3Ｅ－６

学習曲面における ローカルミニマムの十分条件

山口　智　　板倉　秀清
千葉工業大学

１．はじめに

多層のニューラルネットワーク（以下ＮＮと略）を用いたバックプロパゲーション学習法（以下ＢＰと略）におけるローカルミニマムの問題は、問題に含まれる非線形性のため理論的な考察が難かしい。そこで本稿では、有界閉集合上の連続関数のＮＮによる近似的実現において、中間層から出力層への変換を有界閉集合上の連続関数が作るベクトル空間の要素の有限次元部分空間への射影とみることにより、荷重空間上の一点がローカルミニマムとなるための十分条件を導いた。得られた条件は、Ｂａｌｄｉら[1]が線形ユニットからなるＮＮについて導いた条件の拡張になっている。そこで、Ｂａｌｄｉの結果を利用して以前に提案した学習アルゴリズム[2]を非線形ユニットを持つＮＮに適用し、シミュレーションを行ったところ従来のＢＰよりよい結果が得られた。

２．３層ＮＮと直交関数系

入力パターンの集合Ｘをｎ次元ユークリッド空間内の有界閉集合、ｆをＸから実数空間ＲへのＸ上で二乗可積分な連続関数とする。また、入力層ユニット数ｎ個、中間層ユニット数ｐ個、出力層ユニット数１個とし、入力層ユニットおよび出力層ユニットの入出力関数は線形、中間層ユニットの入出力関数を非線形で二乗可積分な関数とする３層ＮＮが与える入出力関数をＮ：Ｘ→Ｒで表す。このとき、ＢＰ学習の目的は所望の関数ｆとＮの二乗誤差

$$E=\int_X \| f(x)-N(x) \|^2 dx \qquad (1)$$

が最小となるＮ（ｘ）を与える３層ＮＮを作り出すことである。

入力層に $x \in X$ が与えられたときの中間層ユニットのそれぞれの出力値を $\phi_1(x),\ldots,\phi_p(x)$ とすると、Ｎは $\phi_1(x)$ の線形結合

$$N(x)=w_1\phi_1(x)+\ldots+w_p\phi_p(x) \qquad (2)$$

として表すことができる。ここで、$\phi_i(x)$ は二乗可積分な関数である。

関数 $\phi_i(x)$ を固定して考えると、（２）式による関数ｆ（ｘ）の近似は、$\phi_1(x),\ldots,\phi_p(x)$ を説明変数としてｘに関する非線形回帰となり、（１）式の最小値を与えるＮ（ｘ）は、Ｘ上の連続関数からなるベクトル空間の要素ｆ（ｘ）を $\phi_1(x),\ldots,\phi_p(x)$ が張るｑ（≦ｐ）次元部分空間への射影したものである。また、$\phi_1(x),\ldots,\phi_p(x)$ がｐ次元空間の正規直交基底となっていれば、$w_1,\ldots,w_p$ は ϕ_i を含む直交関数系によるｆ（ｘ）のフーリエ展開の係数となる。

３層ＮＮとＢＰ学習を中間層と出力層の間の関係に注目して見ると、ＮＮの入出力関数は中間層ユニットから得られるｐ個の関数が張る部分空間へのｆ（ｘ）の射影であることがわかる。このとき、入力層と中間層の間では、最も損失が少なくなるような部分空間を張る $\phi_1(x),\ldots,\phi_p(x)$ を求めているものと考えられる。

３．ローカルミニマムの十分条件

ここでは、ＮＮの結合荷重が作る荷重空間の一点が（１）式により形成される学習曲面のローカルミニマムであるための十分条件について考える。

［定理］

$\psi_i(x)$ $(i=1,\ldots)$ を完全直交関数系とし、ＮＮによって近似しようとする関数ｆ（ｘ）が

$$f(x)=\Sigma_{i=1}^{\infty}\alpha_i\psi_i(x)$$

と表せるとする。このとき、入出力関数Ｎ（ｘ）が $\alpha_i\psi_i(x)$ $(i=1,\ldots)$ の部分列の和

$$N(x)=\Sigma_{i=1}^{p}\alpha_{k_i}\psi_{k_i}(x) \qquad (3)$$

となるような３層ＮＮが得られたならば（１）式における評価関数Ｅのgradientは０となる。このときのＥの値は

$$E=\Sigma_{i=1}^{\infty}\alpha_i^2-\Sigma_{i=1}^{p}\alpha_{k_i}^2 \qquad (4)$$

となる。（４）式の最小値を与えずに（３）式を満たす

A Sufficient Condition for Local Minima on a Learning Surface
Satoshi YAMAGUCHI, Hidekiyo ITAKURA
Chiba Institute of Technology

ような結合荷重を与える荷重空間上の点はEが作る学習曲面のローカルミニマムを与える。 ■

（証明は略）

中間層ユニットの入出力関数として、よく用いられているsigmoid関数を利用するした場合、第２層のユニットで直交関数系を形成することは期待できない。しかしながら、中間層にsigmoid関数を用いた３層ＮＮは任意の連続関数を近似することが知られている[3]。そこで図１のように、出力層の前に線形ユニットを持つ層をいれた４層ＮＮを考て、第３層に直交関数系が現れていると考えればよい。

４．シミュレーション

前節の結果は文献［１］におけるＢａｌｄｉの結果の非線形への拡張である。我々はＢａｌｄｉの結果を用いて、すべてのユニットが線形の入出力関数である３層ＮＮの場合、学習途中で中間層ユニット数を増やすことにより、学習が早く収束することを示した[2]。この方法の論理的根拠は、ローカルミニマムの個数が中間層の個数ｐに依存していることにある。前節のようにＮＮが非線形ユニットを持つ場合、ローカルミニマムの数はｐの値にかかわらず無限に存在するが、ｆ（ｘ）が少ない直交関数系の部分列で十分に近似できると仮定すると、この学習法も有効に働くことが期待できる。このことをシミュレーションによって示す。

通常のＢＰ学習法と、中間層ユニットを学習途中に増やす方法を用いて閉区間［−ａ，ａ］上の関数ｆ（ｘ）

$$f(x) = \begin{cases} x + a & (x < 0) \\ -x + a & (x \geq 0) \end{cases}$$

を近似する学習を行い収束の様子を比較した。図２は入力層ユニット１個、中間層ユニット６個、出力層ユニット１個の３層ＮＮを用いた学習であり、図３は入力層ユニット１個、出力層ユニット１個、中間層ユニットの個数を１個から学習途中で徐々に増やして学習を行ったときの学習の収束の様子である。後者はユニットが増えるごとに一時的に誤差が増加するが、その後急激に収束していることがわかる。この結果を見ると、非線形ユニットが入ったＮＮに対しても、我々の提案した学習が有効であると思われる。

５．おわりに

ＮＮが近似しようとする関数を直交関数系で近似することにより、多層ＮＮの荷重曲面上にできるローカルミニマムの十分条件について考察した。また、ＢＰ学習の高速化をはかるアルゴリズムを提案し、簡単なシミュレーションによりその効果を確認した。

参考文献

［１］Baldi,Horkin:"Neural Networks and Principal Component Analysis: Learning from Examples without Local Minima", Neural Networks, pp.53-58, vol.2 (1988)

［２］山口，板倉：”ニューラルネットにおけるＢＰ学習と中間層ユニット数の関係” １９９０信学秋期全大，Ｄ−１１(1990)

［３］船橋：”ニューラルネットワークのcapabilityについて”，信学技法，ＭＢＥ８８−５２(1988)

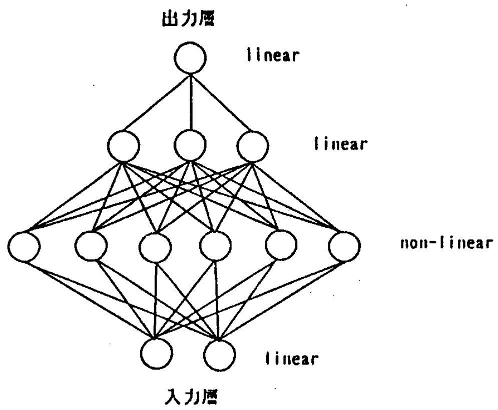

図１

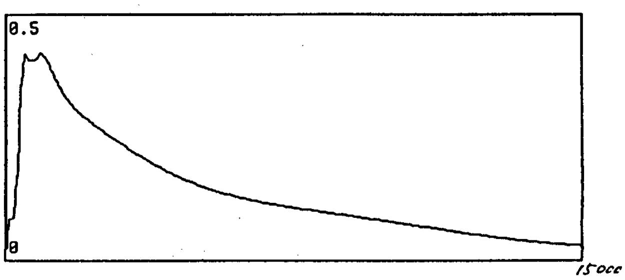

図２

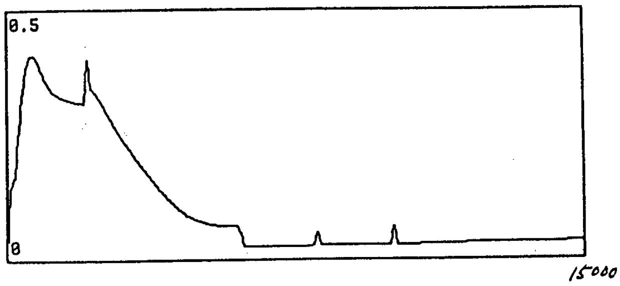

図３