数理解析研究所講究録
第 1611 巻 2008 年 213-220

Quasi-convexity of model function and ordering process in self-organizing maps
(自己組織化マップモデルにおける順序化と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 Kohonen [7] 型アルゴリズムにとして知られている自己組織化マップモデルにおける順序化と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，モデルの吸収状態の存在性の観点により考察する．

ここで扱う自己組織化マップは非常に実用的であり広範囲に応用例を有し，アルゴリズムも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) 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 $m : I \to V$ 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，$m_0 : I \to V$ を初期モデル関数とする．

(iii) $X \subset V$ をインプット集合とする．$x_0, x_1, x_2, \ldots \in X$ 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:

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

$$N_1(i) = \left\{ j \in I \;\middle|\; d(j,i) \le \varepsilon \right\} \quad (i \in I).$$

(b) 学習率: $0 \leq \alpha \leq 1$.
(c) 方法:

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

2. $\mathbb{R}$値ノード，1次元ノード配列モデル

(i) 有限個のノードを仮定する．$I = \{1, 2, \ldots, n\} \subset \mathbb{N}$.
(ii) ノード値の空間を$\mathbb{R}$（ユークリッドノルム）とする．$m_0 = [m_0(1), m_0(2), \cdots, m_0(n)]$ などと記すことにする．
(iii) $x_0, x_1, x_2, \ldots \in X \subset \mathbb{R}$を入力列とする．
(iv) Learning process 2，Learning process 3の一方を仮定する．

Learning process 2

(a) 学習範囲:

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

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

(b) 学習率: $0 \leq \alpha \leq 1$.
(c) 方法:

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

Learning process 3

(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 $1,2,\dots,n$ があり，そのそれぞれに対してノードの値 $m_0(1), m_0(2), \dots, m_0(n)$ 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，$m_0(1), m_0(2), \dots, m_0(n)$ 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 $x_1, x_2, x_3, \dots$ 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 $m_1, m_2, m_3, \dots$ 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．

3. モデルの吸収状態について

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 2* を仮定する．モデル関数 $m_1, m_2, m_3, \dots$ 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，このような状態を自己組織化マップモデルの吸収状態と呼ぶことにする．モデルの吸収状態として後で述べる意味での準凸性，準凹性などがある．

2 次元ノード配列の場合を考える．

(i) ノード集合 $I = I_1 \times I_2$．ただし

$$I_1 = \{1,2,\dots,N_1\}, I_2 = \{1,2,\dots,N_2\}, \quad d_I((i,j),(k,l)) = |i-k| + |j-l|$$

(ii) ノード値空間 V は内積 $\langle \cdot, \cdot \rangle$ をもつ内積空間とする．

(iii) $x_0, x_1, x_2, \dots \in X \subset V$ はインプット列とする．

(iv) Learning process

(a) 学習範囲:

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

$$(m\in M, x\in X),$$

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

(b) 学習率: $0\le\alpha\le 1$.

(c) 方法:

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

Condition 5 すべての $(i,j)\in I$ に対して

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

が成り立つ.

ここで, Condition 5 はモデルの吸収状態にならないが, 考察として以下の性質をもつことがわかる.

Theorem 2 上の *2* 次元ノード配列モデルを仮定する. 等式

$$\|m(i^*,j^*)-x\|=\min_{(i,j)\in I}\|m(i,j)-x\|.$$

を満たす (i^*,j^*) が一意に存在すると仮定する. このとき, 以下が成り立つ.

(1) *(a)* $(i,j),(i+1,j),(i+2,j)\notin N_1(i^*,j^*)$ に対して $\|m(i+1,j)-m(i,j)\|\le\|m(i+2,j)-m(i,j)\|$ ならば, すべての α,x に対して $\|m'(i+1,j)-m'(i,j)\|\le\|m'(i+2,j)-m'(i,j)\|$ が成り立つ.

(b) $(i,j),(i-1,j),(i-2,j)\notin N_1(i^*,j^*)$ に対して $\|m(i-1,j)-m(i,j)\|\le\|m(i-2,j)-m(i,j)\|$ ならば, すべての α,x に対して $\|m'(i-1,j)-m'(i,j)\|\le\|m'(i-2,j)-m'(i,j)\|$ が成り立つ.

(c) $(i,j),(i,j+1),(i,j+2)\notin N_1(i^*,j^*)$ に対して $\|m(i,j+1)-m(i,j)\|\le\|m(i,j+2)-m(i,j)\|$ ならば, すべての α,x に対して $\|m'(i,j+1)-m'(i,j)\|\le\|m'(i,j+2)-m'(i,j)\|$ が成り立つ.

(d) $(i,j),(i,j-1),(i,j-2)\notin N_1(i^*,j^*)$ に対して $\|m(i,j-1)-m(i,j)\|\le\|m(i,j-2)-m(i,j)\|$ ならば, すべての α,x に対して $\|m'(i,j-1)-m'(i,j)\|\le\|m'(i,j-2)-m'(i,j)\|$ が成り立つ.

(2) *(a)* $\|m(i^*,j^*)-m(i^*-1,j^*)\|\le\|m(i^*+1,j^*)-m(i^*-1,j^*)\|$ ならば, すべての α,x に対して $\|m'(i^*,j^*)-m'(i^*-1,j^*)\|\le\|m'(i^*+1,j^*)-m'(i^*-1,j^*)\|$ が成り立つ.

(b) $\|m(i^*,j^*)-m(i^*+1,j^*)\| \le \|m(i^*-1,j^*)-m(i^*+1,j^*)\|$ ならば，すべての α, x に対して $\|m'(i^*,j^*)-m'(i^*+1,j^*)\| \le \|m'(i^*-1,j^*)-m'(i^*+1,j^*)\|$ が成り立つ．

(c) $\|m(i^*,j^*)-m(i^*,j^*-1)\| \le \|m(i^*,j^*+1)-m(i^*,j^*-1)\|$ ならば，すべての α, x に対して $\|m'(i^*,j^*)-m'(i^*,j^*-1)\| \le \|m'(i^*,j^*+1)-m'(i^*,j^*-1)\|$ が成り立つ．

(d) $\|m(i^*,j^*)-m(i^*,j^*+1)\| \le \|m(i^*,j^*-1)-m(i^*,j^*+1)\|$ ならば，すべての α, x に対して $\|m'(i^*,j^*)-m'(i^*,j^*+1)\| \le \|m'(i^*,j^*-1)-m'(i^*,j^*+1)\|$ が成り立つ．

(3) (a) $\|m(i^*+1,j^*)-m(i^*,j^*)\| \le \|m(i^*+2,j^*)-m(i^*,j^*)\|$ ならば，すべての α, x に対して $\|m'(i^*+1,j^*)-m'(i^*,j^*)\| \le \|m'(i^*+2,j^*)-m'(i^*,j^*)\|$ が成り立つ．

(b) $\|m(i^*-1,j^*)-m(i^*,j^*)\| \le \|m(i^*-2,j^*)-m(i^*,j^*)\|$ ならば，すべての α, x に対して $\|m'(i^*-1,j^*)-m'(i^*,j^*)\| \le \|m'(i^*-2,j^*)-m'(i^*,j^*)\|$ が成り立つ．

(c) $\|m(i^*,j^*+1)-m(i^*,j^*)\| \le \|m(i^*,j^*+2)-m(i^*,j^*)\|$ ならば，すべての α, x に対して $\|m'(i^*,j^*+1)-m'(i^*,j^*)\| \le \|m'(i^*,j^*+2)-m'(i^*,j^*)\|$ が成り立つ．

(d) $\|m(i^*,j^*-1)-m(i^*,j^*)\| \le \|m(i^*,j^*-2)-m(i^*,j^*)\|$ ならば，すべての α, x に対して $\|m'(i^*,j^*-1)-m'(i^*,j^*)\| \le \|m'(i^*,j^*-2)-m'(i^*,j^*)\|$ が成り立つ．

Proof. (1),(2) は明らかである．(3) を示す．$m_0 = m(i^*,j^*), m_1 = m(i^*+1,j^*), m_2 = m(i^*+2,j^*)$ とおくとき，$\|m_1 - m_0\| \le \|m_2 - m_0\|$ が成り立つならば，次の不等式が成り立つ．

$$
\begin{aligned}
&\|m_2' - m_0'\|^2 - \|m_1' - m_0'\|^2 \\
=&\|m_2 - (1-\alpha)m_0 - \alpha x\|^2 - \|(1-\alpha)m_1 + \alpha x - (1-\alpha)m_0 - \alpha x\|^2 \\
=&\|(1-\alpha)(m_2 - m_0) + \alpha(m_2 - x)\|^2 - \|(1-\alpha)(m_1 - m_0)\|^2 \\
=&(1-\alpha)^2\|m_2 - m_0\|^2 + 2(1-\alpha)\alpha\langle m_2 - m_0, m_2 - x\rangle + \alpha^2\|m_2 - x\|^2 \\
&\quad - (1-\alpha)^2\|m_1 - m_0\|^2 \\
=&(1-\alpha)^2(\|m_2 - m_0\|^2 - \|(m_1 - m_0)\|^2) \\
&\quad + 2(1-\alpha)\alpha(\|m_2 - x\|^2 - \langle m_0 - x, m_2 - x\rangle) + \alpha^2\|m_2 - x\|^2 \\
\ge&(1-\alpha)^2(\|m_2 - m_0\|^2 - \|m_1 - m_0\|^2) \\
&\quad + 2(1-\alpha)\alpha(\|m_2 - x\|^2 - \|m_0 - x\|\|m_2 - x\|) + \alpha^2\|m_2 - x\|^2 \\
\ge&(1-\alpha)^2(\|m_2 - m_0\|^2 - \|m_1 - m_0\|^2) \\
&\quad + 2(1-\alpha)\alpha(\|m_2 - x\|^2 - \|m_2 - x\|\|m_2 - x\|) + \alpha^2\|m_2 - x\|^2 \ge 0
\end{aligned}
$$

同様に $m_0 = m(i^*,j^*), m_1 = m(i^*-1,j^*), m_2 = m(i^*-2,j^*)$，または $m_0 = m(i^*,j^*), m_1 = m(i^*,j^*+1), m_2 = m(i^*,j^*+2)$，または $m_0 = m(i^*,j^*), m_1 = m(i^*,j^*-1), m_2 = m(i^*,j^*-2)$ とおくことにより，残りの命題が成立することを示すことができる． □

ここで，I, V が一般の場合について言及する．

Theorem 3 ノード集合 I を距離 d によって距離付けされた加算集合とする．ノード値の空間 V をノルム線形空間とする．$\|\cdot\|$ を V におけるノルムとする．任意のモデル関数

m とインプット $x \in X$ に対して，$i^* \in I$ が

$$\|m(i^*) - x\| = \inf_{i \in I} \|m(i) - x\|$$

を満たすならば，m を更新したモデル関数 m' に対して

$$\|m'(i^*) - x\| = \inf_{i \in I} \|m'(i) - x\|$$

が成り立つ．

4.　モデル関数の準凸性と準凹性について

ここで，半順序集合上の準凸性，準凹性を以下のように導入することにより，自己組織化マップモデルのモデル関数に関する準凸性および準凹性について考察する．

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 4　$(Y, \leq)$ 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．

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

以下は準凸関数と準凹関数の基本的性質である．

Theorem 5　$(Y, \leq)$ 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.

Theorem 6 *Learning process 2* 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.

Theorem 6 の証明のアウトラインは [6] を参照.

Remark 1 Theorem 6 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 3 を仮定した場合の単調性保存に関する結果である.

Theorem 7 *Learning process 3* 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.

ここで, 半順序集合上の強い準凸, 強い準凹を以下のように導入する.

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.

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 8 $(Y, \leq)$ 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 9 *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．

Theorem 10 *Learning process 3* 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．

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.