Fifth IFSA World Congress (1993), 1145-1150

Generalized Fuzzy Modeling

Hee Soo Hwang, Young Hoon Joo,
Kwang Bang Woo

Department of Electrical Engineering,
Yonsei University, Seoul 120-749, Korea

Abstract

In this paper, two methods of fuzzy modeling are prsented to describe the input-output relationship effectively based on relation characteristics utilizing simplified reasoning and neuro-fuzzy reasoning. The methods of modeling by the simplified reasoning and the neuro-fuzzy reasoning are used when the input-output relation of a system is 'crisp' and 'fuzzy', respectively. The structure and the parameter identification in the modeling method by the simplified reasoning are carried out by means of FCM clustering and the proposed GA hybrid scheme, respectively. The structure and the parameter identification in the modeling method by the neuro-fuzzy reasoning are carried out by means of GA and BP algorithm, respectively. The feasibility of the proposed methods are evaluated through simulation.

I. Introduction

Fuzzy modeling is the method of describing the input-output characteristics of a system utilizing fuzzy inference rules[7,12,13]. A great deal of research on the fuzzy modeling utilizing local optimization techniques[12,13] and the learning capability of NN(neural network)[9,10] for automatic identification has been done. The methods have the following problems: (1) The determination of structure(the numbers of membership functions and rules) depends on the iterative procedure or human's experience, so it is time consuming. (2) The methods by NN's in are unable to identify the fuzzy inference rules and tune the membership functions simultaneously. (3) The boundaries between the input subspaces partitioned by the identified fuzzy rules are linear because the membership functions are designed variable by variable under the assumption of input-output variables being independent.

In this paper, two methods of fuzzy modeling are proposed to identify the fuzzy rules for describing the input-output relationship effectively based on relation characteristics of a target system utilizing simplified reasoning and neuro-fuzzy reasoning, respectively. The method of modeling by the simplified reasoning is utilized when the relation of each input to outputs of a system can be considered independent and the influences of each input on outputs can be estimated euqal. For the brevity, the case is defined as the input-output relation is 'crisp'. On the other hand, the case that the relation of each input to outputs of a system can not be considered independent, and the influences of each input on outputs can not estimated equal is defined as the input-output relation is 'fuzzy'. In the latter case, the modeling method by neuro-fuzzy reasoning is utilized. However, in the system which requires small memory capacity and fast inference time even though the input-output relation is 'fuzzy', the modeling method by the simplified reasoning is to be employed.

The schematic diagram for the identification of fuzzy rules in the proposed methods is shown in Fig. 1, in which the identification procedure is classified into the structure identification and the parameter identification. The structure identification in the modeling method by the simplified reasoning is carried out FCM(Fuzzy C-Means) clustering. Through the structure identification, the numbers of fuzzy partitions of input variables in their input spaces are identified. The number of fuzzy rules is determined by the numbers of the fuzzy partitions of input variables. The parameter identification is carried out by means of the proposed GA(Genetic Algorithm) hybrid scheme. In this case the parameters which define membership functions of the premise and linear equations of the consequence are identified. The structure identification in the modeling method by the neuro-fuzzy reasoning is carried out by means of GA in which the number of hidden layers, hidden nods in the layers, learning rate of BP(Back Propagation) algorithm, and connection pattern between outputs of nods in input layer and inputs of nods in output layer are identified. The parameter identification is carried by means of BP algorithm in which the connection weights between nods in input layer and nods in hidden and output layers, and nods in hidden layer and nods in output layer are identified.

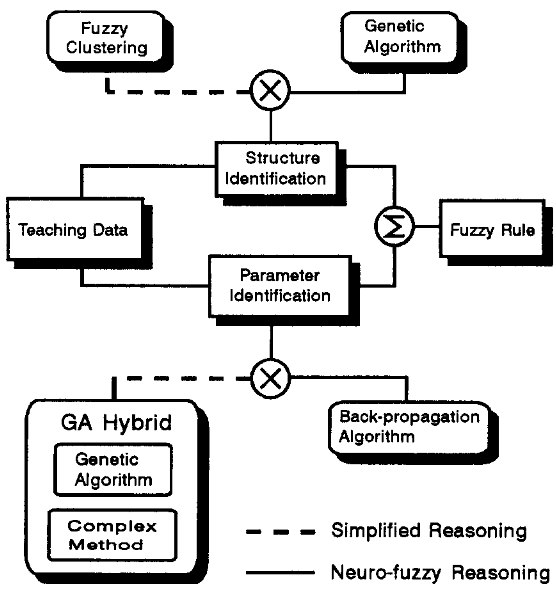


Fig. 1. Schematic diagram for the identification procedure of fuzzy model

The feasibility of the modeling methods by the simplified reasoning and the neuro-fuzzy reasoning are evaluated through modelings of gas furnace[3] and a nonlinear system[12,13], respectively.

II. Fuzzy Model by Simplified Reasoning

Fuzzy model is composed of the fuzzy rules which describe the input-output characteristics of a system. The format of fuzzy rule and reasoning, and the identification method of the fuzzy rules are described.

1. Fuzzy implication and *Reasoning*

Suppose fuzzy rules $R^i(i=1,\ 2)$ of the following format.

R^1: If x_1 is Small & x_2 is Big then $y_1=w_1a_1+b_1$

R^2: If x_1 is Big & x_2 is Medium then $y_2=w_2a_2+b_2$

where Small, Medium, and Big are fuzzy labels of x_1 and x_2, respectively, w_i the degree of fulfillment of the premise, and a_i and b_i consequence parameters.

Fig. 2 shows the procedure of reasoning, where w_1 and w_2 are calculated by eq. (1). Given input data x_1^o and x_2^o, the output y^* inferred from the above two rules is obtained in terms of the average of y_1 and y_2 with the weights w_1 and w_2.

$$w_1=\mu_{Small}(x_1^o)\mu_{Big}(x_2^o),\quad w_2=\mu_{Big}(x_1^o)\mu_{Medium}(x_2^o) \qquad (1)$$

$$y^*=\frac{w_1y_1+w_2y_2}{w_1+w_2}=\frac{w_1(w_1a_1+b_1)+w_2(w_2a_2+b_2)}{w_1+w_2} \qquad (2)$$

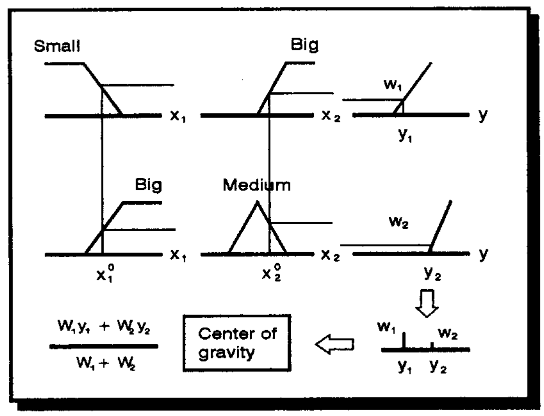


Fig. 2. Procedure of reasoning

2. Structure identification

The identification of structure which corresponds to the decision of the number of fuzzy rules and the shapes of the membership functions is carried out by means of FCM clustering[1] which produces a fuzzy c-partition of data set, via iterative optimization of the squared distances weighted by the mth power of the membership. Our purpose is to find the optimal or sub-optimal numbers of clusters to effectively describe the relation between each input-output value of a system.

For the brevity, we consider a system composed of inputs x_1 and x_2 and an output y. The procedure of the structure identification is summarized as follows:

[Step 1] The data set, $\{x_1,\ x_2,\ y\}$ is partitioned into the data sets(composed of an input and an output),

$\{x_1, y\}$ and $\{x_2, y\}$ as many as the number of input variables.

[Step 2] FCM clustering is carried out for the partitioned data sets as the number of cluster increases from 2 to n. When the clustering converges for each cluster number, fuzzy partitioning measure[2] in eq. (3) is calculated.

$$S = \frac{\sum_{i=1}^{c}\sum_{k=1}^{n}(\mu_{ik})^2 \| x_k - v_i \|^2}{n \min_{i,j} \| v_i - v_j \|^2} \tag{3}$$

[Setp 3] For the partitioned data set, the cluster numbers which have relatively lower values of fuzzy partitioning measure, S and smaller cluster numbers are selected as the numbers of appropriate partitioning of input spaces.

[Step 4] The number of fuzzy rules is determined by the multiplication of the selected cluster numbers for the partitioned data sets because the input–output relation is independent(defined as 'crisp' beforehand), so input space in fuzzy rules is partitioned by variable by variable. For example, if the cluster numbers selected for the partitioned data sets, $\{x_1, y\}$ and $\{x_2, y\}$ in step 3 are c_1 and c_2, respectively, the number of fuzzy rules is c_1c_2.

3. Parameter identification

The identification of parameters which define the membership functions of the premise and the coefficients of the linear equations in the consequence is carried out by the GA hybrid scheme.

GAs are different from conventional optimization and search procedure in four ways[8]: (1) GAs work with a coding of the parameter set, not the parameters themselves; (2) GAs search from a population of points, not a single point; (3) GAs use objective function not derivatives or other auxiliary information; (4) GAs use probabilistic transition rules to guide their search. These four differences contribute to a GA's robustness and resulting advantage over other more commonly used techniques. However, GAs are blind search: they exploit only the coding and the objective function value to determine plausible trials in the next generation. Not using the knowledge available in a particular problem puts GA at a competitive disadvantage with methods that do make use of problem–specific information. Therefore, when probelm–specific information exist it may be advantageous to combine problem–specific information with GAs to improve ultimate genetic search performance.

Genetic Algorithms

GAs are iterative adaptive general purpose search strategies based on the principles of natural population genetics and natural selection. A genetic algorithm that yields good results in many practical problems is composed of three operators: reproduction, crossover, and mutation. Reproduction is a process in which individual strings are copied according to their fitness function values which we want to maximize. After reproduction, simple crossover may proceed in two steps. First, members of the newly reproduced strings in the mating pool are mated at random. Second, each pair of strings is selected uniformly at random between 1 and string length less one, L−1. Two new strings are created by swapping all characters between positions k+1 and L inclusively. Mutation is a secondary operator whose use guarantees that the probability of searching a particular sub–region of the solution is never zero. These operators are simplicity itself, yet, the resulting search performance is wide–ranging and impressive due to implicit parallelism of GA.

In many problems, the objective is stated as the minimization of some cost function rather than the maximization of some utility or profit. In the parameter identification, our purpose is to minimize the cost function, eq. (4) which is defined as the average of squared errors between target output and inferred output.

$$E = \frac{1}{n}\sum_{i=1}^{n}(y_i^o - y_i^*)^2 \tag{4}$$

where n is the total number of data, y_i^o target output, y_i^* output inferred from fuzzy rules.

With GA, we use the following cost to fitness transformation:

Fitness function, f = 1.0/E (5)

One succesful method employed in coding multiparameters of optimization problems is the concatenated, multiparameter, mapped, and fixed point coding. To construct a multiparameter coding, we can simply concatenate many single parameter codings which define the membership functions of the premise and coefficients of the consequence. Each coding has its own sublength L, its own minimum and maximum values, P_{min} and P_{max}, respectively.

GA hybrid scheme

The GA hybrid scheme, where GA runs to substantial convergence and then local optimization procedure takes over searching from the points which display good offline performance in the evolution of GA, is proposed in this paper. In the hybrid scheme, GA finds the hills and the hill-climber goes and climbs them. As a hill-climber the complex method[11] is utilized. In the complex method to solve constrained minimization problems, a sequence of geometric figures each having $k \geq n+1$(n is the dimension of space, the number of parameters) vertices is formed to find the constrained minimum point. The method assumes that initial feasible points which satisfy all the constraints are available. In order to solve this diffculty the k points which have better offline performance in GA are employed as initial feasible points for the complex method. The sequence of GA hybrid scheme is shown in Fig. 3.

```
Generate random populations, P(0);
Evaluate fitness of individuals in P(0);
While( t is less than max generation number) {
   Generate P(t+1) from P(t) as follows:
       select the fittest individuals from P(t);
       recombine them by crossover and mutation;
   Evaluate the fitness of individuals in P(t);
   Retain the fittest individual;
   t = t + 1;
}
Select the fittest individuals as many as initial
feasible solutions for complex method;
Obtain the new feasible solutions by expansion,
reflection and contraction until the convergence
criterion is satisfied;
Find the optimal feasible solution;
```

Fig. 3. Sequence of the proposed GA hybrid scheme

III. Fuzzy Model by Neuro-Fuzzy Reasoning

The neuro-fuzzy network for the case that the input-output relation is not independent and the influences of each input on outputs are not equal(defined as 'fuzzy' beforehand) is proposed. The structure and the parameter identification of the network are described.

1. Neuro-fuzzy network

The proposed neuro-fuzzy network is shown in Fig. 4, in which the network is composed of three layers, input, hidden, and output layer. The input layer, and the hidden and output layer consist of fuzzy neurons and neural neurons, respectively. In this network, system inputs are transferred into fuzzy neurons in input layer. After all the outputs of the fuzzy neurons are combined by the connection weights between the input and hidden layer, they are inputed to the neural neurons in the hidden layer. All the outputs of neural neurons in the hidden layer are combined by the connection weights between the hidden and output layer, and are inputed to the neural neuron in the output layer. Notice that some of the outputs of the fuzzy neurons are combined by the connection weights between the input and output layer, and they are inputed to the neural neuron in the output layer. The fuzzy neurons which have the connection weights to the output layer are selected through the structure identification. The output of the neuron in the output layer corresponds to the value inferred from the neuro-fuzzy network.

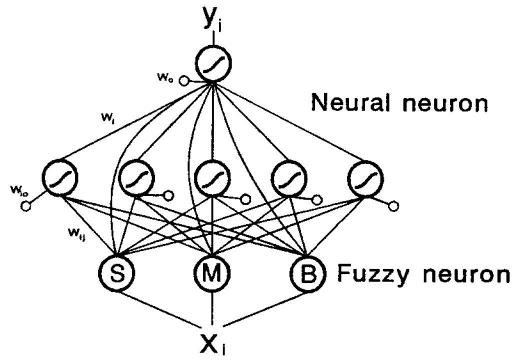


Fig. 4. Neuro-fuzzy network

The neural neuron model in the hidden and output layer of the network is shown in Fig. 5, its input-output relation is expressed by eq. (6). Fuzzy neurons in the input layer are shown in Fig. 6, and their input-output relations are described by eq. (7-9).

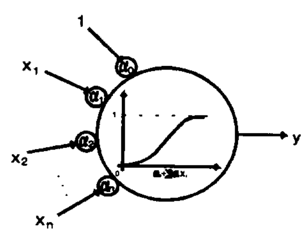
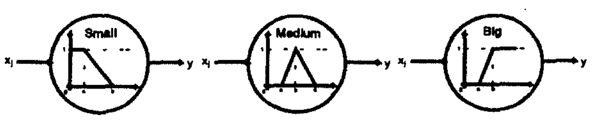


Fig. 5. Neural neuron model

$$y_i = f(\sum_{j}^{N} a_{ji}x_j + a_{oi}) \tag{6}$$

where $f(z) = 1./(1. + e^{-z})$



(a) Small (b) Medium (c) Big

Fig. 6. Fuzzy neuron model

$$\text{Small} \begin{cases} x_j <= a & y_i = 1.0 \\ x_j > b & y_i = 0.0 \\ a < x_j <= b & y_i = (x_j - a)/(b-a)+1. \end{cases} \quad (7)$$

$$\text{Medium} \begin{cases} x_j <= a & y_i = 0.0 \\ x_j > c & y_i = 0.0 \\ a < x_j <= b & y_i = (x_j - a)/(b-a) \\ b < x_j <= c & y_i = (b - x_j)/(c-b)+1. \end{cases} \quad (8)$$

$$\text{Big} \begin{cases} x_j <= a & y_i = 0.0 \\ x_j > b & y_i = 1.0 \\ a < x_j <= b & y_i = (x_j - a)/(b-a) \end{cases} \quad (9)$$

In the proposed network, the different influences of each input on output are coordinated by the connection weights between nods, and the boundaries of the input subspaces partitioned in fuzzy rules are nonlinear.

2. Identification of the network

The identification of the network is classified into the structure and the parameter identification. The schematic diagram of the identification procedure is shown in Fig. 7.

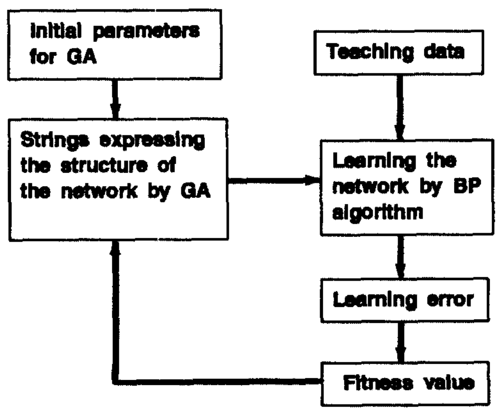


Fig. 7. Schematic diagram for the identification procedure

The format of the binary-coded strings in GA for the optimal structure of the network shown in Fig. 4 is shown in Fig. 8. The parameter identification means the identification of the connection weights between nods.

0	0010	01	0	1	1	00	01	10

Box Code

1st : No. of hidden layers
2nd: No. of neurons in hidden layers
3rd: Learning rate
4th: Connection pattern(two fuzzy neurons, M, B, have connection weights to output layer in Fig. 4.)
5th: Membership parameters, a, b, and c for Small, Medium, and Big in Fig. 6

Fig. 8. Format of strings in GA for the structure identification

The procedure for the identification of the network is summarized as follows:

[Step 1] GA generates initial population, P(0) randomnly. The strings of each individual are decoded and the neuro-fuzzy networks having the structure specified by the decoded values are generated. The networks are learned by BP algorithm until the specified iteration number is reached. Learning errors are calculated by eq. (4), and transformed into fitness value by eq. (5).

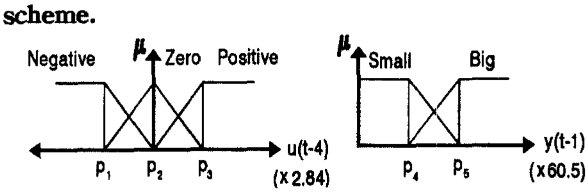
[Step 2] From P(t), P(t+1) is generated by genetic operators. The neuro-fuzzy networks defined by the individuals in P(t+1) are generated. They are learned by BP algorithm until the specified iteration number is reached. Learning errors are calculated and transformed into fitness values.

[Step 3] Retain the fittest individual. Increase t by 1 and go to step2 until t becomes the specified generation number.

IV. Simulation

1. Identification of fuzzy model

Our purpose is to verify the proposed modeling method through the identififcation of the fuzzy model which describes the relation between a gas flow u(t) and the combusted CO_2 concentration y(t) of gas furnace using 299 pairs of data[3]. We consider u(t−4) and y(t−1) as input variables of fuzzy rules, and y(t) as output variable. The numbers of clusters with smaller S and smaller c in order to minimize the number of fuzzy rules, are determined as 3 and 2 for c_1, and c_2. From the clustering results, the membership functions of input variables are defined as Positive, Zero, Negative for u(t−4), and Small and Big for y(t−1), respectively. The fuzzy rules after the structure identification are shown in Fig. 9, in which $p_1, p_2, \ldots, p_5$, and $a_1, b_1, \ldots, b_6$ are the parameters to be identified by the GA hybrid scheme.



If u(t−4) is Negative & y(t−1) is Small, then $y = a_1w_1+b_1$
If u(t−4) is Negative & y(t−1) is Big, then $y = a_2w_2+b_2$
If u(t−4) is Zero & y(t−1) is Small, then $y = a_3w_3+b_3$
If u(t−4) is Zero & y(t−1) is Big, then $y = a_4w_4+b_4$
If u(t−4) is Positive & y(t−1) is Small, then $y = a_5w_5+b_5$
If u(t−4) is Positive & y(t−1) is Big, then $y = a_6w_6+b_6$

Fig. 9. Fuzzy rules after structure identification

The parameters identified by the GA hybrid scheme are shown in Table 1, and errors for the comparison with other results in Table 2.

Table 1. Identified parameters for fuzzy rules

Premise		Consequence			
p_1	−1.635	a_1	8.062	b_1	58.555
p_2	0.622	a_2	4.848	b_2	53.240
p_3	0.94	a_3	0.407	b_3	47.192
p_4	0.79	a_4	9.765	b_4	51.564
p_5	0.968	a_5	2.056	b_5	44.148
		a_6	35.878	b_6	43.912

Table 2. Comparison of identification error with other fuzzy models

Model Name	Error	No. of Rules
Tong[4]	0.469	19
Pedrycz[5]	0.776	20
Xu[6]	0.328	25
Sugeno[7]	0.355	6
Our model	0.166	6

The reason the accuracy of our fuzzy model is superior is that the proposed GA hybrid scheme enables to identify the globally ompitimal parameters in the premise and the consequence of the fuzzy rules simultaneously.

2. Identification of neuro-fuzzy network

Our purpose is to verify the proposed modeling method through the identification of a nonlinear system by the neuro−fuzzy network using the first 20 data of 40 pairs of the input−output data[12,13]. The rest 20 data are used for the evaluation of the identified model.

The identified structure is as follows: The number of hidded layers is 1, the number of neurons in the hidden layer 20, learning rate 0.7, connection pattern 110000(fuzzy neurons, Small and Big of only x_1 have connection weights to output layer), and the membership parameters, a's and b's (defined in eq. (7) and (9)) of x_1, x_2, and x_3 are 0.2, 1.0, 0.2 ,1.0, 0.267, and 0.933, respectively.

The identification error, E_1 and the error, E_2(by eq. (10)) for the evaluation of the identified model are shown in Table 3. The errors, E_1 and E_2 are shown in Table 4 for the comparison with those of other models.

$$E = \frac{1}{20} \cdot \sum_{i=1}^{20} \frac{|y_i^o - y_i^*|}{y_i^o} \times 100(\%) \qquad (10)$$

where y_i^o is target output, and y_i^* output inferred from the learned network.

Table 3. Identification error of the learned network

Error \ Iter.	250	500	1000	2000
E_1	0.059	0.012	0.004	0.0005
E_2	1.314	1.252	1.250	1.251

Table 4. Comparison of errors with other models

Identification Method	Model Name	E_1	E_2
Equation	Linear[13]	12.7	11.1
	GMDH[13]	4.7	5.7
Fuzzy	Sugeno I[13]	1.5	2.1
	Sugeno II[13]	1.1	3.6
Fuzzy−neural	Horikawa I[10]	0.84	1.22
	Horikawa II[10]	0.73	1.28
	Horikawa III[10]	0.63	1.25
Neuro−fuzzy	Our model	0.0005	1.251

The reason that the accuracy of the identified neuro−fuzzy network is superior is that GA enables to identify the optimal structure of the network.

V. Conclusion

In this paper, the methods of modeling by the simplified reasoning and the neuro−fuzzy reasoning were presented for two cases that the input−output relation of a system is 'crisp' and is 'fuzzy'. The proposed methods are able to identify the fuzzy model of a nonlinear system automatically whether the target system characteristics is 'crisp' or 'fuzzy'.

Reference

[1] Bezdek, J. C., Pattern Recognition with Fuzzy Objective Function Algorithms, New York Plenum, 1981.
[2] Xuanli, L. X., "A Validity Measure for Fuzzy Clustering", IEEE Trans. Pattern Anal. Machine Intell., vo. PAMI-13, no. 8, pp. 841-847, 1991.
[3] Box, G. E. P., et al., Time Series Analysis, Forcasting and control. Holden Day, 1970.
[4] Tong, R. M., "The evaluation of fuzzy models derived from experimental data", Fuzzy Sets and Systems., Vol. 13, pp. 1-12, 1980.
[5] Pedrycz, W., "An identification algorithm in fuzzy relational systems", Fuzzy Sets and Systems, Vol. 13, pp. 153-167, 1984.
[6] Chen-Wei Xu and Yong-Zailu, "Fuzzy Model Identification and Self-Learning for Dynamic Systems", IEEE Trans. on Syst. Man Cybern., SMC-17, No. 4, pp. 684-689, 1987.
[7] Sugeno, M. and Yasukawa, T., "Linguistic Modeling Based on Numerical Data", IFSA' 91 Brussels, Computer, Management & System Science., pp. 264-267, 1991.
[8] Goldgerg, D. E., Genetic Algorithms in Search, Optimization, and Machine Learning, Addison Wesley, 1989.
[9] Hideyuki Takagi and Isao Hayashi, "NN-Driven Fuzzy Reasoning", in Fuzzy Models For Pattern Recognition, IEEE Press, pp. 496-512, 1992.
[10] Shin-ichi, et al., "On Fuzzy Modeling Using Fuzzy Neural Networks with the Back-Propagation Algoritjm", IEEE Trans. on Neural Networks, Vo. 3, No. 5, pp. 801-806, 1992.
[11] S. S. Rao, Optimization Theory and Application, JOHN WILEY & SONS, pp. 345-348, 1984.
[12] G. Kang and M. Sugeno, "Fuzzy Modeling," Trans. SICE, Vol. 23, No. 6, pp. 106-108, 1987
[13] M. Sugeno, Fuzzy Control, Nikkan Kogyo Shimbun-sha, Japan, 1988