

2D-4 L V Q を前処理として用いた バックプロパゲーション学習の改良法

田中 武久 中 基孫 吉田 邦夫

松下技研株式会社

1. はじめに

ニューラルネットワークの学習ではバックプロパゲーション(BP)が広く研究され、様々な改良BPが提案されているが、学習パターンの提示の仕方に関するものは少ない。

本稿では、BP改良法のDSCサーチ法に競合学習系のLVQを前処理として用いてDSCサーチ法の学習の効率を高める方法を提案し、実験により確認した。

2. DSCサーチ法^{1), 2)}

DSCサーチ法は、非線形最適化法から提案され、BPの数倍以上の速さで収束する。以下に手順を説明する。

- ①重みベクトルを固定したままで、従来のBPと同様に最急降下方向を全ての学習パターンに対して計算し、それを平均する。

- ②①の平均した最急降下方向で誤差評価関数を最小にする重みを、Davies, Swann, Campeyの提案した線形探索法(DSCサーチ)により求め、重みを変更する。

- ③定めた基準値内に誤差評価関数値が入るまで、①、②をくりかえす。

3. LVQ (Learning Vector Quantization)³⁾

LVQは自己組織的な競合学習を、教師あり学習と組み合わせたもので、LVQ1とLVQ2が提案されている。本稿ではLVQ1を用いた。

LVQ1では、学習サンプルの入力ベクトルと同次元の参照ベクトルが一定数あり、N個のカテゴリのいずれかに属している。動作時には、入力ベクトルに最も近い参照ベクトルの属するカテゴリが、入力ベクトルの属するカテゴリとなる。以下、学習の手順を説明する。

- ①各参照ベクトル m_i の初期値を設定する。ただし m_i の値はなるべく正しい値になるようにする。

- ②学習サンプルから、学習するパターンをランダムに選び、学習パターンの入力ベクトル x と各参照ベクトルとの距離 $\|x - m_i\|$ を計算する。

- ③ $\|x - m_i\|$ の中で最小のものを探し、その m_i を m_c と置く。

- ④ m_c の値を以下のように変更する。

x と m_c が同じカテゴリに属しているとき

$$m_c(t+1) = m_c(t) + \alpha(t)(x(t) - m_c(t))$$

x と m_c が違うカテゴリに属しているとき

$$m_c(t+1) = m_c(t) - \alpha(t)(x(t) - m_c(t))$$

ただしここで α は、 $0 < \alpha < 1$ を満たし、時間とともに単調に減少する係数である。

- ⑤収束するまで②から④を繰り返す。

LVQは学習が簡単で高速な上に、ローカルミニマの

問題が無い。

4. DSCサーチ法、LVQの課題

DSCサーチ法の課題として次の点があげられる。

- ①初期値、学習パターンの提示の仕方により学習の過程が異なり、ローカルミニマに落ち込むこともある。汎化を期待する場合には、学習の結果の状態は直接ネットワークの性能を左右する。

- ②問題に最適なネットワーク規模の決定が困難である。

- ③学習サンプルの数にほぼ比例した時間がかかる。

また、LVQの課題としては次の点があげられる。

- ①参照ベクトルの初期値は、かなり良い状態にしなければならない。

- ②LVQはルックアップテーブルを作るだけなので、参照ベクトルが非常に多くなる場合がある。また、連続関数を出力するには不向きである。

- ③入力ベクトルに最も近い参照ベクトルを探す時、時間がかかる。

5. DSCサーチ法とLVQの組合せ

4. で述べた課題を解決する1つの方向を示すため、本稿ではDSCサーチ法の学習サンプルの前処理にLVQを用いることを提案する。ここで提案するアルゴリズムは以下の通りである。

- ①多数の学習サンプルをカテゴリに分け、LVQで学習する。入力をカテゴリ分けする問題(パターン分類問題)ではそのカテゴリを用いる。

- ②LVQ学習後、LVQの参照ベクトルを学習パターンとして、DSCサーチ法で学習を行なう。

本方法の長所として次の点があげられる。

- ①学習サンプルの数を少なくできるため、DSCサーチ法を高速にできる。

- ②参照ベクトルの分布から学習パターンの構造がわかり、DSCサーチ法を用いるネットワークの規模を見積ることができる。

- ③参照ベクトルの分布から、学習サンプルを事前にスケールリングすることができる。

- ④特に誤差の多いカテゴリを集中して学習するなど、学習サンプルの提示の仕方を工夫しやすい。

また、本方法の問題点として次の点があげられる。

- ①LVQとDSCサーチを両方行なうので、計算量が増加する。

- ②学習サンプルの分布の仕方が、DSCサーチの学習に適している保証はない。

- ③カテゴリ分けされていない学習サンプルをどのようにカテゴリ分けするかが示されていない。

6. シミュレーション実験

5. で述べた方法を検証するため、シミュレーションを行なった。

用いた学習のサンプルは、2次元パターン分類問題で、各カテゴリの学習パターンは表1の平均値と分散で正規分布している。LVQの参照ベクトルは、900個の場合と270個の場合を行い、各カテゴリの参照ベクトルの数は等しくした(カテゴリII、IVは他のカテゴリの倍)。

参照ベクトルは表1の分布により正規乱数で作成し、学習サンプルは計10000個正規乱数で作成した。

LVQの学習では上記学習サンプルを計600回提示した。図1に参照ベクトル(270個の場合)の初期値の分布、図2に学習後の参照ベクトル(270個の場合)の分布を示す。また、学習サンプルに対する誤りの率を表2に示す。

次に初期値の参照ベクトル、100回学習後の参照ベクトル、600回学習後の参照ベクトルを用いて、DSCサーチ学習を入力層2個、中間層10個、出力層7個の3層パーセプトロンで行なった。ただし、出力層の7個の各ユニットを各カテゴリに割り当て、最も出力の大きいユニットのカテゴリを入力パターンに対する出力とした。そして、2000回、4000回、14000回学習後の重みを用いて、LVQで用いた学習サンプルに対する出力を計算し、誤りの率を計算した。その結果を表3に示す。

参照ベクトルの数が900個の場合、DSCサーチ学習後のネットワークの性能は、LVQにより学習した場合とランダムにとった場合ではほとんど差がない。しかし、

参照ベクトルの数が270個の場合は明らかにLVQ学習後の学習サンプルを用いた方が優れている。

7. 今後の課題

今後の課題として以下の点があげられる。

- ① LVQ1だけでなく、自己組織的特徴マップやLVQ2などのモデルを用いた学習を行なう。
- ② LVQ学習後の参照ベクトルの分布から多層パーセプトロンの規模を決める方法を検討する。
- ③ 入力パターンだけでなく、出力パターンを含めた前処理を検討する。

8. まとめ

BPの改良法であるDSCサーチ法の前処理にLVQを用いることにより、DSCサーチ法の学習効率が向上することを示した。

- 1) 高木他、"非線形最適化手法を用いたニューラルネットワーク学習アルゴリズムの高速化"、89'信学全大、SD-1-12
- 2) 阪上他、"ニューラルネットワークの学習における非線形最適化法の評価検討"、89'信学全大、SD-1-13
- 3) T.Kohonen他、"Statistical Pattern Recognition with Neural Networks:Bechmarking Studies"、ICNN Vol.1, San Diego, July24-27, 1988, 61-68

カテゴリ		I	II-1	II-2	III	IV-1	IV-2	V	VI	VII
x	平均	3.0	14.0	11.5	26.0	6.5	22.5	16.5	4.0	22.5
	σ	1.5	4.0	1.75	2.0	3.25	1.25	1.75	2.0	3.75
y	平均	2.5	2.5	0.5	2.5	1.5	1.5	1.5	0.5	0.5
	σ	0.25	0.25	0.25	0.5	0.25	0.25	0.25	0.25	0.25

表1 各カテゴリの学習サンプルの平均と分散

学習回数	0	100	300	600
900個の場合の誤り率(%)	9.07	6.84	5.81	5.44
270個の場合の誤り率(%)	12.0	6.21	5.44	5.42

表2 LVQ学習の結果

参照ベクトルの数	900			270		
	0	100	600	0	100	600
参照ベクトルの学習回数						
2000回学習後の誤り率(%)	36.2	35.3	36.0	36.2	34.0	33.0
4000回学習後の誤り率(%)	27.2	26.1	27.0	25.4	23.4	16.2
14000回学習後の誤り率(%)	6.39	6.03	6.04	12.0	6.92	6.19

表3 DSCサーチ学習の結果

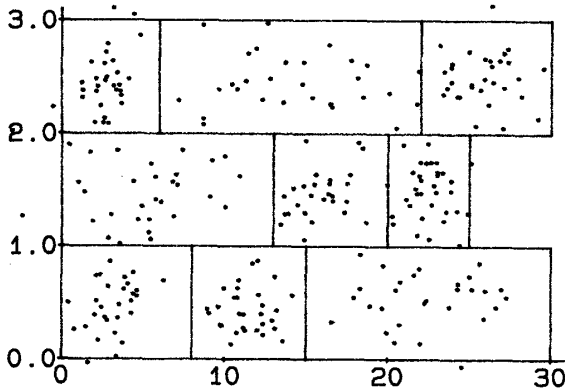


図1 LVQ学習前の参照ベクトルの分布

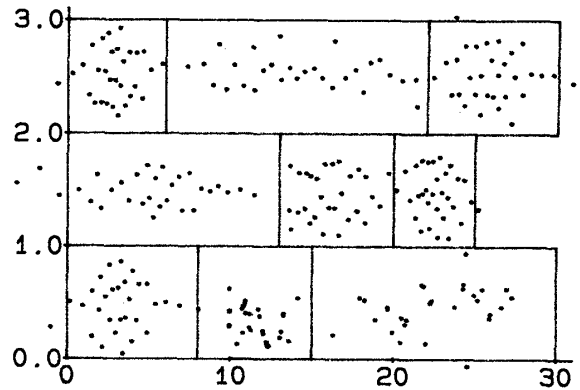


図2 600回LVQ学習後の参照ベクトルの分布