

1 はじめに

人間の情報処理に近いシステムの開発は、高度情報化社会において非常に重要なテーマであり、従来から多くの研究が行われている。特に、複雑な振る舞いやアルゴリズム化できない問題に対して、人間の情報処理機能を模範して解決していくアプローチは非常に有効な手段と考えられている。そのような手段の代表的なものとして、ニューラルネットワークとファジィ理論が挙げられる。ニューラルネットワークは人間の脳の情報処理機能を数理モデル化したものであり、学習機能を備えている。一方、ファジィ理論は人間のあいまいな思考を取り入れた数理モデルである。なかでもファジィ推論 [1] と呼ばれるファジィ理論を用いた推論法は制御分野やパターン認識の分野において多くの成功をおさめている。ファジィ推論を実現するためにはファジィルールと呼ばれる推論ルールを獲得する必要がある。しかし、人間の勘と経験でファジィルールを獲得することは非常に困難であり多大な時間と労力が必要となる。そのため学習機能を備えたニューラルネットワークを用いて数値データからファジィルールの自動獲得を行うファジィ・ニューロシステムの研究が行われている [2][4]。しかしこれらの研究において Jang[2] はファジィルールの数をあらかじめ指定する必要があり、数値データから適応的にファジィルールを獲得できるとは言い難い。また Iyatomi らの [3] では自動的にファジィルールの数が求まるものを提案した。しかし推論したい数値データの入出力関係を十分に考慮していない手法でファジィルールの数を決定している。よって必ずしも推論したい数値データの特徴を反映したファジィルールの数になっていない。またファジィルール数を決定するための判定基準が取り扱いにくい。

本研究では推論したい数値データの入出力関係を考慮し、適応的にファジィルールを生成する自己適応型ニューラルファジィ推論ネットワーク (SANFIN : Self - Adaptive Neural Fuzzy Inference

Network) を提案する。

2 SANFIN の構造と動作

2.1 SAFIN の構造

図 1 に SANFIN の構造を示す。SANFIN の構造はファジィ推論を表すネットワーク構造となっており、入力層、ルール層、出力層の 3 層から構成される。また入力層、出力層はルール層間において全結合となっている。入力層からルール層には、ファジィルールの前件部に相当するメンバシップ関数 (図 2) によって表現される重みが配置され、ルール層から出力層には後件部に相当する重みが配置されている。ルール層おける各ノードはファジィルールを表しており、図中の実線で示された重みに接続されている j 番目のファジィルールは以下の通りである。

if X_1 is A_{1j} and ... and X_i is A_{ij} and ... and X_N is A_{Nj}
then $(Y_1$ is $V_{j1}), \dots, (Y_k$ is $V_{jk}), \dots, (Y_L$ is $V_{jL})$

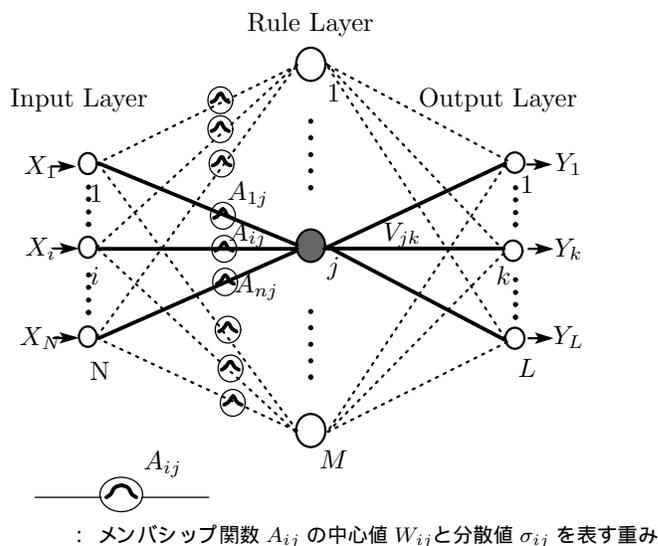
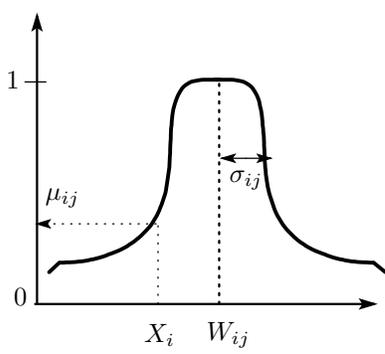


図 1 : SANFIN の構造

Self-Adaptive Neural Fuzzy Inference Network



{ 図 2 : メンバシップ関数 A_{ij} }

2.2 SANFIN の動作

ファジィ推論の演算手法がそのまま SANFIN に適用される。入力する変数の数を N , ルール層のノード数に相当するファジィルールの数を M , 出力する変数の数を L とし, SANFIN への入力 X を

$$\mathbf{X} = \{ X_1 \dots X_i \dots X_N \} \quad (1)$$

とする。

ルール層では, 入力 X が入力層とルール層間で接続された重みを介した値を受け取る。ルール層上のノード j が入力層上のノード i から受け取る値は

$$\mu_{ij} = \exp\left(-\frac{(X_i - w_{ij})^2}{\sigma_{ij}^2}\right) \quad (2)$$

となる。また, ノード j は入力層の各ノードから受け取った値の積

$$\rho_j = \prod_i \mu_{ij} \quad (3)$$

を出力層に接続されている重みを介して出力する。出力層では, ルール層上の全ノードからの値を受け取る。出力層上のノード k が接続されている重みを介して受け取る値は

$$I_k = \sum_j (V_{jk} \rho_j) \quad (4)$$

となる。またノード k の出力は (4) 式の平均をとることによって最終的な推論値 Y_k を出力する。

$$Y_k = \frac{\sum_j (V_{jk} \rho_j)}{\sum_j \rho_j} \quad (5)$$

3 SANFIN の構築手法

入力とそれに対応する望ましい出力の組の集合となっている学習データを用いてファジィルールを生

成していく。ここで, 学習データの入力にあたるデータを入力学習データ, 出力にあたるデータを出力学習データとして区別する。入力学習データに対するファジィ推論値が出力学習データとなるようにファジィルールを構築していくことが目標となる。そのようなファジィルールを生成するため本研究ではファジィルール数を決定するための教師なし学習過程とメンバシップ関数および後件部定数のパラメータ調整を行う教師あり学習過程の 2 段階で構成される。

3.1 教師なし学習過程

教師なし学習過程では学習データの集合をいくつかのクラスタに分類するクラスタリングを行う。各々のクラスタを 1 つのファジィルールとすることでルール数を決定し, メンバシップ関数の中心値 w_{ij} , 分散 σ_{ij} と後件定数 V_{jk} を仮決定する。本研究では学習データの入出力空間の特性を考慮し, クラスタが存在しない状況から逐次に提示された学習データに対して適応的にクラスタの生成または更新を行う手法を用いる。ここで p 番目 ($p = 1 \sim l$) に提示された学習データにおいて入力学習データを \mathbf{x}^p , また \mathbf{x}^p に対応する出力学習データ \mathbf{y}^p ととする。

$$\mathbf{x}^p = (x_1^p, \dots, x_N^p) \quad (6)$$

$$\mathbf{y}^p = (y_1^p, \dots, y_k^p) \quad (7)$$

本手法では学習データの入出力空間の特性を考慮するために, 入力学習データ \mathbf{x}^p と出力学習データ \mathbf{y}^p を個別にクラスタリングする。 \mathbf{x}^p , \mathbf{y}^p は, それぞれ一定の基準値を満たす $\varphi^{in}, \varphi^{out}$

$$\varphi^{in} = \frac{|\mathbf{x}^p \wedge \mathbf{W}^j|}{|\mathbf{x}^p \vee \mathbf{W}^j|} \geq \eta^{in} \quad (8)$$

$$\varphi^{out} = \frac{|\mathbf{y}^p \wedge \mathbf{V}^j|}{|\mathbf{y}^p \vee \mathbf{V}^j|} \geq \eta^{out} \quad (9)$$

を最大とするクラスタに分類させる。ここで η^{in}, η^{out} の値は 1 以下の正の定数であり, これらの値が大きいほど精度の高い分類が行われる。また $\mathbf{W}^j, \mathbf{V}^j$ は入力空間, 出力空間におけるクラスタ j の代表点である。但し, 入力学習データのクラスタリングに関しては, 出力学習データのクラスタリング状況をふまえながらクラスタの生成または更新を行うこととし, その具体的な手順を以下に示す。

(a) クラスタの初期生成

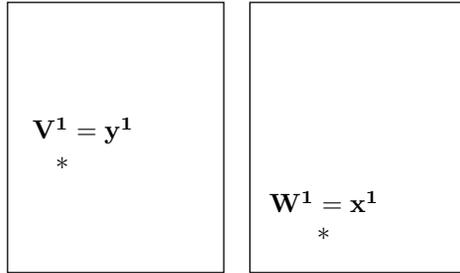
最初に提示された入出力学習データ $\mathbf{x}^1, \mathbf{y}^1$ をそ

それぞれ初期クラスタの代表点 W^1, V^1 として生成する。(図 3)

$$W^1 \leftarrow x^1 \quad (10)$$

$$V^1 \leftarrow y^1 \quad (11)$$

ここで W^1 と V^1 は入力-出力の対応関係となっている。



出力空間上のクラスタ 出力空間上のクラスタ
 { 図 3 : 入出力空間における初期クラスタの生成 }

(b) クラスタの更新

[出力学習データのクラスタリング]

p 番目に提示された出力学習データ y^p に対して式 (9) を満たし、かつその値が最大となる出力空間上のクラスタ J_{max} を検出する。

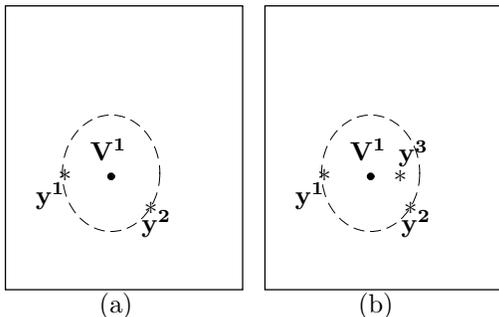
$$J_{max} = \arg \max_j \left(\frac{|y^p \wedge V^j|}{|y^p \vee V^j|} \right) \quad (12)$$

検出されたクラスタ J_{max} の代表点 $V^{J_{max}}$ は y^p を包含するような円の中心点に更新される。(図 4(a))

$$V_{max}^J \leftarrow \frac{(D_{J_{max}}^{out} - R_j^{out})(V^{J_{max}} + y^p)}{D_{J_{max}}^{out}} \quad (13)$$

但し, $D_{J_{max}}^{out}$ は $V^{J_{max}}$ と y^p とのユークリッド距離, $R_{J_{max}}^{out}$ は出力空間におけるクラスタ J_{max} の半径である。

図 4 (b) に示すように, y^p が $V^{J_{max}}$ を中心とする円で包含される位置に検出された場合はクラスタ J_{max} の更新をしない。



{ 図 4 : 出力空間上のクラスタの更新 }

[入力学習データのクラスタリング]

y^p と対応関係にある入出力学習データ x^p と, (12) 式で検出された出力空間上のクラスタの代表点 $V^{J_{max}}$ と対応関係にある入力空間上のクラスタの代表点 $W^{J_{max}}$ を式 (8) に基づいて比較する。すなわち

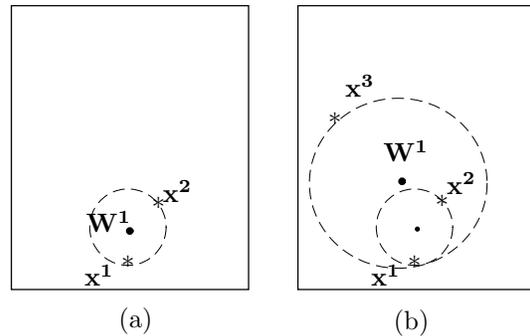
$$\frac{|x^p \wedge W^{J_{max}}|}{|x^p \vee W^{J_{max}}|} \geq \eta^{in} \quad (14)$$

を満たすならば (13) 式と同様に $W^{J_{max}}$ を更新する (図 5, (a)(b)).

$$W^{J_{max}} \leftarrow \frac{(D_{J_{max}}^{in} - R_j^{in})(W^{J_{max}} + x^p)}{D_{J_{max}}^{in}} \quad (15)$$

但し, $D_{J_{max}}^{in}$ は $W^{J_{max}}$ と x^p とのユークリッド距離, $R_{J_{max}}^{in}$ は入力空間におけるクラスタ J_{max} の半径である。

また, x^p が $W^{J_{max}}$ を中心とする円で包含される位置に検出された場合はクラスタ J_{max} の更新をしない。



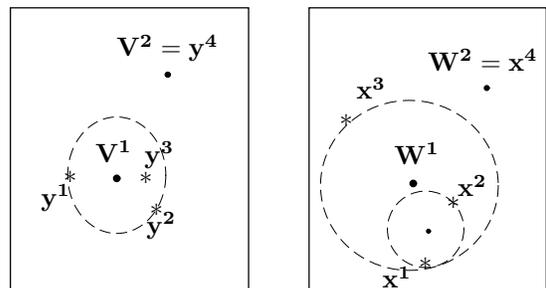
{ 図 5 : 入力空間上のクラスタの更新 }

(c) 新たなクラスタの生成

y^p に対して, (9) 式を満たすクラスタ V^j が存在しない場合, 学習データ y^p, x^p を新たなクラスタの代表点として登録する。

$$W^{new} \leftarrow x^p \quad (16)$$

$$V^{new} \leftarrow y^p \quad (17)$$



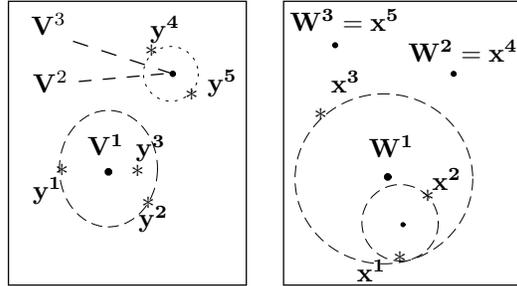
出力空間上のクラスタ 入力空間上のクラスタ
 { 図 6 : クラスタの生成 (1) }

また y^p に対して, (9) 式を満たすクラスタ V^j が存在するが, x^p に対して (14) 式を満たすクラスタの代表点 W^j が存在しない場合

$$W^{new} \leftarrow x^p \quad (18)$$

$$V^{new} \leftarrow V^{J_{max}} \quad (19)$$

とする.



出力空間上のクラスタ 入力空間上のクラスタ

{ 図 7: クラスタの生成 (2) }

全ての学習データを上記で示した (a),(b),(c) のいずれかに適用した時点で教師なし学習過程を終える. このとき生成されたクラスタ数が SANFIN のルール層におけるノード数が決定し, またクラスタの代表点 W^j, V^j が重みの初期値として採用される. 但し, $j = (1 \sim M)$ であり, M は生成されたクラスタ数とする.

3.2 教師あり学習過程

教師なし学習過程では SANFIN の重みに相当するメンバシップ関数の中心値と後見部定数が大まかに決定された. 教師あり学習過程では Back Propagation(BP) 学習則を用いて, メンバシップ関数の分散値を決定し, 同時にその中心値と後件部定数の最終的な調整を行う. BP 学習では学習データに対する SANFIN の出力誤差を最小とするように各パラメータを最適化する. ここで SANFIN の出力誤差 E を以下のように定義する.

$$E = \frac{1}{2} \sum_k^L [Y_k - y_k]^2 \quad (20)$$

Y_k はネットワークの k 番目における出力であり (図 1), y_k は教師信号である. 本研究では計算の効率性を考えて

$$J = \{ j \mid \rho_j \geq \text{Small Value} \} \quad (21)$$

として, 重みの更新式を以下に示す.

$$W_{Jk}(t+1) = W_{Jk}(t) + \alpha \left(-\frac{\partial E}{\partial W_{Jk}} \right) \quad (22)$$

$$\sigma_{Jk}(t+1) = W_{Jk}(t) + \beta \left(-\frac{\partial E}{\partial \sigma_{Jk}} \right) \quad (23)$$

$$V_{Jk}(t+1) = V_{Jk}(t) + \gamma \left(-\frac{\partial E}{\partial V_{Jk}} \right) \quad (24)$$

なお α, β, γ は学習定数である.

4 おわりに

本研究では学習データにおける入出力空間の特性を考慮しながら適応的にファジイルールを生成する SANFIN を提案した. 今後は実用的な例題を用いて [2],[3] 等の手法と比較することで SANFIN の有効性を検証する. 具体的には

- 推論値の精度
- 生成されるファジイルール数
- 大域的に意味のあるファジイルールを生成しているかどうか

の 3 つを検証基準として使用する.

参考文献

- [1] L.A. ザデー・R.R. イエーガ 共著, 浅居 喜代治 訳 (1998) 『ソフトコンピューティング』, 海文堂
- [2] J.S.R.Jang, "ANFIS: Adaptive-Network Based Fuzzy Inference System", IEEE Trans. syst., Man, Cybern., vol.23, pp.665-685, June 1993
- [3] H.Iyatomi, M.Hagiwara, "Adaptive Fuzzy Inference Neural Network", Pattern Recognition, Vol37, No10, pp2049-2057, 2004