情報処理学会論文誌　数理モデル化と応用　Vol.5　No.3　1–13 (Sep. 2012)

サンプルの所属度に応じた可変自己組織化マップ

多賀谷 侑史[1,a)]　安藤 晋[2]　関 庸一[2,b)]

受付日 2010年10月23日，再受付日 2012年1月31日，
採録日 2012年3月26日

概要：自己組織化マップ（SOM）は，2 次元格子平面上のマップ中のノードと，その特徴量を表す参照ベクトルを用いてサンプル構造の可視化と次元縮約を行う手法である．従来の SOM ではマップ形状があらかじめ指定されるため，サンプルが持つ位相構造がその形状に折り畳まれる．このときマップ上での参照ベクトルの連続性が損なわれる，あるいは，少数のサンプルしか適合しない参照ベクトルが生じるといった問題が起こりうる．本研究では十分に大きな格子平面の中で採用するノードの位置を可変とする可変自己組織化マップの算法を提案し，サンプルの位相構造を表現するマップを生成することを目指す．提案手法では採用するノードをその参照ベクトルとサンプルの適合性の累計に基づいて更新することでマップの形状が決定される．人工・実データを使った評価実験により，提案手法が従来の SOM よりも参照ベクトルの連続性と適合性に関する指標を向上させることを示した．

キーワード：自己組織化マップ，次元縮約，クラスタリング

Flexible Self-organizing Maps Using Degree of Membership

Yuji Tagaya[1,a)]　Shin Ando[2]　Yoichi Seki[2,b)]

Received: October 23, 2010, Revised: January 31, 2012,
Accepted: March 26, 2012

Abstract: Self-Organizing Maps (SOM) is a method of visualization and dimensionality reduction using a map consisting of nodes on a two-dimensional lattice space and their reference vectors. In conventional SOM, the shape of the map is pre-defined, and the topological structure of the samples is folded into the given shape. The folding can cause problems such as discontinuity of the reference vectors among the neighboring nodes of map or the occurrence of reference vectors with few fit samples. We propose a Flexible SOM algorithm, in which the location of the nodes are flexible within a sufficiently large lattice space to create a map which naturally represents the topological structure of samples. The shape of the map is determined by iteratively updating the location of the nodes with regards to the cumulative membership of the samples assigned to each reference vectors. We present empirical evaluations using artificial and real-world datasets which show that the proposed method improves the fitness and the continuity of the reference vectors from the conventional SOM.

Keywords: self-organizing map, dimensionality reduction, clustering

1. はじめに

SOM（Self-Organizing Maps，自己組織化マップ）[1] はサンプル間の大域的構造よりも局所的構造を重視して接続することで次元縮約する方法である．一般に有界な二次元格子が次元を縮約する空間として用いられる．その各格子点（ノード）に特徴量空間の代表点（参照ベクトル）を対応させ，各サンプルを特徴量空間内で最近隣である代表点と対応する格子点（最整合ノード）に対応づけることで，格子空間への次元縮約が実現される．SOM は高次元特徴量空間中のサンプル分布を分節化し，二次元マップとして可視化する方法となる．結果が格子上に分節化されたサン

1　サンデンシステムエンジニアリング
Sanden System Engineering Corporation, Isezaki, Gunma, 372–0801, Japan
2　群馬大学大学院工学研究科情報工学専攻
Gunma University, Kiryu, Gunma 376–8515, Japan
a)　tagaya@dml.cs.gunma-u.ac.jp
b)　seki@cs.gunma-u.ac.jp

プルセットとなるため，離散的な取扱いが可能となり，多群の特徴量の関係をモデル化する基盤をあたえる方法としても応用上有用な結果を与えている [2], [3], [4].

しかし，古典的 SOM では事前に指定する二次元格子の領域に全サンプルが対応づけられるため，サンプル分布の持つ自然な位相構造がその領域形状に合うように折り畳まれ，詰め込まれる．この場合大きく異なる参照ベクトルが隣接して配置され，本来異なるクラスタに属すべきサンプル群が隣接することも生ずる．この結果，このような部分では，“マップ上で隣接するノードは類似した参照ベクトルを持つ” という SOM のマップの重要な性質（以下では，**マップ連続性**と呼ぶ）が成立しない．特徴量の関係をマップ上での位相関係を利用してモデル化しようとする場合 [5] などには，マップ不連続性が障害となる．

一方，SOM の更新反復条件を短く設定し，アンニーリングを急冷するようにすれば，前述の折り畳まれた部分でもマップ連続性が成立するマップが得られる場合もある．しかし，このような部分では隣接ノードの中間的な特徴量を表し，適合するサンプルの少ない参照ベクトルを持つノードが生じる．このようなサンプル密度の低い参照ベクトルを採用することは，特徴量空間中のサンプル分布を表すうえで効率的でない．

本論文では上述の問題を解決するため，利用ノード数に対して十分な広さを持つ格子空間中で，任意形状のノード集合の利用を許容する可変自己組織化マップ（FlexibleSOM, FSOM）を提案する．利用ノード数は事前に指定し，格子空間中の利用するノードの位置は反復算法により更新する．ノードの更新手続きは，隣接ノードと参照ベクトルが類似しないノードが廃止され，サンプル密度の高いノードで置換するよう定義する．提案手法の貢献について定量的な議論を行うため，本研究では連続性と適合性を評価する指標を導入する．我々は数値実験により提案手法がこれらの指標を従来の SOM よりも改善し，より直観的に理解しやすいマップを生成することを示す．

以下では，2 章で従来の類似手法を説明する．3 章では Kohonen [1] による SOM の算法とを説明し，SOM の問題点について数値例で示す．また，本研究で用いる評価指標である不適合度とマップ不連続度を定義する．4 章では提案法を与え，そのパラメータ設定について 5 章で実験評価する．6 章で実データへの適用事例を示す．

2. 従来の研究

高次元の特徴量を低次元に次元縮約する古典的な方法としては，主成分分析（PCA）[6] や多次元尺度構成法（MDS）[7] などがある．特徴量空間の直交変換を行う PCA や，サンプル間距離を再現する低次元空間を構成する古典的 MDS では，全サンプルを一括して，大域的距離関係を保存した次元縮約が行われる．そのため，ノイズを除いたサンプル分布が位相的には低次元であっても，それが線形部分空間に収まらない場合には，位相的な次元数までの次元縮約ができなかった．

これに対して，LLE [8] や Isomap [9] は特徴量空間の近傍ごとにその構造を抽出することにより，積極的に大域的構造を捨て，本質的な次元数の空間に縮約することを目指している．これらの方法は，局所的な位相空間を接続することで，データに本質的な次元数の多様体を構成する方法を提供している．しかし，データを分節化する機能はない．

一方，SOM と同様のニューラルネットワークモデルを用い分節化を行う Growing Neural Gas などの方法も提案されているが [10], [11]，これらは，高次元の特徴量空間中のノードの隣接グラフとして結果が表現されるため，高次元特徴量を持つサンプルセットを，人間が理解しやすく可視化することには向いていない．

二次元格子上での表現を与える方法として，追加学習を行った場合でも位相を保持するため，サンプル分布に応じてマップ上にノードを追加する SOM の拡張 [12] も提案されているが，用意した初期格子のノードを破棄する手続きを持たないため，サンプル分布の代表点として必要性の低いノードが残るという問題がある．なお，マップ上で，位相的に離れた参照ベクトルを識別し，クラスタ関係を理解する手法として，U-matrix [13], [14] があるが，クラスタの境界としてノードを用いるため，サンプル分布を代表する効率を落とすという問題は残る．

3. 標準 SOM の問題点と評価基準

本論文では Kohonen [1] による Incremental SOM を標準 SOM と呼ぶ．標準 SOM はサンプル数 × 特徴量次元数の行列を入力とし，低次元格子上のあらかじめ指定された格子点（これを以下ではノードと呼ぶ）にサンプルを分類する．以下ではノード数を K，特徴量次元数を p，サンプル数を N と表す．格子としては有界な四角格子または六角格子を通常用い，格子空間のノード k の座標を $\mathbf{r}_k$ $(k = 1, 2, \ldots, K)$ とする．一方，p 次元特徴量空間内のサンプルを $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})^t$ $(i = 1, \ldots, N)$ とする．また特徴量空間中でノード k と対応づけられる代表点を参照ベクトルと呼び，$\mathbf{m}_k = (m_{k1}, m_{k2}, \ldots, m_{kp})^t$ で表す．これが SOM の出力となる．なお，格子空間上のノルムを $||\mathbf{r}||$ とし，特徴量空間上の距離を $\mathrm{d}(\mathbf{x}, \mathbf{x}')$ と表す．

3.1 標準 SOM の算法

標準 SOM はランダムに選んだサンプルの最整合ノードを決定し，それに近いノードの参照ベクトルを更新する，という手順を繰り返す k-means 法に似た算法となる [15]．標準 SOM の算法を**図 1** に示す．本研究では，最整合ノードが c である場合のノード k の学習率関数 $h_{ck}(t)$ として，次式のガウス関数を用いる．ここで，$t = 0, 1, ..., T-1$ は

SOM($\mathbf{X}$, $\mathbf{L}$, $\mathbf{M}^{(0)}$, T, $\{\alpha, \sigma, \sigma_0\}$)
1　for $t = 0$ to $T-1$
2　　$i =$ 一様乱数 on $\{1, \ldots, N\}$;
3　　$\mathbf{x}^{(t)} = \mathbf{x}_i$;
4　　$c = \arg\min_{k \in \mathbf{L}} \mathrm{d}(\mathbf{x}^{(t)}, \mathbf{m}_k^{(t)})$;
5　　for $k = 1$ to K
6　　　$\mathbf{m}_k^{(t+1)} = \mathbf{m}_k^{(t)} + h_{ck}(t)(\mathbf{x}^{(t)} - \mathbf{m}_k^{(t)})$;
7　return $\mathbf{M}^{(T)}$;

- $\mathbf{X}$; 特徴量データ ($\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_n)^t$)
- $\mathbf{L}$; 最整合ノード選択対象として利用するノード集合
- $\mathbf{M}^{(t)} = (\mathbf{m}_1^{(t)}, \ldots, \mathbf{m}_K^{(t)})$; 参照ベクトル
- T ; 反復回数
- $h_{ck}(t)$; 最整合ノードが c である場合のノード k での学習率関数 ($0 < h_{ck}(t) < 1$)，t の単調減少関数．本論文では式 (1).

図 1　標準 SOM 算法
Fig. 1　Standard algorithm of SOM.

参照ベクトル更新の反復を表す．

$$h_{ck}(t) = \alpha^{(t)} \cdot \exp\left(-||\mathbf{r}_c - \mathbf{r}_k||^2 / 2\sigma^{(t)2}\right) \qquad (1)$$

ただし，近傍半径 $\sigma^{(t)} > 0$ と学習率係数 $\alpha^{(t)} > 0$ は t の単調減少関数である．本研究では，これらを次式の単調減少等差数列とする．

$$\sigma^{(t)} = \sigma_0 + (\sigma - \sigma_0)\left(\frac{T-t}{T}\right) \qquad (2)$$

$$\alpha^{(t)} = \alpha\left(\frac{T-t}{T}\right) \qquad (3)$$

なお，σ_0 は十分小さな正の実数である．これにより学習率関数の指数の発散を避けている．

以下では，標準 SOM の呼び出しを SOM ($\mathbf{X}$, $\mathbf{L}$, $\mathbf{M}$, T, $\{\alpha, \sigma, \sigma_0\}$) で表す．参照ベクトル初期値 $\mathbf{M}^{(0)}$ はランダムサンプルした特徴量で与える．

3.2　標準 SOM によるマップの課題

標準 SOM で 2 次元のマップを構成する際に，特徴量のサンプル分布が，マップより 1 次元低い 1 次元の位相構造を持つ場合に，サンプル分布の折り畳みがどのように行われるかを示すため，簡潔な数値例を与える．用いるサンプルセットを**図 2** に示す．これは $[0, \pi]$ 上の一様乱数に従う θ から得られた半円周上の一様乱数座標 $(\sin\theta, \cos\theta)^t$ に，2 次元正規誤差 $\mathrm{N}(\mathbf{0}, (0.1)^2\mathbf{I}_2)$ を付加したものである ($N = 10{,}000$)．ただし，$\mathbf{0} = (0,0)^t, \mathbf{I}_2$ を 2 次の単位行列とする．

これに対し，標準 SOM を 5×5 の正方格子で適用した結果を**図 3** に示す．このマップ中央から下にかけて，サンプルがほとんど所属しないノードが 3 点存在している．特徴量空間で見ると（**図 4**），この 3 点は円周の内側の 3 点として表れ，サンプル分布（図 2）の密度がない位置に存在することから，データの代表点として不適当な参照ベクトルであることが分かる．

この結果でのマップ連続性を確認するため，すべてのノード組合せについて，参照ベクトル間の距離に対しマッ

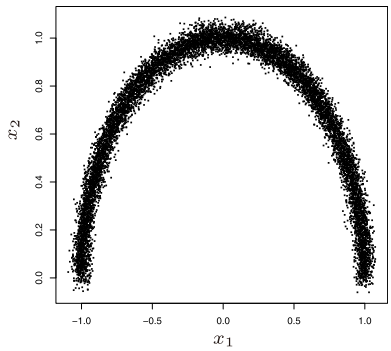
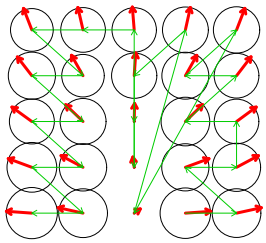


図 2　1 次元位相構造を持つサンプル例
Fig. 2　Example of one-dimensional topology.



太い矢印は参照ベクトルであり，細い矢印は参照ベクトルを極座標表示したときの角 θ の大きさ順にノードを結んでいる．これが元データの 1 次元の位相構造のマップ上での表現となる．円の大きさは，ノードに配分されたサンプル数を表している．$\alpha = 0.02$，$\sigma = 1$，$T = 10^5$，$\sigma_0 = 0.01$，$K = 25$.

図 3　標準 SOM によるマップ例
Fig. 3　Example of standard SOM map.

プ上での距離を求めた結果を図 5 に示す．中央の縦線は，特徴量空間での距離の平均値（0.8997）を示す．マップ上での距離が 1 の水平線上の点は，隣接しているノード対を表すが，このマップでは特徴量空間での距離が平均値を超えていても，マップ上で隣接していているノード対があることが分かる．つまり，マップ連続性が成立していない部分がマップにあることが分かる．

3.3　マップ評価基準

前節の議論に基づき，標準 SOM の問題およびその解決方法について定量的に評価するため本研究では以下で定義する指標を用いる．まず，サンプルの特徴量ベクトルを最整合ノードの参照ベクトルで代表させた場合の，文節化による残差の合計を不適合度とし，定義を下式で与える．

$$D^2 = \sum_{i=1}^{N} \min_k \mathrm{d}(\mathbf{x}_i, \mathbf{m}_k) \tag{4}$$

不適合度が小さいほど，得られた参照ベクトル群がサンプル分布の代表点として代表性が高いといえる．次に，マップ上で隣接するノード対が特徴量空間上でどの程度離れているかを評価するため，マップ不連続度を下式で定義する．

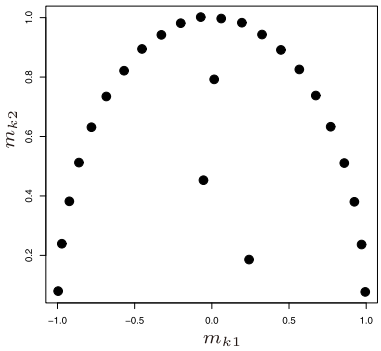
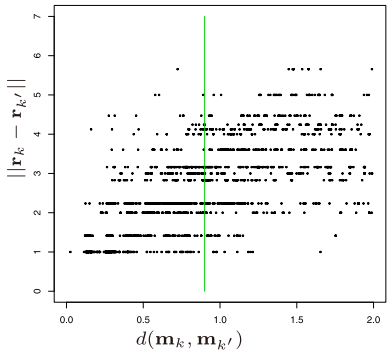


図 3 の参照ベクトル $\{(m_{k1}, m_{k2}), k = 1, ..., 25\}$ を図 2 の特徴量空間にプロットした結果．

図 4　標準 SOM の参照ベクトル

Fig. 4　Reference vectors of standard SOM.



$\alpha = 0.02, \sigma = 1, T = 10^5, \sigma_0 = 0.01, K = 25.$ 乱数を変えた 5 回の実験結果での $\binom{25}{2} \times 5 = 1500$ 個の距離対を示す．

図 5　特徴量空間・マップ上での距離の関係（標準 SOM）

Fig. 5　Distances in feature space and over SOM map.

$$C = \max_{(k,k') \in \mathbf{S}} \mathrm{d}(\mathbf{m}_k, \mathbf{m}_{k'}) \tag{5}$$

なお，$\mathbf{S} = \{(k, k') \in \mathbf{L} \times \mathbf{L} \,|\, k \neq k', ||\mathbf{r}_k - \mathbf{r}_{k'}|| = 1\}$ とし，ここではマップ上で隣接するノード間の距離を 1 とする．マップ不連続度が小さいほど，マップ連続性が保たれたマップであるといえる．

4.　可変自己組織化マップ

4.1　マップの拡張

提案手法では，十分広い格子空間から，位相構造を表すのに適当なノードを選んで用いることができるように標準 SOM を拡張する．この際，利用するノード数は事前に指定する．これを K で表す．この利用ノード集合は次節で述べる算法で反復更新し収束させるが，更新の際，新たな利用ノードの候補とするのは現時点での利用ノード集合の近傍のみとする．これにより，格子空間の全ノードを候補とすることを避ける．この際に用いる格子空間でのノード集合に関する概念を図 6 に示す．まず，ノード k とそれに隣接するノードの集合を k の近傍 $\mathbf{N}_k$ とする．$\mathbf{L}$ を利用ノード集合の暫定解とするとき，$\mathbf{L}$ に含まれる利用ノードの近傍の合併 $\mathbf{U} = \cup_{k \in \mathbf{L}} \mathbf{N}_k$ が，次回の更新における利用ノードの候補集合となる．なお，利用ノード集合 $\mathbf{L}$ の初期値は適当な連結ノード集合とする．

4.2　FSOM 算法の概要

提案する FSOM では，各サンプルがマップの各ノードに所属する程度を定義し，所属度と呼ぶ．FSOM ではこの所属度を各ノードごとに集計し，その累計が大きくなるノードを利用ノードとして新たに採用し，それの小さいノードの利用をやめることで，格子空間の中から利用ノードを選択する．利用ノード集合が変更されると，サンプルの最整合ノードに変更されるものが生じ，その影響で参照ベクトルの変更も必要となるので標準 SOM を実施し，参照ベクトルを更新する．以上を反復して利用ノードを収束させる．

4.3　所属度の定義

サンプル i のノード k に対する所属度を Membership(i, k) とし，サンプルごとの合計が 1 となる非負数とする．所属

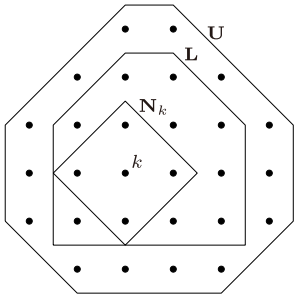


図 6　ノード集合の概念

Fig. 6　Concept of node set.

度のノード k における累計 $\sum_{i=1}^{N}\mathrm{Membership}(i,k)$ を b_k と表す．この b_k は，ノード k の近隣に多くのサンプルが所属するノードがあり，また，ノード k の参照ベクトルに類似するサンプルが近隣ノードに所属している場合に大きくなるものとしたい．このため，$\mathrm{Membership}(i,k)$ は以下の性質を持つものとする．まず，これを最整合ノード c_i とノード k のマップ上での距離 $||\mathbf{r}_{c_i}-\mathbf{r}_k||$ について減少関数とする．ここで c_i はサンプル i の最整合ノードである．さらに，$\mathrm{Membership}(i,k)$ は，サンプル i と参照ベクトル $\mathbf{m}_k$ の特徴量空間での距離 $\mathrm{d}(\mathbf{x}_i,\mathbf{m}_k)$ についても減少関数とする．以上より $\mathrm{Membership}(i,k)$ を下式のように定義する．

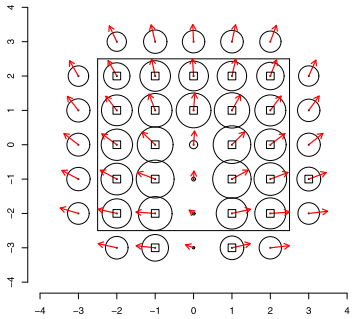
$$\mathrm{Membership}(i,k) = \frac{\mathrm{d}(\mathbf{x}_i,\mathbf{m}_k)^{-2}\cdot\exp\left(-||\mathbf{r}_{c_i}-\mathbf{r}_k||^2/2\sigma_s^2\right)}{\Sigma_{h\in\mathbf{U}}\mathrm{d}(\mathbf{x}_i,\mathbf{m}_h)^{-2}\cdot\exp\left(-||\mathbf{r}_{c_i}-\mathbf{r}_h||^2/2\sigma_s^2\right)} \quad (6)$$

ここで，σ_s は所属度をマップ上でどの程度広く配分するかを定めるパラメータであり，マップ更新の反復 s とともに減少させる．これにより，最終的には最整合ノードにのみ所属度を配分するようにし，マップ更新を収束させる．なお，本論文では σ_s の初期値を，利用する標準 SOM の近傍半径の初期値 σ と同じ値とした．これは，参照ベクトルの最大更新範囲と所属度の配分範囲をおおむね等しく設定していることとなる．

以上のように所属度を定義することにより，隣接ノードと参照ベクトルが類似しない，つまり，マップ不連続度の高いノードで隣接ノードからの所属度配分が少なくなる．このようなノードは，類似した隣接ノードを多く持つノードに比べ，存続に不利となり廃止されマップ不連続度の改善が期待される．また，廃止されたノードに代えて所属度累計の高いノードが採用され，サンプル密度の高い領域により多くの代表点をとることができる．これにより不適合度の改善が期待される．

4.4　FSOM 算法の詳細

算法の詳細を**図 8** に示す．これをマップのノード移動の例である**図 7** を用いて説明する．まず，初期利用ノー



$[-2,2]\times[-2,2]$ の四角内の 25 個のノードが初期利用ノード．ノード中心に□があるノードが更新された利用ノード．矢印が参照ベクトル．円の大きさが所属度累計．

図 7　マップの更新過程

Fig. 7　Update procedure of flexible SOM map.

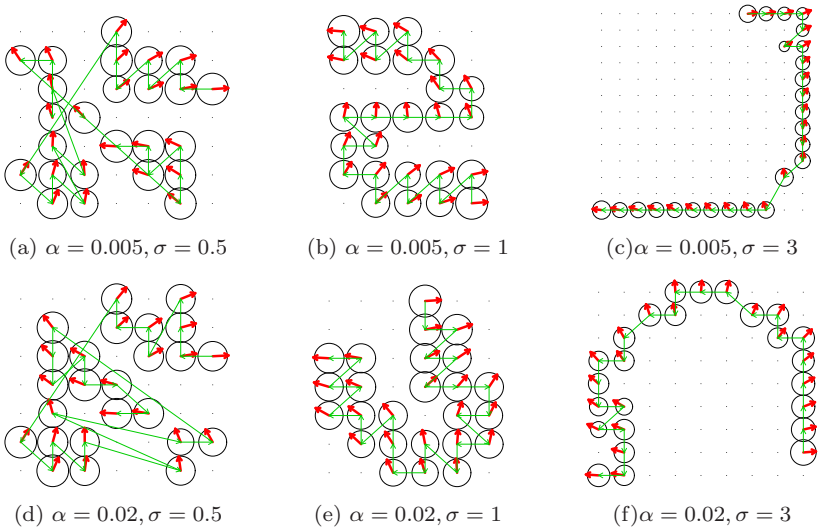
FlexibleSOM($\mathbf{X},\mathbf{L}^*,\mathbf{M},T,S,\{\alpha,\sigma,\sigma_0\}$)

1　$\mathbf{M}^*$ =SOM($\mathbf{X},\ \mathbf{L}^*,\mathbf{M},T,\{\alpha,\sigma,\alpha_0\}$);
2　$D^{2*}=\infty$;
3　for $s=0$ to $S-1$
4　　$\sigma_s=\sigma_0+(\sigma-\sigma_0)\left(\frac{S-s}{S}\right)$;
5　　$\mathbf{U}=\cup_{k\in\mathbf{L}}\mathbf{N}_k$;
6　　for $k\in\mathbf{U}$
7　　　$b_k=\sum_{i=1}^{N}\mathrm{Membership}(i,k)$;
8　　$\mathbf{L}=\{k\in\mathbf{U}\,|\,b_k$ が上位 K 個 $\}$;
9　　$\mathbf{M}=SOM(\mathbf{X},\ \mathbf{L},\mathbf{M}^*,T,\{\alpha,\sigma_s,\alpha_0\})$;
10　　D^2 =deviance($\mathbf{X},\mathbf{M}$);
11　　if $\mathbf{L}\neq\mathbf{L}^*$ or $D^2<D^{2*}$ then
12　　　$D^{2*}=D^2,\mathbf{M}^*=\mathbf{M},\mathbf{L}^*=\mathbf{L}$;
13　return $\mathbf{M}^*,\ \mathbf{L}^*$;

- $\mathbf{L},\mathbf{L}^*$; 利用ノード集合
- $\mathbf{M},\mathbf{M}^*$; 参照ベクトル ($k\in\mathbf{U}$)
- T ; 標準 SOM 反復回数
- S ; 利用ノードの更新回数
- $\mathbf{X}$，$\{\alpha,\sigma,\sigma_0\}$; 3.1 節参照
- deviance($\mathbf{X},\mathbf{M}$)：不適合度．式 (4) の算出関数
- Membership(i,k)：サンプル i のノード k への所属度．式 (6) で算出．

図 8　可変自己組織化マップ算法

Fig. 8　Algorithm of FSOM.



(a) $\alpha = 0.005, \sigma = 0.5$　(b) $\alpha = 0.005, \sigma = 1$　(c)$\alpha = 0.005, \sigma = 3$

(d) $\alpha = 0.02, \sigma = 0.5$　(e) $\alpha = 0.02, \sigma = 1$　(f)$\alpha = 0.02, \sigma = 3$

$T = 10^5$，$\sigma_0 = 0.01$，$K = 25$，$S = 30$．記号は図 3 に同じ．5 回の実験のうちで不適合度が最小の結果を例示している．

図 9　可変自己組織化マップ (1 次元位相構造)

Fig. 9　FSOM map of one-dimensional topology structure.

ド集合 $\mathbf{L}^*$ を図 7 で四角で囲まれたノード群とする．このとき，利用ノード候補集合 $\mathbf{U}$ は，図 7 で参照ベクトルの矢印を持つノード群となる．図 8 の 1 行目の SOM 呼び出しでは，参照ベクトル初期値 $\mathbf{M}^*$ を求めているが，これに用いられる関数 SOM は，$\mathbf{L}^*$ に対応する利用ノード候補集合 $\mathbf{U}$ 全体に対し参照ベクトルを与えるよう，図 1 の算法を拡張したものである．7 行目で算出する所属度累計値 b_k を，図 7 では円の大きさで表す．この場合に次の利用ノードとして選ばれるのは，8 行目で所属度累計値の大きい順に選択されるノード（図 7 でノード中心に □ を表示）となる．この更新された利用ノード集合 $\mathbf{L}$ に対し，9–10 行目で，参照ベクトルと不適合度を与え，11–12 行目では，利用ノード集合が更新されるか，不適合度が改善された場合に，マップを更新している．なお，利用ノード候補集合 $\mathbf{U}$ が拡張された場合には，拡張されて追加されたノードに新たな参照ベクトルが必要となるが，この場合の初期値には，0 ベクトルを与える．

5.　数値実験

5.1　パラメータ設定の評価

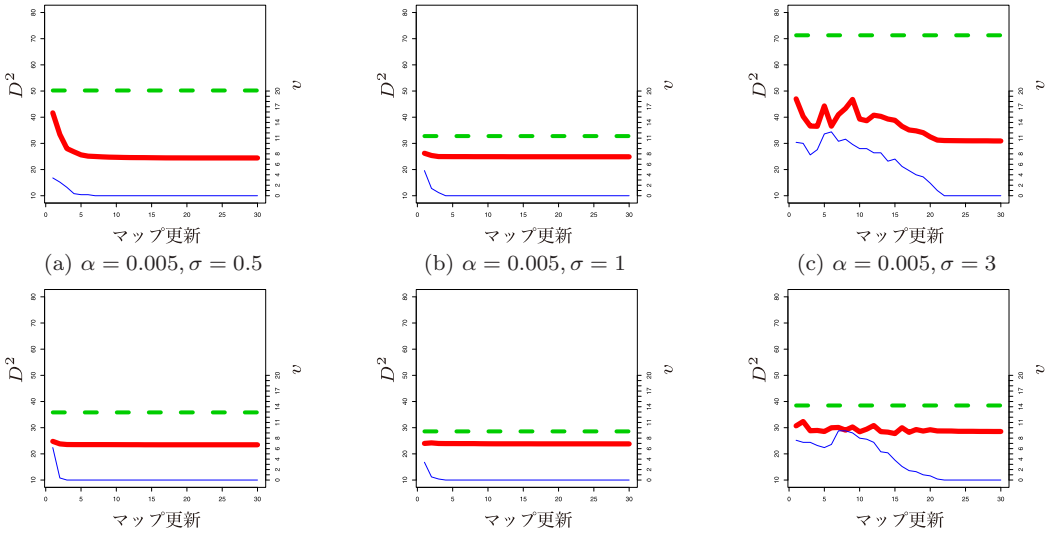
3.2 節の一次元の位相構造を持つサンプルセットを用いて，FSOM のパラメータ設定の評価のため数値実験を行う．提案する FlexibleSOM($\mathbf{X}, \mathbf{L}, \mathbf{M}, T, S, \{\alpha, \sigma, \sigma_0\}$) および標準 SOM($\mathbf{X}, \mathbf{L}, \mathbf{M}, T, \{\alpha, \sigma, \sigma_0\}$) を $\alpha = 0.005, 0.01, 0.02, 0.1, 0.2, 0.3, 0.4, 0.5$，$\sigma = 0.5, 1, 3$ と変化させて実行した．σ_0 と T は $\sigma_0 = 0.01$，$T = 10^5$ と固定している．なお，FSOM のマップ更新回数は $S = 30$ とした．得られたマップを**図 9** に例示する．以下で 3.3 節で定めた不適合度とマップ不連続度での評価結果を示す．

5.1.1　不適合度の評価結果

まずマップ更新にともないどのように不適合度が改善されていたかを確認する．あるマップ更新にともない配置位置が更新された利用ノードの個数を移動ノード数とする．**図 10** に移動ノード個数 v と不適合度の変化を示す．移動ノードは σ が大きい場合，反復の後半まで生起するが，最終的には収束し，不適合度は最後には標準 SOM よりも改善されることが分かる．次に不適合度のパラメータ設定ごとの最終結果を標準 SOM と比較したグラフを**図 11** (a) に示す．FSOM では標準 SOM と比べて，不適合度が改善されていることが分かる．また，σ が大きいときに α が小さいと，不適合度にばらつきが見られることが分かる．σ が大きいときは不適合度も大きくなるが，α も大きくすることで，σ が小さい場合の不適合度に近づくことが分かる．

5.1.2　マップ不連続度の評価結果

参照ベクトルの特徴量空間でのプロット例を**図 12** に示す．図 4 と異なり，FSOM では図 2 の代表点として不適当な参照ベクトルが存在しない．また，図 5 と比較して**図 13** を見ると，特徴量空間での距離が平均以上であるノードどうしは隣接しておらず，標準 SOM よりも位相が保たれていることが分かる．これは各パラメータ設定でのマップ不連続度をプロットした図 11 (b) から α や σ の設定に依存しないことが確認でき，FSOM では，局所的なマップ連続性が保たれることが分かる．また，σ は大きいほど，マップ連続性が保たれることも確認できる．ただし，図 9 を見ると，$\sigma = 0.5$ と σ が小さい場合には大域的な位相構造が分断された断片的ノード集合が生じてしまうことが見てとれる．



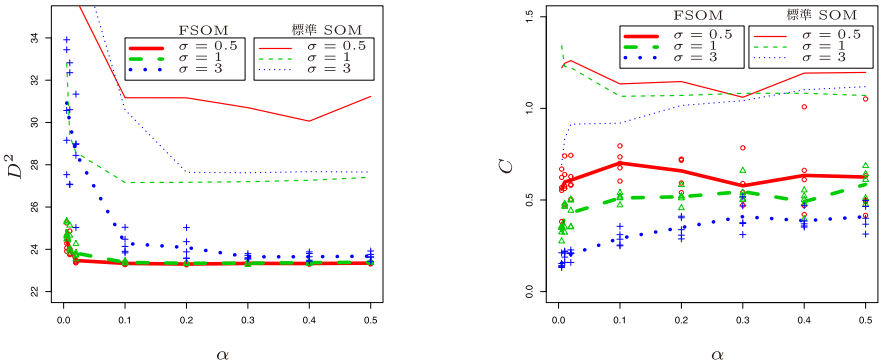
(a) $\alpha = 0.005, \sigma = 0.5$　(b) $\alpha = 0.005, \sigma = 1$　(c) $\alpha = 0.005, \sigma = 3$

(d) $\alpha = 0.02, \sigma = 0.5$　(e) $\alpha = 0.02, \sigma = 1$　(f)$\alpha = 0.02, \sigma = 3$

$T = 10^5$，$\sigma_0 = 0.01$，$K = 25$，$S = 30$．破線が標準 SOM の D^2，太い実線が FSOM の D^2，細い実線が移動ノード数 v である．それぞれ，5 回の実験の平均を示している．なお，標準 SOM は，マップ更新がないので単一の値のレベルを水平線で示す．

図 10　マップ更新にともなう不適合度とマップ更新量の収束（1 次元位相構造）

Fig. 10　Convergence of degree of unfitness against map update iterations.

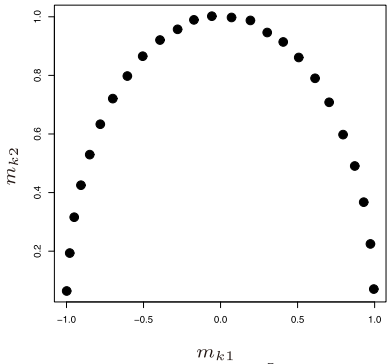


(a) 不適合度の比較　(b) マップ不連続度の比較

○ : $\sigma = 0.5$，△ : $\sigma = 1$，＋ : $\sigma = 3$．折れ線は σ ごとの平均である．標準 SOM はプロットを省略し，平均の折れ線のみ示した．

図 11　α と σ の効果（1 次元位相構造）

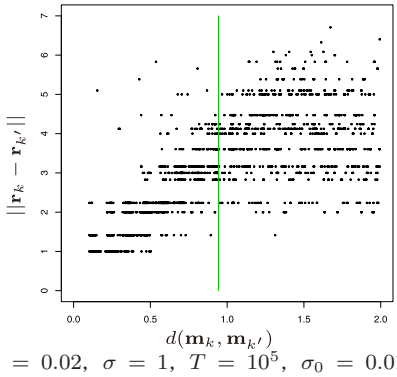
Fig. 11　Effect of α and σ.



$\alpha = 0.02$，$\sigma = 1$，$T = 10^5$，$\sigma_0 = 0.01$，$K = 25$，$S = 30$．図 9(e) の参照ベクトル $\{(m_{k1}, m_{k2}), k = 1, ..., 25\}$ を図 2 の特徴量空間にプロットした結果．

図 12　可変自己組織化マップの参照ベクトル例

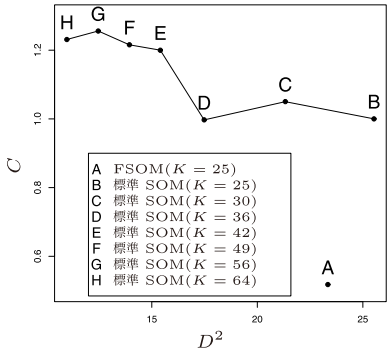
Fig. 12　Examples of reference vectors of FSOM.



$\alpha = 0.02$，$\sigma = 1$，$T = 10^5$，$\sigma_0 = 0.01$，$K = 25$，$S = 30$．乱数を変えた 5 回の実験での $\binom{25}{2} \times 5 = 1500$ 個の距離対を示す．中央の縦線は特徴量空間での距離の平均値（0.9438）を示す．

図 13　特徴量空間・マップ上の距離の関係（FSOM）

Fig. 13　Distances in feature space and over FSOM map.



FSOM は $\alpha = 0.2$, $\sigma = 1$, $T = 10^5$, $\sigma_0 = 0.01$, $S = 30$, 標準 SOM は $\alpha = 0.2$, $\sigma = 1$, $T = 3 * 10^6$, $\sigma_0 = 0.01$. 5 回の実験の平均値をプロットした.

図 14　不適合度 D^2 と不連続度 C のプロット（1 次元位相構造）

Fig. 14　Plot of degree of unfitness and discontinuity for one-dimensional topology structure.

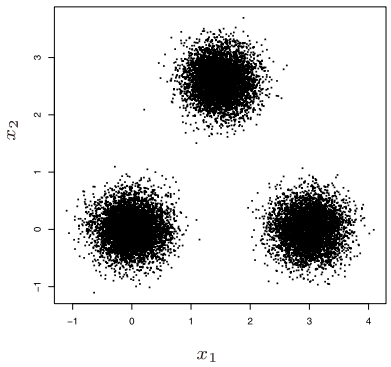


図 15　2 次元位相構造を持つサンプル

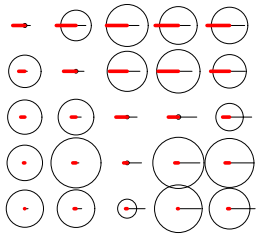
Fig. 15　Two-dimensional topology data.

5.2　標準 SOM でノード数を増やした場合との比較

標準 SOM でノード数を増やした場合の不適合度と不連続度を FSOM と比較する．ノード数 $K = 25, 30, 36, 42, 49, 56, 64$ で実験する．ここでは，FSOM の場合に利用ノードが 7×7 の格子空間（$K = 49$）に収まるため，それを上回る数値として K の最大値を 64 と設定した．また，同じ計算時間で比較するため，標準 SOM の反復回数 T を，FSOM の反復回数 T とマップ更新回数 S の積と等しくして行った．結果を**図 14** に示す．同一ノード数では不適合度，不連続度ともに FSOM が優れており，標準 SOM のノード数を増やした場合，不適合度は FSOM よりも低くなるが，不連続度は改善されないことが分かる．

5.3　クラスタ分離能力の確認

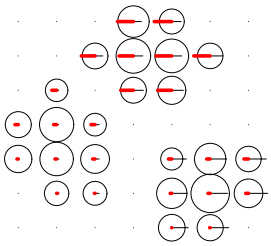
2 次元位相構造を持ち，複数の群が混合したサンプルセットから，その群を識別したマップが得られることを確認する．サンプルセットは，中心 μ_k が相互に標準偏差の 3 倍離れ，識別が容易な次の 3 つの分布 $\mathbf{x} \sim \mathrm{N}(\mu_k, \mathbf{I}_2)$ $(k = 1, 2, 3,\ \mu_1 = (0, 0)^t,\ \mu_2 = (3, 0)^t,\ \mu_3 = (1.5, 1.5\sqrt{3})^t)$ から 5,000 個ずつを生成したものを用いた．サンプルセット例を**図 15** に，標準 SOM によるマップを**図 16** に示す．すべてのノードが隣接しているために，どのノード群がクラスタを形成しているのか明確ではない．次に，図 6 に示す近傍を用いた FSOM によるマップを**図 17** に示す．ノード集合の近傍が，相互に他のノード集合と重ならない 3 つのクラスタに分かれていることが分かる．各ノードクラスタに属すサンプルを調べると，それぞれの群のサンプルは完全に分離できていることが分かった．つまり，予備知識なしで，3 つの分布を，近傍関係で連結である 3 つのクラスタとして分類できたことが分かる．また，この際の不適合度と移動ノード数を**図 18** に示す．不適合度が通常 SOM に比べて改善できていることが分かる．なお，パラメータ $\alpha = 0.005, 0.1, 0.2$ と $\sigma = 0.5, 1, 3$ の 6 組合せで実験し，すべてで同様の結果を得ている．



$\alpha = 0.02, \sigma = 1, \sigma_0 = 0.01, T = 10^5, K = 25.$
細い線が x_1，太い線が x_2.

図 16　標準 SOM によるマップ（2 次元位相構造）

Fig. 16　Standard SOM map of two-dimesional topology data.



$\alpha = 0.02, \sigma = 1, T = 10^5, \sigma_0 = 0.01, K = 25,$
$S = 30$. 細い線が x_1，太い線が x_2.

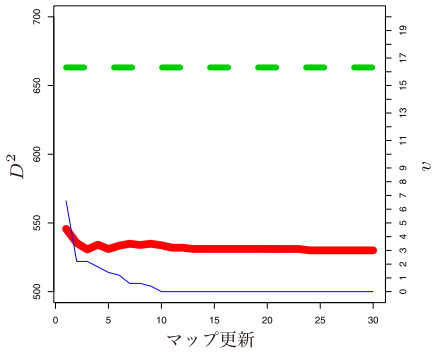
図 17　可変自己組織化マップ（2 次元位相構造）

Fig. 17　FSOM map of two-dimensional topology data.

標準 SOM のノード数と反復回数を増やした場合の不適合度と不連続度の FSOM との比較結果を**図 19** に示す．この場合も 1 次元位相構造の数値例と同様の結果となった．

6.　カード利用履歴への適用

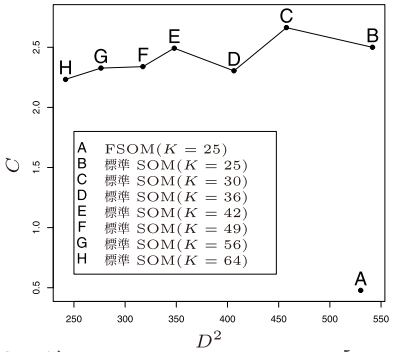
本章では，あるクレジットカードの利用履歴を対象とした適用事例を与える．用いる変量を**表 1** に示す．各変量は金額であるので，対数変換して，過去 6 カ月の月々の値を並べて当月のその利用者の特徴量（11×6 次元）としている．なお，月を並べる順番については利用と返済のバランスを示すキャッシュフローで整序してベクトル化している．参照ベクトルの表現法を**図 20** に示す．個人のサンプリングにあたっては，破産に至った利用者の比率を高めて



$\alpha = 0.02, \sigma = 1, T = 10^5, \sigma_0 = 0.01, K = 25, S = 30$．破線が標準 SOM の D^2，太い実線が FSOM の D^2，細い実線が移動ノード数 v である．5 回の実験の平均を示す．なお，標準 SOM は，マップ更新がないので単一の値のレベルを水平線で示す．

図 18　不適合度の収束

Fig. 18　Convergence of degree of unfitness against map update iterations.



FSOM は $\alpha = 0.02$，$\sigma = 1$，$T = 10^5$，$\sigma_0 = 0.01$，$S = 30$，標準 SOM は $\alpha = 0.02$，$\sigma = 1$，$T = 3 * 10^6$，$\sigma_0 = 0.01$．5 回の実験の平均値をプロットした．

図 19　不適合度 D^2 と不連続度 C のプロット（2 次元位相構造）

Fig. 19　Plot of degree of unfitness and discontinuity for two-dimensional topology structure.

抽出を行った．本実験のサンプルは 5 章で扱った数値例に比べ多様性が高いと想定し，標準 SOM では 9×9 のマップを用い，FSOM では $K = 81$ とする．また，マップが大きいので各ノードの近傍を周辺 8 ノードと広くしている．用いるパラメータは図 11 (b) より，マップ連続性が最も保たれる $\sigma = 3$ を参考とし，5 章の場合よりも大きなマップを考慮して $\sigma = 3, 4, 5, 6$ を選択した．また，図 11 (a) より，$\sigma = 3$ の場合に不適合度が安定する十分な値として $\alpha = 0.5$ を選ぶ．得られたマップ不連続度を σ ごとに比較すると，σ が大きいほどマップ不連続度が改善されるが不適合度が大きくなるという傾向が見られた．ただし，$\sigma = 5$ と $\sigma = 6$ との比較ではマップ不連続度の改善が見られなくなったため，以降では $\sigma = 5$ の実験結果を示す．得られた不適合度を**図 21** に示す．マップの形状を可変とすることで，標準 SOM よりも不適合度を改善できていることが

表 1　クレジットカード利用履歴の変量

Table 1　Variables in credit history.

変量名	説明
SP1 回払い	返却回数が 1 回の月払いの SP 利用金額
SP リボ払い	リボルビング払いの SP 利用金額
SP ボーナス払い	ボーナス一括払い，年 N 回払いをまとめた SP 利用金額
SP 分割払い	上記の SP 利用以外の支払い形態をまとめた SP 利用金額
SP 利用残高	当月における SP 利用残高金額
CS1 回払い	返却回数が 1 回の月払いの CS 利用金額
CS リボ払い	リボルビング払いの CS 利用金額
CS 分割払い	上記の CS 利用以外の支払い形態をまとめた CS 利用金額
CS 利用残高	当月における CS 利用残高金額
入金元金	当月における入金元金
入金手数料	当月における入金手数料

注　SP：ショッピング，CS：キャッシング

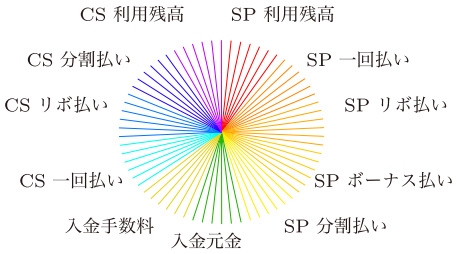
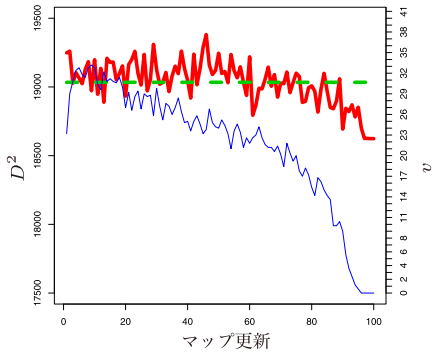


図 20　参照ベクトルの可視化

Fig. 20　Visualization of reference vector.



$\alpha = 0.5$，$\sigma = 5.0$，$K = 81$，$T = 100000$，$S = 100$．破線が標準 SOM の D^2，太い実線が可変自己組織化マップの D^2，細い実線が移動ノード数 v である．5 回の実験の平均である．なお，標準 SOM は，マップ更新がないので単一の値のレベルを水平線で示す．

図 21　不適合度（クレジットカードデータ）

Fig. 21　Convergence of degree of unfitness against map update iterations (credit card data).

分かる．また，マップ不連続度も標準 SOM の 4.166 に対し，3.057 と改善されている．標準 SOM のノード数と反復回数を増やした場合（**図 22**），標準 SOM の不適合度は FSOM よりも改善されるが，不連続度は改善されないこと

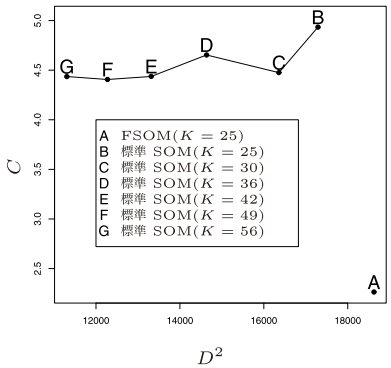
が，この例でも確認された．

6.1　標準 SOM の適用

標準 SOM によるマップを図 **23** に示す．左下にはショッピング利用（SP），右下には低頻度の利用，右上にはキャッシング利用（CS），左上には SP，CS の併用が見られる．

各顧客は月ごとに所属ノードが異なるため，ノード間の顧客の遷移を図 **24** に示す．矢印の太さは，移動元のノードの所属人数に対する，移動人数の割合である．その割合が 3%以上である矢印を記している．矢印の密度から，下側（SP 群）と右上（CS 群）と左上（SP・CS 併用群）のノード群に分かれていることが確認できるが，これらのノード群の間での顧客遷移の特徴をつかむことは難しい．

個々のノード群内の顧客移動を見ると，CS 群の上部で，顧客の遷移が多いノード群があり，CS リボ払いの利用月数が変わる顧客が多いことが分かる．また，SP 群の中央から右下へ向かって顧客が移動するように矢印があり，SP 利用の返済を終えた顧客は利用を休む傾向があると解釈できる．このマップ上で，3 カ月以上の延滞の比率が半分を超えるノードは F6 であるが，図 24 を見ても，重度の延滞がどのようなノードを経由して発生するかを読み取ることは難しい．



FSOM は $\alpha = 0.5$，$\sigma = 5$，$T = 10^5$，$\sigma_0 = 0.01$，$S = 100$，標準 SOM は $\alpha = 0.5$，$\sigma = 5$，$T = 10^7$，$\sigma_0 = 0.01$．5 回の実験の平均値をプロットした．

図 22　不適合度 D^2 と不連続度 C のプロット（クレジットカードデータ）

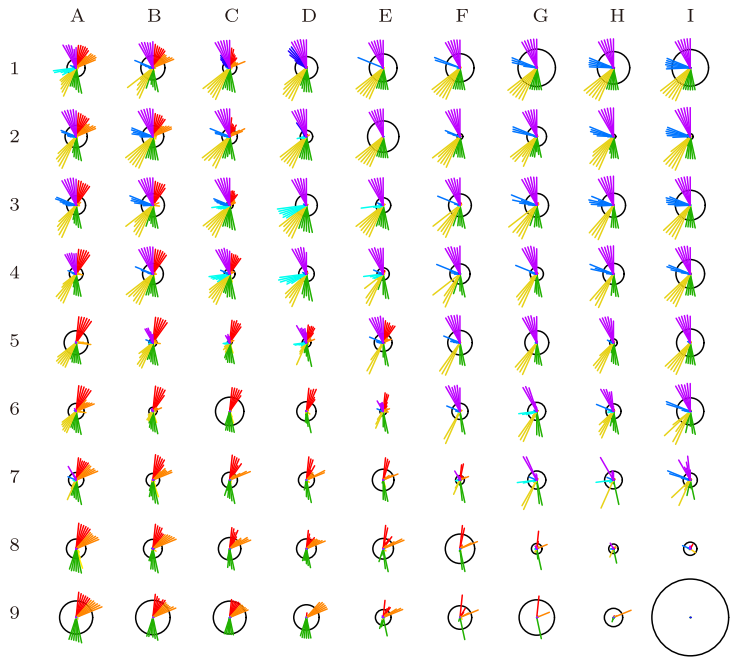
Fig. 22　Plot of degree of unfitness and discontinuity for credit card data.

6.2　可変自己組織化マップの適用

FSOM によるマップを図 **25** に示す．各ノードの周囲 8 カ所の近傍でつながっているノード集合をクラスタとして囲んで示した．

クラスタごとに図 23 と比較する．クラスタ 1 は，SP 利用残高に対して手数料が発生している利用である．クラスタ 2 は利用がなく SP 利用の返済をしており，分割払いによる返済も見られる．クラスタ 3 とクラスタ 4 上側は SP1 回払いを主とする利用であり，クラスタ 4 下側は頻度の低い利用である．クラスタ 3，クラスタ 4 の上側，クラスタ 4 の下側の順に頻度が低くなっていくことがマップ上の位置関係に表れている．クラスタ 5 は頻度が低く手数料が発生している利用である．以上の SP 利用クラスタおよび低頻度クラスタは，標準 SOM では下側のノード群と対応する．

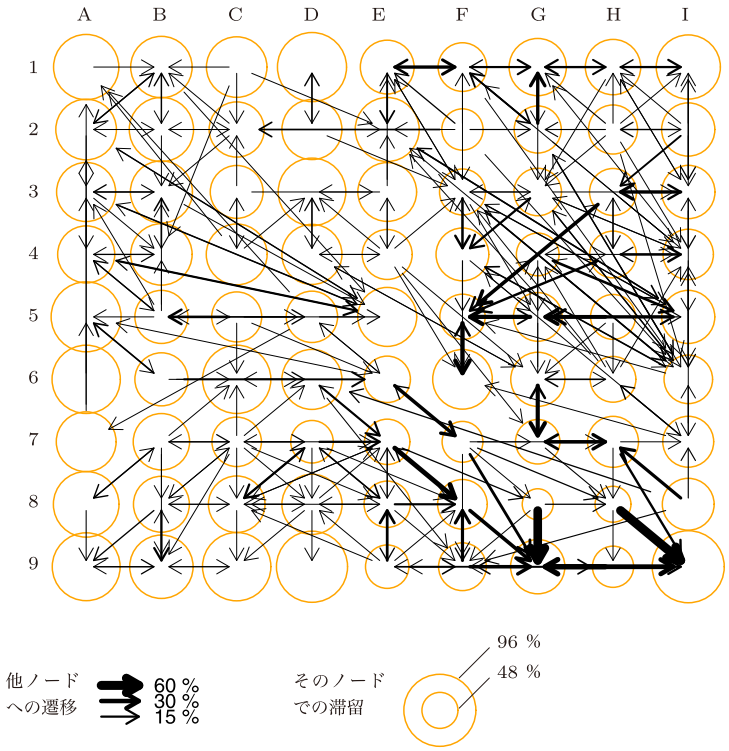
クラスタ 6，7，8，9 は多めの CS 利用残高と手数料を特徴とし，それぞれのパターンで部分的に返済していない月



$\alpha = 0.5$，$\sigma = 5.0$，$\sigma_0 = 0.01$，$K = 81$，$T = 100000$.

図 23　標準 SOM によるマップ（クレジットデータ）

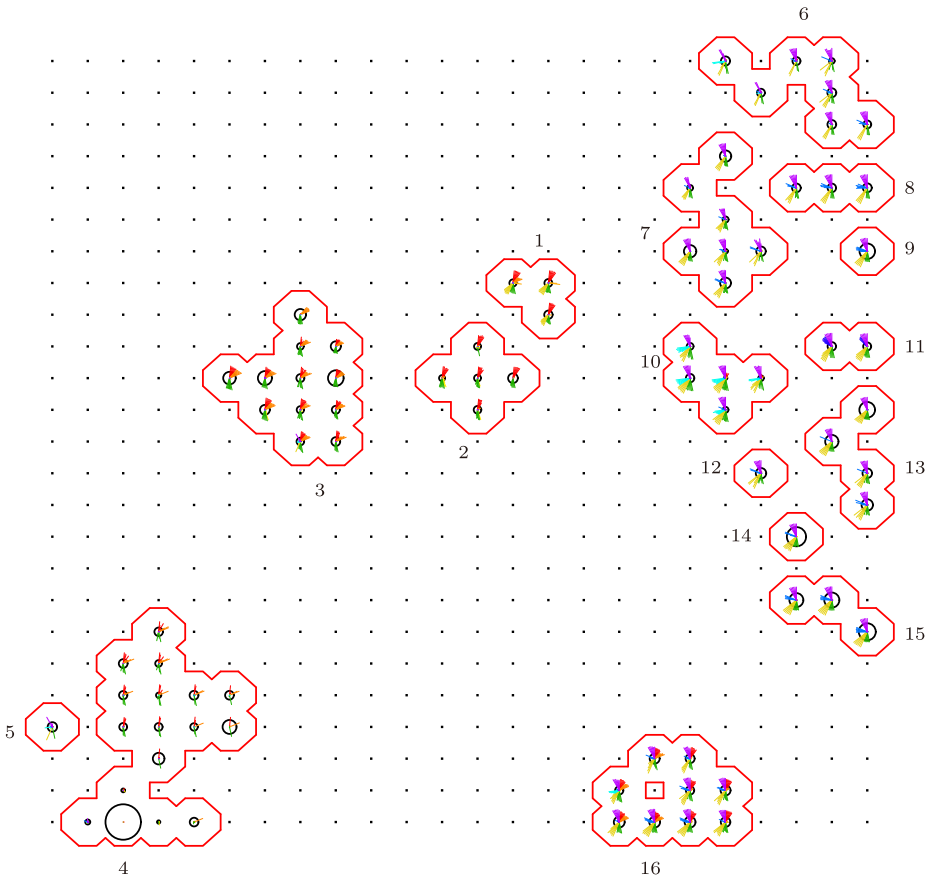
Fig. 23　Standard SOM map for credit card data.



矢印の太さは，各ノードの所属人数のうち，他のノードへ遷移した人数の割合を示す．上の凡例では代表的割合での太さを示す．円は，各ノードの所属人数のうち，そのノードに滞留した人数の割合を示す．

図 24　顧客のノード間遷移（標準 SOM）

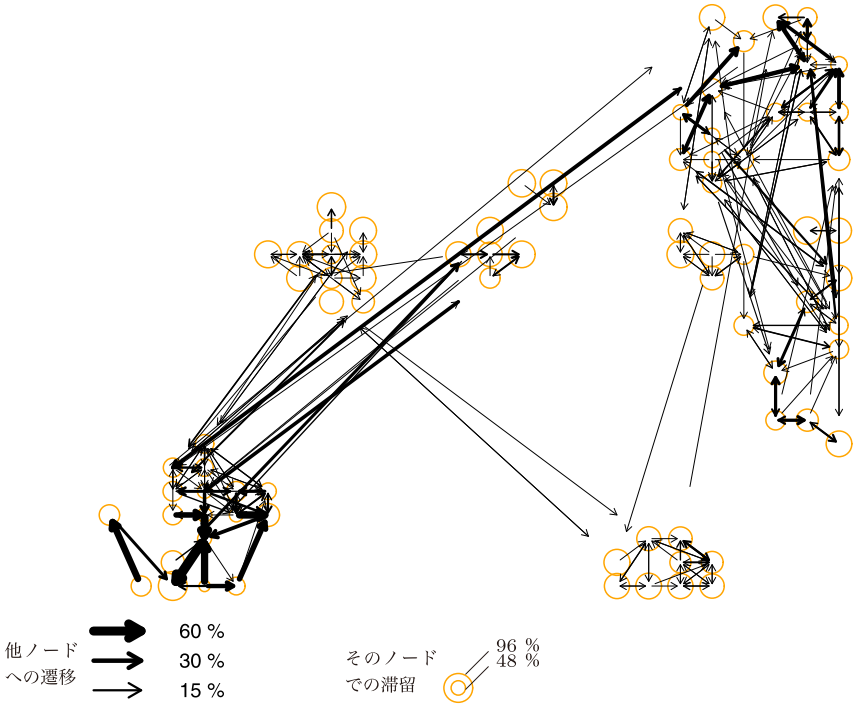
Fig. 24　Node transition of clients over standard SOM map.



$\alpha = 0.5$, $\sigma = 5.0$, $\sigma_0 = 0.01$, $K = 81$, $T = 100000$, $S = 100$.

図 25　可変自己組織化マップ（クレジットデータ）

Fig. 25　FSOM map for credit card data.



矢印の太さは，各ノードの所属人数のうち，他のノードへ遷移した人数の割合を示す．上の凡例では代表的割合での太さを示す．円は，各ノードの所属人数のうち，そのノードに滞留した人数の割合を示す．

図 26　顧客のノード間遷移（可変自己組織化マップ）

Fig. 26　Node transition of clients over FSOM map.

が存在する．クラスタ 10 は CS1 回払いの利用が多い．クラスタ 11，12，13，14，15 は多めの CS 利用残高と手数料の発生があるが，おおむね毎月返済している利用である．これらの CS 利用クラスタは，標準 SOM の中央から右上のノード群に対応する．なお，図 25 上で 3 カ月以上の延滞の比率が半分を超えるのは，こららの中でも特に返済していない月の多いクラスタ 6 の右上 2 ノードである．

クラスタ 16 は CS，SP を併用した利用である．上側に向けて利用残高が多いまま返済月数が少なくなり，左側では SP1 回払いが多くなっている．標準 SOM では左上のノード群に対応する．

FSOM の場合の顧客遷移図を**図 26** に示す．ノード群が SP 群（中央），低頻度群（左下），CS 群（右上），SP・CS 併用群（右下）と，標準 SOM の場合よりも明確に分かれている．また，分かれたノード群の間で，顧客遷移の多い部分を確認することもできる．

標準 SOM の遷移図では右上のノード群の間で顧客移動が多かったが，それに対応する移動は FSOM の遷移図の CS 群下部で確認できる．また，SP 利用の返済を終えた顧客が利用を休む傾向は，FSOM 遷移図の左下に見ることができる．

以下，標準 SOM ではできなかった解釈を記述する．まず，低頻度群と CS 群クラスタ 6 の左側との間で遷移が多く，CS 群上側から CS 群下側に向かう遷移も多いことが分かる．これはほとんど休眠状態の顧客（クラスタ 4）が CS1 回払い（クラスタ 6 左側）から少しずつ利用を始め，その後，クラスタ 8 のリボ払い CS などに遷移し，CS のみを利用するクラスタ 9〜15 の間を遷移すると解釈できる．なお，重度の延滞があるクラスタ 6 右側へは，利用残高のある月数に比べ，返済月数が少なめのクラスタ 7 上側からや，リボ払いのクラスタ 8 から太い遷移がある．

手数料が発生する SP 利用（クラスタ 1）では，他クラスタ間との目立った遷移がなく，安定してそのタイプの使い方をしていると解釈できる．また，SP・CS 併用群への遷移は SP 群と CS 群からのみであり，低頻度利用の顧客が急に SP と CS を同時に使いはじめることはないことが分かる．

以上のように，FSOM では標準の SOM に比べてマップ連続性の高いマップを構成できたため，マップの位相関係を利用してマップ上での分析をする際に有利であることが明らかとなった．

7.　おわりに

本研究では，従来の自己組織化マップを，任意形状の格子点集合の利用を許容するように拡張することで，高次元特徴量データの持つ自然な位相構造を低次元マップとして表現する可変自己組織化マップを提案した．提案手法により，従来の自己組織化マップと同じ個数の参照ベクトルを用いて，参照ベクトルがサンプルへ適合する程度を改善し，参照ベクトルのマップ上での連続性を改善できた．また，

異なった特徴量類型が非連結なノードクラスタとして表現され，位相構造を自然に表すマップが生成できることを数値例や適用事例から示した．正方格子以外の格子を用いた場合の評価や，計算量の低減などが今後の課題となる．

参考文献

[1] Kohonen, T.: *Self-Organizing Maps*, Springer (1995).
[2] 関　庸一，長井　歩，石原純一郎，渡辺亮：自己組織化マップによる行動履歴の類型化—クレジットカード利用履歴を利用したキャッシング移行予測，日本経営工学会誌，Vol.57, No.5, pp.404–412 (2006).
[3] 五反田剛，石井良和，原健一郎，関　庸一：SOM によるファン層の解析に基づく CD 購買予測モデルの作成，オペレーションズ・リサーチ，Vol.52, No.2, pp.87–93 (2007).
[4] 徳高平蔵，藤村喜久郎，山川　烈：自己組織化マップ応用事例集，海文堂出版株式会社 (2002).
[5] Seki, Y. and Okawara, E.: State Diffusion Model on SOM map based on Multinomial Logistic Model, *4th World Conference of the International Association for Statistical Computing*, Dec. 7(5-8), Yokohama, Japan, pp.1391–1396 (2008).
[6] Anderson, T.W.: *An introduction to Multivariate Statistical Analysis*, Wiley (1984).
[7] 高根芳雄：多次元尺度法，東京大学出版会 (1980).
[8] Roweis, S.T. and Saul, L.K.: Nonlinear Dimensionality Reduction by Locally Linear Embedding, *Science*, Vol.290, pp.2323–2326 (2000).
[9] Tenenbaum, J.B., de Silva, V. and Langford, J.C.: A Global Geometric Framework For Nonlinear Dimensionality Reduction, *Science*, Vol.290, pp.2319–2323 (2000).
[10] Fritzke, B.: A growing neural gas network learns topologies, *Advances in Neural Information Processing Systems 7* (*NIPS '94*), pp.625–632, MIT Press (1995).
[11] 須藤明人，佐藤彰洋，長谷川修：自己増殖型ニューラルネットを用いたノイズのある環境下での追加学習が可能な連想記憶モデル，日本神経回路学会誌，Vol.15, No.2, pp.98–109 (2008).
[12] 島田敬士，谷口倫一郎：密度可変型自己組織化マップによる追加学習の実現法，日本神経回路学会誌，Vol.14, No.2, pp.71–78 (2007).
[13] Sammon, J.W.: A Nonlinear Mapping for Data Structure Analysis, *IEEE Trans. Comput.*, Vol.C-18, No.5 (1969).
[14] Vesanto, J.: SOM-Based Data Visualization Methods, *Intelligent Data Analysis*, Vol.3, Issue 2, pp.111–126 (1999).
[15] Hastie, T., Tibshirani, R. and Friedman, J.: *The Elements of Statistical Learning*, Springer (2001).

多賀谷 侑史

2012 年群馬大学大学院工学研究科情報工学専攻前期博士課程修了．同年より（株）サンデンシステムエンジニアリングに勤務．現在に至る．

安藤 晋　（正会員）

2004 年東京大学大学院工学系研究科電子工学専攻博士課程修了．同年東京工業大学総合理工学研究科特別研究員．2005 年横浜国立大学工学研究院助手．2005 年，群馬大学大学院工学研究科助教．現在に至る．データマイニング，知識情報処理研究に従事．電子情報通信学会，人工知能学会，IEEE，ACM 各会員．

関 庸一

1982 年早稲田大学理工学部数学科卒業．1998 年同大学大学院理工学研究科工業経営学専門分野博士後期課程単位取得後退学．1987 年早稲田大学理工学部助手．1990 年群馬大学工学部情報工学科助手．2002 年同大学教授．現在に至る．工学博士．データマイニングおよび応用データ解析の研究に従事．日本経営工学会，日本品質管理学会，応用統計学会，OR 学会，人工知能学会，INFORMS 等会員．