ニューラルネットのＨＡ学習におけるモデル設定法

５Ｑ－０１

河野 芳江　　倉持 裕
ＡＴＲ適応コミュニケーション研究所

１．はじめに

パーセプトロン型ニューラルネットワークの学習の高速化・大規模化を狙い、高次元アルゴリズム（ＨＡ）[1]と呼ばれる大局解探索手法を学習アルゴリズムとする手法を検討している。与えられた問題を解く際、中間層の構成法、すなわち、中間ユニット数およびユニット演算処理タイプの選択法が、学習の高速化や結果として得られるネットワークの単純化などの観点から重要になる。中間ユニット数は、少なすぎると機能を達成できず、逆に、必要以上に多くとも機能は大きく改善しない上に、学習や使用の際に計算コストがかかり、また、その複雑さゆえ結果の直感的な解釈が難しくなる。所望の機能を達成できる必要最小限の規模のネットワークを効率的に構成できる手法が望ましい。そこで、本研究では、ＨＡ学習法において、中間ユニット数をひとつから始め、目標精度が得られるまで、ひとつづつユニットを増やすことにより、適切な中間層構造を自動決定する一手法を試み、その有効性を議論する。

２．ネットワークモデル

３層のネットワーク考える。入力層、中間層、出力層のユニット数を各々n_1、n_2、n_3、入力－中間および中間－出力ユニット間の各結合重みを各々、ω_{ij}とΩ_{jk}、中間層および出力層の各ユニットの閾値を各々、θ_jとΘ_kとする。いま、n_s個の教師信号が与えられているものとし、ｓ 番目の教師信号の入・出力値を各々、X^s_i、 Y^s_kと表す。

ｊ番目の中間ユニットの入力を$\eta_j(x_i;\omega_{ij},\theta_j)$とし、これにシグモイド関数を出力関数として適用する。これにより、方程式$\eta_j=0$を満たす識別曲面を境に、単調でなだらかな遷移領域を経て出力が０と１の領域に入力空間が分けられる。一般に、ユニット入力として超平面識別に対応する積和演算が用いられるが、これは開いた特徴領域の表現には向いているものの、閉領域の表現には多くの中間ユニットが必要となり効率が悪い。もう１つの単純な識別法として、中心$(\omega_{1j},\omega_{2j},..,\omega_{n1j})$、半径$|\theta_j|$の超球面

$$\eta^s_j = -\theta^2_j + \sum_{i=1}^{n_1}(X^s_i - \omega_{ij})^2 \tag{1}$$

を識別表面とする方法がある。現在まで、超球面識別は閉領域を効率的に表現できるばかりでなく、入力空間の定義域において曲面の一部分だけを使うことにより、開領域の表現も無理なく可能であることを示した[2]。そこで、広いパラメータ範囲を解探索する超球面識別ユニットを開領域・閉領域の両方に対応可能な探索ユニットとして採用し、機能実現に要する中間ユニット数の削減を狙う。探索の結果、定義域において近似的に超平面と見なされる場合には、必要ならば、それを積和演算の形に書き直すのは簡単である。これにより、中間層において、閉領域の表現に有利な超球面識別ユニットと開領域の表現に有利な超平面識別ユニットの２つのタイプが混在する混合ネットワークの逐次的構成が実現できる。中間ユニットの出力は、

$$\xi^s_j = f(\eta^s_j) \equiv \frac{1}{1+\exp(-\eta^s_j/T_j)} \tag{2}$$

で与える。ここでは、シグモイド温度T_jも最適化の対象とした。

出力ユニットの入力には積和演算をとり、入力をそのまま出力することとする。

$$\zeta^s_k = -\Theta_k + \sum_{j=1}^{n_2}\xi^s_j \cdot \Omega_{jk} \tag{3}$$

コスト関数を、教師信号出力値と実際の出力の二乗誤差の和

A model determination method in a layered neural network by Hamiltonian algorithm
KOHNO Yoshie, KURAMOCHI Yutaka
ATR Adaptive Communications Research Laboratories

$$V(\omega_{ij},\theta_j,\Omega_{jk},\Theta_k)=\sum_{s=1}^{n_s}\sum_{k=1}^{n_3}\left(\zeta_k^s-Y_k^s\right)^2 \qquad (4)$$

とし、これが与える精度 ε より小さくなるようにネットワークを構成する。

3．計算方法

下記の条件(1)～(4)に基づき、中間ユニット１つから最適化を始め、指定精度を達成するまで中間ユニットをひとつづつ追加する。各回において最適化し終えた中間ユニットのパラメータは固定し、追加したユニットのみを学習することにする。

(1)中間ユニットパラメータの初期化

・超球の中心座標 ω_{ij}：　[0, 1]の範囲の乱数

・超球の半径 θ_j：　[0, 0.5]の範囲の乱数

・シグモイド温度 T_j：　0.5

(2)中間ユニットパラメータの探索

局所解に捕まりにくい最適化手法であるＨＡを用いて、最適解のグルーバル探索を行う。開領域にも対応できるよう、パラメータ探索範囲を下記のように広めに設定する。下記範囲をはずれて運動した場合にはペナルティを課す。

・超球の中心座標 ω_{ij} の探索範囲：　[－1, 2]

・超球の半径 θ_j の探索範囲：　[0, 1]

・シグモイド温度 T_j の探索範囲：　[0.05, 1]

(3)中間ユニット追加の条件

最良解を更新した後の学習回数が指定回数 N_{max} を超えても最良解を更新できなかったら、その最良解をそのユニットの最適パラメータとして固定し、次のユニットを１つ追加する。

(4)学習終了条件

教師信号出力値と実際の出力の二乗誤差の和(4)が指定精度 ε を下回った時。

4．計算機実験

図１（等高線プロットおよび3次元プロット）に示す２入力１出力の関数近似問題を試みた。教師信号数 $n_s=50$、学習終了条件として教師信号ひとつ当たりの平均誤差 0.05、$N\max=100{,}000$ で、計算機実験を行った。

初期値により計算結果にややバラつきがあるため、３通りのランダムな初期値に対して計算を試みた。

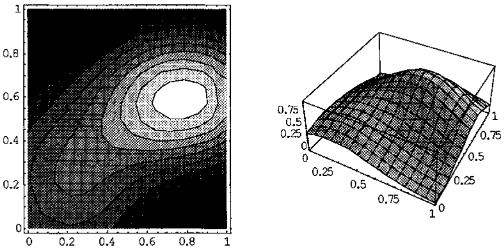

図１

いずれの場合も、最適中間ユニット数は３つであった。３試行の平均学習回数は約 1,870,000 回、平均 CPU 時間（COMPAQ Alpha Station XP1000 CPU: Alpha 21264 (667MHz)）は約 180 秒であった。

比較のため、中間ユニット数をあらかじめ３と与えて、ネットワークの全てのパラメータを一時に最適化する通常の方法でも、３通りのランダムな初期値に対して計算を行ってみた。その結果、平均学習回数は約 14,110,000 回、平均 CPU 時間は約 870 秒であった。

これらの結果から、中間ユニットをＨＡでひとつづつ個別に最適化してゆく方法は、学習時間の観点からも全ネットワークを一時に最適化する方法よりも優れていることが分かった。

5．おわりに

３層ニューラルネットのＨＡ学習法において、超平面識別タイプの中間ユニットをＨＡでひとつづつ個別に最適化してゆく手法を試みた。これにより、重要度の高いユニットから順に結果が得られ、目標精度を達成する必要最小限の規模のネットワークを逐次的に構成することができた。また、学習時間の観点からもネットワークのパラメータを全探索する方法よりも優れていることを計算機実験で示した。

参考文献

[1]新上和正:"高次元アルゴリズム", ｂｉｔVol.31, NO.7, p.2 (1999)

[2]河野芳江, 下川信祐：”高次演算処理ユニット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 63 回全国大会、2-71 (2001)