ニューラルネットワークのＨＡ学習の関数依存特性：高木関数

6K－08

河野 芳江　下川 信祐
（株）ＡＴＲ環境適応通信研究所

１．はじめに

言語、ピクチャー、履歴データなどの多種多様なデータを使って、様々な事象間の相関や規則性を写像として取り出す手法の構築を目指している。筆者らは、写像モデルとして３層 Perceptron 型ニューラルネットワークを採用し、その学習方法として高次元アルゴリズム[1]を用いる手法の検討[2－5]を進めている。一般に、扱う入出力数や教師信号数が増加すると、ニューラルネットの多数の結合強度を最適設計することは大変難しくなる。高次元アルゴリズムは、他の最適解探索法と比べ、局所解に捕まりにくい、変数の多い問題に有利、等の特徴があるため、有効な手段になると期待される。

前回、適切な混合性を導入することにより、教師信号の入力と出力が単純な関数関係―アフィン関数―をもつ場合に対しては、本手法が効率的に機能することを報告した[3]。

今回、入出力間に、より複雑な関数関係がある場合に対して、本手法の性能評価を試みた。一例として、連続であるが至るところ微分不可能な特異関数の１つである高木関数を取り上げ、計算機実験を行った結果を報告する。

２．モデルと学習法

３層 Perceptron 型ネットワークを考える。複雑な関数表現を可能にするため、入力－中間、および、中間－出力ユニット間の各結合に各々２つの結合強度、w^1_{ij}、w^2_{ij}、および、W^1_{jk}、W^2_{jk} を設定し、中間ユニットの入出力 x_j と X_j、および、出力ユニットの入出力 y_k と Y_k に対し、次のような分数型の演算を用いた[2]。

Performance of a Method of Learning in Neural Networks by Hamiltonian Algorithm: Test Results for Takagi-function Case
KOHNO Yoshie, SHIMOGAWA Shinsuke
ATR Adaptive Communications Research Laboratories

$$x_j(s)=\frac{\sum_{i=1}^{n1} w^1_{ij}\cdot d_i(s)}{1+\sum_{i=1}^{n1} w^2_{ij}\cdot d_i(s)}\ \text{、}\ X_j(s)=f(x_j(s)) \qquad (1)$$

$$y_k(s)=\frac{\sum_{j=1}^{n2} W^1_{jk}\cdot X_j(s)}{1+\sum_{j=1}^{n2} W^2_{jk}\cdot X_j(s)}\ \text{、}\ Y_k(s)=f(y_k(s)) \qquad (2)$$

n_1、n_2、n_3を、各々、入力層、中間層、出力層のユニット数、n_sを教師信号数、$d_i(s)$, $e_k(s)$を s 番目の教師信号の入出力値とする。$f(u)$はシグモイド関数 $f(u)=\{1+\exp(-u/T)\}^{-1}$である。

コスト関数を、教師信号と出力の二乗誤差の和

$$V(w^1_{ij},w^2_{ij},W^1_{jk},W^2_{jk})=\sum_{s=1}^{ns}\sum_{k=1}^{n3}\{e_k(s)-Y_k(s)\}^2 \qquad (3)$$

とし、これを最小にする、$2(n_1n_2+n_2n_3)$ 個の結合強度 w^1_{ij}, w^2_{ij}, W^1_{jk}, W^2_{jk} $(i=1,\ldots,n_1;\ j=1,\ldots,n_2;\ k=1,\ldots,n_3)$ を求める最適化問題としてニューラルネットの学習問題を取り扱う。最適化手法として、多変数の場合に有利な高次元アルゴリズムを採用する。手順は次の通り。

(1) 被最適化変数に正準共役な変数を導入し、文献[3]にある方法で混合性を導入したハミルトン関数を設定する。

(2) 変数をランダムな初期値におき、正準方程式に従う運動を追跡することによりコスト関数の最小値を求める。

混合性の増強により、運動は解でない領域をすばやく通り過ぎ、解に相当する領域に長く滞在する性質を持つようになるため、効率的に最適解を探索することが可能になる。

３．テスト関数を用いた評価例 ―高木関数―

本手法の性能評価のため、教師信号の入力と出力

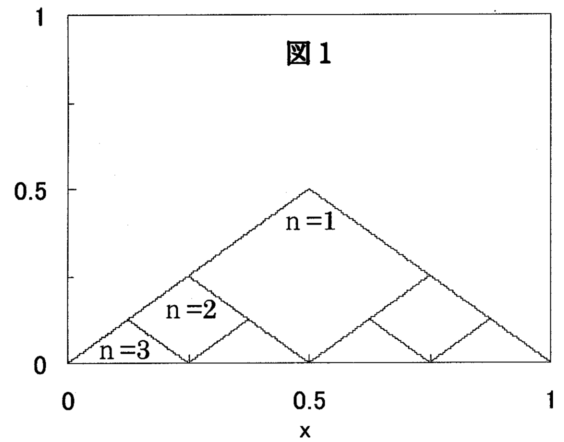
が比較的複雑な関数関係にあるいくつかのテスト関数に対して計算機実験を試みた。今回、その一例として、高木関数に対する結果を報告する。

高木関数(T(x))とは、区間[0,1]において次式で定義される、連続であるが至るところ微分不可能な特異関数である。

$$T(x) = \sum_{n=0}^{\infty} \frac{\tau(2^n x)}{2^n} = \tau(x) + \frac{\tau(2x)}{2} + \frac{\tau(2^2 x)}{2^2} + \cdots \tag{4}$$

$$\tau(u) = \begin{cases} u-[u] & \left(0 \leqq u-[u] < \frac{1}{2}\right) \\ 1-(u-[u]) & \left(\frac{1}{2} \leqq u-[u] < 1\right) \end{cases}$$

ここで、$[u]$は実数 u の整数部分を表す。図１に、区分的に線形な関数 $\tau(2^n x)/2^n$（n＝1,2,3）を示す。



高木関数はこれらの関数の和の n→∞の極限である。

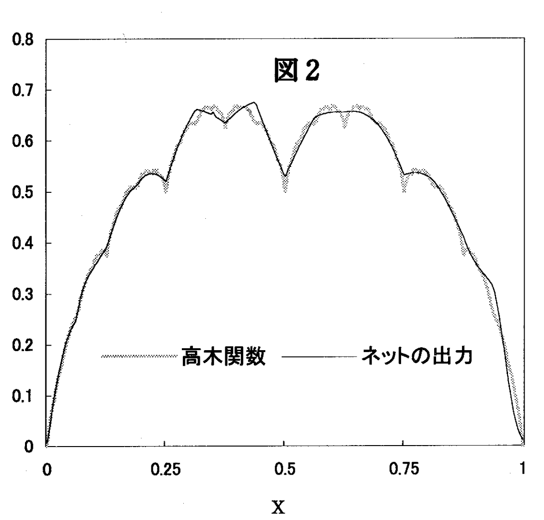


図２に、高木関数の真値と本手法を用いたニューラルネットの出力を示す。高木関数は至るところ凹凸しているので、写像を精度良く近似するためには、要求する精度に相応な数の教師信号を与えることが必要である。図は教師信号数が５０の場合である。教師信号の入力 x ∈[0,1]には一様乱数を使用した。鋭く尖っている個所はやや正確さに欠ける傾向があるが、これは、その先鋭部分に対応する教師信号が与えられていないことが主な原因であり、さらに教師信号数を増やしてゆけば改善されると考えられる。こうした個所を除けば、関数の全体的な様子はかなり良く捉えられていることが分かる。

４．おわりに

高次元アルゴリズムを用いたニューラルネットの学習法の性能評価を行った。滑らかではなく変化が激しい高木関数をテスト関数として取り上げたが、十分な教師信号を与えればかなり精度良く関数を近似できることが分かった。

さらに精度や効率を上げるため、今後、ユニットでの演算処理法として今回用いた分数型以外の高次項も用い、種々の方法を比較･検討する予定である。

【参考文献】

[1]新上：”高次元アルゴリズム”、ｂｉｔ、Vol.31, NO.7, p.2 (1999).

[2]倉持,新上,下川,河野：”拡張型結合をもつニューラルネットワークのＨＡ学習”、情報処理学会第 61 回全国大会、2-57 (2000).

[3]河野,下川, 新上：”高次元アルゴリズムによる効率的な大局解探索法:ニューラルネットワークの学習問題への適用”、情報処理学会第 61 回全国大会、2-63 (2000).

[4]新上, 大田原：”ニューラルネットワークのＨＡ学習の関数依存特性：偏差値関数の場合”、情報処理学会第 62 回全国大会発表予定、6K-01 (2001).

[5]下川, 寺前, 河野：”関数関係とならない教師データを伴う３層ニューラルネットの学習”、情報処理学会第 62 回全国大会発表予定、6K-03 (2001).