

ファジィ c -平均法と DC アルゴリズム

楠木 祥文 (大阪大学)

Fuzzy c -means and DC Algorithm

*Y. Kusunoki (Osaka University)

Abstract: c -means (k -means) and fuzzy c -means are the most popular methods for cluster analysis. Fuzzy c -means methods as well as c -means methods are interpreted as optimization problems, which determine c cluster centers (centroids) minimizing error of the partition of objects (data) given by the centroids. That problems are solved by alternative minimization, where the positions of centroids and the assignments of objects are alternatively and iteratively determined. In this study, we reformulate the optimization problems of fuzzy c -means methods to DC (Difference of Convex functions) problems, and show that the procedures of alternative minimization coincide with an algorithm to solve the DC problems.

Key Words: DC algorithm, DC programming, Fuzzy c -means

1 はじめに

与えられたデータに内在するグループを発見するクラスタリングにおいて、 c -平均法 (k -平均法) やその拡張であるファジィ c -平均法は、最も良く用いられる手法の一つである。それらの手法では、データの分割に関する評価関数を最小化するために、クラスター中心と対象のクラスターへの帰属度を交互に決定する交互最適化のアプローチが取られる。本研究では、3 種類のファジィ c -平均法について、それらの交互最適化の手続きが、凸関数 f から凸関数 g を引いた関数の最小化問題である DC (Difference of Convex functions) 計画問題に対する DCA (DC Algorithm) と一致することを示す。DCA とは、において、現在の実行可能解で凹関数 $-g$ を線形近似した問題を繰り返し解き局所解を求める方法である。

2 準備

2.1 c -平均法

m 個の対象を c 個のクラスターに分割する問題を考える。対象とクラスターの添え字集合をそれぞれ $M = \{1, \dots, m\}$, $C = \{1, \dots, c\}$ とする。各対象 i の表現を $x_i \in \mathbb{R}^n$ とし、各クラスター j の中心を $w_j \in \mathbb{R}^n$ で表す。 c -平均法³⁾ はクラスター中心 w_1, \dots, w_c を用いて、次の目的関数値が小さくなるような対象集合の分割を求める手法である。

$$\begin{aligned} \min \quad & \sum_{j \in C} \sum_{i \in M} u_{ij} \|x_i - w_j\|^2 \\ \text{s.t.} \quad & \sum_{j \in C} u_{ij} = 1; \quad u_i \in \{0, 1\}^c, \quad i \in M \end{aligned} \quad (1)$$

ここで、 $u_i = (u_{i1}, u_{i2}, \dots, u_{ic}) \in \mathbb{R}^c$ であり、 u_{ij} は対象 i がクラスター j に属していれば 1 を取る決定変数である。また、 $\|\cdot\|$ はユークリッドノルムである。各クラスター j について、 $\sum_{i \in M} u_{ij} \|x_i - w_j\|^2$ を誤差と考え、その総和を最小化するように、クラスター中心 w_1, \dots, w_c と分割 u_1, u_2, \dots, u_m を求める。

c -平均法では次の交互最適化の手続きによって、上記の最適化問題が求解される。

- 各 $j \in C$ について $w_j = \bar{w}_j$ に固定し、問題 (1) を $(u_i)_{i \in M}$ について解く。最適解 $(\bar{u}_i)_{i \in M}$ は、各 $i \in M$ と各 $j \in C$ について以下のように定まる。

$$\bar{u}_{ij} = \begin{cases} 1 & \text{if } j = \operatorname{argmin}_{j' \in C} \|x_i - \bar{w}_{j'}\|^2 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

- 各 $i \in M$ について、 $u_i = \bar{u}_i$ に固定し、問題 (1) を $(w_j)_{j \in C}$ について解く。最適解 $(\bar{w}_j)_{j \in C}$ は、各 $j \in C$ について以下ようになる。

$$\bar{w}_j = \frac{\sum_{i \in M} \bar{u}_{ij} x_i}{\sum_{i \in M} \bar{u}_{ij}} \quad (3)$$

この二つの操作を、 $(w_j)_{j \in C}$ と $(u_i)_{i \in M}$ が収束するか、目的関数値が変化しなくなるまで繰り返す。

2.2 ファジィ c -平均法

c -平均法では帰属度 u_{ij} は 0 または 1 を取り、各対象 i はいずれか一つのクラスター j のみに帰属しなければならない。与えられたデータによっては、ある対象を複数のクラスターに帰属させることにより、よりデータの分布を捉えたクラスタリング結果を得ることができる。そのようなファジィ分割を求める方法としてファジィ c -平均法 (FCM: Fuzzy c -Means)³⁾ が提案されている。本節では 3 種類の FCM を導入する。FCM に対応して、前節の c -平均法を HCM (Hard c -Means) と呼ぶ。

一つ目は Bezdek と Dunn³⁾ によって提案された標準的な FCM である。それに対応する最適化問題は以下になる。

$$\begin{aligned} \min \quad & \sum_{i \in M} \sum_{j \in C} u_{ij}^{\nu_n} \|x_i - w_j\|^2 \\ \text{s.t.} \quad & \sum_{j \in C} u_{ij} = 1; \quad u_i \in [0, 1]^c, \quad i \in M \end{aligned} \quad (4)$$

ここで、 $\nu_n \geq 1$ である。HCM の $u_i \in \{0, 1\}^c$ が $u_i \in [0, 1]^c$ に緩和されている。また、目的関数の u_{ij} が $u_{ij}^{\nu_n}$ に置き換わっている。 $u_{ij}^{\nu_n}$ の効果により、 $0 < u_{ij} < 1$ となる解を得られるようになる。

この問題も HCM と同様に交互最適化の手続きで求解される。 $w_j = \bar{w}_j$, $j \in C$ に固定したとき、最適解 $(\bar{u}_i)_{i \in M}$ は以下のように求まる。各 $i \in M$ について、すべての $j \in C$ で $\|x_i - \bar{w}_j\|^2 > 0$ のときは、各 $j \in C$ について、

$$\bar{u}_{ij} = \left(\sum_{j' \in C} \left(\frac{\|x_i - \bar{w}_j\|^2}{\|x_i - \bar{w}_{j'}\|^2} \right)^{1/(\nu_n - 1)} \right)^{-1} \quad (5)$$

となり、それ以外の場合は、式 (2) と同様となる。 $u_i = \bar{u}_i$, $i \in M$ に固定したとき、最適解 $(\bar{w}_j)_{j \in C}$ は以下のように求まる。各 $j \in C$ について、

$$\bar{w}_j = \frac{\sum_{i \in M} \bar{u}_{ij}^{\nu_n} x_i}{\sum_{i \in M} \bar{u}_{ij}^{\nu_n}} \quad (6)$$

二つ目はエントロピー正則化に基づくファジィ c -平均法 (FCME: Fuzzy c -Means with Entropy regularization) である。その最適化問題は以下ようになる。

$$\begin{aligned} \min \quad & \sum_{i \in M} \sum_{j \in C} u_{ij} \|x_i - w_j\|^2 + \nu_e \sum_{i \in M} \sum_{j \in C} u_{ij} \log u_{ij} \\ \text{s.t.} \quad & \sum_{j \in C} u_{ij} = 1; u_i \in [0, 1]^c, i \in M \end{aligned} \quad (7)$$

ここで、 $\nu_e \geq 0$ である。FCM と同様に、 u_i の制約が緩和されており、さらに、目的関数に $\nu_e \sum_{i \in M} \sum_{j \in C} u_{ij} \log u_{ij}$ が加えられている。目的関数に加えられた項は正則化項と呼ばれ、 u_{ij} が極端な値 (端点 0, 1) を取ることを妨げる効果があり、結果として、最適解は $0 < u_{ij} < 1$ を取る傾向となる。

交互最適化の手続きは以下ようになる。 $w_j = \bar{w}_j$, $j \in C$ に固定して、 $(\bar{u}_i)_{i \in M}$ を求めると、各 $i \in M$ と各 $j \in C$ について、

$$\bar{u}_{ij} = \left(\sum_{j' \in C} \exp \left(\frac{-\|x_i - \bar{w}_j\|^2}{\nu_e} \right) \right)^{-1} \exp \left(\frac{-\|x_i - \bar{w}_j\|^2}{\nu_e} \right) \quad (8)$$

$u_i = \bar{u}_i$, $i \in M$ に固定して、 $(\bar{w}_j)_{j \in C}$ を求めると、式 (2) と同様となる。

三つ目は 2 次正則化に基づくファジィ c -平均法 (FCMQ: Fuzzy c -Means with Quadratic regularization) である。その最適化問題は以下ようになる。

$$\begin{aligned} \min \quad & \sum_{i \in M} \sum_{j \in C} u_{ij} \|x_i - w_j\|^2 + \frac{\nu_q}{2} \sum_{i \in M} \|u_i\|^2 \\ \text{s.t.} \quad & \sum_{j \in C} u_{ij} = 1; u_i \in [0, 1]^c, i \in M \end{aligned} \quad (9)$$

ここで、 $\nu_q \geq 0$ である。FCM と同様に、 u_i の制約が緩和されており、さらに、目的関数に正則化項 $\frac{\nu_q}{2} \sum_{i \in M} \|u_i\|^2$ が加えられている。

交互最適化の手続きは以下ようになる。 $w_j = \bar{w}_j$, $j \in C$ に固定して、 $(\bar{u}_i)_{i \in M}$ を求めるが、それは、陽な形式で表現できず、次の条件を満たすものとなる。各 $i \in M$ について、

$$\sum_{j \in C} \bar{u}_{ij} = 1, \bar{u}_{ij} = \max \left\{ 0, \frac{-\|x_i - \bar{w}_j\|^2}{\nu_q} - \mu \right\}, j \in C \quad (10)$$

ここで μ はある定数である。 $u_i = \bar{u}_i$, $i \in M$ に固定して、 $(\bar{w}_j)_{j \in C}$ を求めると、式 (2) と同様となる。

2.3 DC 計画問題と DC アルゴリズム

DC (Difference of Convex functions) 計画問題は、 $x \in \mathbf{R}^n$ を変数とする二つの凸関数 f, g の差で表される関数 $h(x) = g(x) - f(x)$ を最小化する問題である。このような関数 h を DC 関数と呼び、 $g - f$ の表現を h の DC 分解と呼ぶ。つまり、次のような問題を考える。

$$\min h(x) = g(x) - f(x) \quad (11)$$

凸制約 $x \in S$ を考える場合は、標示関数 $\delta_S(x) = 0$ if $x \in S$; $\delta_S(x) = \infty$ otherwise; を考え、 $h(x) = g(x) + \delta_S(x) - f(x)$ とすれば、上記の形式になる。また、ここでは、 f と g は閉真凸関数²⁾であると仮定する。

閉真凸関数 f に対して、 f の共役関数 f^* を

$$f^*(y) = \sup \{x^\top y - f(x) \mid x \in \mathbf{R}^n\}, y \in \mathbf{R}^n$$

と定義する。 f^* の共役関数は f となる²⁾。共役関数を用いて問題 (11) の双対問題 $\min f^*(y) - g^*(y)$ を与えることができる。このとき、 $\inf \{g(x) - f(x) \mid x \in \mathbf{R}^n\} = \inf \{f^*(y) - g^*(y) \mid y \in \mathbf{R}^n\}$ が成り立つ¹⁾。

任意の $x' \in \mathbf{R}^n$ に対して、 $f(x') \geq f(x) + y^\top (x' - x)$ を満たす y を x における f の劣勾配と呼ぶ。 x における f の劣勾配の集合を劣微分と呼び $\partial f(x)$ で表す。 f が x で微分可能なとき、劣微分は単集合となり、その要素は f の勾配 $\nabla f(x)$ と一致する。 x が DC 計画問題 (11) の局所最適解である必要条件として、次のものがある。

$$\partial f(x) \cap \partial g(x) \neq \emptyset \quad (12)$$

このような x は $g - f$ の critical point と呼ばれる¹⁾。DC 計画問題の解法の一つである DCA (DC Algorithm) は、最適性条件 (12) に基づき、critical point を見つける手続きである。DCA は以下のような手続きである。

- 現在の解を \bar{x} とし、 $f(x)$ の \bar{x} における劣勾配を $\bar{y} \in \partial f(\bar{x})$ とする。
- \bar{y} を用いて、 $f(x)$ を線形近似した問題を考え、その最適解を \bar{x} とする。つまり、 g の共役関数を g^* とすると、 $\bar{x} \in \partial g^*(\bar{y})$ となる。または、等価な表現として、

$$\bar{x} = \operatorname{argmin} \{g(x) - (f(\bar{x}) + \bar{y}^\top (x - \bar{x})) \mid x \in \mathbf{R}^n\}$$

上記の反復を \bar{x} と \bar{y} が収束するまで繰り返す。 $\{x^l\}$ と $\{y^l\}$ を DCA で生成された点列とすると、 $\{g(x^l) - f(x^l)\}$ と $\{f^*(y^l) - g^*(y^l)\}$ は各反復ごとに減少する。問題 (11) の最適値が有限で、点列 $\{x^l\}$ と $\{y^l\}$ が有界であれば、すべての $\{x^l\}$ (または $\{y^l\}$) の極限点 \bar{x} (または \bar{y}) は $g - f$ (または $f^* - g^*$) の critical point である¹⁾。

3 c -平均法と DC アルゴリズム

本節では、HCM と DC アルゴリズムとの関係を述べる。HCM の最適化問題 (1) から、制約を一部取り除いた問題を考える。

$$\begin{aligned} \min \quad & - \sum_{i \in M} u_i^\top t_i \\ \text{s.t.} \quad & \|x_i - w_j\|^2 + t_{ij} \leq 0, i \in M, j \in C \\ & \sum_{j \in C} u_{ij} = 1; u_i \geq 0, i \in M \end{aligned} \quad (13)$$

ここで, $t_i = (t_{i1}, \dots, t_{ic}) \in \mathbf{R}^c$ であり, $-\|x_i - w_j\|^2$ を t_{ij} に置き換えている. 制約 $u_i \in \{0, 1\}^c$ が $u_i \geq 0$ に緩和されているが, 交互最適化で \bar{w}_j を固定して u_{ij} を求める場合には, 問題 (1) の最適解は, 上記の (13) の最適解集合に含まれていることが分かる.

$t \in \mathbf{R}^c$ とし, $f_m(t) = \max\{t_j \mid j \in C\}$ を最大値関数とする. その共役関数は次のようになる. $u \in \mathbf{R}^c$ に対して,

$$f_m^*(u) = \begin{cases} 0 & \text{if } \sum_{j \in C} u_j = 1; u \geq 0 \\ \infty & \text{otherwise} \end{cases} \quad (14)$$

f_m^* を用いて, 問題 (13) を書き換えると,

$$\begin{aligned} \min \quad & - \sum_{i \in M} (u_i^\top t_i - f_m^*(u_i)) \\ \text{s.t.} \quad & \|x_i - w_j\|^2 + t_{ij} \leq 0, i \in M, j \in C \end{aligned}$$

共役関数の関係から, $f_m(t) = \sup\{u^\top t - f_m^*(u) \mid u \in \mathbf{R}^c\}$ であるので, HCM の最適化問題は

$$\begin{aligned} \min \quad & - \sum_{i \in M} \max\{t_{ij} \mid j \in C\} \\ \text{s.t.} \quad & \|x_i - w_j\|^2 + t_{ij} \leq 0, i \in M, j \in C \end{aligned} \quad (15)$$

となり, 凸制約上の凹関数最小化問題に帰着される.

関数 g を次のように定義する. $w \in \mathbf{R}^n, t_1, t_1, \dots, t_m \in \mathbf{R}$ に対して,

$$g(w, t_1, \dots, t_m) = \begin{cases} 1 & \|x_i - w\|^2 + t_i \leq 0, i \in M \\ \infty & \text{otherwise} \end{cases}$$

$\tilde{g}((w_j)_{j \in C}, (t_i)_{i \in M}) = \sum_{j \in C} g(w_j, t_{1j}, t_{2j}, \dots, t_{mj})$ と $\tilde{f}_m((t_i)_{i \in M}) = \sum_{i \in M} f_m(t_i)$ を定義すると, 問題 (15) は DC 分解を $\tilde{g}((w_j)_{j \in C}, (t_i)_{i \in M}) - \tilde{f}_m((t_i)_{i \in M})$ とする DC 計画問題である.

DCA の手続きとして, まず $\tilde{f}_m((t_i)_{i \in M}) = \sum_{i \in M} f_m(t_i)$ を線形近似した問題を考える. 各 $i \in M$ について, \bar{t}_i における f の劣勾配 \bar{u}_i は, 関数 $\bar{t}_i^\top u_i - f_m^*(u_i)$ を最大化する u_i として求めることができる. $\bar{t}_{ij} = -\|x_i - \bar{w}_j\|^2, j \in C$ となっている場合, そのような u_i を求めることは, HCM で $(\bar{w}_j)_{j \in C}$ を固定して \bar{u}_i を求める手続きと一致する. つまり, 式 (2) の \bar{u}_i は関数 f_m の \bar{t}_i における劣勾配であることが分かる.

次に, DCA では, 線形化した目的関数 $-\sum_{i \in M} \bar{u}_i^\top t_i + \tilde{g}((w_j)_{j \in C}, (t_i)_{i \in M})$ を最小化する \bar{w}_j と \bar{t}_i を求めるが, これも, HCM の \bar{u}_i を固定して $(\bar{w}_j)_{j \in C}$ を求める手続きと一致する. このとき, $\bar{t}_{ij} = -\|x_i - \bar{w}_j\|^2, i \in M, j \in C$ とすることができる. したがって, HCM と DCA が等価であることが分かる.

4 ファジィc-平均法と DC アルゴリズム

上述の 3 種類のファジィc-平均法も, 適当な DC 計画問題を考えることにより, DCA として見ることできる.

4.1 FCM

FCM の問題 (4) を以下のように変形する.

$$\begin{aligned} \min \quad & - \sum_{i \in M} u_i^\top t_i \\ \text{s.t.} \quad & \|x_i - w_j\|^2 + t_{ij} \leq 0, i \in M, j \in C \\ & \|u_i\|_{1/\nu_n} \geq 1; u_i \geq 0, i \in M \end{aligned} \quad (16)$$

ここで, $u_{ij}^{\nu_n}$ を u_{ij} に置き直すことにより, $\sum_{j \in C} u_{ij} = 1$ は $\sum_{j \in C} u_{ij}^{1/\nu_n} = 1$ に置き換えられる. また, u_{ij} を小さくなるほど, 目的関数値は小さくなるので, $\sum_{j \in C} u_{ij}^{1/\nu_n} = 1$ を $\sum_{j \in C} u_{ij}^{1/\nu_n} \geq 1$ とし, $u_i \in [0, 1]^c$ を $u_i \geq 0$ としても, 最適解集合は大きくなるのみである. さらに, 前者の制約の両辺を ν_n 乗することで, $\|u_i\|_{1/\nu_n} \geq 1$ を得る.

関数 f_n を次のように定義する. $t \in \mathbf{R}_-^c$ に対して,

$$f_n(t) = \begin{cases} -\|t\|_{\nu_n-1} & \text{if } t \in \mathbf{R}_-^c \\ 0 & \text{if } t \in \mathbf{R}_-^c \setminus \mathbf{R}_-^c \end{cases} \quad (17)$$

ここで, \mathbf{R}_- は非正の値をとる実数ベクトルの集合であり, \mathbf{R}_-^c は負の値をとる実数ベクトルの集合である. この関数の共役関数は, $u \in \mathbf{R}^c$ に対して,

$$f_n^*(u) = \begin{cases} 0 & \text{if } \|u\|_{1/\nu_n} \geq 1; u \geq 0 \\ \infty & \text{otherwise} \end{cases} \quad (18)$$

である. したがって, 共役関数の関係を用いると, 問題 (16) を以下のように変形できる.

$$\begin{aligned} \min \quad & - \sum_{i \in M} f_n(t_i) \\ \text{s.t.} \quad & \|x_i - w_j\|^2 + t_{ij} \leq 0, i \in M, j \in C \end{aligned} \quad (19)$$

DCA の手続きから, $f_n(t_i)$ の劣勾配が帰属度 u_i となるが, $t_i \in \mathbf{R}_-^c$ の領域で $f_n(t_i)$ は微分可能であり, その勾配は, $p = 1/(1 - \nu_e)$ とおくと, $u_i = \nabla f_n(t_i) = \|-t_i\|_{\nu_n-1}^p ((1/t_{i1})^p, \dots, (1/t_{ic})^p)$ となり, 式 (5) と一致する.

4.2 FCME

FCME の問題 (7) を以下のように変形する.

$$\begin{aligned} \min \quad & - \sum_{i \in M} u_i^\top t_i + \nu_e \sum_{i \in M} \sum_{j \in C} u_{ij} \log u_{ij} \\ \text{s.t.} \quad & \|x_i - w_j\|^2 + t_{ij} \leq 0, i \in M, j \in C \\ & \sum_{j \in C} u_{ij} = 1; u_i \geq 0, i \in M \end{aligned} \quad (20)$$

ここで, u_{ij} を小さくなるほど, 目的関数値は小さくなるので, $u_i \in [0, 1]^c$ を $u_i \geq 0$ としても最適解集合は大きくなるのみである.

関数 f_e を次のように定義する. $t \in \mathbf{R}^c$ に対して,

$$f_e(t) = \nu_e \log \left(\sum_{j \in C} \exp(t_j/\nu_e) \right) \quad (21)$$

この関数の共役関数は, $u \in \mathbf{R}^c$ に対して,

$$f_e^*(u) = \begin{cases} \nu_e \sum_{j \in C} u_j \log u_j & \text{if } \sum_{j \in C} u_j = 1; u \geq 0 \\ \infty & \text{otherwise} \end{cases} \quad (22)$$

である. したがって, 共役関数の関係を用いると, 問題 (20) を以下のように変形できる.

$$\begin{aligned} \min \quad & - \sum_{i \in M} \nu_e \log \left(\sum_{j \in C} \exp(t_{ij}/\nu_e) \right) \\ \text{s.t.} \quad & \|x_i - w_j\|^2 + t_{ij} \leq 0, i \in M, j \in C \end{aligned} \quad (23)$$

最大値関数 f_m を考えると, $f_e^*(u)$ は $f_m^*(u)$ に, ν_e を乗じて $h_e(u) = \sum_{j \in C} u_j \log u_j$ を加えたものと見ることができ, 共役関数の関係を用いると, $f_e = (f_m^* + \nu_e h_e)^*$ となる. このような操作は関数を滑らかにする効果があり正則化と呼ばれる. FCME は, 最大値関数 f_m を滑らかな関数 f_e に置き換えた問題を解いているとみなすことができる.

DCA の手続きから, $f_e(t_i)$ の劣勾配が帰属度 u_i となるが, 正則化された $f_e(t_i)$ は微分可能であり, その勾配は $\nabla f_e(t_i)$ は式 (8) と一致する.

パラメータ ν_e の値が大きいときの $f_e(t)$ について考える. 関数 $f_e(t)$ は次の最適化問題の最適値で与えられる.

$$\begin{aligned} \max \quad & t^\top u - \nu_e \sum_{j \in C} u_j \log u_j \\ \text{s.t.} \quad & \sum_{j \in C} u_j = 1; u \geq 0, i \in M \end{aligned}$$

ν_e の値が大きいときは, 目的関数の $\sum_{j \in C} u_j \log u_j$ が強く影響するため, 最適解は $(1/c, 1/c, \dots, 1/c)$ に近づく. したがって, 最適値は $(1/c)(\sum_{j \in C} t_j - \nu_e \log c)$ に近づき, $f_e(t)$ は t のアフィン関数に近づく.

4.3 FCMQ

FCMQ の問題 (9) を以下のように変形する.

$$\begin{aligned} \min \quad & - \sum_{i \in M} u_i^\top t_i + \frac{\nu_q}{2} \sum_{i \in M} \|u_i\|^2 \\ \text{s.t.} \quad & \|x_i - w_j\|^2 + t_{ij} \leq 0, i \in M, j \in C \\ & \sum_{j \in C} u_{ij} = 1; u_i \geq 0, i \in M \end{aligned} \quad (24)$$

ここで, u_{ij} を小さくなるほど, 目的関数値は小さくなるので, $u_i \in [0, 1]^c$ を $u_i \geq 0$ としても最適解集合は大きくなるのみである.

関数 f_q^* を次のように定義する. $u \in \mathbb{R}^c$ に対して,

$$f_q^*(u) = \begin{cases} \frac{\nu_q}{2} \|u\|^2 & \text{if } \sum_{j \in C} u_j = 1; u \geq 0 \\ \infty & \text{otherwise} \end{cases} \quad (25)$$

である. この共役関数 f_q は陽に求めることはできないが, 共役関数の関係から, $t \in \mathbb{R}^c$ について $f_q(t) = \sup_{u \in \mathbb{R}^c} \{t^\top u - f_q^*(u)\}$ の関係によって, 数値的に求めることができる. $t^\top u - f_q^*(u)$ を最大にする u は式 (10) で与えられる. また, これは u_i が勾配 $\nabla f_q(t_i)$ である必要十分条件である.

関数 f_q を用いると, 問題 (24) は以下のように表現できる.

$$\begin{aligned} \min \quad & - \sum_{i \in M} f_q(t_i) \\ \text{s.t.} \quad & \|x_i - w_j\|^2 + t_{ij} \leq 0, i \in M, j \in C \end{aligned} \quad (26)$$

FCME のときと同様に, 最大値関数 f_m を考えると, $f_q = (f_m^* + \nu_q h_q)^*$ を得る. h_q による正則化は Moreau-Yoshida 正則化と呼ばれている. パラメータ ν_q の値が大きいときは, $f_e(t)$ と同様に, $f_q(t)$ も t のアフィン関数に近づく.

4.4 FCM, FCME, FCMQ の比較

FCME と FCMQ は FCM の問題の最大値関数 f_m を, それぞれ, 正則化された f_e と f_q に置き換えた問題を考えられている. FCM では, f_m がソルム型の関数 f_n に置き換わっている. FCM, FCME, FCMQ の違いは, f_n, f_e, f_q の違いにのみ依存する. 2 変数 t_1, t_2 の場合, f_m とこの三つの関数の等高線を描くと Fig. 1 のようになる. 図の 1 行目は f_m の等高線を図示している. 2, 3, 4 行目は, それぞれ, $\nu_n = 2, 10, 20, \nu_e = 10, 20, 30, \nu_q = 20, 40, 60$ としたときの, 関数 f_n, f_e, f_q の等高線である. 各図の縦軸と横軸は t_1 と t_2 を表しており, その範囲は -100 から 0 である.

f_n は, ほかの二つの関数と比べると, パラメータによる等高線の変化が小さい. また, 各図の右下と左上では, 等高線が軸と平行になっており, さらに, 勾配が大きくなっている. f_e と f_q では, パラメータが大きくなるほど等高線が丸くなっていくことが分かる. f_e と f_q を比較すると, f_e では図の右下と左上が曲がっているのに対して, f_q では軸に平行となっている. つまり, クラスター間の距離がある程度大きい場合は, FCME と比較して, FCMQ は HCM に近い挙動を取ると考えられる.

Fig. 2 は, 人工データに対して, HCM, FCM, FCME, FCMQ を適用した結果を図示している. データは異なる平均と分散を持つ三つの正規分布から生成されており, 三つのクラスターがあると仮定する. 図の 1 行目に HCM の結果を図示し, 2, 3, 4 行目には, それぞれ, $\nu_n = 2, 10, 20, \nu_e = 10, 20, 30, \nu_q = 20, 40, 60$ としたときの, FCM, FCME, FCMQ の結果を図示している. 各図には, 各手法に対応する最適化問題の目的関数の等高線も描いている.

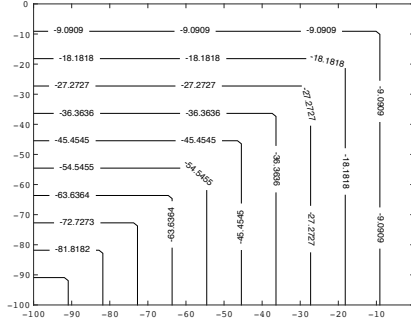
FCM の結果を見ると, 関数 f_n の議論から分かるように, パラメータを変化させてもクラスタリング結果はそれほど変化しない. FCME と FCMQ では, パラメータを大きくすると, 左上の二つのクラスター中心が同じ場所に移動することが分かる. これは, パラメータの大きくすると, $f_e(t)$ と $f_q(t)$ は, t_1, t_2, \dots, t_c の差が小さい場合, 和 $\sum_{j \in C} t_j$ に近づくためであると考えられる. 右下のクラスター中心に注目すると, FCME では, パラメータが大きくなるにつれて, それが左上に移動しているが, FCMQ では, その位置は変化していない. これは, Fig. 1 で, f_e では t_1 と t_2 の差が大きい領域でも, 等高線が曲がっており, あるクラスター中心からある程度遠い対象もその中心の配置に影響するのに対して, f_q では, 等高線は f_m と同様に軸に平行になっており, 遠い対象は中心の配置に影響しないためであると考えられる.

5 おわりに

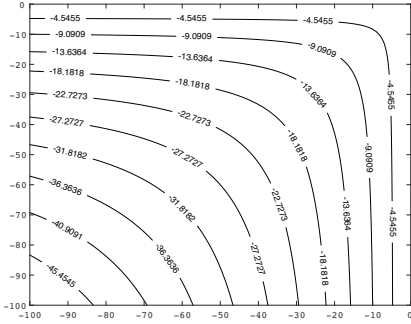
本研究では, 3 種類のアフィン平均法: FCM, FCMQ, FCME に対する交互最適化の手続きが, それぞれ, DC 計画問題 (19), (23), (26) に対する DCA と一致することを示した. 今後の課題として, 混合正規分布に対する EM アルゴリズム, KL 情報量に基づくファジィ c-平均法, ファジィ c-多様体法なども DCA に帰着できることを示す.

参考文献

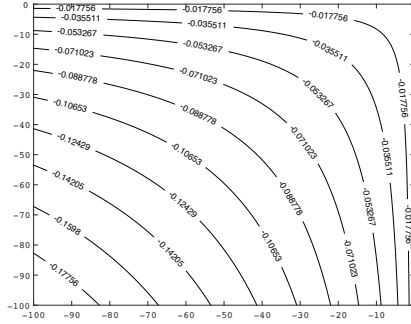
- 1) H. A. L. Thi, T. P. Dinh: The DC Programming and DCA Revisited with DC Models of Real World Nonconvex Optimization Problems, Annals of Operations Research, **133**, 23/46 (2005)
- 2) 福島雅夫: 非線形最適化の基礎, 朝倉書店 (2001)
- 3) S. Miyamoto, H. Ichihashi, K. Honda: Algorithms for Fuzzy Clustering, Springer (2008)



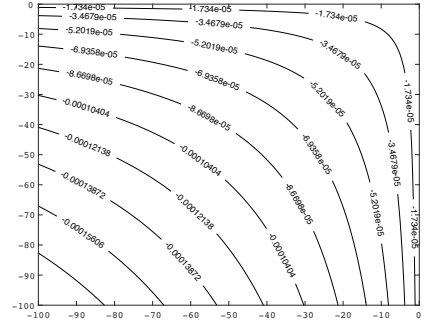
(a) f_m



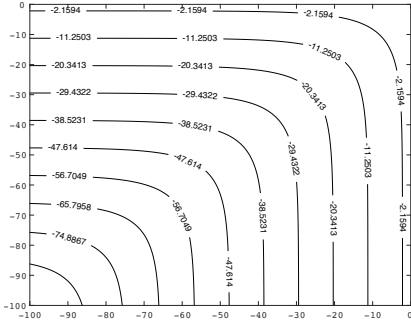
(b) f_n with $\nu_n = 2$



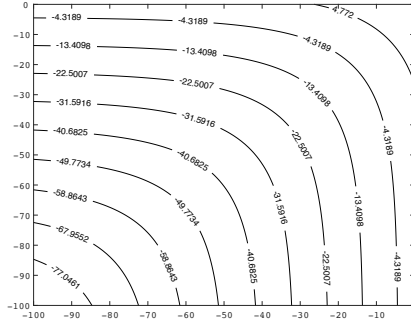
(c) f_n with $\nu_n = 10$



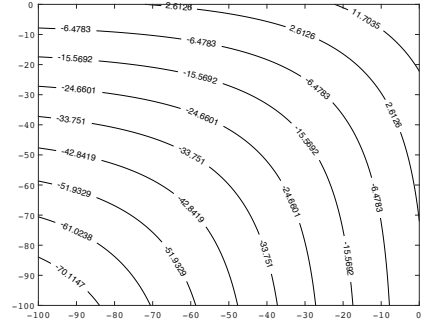
(d) f_n with $\nu_n = 20$



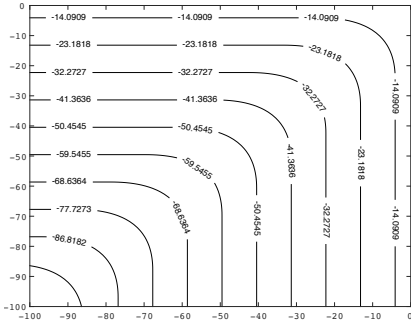
(e) f_e with $\nu_e = 10$



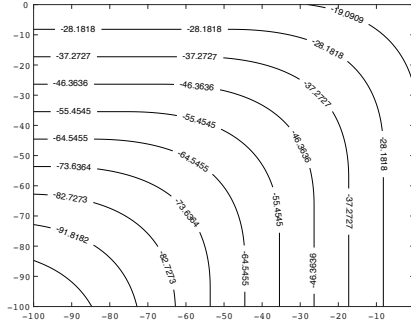
(f) f_e with $\nu_e = 20$



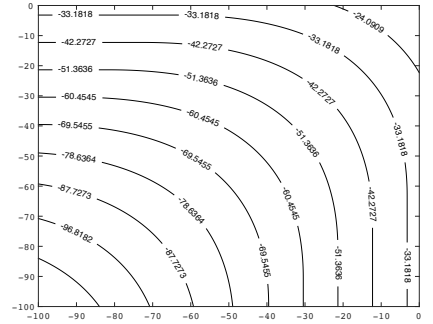
(g) f_e with $\nu_e = 30$



(h) f_q with $\nu_q = 20$



(i) f_q with $\nu_q = 40$



(j) f_q with $\nu_q = 60$

Fig. 1: Comparison among f_m , f_n , f_m and f_q with different parameters. The functions of 2 variables are shown with contour lines.

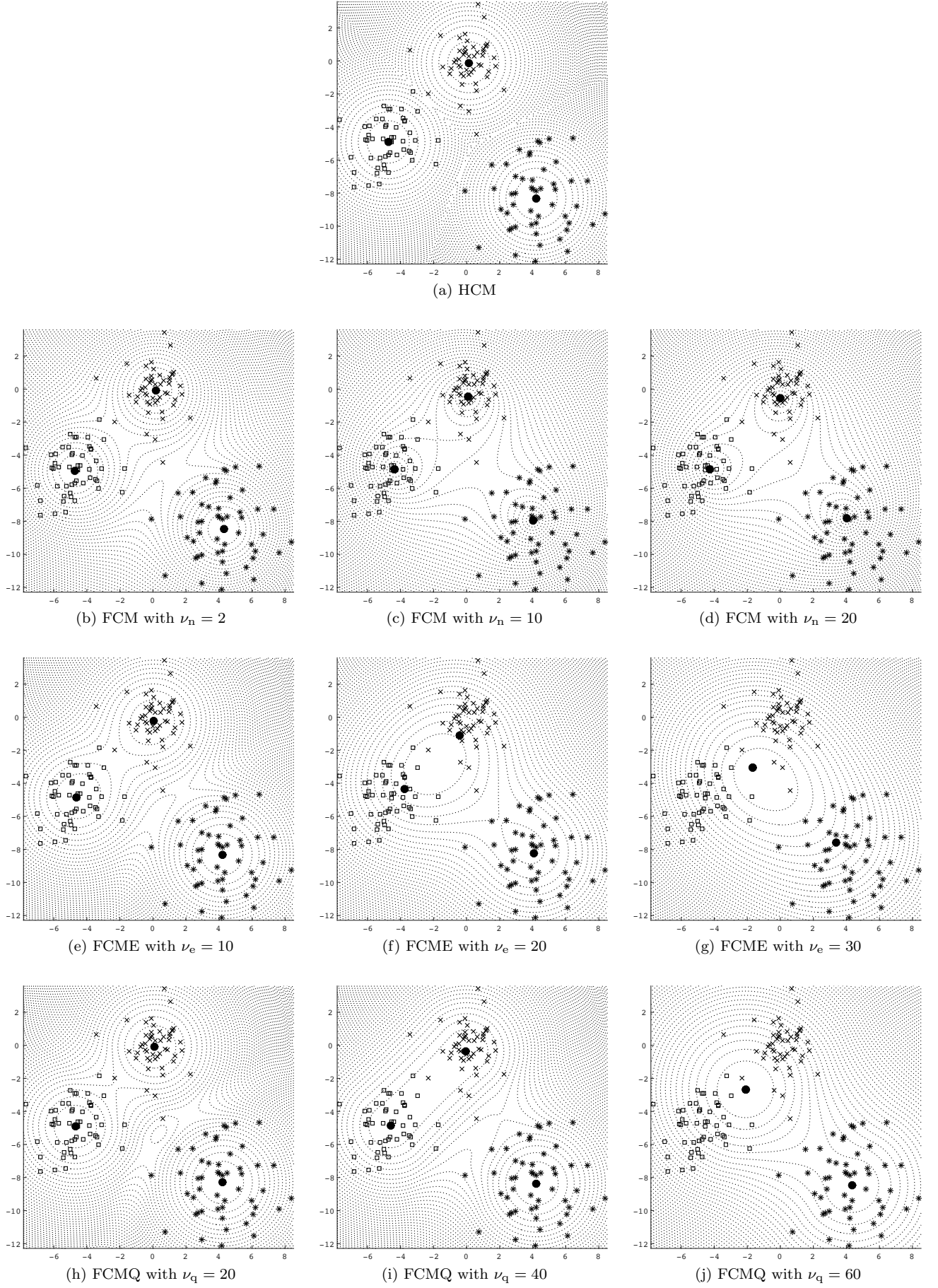


Fig. 2: Comparison among HCM, FCM , FCME and FCMQ with different parameters. Assume that there are 3 true clusters whose objects are marked by \times , $*$ and \square . There are 3 centroids marked by \bullet . The objective functions of each methods are shown by contour lines.