# 第4章

# 2次正則化によるファジィクラスタリング

第1章では，正則化の概念にもとづいて，エントロピー関数を用いたファジィクラスタリングを提案し，標準的な fuzzy $c$-means 法も crisp $k$-means 法の一種の正則化であるという考察を行った。本章では正則化のアイデアを更に推し進めることによって，正則化項としてメンバシップの2次関数を用いる2次正則化法を提案する。この新しい手法では，クラスタの中心の計算は 通常の fuzzy $c$-means 法と似ているが，メンバシップ値の計算は異なった形となる。ここでは，メンバシップ計算のための効率のよいアルゴリズムを提案する。さらに fuzzy $c$-means 法から導かれるクラスタのファジィ分類関数を2次正則化手法によって求める。この分類関数は区分線形型となるという著しい性質があることが示される。

## 4.1 はじめに

ファジィクラスタリングの諸手法は，理論と応用の両面において現在広く研究されている。なかでも，fuzzy $c$-means 法 [11, 12, 10] は，最も良く知られた方法であり，多くの変形や応用例が報告されている。

標準的な fuzzy $c$-means 法は crisp な $k$-means 法 [5, 6] にパラメータを導入してファジィ化したものとみなすことができる。第3章では crisp $k$-means 法を別の方法によりファジィ

化したエントロピー正則化法を提案し，エントロピー正則化法も，標準的な fuzzy $c$-means 法も crisp $k$-means 法の異なる形での正則化となっていることを示した [39]。この意味で，エントロピー正則化法も，広義の fuzzy $c$-means であるといえる。

本章では 先の2種類の正則化とは異なる crisp $k$-means の正則化である2次正則化法を提案する。これにより，正則化の方法によって fuzzy c-means のさまざまなバリエーションが考えられることが示される。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であり，3章ではエントロピー関数を用いた正則化 [39] を考えた。そこでこの章では，メンバシップ値の2次関数を用いた他の正則化を考察し，この方法によりメンバシップを計算するアルゴリズムを開発する。

さらに，標準的な fuzzy $c$-means 法 から導かれるファジィ分類関数 [33] と比べ，2次正則化法から得られるファジィ分類関数が区分線形メンバシップという顕著な特徴があることを示す。

このように，従来の標準的な fuzzy $c$-means 法およびエントロピー正則化法に加えて，ここで述べる第3の fuzzy $c$-means 法が得られるのである。

## 4.2 エントロピー正則化法

正則化という概念はこれら以外の正則化が考えられる点で役に立つ。3章で見たはエントロピー正則化法 [39] は crisp $k$-means 法の典型的な正則化となっている。

関数 $J$ の最適化問題の典型的な正則化は，正則化関数 $K$ を加えた，

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

という最適化問題となり，$\tilde{J}$ の最小化を $J$ の最小化のかわりに考察する。$\alpha(> 0)$ は正則化パラメータである。

エントロピー正則化法では $J = J_{ckm}$，

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

さらに計算の便宜のために，$\alpha = \lambda^{-1}$ とし，

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

という目的関数を用いた。これによってメンバシップの計算アルゴリズム [39] は標準的な fuzzy c-means 法とは異なったものになる。ここでは正則化 [13] においてより広く利用される2次関数を用いることにより，さらに新しいクラスタリングアルゴリズムを考えよう。

## 4.3　2次関数を用いた新たな目的関数

新たな正則化のアイデアでは $J = J_{ckm}(U, V)$,

$$K = \frac{1}{2}\sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^2$$

という2次関数を用いた正則関数を考える。さらに，$\alpha = \lambda^{-1} > 0$ とし，次のような目的関数を得る。

$$J_{quad}(U, V) = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}d_{ik} + \frac{1}{2}\lambda^{-1}\sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^2$$

ここで $K$ は，エントロピー正則化法の $K$ 同様，すべての $u_{ik}$ が $c^{-1}$ のときに最小になり，$\lambda$ が小さくなればなるほどこの項の影響でメンバシップはよりファジィ化されることになる。

最適化の2段階の手続き **FCM** は同様に用いることができる。ただし，評価関数は $J_{quad}$，メンバシップ行列 $U$ の許容集合は $M_{fcm}$ を用いるものとする。

FCM2 の $V$ の最適化では，$K$ がエントロピー関数でも2次関数でも $K$ は $V$ を含まないので，$K$ の違いは最適化には影響せず，エントロピー正則化法と同じとなる。つまり，$\bar{U}$ が固定された時の $V$ に対する $J_{quad}$ の最適化は，

$$\min_{V} J_{ckm}(\bar{U}, V)$$

となり，$\bar{V}$ は以下の式で求められる。

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

次に，FCM3 の $U$ に対する最適化には新しい計算方法が必要である。そのために，最適解の性質を調べよう。なお，証明はすべて補足に記す。

## 4.4　メンバシップ計算の原始的アルゴリズム

メンバシップの計算を考える前に，問題を単純化して考えよう。目的関数 $J_{quad}(U,V)$ は，

$$J_{(k)}(U,V) = \sum_{i=1}^{c} u_{ik} d_{ik} + \frac{1}{2}\lambda^{-1} \sum_{i=1}^{c} u_{ik}^2$$

とおくと，$J_{quad} = \sum_{k=1}^{n} J_{(k)}$ であるから，各 $J_{(k)}$ は他の $J_{(k')}$ とは独立して最小化することができる。

そこで今後，任意に固定された $k$ についての問題を考えることにし，$k$ を省略することがある。

**定義 4.1** 任意の $U(\in M_{fcm})$ に対し，$U$ の第 $k$ 列ベクトルを $U_k$ とする。このとき，

$$I[U_k] = \{i | u_{ik} > 0\}$$

とし，$I[U_k]$ を $U_k$ の**正値添字集合**あるいは単に，**添字集合**と呼ぶ。

**命題 4.2** FCM3 で求める最適解を $\bar{U} = (\bar{u}_{ik})$, $\bar{I}_k = I[\bar{U}_k]$ とすると,

$$\bar{u}_{ik} = \begin{cases} (\lambda \sum_{\ell \in \bar{I}_k} d_{\ell k} + 1)/|\bar{I}_k| - \lambda d_{ik}, & i \in \bar{I}_k \\ 0, & i \notin \bar{I}_k \end{cases} \tag{4.1}$$

を満たす。ただし $|\bar{I}_k|$ は，集合 $\bar{I}_k$ の要素数を表す。

したがって，各 $k$ に対し最適解の正値添字集合 $\bar{I}_k$ が見つかれば (4.1) 式によりメンバシップ値は求まるので，問題は最適な添字集合 $\bar{I}_k$ をいかに見つけるかということになる。

次の定義は適切な添字集合の選択方法について述べたものである。

**定義 4.3** 集合 $I = \{i_1, i_2, \cdots, i_q\}$ $(\subseteq \{1, 2, \cdots, c\})$ が，任意の $i \in I$ に対し,

$$(\lambda \sum_{\ell \in I} d_{\ell k} + 1) \ / \ |I| - \lambda d_{ik} > 0$$

を満たすときこの集合 $I$ を，**許容添字集合**といい，この時の解を $U_k(I)$，目的関数の値を $J_{(k)}(U_k(I))$ と書く。また許容添字集合 $I$ が，**最適添字集合**であるとは，目的関数の値 $J_{(k)}(U_k(I))$ がすべての許容添字集合の目的関数の中で最適解となっていることである。最適添字集合を $\bar{I}$ と書くことにする。

正確には，許容添字集合 $I$ が与えられたとき,

$$u_{ik} = \begin{cases} (\lambda \sum_{\ell \in I} d_{\ell k} + 1)/|I| - \lambda d_{ik}, & i \in I \\ 0, & i \notin I \end{cases}$$

は最適解であるとは限らないが，明らかに実行可能解である。そこでこれを，$U_k(I) = (u_{1k}, u_{2k}, \cdots, u_{ck})^\top$ と書く。

本来，許容添字集合，最適添字集合は $k$ に依存しているが，簡単のために任意に固定された $k$ について考えているものとして，省略することにする。

もし $I$ が許容添字集合でないとすると，命題 4.2 より最適性の必要条件を満たさないから，$I$ は最適解の添字集合ではない。よって，最適解を求めるためには，$\{1, 2, \cdots, c\}$ のすべての部分集合を試し，最適添字集合を発見すればよい。

原始的アルゴリズム

Step 1 $W = \{I_j | I_j \subseteq \{1, 2, \cdots, c\}\}$ とする。

Step 2 各 $I_j \in W$ に対し、以下を実行。

Step 3 $I_j$ が許容添字集合でなければ，Step 2 へ。

Step 4 許容添字集合 $I_j$ のメンバシップ値 $U_k(I_j)$ 及び目的関数 $J_{(k)}(U_k(I_j))$ を計算し，これまでに求められた最良の値より良ければ最良解を更新する。

Step 5 Step 2 へ戻る。

明らかに，このアルゴリズムは非効率的である。次に，最適解がより能率的なアルゴリズムによって求めることができることを示そう。

## 4.5 メンバシップ計算の効率的アルゴリズム

2次正則化法では，メンバシップ値を計算するために，各点ごとに独立してメンバシップを計算することができ，さらに添字集合が分かればメンバシップ値を計算できることを示した。これにより，すべての添字集合に対してメンバシップ値を計算し，最適解が求められることがわかった。

この節ではさらに，個体 $x_k$ に近い中心 $v_i$ に対する添字集合から順次調べれば十分であることを示し，効率的なアルゴリズムを与えよう。

今後、任意に固定された $k$ に対して

$$d_{k_1,k} \leq d_{k_2,k} \leq \cdots \leq d_{k_c,k}$$

となるように添え字 $k_i$ を付け，これを単に，

$$d_{1k} \leq d_{2k} \leq \cdots \leq d_{ck} \tag{4.2}$$

と書くことにする。

**命題 4.4** $I(i)$ を $i$ 以下の正の整数の集合

$$I(i) = \{1, 2, \cdots, i\}$$

とする。このとき，$I(K)$ が最適添字集合となるような $K$ $(\leq c)$ が存在する。

このように，最適解を見つけるためには，添字集合 $I(1), I(2), \cdots, I(c)$ を考えれば十分である事がわかる。次に，さらに最適解を絞り込むことができることを示そう。

**定義 4.5** 任意の $k$ $(1 \leq k \leq n)$，$L$ $(1 \leq L \leq c)$，$\ell$ $(1 \leq \ell \leq L)$ に対し，$f_{(k)\ell}^{L}$ を次のように定義する。

$$f_{(k)\ell}^{L} = \lambda \left[ \sum_{j=1}^{L} (d_{jk} - d_{\ell k}) \right] + 1 \tag{4.3}$$

ここでは特定の $k$ について考えているので，以降 $f_{(k)\ell}^{L}$ を単に $f_{\ell}^{L}$ と書くことにする。

**補題 4.6** 以下の性質 (a1), (a2) が成り立つ。

(a1) 任意の $L$ $(1 \leq L < c)$ に対し，

$$f_{\ell}^{L+1} \geq f_{\ell}^{L}, \quad \ell = 1, 2, \cdots, L$$

(a2) $$f_1^1 \geq f_2^2 \geq \cdots \geq f_c^c$$

**補題 4.7** 以下の性質 (b1), (b2), (b3) が成り立つ。

(b1) $I(1)$ は常に許容添字集合である。

(b2) $I(L-1)$ $(1 < L \leq c)$ が許容添字集合で $f_L^L > 0$ ならば，$I(L)$ もまた許容添字集合である。

(b3) $f_{L'}^{L'} \leq 0$ となる $L'(1 < L' \leq c)$ が存在すれば，$I(L'), I(L'+1), \cdots, I(c)$ は許容集合ではない。

これによって効率的なアルゴリズムを示すことができる。

**命題 4.8** $J_{(k)}$ を最小化する解は次の効率的アルゴリズムによって求められる。

効率的アルゴリズム

**Step 1** (4.3) 式により $f_{L'}^{L}, (L = 1, 2, \cdots, c)$ を計算する。

**Step 2** $\bar{L}$ を $f_{\bar{L}+1}^{\bar{L}+1} \leq 0$ を満たすもっとも小さい値とする。

**Step 3** $u_{ik}$ を，各 $i\,(1 \leq i \leq \bar{L})$ に対し，

$$u_{ik} = \bar{L}^{-1}(\lambda \sum_{\ell=1}^{\bar{L}} d_{\ell k} + 1) - \lambda d_{ik}$$

$i\,(\bar{L} + 1 \leq i \leq c)$ に対し，

$$u_{ik} = 0$$

とする。

## 4.6　ファジィ分類関数

fuzzy $c$-means 法ではクラスタリングの結果，すべての点に対して定義されるファジィ分類関数が得られる。標準的な fuzzy $c$-means 法では，クラスタ $i$ に対し，

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

という分類関数が Keller[33] により導入されている。この関数は $\hat{u}_i(\boldsymbol{x}_k) = u_{ik}$, $k = 1, 2, \cdots, n$ を満たし，$\hat{u}_i(\boldsymbol{x})$, $i = 1, 2, \cdots, c$ が与えられた点だけでなく全空間のファジィ分割を補間的に規定している。この分類関数はメンバシップを求める式の $\boldsymbol{x}_k$ を 変数 $\boldsymbol{x}$ で置き換えるだけで求められる。

2 次正則化によるクラスタリングの分類関数 $u_i(\boldsymbol{x})$ は以下のように求められる。

$\boldsymbol{x}$ と各 $\boldsymbol{v}_j$ の距離の小さい順に添え字 $j_i$ を付け，

$$d(\boldsymbol{x}, \boldsymbol{v}_{j_1}) \leq d(\boldsymbol{x}, \boldsymbol{v}_{j_2}) \leq \cdots \leq d(\boldsymbol{x}, \boldsymbol{v}_{j_c})$$

とする。各 $L = 1, 2, \cdots, c$ に対し，

$$f_L^L(\boldsymbol{x}) = \lambda \sum_{\ell=1}^{L} (d(\boldsymbol{x}, \boldsymbol{v}_{j_\ell}) - d(\boldsymbol{x}, \boldsymbol{v}_{j_L})) + 1$$

を計算し，$\bar{L}$ を $f_{L+1}^{L+1}(\boldsymbol{x}) \leq 0$ を満たす最小の $L$ とする。このとき，$i = 1, 2, \cdots, \bar{L}$ に対し，

$$u_{j_i}(\boldsymbol{x}) = \bar{L}^{-1}(\lambda \sum_{\ell=1}^{\bar{L}} d(\boldsymbol{x}, \boldsymbol{v}_{j_\ell}) + 1) - \lambda d(\boldsymbol{x}, \boldsymbol{v}_{j_i})$$

とし，$i = \bar{L}+1, \bar{L}+2, \cdots, c$ に対し，

$$u_{j_i}(\boldsymbol{x}) = 0$$

とする。この関数は，$u_{j_i}(\boldsymbol{x}) > 0$ のとき，

$$\begin{aligned}
&\lambda^{-1}(\bar{L} u_{j_i}(\boldsymbol{x}) - 1) \\
&= \sum_{\ell=1}^{\bar{L}} (\|\boldsymbol{x} - \boldsymbol{v}_{j_\ell}\|^2 - \|\boldsymbol{x} - \boldsymbol{v}_{j_i}\|^2) \\
&= \sum_{\ell=1}^{\bar{L}} (2\langle \boldsymbol{v}_{j_i} - \boldsymbol{v}_{j_\ell}, \boldsymbol{x} \rangle + \|\boldsymbol{v}_{j_\ell}\|^2 - \|\boldsymbol{v}_{j_i}\|^2)
\end{aligned}$$

を描いている。ただし，$\langle \cdot, \cdot \rangle$ は内積を表す。このように分類関数 $u_{j_i}(\boldsymbol{x})$ は区分的線形という注目すべき特徴があることがわかる。

## 4.7　クラスタリング例

図 4.1 の左側にある平面上に分布した 792 個の点を 2 つのクラスタに 2 次正則化法によってクラスタリングした結果が右に示されている。各点のメンバシップの初期値は乱数によっ

て与えられている。この時のパラメータ $\lambda$ は 1 であり，クラスタリング結果は $\alpha = 0.5$ で $\alpha$-cut した時の結果である。中央にある点線は，$u_1(x) = u_2(x) = 0.5$ となる等高線を表す。また，その両側にある点線はそれぞれ，$u_1(x) = 1, u_2(x) = 1$ となる境界線を表している。2 つの三角はクラスタの中心を表している。

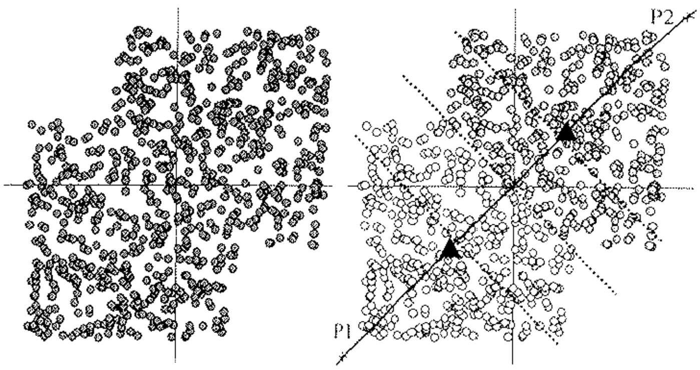


図 4.1: データとそのクラスタリング結果

図 4.2 はクラスタリング結果の上に描かれている P1-P2 を結ぶ直線上のメンバシップ関数である。2 次正則化によって得られたファジィ分類関数が区分的線形となっていることがわかる。

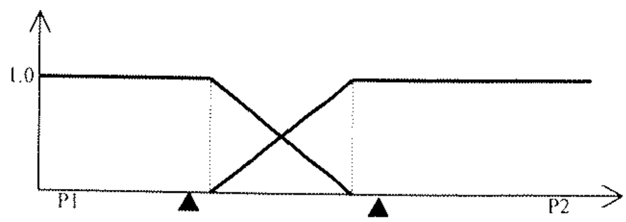


図 4.2: P1-P2 線上の区分的線形メンバシップ関数

## 4.8 おわりに

2次正則化を用いたファジィクラスタリング手法について提案し，そのアルゴリズムについて述べた。さらに区分的線形な分類関数について考察を行った。

このクラスタリング手法は標準的な fuzzy $c$-means 法よりも，より crisp $k$-means 法に近い性質を持つ。$\lambda$ が十分大きい時には最適添字集合が $I(1)$ となり，クリスプなクラスタリングとなり，ある程度 $\lambda$ が小さくなっても，個体 $\boldsymbol{x}_k$ から遠い中心 $\boldsymbol{v}_i$ はメンバシップに影響を与えない。このような性質はこれまでのクラスタリング手法 [10, 39] には見られなかったものである。

このように fuzzy $c$-means 法には標準的な fuzzy $c$-means 法，エントロピー法 [39]，そして今回提案した2次正則化法という3種類が考えられる事がわかった。さまざまな応用に関するこれらの手法の比較については今後の研究としたい。

## 4.9 補足

### 補足 1. 命題 4.2 の証明

次のような新しい変数を導入し，

$$w_{ik}^2 = u_{ik}, \quad i = 1, 2, \cdots, c, \quad k = 1, 2, \cdots, n$$

以下のようなラグランジュ関数を考える。ただし，$\mu_k$ はラグランジュ乗数である。

$$\begin{aligned} L &= J_{quad} - \sum_{k=1}^{n} \mu_k (\sum_{i=1}^{c} w_{ik}^2 - 1) \\ &= \sum_{i=1}^{c} \sum_{k=1}^{n} w_{ik}^2 d_{ik} + \frac{1}{2}\lambda^{-1} \sum_{i=1}^{c} \sum_{k=1}^{n} w_{ik}^4 - \sum_{k=1}^{n} \mu_k (\sum_{i=1}^{c} w_{ik}^2 - 1) \end{aligned}$$

このとき最適性の必要条件より，

$$\frac{\partial L}{\partial w_{ik}} = w_{ik}(d_{ik} - \mu_k + \lambda^{-1} w_{ik}^2) = 0$$

となり，$w_{ik} = 0$ または，$w_{ik}^2 = \lambda(\mu_k - d_{ik})$。よって，

$$u_{ik} = 0 \tag{4.4}$$

または，

$$u_{ik} = \lambda(\mu_k - d_{ik}) \tag{4.5}$$

を得る。これらの式の値は非負となることに注意する。いま，最適解を $\bar{U} = (\bar{u}_{ik})$，$\bar{I}_k = I[\bar{U}_k]$ とすると (4.4), (4.5)，さらに $\sum_{i=1}^{c} u_{ik} = 1$ より，

$$\lambda|\bar{I}_k|\mu_k - \lambda \sum_{\ell \in \bar{I}_k} d_{\ell k} = 1 \tag{4.6}$$

となる。これにより，

$$\bar{u}_{ik} = \begin{cases} (\lambda \sum_{\ell \in \bar{I}_k} d_{\ell k} + 1)/|\bar{I}_k| - \lambda d_{ik}, & i \in \bar{I}_k \\ 0, & i \notin \bar{I}_k \end{cases}$$

を得る。

### 補足 2. 命題 4.4 の証明

証明は，$I(i)$ と異なる許容添字集合が最適添字集合であれば，$I(K)$ が最適添字集合となるような $K$ が存在することにより示す。

まず，$I = \{i_1, i_2, \cdots, i_K\}$，$(i_1 < i_2 < \cdots < i_K)$ を $I(K)$ と異なる最適添字集合とし，$\hat{u}_{ik}$ を添字集合 $I$ から得られたメンバシップであるとする。このとき，次のような別のメンバシップ $y_{ik}$ を考える。

$$y_{1k} = \hat{u}_{i_1 k},\, y_{2k} = \hat{u}_{i_2 k},\ \ \cdots,\ y_{Kk} = \hat{u}_{i_K k}$$

ただし，$i \notin I - I(K)$ なる $i$ に対しては $y_{ik} = 0$ とする。このメンバシップ $y_{ik}$ は明らかに $M_{fcm}$ に入るため，実行可能である。$I$ に対する仮定より，$I(K) = \{i \mid y_{ik} > 0\}$ であり，

$$\begin{aligned} J_{(k)}(\hat{U}_k(I)) &= \sum_{i \in I} \left[ \hat{u}_{ik} d_{ik} + \frac{1}{2}\lambda^{-1}\hat{u}_{ik}^2 \right] \\ &\geq \sum_{i=1}^{K} \left[ y_{ik} d_{ik} + \frac{1}{2}\lambda^{-1} y_{ik}^2 \right] \\ &= J_{(k)}(\{y_{ik}\}) \end{aligned}$$

となる。このように，許容添字集合であるかどうかはわからないが，$I(K) = \{i \mid y_{ik} > 0\}$ の上だけで正となる実行可能解 $y_{ik}$ を得ることができる。もし $I(K)$ が許容添字集合でなければ，$y_{ik}$ は最適性の条件を満たさないので，より最適な解が存在することになり，$I$ が最適添字集合であるという仮定に反する。もし $I(K)$ が許容添字集合であれば，$I(K)$ は最適添字集合である。

したがって，$I(K)$ が最適添字集合となるような $K$ $(\leq c)$ が存在する。

## 補足 3. 補題 4.6 の証明

(a1) は，

$$f_\ell^{L+1} - f_\ell^L = \lambda(d_{L+1,k} - d_{\ell k}) \geq 0$$

より得られる。また，

$$f_L^L - f_{L+1}^{L+1} = \lambda L(d_{L+1,k} - d_{Lk}) \geq 0$$

より，(a2) は明らか。

## 補足 4. 補題 4.7 の証明

**(b1)**: $I(1)$ が許容添字集合であることは明らか。

**(b2)**: $I(L-1)$ の解を $\hat{u}_{ik}$ とし，

$$z_{ik} = L^{-1}(\lambda \sum_{\ell=1}^{L} d_{\ell k} + 1) - \lambda d_{ik} \tag{4.7}$$

とする。もし $z_{ik} > 0$ $(1 \le i \le L)$ ならば，$I(L)$ は許容添字集合であり，$z_{ik} < 0$ となる $i$ が存在すれば，$I(L)$ は許容添字集合ではない。いま $f_L^L > 0$ であるとしよう。$z_{Lk} = f_L^L/L$ より $z_{Lk} > 0$ である。また補題4.6より，$f_i^L \ge f_i^{L-1} (1 \le i \le L-1)$ となるから，

$$z_{ik} \ge \frac{L-1}{L}\hat{u}_{ik} > 0$$

である。したがって，$I(L)$ は許容添字集合である。

**(b3):** 次に，$f_L^L \le 0$ であるとしよう。$z_{Lk} \le 0$ となるので，$I(L)$ は許容添字集合ではない。さらに，命題4.4の2番目の結果より，

$$z_{L+1,k} = f_{L+1}^{L+1}/(L+1) \le f_L^L/(L+1) \le 0$$

よって，$I(L+1), I(L+2), \cdots, I(c)$ は許容添字集合ではない。

## 補足 5. 命題4.8の証明

命題4.4により，最適解を与える添字集合の候補は，$I(1), I(2), \cdots, I(c)$ の中のどれかに限られる。

証明は $I(L)$ の $L$ を構成的に増やすことによる。まず最初に $L=1$，$I = \{1\}$，つまり最も近くにあるクラスタへのクリスプ解であるとする。次に $L$ を $f_L^L \le 0$ すなわち $I(L)$ が許容添字集合でなくなるまで増やして行く。

$I(L-1)$ と $I(L)$ が許容添字集合であり $z_{ik}$ が (4.7) 式で定義されているとする。このとき，$z_{Lk} > 0$ である。

添字集合が $I(L-1)$ のときの目的関数 $J_{(k)}$ の値を $J_{(k)}^1$ とすると，(4.5) および (4.6) を用いて，以下のようになる。ただし，$\sum^{L}$, $\sum^{L-1}$ は，$\sum_{i \in I(L)}$，$\sum_{i \in I(L-1)}$ をそれぞれあらわす。

$$\begin{aligned} J_{(k)}^1 &= \sum^{L-1}\left[u_{ik}(d_{ik} + \lambda^{-1}u_{ik}) - \lambda^{-1}u_{ik}^2/2\right] \\ &= \sum^{L-1}\left[u_{ik}\mu_k - \lambda^{-1}u_{ik}^2/2\right] \\ &= \mu_k - \frac{1}{2\lambda}\sum^{L-1}u_{ik}^2 \end{aligned}$$

$$
\begin{aligned}
&= \mu_k - \frac{\lambda}{2}\sum^{L-1}(\mu_k - d_{ik})^2 \\
&= \mu_k - \frac{\lambda(L-1)}{2}\mu_k^2 + \lambda\mu_k \sum^{L-1} d_{ik} - \frac{\lambda}{2}\sum^{L-1} d_{ik}^2 \\
&= -\frac{\lambda(L-1)}{2}\mu_k^2 + \mu_k(1 + \lambda\sum^{L-1} d_{ik}) - \frac{\lambda}{2}\sum^{L-1} d_{ik}^2 \\
&= -\frac{\lambda(L-1)}{2}\mu_k^2 + \lambda(L-1)\mu_k^2 - \frac{\lambda}{2}\sum^{L-1} d_{ik}^2 \\
&= \frac{\lambda}{2}\left[(L-1)\mu_k^2 - \sum^{L-1} d_{ik}^2\right] \\
&= \frac{\lambda}{2}\left[\frac{1}{L-1}(\sum^{L-1} d_{ik} + \frac{1}{\lambda})^2 - \sum^{L-1} d_{ik}^2\right]
\end{aligned}
$$

いま，

$$
\sum_{i=1}^{L-1} d_{ik} + \frac{1}{\lambda} - (L-1)d_{Lk} = \frac{L}{\lambda} z_{Lk} \tag{4.8}
$$

に注意して，添字集合が $I(L)$ のときの目的関数 $J_{(k)}$ の値を $J_{(k)}^2$ とし，$\Delta J = J_{(k)}^1 - J_{(k)}^2$ とすると，

$$
\begin{aligned}
&\frac{2}{\lambda}\Delta J \\
&= \frac{2}{\lambda}\left(J_{(k)}^1 - J_{(k)}^2\right) \\
&= \frac{1}{L-1}(\sum^{L-1} d_{ik} + \frac{1}{\lambda})^2 - \frac{1}{L}(\sum^{L} d_{ik} + \frac{1}{\lambda})^2 + d_{Lk}^2 \\
&= \frac{1}{(L-1)L}\left[L(\sum^{L-1} d_{ik} + \frac{1}{\lambda})^2 - (L-1)(\sum^{L} d_{ik} + \frac{1}{\lambda})^2\right] + d_{Lk}^2 \\
&= \frac{1}{(L-1)L}\left[(L-1)\left((\sum^{L-1} d_{ik} + \frac{1}{\lambda})^2 - (\sum^{L} d_{ik} + \frac{1}{\lambda})^2\right) + (\sum^{L-1} d_{ik} + \frac{1}{\lambda})^2\right] + d_{Lk}^2 \\
&= \frac{1}{(L-1)L}\left[(L-1)\left(\sum^{L-1} d_{ik} + \sum^{L} d_{ik} + \frac{2}{\lambda}\right)(-d_{Lk}) + (\sum^{L-1} d_{ik} + \frac{1}{\lambda})^2\right] + d_{Lk}^2 \\
&= \frac{1}{(L-1)L}\left[(L-1)(\sum^{L} d_{ik} + \frac{1}{\lambda})(-d_{Lk}) + (L-1)(\sum^{L-1} d_{ik} + \frac{1}{\lambda})(-d_{Lk}) + (\sum^{L-1} d_{ik} + \frac{1}{\lambda})^2\right] + d_{Lk}^2 \\
&= \frac{1}{(L-1)L}\left[(L-1)(\sum^{L} d_{ik} + \frac{1}{\lambda})(-d_{Lk}) + (\sum^{L-1} d_{ik} + \frac{1}{\lambda})(\sum^{L-1} d_{ik} + \frac{1}{\lambda} - (L-1)d_{Lk})\right] + d_{Lk}^2
\end{aligned}
$$

後ろの項に (4.8) を使い,

$$
\begin{aligned}
&= \frac{1}{(L-1)L}\left[(L-1)(\sum^{L} d_{ik} + \frac{1}{\lambda})(-d_{Lk}) + (\sum^{L-1} d_{ik} + \frac{1}{\lambda})\frac{L \cdot z_{Lk}}{\lambda}\right] + d_{Lk}^2 \\
&= \frac{1}{(L-1)\lambda} z_{Lk}(\sum^{L-1} d_{ik} + \frac{1}{\lambda}) + \frac{1}{L}(\sum^{L} d_{ik} + \frac{1}{\lambda})(-d_{Lk}) + d_{Lk}^2 \\
&= \frac{1}{(L-1)\lambda} z_{Lk}(\sum^{L-1} d_{ik} + \frac{1}{\lambda}) - \frac{d_{Lk}}{L}(\sum^{L} d_{ik} + \frac{1}{\lambda} - L d_{Lk}) \\
&= \frac{1}{(L-1)\lambda} z_{Lk}(\sum^{L-1} d_{ik} + \frac{1}{\lambda}) - \frac{d_{Lk}}{L}\left(\sum^{L-1} d_{ik} + \frac{1}{\lambda} - (L-1) d_{Lk}\right)
\end{aligned}
$$

さらに，(4.8) を使い,

$$
\begin{aligned}
&= \frac{1}{(L-1)\lambda} z_{Lk}\left(\sum^{L-1} d_{ik} + \frac{1}{\lambda}\right) - \frac{d_{Lk}}{L}\left(\frac{L \cdot z_{Lk}}{\lambda}\right) \\
&= \frac{1}{(L-1)\lambda} z_{Lk}\left(\sum^{L-1} d_{ik} + \frac{1}{\lambda} - (L-1) d_{Lk}\right)
\end{aligned}
$$

さらに，(4.8) を使い,

$$
\begin{aligned}
&= \frac{1}{(L-1)\lambda} z_{Lk}\left(\frac{L}{\lambda} z_{Lk}\right) \\
&= \frac{L}{(L-1)\lambda^2} z_{Lk}^2
\end{aligned}
$$

これにより，$I(L)$ が許容添字集合ならば，$I(L-1)$ よりも $I(L)$ の方が良い解を与えることがわかる。以上により，このアルゴリズムにより $J_{(k)}$ を最小化する解が得られることがわかる。