Journal of the Operations Research
Society of Japan
Vol. 45, No. 3, September 2002

PREDICTION OF CHAOTIC TIME-SERIES BY USING THE MULTI-STAGE FUZZY INFERENCE SYSTEMS AND ITS APPLICATIONS TO THE ANALYSIS OF OPERATIONAL FLEXIBILITY

Noboru Takagi Shozo Tokinaga
Kyushu University

(Received April 13, 2001; Revised February 8, 2002)

Abstract This paper deals with the prediction of chaotic time series by using the multi-stage fuzzy inference system and its application to the analysis of operating flexibility. Multi-national corporation obtained by shifting manufacturing plants located in different countries is denoted as operating flexibility. Even though the operating flexibility is optimized by the stochastic dynamic programming under the known process of exchange rate, it is usually hard to explicitly predict the change. Therefore, the value of operating flexibility depends on the prediction of time series. Then, we utilize the prediction of exchange rate by using the multi-stage fuzzy inference system. Since we divide the fuzzy inference system and related input variables into several stages, the number of rules included in multi-stage fuzzy inference systems is remarkably smaller compared to conventional fuzzy inference systems. The weight included in inference rules are optimized by the backpropagation algorithm. We also propose a method to optimize the shape of membership function and the appropriate selection of input variables based upon the genetic algorithm (GA). The method is applied to the approximation of typical multi-dimensional chaotic dynamics. The simulation study for multi-dimensional chaotic dynamics shows that the inference system gives a better prediction. The prediction is then applied to estimate the exchange rates by using input variables consisting of economic indicators as well as the exchange rates. The result shows a better performance of the multi-stage fuzzy inference system than when conventional methods are used. Then, we can find effective operating flexibility for shifting the manufacturing plants depending on the predicted exchange rate.

1. Introduction

In the field of economics of multi-national corporations, considerable efforts have been paid to investigate the firm-level advantages and foreign direct investment by inquiring about the costs of managing foreign operations and human resources. However, it is true that a predominant portion of trade stems from the transfer of goods among affiliates are within a corporation rather than investment. The coordination of a network of subsidiaries throughout the world provides so called operating flexibility. This operating flexibility is a kind of advantage by shifting the manufacturing plants depending on the exchange rate and the government policies. It can be considered as owing the option to respond to uncertain events [7][11].

Since the problem of optimizing the operating flexibility is a kind of dynamic programming, we can know the best policy if the process of uncertainty (for example exchange rates) is known [7][11]. Even though the stochastic modeling such as the Brownian motion can represent the typical behavior of exchange rate, it is not real time series. Then, the result of operating flexibility using the stochastic dynamic programming based on the normative modeling of exchange rate does not confirm the applicability. Therefore, the value of operating flexibility depends on the prediction of time series, and we must develop a stable method for time series prediction.

The fuzzy inference systems provide a better result than the multivariate analysis since the fuzzy systems estimate nonlinear functions without requiring mathematical description[12][14][20][24]. Fuzzy inference systems play the role of universal approximators utilizing linguistic information by using a restricted number of observations. They are capable of approximating any real continuous function on a compact set to arbitrary accuracy. Especially, if the system dynamics is chaotic, the system equations are deterministic, and are able to be estimated from observed time series.

Authors have been demonstrating the approximation of system equations for chaotic dynamics by using conventional single stage fuzzy inference systems [8][9][10], but it is also needed to approximate multi-dimensional chaotic systems to describe the performance of complicated phenomena which are familiar to us. For the multi-dimensional dynamics, we must approximate several system equations simultaneously by using a set of multi-variables, and usually the number of input variables including the time-delayed variables becomes relatively large.

However, in conventional fuzzy inference systems, the number of rules is proporsional to the exponential of the number of input variables, and sometime it becomes untractable and unproductive for approximating multi-dimensional chaotic dynamics.

In this paper, we use the approximation of multi-dimensional dynamics by using the multi-stage fuzzy inference system with a smaller number of rules compared to the conventional single stage system [19]. In our system, the inference system is divided into several stages, and only a part of the input variables is used as the input signal to each stage of inference, and the output of each stage is used as an input to the next stage [8][9][10][15]. Therefore, the total number of inference rules becomes remarkably smaller compared to the single stage system, and at the same time the performance of the system does not deteroirate. Even though the basic idea of the multi-stage fuzzy inference system was proposed by Raji and Zhou[15], the number of stages was only two and no algorithm for the optimization of parameters (weight) was given for general cases. In this paper, we show the algorithm to optimize the parameters by using the conventional backpropagation algorithm used in the neural network design. Moreover, we utilize also an optimization for the membership functions based on the GA.

We also utilize a method to optimize the shape of membership functions based upon the genetic algorithm (GA) [6][17]. Other than the GA in the optimization of the membership function, we use the GA to determine the inputs variables and the number of inputs. The capability of the method is proved by the approximation of typical multi-dimensional dimensional discrete and continuous chaotic dynamics.

Among the neurofuzzy systems, the neurofuzzy GMDH model proposed by Ichihashi et.al has a similar scheme to copy with the multi-stage inference [13]. But in the system output is calculated by adding the output of each stage, and the unit of the networks in the RBF functions. Then, it is not simple to compare the performance of the neurofuzzy GMDH model with our systems.

The multi-stage fuzzy inference system is applied to predict the exchange rates of various countries to decide optimal shifting of production locations in the operating flexibility. For the prediction of exchange rate, we used also several economic indicators as well as exchange rates themselves. The prediction error of the exchange rates by using the multi-stage fuzzy inference system is discussed.

Simulation studies, as well as switching costs among plants show that the operating flexibility derived by predicting the future exchange rate ensures the depression of production cost.

In the following discussion, Section 2 summarizes the basic idea of operating flexibility in multi-national corporation, and the optimization using the dynamic programming. Section 3 shows the scheme of the multi-stage fuzzy inference system. In Section 4, we show the algorithm for optimizing the weight of fuzzy inference rules by using the backpropagation, and the shape of membership functions by using the GA. We also show the GA for the selection of input variables and the number of stages in the multi-stage fuzzy inference system. In Section 5, we show simulation studies for approximating the multi-dimensional chaotic time series. Section 6 shows a simulation study of operating flexibility based on the prediction of exchange rate by using the multi-stage fuzzy inference systems.

2. Operating Flexibility

2.1. Formulation by dynamic programming

Suppose some input factors of production are priced in the world market, but other inputs such as the labor are priced in the local currency due to institutional and governmental regulations [7][11]. The wage rate in two countries are then equated not by the nominal exchange rate but by the real exchange rate.

$$w_1 = \theta w_2 \tag{1}$$

where w_1 is the wage rate in country 1, and $w_2 = w_1/\theta$ is the effective real wage in country 2. In addition, taxes, tariffs, trade barriers and transaction costs can affect the currency value of locally sourced input.

Under the assumption that the output price is set in the world market by the currency of country 1(for example, in dollars), the production decision will be based entirely on the lower cost alternative. Suppose the plant in country 1 is facing the prices P_1, and the minimum cost of producing one unit of output is given by $\phi^1 = \phi(P_1)$. Since the technologies are identical, the unit cost function of the plant in country 2 must be $\phi^2 = \phi(P_2\theta(w_1/w_2))$ where P_2 is the input price faced by the plant in country 2 and is expressed in the currency in country 1.

$$\phi(P_2) = \phi(\theta P_1) = \theta\phi(P_1) \tag{2}$$

Since the relative cost of production determines the production location, we must notice that if $\theta < 1, \phi^1 > \phi^2$, then the firm will choose to produce in conutry 2. If $\theta > 1, \phi^1 < \phi^2$, then the firm will produce in the country 2.

Without loss of generality, we normalize the costs of the plant in country 1 as one, namely, $\phi^1 = 1$. If all input prices are locally determined, then the value of costs in country 2 is

$$\phi^2(\theta) = \theta. \tag{3}$$

Assuming the above macroeconomic description, the value obtained by operating flexibility can be solved by a dynamic programming. We assume the length of economic life of production is T periods with the interval Δt, and from the beginning the firm knows the realized values of all relevant variables including the real exchange rate θ for that period.

If switching between locations is costless, then the firm will choose the location with the minimum cost over the last period under the flexible production arrangement as follows.

$$F(\theta_T) = \min[1, \phi^2(\theta_{T-1})] \tag{4}$$

Based on the dynamic programming, at any previous time t, the cost, will be the sum of costs from the optimal operation in the period beginning at time t and the minimized value function at time $t+1$.

$$F(\theta_t) = \min[1, \phi^2(\theta_t)] + \rho E_t(F(\theta_{t+1})), t = 0, 1, ..., T \tag{5}$$

where E_t is the expectation operator using the information at time t, and ρ is the one-period risk-free discount factor.

In practice, it is costly to switch between plants due to costs associated with shutdowns and startups, and labor contracting. The decision by a firm to switch production and the cost for switching multiple times over a life become also a function of the current mode of operation. We denote the cost necessary to switch from location i to j as κ_{ij}. The problem of choosing the location and the evaluation of overall costs becomes more complex than the previous one. If the firm is operating in country 1 during the period $t-1$, the cost at time t is given by

$$F(\theta_t, 1) = \min[1 + \rho E_t(F(\theta_{t+1}, 1)), -k_{12} + \phi^2(\theta_t) + \rho E_t(F(\theta_{t+1}, 2))] \tag{6}$$

where ρ is the discount rate, and $\phi(\theta_t, 1)$ and $\phi(\theta, 2)$ are the values of the flexible project at time t (when θ_t is realized), when the firm is operating in locations 1 and 2 during the period $t-1$, respectively. The first argument in equation (6) is the cost if the firm chooses to use location 1 for the period beginning at time t. The second argument gives the cost when the firm switches to use location 2 with incurred cost κ_{12}.

Similarly, the value of the project when operating in location 2 during the previous period is given by

$$F(\theta_t, 2) = \min[1 - \kappa_{21} + \rho E(F(\theta_{t+1}, 1)), \phi^2(\theta_t) + \rho E_t(F(\theta_{t+1}, 2))]. \tag{7}$$

In general, if there is a set $L_s = (1, 2, ..., L)$ possible production locations with associated cost functions ϕ^m in mth country, the valuation equations can be written as

$$F(\theta_t, l) = \min(-\kappa_{l,m} + \phi^m(\theta_t) + \rho E(F(\theta_{t+1}, m))), l, m, \in L_s. \tag{8}$$

If the number of exchange rate processes becomes large, the model becomes intractable. However, it should be remembered that most currencies are pegged into the dollar, so the model can be reduced to considering two exchange rates. When we use the characterization of uncertainty of exchange rate such as the Brownian motion, the expectation must be computed numerically. Then we would introduce a kind of discretization for the process of exchange rate.

Suppose at any time t, θ can only take one of M discrete values, $\theta^1\theta^2, ..., \theta^M$ (for example limited in the range between 0.5 and 1.5). We would then know the transition probability for $\theta_{t+1} = \theta^j$, and denote it as p_{ij}. In the discrete state-space we can rewrite the equation for optimal cost as follows.

$$F(\theta_t = \theta^j, l) = \min[-\kappa_{lm} + \phi^m(\theta^j) + \rho \sum_{i=1}^{M} p_{ij} F(\theta_{t+1} = \theta^i, m)] \tag{9}$$

$$\Delta\theta_t = \lambda(\hat{\theta} - \theta_t) + \sigma\theta_t \Delta Z \tag{10}$$

where the parameters $\hat{\theta}$, λ and σ are appropriately selected to fit the models to real exchange rates.

From simulation studies, we see that the cost obtained by optimizing the operating flexibility by the dynamic programming is less than the cost necessary for the production without switching. Especially, the initial state of the exchange rate is close to one, and the relative advantage of operating flexibility is high.

2.1.1. Prediction of exchange rate

In the operating flexibility evaluated by the dynamic programming, it is postulated that the process of exchange rate is known before the firm begins to decide the location of the production period T. Of course, we can formalize the process by a stochastic process by introducing the transition probability for the state of θ, but real exchange rate does not necessarily follow the Brownian motion.

Therefore, we must also consider the cases where we can only know the predicted value of the exchange rate, and must decide the location based on the prediction of exchange rate. In these situations, the optimization using the dynamic programming is not more applicable, while the process of exchange rate is not given to the firm. We use a more accurate prediction of exchange rate, so that the costs can be more depressed than operations without switching the locations.

Originally, we only need to use the one-step ahead prediction to estimate the next exchange rate at time $t+1$ so we must decide what are the optimal exchange rates and which productions should be shifted. A firm may decline to switch locations if the possibility of a reversal in the relative cost advantage due to subsequent exchange rate movements is high. The future real exchange rates affect the current choice of technology.

In real applications, we must carefully select the model to fit the process of exchange rate. However, it is not ensured that the exchange rate follows the model such as the stochastic process, and the DP procedure is not more applicable to evaluate the operating flexibility.

Therefore, we treat the operating flexibility as a forward decision making based on the prediction of exchange rate. At first, we predict the exchange rate, and then compare the feasible operations to decide whether to change the site of production by considering the future exchange rate and the switching costs. The procedure of operating flexibility is summarized as follows.

(Step 1)prediction of $\phi^i(t)$

By using the precdition method we estimate the future value of exchange rate.

(Step 2)decision of operation

We choose the site of production by using the decision scheme.

(Step 3) evaluation of operating flexibility

At the final stage, we compare the cost (return) of production obtained by switching the production with the cost of production without switching the site.

2.2. Exchange rate as a chaotic time series

One of the most surprising aspects of experiences with floating exchange rates since 1973 has been the unpredictability and volatility of the exchange markets. Prior to floating, exchange rates were expected to move with changes in interest rate diferentials, relative money supplies, net incomes, the current account or other factors.

There were many cases where applying existing models of systematic change rate behavior could not outperform a simple random walk model in a variety of out-of-sample tests. Finding that the exchange rates appear to follow a random walk is unsatisfactory to explain any real existing movements. Recently, interest has been shown in the possibility that nonlinearities account for the apparent unpredictability of exchange rates. For exmaple, autoregressive conditional heteroscedastic (ARCH) effects are well demonstrated.

Economists have also been encouraged to use nonlinear modeling techniques by the recent development of the chaos theory[1][2][4][18]. The application of the chaos theory to an economic model is new and rapidly expanding. Chaos theory demonstrates that even simple nonlinear deterministic dynamic systems can produce very complex behaviors over time.

In the nonlinear dynamic modeling in the growth theory, such as the overlapping generation models, there have been many fruitful results in formalizing the chaotic behavior [1].

In a variety of Keynesian IS/IM models, it is not too difficult to generate chaotic dynamics for some parameterization [3][18]. In macroeconomics, authors developed a model of deterministic chaos in the foreign exchange market[4]. Only a few authors have utilized techniques developed in the sciences to examine economic data for the presence of chaos. These authors have met with varying degrees of success, and notable results are available in US macroeconomic data, Divisia monetary aggregates, silver and gold returns, and the stock market data.

If the exchange rate is assumed to be generated by a chaotic dynamical system, we would expect it to be unpredictable and subject to sudden shifts. The market would appear to be overacting to unimportant shocks. These are features of the behaviors of exchange rates and other economic variables.

The appeal of chaos theory is that the type of complex dynamic behavior apparent in actual time series is intrinsic to the model and does not require the introduction of exogenous shocks. The possibility that economic time series are chaotic has severe implications for economic modeling and forecasting.

3. Prediction by Using Multi-stage Fuzzy Inference Systems

3.1. Multi-stage fuzzy inference systems

The multi-stage fuzzy inference system treated here is represented by N stages of rule sets each of which includes $n_i (i = 1, 2, ..., N)$ if-then type rules[8][9][10][15]. The overview of the system is shown in Figure.1. In Figure.1 if the input variables $x_{b1}, ..., x_{e1}$ are imposed to the first stage as the input variables, we will have the output y_1 for the first stage. In the second stage, we use the input variables $x_{b2}, ..., x_{e2}$, as well as y_1 to form the first stage, and we obtain the output y_2. The same procedure for the fuzzy inference rules is applied until stage N, and we have the output y_N as the output of the whole system.

If x_1 is A^1_{11} and ... and x_M is A^1_{1M}
then y_1 is w^1_1
If x_1 is A^2_{11} and ... and x_M is A^2_{1M}
then y_1 is w^2_1
......
If x_1 is $A^{n_1}_{11}$ and ... and x_M is $A^{n_1}_{1M}$
then y_1 is $w^{n_1}_1$
If x_1 is A^1_{21} and ... and x_M is A^1_{2M} and y_1 is B^1_2
then y_2 is w^1_2
If x_1 is A^2_{21} and ... and x_M is A^2_{2M} and y_1 is B^2_2
then y_2 is w^2_2
......
If x_1 is $A^{n_2}_{21}$ and ... and x_M is $A^{n_2}_{2M}$ and y_1 is $B^{n_2}_2$
then y_2 is $w^{N_2}_2$

If x_1 is A^1_{N1} and ... and x_M is A^1_{NM} and y_{N-1} is B^1_N
then y_n is w^1_n
If x_1 is A^2_{N1} and ... and x_M is A^2_{NM} and y_{N-1} is B^2_N
then y_N is w^2_N
......
If x_1 is $A^{n_N}_{N1}$ and ... and x_M is $A^{n_N}_{NM}$ and y_{N-1} is $B^{n_N}_N$
then y_N is $w^{n_N}_N$ (11)

where x_j $(j = 1,2,,M)$ are the input variables, A^k_{ij} $(i = 1,2,...,N, j = 1,2,...,M)$are the fuzzy sets assigned to each rule and input variable. The real number w^k_i is the weight used for each inference rule. The output variables $y_i(i = 1,2,...,N-1)$ are the intermediate variables representing the inference of a stage and are also used as the input variable to the next stage. The output y_N yielded by the whole set of rules is obtained at the final stage.

Table 1 shows a comparison of the number of rules included in single-stage and multi-stage fuzzy inference systems where the number of membership functions is assumed to be four, and the stage of the multi-stage fuzzy inference system is chosen to be an appropriate number. In the single stage inference system, the rule is composed in ordinary ways, such as If x_1 is A_{i1} and ... and x_M is A_{iM} then y is w_i for the i−th rule. Then, the number of rules included in the inference system given by M^{M_B} where M_B is the number of membership functions.

As is seen from the result, the number of rules in the single-stage fuzzy inference system becomes very large and untractable if the number of input variables is greater than five, but its number is still even relatively small in the multi-stage fuzzy inference system.

Table 1: Number of rules in single-stage and multi-stage systems.

system	M=3	M=5	M=7
single-stage	64	1024	16384
multi-stage	36	96	192

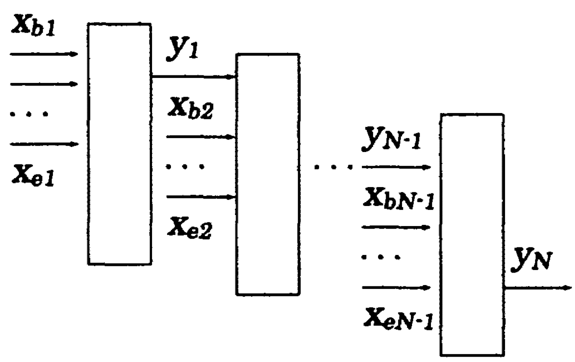

Figure 1: Overview of the inference system

3.2. Defuzzification

We assume that we use an inference where the output of the fuzzy inference is defuzzified (defuzzification) as a real number by using following equations (12) and (13)[12]. Then, the output of the inference in ith stage of inference is represented as

$$\mu^k_i = \prod_{j=1}^{M^*} \mu A^k_{ij}(x_j) \tag{12}$$

$$y_i = \sum_{k=1}^{N_i} \mu_i^k w_i^k / \sum_{k=1}^{N_i} \mu_i^k \tag{13}$$

where μA_{ij}^k is the membership function of the fuzzy set A_{ij}^k, μ_i^k is the fitness of the ith rules. The number M^* is the number of inputs to the ith stage, and its maximum value is $M+1$, while the output of the former stage is used as the input to the next stage. Usually, when only a part of the input variable to the system is used as the input to each stage, then M^* is sufficiently smaller than $M+1$. The optimal value for M^* can be determined by using the GA described later.

By summing up μ_i^k by multiplying the weight w_i^k, it yields the output of the ith stage. The variable y_{i-1} is used as the input to ith stage, and we assign the membership function $\mu B_i^k(y_{i-1})$ to be used in place of $\mu A_{ij}^k(x_j)$.

In the following discussion, the weight w_{ij} and the shapes of membership functions are optimized to improve the result of inference. The weight w_{ij} is optimized by using the backpropagation algorithm developed for the design of neural networks. The shapes of membership functions are optimized by the GA [2][6]. These two optimization techniques are used alternatively. Namely, at first the weight w_{ij} is optimized by assuming that the shapes of membership functions are fixed. Then, the shapes of membership functions are optimized by assuming that the weight w_{ij} is fixed.

It is assumed that the shapes of membership functions used in stages are not necessarily the same. These shapes are changed so that the inference system gives the best prediction.

3.3. Optimization of weight

In the successive optimization of the inference system, the shapes of the membership function are regulated, and the weights are tuned by learning.

The shape of the membership function is optimized by using the GA. Therefore, we assume temporarily that the number and the shape of the membership function are fixed. Then, the weight w_i^k is the only parameters to be optimized.

The evaluation function to be minimized is the difference between the prescribed value y^m and the output $y(=y_N)$ for mth learning data where x_j^m is used for the input.

$$H_m = (y - y^m)^2/2 \tag{14}$$

The weight is updated so that the value in (14) for each learning data set with variables is ${x_j}^m$ and y^m. The optimization is realized by using the backpropagation algorithm (delta rule) which is familiar to us in the synthesis of the neural network [17].

In the case of the multi-stage fuzzy inference system, the optimization formula is relatively complicated, but a similar backpropagation algorithm as the single stage system is able to be derived. The framework for optimizing the parameters by using the backpropagation algorithm is given as follows.

$$\Delta w_{i-2}^k(t) = -\alpha\delta_{i-1}\frac{\mu_{i-2}^k}{\sum_k \mu_{i-2}^k} + \eta\delta w_{i-2}^k(t-1) \tag{15}$$

$$\delta_{i-1} = \sum_k \delta_i \frac{\sum_k w_{i-1}^k - \sum_k w_{i-1}^k \mu_{i-1}^k}{(\sum_k \mu_{i-1}^k)^2} \times \frac{\partial \prod_{j=1}^{m^*} \mu A_{i-1j}^k(q_{i-2}^j)}{\partial q_{i-2}^k} \tag{16}$$

where q_i^k is the kth input to the ith stage, and α and η are the parameters to accelerate the convergence.
The proof of the algorithm is shown in references [9][10].

4. Optimization of Shape of Membership Functions

4.1. Optimization by using the GA

For the fuzzy inference system, it is possible to select the shapes of the membership functions as trapezoidal, triangular, and even in the form of normal distribution. However, in all cases the shapes of the membership functions can be described by using several parameters. Therefore, we restricted ourselves to the cases where the membership functions are triangular, but the method is easily extended to other cases.

Assuming that the input variable x_j is normalized as $0 \leq x_j \leq 1$, Figure 2 shows an ordinary structure of a set of membership functions (fuzzy set) in a rule which is represented as a string used in GA [6][21].

The i-th triangular membership function can be defined in triplet (X_1^i, X_c^i, X_2^i), where X_1^i and X_2^i are the lower and upper bound of the subject fuzzy number, with X_c^i being the argument that achieves the maximum membership grade. Since the numbers X_1^i,X_2^i,X_c^i are less than 1, the integers under decimal point are represented as a set of 5 integers. A string corresponding to a fuzzy rule contains $\sum_{i=1}^{N}[15 \times (M_B - 2) + 20] \times M_B^{M_i}$ integers where M_B is the number of patterns for membership function and M_i is the number of input variables to the ith stage.

Such triplet (X_1^i, X_c^i, X_2^i) for each membership function in a rule (for a set of input variables) is encoded into a series of numbers such as $X_1^1, X_c^1, X_2^1, X_1^2, X_c^2, X_2^2, \ldots$. The series of numbers is called an individual in the GA.

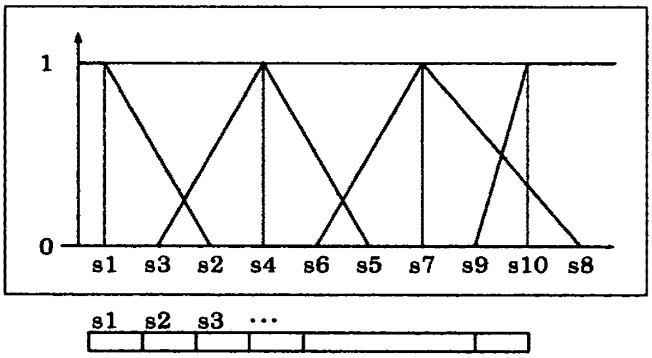

Figure 2: Representation of membership function in strings

In the following we show how to combine the optimization of the membership function by using the GA with the optimization of the weight of fuzzy rules by using the backpropagation algorithm treated in equations (15), (16).

The GA is a kind of numerical optimization method utilizing the genetic operations among the individuals representing the solution of the problem. Usually, the individuals are represented as a string of decimal or binary numbers. The optimization is carried out by the crossover operations on the individuals. For example, if the individuals A and B are relatively close to the optimal solution, then we divide A and B in an arbitrary position and generate two new individuals by exchanging two parts of the individuals A and B (having offsprings). The details of the algorithm applied in this case is shown in Secton 4.2.

Namely, at first the shape of the membership function is determined by using the pool of strings of GA at the previous stage, then the backpropagation is applied to optimize the weight.

The error of approximation is defined as

$$rmse = \sum_{k=1}^{K}[\frac{1}{T}\sum_{t=1}^{T}(y_k(t) - \hat{y}_k(t))^2]^{\frac{1}{2}}/\sigma_k \tag{17}$$

where K is the dimension of system, and $y_k(t)$ and $\tilde{y}_k(t)$ are observed and predicted values for the time series in the k-th dimension. σ_k is the variance of the time series in k-th dimension. The $rmse$ is defined for each individual, and then the ability p_k that corresponds to k-th individual is defined by the inverse of $rmse$.

4.2. Overall algorithm

The overall algorithm of learning including the backpropagation is summarized as follows. The algorithm includes two processes of learning, which are denoted as Learning-I for the optimization of weight for fuzzy inference rules, and Learning-II for the iteration of the GA applied for optimizing the shape of the membership functions. Learning-I is included inside Learning-II. The population size of Learning-II is 50.

(Step 1)

Give initial shape of membership function, namely, triplet (X_1^i, X_c^i, X_2^i) are given at random. For example,

$$X_1^1 = 0, X_c^1 = d + \varepsilon_c, X_2^1 = 2d + \varepsilon_2 \tag{18}$$

$$X_1^I = 1 - 2d + \varepsilon_1, X_c^I = 1 - d + \varepsilon_c, X_2^I = 1 + \varepsilon_2 \tag{19}$$

$$X_1^i = (i-1)d + \varepsilon_1, X_c^i = id + \varepsilon_c, X_2^i = (i+1)d + \varepsilon_2 \tag{20}$$

where $d = 1/(I+1)$, and I is the number of membership functions for a rule. ε_1, ε_2, ε_c are the random variables generated from normal distribution.

(Step 2)

Encode the triplet (X_1^i, X_2^i, X_c^i) into a string for the corresponding set of membership functions. Then, these strings are summarized as a list L.

(Step 3)

At first we optimize the weights, and then apply the GA. Optimize the weights of the rule, by using the backpropagation algorithm. The optimization is iterated 1300 times for each fuzzy inference system (Learning-I). The parameters to accelerate the convergence are selected as $\alpha = 0.2$ and $\beta = 0.3$. The appropriate values of α and β are determined by examining many examples to accelerate the convergence of the optimization while preventing the system from exploding to an unstable state.

(Step 4)

Calculate the $rmse$ for each individual, and obtain the fitness p_i by the fuzzy inference system realized by the string. Sort the individuals according to p_i, and make it as a list L. Apply the GA to the list L (Learning-II).

(Step 5)

After applying the GA algorithm, increase the count of iteration. If the total number of optimizations of the multi-stage fuzzy inference system is greater than the prescribed value (600 iterations), then the terminate the algorithm, otherwise go to step 2 again to optimize the weight for newly generated strings in list L.

In the GA treated in this paper, the string representing the parameter of the membership function is divided into two parts at an arbitrary position. For example, for the string shown in Fig.2 we can select the point of crossover in s_i. It is necessary that the crossover point be located among s_i. It is possible to divide a set of integers in s_i into two parts.

5. Approximation of Multi-dimensional Systems

5.1. Approximation of discrete multi-dimensional system

This section presents several examples of forecasting of chaotic time series by multi-stage fuzzy inference systems. We consider the logistic map as the 1-D discrete dynamic system, and the Ushiki map and the Henon map for 2-D systems[5][22]. The initial condition for a discrete system is chosen as $M = 3, N = 3$.

The forecast error *rmse* is presented for two cases of simulation studies. In the first case, the membership function included in the fuzzy rules is fixed as symmetric triangular, and is denoted as the cases without GA. On the contrary, in cases where we optimize the shape of the membership function by using the GA, a better forecast is expected. We denote the forecast as the cases with GA.

Table 2 shows the approximation error for the discrete multi-dimensional chaotic dynamics by using the multi-stage fuzzy inference systems with and without GA for optimization.

As is seen from the results, the prediction of multi-stage fuzzy inference is relatively good, especially for the estimation of system equations by using the GA, the results are meaningfully improved. The performance of the multi-stage fuzzy inference system also seems to be optimized by the GA for the selection of input variables as well as for the shape of the membership function.

For comparison, the result of approximation by using the single stage fuzzy inference system is shown in Table 3. In the single stage inference system, the number of membership functions is limited to five, and the number of input variables is set at 5. Otherwise, the inference system is intractable.

Table 2: Approximation error (1-D, 2-D)

systems	cases	rmse
Logistic map	without GA	$x(t)$:0.053
	with GA	$x(t)$:0.040
Henon map	without GA	$x(t)$:0.053, $y(t)$:0.069
	with GA	$x(t)$:0.040, $y(t)$: 0.049
Ushiki map	without GA	$x(t)$:0.042,$y(t)$:0.054
	with GA	$x(t)$:0.033,$y(t)$:0.036

Table 3: Approximation error (1-D, 2-D) by single stage inference

systems	rmse
Logistic map	$x(t)$:0.153
Henon map	$x(t)$:0.125, $y(t)$:0.189
Ushiki map	$x(t)$:0.102,$y(t)$:0.108

5.2. Approximation of continuous chaotic system

We will now consider the sampled observation of time series generated by the continuous time chaotic system such as the Roessler equations[16].

Even though, the time series generated by the continuous chaotic system must be predicted by the system describing the continuous time system, we assume here that the time series generated by a continuous system is observed in a discrete time, and the sampled values are used to approximate the system by using the fuzzy inference. The input to the system is discretized, and the output of the inference system is observed in discretized time points.

The parameters for discretization of the continuous system are chosen as $\Delta h = 0.02$, and $M = 3, N = 3$.

Table 4 shows the result of prediction error for the chaotic dynamics generated by using the Roessler equation and the Lorenz equations obtained by the multi-stage fuzzy inference system. The results obtained by using the GA are compared ones derived without GA.

As is seen from the result, the prediction of multi-stage fuzzy inference is also applicable to the approximation of continuous chaotic dynamics. The GA method contributes to the optimization of the selection of input variables as well as for the shape of the membership function.

Table 4: Prediction error for 3-D system.

systems	cases	rmse
Roessler equation	without GA	$x(t)$:0.022, $y(t)$:0.024 $z(t)$:0.026
	with GA	$x(t)$:0.015, $y(t)$:0.020 $z(t)$:0.021
Lorentz map	without GA	$x(t)$:0.033,$y(t)$:0.043 $z(t)$:0.043
	with GA	$x(t)$:0.023, $y(t)$:0.033 $z(t)$:0.034

6. Applications

6.1. Operation under artificially generated exchange rates

At first, we consider the effect of operating flexibility for the case where the exchange rate can be approxiamted by artificially generated chaotic time series[19]. As was mentioned earlier, there is no problem if we postulate two contries, and change the production among these two contries (for example, Japan and the US).

For simplicity, the exchange rate of yen versus US dollar $\phi(t)$ is represented by

$$\theta(t) = \theta_0 + 0.5(x(t) - x_{mean}) \tag{21}$$

where $x(t)$ is generated by the logistic map, and θ_0 is the initial value of $x(t)$.

We assume that $\theta(t)$ are quarterly data, and are calculated at 100 time points. The time series is estimated by using the multi-stage fuzzy inference system whose input at time t is $\theta(t-1), \theta(t-2), \ldots$. Then we get the predicted value for $\theta(t)$ and it is denoted as $\hat{\theta}(t)$.

Then, we change the site of production depending on the predicted $\hat{\theta}(t)$. Namely, if we have a higher return by changing the site of production, then we shift the production into another country.

Since it has been shown that the prediction error for the time series generated by the logistic map by using the multi-stage fuzzy inference system is very small, the operating flexibility in this case is expected.

Figure 3 shows the relative saving of production cost by changing the production site depending on the prediction $\hat{\theta}(t)$. Figure 3 shows the following rates of cost saving V_2, V_3 obtained by the production costs r_1, r_2 and r_3 where the ratios V_2 and V_3 are defined as the relative value of cost saving for r_1 with respect to r_2 and r_3.

$$r_i = \sum_t^T \rho^{(T-t)} c_i(t), i = 1,2,3 \tag{22}$$

$$V_2 = (r_2 - r_1)/r_2 \times 100, V_3 = (r_3 - r_1)/r_3 \times 100 \tag{23}$$

where $c_2(t), c_3(t)$ are the production cost necessary for producing goods without changing the site of production in a foreign country and homeland, and $c_1(t)$ is the cost of the production where the site is switched depending on the predicted $\hat{\theta}(t)$. The coefficient ρ means the deflator in the economy. The horizontal axis of Figure 3 is the initial value of $\theta(t)(= \theta_0)$.

The ratio V_2 is the relative value of cost saving of operating flexibility against the continuous production in a foreign country. The ratio V_3 is the relative value of operating flexibility against the continuous production in home land.

The net value of operating flexibility (cost saving) is defined by

$$V_m = \min(V_2, V_3) \tag{24}$$

As is seen in Figure 3, the operating flexibility works Effectively, especially in the region where the initial value of exchange rates θ_0 is close to 1.

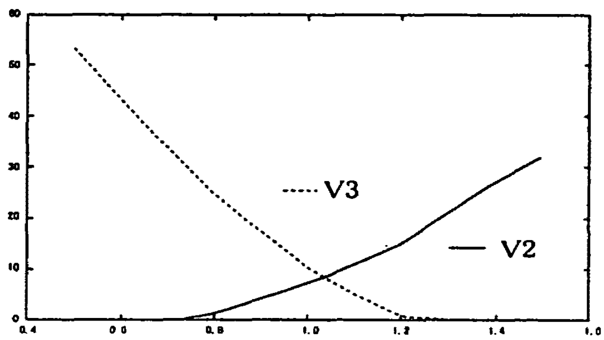

Figure 3: Relative value of operating flexibility

6.2. Operation under real exchange rate

Then, we apply the prediction method by using the multi-stage fuzzy inference system to real exchange rates, and consider the operating flexibility based on the prediction. We assume that the production is switched between Japan and the US depending on the exchange rate $\theta(t)$ of yen versus US dollar.

The exchange rate is predicted by using the multi-stage fuzzy inference system whose input variables are the several Japanese economic indicators listed in Table 5. These economic indicators, including exchange rate, were obtained from public database, and have the following properties.

Observed period:form 1970 to 1998

Sampling rate:quarterly data

Adjustment:no adjustment for seasonality, trends

Table 5: Variables corresponding to economic indicators

variables	meaning	variable	meaning
C	consumption	IM	import
CPI	consumers price index	W	wage
EX	export	R	short term interest rate
RATE	exchange rate(vs. US dollar)	T	tax
GDP	gross domestic product	U	unemployment
I	investment		

The exchange rate is predicted by the inference system, and the prediction is denoted as $\hat{\theta}(t)$. We consider two cases of prediction for exchange rate.

(Case A) using 11 indicators in time $t-1$

In this case we use 11 indicators in Table 4 but only at time point $t-1$ to predict $\theta(t)$. These indicators are used as the input variables for three stages fuzzy inference systems.

(Case B) using 3 indicators in time $t-1, t-2, t-3$

By using the correlation analysis we select three indicators that are regarded as affecting the exchange rate, and use them at time points $t-1, t-2, t-3$. The number of stages of the fuzzy inference system is three.

Figures 4 and 5 show the prediction of exchange rate by using the multi-stage fuzzy inference system for Case A and Case B. The dashed lines mean the prediction, and the solid lines denote the observations. The value of exchange rates are normalized between 0.2 and 0.8. The prediction error defined in equation (17) are 0.02 (Case A) and 0.03(Case B). The prediction error is remarkably small to apply the prediction for the operating flexibility.

We then consider the value of operating flexibility based on the prediction of $\theta(t)$. Since the exchange rate possesses a kind of long trend from 1970 to 1998, it is not relevant to use a fixed level (for example $\theta(t) = 1$) of exchange rate to switch the production. In this case we use the line as the nominal level $\theta_n(t)$ to switch the predction in place of $\theta_n(t) = 1$ as follows.

$$\theta(k) = 136.9 - 0.39k \tag{25}$$

where k is the number of quarterly data beginning at the first quarter in 1970.

Figure 6 shows the relative return by changing the production That depends on the prediction $\hat{\theta}(t)$ for Cacs A (a very similar result is obtained for Case B). We define the relative value of cost saving of operating flexibility as shown in equations (22)-(24). The horizontal axis κ of Figure 6 is the coefficient of switching cost from each site of production (we define for simplicity $\kappa_{21} = \kappa_{12}$). We define the switching cost by $\kappa\theta(t)$, while the cost seems to be dependent on the current exchange rate. The ratio V_2 is the relative value of cost saving against the continuous production in a foreign country. The ratio V_3 is the relative value of operating flexibility against the continuous production in home land. The net value of operating flexibility (cost saving) is defined by $V_m = \min(V_2, V_3)$.

As is seen in Figure 6, the operating flexibility works Effectively, especially in the region where the switching cost is relatively small. In all regions of κ the value V_2 is large, Which indicates the effectiveness of operating flexibility. Even though, as seen in Figure 6, the relative value V_3 decreases to under zero in the region $\kappa > 0.005$, we must note that the nominal exchange rate is assumed to follow the linear regression line, and in the real world, it is usually very hard to know such kind of trend.

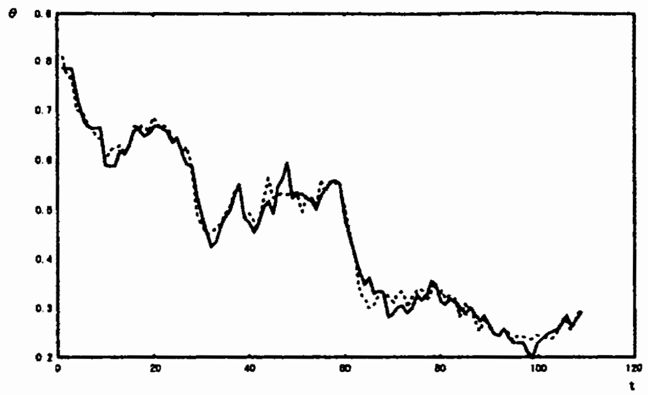

Figure 4: Prediction of exchange rate (Case A)

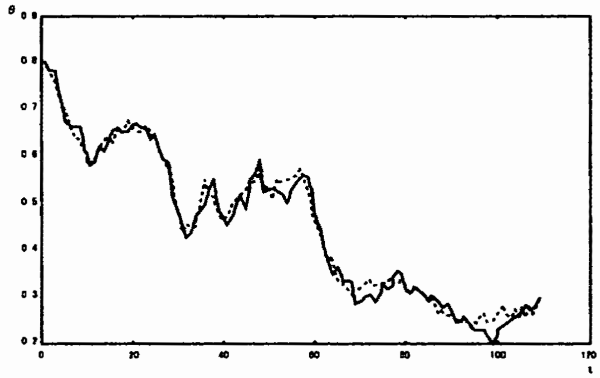

Figure 5: Prediction of exchange rate (Case B)

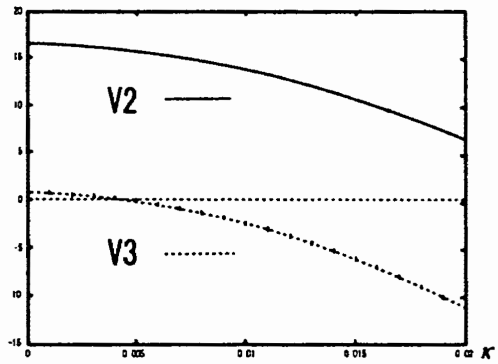

Figure 6: Relative return of operating flexibility(Case A)

7. Conclusion

This paper treated the approximation of multi-dimensional chaotic dynamics by using the multi-stage fuzzy inference system. To suppress the whole number of inference rules, the system consists of several stages of fuzzy inference and the input variables are used in a distributed manner. The backpropagation algorithm was used to optimize the weight of each rule as in the single stage inference system. The shape of the membership function was optimized by using the GA. The inference system was applied to the approximation of known discrete and continuous chaotic dynamics. The prediction was applied to estimate the exchange rate which follows the chaotic dynamics of logistic map. The prediction was then applied to estimate the exchange rates by using input variables consisting of economic indicators as well as the exchange rates. The operating flexibility using the predicted exchange rates was evaluated. We finally found effective operating flexibility for shifting the manufacturing plants depending on the predicted exchange rate.

As further work remains in applying the inference system to real chaotic dynamics and to other fuzzy inference systems, the authors plan to continue this study.

References

[1] J. D. Farmer and J. J. Sidrowich: Exploiting chaos to predict the future and reduce noise. in Y.C.Lee (eds.): *Evolution, Learning and Cognition* (World Scientific, Singapore, 1988), 277-330,

[2] L. E. Goldberg: *Genetic Algorithm in Search, Optimization and Machine* (Addison-Wesly, Reading, Massachusetts, 1989).

[3] J. M. Grandmomt: On endogenous competetive business cycles. *Econometrica*, **45** (1985) 995-1045.

[4] P. De Grauwe and K. Vansanten: Deterministic chaos in foreign exchange market. *Center for Economic Policy Research Discussion Paper*, 370(1990).

[5] M. Henon: A two-dimensional mapping with a strange attractor. *Communications in Mathematical Physics*, **50**(1976) 69-77.

[6] J. H. Holland: *Adaption in Neural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence* (University of Michigan Press, 1975).

[7] A.Huchzermeier and M.A.Cohen: Valuing operational flexibility under exchange rate risk. *Opèrations Research*, **44**-1, (1996) 100-113.

[8] Y. Kishikawa and S. Tokinaga: Prediction of chaotic time series by using the multi-stage fuzzy inference systems and its applications (in Japanese). *Proceedings of 1999 Fall National Conference of ORSJ* (1999), 26-27.

[9] Y. Kishikawa and S. Tokinaga: Prediction of stock trend by using the wavelet transform and the multi-stage fuzzy inference system optimized by the GA. *IEICE Transaction.*, **E83-A**-2(2000), 357-366.

[10] Y. Kishikawa and S. Tokinaga: Approximation of multi-dimensional chaotic dynamics by using multi-stage fuzzy inference systems and the GA. *IEICE Transaction.*, **E84-A**-9(2001), 1204-1215.

[11] B. Kogut and N. Kulatilaka: Operational flexibility, global manufacturing and the option value of a multinational networks. *Management Science*, **40**-1(1994), 123-139.

[12] S. Murakami and M. Maeda: Automobile speed control system using a fuzzy logic controller. in M.Sugeno(eds.): *Industrial Applications of Fuzzy Control* (Elsvier Science Publishers, 1985), 105-123.

[13] T. Ohtani,H. Ichihasi, T. Miyoshi and K. Nagasaka: Orthogonal and successive projection method for the learning of neurofuzzy GMDH. *Information Science*, **110**(1998) 5-24.

[14] T. J. Procyk and E. H. Mandami: A linguistic self-organizing process controller. *Automatica*, **15**(1979) 15-30.

[15] G. V. S Raju, Jun Zhou: Adaptive hierarchical fuzzy controller. *IEEE Transaction on Systems, Man and Cybernetics*, **23**-4(1993).

[16] O. E. Roessler: Continuous chaos for prototype equations, bifurcation theory and application in scientific disciplines. *Annals of the New York Academy of Sciences*, **316**(1979) 376-392.

[17] D. E. Rumelhart,G. E. Hinton and R. J. Williams: Learning internal representation by error propagation, Paralell distributed processing. in McClelland and the PDP Research Group (eds.) (MIT Press 1989).

[18] W. Shafer: Keynesian chaos. *Journal of Macroeconomic*, **7** (1985) 227-295.

[19] N. Takagi and S. Tokinaga: Evaluation of switching the production sites by using the time series prediction based on the multi-stage fuzzy inference systems (in Japanese) . *Proceedings of 2001 Spring National Conference of ORSJ*(2001) 258-259.

[20] T. Takagi and M. Sugeno: Fuzzy identification of system and its application to modeling and control. *IEEE Transaction on System Man and Cybenetics*, **SMC-15**(1985) 116-132.

[21] K. Tan and S. Tokinaga: Optimization of fuzzy inference rules by using the genetic algorithm and its applications to the bond rating. *Journal of the Operations Research Society of Japan*, **42**-3(1999) 302-315.

[22] S. Ushiki, M. Yamaguti and H.Matano: Discrete population models and chaos. *Lecture Notes in Numerical Applications and Analysis*, **2** (1980).

[23] L. Wang and J. Mendel: Generating fuzzy rules by learning from examples. *IEEE Transaction on System, Man and Cybernetics*, **22**-6(1993) 1414-1427.

[24] A. Zadeh: Fuzzy sets. *Information and Control*, **8**(1965) 338-353.

Shozo Tokinaga
6-19-1 Hakozaki, Higashi-ku
Fukuoka 812-8581, JAPAN
E-mail:
tokinaga@en.kyushu-u.ac.jp