数理解析研究所講究録
第1544巻 2007年 196-205

Quasi-convexity of model function in self-organizing maps
(自己組織化マップにおけるモデル関数の準凸性について)

秋田県立大学 システム科学技術学部　星野満博 (Mitsuhiro Hoshino)
秋田県立大学 システム科学技術学部　木村 寛 (Yutaka Kimura)
Faculty of Systems Science and Technology, Akita Prefectural University

1. 基本的な自己組織化マップモデル

本報告は自己組織化マップとよばれるアルゴリズムにおけるモデル関数の性質に関する一つの理論的考察である．自己組織化マップモデルにおけるノード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考察する．また，基本的な自己組織化マップモデルにおいて，そのモデル関数に対して準凸性および準凹性の概念を導入する．

ここで扱う自己組織化マップは Kohonen [7] 型アルゴリズムとして有名であり，非常に実用的であり広範囲に応用例をもつ．アルゴリズム自身は非常にシンプルであるが，その数学的構造はあまり明らかではない．

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つの要素によって定義する．

(i) 有限個のノードを仮定する．$I=\{1,2,\ldots,n\}\subset\mathbb{N}$ をノードの集合とする．

(ii) 各ノードは，それぞれ1つの値（ここでは実数値）をもつ．ノードの値の空間を $\mathbb{R}$ とする．$m(i)$ をノード i の値として，その対応 $m:I\to\mathbb{R}$ をモデル関数 (model function, reference function) と呼ぶことにする．また，M をモデル関数の全体，$m_0:I\to\mathbb{R}$ を初期モデル関数とする．モデル関数を

$$m_0=[m_0(1),m_0(2),\cdots,m_0(n)]$$

などと記すことにする．

(iii) X をインプット集合とする．$x_0,x_1,x_2,\ldots\in X\subset\mathbb{R}$ をインプット列とする．

(iv) 学習プロセスとして以下の2つを仮定する．

Learning process 1

(a) 学習範囲:

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

(b) 学習率: $0 \leq \alpha \leq 1$.

(c) 方法:

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

Learning process 2

(a) 学習範囲:

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

$$N_1(i) = \left\{ j \in I \;\middle|\; |j - i| \leq 1 \right\} \quad (i \in I).$$

(b) 学習率: $0 \leq \alpha \leq 1$.

(c) 方法:

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

今，n 個のノード $1, 2, \ldots, n$ があり，そのそれぞれに対してノードの値 $m_0(1)$, $m_0(2)$, ..., $m_0(n)$ が与えられている．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．$x_0 \in X$ が入力されたならば，$m_0(1), m_0(2), \ldots, m_0(n)$ のなかで x_0 と最も近いものを選び，その値に対応するノード i^* とその周囲のノード i に対して学習

$$m_1(i) = (1-\alpha)m_1(i) + \alpha x_1$$

が適用され，それ以外のノードに対しては学習が適用されず，$m_1(i) = m_0(i)$ となる．インプット $x_1, x_2, x_3, \ldots$ 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 $m_1, m_2, m_3, \ldots$ 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，多くの実問題へ応用されている．

本報告では，1次元ノード配列，実数値ノードのモデルにおいて，モデル関数が単調化に至るまでのプロセスについて考察するとともに，モデル関数の準凸性，準凹性とその保存について言及する．

2. モデル関数の準凸性と準凹性について

次の定理は，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 *Learning process 1* を仮定する．モデル関数 $m_1, m_2, m_3, \ldots$ 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，自己組織化マップモデルにおいて，モデル関数に対して準凸性および準凹性を導入する．一般に，凸集合上で定義された関数に対して凸，凹，準凸，準凹などを定義するが，ここでのモデル関数は線形空間上で定義されていないので，凸集合の代わりに半順序集合上で定義される関数に対して，準凸性，準凹性を以下のように定義することにする．通常の意味での準凸関数，準凹関数の定義および性質に関しては [8][9] 等において詳しく述べられている．

Definition 1 $(Y, \leq)$ を半順序集合とし，f を Y 上の実数値関数とする．このとき，f が準凸 (quasi-convex) であるとは，$y_1 \leq y_2 \leq y_3$ である任意の $y_1, y_2, y_3 \in Y$ に対して

$$f(y_2) \leq \max\{f(y_1), f(y_3)\} \tag{1}$$

が成り立つときをいう．また，f が準凹 (quasi-concave) であるとは $y_1 \leq y_2 \leq y_3$ である任意の $y_1, y_2, y_3 \in Y$ に対して

$$f(y_2) \geq \min\{f(y_1), f(y_3)\} \tag{2}$$

が成り立つときをいう．

以下の結果は半順序集合上の関数の準凸性に関する必要十分条件を与える．

Theorem 2 $(Y, \leq)$ を半順序集合とし，f を Y 上の実数値関数とする．f のレベル集合 $L_a(f)$ を以下のように定義する．各 $a \in \mathbb{R}$ に対して，

$$L_a(f) = \{y \in Y \mid f(y) \leq a\}.$$

このとき，次の 2 つの条件は同値である．

(i) f は準凸関数である.

(ii) すべての $a \in \mathbb{R}$ に対して以下が成り立つ.

$$y_1, y_3 \in L_a(f),\ y_1 \leq y_2 \leq y_3 \text{ ならば } y_2 \in L_a(f) \text{ である.}$$

PROOF. (i) $\Rightarrow$ (ii) $y_1, y_3 \in L_a(f)$, $y_1 \leq y_2 \leq y_3$ を仮定する. このとき $f(y_1) \leq a$, $f(y_3) \leq a$ が成り立つので, f の準凸性から

$$f(y_2) \leq \max\{f(y_1), f(y_3)\} \leq a$$

が得られる. これにより $y_2 \in L_a(f)$ が成り立つ.

(ii) $\Rightarrow$ (i) $y_1 \leq y_2 \leq y_3$ とする. $a = \max\{f(y_1), f(y_3)\}$ と置くと, $f(y_1) \leq a$, $f(y_3) \leq a$ であるので $y_1, y_3 \in L_a(f)$ となる. したがって, 条件 (ii) より $y_2 \in L_a(f)$ が得られる. これにより, $f(y_2) \leq a = \max\{f(y_1), f(y_3)\}$ が成り立つので f は準凸関数である. □

以下は準凸関数と準凹関数の基本的性質である.

Theorem 3 $(Y, \leq)$ を半順序集合とし, f を Y 上の実数値関数とする. このとき, 以下が成り立つ.

(i) f が準凸関数であるための必要十分条件は $-f$ が準凹関数であることである.

(ii) f が単調関数であるための必要十分条件は f が準凸かつ準凹な関数であることである.

次の定理が示すように, モデル関数に対して, 学習の前後において準凸性および準凹性が保存される.

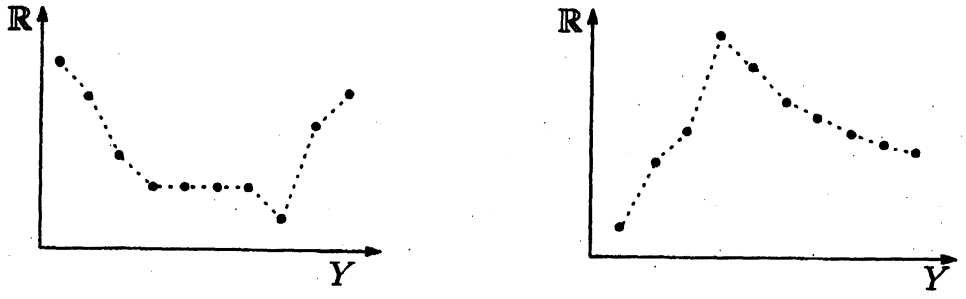

図 1: 順序集合上で定義された準凸関数 (左) と準凹関数 (右)

Theorem 4 *Learning process 1*を仮定する．このときモデル関数$m_1, m_2, m_3, \ldots$に対して以下が成り立つ．

(i) モデル関数m_kがI上で準凸であるならば，モデル関数m_{k+1}もI上で準凸である．

(ii) モデル関数m_kがI上で準凹であるならば，モデル関数m_{k+1}もI上で準凹である．

以下は，Theorem 4の証明のアウトラインである．証明の詳細は[6]を参照．

証明のアウトライン

(i): m_kがI上で準凸であることを仮定する．$i_1 < i_2 < i_3$であるノード$i_1, i_2, i_3 \in I$を任意にとる．以下の(A)-(H)の場合に対して

$$Q = \max\{m_{k+1}(i_1), m_{k+1}(i_3)\} - m_{k+1}(i_2) \geq 0$$

が成り立つことを示す．

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

Case (B): $i_1, i_2 \in \underset{i^* \in I(m_k, x_k)}{\cup} N_1(i^*)$, $i_3 \notin \underset{i^* \in I(m_k, x_k)}{\cup} N_1(i^*)$.

Case (C): $i_1, i_3 \in \underset{i^* \in I(m_k, x_k)}{\cup} N_1(i^*)$, $i_2 \notin \underset{i^* \in I(m_k, x_k)}{\cup} N_1(i^*)$.

Case (D): $i_1 \in \underset{i^* \in I(m_k, x_k)}{\cup} N_1(i^*)$; $i_2, i_3 \notin \underset{i^* \in I(m_k, x_k)}{\cup} N_1(i^*)$.

Case (E): $i_2, i_3 \in \underset{i^* \in I(m_k, x_k)}{\cup} N_1(i^*)$, $i_1 \notin \underset{i^* \in I(m_k, x_k)}{\cup} N_1(i^*)$.

Case (F): $i_2 \in \underset{i^* \in I(m_k, x_k)}{\cup} N_1(i^*)$, $i_1, i_3 \notin \underset{i^* \in I(m_k, x_k)}{\cup} N_1(i^*)$.

Case (G): $i_3 \in \underset{i^* \in I(m_k, x_k)}{\cup} N_1(i^*)$, $i_1, i_2 \notin \underset{i^* \in I(m_k, x_k)}{\cup} N_1(i^*)$.

Case (H): $i_1, i_2, i_3 \notin \underset{i^* \in I(m_k, x_k)}{\cup} N_1(i^*)$.

(ii)に関しても(i)と同様に示すことができる． □

Remark 1 Theorem 4から，モデル関数が単調な状態に到達する前の状態として準凸または準凹の状態が存在することがわかる．

以下は，Learning process 2を仮定した場合の単調性保存に関する結果である．

Theorem 5 *Learning process 2*を仮定する．モデル関数$m_1, m_2, m_3, \ldots$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．

Definition 2 $(Y, \leq)$ を半順序集合とし，f を Y 上の実数値関数とする．このとき，f が強い準凸 (strongly quasi-convex) であるとは，$y_1 < y_2 < y_3$ である任意の $y_1, y_2, y_3 \in Y$ に対して

$$f(y_2) < \max\{f(y_1), f(y_3)\} \tag{3}$$

が成り立つときをいう．また，f が強い準凹 (strongly quasi-concave) であるとは $y_1 < y_2 < y_3$ である任意の $y_1, y_2, y_3 \in Y$ に対して

$$f(y_2) > \min\{f(y_1), f(y_3)\} \tag{4}$$

が成り立つときをいう．

Remark 2 図 1 の左側の関数は準凸であるが強い準凸ではない．また，図 1 の右側の関数は強い準凹である．

以下は，半順序集合上の関数の強い準凸性と強い準凹性に関する基本的な性質である．

Theorem 6 $(Y, \leq)$ を半順序集合とし，f を Y 上の実数値関数とする．このとき，以下が成り立つ．

(i) f が強い準凸関数ならば，f は準凸関数である．

(ii) f が強い準凹関数ならば，f は準凹関数である．

次の 2 つの定理は，学習の前後におけるモデル関数の強い準凸性および強い準凹性の保存に関する結果を与える．

Theorem 7 *Learning process 1* を仮定する．モデル関数 $m_1, m_2, m_3, \ldots$ に対して，以下が成り立つ．

(i) モデル関数 m_k が I 上で強い準凸であるならば，モデル関数 m_{k+1} も I 上で強い準凸である．

(ii) モデル関数 m_k が I 上で強い準凹であるならば，モデル関数 m_{k+1} も I 上で強い準凹である．

Theorem 8 *Learning process 2* を仮定する．モデル関数 $m_1, m_2, m_3, \ldots$ に対して，以下が成り立つ．

(i) モデル関数 m_k が I 上で強い準凸であるならば，モデル関数 m_{k+1} も I 上で強い準凸である．

(ii) モデル関数 m_k が I 上で強い準凹であるならば，モデル関数 m_{k+1} も I 上で強い準凹である．

3. EXAMPLE

Example 1 次のような，実数の値をもつ6個の1次元配列されたノードからなる自己組織化マップモデルを考える．ノード集合を

$$I = \{1, 2, 3, 4, 5, 6\}$$

とする．ノード$1, 2, 3, 4, 5, 6$の初期値は，それぞれ，$m_0(1) = 2$，$m_0(2) = 4$，$m_0(3) = 2$，$m_0(4) = 2$，$m_0(5) = 5$，$m_0(6) = 0$ である．したがって，初期モデル関数は

$$m_0 = [2, 4, 2, 2, 5, 0]$$

である．次のようなインプットがもたらされたと仮定する．

$$\begin{array}{llllll} x_0 = 5, & x_1 = 4, & x_2 = 2, & x_3 = 1, & x_4 = 2, & x_5 = 4, \\ x_6 = 0, & x_7 = 2, & x_8 = 1, & x_9 = 1, & x_{10} = 1, & x_{11} = 4, \\ x_{12} = 3, & x_{13} = 3, & x_{14} = 1, & x_{15} = 1, & \ldots. & \end{array}$$

また，学習プロセスとして学習率$\alpha = \frac{1}{2}$をもつLearning process 1 を仮定する．このとき，各ノードの値は次のように更新される．

(1) モデル関数 $m_0 = [2, 4, 2, 2, 5, 0]$

(2) m_1: インプット $x_0 = 5$

$I(m_0, x_0) = \{5\}$, $N_1(5) = \{4, 5, 6\}$,
モデル関数 $m_1 = [2,\ 4,\ 2,\ 3.5,\ 5,\ 2.5]$.

(3) m_2: インプット $x_1 = 4$

$I(m_1, x_1) = \{2\}$, $N_1(2) = \{1, 2, 3\}$,
モデル関数 $m_2 = [3,\ 4,\ 3,\ 3.5,\ 5,\ 2.5]$.

これらの更新を繰り返すことにより，次のようなモデル関数が得られる．

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

モデル関数 m_k は，$k \geq 5$ のとき I 上で準凹であり，$k \geq 13$ のとき I 上で単調減少となっていることがわかる． □

Example 2 次のような，実数の値をもつ100個の1次元配列されたノードからなる自己組織化マップモデルを考える．ノード集合を

$$I = \{1, 2, 3, \ldots, 100\}$$

とする．ノード $1, 2, 3, 4, 5, 6$ の初期値は，以下の初期モデル関数によって与えられる．

$$
\begin{aligned}
m_0 = [&5, 1, 6, 6, 3, 1, 0, 3, 0, 7, 9, 2, 2, 10, 5, 7, 9, 5, 6, 1, 7, 6, 8, 5, 9, 3, 9, 1, 9, \\
&2, 4, 9, 9, 10, 3, 9, 1, 9, 8, 10, 0, 7, 2, 1, 3, 0, 9, 6, 4, 10, 4, 1, 8, 0, 0, 9, 6, \\
&8, 0, 10, 3, 6, 4, 8, 0, 10, 3, 9, 9, 0, 4, 10, 6, 9, 1, 7, 8, 5, 9, 5, 1, 9, 6, 3, 7, \\
&5, 2, 2, 3, 5, 0, 7, 0, 2, 2, 4, 3, 1, 10, 3].
\end{aligned}
$$

インプット列は $[0, 10]$ 上の一様分布に従う確率変数によって生成させる．

$$
\begin{aligned}
x = &6.17655, 5.74143, 3.09101, 8.82768, 0.419905, 5.44219, 2.87489, 9.34485, \\
&2.83286, 8.54906, 4.73626, 0.181078, 2.97653, 4.9316, 5.73355, \ldots .
\end{aligned}
$$

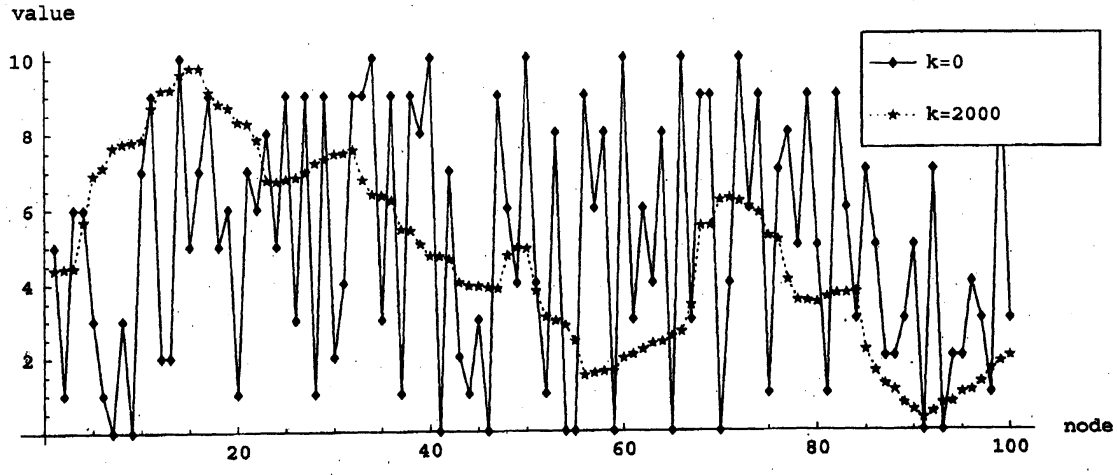

図 2:　初期モデル関数と 2000 回更新した後のモデル関数 (k は更新回数)

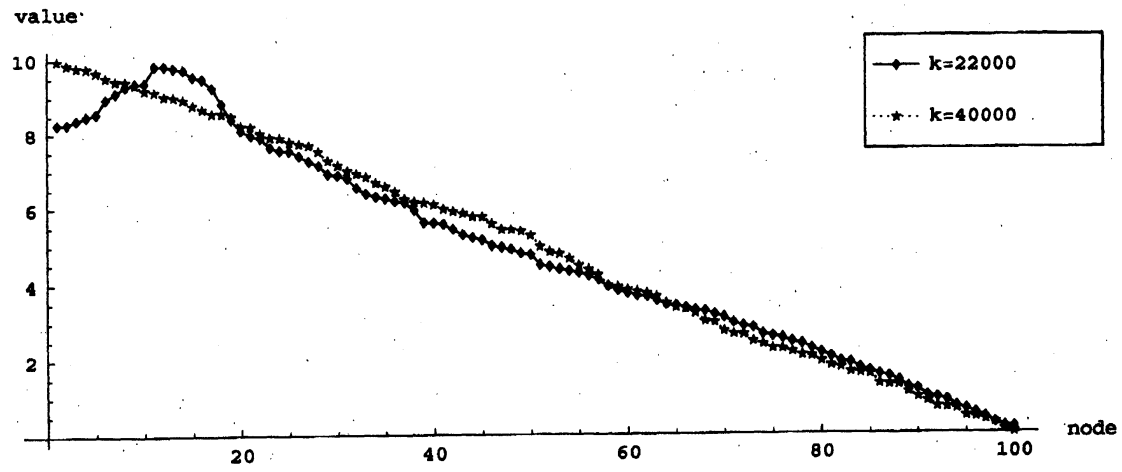

図 3:　22000 回および 40000 回更新した後のモデル関数 (k は更新回数)

また，学習プロセスとして学習率 $\alpha = \frac{1}{2}$ をもつ Learning process 1 を仮定する．

このとき，初期モデル関数および更新後のモデル関数は図 2，3 のようになる．2000 回の更新によりモデル関数が局所的に単調化されていることが観測される．また，22000 回の更新後は準凹になっているが，まだ単調になっていないことがわかる．さらに，40000 回の更新後は完全に単調の様態の到達していることがわかる（正確な観測値としては，$k \geq 20295$ のとき m_k は準凹となり，$k \geq 38276$ のとき m_k は単調減少）．□

参考文献

[1] P. Billingsley, *Probability and Measure*, John Wiley & Sons, 1979.
[2] R. M. Dudley, *Real Analysis and Probability*, Wadsworth & Brooks, 1989.
[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, Y. Kimura and I. Kaku, *Quasi-convexity and monotonicity of model function of node in self-organizing maps*, Proceedings of 8th ICIM (2006), pp. 1173–1177

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

[8] W. Takahashi, *Nonlinear Functional Analysis*, Yokohama publishers, 2000.

[9] K. Tanaka, 凸解析と最適化理論, 牧野書店, 1994.

[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.