

On a measure of ordering and states with absorbing tendencies in self-organizing maps

(自己組織化マップにおける整列化の評価と準吸収状態について)

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とモデル関数の性質に関する理論的考察である。自己組織化マップは非常に実用的であり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、理論面においても、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について、モデルの吸収状態と整列化の形成過程に注目して考察する。特に、内積空間に入力値をもつ 1 次元配列モデルにおける準吸収状態の存在とその特性について、数値例とともに述べる。

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 $m: I \rightarrow V$ 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、 $m_0: I \rightarrow V$ を初期モデル関数とする。
- (iii) $X \subset V$ を入力集合とする。 $x_0, x_1, x_2, \dots \in X$ を入力列とする。
- (iv) 学習プロセスとして以下の 2 つの内の一方を仮定する。

学習プロセス L_A

(a) 学習範囲:

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

(b) 学習率: $0 \leq \alpha \leq 1$.

(c) 更新後の値:

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

学習プロセス L_m

(a) 学習範囲:

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

(b) 学習率: $0 \leq \alpha \leq 1$.

(c) 更新後の値:

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

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.

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

- (i) 有限個のノードを仮定する. $I = \{1, 2, \dots, n\} \subset \mathbb{N}$.
- (ii) ノード値の集合を $\mathbb{R}$ (ユークリッドノルム) の部分集合とする.
- (iii) $x_0, x_1, x_2, \dots \in X \subset \mathbb{R}$ 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:

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

(b) 学習率: $0 \leq \alpha \leq 1$.

(c) 更新後の値:

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

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:

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

(b) 学習率: $0 \leq \alpha \leq 1$.

(c) 更新後の値:

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

今, n 個のノード $1, 2, \dots, n$ があり, そのそれぞれに対してノードの値 $m_0(1), m_0(2), \dots, m_0(n)$ が与えられている. $x_0 \in X$ 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 $m_1(1), m_1(2), \dots, m_1(n)$ が得られる. 入力 $x_1, x_2, x_3, \dots$ 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 $m_1, m_2, m_3, \dots$ 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 ($\{1, 2, \dots, n\}, V = \mathbb{R}, X \subset \mathbb{R}, \{m_k(\cdot)\}_{k=0}^\infty$) に対して, 学習プロセス L_A を仮定する. モデル関数 $m_1, m_2, m_3, \dots$ 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 ($\{1, 2, \dots, n\}, V = \mathbb{R}, X \subset \mathbb{R}, \{m_k(\cdot)\}_{k=0}^\infty$) に対して, 学習プロセス L_m を仮定する. モデル関数 $m_1, m_2, m_3, \dots$ に関して, 次が成り立つ.

- (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, モデル関数が一度その状態になると, その状態が保存されるという意味において, このような状態を自己組織化マップモデルの吸収状態と呼ぶことにする. モデルの吸収状態として準凸性, 準凹性などがある [7].

4. 1次元ノード配列, $\mathbb{R}^2$ -値ノードの場合の数値例

ここでは, 以下のモデルを仮定する.

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

- (i) ノード集合 $I = \{1, 2, \dots, n\}$. ただし, 距離を $d_I(i, j) = |i - j|$ によって定義する.
- (ii) モデル関数 $m : I \rightarrow \mathbb{R}^2$. ここで, 各 $x = (x_1, x_2), y = (y_1, y_2)$ に対して, $\|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$ とする.
- (iii) 入力 $x_0, x_1, x_2, \dots \in X \subset \mathbb{R}^2$.
- (iv) 学習プロセス L_m (1次元配列, $\varepsilon = 1$)

(a) 学習範囲:

$$J(m, x) = \min \left\{ i^* \in I \mid \|m(i^*) - x\| = \inf_{i \in I} \|m(i) - x\| \right\}$$

$$(m \in M, x \in X),$$

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

(b) 学習率: $0 \leq \alpha \leq 1$.

(c) 更新後の値:

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

1次元ノード配列, $\mathbb{R}^2$ -値ノードをもつ自己組織化マップの数値例を与える.

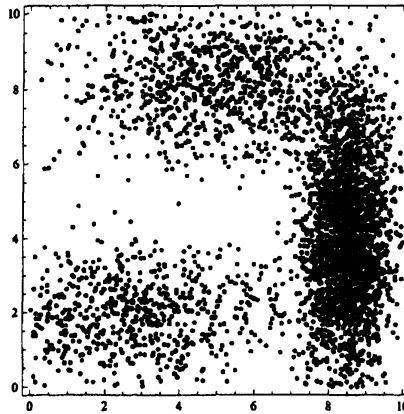


図 1: 入力分布: 横軸が第 1 成分の値, 縦軸が第 2 成分の値を意味する.

Example 1 次のような 35 個のノードをもつ自己組織化マップを考える.

- (1) ノード集合 $I = \{1, 2, 3, \dots, 35\}$. 距離を $d_I(i, j) = |i - j|$ によって定義する.
- (2) $V = [0, 10] \times [0, 10]$ とする. 初期モデル関数はランダムに生成した値

$$m_0(1) = (3, 1), m_0(2) = (9, 8), \dots, m_0(35) = (0, 7)$$

とする. $x = (x_1, x_2), y = (y_1, y_2)$ に対して, $\|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$ とする.

- (3) 図 1 における点を一様にランダムに発生させた値を入力値とする.
- (4) 学習過程 L_m を仮定する. ここで, 学習率は $\alpha = \frac{1}{3}$ とする.

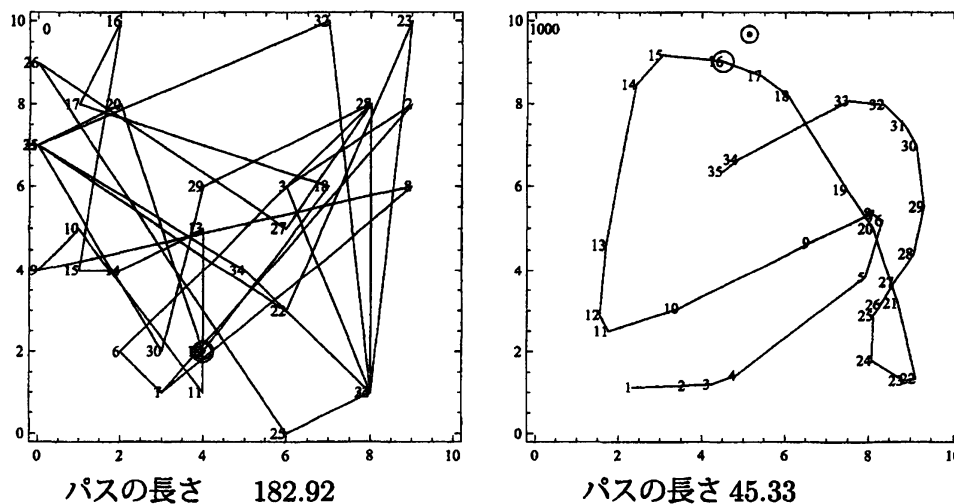


図 2: m_0 (左), m_{1000} (右)

図 2, 3 は学習回数 0 回, 1000 回, 7000 回, 16000 回のときのノードとノード値を 2 次元座標として図示したものである. また, パスの長さは $\sum_{i=1}^{34} \|m(i) - m(i+1)\|$ によって

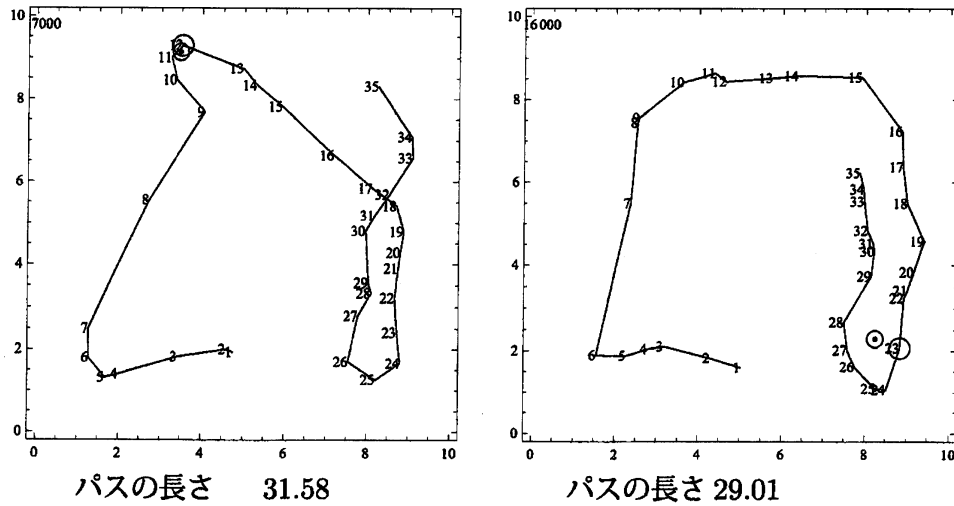


図 3: m_{7000} (左), m_{16000} (右)

定義している. 学習回数が増加するとパスの長さが短くなり, 更に図1において点の密度が高いエリアでノードが多くなっていることがわかる. $\square$

5. 準吸収状態について

ここでは, 内積空間上の値をもつ1次元ノード配列モデルの場合について考える.

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

- (i) ノード集合 $I = \{1, 2, \dots, n\}$, $d_I(i, j) = |i - j|$.
- (ii) ノード値空間 V は内積 $\langle \cdot, \cdot \rangle$ をもつ内積空間のある部分集合とする.
- (iii) $x_0, x_1, x_2, \dots \in X \subset V$ は入力列とする.
- (iv) 学習プロセス L_m

(a) 学習範囲:

$$J(m, x) = \min \left\{ i^* \in I \mid \|m(i^*) - x\| = \inf_{i \in I} \|m(i) - x\| \right\}$$

$$(m \in M, x \in X),$$

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

(b) 学習率: $0 \leq \alpha \leq 1$.

(c) 更新後の値:

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

Condition $S_{\text{inn}}(i)$

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

$$(\{1, 2, \dots, n\}, V, X, \{m_k(\cdot)\}_{k=0}^\infty) \quad (V \text{ は内積空間のある部分集合})$$

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, 単調性のような吸収状態にはなっていないが吸収的傾向を有する準吸収状態となっている. 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.

Example 2 図1における数値例を用いる. $S_{\text{inn}}[m, m', S \rightarrow S]$ をモデル関数 m とその更新されたモデル関数 m' の両方に対して, 条件 $S_{\text{inn}}(i)$ が成り立つようなノード i の個数とし, $S_{\text{inn}}[m, m', S \rightarrow N]$ をモデル関数 m に対して $S_{\text{inn}}(i)$ が成り立つが, その更新されたモデル関数 m' に対しては $S_{\text{inn}}(i)$ が成り立たないようなノード i の個数とする. また $S_{\text{inn}}[m, m', N \rightarrow S]$ をモデル関数 m に対して $S_{\text{inn}}(i)$ が成り立たないが, その更新されたモデル関数 m' に対しては $S_{\text{inn}}(i)$ が成り立つようなノード i の個数とし, $S_{\text{inn}}[m, m', N \rightarrow N]$ をモデル関数 m とその更新されたモデル関数 m' の両方に対して, 条件 $S_{\text{inn}}(i)$ が成り立たないようなノード i の個数とする. このとき, 学習回数とこれらの値について, 図4のような推移が観測された. また, 条件 $S_{\text{inn}}(i)$ を満たす状態への吸収的傾向が確認できる. $\square$

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

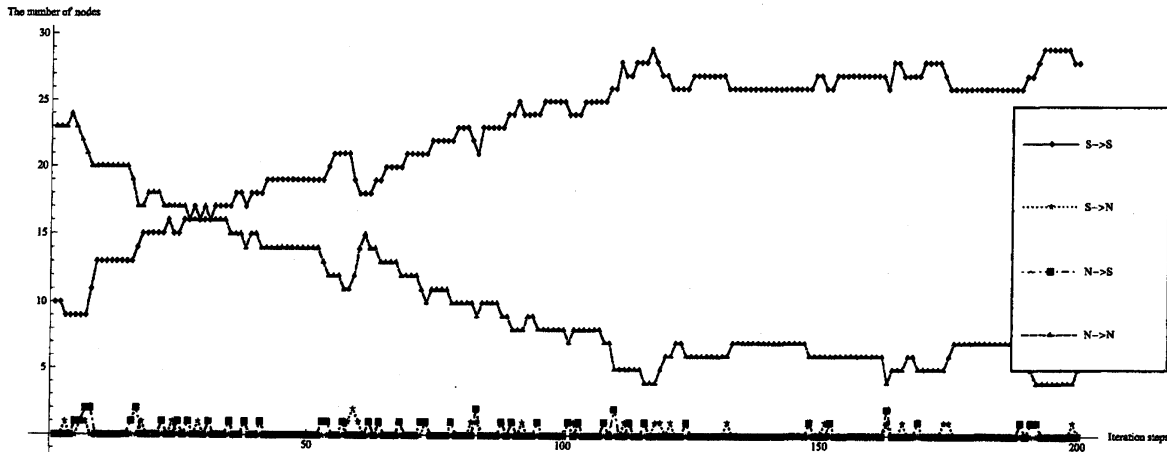


図 4: $S_{\text{inn}}[m, m', S \rightarrow S]$, $S_{\text{inn}}[m, m', S \rightarrow N]$, $S_{\text{inn}}[m, m', N \rightarrow S]$, $S_{\text{inn}}[m, m', N \rightarrow N]$ の推移

- [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.