

BP学習アルゴリズムを模倣するニューラルネットワークの構築

久保博隆 橋本周司

早稲田大学理工学部応用物理学科

1. はじめに

ニューラルネットワーク（以下NN）を学習させる手段として誤差逆伝播法（以下BP法）が一般的に用いられている[1][2][3]. しかしBP法では学習のための機構がNNとは独立しているばかりではなく、単純な素子の結合では実現が難しい計算が必要である.

本稿では訓練対象のNNとは別に、BP法を模倣するNNを作成し、それを用いて訓練対象のNNを学習させることを検討する.

学習法にNNを用いることにより、分散コンピューティング等のNN自体の利点を学習に取り入れることが出来る.

2. 手法

BP法においては、出力関数が

$$f(x)=1/(1+\exp(-x)) \quad (\text{シグモイド関数})$$

である時、各ニューロンの出力O、教師信号y、結合荷重wによって、以下の式のように結合荷重の修正値Δwが計算されている.

$$\Delta w^{k-1,j,k_i} = \epsilon d k_i O^{k-1,j} \quad (\epsilon \text{ は学習定数})$$

$$d k_i = (y_i - O k_i) O k_i (1 - O k_i) \quad (\text{第}k\text{層が出力層}) \quad (1)$$

$$= (\sum_p w^{k-1,p,k_i+1} d k_{i+1,p}) O k_i (1 - O k_i) \quad (\text{第}k\text{層が出力層以外})$$

- k 現在注目している層
- j 第 k-1 層のニューロンの番号
- i 第 k 層のニューロンの番号
- p 第 k+1 層のニューロンの番号
- $O^{k-1,j}$ 第 k 層の i 番目のニューロンの出力
- w^{k-1,j,k_i} 第 k-1 層の j 番目のニューロンのから第 k 層の i 番目のニューロンへの結合荷重

つまりO、y、wから、結合荷重を更新するのが訓練器の役割である.

そこで、学習したいNNの各O、y、wを入力すると、上の修正値Δwを出力するような「BP法を模倣するNN」を作成することを試みた.

提案手法全体の流れを、図1に示す.

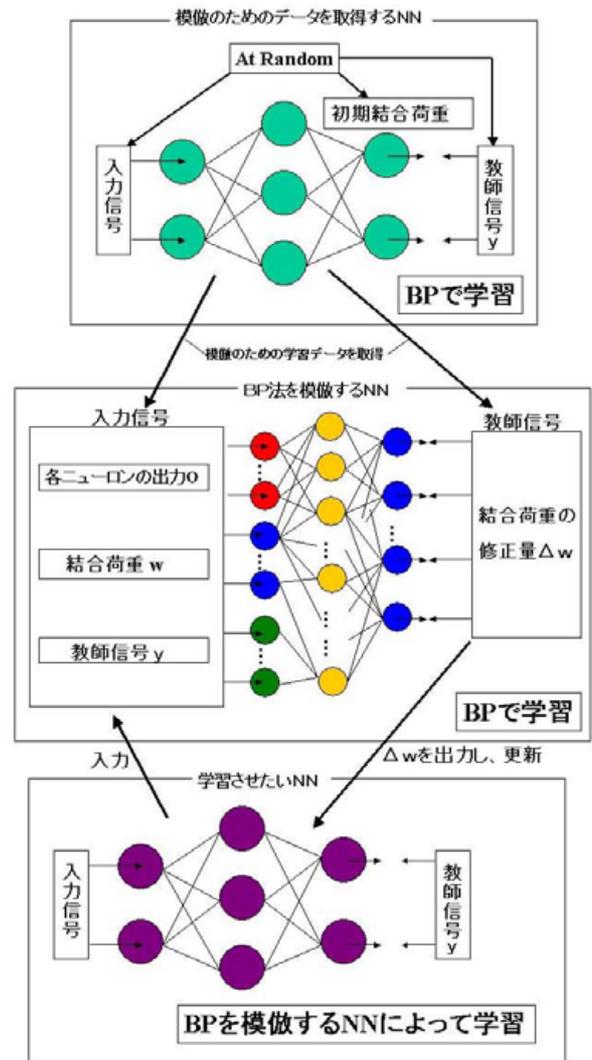


図1. 学習全体の流れ

「BP法を模倣するNN」自体の学習には、慣性項のあるBP法を用いた.

はじめに、「模倣のためのデータを取得するNN」に

- ・ランダムに生成した初期結合荷重
- ・ランダムに生成した入力信号と教師信号の対

を与えてBP法で学習を行った.

つまり、「模倣のためのデータを取得するNN」の各ニューロンの出力O、教師信号y、および結合荷重wを正規化した値を「BP法を模倣するNN」の入力信号とし、又、(1)式で求めたΔwを正規化した値を教師信号とする. 平均二乗誤差が 10^{-5} 以下に収束するまで「BP法を模倣するNN」を学習させた.

「BP法を模倣するNN」の学習が完了後、それを用いて「学習させたいNN」（「模倣のためのデータを取得するNN」と同型）を学習させた.

Construction of the neural network to imitate Back Propagation learning algorithm
 Hiroataka Kubo kubo@shalab.phys.waseda.ac.jp
 Shuji Hashimoto shuji@shalab.phys.waseda.ac.jp
 Department of Applied Physics
 Waseda University
 3-4-1 Okubo, Shinjuku, Tokyo 169-8555 Japan

3. 評価実験

3.1 実験条件

学習に必要な3つのNNを図2～4と表1, 2に示す. 簡単のためここでは, 2入力1出力の論理関数の学習を試みた.

尚, シグモイド関数の特性を考慮して, 「模倣のためのデータを取得するNN」の入力信号と教師信号には, 0/1ではなく0.1/0.9を与えた.

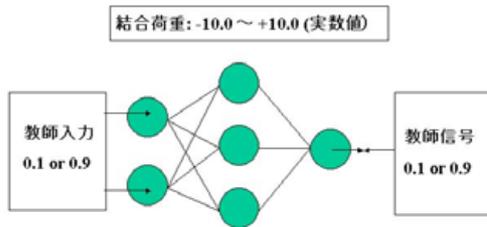


図2. 「模倣のためのデータを取得するNN」の概形

表1. 「模倣のためのデータを取得するNN」の詳細

| | |
|-----------|------------------|
| 入力層 | 2 |
| 中間層 | 3 |
| 出力層 | 1 |
| 総ニューロン数 | 6 |
| 結合荷重数 | 9 |
| 出力関数 | $1/(1+\exp(-x))$ |
| B P法の学習係数 | 0.8 |

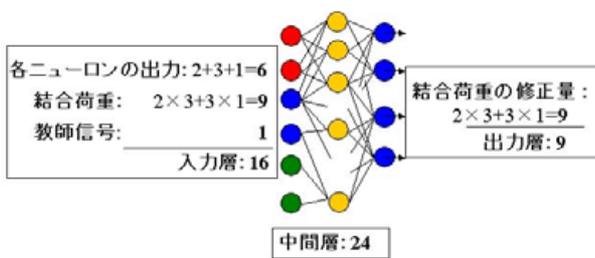


図3. 「BP法を模倣するNN」の概形

表2. 「BP法を模倣するNN」の詳細

| | |
|---------|----------------------|
| 入力層 | 16 |
| 中間層 | 24 |
| 出力層 | 9 |
| 出力関数 | $1/(1+\exp(-x/100))$ |
| 学習係数 | 0.2 |
| モーメント係数 | 0.8 |



図4. 「学習させたいNN」の概形
(「模倣のためのデータを取得するNN」と同型)

3.2 実験結果

「BP法を模倣するNN」で「学習させたいNN」にOR, 及びNORを学習させ, 直接BP学習を行った場合と比較した. ORを学習した場合の出力誤差の推移を図5, NORを学習した場合の推移を図6に示す. 表3, 表4はそれぞれの学習終了後の入出力関係である. 学習試行回数は, 入力4パターンを1回とカウントした.

「BP法を模倣するNN」により, 直接のBPと同程度にOR, NORが学習可能であることがわかる.

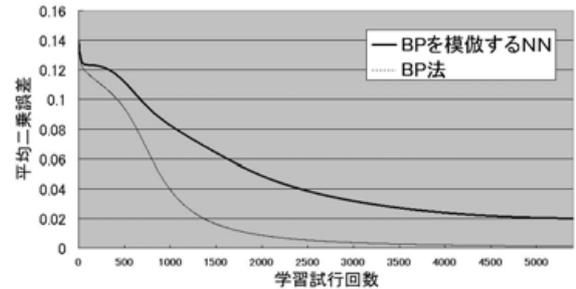


図5. 試行回数と平均二乗誤差の推移 (OR)

表3. BP法を模倣するNNを用いORを学習した時の出力

| 入力 | 出力 | 教師信号 |
|-----|-----|-------|
| 0.1 | 0.1 | 0.204 |
| 0.1 | 0.9 | 0.954 |
| 0.9 | 0.1 | 0.649 |
| 0.9 | 0.9 | 0.943 |

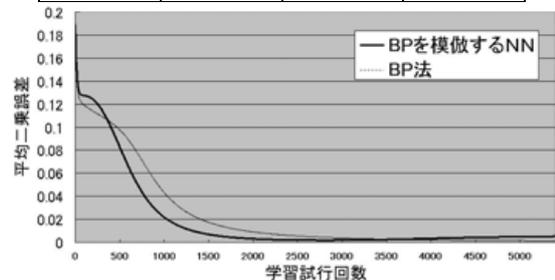


図6. 試行回数と平均二乗誤差の推移 (NOR)

表4. BP法を模倣するNNを用いNORを学習した時の出力

| 入力 | 出力 | 教師信号 |
|-----|-----|-------|
| 0.1 | 0.1 | 0.843 |
| 0.1 | 0.9 | 0.152 |
| 0.9 | 0.1 | 0.062 |
| 0.9 | 0.9 | 0.106 |

4. まとめ

BP学習アルゴリズムを模倣するニューラルネットワークを作成する方法を提案し, 簡素なモデルでその効果を確認した. 今後は, 構造がより複雑なニューラルネットワークにおいて本学習方法の適用可能性とネットワーク規模について検討したい.

参考文献

- [1] D. Rumelhart and J. McClelland 1986, "Learning internal representation by error propagation," Parallel Distributed Processing 1, 318-362, MIT Press.
- [2] 長尾智晴, 最適化アルゴリズム, 昭晃堂 (2000)
- [3] Metal Brain, <http://metalbrain.piroo.com/>