ファジィ領域の解析による特徴量の選択

3K－2

Ruck Thawonmas　阿部重夫
日立製作所　日立研究所

1. はじめに

パターン認識の分野では最適な特徴量の決定は、最小規模で最良の認識系を実現するうえから重要な課題である．特徴量を決定する方法としては元の入力を低次元の特徴量に変換する特徴抽出と元の入力から必要な特徴量を選択する特徴選択とがある．本論文ではファジィルールを生成するときに得られるクラス間の重なりの度合いを定量化する指標により不要な特徴量を削除する方法を提案する．そのために以下ではまず数値データよりファジィルールを抽出する方法について述べ，ついで特徴量を選択する方式について述べる

2. ファジィルールの抽出

まず、教師データから直接ファジィルールを抽出する方法を文献1)にしたがって述べる．

m次元の教師データを用いてn個のクラスに分離するファジィルールを生成するとする．X_iをクラスiに属する教師データの集合とする．最初にX_iを用いて，レベル1の活性領域を次のように定義する．

$$A_{ii}(1) = \{x \mid v_{iik}(1) \le x_k \le V_{iik}(1),\ k = 1,\ldots,\ m\} \quad (1)$$

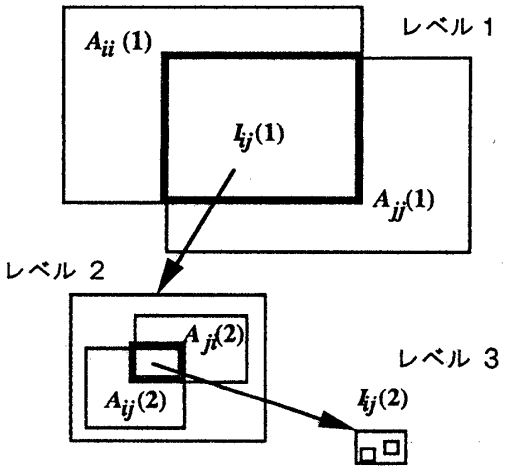

図1　ファジィルールの抽出

Input Feature Selection by Analysis of Fuzzy Regions
Ruck Thawonmas and Shigeo Abe, Hitachi, Ltd.

ただしx_k: 入力ベクトル xのk番目の要素;

$v_{iik}(1)$: x_k ($x \in X_i$) の最小値

$V_{iik}(1)$: x_k ($x \in X_i$) の最大値　(2)

このとき，クラスiと他のクラスの活性領域に重なりがないときは，次のレベル1のファジィルールをクラスiについて定義する．

If x is $A_{ii}(1)$ then x is class i　(3)

もし，重なりがあるときはそれを，図1に示すように再帰的に解消する．図は，2変数のデータに対して，ファジィルールを定義する方法を示したものでありクラスi, j間のレベル1の禁止領域を次式で定義する．

$$I_{ij}(1) = \{x \mid w_{ijk}(1) \le x_k \le W_{ijk}(1),\ k = 1,\ldots,\ m\} \quad (4)$$

ただし $v_{iik}(1) \le w_{ijk}(1) \le W_{ijk}(1) \le V_{iik}(1)$

これより禁止領域を持ったレベル1のファジィルールは次のようになる．

If x is $A_{ii}(1)$ and x is not $I_{ij}(1)$ then x is class i　(5)

もし，$I_{ij}(1)$の中にX_iに属するデータがあれば，レベル2の活性領域を次のように定義する．

$$A_{ij}(2) = \{x \mid v_{ijk}(2) \le x_k \le V_{ijk}(2),\ k = 1,\ldots,\ m\} \quad (6)$$

ただし　$x \in X_i$ で x は$I_{ij}(1)$の中にあり次式が成り立つ．

$$w_{ijk}(1) \le v_{ijk}(2) \le x_j \le V_{ijk}(2) \le W_{ijk}(1)$$

もしレベル2の活性領域に重なりがなければ，クラスiに対する次のレベル2のファジィルールが求まる．

If x is $A_{ij}(2)$ then x is class i　(7)

もし，$A_{ij}(2)$ and $A_{ji}(2)$の間に重なりがあると，次のようにレベル2の禁止領域を定義する．

$$I_{ij}(2) = \{x \mid w_{ijk}(2) \le x_k \le W_{ijk}(2),\ k = 1,\ldots,\ m\} \quad (8)$$

ただし $v_{ijk}(2) \le w_{ijk}(2) \le W_{ijk}(2) \le V_{ijk}(2)$

従って，禁止領域を持ったレベル2の次のファジィルールが得られたことになる．

If x is $A_{ij}(2)$ and x is not $I_{ij}(2)$ then x is class i　(9)

同様にして，もし重なりが残っていればレベル3以上のファジィルールを定義できる．ここで，一

般にレベルlのファジィルールを次のように定義する．

$$\text{If x is } A_{ij}(l) \text{ then x is class i} \tag{10}$$

あるいは，禁止領域を持ったレベルlのファジィルールは

$$\text{If x is } A_{ij}(l) \text{ and x is not } I_{ij}(l) \text{ then x is class i} \tag{11}$$

ただしl=1のときi=jで$l \geq 2$のときi≠jである．

3. ファジィ領域に基づく特徴量の選択

3.1 分離の複雑度

ここでm次元入力のときにクラスi, jの重なりがないとき，すなわち禁止領域$I_{ij}(1)$が定義されなかったときは，クラスi, jの分離は$A_{ii}(1)$, $A_{jj}(1)$により，あるいはこれらの活性領域の間の任意に位置に超平面を設定することにより容易に実現できる．そこで，このときのクラスiのクラスjに対する複雑度を0と考える．このとき，ある入力を削除してもクラスi, jの間に禁止領域$I_{ij}(1)$が定義されなかったときは，2つのクラスの分離のしやすさはm次元入力のときと変わらず，この2つのクラスに関してはこの入力は削除可能な入力と考えるわけである．

つぎに，レベルlの活性領域$A_{ij}(l)$に対して禁止領域$I_{ij}(l)$が定義された場合は，活性領域$A_{ij}(l)$と禁止領域$I_{ij}(l)$の体積比をクラスiのクラスjに対するレベルlの複雑度と考える．レベルが深くなるにつれ，定義される領域の体積が減るから，レベルlの複雑度はレベル$l+1$の複雑度より大きな重み（あるいは出現確率）をつける必要がある．複雑度は禁止領域の大きさに依存するからこの重みは禁止領域に比例したものとする必要がある．その一つの方法として，レベルlにおけるクラスiのクラスjに対す重み$p_{ij}(l)$を

$$p_{ij}(l) = (\text{禁止領域} I_{ij}(l) \text{におけるクラスiのデータの数}) / (\text{全体のデータの数}) \tag{12}$$

とする．ここで全体のデータの数で割るのは，各々のクラスのデータの数が実際の出現確率に対応すると考えていることによる．もし，各々のクラスのデータ数が等しいときは(12)式はクラスのデータ数で割ったのと同じになる．(12)式の代わりに禁止領域の体積を重みとしてもよい．

以上より，入力の集合Fに対するクラスiのクラスjに対する複雑度$o_{ij}(F)$はつぎのように定義される．

$$o_{ij}(F) = \sum_{l} p_{ij}(l) \frac{B_{I_{ij}}(F, l)}{B_{A_{ij}}(F, l)} \tag{13}$$

ただし

$$B_{X_{ij}}(F, l) = \prod_{f \in F} b_{X_{ij}}(f, l)$$

$$b_{I_{ij}}(f, l) = \begin{cases} W_{ijf}(l) - w_{ijf}(l), & W_{ijf}(l) - w_{ijf}(l) > \varepsilon \\ \varepsilon, & W_{ijf}(l) - w_{ijf}(l) \leq \varepsilon \end{cases}$$

$$b_{A_{ij}}(f, l) = \begin{cases} V_{ijf}(l) - v_{ijf}(l), & V_{ijf}(l) - v_{ijf}(l) > \varepsilon \\ \varepsilon, & V_{ijf}(l) - v_{ijf}(l) \leq \varepsilon \end{cases} \tag{14}$$

でεは小さな正の数である．

これより入力の集合Fに対する複雑度O(F)は

$$O(F) = \sum_{\substack{i, j = 1, \ldots, n \\ i \neq j}} o_{ij}(F) \tag{15}$$

で与えられる．

3.2 特徴選択

m次元の入力の番号1,..., mを要素として含む集合をIとして，m個全ての入力を用いたときの複雑度を求めこれをO(I)とする．次に集合Iを集合Fに設定して，添字iを1に初期化する．集合Fからi番目の変数を削除した集合をF(i)として，入力のうち集合F(i)に対応する複雑度を計算し，O(F(1)),..., O(F(|F|))の最小値を求めそれをO(F(k))とする．ただし|・|は集合の要素数を示す．最小値をとるのは複雑度が最も増加しない入力を削除するためである．つぎに

$$|O(I) - O(F(k))| / O(I) \geq \alpha \tag{16}$$

を計算する．ここで，αは入力の削除処理を打ち切るための正の数である．(16)式の左辺は全ての入力変数を使った場合からどの程度複雑度が相対的に増加したかを示す尺度になる．(16)式を満たすときは処理を終了し，満たさないときはさらに入力変数をへらす処理を行う．

4. 特徴量削減の効果

あやめのデータ，車番認識データ，サイロイドデータ，血球データで汎化能力を低下することなく特徴量が削減できた．詳細は講演時に示す．

5. 結言

クラス間のファジィ領域の重なりの複雑度を定量化して，複雑度が低下しない範囲で特徴量を選択する方式を述べた．

参考文献

1) S. Abe and M.-S. Lan, "A Method for Fuzzy Rules Extraction Directly from Numerical Data and Its Application to Pattern Classification," IEEE Trans. Fuzzy Systems, Vol. 3, No. 1, pp. 18-28, 1995.