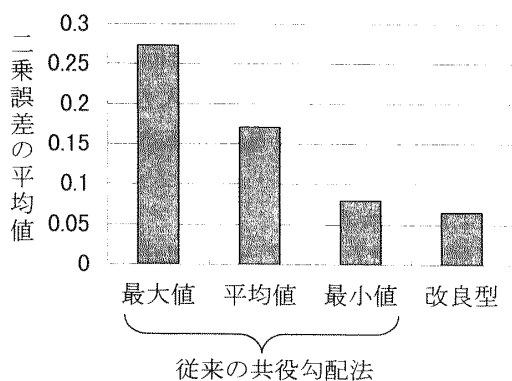


図3 二乗誤差の平均値の比較 (data 1)

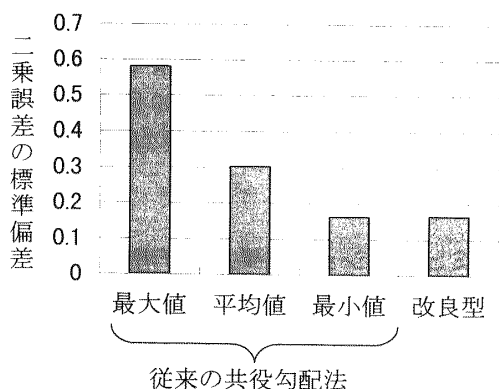


図4 二乗誤差の標準偏差の比較 (data 1)

表1 シミュレーション結果 (data 2)

再出発までの学習回数	平均値	標準偏差
20	0.231899	0.728331
40	0.000171	0.000247
60	0.009812	0.016171
80	0.244986	0.773399
100	0.607043	1.904898
改良型	0.000059	0.000055

最小値および、改良型共役勾配法 BP 学習による data 2 学習 2000 回終了時に得られた二乗誤差の初期値による平均値と標準偏差を示した。また、表 1 には data 1 による共役勾配法 BP 学習と改良型共役勾配法 BP 学習の学習 300 回終了時に得られた各々の二乗誤差の初期値による平均値と標準偏差を示す。

二乗誤差の平均値が小さいことはそれだけ学習の効率がよいことを示し、逆にそれが大きいと学習の効率が悪いことを示す。また、二乗誤差の標準偏差が小さいということは、初期値によるばらつきが抑えられているので、初期値の影響が少ない安定した学習といえ

て、逆にそれが大きいと初期値の影響が大きい不安定な学習といえる。

図 1, 図 2 および表 1 より、従来の共役勾配法による BP 学習は、再出発までの学習回数によって学習の効率が良くなる場合と悪くなる場合があり、学習の効率は再出発までの学習回数に大きく影響することがわかる。また、従来の共役勾配法 BP 学習は再出発までの学習回数によって初期値の影響が少ない安定した学習になる場合と初期値の影響が大きい不安定な学習になる場合があり、学習の安定性も再出発までの学習回数に大きく影響することがわかる。以上のことより、再出発までの学習回数によって学習結果は大きくばらつき、その結果と再出発までの学習回数の関係には規則性がみられない。

図 3, 図 4 および表 1 より従来の共役勾配法 BP 学習と改良型共役勾配法 BP 学習の二乗誤差を比較すると、二乗誤差の平均値は、改良型共役勾配法 BP 学習の方が共役勾配法 BP 学習の最小値より小さい値となっているので、改良型共役勾配法 BP 学習の方が共役勾配法 BP 学習よりも学習の効率がよいことがわかる。また、data 1 のシミュレーション結果では改良型共役勾配法 BP 学習の方が共役勾配法 BP 学習の最小値よりも下回っており、二乗誤差の標準偏差は、data 2 のシミュレーション結果では改良型共役勾配法学習の方が共役勾配法学習の最小値よりも僅かに上回っているが、その平均値と比べると非常に小さい値となっているので、改良型共役勾配法 BP 学習の方が共役勾配法 BP 学習よりも初期値の影響が小さい安定した学習であることがわかる。以上のことより、改良型共役勾配法 BP 学習は、共役勾配法 BP 学習アルゴリズムよりも学習の効率が良く、かつ、ウェイトの初期値に依存しない安定性のある学習であることがいえる。

7. ま と め

二次近似の精度を探索点ごとに計算を行い、その結果を再出発を行う判定の評価値としてみていく改良型共役勾配法を提案し、これを適用した改良型共役勾配法 BP 学習アルゴリズムを提案した。この学習法は、従来の共役勾配法の学習速度を維持しつつ、学習の効率を向上することができた。さらに、この学習法により、共役勾配法のもつ不安定な面も克服することが可能となり、その有効性がシミュレーション実験により検証できた。

参 考 文 献

[1] Branko Soucek and The IRIS Group: "Neural

- and Intelligent Systems Integration”, John Wiley & Sons, Inc. (1991)
- [2] Charalambous, C.: “Conjugate Gradient Algorithm for Efficient Training of Neuralnetworks”, IEE, Proc., Vol. 139, No. 3, pp. 301-310 (1992)
- [3] 吉田利信: “2次元探索共役勾配法を用いた多層神経回路網の高速学習法”, 情報処理学会論文誌, Vol. 31, No. 7, pp. 988-995 (1990)
- [4] 斉藤和己, 中野良平: “3層ニューラルネットワークにおける2階導関数を用いた学習アルゴリズムの高速化”, 信学技報, NC 94-7 (1994)
- [5] 後藤正幸, 俵 信彦: “二次近似範囲を調整する共役勾配法とBP学習への適用”, 電子情報通信学会研究技術報告 NC 94-37 (1994)
- [6] 後藤正幸, 俵 信彦: “共役勾配法を導入したBP学習における安定化に関する研究”, 日本経営工学会誌, Vol. 46, No. 1, pp. 70-77 (1995)
- [7] 後藤正幸, 開沼泰隆, 俵 信彦: “変傾共役勾配法によるBP学習の安定化と高速化”, 日本経営工学会誌, Vol. 46, No. 2, pp. 152-158 (1995)
- [8] 開沼泰隆: “ニューラルネットワークの学習時間低減に関する一考察”, 日本経営工学会平成4年度秋季予稿集, pp. 208-209 (1992)
- [9] 開沼泰隆, 後藤正幸, 俵 信彦: “バックプロパゲーション学習アルゴリズムについて”, 日本経営工学会平成5年度春季予稿集, pp. 76-77 (1993)
- [10] 後藤正幸, 開沼泰隆, 俵 信彦: “共役勾配法によるBP学習について”, 日本経営工学会平成5年度秋季予稿集, pp. 138-139 (1993)
- [11] 板和正敏: 「非線形システムの最適化」, 森北出版 (1986)
- [12] 今野 浩, 山下 浩: 「非線形計画法」, 日科技連 (1978)