

再結合する自己組織化マップとその応用

REUNIFYING SELF-ORGANIZING MAP AND ITS APPLICATION

松下 春奈
Haruna Matsushita

西尾 芳文
Yoshifumi Nishio

徳島大学 工学部 電気電子工学科
Department of Electrical and Electronic Engineering, Tokushima University

1 はじめに

近年、自己組織化マップ (Self-Organizing Map: SOM) がクラスタリングの分野で非常に注目されている。しかし、複数のクラスタが離れた場所に位置するような入力データに対して SOM を用いた場合、クラスタ間に不活性ニューロンが存在してしまうことが知られている。そこで本研究では、再結合する自己組織化マップ (Reunifying SOM: RSOM) を提案する。RSOM の全ニューロンは、各自の 2 次元グリッド上の位置関係を記憶してはいるが、初期状態ではどのニューロンとも結合されていない。しかしながら、学習が進むに従って、徐々に他のニューロンと結合されていく。クラスタリングへ応用することで、提案手法の振る舞いを調査する。また、その結果に不活性ニューロンが存在しないことから、RSOM の有効性を確認する。

2 再結合する自己組織化マップ (RSOM)

各ニューロン i は、 d 次元重みベクトル $w_i = (w_{i1}, w_{i2}, \dots, w_{id})$ ($i = 1, 2, \dots, m$) を持っている。RSOM の全ニューロンは、各自の 2 次元グリッド上の位置関係を覚えてはいるが、初期状態ではどのニューロンとも結合されていない。入力データ $x_j = (x_{j1}, x_{j2}, \dots, x_{jd})$ ($j = 1, 2, \dots, N$) の範囲は、0 から 1 と仮定する。

(RSOM1) 入力データ x_j を全ニューロンに入力する。

(RSOM2) 最も入力に近い重みベクトルを持つニューロンを勝者ニューロン c とする。

(RSOM3) 各ニューロンの重みベクトルを更新する。

$$w_i(t+1) = w_i(t) + h_{Rc,i}(t)(x_j - w_i(t)) \quad (1)$$

ここで、 $h_{Rc,i}(t)$ は式 (2) で与えられる近傍関数である。

$$h_{Rc,i}(t) = \alpha(t) \exp\left(-\frac{(n_{c,i}/m + \|w_i - x_j\|)^2}{2\sigma^2(t)}\right) \quad (2)$$

$\|\cdot\|$ はユークリッド距離を表し、 $n_{c,i}$ は c と各ニューロン i との近傍距離を示す。近傍距離は、結合されたマップノード間の最短経路と定義する。もし、ニューロン i が直接的にも間接的にも c と結合されていないとき、 $n_{c,i}$ はニューロン数 m と等しいとする。 $\alpha(t)$ はラーニングレート、 $\sigma(t)$ は近傍関数の幅に対応する。 $\alpha(t)$ と $\sigma(t)$ は時間と共に減少する。

(RSOM4) 全てのニューロンが結合されていると仮定したときの、 c の 1 近傍ニューロンの集合を N_{c1} とする。 N_{c1} の中で最も近傍距離が長いニューロンの集合を Sq とする。

もし $n_{c,Sq} = 1$ ならば (RSOM7) を、違うならば (RSOM5) を行なう。

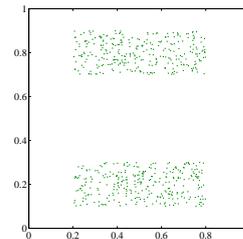
(RSOM5) Sq の中で最も x_j に近い重みベクトルを持つニューロンを、結合ニューロン q とする。

もし $\|w_q - w_c\| \leq D(t)$ ならば (RSOM6) を行なう。違うならば N_{c1} から q を除き、(RSOM4) からシミュレーションをやり直す。 $D(t)$ は時間と共に増加する。(RSOM6) c と q を直接結合する。つまり、 $n_{c,q} = 1$ となる。

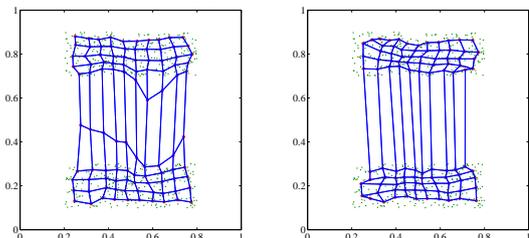
(RSOM7) (RSOM1) に戻り、全ての入力に対して繰り返す。

3 クラスタリングへの応用

図 1(a) に示される、2 次元入力データに対してシミュレーションを行う。従来の SOM の結果を図 1(b) に示す。この結果を見ると、2 つのクラスタ間に多くの不活性ニューロンが存在している。それに比べて、図 1(c) の提案手法の結果からは、不活性ニューロンが全く無いことが確認できる。これらの結果より、RSOM を用いたクラスタリングは有効であるといえる。



(a) 入力データ。



(b) 従来の SOM の結果。(c) RSOM の結果。

図 1 2 次元データのクラスタリング。

4 まとめ

本研究では、再結合する自己組織化マップを提案した。また、クラスタリングへ応用することで提案手法の振る舞いを調査し、有効性を確認した。