教師データから抽出したファジィルールのチューニング法

1P－9

阿部重夫　日立製作所 日立研究所
Min-Shong Lan　Rockwell International Science Center
Ruck Thawonmas　東北大学 電気通信研究所

1．はじめに

多層ネットトは入出力データの複雑な関係を学習できるが、学習済みのネットの解析は難しい。これに対して、ファジィシステムの知識獲得は難しいが、一旦知識が獲得されればシステムの解析あるいは修正は比較的容易である。これらの両者の方式の溝を埋めるものとして、文献1, 2）で我々は、教師データから直接抽出したファジィルールを用いてファジィ推論する方法を提案した。本論文では汎化能力を向上するためのパターン認識用のファジィルールのチューニング法について述べる。

2. ファジィルールの抽出と推論方法

2.1 ルール抽出法

m次元の入力ベクトルxをn個のクラスに認識するファジィルールをクラスiに対する入力ベクトルの集合X_iから抽出することを考える。まず集合X_iに属する入力データの最小値と最大値とを計算してクラスiに対応する入力領域を定義するレベル１の活性領域$A_{ii}(1)$を求める。もしクラスiに対応する活性領域$A_{ii}(1)$がクラスjに対応する活性領域$A_{jj}(1)$と重なるときは重なった領域を禁止領域$I_{ij}(1)$として定義する。この禁止領域$I_{ij}(1)$を$A_{ii}(1)$と$A_{jj}(1)$との中で拡張することにより汎化能力の向上を図る。$A_{ii}(1)$と$A_{jj}(1)$の各々に対応する拡張禁止領域を$J_{ij}(1)$と$J_{ji}(1)$と定義する。拡張禁止領域に クラスiあるいはjのデータが存在するときは, 拡張禁止領域の中にさらに活性領域$A_{ii}(2)$および/あるいは$A_{jj}(2)$を定義する。さらに２つの活性領域が定義されそれが重なるときは、禁止領域$I_{ij}(2)$、拡張禁止領域$J_{ij}(2)$, $J_{ji}(2)$を定義する。この処理は重なりが解消されるまで続ける。

クラスjとの重なりを解消することにより得られたレベルlの クラスiのファジィルールを$r_{ij}(l)$とする。禁止領域なし、およびありのときのファジィルールは各々(1), (2)式で与えられる。

If x が$A_{ij'}(1)$ にあるthen x はクラス iに属す (1)

If x が $A_{ij'}(l)$ にあり x は$J_{ij}(l)$ にない
then x はクラス iに属す　(2)

ただし$l=1$で$j'=i$, $l\geq 2$ で$i\neq j'$である。

2.2 ファジィルールを用いた推論

(1)式で与えられるファジィルールに対する入力ベクトルxの成立度は、xが活性領域$A_{ij'}(l)$ の内側にあるときは1で、xが活性領域$A_{ij'}(l)$ の外側にあるときは同一の成立度の等高線が$A_{ij'}(l)$ の表面に平行で表面から等距離にあるとする。すなわち

$$m_{A_{ij'}(l)}(x) = \min_{k=1,\dots,m} m_{A_{ij'}(l)}(x, k), \quad (3)$$

$$m_{A_{ij'}(l)}(x, k) = [1-\max(0, \min(1, \gamma_i (v_{ij'k}(l) - x_k)))] \times [1 - \max(0, \min(1, \gamma_i (x_k - V_{ij'k}(l))))] \quad (4)$$

但し、 γ_i は感度定数 である。

以上よりファジィルール$r_{ij}(l)$に対するxの成立度は次式で与えられる。

$$d_{ij(l)}(x) = m_{A_{ij'}(l)}(x) \quad (5)$$

(2)式で与えられるファジィルールに対する入力ベクトルxの成立度は、xが活性領域$A_{ij'}(l)$ の内側にあるが$\overline{A_{ij'}(l) - J_{ij}(l)}$にないときは1とする。ただし、$\overline{S}$はSの閉集合を示し、$l=1$のとき$j'=i$で$l>1$のとき$j'=j$である。xがこの外側にあるときは同一の成立度の等高線は$\overline{A_{ij'}(l) - J_{ij}(l)}$ の表面に平行で表面から等距離にあるとする。このときの成立度は(3), (4)式と同様に計算される。

従ってファジィルールの集合$\{r_{ij}(l) \mid l=1,..\}$ に対する成立度$d_{ij}(x)$は次式で与えられる。

$$d_{ij}(x) = \max_{l=1,\dots} (d_{rij(l)}(x)) \quad (6)$$

ここで最大値をとるのは活性領域$A_{ij}(l+1)$が存在すれば, 禁止領域 $I_{ij}(l)$に含まれるため$\{r_{ij}(l) \mid l=1,..\}$の各々のルールは独立であるからである。

クラスiに対する入力 x の成立度$d_i(x)$ は次式で与えられる。

$$d_i(x) = \min_{\substack{j\neq i, j=1,\dots,n,\\ A_{ii}(1)\cap A_{jj}(1)\neq\emptyset}} (d_{rij}(x)). \quad (7)$$

Tuning Fuzzy Rules Derived From Data
Shigeo Abe, Hitachi, Ltd., Ming-Shong Lan
Rockwell International Corp. Ruck Thawonmas,
Tohoku University

クラスiの活性領域が1つ以上のクラスの活性領域と重なるときは、例えばこれをj, kとすると重なりを独立に解消している。これは(7)式において最小値をとることに対応している。

3. 最小二乗法によるチューニング

教師データを用いてファジィシステムが作られたとする。与えれた入力**x** に対してクラスiの成立度が1以下とする。すなわち$1 > d_i(\mathbf{x})$。これは入力**x**がクラスiの活性領域$A_{ii'}(l)$の外か、拡張禁止領域$J_{ij}(l)$の中にあることを意味している。この場合、ファジィルール$r_{ij}(l)$の成立度$d_{ij(l)}(\mathbf{x})$は**x**と$A_{ij'}(l)$あるいは$A_{ij'}(l)$ - $J_{ij}(l)$の表面を含む超平面からの距離の最大値あるいは最小値で決まる。このとき超平面は入力変数のひとつに平行である。$d_{ij(l)}(\mathbf{x})$が決定される超平面は同一クラスには同じ感度定数が与えられているためγ_iを変えても変化しない。ここで$d_{ij(l)}(\mathbf{x})$を決める超平面をk番目の入力変数に平行であるとして、k番目の入力変数の切片をb_kとする。このとき成立度$d_{ij(l)}(\mathbf{x})$は$\min(0, \gamma_i|x_k - b_k|)$で与えられる。またクラスiの全ての入力変数に対して同一の感度定数γ_iを使っているため、成立度$d_i(\mathbf{x})$を決める超平面は感度定数γ_iを変えても変わらない。

以上の議論より感度定数を決める最小二乗法を求めることができる。$1 > d_i(\mathbf{x}), i = 1,\ldots, n$を満たす入力**x**に対して距離ベクトル$\mathbf{t} = (t_1,\ldots, t_n)$を定義する。ここで$t_k = |x_l - b_l|$で$d_i(\mathbf{x})$を決める超平面は$l$番目の軸に平行で切片を$b_l$とする。クラスiのj番目の距離ベクトルを$\mathbf{t}_{ij} = (t_{ij,1},\ldots, t_{ij,n})$とする。ここで$1 \leq j \leq N_i$ でN_i はクラスiのチューニング用データの数である。もし

$$t_{ij,i} < t_{ij,k} \qquad k \neq i \text{を満たす全ての}k \tag{8}$$

が成立すると、全ての感度定数を同一にしたとき正しく認識できる。与えられたデータ$\mathbf{x}_{ij}$に対して、$k \neq i$ で$d_k(\mathbf{x}_{ij}) = 1$のときファジィ推論では正しくそのデータを認識できないが、$d_i(\mathbf{x}_{ij}) = 1$であれば感度定数の如何に係わりなく正しく認識できる。このため、これらふたつの条件を満たすデータは$\mathbf{x}_{ij}$ $(i = 1,\ldots, n, j = 1,\ldots, N_i)$には含まない、したがって$\mathbf{t}_{ij}$ $(i = 1,\ldots, n, j = 1,\ldots, N_i)$にも含まないとする。

$d_k(\mathbf{x}_{ij})$の目標値は$k = i$のとき1でそれ以外では0であるから$d_i(\mathbf{x}_{ij}) = 1 - \gamma_i t_{ij,i}$は正であると仮定できる。しかしながら、$d_k(\mathbf{x}_{ij}) = 1 - \gamma_k t_{ij,k}$は正にも負にもなりうる。負のときは感度定数を求めるときに0で置き換える必要がある。このため次のように誤認識のデータに対する削除定数δ_1と正しく認識されたデータに対する削除定数 δ_2を定義する。誤認識したデータ$\mathbf{x}_{ij}$に対して

$$t_{ij,i} - t_{ij,k} > \delta_1 \tag{9}$$

が全ての$k \neq i$ となるkに対して成立すると$\mathbf{x}_{ij}$を$\mathbf{x}_{ik}$ $(k = 1,\ldots, N_i)$から削除する。また正しく認識されたデータ$\mathbf{x}_{ij}$に対して

$$t_{ij,k} - t_{ij,i} > \delta_2 \tag{10}$$

が全ての$k \neq i$ となるkに対して成立すると$\mathbf{x}_{ij}$を$\mathbf{x}_{ik}$ $(k = 1,\ldots, N_i)$から削除する。

したがって、各々のクラスに対する感度定数は次の誤差関数を最小化することにより求まる。

$$E = \frac{1}{2}\{\sum_{i=1}^{n}\sum_{j=1}^{N_i}(\gamma_i t_{ij,i})^2 + \sum_{k=1}^{n}\sum_{\substack{i=1\\ i\neq k}}^{n}\sum_{j=1}^{N_i}(1 - \gamma_k t_{ij,k})^2\} \tag{11}$$

(11)式は次式が成立するとき最小化される。

$$\frac{\partial E}{\partial \gamma_i} = \{\sum_{k=1}^{n}\sum_{j=1}^{N_i}(t_{kj,i})^2\}\gamma_i - \sum_{\substack{k=1\\ k\neq i}}^{n}\sum_{j=1}^{N_i} t_{kj,i} = 0 \tag{12}$$

したがって、最適な感度定数は次式により求まる。

$$\gamma_i = \frac{\sum_{\substack{k=1\\ k\neq i}}^{n}\sum_{j=1}^{N_i} t_{kj,i}}{\sum_{k=1}^{n}\sum_{j=1}^{N_i}(t_{kj,i})^2} \tag{13}$$

削除定数δ_1, δ_2の最適値はわからないためδ_1, δ_2の異なる値で(13)式を計算する必要があるが、繰り返し計算でないため極めて高速に行なうことができる。

10. 結言

教師データから直接に抽出したファジィルールを最小二乗法によりチューニングする方法を開発した。この方式の数字認識での評価結果については講演時に示す。

11. 参考文献

1) S. Abe and M.-S. Lan, "A Classifier Using Fuzzy Rules Extracted Directly from Numerical Data," Proc. 2nd IEEE International Conference on Fuzzy Systems, pp. 1191-1198, March 1993.

2) S. Abe and M.-S. Lan, "A Function Approximator Using Fuzzy Rules Extracted Directly from Numerical Data," IJCNN'93 Nagoya, pp. 1887-1892, October, 1993.