

ダイナミックプログラミングを用いた ファジィメトリッククラスタリング Fuzzy Metric Clustering Based on Dynamic Programming

城西大学・数学科 岩村 覚三 (Kakuzo IWAMURA)

Department of Mathematics, Josai University,
kiwamura@josai.ac.jp

神奈川大学・工学部 堀口 正之 (Masayuki HORIGUCHI)

Faculty of Engineering, Kanagawa University,
horiguchi@kanagawa-u.ac.jp

帝京大学・経済学部 堀池 真琴 (Makoto HORIIKE)

Faculty of Economics, Teikyo University,
hori-ike@main.teikyou-u.ac.jp

要旨

多次元線形単調ファジィ数に基づくファジィメトリックとダイナミックプログラミング (DP) を用いたファジィメトリッククラスタリングプログラム (FMCP) を C 言語で開発した. その特徴は, 与えられた多次元線形単調ファジィ数データを直接用いてファジィメトリッククラスタリングアルゴリズムを実行できることである. IBM NetVista メモリ 1.50GB クロックタイム Pentium4 2.00GHz を用いてファジィ感覚による食品嗜好調査の結果をもとに 20 品目の食品を 3 つにクラスタリングする計算実行結果も報告する.

§1 はじめに

N 個の元からなる集合 $\mathcal{N}$ と $i, j \in \mathcal{N} (i \neq j)$ である全ての点対の距離 d_{ij} が与えられた時, この集合 $\mathcal{N}$ を何らかの意味で最適に分割したいことがある. 例えば $\mathcal{N}$ を M 個の互いにオーバーラップしない部分集合に, 各々の部分集合内平均距離和の M 個合計した値を最小になるように分割したいことがある. このような最適分割問題は**クラスタリング問題**と呼ばれている. ここでオーバーラップしない各部分集合を**クラスタ**とも呼ぶ.

ここでは, 次のクラスタリング問題 $\mathbb{P}$ を考察する. すなわち, 集合 $\mathcal{N}$ の M 個への分割を $g_k (g_k \neq \emptyset, 1 \leq k \leq M)$ とするとき

$$\bigcup_{k=1}^M g_k = \mathcal{N} \quad (1)$$

$$g_k \cap g_{k'} = \phi (1 \leq k \neq k' \leq M) \quad (2)$$

の下で

$$D(g_1, \dots, g_M) = \sum_{k=1}^M \frac{1}{|g_k|} \sum_{i < j \in g_k} d_{ij} \quad (3)$$

を最小にする $g_k (1 \leq k \leq M)$ を求める問題を扱う. d_{ij} が実数値である時のクラスタリング問題は Arthanari and Dodge(1981) [1] に詳しい. 2 要素間の距離が 1 次元単調ファジィ数 (cf. [3, 10]) で与えられている場合のクラスタリング問題は Kamimura and Kurano(2000,2001)[8, 9] で研究された. しかしながら, そこでは最適なクラスター分割を求めるアルゴリズムや具体的計算方法 (コンピュータプログラム) は示されていない. 我々は, 多次元線形単調ファジィ数に基づくファジィメトリックとダイナミックプログラミング (DP) を用いたファジィメトリッククラスタリングプログラム (FMCP) を C 言語で開発した. その特徴は, 与えられた多次元線形単調ファジィ数データを直接用いてファジィメトリッククラスタリングアルゴリズムを実行できることである. IBM NetVista メモリ 1.50GB クロックタイム Pentium4 2.00GHz を用いてファジィ感覚による食品嗜好調査の結果をもとに 20 品目の食品を 3 つにクラスタリングする計算実行結果も報告する.

ファジィ理論のクラスタリング問題への適用事例や応用については文献 [11, 12, 13] を参考にされたい.

§2 多次元線形単調ファジィ数

一般にクラスタリング問題では, 各要素 $i \in \mathcal{N}$ を特徴付ける多次元データ $a_i = (a_{i1}, a_{i2}, \dots, a_{il})$ ($i \in \mathcal{N}$) が与えられ, これをもとに 2 要素間の距離 d_{ij} が定義される. この距離 d_{ij} を用いて (1)-(3) によって記述される最適クラスター問題 $\mathcal{P}$ が考えられる (cf. [1]). ここでは, 各要素 $i \in \mathcal{N}$ に多次元線形単調ファジィによって表現されたファジィデータが与えられている場合を取り扱う. 最初に線形ファジィ数について解説し, 次にファジィ距離によるクラスター問題とその解析方法を述べる.

§2.1 1 次元線形単調ファジィ数

1 次元単調ファジィ数 $\tilde{a}$ を定義する. R_+ を非負実数とする. R_+ から区間 $[0, 1]$ への写像 $\tilde{a}$ は, 次の 3 つの性質を満たす時 1 次元単調ファジィ数と呼ぶ:

- (i) $R_+ \ni t_1, t_2$ に対し, $t_1 \leq t_2$ ならば $\tilde{a}(t_1) \leq \tilde{a}(t_2)$,
- (ii) $\tilde{a}(t)$ は右連続であり $y = \tilde{a}(t)$ の不連続点は高々有限個である,
- (iii) $\tilde{a}(t) = 1$ となる $t \in R_+$ が存在する.

1 次元単調ファジィ数の全体を $\mathcal{F}_+$ で表わす.

$\alpha \in [0, 1]$ に対して $\tilde{a} \in \mathcal{F}_+$ の α -カット $\tilde{a}_\alpha$ は次で定義される (cf. [2]):

$$\begin{aligned} \alpha > 0 \text{ のとき } \tilde{a}_\alpha &= \{x \in R_+ | \tilde{a}(x) \geq \alpha\} = [\tilde{a}^{-1}(\alpha), \infty), \\ \alpha = 0 \text{ のとき } \tilde{a}_0 &= cl\{x \in R_+ | \tilde{a}(x) > 0\} = [\tilde{a}^{-1}(0), \infty). \end{aligned}$$

ただし, clA は集合 $A \subset R_+$ の閉包を表し $\tilde{a}^{-1}(\alpha) = \min\{x \in R_+ | \tilde{a}(x) \geq \alpha\}$, $\tilde{a}^{-1}(0) = \inf\{x \in R_+ | \tilde{a}(x) > 0\}$ とする.

次に, $\mathcal{F}_+$ に加法 (+) とスカラー倍を次のように定義する (cf. [5]). 任意の $\tilde{a}, \tilde{b} \in \mathcal{F}_+$ に対して

$$(\tilde{a} + \tilde{b})(x) = \sup_{\substack{x_1 + x_2 = x \\ x_1, x_2 \geq 0}} \min\{\tilde{a}(x_1), \tilde{b}(x_2)\} \quad (x \in R_+),$$

$$\lambda \tilde{a}(x) = \begin{cases} \tilde{a}(x/\lambda) & (\lambda > 0 \text{ のとき}), \\ 1_{R_+}(x) & (\lambda = 0 \text{ のとき}). \end{cases}$$

ただし, 1_A は集合 A の指示関数で $1_A(x) = 1 \ (x \in A), = 0 \ (x \notin A)$ によって定義される. このとき, $\tilde{a} + \tilde{b} \in \mathcal{F}_+, \lambda \tilde{a} \in \mathcal{F}_+$ かつ α -カットに関して $(\tilde{a} + \tilde{b})_\alpha = [\tilde{a}^{-1}(\alpha) + \tilde{b}^{-1}(\alpha), \infty), (\lambda \tilde{a})_\alpha = [\lambda \tilde{a}^{-1}(\alpha), \infty) \ (\lambda \geq 0)$ が成り立つ.

図 1 と図 2 に単調ファジィ数の例を挙げる.

図 1: (連続でない 1 次元単調ファジィ数の例.)

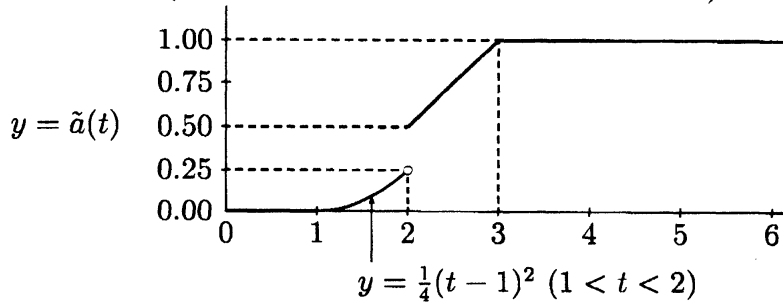
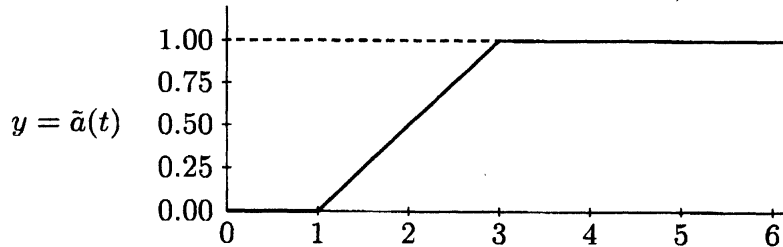


図 2: (線形な 1 次元単調ファジィ数の例.)



一般に

$$\tilde{a}(t) = \begin{cases} 0 & (0 \leq t < \underline{a}) \\ (t - \underline{a})/(\bar{a} - \underline{a}) & (\underline{a} \leq t \leq \bar{a}) \\ 1 & (\bar{a} < t) \end{cases} \quad (4)$$

で決められる $\tilde{a}(t)$ は, 1 次元単調ファジィ数であり, $\underline{a}$ と $\bar{a}$ の 2 つの値によって一意に決まる. 2 つの数 $\underline{a}, \bar{a}$ から決まる 1 次元単調ファジィ数 $\tilde{a}$ を $[\underline{a}, \bar{a}]$ で表わす. 図 2 の 1 次元単調ファジィ数は $\underline{a} = 1, \bar{a} = 3, \tilde{a} = [1, 3]$ と書ける. 式 (4) で決まる 1 次元単調ファジィ数は $\underline{a} \leq t \leq \bar{a}$ において $\tilde{a}(t)$ の値が t に関し線形であるので, これを **1 次元線形単調ファジィ数** と呼ぶ. 1 次元線形単調ファジィ数 $\tilde{a}$ に対しその α -カット $\tilde{a}_\alpha = [\tilde{a}^{-1}(\alpha), \infty)$ について明ら

かに次が成り立つ:

$$\tilde{a}_\alpha = \begin{cases} 0 & (\alpha = 0) \\ (\bar{a} - \underline{a})\alpha + \underline{a} & (0 < \alpha \leq 1) \end{cases} \quad (5)$$

従って2つの1次元線形単調ファジィ数 $\tilde{a} = [\underline{a}, \bar{a}]$, $\tilde{b} = [\underline{b}, \bar{b}]$ ($0 \leq \underline{a} \leq \bar{a} < \infty, 0 \leq \underline{b} \leq \bar{b} < \infty$) に対し

$$|\tilde{a}^{-1}(\alpha) - \tilde{b}^{-1}(\alpha)| = \begin{cases} 0 & (\alpha = 0) \\ |\{(\bar{b} - \underline{a}) + (\bar{a} - \bar{b})\}\alpha + \underline{a} - \underline{b}| & (0 < \alpha \leq 1) \end{cases} \quad (6)$$

であることがわかる.

§2.2 多次元線形単調ファジィ数

l 項目からなる l 次元線形単調ファジィ数 $\tilde{a} \in \mathcal{F}_+^l$ は $\tilde{a} = (\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_l)$, $\tilde{a}_k = [\underline{a}_k, \bar{a}_k]$ ($1 \leq k \leq l$) と書ける. さて, 2つの l 次元線形単調ファジィ数 $\tilde{a} = (\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_l)$, $\tilde{a}_k = [\underline{a}_k, \bar{a}_k]$ ($1 \leq k \leq l$) と $\tilde{b} = (\tilde{b}_1, \tilde{b}_2, \dots, \tilde{b}_l)$, $\tilde{b}_k = [\underline{b}_k, \bar{b}_k]$ ($1 \leq k \leq l$) に対しノルム $\|\cdot\|$ は次の (7)-(9) によって与えられる $\|\tilde{a} - \tilde{b}\|$ の α -カット ($\alpha \in [0, 1]$) を用いて (10) により与えられる.

$$\|\tilde{a} - \tilde{b}\|_\alpha = [Q(\alpha), \infty) \quad (7)$$

$$Q(\alpha) := \sup_{0 \leq \alpha' \leq \alpha} q(\alpha') \quad (8)$$

$$\begin{aligned} q(\alpha') &:= \sum_{k=1}^l (\tilde{a}_k^{-1}(\alpha') - \tilde{b}_k^{-1}(\alpha'))^2 \\ &= \sum_{k=1}^l \{(\bar{a}_k - \underline{a}_k - \bar{b}_k + \underline{b}_k)\alpha' + \underline{a}_k - \underline{b}_k\}^2 \end{aligned} \quad (9)$$

$$\|\tilde{a} - \tilde{b}\|(x) = \sup_{\alpha \in [0, 1]} \{\alpha, 1_{\|\tilde{a} - \tilde{b}\|_\alpha}(x)\}. \quad (10)$$

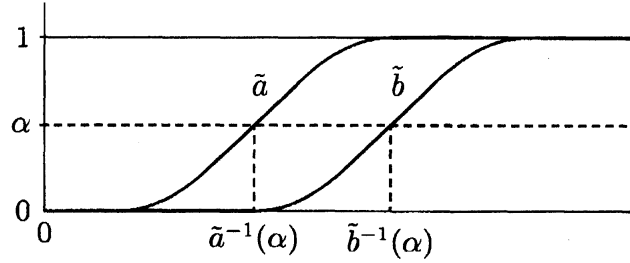
このとき, $\|\tilde{a} - \tilde{b}\| \in \mathcal{F}_+$ かつ $\|\cdot\|$ は $\mathcal{F}^l$ 上のファジィノルム (cf. [10]) となることが示される (cf. [5, 8]).

ファジィ距離によるクラスタリング問題を定義するために, ファジィマックスオーダーといわれる $\mathcal{F}_+$ 上の半順序を考える. 任意の $\tilde{a}, \tilde{b} \in \mathcal{F}_+$ に対して $\tilde{a}^{-1}(\alpha) \leq \tilde{b}^{-1}(\alpha)$ ($\alpha \in [0, 1]$) が成り立つとき $\tilde{b}$ は $\tilde{a}$ を優越すると呼び, $\tilde{a} \preceq \tilde{b}$ と記す (図3参照).

今, 各 $i \in \mathcal{N}$ に対して, l 次元線形単調ファジィ数によって表されたファジィデータ

$$\tilde{a}_i = (\tilde{a}_{i1}, \tilde{a}_{i2}, \dots, \tilde{a}_{il}) \in \mathcal{F}_+^l, \quad \tilde{a}_{ik} = [\underline{a}_{ik}, \bar{a}_{ik}] \quad (1 \leq k \leq l, 1 \leq i \leq N)$$

が与えられたとする. これを用いて, 要素 $i, j \in \mathcal{N}$ の間のファジィ距離 $\tilde{d}_{ij}$ を $\tilde{d}_{ij} = \|\tilde{a}_i - \tilde{a}_j\|$

図 3: $\tilde{a} \preccurlyeq \tilde{b}$ の例示

で定義する. $\tilde{d}_{ij}$ の α -カットを $\tilde{d}_{ij,\alpha} = [d_{ij}(\alpha), \infty)$ とおくと (8)-(10) により

$$d_{ij}(\alpha) = \sqrt{Q(\alpha)} \quad (11)$$

$$Q(\alpha) = \sup_{0 \leq \alpha' \leq \alpha} q(\alpha') \quad (12)$$

$$q(\alpha') = \sum_{k=1}^l (p_k \alpha' + \underline{a}_{ik} - \underline{a}_{jk})^2 \quad (13)$$

と書ける. ここで,

$$p_k := (\underline{a}_{jk} - \underline{a}_{ik}) + (\bar{a}_{ik} - \bar{a}_{jk}) \quad (1 \leq k \leq l)$$

である. (12) は

$$g(\alpha') = \left(\sum_{k=1}^l p_k^2 \right) \alpha'^2 + 2 \sum_{k=1}^l p_k (\underline{a}_{ik} - \underline{a}_{jk}) \alpha' + \sum_{k=1}^l (\underline{a}_{ik} - \underline{a}_{jk})^2$$

と書けるので

$$\bar{\alpha} = \frac{-\sum_{k=1}^l p_k (\underline{a}_{ik} - \underline{a}_{jk})}{\sum_{k=1}^l p_k^2} \quad (14)$$

において

$$\begin{cases} \bar{\alpha} \leq 0 \text{ なら} & Q(\alpha) = q(\alpha) \quad (0 \leq \alpha \leq 1) \\ 0 < \bar{\alpha} < \frac{1}{2} \text{ なら} & \begin{cases} 0 \leq \alpha \leq 2\bar{\alpha} \text{ で } Q(\alpha) = q(0) \\ 2\bar{\alpha} \leq \alpha \leq 1 \text{ で } Q(\alpha) = q(\alpha) \end{cases} \\ \frac{1}{2} \leq \bar{\alpha} \text{ なら} & Q(\alpha) = q(0) \quad (0 \leq \alpha \leq 1) \end{cases} \quad (15)$$

より $Q(\alpha)$ が求められる. ここで, ファジィ距離 $\tilde{d}_{ij}$ ($i, j \in \mathcal{N}$) によるクラスタリング問題 $\tilde{\mathcal{P}}$ を定義することができる. すなわち, 与えられたクラスターの数 M ($1 \leq M \leq N$) に対して §1 の (1), (2) を満たすクラスター分割 $\{g_1, g_2, \dots, g_M\}$ の中で対応する単調ファジィ数

$$\tilde{D}(g_1, g_2, \dots, g_M) = \sum_{k=1}^M \frac{1}{|g_k|} \sum_{\substack{i < j \\ i, j \in g_k}} \tilde{d}_{ij} \quad (16)$$

を半順序 $\preceq$ において最小 (パレートの意味で最適) にするクラスター分割を求めることである。ところが, この §2 での α -カットの議論から $\tilde{D}(g_1, g_2, \dots, g_M)$ の α -カットは次で与えられる。

$$\tilde{D}(g_1, g_2, \dots, g_M)_\alpha = [D(g_1, g_2, \dots, g_M)_\alpha, \infty), \quad (17)$$

ただし,

$$D(g_1, g_2, \dots, g_M)_\alpha = \sum_{k=1}^M \frac{1}{|g_k|} \sum_{\substack{i < j \\ i, j \in g_k}} d_{ij}(\alpha). \quad (18)$$

従って, 半順序 $\preceq$ の定義よりファジィ距離によるクラスタリング問題 $\tilde{\mathbb{P}}$ は各 $\alpha \in [0, 1]$ に対して (18) 式による $D(g_1, g_2, \dots, g_M)_\alpha$ を最小にする通常のクラスタリング問題 $\mathbb{P}(\alpha)$ を解くことに帰着される。

§2.3 α -カットによる2つの多次元線形単調ファジィ数の距離の計算の詳細

与えられた N 個の l 次元線形単調ファジィ数 $\tilde{a}_i = (\tilde{a}_{i1}, \tilde{a}_{i2}, \dots, \tilde{a}_{il}) \in \mathcal{F}_+^l$, $\tilde{a}_{ik} = [\underline{a}_{ik}, \bar{a}_{ik}]$ ($1 \leq k \leq l, 1 \leq i \leq N$) が与えられたとする。カット値 α ($0 \leq \alpha \leq 1$) が与えられた時, $\tilde{a}_i$ と $\tilde{a}_j$ の距離 $d_{ij}(\alpha)$ は §2.2 より

$$d_{ij}(\alpha) = \sqrt{Q(\alpha)} \quad (19)$$

$$Q(\alpha) = \sup_{0 \leq \alpha' \leq \alpha} q(\alpha') \quad (20)$$

$$q(\alpha') = \sum_{k=1}^l [p_k \alpha' + \underline{a}_{ik} - \underline{a}_{jk}]^2 \quad (21)$$

と書ける。ここで

$$p_k := (\underline{a}_{jk} - \underline{a}_{ik}) + (\bar{a}_{ik} - \bar{a}_{jk}) \quad (1 \leq k \leq l)$$

である。(12) は

$$g(\alpha') = \left(\sum_{k=1}^l p_k^2 \right) \alpha'^2 + 2 \sum_{k=1}^l p_k (\underline{a}_{ik} - \underline{a}_{jk}) \alpha' + \sum_{k=1}^l (\underline{a}_{ik} - \underline{a}_{jk})^2$$

と書けるので

$$\bar{\alpha} = \frac{-\sum_{k=1}^l p_k (\underline{a}_{ik} - \underline{a}_{jk})}{\sum_{k=1}^l p_k^2} \quad (22)$$

とにおいて

$$\begin{cases} \bar{\alpha} \leq 0 \text{ なら} & Q(\alpha) = q(\alpha) \quad (0 \leq \alpha \leq 1) \\ 0 < \bar{\alpha} < \frac{1}{2} \text{ なら} & \begin{cases} 0 \leq \alpha \leq 2\bar{\alpha} \text{ で } Q(\alpha) = q(0) \\ 2\bar{\alpha} \leq \alpha \leq 1 \text{ で } Q(\alpha) = q(\alpha) \end{cases} \\ \frac{1}{2} \leq \bar{\alpha} \text{ なら} & Q(\alpha) = q(0) \quad (0 \leq \alpha \leq 1) \end{cases} \quad (23)$$

と計算させればよいことが解る。与えられたカット値 α に対し2つの l 次元線形単調ファジィ数の距離が得られたので Jensen(1969) [6] の DP アルゴリズムを採用してファジィメトリッククラスタリング問題を解くことが出来る。

§3 ダイナミックプログラミングアルゴリズムとファジィダイナミッククラスタリングアルゴリズム

§2において、ファジィ距離によるクラスタリング問題 $\tilde{P}$ を解くことは、各レベル $\alpha \in [0, 1]$ に対する通常のクラスタリング問題 $P(\alpha)$ を解けばよいことが示された。ここでは、 $P(\alpha)$ ($\alpha \in [0, 1]$) を解くためのダイナミックプログラミングによるアルゴリズムと開発プログラムについて述べる。

Jensen(1969) [6] は任意の $Z \subseteq N$ に対し $\frac{1}{|Z|} \sum_{i < j \in Z} d_{ij}^2$ を $T(Z)$ と書いて次のようなダイナミックプログラミングアルゴリズムを作成した。

$$W_K^*(Z) = \begin{cases} 0 & (K = 0 \text{ のとき}) \\ \min_y [T(Z \cap y^c) + W_{K-1}^*(y)] & (1 \leq K \leq M \text{ のとき}) \end{cases} \quad (24)$$

なお、ここで

K : ステージ変数添字
 Z : ステージ K での状態変数
 y : ステージ $K - 1$ での状態変数

である。

この Jensen のアルゴリズムを下部アルゴリズムとして用いて我々は下の様なファジィダイナミッククラスタリングアルゴリズムを作成し、C 言語による実行可能プログラムを作成した。

```
for (1 ≤ l ≤ 10)
  { // カット値  $\alpha = \frac{l}{10.0}$  に対し
    for (1 ≤ i < j ≤ N)
      {
         $d_{ij}(\alpha) \leftarrow (15)$  と (11) から求めた値 //  $d_{ij}(\alpha) = d_{ji}(\alpha)$  に注意。
      }
    この  $d_{ij}(\alpha)$  に Jensen のダイナミックプログラミングアルゴリズムを適応して
    最適クラスタリングを求めファイルに出力する。
  }
```

実際にこのアルゴリズムを C 言語に直すプログラム化には細心の注意が必要であった。我々は $2 \leq N \leq 32, 2 \leq M < N$ に対し正常計算できるファジィダイナミッククラスタリング C プログラムを作成した。32 に限定したのは C 言語の unsigned integer の 4 バイト (32 ビット) で部分集合 $Z (Z \subseteq N)$ を表したためである。また同一の α 値に対し

$$T(Z) = \frac{1}{|Z|} \sum_{i < j \in Z} d_{ij}^2(\alpha) \quad (Z \subseteq N)$$

の計算が何度も必要になるので計算高速化のために $T(Z)$ の値をメモリ内に配列で保持した。

図 4: 食品嗜好調査(アンケート)票

食品の嗜好調査

ファジィクラスタリングの研究のために次のアンケート調査にご協力下さい。

嗜好尺度は次のとおりです。

1	2	3	4	5	6	7	8	9
も	そ	嫌	す	好	す	好	そ	も
っ	う	い	こ	き	こ	き	う	っ
と	と		し	で	し		と	と
も	も		い	も	好		も	も
嫌	嫌		い	嫌	き		好	好
				い				
				な				

好き嫌いを単刀直入に表すために、その程度がだいたいこの範囲にあると思われるところを下記の記入例のように、幅を持って で記入してください。

記入例

(1) ご飯

1	2	3	4	5	6	7	8	9

好きであるが、その程度はこの範囲にあると思われる場合

(2) お茶漬け

1	2	3	4	5	6	7	8	9

好きでも嫌いでもないが、どちらかといえば少し好きと思われる場合

(3) みそ汁

1	2	3	4	5	6	7	8	9

嫌いであるが、その程度はこの範囲にあると思われる場合

§4 ファジィ感覚による食品嗜好調査

好き嫌いなどの感覚的な判断を数量化する場合、“およそ5くらい”、“だいたい8以上”などのようにファジィ感覚的な記述が許される枠組みを設定することが必要である。我々は、ファジィ感覚による食品嗜好調査(アンケート調査)を行い、§2-§3で議論したファジィ距離によるクラスタリング問題の解析法のもとで分析を行った。我々は、図4のようなアンケート用紙を用意した。嗜好尺度を“もっとも嫌い”から“もっとも好き”までの9段階として、図4の記入例のように好き嫌いの範囲を幅を持って囲んでもらった。数量化においては、例えば図4の記入例(1)ご飯については線形単調ファジィ数 $[6.6, 8.4]$ と対応付けた。この数量化の規則を当てはめると、図4の記入例(2)お茶漬けの好き嫌いの程度は、 $[4.2, 7.6]$ となる。

具体的調査として、20品目(図5の右欄の食品)に対してある大学の新生47名に対してアンケート調査を行い上記の規則に従って数量化し、各食品の特徴を示す47次元線形単調ファジィ数が得られる。それらを表にまとめ一部を抜粋したものが表1である。詳細は、堀

口の web ページ [4] にある. 表 1 のデータをインプットとして, 我々が開発したソフトウェア (PMCP) により $N = 20$ 品目を $M = 3$ のクラスタに分割する問題を解き, その結果をまとめたものが表 2 及び図 5 で与えられている.

レベル α は個人の食品に対する好き嫌いの判断を行うときの比重の置き方を表すとすれば, α が 0 から 1 に変化するとき一般に食品を嫌いと判断したがる状態 (飽食あるいは食欲減退) から好きと判断したがる傾向の状態 (空腹あるいは食欲旺盛) に移行することを表していると解釈できる. 表 2 を見ると, α の値には関係しない嗜好の似通った食品の 3 つのグループとして次が挙げられる.

A グループ (日常的に楽しみをもつ食品群): ご飯, カレーライス, 焼きそばなど

B グループ (庶民的な食品群): 納豆, 焼き魚, みそ汁

C グループ (補助的な食品群): 梅干し, たくわん, ひやむぎ

次に, 個人の嗜好の傾向に左右されて所属するグループが変わる食品群をあげることができる. 例えば, トーストは飽食気味のときは A グループに属するが空腹になるにしたがってグループを A から C へ, C から B へと変える食品であることがわかる.

表 1: ある大学の新入生 47 名のアンケート調査結果による線形単調ファジィ数のデータ (一部抜粋)

	ご飯	カレー	...	ハンバーグ	焼き魚	さしみ	...	冷や奴	納豆
1	[8.6,9]	[5.7,8]	...	[6.6,8.4]	[6.7,8.2]	[6.7,8.4]	...	[6.3,9]	[7.7,9]
2	[6.4,8.4]	[5.7,8.3]	...	[5.8,8.4]	[4.8,7.2]	[6.3,8.4]	...	[5.5,8.4]	[5.6,8.5]
3	[2.5,3.3]	[6.5,7.4]	...	[6.4,7.3]	[6.3,7.1]	[7.4,8.3]	...	[8.4,9]	[4.5,5.3]
4	[4.8,5.3]	[7.7,8.3]	...	[7.7,9]	[6.7,7.2]	[7.7,8.2]	...	[4.8,6.3]	[5.8,6.2]
5	[3.8,4.1]	[2.8,3.2]	...	[6.4,7.1]	[6.7,7.4]	[8.6,9]	...	[4.7,5.2]	[2.8,3.3]
6	[7.1,9]	[5.7,9]	...	[6.7,9]	[6.5,8.8]	[5.6,7]	...	[6.1,7.8]	[6.1,8]
7	[6.7,9]	[8.6,9]	...	[7.9,9]	[7.8,9]	[7.3,9]	...	[7.9,9]	[8.3,9]
8	[6.6,8.5]	[6.8,8.9]	...	[5.7,7.3]	[4.5,5.4]	[5.5,7.4]	...	[5.3,6.6]	[5.6,7.3]
...	...	...	...	...	...	...	...	...	...
43	[4.8,5.2]	[5.8,6.2]	...	[8.7,9]	[5.8,6.2]	[6.8,7.2]	...	[5.8,6.2]	[7.8,8.2]
44	[8.7,9]	[8.7,9]	...	[8.5,9]	[6.7,8.3]	[4.8,6.4]	...	[6.7,8.4]	[8.5,9]
45	[7.4,8.7]	[7.5,9]	...	[5.7,8.3]	[6.8,8.5]	[5.6,8.3]	...	[3.5,5.2]	[3.5,7.3]
46	[4,6]	[6.2,8.5]	...	[6.2,7.9]	[2.2,4]	[6.5,8.3]	...	[6.1,7.8]	[4.2,5.7]
47	[8.5,9]	[7.7,9]	...	[5.7,7.3]	[5.7,7.3]	[4.7,6.4]	...	[4.7,7.2]	[6.7,9]

§5 結語

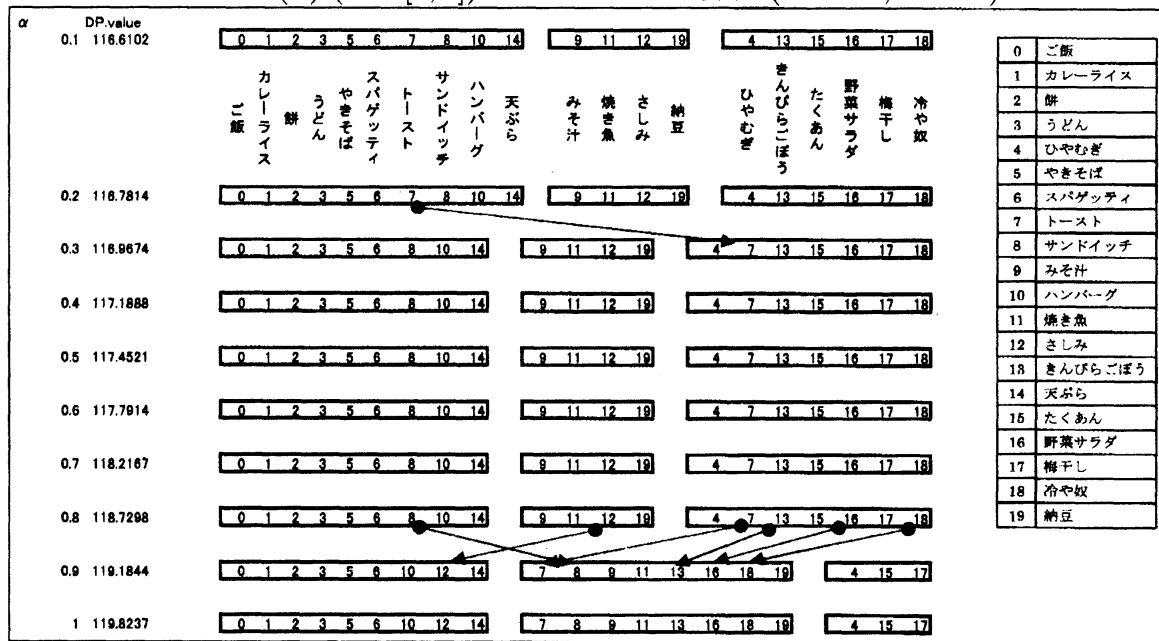
多次元線形単調ファジィ数に基づくファジィメトリック入力データを扱えるファジィメトリッククラスタリングプログラムを C 言語で開発した. Kamimura and Kurano(2000,2001) [8] [9] と異なり与えられた多次元線形単調ファジィ数データを直接用いてファジィメトリッククラスタリングアルゴリズムが実行できる. $N \leq 25$ までの N に対し実用性がある. 我々の C プログラムは線形性を仮定しない多次元単調ファジィ数にも拡張適応可能である. この方向への拡張は将来の課題となる.

謝辞

本研究に対して, 福岡教育大学の上村英樹氏及び千葉大学教育学部の蔵野正美氏より有益な助言を頂いた. ここに感謝の意を表す.

表 2: α の値を 0.1 から 1 の間で変化させたときのファジィダイナミッククラスタリングアルゴリズムの実行例

N = 20 M = 3	
alpha	partition
alpha = 0.100000 min. DP value = 116.610245	
pantition = 18 17 16 15 13 4	19 12 11 9 14 10 8 7 6 5 3 2 1 0
alpha = 0.200000 min. DP value = 116.781425	
pantition = 19 12 11 9	18 17 16 15 13 4 14 10 8 7 6 5 3 2 1 0
alpha = 0.300000 min. DP value = 116.967422	
pantition = 19 12 11 9	14 10 8 6 5 3 2 1 0 18 17 16 15 13 7 4
alpha = 0.400000 min. DP value = 117.188759	
pantition = 19 12 11 9	14 10 8 6 5 3 2 1 0 18 17 16 15 13 7 4
alpha = 0.500000 min. DP value = 117.452110	
pantition = 19 12 11 9	14 10 8 6 5 3 2 1 0 18 17 16 15 13 7 4
alpha = 0.600000 min. DP value = 117.791405	
pantition = 19 12 11 9	14 10 8 6 5 3 2 1 0 18 17 16 15 13 7 4
alpha = 0.700000 min. DP value = 118.216721	
pantition = 19 12 11 9	14 10 8 6 5 3 2 1 0 18 17 16 15 13 7 4
alpha = 0.800000 min. DP value = 118.729797	
pantition = 19 12 11 9	18 17 16 15 13 7 4 14 10 8 6 5 3 2 1 0
alpha = 0.900000 min. DP value = 119.184380	
pantition = 17 15 4	14 12 10 6 5 3 2 1 0 19 18 16 13 11 9 8 7
alpha = 1.000000 min. DP value = 119.823685	
pantition = 17 15 4	14 12 10 6 5 3 2 1 0 19 18 16 13 11 9 8 7

図 5: $P(\alpha)$ ($\alpha \in [0, 1]$) の最適クラスタ分割 ($N = 20, M = 3$)

参考文献

- [1] T.S. Arthanari, Y. Dodge, Mathematical Programming in Statistics, Wiley, New York, 1981.
- [2] D.Dubois, H. Prade, Fuzzy Sets and Systems: Theory and Applications, Academic Press, New York, 1979.
- [3] A. George, P. Veeramani, On some results in fuzzy metric spaces, Fuzzy Sets and Systems 64(1994)395-399.
- [4] M. Horiguchi, Interval Data of Monotone Fuzzy Number, URL <http://www.math.kanagawa-u.ac.jp/~horiguchi/>
- [5] K. Iwamura, M. Kurano, M. Horiguchi, Fuzzy Metric Clustering and Dynamic Programming, January 2004, Hand writing.
- [6] R.E. Jensen, A dynamic programming algorithm for cluster analysis, Oper. Res. 17(1969)1034-1057.
- [7] O. Kabva, S. Seikkala, On fuzzy metric spaces, Fuzzy Sets and Systems 12(1984)215-229.
- [8] H. Kamimura, M. Kurano, Clustering by a Fuzzy Metric: Applications to the Cluster-Median Problem, Bulletin of Informatics and Cybernetics 32(2000)11-21.

- [9] H. Kamimura, M. Kurano, Clustering by a Fuzzy Metric, *Fuzzy Sets and Systems* 120(2001)249–254.
- [10] O. Kramosil, J. Michalek, Fuzzy metric and statistical metric spaces, *Kybernetika* 11(1975)326–334.
- [11] Enrique H. Ruspini, New experimental results in fuzzy clustering, *Information Sciences* 6(1973)273–284.
- [12] M. Sato, Y. Sato, A general fuzzy clustering model based on aggregation operators, *Behaviormetrika* 22(1995)115–128
- [13] J. Watada, H. Tanaka, K. Asai, A heuristic method of hierarchical clustering for fuzzy intransitive relations, In: R. R. Yager(Eds.), *Fuzzy set and possibility theory*, Pergamon Press, New York, 1982(pp. 148–166).