

日本語母音認識ニューラルネットワークにおける
局所的最小値を抜け出すための前進後退学習法*

1W-5

飛弾 和臣*、北澤 茂良*

静岡大学 理工学研究科

1 はじめに

本稿では、局所的最小値 (local minimum) を抜け出す一提案を日本語母音を認識を例として報告する。学習則として広く利用されている逆誤差伝搬法は最急降下法を基本とすることから、地域的最低値に陥ってしまい、それ以上学習が進まない場合がある。仮に上手く学習できた場合でも、重みの初期設定を変えてやると、収束しない場合もありえる。そこで地域的最低値を陥りにくくするのではなく、上手に抜け出す方法を提案し検討する。

2 学習則

多層ネットワークの学習則として最も広く使われている逆誤差伝搬法学習法はRumelhartにより提案されたもので、教師付き学習方法である。

これは、多層ネットワークで、与えられた入出力関係を満たすようにニューラルネットワークを組織化させる一つの手法である。実際の出力信号と教師信号の平均二乗誤差を評価関数とし、これが最小となるように、ネットワークの結合重み係数を調整する。この誤差を最小にする解法のアプローチは、最急降下法を元にして重み変更量を決定する。この逆誤差伝搬法では、誤差値による重み係数の修正量の幅を決定する学習係数 η や、前回の修正量を調整するモーメント係数 α といったパラメータを決定しておく必要がある。このパラメータの値は、対象としているタスク、サンプルデータの量、学習進行の程度によっても、エラー空間の形状が、変化するために、経験的に決める以外に方法が無い。

そのため、このバックプロパゲーション法による学習の性能向上を目指して、さまざまな工夫が提案されてきた。例えば、収束性能向上のための重み初期値の決め方やパラメータの調整方法等がそれである。しかしながら、逆誤差伝搬法は最急勾配法を基本として学習を進めるため、本質的に、局所的最低値に陥ることなく大域的最低値 (Global minimum) への収束を保証した訳でない。そこで、地域的最低値に陥るのを避けることは出来ないの、地域的最低値に陥った状態から、一時的に学習を後退させることで、地域的最低値を抜け出す手法を提案する。以後、前進後退学習法 (Forward Backward Learning Method) と呼ぶことにする。

3 局所的最低値での学習

前回の学習の誤差値と今回の学習の誤差値がある程度変わらない状態では、学習が収束しているの、この後いくら学習しても学習は進まない。そこでさらに学習を進める為、式 (1) を満たす状態になったら、学習本来の誤差値が減少する方向に結合重みを変更するのではなく、一時的に誤差値が増加する方向に結合重みを変更するように学習方法を切替えるのである。

$$0 \leq E_2 - E_1 < \epsilon \ll 1 \quad (1)$$

E_2 : 前回学習した結果の誤差値

E_1 : 今回学習した結果の誤差値

学習を後退させる際、全てのサンプルパターンにこの手法を適用するのはあまり良い方法とはいえない。式 (1) を満たすような局所的最低値に陥った状態では、当然、教師信号と同じ出力値を出すことの出来るサンプルパターンと、教師信号とは違った出力値を出すサンプルパターンが存在する。その為、この両サンプルパターンを同じ様に学習対象として扱っては、期待しているような、違った場所への収束は出来ない。

そこで、教師信号値と同じ出力を出せないものに関しては依然誤差値が減少する方向に結合重みを変更し、教師信号と一致するものみ、あたかもエネルギー曲面の坂を登るかの様に、誤差値が増加する方向に変更するのである。この誤差値を増加させる後退学習も式 (1) を満たし学習が進まなくなる。その時は、学習方針を誤差値が増加する方向への後退学習から、本来の誤差値を減少する方向への前進学習へと切替える。場合によっては式 (1) を満たさず、誤差値が増え続ける場合も考えられるため、ある程度の誤差値に達した場合も同様の処理を行なう。以上、一連の処理を繰り返すことによりさまざまな収束場所に収束を繰り返すことが可能となる。

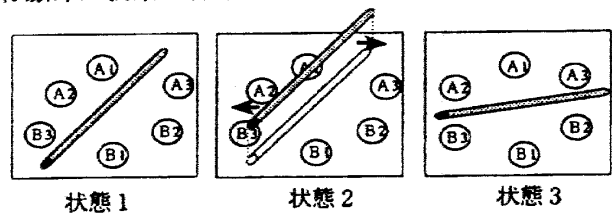


図 1: イメージ図

一度局所的最低値に収束しても (状態 1)、反対に学習をすすめることで収束状態を抜け出し且つ、教師信号と違う出力値を出力するサンプルパターン (A 3 と B 3) が依然学習を進める為 (状態 2)、A 3 と B 3 を満たす様な収束場所へ収束する (状態 3)。

*Foward Backward Learning Method Extricable Local Minimum applied to a Neural Network for Japanese Vowel Recognition. By Hida Kazuomi and Kitazawa Shigeyoshi (Shizuoka Univ.)

4 実験

今回提案した局所的な最小値から抜け出して、新たな場所へ収束することを確認するために、予め、局所的な最小値に陥ることが分かっている初期重みを用いて、以下の条件で実験を行ない、提案した方法が期待通りの結果をだすかを調べる。

4.1 日本語母音認識ニューラルネットワーク

この効果を調べる為に図の様な4層の階層型ニューラルネットワークを用いる。中間層2は3ユニットから成り、弁別素性(左のユニットから compact/diffuse, grave/acute, flat/plain) に対応させることを意図している。出力層は、日本語の5母音(左のユニットから /a,i,u,e,o/) に対応している。

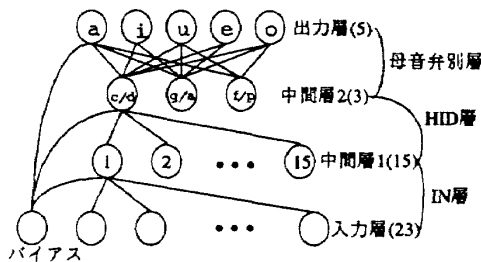


図 2: 日本語の母音認識の階層ニューラルネットワーク

4.2 音声資料

入学習データは、日本人成人男性 82 人が発声した、5 母音 /a,i,u,e,o/ の 5 種類に加えて、5 母音と破裂子音 /p,t,k,b,d,g/ の組合せ CV 単音節 30 種類の単音節を、サンプリング周波数 16kHz で採集した。そして最大振幅の 18% 以上の部分を切りだし、窓長 256 点を 15 フレーム取り出し、23 次の線形予測係数から 256 個のパワースペクトルを作って dB 単位に直した。次にこれを 22 区間の臨界帯域幅に区切り、60dB 等ラウドネス曲線の補正を行なった。最後に -0.5 ~ 0.5 以内に正規化したものを使用した。結果として 43050 の学習データを作成する [2][3]。

4.3 実験結果

以下の条件で実験を行なった。

- 条件 学習パラメータ $\alpha \cdot \eta$ の組合せ数 6 通り
- α (0.1,0.9) η (0.01,0.1,0.5)
- サンプル数 100 個 (作成した 43050 個のサンプルパターンからランダムに選択)
- 初期値は共通のものを使用
- 学習回数 1000 回

表 1: 各要素の認識表

		出力値				
		あ	い	う	え	お
教師信号	あ	14/14	0/0	0/0	0/0	0/0
	い	0/0	24/24	0/0	0/0	0/0
	う	0/0	0/0	23/23	0/0	0/0
	え	0/0	0/0	1/0	21/22	0/0
	お	0/0	0/0	0/0	7/0	10/17

前進のみの学習則 / 前後進学習則

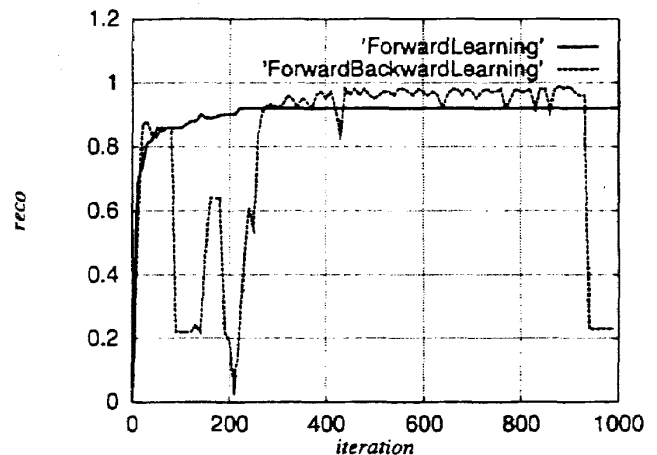


図 3: 学習回数と認識率との関係

4.4 評価

学習回数が約 50 回付近では、ともにある収束場所に向かって学習が進んでいることが分かる、それ以降従来の局所的な最小値を抜け出す方法を探り入れていないものはそのまま収束に向かってのに対して、新しく提案したものは、100 回 ~ 300 回辺りまで、収束場所から離れ、再学習を行なっていること確認出来る (図 3)。最終的に、従来のアルゴリズムでは、最初収束した場所付近を探索し続けた結果 92.0% で学習を終了した。これに対し、新しいアルゴリズムでは、1000 回の学習の過程の中で認識率 99.0% という結果を得た。

また、各サンプルデータの教師信号と、実際の出力結果との対応表から、そのまま局所的な最小値に収束してしまった結果が、日本語母音の「お」を 17 サンプル中 10 サンプルデータしか正確に認識しなかったのに対して、この新しいアルゴリズムでは、17 サンプルデータ全て認識することが出来た。このことから、最初に収束した収束場所から別の新たな収束場所へ収束したことが考えられる。

5 まとめ

今回はこの提案した手法の効果が顕著に現れるよう、サンプルパターンの数を減らしたが、本来この研究で使用している学習サンプルパターン 43050 個でも試してみたい。

謝辞

ATR で開発された DCP 2 [1] を使用させて頂いた。

参考文献

- [1] ATR, DCP2, Technical Report '89(1989.9).
- [2] 新村貴彦, 上村大一郎, 大池盛一郎, 北澤茂良, "母音と子音の不変な音響特徴の抽出", 電子通信学会 SP90-80, 1991.
- [3] 中尾充宏, 北澤茂良, "母音認識ニューラルネットワークによる母音の弁別素性の抽出", 電子情報通信学会 SP94-70, 1994.