



Title	重み乗数の性質を利用したファジィクラスタリングについて
Author(s)	今井, 英幸; Imai, Hideyuki; 宮澤, 安夫 他
Citation	北海道大學工学部研究報告, 167, 11-18
Issue Date	1994-01-14
Doc URL	https://hdl.handle.net/2115/42393
Type	departmental bulletin paper
File Information	167_11-18.pdf



重み乗数の性質を利用したファジィクラスタリングについて

今井 英幸 宮澤 安夫
宮腰 政明 伊達 惇

(平成5年8月31日受理)

The Fuzzy Clustering Algorithm Based on the Properties of the Weighting Exponent

Hideyuki IMAI, Yasuo MIYAZAWA,
Masaaki MIYAKOSHI and Tsutomu DA-TE

(Received August 31, 1993)

Abstract

In the fuzzy c-means algorithm, proposed by Bezdek and Dunn, the fuzziness of the grade of membership depends on the weighting exponent. Though one must choose the exponent to use the fuzzy c-mean algorithm, no theoretical basis for the optimal exponent is known.

In this paper, we discuss the properties of the weighting exponent by numerical experiments. And we propose the fuzzy clustering algorithm based on these properties. Moreover we propose the graphical representation method for clustering.

1. はじめに

Bezdek[1], Dunn[2]等によって提案されたファジィクラスタリングの代表的手法である fuzzy c-means (FCM)法においては、重み乗数(weighting exponent)と呼ばれるクラスタリングの曖昧さを決めるパラメータを適当に決定する必要があるが、その決定方法には明確な基準はない。本報告ではこの重み乗数の性質について検討し、さらにこの性質を利用したクラスタリングアルゴリズムを提案する。この手法ではより集中している個体群から一つのクラスターを形成していく興味深い過程が見られる。そこで分析の結果のみを出力する従来の方法に対し、クラスターの生成過程を図的に表現する方法についても合わせて提案する。

2. クラスタ分析

有限個の個体(分類対象)をいくつかのある意味で似た個体の集団に分類することを考える。この集団のことをクラスターと言う。また、それぞれの個体がある一つのクラスターにのみ所属する場合これをハードクラスタリングと言い、これに対しそれぞれの個体がある帰属度により複数のクラスターに所属する場合をファジィクラスタリングと言う。

2.1 k-means 法

ハードクラスタリングの手法としてよく用いられるものに k-means 法がある。いま、 n 個の個体について各個体の p 変量観測ベクトルを

$$\mathbf{x}_i = (x_{i1}, \dots, x_{ip})' \in \mathbf{R}^p, \quad i = 1, \dots, n,$$

その多変量データ行列 X を

$$X = (x_{i\ell}), \quad i = 1, \dots, n, \quad \ell = 1, \dots, p$$

とする。ここで個体の集合を互いに排反な空でない k 個のクラスター

$$G = \{C_1, \dots, C_k\}$$

に分割することを考える。このときクラスターの等質性基準として考えられるのは、クラスター内でのばらつきをできるだけ小さく、クラスター間のばらつきをできるだけ大きくしようというものである。ここで平方和積和行列（全変動）を

$$T = \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}}) (\mathbf{x}_i - \bar{\mathbf{x}})'$$

クラスター内平方和積和行列の和（クラスター内変動の和）を

$$W(G) = \sum_{j=1}^k \sum_{i \in C_j} (\mathbf{x}_i - \mathbf{v}_j) (\mathbf{x}_i - \mathbf{v}_j)'$$

クラスター間平方和積和行列の和（クラスター間変動の和）を

$$B(G) = \sum_{j=1}^k n_j (\mathbf{v}_j - \bar{\mathbf{x}}) (\mathbf{v}_j - \bar{\mathbf{x}})'$$

（ n_j はクラスター C_j 内の個体数）

とすると、次の関係が成り立つ。

$$T = W(G) + B(G)$$

ただし

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i, \quad \mathbf{v}_j = \frac{1}{n_j} \sum_{i \in C_j} \mathbf{x}_i$$

である。 T は一定であるから、 $W(G)$ を小さくすれば $B(G)$ は大きくなる。そこで通常用いられるのは次のような基準である。

- (i) $\text{tr}(W(G)) \rightarrow \min$
- (ii) $\text{det}(W(G)) \rightarrow \min$

k-means 法では基準(i)を用いる。この評価関数 $\text{tr}(W(G))$ の大域的な最小値を与える分割を得るためには組み合わせ最適化問題を解く必要があるが、これは個体数が多くなった場合、現実には不可能である。そのためいくつかの近似解法が提案されている。k-means 法はそのような近似解法の一つであり、計算時間が短いという特徴がある。この手法はある初期分割が与えられたとき、ある個体をそれがいま所属するクラスターから別のクラスターに移動させてみて、クラスタリングが改良されるか（移動前より $\text{tr}(W(G))$ が小さくなるか）どうかを調べ、改良されていない

ば移動を行わず、改良されていれば実際に移動を行なうという操作を $\text{tr}(W(G))$ の値が収束するまで反復する手法である。

2.2 Fuzzy c-means 法

Bezdek[1], Dunn[2]等はこの k-means 法をファジィ理論で説明しようと、これを重み付き最小二乗の問題として扱い FCM 法を提案した。いま、 n 個の個体を互いに排反な空でない c 個のクラスターに分類することを考える。このとき以下の関数を最小にする U を求めることによりクラスタリングを行なうのが FCM 法である。

$$J_m(U, V) = \sum_{i=1}^n \sum_{j=1}^c (u_{ij})^m (d_{ij})^2$$

ここで V は

$$V = (v_1, \dots, v_c), \quad v_j \in \mathbf{R}^p$$

で、各クラスターの重心を表しており d_{ij} はこの重心 v_j と個体 x_i との距離を表す。ここでは最も一般的なユークリッド距離

$$d_{ij} = \sqrt{\sum_{l=1}^p (x_{il} - v_{jl})^2}$$

を用いる。また各個体が各クラスターに所属する度合を次の行列で表す。

$$U = (u_{ij}), \quad i = 1, \dots, n, \quad j = 1, \dots, c$$

つまり u_{ij} は、個体 x_i が第 j 番目のクラスターに所属する度合を表している。さらに m は重み乗数と呼ばれる $1 < m < \infty$ の値である。このとき、 J_m を最小化する U, V は以下の式によって与えられる。

$$v_j = \frac{\sum_{i=1}^n (u_{ij})^m x_i}{\sum_{i=1}^n (u_{ij})^m}, \quad j = 1, \dots, c$$

$$u_{ij} = 1 / \left(\sum_{k=1}^c \left(\frac{d_{ij}}{d_{ik}} \right)^{2/(m-1)} \right), \quad i = 1, \dots, n, \quad j = 1, \dots, c$$

ただし $d_{ik} = 0$ のときは

$$u_{ij} = \begin{cases} 1 & j=k \\ 0 & j \neq k \end{cases}$$

以上のことを利用し、交互最小二乗法を用いて最小化を行なう。

FCM アルゴリズム

- 初期化：ある適当な初期分割 $U^{(0)}$ 、およびクラスター数 c を選び $t=1$ とする。
- Step 1： $U^{(t)}$ を用いて $v^{(t)}$ を計算する。
- Step 2： $v^{(t)}$ を用いて $U^{(t+1)}$ を計算する。
- Step 3： $\epsilon > 0$ を満たす十分小さな ϵ に対し $\|U^{(t+1)} - U^{(t)}\|$ なら計算を終了し、そうでなければ $t \leftarrow t+1$ とし Step 1 に戻りこれを繰り返す。

2.3 数値実験

実験で用いたデータは $p=2$ とした 9 種類で以下のようなものである。クラスター数 c に関しては図 1 の 3 つの場合を考え、それぞれに対しクラスターの重心間距離 d を $d=2, 3, 4$ とした。これら 9 通りのデータについてそれぞれの重心を平均ベクトル、 I_2 を分散共分散行列とする正規乱数を各クラスターについて 10 個ずつ発生させたものをデータとして用いた。

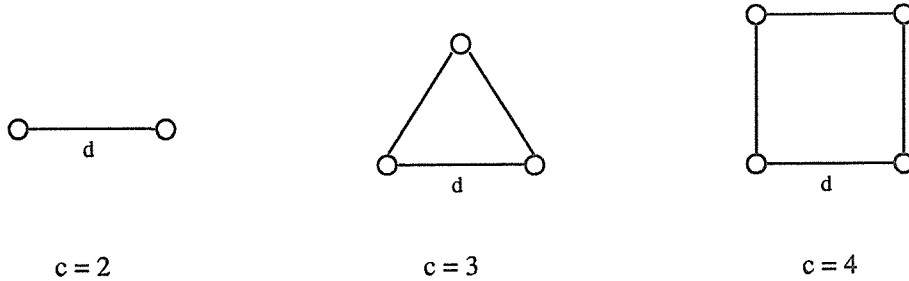


図 1 数値実験データのクラスター重心の配置

実験は重み乗数の値として 4.0, 3.0, 2.0, 1.5, 1.25, 1.05 の 6 種類を用い、それぞれに対しランダムに初期分割を与えて FCM 法によるクラスタリングを上記の 9 種類のデータについてそれぞれ 100 回ずつ行ない、その時のクラスタリングの様子について調べた。その結果が表 1 である。

表 1 m を変化させたときの局所解の数

	$m=4.0$	$m=3.0$	$m=2.0$	$m=1.5$	$m=1.25$	$m=1.05$
$c=2, d=2$	1	1	1	1	1	1
$c=2, d=3$	1	1	1	1	1	1
$c=2, d=4$	1	1	1	1	1	1
$c=3, d=2$	1	1	1	1	1	2 (92)
$c=3, d=3$	1	1	1	1	1	1
$c=3, d=4$	1	1	1	1	1	1
$c=4, d=2$	1	1	1	1	1	5 (93)
$c=4, d=3$	1	1	1	1	2 (96)	3 (72)
$c=4, d=4$	1	1	1	1	2 (99)	5 (90)

表 1 の数値は局所解の数であり、() 内の数値は 100 回のシミュレーションのうち何回が大域的最適解に収束したかを表している。つまり表中の数値 1 は、すべての初期分割に対してただ一つの解に収束したことを表しており、2 (92) は 100 回のうち 92 回は大域的最適解に収束したが、残りの 8 回は他の局所解に収束したことを表している。

この結果から、クラスター数が 2 の場合にはいかなる初期分割に対しても大域的最適解に収束すると考えられるが、クラスター数が 3 あるいは 4 で重み乗数 m が 1 に近い場合は初期分割に依存して複数の局所解が得られる場合もあるということがわかる。しかし、この場合も重み乗数 m を十分大きくするとやはりただ一つの大域的最適解と思われる解に収束する。

またこの手法は $m \rightarrow \infty$ とすると $u_{ij}=1/c$, ($\forall i, j$) となり、 $m \rightarrow 1$ とすると $u_{ij}=\{0, 1\}$ (この

とき FCM 法は k-means 法に一致する) となることから重み乗数 m は曖昧さを決定する要因であると考えられる。

3. 重み乗数 m の性質を利用したファジィクラスタリング

前節で述べたように、重み乗数 m が 1 に近い場合は初期分割に依存して複数の局所解を得る場合があるが、その場合も重み乗数 m を十分大きくするとただ一つの大域的最適解のみを得る。そこで、十分大きな重み乗数 m のときに得られた結果をそれより少しだけ小さい重み乗数 $m - \Delta m$ のときの初期分割として、この操作を繰り返し行なえば重み乗数が十分 1 に近いときも初期分割に依らず大域的最適解に収束することが予想される。またこの操作は帰属度を曖昧な状態からより曖昧でない状態に連続的に移行することでもある。

3.1 重み乗数の性質を利用した FCM アルゴリズム

- 初期化： c を固定し、初期分割 U と最終的な重み乗数 m_0 および十分大きな m を与える。
- Step 1： FCM アルゴリズムの Step 1, Step 2 を行なう。
- Step 2： $\epsilon > 0$ を満たす十分小さな ϵ に対し $\|U^{(t+1)} - U^{(t)}\| < \epsilon$ なら Step 3 へ、そうでなければ $t \leftarrow t + 1$ とし Step 1 に戻る。
- Step 3： $m = m_0$ ならば計算を終了し、そうでなければ $m \leftarrow m - \Delta m$ とし Step 1 に戻りこれを繰り返す。ただし直前の m で得られたクラスタリングの結果を次の初期分割とする。

実際に 2.3 で用いた 9 種類のデータに対し、重み乗数を $4.0 \rightarrow 3.0 \rightarrow 2.0 \rightarrow 1.5 \rightarrow 1.25 \rightarrow 1.1 \rightarrow 1.05$ と変化させ上のアルゴリズムによるクラスタリングを 100 回ずつ行なった結果、100 回発生させたどのランダムな初期分割に対しても k-means 法の大域的最適解と一致したクラスタリングが 9 種類のデータすべてについて得られた。その一例が以下の図 2 で、クラスター数およびクラスターの重心間距離はそれぞれ $c = 3$, $d = 4$ である。ただしここではあるデータ x_i のクラスター j への帰属度 u_{ij} が 0.9 以上となったときにはデータ x_i はクラスター j に分類されたとして各クラスターに応じて \circ , \triangle , \square で表し、どのクラスターに関しても帰属度が 0.9 を超えないデータは \times 印で表す。また \bullet , \blacktriangle , \blacksquare はそれぞれのクラスターの重心を表している。

この結果より重み乗数の値が小さくなるにつれて分割がはっきりしていくのがよくわかる。

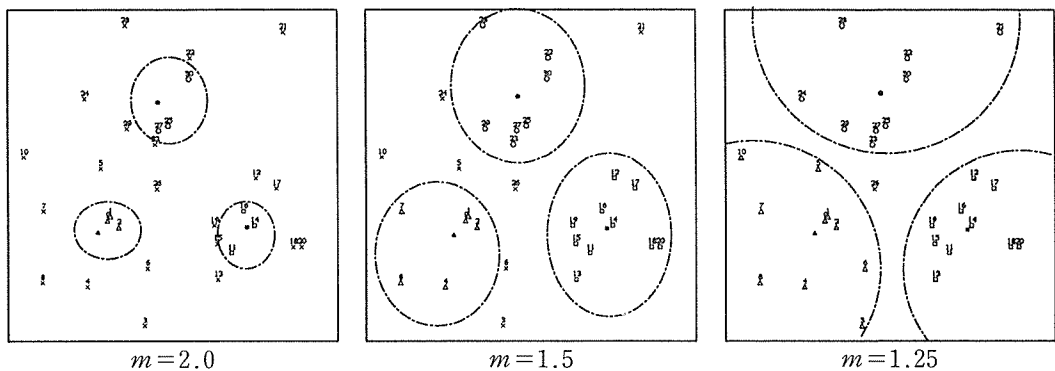


図 2 重み乗数の性質を利用したファジィクラスタリング (2次元データ)

4. クラスターの生成過程とその視覚的表現

前章で提案した手法は重み乗数の値が小さくなるにつれて、より曖昧な個体を徐々にクラスターに分類していく。このようなクラスターの生成過程を視覚的に表現することは結果の解釈のためにも必要であると考えられる。ここで2次元データに関する場合は問題とはならないので、多次元データの視覚的表示方法について考察する。

多次元データを図示する場合、データを低次元空間（1次元もしくは2次元）へ射影する方法が用いられる。以下では2次元空間への射影について述べるが、これらの結果は容易に一般の射影に拡張できる。多くの場合、データのばらつき（分散）を最大にするという基準のもとで主成分分析による次元の縮約が行なわれるが、本報告ではクラスター間のばらつきを最大にするという基準のもとで次元の縮約を行なう。クラスター間のばらつきを表す量としては2.1で示したクラスター間平方和積和行列の和を用いる。この式をファジィクラスタリングに適用できるように拡張すると

$$B = \sum_{j=1}^c \left(\sum_{i=1}^n u_{ij} \right) (v_j - \bar{x})(v_j - \bar{x})'$$

となる。クラスター間のばらつき B を最大にするという基準で次元を2次元に射影するためには、 $a'a = 1$ のもとで $a'Ba$ を最大にするような $a \in R^p$ と、それと直交し、かつ $b'b = 1$ のもとで $b'Bb$ を最大にするような $b \in R^p$ とで張られる平面を求めればよく、これは B の最大固有値に対応する固有ベクトル a と、2番目に大きい固有値に対応する固有ベクトル b を求めることに帰着される。またそのアルゴリズムは以下の通りである。

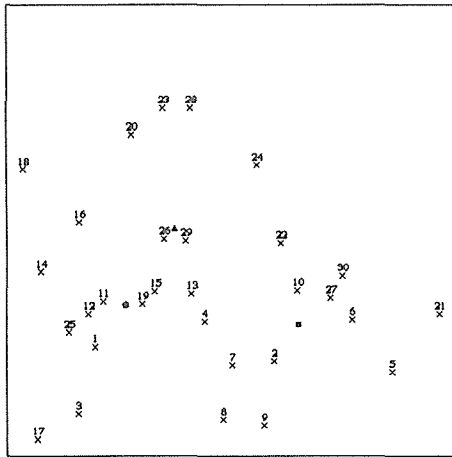
- 初期化：クラスター数 c 、初期分割 $U^{(1)}$ 、最終的な重み乗数 m_0 および十分大きな m を与え、 $t = 1$ とする。
- Step 1： $U^{(t)}$ を用いて $v^{(t)}$ を計算する。
- Step 2： $v^{(t)}$ を用いて $U^{(t+1)}$ を計算する。
- Step 3： $\epsilon > 0$ を満たす十分小さな ϵ に対し $\|U^{(t+1)} - U^{(t)}\|$ なら Step 4 へ、そうでなければ $t \leftarrow t + 1$ とし Step 1 に戻る。
- Step 4： Step 3 で得られた結果より B を求め、その最大固有値に対応する固有ベクトル a と、2番目に大きい固有値に対応する固有ベクトル b を求める。
- Step 5： データ x_i および重心ベクトル v_j を $a-b$ 平面に射影し、射影されたデータ ($a'x_i, b'x_i$) および重心ベクトル ($a'v_j, b'v_j$) を図示する。
- Step 6： $m = m_0$ ならば計算を終了し、そうでなければ $m \leftarrow m - \Delta m$ とし Step 1 に戻りこれを繰り返す。ただし直前の m で得られたクラスタリングの結果を次の初期分割とする。

5. 数値実験

$p = 5$ とし、平均が $\mu_1 = (2, 2, 2, 2, 0)'$ 、 $\mu_2 = (0, 2, 2, 2, 2)'$ 、 $\mu_3 = (2, 2, 0, 2, 2)'$ である3個の正規母集団からそれぞれ10個、正規乱数を発生させて30個の標本を作った。ただしそれぞれの母集団の分散共分散行列はすべて I_5 である。図3はこのデータに関して重み乗数の性質を利用したファジィクラスタリングを行なった結果を上アルゴリズムに従い図示したものである。1から10までは平均 μ_2 、11から20までは平均 μ_3 、21から30までは平均 μ_1 の母集

図からのデータ点である。ただし○, △, □などの記号は図2と同様である。

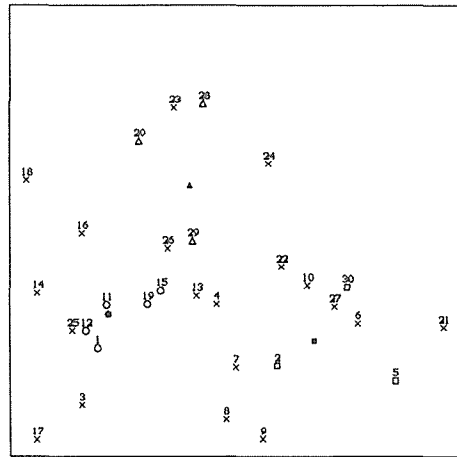
この結果より単なる第1主成分と第2主成分による散布図(図4参照)ではわかりにくかったクラスター分析の結果が今回の手法によるとはっきりとわかるようになり, また重み乗数の値が小さくなるにつれてより曖昧な個体を徐々にクラスターに分類していくという生成過程が視覚的に表現されることによってよりわかりやすくなった。



$m = 2.00$

$$\alpha = (0, 820, 0, 000, 0, 000, 0, 572, 0, 000)'$$

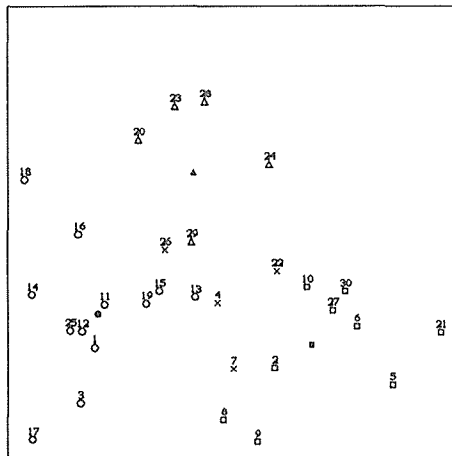
$$b = (0, 234, 0, 000, 0, 000, -0, 335, 0, 913)'$$



$m = 1.50$

$$\alpha = (0, 854, 0, 000, 0, 000, 0, 521, 0, 000)'$$

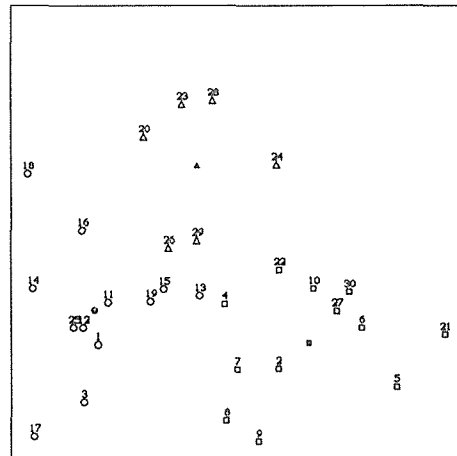
$$b = (0, 284, 0, 000, 0, 000, -0, 465, 0, 838)'$$



$m = 1.25$

$$\alpha = (0, 867, 0, 000, 0, 000, 0, 498, 0, 000)'$$

$$b = (0, 279, 0, 000, 0, 000, -0, 486, 0, 828)'$$

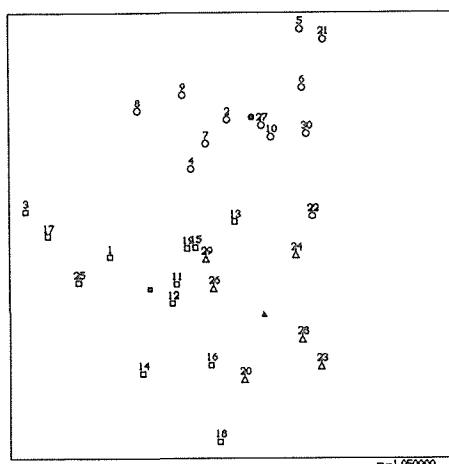


$m = 1.05$

$$\alpha = (0, 881, 0, 000, 0, 000, 0, 474, 0, 000)'$$

$$b = (0, 260, 0, 000, 0, 000, -0, 483, 0, 836)'$$

図3 クラスターの生成過程



$m=1.05$

第1主成分 = (0.599, 0.000, 0.296, 0.338, 0.663)'

第2主成分 = (0.517, 0.000, 0.000, 0.478, -0.710)'

図4 主成分分析による表示

6. ま と め

FCM法によるクラスタリングにおいては、重み乗数の値が1に近くなると初期分割に依存して複数の局所解が得られる場合があるが、重み乗数の性質を利用した今回のクラスタリングでは、ほとんどの場合初期分割に依らず大域的最適解に収束することを確認した。またその際のクラスターの生成過程をより理解しやすい形で視覚的に表現した。

参考文献

- 1) Bezdek, James C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. Prentice Hall Press.
- 2) Dunn, J.C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. Journal of Cybernetics, 3, 32-57.
- 3) 佐藤義治(1992). クラスタ分析における焼き鈍し法の有効性. 計算機統計学, 5 (1), 11-18.
- 4) 宮澤安夫・今井英幸・宮腰政明・伊達惇(1993), ファジィクラスタリングにおける重み乗数の性質について. 情報処理北海道シンポジウム '93講演論文集