

ニューラルネットワークの重みとファジィルールの選択

山岡 正人 向殿 政男
明治大学 理工学部

本手法は、与えられた入出力データからニューラルネットワークの学習により、推論規則に於ける前件部と後件部を同定していき、ファジィルールの選択を行うものである。その際、ニューラルネットワークの学習を通じて得た重み(ウェイト)から構築モデルに対する入力変数の相対的重要度を求め、不必要と思われる入力変数を削除していく。従って、本手法は、ニューラルネットワークを用いたファジィモデリングの一手法であるといえる。本論文では、このファジィモデリングの一手法を述べ、終わりに具体例を挙げ本手法の有効性を確かめる。

The Connection Weights after Learning Neural Network and a Selection of Fuzzy Inference Rules

Masato Yamaoka Masao Mukaidono

Department of Computer Science
Meiji University
1-1-1, Higashi-Mita, Tama-ku, Kawasaki-si
JAPAN 214

We propose a new method for fuzzy modeling with neural network. Usually in fuzzy modeling inference rules have to be determined by a discovery method. Therefore acquiring desirable rules efficiently is a very important problem to identify the fuzzy model.

It is well known that neural network has a learning function and non-linearity properties. So we give the acquired input-output data to neural network to make it learn with the error back propagation algorithm. In the built network, we reduce the unnecessary input variables by calculating the importance degree of each input variable with the connection weights between the input layer and the hidden layer.

Selecting the input variable, we also identify the antecedent part and the consequent part with neural networks.

In this paper, we show the method by which we can get the most desirable inference rules by taking advantage of properties of neural network and we also provide an example with this method.

1. はじめに

ファジィモデリングとは、あるプラントをif-thenルールの形式で記述することであるが、具体的には与えられた入出力データから、各ルールに於ける諸変数を同定することなのである。そして、一般的にこれは前件部及び後件部の構造同定とパラメータ同定を意味している。ファジィモデルの代表的なものに、菅野等^[1]が提案したモデルがある。これは、後件部を線形方程式で表現した推論法で、現在のファジィモデリングの分野に大きく貢献している。

ファジィモデリングの問題として、計算アルゴリズムが統計的手法や発見的手法に基づいているため、非常に複雑で、諸変数の決定にかなりの手間がかかることなどが挙げられる。

本論文では、与えられた入出力データからニューラルネットワークを使って推論規則に於ける前件部と後件部を同定していく手法を述べる。特に後件部の同定に際し、ニューラルネットワークの学習を通じて得た重み(シナプス結合係数)により、直接構築するモデルに関与しない入力変数を削除し、最終的に最適なファジィルールを獲得していく。そして、具体的な例を挙げ、本手法の有効性を確かめる。

2. バックプロパゲーション法

本手法では代表的なニューラルネットワークの学習法であるバックプロパゲーション法を用いる。そこで、以下にバックプロパゲーション法の概要を述べる。バックプロパゲーション法とは、Rumelhartにより提案された学習方法で、3層の階層型ネットワークを用いてこの学習法を説明する。いま、 x_k, y_j, z_i, d_i を

x_k : 入力層の第k番目ニューロンの出力値

y_j : 中間層の第j番目ニューロンの出力値

z_i : 出力層の第i番目ニューロンの出力値

d_i : 出力層の第i番目ニューロンに対する要求出力値

のように定義すると z_i と d_i との誤差が0になる状態が望ましいのである。従って以下の式が0になればよいわけである。

$$I = \sum_{i=1} (z_i - d_i)^2 \quad (1)$$

この誤差信号Iを極小化するために、中間-出力層間の結合係数 w_{ij} と入力-中間層間の結合係数 w_{jk} を以下のように変化させることが出来る。

$$\delta w_{ij} = -a \cdot (z_i - d_i) \cdot f' \left(\sum_{i \neq j} \sum w_{ij} \cdot y_j - p_i \right) \cdot y_j \quad (2)$$

$$\delta w_{jk} = -a \cdot (z_i - d_i) \cdot f' \left(\sum_{i \neq j} \sum w_{ij} \cdot y_j - p_i \right) \cdot f' \left(\sum_{j \neq k} \sum w_{jk} \cdot x_k - q_j \right) \cdot x_k \quad (3)$$

ここで、 a, p_i, q_j は

a : 学習係数

p_i : 出力層の第i番目ニューロンの閾値

q_j : 中間層の第j番目ニューロンの閾値

である。

関数 f はシグモイド関数で次式で表せる。

$$f(x) = 1 / (1 + \exp(-x)) \quad (4)$$

また、 f' は f の微分型である。式(2)と(3)から誤差の修正量が出力層から中間層へと計算されるばかりではなく中間層から入力層へまでも計算されていることが分かる。それ故にこの学習アルゴリズムはバックプロパゲーション法(誤差逆伝播法)と呼ばれている。

入力層 中間層 出力層

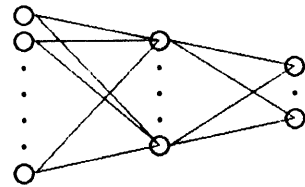


図1 3層階層型ネットワーク

3. ニューラルネットワークを用いた

入力変数選択

与えられている入出力データより、出力に関連のある入力変数を選択する事がまず初期の段階として必要になってくる。つまり、外乱となる入力変数を取り除き、出力と相互関係の大きい入力変数を選択する事に相当する。従来は、変数減少法により変数の選択を行った。これは、すべての入

力変数の組から一つを取り除き、評価指標が減少すればその入力変数の重要度は低いと考えられるので、その入力変数を捨てていく方法である。しかし、変数減少法を用いる場合もし入力変数の数が多いと、一つ一つの入力変数を調べていくのはかなりの手間がかかるといった問題が起こる。

そこで、与えられている入出力データをニューラルネットワークに与え、学習後の構築されたネットワークに於ける入力層-中間層間のシナプス結合係数を調べることで、入力変数の重要度を判断し、最終的に重要度が極めて低いものを除去していく。以下に、変数減少法の代わりとなる入力変数選択法を提案する。

いま、入力層第 j 番目ニューロンから中間層第 i 番目ニューロンのシナプス結合係数を w_{ij} とすると、 P_j を次式のように定義する。

$$P_j = \sum_{i=1}^n |w_{ij}| \quad (i=1 \dots n, j=1 \dots m) \quad (5)$$

ここで、 P_j は学習後のネットワークに於ける入力層第 j 番目ニューロンから中間層ニューロンへの結合係数の絶対値の和を表している。結合係数 w_{ij} が前層からの影響度を表していることは式(2)、(3)より分かる。ここで α を以下のように定義する。

$$\alpha_j = \frac{P_j}{MP_j} \quad (6)$$

但し $MP_j = \max_{1 \leq j \leq m} P_j$

α は入力変数の相対的重要度を表しているといえる。そこで相対的重要度が極めて低くなっている入力変数を削除することが出来る。

4. ニューラルネットワークによる前件部の同定

前件部の同定を行う前に、与えられている入出力データ (x_i, y_i) をニューラルネットワークの学習用データ(k 個)と評価用データ(l 個)とに分類する。そして、学習用データを更にクラスタリング手法を用いて r 分割する。ここで推論規則は r 個となるわけで、いま s 番目の推論規則を L^s ($s=1 \dots r$)とすると、 L^s の学習データは (x_i^s, y_i^s) ($i=1 \dots k^s$)で表すことが出来る。

但し、 k^s は各推論規則での学習用データ数を表している。

前件部の構造同定を行う。その際、3. で述べた手法を用いて既に入力変数の選択を行ったものとする。そして、今後その取り除くべき入力変数は、後件部を同定する為のニューラルネットワークにも用いないとする。 x_i を入力層のニューロンに与え、出力層のニューロンには以下の数値を割り付ける。

$$w_i^s = \begin{cases} 1 & (x_i, y_i) \in L^s \\ 0 & (x_i, y_i) \notin L^s \\ & \text{for } i=1 \dots k, s=1 \dots r \end{cases} \quad (7)$$

そしてバックプロパゲーション法により入出力用データ (x_i, y_i) の各推論規則 L^s に対する度合い \bar{w}_i^s を推定していく。いま、各推論規則 L^s の前件部のファジイ変数を A^s とする。 \bar{w}_i^s は、各推論規則 L^s に於ける前件部の適合度となる。 \bar{w}_i^s は以下のようにも表現できる。但し、 $A^s = (A_1^s \dots A_{m^s}^s)$

$$\mu_{A^s}(x_i, y_i) = \bar{w}_i^s \quad (i=1 \dots k) \quad (8)$$

ここで更に、構築されたネットワークに評価用データを与え同様に前件部の適合度を求める。^[2]

$$\mu_{A^s}(x_i, y_i) = \bar{w}_i^s \quad (i=1 \dots l) \quad (9)$$

5. ニューラルネットワークによる後件部の同定

菅野等は以下の式で表されるファジイモデルの同定を行った。

$$L^s : \text{if } x_1 \text{ is } A_1^s, x_2 \text{ is } A_2^s, \dots, x_m \text{ is } A_m^s \\ \text{then } y^s = a_0^s + a_1^s x_1 + \dots + a_m^s x_m \quad (10)$$

上式では後件部は線形方程式で表されているが、本論文では線形方程式を使用せず、ニューラルネットワークの入出力関係式(11)を用いて y^s を推定する。

$$y^s = N^s(x_1, x_2, \dots, x_m) \quad (s=1 \dots r) \quad (11)$$

後件部の同定に関しても入力変数の選択を行わなくてはならない。そこで、3. で提案した手法をここでも使用する。また、入出力データの取り扱いについては、4. のはじめに述べた方法で行うとする。以上のことより、後件部の入力変数は前件部の入力変数とは必ずしも一致するとは限らないことが分かる。以上述べた方法で推論規則の前件部及び後件部を同定した後、最終的な推定値 y^s

を以下の式から求める。

$$y^*_1 = \frac{\sum_{s=1}^r \overline{w}_1^s y_1^s}{\sum_{s=1}^r \overline{w}_1^s} \quad (12)$$

6. 具体例

本手法の有効性を確かめるために、以下に例を示す。例として(13)式で示されるような非線形システムを用いた。

$$y = (1 + x_1^{0.5} + x_2^{-1} + x_3^{-1.5})^2 \quad (13)$$

表1は(13)式によって得られた40個の入出力データを表している。表1のデータのうち左側20個を学習用データ、右側20個を評価用データとしてそれぞれ使用した。

① まず初期段階として出力に対し関係のある入力変数の選択を3. で述べた手法により求める。

表1の入力変数 x_4 は実際、式(13)には関与しておらずダミー変数と分かるが、これを構築されたニューラルネットワークにより確かめる。ネットワークの形は3層の階層型を用いる。また、入力層、中間層、出力層のニューロン数をそれぞれ4, 5, 1とした。表2に式(4), (5)を使って求められる入力変数の相対的重要度 α を示す。表2より、入力変数 x_4 を削除してよいことが分かる。従って、以後入力変数 x_4 は用いないことにする。

入力	α
x_1	1.00
x_2	0.68
x_3	0.77
x_4	0.30

表2 入力変数の相対的重要度

② 20個の学習用データをクラスタリング手法を使って分割する。分割されたデータを表3に示す。

推論規則	学習用データ番号
L^1	1, 2, 3, 4, 5, 11, 12, 13, 14, 15
L^2	6, 7, 8, 9, 10, 16, 17, 18, 19, 20

表3 分割された学習用データ

③ 学習用データ (x_i, y_i) ($i=1 \dots 20$)の推論規則に於ける前件部の適合度を求めるために4層の階層型ネットワークを用いる。学習方法は4. で述べた通りとする。また、入力層、中間層1、中間層2、出力層のニューロン数はそれぞれ3, 4, 4, 2とした。学習後の学習用データ及び評価用データの推論規則に於ける前件部の適合度を表4に示す。

④ 推論規則に於ける後件部を同定する為に、4層の階層型ネットワークを2つ用いる。入力層、中間層1、中間層2、出力層のニューロン数は共にそれぞれ3, 8, 8, 2とした。学習データを用いて学習した後、入力変数の選択をそれぞれのネットワークで行う。各々のネットワークに於ける入力変数の相対的重要度をそれぞれ表5, 表6に示す。結果的に2つのネットワークに於いて、共に入力変数 x_1 を削除できることが分かる。

入力	α
x_1	0.05
x_2	0.89
x_3	1.00

表5 L^1 に於ける入力変数の相対的重要度

入力	α
x_1	0.18
x_2	0.91
x_3	1.00

表6 L^2 に於ける入力変数の相対的重要度

⑤ 以下に同定された推論規則を示す。

L^1 : if $r = (x_2, x_3)$ is A^1
then $y^1 = N^1(x_2, x_3)$

L^2 : if $r = (x_2, x_3)$ is A^2
then $y^2 = N^2(x_2, x_3)$

⑥ ⑤の推論規則より(12)式を使って推定値 y^*_1 を求めそれを表7に示す。

⑦ 本手法を評価するため以下の式を使ってGMDH及び菅野等が提案したファジィモデルと比較を行う。その結果を表8に示す。

	学習用データ	評価用データ
GMDH	4.7	5.7
菅野行' #1	1.5	2.1
菅野行' #2	1.1	3.6
本手法	3.1	6.4

表8 本手法の性能評価

7. 考察

表8より本手法による学習用データに対する評価値はGMDHの4.7より小さいけれども菅野等のモデルに比べると大きい値となっているのでまだ改善の余地がある。一方、評価用データに対する評価値であるが、学習用データに対する評価値に比べてそれ程のひらきがないのでまずまずの結果ではないかと思う。何れにしてもニューラルネットワークの構築方法に主だったものがない為に、ネットワークの精度に多少のばらつきが生ずるのは致し方ない。従って体型立ったネットワークの構築方法の確立、更に前件部及び後件部に関するより細かなパラメータの調整なども今後の課題としたい。

8. まとめ

本論文ではニューラルネットワークを用いて、推論規則に於ける前件部及び後件部を同定する手法を述べた。特にその中で、変数減少法に代わって入力変数を選択する方法を提案した。今後、より複雑なシステムにも対応できるように更に精度の高いファジィモデルの構築に向けて努力していきたい。

9. 参考文献

- [1]菅野：「ファジィ制御」, 日刊工業新聞社 (1988)
- [2]林, 高木：「神経回路網モデルによるファジィ推論の定式化」, 第4回ファジィシステムシンポジウム講演論文集, pp.55-60, (1988)

学習用データ					評価用データ						
No.	X ₁	X ₂	X ₃	X ₄	y ₁	No.	X ₁	X ₂	X ₃	X ₄	y ₁
1	1	3	1	1	11.110	21	1	1	5	1	9.545
2	1	5	2	1	6.521	22	1	3	4	1	6.043
3	1	1	3	5	10.190	23	1	5	3	5	5.724
4	1	3	4	5	6.043	24	1	1	2	5	11.250
5	1	5	5	1	5.242	25	1	3	1	1	11.110
6	5	1	4	1	19.020	26	5	5	2	1	14.360
7	5	3	3	5	14.150	27	5	1	3	5	19.610
8	5	5	2	5	14.360	28	5	3	4	5	13.650
9	5	1	1	1	27.420	29	5	5	5	1	12.430
10	5	3	2	1	15.390	30	5	1	4	1	19.020
11	1	5	3	5	5.724	31	1	3	3	5	6.380
12	1	1	4	5	9.766	32	1	5	2	5	6.521
13	1	3	5	1	5.870	33	1	1	1	1	16.000
14	1	5	4	1	5.406	34	1	3	2	1	7.219
15	1	1	3	5	10.190	35	1	5	3	5	5.724
16	5	3	2	5	15.390	36	5	1	4	5	19.020
17	5	5	1	1	19.680	37	5	3	5	1	13.390
18	5	1	2	1	21.060	38	5	5	4	1	12.680
19	5	3	3	5	14.150	39	5	1	3	5	19.610
20	5	5	4	5	12.680	40	5	3	2	5	15.390

表1 入力カデータ

学習用データ					評価用データ						
No.	X ₁	X ₂	X ₃	\bar{w}_1^1	\bar{w}_1^2	No.	X ₁	X ₂	X ₃	\bar{w}_1^1	\bar{w}_1^2
1	1	3	1	0.984944	0.015080	21	1	1	5	0.993258	0.008746
2	1	5	2	0.991265	0.008739	22	1	3	4	0.993057	0.006948
3	1	1	3	0.989787	0.010218	23	1	5	3	0.992836	0.007188
4	1	3	4	0.993057	0.006948	24	1	1	2	0.985969	0.014035
5	1	5	5	0.994484	0.005520	25	1	3	1	0.984944	0.015080
6	5	1	4	0.013920	0.986076	26	5	5	2	0.013537	0.986459
7	5	3	3	0.013719	0.986276	27	5	1	3	0.013025	0.986970
8	5	5	2	0.013537	0.986459	28	5	3	4	0.015019	0.984976
9	5	1	1	0.012042	0.987853	29	5	5	5	0.020064	0.979932
10	5	3	2	0.012896	0.987089	30	5	1	4	0.013920	0.986076
11	1	5	3	0.992836	0.007168	31	1	3	3	0.991623	0.006382
12	1	1	4	0.991946	0.008059	32	1	5	2	0.991265	0.008739
13	1	3	5	0.993971	0.006034	33	1	1	1	0.978597	0.021405
14	1	5	4	0.993827	0.006177	34	1	3	2	0.989234	0.010771
15	1	1	3	0.989787	0.010218	35	1	5	3	0.992836	0.007168
16	5	3	2	0.012896	0.987089	36	5	1	4	0.013920	0.986076
17	5	5	1	0.012777	0.987218	37	5	3	5	0.017173	0.982823
18	5	1	2	0.012439	0.987556	38	5	5	4	0.016675	0.983321
19	5	3	3	0.013719	0.986276	39	5	1	3	0.013025	0.986970
20	5	5	4	0.016675	0.983321	40	5	3	2	0.012896	0.987089

表4 学習用データ及び評価用データの推論規則に於ける前件部適合度

学習用データ			学習用データ		
No.	y ₁	y* ₁	No.	y ₁	y* ₁
1	11.110	10.965	21	9.545	7.550
2	6.521	6.827	22	6.043	6.545
3	10.190	10.744	23	5.724	5.953
4	6.043	6.545	24	11.250	12.408
5	5.242	5.458	25	11.110	10.965
6	19.020	18.537	26	14.360	14.213
7	14.150	13.786	27	19.610	19.081
8	14.360	14.213	28	13.650	13.692
9	27.420	28.019	29	12.430	12.581
10	15.390	15.309	30	19.020	18.537
11	5.724	5.953	31	6.380	7.824
12	9.766	9.034	32	6.521	6.825
13	5.870	5.921	33	16.000	13.895
14	5.406	5.618	34	7.219	9.158
15	10.190	10.744	35	5.724	5.953
16	15.390	15.309	36	19.020	18.537
17	19.680	19.392	37	13.390	13.930
18	21.060	21.552	38	12.680	12.561
19	14.150	13.783	39	19.610	19.081
20	12.680	12.561	40	15.390	15.309

表7 同定された推論規則による推定値