Journal of Nonlinear and Convex Analysis
Volume 16, Number 11, 2015, 2255–2261
Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

ON STATE PRESERVING PROPERTY AND NONEXPANSIVITY IN SELF-ORGANIZING MAPS

MITSUHIRO HOSHINO

Dedicated to Professor Wataru Takahashi on the occasion of his 70th birthday.

Abstract. We deal with self-organizing map models referred to as Kohonen type algorithm. By repeating learning in self-organizing map models, some model functions have important properties such as ordering which appears in the relation between the array of nodes and the values of nodes. We investigate closed classes of states in self-organizing maps with a one dimensional array of nodes and variable learning sets. We give a condition that the learning mapping is non-expansive in general input self-organizing maps. The learning adopted in this result can be used as a measure for the learning process to converge.

1. Formulation of self-organizing map models

We consider self-organizing map models referred to as Kohonen [7] type algorithm. A self-organizing map algorithm is very practical and has many useful applications, such as a semantic map, a diagnosis of speech voicing, the traveling-salesman problem, and so on. There are some interesting phenomena between the array of nodes and the values of nodes in these models. Indeed practical properties in self-organizing map models are easy to observe, but they still remain without mathematical proofs in general cases. Firstly, a proof of the convergence of the learning process in the one-dimensional case was given by Cottrell and Fort [1]. Subsequently, convergence properties are more generally studied, e.g., in Erwin, Obermayer, and Schulten [2, 3, 4].

The purpose of this paper is to make a study of closed classes of states and their characterization in the model.

In this paper, we investigate closed classes of states in one dimensional arrayed self-organizing maps with a one dimensional input space. These properties can be used as a measure of estimation of the extent of ordering and the degree of converging for the learning process in order to tune a practical self-organizing map algorithm. Moreover, we give a condition that the learning mapping is non-expansive in self-organizing maps with inputs in an inner product space. In applications, it is needed to make learning processes in self-organizing map converge or stop in an appropriate state of nodes and a suitable step of process. This result can be used as a useful instrument for the learning process not to expand the difference between the values of two neighbor nodes and to converge in a practical problem such as a type of problem similar to the shortest path problem. And a self-organizing map

2010 *Mathematics Subject Classification.* 68T05.
Key words and phrases. Self-organizing maps, learning-rate factor.

with one dimensional array and two dimensional inputs is a particular case of the result in this paper, which is applied to this problem.

We consider to characterize a model $(I, V, X, \{m_k(\cdot)\}_{k=0}^{\infty})$ with four elements which consist of the *nodes*, the *values of nodes*, *inputs* and *model functions* with some *learning processes*, in this paper. There are several types of models with various spaces of nodes, spaces of their values and ways of learning for nodes. We suppose the following.

(i) We suppose an array of *nodes*. Let I denote the set of all nodes, which is called the *node set*. We assume that I is a countable set metrized by a metric d. Usually, the following are used in many applications, a finite subset of the set $\mathbb{N}$ of all natural numbers, or a finite subset of $\mathbb{N}^2$.

(ii) We suppose that each node has its *value*. V is the space of values of nodes. We assume that V is a real linear normed space with a norm $\|\cdot\|$. A mapping $m : I \to V$ transforming each node i to its value $m(i)$ is called a *model function*. Let M be the set of all model functions.

(iii) X is the *input set*. Let X be a subset of V. $x \in X$ is called an *input*.

(iv) The *learning process* is defined by the following. If an input is given, then the value of each node is renewed to a new value by the input. If an input x is given, node i learns from x and its value $m(i)$ changes to a new value $m'(i)$ determined by $m'(i) = (1 - \alpha_{m,x}(i))m(i) + \alpha_{m,x}(i)x$ according to the rate $\alpha_{m,x}(i) \in [0, 1]$. If an initial model function m_0 and a sequence $x_0, x_1, x_2, \ldots \in X$ of inputs are given, then the model functions $m_1, m_2, m_3, \ldots$ are generated sequentially according to

$$m_{k+1}(i) = (1 - \alpha_{m_k,x_k}(i))m_k(i) + \alpha_{m_k,x_k}(i)x_k, \quad k = 0, 1, 2, \ldots.$$

There are several types of models with various spaces of nodes, spaces of their values and ways of learning for nodes.

2. A fundamental self-organizing map and an absorbing class

In this paper, we restrict our considerations to a basic self-organizing map with real-valued nodes and a one-dimensional array of nodes. We suppose that a set V of values of nodes is identified with $\mathbb{R}$ which is the set of all real numbers.

We consider a model

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

(i) Let $I = \{1, 2, \ldots, n\}$ be the node set with metric $d(i, j) = |i - j|$. (ii) Assume $V = \mathbb{R}$, that is, each node is $\mathbb{R}$-valued. (iii) $x_0, x_1, x_2, \ldots \in X \subset \mathbb{R}$ is an input sequence. (iv) We assume a learning process defined by the following procedures.

Learning process $\mathrm{L_A}$ with a learning radius $\varepsilon = 1$ is as follows.
(a) Areas of learning:

$$I(m_k, x_k) = \{i^* \in I \mid |m_k(i^*) - x_k| = \inf_{i \in I} |m_k(i) - x_k|\} \tag{2.1}$$

and $N_1(i) = \{j \in I \mid |j - i| \leq 1\}$. (b) Learning-rate factor: $0 \leq \alpha \leq 1$. (c) Learning: let $N_1(I(m_k, x_k)) = \cup_{i^* \in I(m_k,x_k)} N_1(i^*)$ and $\{m_k\}$ is defined by the following, for each $k = 0, 1, 2, \ldots$, if $i \in N_1(I(m_k, x_k))$ then

$$m_{k+1}(i) = (1 - \alpha)m_k(i) + \alpha x_k, \tag{2.2}$$

otherwise $m_{k+1}(i) = m_k(i)$.

If an input $x_0 \in X$ is given, then we choose node i^* which has the most similar value to x_0 within $m_0(1), m_0(2), \ldots, m_0(n)$. Node i^* and the nodes which are in the neighborhood of i^* learn x_0 and their values change to new values $m_1(i) = (1 - \alpha)m_0(i) + \alpha x_0$. The nodes which are not in the neighborhood of i^* do not learn and their values do not change. Repeating these updating for the inputs $x_1, x_2, x_3, \ldots$, the value of each node is renewed sequentially. Simultaneously, model functions $m_1, m_2, m_3, \ldots$ are also generated sequentially. By repeating learning, some model functions have properties such as monotonicity and a certain regularity which may appear in the relation between the array of nodes and the values of nodes. Self-organizing maps apply to many practical problems by using these properties.

The following is a well-known property [7].

Theorem 2.1. *We consider a self-organizing map model*

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

with Learning process $L_A(\varepsilon = 1)$. For model functions m_1, m_2, $\ldots$, the following statements hold:

(i) *if m_k is increasing on I, that is $m_k(i) \leq m_k(i+1)$ for all i, then m_{k+1} is increasing on I;*
(ii) *if m_k is decreasing on I, that is $m_k(i) \geq m_k(i+1)$ for all i, then m_{k+1} is decreasing on I;*
(iii) *if m_k is strictly increasing on I, that is $m_k(i) < m_k(i+1)$ for all i, then m_{k+1} is strictly increasing on I;*
(iv) *if m_k is strictly decreasing on I, that is $m_k(i) > m_k(i+1)$ for all i, then m_{k+1} is strictly decreasing on I.*

The class of states with monotone in this self-organizing map is a closed class in the sense that once model function leads to increasing state, it never leads to other states for the learning by any input. Such properties as monotone are called *absorbing states* of self-organizing map models.

3. An extended learning process and an absorbing class

We give a results for preserving monotone of model functions.

Theorem 3.1. *We consider a self-organizing map model*

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

Assume Learning process L_A $(\varepsilon = 1, 2, \ldots)$ with learning rates α_i depending on node i. For learning, let $N_\varepsilon(I(m_k, x_k)) = \cup_{i^* \in I(m_k, x_k)}\{i \in I \mid |i - i^*| \leq \varepsilon\}$ and suppose that

$$m_{k+1}(i) = \begin{cases} (1 - \alpha_i)m_k(i) + \alpha_i x, & \text{if } i \in N_\varepsilon(I(m_k, x_k)), \\ m_k(i), & \text{otherwise}, \end{cases}$$

where we assume that $\{\alpha_i\} \subset [0,1)$ *satisfies, for each* $i^* \in I(m_k, x_k)$,

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

and

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

Then, for model functions m_1, m_2, $\dots$, *if* m_k *is increasing on* I, *then* m_{k+1} *is increasing on* I.

Note that similar statements hold for strictly increasing state, decreasing state and strictly decreasing state.

Proof. We show the statement for a singleton $I(m_k, x_k) = \{i^*\}$. If it is not a singleton, the theorem can also be shown by using the same argument. Suppose that model function m_k is increasing.

(a) For $i \le i^* - \varepsilon - 2$, $i \ge i^* + \varepsilon + 1$, we have $m_{k+1}(i+1) = m_k(i+1)$ and $m_{k+1}(i) = m_k(i)$. So $m_{k+1}(i) \le m_{k+1}(i+1)$.

(b) For $i = i^* - \varepsilon - 1$, we have

$$\begin{aligned} m_{k+1}(i+1) - m_{k+1}(i) &= (1 - \alpha_{i+1}) m_k(i+1) + \alpha_{i+1} x - m_k(i) \\ &= m_k(i+1) - m_k(i) + \alpha_{i+1}(x - m_k(i+1)). \end{aligned}$$

Suppose $m_k(i+1) \ge x$. Since m_k is increasing, $x \le m_k(i+1) \le m_k(i^*)$. This contradicts $i+1 \notin I(m,x)$ and $|x - m_k(i^*)| < |x - m_k(i+1)|$. So $m_k(i+1) < x$. Thus, we have $m_{k+1}(i+1) - m_{k+1}(i) \ge 0$.

(c) For $i = i^* - \varepsilon, i^* - \varepsilon + 1, \dots, i^* - 1$, we have

$$\begin{aligned} & m_{k+1}(i+1) - m_{k+1}(i) \\ &\quad = (1 - \alpha_{i+1}) m_k(i+1) + \alpha_{i+1} x - (1 - \alpha_i) m_k(i) - \alpha_i x \\ &\quad = (1 - \alpha_{i+1})(m_k(i+1) - m_k(i)) + (\alpha_{i+1} - \alpha_i)(x - m_k(i)) \ge 0 \end{aligned}$$

(d) For $i = i^*, i^* + 1, \dots, i^* + \varepsilon - 1$, we obtain $m_{k+1}(i+1) - m_{k+1}(i) \ge 0$ by using the similar argument to (c).

(e) For $i = i^* + \varepsilon$, we have

$$\begin{aligned} m_{k+1}(i+1) - m_{k+1}(i) &= m_k(i+1) - (1 - \alpha_i) m_k(i) + \alpha_i x \\ &= m_k(i+1) - m_k(i) + \alpha_i (m_k(i) - x) \ge 0. \end{aligned}$$

Thus, the renewed model function m_{k+1} is also increasing. □

Theorem 3.2. *We consider a self-organizing map model*

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

Assume Learning process L_A *for any* ε $(\varepsilon = 1, 2, \dots)$ *with a constant learning-rate factor* $0 \le \alpha < 1$. *Then, for model functions* m_1, m_2, $\dots$, *the following statements hold:*

(i) m_k *is increasing on* I, *then* m_{k+1} *is increasing on* I;
(ii) m_k *is decreasing on* I, *then* m_{k+1} *is decreasing on* I;
(iii) m_k *is strictly increasing on* I, *then* m_{k+1} *is strictly increasing on* I;
(iv) m_k *is strictly decreasing on* I, *then* m_{k+1} *is strictly decreasing on* I.

Proof. For any constant learning-rate factor $\alpha \in [0, 1)$, $\{\alpha_i\}$ given by $\alpha_i = \alpha$ for all i satisfies the condition in Theorem 3.1. The statement follows from Theorem 3.1. □

4. A condition for the learning to be a non-expansive mapping

In this section, we consider a self-organizing map with a one dimensional array and a general input space,

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

(i) The node set. Let $I = \{1, 2, \dots, n\}$ with metric $d_I(i, j) = |i - j|$, $i, j \in I$.
(ii) The values of nodes. Let $m : I \to V$, where V is a normed linear space with an inner product $\langle \cdot, \cdot \rangle$.
(iii) $x_0, x_1, x_2, \dots \in X \subset V$ is an input sequence.
(iv) Assume Learning process $\mathrm{L_m}$ with 1-dimensional array and $\varepsilon = 1$.
   (a) Areas of learning:

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

   and $N_1(i) = \{j \in I \mid d_I(i, j) \le 1\}$.
   (b) Learning-rate factor: $0 \le \alpha \le 1$.
   (c) Learning: if $i \in N_1(J(m, x))$ then $m'(i) = (1 - \alpha)m(i) + \alpha x$, otherwise $m'(i) = m(i)$.

For self-organizing maps with inputs in an inner product space, we provide a condition that the learning mapping $m(i) \mapsto m'(i)$, transforming value $m(i)$ of node i to its renewed value $m'(i)$, is non-expansive on neighboring nodes, in the sense of the following theorem.

Theorem 4.1. *We consider a self-organizing map*

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

Assume Learning process $L_m(\varepsilon = 1)$ with variable learning rates α defined by the following. Let m be an arbitrary model function and x an arbitrary input. Let m' be the renewed model function of m by x. Let $i^* = J(m, x)$,

$$\alpha_1 = \begin{cases} \dfrac{2\langle m(i^* - 2) - m(i^* - 1)\,,\, x - m(i^* - 1)\rangle}{\|x - m(i^* - 1)\|^2}, & \text{if } i^* \ge 3 \text{ and } x \ne m(i^* - 1), \\ 1, & \text{if } i^* = 1, 2 \text{ or } x = m(i^* - 1), \end{cases}$$

and

$$\alpha_2 = \begin{cases} \dfrac{2\langle m(i^* + 2) - m(i^* + 1)\,,\, x - m(i^* + 1)\rangle}{\|x - m(i^* + 1)\|^2}, & \text{if } i^* \le n - 2 \text{ and } x \ne m(i^* + 1), \\ 1, & \text{if } i^* = n - 1, n \text{ or } x = m(i^* + 1). \end{cases}$$

Then, for any α satisfying that $0 \le \alpha \le \max\{0, \min\{\alpha_1, \alpha_2, 1\}\}$,

$$\|m'(i + 1) - m'(i)\| \le \|m(i + 1) - m(i)\|, \quad i = 1, 2, \dots, n - 1. \tag{4.1}$$

Note that if $\min\{\alpha_1, \alpha_2\} \le 0$, then the model function is not renewed.

Proof. (a) For $i \leq i^*-3, i \geq i^*+2$, we have $m'(i+1) = m(i+1)$ and $m'(i) = m(i)$. Therefore (4.1) holds.

(b) We verify (4.1) for $i = i^* - 2$. We have

$$\begin{aligned}
&\|m'(i^*-1) - m'(i^*-2)\|^2 \\
&\quad = \|(1-\alpha)m(i^*-1) + \alpha x - m(i^*-2)\|^2 \\
&\quad = \|m(i^*-1) - m(i^*-2) - \alpha(m(i^*-1) - x)\|^2 \\
&\quad = \|m(i^*-1) - m(i^*-2)\|^2 - 2\alpha\langle m(i^*-1) - m(i^*-2)\, ,\, m(i^*-1) - x\rangle \\
&\qquad + \alpha^2\|m(i^*-1) - x\|^2.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\|m'(i^*-1) - m'(i^*-2)\|^2 - \|m(i^*-1) - m(i^*-2)\|^2 \\
&\quad = \alpha\left\{\|m(i^*-1) - x\|^2\alpha - 2\langle m(i^*-1) - m(i^*-2)\, ,\, m(i^*-1) - x\rangle\right\}.
\end{aligned}$$

If $x = m(i^*-1)$, then $\|m'(i^*-1) - m'(i^*-2)\| = \|m(i^*-1) - m(i^*-2)\|$ holds. Hence, for any α satisfying that $0 \leq \alpha \leq \max\{0, \min\{\alpha_1, \alpha_2, 1\}\}$,

$$\|m'(i^*-1) - m'(i^*-2)\|^2 - \|m(i^*-1) - m(i^*-2)\|^2 \leq 0.$$

So (4.1) holds for $i = i^* - 2$.

(c) For $i = i^* - 1, i^*$, we have

$$\begin{aligned}
\|m'(i+1) - m'(i)\| =& \|(1-\alpha)m(i+1) + \alpha x - (1-\alpha)m(i) - \alpha x\| \\
=& (1-\alpha)\|m(i+1) - m(i)\| \\
\leq& \|m(i+1) - m(i)\|.
\end{aligned}$$

(d) We verify (4.1) for $i = i^* + 1$. We have

$$\begin{aligned}
&\|m'(i^*+2) - m'(i^*+1)\|^2 \\
&\quad = \|m(i^*+2) - (1-\alpha)m(i^*+1) - \alpha x\|^2 \\
&\quad = \|m(i^*+2) - m(i^*+1)\|^2 - 2\alpha\langle m(i^*+2) - m(i^*+1)\, ,\, x - m(i^*+1)\rangle \\
&\qquad + \alpha^2\|x - m(i^*+1)\|^2.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\|m'(i^*+2) - m'(i^*+1)\|^2 - \|m(i^*+2) - m(i^*+1)\|^2 \\
&\quad = \alpha\left\{\|x - m(i^*+1)\|^2\alpha - 2\langle m(i^*+2) - m(i^*+1)\, ,\, x - m(i^*+1)\rangle\right\}.
\end{aligned}$$

If $x = m(i^*+1)$, then $\|m'(i^*+2) - m'(i^*+1)\| = \|m(i^*+2) - m(i^*+1)\|$ holds. Hence, for any α with $0 \leq \alpha \leq \max\{0, \min\{\alpha_1, \alpha_2, 1\}\}$,

$$\|m'(i^*+2) - m'(i^*+2)\|^2 - \|m(i^*+2) - m(i^*+1)\|^2 \leq 0.$$

So (4.1) holds for $i = i^* + 1$.

Thus (4.1) holds for any α with $0 \leq \alpha \leq \max\{0, \min\{\alpha_1, \alpha_2, 1\}\}$. □

5. Conclusions

In this paper, we investigated closed classes of state in essential one dimensional arrayed self-organizing maps to make a contribution of the theoretical properties of self-organizing map algorithm. We presented a condition that the learning mapping is non-expansive in self-organizing maps with a one dimensional array and general inputs in an inner product space. This result can be used as an instrument for the learning process not to expand the difference between the values of two neighbor nodes and to converge.

Acknowledgment

Sincere thanks go to the referees' valuable suggestions and comments.

References

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

[2] E. Erwin, K. Obermayer and K. Schulten, *Convergence properties of self-organizing maps*, in: Artificial Neural Networks, T. Kohonen, K. Mäkisara, O. Simula and J. Kangas, (eds.), 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), 35–45.

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

[5] M. Hoshino and Y. Kimura, *Absorbing states and quasi-convexity in self-organizing maps*, J. Nonlinear Convex Anal. **10** (2009), 395–406.

[6] M. Hoshino and Y. Kimura, *State preserving properties in self-organizing maps with inputs in an inner product space*, in: Nonlinear Analysis and Convex Analysis (Tokyo, 2009), S. Akashi, Y. Kimura, T. Tanaka (eds.), Yokohama Publishers, Yokohama, pp. 65–76.

[7] T. Kohonen, *Self-Organizing Maps, Third Edition*, Springer-Verlag, Berlin, Hidelberg, New York, 2001.

Manuscript received May 31, 2014
revised October 24, 2015

Mitsuhiro Hoshino
Department of Management Science and Engineering, Faculty of Systems Science and Technology, Akita Prefectural University, 84-4 Ebinokuchi Tsuchiya Yurihonjo, Akita 015-0055, Japan
E-mail address: hoshino@akita-pu.ac.jp