

階層型ニューラルネット学習効率の入力パターン分解能依存性評価

1J-10

西村 治彦 小山 宣樹

兵庫教育大学

1. はじめに

階層型ニューラルネットの学習アルゴリズム「バックプロパゲーション(BP)法^[1]」では、パターン認識の学習能力を左右する重要な問題として、これまでアルゴリズム自体がもつパラメータ自由度の決定法が注目されてきた^[2]。しかし、入力パターンデータを含む全体系でネットワークを把えたとき、入力パターンデータの分解能の選択は、入力ユニット数の変化を通じて、ニューラルネットの構造自体を変化させることになる。同一学習対象に対する分解能変化が全体の学習効率にどのような影響を与えるのか。アルゴリズムに内在する他のパラメータ同様、検討を加える必要がある。

我々はこのような観点から、これまで研究を進めてきたが^[3]、ここでは、定量的評価の精密化のために新たに学習効率評価指数を導入し、それに基づく検討結果について報告する。

2. 実験方法および評価指数の導入

入力パターンを○□△×の4種類とし、それぞれをN行N列の格子上升目の■、□状態の集合として表現する。これらは N^2 個の入力層ユニットの1, 0状態と対応している。分解能Nとしては10, 20, 30, 40, 50の5種類を選択し、各分解能毎に3層構造ネット(各層ユニット数: N^2 , 5, パターン数)でBP法のアルゴリズム(逐次修正法)に従って学習させる。その他のパラメータについては、シグモイド関数の傾き係数は0.75、荷重・しきい値の修正/忘却係数は1:1、荷重・しきい値の初期化は一様乱数($\leq |0.2|$)とした。

実験環境としてはニューロエミュレータ

Neuro-07(NEC PC-9801上)^[4]をベースにシステム構築している。

図1に100回(図形1セットで1回)学習における誤差2乗和E(出力ユニット値と教師信号値の差の2乗)の経時変化の例を示す。縦軸はログスケールである。

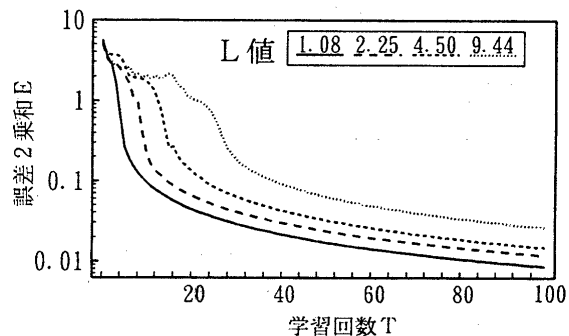


図1 誤差2乗和Eの経時変化の例

これらの学習曲線から、それぞれの学習効率の良否を定量的に評価するため、我々は、これまでのEの経時的加重平均値Fの考え^[3]の一般化に加え、さらにEの変化率に対する経時的加重平均値Gを導入する。

$$F_n = \frac{\sum_{i=1}^n E_i \sqrt{T_i}}{\sum_{i=1}^n \sqrt{T_i}} \quad (n=1,2,3,\dots)$$

$$G_m = \frac{\sum_{i=1}^m \{(\log_{10} E_{i-1} - \log_{10} E_i) / (1/\sqrt{T_i})\}}{\sum_{i=1}^m (1/\sqrt{T_i})} \quad (m=1,2,3,\dots)$$

図2は、100種類の荷重初期値(乱数発生による)サンプルに対する学習結果をもとに、そのF-G相関($n=2, m=3, r=100$ の場合)を図示したものである。G→大、F→小のとき学習効率→良という負の相関がよく現れている。このFとGの比から最

Dependence of Learning Efficiency on Pixel Size for Input Pattern Data in Multilayered Neural Network

Haruhiko NISHIMURA, Nobuki KOYAMA
Hyogo University of Teacher Education

終的な学習効率の評価指数

$$L = F / G$$

を定義する。図1の学習曲線のL値はそれぞれ1.08, 2.25, 4.50, 9.44となり、その全体的な学習効率が定量的に評価(L→小ほど良)されている。

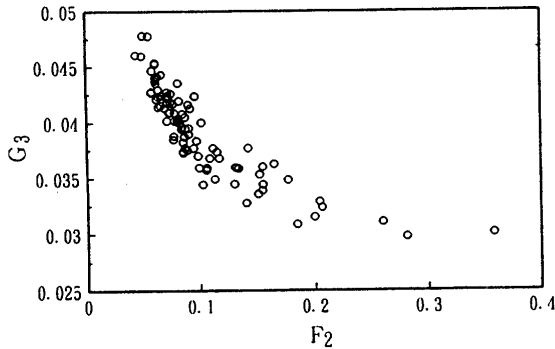


図2 F-G相関図(100回学習, N=30の場合)

3. 結果

図3は、分解能毎に100サンプルの荷重初期値に対して学習を行なわせ、そのL値の累積頻度をグラフ化したものである。N=6, 10, 50に比べてN=30のL値は明かに低域に多く分布している。また、N=50では、他に比べて累積頻度が85あたりで頭打ちとなり、学習収束性の悪い場合の頻度が増えてきている。これによりN=30の場合に最も学習効率がよいことが確認できる。我々は、この他にも、学習後のネットワークの

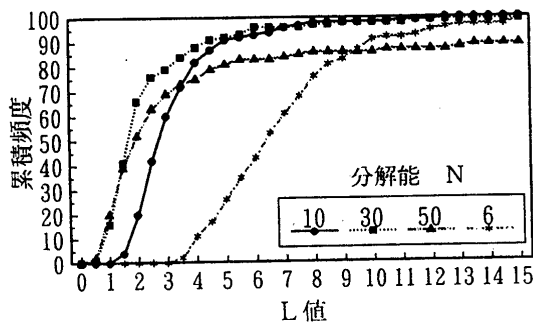


図3 分解能毎の学習効率評価指数Lの累積頻度(サンプル数各100)

ノイズに対するロバストネスについても、その分解能依存性を検討している。

4. おわりに

以上、今回のL値による検討の結果からも、学習効率の入力分解能依存性と最適分解能の存在が確認された。同様の結果は入力パターンの種類や個数の若干の変更に対しても確認されている。

この性質は、図4に示すように、分解能向上に伴う両者の競合によって説明できると思われる。この傾向の定量的および一般的定式化と適正分解能の決定アルゴリズムの検討が今後の課題として残されている。

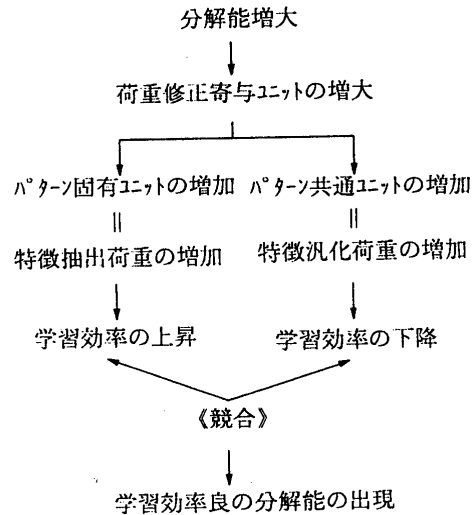


図4 分解能増大に伴うユニット間競合

参考文献

[1] Rumelhart, D. E., Hinton, G. E., and Williams, R., Learning Representations by Back Propagating Errors, Nature, 323, 9(1986) 533-536.
 [2] Hecht-Nielsen, R., Neurocomputing, Addison Wesley(1990).
 [3] 西村, 小山: BP法パターン認識効率の入力データ分解能依存性, 情報処理学会第44回全国大会講演論文集(2), (1992)125-126.
 [4] Neuro-07 ユーザーズマニュアル, 日本電気インフォメーションテクノロジー(株), (1990).