Journal of Nonlinear and Convex Analysis
Volume 10, Number 3, 2009, 395–406

Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

ABSORBING STATES AND QUASI-CONVEXITY IN SELF-ORGANIZING MAPS

MITSUHIRO HOSHINO AND YUTAKA KIMURA

Abstract. The purpose of this paper is to make a study of absorbing states and their characterization in self-organizing map models. Self-organizing map algorithm is very practical and has many useful applications. However, its theoretical and mathematical structure is not clear. We introduce quasi-convexity for model function in basic self-organizing map models.

1. Formulation of self-organizing map models

We consider self-organizing map models referred to as Kohonen [5] type algorithm. Self-organizing map algorithm is very practical and has many useful applications, semantic map, diagnosis of speech voicing, solving traveling-salesman problem, and so on. There are some interesting phenomena between the array of nodes and the values of nodes in these models.

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 followings.

(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. In many applications, usually, we use the following ones, 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 normed linear space with $\|\cdot\|$. A mapping $m : I \to V$ transforming each node i to its value $m(i)$ is called a *model function* or a *reference 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 according to the input. If an input x is given, node i learns x and its value $m(i)$ changes to a new value $m'(i)$ determined by

$$m'(i) = (1 - \alpha_{m,x}(i))m(i) + \alpha_{m,x}(i)x \tag{1.1}$$

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

2000 *Mathematics Subject Classification.* Primary 47H05; Secondary 41A65.
Key words and phrases. Self-organizing map.

$m_1, m_2, m_3, \dots$ are generated sequentially according to the following equation.

$$m_{k+1}(i) = (1 - \alpha_{m_k,x_k}(i))m_k(i) + \alpha_{m_k,x_k}(i)x_k. \tag{1.2}$$

There are several types of models with various spaces of nodes, spaces of their values and ways of learning for nodes. In this paper, we use two types of learning processes defined by the following.

Learning process A.

(a) Areas of learning: for each $m_k \in M$ and $x_k \in X$,

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

$$N_\varepsilon(i) = \left\{ j \in I \;\middle|\; d(j, i) \le \varepsilon \right\} \quad (\varepsilon > 0 \text{ is the learning radius, } i \in I).$$

(b) Learning-rate factor: $0 \le \alpha \le 1$.

(c) Learning:

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

We note that, if $I(m_k, x_k) = \emptyset$ then $m_{k+1} = m_k$.

Learning process B. This learning process is the same as Learning process A except that the one node $J(m_k, x_k)$ is selected from $I(m_k, x_k)$ by the given rule. If $i \in N_\varepsilon(J(m_k, x_k))$ then $m_{k+1}(i) = (1-\alpha)m_k(i) + \alpha x_k$, otherwise $m_{k+1}(i) = m_k(i)$. For example, if I is a totally ordered finete set, $J(m_k, x_k)$ may be defined by

$$J(m_k, x_k) = \min\left\{ i^* \in I \;\middle|\; \|m_k(i^*) - x_k\| = \inf_{i \in I} \|m_k(i) - x_k\| \right\}.$$

2. The fundamental type of self-organizing map

In Sections 2 and 3, we restrict our considerations to basic self-organizing maps with real-valued nodes and one dimensional array of nodes. Now, we suppose that set V of values of nodes is identified with $\mathbb{R}$ which is the set of all real numbers.

(i) Let $I = \{1, 2, \dots, n\}$ be the node set with metric $d(i, j) = |i - j|$. (ii) Assume $V = \mathbb{R}$, that is, each node is $\mathbb{R}$-valued. A model function m is written as $m = [m(1), m(2), \dots, m(n)]$. (iii) $x_0, x_1, x_2, \dots \in X \subset \mathbb{R}$ is an input sequence. (iv) In Sections 2 and 3, we assume two learning processes defined by the followings. These learning processes are essential in both theoretical study and application of self-organizing map models.

Learning process A. (1-dimensional array, $\mathbb{R}$-valued nodes and $\varepsilon = 1$) (a) areas of learning: $I(m_k, x_k) = \{i^* \in I \mid |m_k(i^*) - x_k| = \inf_{i \in I} |m_k(i) - x_k|\}$ and $N_1(i) = \{j \in I \mid |j - i| \le 1\}$; (b) learning-rate factor: $0 \le \alpha \le 1$; (c) learning: for each $k = 0, 1, 2, \dots$, if $i \in \cup_{i^* \in I(m_k,x_k)} N_1(i^*)$ then $m_{k+1}(i) = (1-\alpha)m_k(i) + \alpha x_k$, otherwise $m_{k+1}(i) = m_k(i)$.

Learning process B. (1-dimensional array, $\mathbb{R}$-valued nodes and $\varepsilon = 1$) (a) areas of learning: $J(m_k, x_k) = \min\{i^* \in I \mid |m_k(i^*) - x_k| = \inf_{i \in I} |m_k(i) - x_k|\}$ and $N_1(i) = \{j \in I \mid |j - i| \le 1\}$; (b) learning-rate factor: $0 \le \alpha \le 1$; (c) learning:

for each $k = 0, 1, 2, \ldots$, if $i \in N_1(J(m_k, x_k))$ then $m_{k+1}(i) = (1-\alpha)m_k(i) + \alpha x_k$, 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 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.

3. Quasi-convexity and quasi-concavity of model function

In this section, we deal with the case of the 1-dimensional array, real-valued nodes and the learning radius of $\varepsilon = 1$ defined in Section 2. The following properties are well-known results.

Theorem 3.1. *We consider a self-organizing map model* $(I = \{1, 2, \ldots, n\}, V = \mathbb{R}, X \subset \mathbb{R}, \{m_k(\cdot)\}_{k=0}^{\infty})$ *with Learning process* $A(\varepsilon = 1)$*. For model functions* m_1*,* m_2*,* m_3*, …, the following statements hold:*

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

Such properties as monotonicity may be called *absorbing states* of self-organizing map models in the sense that once model function leads to increasing state, it never leads to other states for the learning by any input. The purpose of this paper is to make a study of absorbing states and their characterization.

We introduce quasi-convexity and quasi-concavity of model function in fundamental self-organization maps. Generally, we use convexity and quasi-convexity for functions on convex sets. However, model functions in self-organizing maps are not defined on a linear space, therefore, are not defined on a convex set in usual sense. Now, we define quasi-convexity and quasi-concavity for a function on a totally ordered set instead of a convex set.

Definition 3.2. Let $(Y, \leq)$ be a totally ordered set and f a real-valued function on Y. Then f is said to be *quasi-convex* on Y if for any $y_1, y_2, y_3 \in Y$ with $y_1 \leq y_2 \leq y_3$,

$$f(y_2) \leq \max\{f(y_1), f(y_3)\}. \tag{3.1}$$

Also, f is said to be *quasi-concave* on Y if for any $y_1, y_2, y_3 \in Y$ with $y_1 \leq y_2 \leq y_3$,

$$f(y_2) \geq \min\{f(y_1), f(y_3)\}. \tag{3.2}$$

We give a necessary and sufficient condition for the quasi-convexity defined on a totally ordered set.

Theorem 3.3. *Let Y be a totally ordered set. Let f be a real-valued function on Y. For each $a \in \mathbb{R}$, we put $L_a(f) = \{y \in Y \mid f(y) \le a\}$, which is called a level set of f. Then the following statements are equivalent.*

(i) *f is a quasi-convex function on Y;*
(ii) *for all $a \in \mathbb{R}$, if $y_1, y_3 \in L_a(f)$ and $y_1 \le y_2 \le y_3$, then $y_2 \in L_a(f)$.*

Proof. (i) $\Rightarrow$ (ii) If suppose $y_1, y_3 \in L_a(f)$ and $y_1 \le y_2 \le y_3$. Then, we have $f(y_1) \le a$ and $f(y_3) \le a$. By the quasi-convexity of f, $f(y_2) \le \max\{f(y_1), f(y_3)\} \le a$. This implies $y_2 \in L_a(f)$.

(ii) $\Rightarrow$ (i) Let $y_1 \le y_2 \le y_3$. Putting $a = \max\{f(y_1), f(y_3)\}$, we have $y_1, y_3 \in L_a(f)$. By condition (ii), we have $y_2 \in L_a(f)$. Therefore,

$$f(y_2) \le a = \max\{f(y_1), f(y_3)\}.$$

Thus f is a quasi-convex function. □

The following theorem is a result about properties of quasi-convex functions.

Theorem 3.4. *Let Y be a totally ordered set. Let f be a real-valued function on Y. Then the following statements hold:*

(i) *f is a quasi-convex function on Y if and only if $-f$ is a quasi-concave function on Y;*
(ii) *f is quasi-convex and quasi-concave on Y if and only if f is monotone on Y, that is, either $f(y_1) \le f(y_2)$ for every $y_1 \le y_2$, or $f(y_1) \ge f(y_2)$ for every $y_1 \le y_2$.*

Theorem 3.5. *We consider a self-organizing map model $(I = \{1, 2, \dots, n\}, V = \mathbb{R}, X \subset \mathbb{R}, \{m_k(\cdot)\}_{k=0}^{\infty})$ with Learning process A($\varepsilon = 1$). For model functions m_1, m_2, m_3, ..., the following statements hold:*

(i) *if m_k is quasi-convex on I, then m_{k+1} is quasi-convex on I;*
(ii) *if m_k is quasi-concave on I, then m_{k+1} is quasi-concave on I.*

Proof. (i): Suppose that m_k is quasi-convex on I. Take any $i_1, i_2, i_3 \in I$ with $i_1 < i_2 < i_3$. Let x_k be the current input. We put

$$Q = \max\{m_{k+1}(i_1), m_{k+1}(i_3)\} - m_{k+1}(i_2).$$

In order to prove that m_{k+1} is quasi-convex, we show $Q \ge 0$ in eight cases (A)-(H).

Case (A): $i_1, i_2, i_3 \in \cup_{i^* \in I(m_k, x_k)} N_1(i^*)$. We have

$$\begin{aligned} Q &= \max\{(1-\alpha)m_k(i_1) + \alpha x_k,\ (1-\alpha)m_k(i_3) + \alpha x_k\} - ((1-\alpha)m_k(i_2) + \alpha x_k) \\ &= (1-\alpha)\left(\max\{m_k(i_1),\ m_k(i_3)\} - m_k(i_2)\right) \ge 0. \end{aligned}$$

Case (B): $i_1, i_2 \in \cup_{i^* \in I(m_k, x_k)} N_1(i^*)$ and $i_3 \notin \cup_{i^* \in I(m_k, x_k)} N_1(i^*)$. We have

$$\begin{aligned} Q &= \max\{(1-\alpha)m_k(i_1) + \alpha x_k,\ m_k(i_3)\} - ((1-\alpha)m_k(i_2) + \alpha x_k) \\ &= \max\{(1-\alpha)(m_k(i_1) - m_k(i_2)),\ (1-\alpha)(m_k(i_3) - m_k(i_2)) + \alpha(m_k(i_3) - x_k)\}. \end{aligned}$$

(B1): If $m_k(i_1) \ge m_k(i_2)$, then the first element in the maximum operation of Q is nonnegative. Hence we have $Q \ge 0$. (B2): If $m_k(i_1) < m_k(i_2)$, then $m_k(i_3) \ge$

$m_k(i_2)$. We show $m_k(i_2-1) \le m_k(i_2)$. Suppose $m_k(i_2-1) > m_k(i_2)$. Then we have

$$m_k(i_2-1) > \max\{m_k(i_1), m_k(i_2)\}.$$

This inequality contradicts that m_k is quasi-convex. Similarly, we have

$$m_k(i_2) \le m_k(i_2+1) \le m_k(i_3)$$

by the quasi-convexity of m_k. Now, we show $m_k(i_3) \ge x_k$. Suppose $m_k(i_3) < x_k$. Since

$$m_k(i_2-1) \le m_k(i_2) \le m_k(i_2+1) \le m_k(i_3) < x_k$$

and $i_2 \in \cup_{i^* \in I(m_k,x_k)} N_1(i^*)$, we have $i_3 \in I(m_k, x_k)$. This contradicts the condition of Case (B). Therefore, $m_k(i_3) \ge m_k(i_2)$ and $m_k(i_3) \ge x_k$ imply that the second element in the maximum operation of Q is nonnegative. Thus, we have $Q \ge 0$.

Case (C): $i_1, i_3 \in \cup_{i^* \in I(m_k,x_k)} N_1(i^*)$ and $i_2 \notin \cup_{i^* \in I(m_k,x_k)} N_1(i^*)$. We have

$$\begin{aligned} Q &= \max\{(1-\alpha)m_k(i_1) + \alpha x_k,\ (1-\alpha)m_k(i_3) + \alpha x_k\} - m_k(i_2) \\ &= \max\{(1-\alpha)(m_k(i_1) - m_k(i_2)) + \alpha(x_k - m_k(i_2)), \\ &\qquad (1-\alpha)(m_k(i_3) - m_k(i_2)) + \alpha(x_k - m_k(i_2))\}. \end{aligned}$$

(C1): If $m_k(i_1) \ge m_k(i_2)$ and $m_k(i_3) \ge m_k(i_2)$, then it follows from the quasi-convexity of m_k that $m_k(i_1-1) \ge m_k(i_1)$ and $m_k(i_3+1) \ge m_k(i_3)$. Moreover, by the quasi-convexity of m_k, we have

$$m_k(i_2) \le m_k(i_1+1) \le m_k(i_1)$$

or

$$m_k(i_2) \le m_k(i_3-1) \le m_k(i_3).$$

Now, we show $x_k \ge m_k(i_2)$. Suppose $x_k < m_k(i_2)$. Since

$$x_k < m_k(i_2) \le m_k(i_1+1) \le m_k(i_1) \le m_k(i_1-1)$$

or

$$x_k < m_k(i_2) \le m_k(i_3-1) \le m_k(i_3) \le m_k(i_3+1),$$

$i_1, i_3 \in \cup_{i^* \in I(m_k,x_k)} N_1(i^*)$ implies $i_2 \in I(m_k, x_k)$. This contradicts the condition of Case (C). Therefore, $Q \ge 0$ holds in Case (C1). (C2): If $m_k(i_1) < m_k(i_2) \le m_k(i_3)$, then it follows from the quasi-convexity of m_k that

$$m_k(i_2) \le m_k(i_3-1) \le m_k(i_3) \le m_k(i_3+1).$$

Now, suppose $x_k < m_k(i_2)$. Then $i_3 \in \cup_{i^* \in I(m_k,x_k)} N_1(i^*)$ implies $i_2 \in I(m_k, x_k)$. This contradicts the condition of Case (C). Therefore, $x_k \ge m_k(i_2)$. Hence the second element in the maximum operation of Q is nonnegative and $Q \ge 0$ holds in Case (C1). (C3): If $m_k(i_3) < m_k(i_2) \le m_k(i_1)$, then, from the proof of Case (C2) and the symmetry of i_1 and i_3, it follows that the first element in the maximum operation of Q is nonnegative and $Q \ge 0$ holds in Case (C3).

Case (D): $i_1 \in \cup_{i^* \in I(m_k,x_k)} N_1(i^*)$ and $i_2, i_3 \notin \cup_{i^* \in I(m_k,x_k)} N_1(i^*)$. We have

$$\begin{aligned} Q &= \max\{(1-\alpha)m_k(i_1) + \alpha x_k,\ m_k(i_3)\} - m_k(i_2) \\ &= \max\{(1-\alpha)(m_k(i_1) - m_k(i_2)) + \alpha(x_k - m_k(i_2)),\ m_k(i_3) - m_k(i_2)\}. \end{aligned}$$

(D1): If $m_k(i_3) \geq m_k(i_2)$, then the second element in the maximum operation of Q is nonnegative. Hence we have $Q \geq 0$. (D2): If $m_k(i_3) < m_k(i_2)$, then $m_k(i_2) \leq m_k(i_1)$. Moreover, by the quasi-convexity of m_k, we have

$$m_k(i_1 - 1) \geq m_k(i_1) \geq m_k(i_1 + 1) \geq m_k(i_2).$$

Now, suppose $x_k < m_k(i_2)$. Then, $i_1 \in \cup_{i^* \in I(m_k, x_k)} N_1(i^*)$ implies $i_2 \in I(m_k, x_k)$. This contradicts the condition of Case (D). Therefore, $x_k \geq m_k(i_2)$. Hence the first element in the maximum operation of Q is nonnegative and $Q \geq 0$ holds in Case (D2).

Case (E): $i_2, i_3 \in \cup_{i^* \in I(m_k, x_k)} N_1(i^*)$ and $i_1 \notin \cup_{i^* \in I(m_k, x_k)} N_1(i^*)$. By the symmetry of i_1 and i_3, it follows from the proof of Case (B) that $Q \geq 0$.

Case (F): $i_2 \in \cup_{i^* \in I(m_k, x_k)} N_1(i^*)$ and $i_1, i_3 \notin \cup_{i^* \in I(m_k, x_k)} N_1(i^*)$. We have

$$\begin{aligned} Q &= \max\{m_k(i_1),\ m_k(i_3)\} - ((1-\alpha)m_k(i_2) + \alpha x_k) \\ &= \max\{(1-\alpha)(m_k(i_1) - m_k(i_2)) + \alpha(m_k(i_1) - x_k), \\ &\qquad (1-\alpha)(m_k(i_3) - m_k(i_2)) + \alpha(m_k(i_3) - x_k)\}. \end{aligned}$$

(F1): If $m_k(i_1) \geq m_k(i_2)$ and $m_k(i_3) \geq m_k(i_2)$, then, from the quasi-convexity of m_k, it follows that $m_k(i_2 - 1) \leq m_k(i_1)$ and $m_k(i_2 + 1) \leq m_k(i_3)$. Now, suppose $m_k(i_1) < x_k$ and $m_k(i_3) < x_k$. Then $i_2 \in \cup_{i^* \in I(m_k, x_k)} N_1(i^*)$ implies $i_1, i_3 \in I(m_k, x_k)$. This contradicts the condition of Case (F). Therefore, we have $m_k(i_1) \geq x_k$ or $m_k(i_3) \geq x_k$. Hence, $Q \geq 0$ holds in Case (F1). (F2): If $m_k(i_1) \geq m_k(i_2) > m_k(i_3)$, then, by using the quasi-convexity of m_k, we have

$$m_k(i_1) \geq m_k(i_2 - 1) \geq m_k(i_2) \geq m_k(i_2 + 1).$$

Now, suppose $m_k(i_1) < x_k$. Then, by $i_2 \in \cup_{i^* \in I(m_k, x_k)} N_1(i^*)$, we have $i_1 \in I(m_k, x_k)$, which contradicts the condition of Case (F). Therefore, $m_k(i_1) \geq x_k$. It follows that the first element in the maximum operation of Q is nonnegative and $Q \geq 0$ holds in case (F2). (F3): If $m_k(i_3) \geq m_k(i_2) > m_k(i_1)$, then, by the proof of case (F2) and the symmetry of i_1 and i_3, the first element in the maximum operation of Q is nonnegative and $Q \geq 0$ holds in Case (F3).

Case (G): $i_3 \in \cup_{i^* \in I(m_k, x_k)} N_1(i^*)$ and $i_1, i_2 \notin \cup_{i^* \in I(m_k, x_k)} N_1(i^*)$. By the symmetry of i_1 and i_3, it follow from the proof of Case (D) that $Q \geq 0$.

Case (H): $i_1, i_2, i_3 \notin \cup_{i^* \in I(m_k, x_k)} N_1(i^*)$. In this case, $Q \geq 0$ is clear.

Thus, m_{k+1} is quasi-convex. Similarly, (ii) is also proved. □

We note that, by Theorem 3.5, we observe that the state model functions are quasi-convex or quasi-concave before the state model functions are monotone.

Theorem 3.6. *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* $B(\varepsilon = 1)$*. For model functions* $m_1, m_2, m_3, \ldots$, *the following statements hold:*

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

Definition 3.7. Let $(Y, \leq)$ be a totally ordered set and f a real-valued function on Y. Then f is said to be *strongly quasi-convex* on Y if for any $y_1, y_2, y_3 \in Y$ with

$y_1 < y_2 < y_3$,

$$f(y_2) < \max\{f(y_1), f(y_3)\}. \tag{3.3}$$

Also, f is said to be *strongly quasi-concave* on Y if for any $y_1, y_2, y_3 \in Y$ with $y_1 < y_2 < y_3$,

$$f(y_2) > \min\{f(y_1), f(y_3)\}. \tag{3.4}$$

The following theorem gives elementary properties for strongly quasi-convex functions and strongly quasi-concave functions.

Theorem 3.8. *Let Y be a totally ordered set. Let f be a real-valued function on Y. Then the following statements hold:*

(i) *if f is a strongly quasi-convex function on Y, then f is a quasi-convex function on Y;*

(ii) *if f is a strongly quasi-concave function on Y, then f is a quasi-concave function on Y.*

Proof. These statements are directly proved by their definitions. □

The following theorems show that strongly quasi-convexity or strongly quasi-concavity is also an absorbing state of self-organizing map models.

Theorem 3.9. *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 A($\varepsilon = 1$). For model functions $m_1, m_2, m_3, \ldots$, the following statements hold:*

(i) *if m_k is strongly quasi-convex on I, then m_{k+1} is strongly quasi-convex on I;*

(ii) *if m_k is strongly quasi-concave on I, then m_{k+1} is strongly quasi-concave on I.*

Theorem 3.10. *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 B($\varepsilon = 1$). For model functions $m_1, m_2, m_3, \ldots$, the following statements hold:*

(i) *if m_k is strongly quasi-convex on I, then m_{k+1} is strongly quasi-convex on I;*

(ii) *if m_k is strongly quasi-concave on I, then m_{k+1} is strongly quasi-concave on I.*

Strongly quasi-convexity of m_k implies strongly quasi-convexity of m_{k+1} in Theorems 3.9 and 3.10 by the arguments used in the proof of Theorem 3.5.

4. 2-dimensional array case

In this section, we suppose the case of nodes with values in a normed linear space and 2-dimensional array. (i) The node set. Let $I = I_1 \times I_2$ with a metric d_I, where $I_1 = \{1, 2, \ldots, N_1\}$, $I_2 = \{1, 2, \ldots, N_2\}$ and $d_I((i, j), (k, l)) = |i - k| + |j - l|$, $(i, j), (k, l) \in I$. (ii) The values of nodes. Let $m : I_1 \times I_2 \to V$, where V is a normed linear space with an inner product $\langle \cdot, \cdot \rangle$. (iii) $x_0, x_1, x_2, \ldots \in X \subset V$ is an input sequence. (iv) Assume Larning process A (2-dimensional array and $\varepsilon = 1$) (a) areas of learning: $I(m,x) = \{(i^*,j^*) \in I \mid \|m(i^*,j^*) - x\| = \inf_{(i,j)\in I} \|m(i,j) - x\|\}$, $m \in M$, $x \in X$ and $N_1(i,j) = \{(k,l) \in I \mid d_I((i,j),(k,l)) \leq 1\}$; (b) learning-rate factor: $0 \leq \alpha \leq 1$; (c) learning: if $(i,j) \in \cup_{(i^*,j^*)\in I(m,x)} N_1(i^*,j^*)$ then $m'(i,j) = (1-\alpha)m(i,j) + \alpha x$, otherwise $m'(i,j) = m(i,j)$.

We introduce the following condition.

Condition E. For all $(i,j) \in I$,

$$\|m(i+1,j) - m(i,j)\| \leq \|m(i+2,j) - m(i,j)\|,$$
$$\|m(i-1,j) - m(i,j)\| \leq \|m(i-2,j) - m(i,j)\|,$$
$$\|m(i,j+1) - m(i,j)\| \leq \|m(i,j+2) - m(i,j)\|,$$
$$\|m(i,j-1) - m(i,j)\| \leq \|m(i,j-2) - m(i,j)\|.$$

In this situation, if the learning-rate factor is taken sufficiently small, Condition E is preserved. However, E is not preserved for a quick or large change by the learning. The next theorem gives a result with respect to the preservation of Condition E for an arbitrary α.

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

$$\big(\{1,2,\ldots,N_1\} \times \{1,2,\ldots,N_2\},\ V,\ X,\ \{m_k(\cdot,\cdot)\}_{k=0}^{\infty}\big)$$

with Learning process $A(\varepsilon = 1)$. Let m be an arbitrary model function and x an arbitrary input. Let m' be the renewed model function of m by x.

(i) *For $(i,j),(i+1,j),(i+2,j) \notin \cup_{(i^*,j^*)\in I(m,x)} N_1(i^*,j^*)$, if*

$$\|m(i+1,j) - m(i,j)\| \leq \|m(i+2,j) - m(i,j)\|,$$

then

$$\|m'(i+1,j) - m'(i,j)\| \leq \|m'(i+2,j) - m'(i,j)\|.$$

Moreover, for $(i,j),(i-1,j),(i-2,j) \notin \cup_{(i^*,j^*)\in I(m,x)} N_1(i^*,j^*)$, if

$$\|m(i-1,j) - m(i,j)\| \leq \|m(i-2,j) - m(i,j)\|, \tag{4.1}$$

then (4.1) also holds for m'.

For $(i,j),(i,j+1),(i,j+2) \notin \cup_{(i^*,j^*)\in I(m,x)} N_1(i^*,j^*)$, if

$$\|m(i,j+1) - m(i,j)\| \leq \|m(i,j+2) - m(i,j)\|, \tag{4.2}$$

then (4.2) also holds for m'.

For $(i,j),(i,j-1),(i,j-2) \notin \cup_{(i^*,j^*)\in I(m,x)} N_1(i^*,j^*)$, if

$$\|m(i,j-1) - m(i,j)\| \leq \|m(i,j-2) - m(i,j)\|, \tag{4.3}$$

then (4.3) also holds for m'.

(ii) *Let $(i^*,j^*) \in I(m,x)$. If*

$$\|m(i^*,j^*) - m(i^*-1,j^*)\| \leq \|m(i^*+1,j^*) - m(i^*-1,j^*)\|, \tag{4.4}$$

then (4.4) also holds for m'. If

$$\|m(i^*,j^*) - m(i^*+1,j^*)\| \leq \|m(i^*-1,j^*) - m(i^*+1,j^*)\|, \tag{4.5}$$

then (4.5) also holds for m'. If

$$\|m(i^*,j^*) - m(i^*,j^*-1)\| \leq \|m(i^*,j^*+1) - m(i^*,j^*-1)\|, \tag{4.6}$$

then (4.6) also holds for m'. If

$$\|m(i^*,j^*) - m(i^*,j^*+1)\| \leq \|m(i^*,j^*-1) - m(i^*,j^*+1)\|, \tag{4.7}$$

then (4.7) also holds for m'.

(iii) *Let* $(i^*, j^*) \in I(m, x)$. *If*

$$\|m(i^* + 1, j^*) - m(i^*, j^*)\| \leq \|m(i^* + 2, j^*) - m(i^*, j^*)\|, \tag{4.8}$$

then (4.8) also holds for m'. *If*

$$\|m(i^* - 1, j^*) - m(i^*, j^*)\| \leq \|m(i^* - 2, j^*) - m(i^*, j^*)\|, \tag{4.9}$$

then (4.9) also holds for m'. *If*

$$\|m(i^*, j^* + 1) - m(i^*, j^*)\| \leq \|m(i^*, j^* + 2) - m(i^*, j^*)\|, \tag{4.10}$$

then (4.10) also holds for m'. *If*

$$\|m(i^*, j^* - 1) - m(i^*, j^*)\| \leq \|m(i^*, j^* - 2) - m(i^*, j^*)\|, \tag{4.11}$$

then (4.11) also holds for m'.

Proof. (i) For $(i, j), (i + 1, j), (i + 2, j) \notin \cup_{(i^*,j^*)\in I(m,x)} N_1(i^*, j^*)$,

$$\begin{aligned}&\|m'(i + 2, j) - m'(i, j)\| - \|m'(i + 1, j) - m'(i, j)\|\\ =&\|m(i + 2, j) - m(i, j)\| - \|m(i + 1, j) - m(i, j)\|.\end{aligned}$$

Therefore, the first statement holds. By the same argument, other statements of (i) also hold.

(ii) If (4.4) holds, then

$$\begin{aligned}&\|m'(i^* + 1, j^*) - m'(i^* - 1, j^*)\| - \|m'(i^*, j^*) - m'(i^* - 1, j^*)\|\\ =&\|(1 - \alpha)m(i^* + 1, j^*) + \alpha x - (1 - \alpha)m(i^* - 1, j^*) - \alpha x\|\\ &\quad - \|(1 - \alpha)m(i^*, j^*) + \alpha x - (1 - \alpha)m(i^* - 1, j^*) - \alpha x\|\\ =&(1 - \alpha)\big\{\|m(i^* + 1, j^*) - m(i^* - 1, j^*)\| - \|m(i^*, j^*) - m(i^* - 1, j^*)\|\big\} \geq 0.\end{aligned}$$

By the same argument, we obtain other statements of (ii).

(iii) Let $m_0 = m(i^*, j^*)$, $m_1 = m(i^*+1, j^*)$, $m_2 = m(i^*+2, j^*)$, $m_0' = m'(i^*, j^*)$, $m_1' = m'(i^* + 1, j^*)$ and $m_2' = m'(i^* + 2, j^*)$. If $\|m_1 - m_0\| \leq \|m_2 - m_0\|$, then

$$\begin{aligned}&\|m_2' - m_0'\|^2 - \|m_1' - m_0'\|^2\\ =&\|m_2 - (1 - \alpha)m_0 - \alpha x\|^2 - \|(1 - \alpha)m_1 + \alpha x - (1 - \alpha)m_0 - \alpha x\|^2\\ =&\|(1 - \alpha)(m_2 - m_0) + \alpha(m_2 - x)\|^2 - \|(1 - \alpha)(m_1 - m_0)\|^2\\ =&(1 - \alpha)^2\|m_2 - m_0\|^2 + 2(1 - \alpha)\alpha\langle m_2 - m_0, m_2 - x\rangle + \alpha^2\|m_2 - x\|^2\\ &\quad - (1 - \alpha)^2\|m_1 - m_0\|^2\\ =&(1 - \alpha)^2\big(\|m_2 - m_0\|^2 - \|m_1 - m_0\|^2\big)\\ &\quad + 2(1 - \alpha)\alpha\big(\|m_2 - x\|^2 - \langle m_0 - x, m_2 - x\rangle\big) + \alpha^2\|m_2 - x\|^2\\ \geq&(1 - \alpha)^2\big(\|m_2 - m_0\|^2 - \|m_1 - m_0\|^2\big)\\ &\quad + 2(1 - \alpha)\alpha\big(\|m_2 - x\|^2 - \|m_0 - x\|\|m_2 - x\|\big) + \alpha^2\|m_2 - x\|^2.\end{aligned}$$

Since $(i^*, j^*) \in I(m, x)$, $\|m_0 - x\| \leq \|m(i, j) - x\|$ for all $(i, j) \in I$. Therefore,

$$\begin{aligned}\|m_2' - m_0'\|^2 - \|m_1' - m_0'\|^2 \geq&(1 - \alpha)^2\big(\|m_2 - m_0\|^2 - \|m_1 - m_0\|^2\big)\\ &+ \alpha^2\|m_2 - x\|^2 \geq 0.\end{aligned}$$

Thus $\|m'_2 - m'_0\| \geq \|m'_1 - m'_0\|$, so (4.8) holds for m'. By the same argument, three other statements also hold. If $(i^*+1, j^*) \in I(m,x)$, $(i^*+2, j^*) \in I(m,x)$, or $(i^*+3, j^*) \in I(m,x)$, then these cases are reduced to (ii). $\square$

If $I(m,x)$ is a singleton, the following seven cases are in a series of three nodes $(i,j),(i+1,j),(i+2,j)$ (or $(i,j),(i,j+1),(i,j+2)$).

	(i,j)	$(i+1,j)$	$(i+2,j)$
(I)	$\circ$	$\circ$	$\circ$
(II)	$\triangle$	$\circ$	$\circ$
(III)	$\circ$	$\triangle$	$\circ$
(IV)	$\circ$	$\circ$	$\triangle$
(V)	$\circ$	$\triangle$	$*$
(VI)	$\triangle$	$*$	$\triangle$
(VII)	$*$	$\triangle$	$\circ$

$*$ is the nearest node to x. $\triangle$ is a neighbor of the nearest node to x. $\circ$ is not in the neighborhood of the nearest node to x.

Statements (i), (ii) and (iii) of the previous theorem are cases of type (I), (VI) and (VII), respectively.

5. A property in general case

We give a property in general case defined in Section 1.

Theorem 5.1. *We consider a self-organizing map model $(I, V, X, \{m_k\}_{k=0}^{\infty})$ with Learning process A, where we assume that I is a countable metrizable space and V is a normed linear space with $\|\cdot\|$. For any model function m and any input $x \in X$, if $i^* \in I(m,x)$, then $i^* \in I(m',x)$ for the renewed model function m' of m by input x.*

Proof. For any $i \in \cup_{j \in I(m,x)} N_\varepsilon(j)$,

$$\|m'(i) - x\| = \|(1-\alpha)m(i) + \alpha x - x\| = (1-\alpha)\|m(i) - x\|.$$

For any $i \notin \cup_{j \in I(m,x)} N_\varepsilon(j)$, $\|m'(i) - x\| = \|m(i) - x\|$. Moreover, if $i^* \in I(m,x)$,

$$\|m'(i^*) - x\| = (1-\alpha)\|m(i^*) - x\| \leq (1-\alpha)\|m(i) - x\|$$

for any $i \in I$. Therefore, for any $i \in \cup_{j \in I(m,x)} N_\varepsilon(j)$,

$$\|m'(i^*) - x\| \leq (1-\alpha)\|m(i) - x\| = \|m'(i) - x\|.$$

For any $i \notin \cup_{j \in I(m,x)} N_\varepsilon(j)$,

$$\|m'(i^*) - x\| \leq (1-\alpha)\|m(i) - x\| = (1-\alpha)\|m'(i) - x\| \leq \|m'(i) - x\|.$$

Thus, we obtain $\|m'(i^*) - x\| = \inf_{i \in I} \|m'(i) - x\|$. Hence $i^* \in I(m',x)$. $\square$

6. Numerical example

We give a simple numerical example of the case of $\mathbb{R}$-valued nodes and the 1-dimensional array of nodes.

Example 6.1. Consider 6 nodes model with $I = \{1, 2, 3, 4, 5, 6\}$. The initial model function is $m_0 = [2, 4, 2, 2, 5, 0]$. Assume that we observe sequentially $x_0 = 5$, $x_1 = 4$, $x_2 = 2$, $x_3 = 1$, $x_4 = 2$, $x_5 = 4$, $x_6 = 0$, $x_7 = 2$, $x_8 = 1$, $x_9 = 1$, $x_{10} = 1$, $x_{11} = 4$, $x_{12} = 3$, $x_{13} = 3$, $x_{14} = 1$, $x_{15} = 1$, ... as inputs. Suppose Learning process A with $\varepsilon = 1$ and $\alpha = \frac{1}{2}$. First, it follows from m_0 and $x_0 = 5$ that $I(m_0, x_0) = \{5\}$ and $N_1(5) = \{4, 5, 6\}$. So $m_0(1), \ldots, m_0(6)$ are updated to $m_1(i) = m_0(i)$ for $i = 1, 2, 3$ and $m_1(i) = \frac{m_0(i)+x_0}{2}$ for $i = 4, 5, 6$. Thus $m_1 = [2,\ 4,\ 2,\ 3.5,\ 5,\ 2.5]$. Repeating these update, we sequentially obtain model functions.

$$
\begin{array}{rllllllll}
m_0 = [& 2, & 4, & 2, & 2, & 5, & 0 &] \\
m_1 = [& 2, & 4, & 2, & 3.5, & 5, & 2.5 &] \\
m_2 = [& 3, & 4, & 3, & 3.5, & 5, & 2.5 &] \\
m_3 = [& 3, & 4, & 3, & 3.5, & 3.5, & 2.25 &] \\
m_4 = [& 3, & 4, & 3, & 3.5, & 2.25, & 1.625 &] \\
m_5 = [& 3, & 4, & 3, & 2.75, & 2.125, & 1.8125 &] \\
m_6 = [& 3.5, & 4, & 3.5, & 2.75, & 2.125, & 1.8125 &] \\
m_7 = [& 3.5, & 4, & 3.5, & 2.75, & 1.0625, & 0.90625 &] \\
m_8 = [& 3.5, & 4, & 2.75, & 2.375, & 1.53125, & 0.90625 &] \\
m_9 = [& 3.5, & 4, & 2.75, & 2.375, & 1.26563, & 0.953125 &] \\
m_{10} = [& 3.5, & 4, & 2.75, & 2.375, & 1.13281, & 0.976563 &] \\
m_{11} = [& 3.5, & 4, & 2.75, & 2.375, & 1.06641, & 0.988281 &] \\
m_{12} = [& 3.75, & 4, & 3.375, & 2.375, & 1.06641, & 0.988281 &] \\
m_{13} = [& 3.75, & 3.5, & 3.1875, & 2.6875, & 1.06641, & 0.988281 &] \\
m_{14} = [& 3.75, & 3.25, & 3.09375, & 2.84375, & 1.06641, & 0.988281 &] \\
m_{15} = [& 3.75, & 3.25, & 3.09375, & 2.84375, & 1.0332, & 0.994141 &] \\
 & \cdots\cdots\cdots. & & & & & &
\end{array}
$$

We notice that the model function m_k is quasi-concave on I for every $k \geq 5$ and decreasing on I for every $k \geq 13$.

Acknowledgments

The authors are thankful to Professor Tamaki Tanaka of Niigata University for his valuable comments and encouragement, and to the anonymous referees for their valuable suggestions on the original draft.

References

[1] R. M. Dudley, *Real Analysis and Probability*, Cambridge Studies in Advanced Mathematics, Vol. 74, Cambridge University Press, Cambridge, 2002 (Revised reprint of the 1989 original).

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

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

[4] 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*, in Proceedings of 6th ICIM, 2002, pp. 369–375

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

[6] W. Takahashi, *Nonlinear Functional Analysis*, Yokohama publishers, Yokohama, 2000.
[7] P. L. Zador, *Asymptotic quantization error of continuous signals and the quantization dimension*, IEEE Trans. Inform. Theory, **IT-28** (1982), 139–149.

Manuscript received October 1, 2007
revised April 14, 2009

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

Yutaka Kimura
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: yutaka@akita-pu.ac.jp