楕円領域をもつファジィクラシファイアの学習方式

1C－9

阿部重夫　Ruck Thawonmas
日立製作所　日立研究所

1.はじめに

多層ネットの解析の難しさを解決する学習機能を持つファジィクラシファイア[1]が開発されている．これらの入力領域は，1) 楕円領域，2) 超直方体，3) 多面体で近似される．楕円領域を持つ代表的なクラシファイアに３層のラディアル基底関数クラシファイアがある．

本論文では，高速学習と高い汎化能力を実現するために最も単純な構成の２層の構成の楕円領域をもつファジィクラシファイアについて述べる．

2.クラシファイアの構成

m次元の入力ベクトルxをn個のクラスに分離することを考える．クラスタijをクラスiのj番目のクラスタとしてクラスiがいくつかのクラスタijに分割されているとする．このとき各々のクラスタijに対して次のルールR_{ij}が定義する．

$$R_{ij}\text{: If } \mathbf{x} \text{ is } \mathbf{c}_{ij} \text{ then } \mathbf{x} \text{ belongs to class } i \quad (1)$$

ただし，$\mathbf{c}_{ij}$はクラスタijの中心で，入力xに対するメンバーシップ関数$m_{ij}(\mathbf{x})$を次式とする．

$$m_{ij}(\mathbf{x}) = \exp\left(-h_{ij}^2(\mathbf{x})\right) \quad (2)$$

$$h_{ij}^2(\mathbf{x}) = \frac{d_{ij}^2(\mathbf{x})}{\alpha_{ij}} \quad (3)$$

$$d_{ij}(\mathbf{x}) = \sum_{k=1}^{m} \frac{(x_k - c_{ij,k})^2}{r_{ij,k}} \quad (4)$$

ただし$d_{ij}(\mathbf{x})$はxと$\mathbf{c}_{ij}$との重みつき距離，$h_{ij}(\mathbf{x})$は調整された距離，α_{ij} (> 0)はクラスタijの調整パラメータ，$\mathbf{c}_{ij} = (c_{ij,1}, \ldots, c_{ij,m})^t$で$r_{ij,k}$はクラスタijのk番目の半径である．入力xに対して，(2) - (4)式を計算してメンバーシップ関数の値を計算する．このときメンバーシップ関数$m_{kl}(\mathbf{x})$が最大のとき入力xがクラスkに属すると判定する．学習の手順は，まず各々のクラスをクラスタに分割し，各々のクラスタに属するデータを用いて中心$\mathbf{c}_{ij}$，および半径$r_{ij,k}$を推定する．ついで，正しく認識されたデータが誤認識するのを許してα_{ij}を調整して認識率を高める．

以下では紙面の関係でα_{ij}のチューニング法についてのみ述べる．

2.ファジィルールのチューニング

クラス間の重なりを考えないで定義したファジィルール(1)のα_{ij}を，教師データに対して認識率が最大になるように調整する．α_{ij}を増やすと(2)式で与えられるメンバーシップ関数の成立度は増加し，減らすと成立度は低下する．調整の考え方を示すために，図１の2つのクラスの各々に1つのルールが定義されたときを考える(簡単のためガウス関数のかわりに区分線形関数を用いている)．データ1はクラス2に正しく分類されているが，データ2, 3, 4はクラス2に間違って分類されている．もしα_{11}を増加あるいはα_{21}を減少するとデータ1が最初に誤認識するが，データ1の誤認識を許すとすると，データ2, 3, 4を正しく認識するようにできる．すなわち，クラス2のメンバーシップ関数を図の斜線の中に入るようにα_{21}を減少すると，データ1が誤認識するがデータ2, 3, 4は正しく認識される．このように，各々のファジィルールR_{ij}に対して認識率が改善されるときは正しく認識されるデータが誤認識されることを許して最適なα_{ij}の修正量を決める．

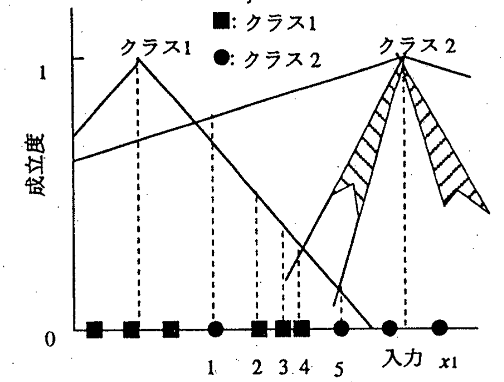

図１　チューニングの考え方

2.1 正しく認識されたデータから求まるα_{ij}の上下限

正しく認識されていたl-1個のデータの誤認識を許すときのα_{ij}の上限$U_{ij}(l)$および下限$L_{ij}(l)$を求める．入力教師データをファジィルール$\{R_{ij}\}$を用いて正しく分類されたデータの集合Xと誤認識したデータの集合Yに分割する．ついでクラスiに属するx (∈ X)で次式を満たすものを求める．

$$h_{ij}(\mathbf{x}) \le \min_{k \ne j} h_{ik}(\mathbf{x}) \quad (5)$$

もし(5)式が成立しないときはα_{ij}を変化させてもxは正しく認識されたままである．xがさらに

$$h_{ij}^2(\mathbf{x}) = \frac{d_{ij}^2(\mathbf{x})}{\alpha_{ij}} < \min_{o \ne i} h_{op}^2(\mathbf{x}) \le \min_{k \ne j} h_{ik}^2(\mathbf{x}), \quad (6)$$

を満足するとするとxを正しく認識させるための

Training of a Fuzzy Classifier with Ellipsoidal Regions
Shigeo Abe and Ruck Thawonmas, Hitachi, Ltd.

次式の下限$L_{ij}(x)$が存在することになる．

$$L_{ij}(\mathbf{x}) = \frac{d_{ij}^2(\mathbf{x})}{\min_{o \neq i} h_{op}^2(\mathbf{x})} < \alpha_{ij} \quad (7)$$

(6)式 が満たされないとき，すなわち

$$h_{ij}^2(\mathbf{x}) = \frac{d_{ij}^2(\mathbf{x})}{\alpha_{ij}} < \min_{k \neq j} h_{ik}^2(\mathbf{x}) < \min_{o \neq i} h_{op}^2(\mathbf{x}) \quad (8)$$

のときはα_{ij}を減少してもxに誤認識は生じない．

これより正しく認識されたデータを誤認識しない下限 $L_{ij}(1)$は

$$L_{ij}(1) = \max_{\mathbf{x} \in X} L_{ij}(\mathbf{x}) \quad (9)$$

で与えられる．ここで議論を簡単にするために異なるxに対して$L_{ij}(x)$は異なるとする．このとき (9) 式 は1つのxで満足される．一般に下限は次式で与えられる．

$$L_{ij}(l) = \max_{\mathbf{x} \in X,\, L_{ij}(\mathbf{x}) \neq L_{ij}(1), \ldots, L_{ij}(l-1)} L_{ij}(\mathbf{x}) \quad (10)$$

同様に上限$U_{ij}(l)$を求めることができる．これより α_{ij} は次のように制約される．

$$\ldots < L_{ij}(l) < L_{ij}(l-1) < \ldots < L_{ij}(1) < \alpha_{ij} < U_{ij}(1) < \ldots$$
$$< U_{ij}(l-1) < U_{ij}(l) < \ldots \quad (11)$$

ここでα_{ij} を $(L_{ij}(1), U_{ij}(1))$の範囲で変化させると，正しく認識されたデータ x (∈ X) は正しく認識されたままである．また $[U_{ij}(l-1), U_{ij}(l))$ あるいは $(L_{ij}(l), L_{ij}(l-1)]$の区間で変化させると，$l-1$ 個の正しく認識されたデータ x (∈ X) に誤認識を生じる．

2.2 誤認識を解消するα_{ij}の範囲

クラスiに属し，クラス o (≠i)に間違ったデータ x (∈ Y)について考える．このデータは次式が成立すると

$$\alpha_{ij} > V_{ij}(\mathbf{x}) = \frac{d_{ij}^2(\mathbf{x})}{\min_{p} h_{op}^2(\mathbf{x})} \quad (12)$$

$h_{ik}(x)$ (k ≠ i)の値にかかわらず正しく認識される．ここで$V_{ij}(x)$はα_{ij}の下限である．

α_{ij} を区間 $(\alpha_{ij}, U_{ij}(l))$の間の値に設定したとき，Inc(l)を正しく認識される誤認識のデータの数とする．もし $V_{ij}(x)$ が区間 $(\alpha_{ij}, U_{ij}(l))$ に含まれているときにInc(l)を1つカウントアップする．さらに

$$\beta_{ij}(l) = \max_{V_{ij}(\mathbf{x}) < U_{ij}(l)} V_{ij}(\mathbf{x}) \quad (13)$$

と定義する． α_{ij} を $\max(\beta_{ij}(l), U_{ij}(l-1))$より大きく設定すると，$l-1$個のデータが誤認識するが，誤認識していたInc($l$) 個のデータが正しく認識される．

同様にクラスoに属するデータxがクラスiに誤認識されたとする． α_{ij}を減少することにより xが正しく認識されるためにはクラスoの最小の調整された距離がn個のクラスの間で第2の最小値である必要がある．即ち次式の q は oである必要がある．

$$\min_{k} h_{ik}(\mathbf{x}) < \min_{q \neq i} h_{qr}(\mathbf{x}) . \quad (14)$$

つぎに $h_{ij}(x)$ はクラス iで最小で，クラスiの第2の最小値はクラスoの調整された距離の最小値より大きい必要がある．即ち

$$h_{ij}(\mathbf{x}) < \min_{p} h_{op}(\mathbf{x}) < \min_{k \neq j} h_{ik}(\mathbf{x}) \quad (15)$$

このときxは次式が成立すれば正しく認識される．

$$\alpha_{ij} < K_{ij}(\mathbf{x}) = \frac{d_{ij}^2(\mathbf{x})}{\min_{p} h_{op}^2(\mathbf{x})} \quad (16)$$

ここで $K_{ij}(x)$ はα_{ij}の上限である．

Dec(l) を α_{ij} を区間 $(L_{ij}(l), \alpha_{ij})$に設定したときに誤認識のデータが正しく認識される数とする． このとき$K_{ij}(x)$ が区間$(L_{ij}(l), \alpha_{ij})$に含まれるときにDec(l)を1つカウントアップする．ここで

$$\gamma_{ij}(l) = \min_{K_{ij}(\mathbf{x}) > L_{ij}(l)} K_{ij}(\mathbf{x}). \quad (17)$$

と定義する．もし α_{ij} を $\max(\gamma_{ij}(l), L_{ij}(l-1))$より小さく設定すると$l$ -1個の正しく認識されていたデータに誤認識が起こるがDec(l) 個のデータが正しく認識されることになる．

2.3 α_{ij}のチューニング

Inc(l) ($l = 1, \ldots, l_M$) に対して次式を満たすlを求める．ただしl_M は正の整数である．

$$\max_{l} \left(\mathrm{Inc}\,(l) - l + 1\right) \quad (18)$$

同様に Dec(l) ($l = 1, \ldots l_M$)に対して次式を満たすlを求める．

$$\max_{l} \left(\mathrm{Dec}\,(l) - l + 1\right) \quad (19)$$

(18) が (19)より大きいか等しい場合に，もし α_{ij} を区間$(\alpha_{ij}, U_{ij}(l))$において$\beta_{ij}(l)$より大きくすると，正しく 認識されるデータの増加数はInc(l) - l + 1個となる．そこで α_{ij} を区間 $[\beta_{ij}(l), U_{ij}(l))$ で次のように設定する．

$$\alpha_{ij} = \beta_{ij}(l) + \delta\left(U_{ij}(l) - \beta_{ij}(l)\right) \quad (20)$$

ここで δ は 0 ≤ δ < 1を満たす．

(18)式 が(19)式より小さいときは，次のように α_{ij} を $(L_{ij}(l), \gamma_{ij}(l)]$内で$\gamma_{ij}(l)$より小さく修正する．

$$\alpha_{ij} = \gamma_{ij}(l) - \delta\left(\gamma_{ij}(l) - L_{ij}(l)\right) \quad (21)$$

3. おわりに

楕円領域を有するファジィクラシファイアの学習方法について述べた．詳細は講演時に示すが，種々の問題で短い学習時間で多層ネットの最大性能を実現できた．

参考文献

1) 阿部重夫, "ニューラルネットとファジィシステム一理論と応用," 近代科学社, 1995年10月