拡張型結合を持つニューラルネットワークのＨＡ学習

4N－3

倉持　裕　　新上和正　　下川信祐　　河野芳江
（株）ＡＴＲ環境適応通信研究所

１．はじめに

筆者らは、発見的解探索手法として高次元アルゴリズム（ＨＡ）を提案[1]し、各種組み合わせ最適化問題への適用を推進している。現在までに、ニューラルネット（ＮＮ）の学習への適用も示した[2]。これらをニューラルネットで画像をはじめとした人間の印象の分類に適応し、ＱｏＳ制御の指針として検討している[3][4]。ＨＡは多変数の解空間においても局所解に捕まり難いとうい特徴を持つ。本報告では、より高い能力を持つとされる積結合のように結合を拡張したＮＮ学習へのＨＡ適用例を示す。今回検討したＮＮは、数値の分散などを扱い易くする目的で、結合の和を分母分子に持つ分数型を考える。

２．神経素子の結合

Fig.1 に今回考えた神経素子を示す。この場合、結合荷重を表す変数の個数は通常の２倍となる。この神経素子を３層の Perceptron 型のネットワークとして構成し、その学習に高次元アルゴリズムを適用する場合を考える。

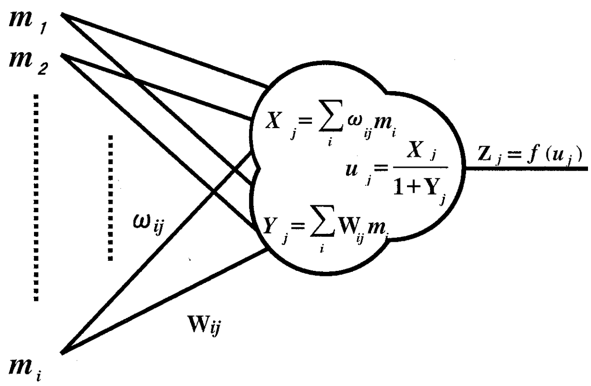

Fig.1 拡張型結合を持つ神経素子

A Method of Learning in Extended Neural Nets
by Hamiltonian Algorithm

KURAMOCHI Yutaka , SHINJO Kazumasa,
SHIMOGAWA Sinsuke, KOHNO Yosie

ATR　Adaptive Communications Research Laboratories

ユニットの入出力関数をシグモイド関数 $f(u)$とし、学習を定めるハミルトン関数を式(1)とする。

$$H = \tfrac{1}{2}(AP, P) + Q_1 + Q_2 + Q_3 \tag{1}$$

P は解探索軌道を決める力学変数で、入力－中間層間の結合荷重 ω, W 、中間－出力層間の結合荷重 v, V とすると、それぞれと正準共役な変数（運動量）を s , S , t , T として式(2)のベクトルとして表現される。ただし、入力ユニット数 i、中間ユニット数 j、出力ユニット数 k とする。

$$P = (s_{11},\cdots,s_{ij},S_{11},\cdots,S_{ij},t_{11},\cdots,t_{jk},T_{11},\cdots,T_{jk})^{\mathrm{T}} \tag{2}$$

A は運動量のミキシング[5][6]を決める行列である。ミキシングは、多変数ほど解探索効果が上がるＨＡの特徴をより増強するための手法である。

また、他の学習法と同様に、出力と教師信号 Yとの自乗誤差を最小とするように評価関数 Q_1 を構成する。L は教師信号数を表わし、u_k は出力ユニットの入力である。

$$Q_1 = \sum_L \sum_k \{Y_{kL} - f(u_{kL})\}^2 \tag{3}$$

さらに Q_2, Q_3 は、通常のＮＮにＨＡを適応[2]した際に導入した解探索軌道の発散を防止するための項であり、g は運動を制限する閾値である。θ を、ステップ関数とする。

$$Q_2 = \alpha\left\{\sum_i\sum_j D(\omega_{ij}) + \sum_i\sum_j E(W_{ij})\right\} \tag{4}$$

$$Q_3 = \alpha\left\{\sum_j\sum_k D(v_{jk}) + \sum_j\sum_k E(V_{jk})\right\} \tag{5}$$

$$D(c_{ab}) = (\pm g - c_{ab})^2 \theta(\pm g - c_{ab}) \tag{6}$$

$$E(c_{ab}) = (g - c_{ab})^2 \theta(g - c_{ab}) + (-g - c_{ab})^2 \theta(-c_{ab}) \tag{7}$$

３．学習実験

Fig.1 の神経素子で３層のネットワークを構成し、入力層(i)を 48 ユニット、出力層(k)を６ユニットとして、中間層のユニット数を様々に変化させて評価関数 Q_1 の様子を観察する。教師データは、既知の式(8)からランダムに発生させた入力ベクトル x と、その出力ベクトル y との組を L=10 組用いた。教師入力は３次元ベクトルを 48 次元へ拡張している。

$$y = Bx + b \tag{8}$$

式(1)を用いて学習させた場合の出力と教師信号との自乗誤差和 Q_1 の一例を学習の更新回数を横軸として Fig.2 に示す。ただし、中間層(j)を 10 ユニットとし、初期結合荷重 ω, W, v, V および初期運動量 s, S, t, T は、それぞれランダムに設定している。

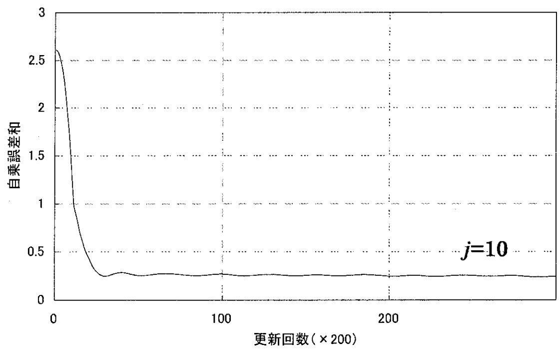

Fig.2 拡張型結合 NN への HA 適用例

比較として、式(9)の神経素子についての自乗誤差和 Q_1 の様子(j=10,20)を Fig.3 に示す。

$$u_j = \sum_i \omega_{ij} m_i \tag{9}$$

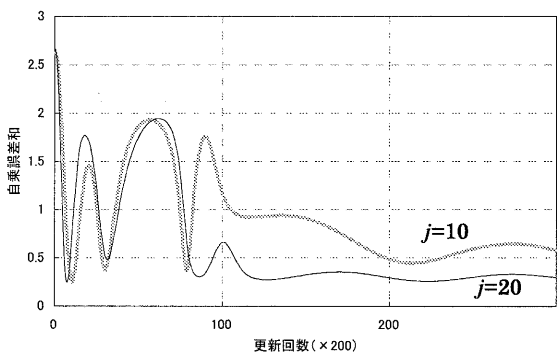

Fig.3 従来 NN への HA 適用例

変数（結合荷重）が増えたため計算量はかなり多くなるが、結合荷重の更新回数はそれほど増加しない上に、より誤差が少なくなっている。このことから、変数の次元が増えるほど効率的に解探索が行えるというＨＡの特徴と拡張型ＮＮ能力の相乗効果が伺える。

４．おわりに

前回の発表[2]で解探索手法である高次元アルゴリズム（HA）を用いたＮＮの学習法について検討し、今回、より複雑な構成の拡張型結合 NN についての学習について検討を行った。変数の増加による解空間の複雑さに対応するために、HA の運動方程式にミキシングの手法[5]を導入した。その結果、１回の更新に要する計算時間は多いが、局所解に捕らわれ難くいという高次元アルゴリズムの特徴が効果的に働いていることが観測された。また、計算量の増加は並列性が高い計算であることから容易に克服できるものと考える。問題点としては、ミキシング係数という新たなパラメータの導入が結合荷重の初期値依存性を軽減するという利点を持つ反面、ミキシングパラメータの設定に対する予備実験[6]という手間の増加は否めない。ただし、NN の構造が決定されれば、以後の学習には異なる教師データセットであっても同一のミキシングパラメータが使用できると考えられる。現在、異なるモノの間を繋ぐ写像の構成法を検討[4]しており、複雑さに柔軟に対応できる手法として、高次元アルゴリズムによる学習を利用して行く予定である。

【参考文献】

[1] 新上,山田,下川：”高次元化によるシステムの制御法”, 信学会総合大会 A-1-14,p14 (1997).

[2] 倉持,新上：”高次元アルゴリズムによるニューラルネットワークの学習”,情報処理学会第 60 回全国大会, p2-143(2000).

[3] 倉持,新上：”画像処理システムに要求される処理コストに関する考察”,情報処理学会第 59 回全国大会, p2-269(1999).

[4] 内田,新上,下川,石野：”動的過程における QoS 分析:車のﾄﾞﾗｲﾊﾞｰ特性と評価尺度の相関関係”, 情報処理学会第 60 回全国大会 1X-05(2000).

[5] 倉持,新上,下川：”高次元アルゴリズムの高速化に関する検討”, 信学会総合大会 A-1-9,p9 (1999).

[6] 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 4N-06 発表予定(2000).