数理解析研究所講究録
第 1821 巻 2013 年 219-225

On states with absorbing tendencies in self-organizing maps with a two-dimensionally indexed array
(2次元配列自己組織化マップにおける準吸収特性について)

秋田県立大学 システム科学技術学部　星野満博 (Mitsuhiro Hoshino)
Faculty of Systems Science and Technology, Akita Prefectural University

1. 基本的な自己組織化マップ

本報告はKohonen型アルゴリズム [8] として知られている自己組織化マップにおける整列化に関する数理的考察である．自己組織化マップは広範囲に応用例を有し，アルゴリズムも非常にシンプルであるが，数理的観点においても幾つかの興味深い性質をもつ．自己組織化マップの学習過程におけるノードの配列とノードの値との間に現れるある種の規則性と整列化の形成過程について，ノード値の状態クラスが閉クラスを形成する吸収特性に注目して考察する．特に，入力値が内積空間に値をもち，配列次元が1次元および2次元の場合における準吸収特性について言及する．

本報告では，自己組織化マップをノード，ノードの値，入力，学習プロセスの4つの要素によって，以下の様に定義する．

$$(I, V, X, \{m_k(\cdot)\}_{k=0}^{\infty})$$

(i) I をすべてのノードの集合とする．I は，距離 d をもつ距離空間の加算部分集合とする．

(ii) 各ノードは，それぞれ1つの値をもつ．V をノードの値の集合とする．V はノルム空間であると仮定する．V におけるノルムを $\|\cdot\|$ とする．$m(i)$ をノード i の値として，その対応 $m: I \to V$ をモデル関数 (model function) と呼ぶことにする．また，M をモデル関数の全体，$m_0: I \to V$ を初期モデル関数とする．

(iii) $X \subset V$ を入力集合とする．$x_0, x_1, x_2, \ldots \in X$ を入力列とする．

(iv) 本報告では，学習プロセスとして以下の2つのタイプの内の一方を仮定する．

学習プロセス L_A

(a) 学習範囲:

$$I(m_k, x_k) = \left\{ i^* \in I \;\middle|\; \|m_k(i^*) - x_k\| = \inf_{i \in I} \|m_k(i) - x_k\| \right\} \quad (m_k \in M, x_k \in X),$$

$$N_1(i) = \left\{ j \in I \;\middle|\; d(j, i) \leq \varepsilon \right\} \quad (i \in I).$$

(b) 学習率: $0 \leq \alpha \leq 1$.

(c) 更新後の値:

$$m_{k+1}(i) = \begin{cases} (1-\alpha)m_k(i) + \alpha x_k & \text{if } i \in \bigcup_{i^* \in I(m_k, x_k)} N_\varepsilon(i^*), \\ m_k(i) & \text{if } i \notin \bigcup_{i^* \in I(m_k, x_k)} N_\varepsilon(i^*), \end{cases} \quad k = 0, 1, 2, \ldots.$$

学習プロセス L_m

(a) 学習範囲:

$$J(m_k, x_k) = \min\left\{ i^* \in I \;\middle|\; \|m_k(i^*) - x_k\| = \inf_{i \in I} \|m_k(i) - x_k\| \right\}$$
$$(m_k \in M, x_k \in X),$$
$$N_\varepsilon(i) = \{ j \in I \mid d(j, i) \leq \varepsilon \} \quad (i \in I).$$

(b) 学習率: $0 \leq \alpha \leq 1$.

(c) 更新後の値:

$$m_{k+1}(i) = \begin{cases} (1-\alpha)m_k(i) + \alpha x_k & \text{if } i \in N_\varepsilon(J(m_k, x_k)), \\ m_k(i) & \text{if } i \notin N_\varepsilon(J(m_k, x_k)), \end{cases} \quad k = 0, 1, 2, \ldots.$$

2. $\mathbb{R}$ 値ノード，1 次元ノード配列モデル

ここでは，最も単純な自己組織化マップである，$\mathbb{R}$ 値ノード，1 次元ノード配列の場合について述べる．

(i) 有限個のノードを仮定する．$I = \{1, 2, \ldots, n\} \subset \mathbb{N}$.

(ii) ノード値の空間を $\mathbb{R}$（ユークリッドノルム）とする．

(iii) $x_0, x_1, x_2, \ldots \in X \subset \mathbb{R}$ を入力列とする．

(iv) 以下の学習プロセスの一方を仮定する．

学習プロセス L_A (1 次元配列，$\mathbb{R}$-値ノード，$\varepsilon = 1$)

(a) 学習範囲:

$$I(m_k, x_k) = \left\{ i^* \in I \;\middle|\; |m_k(i^*) - x_k| = \inf_{i \in I} |m_k(i) - x_k| \right\}$$
$$(m_k \in M, x_k \in X),$$
$$N_1(i) = \{ j \in I \mid |j - i| \leq 1 \} \quad (i \in I).$$

(b) 学習率: $0 \leq \alpha \leq 1$.

(c) 更新後の値:

$$m_{k+1}(i) = \begin{cases} (1-\alpha)m_k(i) + \alpha x_k & \text{if } i \in \bigcup_{i^* \in I(m_k, x_k)} N_1(i^*), \\ m_k(i) & \text{if } i \notin \bigcup_{i^* \in I(m_k, x_k)} N_1(i^*), \end{cases} \quad k = 0, 1, 2, \ldots.$$

学習プロセス L_m (1 次元配列, $\mathbb{R}$-値ノード, $\varepsilon = 1$)

(a) 学習範囲:

$$J(m_k, x_k) = \min\left\{ i^* \in I \;\middle|\; |m_k(i^*) - x_k| = \inf_{i \in I} |m_k(i) - x_k| \right\}$$
$$(m_k \in M, x_k \in X),$$
$$N_1(i) = \left\{ j \in I \;\middle|\; |j - i| \leq 1 \right\} \quad (i \in I).$$

(b) 学習率: $0 \leq \alpha \leq 1$.

(c) 更新後の値:

$$m_{k+1}(i) = \begin{cases} (1-\alpha)m_k(i) + \alpha x_k & \text{if } i \in N_1(J(m_k, x_k)), \\ m_k(i) & \text{if } i \notin N_1(J(m_k, x_k)), \end{cases} \quad k = 0, 1, 2, \ldots.$$

学習のプロセスは次のように進められる. n 個のノード $1, 2, \ldots, n$ があり, そのそれぞれに対してノードの値 $m_0(1)$, $m_0(2)$, ..., $m_0(n)$ が与えられているものとする. $x_0 \in X$ が入力されたならば, 上記の学習プロセスにより各のノードの値が更新され $m_1(1)$, $m_1(2)$, ..., $m_1(n)$ が得られる. 入力 $x_1, x_2, x_3, \ldots$ に対して, これを繰り返すことにより, 逐次にノードの更新がおこなわれ, 同時にモデル関数 $m_1, m_2, m_3, \ldots$ が逐次に生成される.

このような学習を十分な回数, 繰り返したとき, モデル関数において, 単調性等, 各ノードの値の配列にある種の規則性が現れることがある. 実際, 様々なノード集合, ノードの値の空間, 学習方法において, 単調化等の幾つかの興味深い現象が現れる. また, これらの性質を利用することにより, 多くの実用的な問題へ応用されている.

3. 吸収状態について

次の定理は, 自己組織化マップにおけるモデル関数の単調性保存に関する基本的な結果である.

Theorem 1 学習プロセス L_A を仮定するとき, モデル関数 $m_1, m_2, m_3, \ldots$ に対して, 以下が成り立つ.

(i) モデル関数 m_k が I 上で単調増加であるならば, モデル関数 m_{k+1} も I 上で単調増加である.

(ii) モデル関数 m_k が I 上で単調減少であるならば，モデル関数 m_{k+1} も I 上で単調減少である．

(iii) モデル関数 m_k が I 上で狭義単調増加であるならば，モデル関数 m_{k+1} も I 上で狭義単調増加である．

(iv) モデル関数 m_k が I 上で狭義単調減少であるならば，モデル関数 m_{k+1} も I 上で狭義単調減少である．

Theorem 2 学習プロセス L_m を仮定するとき，モデル関数 $m_1, m_2, m_3, \ldots$ に関して，以下が成り立つ．

(i) モデル関数 m_k が I 上で狭義単調増加であるならば，モデル関数 m_{k+1} も I 上で狭義単調増加である．

(ii) モデル関数 m_k が I 上で狭義単調減少であるならば，モデル関数 m_{k+1} も I 上で狭義単調減少である．

ここでの単調増加性，単調減少性のように，モデル関数の状態が一度ある状態のクラスに到達すると，その性質が保存されるという意味において，このような状態クラスは閉クラスを形成する．自己組織化マップの吸収特性と呼ぶことにする．モデルの吸収特性として準凸性，準凹性などがある [7].

4. 準吸収的特性について

1 次元ノード配列の場合

ここでは，以下の自己組織化マップモデルを考える．

$$(\{1,2,\ldots,n\}, V, X, \{m_k(\cdot)\}_{k=0}^{\infty})$$

(i) ノード集合 $I = \{1,2,\ldots,n\}$, $d_I(i,j) = |i-j|$.

(ii) ノード値空間 V は内積 $\langle\cdot,\cdot\rangle$ をもつ内積空間とする．

(iii) $x_0, x_1, x_2, \ldots \in X \subset V$ は入力列とする．

(iv) 学習プロセス L_m

(a) 学習範囲:

$$J(m,x) = \min\left\{i^* \in I \;\middle|\; \|m(i^*) - x\| = \inf_{i\in I} \|m(i) - x\|\right\} \qquad (m \in M, x \in X),$$

$$N_1(i) = \left\{j \in I \mid d_I(j,i) \le 1\right\} \quad (i \in I).$$

(b) 学習率: $0 \le \alpha \le 1$.

(c) 更新後の値:

$$m'(i) = \begin{cases} (1-\alpha)m(i) + \alpha x & \text{if } i \in N_1(J(m,x)), \\ m(i) & \text{if } i \notin N_1(J(m,x)), \end{cases} \quad k = 0, 1, 2, \ldots.$$

Condition $\mathbf{S}_{\mathbf{inn}}(i)$
ノード $i, i+1, i+2$ に対して

$$\langle m(i) - m(i+1), m(i+2) - m(i+1)\rangle \leq 0$$

が成り立つ.

Theorem 3 モデル

$$\bigl(\{1, 2, \ldots, n\},\ V,\ X,\ \{m_k(\cdot)\}_{k=0}^{\infty}\bigr)$$

において，学習過程 $L_m(\varepsilon = 1)$ を仮定する．m を任意のモデル関数として，x を任意の入力とする．m' を m の入力 x により更新されたモデル関数とする．2 つのノードを除くすべての i に対して

$$\langle m(i) - m(i+1), m(i+2) - m(i+1)\rangle \leq 0$$

が成り立つならば

$$\langle m'(i) - m'(i+1), m'(i+2) - m'(i+1)\rangle \leq 0$$

が成り立つ.

上の性質は，単調性のような完全な吸収特性をもたないが吸収的傾向を有する準吸収特性を有する状態クラスとなっている．この性質は，[7] における 2 次元配列モデルにおける結果と同様の議論により証明される．

2 次元ノード配列の場合

ここでは，以下の自己組織化マップモデルを考える．

$$\bigl(\{1, 2, \ldots, n_1\} \times \{1, 2, \ldots, n_2\},\ V,\ X,\ \{m_k(\cdot,\cdot)\}_{k=0}^{\infty}\bigr)$$

(i) ノード集合 $I = \{1, 2, \ldots, n_1\} \times \{1, 2, \ldots, n_2\}$．次の距離を用いる．

$$d_I((i,j),(k,l)) = \sqrt{(i-k)^2 + (j-k)^2}, \quad (i,j),(k,l) \in I.$$

(ii) ノード値空間 V は内積 $\langle\cdot,\cdot\rangle$ をもつ内積空間とする．

(iii) $x_0, x_1, x_2, \ldots \in X \subset V$ は入力列とする．

(iv) 学習プロセス L_m (2 次元配列，$\varepsilon = \sqrt{2}$)

(a) 学習範囲:

$$I(m,x) = \left\{ (i^*,j^*) \in I \mid \|m(i^*,j^*) - x\| = \inf_{(i,j)\in I} \|m(i,j) - x\| \right\}, \quad m \in M, x \in X,$$

$J : M \times X \to I$ は $J(m,x) \in I(m,x)$ を満たすものとする．

$$N_{\sqrt{2}}(i,j) = \left\{ (k,l) \in I \mid d_I((i,j),(k,l)) \leq \sqrt{2} \right\}$$

とする．

(b) 学習率: $0 \leq \alpha \leq 1$.

(c) 更新後の値:

$$m'(i,j) = \begin{cases} (1-\alpha)m(i,j) + \alpha x & \text{if } (i,j) \in N_{\sqrt{2}}(J(m,x)), \\ m(i,j) & \text{if } (i,j) \notin N_{\sqrt{2}}(J(m,x)), \end{cases} \quad k = 0,1,2,\ldots.$$

1 次元の場合の定理 3 を同様に，以下が得られる．

Theorem 4 モデル

$$\left(\{1,2,\ldots,n_1\} \times \{1,2,\ldots,n_2\},\ V,\ X,\ \{m_k(\cdot,\cdot)\}_{k=0}^{\infty} \right)$$

において，学習過程 $L_m(\varepsilon = \sqrt{2})$ を仮定する．m を任意のモデル関数として，x を任意の入力とする．m' を m の入力 x により更新されたモデル関数とする．$d_I((i,j), J(m,x)) \neq \sqrt{2}, 2, \sqrt{5}$ である $(i,j) \in I$ に対して

$$\langle m(i-1,j) - m(i,j), m(i+1,j) - m(i,j) \rangle \leq 0,$$
$$\langle m(i,j-1) - m(i,j), m(i,j+1) - m(i,j) \rangle \leq 0$$

が成り立つならば

$$\langle m'(i-1,j) - m'(i,j), m'(i+1,j) - m'(i,j) \rangle \leq 0$$
$$\langle m'(i,j-1) - m'(i,j), m'(i,j+1) - m'(i,j) \rangle \leq 0$$

が成り立つ．

参考文献

[1] M. Cottrell and J.-C. Fort, *Étude d'un processus d'auto-organisation*, Annales de l'Institut Henri Poincaré, 23(1) (1987), pp.1–20 (in French)

[2] E. Erwin, K. Obermayer, and K. Schulten, *Convergence properties of self-organizing maps*, In T. Kohonen, K. Mäkisara, O. Simula, and J. Kangas, editors, Artificial Neural Networks, Amsterdam Netherlands Elsevier (1991), pp.409–414

[3] E. Erwin, K. Obermayer and K. Schulten, *Self-organization maps: stationary states, metastability and convergence rate*, Bio. Cybern., 67 (1992), pp. 35–45.

[4] E. Erwin, K. Obermayer and K. Schulten, *Self-organization maps: ordering, convergence properties and energy functions*, Bio. Cybern., 67 (1992), pp. 47–55.

[5] W. Fujiwara, E. Itou, M. Hoshino, I. Kaku, A. Sakusabe, M. Sasaki and H. Kosaka, *A study on the effective method of external inspecting using a neural network approach*, Proceedings of 6th ICIM (2002), pp. 369–375

[6] M. Hoshino, and Y. Kimura, *Absorbing states and quasi-convexity in self-organizing maps*, *J. Nonlinear Convex Anal.*, Vol.10, No.3 (2009), pp. 395–406

[7] M. Hoshino and Y. Kimura, *Ordered states and probabilistic behavior of self-organizing maps*, Nonlinear Analysis and Optimization (Shimane, 2008), Yokohama Publishers, pp. 31–44

[8] T. Kohonen, *Self-Organizing Maps, Third Edition*, Springer, 2001.

[9] W. Takahashi, *Nonlinear Functional Analysis*, Yokohama publishers, 2000.

[10] P. L. Zador, *Asymptotic quantization error of continuous signals and the quantization dimension*, IEEE Trans. Inform. Theory, vol. IT-28, No. 2, March (1982), pp. 139–149.