Transactions on Ecology and the Environment vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3541

Acquisition of fuzzy rules using turbo algorithm

K. Takezawa
Division of Changing Earth and Agro-Environment, National Institute of Agro-Environmental Sciences, Kannondai 3-1-1, Tsukuba, Ibaraki 305, Japan

Abstract

A new method of acquiring fuzzy rules using observational data is proposed. The Turbo Algorithm, which has attracted much attention as an efficient procedure for deriving the nonparametric functions from data, plays a major role in this algorithm. Since the algorithm selects the best combination of rules from various candidates automatically, reliable estimation can be achieved through only a small number of rules. To confirm its validity, this method is applied to the analysis of environmental effects on rice production.

1. Introduction

Identification of the relationship between input and output using observed data is a useful technique for estimation, prediction and control. To match the recent progress in computer technology, flexible algorithms for this purpose are required.

In the present paper, the Turbo Algorithm (Friedman and Silverman, 1989) — a distinguished scheme developed as a regression method — is combined with simplified fuzzy logic. Since both of these methods are best known for their practical efficiency, the combination of the two leads to a new dimension in the practical application of regression and fuzzy logic. The relationship between rice production and climate is analyzed to exemplify the performance of this method.

In discussing the commonality of linear regression and simplified fuzzy logic, it is worth describing the common framework of the two fields:
Given $\{\mathbf{x}_i, y_i\}$ ($\mathbf{x}_i$ is a vector which is an independent variable of i–th data, y_i is a scalar which is a dependent variable of i–th data), derive

Transactions on Ecology and the Environment vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3541

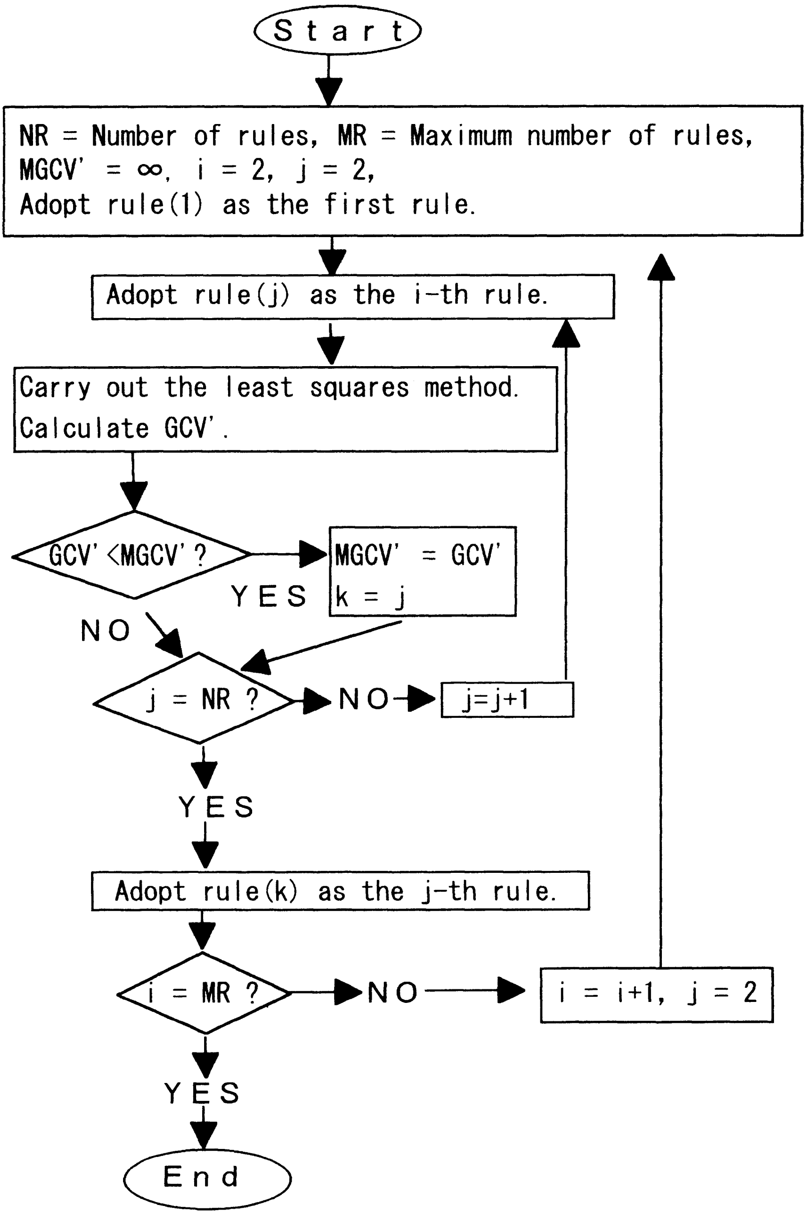

Figure 1. Flow chart of the proposed algorithm

$$Y=\sum_{j=1}^{M} a_j b_j(\mathbf{X}), \tag{1}$$

where a_i is a regression coefficient and $b_i(\mathbf{X})$ is a basis function.

Techniques useful for deriving this form from observed data contribute to both areas. Several algorithms based on multiple regression are proposed: ACE (Breiman and Friedman, 1985), CART (Breiman et al.1984), MARS (Friedman, 1991), Π−method (Breiman, 1991), and ANOVA spline (Gu and Wahba, 1993). The features of these methods have been discussed from both the theoretical and practical points of view. On the other hand, the algorithms used to generate this equation are pursued also in the context of fuzzy logic (Sugeno and Kang, 1986; Sugeno and Kang, 1988). In the recent studies, neural networks, the steepest descent method, and genetic algorithms are applied for practical purposes. The time has thus come when we should seek efficient methods for deriving this equation by integrating these areas; there is no longer any point in simply adhering to a context (Takezawa, 1993).

2. Fundamental concepts

A great amount of practical experience using fuzzy logic has shown that simplified fuzzy logic is the most efficient form in which to represent fuzzy knowledge for the purpose of estimation and control. The typical form of simplified fuzzy logic is

R_1: *If* x_1 *is close to* a_1 *and* x_2 *is close to* b_1, *then* $y=c_1$

$\vdots$

R_r: *if* x_1 *is close to* a_r *and* x_2 *is close to* b_r, *then* $y=c_r$, (2)

where r is the number of rules. The knowledge "x_1 is close to a_i and x_2 is close to b_i" is represented by the membership function

$$\mu_i(x_1, x_2)=\exp(-p(x_1-a_i)^2-q(x_2-b_i)^2), \tag{3}$$

where p and q are constants for indicating the shape of the membership function.

The estimated value (y) is written as

$$y=\frac{\Sigma\, \mu_i(x_1, x_2)c_i}{\Sigma_i(x_1, x_2).} \tag{4}$$

This is a typical weighted average in terms of linear regression. If the normalization factor in (4) is omitted, we have

$$y=\Sigma\, \mu_i(x_1, x_2)c_i. \tag{5}$$

This equation is known as the Radial Basis Function Network (Piggio and Girosi, 1990), an efficient neural network based on simulation of the function of the human eye.

In a nutshell, the practical pursuit of a flexible estimation has been reduced to a simple concept: the weighted average. This result indicates that the association of regression methods with fuzzy logic and neural networks is of considerable importance in pursuing more adaptable algorithms for analysis, prediction and control. In particular, researchers working in the field of artificial intelligence tend to make little of the regression method because of their belief that regression is obsolete. However, the recent progress made in this viable area, especially nonparametric regression (Eubank, 1988; Hastie and Tibshirani, 1990; Wahba, 1990) has substantive implications for artificial intelligence. The algorithm presented here is one which closely adheres to this line. An application of this scheme to the analysis of the influence of climatological factors on rice yield exemplifies some of this method's fascinating features.

3. Proposed algorithm

The number of rules should be small to realize efficient fuzzy estimation or control, while the number of candidates for the rules should be large. Hence, the procedure proposed here selects several rules from many candidates of rules, and the constants in the consequences of the rules are calculated to fit the observed data.

The following rule is always adopted:

Whatever the values of x_1, x_2 *may be,* $y = c_1$. (6)

Since knowledge about one variable may exist, the rules given as below are included among the applicants for rules:

If x_1 *is close to* a_2, *then* $y = c_2$. (7)

The premise of a rule is fixed, and the constants of the consequences are derived by linear calculation. Rules having various shapes are therefore treated as different rules in applicants for rules:

If x_1 *is very close to* a_3, *then* $y = c_3$, (8)

If x_1 *is somewhat close to* a_4, *then* $y = c_4$. (9)

Furthermore, a premise may include a functional relationship between independent variables:

If (x_1+2x_2) *is close to* b_5, *then* $y = c_5$. (10)

Thus, an algorithm is required to select the best combination of rules from these applicants for rules. The Turbo Algorithm, which was originally developed to combine spline functions to represent a relationship between input and output, is suggested here for this purpose. The simplicity of the Turbo Algorithm enables easy extension of the method to the acquisition of fuzzy rules. Figure 1 presents a flow chart of the Turbo Algorithm procedure. The basic concept of the method is that one rule is added and whether or not the rule is adopted is determined by a statistic. GCV' (Friedman, 1991) as defined below is used here as a statistic for determining whether or not a rule among applicants for rules is employed; a rule is adopted if addition of the rule reduces GCV':

$$\mathrm{GCV'} := \mathrm{GCV}\times(N-\mathrm{trace}(\mathbf{H}))\ /(N-\ \mathrm{trace}(\mathbf{H})-M\times d) \quad (11)$$

GCV: Generalized Cross-Validation (Craven and Wahba, 1979; Eubank, 1988)($(\mathrm{GCV})^{1/2}$ approximates the estimation error which a combination of rules generates.)

$$= \mathrm{MSE}*N/\mathrm{trace}(\mathbf{I}-\mathbf{H})^2 \quad (12)$$

MSE: mean square error

H: Hat matrix (Eubank, 1988)
N: number of data
M: number of rules
d: $2 \leqq d \leqq 4$ (Friedman, 1991). The value of d is fixed at 2.0 in the following application.

GCV is a widely used statistic for estimating the error of a linear regression equation. The slight difference between GCV and GCV' is due to the fact that the Turbo Algorithm is a nonlinear procedure because it includes trial and error.

4. Example

The method proposed in the previous section constructs a combination of fuzzy rules to represent the relationship between climatic factors and rice production. The independent variables are the year in which rice was grown, monthly average temperature (℃), and the monthly average of duration of sunlight (hours/day) in the Ibaraki area, Japan. Monthly data are taken from May through October, the period of rice growth in the Ibaraki area. Rice yield per 1,000 square meters in the area is used as the dependent variable. These data are annually published by the Ministry of Agriculture, Forestry and Fisheries.

Two combinations of independent variables are used: year and average temperature (RUN1), and year and average duration of sunlight (RUN2), such that:

R_1: *Whatever the values of* x_1 *and* x_2 *may be,* $y = c_1$;
R_2: *If* x_1 *is close to* a_2, *then* $y = c_2$;
⋮
R_{12}: *If* x_1 *is close to* a_{12}, *then* $y = c_{12}$;
R_{13}: *If* x_2 *is close to* b_{13}, *then* $y = c_{13}$;
⋮
R_{23}: *If* x_2 *is close to* b_{23}, *then* $y = c_{23}$;
R_{24}: *If* x_1 *is close to* a_{24} *and* x_2 *is* b_{24}, *then* $y = c_{24}$;
⋮
R_{144}: *If* x_1 *is close to* a_{144} *and* x_2 *is close to* b_{144}, *then* $y = c_{144}$;
R_{145}: *If* x_1 *is very close to* a_{145}, *then* $y = c_{145}$;
⋮
R_{155}: *If* x_1 *is very close to* a_{155}, *then* $y = c_{155}$;
R_{156}: *If* x_2 *is very close to* b_{156}, *then* $y = c_{156}$;
⋮
R_{166}: *If* x_2 *is very close to* b_{166}, *then* $y = c_{166}$;
R_{167}: *If* x_1 *is very close to* a_{167} *and* x_2 *is very close to* b_{167}, *then* $y = c_{167}$;
⋮
R_{287}: *If* x_1 *is very close to* a_{287} *and* x_2 *is very close to* b_{287}, *then* $y = c_{287}$;
(14)

Transactions on Ecology and the Environment vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3541

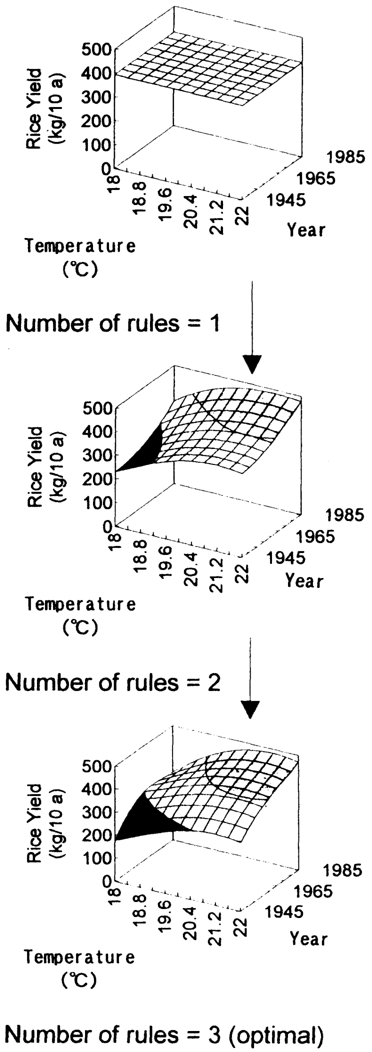

Figure 2: Input and output surface obtained by each step of the procedure; independent variables are year and average temperature.

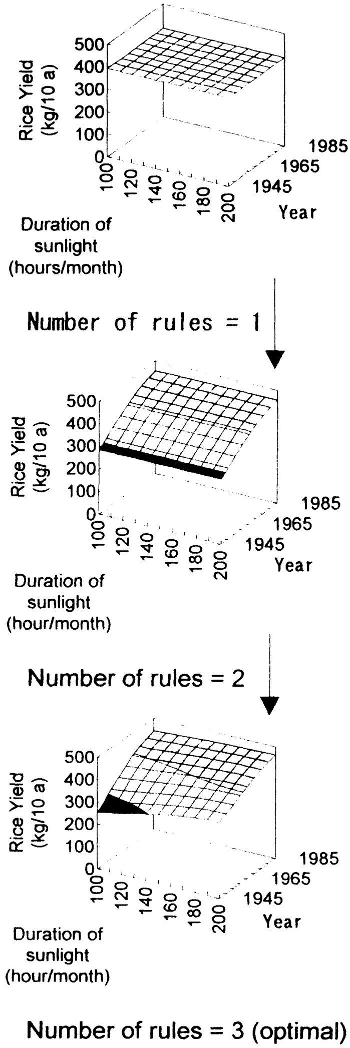

Figure 3: Input and output surface obtained by each step of the procedure; independent variables are year and average duration of sunlight.

Transactions on Ecology and the Environment vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3541

$\{a_i, b_i\}$ ($1 \leq i \leq 10$) is equi-spaced; the spans are slightly wider than those from maximum to minimum of the data. The values of p, q which indicate "close" are reciprocal numbers of the span of a_i, and b_i, respectively, and those indicating "very close" are twice a_i and b_i for "close."

Figures 2 and 3 present the results of the calculations of RUN1 and RUN2. These results show that the three rules are the most appropriate for both combinations of independent variables; these smooth surfaces which enable highly precise estimation are realized by only three rules. And since the number of rules should be small to facilitate realization of a quick control system, this approach is promising for the construction of control logic.

It should also be noted that the surface drawn by the two rules in Fig. 3 depends on the year only. If this surface produces the smallest GCV', the independent variable of the resultant fuzzy model is the year only. This is a simple example confirming that this algorithm includes the variable selection procedure.

5. Conclusions

A forward selection procedure based on the Turbo Algorithm is suggested to acquire a set of fuzzy rules to represent the relationship between independent variables and a dependent variable. An example using rice production and climatological data confirmed the validity of its performance.

This method incorporates the following characteristics:

(1) Various kinds of rules are used as applicants: local rules, global rules, rules for one independent variable, rules for several independent variables, and rules in which independent variables are functions.

(2) This algorithm is free from problems generated by nonlinear optimization because the linear least squares method constitutes the core of this procedure.

(3) The number of rules is optimized.

(4) Variable selection is performed.

(5) The procedure is completely automatic: no human decision is required.

(6) $(GCV')^{1/2}$ gives an estimate of error.

(7) A large number of independent variables of data may increase the amount of calculation.

Fuzzy logic and nonparametric regression have different origins: the former aims to represent obscure knowledge and the latter is created by generalization of traditional linear regression. The two frameworks, however, have much in common in their practical purpose: realization of a smooth surface to represent the relationships between independent variables and a dependent variable using observational data. The superficial difference should not disguise their commonality. Creative development in these areas will be accomplished by overcoming the differences in terminology and mathematical formulae.

References

Breiman, L., Friedman, J.H., Olshen, R. & Stone, C.J. *Classification and Regression Trees*. Wadsworth, Belmont, CA, 1984.

Breiman, L. & Friedman, J.H. Estimating optimal transformations for multiple regression and correlation (with discussion), *Journal of American Statistical Association*, 1985, **80**, pp.580–619.

Breiman, L. The $\prod$–method for estimating multivariate functions from noisy data (with discussion), *Technometrics*, 1991, **33**, No. 2, pp. 125–160.

Craven, P. & Wahba, G. Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation, *Numerische Mathematik*, 1979, **31**, 377–403

Eubank, R. *Spline Smoothing and Nonparametric Regression*, Marcel Dekker, New York, 1988.

Friedman, J.H. & Silverman, B.W. Flexible parsimonious smoothing and additive modeling (with discussion). *Technometrics*, 1989, **31**, No. 1, 3–39.

Friedman, J.H. Multivariate additive regression splines (with discussion), *The Annals of Statistics*, 1991, **19**, 1–141.

Gu, C. & Wahba, G. Semiparametric analysis of variance with tensor product thin plate, *Journal of Royal Statistical Society*, B, 1993, **55**, 2, 353–368.

Hastie, T.J. & Tibshirani, R. J. *Generalized Additive Model*, Chapman and Hall, London, New York, Tokyo, Melbourne, Madras, 1990.

Piggio, T. & Girosi, F. Regularization algorithms for learning that are equivalent to multilayer networks, *Science*, 1990, 247, 978–982.

Sugeno, M. & Kang, G.T. Fuzzy modelling and control of multilayer incinerator, *Fuzzy Sets and Systems*, 1986, **18**, 329–346.

Sugeno, M. and Kang, G.T. Structure identification of fuzzy model, *Fuzzy Sets and Systems*, 1988, **28**, 15–33.

Takezawa, K. Fuzzy logic from the viewpoint of nonparametric regression (in Japanese), *Journal of Japan Society for Fuzzy and Systems*, 1993, **5**, No. 6, 1273–1279.

Wahba, G. *Spline models for observational data*, Regional conference series in applied mathematics, Philadelphia, Pennsylvania: Society for Industrial and Applied Mathematics, 1990.