Kohonen自己組織化ニューラルネットワークにおける学習短縮化アルゴリズム

○中野 峰司, 村田 厚生 (福岡工大)

Algorithm for Learning Reduction in Kohonen's Self-Organizing Neural Network
Takashi NAKANO and Atsuo MURATA
(Fukuoka Institute of Technology)

1. まえがき

Kohonen特徴写像は、教師信号無しで学習する2層ニューラルネットワーク (NN: Neural Network) の一つであり、入力信号のパターンを反映した特徴写像を生成する。しかし、この自己組織化NNでは学習収束までに、多くの試行回数を要する。本研究では、このNNの学習過程を分析し、学習収束にまでの演算回数を減少させるアルゴリズムを提案し、その有効性を示した。

2. 自己組織化ニューラルネットワーク

基礎となるKohonen 特徴写像として、入力ユニットの個数が 2 で、 8×8 の格子状の競合 (出力) 層を有する2層NNを取り上げた。学習率 α と近傍距離 d は、試行回数の単調減少の関数とし、入力ベクトルは、一様分布に従うと仮定した。このNNを計算機上で実現した結果、学習収束に至るまでかなりの計算時間を要した。また、学習率と近傍距離の初期値 α_0 と d_0 がそれぞれ 0.2 と 4 の場合には、非常に高い確率でNNが安定した学習を行うことが確認された。そこで、これらの初期値を持つNNを対象とし、総学習回数 $T=20000$ 回として、入力ベクトルとwinning unitの荷重値の距離、すなわち誤差の時間的変化に基づいて学習過程を分析した。各試行での誤差はかなりばらつくので、誤差の変化の特徴を明らかにするために、40回の繰り返し毎に誤差の平均値を計算し、その経時的変化を観察した。この結果を図1に示す。誤差は、繰り返し回数 $t=5000$ まで ($d=4$)、 $t=10000$ まで ($d=3$)、 $t=15000$ まで ($d=2$)、 $t=20000$ まで ($d=1$) の4段階で推移し、 d によって大きく影響を受けていることがわかる。また、図1より平均誤差がほぼ安定した値をとるようになっても、 d はある回数行われるまで減少しないため、各 d の値での試行が必要以上に行われていると判断される。実際のNNの状態をさらに詳細に検討した結果、ある試行回数 t で安定した学習が行われていたとしても、 d が減少しないために、学習状態が保たれない場合が観察された。以上より、NNでは d の減少するタイミングを適切に把握することが、学習収束までの演算回数の短縮化につながるのではないかと考えられる。

3. 学習短縮化アルゴリズム

まず、平均誤差の変化に伴う学習率の変化を観察した。平均誤差と α の関係を図2に、平均誤差と d の関係を図3に示す。近傍距離と学習率の関係に基づいて、

図1 t と平均誤差の関係 (Kohonenの方法)

図2 平均誤差と近傍距離 d の関係

図3 平均誤差と学習率 α の関係

dが4から3に変化する時(点d)の α が0.15(点l)、dが3から2に変化する時(点c)の α が0.1(点k)、dが2から1に変化する時(点b)の α が0.05(点j)となるように α を選択した。各dでの平均誤差とその標準偏差を求めた。実際には、10回のシミュレーションを実施し、各ブロックで40個の平均誤差に基づく125個の平均値を求め、これを総シミュレーション回数で平均化した。各ブロックでの平均誤差の標準偏差も同様に求めた。dの値を変化させる平均誤差の境界の値を平均誤差の値にすれば、学習状態が多少悪くてもdは急速に減少すると考えられるため、得られた平均誤差から標準偏差を引いた値(正規性を仮定すればdが減少する確率は16.7%)を各dに対する境界の平均誤差とした。例えば点c(図2)、k(図3)はdが3から2に推移する境界であるため、図1のdが3の時の平均誤差の平均0.141622-標準偏差0.012371=0.129251を平均誤差とした。本研究では、以上のようにしてパラメータ α とdを求めた。シミュレーションでは、40次の移動平均を平均誤差とした。繰り返し回数 $t \leq 0$ では、誤差を0.2とした。

4. シミュレーション結果および考察

d=2~4では、予想以上に速い学習が行われたが、d=1に入ってからNNはかなり不安定で、dの値が2になる状態も観測された。d=1での α と平均誤差の値は、それぞれ0.05と0.082597(点j)で、以上のアルゴリズムでは誤差平均の値が0で α が0(点g)としていた。しかし、図1のd=1での平均誤差と標準偏差は、それぞれ0.062168と0.006783で、平均誤差は0にならない。そこで、d=1における α の値を、平均誤差=0.082597で0.050(点j)、平均誤差=0.062168で0.010(点i)、平均誤差=0.05で0.005(点h)、平均誤差=0で0(点g)となるように線形補間してシミュレーションを実施した結果、d=1での学習の安定性が改善された。また、平均誤差が0.200以上では、平均誤差=1.00(実際にはほとんどあり得ない)までd=5(点f)とした。 α に関しては、平均誤差が0.200で0.200、平均誤差が1.00(実際にはほとんどあり得ない)で0.300(点n)となるように線形補間した。もし、この平均誤差が0.200を超える場合に平均誤差と α やdの関係を表す直線mnとefの傾きが大きければ、NNは非常に密集した点の塊となって動きまわって発散し、仮に発散しなかったとしても、d=3や4に入るまでに時間を要し、その間不安定な動きをする。上述のように、平均誤差が0.2以上での α とdを設定すれば、安定した学習が可能になる。

本アルゴリズムに従ってシミュレーションを実施し

図4 tと平均誤差の関係(提案法)

図5 tと平均誤差の関係(提案法の一例)

た結果得られた試行回数と平均誤差の関係を図4に示す。学習は5000回の時点では、終了していることがわかる。また、 $t=0 \sim 3000$ の平均誤差は指数関数的に減少し、学習が速く行われた。d=1での平均誤差と標準偏差は、Kohonenのアルゴリズムでは $t=18001 \sim 20000$ で、それぞれ0.060186、0.004931で、本アルゴリズムでは $t=5001 \sim 20000$ で、それぞれ0.059248、0.004981であり、ほぼ同様の成績が観察された。

図5は図4を求める前の個々のtと平均誤差の関係を示した一例である。図2より平均誤差が0.147595を超えるときd=3、超えないときd=2となるため、d=2に推移するのに図4の平均より多くの時間を要することが観察される。これは、d=3において良い学習状態になりにくかったためか、もしくは入力の変りが大きかったためではないかと推察される。また、同時にdの値が数回増減する現象が観察される。これは、学習状態があまり良くないのにdが減少したためか、入力の変りが大きかったためではないかと考えられる。以上のように、本アルゴリズムでは、よい学習状態でdが減少する確率が高いため、状態の悪いまま学習が進みにくくなり、悪い学習状態では、dを増加させ、不安定な状態を速やかに修正するようにNNが機能することがわかった。