K-L 情報量正則化を用いた 線形ファジィクラスタリング法*

(Linear Fuzzy Clustering with Regularization by K-L information)

本多　克宏 †・神田　章裕 †・市橋　秀友 †
† 大阪府立大学　大学院　工学研究科
Graduate School of Engineering,
Osaka Prefecture University;
1-1 Gakuen-cho, Sakai city, Osaka 599-8531, JAPAN

概 要

データを球状のクラスターに分類する Fuzzy c-Means（FCM）法やその派生手法とガウス混合モデル（GMMs）との類似性に着目して，EM アルゴリズムと良く似たアルゴリズムによりデータの分類が行われるファジィクラスタリング法として K-L 情報量正則化を用いる FCM（KFCM）法が提案されている．本論文では，局所的な主成分分析法と線形ファジィクラスタリング法との関連を議論した後に，確率モデルに基づく混合主成分分析法を線形ファジィクラスタリングに拡張した手法を提案する．提案手法は KFCM 法にパラメータ数の制約を加えることにより定義されているが，特別な場合には代表的な線形ファジィクラスタリング法である Fuzzy c-Varieties（FCV）法とほぼ等しい結果を与えることから，K-L 情報量正則化を用いることで FCV 法を適応的な形状調整機能を有する手法に拡張したモデルであるとみなすことができる．

(FCM-type fuzzy clustering approaches are closely related to Gaussian Mixture Models (GMMs) and the objective function of Fuzzy c-Means with regularization by K-L information (KFCM) is optimized by an EM-like algorithm. In this paper, we propose to apply probabilistic PCA mixture models to linear clustering following the discussion on the relationship between Local PCA and linear fuzzy clustering. Although the proposed method is a kind of the constrained model of KFCM, the algorithm includes a similar formulation with the Fuzzy c-Varieties (FCV) algorithm as a special case. Then the algorithm can be regarded as a modified FCV algorithm with regularization by K-L information, which makes it possible to tune the cluster shapes adaptively.)

keywords: Fuzzy c-varieties, Probabilistic principal component analysis, Regularization by K-L information.

1　はじめに

大規模なデータベースが有する潜在的な特徴を抽出するための手法として，局所領域における主部分空間をとらえる局所的な主成分分析（Local Principal Component Analysis: Local PCA）が注目されている．局所的な主部分空間の抽出は，データ集合をいくつかの部分集合に分割する過程と部分集合ごとにデータの特徴を良く表す線形モデルを推定する過程からなり，様々な統計的手法が提案されている．Fukunaga ら [1] は，前処理としてクラスタリング手法により標本データの類似度に基づいてデータ集合を分割した後に，クラスターごとに局所的な Karhunen-Loéve 展開を施す手法を提案している．また，Kambhatla ら [2] や Hinton ら [3] は，主部分空間からの復元誤差を用いることにより，局所的な主成分の抽出に適したデータ分割を得るための繰り返しアルゴリズムを提案している．特に，Hinton らの “soft” な局所的主成分分析 [3] では，EM（Expectation-Maximization）アル

*知能と情報（日本知能情報ファジィ学会誌）, 15, 6, 682-692 (2003)

ゴリズム [4] と類似したアルゴリズムが用いられている．そこでは，データの分類に関する情報を "欠測値" とみなし，おのおのの標本データの局所的な主成分分析モデルへの貢献度を復元誤差に基づいて推定している．したがって，厳密には確率モデルが規定されていないものの，擬似的な尤度関数の最大化の仮定の下でモデルが推定される．

Roweis [5] や Tipping ら [6] は主成分分析のための確率モデルを提案している．確率的主成分分析（Probabilistic Principal Component Analysis: PPCA） [6] では，確率分布が厳密に定義されており，すべての未知パラメータが単一の尤度関数最大化に基づいて推定されるために，確率論に基づく様々な性質が利用可能となっている．これらの主成分モデルは，その混合分布を考えることで，データの分割と局所的な主成分分析を同時に考慮する混合主成分分析（Mixture of probabilistic PCA: MPCA）を与える．ここで，MPCA はガウス混合モデル（Gaussian Mixture Models: GMMs）にパラメータ数の制約を加えたモデルとみなされるものであり，データ数の少ない多次元データ集合を扱う際には，分散共分散行列のすべての要素を未知パラメータとして用いる GMMs よりも高い汎化能力を持つことが示されている [7]．

一方，クラスター分析法もデータの部分構造をとらえるための方法として様々な応用が試みられている．Bezdek らの Fuzzy c-Means（FCM） [8] やその派生手法 [9, 10, 11] は GMMs との類似性が指摘されており [12]，宮岸ら [13] は K-L 情報量（Kullback-Leibler 情報量）に基づく正則化を用いることで GMMs における EM アルゴリズムと類似したクラスタリングアルゴリズムが得られる KFCM 法を提案している．また，線形ファジィクラスタリング手法である Fuzzy c-Varieties（FCV） [14] は，クラスターのプロトタイプとして線形多様体を用いることで局所的な線形構造をとらえる手法であるが，プロトタイプを張るベクトルをファジィ散布行列の固有ベクトルとして算出することから，局所的な主成分分析との関連が深い．一般に，ファジィクラスタリングでは標本データ点とプロトタイプとの距離の最小化が目的となるが，線形クラスタリング法のアルゴリズムには局所的な主成分分析法との類似点が多い．

そこで本論文では，様々な局所的主成分分析法と線形ファジィクラスタリング手法との関連を議論した後に，クラスター形状の調整機能を持つ新たな線形ファジィクラスタリング法を提案する．提案手法は MPCA のモデルを拡張したものであり，KFCM 法にパラメータ数の制約を付加したモデルとなっている．また，特別な場合には FCV 法とほぼ等しい結果を与える提案手法は，K-L 情報量正則化を用いることで FCV 法を適応的な形状調整機能を有する手法に拡張したモデルであるとみなすことができる．

以下では，確率分野における混合モデルとファジィクラスタリングとの関連性を議論した後に，確率的主成分分析法の拡張とみなされる線形ファジィクラスタリング手法を提案する．

2　混合分布モデルとファジィクラスタリング

2.1　ガウス混合モデル

m 次元の観測値からなる n 個の標本データ $\boldsymbol{x}_i = (x_{i1}, \cdots, x_{im})^\top$, $i = 1, \cdots, n$ が与えられたとする．データ $\boldsymbol{x}$ が複雑な形状の密度分布を持つ場合には，その確率密度分布をいくつかの確率密度分布の重み付き線形結合で表す混合分布モデルが考えられる．おのおの $p_c(\boldsymbol{x})$ なる密度分布を持つ確率モデルを C 個用いる場合，その混合分布モデルは以下のように定義される．

$$p(\boldsymbol{x}) \quad = \quad \sum_{c=1}^{C} \pi_c p_c(\boldsymbol{x}) \tag{1}$$

ここで，重み係数 π_c は事前確率を表しており，条件

$$\sum_{c=1}^{C} \pi_c = 1, \qquad 0 \leq \pi_c \leq 1 \tag{2}$$

を満たすものとする．同様に，おのおのの確率密度分布 $p_c(\boldsymbol{x})$ は，

$$\int p_c(\boldsymbol{x}) d\boldsymbol{x} = 1 \tag{3}$$

を満たす．$p_c(\boldsymbol{x})$ としては，しばしばガウス分布（正規分布），

$$\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{b}_c, \Sigma_c) \tag{4}$$

が用いられ，この場合の混合分布モデルはガウス混合モデル（Gaussian Mixture Models: GMMs）または正規混合分布と呼ばれる（たとえば [15, 16]）．ここで，分散共分散行列 Σ_c の推定の際には，すべての要素を未知パラメータとする場合（full）の他，対角要素のみを用いてそれ以外をすべて 0 に固定する場合（diagonal）や単位行列を用いる場合（spherical）などがある．おのおのの未知パラメータを推定する際の目的関数としては，負の対数尤度

$$L_{gmm} = -\sum_{i=1}^{n} \log\left\{\sum_{c=1}^{C} \pi_c p_c(\boldsymbol{x}_i)\right\} \tag{5}$$

が用いられる．(5) 式を最小とする未知パラメータは，E-step （Expectation step）および M-step（Maximization step）からなる EM アルゴリズムにより導かれる．分散共分散行列のすべての要素を未知パラメータとして用いる場合には，上記の二つのステップは以下のように表される．

- E-step　おのおのの標本データの各クラスへの貢献度（事後確率）を推定する．

$$\begin{aligned} u_{ci} &= \frac{\pi_c p_c(\boldsymbol{x}_i)}{\sum_{l=1}^{C} \pi_l p_l(\boldsymbol{x}_i)} \\ &= \frac{\pi_c \exp(-\frac{1}{2}E_{ci})|\Sigma_c|^{-\frac{1}{2}}}{\sum_{l=1}^{C} \pi_l \exp(-\frac{1}{2}E_{li})|\Sigma_l|^{-\frac{1}{2}}} \end{aligned} \tag{6}$$

ただし，

$$E_{ci} = (\boldsymbol{x}_i - \boldsymbol{b}_c)^\top \Sigma_c^{-1} (\boldsymbol{x}_i - \boldsymbol{b}_c) \tag{7}$$

である．

- M-step　各クラスにおけるパラメータを推定する．

$$\pi_c = \frac{1}{n}\sum_{i=1}^{n} u_{ci} \tag{8}$$

$$\boldsymbol{b}_c = \frac{\sum_{i=1}^{n} u_{ci}\boldsymbol{x}_i}{\sum_{i=1}^{n} u_{ci}} \tag{9}$$

$$\Sigma_c = \frac{\sum_{i=1}^{n} u_{ci}(\boldsymbol{x}_i - \boldsymbol{b}_c)(\boldsymbol{x}_i - \boldsymbol{b}_c)^\top}{\sum_{i=1}^{n} u_{ci}} \tag{10}$$

2.2　FCM 法とその他の派生手法

FCM 法やその他の派生手法のアルゴリズムには，混合分布モデルとの類似点が多い．Bezdek らの標準的な FCM 法 [8] では，クラスタリング基準 d_{ci} としてクラスターのプロトタイプ（中心）と標本データ点との距離が用いられ，目的関数は以下のように定義される．

$$\begin{aligned} L_{fcm1} &= \sum_{c=1}^{C}\sum_{i=1}^{n} (u_{ci})^\theta d_{ci} \\ &= \sum_{c=1}^{C}\sum_{i=1}^{n} (u_{ci})^\theta ||\boldsymbol{x}_i - \boldsymbol{b}_c||^2 \end{aligned} \tag{11}$$

ここで，u_{ci} は第 i 標本データの第 c クラスターへのメンバシップ (帰属度) を表している．指数 $(\theta > 1)$ はメンバシップのファジィ化のために考慮されたもので，θ が大きくなるほど標本データの帰属度が {0,1} の値をとらなくなり，所属するクラスターがあいまいとなるファジィな分割が得られるようになる．最適なデータ分割は，繰り返しアルゴリズムにより導かれる．最適性の必要条件から，新たなプロトタイプは一般化されたメンバシップの重み付きの平均値となり，おのおのの標本データのメンバシップはクラスタリング基準を用いて算出される．通常，c に関する u_{ci} の和は 1 となる制約が付加される．この制約は “確率的制約” [17] と呼ばれ，ガウス混合モデルを用いた EM アルゴリズムにおける事後確率の計算式と類似性がある．

メンバシップのファジィ化のためには，メンバシップのべき乗の他に，エントロピー正則化 [11] が提案されている．そこでは，メンバシップが {0,1} の値を持つクリスプなデータ分割を “特異” と

みなし，ファジィ化された解を正則であるかのように考えてエントロピーを用いた正則化を行っており [18]，FCM 法の目的関数は，

$$
\begin{aligned}
L_{fcm2} &= \sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci}d_{ci} + \lambda\sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci}\log u_{ci} \\
&= \sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci}||\boldsymbol{x}_i - \boldsymbol{b}_c||^2 \\
&\quad +\lambda\sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci}\log u_{ci} \qquad (12)
\end{aligned}
$$

のように定義される．ここで，エントロピー項が標準的な手法におけるメンバシップのべき乗と同様の働きを担い，λ が大きくなるほど，あいまいなデータ分割が得られるようになる．最適性の必要条件から，メンバシップとクラスター中心の更新則は，おのおの，

$$
u_{ci} = \frac{\exp(-\frac{1}{\lambda}||\boldsymbol{x}_i - \boldsymbol{b}_c||^2)}{\sum_{l=1}^{C}\exp(-\frac{1}{\lambda}||\boldsymbol{x}_i - \boldsymbol{b}_l||^2)} \qquad (13)
$$

$$
\boldsymbol{b}_c = \frac{\sum_{i=1}^{n} u_{ci}\boldsymbol{x}_i}{\sum_{i=1}^{n} u_{ci}} \qquad (14)
$$

のように求められる．これら二つの更新則は，ガウス混合分布において未知パラメータを平均 $\boldsymbol{b}_c$ のみに限定し，$\pi_c = 1/C$（定数）とおいた場合に等しい．ただし，分散共分散行列は単位行列を λ 倍したものと等しく，ファジィ度を決める係数 λ がおのおののガウス分布の分散を表しているととらえられる [12]．

宮岸ら [13] はすべての変量を未知パラメータとして用いるガウス混合分布に対応するファジィクラスタリング法として，K-L 情報量正則化を用いた FCM（KFCM）法を提案している．(12) 式のエントロピー項を K-L 情報量に置き換え，ガウス分布における分散共分散行列の行列式の大きさを制限するための項を付加することにより，KFCM 法の目的関数は，以下のように定義される．

$$
\begin{aligned}
L_{kfcm} &= \sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci}d_{ci} + \lambda\sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci}\log\frac{u_{ci}}{\pi_c} \\
&\quad +\sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci}\log|\Sigma_c| \qquad (15)
\end{aligned}
$$

ただし，d_{ci} はマハラノビス距離であり，

$$
d_{ci} = (\boldsymbol{x}_i - \boldsymbol{b}_c)^{\top}\Sigma_c^{-1}(\boldsymbol{x}_i - \boldsymbol{b}_c) \qquad (16)
$$

で表され，分散共分散行列 Σ_c も未知パラメータとして扱われる．また，π_c は第 c クラスターの容量を表す．(15) 式の最小化は，c に関して u_{ci} と π_c のそれぞれの和が 1 となる制約のもとで行われる．ここで，エントロピー正則化を用いる (12) 式の目的関数では，エントロピー項の最小化を目指してすべてのクラスターにおいて u_{ci} が等しい値（$1/C$）に近づくようにすることでメンバシップのファジィ化を行うのに対して，(15) 式の KFCM 法の目的関数では，第 c クラスターにおいてすべての標本データのメンバシップ u_{ci} がクラスター容量 π_c に近づくようにすることで，メンバシップのファジィ化を行っている．したがって，K-L 情報量正則化ではクラスター容量を考慮したメンバシップのファジィ化が行われており，係数 λ によりファジィ度を調整している．最適性の必要条件から，不動点繰り返しアルゴリズムにおける更新則は以下のように求められる．

$$
u_{ci} = \frac{\pi_c\exp\bigl(-\frac{1}{\lambda}d_{ci}\bigr)|\Sigma_c|^{-\frac{1}{\lambda}}}{\sum_{l=1}^{C}\pi_l\exp\bigl(-\frac{1}{\lambda}d_{li}\bigr)|\Sigma_l|^{-\frac{1}{\lambda}}} \qquad (17)
$$

$$
\boldsymbol{b}_c = \frac{\sum_{i=1}^{n} u_{ci}\boldsymbol{x}_i}{\sum_{i=1}^{n} u_{ci}} \qquad (18)
$$

$$
\pi_c = \frac{1}{n}\sum_{i=1}^{n} u_{ci} \qquad (19)
$$

$$\Sigma_c = \frac{\sum_{i=1}^n u_{ci}(\boldsymbol{x}_i - \boldsymbol{b}_c)(\boldsymbol{x}_i - \boldsymbol{b}_c)^\top}{\sum_{i=1}^n u_{ci}} \tag{20}$$

上記のアルゴリズムは，K-L 情報量の項の係数を $\lambda = 2$ とおいたときには，ガウス混合モデルにおける EM アルゴリズムと等しくなるが，$\lambda \neq 2$ のときには，(3) 式を満たす $p_c(\boldsymbol{x})$ を定義できないことから，対応する確率密度分布は存在しない．

3　局所的な主成分分析と線形ファジィクラスタリング

3.1　局所的な主成分分析と確率モデル

分散共分散行列の要素のすべてを未知パラメータとして用いるガウス混合モデルは，多様な形状の部分集合をとらえられるものの，局所的な線形構造の抽出においては必ずしも有用であるとは限らない．そこで，非線形な分布形状を有する多次元データの次元圧縮には，局所的な主成分分析が用いられることが多い．Kambhatla ら [2] は，データ集合の復元誤差を用いた（クリスプな）クラスタリングとおのおののクラスターにおける局所的な主成分の抽出を繰り返すアルゴリズムを提案している．また，Hinton ら [3] はその考えを “soft” なデータ分割に拡張している．“soft” なデータ分割では，おのおのの標本データ点は唯一の主成分分析器に帰属するのではなく，復元誤差の大きさに応じてすべての分析器で共有される．そして，最小化すべき目的関数は，以下の擬似的な負の対数尤度により与えられる．

$$L_{lpca} = \frac{1}{\lambda} \sum_{c=1}^C \sum_{i=1}^n u_{ci} E_{ci} + \sum_{c=1}^C \sum_{i=1}^n u_{ci} \log u_{ci} \tag{21}$$

ここで，E_{ci} は復元誤差，すなわち第 i 標本データと第 c 主部分空間との距離の 2 乗を表している．標本データ $\boldsymbol{x}_i$ の第 c 分析器への貢献度合いは，復元誤差を用いて，

$$u_{ci} = \frac{\exp(-\frac{1}{\lambda} E_{ci})}{\sum_{l=1}^c \exp(-\frac{1}{\lambda} E_{li})} \tag{22}$$

のように求められる．ただし，Hinton らの手法では，単一の擬似的な負の尤度関数の最小化に基づいてパラメータが推定されるものの，確率密度分布は定義されていない．

そこで近年，主成分分析のための確率モデルが提案されている [5, 6]．これらの潜在変量モデルでは，観測データ $\boldsymbol{x}$ がガウス分布に従う低次元の潜在変量の線形変換であると仮定されており，$\boldsymbol{x}$ もガウス型の分布で表される．したがって，その混合分布モデルを考えることで，単一の尤度関数最大化に基づいて局所的な主成分分析を議論することが可能である．Tipping らの混合主成分分析（Mixture of probabilistic PCA: MPCA） [6] では，m 次元の観測ベクトル $\boldsymbol{x}$ に対応する p 次元の潜在変量 $\boldsymbol{f}_c$ をおのおのの確率モデルで仮定し，以下の線形関係を考える．

$$\boldsymbol{x} = A_c \boldsymbol{f}_c + \boldsymbol{b}_c + \boldsymbol{\epsilon}_c \quad ; c = 1, \cdots, C \tag{23}$$

ここで，$(m \times p)$ 行列 A_c は p 本の主成分ベクトルを並べた主成分行列であり，$\boldsymbol{b}_c$ は第 c 確率モデルにおける中心を表す．また，仮定から，p 次元の零ベクトル $\mathbf{0}_p$ および $(p \times p)$ の単位行列 I_p を用いて，$\boldsymbol{f}_c \sim \mathcal{N}(\mathbf{0}_p, I_p)$ と表されるものとする．(23) 式の潜在変量モデルは，誤差のモデル $\boldsymbol{\epsilon}_c \sim \mathcal{N}(\mathbf{0}_m, R_c)$ において R_c を対角行列とした場合には因子分析のモデルとなっているが，もし $R_c = \sigma_c^2 I_m$ と制限される場合には，主成分分析と類似したモデルとなり，$\sigma_c^2 \to 0$ のときには通常の主成分分析に等しくなる．

誤差のモデルとして $\boldsymbol{\epsilon}_c \sim \mathcal{N}(\mathbf{0}_m, \sigma_c^2 I_m)$ を用いる場合，潜在変量 $\boldsymbol{f}_c$ が与えられた条件のもとでの $\boldsymbol{x}$ の条件付き確率密度分布 $p_c(\boldsymbol{x}|\boldsymbol{f}_c)$ は，

$$\boldsymbol{x}|\boldsymbol{f}_c \sim \mathcal{N}(A_c \boldsymbol{f}_c + \boldsymbol{b}_c, \sigma_c^2 I_m) \tag{24}$$

のように与えられることから，$\boldsymbol{x}$ の周辺分布 $p_c(\boldsymbol{x})$ もまたガウス分布となり，

$$\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{b}_c, W_c) \tag{25}$$

のように与えられる．ただし，$W_c = A_c A_c^\top + \sigma_c^2 I_m$ とおいた．したがって，最小化すべき負の対数尤度は，

$$L_{mpca} \quad = \quad -\sum_{i=1}^n \log \left\{ \sum_{c=1}^C \pi_c p_c(\boldsymbol{x}_i) \right\} \tag{26}$$

のように定められる．

この潜在変量モデルにおける未知パラメータは，以下の EM アルゴリズムによって推定される．

- E-step　おのおのの標本データの各クラスへの貢献度（事後確率）を推定する．

$$u_{ci} = \frac{\pi_c \exp(-\frac{1}{2}E_{ci})|W_c|^{-\frac{1}{2}}}{\sum_{l=1}^{C} \pi_l \exp(-\frac{1}{2}E_{li})|W_l|^{-\frac{1}{2}}} \tag{27}$$

ただし，

$$E_{ci} = (\boldsymbol{x}_i - \boldsymbol{b}_c)^\top W_c^{-1}(\boldsymbol{x}_i - \boldsymbol{b}_c) \tag{28}$$

である．

- M-step　各クラスにおけるパラメータを推定する．

$$\pi_c = \frac{1}{n}\sum_{i=1}^{n} u_{ci} \tag{29}$$

$$\boldsymbol{b}_c = \frac{\sum_{i=1}^{n} u_{ci}\boldsymbol{x}_i}{\sum_{i=1}^{n} u_{ci}} \tag{30}$$

$$A_c = U_{pc}(\Delta_{pc} - \sigma_c^2 I_p)^{1/2} V \tag{31}$$

$$\sigma_c^2 = \frac{1}{m-p}\sum_{j=p+1}^{m} \delta_{cj} \tag{32}$$

ここで，Δ_{pc} は Σ_c の大きな p 個の固有値 $\delta_{c1}, \cdots, \delta_{cp}$ を対角要素に持つ $(p \times p)$ の対角行列を，U_{pc} は対応する p 本の固有ベクトルを並べた $(m \times p)$ 行列を，V は $(p \times p)$ の任意の直交行列を表す．また，$\delta_{c,p+1}, \cdots, \delta_{cm}$ は Σ_c の小さな $m-p$ 個の固有値を表す．ただし，Σ_c は標本データの貢献度合いを考慮した重み付きの分散共分散行列

$$\Sigma_c = \frac{1}{\pi_c n}\sum_{i=1}^{n} u_{ci}(\boldsymbol{x}_i - \boldsymbol{b}_c)(\boldsymbol{x}_i - \boldsymbol{b}_c)^\top \tag{33}$$

である．

ここで，MPCA のモデルは，ガウス混合モデルにおいて $m(m+1)/2$ 個の未知パラメータを持つ分散共分散行列の代わりに $(m \times p + 1)$ の要素からなる $A_c A_c^\top + \sigma_c^2 I_m$ を用いたものとみなされる．したがって，MPCA のモデルはガウス混合モデルにおいてパラメータ数に制約を加えたモデルであると考えられ，モデルの複雑さが潜在変量の次元 p で調整されている．そこで，Moerland [7] は，確率密度推定の問題において混合モデルのパラメータ数の違いによる汎化能力の違いを検証するための比較実験を行い，モデルの複雑さを適当に定める必要があるものの，パラメータ数に制約を加えたモデルの方が高い汎化能力があることを示している．

3.2　エントロピー正則化を用いた線形ファジィクラスタリング法

ファジィクラスタリング法を局所的な部分空間の抽出に応用する研究も盛んに行われている．Fuzzy c-Varieties（FCV）法 [14] はクラスターのプロトタイプとして中心点の代わりに線形多様体を用いることで局所的な線形構造をとらえるための線形ファジィクラスタリング法である．FCV アルゴリズムでは，クラスタリング基準 d_{ci} として第 i 標本データ点と第 c 線形多様体との距離の 2 乗，

$$d_{ci} = ||\boldsymbol{x}_i - \boldsymbol{b}_c||^2 - \sum_{k=1}^{p} |\boldsymbol{a}_{ck}^\top(\boldsymbol{x}_i - \boldsymbol{b}_c)|^2 \tag{34}$$

が用いられる．ただし，互いに直交する p 本の単位ベクトル $\boldsymbol{a}_{ck}$ により張られる線形多様体が第 c クラスターのプロトタイプを表す．このクラスタリング基準は，Kambhatla らや Hinton らの局所的な主成分分析においてデータの帰属を定める基準である復元誤差に等しい．ここで，エントロピー正則化を用いたファジィクラスタリングのクラスタリング基準として (34) 式を代入した場合には，FCV

法の目的関数は Hinton らの "soft" なデータ分割を用いた局所的主成分分析における擬似的な負の対数尤度を λ 倍したものに等しくなる．最適性の必要条件から，最適な $\boldsymbol{a}_{ck}$ はファジィ散布行列

$$S_{fc} = \sum_{i=1}^{n} u_{ci}(\boldsymbol{x}_i - \boldsymbol{b}_c)(\boldsymbol{x}_i - \boldsymbol{b}_c)^\top \tag{35}$$

の固有値問題の解として与えられることから，これらの $\boldsymbol{a}_{ck}$ はおのおのの標本データが帰属するクラスターとその貢献度合いを考慮しながら推定される局所的な主成分ベクトルであるとみなすことができる．このことは，(34) 式のクラスタリング基準の最小化を考えた際には，第 2 項目が主成分ベクトルへの射影の長さの最大化，すなわち主成分得点の分散の最大化のための項であり，主成分分析の抽出に寄与しているととらえられることにも通じている．このように，" 正則化" の概念を導入した線形ファジィクラスタリング法と局所的な主成分分析法との間には密接な関連がある．そこで，次小節では MPCA のモデルを拡張することにより，新たな線形ファジィクラスタリング法を提案する．

3.3 K-L 情報量正則化を用いた線形ファジィクラスタリング法

Tipping らの確率的主成分分析法は潜在変量モデルに基づいた主成分分析法の拡張として定式化されていたが，その混合分布モデルを考えた場合にはガウス混合モデルの分散共分散行列のパラメータ数に制約を加えたモデルと理解することができた．一方，KFCM 法では FCM 法に K-L 情報量を用いた正則化を導入することで，ガウス混合モデルに対応するファジィクラスタリング法を提案した．そこで本小節では，KFCM 法の分散共分散行列のパラメータ数に制約を加えることにより，新たな線形ファジィクラスタリング法を提案する．(15) 式の目的関数において，$(m \times p)$ 行列 A_c を用いて，分散共分散行列を少ないパラメータからなる行列 $W_c = A_c A_c^\top + \sigma_c^2 I_m$ に置き換えることにより，パラメータ数に制約を加えた KFCM 法の目的関数は，以下のように定義できる．

$$\begin{aligned} L_{kfcv} &= \sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci} d_{ci} + \lambda \sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci} \log \frac{u_{ci}}{\pi_c} \\ &\quad + \sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci} \log |A_c A_c^\top + \sigma_c^2 I_m| \end{aligned} \tag{36}$$

ただし，d_{ci} は以下の擬似的なマハラノビス距離である．

$$d_{ci} = (\boldsymbol{x}_i - \boldsymbol{b}_c)^\top (A_c A_c^\top + \sigma_c^2 I_m)^{-1} (\boldsymbol{x}_i - \boldsymbol{b}_c) \tag{37}$$

最適性の必要条件から，メンバシップの更新則は次のように求められる．

$$u_{ci} = \frac{\pi_c \exp\left(-\frac{1}{\lambda} d_{ci}\right) |A_c A_c^\top + \sigma_c^2 I_m|^{-\frac{1}{\lambda}}}{\sum_{l=1}^{C} \pi_l \exp\left(-\frac{1}{\lambda} d_{li}\right) |A_l A_l^\top + \sigma_l^2 I_m|^{-\frac{1}{\lambda}}} \tag{38}$$

同様に，$\boldsymbol{b}_c$ および π_c は (18) 式および (19) 式を用いて更新される．つぎに，最適な A_c および σ_c^2 を求める際には，目的関数を，

$$\begin{aligned} L_{kfcv'} &= \sum_{c=1}^{C} \left\{ \sum_{i=1}^{n} u_{ci} \right\} \mathrm{tr}(W_c^{-1} \Sigma_c) \\ &\quad + \lambda \sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci} \log \frac{u_{ci}}{\pi_c} \\ &\quad + \sum_{c=1}^{C} \left\{ \sum_{i=1}^{n} u_{ci} \right\} \log |W_c| \end{aligned} \tag{39}$$

のように変形する．ここで，Σ_c は第 c クラスターにおけるファジィ分散共分散行列 [19] を表し，(33) 式と同様に求められるものとする．すべての要素が 0 の行列を O とおくと，最適性の必要条件 $\partial L_{kfcv'} / \partial A_c = O$ から，

$$-W_c^{-1} \Sigma_c W_c^{-1} A_c + W_c^{-1} A_c = O \tag{40}$$

が得られ，最適な主成分行列 A_c は文献 [6] と同様に以下のように求められる．

$$A_c = U_{pc} (\Delta_{pc} - \sigma_c^2 I_p)^{1/2} V \tag{41}$$

また，σ_c^2 も，

$$\sigma_c^2 = \frac{1}{m-p}\sum_{j=p+1}^{m}\delta_{cj} \tag{42}$$

となる．ただし，記号の定義は MPCA と同様である．このように，MPCA における M ステップと等しい更新式となっており，最適な A_c は Σ_c の大きな固有値に対応する固有ベクトルを並べることで得られる．

ここで，上記の修正された KFCM アルゴリズムは，$\lambda = 2$ のときには MPCA における EM アルゴリズムに一致するが，$\lambda \neq 2$ のときには対応する確率モデルは存在しない．したがって，提案手法は確率モデルに基づく手法ではなく，ファジィモデリングの一種であり，λ によりモデルのファジィ度を調節している．

つぎに，エントロピー正則化を用いた FCV 法と提案法との関係について考察する．3.1 節の記号の定義に従うと，(41) 式から，行列 W_c は，

$$\begin{aligned} W_c &= U_{pc}(\Delta_{pc} - \sigma_c^2 I_p)U_{pc}^\top + U_{mc}(\sigma_c^2 I_m)U_{mc}^\top \\ &= U_{mc}\Delta_{pc}^* U_{mc}^\top \end{aligned} \tag{43}$$

と表される．ただし，

$$\Delta_{pc}^* = \begin{bmatrix} \Delta_{pc} & O \\ O & \sigma_c^2 I_{m-p} \end{bmatrix} \tag{44}$$

とおいた．このように，W_c は分散共分散行列の固有値分解を p 個の固有値と σ_c^2 を用いて近似した行列となっている．また，$p = m-1$ のときは，W_c は分散共分散行列の固有値分解に等しく，提案法は KFCM 法に一致する．逆行列 W_c^{-1} は，

$$W_c^{-1} = U_{mc}(\Delta_{pc}^*)^{-1}U_{mc}^\top \tag{45}$$

となり，局所的な主成分ベクトル $\boldsymbol{a}_{ck}$ を用いて，d_{ci} は，

$$\begin{aligned} d_{ci} &= (\boldsymbol{x}_i - \boldsymbol{b}_c)^\top U_{mc}(\Delta_{pc}^*)^{-1}U_{mc}^\top(\boldsymbol{x}_i - \boldsymbol{b}_c) \\ &= \sum_{k=1}^{p}\frac{1}{\delta_k}|\boldsymbol{a}_{ck}^\top(\boldsymbol{x}_i - \boldsymbol{b}_c)|^2 \\ &\quad + \sum_{k=p+1}^{m}\frac{1}{\sigma_c^2}|\boldsymbol{a}_{ck}^\top(\boldsymbol{x}_i - \boldsymbol{b}_c)|^2 \\ &= \frac{1}{\sigma_c^2}\Big(\sum_{k=1}^{p}\frac{\sigma_c^2}{\delta_k}|\boldsymbol{a}_{ck}^\top(\boldsymbol{x}_i - \boldsymbol{b}_c)|^2 \\ &\quad + \sum_{k=p+1}^{m}|\boldsymbol{a}_{ck}^\top(\boldsymbol{x}_i - \boldsymbol{b}_c)|^2\Big) \end{aligned} \tag{46}$$

と表される．ここで，誤差のモデルにおいて $\sigma_c^2 \to 0$ とした場合には，$\sigma_c^2/\delta_k \to 0$ となるので，

$$\begin{aligned} d_{ci} &\cong \frac{1}{\sigma_c^2}\sum_{k=p+1}^{m}|\boldsymbol{a}_{ck}^\top(\boldsymbol{x}_i - \boldsymbol{b}_c)|^2 \\ &= \frac{1}{\sigma_c^2}\big(||\boldsymbol{x}_i - \boldsymbol{b}_c||^2 - \sum_{k=1}^{p}|\boldsymbol{a}_{ck}^\top(\boldsymbol{x}_i - \boldsymbol{b}_c)|^2\big) \end{aligned} \tag{47}$$

のように，FCV 法と同様の線形多様体からの距離のみを重視するクラスタリング基準と考えられるようになる．また，行列式も，

$$\begin{aligned} |A_c A_c^\top + \sigma_c^2 I_m| &= |U_{mc}\Delta_{pc}^* U_{mc}^\top| \\ &= |U_{mc}^\top U_{mc}||\Delta_{pc}^*| \\ &= |\Delta_{pc}| \times (\sigma_c^2)^{m-p} \to 0 \end{aligned} \tag{48}$$

となるので，目的関数の大小に影響を与えないものと考えて，暫時 (36) 式から削除し，クラスター容量を $\pi_c = 1/C$（定数）と固定すると，(38) 式のメンバシップの更新則は，

$$u_{ci} = \frac{\exp\left(-\frac{1}{\lambda\sigma_c^2}d_{ci}^*\right)}{\sum_{l=1}^{C}\exp\left(-\frac{1}{\lambda\sigma_l^2}d_{li}^*\right)} \tag{49}$$

となり，エントロピー正則化を用いた FCV 法において $\lambda \to 0$ としたモデル，すなわち，メンバシップのファジィ度を極端に小さくしてハードクラスタリングに近づけたモデルに類似する．ただし，

$$d_{ci}^* = ||\boldsymbol{x}_i - \boldsymbol{b}_c||^2 - \sum_{k=1}^{p}|\boldsymbol{a}_{ck}^\top(\boldsymbol{x}_i - \boldsymbol{b}_c)|^2 \tag{50}$$

とおいた．したがって，すべてのデータ点がプロトタイプとなる線形多様体上にほぼ乗っており，なおかつクラスター容量に差がない場合には，提案のアルゴリズムはエントロピー正則化を用いる FCV 法においてファジィ度を小さくしたモデルに類似するといえる．以上の考察は，Tipping らの確率的主成分分析法が $\sigma_c^2 \to 0$ の場合には通常の主成分分析を表すようになることと対応している．また，クラスタリング手法としての意味合いからは，A_c のおのおのの列ベクトルは，FCV 法と同様に，プロトタイプとなる線形多様体を張るベクトルととらえられる．一方，σ_c^2 は線形多様体上では表すことができないデータの広がりの度合いを示しており，楕円体状のクラスターにおける短軸方向の広がりと解釈できることから，Adaptive Fuzzy c-Elliptotypes（AFC）法 [20] と同様の適応的にクラスター形状を変化させる機能を担っていると考えられる．ただし，AFC 法ではクラスター形状を適応的に決定する際に目的関数が必ずしも最適化されていないという問題点がある [21] が，提案法ではすべてのパラメータが (36) 式の最小化に基づいて更新されており，単一の目的関数の最小化の下でモデルが同定されている．

以上から，提案手法はエントロピー正則化の代わりに K-L 情報量正則化を導入することで，FCV 法を形状調整機能を持つ分析手法に拡張したものであるととらえられる．その観点から，本論文では提案手法を K-L 情報量正則化を用いた FCV（KFCV）法と呼ぶこととする．

以下に KFCV 法のアルゴリズムを記す．

K-L 情報量正則化を用いた Fuzzy c-Varieties（KFCV）アルゴリズム

Step 1 乱数により u_{ci} を初期化し，“確率的制約” を満たすように基準化する．

Step 2 (18) 式を用いて $\boldsymbol{b}_c$ を更新する．

Step 3 (19) 式を用いて π_c を更新する．

Step 4 クラスターごとにファジィ分散共分散行列の固有値問題を解き，A_c および σ_c^2 を求める．

Step 5 (38) 式を用いて u_{ci} を更新する．

Step 6 小さな正数 ξ に対して，収束判定条件

$$\max_{c,i} \mid u_{ci}^{NEW} - u_{ci}^{OLD} \mid < \xi$$

を満たせば終了．それ以外は Step 2 へ戻る．

4 数値実験

4.1 データ分割法としての確率モデルとファジィクラスタリング法の比較

まず，人工データを用いて，最尤推定法に基づく確率モデルと目的関数最小化に基づくファジィクラスタリング法との比較を行った．用いたデータは図 1 のデータであり，右下がりの直線状に分布した標本データと右上がりの長方形状に分布した標本データからなっている．直線状に分布したデータの方が数が少ないのに対して，長方形状のデータは数が多く，なおかつ中心線からのデータのばらつきが大きい．このデータ集合に $\theta = 2.0$ として Bezdek らの標準的な FCV 法を適用した場合には，メンバシップが最大となるクラスターに分類した結果，図 2 の○と×に分類された．このように，クラスターごとの容量や形状の違いを考慮しない線形ファジィクラスタリング法では，データ集合ごとの特徴をとらえることができない．そこで，クラスターごとの容量や形状の違いを考慮しながら二つのデータ集合をとらえる実験を，MPCA 法および KFCV 法を用いて行った．おのおのの分析手法で得られたデータ分割の結果を図 3 に示す．ただし，クラスター数は 2 とし，潜在変量の次元は $p = 1$,

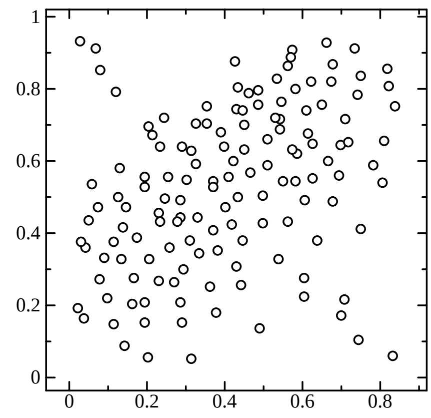


図 1: 実験に用いた人工データ

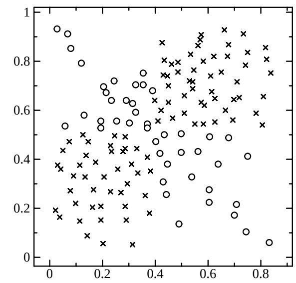
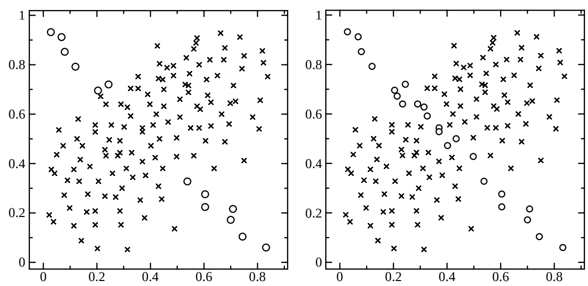


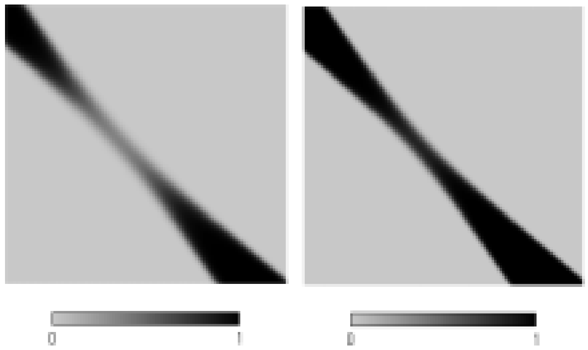
図 2: 標準的な FCV 法（$\theta = 2$）による分類結果



(a) MPCA　　(b) KFCV（$\lambda = 1$）

図 3: MPCA および KFCV（$\lambda = 1$）による分類結果

KFCV 法におけるファジィ度を決めるパラメータは $\lambda = 1$ とした．メンバシップ（事後確率）が最大となるクラスター（クラス）に分類した結果，標本データは○と×に分類され，容量 π_c は MPCA では $\pi_1 = 0.1, \pi_2 = 0.9$，KFCV 法では $\pi_1 = 0.15, \pi_2 = 0.85$ となった．いずれの手法でも容量の差を考慮した分析が行われているが，長方形状に分布したデータの数が圧倒的に多いために，MPCA 法では二つのデータ集合が重なっている領域が×に分類されている．一方，KFCV 法でファジィ度を小さくした場合には，直線状に分布したデータを強調した分類結果が得られている．この結果をより詳しく分析するために，分類関数として用いられる (27) 式および (38) 式の値を比較した．○で表されるクラスター（クラス）の分類関数のデータ領域における値の比較を図 4 に示す．$\lambda = 1$ としてファジィ度を小さくしているために，メンバシップが 0 か 1 をとるクリスプな分割に近づいており，二つのデータ集合の違いを強調した分類関数となっていることが分かる．したがって，ファジィ度を調節することで，提案手法では確率モデルが与えるものとは異なる柔軟なデータ分割を得られることが分かる．このような性質は，データ分類においてパラメータを調整することで分析者の意思を反映させた分類関数を作成し得る可能性を示しているといえる．ただし，ここで用いた 2 次元データの例では，行列 W_c が分散共分散行列に等しくなり，MPCA および KFCV 法で得られる結果はそれぞれ GMMs および KFCM 法の結果と同じものとなっているため，上記の性質は KFCM 法および KFCV 法に共通したものである．



(a) MPCA　　　　(b) KFCV（$\lambda = 1$）

図 4: 分類関数の値の比較

表 1: 最適解が得られた頻度の比較

潜在変量の次元	MPCA	KFCV 法
10	97	99
15	76	80
20	23	40
25	14	26

4.2　アニーリングによる初期値依存の改善

一般に混合分布モデルの同定においては，結果が初期値に影響されやすいという欠点がある．宮岸ら [13] は目的関数の最小化に基づく KFCM 法の利点を利用して，ファジィ度を決定する係数の決定論的アニーリングにより初期値依存を改善する方法を提案している．係数 λ の値を温度とみなして，反復アルゴリズムの中で減少させることで，分析の初期段階ではおのおののデータ点の所属をあいまいとし，分析が進むにつれてクリスプな分割に近づくようになることから，初期分割の影響を減らして局所解に収束することを避けられる．そこで本小節では，KFCV 法においてファジィ度を決定する係数のアニーリングを行うことにより，MPCA の初期値依存を改善することを試みた．実験には UCI Machine Learning Repository [22] で公開されている電離層データを用いた．2 クラスのデータで構成されるこのデータ集合は，おのおの 34 個の観測値からなる 351 個の標本データを含むもので，線形のパーセプトロンを用いて 90%以上の正確さで分類が可能であることが報告されている [23]．したがって，多次元空間におのおののクラスが二つの主要な塊を形成していると考えられることから，本実験ではクラス情報を用いずにそれらの二つの塊の特徴をとらえることを試みた．ただし，34 個の観測値のうち，名義データを除いた 32 次元の数値データのみを使用した．MPCA および KFCV 法を用いて初期分割を乱数で変化させながら行った 100 回の試行において，目的関数最小の解が得られた頻度を表 1 に示す．KFCV 法ではファジィ度を決める係数 λ のアニーリングを行い，アニーリングスケジュールとしては，

$$\lambda(t) = \lambda^* / \ln(2 + t) \tag{51}$$

を使用した．ただし，t は繰り返し回数であり，λ^* を 8 とした．また，λ は 2 になれば固定するように設定した．したがって，分析の初期段階では $\lambda > 2$ となり対応する確率モデルは存在しないが，$\lambda = 2$ と固定された後は MPCA に対応するモデルとなっており，得られる分類結果は確率モデルである MPCA と同等となっている．データの次元が大きいにもかかわらず標本データ数があまり多くないために，潜在変量の数，すなわち未知パラメータの数が多くなるにつれて局所解に陥る頻度が高くなっているが，アニーリングを行うことで初期値への依存度を抑えることができている．このように，(3) 式の制約を用いない提案法の柔軟性を利用することで，初期値依存を改善することができた．

4.3　潜在変量の次元を変化させた場合の分割結果の妥当性の比較

前記の数値例では，確率モデルに基づく MPCA と KFCV 法との比較を行ったが，KFCV 法は KFCM 法にパラメータ数の制約を加えたモデルともとらえられることから，本小節では，KFCM 法

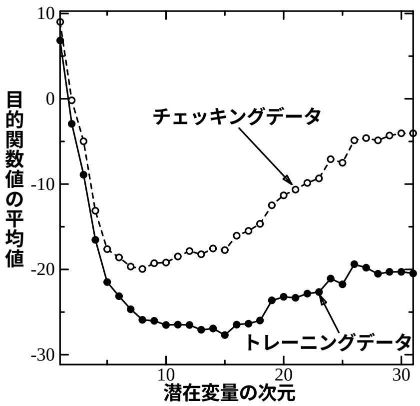


図 5: 潜在変量の次元の違いによる目的関数値の変化

のパラメータ数を減少させる効果を検証する．4.2 節に引き続いて，実験には電離層データのうちの 32 次元の数値データのみを用いた．データ分割結果の妥当性を検証するために，5-fold cross-validation 法を用いて，トレーニングセットから推定された (38) 式のファジィ分類関数によりチェッキングセットでのメンバシップを算出した際の目的関数値を調べた．目的関数値（(36) 式）の標本データ 1 件あたりの平均値の変化を，図 5 に示す．ただし，なるべく局所解に陥ることを避けるために，4.2 節で提案した λ のアニーリングを用いた．ここで，潜在変量が 31 次元の場合は，行列 W_c が分散共分散行列に一致し，KFCM 法の結果に対応している．図 5 から，本実験では潜在変量の次元を 8 としたときがチェッキングセットの目的関数値が最小となっており，次元を大きくするにしたがってトレーニングセットとの乖離が大きくなる傾向，すなわちトレーニングセットへの過学習が見られた．また，次元が大きい場合にはトレーニングセットについても目的関数値が大きくなっているが，これは 4.2 節で取り上げた局所解に陥った試行があったことに起因している．このようにファジィクラスタリングにおいても，標本データ数の少ない多次元データ集合を取り扱う際には，未知パラメータ数を適当に定めることでより適切にデータ集合の局所的な特徴を抽出できることが分かった．

5　おわりに

本論文では，局所的な主成分分析手法と線形ファジィクラスタリング手法との類似性に着目し，確率モデルに基づく混合主成分分析法（MPCA）に類似したアルゴリズムが導かれる新たな線形ファジィクラスタリング法を提案した．提案手法はクラスターのプロトタイプとして線形多様体を用いる FCV 法にクラスター容量と形状調整用のパラメータを導入したモデルであり，K-L 情報量正則化を用いる FCM（KFCM）法にパラメータ数の制約を加えたモデルともみなされる．数値実験では MPCA との比較を通して目的関数の最小化に基づく提案法の柔軟性を示したのちに，KFCM 法の未知パラメータ数を減少させることでより適切なデータ分割モデルが得られることを議論した．ただし，潜在変量の次元をどのように定めればよいかは今後の課題である．

参考文献

[1] K. Fukunaga and D. R. Olsen: An Algorithm for Finding Intrinsic Dimensionality of Data; *IEEE Trans. Computers*, Vol. C-20, pp. 176-183 (1971)

[2] N. Kambhatla and T. K. Leen: Dimension Reduction by Local Principal Component Analysis; *Neural Computation*, Vol. 9, No. 7, pp. 1493-1516 (1997)

[3] G. E. Hinton, P. Dayan and M. Revow: Modeling the Manifolds of Images of Handwritten Digits; *IEEE Trans. Neural Networks*, Vol. 8, No. 1, pp. 65-74 (1997)

[4] A. P. Dempster, N.M. Laird and D.B. Rubin: Maximum Likelihood from Incomplete Data via the EM Algorithm; *J. of the Royal Statistical Society*, Series B, Vol. 39, pp. 1–38 (1977)

[5] S. Roweis: EM Algorithms for PCA and SPCA; *Advances in Neural Information Processing Systems 10*, Eds. M. I. Jordan, M. J. Kearns and S. A. Solla, MIT Press, pp. 626–632 (1998)
[6] M. E. Tipping and C. M. Bishop: Mixtures of Probabilistic Principal Component Analysers; *Neural Computation*, Vol. 11, No. 2, pp. 443-482 (1999)
[7] P. Moerland: A Comparison of Mixture Models for Density Estimation; *Proc. of 9th Int. Conf. Artificial Neural Networks (ICANN'99)*, Vol. 1, pp. 25-30 (1999)
[8] J. C. Bezdek: *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press (1981)
[9] D. E. Gustafson and W. C. Kessel: Fuzzy Clustering with a Fuzzy Covariance Matrix; *Proc. of the IEEE Conf. Decision and Control*, Vol. 2, pp. 761-766 (1979)
[10] I. Gath and A. B. Geva: Unsupervised Optimal Fuzzy Clustering; *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, pp. 773-781 (1989)
[11] 宮本 定明, 馬屋原 一孝, 向殿 政男: ファジィc-平均法とエントロピー正則化法におけるファジィ分類関数; 日本ファジィ学会誌, Vol. 10, No. 3, pp.156–164 (1998)
[12] 赤穂 昭太郎: EM アルゴリズムークラスタリングへの適用と最近の発展ー; 日本ファジィ学会誌, Vol. 12, No. 5, pp.594-602, (2000)
[13] 宮岸 聖高, 市橋 秀友, 本多 克宏: K-L 情報量正則化 FCM クラスタリング法; 日本ファジィ学会誌, Vol. 13, No. 4, pp. 406-417 (2001)
[14] J. C. Bezdek, C. Coray, R. Gunderson and J. Watson: Detection and Characterization of Cluster Substructure 2. Fuzzy c-Varieties and Convex Combinations Thereof; *SIAM J. Appl. Math.*, Vol. 40, No. 2, pp 358-372 (1981)
[15] R. O. Duda and P. E. Hart: *Pattern Classification and Scene Analysis*, Jhon Wiley & Sons (1973)
[16] C. M. Bishop: *Neural Networks for Pattern Recognition*, Clarendon Press (1995)
[17] F. Höppner, F. Klawonn, R. Kruse and T. Runkler: *Fuzzy Cluster Analysis*, Jhon Wiley & Sons (1999)
[18] 宮本 定明: クラスター分析入門, 森北出版 (1999)
[19] Y. Yabuuchi and J. Watada: Fuzzy Principal Component Analysis and Its Application; *Biomedical Fuzzy and Human Sciences*, Vol. 3, No. 1, pp. 83-92 (1997)
[20] R. N. Dave: An Adaptive Fuzzy c-Elliptotype Clustering Algorithm; *Proc. of the North American Fuzzy Information Processing Society: Quater Century of Fuzziness*, Vol. 1, pp. 9–12 (1990)
[21] 馬屋原 一孝, 宮本 定明: 次元係数の正則化による線形ファジィクラスタリング; 日本ファジィ学会誌, Vol. 12, No. 4, pp. 552-561 (2000)
[22] P. M. Murphy and D. W. Aha: *UCI Repository of machine learning databases*, http://www.ics.uci.edu/˜mlearn/MLRepository.html. Irvine, CA: University of California, Department of Information and Computer Science (1994)
[23] V. G. Sigillito, S. P. Wing, L. V. Hutton and K. B. Baker: Classification of Radar Returns from the Ionosphere Using Neural Networks; *Johns Hopkins APL Technical Digest*, Vol. 10, pp. 262-266 (1989)

［問い合わせ先］
〒 599-8531　大阪府堺市学園町 1-1
大阪府立大学大学院工学研究科
電気・情報系専攻経営工学分野
本多 克宏
TEL : 072-254-9355
FAX : 072-254-9915
E–mail: honda@ie.osakafu-u.ac.jp

著者略歴

本多　克宏　(ほんだ　かつひろ) [正会員]
1999年大阪府立大学大学院工学研究科博士前期課程電気・情報系専攻修了．同年日本電信電話（株）入社，同年大阪府立大学工学部経営工学科助手，現在に至る．ニューラルネットワーク，ファジィクラスタリングの研究に従事．IEEE，日本ファジィ学会，システム制御情報学会，日本経営工学会の会員．

神田　章裕　(かんだ　あきひろ)
2003年大阪府立大学大学院工学研究科博士前期課程電気・情報系専攻修了．同年日本生命保険相互会社入社．在学中はファジィクラスタリングの研究に従事．

市橋　秀友　(いちはし　ひでとも) [正会員]
1971年大阪府立大学工学部経営工学科卒業．同年松下電器産業（株）入社，1981年大阪府立大学工学部経営工学科助手，1987年同講師，1989年同助教授，1993年同教授，現在に至る．工学博士．ファジィクラスタリングやニューラルネットワークなどでのデータ解析法，その知的システムや人間機械システムへの応用研究に従事．IEEE，日本ファジィ学会，システム制御情報学会，電子情報通信学会，日本経営工学会などの会員．