

On conservation of states and learning processes in basic self-organizing maps with one-dimensionally indexed array

(1次元配列SOMにおける状態保存性と学習プロセスについて)

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

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の学習プロセスモデルにおける数理的特性についての一つの理論的考察である。自己組織化マップは、アルゴリズムとして非常にシンプルである。また、非常に実用的であり広範囲に応用例を有する。

自己組織化マップの学習プロセス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数理的特性について、ノードからノード値への写像としてのモデル関数とその推移について論ずる。特に、整列化の形成過程と状態クラスの閉性に注目する。

本報告では、1次元配列1次元入力を有する自己組織化マップモデルにおける吸収状態クラスである単調性の保存則に関して、より一般的な不均一な学習率をもつ学習プロセスに拡張した場合について言及する。また、そこでの数値例において、不均一な学習率の違いによって、その収束の速さがどのように変化するかについての数値計算結果を与える。

本報告では、自己組織化マップモデルをノード、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) I をすべてのノードの集合とする。 I は、距離 d をもつある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 をノードの値の空間とする。 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) 学習プロセスが次のように定義する.

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

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 $1, 2, \dots, n$ がある場合を考える. そのそれぞれに対してノードの値 $m_0(1), m_0(2), \dots, m_0(n)$ が与えられているとする. このとき, 入力とこれに伴う学習により各ノードの値が更新される. $x_0 \in X$ 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. 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, モデル関数 $m_1, m_2, m_3, \dots$ が逐次に生成され, ノードの更新が更新される. 学習を繰り返したとき, 各ノードとそれらの値との間に, ある種の規則性が高い頻度で現れる.

本報告では, この現象をモデル関数の推移とその数理的な特性として捉え, 議論する.

2. 1次元ノード配列 $\mathbb{R}$ 値モデルと状態保存

次のような設定の下での本質的かつ単純な自己組織化マップモデルを考える.

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

学習プロセス以外を以下のように仮定する.

- (i) ノード. $I = \{1, 2, \dots, N\} \subset \mathbb{N}$, $d(i, j) = |i - j|$ とする.
- (ii) ノード値. $V \subset \mathbb{R}$, $m: I \rightarrow V$ とする. ここで, ノード値の空間 V におけるノルムとして絶対値を用いることとする.
- (iii) 入力. $\{x_k\}_{k=0}^\infty \subset X \subset \mathbb{R}$ とする.

基本的な学習プロセスとして次のタイプが挙げられる. この学習プロセスに対して以下の基本的な性質が成り立つ.

Theorem 1 1次元入力型自己組織化マップ

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

において, 次の学習プロセスを仮定する.

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:

$$I(m_k, x_k) = \left\{ i^* \in I \mid i^* = \operatorname{argmin}_{i \in I} |m_k(i) - x_k| \right\}, \quad (m_k \in M, x_k \in X),$$

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

(b) 学習率: $0 \leq \alpha \leq 1$.

(c) 更新後の値:

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

$$k = 0, 1, 2, \dots$$

このとき、モデル関数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.

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, このような状態のクラスを自己組織化マップモデルの閉じた状態クラスと呼ぶことにする.

3. 不均一学習率モデルにおける閉状態クラス

本報告では, 1 次元配列 1 次元入力をもつ自己組織化マップモデルにおける吸収状態クラスである単調性の保存則に関する結果を, より一般的な不均一な学習率を伴う学習プロセスを有する場合に拡張すると共に, その下で閉性を有するための十分条件を与える. 証明は [7] を参照.

Theorem 2 1 次元入力型自己組織化マップ

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

において, 次の学習プロセスを仮定する.

学習プロセス L_A ($\varepsilon = 1, 2, \dots, \alpha_i$)

(a) 学習範囲:

$$I(m_k, x_k) = \left\{ i^* \in I \mid i^* = \operatorname{argmin}_{i \in I} |m_k(i) - x_k| \right\}, \quad (m_k \in M, x_k \in X),$$

$$N_\varepsilon(I(m_k, x_k)) = \bigcup_{i^* \in I(m_k, x_k)} \{i \in I \mid |i - i^*| \leq \varepsilon\}.$$

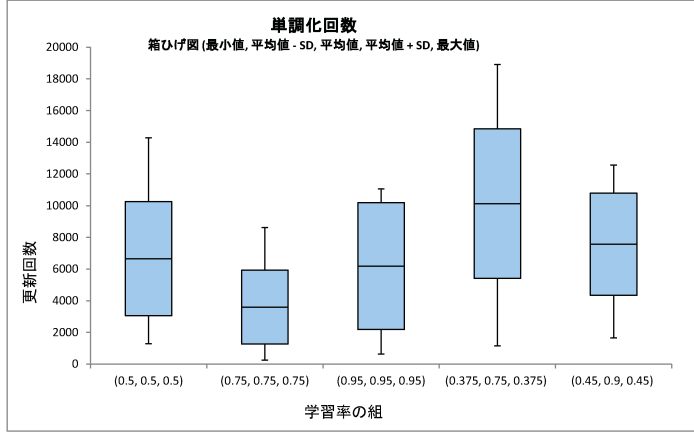


図 1: 箱ひげ図 (最小値, 平均値 - SD, 平均値, 平均値 + SD, 最大値). $(\alpha_{i^*-1}, \alpha_{i^*}, \alpha_{i^*+1})$ は学習率の組. 外れ値を考慮せず. 同一入力列. 分散分析を実施した結果, 有意な差が認められる.

(b) 学習率 α_i : $\{\alpha_i\} \subset [0, 1)$ かつ各 $i^* \in I(m_k, x_k)$ に対して

$$\alpha_i \leq \alpha_{i+1}, \quad i = i^* - \varepsilon, i^* - \varepsilon + 1, \dots, i^* - 1,$$

$$\alpha_i \geq \alpha_{i+1}, \quad i = i^*, i^* + 1, \dots, i^* + \varepsilon - 1$$

を満たす.

(c) 値の更新:

$$m_{k+1}(i) = \begin{cases} (1 - \alpha_i)m_k(i) + \alpha_i x, & i \in N_\varepsilon(I(m_k, x_k)) \text{ のとき,} \\ m_k(i), & \text{上記以外 のとき.} \end{cases}$$

このとき, 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.

上記の結果から以下が成り立つ.

Theorem 3 *Theorem 1* で, 学習プロセス L_A の学習範囲を $\varepsilon = 1, 2, \dots$ とし, 学習率を $0 \leq \alpha < 1$ とした場合においても, 同様に単調性が保存される.

4. 不均一学習率の違いによる収束性

ここでは, 数値例を与え, 不均一な学習率の違いによって, その収束の速さがどのように変化するのかについて数値計算を実施し, 検証する.

Example 1 学習プロセス $L_A(\varepsilon = 1)$ をもつノードの個数が 30 である場合を考える. ノードの初期値は $\{0, 1, \dots, 10\}$ 上の一様分布により発生させた値を用い, 入力 $\{0, 10\}$ 上の

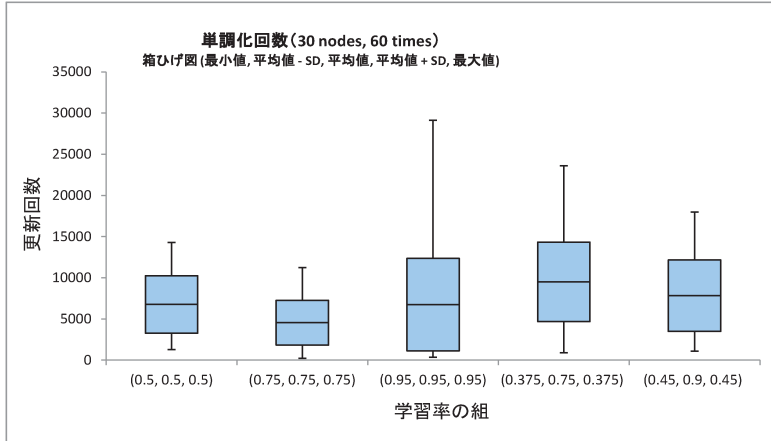


図 2: 箱ひげ図 (最小値, 平均値 - SD, 平均値, 平均値 + SD, 最大値). $(\alpha_{i^*-1}, \alpha_{i^*}, \alpha_{i^*+1})$ は学習率の組. 外れ値を考慮せず.

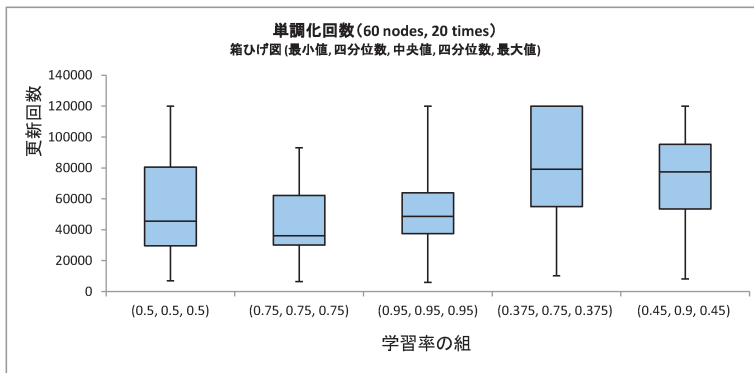


図 3: 箱ひげ図 (最小値, 四分位数, 中央値, 四分位数, 最大値). $(\alpha_{i^*-1}, \alpha_{i^*}, \alpha_{i^*+1})$ は学習率の組. 外れ値を考慮せず. 同一入力列. 更新回数の上限を 120,000 回として実施したが, 上限回数で単調化していないものは, 上限値として作成している.

一様分布によって発生させたものを用いている. 図 1 は, 5 種類の学習率の組のそれぞれに対して, 20 回の実験を実施し単調化するまでの回数を比較したものである. □

Example 2 図 2 は, 上記の例と同じ設定のノードの個数が 30 の場合において, 5 種類の学習率の組のそれぞれに対して, 入力列も変えて 60 回の実験を実施し単調化するまでの回数を比較したものである. □

Example 3 図 3 は, 前述の例と同じ設定で, ノードの個数が 60 の場合における結果である. 5 種類の学習率の組のそれぞれに対して, 20 回の実験を実施し単調化するまでの回数

を比較したものである。三つの例のそれぞれの結果において、同様の傾向をみることができる。□

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] 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
- [6] 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.
- [7] M. Hoshino, On state preserving property and nonexpansivity in self-organizing maps, to be published in J. Nonlinear and Convex Analysis, Vol.16 (2015).
- [8] T. Kohonen, *Self-Organizing Maps, Third Edition*, Springer, 2001.