

ミニバッチ学習 SOINN-AML の提案

日大生産工 ○小村 哲矢 日大生産工 山内 ゆかり

1. まえがき

実世界のデータは非定常な場合があり、それらに対応するために、追加学習を行う機能が必要となる。そこで自己増殖型ニューラルネットワーク(SOINN)[1]を提案した。

SOINNはSelf-Organizing Map(SOM)とGrowing Neural Gas(GNG)に着想を得て構築した教師なし学習手法の一種である。この手法は事前にネットワーク構造を決定する必要がないほか、高いノイズ耐性を有している。また、ラベル付けされていない分布を表現することができ、適切なノード数などの事前条件なしに各クラスタの典型的なプロトタイプパターンも与えることができる。このような特徴からSOINNは教師なし学習の様々な応用に用いられている。また、SOINNの発展としてDistance Metric Learning (DML)[2]とSOINNを組み合わせるSOINNの欠点である高次元の入力データを扱う場合、性能が落ちる点を改良しているSOINN-AML[3]が提案されている。

しかしながら、SOINN-AMLでは距離学習とクラスタリングを同時に行っていくと学習結果が安定しない問題がある。本研究では、SOINN-AMLにミニバッチ学習を取り入れることで安定した学習をできることを目指す。

2. 従来手法

SOINN-AMLは岡田らによって開発された。SOINN-AMLではDMLアルゴリズムの1つであるAdaptive Metric Learning(AML)[4]を用いて、データの入力空間を学習する。SOINNではユークリッド距離を用いて距離計算を行うが、入力空間が高次元の場合は、ユークリッド距離は必ずしも正確な密度の推定値を与えるとは限らないため、AMLにより距離指標を逐次学習する。距離指標を学習することでクラスタリングの性能を向上させる。AMLはクラスタ分類に基づいて距離指標を学習する。そのため、SOINNとAMLを組み合わせることでうまく機能する。

SOINN-AMLのアルゴリズムはSOINNアルゴリズムをベースとしている。次にSOINN-AMLアルゴリズムの概要を説明する。変換行列 G と共分散行列 Σ が恒等行列 I_m として初期化される。したがってAMLが実施される前の距離

計算はユークリッド距離と等しい。AMLの結果、線形変換を行う G と共分散行列 Σ からなる最適化された(1)の式のようにマハラノビス距離が出力される。マハラノビス距離はノード間の距離と閾値 t_i を計算するために使用される。 $\hat{x}_i$ は線形変換用の行列であり、 $m \times 1$ 次元の G と変換前の m 次元の x_i を計算して1次元に変換したものである。

$$D(\hat{x}_i, \hat{x}_j) = \sqrt{(\hat{x}_i - \hat{x}_j)^T \hat{\Sigma}^{-1} (\hat{x}_i - \hat{x}_j)} \quad (1)$$

マハラノビス距離の計算で使用される共分散行列 $\hat{\Sigma}$ は以下の(2)式で示す。 G^T は G を転置したものであり、 μ は x を平均したものである。そして、 I_m は m 次元の単位行列である。 σ は正則化パラメータである。正則化パラメータは事前に設定する。

$$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n G^T (x_i - \mu)(x_i - \mu)^T G + \sigma G^T I_m G \quad (2)$$

($\sigma > 0$)

次にAMLを最適化する評価値は(3)の式の二乗誤差(SSE)で定義される。 C_j はクラスタであり、 $\hat{\mu}_j$ はクラスタの平均である。

$$SSE(\{C_j\}_{j=1}^K) = \sum_{j=1}^K \sum_{\hat{x}_i \in C_j} D(\hat{x}_i, \hat{\mu}_j)^2 \quad (3)$$

次に、第1勝者、第2勝者を探索する。第1勝者、第2勝者の探索のやり方と閾値の計算はSOINNと同じ方法であるため、式は省略する。その後、計算した閾値を超える場合、ノードの分布が不十分であり、入力パターンを新たなノードとして挿入する。両者の閾値を超えない場合は、勝者ノード間の重みを更新する。エッジの年齢を0にリセットし、第1勝者に接続されているエッジの年齢を1加算する。このとき、第1勝者、第2勝者間にエッジが存在しない場合、エッジを挿入してエッジ年齢を更新する。(4)、(5)の式に示す更新を行い、勝者ノードとその近傍ノードの重みベクトルを更新する。

$$\Delta W_{s1} = \frac{1}{T_{s1}} \Sigma^{-1} (\xi - W_{s1}) \quad (4)$$

$$\Delta W_{s2} = \frac{1}{100T_i} \Sigma^{-1} (\xi - W_i) \quad (5)$$

ただし T_{s1} は第1勝者ノードに選ばれた回数、

T_i は勝者ノードに選ばれた回数である。また、更新されたことによりエッジの年齢が事前に定義した age_{max} を超えたエッジを削除し、削除されたエッジのうち、エッジを1つも持たなくなったノードを削除する。入力パターンが事前に設定している λ の整数倍である場合は、隣接するノードが存在しない場合はそのノードは削除される。その後、(6)の式を満たすとき AMLを実施する。AMLは入力パラメータとしてクラスタ数を必要とする。一方で、SOINNはオンライン学習であるため、データを1つずつ入力するため、クラスタ数の増減を繰り返す。そのため、クラスタ数の変動が小さいときに AMLを行うのが理想的であり、AMLが実行させる間隔を制御するパラメータ T_e を設定する必要がある。

$$\begin{aligned} & \left| \left(\max_{0 \leq i \leq T_e} CN(t+i) - CN(t) \right) \right| \leq 1 \\ & \left| \left(\min_{0 \leq i \leq T_e} CN(t+i) - CN(t) \right) \right| \leq 1 \quad (6) \\ & CN(t+T_e) \neq T_{AML} \end{aligned}$$

AMLではk-means法を用いて学習データを分類する。次にQR分解を用いて線形変換で使用するGを更新する。具体的な方法としてはQR分解で直交行列Qと上三角行列Rに分解して直交行列QをGとする。

次に Kernel K-means 法を用いてLの更新をし、分類したデータのラベル付けを行う。ラベル情報は(7)の式により算出する。

$$LS(i) = \arg \max_{1 \leq k \leq K} L(i, k) \quad (7)$$

ラベリング後、エッジの削除と作成を行う。図1のようにラベル情報が異なるノード間のエッジの削除を行う。削除を行った後、最近傍ノードと第2最近傍ノードを(1)の式により計算する。最近傍ノードと第2最近傍ノードのペア間にエッジを作成する。サブクラスタ集合が存在しなくなるまでエッジの削除や作成を繰り返す。

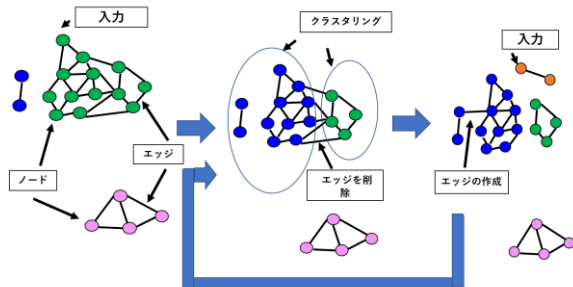


図1. エッジの削除と生成

3. 提案手法

従来研究のSOINN-AMLの問題点としては、学習結果が安定しないということが挙げられる。SOINN-AMLはオンライン学習で学習を行な

っている。オンライン学習は1データごとにパラメータを更新するため学習が安定しにくい。しかし、ミニバッチ学習は小さなグループごとに学習するのでオンライン学習よりも学習結果が安定する特徴を持っている。そこでSOINNアルゴリズムの部分ミニバッチ学習にすることで学習の安定化を図る。

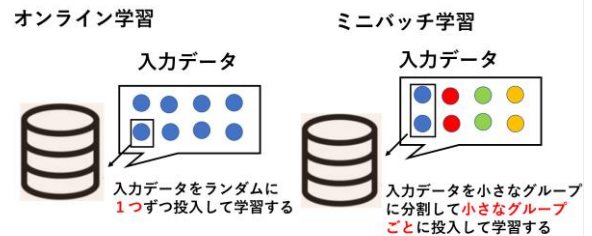


図2. ミニバッチ学習について

4. 実験および検討

オンライン学習は1データごとにパラメータを更新するため学習が安定しにくい。そこで、ミニバッチ学習は小さなグループごとに学習するのでオンライン学習よりも学習結果が安定する。そのため、ミニバッチ学習を導入することで学習が安定すると考えられる。

5. まとめ

本研究では、SOINN-AMLで学習する過程で、本来オンライン学習を行う処理をミニバッチ学習に変更することで学習結果の安定化を試みた。

参考文献

- [1] 山崎和博, 巻渕有哉, 申富饒, 長谷川修, 「SOINNとその実践」, 日本神経回路学会誌 Vol.17, No.4(2010), 187-196
- [2] Shufeng Xiong, Yue Zhang, Donghong Ji, and Yinxia Lou, "Distance Metric Learning for Aspect Phrase Grouping" In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pages 2492–2502, Osaka, Japan. The COLING 2016 Organizing Committee.
- [3] S. Okada and T. Nishida, "Online incremental clustering with distance metric learning for high dimensional data," The 2011 International Joint Conference on Neural Networks, 2011, pp. 2047-2054, doi: 10.1109/IJCNN.2011.6033478.
- [4] J. Ye, Z. Zhao, and H. Liu, "Adaptive distance metric learning for clustering" in Proc. IEEE Conference on CVPR, 2007.