

OPTIMIZATION OF FUZZY INFERENCE RULES BY USING THE GENETIC ALGORITHM AND ITS APPLICATION TO THE BOND RATING

Kangrong Tan Shozo Tokinaga
Kyushu University

(Received March 16, 1998; Final March 8, 1999)

Abstract This paper deals with the optimization of membership function in the rules of fuzzy inference system by using the Genetic Algorithm, and also shows its application to the classification of corporate bond. In the fuzzy inference system the parameters such as the weight are determined to minimize the difference between the prescribed value and the output of the system. However, the output of the system also depends on the shape of the membership function. We utilize the Genetic Algorithm to select better shape of the membership function. The method is applied to the automatic classification of corporate bond (so called bond rating) of Japanese firms. The result shows about 5% improvement of the bond rating compared to the conventional fuzzy inference system.

1. Introduction

The fuzzy inference system has been applied to the wide area of engineering such as the automatic control and the decision support system because of the better performance than the conventional methods [1]. Especially, in the business decision the information system to evaluate the management status of firms by using the financial statement plays very important role. For example, the fuzzy inference provides better classification of corporate bond than the multivariate analysis since the fuzzy system realizes the nonlinear and smooth discriminant function [2].

In the fuzzy inference, the system is usually realized as the if-then rules which include a set of membership function in the antecedent and the weight in each rule. In the learning process of fuzzy rules, the parameters of system are determined to minimize the difference between the system output and the prescribed value. The back propagation algorithm usually applied to the design of neural network is utilized to optimize the weight [3].

In cases where the membership function are symmetric triangle, the back propagation algorithm is also available [4]. However, it is expected that we have better result if we use non-symmetric membership function. In these cases, the optimization of the membership function is not straightforward, while we must determine the shape of the membership function and solve a nonlinear optimization problem.

This paper deals with the optimization of membership function in the rules of fuzzy inference system by using the Genetic Algorithm (called GA) [5]. In our system, the shape of the membership function of the fuzzy rules is represented as a string of GA, and better shape of membership function is selected by using the crossover and mutation operations. The method is applied to the automatic classification of corporate bond (so called bond rating) of Japanese firms. The input variables for fuzzy rules are calculated from the financial statement, and are transformed into the principal components so as to decrease the number

of rules. About 300 GA operations are iterated to get the optimal set of membership function. The result shows about 5% improvement of correct bond rating compared to the conventional method, and proves higher efficiency of the system.

In the following, in Section 2 we show the optimization of parameters of fuzzy inference rules by using the back propagation algorithm. In Section 3 the GA is applied to optimize the shape of membership function. In Section 4, the method of the paper is applied to the automatic bond rating of the Japanese firms, and the result shows better performance than conventional methods.

2. Fuzzy Inference

2.1. Optimization of weight

A fuzzy inference system, in general, is represented as the *if...then...* rules as follows:

If x_1 is A_{i1} and ...and x_m is A_{im} then y is w_i

where, x_j ($j = 1, 2, \dots, m$) is the input variable, A_{ij} ($i = 1, 2, \dots, n, j = 1, 2, \dots, m$) is the fuzzy set, w_i is the weight in i th rule, and n is the number of the rules.

We use the simple fuzzy inference method, the fitness of the rule is computed as in equation (1), and the result of inference obtained by using the defuzzification, in equation (2).

$$\mu_i = \prod_{j=1}^m \mu A_{ij}(x_j) \quad (1)$$

$$y = \frac{\sum_{i=1}^n \mu_i w_i}{\sum_{i=1}^n \mu_i} \quad (2)$$

Where μA_{ij} is the membership function for fuzzy set A_{ij} , and μ_i is the fitness of the rule.

As is seen from equation (1) and (2), if the membership function and the weight are not optimized, the output of the system does not give proper result. Therefore, it is necessary to determine the shape of the membership function and the weight by learning.

At first we formalize the optimization of the weight, and we assume temporally the shapes of the membership functions are fixed.

Suppose we have N sets of learning data containing input variables x_j ($j = 1 \sim m$) and the prescribed output y^p . Then we define the evaluation function as square of difference between the prescribed output y^p and the output y of the system.

$$H_p = (y - y^p)^2 / 2 \quad (3)$$

The weight w_i is updated to minimize the value in (3) for each learning data set with variables x_j^k and y^p . It is shown that the optimization is realized by using the back propagation algorithm (delta rule) which is familiar to us in the synthesis of the neural network. The delta rule is described as follows.

$$w_i(t+1) = w_i(t) - \alpha \frac{\partial H_p}{\partial w_i} \quad (4)$$

$$w_i(t+1) = w_i(t) - \alpha \frac{\mu_i^p}{\sum \mu_i^p} (y - y^p) \quad (5)$$

where t is the step of iteration of calculation, α is a constant to accelerate the convergence to the minimum.

If the membership functions are symmetric triangles, then the back propagation algorithm

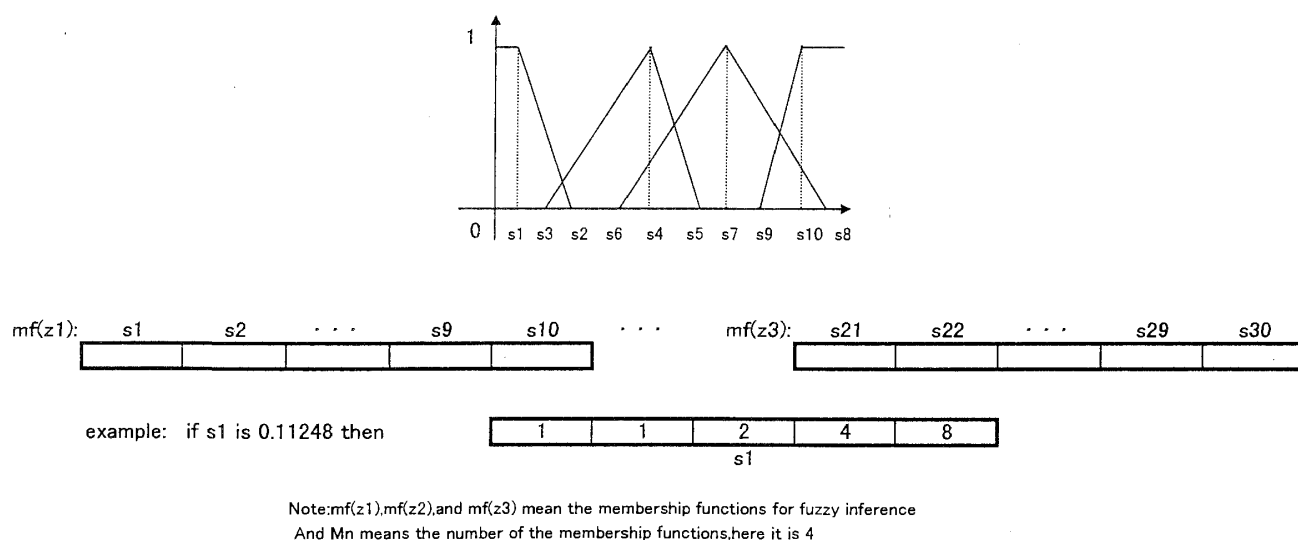


Figure 1: Membership function

is also available. The algorithm is found, for example, in Reference [4]. The algorithm is used later to compare with our optimization algorithm of membership functions where the shapes of the membership functions are triangles, but are not symmetric.

But, it is shown that the inference rules by using non-symmetric membership function give better result than the cases where the membership function are symmetric.

2.2. Discretization of output

In the optimization of fuzzy rules, the prescribed value y^p and the output of the fuzzy inference y are compared according to the equation (4). We treat mainly the classification problem by using the fuzzy inference value in equation(3). Then, in general, the prescribed value y^p is represented as the categorical value. Therefore, at first the categorical value y^p should be transformed into a real number, so that the difference $y - y^p$ in equation (4) can be evaluated.

On the other hand, the output y of the fuzzy inference is a real number, and should be transformed into a categorical value, while the output is compared to the prescribed value y^p (category) to evaluate the ability of fuzzy inference.

Even though, there are many ways to make these transformation, we classify and categorize the output by dividing the real number from zero to 1 into the same interval. Namely, if there are M categories in the prescribed value y^p , then the i th rank of y^p is transformed into a real number K_i as follows.

$$K_i = (i - 0.5)/M \tag{6}$$

On the other hand, the output y of the fuzzy inference is transformed into i th category if y exits between the upper limit U_i and the lower limit L_i .

$$L_i = (i - 1)/M$$

$$U_i = i/M \quad (7)$$

By using these transformation, the ability of fuzzy inference is evaluated as the ratio of correct classification of y . Namely, if the categorical value of y given by using equation (6),(7) is the same as the category of prescribed y^p , then the classification is correct. The ratio of correct classification is obtained by dividing the number of correct classification defined above by the total number of data for testing fuzzy inference.

2.3. Optimization of membership function

There are several kinds of shape of membership function, namely, bell, trapezoidal, and triangular. However, many works show that the ability of fuzzy inference do not clearly depend on shape of the membership function. Therefore, we assume the shape of the membership function is triangle.

As shown in Figure 1, a set of membership function (fuzzy set) in a rule is represented as a string used in GA. Namely, the range of i th triangle of membership function X_1^i, X_2^i , and the point of X_c^i where the membership function has value 1, are represented as the real number with finite length. Since the number X_1^i, X_2^i, X_c^i are less than 1, the integers under decimal point in represented as a set of 5 integers. Therefore, a string corresponding to a fuzzy rule contains $M * 4 + 1 * n$ integers. It is possible to represent the string as a set of bit pattern as in conventional GA. However, it is important to preserve the information containing the range of membership function (basis of function). Therefore we use the string which directly represents the data of membership function.

Now, we combine the optimization of membership function by using the GA with the optimization of weight of fuzzy rules by using the back propagation treated in equations (1)-(5).

Namely, at first the shape of membership function is determined by using the pool of strings of GA at the previous stage.

We have N_s sets of fuzzy inference rules corresponding to each string in the pool of GA string. Then, the weight of fuzzy rules are optimized by using the back propagation algorithm for each set of fuzzy rules. The ability of i th string in the pool is determined the rate of correct classification denoted by p_i . Then, we apply the GA operation to the pool of string.

As the procedure of GA, we use the crossover and mutation operation. Namely, for each string representing the shape of membership function we calculate the rate of classification of fuzzy inference (denoted as p_i) by using the equation (1),(2). We call p_i as the fitness of i th string. Before we apply the classification, the parameter w_i is optimized by using the back propagation algorithm by equation (4),(5). Then, we select a pair of string from the pool of string having higher p_i and then apply the crossover operation as in Figure 2. Newly generated two strings replace two strings having lower p_i in the pool. The mutation operation is applied to a string to generate a quite unique string in the pool.

The parameters included in the GA are determined as follows:

i) the generation of GA is 1000

The GA is applied to the optimization of membership function at least 1000 times.

ii)the probability of mutation is 0.15

For each string we select a random point denoting a integer, and then replace the integer by another integer(random number) at the probability 0.15.

iii)the probability of crossover is 0.8

Strings having higher fitness p_i can generate more descendants by using the so-called

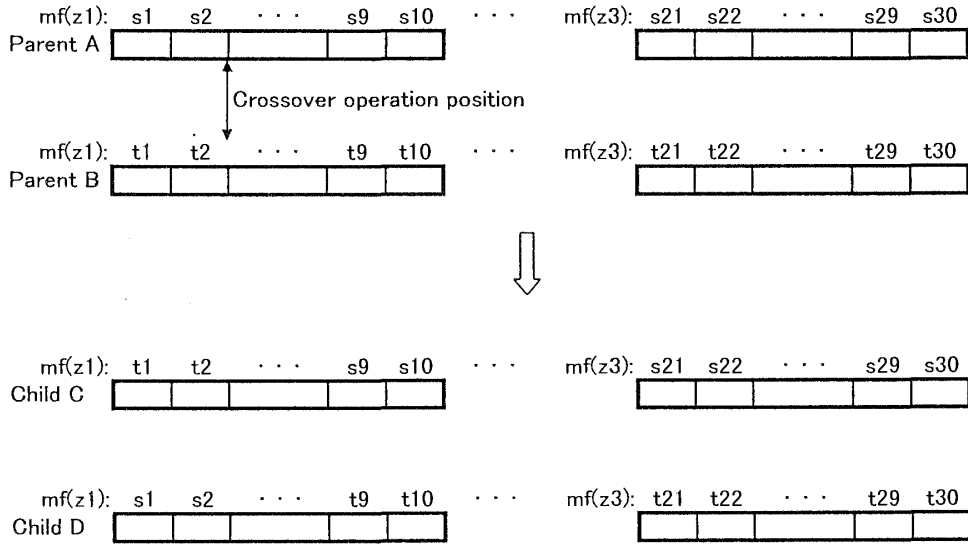


Figure 2: Crossover operation($M_n=4$)

roulette strategy. The total number of descendants N_d is 0.8 times of the number of strings N_s . We apply the crossover operation and mutation to N_d strings by replacing strings with lower p_i .

iv)the number of the individuals of one generation is 50

The number of strings stored in the pool is 50, each of which represents the whole rule of fuzzy inference.

The roulette strategy to select the candidates for the generation of descendants is explained as follows. At first, sort the strings (individuals) according to the fitness p_i , and normalize the fitness p_i in such a way that the total value of fitness is equal to 1. Then, select a random number x from uniform distribution between 0 to 1. Select i th string from the pool L if x is between the accumulated fitness $p_1 = \sum_{i-1} p_i$ and $p_2 = \sum_i p_i$. Repeat the procedure until we have N_d strings (called as L_d) from the pool. Then, we apply the crossover operation to N_d strings already selected by using the operation shown in Figure 1. Namely, from N_d strings we select two strings and apply the crossover operation in Figure 1. Then we have two newly generated strings. These two strings are used in the next stage. Repeat the same procedure to the pairs of strings remained in L_d .

On the other hand, we select i th string from the pool L if x is between p_1 and p_2 , then we directly use the string in the next step without crossover operation. In this way, we select 0.2 times of the number of strings N_s .

It is assumed that the input variable x_j is normalized as $0 \leq x_j \leq 1$. The overall procedure of optimization is summarized as follows.

(step 1)

Give symmetric shape of membership function as the initial value. Namely, the range of i th triangle of membership function X_1^i, X_2^i , and the point X_c^i are given as namely, the range of i th triangle of membership function X_1^i, X_2^i where $\varepsilon_1, \varepsilon_2, \varepsilon_c$ are the random variables generated from normal distribution and the point X_c^i are given as

$$X_1^i = (i - 2)/(I - 1) + \varepsilon_1(i \neq 1, I), X_1^i = (I - 2)/(I - 1) + \varepsilon_1(i = I) \quad (8a)$$

$$X_2^i = i/(I - 1) + \varepsilon_2(i \neq 1, I), X_2^i = 1/(I - 1) + \varepsilon_2(i = 1) \quad (8b)$$

$$X_c^i = (i - 1)/(I - 1) + \varepsilon_c(i \neq 1, I) \tag{8c}$$

and I is the number of membership function.

(step 2)

Represent the range and the point of X_1^i, X_2^i, X_c^i in a string for corresponding set of membership function. Then, these strings are summarized as a list L .

(step 3)

Optimize the weight of the rule, by using the back propagation algorithm and estimate the rate of correct classification p_i by the fuzzy inference system realized by the string.

Sort the string according to p_i , and make it as a list L . Select a pair of string from the list L , and apply the crossover operation to this pair. Then, replace two string located in the bottom of the list L by newly generated two strings. If there is no string to be applied by the crossover operation, go to next step.

(step 4)

If the difference between maximum and minimum probability p_i for each string becomes less than a sharehold value, or the total number of GA operation exceeds a constant value, then terminate the algorithm, otherwise go to step 3.

3. Application to Bond Rating

3.1. Bond rating

The purpose of bond rating is to provide information of credit risk of firms for the investors in the capital market. Generally, the information of bond rating is published by bond rating services such as Moody's Investors, Standards and Poor in the USA, and the Japan Bond Research Institute in Japan. The levels of the credit risk (bond rating) are usually classified into several categories. For example, Standard and Poor's Corporation uses 9 categories labeled from AAA as the best firm to C as the worst firm.

In the following, we focus our analysis on the bond rating of Japanese firms. Then the information of bond rating is obtained from the data of the Japan Bond Research Institute as shown in Table 1.

Table 1 Bond Rating Categories

Category	1	2	3	4	5	6	7	8	9
Grade	AAA	AA	A	BBB	BB	B	CCC	CC	C

The bond rating services publish frequently the rating depending on the financial status of firms or the fundamentals of the economy. However, the details of the evaluation of rating are not published and made by the skilled human experts.

Therefore, it may be beneficial if the fuzzy inference system imitates the human expert where the input data are obtained from the published information.

3.2. Financial ratio

Since every institute of bond rating service uses the financial statement to determine the rank of a firm, we use the financial ratio as the input variable to the fuzzy inference system [6]. The definition of the financial ratio is shown in Table 2. Even though there are many kinds of financial ratios, we use the financial ratios listed in Table 2 by comparing many works. After selecting these financial ratios as the candidates for input variables, we apply a kind of statistical analysis to preserve the balance of distribution of financial ratios. Namely, we remove several financial ratios from the candidates, and finally we obtain a set

of financial ratios used as the input variable for the inference system.

In the real bond rating by human experts, they usually use the financial ratios representing the size of firm such as the net cash flow and net worth [7]. But, in this paper we use only ordinary financial ratios so that we can test whether our method is available to any case of bond rating. Furthermore, in our method, each financial ratio calculated from the financial statement is tested to remove the outlier. Then, it may happen in the final stage we exclude the cash flow and net worth from the set of financial ratios used for bond rating.

In contrast to the conventional methods, such as the multivariate analysis, we must design the fuzzy inference system by using fewer number of input variables. Because the total number of the fuzzy inference rule is proportional to the following number.

$$\text{Number of Rules} = I^m \tag{9}$$

Where m and I are the number of input variable and the number of membership function, respectively. Therefore, we do not use directly the financial ratios as the input variables of the fuzzy inference, but the principal component of the financial ratios. Namely, we use three principal components as the input variables.

At first, we select the firms and the financial ratios to be used for learning and testing by considering the balance between the variety of category of firms and the statistical property of financial ratios. Namely, if the category of firm is restricted, the experiment of the method does not confirm the ability of the inference system. On the other hand, if the firms are selected at random from the pool of database, the convergence of the optimization may not be guaranteed. Therefore, we define two distinct categories of firms and two categories mixed with various industries. Moreover, we apply the statistical testing to preserve the statistical balance of data set.

The overall procedure to calculate the input variable is summarized as follows.

(step 1)

Select the category of industry to be analyzed. Then, calculate a set of financial ratios defined in Table 2 for each firm from the database.

(step 2)

By using the statistical analysis, we obtain the distribution of each financial ratio. If the distribution is far apart from the normal distribution, then the financial ratio is removed from the candidates of input variables for the system.

(step 3)

By using the distribution of financial ratios which are still remained as the candidate for the input variable, apply the statistical analysis to select the firm for learning and inference. Namely, if the financial ratio of a firm exists outside of 5σ of a distribution, then remove the firm from the data set for learning and inference.

(step 4)

By using the procedure from step 1 to step 3, we have a data set of input variables for the system. Then apply the principal component analysis to transform the financial ratio into three principal components. These three principal components calculated for each firm are used as the input variable for the inference system.

(step 5)

Transform these three principal components into a real number between 0 and 1.

Even though we define 5σ to exclude the outlier, but it is not necessary that the distribution of financial ratio is normal.

We usually use the logarithm of financial ratios to fit the distribution of financial ratios to the normal distribution. But in this case, these operation play the same role as the direct usage of financial ratio.

As a result, we select 7 financial ratios from 19 in Table 2 [8]. These 7 financial ratios

are transformed into three principal components, and the cumulative eigenvalue is 74%.

It may be also possible to use the principal component having higher rank correlation with bond rating rather than cumulative. But it is known that the cumulative of principal components of financial ratios becomes 70% already in third principal component. Then we use three principal component obtained from the financial ratios.

So we select 39 firms of the electronics machinery industry (called as Group 1 in the following) , 30 firms of the machine industry (Group 2), and 31 firms form the total pool of firms with various categories of industry (Group 3). The total set of firms listed above is classified as Group 4 (100 firms).

Table 2 Financial Ratio

No.	Financial Ratio	Formula
1	Net Return of Total Capital Ratio	Net Return/the Avg. of Beginning /Ending Total Capital
2	Ordinary Profit of Equity Capital Ratio	Ordinary Profit/the Avg. of Beginning /Ending Equity Capital
3	Ordinary Profit of Stated Capital	Ordinary Profit/the Avg. of Beginning /Ending Stated Capital
4	Gross Profit Ratio	Gross Profit/Total Sales
5	Operating Profit Ratio	Operating Profit/Total Sales
6	Ordinary Profit Ratio	Ordinary Profit/Total Sales
7	Profit for Current Term Ratio	Profit for Current Term/Total Sales
8	Net Worth Turnover Ratio	Total Sales/Avg. of Beginning /Ending Owned Capital
9	Fixed Assets Turnover Ratio	Total Sales/the Avg. of the Beginning /Ending Fixed Assets
10	Current Assets Ratio	Total Sales/the Avg. of beginning /Ending Current Assets
11	Receivables Turnover Ratio	Total Sales/the Avg. of Beginning /Ending Receivables
12	Current Ratio	Current Assets/Current Liability
13	Quick Ratio	Quick Assets/Current Liability
14	Receivables to Payables Ratio	Receivables/Payables
15	Fixed Asset Ratio	Fixed Asset /Owned Capital
16	liability Ratio	Liability/Owned Capital
17	Owned Capital Ratio	Owned Capital/Total Capital
18	Borrowed Capital Distributable Profit Ratio	Interest Expense/Added Value
19	Value-added Ratio	Added Value/Modified Total Sales

3.3. Discretizing the bond rating

The information of rating for the forms in Group 1, 2, 3 and 4 are obtained from the published data of Japan Bond Research Institute. Even though the range of rating varies from AAA to CCC, usually it is very rare to categorize the firm to the lower rank such as CCC or BBB. The range of rating of the experimental data is varies from AAA to BBB, and there are 4 ranks of rating.

In the fuzzy inference system, the result of the classification (output of the system) is given as a real number. Then, the real number should be transformed into a categorical

value. On the same way, the category of rating for each firm should be presented as a real number y^p so that y^p could be compared to the output y of fuzzy rules.

The discretization and categorization of these numbers are shown in Table 3.

Table 3 Discretizing the rating

Grade	Assigned real number y^p for category	Range of output y for category
AAA	0.875	0.75 ~ 1.00
AA	0.625	0.50 ~ 0.75
A	0.375	0.25 ~ 0.50
BBB	0.125	0.00 ~ 0.25

4. Result of Inference

The result of inference by the fuzzy rules depends on the number of data used for learning, and the combination of data used for learning and inference. Then, we consider two cases of experiment. In the first case (case I), the data for learning and for inference are exclusively divided, and the number of data for learning is comparable to that of the inference. In the second case (case II), the data used for learning are also utilized for inference, and the number of data for inference is about 1.5 times larger than that for learning.

4.1. Case I (using exclusive data)

In the experiment the number of membership function M_n of the fuzzy rules are selected to be 4 and 5. The result of inference is summarized in Table 4(a), 4 (b), where the value means the average rate p_i (percentage) of correct classification of bond rating to the total number of data for inference. Table 4(a) and 4(b) show the result of inference without the optimization for the membership function by using the GA. Table 4(a) and 4(b) correspond to the cases where the number of membership function is 4 and 5, respectively.

Table 5(a) and 5(b) represent the result of inference where the rules are optimized by the GA. Table 5(a) and 5(b) correspond to the cases where the number of membership function is 4 and 5, respectively.

Figure 3 depicts the convergence of the GA operation for a case of learning. As is seen from the result, by the optimization of membership function by the GA improve the proper classification by about 5%. This proves the effectiveness of the optimization.

The number of membership function has the influence on the inference, and it improves the proper classification about 2%. However, the computation for the back propagation of w_i and the optimization of the membership function increases.

It is also found that the inference for the case of single category of firm is slightly better than the case with mixed categories. However, the difference is not remarkable, and the inference system is still effective for general bond rating with arbitrary category of industry.

Table 4(a) Result of inference without GA($M_n=4$)

Group	Samples for learning	Samples for Inference	Inference Accuracy
1	20	19	68%
2	20	10	60%
3	20	11	55%
4	70	30	60%

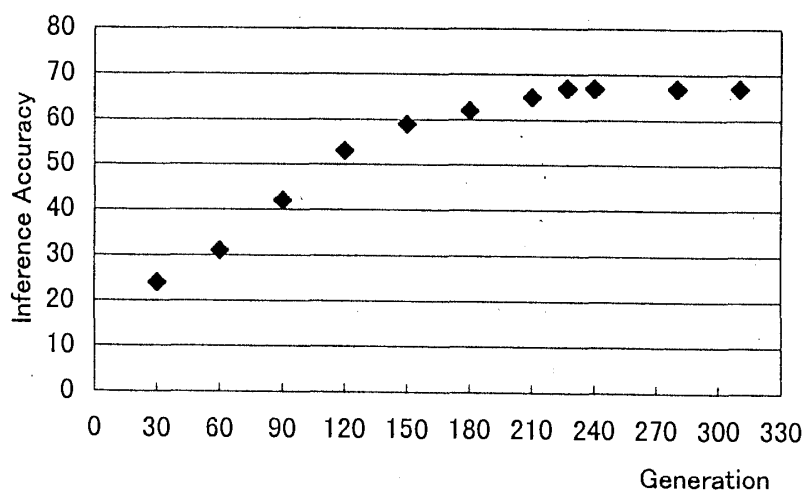


Figure 3: Generation and inference accuracy

Table 4(b) Result of inference without GA($M_n=5$)

Group	Samples for learning	Samples for Inference	Inference Accuracy
1	20	19	70%
2	20	10	60%
3	20	11	55%
4	70	30	61%

Table 5(a) Result of inference with GA($M_n=4$)

Group	Samples for learning	Samples for Inference	Inference Accuracy
1	20	19	74%
2	20	10	70%
3	20	11	64%
4	70	30	67%

Table 5(b) Result of inference with GA($M_n=5$)

Group	Samples for learning	Samples for Inference	Inference Accuracy
1	20	19	75%
2	20	10	70%
3	20	11	65%
4	70	30	69%

4.2. Case II (using partly same data)

The result of inference is summarized in Table 6(a), 6(b), Table 6(a) and 6(b) show the result of inference without the optimization for the membership function by using the GA. Table 6(a) and 6(b) correspond to the cases where the number of membership function is 4 and 5, respectively. Table 7(a) and 7(b) present the result of inference by using the optimization by the GA. Table 7(a) and 7(b) correspond to the cases where the number of membership function is 4 and 5, respectively.

As is seen from the result, by the optimization of membership function by the GA improve the proper classification by about 5%. In each case of classification, the result is better than that given in Table 6. That means if the data for learning is included or if the distribution of the data for learning and inference is close and concentrated, we have better

result of classification.

Table 6(a) Result of inference without GA($M_n=4$)

Group	Samples for learning	Samples for Inference	Inference Accuracy
1	20	39	72%
2	20	30	70%
3	20	31	61%
4	70	100	69%

Table 6(b) Result of inference without GA($M_n=5$)

Group	Samples for learning	Samples for Inference	Inference Accuracy
1	20	39	74%
2	20	30	71%
3	20	31	63%
4	70	100	72%

Table 7(a) Result of inference with GA($M_n=4$)

Group	Samples for learning	Samples for Inference	Inference Accuracy
1	20	39	79%
2	20	30	77%
3	20	31	71%
4	70	100	74%

Table 7(b) Result of inference with GA($M_n=5$)

Group	Samples for learning	Samples for Inference	Inference Accuracy
1	20	39	81%
2	20	30	78%
3	20	31	74%
4	70	100	79%

4.3. Comparison with other methods

(1) Multivariate Analysis

To compare the result with the conventional methods of classification, we apply the discriminant analysis of statistical software package [9]. The condition of comparison (the number of firms and the rank of bond rating) is the same as our method. However, it is assumed to simplify the analysis the data for learning and the data for inference is the same. Because, in general, the statistical package uses the same data set for calculating the discriminant function and the testing the ability of discriminant function. Even More, as is seen from the following comparison, the classification by the discriminant analysis is always poor compared to our method even though the data set for learning and classification are the same.

Table 8 shows the result of classification by using the multivariate analysis. As the result shows, the rate of correct classification by the multivariate analysis is worse than that of fuzzy inference.

Table 8 Result of discriminant analysis

Group	Samples	Discriminant accuracy
1	39	69%
2	30	61%
3	31	63%
4	100	71%

(2)Fuzzy Inference with symmetric membership function.

In this case, we assume that the membership functions are symmetric triangles, and defined as follows.

we assumed the membership functions are as following.

$$A_{ij}(x_j) = \begin{cases} 1 - \frac{2|x_j - a_{ij}|}{b_{ij}} & a_{ij} - \frac{b_{ij}}{2} < x_j < a_{ij} + \frac{b_{ij}}{2} \\ 0 & x_j < a_{ij} - \frac{b_{ij}}{2}, x_j > a_{ij} + \frac{b_{ij}}{2} \end{cases} \quad (10)$$

$$w_i(t+1) = w_i(t) - \alpha \frac{\partial H_p}{\partial w_i} \quad (11)$$

$$\frac{\partial H_p}{\partial w_i} = \frac{\mu_i^p}{\sum \mu_i^p} (y - y^p) \quad (12)$$

$$a_{ij}(t+1) = a_{ij}(t) - \beta \frac{\partial H_p}{\partial a_{ij}} \quad (13)$$

$$\frac{\partial H_p}{\partial a_{ij}} = \frac{\mu_i^p}{\sum \mu_i^p} (y - y^p) (w_i - y) \operatorname{sgn}(x_j - a_{ij}) \frac{2}{b_{ij} A_{ij}(x_j)} \quad (14)$$

$$b_{ij}(t+1) = b_{ij}(t) - \gamma \frac{\partial H_p}{\partial b_{ij}} \quad (15)$$

$$\frac{\partial H_p}{\partial b_{ij}} = \frac{\mu_i^p}{\sum \mu_i^p} (y - y^p) (w_i - y) \frac{1 - A_{ij}(x_j)}{A_{ij}(x_j)} \frac{1}{b_{ij}} \quad (16)$$

Especially,

$$\begin{aligned} & \text{if } z < 0 \text{ then } \operatorname{sgn}(z) = -1 \\ & \text{and if } z > 0 \text{ then } \operatorname{sgn}(z) = 1 \\ & \text{if } x_j = a_{ij} \text{ then } \operatorname{sgn}(z) = 0 \end{aligned} \quad (17)$$

where a_{ij} and b_{ij} are the support of membership function and a_{ij} is the point of the center of triangle, and α, β, γ are constants to accelerate the convergence to the minimum.

The functions include undetermined parameters, a_{ij} and b_{ij} . Learning process and the data set for learning are the same as the fuzzy inference with non-symmetric membership function. However, the shape of membership function is linear, then much simpler algorithm is available for optimization of parameters. Then, to optimize these parameters a_{ij} and b_{ij} , we can use the back propagation algorithm as follows. The details of the algorithm are found in Reference [4].

Table 9 shows the result of classification by using the fuzzy inference with symmetric membership function. As the result shows, the rate of correct classification are worse than that of the fuzzy inference by using non-symmetric membership function.

(3)Neural Network

We use the multi-layer neural network for the classification of corporate bonds. The input to the neural network are the seven financial ratios of firms (seven input node). The node of output is one, and the output level is used to classify the corporate bond. We use one middle layer having five nodes. Learning process is similar to the fuzzy inference rules. Namely, We assign the real number y^p to the category of bond rating as in Table 3. By providing financial ratios to the input layer, we calculate the output y . The difference between y and y^p is backpropagated to the input layer by optimizing the weights of network.

Table 9(a) Result of inference with Delta Rule($M_n=4$)

Group	Samples for learning	Samples for Inference	Inference Accuracy
1	20	19	63%
2	20	10	52%
3	20	11	50%
4	70	30	63%

Table 9(b) Result of inference with Delta Rule($M_n=5$)

Group	Samples for learning	Samples for Inference	Inference Accuracy
1	20	19	66%
2	20	10	60%
3	20	11	52%
4	70	30	64%

Table 9(c) Result of inference with Neural Network

Group	Samples for learning	Samples for Inference	Inference Accuracy
1	20	19	69%
2	20	10	62%
3	20	11	61%
4	70	30	67%

5. Conclusion

This paper treated the optimization of membership function in the rules of fuzzy inference system by using the Genetic Algorithm, and also shows its application to the classification of corporate bond. In the process of learning in fuzzy inference system the parameters were determined to minimize the difference between the prescribed value and the output of the system. The weight was optimized by the back propagation algorithm. On the other hand the shape of the membership function was optimized by the Genetic Algorithm. The method is applied to the bond rating of Japanese firms. The result shows about 5% improvement of the bond rating compared to the conventional fuzzy inference system.

It is also expected that by using another set of financial ratios such as cash flow and net worth we have better result of classification. Much more precise rating may be obtained by categorizing the bond whether the bond is accompanied with better mortgage or not.

As the further works, it still remains to be solved to design more general fuzzy system for business decision, or to apply multistage fuzzy inference to reduce the number of rules. It is also necessary to test whether the method of fuzzy inference system treated here is applicable to firms of various categories, and to prove the robustness of the method even if we use another set of initial values for learning.

References

- [1] M.Kanno: *Control on fuzzy* (Nikkan Shinbunsya,1989).
- [2] K.Tan and S.Tokinaga: A bond rating method based upon fuzzy inference system. *Proc. of National Conference of the Japan Society of Management and Information, Autumn* ,(1997)104-107.
- [3] D.E.Rumelhart *et al.*: Learning Internal Representation by Error Propagation, *Parallel Distributed Processing*. McClelland and the PDP Research Group(ed.),318-362(MIT Press,1986).

- [4] D.Nauck *et al.*: *Neuronale Netze und Fuzzy-Systeme*(Vieweg,1994).
- [5] D.E.Goldberg: *Genetic Algorithms in Search, Optimization and Machine Learning* (Addison Wesley,1989).
- [6] K.Tan and S.Tokinaga: Optimization of fuzzy inference rules by using genetic algorithm and its application. *Proc. of National conference of the Operation Research Society of Japan, Spring*,(1998).
- [7] Moody's Investors Service: *Global Credit Analysis*(IFR Publishing Ltd.,1991), (Ministry of Finance,1995,1996).
- [8] Ministry of Finance: *Financial Statements of Listed Company, CD-ROM Version*(Ministry of Finance,1995,1996).
- [9] SAS Institute: *The Manual of Discriminance* (SAS Institute Inc,1989).

Shozo Tokinaga
Faculty of Economics
Kyushu University
6-19-1 Hakozaki,Fukuoka,812 Japan
E-mail:tokinaga@en.kyushu-u.ac.jp