

On absorbing property and learning rates factor in self-organizing maps on one-dimensional input space

(1次元入力型自己組織化マップにおける状態吸収性について)

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. 自己組織化マップモデルとそのプロセス

本研究は Kohonen 型アルゴリズム [4] として知られている自己組織化マップとその更新プロセスに関する数理モデル及びその性質に関する一つの理論的考察である。自己組織化マップ・アルゴリズムは非常にシンプルであるが、非常に実用的であり広範囲に応用例をもつ。ここでは、自己組織化マップモデル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そこへ至る形成過程に注目する。本報告では、モデル関数から更新モデル関数への写像が縮小型になるような可変学習率を有する自己組織化マッププロセスとその数値計算例を与える。

本報告では、自己組織化マップモデルとそのプロセスをノード、ノードの値、インプット、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) 幾つかのノードとその全体から成る集合 I を考える。 I は距離 d をもつ距離空間の加算集合とする。
- (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し、 $(V, \|\cdot\|)$ として、ノルム空間を仮定する。 $m(i)$ をノード i の値として、その対応 $m: I \rightarrow V$ をモデル関数と呼ぶことにする。また、 M をモデル関数の全体とする。
- (iii) $X \subset V$ をインプット集合とする。 $x_0, x_1, x_2, \dots \in X$ をインプットの列とする。
- (iv) 学習プロセス
初期モデル関数 $m_0: I \rightarrow V$ とその更新ルール

$$m_{k+1}(i) = (1 - \alpha_{m_k, x_k})m_k(i) + \alpha_{m_k, x_k} x_k \quad (1)$$

が与えられているものとする。ここで、 α_{m_k, x_k} は、 $0 \leq \alpha_{m_k, x_k} \leq 1$ を満たす学習率を表す。

例えば, n 個のノード $1, 2, \dots, n$ から成る場合を考える. そのそれぞれに対してノードの初期値 $m_0(1), m_0(2), \dots, m_0(n)$ が与えられているものとする. このとき, インプットとこれに伴う学習により各ノードの値が更新される. インプット $x_0 \in X$ が入力されたならば,

$$m_1(i) = (1 - \alpha_{m_0, x_0})m_0(i) + \alpha_{m_0, x_0}x_0$$

により, ノードの値を更新する. インプット $x_1, x_2, x_3, \dots$ に対して, これを繰り返すことにより, モデル関数 $m_1, m_2, m_3, \dots$ が逐次に生成され, ノードの値が更新される.

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, 各ノードの値にある種の興味深い規則性が現れることがある.

2. 可変型学習を有する一般入力型自己組織化マップ

ここでは, 下記のような可変型学習を有する一般入力型自己組織化マップを導入し, モデル関数の推移を与える写像が下記の意味で縮小型になるようなタイプのプロセスを紹介する.

1次元配列ノード, 一般入力型自己組織化マッププロセス

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

を次のように定義する.

- (i) $I = \{1, 2, \dots, n\}$. $d_I(i, j) = |i - j|$, $i, j \in I$.
- (ii) $m: I \rightarrow V$. V は内積 $\langle \cdot, \cdot \rangle$ をもつノルム空間.
- (iii) インプット $x_0, x_1, x_2, \dots \in X \subset V$.
- (iv) 学習プロセス L_m ($\varepsilon = 1$) を以下のように定義する.

(a) 学習範囲:

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

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

(b) 学習率: $0 \leq \alpha \leq 1$;

(c) 学習:

$$m'(i) = \begin{cases} (1 - \alpha)m(i) + \alpha x & (i \in N_1(j(m, x)) \text{ のとき}) \\ m(i) & (i \notin N_1(j(m, x)) \text{ のとき}) \end{cases}$$

Theorem 1 一般入力型自己組織化マッププロセス

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

において, $L_m(\varepsilon = 1)$ と次のような可変学習率 α を仮定する. $m \in M$ を任意のモデル関数, $x \in X$ を任意のインプットとする. インプット x による m の更新されたモデル関数を m' とする. $0 < \alpha_0 \leq 1$ とする. また, $i^* = j(m, x)$ とし, α_1, α_2 を次のように定義する.

$$\alpha_1 = \begin{cases} \frac{2\langle m(i^* - 2) - m(i^* - 1), x - m(i^* - 1) \rangle}{\|x - m(i^* - 1)\|^2} & (i^* \geq 3 \text{ かつ } x \neq m(i^* - 1) \text{ のとき}) \\ 1 & \left(\begin{array}{l} i^* = 1, 2 \text{ または} \\ x = m(i^* - 1) \text{ のとき} \end{array} \right) \end{cases}$$

$$\alpha_2 = \begin{cases} \frac{2\langle m(i^* + 2) - m(i^* + 1), x - m(i^* + 1) \rangle}{\|x - m(i^* + 1)\|^2} & (i^* \leq n - 2 \text{ かつ } x \neq m(i^* + 1) \text{ のとき}) \\ 1 & \left(\begin{array}{l} i^* = n - 1, n \text{ または} \\ x = m(i^* + 1) \text{ のとき} \end{array} \right) \end{cases}$$

このとき, $0 \leq \alpha \leq \max\{0, \min\{\alpha_1, \alpha_2, \alpha_0\}\}$ を満たす任意の α に対して

$$\|m'(i+1) - m'(i)\| \leq \|m(i+1) - m(i)\|, \quad i = 1, 2, \dots, n-1 \quad (3)$$

が成り立つ.

Example 1 以下のような2次元実数値入力を有するノードの個数を30とする1次元配列モデルの数値例を考える. 次のように仮定する.

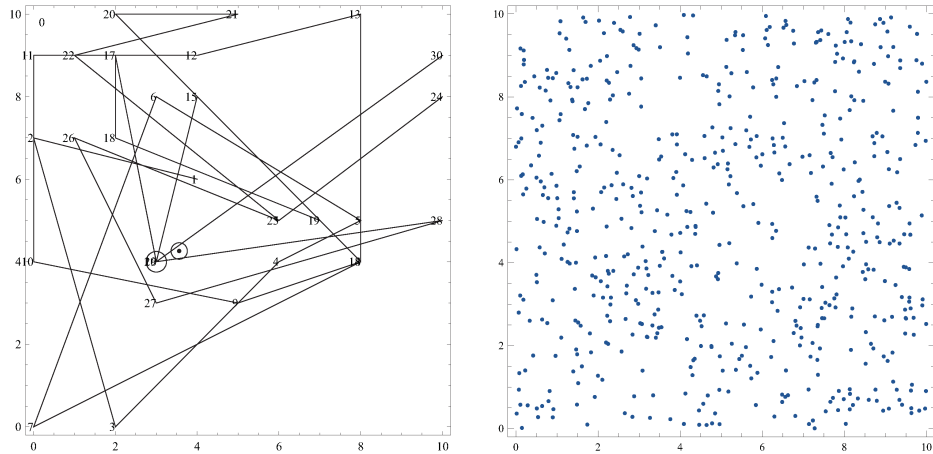


図 1: ノードの初期値 (m_0) と入力

1. $I = \{1, 2, 3, \dots, 30\}$ をノード集合, そこでの距離を $d_I(i, j) = |i - j|$ とする.
2. 初期モデル関数は図 1 により与える. 各ノードの位置は, その値を示す. 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.

3. 各インプットは $[0, 10] \times [0, 10]$ 上の一様分布に従う確率変数により生成されたものである (図1の右側に示されたサンプル点より生成. 個数は600個).
4. 学習プロセスは定理1における可変学習率 α をもつ $L_m(\varepsilon = 1)$ とする. ただし, $\alpha_0 = \frac{1}{3}$ を用いる.

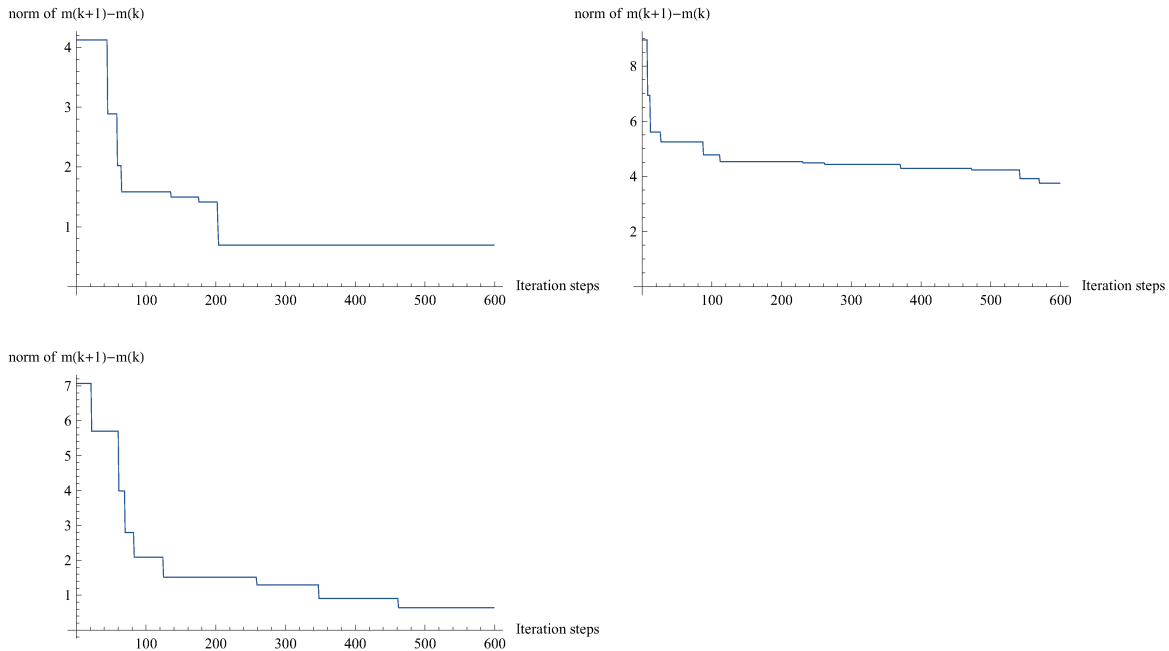


図 2: 2つのノード k と $k + 1$ の値の距離とその推移 (左上: $k = 1$, 右上: $k = 7$, 下: $k = 19$). 横軸は学習回数.

図2, 図3は, 数値計算によるモデル関数の600回の更新結果から作成されたものである. 図2より, 各ノード間の値の差が減少していることがわかる. また, 図3は, パスの長さ, すなわち隣接ノード間の値の距離の総和の推移を表している. $\square$

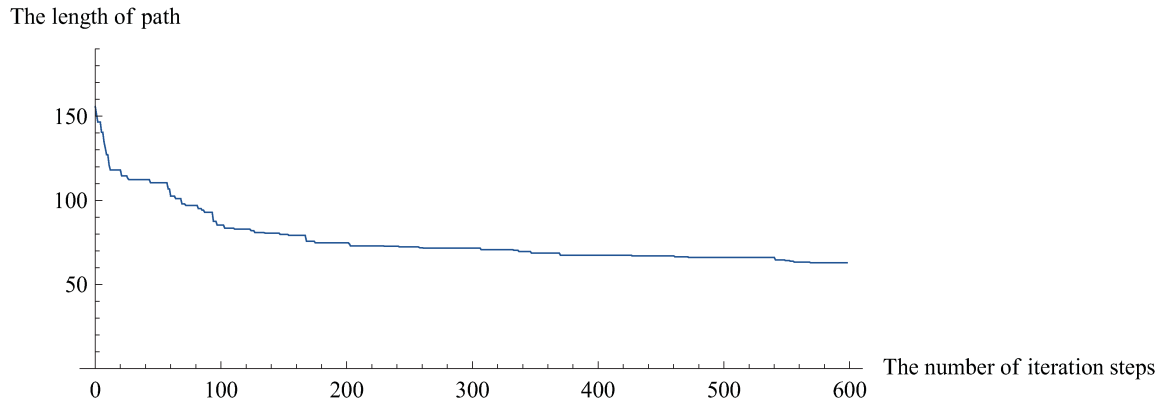


図 3: パスの長さの推移. 横軸は学習回数.

参考文献

- [1] E. Erwin, K. Obermayer and K. Schulten, *Self-organization maps: ordering, convergence properties and energy functions*, Bio. Cybern., 67 (1992), pp. 47–55.
- [2] 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
- [3] M. Hoshino and Y. Kimura, *State preserving properties in self-organizing maps with inputs in an inner product space*, Nonlinear Analysis and Convex Analysis (Tokyo, 2009), pp. 65–76.
- [4] T. Kohonen, *Self-Organizing Maps, Third Edition*, Springer, 2001.