

Particle Swarm Optimization を用いた 階層型ニューラルネットワークの一構成

加藤 暢 恵

(2006年10月5日受理)

Design of a Multi-Layered Neural Network Using Particle Swarm Optimization

Mitsue Kato

Particle Swarm Optimization (PSO), whose concept has been established as a simulation of a simplified social milieu, is known as one of the most useful optimization methods for solving non-convex continuous optimization problems. This paper describes a new learning algorithm to simultaneously adjust connection weights included in neural networks and some user-specified parameters included in units. According to the proposed algorithm, it is possible to improve the learning properties of the neural networks, *e.g.*, the learning cost and/or adaptability. The behavior of the proposed algorithm is examined on a numerical simulation example.

Key words: neural network, evolutionary computation, learning, particle swarm optimization, meta heuristics,

キーワード：ニューラルネットワーク, 進化計算, 学習, PSO, メタヒューリスティクス

1. 緒言

ニューラルネットワーク^{1),2)}は、生物の神経系の特徴的な機能をモデル化したものである。生物の神経系は、多数のニューロンが複雑に結合され、それぞれが並列処理を行っていると考えられている。1943年、McCulloch-Pitts は、神経系内のニューロンの動作原理に基づいたニューロンモデルを提案した¹⁾。

ニューロンモデルに基づいたニューラルネットワークの一つに階層型ニューラルネットワークがあり、これは、近年、工学の分野をはじめ、さまざまな分野の研究に広く応用されている^{2),3)}。この階層型ニューラルネットワークの学習には多くの方法が用いられているが、よく用いられる方法の一つに、バックプロパゲーション法（以下、BP法）がある。

シオン法（以下、BP法）がある。

BP法の学習はニューラルネットワークの良否を判定する評価規範を最小にすることを目的として行われ、勾配法の原理に基づいて評価規範値が減少する方向に最適値を探索する手法である。BP法に基づいた学習法では、評価関数が微分可能でなければならず、また、システムヤコビアンに関わる情報を予め把握しておく必要がある。これらは、ニューラルネットワークを構成する上で、非常に煩わしい問題となっている。したがって、ニューラルネットワークは非線形システムの記述に適しているが、実用上、これらの問題を回避する学習法を確立する必要がある。

一方、Particle Swarm Optimization（以下、PSO）⁴⁻⁶⁾は、非常に簡潔なアルゴリズムで構成されており、目的関数の連続性や微分可能性を必要としない。そのため、目的関数情報のみを用いて多峰構造の無制約最適化問題を解くことが可能である。このことは、工学分野において詳細に記述された多くの複雑な問題に対して適用できる可能性が高いことを示唆しており、これにより、

本論文は、課程博士候補論文を構成する論文の一部として、以下の審査委員により審査を受けた。

審査委員：山本 透（主任指導教員）、門田良実、上田邦夫、前原俊信、雛元孝夫（工学研究科）

PSOは近年注目されるようになっており⁷⁾、ニューラルネットワークの荷重係数を、PSOを用いて決定する試みもなされている⁸⁻¹⁰⁾。ところが、BP法との比較において、特段の優位性は示されていない⁷⁾。これは、ニューラルネットワークの荷重係数の調整も重要であるが、ユニットに含まれるシグモイド関数のパラメータ調整が、学習能力(学習速度や汎化能力など)を大きく左右しており、これらのパラメータの調整も併せて最適化する必要があることによるものと考えられる。

そこで本論文では、PSOを用いて、ニューラルネットワークのユニットに含まれるパラメータと、ニューラルネットワークの荷重係数とを対にし、これらを同時に最適調整する新しいニューラルネットワークの学習方法を提案する。また、本手法の有効性を検証するために、代表的な非線形システムの一つであるHammersteinモデルを取り上げ、そのモデリング性能について、従来法との比較を行う。

本論文の構成は以下の通りである。まず、一般的なニューラルネットワークとPSOの概要を述べたあと、PSOを用いたニューラルネットワークの学習アルゴリズムについて考察する。最後に、提案手法の有効性を数値例により検証する。

2. ニューラルネットワークの構成

生物の神経系の特徴的な機能をモデル化したニューロンは、入力端子である樹状突起、出力端子である軸索と本体である細胞体から構成されており、他の複数のニューロンと軸索でつながっている。この結合部分をシナプスと呼び、このシナプスを通して、他のニューロンからの入力を受け取る。ニューロン間の信号は、物理学的または生理学的には細胞の外部に対する内部の電位差としてとらえられ、その電位差のことを膜電位という。この膜電位がある一定のしきい値を超えると、ニューロンは発火、すなわちパルスが発生し、その信号が軸索を通じて他のニューロンに刺激を及ぼす。

神経系などでは、多層回路を単位とし、それぞれの結合によって情報処理を行っている。本論文で用いる階層型ニューラルネットワークを図1に示す。階層型ニューラルネットワークの構成としては、入力信号が与えられる入力層、出力信号を出力する出力層、これらの間に中間層をもつ階層構造を有している。

I_i : 入力層第*i*ユニットの出力を示す。ただし、□は閾値ユニットを表しており、常に1が入力される。

このユニットを便宜上 H_0 とする。

H_j : 中間層第*j*ユニットの出力を示す。ただし、□は

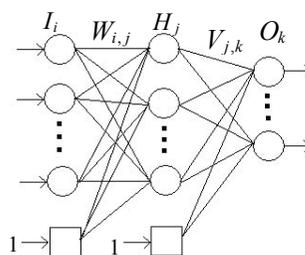


図1 ニューラルネットワークの概要図

閾値ユニットを表しており、常に1が入力される。

このユニットを便宜上 H_0 とする。

O_k : 出力層の出力を示す。

$W_{i,j}$: 入力層第*i*ユニットと中間層第*j*ユニット間の結合係数を示す。

$V_{j,k}$: 中間層第*j*ユニットと出力層第*k*ユニット間の結合係数を示す。

また、 H_j 、 O_k は次式により計算する。

$$H_j = f\left(\sum_{i=0}^p I_i W_{i,j}\right) \quad (1)$$

$$O_k = g\left(\sum_j H_j V_{j,k}\right) \quad (2)$$

$f(\cdot)$ 、 $g(\cdot)$ はシグモイド関数であり、次式により与えられる。

$$f(x) = \frac{2\alpha_f}{1 + e^{-\beta_f x}} - \alpha_f \quad (3)$$

$$g(x) = \frac{2\alpha_g}{1 + e^{-\beta_g x}} - \alpha_g \quad (4)$$

階層型ニューラルネットワークを構成する際、各ユニット間の結合係数 $W_{i,j}$ 、 $V_{j,k}$ 、ならびにシグモイド関数に含まれるパラメータ (α, β) の調整が、非線形近似能力の観点から、極めて重要である。とくに、後者のシグモイド関数のパラメータは、ニューラルネットワークの学習効率を大きく左右するため、慎重に設定しなければならない。そのため、試行錯誤を繰り返さざるを得ず、設計者の大きな負担となっている。

一方、結合係数の学習においても、通常、BP法が適用されるが、BP法に基づいた学習法には、以下のような問題がある。

- (1) 微分に基づく学習則であるために、評価関数が微分可能でなければならない。
- (2) ニューラルネットワークを用いて制御系を設計する場合は、システムやコピアンに関わる情報を予め把握する必要がある。

(1) は不連続な評価関数が扱えないことを意味している。そのため、システムモデリングなどの問題においては、評価関数として、予測誤差信号の2乗がよく用いられる。2乗規範の場合、予測誤差が小さくなる場所では、2乗することでさらに小さくなり、学習能力が極端に劣化する場合がある。また、制御系設計の問題においては、設計仕様にあわせて、微分可能性の制約にとらわれることなく、自由に評価規範を構成できる方が、より実用的である。一方、(2) はシステムヤコビアンが符号が次ぎ次ぎと変化するような系に対しては、エミュレータ等を別途構成し、これによりシステムヤコビアンを常に計算する必要がある²⁾。これは、構造をより複雑化し、あまり実用的ではないと考えられる。

そこで、これらの問題を解決するために、進化的計算手法の一つである PSO を導入し、PSO によって、シグモイド関数に含まれるパラメータと各ユニット間の結合係数を同時に最適化する方法を提案する。

3. Particle Swarm Optimization

3.1. PSO の概要

PSO は、集団を構成する個々の情報を共有しながら進化する鳥や魚などの群れの行動や人間の社会活動などの社会モデルを単純化し、シミュレーションすることによって開発されたメタヒューリスティクスの1つであり、最適化問題を解くための有力な手法の1つとして知られている⁴⁾。

その概念は、群れ状の探索点が情報を共有しながら解空間を探索するというものであり、個体もつ最良の情報とその個体によって形成される集団の最適値から、過去の探索履歴を考慮して大域的最適解、もしくはそれに相当する準最適解を求める手法である。これまでに数多くの数値実験によって、連続型多峰性関数の大域的最適解を実用的な計算時間内に高い精度で求めることが可能なことが明らかにされている^{5),6)}。

また、PSO は、遺伝的アルゴリズム (Genetic Algorithm: GA)¹¹⁾ と同じ確率的探索手法であり、個体の最良値と集団の最良値の調整は GA における交叉に似ている。しかし、PSO は GA のように突然変異、交叉などの複雑な操作が不要で、非常にシンプルな構造を有している。また、パラメータについても、安定性や収束性の数学的解析によって適切な値を導出するという理論的な研究が行われている¹²⁾。

GA などが、遺伝的操作や突然変異などによって、多次元空間内の位置を直接決定するのに対し、PSO は、集団内の相互作用的な力によって位置を決定するた

め、多次元空間内を動きまわるといった特性を持っている。

3.2. PSO の基本アルゴリズム

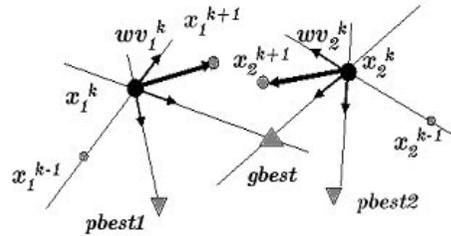


図2 PSOにおける探索点の移動

上述のアルゴリズムを、わかりやすく説明すると、まず、PSO において各探索点は、図2に示すように、自分の持つ最良の位置情報 ($pbest$) と、集団で共有している最良の位置情報 ($gbest$)、および前回の速度ベクトルの線形結合として、新たな探索点を生成する。つまり、 $k+1$ 回目の探索における個体 d の位置 x_d^{k+1} と速度 v_d^{k+1} は、 k 回目の位置 x_d^k と速度 v_d^k を用いて次式により更新される。

$$x_d^{k+1} = x_d^k + v_d^{k+1} \quad (5)$$

$$v_d^{k+1} = \gamma v_d^k + c1 \text{rand} (pbest_d - x_d^k) + c2 \text{rand} (gbest_d - x_d^k) \quad (6)$$

ここで、 $pbest$ は個体 d の k 回目までの探索における個体の最良解を表し、 $gbest$ は k 回目までの探索における、集団の最良解を表す。また、 γ は慣性項とする。

PSO では、探索が進むにつれて慣性項は次の式に従って、徐々に小さくなる。

$$\gamma = \gamma_{\max} - \frac{\gamma_{\max} - \gamma_{\min}}{k_{\max}} k \quad (7)$$

ここで、 γ_{\max} と γ_{\min} はそれぞれ慣性項の最大値と最小値を表し、 k_{\max} は最大探索回数である。多くの数値計算の結果から $\gamma_{\max} = 0.9$ 、 $\gamma_{\min} = 0.4$ を用いることが推奨されている。

慣性項は、個体 d の現在の位置 x_d からの平行移動量が探索が進むにつれて、小さくなることを意味する。また、式 (5) において、 rand は $[0,1]$ の乱数とする。なお、 $c1$ と $c2$ はそれぞれの項に対する重み係数であり、次式のように決められることが多い。

$$\begin{aligned} c1 &= 2, \\ c2 &= 4 - c1 \end{aligned} \quad (8)$$

以上をまとめると、PSO は以下の手順にしたがって

処理される。

- [step 1] 個体数, 最大探索回数を決定する。
- [step 2] 各個体の位置と速度を初期化する。
- [step 3] 各個体に対して, 評価値を計算する。
- [step 4] 個体の最良値 $pbest$ と集団の最良値 $gbest$ を求める。
- [step 5] 各個体の速度と位置は, 式 (5), (6) に従い更新する。
- [step 6] 探索回数 k が最大探索回数以下なら $k=k+1$ として探索を継続, そうでなければ, 探索を終了する。

4. PSO を用いたニューラルネットワークの構成

2 節の階層型ニューラルネットワークの学習に, 3 節の PSO を適用し, 併せてシグモイド関数の設計パラメータも調整する新しい学習アルゴリズムを考察する。

ここで, ニューラルネットワークの各層の数を, 入力層 $p+1$ 個 (閾値ユニットを含む), 中間層 $q+1$ 個 (閾値ユニットを含む), 出力層 r 個とすると, PSO における 1 つの個体 X が受けもつ探索点の総数は, $((p+1) \cdot q + (q+1) \cdot r + 4)$ 個となる。また, ニューラルネットワークの各パラメータは, それぞれを以下のように, PSO の要素に変換される。ここで, l ($l=1, 2, \dots, L$) は, 個体数を表すものとする。

$$X_l = [x_{1,l}, x_{2,l}, \dots, x_{(p+1),l}, \dots, x_{((p+1) \cdot q),l}, \dots, x_{((p+1) \cdot q + (q+1) \cdot r + 4),l}]$$

$$\begin{aligned} x_{1,l} &= W_{1,1} \\ x_{2,l} &= W_{1,2} \\ &\vdots \\ x_{q,l} &= W_{1,q} \\ x_{(q+1),l} &= W_{2,1} \\ &\vdots \\ x_{((p+1) \cdot q),l} &= W_{(p+1),q} \\ x_{((p+1) \cdot q + 1),l} &= V_{1,1} \\ &\vdots \\ x_{((p+1) \cdot q + (q+1) \cdot r)} &= V_{(q+1),r} \\ x_{((p+1) \cdot q + (q+1) \cdot r + 1)} &= \alpha_f \\ &\vdots \\ x_{((p+1) \cdot q + (q+1) \cdot r + 4)} &= \beta_g \end{aligned}$$

5. 数値例

ここでは, 提案手法の有効性を検証するために, 次式で示す, Hammerstein モデルを用いる。

$$y(t) = 0.6 y(t-1) - 0.1 y(t-2) + 1.2 x(t-1) - 0.1 x(t-2) \quad (9)$$

$$x(t-1) = 1.5 u(t-1) - 1.5 u^2(t-1) + 0.5 u^3(t-1) \quad (10)$$

まず, 上記モデルを用いて, BP 法と提案手法の比較を行う。このとき, 対象とするシステムの静特性は, 図 2 に示す通りである。この比較に用いた PSO のパラメータは次に述べるとおりである。 γ は 0.9~0.4 に線形減少, c_1 と c_2 はそれぞれ 2.0 とする。PSO における個体数は 20 個体とし, 最大探索回数を 500 回とする。

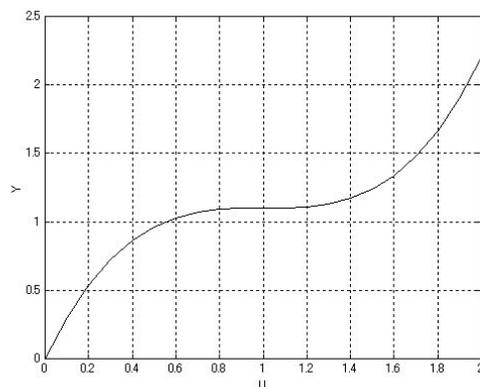


図3 システムの静特性

また, 速度の制限値は 1.0 とし, 結合荷重の初期値は, ± 1.0 の範囲で初期化する。なお, シグモイド関数のパラメータは, PSO のパラメータ固定の場合, $a_f, a_g = 1.5$, $b_f, b_g = 1.0$, BP 法の場合, $a_f, a_g, b_f, b_g = 1.0$ とする。BP 法, PSO 法とも, 500 回繰り返した後, 学習を終了する。NN の構造は, 入力層 6, 中間層 10, 出力層 1 とする。また, このときの評価規範は, 教師信号とニューラルネットワークからの誤差を用いて, (11) 式により与える。

ここで, $y(t)$ は, (9), (10) 式によるモデル出力を示し, $\hat{y}(t)$ は, ニューラルネットワークからの出力 (予測値) を示す。

$$E(i) = \sum_{t=1}^{600} |y(t) - \hat{y}(t)| \quad (i=1, 2, \dots, 500) \quad (11)$$

以上の条件の下で学習を行った後、各手法によりモデリングを行った。その結果を、図4, 5, 6に示す。これらはそれぞれ、提案手法、パラメータ固定 PSO, BP 法によるモデリング結果を示している。また、図7には、(11) 式により与えられる各手法の評価値の推移を示し、表1には、各手法におけるモデル化誤差の統計量を示す。

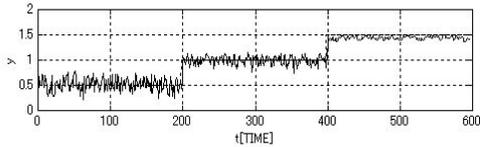


図4 提案手法のモデリング結果

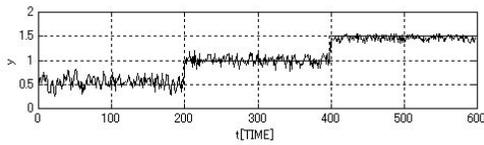


図5 パラメータ固定 PSO のモデリング結果

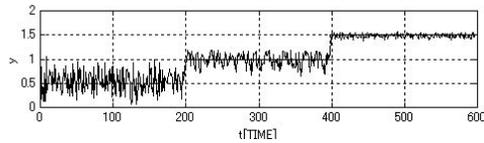


図6 BP 法のモデリング結果

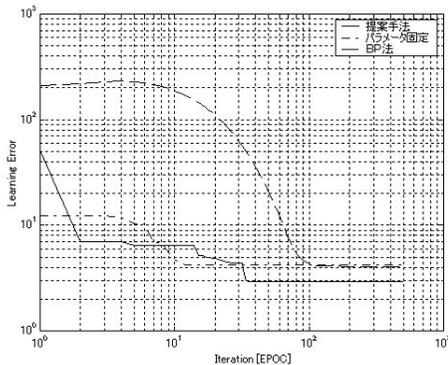


図7 BP 法、パラメータ固定 PSO 及び提案手法 (PSO) の評価値の推移

表1 モデル化誤差の統計量

	提案手法	パラメータ固定PSO	BP 法
平均	0.022	0.075	0.112
分散	0.0041	0.0042	0.0232

以上の結果より、PSO 法は BP 法に比べ学習の収束が早く、また、提案手法は、パラメータを固定した手法よりも学習の収束が早く、学習結果も良好であることがいえる。

6. 結 言

本論文では、従来、BP 法によって行われているニューラルネットワークの結合荷重の更新と、中間層、出力層の出力の計算時に使用するシグモイド関数に含まれる各パラメータ（設計パラメータ）の更新に、PSO を用いた新しいニューラルネットワークの学習法を提案し、その手法の有効性について検討を行った。また、PSO 法においては、シグモイド関数のパラメータを固定する手法と提案手法との比較も行った。その結果、PSO 法は、BP 法よりも早く学習することが可能であり、学習結果も良好であることがわかった。また、提案手法は、パラメータ固定の PSO に比べて、学習の収束が早く、学習結果も良好であることがわかった。これにより、PSO によるニューラルネットワークのパラメータの更新は、BP 法と比較して有効であるといえる。

今後は、このアルゴリズムを実用的な問題に適用し、有効性について検討を行う予定である。

【参考文献】

- 1) S. Haykin: Neural Networks, Macmillian College Publishing, New York, 1994
- 2) 大松 繁, 山本 透: セルフチューニングコントロール, 計測自動制御学会学術図書, 1996
- 3) 渡辺桂吾: ニューラルネットワーク計算知能, 森北出版, 2006
- 4) J. Kennedy, R. Eberhart: "Particle Swarm Optimization", IEEE International Conference on Neural Networks, 1995, pp.1942-1948
- 5) M. Clerc, J. Kennedy: "The Particle Swarm, Explosion, Stability, and Convergence in a Multi-Dimensional Complex Space", IEEE Trans. on Evolutionary Computation, vol.6, pp.58-73, 2002.
- 6) J. Kennedy, R. C. Eberhart: "Swarm Intelligence", Morgan Kaufmann Publishers, 2001.
- 7) 井出 東, 安田恵一郎: "適応型 Particle Swarm Optimization に関する基礎的検討", 電気学会電子・情報・システム部門誌, Vol.124, No.2, 2004, pp.550-557
- 8) 伊藤 稔, 田中正博: "Particle Swarm Optimization

- によるニューラルネットワークの学習に関する研究”, 平成17年電気学会電子・情報・システム部門大会論文誌, 2005, pp.1087-1089
- 9) Magy M. Kandil, et.al: “A New Approach for Optimizing Backpropagation Training with Variable Gain Using PSO”, GVIP 05 Conference, 19-21 December 2005, CICC, Cairo, Egypt.
- 10) Alex v. E. Conradie, et.al, “Adaptive Control utilizing Neural Swarming”, Proceedings of the Genetic and Evolutionary Computation Conference, New York, USA, 2002.
- 11) D.E. Goldberg: Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley Publishing, 1989.
- 12) K.E. Parsopoulos, M.N. Vrahatis: “Recent Approaches to Global Optimization Problems through Particle Swarm Optimization”, Natural Computing, Vol.1, No.2-3, 2002, pp.235-306