

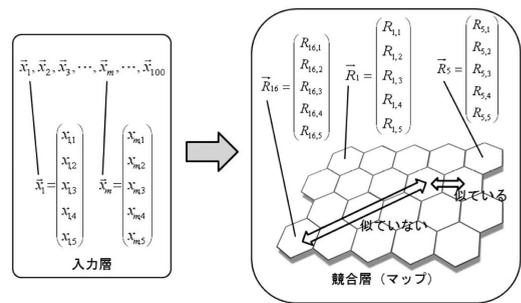
## 自己組織化マップ

本稿は、自己組織化マップ (self-organizing map 以降 SOM とよぶ) について解説する。SOM (Kohonen 1990) は、神経回路を模倣した数理モデルである人工ニューラルネットワークの一種であるが、多次元 (多変数) データのパターン認識や、抽出したパターンを基にしたデータ分類などを行うことができ、多変量解析手法の一つともみなせる。また、SOM はデータ圧縮の手段として優れた性質を持ち、多次元データの位相関係を保持しつつ、低次元 (通常は 2 次元平面) 上に表示できる。

基本的な SOM は入力層と競合層の 2 層構造から成り (第 1 図)、教師なし学習によって入力データのパターン抽出を行う。なお、SOM は学習アルゴリズムや構造によって、バッチ学習型 SOM や階層型 SOM などもあるが、ここでは基本的な SOM の構造について述べる。入力層は入力ベクトルからなり、入力ベクトルはパターン抽出対象の多次元データ (例えば過去 50 年間の太平洋の海面水温データ) に相当する。競合層は二次元平面に配置されたユニット (ノードまたはニューロンとも呼ばれる) から構成されており、各ユニットは参照ベクトルという、入力ベクトルと同じ次元を持つベクトルと結び付けられている。参照ベクトルは入力データから抽出されたパターンを表しており、マップ上で近くに位置するユニットの参照ベクトルは互いに似ているという性質がある。また、各入力データは、各々と最もよく似た参照ベクトルの位置するユニットへと分類される。

第 1 図のように、競合層で縦方向に 4 個、横方向に 5 個の計 20 個のユニットが配置されている場合、入力データは 20 個のパターンに集約されることになる。ユニットの数 (マップサイズ) を大きくすると一般に、パターンのより細かな違いまで抽出される。なお SOM の具体的なアルゴリズム等については、伊藤・見延 (2010) 等を参照されたい。

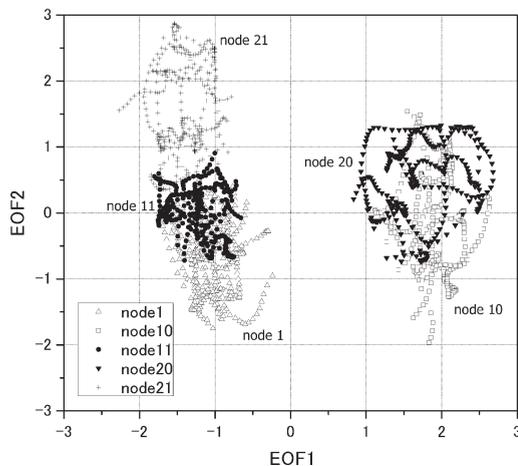
自己組織化マップの気象学分野への応用は盛んに行



第 1 図 基本 SOM の概念図。  $R_i$  ( $i = 1, 2, \dots, 20$ ) は参照ベクトル、  $x_j$  ( $j = 1, 2, \dots, 100$ ) は入力ベクトルを表す。この例では、入力ベクトルの次元 (要素数) は 5 であり、計 100 個の入力ベクトルがある。競合層 (マップ) は縦 4 × 横 5 = 20 個のユニットで構成されている。

われている。ここで多変量解析手法として、しばしば用いられている経験的直交関数展開 (EOF) と、SOM との相違について幾つか述べる。まず両手法のパターン抽出の原理であるが、EOF の原理はデータの分散を最も効率よく表現する合成変数を作成することにあり、それ故、合成変数の分散が最大となる空間モード (軸) が抽出される。また、EOF で抽出される空間モードは互いに直交するという制約をもつ。一方 SOM で抽出されるパターンには直交性の制約はなく、大まかに言うと入力データの反復学習に基づいてパターンの抽出が行われる。また、SOM は、入力データのノイズを除去したパターンを抽出する能力がある事も指摘されている (Reusch *et al.* 2005)。

以下に、西太平洋～インド洋地域の外向長波放射量 (OLR) 偏差場に両手法を適用した例を挙げる。エルニーニョ・南方振動 (ENSO) 現象やマッデン・ジュリアン振動 (MJO) の研究に関連して、OLR に EOF 解析を用いている論文が多い。元々、OLR は背の高い対流雲や上層雲の有無の判別が可能なので、熱帯対流活動の指標として多用されてきたが、雲がなければ地表面あるいは海表面からの赤外放射に大きく左



第2図 EOF1とEOF2の位相空間に投影された、ENSO現象と密接に関連する五つのノード。EOF1の正(負)の領域に位置しているノードがエルニーニョ(ラニーニャ)型である。

右されるため、降水量データ程ではないにしても、その標本集団は正規分布からずれている場合が多い。そのため、EOF解析を適用するのが妥当かどうか本来ならば吟味する必要がある。一方、SOMはその点は特に気にする必要はない。

第2図はEOF1とEOF2のスコアで構成される2次元位相空間に、ENSOと密接に関連するSOMのノード五つを投影させたものである(Sakai *et al.* 2010)。SOMのマップサイズは $5 \times 5$ を採用した(各ノードの空間パターンは上述の原著論文を参照されたい)。ここで、EOF1はENSOを説明する主要な変動モードで、正(負)のスコアはエルニーニョ(ラニーニャ)の位相に対応する。投影結果から、エルニーニョ現象時には主に二つのノード、ラニーニャ現象時には主に三つのノードの出現頻度が高く、特に注目すべきは、ラニーニャ現象時のノード1、ノード21はEOF2のスコアも大きくなっている点である。通常のEOF解析では直交性という制約条件から、EOF2

はEOF1とは全く無相関であるため、EOF2はENSOとは独立した別の現象を説明していると誤って解釈される可能性があるが、SOM解析からはEOF2にもENSOシグナルが混在していると判断できる。また、エルニーニョとラニーニャの位相において、西太平洋～インド洋の熱帯対流活動は強い非対称性をもっていることも、この図から読み取れる。

このように、SOMは気候システムに内在する非線形現象などへの適用可能性が高いという大きな長所を持っている。ただしSOMを適用する際は、データのパターン学習の際に用いる学習率係数や近傍関数、マップサイズ等のパラメータ設定が必要となるが、EOFはその様なパラメータ設定は必要ない。

SOM、EOFには手法としての原理に基づいてそれぞれの長所・短所があるが、一方では、EOF解析等の他の解析手法とSOMを併用することで、互いの長所を生かしたシナジー効果が期待でき、今後新しい適用事例が益々増加していくものと思われる。

#### 参考文献

- 伊藤久徳, 見延庄士郎, 2010: 気象学と海洋物理学で用いられるデータ解析法. 気象研究ノート, (221), 184-191.
- Kohonen, T., 1990: The Self-Organizing Map. Proc. IEEE, 78, 1464-1480.
- Reusch, D. B., R. B. Alley and B. C. Hewitson, 2005: Relative performance of self-organizing maps and principal component analysis in pattern extraction from synthetic climatological data. Polar Geogr. 29, 188-212.
- Sakai, K., R. Kawamura and Y. Iseri, 2010: ENSO-induced tropical convection variability over the Indian and the western Pacific oceans during the northern winter as revealed by a self-organizing map. J. Geophys. Res., 115, D19125, doi:10.1029/2010JD014415.

(富山大学大学院理工学研究部 川村隆一)  
(東京工業大学大学院情報理工学研究科 井芹慶彦)