

# ファジィクラスタリングを用いた スイッチング線形判別

Switching Linear Discriminant Using Fuzzy Clustering

日本ファジィ学会誌, 13, 1, 111-118 (2001)

吳 志賢  
大迫 真実  
本多 克宏  
市橋 秀友

大阪府立大学大学院 工学研究科  
電気・情報系専攻 経営工学分野  
〒 599-8531 大阪府堺市学園町 1-1

Chi-Hyon OH  
Masami OSAKO  
Katsuhiro HONDA  
Hidetomo ICHIHASHI

Graduate School of Engineering,  
Electrical Engineering and Information Science,  
Industrial Engineering, Osaka Prefecture University,  
1-1 Gakuen-cho, Sakai, Osaka 599-8531, Japan

## 要 約

Fisher に始まる線形判別（正準判別分析）は，データ集合をある直線上に射影することによって判別分析を行うもので，その線形判別関数は判別のための一つの総合指標として用いることができ，多くの属性を直接用いるルール形式の判別法よりも構造が単純で理解しやすい．一方，異なるクラスのデータが混在して，境界が複雑な非線形性を持つ場合には識別率が低くなる．そのような場合，非線形な領域を細かく分割し，それぞれに対して線形判別を行うことで識別率を向上させることが出来る．本研究では，判別境界の非線形性を考慮するために，境界付近にクラスター中心を配置するファジィクラスタリング法を提案し，クラスターごとに線形判別関数を用いるスイッチング線形判別法について述べる．提案法により，線形判別分析における構造の簡便さ，利便性を残して，判別境界の非線形性を考慮したルール形式でかつ簡便な指標を用いる判別分析を行うことが出来る．

## キーワード

線形判別，ファジィクラスタリング，パターン識別

## Abstract

Linear discriminant analysis by R. A. Fisher is based on the single linear projection of the whole data set, that provides a convenient index for classification. If the data set to be classified has a region where the data of the different classes overlap and has nonlinear decision boundaries, Fisher's linear discriminant analysis may not be able to demonstrate good classification performance. In this paper, we propose a new linear discriminant approach (switching linear discriminant) which derives the linear discriminant function for each cluster obtained by a fuzzy clustering algorithm which takes local nonlinearity into consideration. The fuzzy clustering algorithm locates cluster centers around the nonlinear boundaries for the switching linear discriminant, which thus provides several indices for classification in a form of simplified rules. The effectiveness of our method is shown in the numerical examples.

## Keywords

Linear discriminant, fuzzy clustering, pattern recognition

## 1. はじめに

Fisher に始まる線形判別（正準判別分析）[1] は、データ集合をある直線上に射影することによって判別分析を行うもので、その判別関数は判別のための総指標を与えるものである。たとえば、体重と身長などの属性を考慮して肥満を判定するには、それぞれ値の組み合わせに対して肥満であるかどうかをルール形式、もしくは表形式で表現する必要がある。しかし一般には BMI（Body Mass Index）と呼ばれる身長体重を 1 つに総合した指標の値で判別される。線形判別関数はそのような総合指標の一つを提供するものである。Fisher の線形判別関数は、パラメトリックな分布の仮定を必要とせず共分散行列のみから求まり、その構造は単純で容易に理解することができる。一方、異なるクラスのデータが混在している領域が存在するような局所的な非線形性を伴ったデータ集合に対しては識別率が低下する。そのような場合、非線形性の強い領域を細かく分割し、それぞれの分割部分に対して線形判別を行うことによって識別率の向上を促すことができる。すなわち、図 1 のように異なったクラスが混在する境界付近をクラスタリングし、それぞれのクラスターに対して線形判別関数を導出すること（スイッチング線形判別）によって識別率が上がる。

データ集合をいくつかのグループに分割するクラスター分析は、これまで多くの研究が行われてきた。ファジ理論を用いたクラスター分析としては、Fuzzy *c*-Means（FCM）法 [2] や Fuzzy *c*-Lines 法 [3]、Fuzzy *c*-Varieties（FCV）法 [4] などが提案されている。特に FCV 法はプロトタイプが線形多様体であり、これらの線形空間を張るベクトルを主成分ベクトルと考えると、クラスター分析と主成分分析の同時分析法とみなすことができる。他にもクラスター分析と多変量解析手法を組み合わせたものとして、局所的な線形構造を発見するスイッチング回帰 [5] への応用である Fuzzy *c*-Regression Models [6] などがあるが、それらはクラスター分析やモデル同定法の一手法として提案されており、多変量解析的な意味合いで用いられることはあまり多くない。多変量解析を主たる目的とした同時解析手法としては、数量化分析 3 類（対応分析）や正準相関分析と FCM 法を組み合わせた手法 [7] [8] が提案

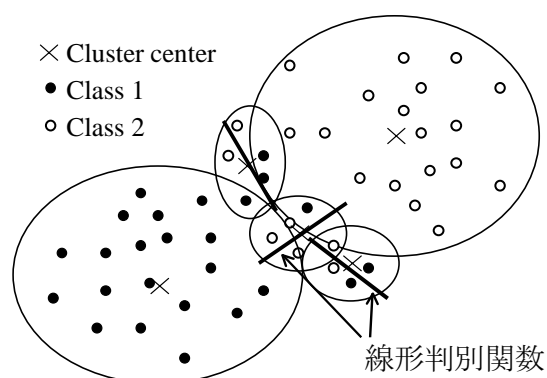


図 1: スwitchング線形判別

されている。また、回帰分析と主成分分析を FCM 法に組み合わせた手法 [9] も提案されている。

本研究では、データ集合の非線形性を考慮するために、クラスの境界付近にクラスター中心を配置するファジィクラスタリング法を提案し、クラスターごとに線形判別関数を用いるスイッチング線形判別法について述べる。ここでのファジィクラスタリング法は、ある係数によるクラス指標と特徴量の線形和の相関を最大にし、データとクラスター中心との距離を最小にすることによってクラスターを形成する。提案するファジィクラスタリング法は、クラスタリング規準であるクラス指標と特徴量の線形和の相関係数を正準相関係数と考え、山川ら [8] による正準相関分析と FCM 法の同時分析法とみなすこともできる。提案法により、線形判別分析における構造の簡便さ、利便性を残して、かつ、データ集合の非線形性を考慮した判別分析を行うことが出来る。数値実験では、提案法をアヤメ [1] とワイン [10] の識別問題に適用し有効性の検証を行う。

## 2. スwitchング線形判別

### 2.1 スwitchング線形判別のためのファジィクラスタリング法

スイッチング線形判別はデータをいくつかのクラスターに分け、それぞれのクラスターに対して線形判別関数を導出しようとするものである。特に、データ集合において異なったクラスのデータが混在しているような領域を細かく分割することによって識別性能は向上すると考えられる。すなわち、クラスの境界付近に

クラスター中心を配置することでそのような分割を実現することが出来る．境界付近にクラスター中心を配置するには，そのためのあるクラスタリング規準を設定する必要がある．

いま， $J$  個のデータ  $x_j$  が与えられており，それぞれ  $I$  種類のクラスのいずれかに属しているとする．データがどのクラスに属しているかを，あるクラスに属していれば 1，そうでなければ 0 とするような  $I$  次元のベクトルで表しクラス指標と呼ぶ．データ  $x_j$  が  $K$  次元の特徴量を持つとすると，ベクトル  $x_j^a, x_j^b$  を用いて以下のように表すことができる．

$$x_j = (x_j^a, x_j^b)^T, \quad j = 1, 2, \dots, J \quad (1)$$

$$x_j^a = (x_{j1}, x_{j2}, \dots, x_{jI})^T \quad (2)$$

$$x_j^b = (x_{jI+1}, x_{jI+2}, \dots, x_{jI+K})^T \quad (3)$$

ここで，

$$x_j^a \in \{(x_{ji}) \mid \sum_{i=1}^I x_{ji} = 1, x_{ji} \in \{0, 1\}\} \quad (4)$$

であり，上付き文字  $T$  は転置を表す．

クラス間の境界付近にクラスターを形成するようなクラスタリング規準を導入するために，まず，ある係数ベクトル  $a_c, b_c$  を用いて，クラス指標と特徴量それぞれに対して線形結合  $y_{cj}, z_{cj}$  を考える．

$$\begin{aligned} y_{cj} &= a_{c1}(x_{j1} - v_{c1}) + \dots + a_{cI}(x_{jI} - v_{cI}) \\ &= a_c^T (x_j^a - v_c^a) \end{aligned} \quad (5)$$

$$\begin{aligned} z_{cj} &= b_{c1}(x_{jI+1} - v_{cI+1}) + \dots + b_{cK}(x_{jI+K} - v_{cI+K}) \\ &= b_c^T (x_j^b - v_c^b) \end{aligned} \quad (6)$$

ここで，ベクトル  $v_c$  はクラスター  $c$  のクラスター中心（平均）を表し，

$$v_c = (v_c^a, v_c^b)^T, \quad c = 1, 2, \dots, C \quad (7)$$

$$v_c^a = (v_{c1}, v_{c2}, \dots, v_{cI})^T \quad (8)$$

$$v_c^b = (v_{cI+1}, v_{cI+2}, \dots, v_{cI+K})^T \quad (9)$$

である．また， $C$  はクラスターの個数を表す．

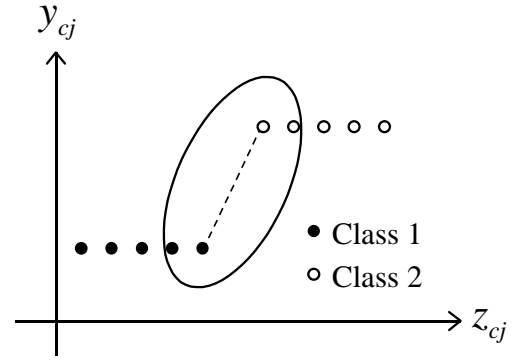


図 2: クラス指標と特徴量の線形和

いま，クラス指標と特徴量の線形和  $y_{cj}, z_{cj}$  の相関を最大にするように  $a_c, b_c$  を決定するとする．ただし，同一クラスのデータだけでクラスターが形成された場合， $y_{cj}$  の分散が 0 になり相関係数を算出することが出来ないで，あらかじめ微小な値の乱数をクラス指標に加えているものとする．算出された  $a_c, b_c$  を用いた  $y_{cj}-z_{cj}$  空間というのは図 2 のようになると考えられる．図 2 からわかるようにクラス間の境界付近だけに着目すると， $y_{cj}, z_{cj}$  の相関が大きいことがわかる．すなわち， $y_{cj}, z_{cj}$  の相関を最大にすることをクラスタリング規準とすると，境界付近に小さなクラスターが形成されることになる．しかし， $y_{cj}, z_{cj}$  の相関を最大にするというクラスタリング規準だけでは，同一クラスのデータが密集しているような空間に対してクラスターが形成されにくい．また，遠方の異なったクラス同士のデータでクラスターを形成する恐れもある．

同一クラスからなるクラスターを形成することは，判別が容易なデータ群を判別が困難な非線形境界領域から分離することを表し，元来クラスタリングを用いなくても線形判別が可能なデータ群を不要なクラスターに分割することを妨げることから，それだけで意味のあるクラスターを形成することになる．そのようなクラスターを形成するために，ファジィクラスタリングの代表的な手法である FCM 法で用いられる，クラスター中心とデータの距離の総和を最小にするというクラスタリング規準を導入する．FCM 法では，一般にクラスター中心からの距離としてユークリッド距離が

用いられ、円形のクラスターが形成されやすい。

上述の2つのクラスタリング規準に沿ったクラスターを形成するファジィクラスタリング法は、ラグランジュ未定乗数法を用いた目的関数の最適化問題に帰着させることができる。ここで求めるべき変数は、 $a_c$  と  $b_c$ 、クラスター中心  $v_c$ 、クラスターへの帰属の度合いを表すメンバシップ値  $u_c$  であり、

$$u_c \in \{(u_{cj}) \mid \sum_{c=1}^C u_{cj} = 1, u_{cj} \in [0, 1]\} \quad (10)$$

である。

最大化すべき目的関数  $L$  を以下のように定める。

$$\begin{aligned} L = & \sum_{c=1}^C \left[ \alpha \left\{ \sum_{j=1}^J \sum_{i=1}^I \sum_{k=I+1}^{I+K} u_{cj} (x_{ji} - v_{ci}) \right. \right. \\ & \times (x_{jk} - v_{ck}) a_{ci} b_{ck} \\ & - \frac{1}{2} \lambda_c^a \left( \sum_{j=1}^J \sum_{i=1}^I \sum_{k=1}^I u_{cj} (x_{ji} - v_{ci}) \right. \\ & \times (x_{jk} - v_{ck}) a_{ci} a_{ck} - 1 \Big) \\ & - \frac{1}{2} \lambda_c^b \left( \sum_{j=1}^J \sum_{i=I+1}^{I+K} \sum_{k=I+1}^{I+K} u_{ci} (x_{ji} - v_{ci}) \right. \\ & \times (x_{jk} - v_{ck}) b_{ci} b_{ck} - 1 \Big) \Big\} \\ & - (1 - \alpha) \sum_{j=1}^J \sum_{i=1}^{I+K} u_{cj} (x_{ji} - v_{ci})^2 \\ & - \beta \sum_{j=1}^J u_{cj} \log u_{cj} \Big] - \sum_{j=1}^J \gamma_j \left( \sum_{c=1}^C u_{cj} - 1 \right) \\ = & \sum_{c=1}^C \left[ \alpha \left\{ a_c^T \sum_c^{ab} b_c \right. \right. \\ & - \frac{1}{2} \lambda_c^a (a_c^T \sum_c^{aa} a_c - 1) \\ & - \frac{1}{2} \lambda_c^b (b_c^T \sum_c^{bb} b_c - 1) \Big\} \\ & - (1 - \alpha) \sum_{j=1}^J u_{cj} \|x_j - v_c\|^2 \\ & - \beta \sum_{j=1}^J u_{cj} \log u_{cj} \Big] - \sum_{j=1}^J \gamma_j \left( \sum_{c=1}^C u_{cj} - 1 \right) \quad (11) \end{aligned}$$

ただし、

$$\begin{aligned} \sum_c^{ab} &= (\sigma_{cjk}^{ab}) \\ \sigma_{cjk}^{ab} &= \sum_{j=1}^J u_{cj} (x_{ji} - v_{ci}) (x_{jk} - v_{ck}) \\ i &= 1, 2, \dots, I, \quad k = I+1, I+2, \dots, I+K \quad (12) \end{aligned}$$

$$\begin{aligned} \sum_c^{aa} &= (\sigma_{cjk}^{aa}) \\ \sigma_{cjk}^{aa} &= \sum_{j=1}^J u_{cj} (x_{ji} - v_{ci}) (x_{jk} - v_{ck}) \\ i &= 1, 2, \dots, I, \quad k = 1, 2, \dots, I \quad (13) \end{aligned}$$

$$\begin{aligned} \sum_c^{bb} &= (\sigma_{cjk}^{bb}) \\ \sigma_{cjk}^{bb} &= \sum_{j=1}^J u_{cj} (x_{ji} - v_{ci}) (x_{jk} - v_{ck}) \\ i &= I+1, I+2, \dots, I+K, \\ k &= I+1, I+2, \dots, I+K \quad (14) \end{aligned}$$

である。ここで、 $\alpha$  ( $0 \leq \alpha \leq 1$ ) と  $\beta$  ( $0 < \beta$ ) は、それぞれ2つのクラスタリング規準のトレードオフを決定するパラメータと、形成されるクラスターのファジィ度を決定するパラメータである。また、 $\lambda_c^a$ 、 $\lambda_c^b$ 、 $\gamma_j$  は、ラグランジュ乗数である。

式 (11) の  $a_c^T \sum_c^{ab} b_c$  は、係数ベクトル  $a_c$ 、 $b_c$  を用いたクラス指標と特徴量の線形結合  $y_{cj}$  と  $z_{cj}$  の相関係数を表し、 $\lambda_c^a (a_c^T \sum_c^{aa} a_c - 1)$  と  $\lambda_c^b (b_c^T \sum_c^{bb} b_c - 1)$  は線形結合  $y_{cj}$ 、 $z_{cj}$  の分散がそれぞれ1である制約を表す。 $\sum_{j=1}^J u_{cj} \|x_j - v_c\|^2$  は各クラスターのデータとクラスター中心との距離の総和であり、FCM法のクラスタリング規準を表す。 $\sum_{j=1}^J u_{cj} \log u_{cj}$  の項は、エントロピー正則化 [11] [12] の項であり、従来から用いられているメンバシップ値のべき乗の代わりに採用したものである。エントロピー正則化を用いると、従来のメンバシップ値のべき乗を用いる場合の例外処理を必要とせず、アルゴリズムを簡便にすることができる。最後の  $\sum_{j=1}^J \gamma_j (\sum_{c=1}^C u_{cj} - 1)$  の項は、データの各クラスターに対するメンバシップ値の和が1である制約を表す。

目的関数  $L$  を最大にするような  $v_c$ 、 $a_c$ 、 $b_c$ 、 $u_c$  は

FCM 法と同様に、最適性の必要条件を繰り返し適用するアルゴリズムによって計算される。

クラスター中心  $v_c$  に関する必要条件、

$$\begin{aligned} \frac{\partial L}{\partial v_{ci}} &= -\alpha \sum_{k=I+1}^{I+K} a_{ci} b_{ck} \sum_{j=1}^J u_{cj} (x_{jk} - v_{ck}) \\ &+ \frac{1}{2} \lambda_c^a \sum_{k=1}^I a_{ci} a_{ck} \sum_{j=1}^J u_{cj} (x_{jk} - v_{ck}) \\ &+ 2(1 - \alpha) \sum_{j=1}^J u_{cj} (x_{ji} - v_{ci}) \\ &= 0 \end{aligned} \quad (15)$$

を満たす解は、

$$v_{ci} = \frac{\sum_{j=1}^J u_{cj} x_{ji}}{\sum_{j=1}^J u_{cj}}, \quad i = 1, 2, \dots, I + K \quad (16)$$

である。

$a_c$ 、 $b_c$  についての最適性の必要条件  $\partial L / \partial a_c = 0$ 、 $\partial L / \partial b_c = 0$  より、

$$\sum_c^{ab} b_c - \lambda_c^a \sum_c^{aa} a_c = 0 \quad (17)$$

$$a_c^T \sum_c^{ab} - \lambda_c^b b_c^T \sum_c^{bb} = 0 \quad (18)$$

を得る。式 (17)、式 (18) より、

$$\lambda_c^a = \lambda_c^b = a_c^T \sum_c^{ab} b_c \quad (19)$$

が導かれる。 $\sum_c^{aa}$ 、 $\sum_c^{bb}$  が正則であるとする、 $\lambda_c = \lambda_c^a = \lambda_c^b$  として、

$$(\sum_c^{aa})^{-1} \sum_c^{ab} (\sum_c^{bb})^{-1} (\sum_c^{ab})^T a_c = \lambda_c^2 a_c \quad (20)$$

を得る。したがって、目的関数  $L$  を最大にするような  $a_c$  を求める問題は式 (20) の固有値問題に帰着し、解は最大固有値  $\lambda_{\max}^2$  に対応する固有ベクトル  $a_c^{\max}$  によって求まる。ただし、線形結合  $y_c$  の分散が 1 である制約より、係数  $a_c$  は、

$$a_c = \frac{a_c^{\max}}{\sqrt{a_c^{\max T} \sum_c^{aa} a_c^{\max}}} \quad (21)$$

となる。係数  $b_c$  は求まった  $a_c$  を用いて式 (18) より求めることができる。

メンバシップ値  $u_c$  は最適性の必要条件  $\partial L / \partial u_{cj} = 0$  より、

$$u_{cj} = \frac{\exp(A_{cj})}{\sum_{m=1}^C \exp(A_{mj})} \quad (22)$$

$$\begin{aligned} A_{mj} &= \frac{\alpha}{\beta} \left\{ \sum_{i=1}^I \sum_{k=I+1}^{I+K} (x_{ji} - v_{mi}) \right. \\ &\quad \times (x_{jk} - v_{mk}) a_{mi} b_{mk} \\ &\quad - \lambda_m \sum_{i=1}^I \sum_{k=1}^I (x_{ji} - v_{mi}) \\ &\quad \times (x_{jk} - v_{mk}) a_{mi} a_{mk} \\ &\quad - \lambda_m \sum_{i=I+1}^{I+K} \sum_{k=I+1}^{I+K} (x_{ji} - v_{mi}) \\ &\quad \times (x_{jk} - v_{mk}) b_{mi} b_{mk} \left. \right\} \\ &\quad - \frac{1 - \alpha}{\beta} \sum_{i=1}^{I+K} (x_{ji} - v_{mi})^2 \end{aligned} \quad (23)$$

となる。

以下にスイッチング線形判別のためのファジィクラスタリング法のアルゴリズムを記述する。

*Step 1:* 定数  $\alpha$ ,  $\beta$ ,  $C$  を定め、メンバシップ値  $u_c$  の初期値を乱数により与える。

*Step 2:* 式 (16) よりクラスター中心ベクトル  $v_c$  を求める。

*Step 3:* 式 (20) の最大固有値に対応する固有ベクトル  $a_c^{\max}$  を求め、式 (21) より係数  $a_c$  を求める。

*Step 4:* 求まった係数  $a_c$  を用いて式 (18) より係数  $b_c$  を求める。

*Step 5:* 式 (22) よりメンバシップ値  $u_c$  を求める。

*Step 6:* 終了判定条件  $\max_{c,k} |u_{ck}^{NEW} - u_{ck}^{OLD}| < \varepsilon$  を満たせば終了。そうでなければ *Step 2* へ。

## 2.2 線形判別手法



スイッチング線形判別ではデータ集合をいくつかのクラスターに分割した後、各クラスターで線形判別を行う。スイッチング線形判別のためのクラスタリング手法はファジィクラスタリングであるので、クラスターへの帰属はあいまいに表現されている。クラスターごとに線形判別関数を導出するには、どのデータがどのクラスターに属するかをクリスプに表されなければならない。本研究では、データは最大のメンバシップ値を持つクラスターに属するとする。

各クラスターに適用する線形判別手法として2種類の方法を採用する。1つは簡便な手法で簡略判別法と呼ぶ。簡略判別法は、ある2つのクラス間で中心(平均)を求め、それらを通る直線上にデータを射影し判別を行うものである。もう一つの線形判別法としては、従来のFisherの線形判別[1]を採用する。Fisherの線形判別は、異なるクラスのデータ同士はできるだけ離れ、同一クラスのデータ同士はなるべくかたまって分布するように、クラスター中心を通るベクトルにデータを射影して判別分析を行うものである。2クラスのデータ集合を判別するときの線形判別関数は以下のよう求められる。

$$y_c = \mathbf{w}_c^T \mathbf{x}_j^b, \quad c = 1, 2, \dots, C, \quad \mathbf{x}_j \in \text{cluster } c \quad (24)$$

ここで、

$$\mathbf{w}_c^T = \sum \mathbf{w}_c^{-1} (\bar{\mathbf{x}}_{c1} - \bar{\mathbf{x}}_{c2}) \quad (25)$$

$$\sum \mathbf{w}_c = \frac{1}{N_c} (N_{c1} \sum_{c1} + N_{c2} \sum_{c2}) \quad (26)$$

$$N_c = N_{c1} + N_{c2} \quad (27)$$

$\bar{\mathbf{x}}_{c1}$ ,  $\bar{\mathbf{x}}_{c2}$  は、クラスター  $c$  におけるクラス1とクラス2それぞれの平均値ベクトルである。 $N_{c1}$ ,  $N_{c2}$  はクラスター  $c$  におけるそれぞれのクラスのデータ数であり、 $N_c$  は2つのクラスの総データ数である。また、 $\sum_{c1}$ ,  $\sum_{c2}$  はクラス1とクラス2の共分散行列である。Fisherの線形判別では、各データが平均値ベクトル間を結ぶベクトルを各クラスの共分散行列のデータ数による加重和の行列の逆行列で調整したベクトルに射影されることになる。そして、 $y_c$  の値があるしきい値より大きい小さいかでクラスを判別する。一方の簡略判別法も同様に、しきい値を用いて判別を行う。

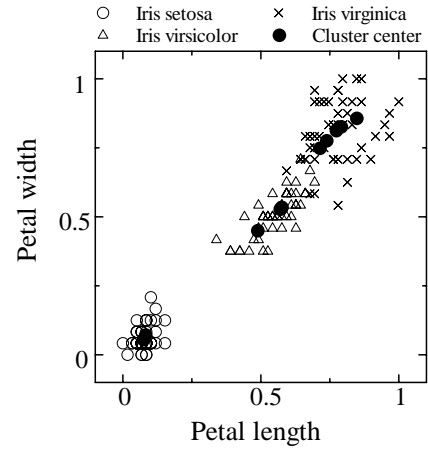


図 3:  $\alpha = 0.3$  の場合のクラスター中心

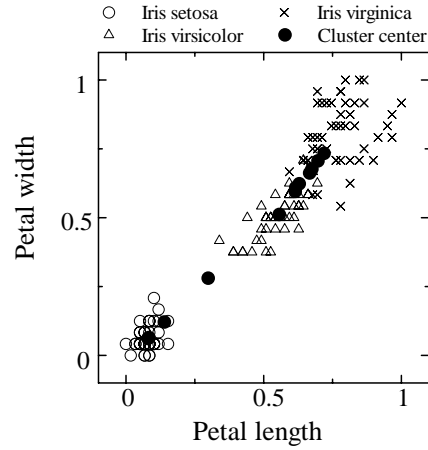


図 4:  $\alpha = 0.9$  の場合のクラスター中心

### 3. 数値実験

#### 3.1 アヤメのデータを用いた実験

数値実験では、提案手法を良く知られた Anderson のアヤメのデータ [1] に適用し有効性を検証する。アヤメのデータは、4次元パターン空間内での3群識別問題である。各クラスは、それぞれ50個のデータが与えられていて、総計150個のデータからなる。

まず初めに、クラスタリング規準の重視度  $\alpha$  の相違によって形成されるクラスターがどのように異なるかを調べた。 $\alpha$  の値を0.3, 0.9として数値実験を行った。結果を図3, 4に示す。ここで、クラスターの数を12個、ファジィ度を決定するパラメータ  $\beta$  を0.8とした。

各図において判別されるべきクラス *Iris setosa*, *Iris versicolor*, *Iris virginica* は、それぞれ  $\circ$ ,  $\triangle$ ,  $\times$  で示されている。また、クラスター中心は  $\bullet$  で示されている。アヤメのデータは、2つの特徴量によって3つのクラスがうまく分類されることが知られている。したがって、それら2つの特徴量、*Petal length*と*Petal width*の2次元空間上のプロットを図3, 4に示している。どのようなクラスターが形成されたかを、各クラスターに含まれるデータの平均ベクトルである黒丸のクラスター中心の分布によって示す。なお、実験結果は局所最適解である可能性があるが、いくつか初期値を変えて実験を行ったところ、ほとんど差異が見られなかったので最適解が得られていると考えられる。

図4に示すように、 $\alpha$ の値が大きいとき、すなわちクラス指標と特徴量の線形和との相関に重きをおいた場合、クラスの決定境界付近にクラスター中心が集中している。境界付近の非線形性が強い場合では、その近傍をより細かく分割することによって識別率が向上すると考えられるので、 $\alpha$ の値を高く設定すればよいことになる。

スイッチング線形判別においてデータは、最大のメンバシップ値を持つクラスターに属するようにクリスプに分割される。得られたクラスターの例を図5に示す。太い $\triangle$ と $\times$ で示されたデータが、クラスターに含まれるデータである。図5に示すように判別境界を定めると、クラスターに含まれる2つのクラスはうまく分割されることになる。本研究で採用した2つの線形判別手法とも、図6, 7に示すように、線形判別関数上でうまく2つのクラスを判別している。

表1に、従来の線形判別と2つの異なる線形判別法を用いたスイッチング線形判別のアヤメのデータに対する識別率を示す。線形判別手法として簡略判別法を用いたものをスイッチング線形判別1、Fisherの線形判別を用いる方法をスイッチング線形判別2と称した。スイッチング線形判別のためのファジィクラスタリング法に用いたパラメータは、 $\alpha = 0.9$ ,  $\beta = 0.8$ とし、クラスター数は12とした。 $\alpha$ を0.9としたのは、決定境界付近に多くのクラスターが形成され、高い識別率が期待できるからである。表1より、従来のFisherの線形判別よりも、スイッチング線形判別の方が識別率

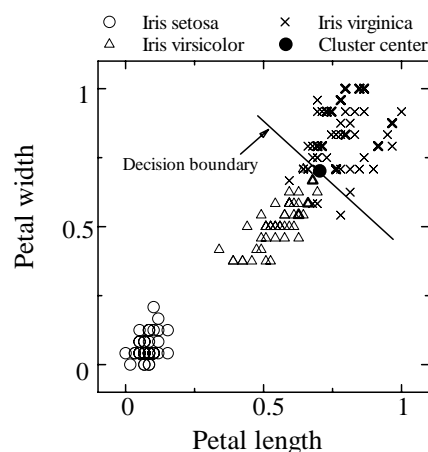


図 5: 形成されるクラスターの例

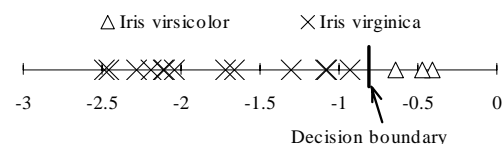


図 6: 簡略判別法

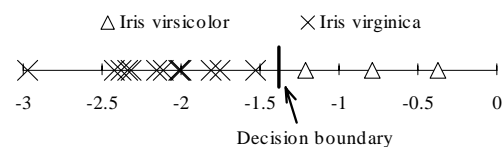


図 7: Fisher の線形判別

が高いことがわかる。さらに、スイッチング線形判別に関しては、クラスターごとにFisherの線形判別を用いることでより高い識別率が得られている。このことにより、非線形な境界領域をいくつかのクラスターに分割して線形判別を行うことは有効であるといえる。また、 $\alpha$ を0.0に設定すると従来のFCM法を行うことになるが、そのときの識別率は98%でクラスタリングを行わないFisherの線形判別と同じ結果となった。このことから、提案したクラスタリング手法の有効性がうかがえる。

### 3.2 ワインのデータを用いた実験

アヤメのデータは比較的特徴ベクトルの次元数が小

表 1: アヤメのデータに対する識別率

手法	識別率 (%)
Fisher の線形判別	98.000
スイッチング線形判別 1	98.667
スイッチング線形判別 2	99.333

さい識別問題であるが、高次元の識別問題に対しても数値実験を行った。高次元の識別問題として 13 次元 3 群判別問題のワインデータ [10] を用いた。ワインデータの各クラスはそれぞれ 59, 71, 48 個のデータからなり、総計 178 個のデータからなる。

本数値実験では、クラスター数と識別率の関係を調べる。表 2 に結果を示す。用いたパラメータは、前述の実験と同様  $\alpha = 0.9$ ,  $\beta = 0.8$  とした。クラスター数を 9, 12, 15, 18, 20 個として実験を行った。また、判別手法として 3.1 節で性能が高かった Fisher の線形判別を用いた。表 2 より、クラスターの数にかかわらず概して良い結果が得られていることがわかる。特にクラスター数を 18 個とすると、全てのクラスの識別が可能となっている。ただし、クラスター数が増えれば増えるほど識別率が良くなるとは限らない。これは、クラスター数によって形成されるクラスターが異なるからである。したがって、スイッチング線形判別を行うときのクラスター数の設定には試行錯誤的調整が必要となる。このことは、2 つのクラスタリング規準のトレードオフパラメータ  $\alpha$  やファジィ度を決定する重み係数  $\beta$  についても同様にいえる。

表 2: ワインのデータに対する識別率

クラスター数	識別率 (%)
9	99.438
12	99.438
15	98.876
18	100.00
20	99.438

## 4. おわりに

本研究では、データ集合をいくつかのクラスターに

分割し、クラスターごとに線形判別関数を適用するスイッチング線形判別を提案した。そのためのファジィクラスタリング法は、データ集合の非線形性の強い境界付近にクラスターを形成する。数値実験では、パターン識別問題でよく用いられるベンチマークであるアヤメとワインのデータに提案法を適用した。アヤメのデータを用いた数値実験では、2 つのクラスタリング規準のトレードオフパラメータと形成されるクラスターの性質を調べた。ワインデータを用いた数値実験では、クラスターの数とスイッチング線形判別の性能の関係を調べた。

クラスター分割を用いているので当然のことであるが、本研究における提案手法は既知のデータに対する線形判別の性能を向上させることができる。線形判別関数は判別のための一つの総合指標を与えるが、提案手法では少数の指標（線形判別関数）を用いることで、一つの指標では困難な問題に有効である。どのような問題に適用し、どのように指標を与えるかについては今後の課題としたい。

## 参考文献

- [1] R. A. Fisher : The use of multiple measurements in taxonomic problems, *Annals of Eugenics London*, Vol. 7, pp. 179-188 (1936)
- [2] J. C. Bezdek : *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York (1981)
- [3] J. C. Bezdek, C. Coray, R. Gunderson and J. Watson : Detection and Characterization of Cluster Substructure. I. Linear Structure: Fuzzy  $c$ -Lines, *SIAM J. Appl. Math.*, Vol. 40, No. 2, pp. 339-357 (1981)
- [4] J. C. Bezdek, C. Coray, R. Gunderson and J. Watson : Detection and Characterization of Cluster Substructure. II. Fuzzy  $c$ -Varieties and Convex Combinations Thereof, *SIAM J. Appl. Math.*, Vol. 40, No. 2, pp. 358-372 (1981)
- [5] D. W. Hosmer, Jr. : Maximum likelihood estimates of the parameters of a mixture of two regression lines, *Communications in Statistics*, Vol. 3, No. 10, pp. 995-1005 (1974)
- [6] R. J. Hathaway and J. C. Bezdek : Switching regression models and fuzzy clustering, *IEEE Trans. on Fuzzy Systems*, Vol. 1, No. 3, pp. 195-204 (1993)
- [7] 金海好彦, 山川あす香, 市橋秀友, 三好哲也 : ファジィ・クラスタリングと数量化 3 類の同時分析法, システム制御情報学会論文誌, Vol. 12, No. 3, pp. 191-198 (1999)

- [8] 山川あす香, 市橋秀友, 三好哲也: 正準相関係数を最大とするファジィ  $c$ -Means クラスタリング法, 日本経営工学会論文誌, Vol. 51, No. 1, pp. 17-26 (2000)
- [9] 本多克宏, 山川あす香, 市橋秀友, 三好哲也, 奥山哲史: ファジィ・クラスタリングと回帰と主成分の同時分析法, システム制御情報学会論文誌, Vol. 13, No. 5, pp. 236-243 (2000)
- [10] M. Forina, *et al.*: Wine recognition data, Available via anonymous FTP from ics.uci.edu in directory /pub/machine-learning-databases/wine (1992)
- [11] S. Miyamoto and M. Mukaidono: Fuzzy  $c$ -means as a regularization and maximum entropy approach, *Proc. of the 7th International Fuzzy Systems Association World Congress*, Vol. II, pp. 86-92 (1997)
- [12] 宮本 定明, 馬屋原 一孝, 向殿 政男: ファジィ  $c$ -平均法とエントロピー正則化法におけるファジィ分類関数, 日本ファジィ学会誌, Vol. 10, No. 3, pp. 548-557 (1998)

#### [ 問い合わせ先 ]

〒 599-8531

大阪府堺市学園町 1-1

大阪府立大学大学院工学研究科

電気・情報系専攻経営工学分野

市橋 秀友

TEL : 0722-54-9352

FAX : 0722-54-9915

E-mail: ichi@ie.osakafu-u.ac.jp

本多 克宏 (ほんだ かつひろ)[正会員]

1999 年 3 月大阪府立大学大学院工学研究科博士前期課程電気・情報系専攻修了。同年 4 月日本電信電話(株)入社。同年 10 月大阪府立大学工学部助手となり現在に至る。ニューラルネットワーク, ファジィクラスタリングの研究に従事。日本経営工学会, 日本ファジィ学会の会員。

市橋 秀友 (いちはし ひでとも)[正会員]

1971 年大阪府立大学工学部経営工学科卒業。同年松下電器産業(株)入社, 1981 年大阪府立大学工学部経営工学科助手, 1987 年同講師, 1989 年同助教授, 1993 年同教授, 現在に至る。工学博士。ファジィ多目的計画法, ファジィ制御, ニューロ・ファジィなどのファジィ理論の応用研究に従事。IEEE, 日本ファジィ学会, システム制御情報学会, 電子情報通信学会, 日本経営工学会などの会員。

#### 著者略歴

呉 志賢 (お ちひょん) [学生会員]

1999 年大阪府立大学大学院工学研究科博士前期課程電気・情報系専攻修了。同年同大学院工学研究科博士後期課程電気・情報系専攻入学, 現在に至る。ファジィクラスタリング, 強化学習などに興味を持つ。日本ファジィ学会, システム制御情報学会, 電子情報通信学会, IEEE の会員。

大迫 真実 (おおさこ まさみ)[非会員]

2000 年 3 月大阪府立大学大学院工学研究科博士前期課程電気・情報系専攻修了。同年 4 月(株)豊田自動織機製作所入社。現在に至る。在学中はファジィクラスタリングの研究に従事。