

データのクラス近傍を保存する SOM とそのパターン識別への応用

SOM that preserves class-proximity and its application to pattern classification

齊藤 綾
Aya Saito

ピトヨ ハルトノ
Pitoyo Hartono

公立はこだて未来大学大学院 システム情報科学研究科
Graduate School of Systems Information Science, Future University-Hakodate

This paper presents a modified Self-Organizing Map algorithm which can preserve in addition to proximity-relation also the so called class-proximity in addition. In the learning process, CPSOM refers not only to the feature vectors of input data but also their classes. As a result, CPSOM is capable of building a clustered map of the data. Therefore, CPSOM can be used for data classification using nearest neighbor method. The proposed CPSOM is applied to several benchmark problems and compare its performances with MLP.

1. 背景と目的

本研究の目的は、高精度な識別能力と同時に解析結果を視覚的に提示できるパターン識別手法を構築することである。そのために、高次元データの類似性を可視化する手法である自己組織化マップ(Self-Organizing Map, SOM) [Kohonen 1997]の改良を行った。SOMとは T.Kohonen により提案された教師なしのニューラルネットワークの一種であり、その最も特徴的な能力は与えられる高次元データをその位相関係を保持したまま1～3次元の競合層に射影することである。高次元データの可視化能力に優れているが、SOMはパターン識別に不向きである。その理由の一つは、競合層にデータを射影する際にデータの所属クラスを考慮していないからである。そこで、本研究ではクラス近傍も保持できる Class-Proximity SOM(CPSOM)を提案し、それをパターン識別へ応用した。CPSOMの目的は、同一のクラスに属し、さらに類似するデータ同士を競合層上で近接に配置することである。これにより、マップ上ではクラス別のクラスタが形成され、その上で LVQ や近傍法などの識別アルゴリズムを実装することができ、可視化能力とパターン識別能力を備える SOMを構築することができる。

SOM をクラスタリングに用いる研究は近年盛んに行われてきた。例えば Self-Organizing Map considering False Neighboring Neuron, FNN-SOM [松下 2007]は競合層上のニューロン間の幾何学的な距離とそれらが代表する入力ベクトル間の距離の整合性をより良く反映できる学習アルゴリズムを提案した。しかし、この手法では入力データの位相関係をマップ化できるが、データのクラスを考慮しないため、識別には不向きである。それに対して、本研究で構築した CPSOM は入力サンプルのクラスも特徴ベクトルの要素として扱うことにより、マップ上でのクラスタを形成することが出来る。実験では本研究で提案した CPSOM を様々なベンチマーク問題に適用し、評価を行った。

2. CPSOM

2.1 構造と学習アルゴリズム

CPSOM は従来の SOM と異なり、サンプルのクラスも特徴として扱う。そのため入力層はサンプルの特徴ベクトルの受容部分とクラスベクトルの受容部分から構成される(図1)。ここでは、サンプル s の特徴ベクトル、 X^s とそのクラスベクトル、 C^s を式(1)と(2)に示す。

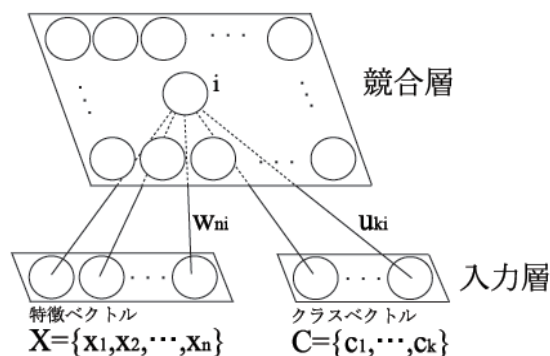


図1. CPSOM の構造

$$X^s = \{x_1^s, x_2^s, \dots, x_n^s\} \in R^n \quad (1)$$

$$C^s = \{c_1^s, \dots, c_k^s\} \in \{0,1\}^k \quad (2)$$

上の式では、 n はデータの次元を示し、 k はクラス数を示す。このとき、 C^s は k 次元のベクトルで表現している。このベクトルはサンプル s のクラスが j ならば、 j 番目の要素のみが 1 の値を取り、それ以外は 0 である。競合層の i 番目のニューロンは結合負荷ベクトル W^i と U^i で入力層の特徴ベクトル受容部分とクラスベクトル受容部分と結合し、式(3)、(4)に示される。

$$W^i = \{w_{1i}, w_{2i}, \dots, w_{ni}\} \quad (3)$$

$$U^i = \{u_{1i}, \dots, u_{ki}\} \quad (4)$$

CPSOM は、高次元空間において、互いの特徴ベクトルが類似し、かつ所属クラスが同一のデータをマップ上で近接に配置されるよう学習を実行する。

学習過程で、サンプル s が与えられたとき、競合層上のニューロン i との距離を式(4)に表す。

$$D^T(s, i) = (1.0 - \eta(T))dist_f(s, i) + \eta(T)dist_c(s, i) \quad (4)$$

$$dist_f(s, i) = \frac{1}{n} \sum_{j=1}^n |x_j^s - w_{ji}| \quad (5)$$

$$dist_c(s, i) = \frac{1}{k} \sum_{j=1}^k |c_j^s - u_{ij}| \quad (6)$$

$$\eta(T) = \frac{0.5}{T+1} + 0.5 \quad (7)$$

ここでは T は学習 epoch を表し、1 epoch では全てのサンプルを任意な順番で学習する。式(4)で定義した距離は従来の SOM と異なり、時間の関数となる。ここでは、学習の初期段階ではクラスベクトルに大きな重みを置き、学習が進行するに従い、クラスベクトルと特徴ベクトルによる距離の重みを同等な大きさに徐々に変化させる。これにより、学習の初期段階では特徴ベクトルの値にかかわらず、同一クラスのデータを競合層上の特定のニューロンに射影し、学習が進むに従い、そのニューロンの周りに同一クラスに属するデータに対応するニューロンが集中し、クラスタを形成する。

入力サンプルは式(8)に従い、最も距離 $D^T(s, i)$ が短い競合層上のニューロンを勝者ニューロン、 i^* とする。

$$i^* = \arg \min D^T(s, i) \quad (8)$$

CPSOM では、競合層上の全てのニューロンはクラスカウンターというパラメータを持つ。入力データに k 種類のクラスが存在するとき、 i 番目のニューロンのクラスカウンター P_i は、式(9)に示される。

$$P_i = \{p_{i1}, p_{i2}, \dots, p_{ic}, \dots, p_{ik}\} \quad (9)$$

ベクトルの要素 p_{ic} はニューロン i のクラス c に対するクラスカウンターとする。

クラスカウンターは 1 epoch の学習が終了する度に、式(10)に従い更新される。ここでは、 $dist_c(s(c), i)$ の値は $[0, 1]$ に正規化する。また、 $s(c)$ はクラス c に対応するクラスベクトルである。

$$p_{ic}(T+1) = p_{ic}(T) + \eta(T)(1 - dist_c(s(c), i)) \quad (10)$$

$$dist_c \leq 1$$

クラスカウンターを用いて、ニューロン i がクラス c に所属する確率、尤度 λ_{ic} を式(11)に定義する。

$$\lambda_{ic}(T) = \frac{p_{ic}(T)}{\sum_{j=1}^k p_{ij}(T)} \quad (11)$$

CPSOM では、クラス別のクラスタを形成するために、式(11)で定義した尤度が重要となる。この尤度は式(12)、(13)に示すように結合負荷ベクトルの更新時に起用する。クラス c に属する入力サンプルに対し、勝者ニューロンとその近傍ニューロンの更新強度は、式(12)、(13)に示すよう尤度の影響を受ける。

$$w_{ji}(t+1) = w_{ji}(t) + \alpha \Lambda(i, i^*) \lambda_{ic}(t) (1 - \eta(t)) (x_j^s - w_{ji}(t)) \quad (12)$$

$$u_{ii}(t+1) = u_{ii}(t) + \alpha \Lambda(i, i^*) \eta(t) (c_i^s - u_{ii}(t)) \quad (13)$$

上の式では α を学習係数、 Λ を近傍関数とする。

2.2 CPSOM が形成するマップの特性

CPSOM で生成されるマップにはクラス毎のクラスタが形成され、その中でサンプル同士の位相関係が保たれる。

CPSOM で形成されるマップの特徴を理解するために、図2に示す人工的なデータに対し、学習を実行した。この 2 次元データでは、各点の座標を特徴とし、クラスを色で表現した。このデータに対し、従来の SOM で生成したマップを図4(a)、提案する CPSOM によるマップを図4(b)に示す。図4(a)のマップには、データ本来の位相関係が反映されているのに対し、図4(b)は位相関係を保存すると同時にクラスタを形成していることが明らかである。図3は CPSOM の学習時のエラーを表し、ここから学習の収束状況がわかる。CPSOM が形成するマップをより明確に説明するため、9 次元データである、glass 問題[UCI]に対し、従来の SOM と CPSOM による学習を実行し、形成されたマップをそれぞれ図5(a)と図5(b)に示す。この2つのマップから明らかであるように、CPSOM はクラスに基く分離されたクラスタを形成し、クラスタ内ではデータの位相関係が保持される。

CPSOM で形成されたマップの特性を利用し、近傍法[Cover 1967]によるパターン識別をマップ上で実装することができる。

2.3 パターン識別への応用

本研究では、クラス付きの学習データに対し、CPSOM の学習を行い、マップを形成する。その後、そのマップを用いて、クラスが不明であるデータに対し、識別処理を実行する。クラスが不明なサンプル s が与えられたとき、まずそのサンプルがクラス、 c に属すると仮定し、クラスベクトルをサンプルに付ける。そのときのサンプル s とニューロン i の距離を式(14)に示す。

$$D(s, i) = \frac{1}{2} dist_f(s, i) + \frac{1}{2} dist_c(s, i) \quad (14)$$

上で説明した仮定クラスに対し、勝者ニューロンとサンプル s の距離を $D_c(s)$ とする。上記で説明した過程を全てのクラスに対して実行し、(15)式に示すように勝者ニューロンとの距離が最小となる仮定クラス j をサンプル s のクラスとする。

$$class(s) = \arg \min_j D_j(s, i) \quad (15)$$



図2. 螺旋の画像

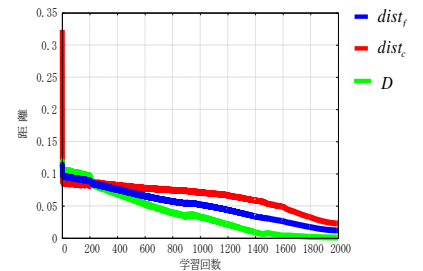
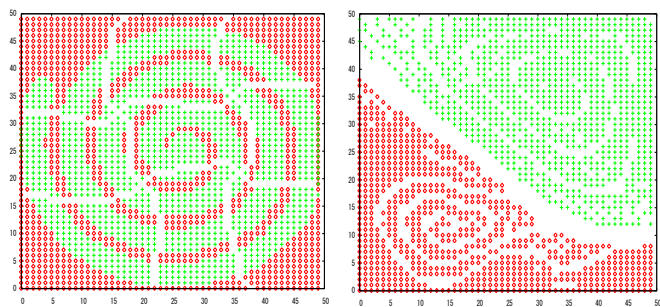
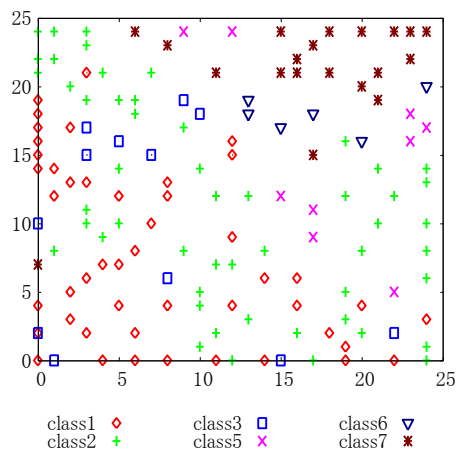


図3. CPSOM の学習曲線



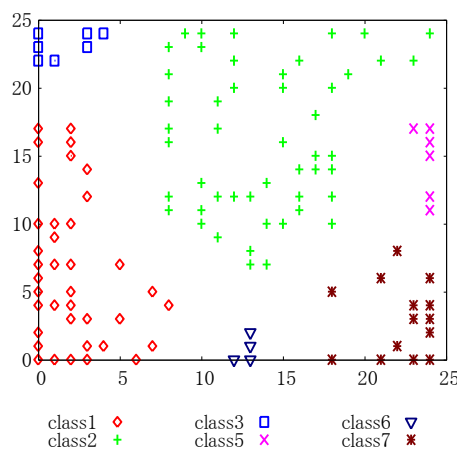
(a) SOM (b) CPSOM

図4. 2種類の SOM による螺旋のマップ



class1 \diamond class2 $+$ class3 \square class5 \times class6 ∇ class7 $*$

(a) SOM



class1 \diamond class2 $+$ class3 \square class5 \times class6 ∇ class7 $*$

(b) CPSOM

図5. 2種類の SOM による glass 問題のマップ

3. UCIのベンチマーク問題を用いた実験

実験では、上で説明した識別処理をマップ上で実装する。

識別能力の評価を行うため、MLP[Rumelhart 1986]との比較実験を行った。実験では性質の異なる5つのベンチマーク問題に対し、評価を行った。各問題に対し、30 実験を繰り返し、学習に用いないデータの平均識別率を評価の尺度とする。」

CPSOM の学習時のパラメータは全ての問題に共通し、マップサイズを 25×25 、学習係数 $\alpha = 0.9$ 、近傍係数 $\sigma = 15$ 、学習回数は 2000 回とした。

MLP に対しては、10 通りの学習パラメータで、学習と識別を 30 試行ずつ行い、最も平均識別率の高いものを選んだ。また、本実験では学習回数を経験的に 5,000 回とし、エラーが 0.001 を下回った場合はその時点で学習を終了した。

実験の結果は図5に示すとおりである。

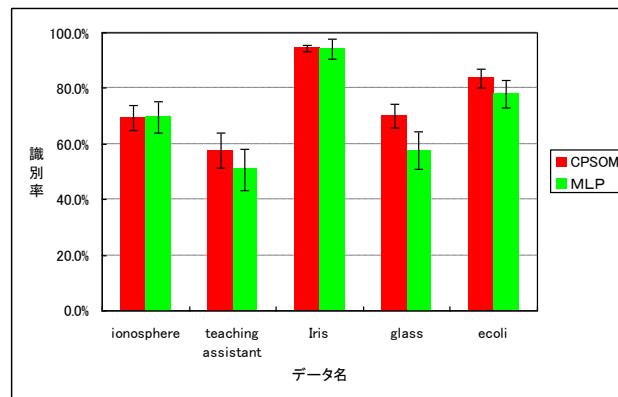


図5. 未学習データに対する識別能力の比較

4. 考察と今後の展望

図5より、CPSOM は MLP に匹敵する高精度の識別率を持つことが確認できた。さらに、本手法は定量的なデータだけでなく、定性的なデータにも適用できるため、感性情報処理への応用も検討する。

参考文献

[Kohonen 1997]Teuvo:Kohonen:Self-Organizing Maps(2nd ed), Springer Verlag,1997.
 [松下 2007] 松下春奈, 西尾芳文: 誤り近傍ニューロンを考慮する自己組織化マップとその応用, 電子情報通信学会技術研究報告. NLP, 非線形問題, Vol.106 , No.451,pp.1-4, 2007.
 [UCI]UCI Machine Learning Repository Content Summary, <http://mllearn.ics.uci.edu/MLSummary.html>
 [Cover 1967]T.Cover and P.Hart, Nearest: Neighbor Pattern Classification, *IEEE Trans on Information Theory*, Vol.IT-13,No.1,pp.21-27,1967
 [Rumelhart 1986] D. E. Rumelhart, G. E. Hinton and R. J. Williams:Learning Internal Representation by Error Propagation, in D. E. Rumelhart, J. L. McClelland and the PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, No.1, Foundations, MIT Press, 1986.