第2章

crisp k-means法とfuzzy c-means法

fuzzy c-means法[10, 11, 12]は，非階層的なファジィクラスタリング手法の中で最もよく知られ，かつ多くの変形と応用が研究されている[10, 32]。本論文は，このfuzzy c-means法を中心とした研究について述べている。そこで，本章ではこのfuzzy c-means法とその前身である，ファジィ化される前のcrisp k-means法について簡単に紹介し，fuzzy c-means法がどのようにしてcrisp k-means法からファジィ化されたかについて述べる。以下の内容は主に宮本[8]に詳しく述べられている。また，それぞれの手法によるクラスタリング結果を簡単な例題を用いて示しす。

2.1 crisp k-means法

crisp k-means法はクリスプk-平均法とも呼ばれ，McQueen[5]やDuda and Hart[6], Forgy[7]らによって1960年代から1970年代に示された非階層的なクラスタリング手法である[1, 8]。クラスタリングの対象となる個体を，k個のクラスタの中でもっとも近い平均値（クラスタの重心）をもつクラスタに割り当てるように分割を行うクラスタリング手法である。以下に，本論文を通して用いる記号を定め，crisp k-means法について簡単に述べる。

いま**個体** $x_1, x_2, \cdots, x_n$ を p 次元ユークリッド空間の点 $(x_k \in R^p)$ とし，その成分は $x_k = (x_k^1, x_k^2, \cdots, x_k^p)$ と表されるとする。また，

$$X = \{x_1, x_2, \cdots, x_n\}$$

をクラスタリングすべき**個体の集合**であるとする。crisp k-means 法では一般に k 個のクラスタへの分割を考えるが，論文全体で統一するために慣例によって，**クラスタの数**を c とする。そして各**クラスタの中心**を $v_1, v_2, \cdots, v_c$ とし，成分を $v_i = (v_i^1, v_i^2, \cdots, v_i^p)$ と表わす。また，$V = (v_1, v_2, \cdots, v_c)$ とする。

このとき，個体 x_k がクラスタ i に**所属するか否か**を変数 u_{ik} によって，

$$u_{ik} = \begin{cases} 1, & (x_k \text{ がクラスタ } i \text{ に所属する}) \\ 0, & (x_k \text{ がクラスタ } i \text{ に所属しない}) \end{cases}$$

とし，行列 $U = (u_{ik})$ を定める。すべての個体はただ1つのクラスタに所属するので，$\sum_{i=1}^{c} u_{ik} = 1$ とする。

このとき crisp k-means 法は，行列 U とクラスタの中心 V を適切に選ぶことにより，

$$J_{ckm}(U, V) = \sum_{i=1}^{c} \sum_{k=1}^{n} u_{ik}\, d(v_i, x_k)$$

の最小化問題として定式化できる。ただしユークリッド距離の場合，

$$d(v_i, x_k) = \|x_k - v_i\|^2 = \sum_{\ell=1}^{p} (x_k^\ell - v_i^\ell)^2$$

を用いる。今後 $d(v_i, x_k)$ を単に d_{ik} とする事もある。ここで U の**許容集合**

$$M_{ckm} = \left\{ (u_{ik}) \,\middle|\, \sum_{i=1}^{c} u_{ik} = 1,\ u_{ik} \in \{0, 1\},\ k \le n \right\} \tag{2.1}$$

を用いれば，この最適化問題は次のように書ける。

$$\min_{U \in M_{ckm},\, V \in \boldsymbol{R}^{pc}} J_{ckm}(U, V)$$

ただし，$\boldsymbol{R}^{pc}$ は p 次元空間 $\boldsymbol{R}^p$ の c 個の直積である。

この目的関数を U と V について同時に最小化するのは困難である。そこで，最初に U を固定し V について，次に V を固定し U について $J_{ckm}(U, V)$ を最小化する**2段階アルゴリズム**を考える [10]。

crisp k-means 法による CKM アルゴリズム

CKM 1. $\bar{U}$ と $\bar{V}$ の初期値を適当に与える。

CKM 2. $\min_{V \in R^{pc}} J_{ckm}(\bar{U}, V)$ の最適解を $\bar{V}$ とする。

CKM 3. $\min_{U \in M_{ckm}} J_{ckm}(U, \bar{V})$ の最適解を $\bar{U}$ とする。

CKM 4. 最適解 $(\bar{U}, \bar{V})$ が収束していれば終了，そうでなければ CKM 2 に戻る。

この手続きは Forgy の方法 [1, 7] として引用される crisp k-means 法の標準的なアルゴリズムである。

ここで CKM 3. の U に関する最適解は，

$$\bar{u}_{ik} = \begin{cases} 1, & (\|\boldsymbol{x}_k - \boldsymbol{v}_i\| \leq \|\boldsymbol{x}_k - \boldsymbol{v}_j\|, \quad 1 \leq j \leq c) \\ 0, & (\text{その他}) \end{cases}$$

となり，V に関する最適解は，各クラスタの重心（平均値）となる。

2.2 fuzzy c-means 法

fuzzy c-means 法 [10, 11, 12] はファジィc-平均法とも呼ばれ，L.A.Zadeh によって 1965 年に提案されたしたファジィ理論 [9] を crisp k-means 法に適用し，ファジィ化したクラスタリング手法である。各個体はクラスタに対してはっきりと所属が決定されず，各クラスタに対する所属の度合いを与えるような分割を考える。

fuzzy c-means 法では，この crisp k-means 法をファジィ化するために 2 値マトリックス U をファジィに一般化する。これによって u_{ik} は 0 から 1 の任意の値をとれることになり，許容集合は，

$$M_{fcm} = \left\{ (u_{ik}) \;\middle|\; \sum_{i=1}^{c} u_{ik} = 1,\ u_{ik} \in [0,1],\ k \leq n \right\} \tag{2.2}$$

と一般化される。この M_{fcm} の制約を満たす U によって定められる X 上のファジィ集合

$$\mu_i(\boldsymbol{x}_k) = u_{ik}, \quad \boldsymbol{x}_k \in X, \quad i = 1, 2, \cdots, c$$

を X のファジィ分割と呼ぶ。また，変数 u_{ik} を個体 $\boldsymbol{x}_k$ のクラスタ i へのメンバシップ値，または帰属度と呼び，行列 U をメンバシップ行列と呼ぶ。

以上のようにファジィ分割は導入されたが，一般化はまだ十分ではない。この手続きで，J_{ckm}, M_{fcm} を用いたとしても，クラスタリング結果は crisp k-means 法の結果と同じ結果となり，ファジィ化されないからである。この目的関数は u_{ik} に対して線形となり，ステップ CKM3 は線形計画問題となる。そして線形計画問題の最適解は端点，すなわち u_{ik} が 0 か 1 となり，メンバシップはクリスプ解となるのである。

そのため，Dunn [11, 12] と Bezdek [10] はクラスタリング結果にファジィ分割が有効となるように，スムージングパラメータ $q(>1)$ を導入し，次のような評価規範を用いることを提案している。

$$J_{fcm}(U,V)=\sum_{i=1}^{c}\sum_{k=1}^{n}(u_{ik})^q\, d(\boldsymbol{v}_i,\boldsymbol{x}_k)$$

そして 2 段階アルゴリズムは，この J_{fcm} を用いることにより次のようになる。

fuzzy c-means 法による FCM アルゴリズム

FCM 1. $\bar{U}$ と $\bar{V}$ の初期値を適当に与える。

FCM 2. $\min_{V\in R^{pc}} J_{fcm}(\bar{U},V)$ の最適解を $\bar{V}$ とする。

FCM 3. $\min_{U\in M_{fcm}} J_{fcm}(U,\bar{V})$ の最適解を $\bar{U}$ とする。

FCM 4. 最適解 $(\bar{U},\bar{V})$ が収束していれば終了，そうでなければ FCM 2 に戻る。

これによって，クラスタリング結果のメンバシップ u_{ik} はファジィ化され，$\boldsymbol{x}_k$ が $\boldsymbol{v}_i$ に一致しなければ，$0<u_{ik}<1$ となる。また，この u_{ik} はスムージングパラメータ q に応じて変化することになる。一般に q の値が 1 に近いときにはクラスタリング結果は crisp k-means 法の結果に近いものとなり，q の値が大きくなればなるほど，クラスタリング結果はあいまいな結果が得られるようになる。

$\bar{V}$ が与えられたとき，ステップ FCM3 の解 $\bar{U}$ は

$$\bar{u}_{ik}=\left[\sum_{j=1}^{c}\left(\frac{d(\bar{\boldsymbol{v}}_i,\boldsymbol{x}_k)}{d(\bar{\boldsymbol{v}}_j,\boldsymbol{x}_k)}\right)^{\frac{1}{q-1}}\right]^{-1} \tag{2.3}$$

となり，$\bar{U}$ が与えられたとき，ステップ FCM2 の解 $\bar{V}$ は

$$\bar{v}_i = \frac{\sum_{k=1}^{n} (\bar{u}_{ik})^q x_k}{\sum_{k=1}^{n} (\bar{u}_{ik})^q} \tag{2.4}$$

で得られることが示される。

以上が，crisp k-means 法をファジィ化した fuzzy c-means 法である。見てきたように，fuzzy c-means 法は crisp k-means 法にスムージングパラメータ q を導入することにより，crisp k-means 法の解を近似する形で変形されていると見ることができる。第 1 部では，これを crisp k-means 法の正則化であると考え，crisp k-means 法の別の形の正則化（ファジィ化）について考える。

2.3　クラスタリング結果の比較

図 2.1 は，一様乱数によって図のような 2 次元上の領域に点を生成した人工的なデータである。このデータに対し crisp k-means 法と fuzzy c-means 法とによる 2 つのクラスタへの分割結果を図 2.2，図 2.3 に示す。fuzzy c-means 法のクラスタリング結果は $\alpha = 0.6$ で α-cut した時の結果であり，白い点はいずれのクラスタへのメンバシップ値も 0.6 以下である点である。このように crisp k-means 法では，すべての点は必ずいずれかのクラスタに所属し，同じクラスタに属する点にはその属性に違いはない。それに対し，fuzzy c-means 法のクラスタリング結果はメンバシップの値によって与えられる。同じクラスタに属する点でもそのメンバシップ値によって属性には違いがあり，各点のクラスタへの帰属度にあいまいさを含んでいる。実際，図 2.3 の 2 つのクラスタの境界近くの点はどちらのクラスタへのメンバシップ値も 0.5 に近く，かなりあいまいな結果となっていることが分かる。さらに，このクラスタリング結果のあいまいさの度合いをスムージングパラメータ $q(> 1)$ によって与えることができる。q を 1 に近い値にするとクリスプなクラスタリング結果と近い結果が得られ，q の値を大きくしてゆくとクラスタリング結果はしだいにあいまいなものになってゆく。

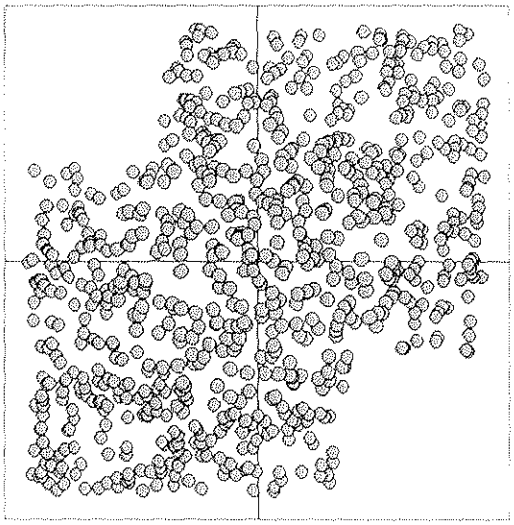

図 2.1: 2 次元上の人工的なデータ

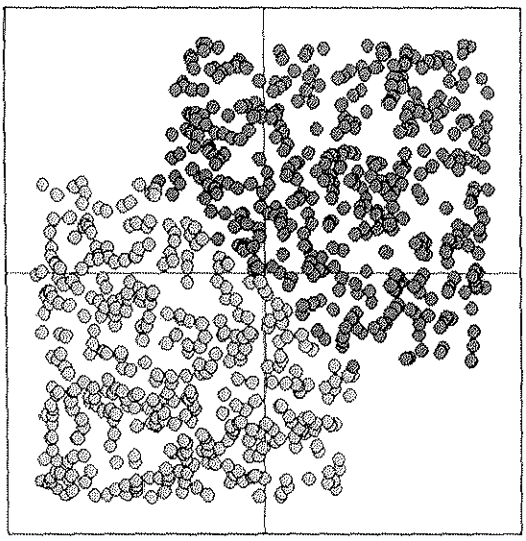

図 2.2: crisp k-means 法によるクラスタリング結果

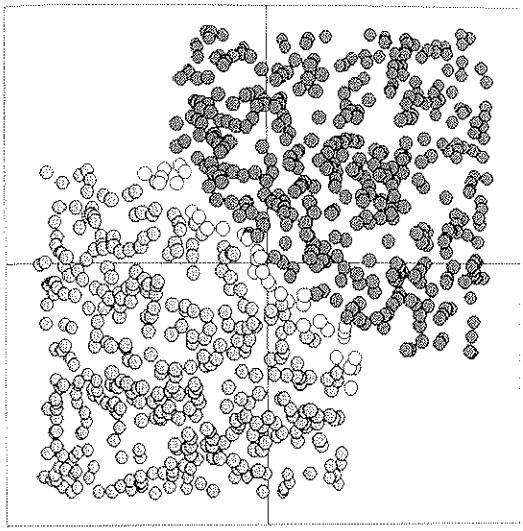

図 2.3: fuzzy c-means 法によるクラスタリング結果