

## モジュラーネットワーク SOM : 教師あり学習と教師なし学習の融合

Modular network SOM : Fusion of supervised and unsupervised learning

徳永 憲洋

Kazuhiro TOKUNAGA

九州工業大学 大学院生命体工学研究科

Kyushu Institute of Technology

古川 徹生

Tetsuo FURUKAWA

九州工業大学 大学院生命体工学研究科

Kyushu Institute of Technology

**Abstract:** A modular network SOM (mnSOM) proposed by authors is an extension and a generalization of a conventional SOM in which each nodal unit is replaced by a module such as a neural network. The mnSOM has the learning algorithm in which both the supervised and the unsupervised learning are combined. Thus, each module of the mnSOM implements the supervised learning, while, roles in the modules are allocated self-organizingly. In this paper, the architecture and the algorithm of the mnSOM are shown, moreover, the applications are presented.

### 1 はじめに

われわれ人間の脳内における学習の情報処理メカニズムは知的行動を獲得するために重要な要素であり、今まで学習メカニズムのモデルとして多くのニューラルネットワークが提案されている。ニューラルネットワークは学習の方法によって「教師あり学習」「教師なし学習」に大別できる。教師あり学習は、お手本である望ましい出力をネットワークに提示しながらネットワークを学習させる方法であり、主なものとして多層パーセプトロン (Multi-Layer Perceptron:MLP) やリカレントニューラルネットワークなどが挙げられる。一方、教師なし学習は、目標とする出力をネットワークに提示せず自己組織的に学習させる方法であり、主なものとして Kohonen の自己組織化マップ (Self-Organizing Map:SOM) が挙げられる。これらは互いに独立したモデルであり、使用の目的も違うために、用途に応じて使い分ける必要がある。

しかし、ロボットなどが人間のように知的行動を獲得するためには、教師あり学習あるいは教師なし学習のどちらかを単独で用いるだけでは限界がある。つまり、ロボットに対して人間側から一方的に望ましい行動規範、すなわち教師信号を与える教師あり学習だけでは限界があるし、お手本の無い教師なし学習だけでも学習できることが限られてくる。また使用目的に合わせて柔軟に適應できるアーキテクチャ、アルゴリズムが必要となる。

一方、われわれは一般化した SOM としてモジュラーネットワーク SOM (Modular network SOM:mnSOM) を

提案してきた [1]。mnSOM は SOM のベクトルユニットを MLP やリカレントニューラルネットワークなどのニューラルネットワークモジュールに置き換えたものである。mnSOM の特徴として主に次の二点が挙げられる。まず一点目は、mnSOM を扱うユーザは mnSOM の使用目的に応じて各モジュールを自由にデザインすることである。例えば、都市ごとの気象ダイナミクスの類似度をマップにするならば、mnSOM の各モジュールは気象変動を予測できるニューラルネットワークにデザインすればよい [1]。我々はこれまでの研究で、mnSOM のモジュールとして様々なニューラルネットワークを採用し、さらにいくつかの応用課題にも取り組んでいる [1, 2, 3]。これにより通常の SOM はベクトルデータしか扱えないが、mnSOM は広範な領域で利用することが可能である。次に二点目は、mnSOM の学習メカニズムには教師あり学習と教師なし学習が混在していることである。つまり mnSOM の各モジュールは教師あり学習を行うが、どのモジュールが学習すべきかは SOM のアルゴリズムによって mnSOM 自身が教師なしで自己組織的に判断する。気象の課題を例にすると、各モジュールはある一都市の気象変動の予測モデルを気象データを教師信号として学習によって獲得する。このとき、隣り合うモジュールは似た気象予測モデルを獲得するように自己組織的に学習が行われる。さらに、気象データの観測されていない中間都市の気象予測モデルも自己組織的に発現する。以上の特徴から、mnSOM は知的情報処理メカニズムの基盤となるアーキテクチャ、アルゴリズムになると考えられる。

本稿では, mnSOM のアーキテクチャ, アルゴリズムを説明し, さらに応用例も示す.

## 2 mnSOM のアーキテクチャ, アルゴリズム

### 2.1 アーキテクチャ

mnSOM はニューラルネットワークで構成されるモジュールが格子状に配置されたモジュラーネットワークの構造を持つ (図 1(a)). 本稿では, mnSOM の各モジュールが多層パーセプトロン (Multi-Layer Perceptron: MLP) で構成された MLP-mnSOM について説明する. MLP-mnSOM は, mnSOM における原理, アルゴリズムの理解が容易であるためである.

まず, MLP-mnSOM の問題の枠組みから説明する. 今,  $d_{in}$  次元入力,  $d_{out}$  次元出力のシステムが  $I$  個あり, 各システムは入出力関係が関数  $f_i(\cdot)$   $i = 1, \dots, I$  で記述されるとする. さらに各システムからは入出力ベクトルの集合  $D_i = \{(x_{ij}, y_{ij})\}$   $j = 1, \dots, J$  が観測されているとする. つまり, 各システムの入出力関係は関数  $y_{ij} = f_i(x_{ij})$  で記述される. 既知である情報は入出力ベクトル集合  $D_i$  だけであり, 各システムの関数  $f_i(\cdot)$  は未知である. なお関数と観測データ集合は 1 対 1 である必要はない. しかしここでは話を簡単にするために, 1 つの関数から 1 つの観測データ集合が得られたと仮定しよう. 一方, MLP-mnSOM は  $K$  個の MLP モジュールで構成されるとする. 以上の条件下で, MLP-mnSOM は以下のタスクを同時に実行する. (1)  $I$  個の入出力ベクトル集合  $D_i$  を  $K$  個のモジュールを使って学習し, 関数  $f_i(\cdot)$  を獲得・再現する.  $K$  個のモジュールのうち  $I$  個が各関数に対する勝者となる. 関数の獲得は教師あり学習である. 入出力ベクトル集合  $D_i$   $i = 1, \dots, I$  から関数  $f_i(\cdot)$  を推定する. (2) 勝者以外のモジュールは SOM における協調学習によって中間的な関数を獲得し, 関数間を補間しながらマップを生成する. 関数間の補間および関数の各モジュールへの割り当ては教師なし学習で行われる.

(2) において, 関数間の類似度は関数空間における距離を元にマップは作られる. つまり関数  $g(x)$  と関数  $f(x)$  の間の類似度は次式に示す関数空間における距離  $L^2(g, f)$  で定義される.

$$L^2(g, f) = \int \|g(x) - f(x)\|^2 p(x) dx. \quad (1)$$

ここで  $p(x)$  は  $x$  の確率密度関数である. なお, 話を簡単にするため, 本稿では  $p(x)$  は全関数において等しいと仮定する.

### 2.2 アルゴリズム

次に MLP-mnSOM のアルゴリズムを示す. mnSOM の学習アルゴリズムは通常の SOM におけるアルゴリズムと等価である. ここでは mnSOM の学習アルゴリズムを評価, 競合, 協調, 適応プロセスの 4 つのプロセスに分けて説明する.

#### 評価プロセス

まず評価プロセスでは各 MLP モジュールの実現する関数  $g^k(\cdot)$  とマップする関数  $f_i(\cdot)$  との距離  $L^2(g^k, f_i)$  を (1) で評価する. ただし, 実際には関数  $f_i(\cdot)$  は未知なので, (1) で直接評価することはできない. そこで観測されている  $D_i$  を用いて近似的に評価を行う. つまり  $D_i$  におけるひとつの入出力ベクトル  $x_{ij}$  を全モジュールに与え, 各モジュールの出力  $\tilde{y}_{ij}^k$  を求め, 教師信号  $y_{ij}$  との誤差  $e_{ij}^k$  を計算する. この過程を全ベクトルに対して行い,  $e_{ij}^k$  のアンサンブル平均  $E_i^k$  で (1) を近似的に評価する. すなわち

$$L^2(g^k, f_i) \simeq E_i^k = \frac{1}{J} \sum_{j=1}^J \|\tilde{y}_{ij}^k - y_{ij}\|^2 \quad (2)$$

#### 競合プロセス

競合プロセスでは, 全クラスの関数に対してそれぞれ勝者モジュールを決める. つまり次式に示すように  $i$ -th 関数に対して最も評価誤差  $E_i^k$  が小さかったモジュールを勝者モジュール  $k_i^*$  とする.

$$k_i^* = \arg \min_k E_i^k \quad (3)$$

#### 協調プロセス

協調プロセスでは, 各モジュールの学習量を求める.  $i$ -th 関数に対する  $k$ -th モジュールの学習量  $\psi_i^k$  は以下の式で求められる.

$$\psi_i^k = \frac{\phi_i^k(l(k, k_i^*), T)}{S(T)} \quad (4)$$

$$\phi_i^k(l; T) = \exp\left[-\frac{l^2}{2\sigma^2(T)}\right]. \quad (5)$$

$l(k, k_i^*)$  は  $k$ -th モジュールと  $i$ -th オブジェクトに対する勝者モジュールとのマップ上における距離 (mnSOM モジュールの格子上における距離) を示す.  $\phi_i^k(l; T)$  と  $\sigma^2(T)$  はそれぞれ学習ステップ  $T$  における近傍関数と近傍関数の広さを示す. 一般に,  $\sigma^2(T)$  は学習ステップ  $T$  に対して単調減少する関数を用いる. また  $S(T)$  は学習量を正規化するために必要とする. 一括学習の場合は,  $S(T)$  は次式のように, 近傍関数の総和として与えら

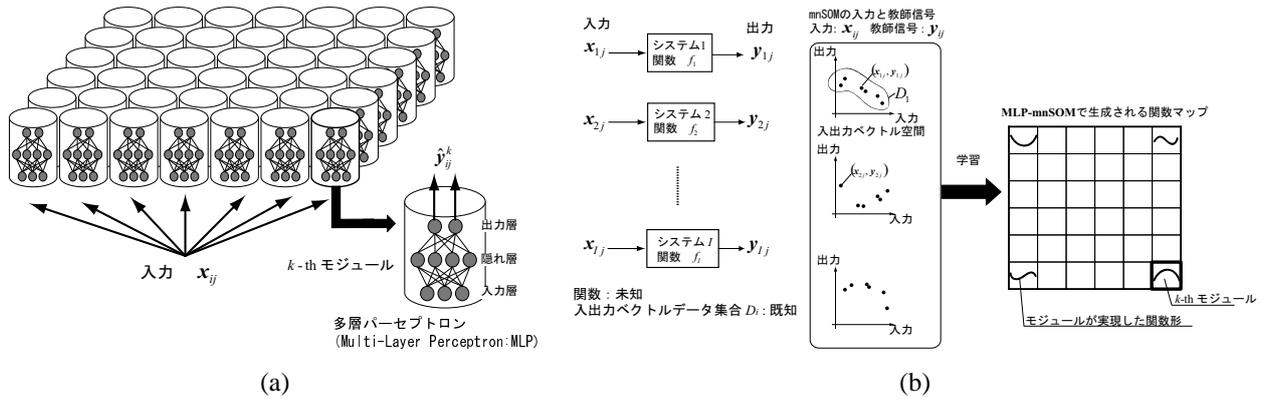


図1 (a)MLP-mnSOM のアーキテクチャ。(b)MLP-mnSOM の枠組み。

れる。

$$S(T) = \sum_{i=1}^I \phi_i^k(l; T). \quad (6)$$

すなわち、各モジュールの全入出力ベクトル集合に対する学習量  $\psi_i^k$  の和が1になるように正規化されている。一方、逐次学習の場合は、 $S(T)$  は次式のように時間平均で近似する。

$$S(T) = (1 - \epsilon)S(T - 1) + \phi^k(l; T) \quad (7)$$

$\epsilon$  は  $0 < \epsilon \leq 1$  の定数である。

#### 適応プロセス

適応プロセスでは全モジュールを次式のように誤差逆伝播法によって学習させる。

$$\Delta w^k = -\eta \sum_{i=1}^I \psi_i^k \frac{\partial E_i^k}{\partial w^k} = -\eta \frac{\partial E^k}{\partial w^k}. \quad (8)$$

ここで、 $w^k$  は  $k$ -th モジュールの結合加重ベクトルである。また  $E^k = \sum_i \psi_i^k E_i^k$  である。つまり  $E^k$  は各モジュールの誤差エネルギーであり、この誤差エネルギーが大域的な最小値となる部分が各モジュールの実現する関数  $g^k$  となる。このとき  $g^k$  は  $\{f_i\}$  の重みつき内分点となる。つまり、 $g^k = \sum_{i=1}^I \psi_i^k f_i(x)$  である。(8) による  $w^k$  の更新は  $\psi_i^k$  を固定したまま十分な回数を繰り返す。

### 3 mnSOM のアプリケーション

本稿では mnSOM の応用例として、MLP-mnSOM で生成した関数マップと気象マップについて示す。

#### 関数マップ

本シミュレーションでは6個の1入力1出力関数システムから観測された入出力データ集合のマップ生成を試みた。各システムから観測された入出力データ集合を図2(a)に示す。一方、mnSOMのネットワークは100個

のモジュールを  $10 \times 10$  の正方格子状に配置したものであり、各モジュールは1入力1出力の3層MLPである。

実験結果を図2(b)に示す。各マス目は1つのモジュールに対応し、各マス目内に描かれた曲線は対応するモジュールが学習によって獲得した関数形である。また本シミュレーションで与えた6個の関数に対応する勝者モジュールを太枠で示している。各モジュールは様々な形の関数を実現した。さらに隣り合うモジュールは似た関数形を実現し、対角に位置するモジュールは性質の異なる関数を実現した。また6個の関数から観測された入出力データ集合だけを教師データとして提示しているにもかかわらず、mnSOMは6個の関数間を内挿する関数形も獲得した。

#### 気象マップ

次に mnSOM による気象ダイナミクスのマッピングについて示す。本シミュレーションでは、九州各地20都市(図3(a))の2000年100日分の気圧、気温、湿度、日照時間を要素とした気象データを用いた。図3(a)中のゴシック体の10都市を学習クラスとして用い、斜体文字の10都市をテストクラスとして用いた。一方、mnSOMの各モジュールは一昨日、昨日、今日の気象を入力とし、明日の気象を予測するように学習が行われる。

図3(b)に結果を示す。都市名の書いてある格子は各都市に対するの勝者モジュールを示す。結果から、気象データを学習させただけでもかかわらず、九州都市の地理的位置のトポロジーを保存したマップが生成された。さらに mnSOM にはゴシック体の都市の気象データしか学習させていないにもかかわらず、mnSOMは斜体文字の都市の気象予測モデルも獲得した。つまり、mnSOMは“九州地区の気象マップ”の生成に成功したといえる。

以上のように mnSOM は教師あり学習と教師なし学

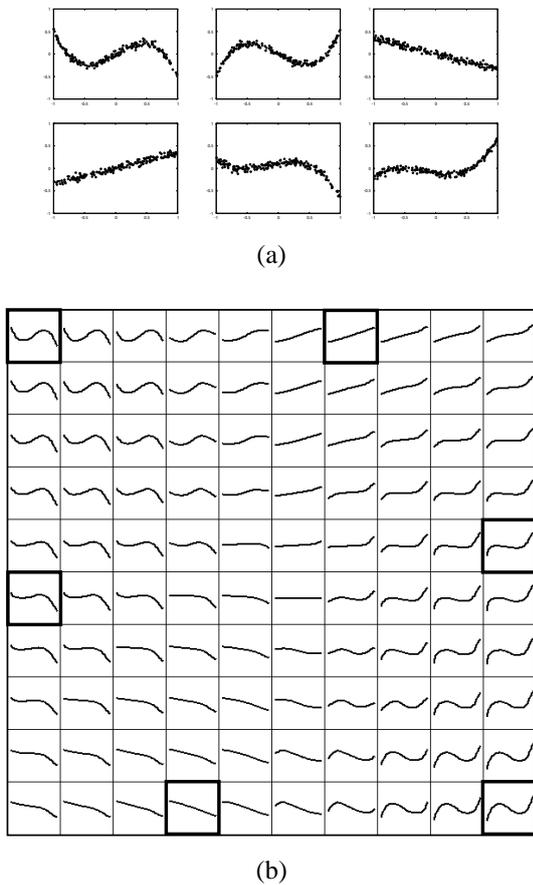


図 2 MLP-mnSOM による関数集合マッピング . (a) 入出力データ集合例 . (b) 関数マップ .

習を兼ねた学習アルゴリズムを持っているおり，さらにモジュールを自由にデザインできることから広範囲の領域で工学的な応用が可能なので，知的情報処理メカニズムの基盤となるアルゴリズムになるだろうと期待できる .

謝辞

本研究は九州工業大学 21 世紀 COE プログラムの一環として行われた . また，本研究は財団法人大川情報通信基金の援助を受けた .

連絡先

徳永 憲洋

E-mail: tokunaga@brain.kyutech.ac.jp

古川 徹生

E-mail: furukawa@brain.kyutech.ac.jp

参考文献

- [1] 徳永, 肝付, 古川, 安井 : “関数空間型 SOM”, 日本神経回路学会誌, **12**, 1, pp. 39-51 (2005).
- [2] T. Furukawa, K. Tokunaga, S. Kaneko and S. Ya-

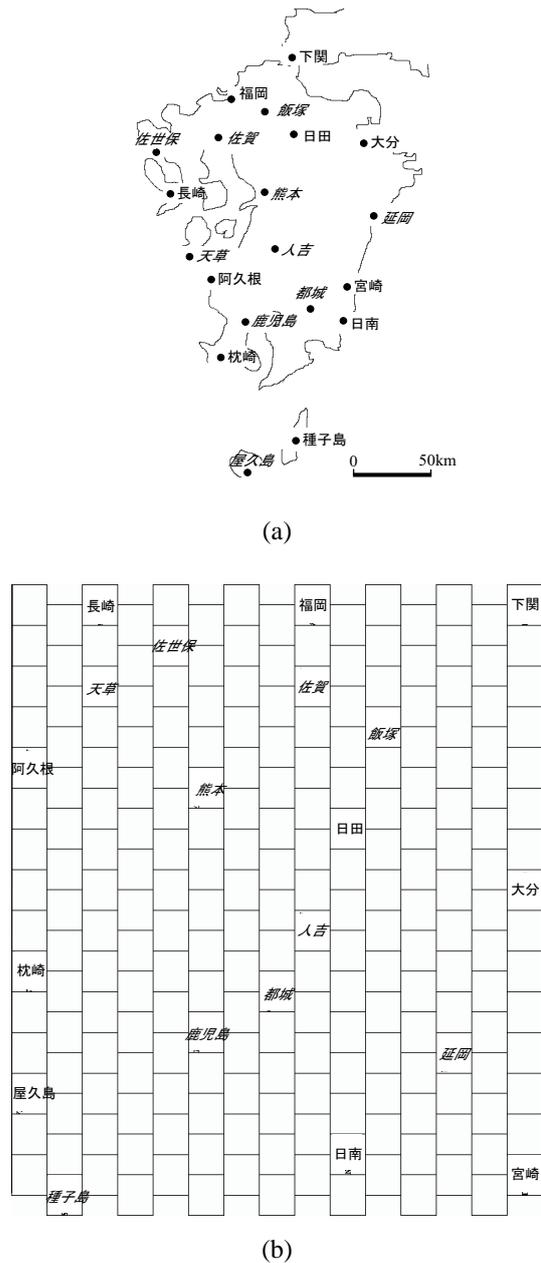


図 3 mnSOM による気象マップ生成 . (a) 九州地図 . (b) 九州気象マップ .

sui: “Generalized self-organizing maps (mnSOM) for dealing with dynamical systems”, Proc. of International Symposium on Nonlinear Theory and its Applications, pp. 231-234 (2004).

- [3] T. Minatohara and T. Furukawa: “Self-organizing adaptive controllers: Application to the inverted pendulum”, Proc. of Workshop on Self-Organizing Maps (2005).