

第I部

正則化によるファジィクラスタリング

第3章

fuzzy c -means 法とエントロピー正則化法

本章では、fuzzy c -means 法における新しい手法としてエントロピー関数を正則化項としてもつエントロピー正則化法を提案する。また、ファジィ分類関数を導入し、従来の fuzzy c -means 法とエントロピー正則化法により生成されるファジィ分類関数を比較する。エントロピーにもとづく方法として、エントロピー最大化法が既に提案されているが、正則化の概念を導入することによって、エントロピーを用いる方法を、一般的な fuzzy c -means 法の交互最適化アルゴリズムの枠において議論できるようになる。このことは、fuzzy c -varieties 法などの fuzzy c -means 法の変形をエントロピー正則化を用いて行うことができることを意味している。また、従来の fuzzy c -means 法から生成されるファジィプロトタイプ分類関数に対応するプロトタイプ分類関数がエントロピー正則化から得られる。これら2種類の分類関数の理論的性質を調べ、計算幾何学における Voronoi 図との関連を明らかにする。数値例によって、これら2つの方法によるクラスタリング結果とファジィ分類関数を比較する。

3.1 はじめに

最近、従来の fuzzy c -means 法とは異なるアイデアにもとづくエントロピー最大化法が李と向殿 [14, 15] によって提案されている。ここではこの方法にまず注目する。そして、このエ

エントロピー最大化法は正則化の概念を用いることによって crisp k -means 法の fuzzy c-means 法とは異なるファジィ化であり, fuzzy c-means 法の2段階アルゴリズムによって考えることができることを示す。

3章の目的は, エントロピー関数を用いる方法の意味について考察を進め, fuzzy c-means 法とこのエントロピー法との方法を比較することによって, 両者の特徴を明らかにすることである。議論は数値例にもとづいて行うのではなく, 理論的な性質を明らかにするため, 得られた結果は一般性をもっている。

これらの考察を行うにあたって2つの要点がある。1つめは, エントロピー最大化法をエントロピー正則化法として再定式化する点である。正則化[13]によって, エントロピー関数を用いる方法は fuzzy c-means 法の交互最適化アルゴリズムの枠組みで議論することができ, 様々な変形が可能となる。第2の要点となるのは, fuzzy c-means 法におけるメンバシップを決定するファジィ分類関数である。この分類関数はファジィクラスタリングにかかわる従来の研究でも暗に言及されてきたが[33], この関数の性質を詳細に調べたものはないと思われる。上述の2つの方法についてそれぞれ定められる分類関数は, 異なるファジィ判別の方法を与える。

また, fuzzy c-means 法とエントロピー正則化法の分類関数を比較するにあたって, 計算幾何学で研究されてきた Voronoi 図[34]に言及することによって, 両者の共通点と相違を明確にすることができる。

なお, これら2つの方法は, ともにクリスプな k -means のファジィ化であるという意味で, 広義の fuzzy c-means 法であるといえる。しかしながら, 便宜上ここでは, 従来の標準的な方法を fuzzy c-means 法と呼び, 後の方法をエントロピー正則化法あるいはエントロピー法と呼ぶことにする。

3.2 エントロピー最大化法

エントロピー最大化法は適切でない問題 (ill-posed problems) を解くための有効な手段として, スペクトル解析や画像復元などにおいて用いられている [35]。李と向殿 [14, 15] はこ

の概念にもとづいて従来の fuzzy c-means 法とは異なるファジィクラスタリングの方法を提案した。ここでは、2.2節で述べた FCM アルゴリズムをさらに一般化した次の手続きが用いられている。

エントロピー最大化法による FC アルゴリズム

- FC1. \bar{U} と \bar{V} の初期値を適当に決める。
- FC2. ある基準によって V を求め、解を \bar{V} とする。
- FC3. ある基準によって U を求め、解を \bar{U} とする。
- FC4. 解 (\bar{U}, \bar{V}) が収束していれば終了。そうでなければステップ FC2 に戻る。

エントロピー最大化法による FC3 では

$$\begin{aligned} \max \quad & - \sum_{i=1}^c \sum_{k=1}^n u_{ik} \log u_{ik} \\ \text{subject to} \quad & \sum_{i=1}^c u_{ik} = 1, \quad k = 1, 2, \dots, n, \\ & \sum_{i=1}^c \sum_{k=1}^n u_{ik} d_{ik} = L \end{aligned}$$

の最適解を $\bar{U} = (\bar{u}_{ik})$ としている。また、FC2 では

$$\bar{v}_i = \frac{\sum_{k=1}^n \bar{u}_{ik} \mathbf{x}_k}{\sum_{k=1}^n \bar{u}_{ik}}$$

によって解を求めている。

ここで問題となるのは、このアルゴリズムにおいて、FC2 と FC3 で異なる基準が用いられていることである。2.2節の FCM では、交互最適化 FCM2 と FCM3 において同一の目的関数が用いられているため、目的関数の値は単調に減少するのに対し、基準が異なる場合には、目的関数が減少する保証がない。このように、エントロピー最大化によるアルゴリズムはややアド・ホックであると批判される余地がある。

しかしながら、正則化の概念 [13] を導入することによって、エントロピー関数を利用する方法を交互最適化アルゴリズム FCM の枠組みでとらえることが可能である。そこで、次に正則化によるエントロピー法の定式化を考える。

3.3 エントロピー正則化法

正則化の概念は関数方程式の「適切でない問題 (ill-posed problem)」を解くための手法 [13] として、従来より多くの現実的問題に用いられてきた。一般に正則化とは、ある意味で特異な問題から正則なものへの修正を意味する。特異な問題は解くことが難しいが、正則なものは扱いやすい。正則化された問題の解が元の問題の解の近似となっているとき、その問題を元の問題の正則化という。

クラスタリングの場合 crisp k -means 法は通常の意味での「適切でない問題」とはいえないが、crisp 解をあたかも特異であるかのようにみなし fuzzy 解を正則であるように考えることによって、fuzzy c -means 法は crisp k -means 法の正則化であると考えられることができる。c-means 法の定式化において、メンバシップがファジィに一般化されたとき、クリस्प解は許容領域の端点によって特徴付けられる [10]。したがって端点を特異点と考え、中間の点を正則な解と考えるのである。

この正則化の考え方に基づけば、Dunn [11, 12] や Bezdek [10] による fuzzy c -means 法以外の正則化によるファジィクラスタリングを考えることが可能になる。

具体的には、ファジィメンバシップ値の許容集合 M_{fcm} を仮定したとき、2 値解を極値的、つまり特異であると考え、fuzzy c -means 法では、正則化パラメータは q であり、目的関数はこのパラメータによって解がファジィ化されるよう変形されている。また $q \rightarrow +1$ となるとき、クラスタはクリस्पなクラスタに近づくことが Bezdek [10] によって示されている。したがってここでは、正則化された関数 J_{fcm} は元の関数 J_{ckm} に似ていて、正則化された問題の解は元の解の近似と考えることができるのである。

それでは次に、目的関数として、エントロピーにより正則化された関数を導入しよう。典型的な正則化では、ある目的関数 J の最適化のかわりに、正則化パラメータ α と正則化項 K を用いて正則化された目的関数

$$\hat{J} = J + \alpha K$$

の最適化をその近似と考える。ここでは、

$$J = \sum_{i=1}^c \sum_{k=1}^n u_{ik} d_{ik},$$

$$K = \sum_{k=1}^n \sum_{i=1}^c u_{ik} \log u_{ik}$$

として、正則化された目的関数は次式で与えられる。

$$J_\lambda(U, V) = \sum_{i=1}^c \sum_{k=1}^n u_{ik} d_{ik} + \lambda^{-1} \sum_{k=1}^n \sum_{i=1}^c u_{ik} \log u_{ik} \quad (3.1)$$

ただし、今後の計算の便宜のために、 $\alpha = \lambda^{-1} (\geq 0)$ と置いている。このパラメータが $\lambda \rightarrow +\infty$ のとき、得られるクラスタはクリスプなクラスタに近づく。これらについては、補足2を参照されたい。この目的関数をアルゴリズム FCM の評価関数として用いる方法をエントロピー関数による正則化と呼ぶ。

エントロピー正則化による FCM2 の解は

$$\bar{v}_i = \frac{\sum_{k=1}^n \bar{u}_{ik} \boldsymbol{x}_k}{\sum_{k=1}^n \bar{u}_{ik}} \quad (3.2)$$

であることは (3.1) の第2項 $\lambda^{-1} \sum_{k=1}^n \sum_{i=1}^c u_{ik} \log u_{ik}$ が V に依存しないことから容易にわかる。また、FCM3 の解は

$$\bar{u}_{ik} = \frac{e^{-\lambda d_{ik}}}{\sum_{j=1}^c e^{-\lambda d_{jk}}} \quad (3.3)$$

と計算される。なお、計算の詳細は補足1に示す。

目的関数 (3.1) の最適化により得られた (3.2), (3.3) は李と向殿によるエントロピー最大化法の最適解と同一の形をもっている [14, 15]。従って、正則化法は、一見、エントロピー関数を用いる方法に新たなアルゴリズムをもたらしてはいないように見える。しかしながら、エントロピーを用いる方法がアルゴリズム FCM の交互最適化の枠組みで厳密に定式化されるということが示されたため、fuzzy c-varieties 法など、fuzzy c-means 法の変形として交互最適化のなかで開発されてきたアルゴリズムを、エントロピーを用いて新たに考察し直すことができるということがわかる。

fuzzy c-means 法の変形は多岐にわたるが、それらに対応する新たなアルゴリズムがすべて検討の対象となる。ここでは、これらの変形について詳細に述べることはしないが、後の章で簡単に議論を行う。

3.4 fuzzy c-means 法とエントロピー法の解の比較

fuzzy c-means 法とエントロピー法では、解の形が大きく異なるように見える。fuzzy c-means 法の u_{ik} に関する最適解は、

$$u_{ik} = \left[\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{1}{q-1}} \right]^{-1} \quad (3.4)$$

で与えられるが、これをを变形すると、(3.3) と類似した形になることがわかる。いま、(3.4) において $\beta = \frac{1}{q-1} > 0$ とおけば、 $a^z = e^{z \log a}$ より、

$$\begin{aligned} u_{ik} &= \left[\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{\beta} \right]^{-1} \\ &= \left[\sum_{j=1}^c e^{\beta(\log d_{ik} - \log d_{jk})} \right]^{-1} \\ &= \frac{e^{-\beta \log d_{ik}}}{\sum_{j=1}^c e^{-\beta \log d_{jk}}} \end{aligned} \quad (3.5)$$

すなわち、(3.3) と (3.5) は、 $\lambda = \beta$ と対応させ、 $\log d_{jk}$ を d_{jk} で置き換えれば同一の形となる。

3.5 エントロピー正則化法によるいくつかのアルゴリズム

ここで報告した正則化の概念によれば、従来試みられた研究に並行した形で、次の考察が可能であることがわかる。

- (I) FCMの交互最適化アルゴリズムにもとづく fuzzy c -varieties 法[10]や fuzzy c -regression Models[36, 37]などの方法をエントロピー正則化によって行うことができる。
- (II) エントロピー正則化によって、ユークリッド空間以外の空間、たとえば、 L_1 空間における fuzzy c -means 法に対応するアルゴリズムを得ることができる。
- (III) ファジィk-最近隣法やファジィプロトタイプ法 [33]などのファジィ分類の方法が、新たにエントロピー正則化から得られる。

これらのうち、(II)については、別の機会にゆずり、(III)については3.6節で述べる。ここでは、(I)のうち、エントロピー法から得られる fuzzy c -varieties 法について簡単に紹介しよう。簡単のため、線形多様体としては1次元のものに限って述べる。

クラスタ i を示す1次元多様体、すなわち直線はスカラー変数 α によって $v_i + \alpha s_i$ と表されると仮定する。ここで、 v_i はクラスタ i の中心で、 s_i は直線の方法を示す単位ベクトルである。このとき、ユークリッド距離の代わりに点と多様体との間の距離

$$D_{ik} = \|x_k - v_i\|^2 - \langle x_k - v_i, s_i \rangle^2$$

を用い、目的関数として次式を用いる。ただし、 $V = (v_1, v_2, \dots, v_c)$ 、 $S = (s_1, s_2, \dots, s_c)$ とする。

$$J(U, V, S) = \sum_{i=1}^c \sum_{k=1}^n u_{ik} D_{ik} + \lambda^{-1} \sum_{k=1}^n \sum_{i=1}^c u_{ik} \log u_{ik} \quad (3.6)$$

従来の fuzzy c -varieties 法に並行する議論によって次のアルゴリズムが得られる。

エントロピー正則化法による FCV アルゴリズム

- FCV1. $\bar{U}, \bar{V}, \bar{S}$ の初期値を設定する。
- FCV2. $J(\bar{U}, V, \bar{S})$ を V について最適化し、最適解を \bar{V} とする。
- FCV3. $J(\bar{U}, \bar{V}, S)$ を S について最適化し、最適解を \bar{S} とする。
- FCV4. $J(U, \bar{V}, \bar{S})$ を U について最適化し、最適解を \bar{U} とする。
- FCV5. 収束していれば終了、そうでなければ FCV2 に戻る。

ここで、FCV2における最適解は

$$\bar{v}_i = \frac{\sum_{k=1}^n \bar{u}_{ik} x_k}{\sum_{k=1}^n \bar{u}_{ik}}$$

FCV3では、 \bar{s}_i を次式で与えられる行列の最大固有値に対応する固有ベクトルとする。ただし、 \bar{s}_i は正規化するものとする。

$$A^i = \sum_{\ell=1}^n \bar{u}_{i\ell} (x_\ell - \bar{v}_i)(x_\ell - \bar{v}_i)^T$$

FCV4では、

$$\bar{u}_{ik} = \frac{e^{-\lambda D_{i,k}}}{\sum_{j=1}^c e^{-\lambda D_{j,k}}}$$

これらの式の導出の詳細は省略する。従来のFCVに関する議論と、エントロピー法に関する議論とを組み合わせることによって容易に行われるからである。

3.6 ファジィ分類関数

クラスタリングは外的基準なしにクラス分けを行う方法であるが、これに対して、既に存在する基準のもとで、新たに与えられた個体を分類する方法を判別あるいは教師付き分類と呼ぶ。この章では、クラスタリングによって生成されたファジィなクラスを所与のクラスと見なしたとき、新たに観測されたデータ点をどのように分類するのが合理的であるかを考察する。

Keller [33]はファジィプロトタイプ分類とファジィk最近隣分類の方法を提案している。すなわち、ファジィプロトタイプ分類として、プロトタイプベクトル z_1, z_2, \dots, z_c が与えられたとき、観測 x をクラス i に割り当てるファジィ分類関数 B_i を

$$B_i(x) = \left[\sum_{j=1}^c \left(\frac{\|x - z_i\|^2}{\|x - z_j\|^2} \right)^{\frac{1}{q-1}} \right]^{-1} \quad (3.7)$$

(ただし、 $\mathbf{x} = \mathbf{z}_i$ のとき、 $B_i(\mathbf{x}) = 1, B_j(\mathbf{x}) = 0$ for all $j \neq i$) で定義している。また、ファジィ k 最近隣分類関数 C_i は、既に存在する個体の c 個のファジィクラスが与えられたとして、観測 \mathbf{x} に最も近いほうから順に個体 $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k$ をとるとき、次式で定義される。

$$C_i(\mathbf{x}) = \sum_{\ell=1}^k u_{i\ell} \left[\sum_{j=1}^k \left(\frac{\|\mathbf{x} - \mathbf{y}_\ell\|^2}{\|\mathbf{x} - \mathbf{y}_j\|^2} \right)^{\frac{1}{q-1}} \right]^{-1} \quad (3.8)$$

ここで、 $u_{i\ell}$ は個体 \mathbf{y}_ℓ がクラス i に分類されるメンバシップで、ファジィ分割 $\sum_{i=1}^c u_{i\ell} = 1$ を満たすものとする。

これからすぐわかるのは、 $B_i(\mathbf{x})$ が (3.4) の右辺の d_{ik} を $\|\mathbf{x} - \mathbf{z}_i\|^2$ で置き換えた式で定義されていることである。また、 $C_i(\mathbf{x})$ についても、(3.4) の形を利用していることが明らかである。このことより、エントロピー法を上記の議論にあてはめることができ、(3.3) の右辺の d_{ik} を $\|\mathbf{x} - \mathbf{z}_i\|^2$ で置き換えた式によってファジィプロトタイプ分類関数が新たに定義され、ファジィ k 最近隣分類関数も同様に定義されることがわかる。それらをそれぞれ $D_i(\mathbf{x})$ および $E_i(\mathbf{x})$ とすると、次式が得られる。

(a) ファジィプロトタイプ分類法

$$D_i(\mathbf{x}) = \frac{e^{-\lambda\|\mathbf{x} - \mathbf{z}_i\|^2}}{\sum_{j=1}^c e^{-\lambda\|\mathbf{x} - \mathbf{z}_j\|^2}} \quad (3.9)$$

(b) ファジィ k 最近隣分類法

$$E_i(\mathbf{x}) = \frac{\sum_{\ell=1}^k u_{i\ell} e^{-\lambda\|\mathbf{x} - \mathbf{y}_\ell\|^2}}{\sum_{j=1}^k e^{-\lambda\|\mathbf{x} - \mathbf{y}_j\|^2}} \quad (3.10)$$

ここで、特に考察の対象とするのは、プロトタイプ分類法である。いま、fuzzy c -means 法あるいはエントロピー法の交互最適化が収束して終了したとする。そこで、プロトタイプ分類関数における \mathbf{z}_i を収束したときのクラスター中心 \mathbf{v}_i で置き換える。このとき得られるファジィプロトタイプ分類関数はクラスタリングされた個体のデータ点 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ を内挿している。すなわち、次の命題が成り立つ。

命題 3.1 $z_i = v_i, i = 1, 2, \dots, c$ とする。分類関数 $B_i(x)$ と (3.4) で与えられるメンバシップに対して

$$B_i(x_k) = \bar{u}_{ik}, \quad k = 1, 2, \dots, n$$

が成り立つ。また、分類関数 $D_i(x)$ と (3.3) で与えられるメンバシップに対して

$$D_i(x_k) = \bar{u}_{ik}, \quad k = 1, 2, \dots, n$$

が成り立つ。

証明は容易であるので省略する。

このように、ファジークラスタリングが行われた結果、クラスタリングされた個体のデータ点を内挿する分類関数 $B_i(x)$ と $D_i(x)$ が自然に生じている。このうち、関数 $B_i(x)$ が Keller ら [33] によって用いられたことは上に述べた通りである。そこで、 $B_i(x)$ と $D_i(x)$ の性質を比較検討しよう。なお、以下では、プロトタイプ z_i と v_i は同一であると仮定する。

ファジークラスタリングの結果を利用するとき、メンバシップの値をそのまま用いる場合とクリスプなクラスタに変換して用いる場合の2通りがある。クリスプなクラスタに変換する場合、命題 3.1 から、領域

$$Q_i = \{x \in R^p \mid B_i(x) > B_j(x), \text{ for all } j \neq i\}$$

および

$$R_i = \{x \in R^p \mid D_i(x) > D_j(x), \text{ for all } j \neq i\}$$

内の点がクリスプにされたクラスタ i に属すると考えるのが妥当である。 Q_i および R_i は、クラスタ i に対応する分類関数の値が他のクラスタに対する値よりも大きい領域を意味しているからである。また、これらの領域はクラスタのみならず、新たに得られた観測 x がどのクラスに入れられるべきかを示している。(ただし、 Q_i, R_i は境界を含まないので、境界上の点は、クリスプに分類できない。)

これとは別に、最近隣分類が基本的な分類の方法として知られている。これは、観測 x をプロトタイプ $z_i = v_i$ のうち最も近くのものに対応させる分類である。この分類によれば

$$V_i = \{x \in R^p \mid \|x - v_i\| < \|x - v_j\|, \text{ for all } j \neq i\}$$

がクラス i に分類される点の集合となる。

次の命題が得られる。

命題 3.2 fuzzy c -means 法とエントロピー正則化法による分類関数から得られる領域と最近隣分類によって得られる領域は等しい。すなわち、以下の関係が成り立つ。

$$Q_i = R_i = V_i, \quad i = 1, 2, \dots, c.$$

証明は容易であるが、補足 3 に記す。

V_i は R^p における Voronoi 図 [34] の v_i を含む領域である。命題 3.2 はファジィクラスタリングの結果をクリस्प化したとき、Voronoi 図の領域が生じることを意味している。

後の結果を述べるために Voronoi 図の領域の性質について簡単に言及しよう。 v_i を含む領域 V_i には、有界なものとは非有界なものがある。 V_i が非有界であれば、 v_i を通り、かつすべての $j \neq i$ について v_j をその片側に含む超平面が存在する。いいかえれば、すべての $j \neq i$ についてスカラー積 $\langle x - v_i, v_j - v_i \rangle$ が負となるベクトル $x - v_i$ をとることができる。また、 V_i が有界であれば、このような超平面（およびベクトル）は存在しない。図 3.1 は Voronoi

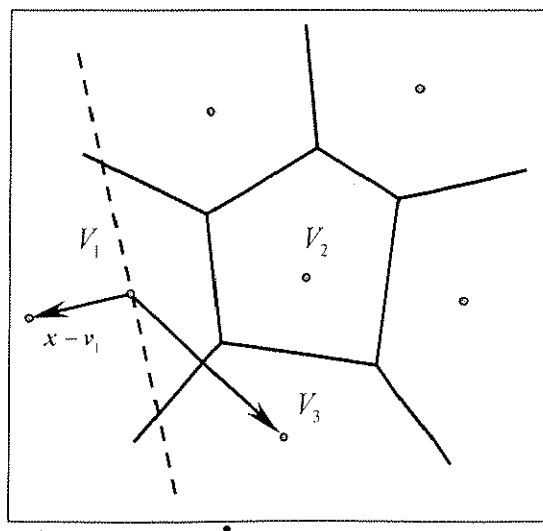


図 3.1: Voronoi 図の例

図の例である。 V_1 は非有界であり、 V_2 は有界である。 V_1 を横切っている破線は v_1 を通り、 $v_j, j \neq i$ をその片側に含んでいる。従って、ベクトル $x - v_1$ は $\langle x - v_1, v_j - v_1 \rangle < 0, j \neq i$,

を満たしている。ファジィ分類関数は値の大きさに観測のクラスに対する所属度を表すため、最近隣分類よりもはるかに情報量が多い。また、このことは、ファジィ分類関数の性質を調べておくことが有用であることを意味している。そこで、 $B_i(\mathbf{x})$ と $D_i(\mathbf{x})$ の最大値がどこでとられるかを考えてみよう。 $B_i(\mathbf{x})$ については、最大値は $\mathbf{x} = \mathbf{v}_i$ でとられることは容易にわかる。しかしながら、 $D_i(\mathbf{x})$ については、一般に最大値は $\mathbf{x} = \mathbf{v}_i$ ではとられない。さらに、関数 $B_i(\mathbf{x})$ と $D_i(\mathbf{x})$ の \mathbf{x} が原点から限りなく遠ざかるときの挙動を観察しよう。このことは、他の点から遠く離れた点がどのように分類されるかを知る上で重要である。このためには、これらの関数値の逆数を考察するのが便利である。まず、

$$\{B_i(\mathbf{x})\}^{-1} = \sum_{j=1}^c \left(\frac{\|\mathbf{x} - \mathbf{v}_i\|^2}{\|\mathbf{x} - \mathbf{v}_j\|^2} \right)^{\frac{1}{c-1}}$$

(ただし、 $\mathbf{x} \neq \mathbf{v}_i$) より、 $\|\mathbf{x}\| \rightarrow \infty$ となるに従って $\{B_i(\mathbf{x})\}^{-1}$ は c に近づくことがわかる。一方、

$$\begin{aligned} \{D_i(\mathbf{x})\}^{-1} &= \sum_{j=1}^c e^{-\lambda(\|\mathbf{x} - \mathbf{v}_j\|^2 - \|\mathbf{x} - \mathbf{v}_i\|^2)} \\ &= 1 + \sum_{j \neq i} e^{-\lambda(2\langle \mathbf{x}, \mathbf{v}_i - \mathbf{v}_j \rangle + \|\mathbf{v}_j\|^2 - \|\mathbf{v}_i\|^2)} \end{aligned}$$

における $e^{-\lambda(2\langle \mathbf{x}, \mathbf{v}_i - \mathbf{v}_j \rangle + \|\mathbf{v}_j\|^2 - \|\mathbf{v}_i\|^2)}$ は、ベクトル \mathbf{x} がすべての $j \neq i$ についてスカラー積 $\langle \mathbf{x}, \mathbf{v}_i - \mathbf{v}_j \rangle$ が正となる方向に原点から無限に遠ざかるならば 0 に近づき、そうでない方向で原点から遠ざかるならば発散する。このような方向がとれるかどうかは、先に述べた Voronoi 図の領域の非有界/有界性に依存していることは明らかである。これらをまとめて次の命題 3.3, 命題 3.4 として述べることができる。

命題 3.3 fuzzy c-means 法より生じるファジィ分類関数 $B_i(\mathbf{x})$ は最大値 1 を $\mathbf{x} = \mathbf{v}_i$ においてとり、 $\|\mathbf{x}\| \rightarrow \infty$ のとき、

$$B_i(\mathbf{x}) \rightarrow \frac{1}{c}$$

となる。

命題 3.4 エントロピー法から生じるファジィ分類関数 $D_i(\mathbf{x})$ について次の (i), (ii) が成り立つ。

(i) v_i を含む Voronoi 図の領域 V_i が有界のとき, $D_i(\mathbf{x})$ は最大値を領域のどこかでとるが, 最大値は 1 よりも小さい。また, $\|\mathbf{x}\| \rightarrow \infty$ のとき, $D_i(\mathbf{x}) \rightarrow 0$ となる。

(ii) V_i が非有界のとき, その領域のなかで $\|\mathbf{x}\| \rightarrow \infty$ と原点から遠ざかるに従って $D_i(\mathbf{x})$ は増加し,

$$D_i(\mathbf{x}) \rightarrow 1$$

となる。

3.7 数値例

3 章の主な目的は fuzzy c -means 法とエントロピー法を理論的に比較することであるので, 数値例は簡単なものにとどめ, エントロピー法による fuzzy c -varieties 法などの計算例については, 別の機会にゆずるものとする。

図 3.2 は三角形領域のなかにランダムに点を配置した例である。三角形は一辺の長さが $\sqrt{3}$ で, 重心は原点にある。点の数は 294 個である。図 3.3 と図 3.4 はそれぞれ従来の fuzzy

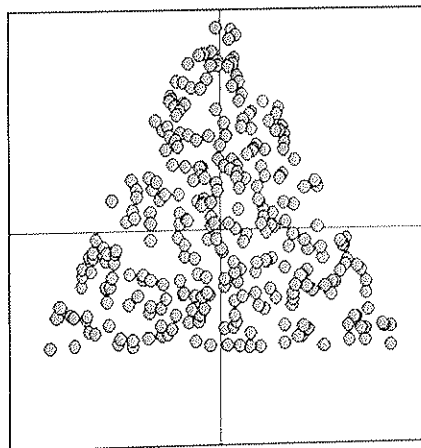


図 3.2: 三角形領域上の点集合

c-means 法とエントロピー法を用いてこれらの点を4つのクラスタに分類した結果を示している。ただし、 $u_{ik} \geq 0.60$ を満たす点について、中央のクラスタは灰色で表し、周囲の3つのクラスタについては黒色で表している。また、すべてのクラスタについて $u_{ik} < 0.60$ となる点は白く表している。パラメータ q および λ はそれぞれ $q = 1.8$, $\lambda^{-1} = 0.06$ である。

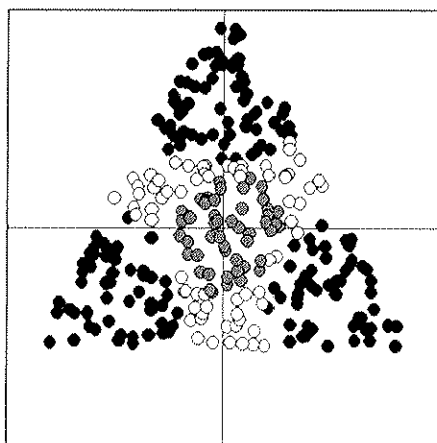


図 3.3: 図 3.2 に fuzzy c-means 法を用いて形成した4つのクラスタ

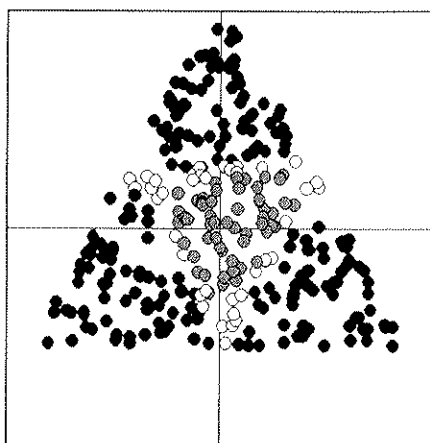


図 3.4: 図 3.2 にエントロピー法を用いて形成した4つのクラスタ

次に、これらから生成される分類関数 $B_i(x)$ と $D_i(x)$ をみよう。図 3.5 は従来の fuzzy c-means 法において上方及び、中央のクラスタに対する分類関の等高線を示している。

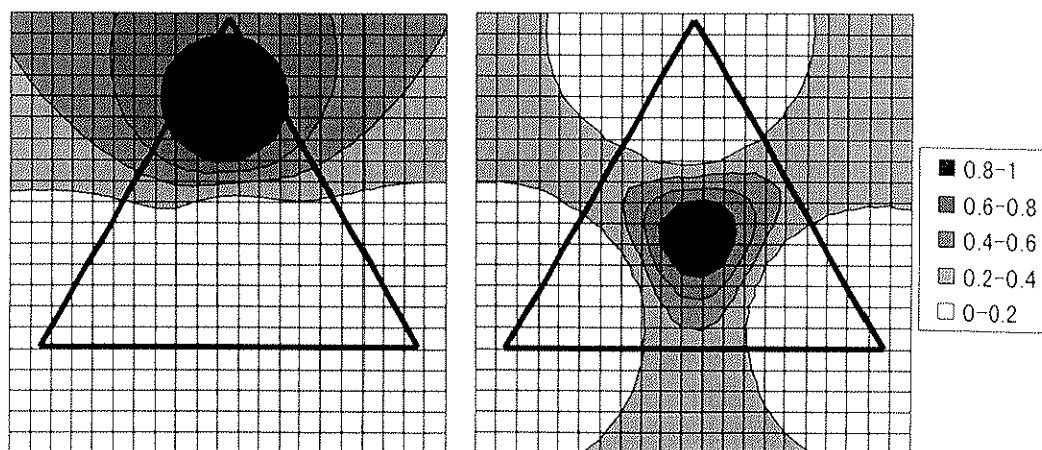


図 3.5: fuzzy c-means 法による上方・中央のクラスタに対する分類関数

さらに、エントロピー法による分類関数の等高線を図 3.6 に示した。これらはそれぞれ上方と中央のクラスタに対応している。

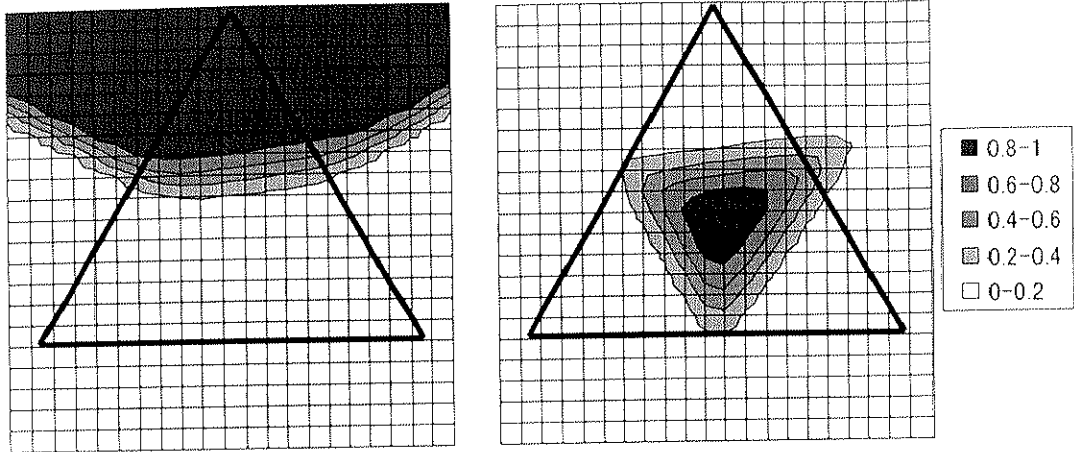


図 3.6: エントロピー法による上方・中央のクラスタに対する分類関数

上方のクラスタに対応する Voronoi 図の領域は非有界であり、中央のクラスタには有界領域が対応している。上方の領域について、図 3.6 の上方のクラスタの分類関数は、点が原点から遠ざかるに従って値が単調に大きくなる様子が見られる。これに対して、図 3.5 ではこのような単調性は見られない。また中央の領域についても、図 3.5 と図 3.6 における中央のクラスタの分類関数の形状が大きく異なっているのがわかる。図 3.5 の等高線が無限に遠くまで伸びているのに対し、図 3.6 では、有界領域に限定されている。

このように、3.6 節の命題 3.3、命題 3.4 と数値例の観察とは合致している。

3.8 おわりに

ここでは、エントロピー最大化法に正則化の概念を導入することによって、エントロピー関数を用いる方法が従来の fuzzy c -means 法の交互最適化の枠組みで論じられることを示し、そこから、従来提案されてきた fuzzy c -means 法の様々な変形がエントロピー正則化法によって扱えることを示した。

また、クラスタリングの結果として自然に生じるファジィ分類関数を考察した。すなわち、fuzzy c-means 法とエントロピー法による分類関数を比較することによって、それぞれの特徴が対照的であることを示した。要約すれば、従来の fuzzy c-means 法が中心集中型で、中心にメンバシップ値 1 をもつのにに対し、エントロピー法は分離型で、中心は必ずしもメンバシップ値 1 をもたない。また、 x が無限に遠ざかるとき、従来の fuzzy c-means 法では、メンバシップ値は必ず $1/c$ になるが、エントロピー法では、Voronoi 図の非有界領域において x が原点から遠ざかるに従って、メンバシップ値は 1 に近づく。

この考察において最も強調したい主張は、これら 2 つの方法からなる広義の fuzzy c-means 法においては、アルゴリズムが収束したとき定義される分類関数が重要であるということである。また、これら 2 種類の分類関数を対照させることによって、それぞれを単独に考察するときよりも多くの事実が知られるようになる。この意味で、従来の fuzzy c-means 法とエントロピー法は共に欠かせないファジィクラスタリングの両輪であるといえよう。

エントロピー法にもとづいた広義の fuzzy c-means 法の変形については、fuzzy c-varieties 法だけを簡単に紹介するにとどめた。これらについて、ここで行ったのと同様の考察を試みることは興味深い。たとえば、 L_1 fuzzy c-means 法 [38] についてアルゴリズムを開発し、分類関数の性質を調べることが考えられる。また、fuzzy c-regression Models についても考察が可能であろう。このように、fuzzy c-means 法とエントロピー法との比較、さらに分類関数の考察はファジィクラスタリングの研究に新たな局面を開くと思われる。

3.9 補足

補足 1. エントロピー法の解

FCM における最適解 (3.3) を導出しよう。制約条件を考慮すれば、ラグランジアン L は次式となる。

$$L = J_\lambda(U, V) + \sum_{k=1}^n \nu_k \left(\sum_{i=1}^c u_{ik} - 1 \right)$$

ただし、 ν_k はラグランジュ乗数である。

$$\frac{\partial L}{\partial u_{ik}} = d_{ik} + \lambda^{-1}(1 + \log u_{ik}) + \nu_k = 0$$

を解いて

$$u_{ik} = \exp(-1 - \lambda\nu_k) \exp(-\lambda d_{ik})$$

制約条件 $\sum_{i=1}^c u_{ik} = 1$ より

$$\exp(-1 - \lambda\nu_k) = \left(\sum_{j=1}^c \exp(-\lambda d_{jk}) \right)^{-1}$$

となるから、(3.3) が得られる。

補足 2. 正則化について

fuzzy c-means 法の場合、

$$u_{ik}^{-1} = \sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{1}{q-1}}$$

である。 \mathbf{x}_k が \mathbf{v}_i に最も近いとすると、上式の \sum_j の各項のうち一つが 1 で他の項は 1 より小さい。ただし、最も近い \mathbf{v}_i は一意に定まると仮定する。このとき、 $q \rightarrow +1$ とすると、明らかに、1 より小さい項は 0 に近づき、 u_{ik} は 1 に近づく。 \mathbf{x}_k が \mathbf{v}_i に最も近くはないとすると、 \sum_j の項のうちに 1 より大きいものが存在する。この項は $q \rightarrow +1$ とすると無限大に発散するので、 u_{ik} は 0 に近づく。

エントロピー法の場合、

$$u_{ik}^{-1} = \sum_{j=1}^c e^{-\lambda(d_{jk} - d_{ik})}$$

である。同様に、 \mathbf{x}_k が \mathbf{v}_i に最も近いとすると、上式の \sum_j の各項における $d_{jk} - d_{ik}$ のうち一つが 0 で他の項は正である。(ただし、最も近い \mathbf{v}_i は一意に定まると仮定する。) $\lambda \rightarrow +\infty$ とすると、明らかに u_{ik} は 1 に近づく。 \mathbf{x}_k が \mathbf{v}_i に最も近くはないとすると、 \sum_j の項における $d_{jk} - d_{ik}$ のうちに負のものが存在する。この項は $\lambda \rightarrow +\infty$ とすると無限大に発散するので、 u_{ik} は 0 に近づく。

補足 3. 命題 3.2 の証明

$j \neq i$ とする。(3.5) より

$$\begin{aligned} B_i(\boldsymbol{x}) > B_j(\boldsymbol{x}) &\Leftrightarrow e^{-\beta \log \|\boldsymbol{x} - \boldsymbol{v}_i\|^2} > e^{-\beta \log \|\boldsymbol{x} - \boldsymbol{v}_j\|^2} \\ &\Leftrightarrow \|\boldsymbol{x} - \boldsymbol{v}_i\| < \|\boldsymbol{x} - \boldsymbol{v}_j\| \end{aligned}$$

また,

$$\begin{aligned} D_i(\boldsymbol{x}) > D_j(\boldsymbol{x}) &\Leftrightarrow e^{-\lambda \|\boldsymbol{x} - \boldsymbol{v}_i\|^2} > e^{-\lambda \|\boldsymbol{x} - \boldsymbol{v}_j\|^2} \\ &\Leftrightarrow \|\boldsymbol{x} - \boldsymbol{v}_i\| < \|\boldsymbol{x} - \boldsymbol{v}_j\| \end{aligned}$$