

新上 和正 大田原 一成
(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、様々な事象間の因果関係や相関関係を写像として取り出す手法の構築を目指している。ここでは、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を用いる手法の検討 [1] 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。今回、入力と出力との間により複雑な関数関係が成り立っている（例えば、生徒のテストの点数を入力とし、その生徒の偏差値を出力とする）偏差値関数に対して、本手法の性能評価を試みた。

2. モデルと学習法

3 層 Perceptron 型ネットワークを考える。より複雑な関数関係の表現を可能にするため、入力—中間、および、中間と出力ユニット間の各結合に各々 2 種類の結合強度、 $\omega_{ij}^1, \omega_{ij}^2$, および、 $\Omega_{jk}^1, \Omega_{jk}^2$ 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。

$$x_j(s) = \frac{\sum_{i=1}^{n_1} \omega_{ij}^1 d_i(s)}{1 + \sum_{i=1}^{n_1} \omega_{ij}^2 d_i(s)}, \quad y_k(s) = \frac{\sum_{i=1}^{n_2} \Omega_{ik}^1 X_i(s)}{1 + \sum_{i=1}^{n_2} \Omega_{ik}^2 X_i(s)}. \quad (1)$$

ここで、 $X_j(s) = f(x_j(s))$, $Y_k(s) = f(y_k(s))$, 但し $f(x)$ はシグモイド関数で、 $f(x) = (1 + e^{-x/T})^{-1}$. 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。

コスト関数を、

$$V(\omega_{ij}^1, \omega_{ij}^2, \Omega_{jk}^1, \Omega_{jk}^2) = \sum_{s=1}^{n_s} \sum_{k=1}^{n_3} \{e^k(s) - Y_k(s)\}^2 \quad (2)$$

として、これを最小とする $\omega_{ij}^{1*}, \omega_{ij}^{2*}, \Omega_{jk}^{1*}, \Omega_{jk}^{2*}$ を求める。

入力 $d_i(s)$ と $e_k(s)$ は、点数とその偏差値との関係を持っているとする仮定する：

$$e_i(s) = \frac{1}{\sigma(s)} (d_i(s) - \bar{d}(s)), \quad (\bar{d}(s) = \frac{\sum_{i=1}^{n_1} d_i(s)}{n_1}, \sigma(s) = \sqrt{\frac{\sum_{i=1}^{n_1} (d_i(s) - \bar{d}(s))^2}{n_1}}). \quad (3)$$

この問題は、「式 (3) のように入力と出力との（隠れた）複雑な関数関係に対しても、3 層のネットワークがその関数関係をどの程度学習（表現）してくれるか？」である。

3. テスト関数を用いた評価結果

式 (2) のコスト関数 V を最小とする方法として高次元アルゴリズムを用いた。また、本手法の性能評価のため、 $n_1(n_3) = 20, 30, 60, n_2 = 60, 90, n_s = 30, 60, 90, 150$, また、学習した関数をテストする生徒の（写像を構築するために使った集団（その数は n_s ）と異なる）集団数を 1000 程度取る。計算の結果、

(1) 一般的に n_s を多くすれば写像の精度が上がること、

(2) 収束は比較的良いと思われる — コスト関数 V が小さい値を取る — 場合でも、逆に写像の近似が悪い場合が（予想に反して）起こること、などを報告する。

今後、入力として言語（数量化して）を使い、これによって引き起こされる事象に対応してゆくための、隠れた因果関係を写し取る写像の構築の検討を行う。

参考文献

[1] 河野、下川：「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」, 情報処理学会第 62 回全国大会発表予定 (2001)6K-08.

Performance of a Method of Learning in Neural Networks: Results for a case of deviation (dispersion) functions

SHINJO K., OHTAWARA Kazushige

ATR Adaptive Communications Research Laboratories