

学習型トラピゾイドファジィ推論のハイパーパラメータの特性評価

Characterization of Hyperparameter of Learning Type Trapezoidal Fuzzy Inference

○ 入江 穂乃香
○ Honoka Irie
関西大学
Kansai University

林 勲
Isao Hayashi
関西大学大学院
Kansai University

Abstract: Learning type trapezoidal fuzzy inference is a general form of triangular fuzzy inference, and fuzzy clustering using fuzzy inference can identify the class boundary area by learning. It is necessary to adjust the trapezoidal membership function of the antecedent part and the singleton real value of the consequent part with the learning mechanism. In addition, initial value setting of parameters is an important factor for learning for minimizing error. In particular, the initial value of the singleton in the consequent part greatly changes the accuracy of the estimation of fuzzy inference depending on the set value. In this paper, we will discuss the characteristics of the hyperparameters to determine fuzzy rules of class discrimination with minimizing error. In addition, we discuss the influence of the learning parameters of the antecedent part and the consequent part on error accuracy, and the procedure of learning for improving learning. By the numerical examples, several combinations of extensive hyperparameters are discussed by significance test in order to acquire the optimal fuzzy rule of fuzzy clustering.

1. はじめに

ファジィ推論 [1,2] は、専門家の知識を if-then 形式で表現するモデルとして 1990 年代に多用され、ファジィ制御が組み込まれた製品はファジィ家電 [3] と呼ばれるなど、ファジィ制御は、家電製品や工場プラント制御、列車自動制御、自動車制御、ロボット制御、人工衛星制御等で質の高い有用性を示した。ファジィ推論は、メンバシップ関数を用いて入力値を度合付きとし、その度合で入力に対応する出力値を補間する非線形モデルといえる。しかし、入出力データが与えられた場合に、ファジィルールのチューニング問題 [4] が存在し、前件部のメンバシップ関数の形状と後件部の出力関数を調整する必要がある。特に、ファジィルールの実際上の設計問題として、学習係数や初期値の設定、学習順序の決定などは重要な課題である。

本論文では、ファジィ推論の一般形として学習型トラピゾイド(台形型)ファジィ推論を採用し、誤差が最小となるクラス判別のファジィルールを決定するための前件部と後件部のハイパーパラメータの特性について議論する。パラメータの調整を単なるチューニング問題として捉えるのではなく、具体的なファジィルールの設計問題として捉え、精度効率が良く高い有用性をもつファジィルールを構成するための多次元のパラメータの決定方法を議論する。この学習の最適性を目的とした多次元パラメータの学習をハイパーパラメータ学習と呼ぶ。

チューニング問題では、簡略ファジィ推論の学習において、市橋ら [5] は最急降下法を用いた後件部の学習効率化を提案し、野村ら [4] は最急降下法による前件部の

メンバシップ関数の調整法を提案している。石淵ら [6,7] はファジィ推論によるファジィ識別システムにおいて、パターン空間の分割数やメンバシップ関数の形状、メンバシップ値の演算方法などの条件と認識率との関係を議論している。さらに、パターン識別境界線の平滑化法を提案している。一方、馬野ら [8] は、パターン分割数を逐次的に増加させるファジィ推論に忘却型の学習則を導入し、ルール数の最適化を可能としている。

ここでは、これらの提案手法を参考に、三角形ファジィ推論の一般形であるトラピゾイド(台形型)ファジィ推論を用いて、後件部のシングルトンの学習係数とその初期値の設定、後件部と前件部の学習順序について議論する。特に、学習の初期値の設定は誤差最小化のための重要な課題である。通常は、乱数などで初期値を設定するが、適度な初期値から学習を開始しなければ誤差が減少しない。そこで、観測データに依存して初期値を設定する方法や観測データとは独立して初期値を設定する方法など、新たな 6 種類の初期値の設定方法を提案する。また、前件部や後件部のみの一方学習や前件部と後件部を交互に学習する交互学習など、学習順序も結果に大きく影響を与える。そこで、初期配置と前件部、後件部の学習順序についても、誤差への影響を有意性検定等を用いて議論する。

まず、トラピゾイドファジィ推論を定義し、ファジィルールの調整のための最急降下法の定式化を行う。次に、新たな 6 種類の初期配置法を提案し、学習順序について議論する。最後に、数値例を用いた誤差の評価により、これらのハイパーパラメータの特性について議論する。

2. ファジィ推論

簡易型ファジィ推論 [5] は, if-then 型のルールを用いた推論メカニズムである. 前件部では, メンバシップ関数のファジィ集合を定義し, 後件部では, シングルトンを定義する. ここでは, 三角型のメンバシップ関数を一般化した正規なトピゾイド (台形) 型ファジィ集合を用いる. 正規なファジィ集合とは, メンバシップ関数が連続で, メンバシップ値が 1.0 である定義域が必ず存在する凸ファジィ集合である.

いま, 入力変数 $x_j, j = 1, 2, \dots, n$ の台形型ファジィ集合 $\mu_F(x_j)$ を次の $L-R$ ファジィ集合で定義する.

$$\mu_F(x_j) = \begin{cases} L(\frac{x_b-x_j}{\alpha}) & ; x_j \leq x_b \\ 1.0 & ; x_b \leq x_j \leq x_c \\ R(\frac{x_j-x_c}{\beta}) & ; x_j \geq x_c \end{cases} \quad (1)$$

ただし, 下底の左端, 上底の左端, 上底の右端, 下底の右端の座標をそれぞれ, $(x_a, y_a), (x_b, y_b), (x_c, y_c), (x_d, y_d)$ で表し, 左の広がり $\alpha = |x_b - x_a|$ で, 右の広がり $\beta = |x_d - x_c|$ で表す. なお, $L(), R()$ は型関数と呼ばれ, $\max(0, 1 - |x|)$ なる関数である.

図 1 に正規なトラピゾイド型ファジィ集合を示す.

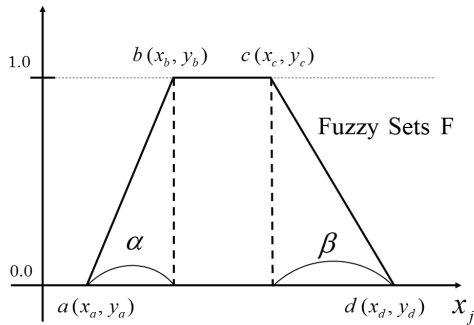


Fig. 1: Trapezoidal Fuzzy Sets

いま, メンバシップ関数を $\mu_F(x_j)$ で表し, 出力変数を z , 後件部のシングルトンを p_i で表すと, i 番目のファジィルール $r_i, i = 1, 2, \dots, R$ は次のようになる.

$$\begin{aligned} r_1: & \text{ if } x_1 \text{ is } \mu_{F_{11}}(x_1) \text{ and } \dots \text{ and } x_n \text{ is } \mu_{F_{1n}}(x_n) \\ & \text{ then } C = \{C_{1k} \mid z = p_1, \\ & \quad p_1 = g(p_{11}, \dots, p_{1d}, \dots, p_{1W_1})\} \\ & \vdots \\ & \vdots \\ r_i: & \text{ if } x_1 \text{ is } \mu_{F_{i1}}(x_1) \text{ and } \dots \text{ and } x_n \text{ is } \mu_{F_{in}}(x_n) \\ & \text{ then } C = \{C_{ik} \mid z = p_i, \\ & \quad p_i = g(p_{i1}, \dots, p_{id}, \dots, p_{iW_i})\} \\ & \vdots \\ & \vdots \\ r_R: & \text{ if } x_1 \text{ is } \mu_{F_{R1}}(x_1) \text{ and } \dots \text{ and } x_n \text{ is } \mu_{F_{Rn}}(x_n) \\ & \text{ then } C = \{C_{Rk} \mid z = p_R, \\ & \quad p_R = g(p_{R1}, \dots, p_{Rd}, \dots, p_{RW_R})\} \end{aligned}$$

ただし, C は出力クラスの変数であり, C_{ik} はルール r_i のクラス値が C_k であることを示す. また, $p_{id}, d = 1, 2, \dots, W_i$ は当該ルール r_i の前件部でメンバシップ値

の積演算が正值をもつデータ d のシングルトン, W_i はその個数, g は p_i への変換関数である.

いま, 入力データ $\mathbf{x} = (x_1, x_2, \dots, x_n)$ が得られたとしよう. 第 i 番目のファジィルール r_i の前件部に入力データ \mathbf{x} を入力し, 前件部の適合度 $\mu_i(\mathbf{x})$ を計算する.

$$\mu_i(\mathbf{x}) = \mu_{F_{i1}}(x_1) \cdot \mu_{F_{i2}}(x_2) \cdot \dots \cdot \mu_{F_{in}}(x_n) \quad (2)$$

ファジィ推論の結果 \hat{z} とクラス C は次式から求める.

$$\begin{aligned} \hat{z} &= \frac{\sum_{i=1}^R \mu_i(\mathbf{x}) \cdot p_i}{\sum_{i=1}^R \mu_i(\mathbf{x})} \\ C &= \{C_k \mid \min |\hat{z} - z|\} \end{aligned}$$

3. ハイパーパラメータの学習

入出力データが与えられたとき, ファジィ推論の結果と出力値との誤差を最小にするように, ファジィルールのパラメータを学習する. ただし, パラメータ学習を単なるチューニング問題として捉えるのではなく, 具体的なファジィルールの設計問題として捉え, 多次元パラメータの決定方法を議論する. そのため, ここでは, 学習の最適性を目的とした多次元パラメータの学習をハイパーパラメータ学習と呼ぶ. ハイパーパラメータ学習は次の 3 項目からなる.

学習係数: 前件部のメンバシップ関数と後件部のシングルトンは最急降下法で学習され, その学習係数を最適に決定する.

初期配置: 前件部のメンバシップ関数と後件部のシングルトンの初期値は最適に設定される.

学習順序: パラメータの学習順序を最適に決定する,

入出力データ $(\mathbf{x}, z) = (x_1, x_2, \dots, x_n, z)$ が得られた場合, 前件部のメンバシップ関数と後件部のシングルトンは次式の評価値 E が最小となるように学習される.

$$E = \frac{1}{2}(\hat{z} - z)^2 \quad (3)$$

ここでは, 前件部のメンバシップ関数と後件部のシングルトンを最急降下法で学習する. ファジィルール r_i の前件部のメンバシップ関数のパラメータ $(b_{ij}, \alpha_{ij}, c_{ij}, \beta_{ij})$ と後件部のシングルトン p_i は次式で更新される.

$$\begin{aligned} b_{ij}(t+1) &= b_{ij}(t) + \Delta b_{ij} \\ \alpha_{ij}(t+1) &= \alpha_{ij}(t) + \Delta \alpha_{ij} \\ c_{ij}(t+1) &= c_{ij}(t) + \Delta c_{ij} \\ \beta_{ij}(t+1) &= \beta_{ij}(t) + \Delta \beta_{ij} \\ p_i(t+1) &= p_i(t) + \Delta p_i \end{aligned}$$

ここで, t は学習回数であり, Δp_i 等はパラメータ p_i の更新量である.

$$\begin{aligned} \Delta b_{ij} &= -K_b \cdot \frac{\partial E}{\partial b} \\ \Delta \alpha_{ij} &= -K_\alpha \cdot \frac{\partial E}{\partial \alpha} \end{aligned}$$

$$\begin{aligned}\Delta c_{ij} &= -K_c \cdot \frac{\partial E}{\partial c} \\ \Delta \beta_{ij} &= -K_\beta \cdot \frac{\partial E}{\partial \beta} \\ \Delta p_i &= -K_p \cdot \frac{\partial E}{\partial p}\end{aligned}$$

ただし, $K_b, K_\alpha, K_c, K_\beta, K_p$ は学習係数である. ハイパーパラメータの個数は $(4n+1)R$ 個となる.

次に, 初期値の設定と学習誤差との関係について議論する. 前件部のメンバシップ関数と後件部のシングルTONは, 最急降下法によりデータが入力されるごとに逐次的に学習される. この逐次学習はメンバシップ関数とシングルTONの初期値から連続的に学習されるので, 学習結果は初期値に強く依存する. ここでは, この初期値の設定方法を初期配置と呼ぶ. ただし, 前件部のメンバシップ関数は定義域で均等配置とし, 後件部のシングルTONの初期配置のみを議論する.

後件部のシングルTONの初期配置として, 次の6種類を提案する.

(1) 総和法 (データ依存型)

当該ファジイルール r_i において, データ d の適合度 $\mu_i(\mathbf{x}_d)$ をクラスごとに総和し, 最大値となるクラスを後件部のシングルTONの初期値とする. シングルTONの初期値は観測データと強い依存関係を有する.

$$\begin{aligned}p_i &= g(p_{i1}, \dots, p_{iW_i}) \\ &= v(C_{idk}) \text{ for } \{C_{idk} | \max_k \sum_{d=1}^{W_i} \mu_i(\mathbf{x}_d)\}\end{aligned}$$

(2) 最大法 (データ依存型)

当該ファジイルール r_i において, データ d の前件部の適合度 $\mu_i(\mathbf{x}_d)$ が最大値となるデータのクラスを後件部のシングルTONの初期値とする. シングルTONの初期値は観測データと強い依存関係を有する.

$$\begin{aligned}p_i &= g(p_{i1}, \dots, p_{iW_i}) \\ &= v(C_{idk}) \text{ for } \{C_{idk} | \max_d \mu_i(\mathbf{x}_d)\}\end{aligned}$$

ただし, $v(C_{idk})$ はクラス C_{idk} の数値化を表す.

(3) 優勢法 (データ依存型)

当該ファジイルール r_i において, データ d の適合度 $\mu_i(\mathbf{x}_d)$ がしきい値 θ 以上となるデータ個数をクラスごとに求め, 最大個数となるクラスを後件部のシングルTONの初期値とする. シングルTONの初期値は観測データと弱い依存関係を有する.

$$\begin{aligned}p_i &= g(p_{i1}, \dots, p_{iW_i}) \\ &= v(C_{idk}) \text{ for } \{C_{idk} | \max_k \text{card}(\mu_i(\mathbf{x}_d) \geq \theta)\}\end{aligned}$$

ただし, $\text{card}(\mu_i(\mathbf{x}_d) \geq \theta)$ は $\mu_i(\mathbf{x}_d) \geq \theta$ となるデータ d の個数を表す. なお, $\mu_i(\mathbf{x}_d) < \theta$ の場合には, あらかじめ設定された値を初期値とする.

(4) 既定法 (データ独立型)

離散値である後件部のクラス値の中間値をシングルTONの初期値として設定する. シングルTONの初期値は観測データと弱い独立関係を有する.

(5) クラス乱数法 (データ独立型)

乱数によって, 離散値である後件部のクラス値をシングルTONの初期値として設定する. シングルTONの初期値は観測データと強い独立関係を有する.

(6) 乱数法 (データ独立型)

乱数によって, 後件部のシングルTONの初期値を実数で設定する. シングルTONの初期値は観測データと強い独立関係を有する.

シングルTONの初期値配置法を観測データへの関連性の強さ τ を用いて並べると次のようになる. この強さはデータ依存型からデータ独立型への順序となっている.

総和法 τ 最大法 τ 優勢法 τ 既定法 τ クラス乱数法 τ 乱数法

すなわち, 観測データとの依存性が最も強い手法が総和法であり, 観測データとの依存性が最も弱い手法 (観測データとの独立性が最も強い手法) が乱数法である.

最後に, 学習順序について説明する. ハイパーパラメータには, 前件部のメンバシップ関数のパラメータ ($b_{ij}, \alpha_{ij}, c_{ij}, \beta_{ij}$) と後件部のシングルTON p_i の合計 $(4n+1)R$ 個のパラメータがある. 初期配置を含めて, これらのハイパーパラメータの学習順序を議論する必要がある. 学習順序としては, まずは初期配置を行い, その後前件部と後件部を学習する.

順序1: 初期配置 \rightarrow (前件部, 後件部)

順序2: 初期配置 \rightarrow 前件部 \rightarrow 後件部

順序3: 初期配置 \rightarrow 後件部 \rightarrow 前件部

順序4: 初期配置 \rightarrow 前件部 \rightarrow (前件部, 後件部)

順序5: 初期配置 \rightarrow 後件部 \rightarrow (前件部, 後件部)

順序1は, 初期配置の後で, 前件部のメンバシップ関数と後件部のシングルTONが交互に学習される. 順序2と順序3は, 一定回数のエポック数の学習の後で, 学習対象を切り替える. 順序4と順序5は, 一定回数のエポック数の学習の後で, 前件部と後件部が交互に学習される. なお, エポックとは, パラメータ学習のためにデータ集合を投入する回数をいう.

4. 数値データの結果と考察

ここでは, 数値データを用いてファジィ推論のハイパーパラメータの特性を検証する. 数値データを図2に示す. 数値データのデータ数は200個である. 2入力1出力の2群判別問題として, 乱数により各入力値を $[0, 1]$ 内で発生させ, 2群クラスの実数値を2.0(赤・○印)と3.0(青・△印)に設定した. データの生成では, まず, データの偏りをファジイルールで表現することを仮定し, 各入力 $[0, 1]$ を5領域に分割し, 5種類のメンバシップ関数を設定した. 次に, 異なるクラスの入力データが各

Table 1: Error by Learning Parameters and Initial Setting Methods

初期配置法	誤差	初期値設定				後件部学習				前件部・後件部学習				全体学習			
		0.1	0.4	0.7	1.0	0.1	0.4	0.7	1.0	0.1	0.4	0.7	1.0	0.1	0.4	0.7	1.0
総和法	平均	41.55	41.55	41.55	41.55	9.034	9.031	9.040	9.054	8.770	8.653	8.574	8.460	8.676	8.527	8.505	8.442
	分散	1.2E-32	1.2E-32	1.2E-32	1.2E-32	2.5E-12	7.8E-11	8.9E-10	5.4E-9	3.0E-5	3.9E-5	3.0E-5	1.2E-5	6.0E-6	1.0E-5	9.4E-5	9.1E-6
最大法	平均	44.51	44.51	44.51	44.51	9.033	9.031	9.040	9.058	8.644	8.458	8.592	8.456	8.505	8.432	8.419	8.422
	分散	1.1E-31	1.1E-31	1.1E-31	1.1E-31	1.9E-12	9.6E-11	1.1E-9	7.2E-9	4.3E-5	1.1E-5	8.6E-5	1.6E-5	1.1E-5	7.0E-6	1.0E-5	9.6E-6
優勢法	平均	56.29	56.29	56.25	56.25	9.035	9.031	9.039	9.055	8.978	8.742	8.499	8.583	8.721	8.615	8.472	8.439
	分散	2.4E-6	3.2E-6	2.5E-6	2.5E-6	4.2E-11	1.4E-10	1.3E-9	6.5E-9	1.4E-4	4.5E-5	1.2E-5	8.7E-5	4.6E-6	8.4E-6	8.0E-6	9.6E-6
既定法	平均	56.35	56.35	56.35	56.35	9.033	9.031	9.040	9.056	8.923	8.405	8.435	8.530	8.658	8.405	8.434	8.512
	分散	4.9E-32	4.9E-32	4.9E-32	4.9E-32	2.7E-12	9.5E-11	1.2E-9	9.7E-9	1.1E-4	5.7E-6	9.3E-6	9.5E-6	4.8E-6	5.7E-6	9.1E-6	7.1E-6
クラス 乱数法	平均	59.82	59.51	60.64	59.62	9.036	9.031	9.039	9.058	8.674	8.693	8.530	8.516	8.619	8.586	8.530	8.447
	分散	7.4E-4	4.5E-4	4.8E-4	5.7E-4	1.2E-9	6.5E-11	1.4E-9	5.5E-9	1.6E-5	3.9E-5	8.0E-6	2.8E-5	6.0E-6	1.1E-5	8.0E-6	9.0E-6
乱数法	平均	61.71	60.86	61.51	61.38	9.035	9.031	9.039	9.056	9.092	8.738	8.476	8.456	8.688	8.588	8.455	8.456
	分散	6.6E-4	2.7E-4	5.0E-4	3.6E-4	1.2E-9	1.0E-10	1.5E-9	6.5E-9	2.4E-4	4.8E-5	1.3E-5	7.0E-6	4.4E-6	7.9E-6	9.6E-6	7.0E-6

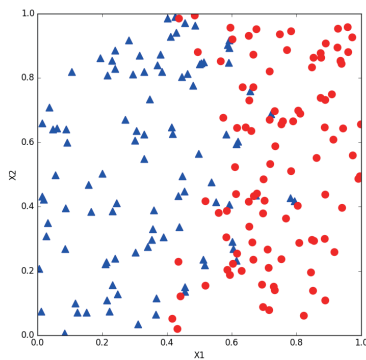
誤差: $\times 10^{-2}$ 

Fig. 2: Numerical Example Data

メンバーシップ関数の重複領域に存在するように、入力 x_1 の 5 領域の各境界にデータを 10.0% の重複率で発生させ、入力 x_2 方向では、異なるクラスを 20.0% の重複率で発生させた。

実験の設定条件として、各入力に 5 種類のメンバーシップ関数を設定し、学習係数 K_b , K_α , K_c , K_β は 0.01 として、後件部のシングルトンの学習係数 K_p は、0.1 から 1.0 まで 4 種類に設定した。ルール数は 25 個であるので、学習パラメータの総数は 225 個となる。前件部と後件部の学習には、それぞれ 300 エポック回の学習を設定した。したがって、前件部と後件部の両方を学習する場合には、合計 600 エポック回となり、さらに、前件部か後件部を再学習する場合には、合計 900 エポック回となる。実験では、1 エポックごとに乱数によりデータの並び順序を変更して、600 エポック回もしくは 900 エポック回の学習を 1 試行として、30 試行の誤差平均値と分散を計算した。

ここでは、次の 3 項目でハイパーパラメータの特性を検証する。

学習係数: 後件部のシングルトンの学習係数を 0.1 から 1.0 まで 4 種類に設定し、誤差への影響を有意性検定で検証する。

初期配置: 前件部のメンバーシップ関数は $[0, 1]$ で均等配置し、後件部のシングルトンの初期値は 6 種類の初期配置法で設定し、誤差への影響を有意性検定で検証する。

学習順序: 初期配置、前件部学習、後件部学習の順序を変更し、それぞれの過程で誤差への影響を有意性検定で検証する。

ハイパーパラメータを次の手順で検証した。

Step 1 数値データのデータ数を 100 個とする。

Step 2 前件部のメンバーシップ関数は均等配置とし、後件部のシングルトンを 6 種類の初期配置法により設定する。

Step 3 学習順序を 5 種類の設定法により設定する。ただし、ここでは、トラpezoidal フォジィ推論の学習の限界を検証するため、学習順序は「初期配置→後件部」学習を基本として、順序 1, 順序 3, 順序 5 とした。

Step 4 学習係数を 4 種類で変更し、前件部と後件部をそれぞれ 300 エポック回で学習した。ただし、1 エポックごとに乱数によりデータの並び順序を変更し、600 エポック回もしくは 900 エポック回の学習を 1 試行として、30 試行の誤差平均値と分散を計算した。なお、前件部と後件部の各学習で誤差が最小となるルールを最適ルールとしてそれぞれ保持し、全体学習の誤差平均値を計算した。式 (3) の誤差値を用いた t 検定により、ハイパーパラメータを検証した。

結果を表 1 から表 3 に示す。表 1 は、6 種類の初期配置法に対して、初期値設定、後件部、前件部・後件部を学習した際の 30 試行の誤差平均値と分散を表す。最適ルールによる全体学習の誤差平均値と分散も示す。表 2 は、初期値設定、後件部学習、前件部・後件部学習、全体学習での誤差平均値に対して、初期配置法の比較の t 検定の結果である。有意水準 0.05 に対して、有意な差が認められた場合は * 印で示した。表 3 は、順序 5 において、初期値設定と後件部学習、後件部学習と前件部・後件部学習のそれぞれの誤差平均値を比較するため、そ

Table 2: Comparison between Initial Setting Methods

初期配置法		初期値設定				後件部学習				前件部・後件部学習				全体学習			
		0.1	0.4	0.7	1.0	0.1	0.4	0.7	1.0	0.1	0.4	0.7	1.0	0.1	0.4	0.7	1.0
総和法	最大法	*	*	*	*	*	*	*	*					*	*	*	*
	優勢法	*	*	*	*	*	*	*	*					*	*	*	*
	既定法	*	*	*	*	*	*	*	*	*				*	*	*	*
	クラス乱数法	*	*	*	*	*	*	*	*					*	*	*	*
	乱数法	*	*	*	*	*	*	*	*					*	*	*	*
最大法	総和法	*	*	*	*	*	*	*	*					*	*	*	*
	優勢法	*	*	*	*	*	*	*	*	*				*	*	*	*
	既定法	*	*	*	*	*	*	*	*					*	*	*	*
	クラス乱数法	*	*	*	*	*	*	*	*					*	*	*	*
	乱数法	*	*	*	*	*	*	*	*					*	*	*	*
優勢法	総和法	*	*	*	*	*	*	*	*					*	*	*	*
	最大法	*	*	*	*	*	*	*	*	*				*	*	*	*
	既定法	*	*	*	*	*	*	*	*	*				*	*	*	*
	クラス乱数法	*	*	*	*	*	*	*	*					*	*	*	*
	乱数法	*	*	*	*	*	*	*	*					*	*	*	*
既定法	総和法	*	*	*	*	*	*	*	*					*	*	*	*
	最大法	*	*	*	*	*	*	*	*	*				*	*	*	*
	優勢法	*	*	*	*	*	*	*	*	*				*	*	*	*
	クラス乱数法	*	*	*	*	*	*	*	*	*				*	*	*	*
	乱数法	*	*	*	*	*	*	*	*	*				*	*	*	*
クラス乱数法	総和法	*	*	*	*	*	*	*	*					*	*	*	*
	最大法	*	*	*	*	*	*	*	*					*	*	*	*
	優勢法	*	*	*	*	*	*	*	*					*	*	*	*
	既定法	*	*	*	*	*	*	*	*	*				*	*	*	*
	乱数法	*	*	*	*	*	*	*	*	*				*	*	*	*
乱数法	総和法	*	*	*	*	*	*	*	*					*	*	*	*
	最大法	*	*	*	*	*	*	*	*					*	*	*	*
	優勢法	*	*	*	*	*	*	*	*					*	*	*	*
	既定法	*	*	*	*	*	*	*	*	*				*	*	*	*
	クラス乱数法	*	*	*	*	*	*	*	*	*				*	*	*	*

それぞれの学習法の比較の p 値を示し、有意水準 0.05 に対して有意な差が認められた場合に、*印を示した結果である。

まず、「学習係数」について議論する。表 1 において、後件部学習では、誤差平均値の最小値は、全ての初期配置法の学習係数が 0.4 の 0.09031 であった。その次に小さな誤差平均値は、全ての初期配置法の学習係数が 0.1 の 0.09033~0.09036 であった。誤差平均値の最大値は、全ての初期配置法の学習係数が 1.0 の 0.09054~0.09058 であった。また、分散も大きい。学習係数が 0.7 の場合も、誤差平均値は大きく分散も大きい。数値データでは、2 群クラスの実数値の差は $3.0 - 2.0 = 1.0$ であるので、これらの結果から、後件部のシングルトン学習では、学習係数は、クラス間の実数値の差の 1/2 よりも小さな値に設定することが推奨される。一方、前件部・後件部の交互学習では、誤差平均値の最小値は、学習係数が 0.7 や 1.0 の場合が多い。交互学習では、前件部のメンバシップ関数と後件部のシングルトンを交互に学習するので、後件部の学習係数は、前件部のメンバシップ関数の振る舞いを含めるため、後件部学習よりも比較的、大きな値とすることが推奨される。

次に、「初期配置」について議論する。初期配置法の比較では、表 1 の全体学習の誤差平均値の最小値は既定法の 0.08405 であったが、全ての学習係数では、最小値は最大法である。これは前件部のメンバシップ関数がトラ

ピゾイド型であるため、メンバシップ値が 1.0 となる定義域が広く、その定義域に含まれるデータの出力クラスを後件部のシングルトンの初期値として設定することによる。また、初期値設定の終了時の誤差平均値は、総和法が最小値となる 0.4155 であり、最大法が 0.4451 でそれに次いで小さい。最大値は、学習係数が 0.1 の乱数法が 0.6171 と大きく、クラス乱数法がそれに次いで大きい、優勢法と既定法はその中間に属する。したがって、誤差平均値の小さい順に初期配置法を並べると、総和法 \leq 最大法 \leq 優勢法 \leq 既定法 \leq クラス乱数法 \leq 乱数法となる。この順序は、データ依存型からデータ独立型手法の順序と一致している。また、表 3 において、すべての場合で有意性が認められたが、データ依存型の総和法や最大法の初期値設定と後件部学習の比較の p 値は極めて小さい。また、初期値設定と後件部学習の比較の p 値を後件部学習と前件部・後件部学習の比較の p 値で除算すると、総和法や最大法は、クラス乱数法や乱数法よりも遥かに小さい値を示した。また、既定法もかなり小さい値を示した。これらの結果から、データ依存型手法を用いた初期値設定は、学習に効果的に有用であることがわかる。また、初期配置法の有用性を示すと、(総和法, 最大法) \leq (優勢法, 既定法) \leq (クラス乱数法, 乱数法) となる。

最後に、「学習順序」について議論する。表 2 の後件部学習において、初期配置法を誤差平均値で比較する

Table 3: Comparison between Learning Processes

初期配置法	比較	0.1		0.4		0.7		1.0	
		p 値	有意性	p 値	有意性	p 値	有意性	p 値	有意性
総和法	初期値設定-後件部学習	1.05E-155	*	6.01E-134	*	1.26E-118	*	2.78E-107	*
	後件部学習-前後件部学習	1.15E-8	*	2.57E-9	*	2.25E-10	*	5.04E-12	*
最大法	初期値設定-後件部学習	1.55E-158	*	8.92E-134	*	3.48E-118	*	1.31E-106	*
	後件部学習-前後件部学習	2.60E-9	*	5.21E-13	*	1.48E-11	*	3.33E-12	*
優勢法	初期値設定-後件部学習	9.36E-74	*	7.17E-72	*	1.27E-73	*	5.50E-74	*
	後件部学習-前後件部学習	1.11E-8	*	1.53E-8	*	8.46E-12	*	6.91E-12	*
既定法	初期値設定-後件部学習	6.10E-160	*	1.78E-137	*	2.35E-121	*	2.51E-108	*
	後件部学習-前後件部学習	3.40E-10	*	1.37E-14	*	8.08E-12	*	3.09E-12	*
クラス乱数法	初期値設定-後件部学習	1.98E-38	*	2.00E-41	*	2.81E-41	*	4.33E-40	*
	後件部学習-前後件部学習	3.48E-10	*	5.27E-8	*	9.77E-11	*	4.53E-12	*
乱数法	初期値設定-後件部学習	1.37E-39	*	6.24E-45	*	2.20E-41	*	7.84E-44	*
	後件部学習-前後件部学習	5.83E-10	*	2.28E-9	*	3.36E-11	*	2.74E-13	*

と、有意な差が認められた。しかし、前件部・後件部学習では、有意な差が認められなかった。一方、全体学習では、初期配置法の誤差平均値に有意な差が認められた。これらの結果から、初期配置法の違いは、後件部学習には影響を与えるが、前件部学習にはあまり影響を与えず、全体の誤差平均値にも影響を与えることを示している。さらに、表3において、初期値設定と後件部学習の比較の p 値、及び、後件部学習と前後件部学習の比較の p 値を比較すると、どちらも t 検定で有意であるが、初期値設定と後件部学習の比較の p 値の方が小さい。特に、総和法や最大法のデータ依存型手法や既定法の初期値設定と後件部学習の比較の p 値が極めて小さい。これらの結果から、初期値設定から後件部学習の段階でファジールールが大きく学習されていることがわかる。すなわち、ファジールールの学習では、後件部のシングルトン学習が誤差減少へ与える影響が大きく、前件部のメンバシップ関数の学習は、むしろルールの微調整の意味が強いといえる。

5. おわりに

本論文では、学習型トラピゾイド(台形型)ファジィ推論のハイパーパラメータの特性について議論した。数値例から、学習係数と初期配置法、学習順序の特性を明らかにし、後件部のシングルトン学習が誤差減少へ与える影響が大きく、前件部のメンバシップ関数の学習はむしろルールの微調整の意味が強いことを示した。

今後、議論した特性を活かし、実際のパターン分類に 응용して、学習型トラピゾイドファジィ推論の応用可能性を確認する必要がある。

参考文献

- [1] L.A.Zadeh et al: *Fuzzy Sets, Fuzzy Logic, Fuzzy Systems*, World Scientific Press (1996).
- [2] E.H.Mamdani: Application of Fuzzy Algorithms for Control of Simple Dynamic Plant, *Proc. of Institution of Electrical Engineers*, Vol.121, No.12, pp.1585-1588 (1974).

- [3] 林:家電製品へのファジィ制御の応用状況, 電気学会論文誌 C 部門, Vol.113, No.7, pp.466-473 (1993)
- [4] 野村, 林, 若見:デルタルールによるファジィ推論の自動チューニング手法と障害物回避への応用, 日本ファジィ学会誌, Vol.4, No.2, pp.379-388 (1992)
- [5] 市橋, 渡邊:簡略ファジィ推論を用いたファジィモデルによる学習型制御, 日本ファジィ学会誌, Vol.2, No.3, pp.429-437 (1990)
- [6] 石淵, 野崎, 田中:ファジールールを用いたパターン識別, 日本ファジィ学会誌, Vol.5, No.1, pp.74-84 (1993)
- [7] 石淵, 森澤:ファジィ識別システムにおけるメンバシップ関数の調整, システム制御情報学会誌, Vol.10, No.5, pp.223-235 (1997)
- [8] 馬野, 福中, 鳩野, 田村:忘却の概念を用いたファジィ・ニューラルネットワークによるファジールールの抽出, 計測自動制御学会論文集, Vol.32, No.3, pp.409-416 (1996)

[連絡先]

林 勲 関西大学大学院 総合情報学研究所
〒569-1095 大阪府高槻市霊仙寺町 2-1-1
tel. 072-690-2448
fax. 072-690-2491
e.mail ihaya@cbii.kutc.kansai-u.ac.jp