Cybernetics and Systems Analysis, Vol. 42, No. 2, 2006

IDENTIFICATION OF A NONLINEAR DEPENDENCE BY A FUZZY KNOWLEDGEBASE IN THE CASE OF A FUZZY TRAINING SET

A. P. Rotshtein[a] and S. D. Shtovba[b]

UDC 658.012

This paper generalizes the method of identification of nonlinear dependences by a fuzzy knowledgebase to the case of fuzzy training sets. In such a set, terms are used to estimate inputs. The computer experiments performed showed that the fuzziness of experimental data is no obstacle to identification. The use of fuzzy training sets allows one to apply the proposed method to the identification of "input-output" dependences in medicine, economics, sociology, politics, and other areas in which experimental data are based on expert judgments.

Keywords: *identification, fuzzy inference, fuzzy knowledgebase, fuzzy training set.*

INTRODUCTION

Identification of nonlinear dependences, i.e., the construction of their models from the results of observations, is an important problem and finds use in engineering, economy, medicine, sociology, and other domains [1]. In [2–5], a method of two-stage identification of nonlinear dependences with the help of fuzzy knowledgebases is proposed. At the first stage, the structural identification is realized. It consists of the formation of a fuzzy knowledgebase that roughly reflects a nonlinear interrelation "inputs–output" on the basis of linguistic rules of the form "if–then." These rules are generated by an expert or are obtained as a result of extraction of fuzzy knowledge from experimental data [6]. At the second stage, the dependence being investigated is parametrically identified by finding weights of the linguistic rules and membership functions of fuzzy terms that minimize the deviation of the results of modeling from the experimental data represented by a training set.

A distinctive feature of the method considered in [2–5] lies in the use of the so-called crisp training set consisting of a collection of quantitative pairs "inputs–output." In modern packages of fuzzy modeling [7], logical inference is also realized only for crisp values of input variables but, in many applied identification problems, the experimental data accessible for tuning a model contain nonnumerical (linguistic) estimates of input variables. For example, in problems of medical diagnostics, the values of inputs are estimated by the following linguistic terms: "If a pain is knife-like acute, the localization of the pain is the upper abdomen, skin is pale, mucous coat of stomach is pale, impairment of consciousness is insignificant, anamnesis includes stomach ulcer, then the diagnosis is a perforated ulcer." In the problem of diagnosis of cracks of buildings, we can have the following terms: "If the type of a construction is a blank wall, the form a crack is diagonal, the length of the crack equals 8 m, the width of the crack equals 10 mm, the blind area is of inadequate quality, then the cause of the crack is a differential settlement of walls." In the problem of financial forecasting, the terms being used can be as follows: "If the inflationary expectations are high, the political situation is unstable, and the general crisis of economy is observed, then the rate of the national currency will decrease by 50%."

The objective of this paper is the generalization of the method from [2–5] to the case of fuzzy training sets in which values of inputs are estimated by linguistic terms. It is assumed that identical linguistic terms that enter into a knowledgebase and into a training set are specified by the same fuzzy sets. At the same time, the membership functions of these fuzzy sets are tuned simultaneously as a result of solution of the corresponding optimization problem.

---

[a]Jerusalem Polytechnic Institute Makhon Lev, Jerusalem, Israel, *rot@mail.ict.ac.il*. [b]Vinnitsa National Technical University, Vinnitsa, Ukraine, *shtovba@ksu.vstu.vinnica.ua.* Translated from Kibernetika i Sistemnyi Analiz, No. 2, pp. 17-24, March-April 2006. Original article submitted March 16, 2005.

 1060-0396/06/4202-0176

FUZZY KNOWLEDGEBASE

We will consider an object of the type $y = f(x_1, x_2, \ldots, x_n)$ with n inputs and one output. We describe a nonlinear relation "inputs–output" by the following fuzzy knowledgebase in the format from [2]:

$$
\begin{aligned}
&\text{IF } [(x_1 = a_1^{j1}) \text{ AND } (x_2 = a_2^{j1}) \text{ AND } \ldots \text{ AND } (x_n = a_n^{j1})] \text{ (with a weight } w_{j1}),\\
&\text{OR } [(x_1 = a_1^{j2}) \text{ AND } (x_2 = a_2^{j2}) \text{ AND } \ldots \text{ AND } (x_n = a_n^{j2})] \text{ (with a weight } w_{j2}),\\
&\ldots\\
&\text{OR } [(x_1 = a_1^{jk_j}) \text{ AND } (x_2 = a_2^{jk_j}) \text{ AND } \ldots \text{ AND } (x_n = a_n^{jk_j})] \text{ (with a weight } w_{jk_j}),\\
&\text{THEN } y = d_j \quad \text{for all } j = \overline{1, m}.
\end{aligned} \tag{1}
$$

Here, a_j^{jp} is a fuzzy term that estimates an input variable x_i in the rule with a number jp, $j = \overline{1, m}$, $p = \overline{1, k_j}$, d_j is a fuzzy term that is used to linguistically estimate the output variable y on an interval $[\underline{y}, \overline{y}]$, $j = \overline{1, m}$, k_j is the number of rules in which the output y is estimated by the term d_j, $j = \overline{1, m}$, m is the number of fuzzy values of the output variable y, and $w_{jp} \in [0, 1]$ is the weight coefficient of the rule whose number is jp and that characterizes the subjective measure of confidence of an expert in its truth.

We denote by $\mu^{jp}(x_i)$ the function of membership of an input $x_i \in [\underline{x}_i, \overline{x}_i]$ in a fuzzy term a_i^{jp}, i.e., we have $a_i^{jp} = \int_{[\underline{x}_i, \overline{x}_i]} \mu^{jp}(x_i) / x_i$, $i = \overline{1, n}$, $j = \overline{1, m}$, $p = \overline{1, k_j}$. As in [2–5], we use the following bell-shaped membership function:

$$
\mu^T(x) = \frac{1}{1 + \left(\dfrac{x - b}{c}\right)^2}, \tag{2}
$$

where b and c are, respectively, the coordinate of the maximum and the coefficient of concentration of the membership function of a fuzzy set T.

FUZZY MODEL

According to [2], the degrees of membership of the output y in terms $d_1, d_2, \ldots, d_m$ for current input values $X^* = (x_1^*, x_2^*, \ldots, x_n^*)$ are defined as follows:

$$
\mu^{d_j}(X^*) = \bigvee_{p=\overline{1, k_j}} \left(w_{jp} \cdot \bigwedge_{i=\overline{1, n}} \mu^{jp}(x_i^*) \right), \quad j = \overline{1, m}. \tag{3}
$$

Here, $\mu^{jp}(x_i^*)$ is the degree of membership of the current value of the ith input $x_i^* \in [\underline{x}_i, \overline{x}_i]$ in a term a_i^{jp}, $i = \overline{1, n}$, $j = \overline{1, m}$, $p = \overline{1, k_j}$.

Relationship (3) is determined from knowledgebase (1) by the replacement of terms by membership functions and the logical operations AND ($\bigcap$) and OR ($\bigcup$) by the operations of minimum ($\wedge$) and maximum ($\vee$). Substituting the vector $X^* = (x_1^*, x_2^*, \ldots, x_n^*)$ of values of input variables in (3), we obtain the following fuzzy set of the output variable:

$$
\tilde{y}^* = \left(\frac{\mu^{d_1}(X^*)}{d_1}, \frac{\mu^{d_2}(X^*)}{d_2}, \ldots, \frac{\mu^{d_m}(X^*)}{d_m} \right) \tag{4}
$$

whose carrier is $\{d_1, d_2, \ldots, d_m\}$. To pass to a fuzzy set whose carrier is $[\underline{y}, \overline{y}]$, the membership functions of terms d_j $(j = \overline{1, m})$ should be "cut off" at the level $\mu^{d_j}(X^*)$ and the obtained fuzzy sets should be united (aggregated). According to [8], we have

$$
\tilde{y}^* = \underset{j=\overline{1, m}}{\operatorname{agg}} \left(\int_{[\underline{y}, \overline{y}]} \min\left(\mu^{d_j}(X^*), \mu^{d_j}(y)\right) / y \right), \tag{5}
$$

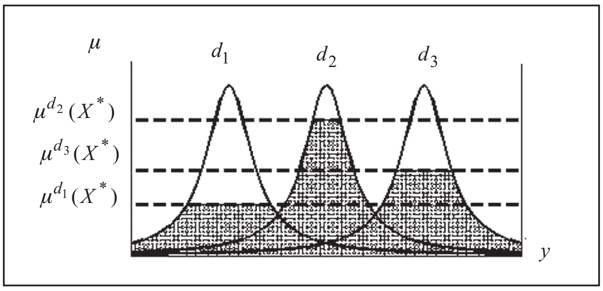

Fig. 1

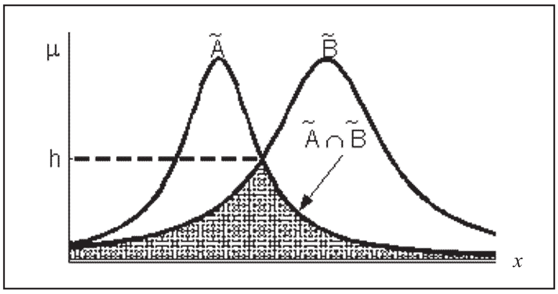

Fig. 2

where agg is the operation of aggregation of fuzzy sets that is realized by the operation of maximum and $\mu^{d_j}(y)$ is the membership function of a term d_j, $j=\overline{1,m}$, i.e., we have $d_j = \int\limits_{[\underline{y},\, \overline{y}]} \mu^{d_j}(y)/\,y.$

In Fig. 1, three fuzzy sets are aggregated. The crisp value of the output y that corresponds to the input vector $X^* = (x_1^*, x_2^*, \ldots, x_n^*)$ is determined with the help of the operation of defuzzification [8]. We use the defuzzification based on the method of center of gravity [8] since it provides the best figures of accuracy and rate of tuning a fuzzy model [9].

FUZZY TRAINING SET

We define a fuzzy training set as M pairs of experimental (or expert) data,

$$(X^r,\ y^r),\ \ r=\overline{1,M}, \tag{6}$$

where $X^r = (x_1^r, x_2^r, \ldots, x_n^r)$ is the input vector in the rth pair and y^r is the corresponding output.

In training set (6), the values of input variables can be specified not only by numbers but also by terms such as "low," "mean," "high," etc. These terms are formalized in terms of fuzzy sets with the help of membership function (2). We assume that, in a training set, fuzzy values of input variables are chosen from the same term-sets as in knowledgebase (1). Hence, the coordinates of the input vector in the rth pair of such a set can be crisp numbers $x_i^r \in [\underline{x}_i, \overline{x}_i]$ or fuzzy terms $x_i^r \in \{a_i^{jp}\}$, $i=\overline{1,n}$, $j=\overline{1,m}$, $p=\overline{1,k_j}$, $r=\overline{1,M}$.

The degrees of membership of inputs in terms from knowledgebase (1) are calculated taking into account crisp and fuzzy values. In the first case, the degree of membership is calculated by the substitution of the current value of a variable in formula (2). For fuzzy initial data, the degree of membership of one fuzzy set (the value of an input variable $\widetilde{x}_i^*$) in another fuzzy set (a term a_i^{jp} from knowledgebase (1)) must be determined. According to [8], the degree of membership is equal to the height of intersection of these fuzzy sets,

$$\mu^{jp}(\widetilde{x}_i^*) = h(a_i^{jp} \cap \widetilde{x}_i^*) = \sup_{x_i \in [\underline{x}_i, \overline{x}_i]} \min\left(\mu^{jp}(x_i), \mu^{\widetilde{x}_i^*}(x_i)\right). \tag{7}$$

Figure 2 illustrates the finding of the degree of membership of a fuzzy set $\widetilde{A}$ in a fuzzy set $\widetilde{B}$ by formula (7).

Logical inference based on fuzzy values of inputs is realized by the formula

$$\mu^{d_j}(X^*) = \bigvee_{p=\overline{1,k_j}} \left(w_{jp} \cdot \bigwedge_{i=\overline{1,n}} \sup_{x_i \in [\underline{x}_i, \overline{x}_i]} \min\left(\mu^{jp}(x_i), \mu^{\widetilde{x}_i^*}(x_i)\right)\right),\ \ j=\overline{1,m}, \tag{8}$$

that generalizes formula (3) to the case of fuzzy inputs.

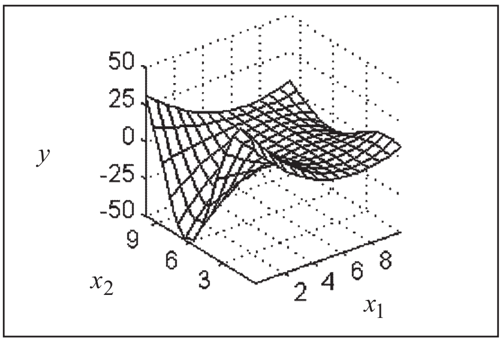

Fig. 3

TUNING A FUZZY MODEL

The tuning of a fuzzy model consists of finding its parameters that minimize the distinction between a desirable and the actual behavior of the model. In this case, the desirable behavior of the model is assumed to be specified by a fuzzy training set.

We specify the fuzzy model of an object $y = f(x_1, x_2, \ldots, x_n)$ by the formula $y = F(X, B, C, W)$, where $X = (x_1, x_2, \ldots, x_n)$ is the input vector, $B = (b_1, b_2, \ldots, b_q)$ and $C = (c_1, c_2, \ldots, c_q)$ are vectors of the parameters of membership functions (2) of fuzzy terms from knowledgebase (1), $W = (w_1, w_2, \ldots, w_N)$ is the vector of weight coefficients of fuzzy rules in (1), N is the total number of rules (lines) in (1), q is the total number of terms, and F is the "inputs–output" relation operator that corresponds to relationships (2)–(5) and (8).

Following [2–5], we formulate the problem of tuning a fuzzy model in the form of the following optimization problem: find a vector (B, C, W) such that the following condition is true:

$$R = \sqrt{\frac{1}{M} \sum_{r = \overline{1, M}} [y^r - F(X^r, B, C, W)]^2} \to \min. \tag{9}$$

It is assumed that the parameters of membership functions must be selected so that the linear ordering of the terms being considered remains valid.

COMPUTER EXPERIMENT

Let us consider an object that has two inputs $x_1, x_2 \in [0, 10]$ and one output y and that is specified by the dependence

$$y = (x_1 - 7)^2 \cdot \sin(0.61 \cdot x_2 - 5.4). \tag{10}$$

The problem was put as follows. Using the plot of reference dependence (10) (see Fig. 3), it is required to synthesize a fuzzy model and to tune it with the help of a fuzzy training set. The adequacy of the fuzzy model should be checked using criterion (9) with the help of a crisp test set consisting of 1000 randomly generated pairs "inputs–output," and the results of identification with the help of crisp and fuzzy training sets should be compared. These data sets are given in [10].

The fuzzy knowledgebase was visually constructed by an expert from Fig. 3. It consists of seven rules presented in Table 1. To linguistically estimate input variables, the terms Low, Average, and High are used. The output variable is estimated by the terms Low (L), Below Average (BA), Average (A), Above Average (AA), and High (H). The initial membership functions of these terms are shown in Fig. 4a. Before its tuning, the fuzzy model reflects only key distinctive features of the dependence being identified. In this case, the value obtained for criterion (9) with the help of the test set equals 10.3 (Fig. 5a). As an illustration, the membership functions tuned with the help of fuzzy and crisp training sets consisting of 80 points are shown in Figs. 4b and 4c, respectively. The weight coefficients (W) of the rules of these fuzzy

TABLE 1

x_1	x_2	y	***W*** before tuning	***W*** after tuning based on a fuzzy set	***W*** after tuning based on a crisp set
Low	Low	High	1	1	1
Low	Average	Low	1	1	1
Low	High	High	1	1	1
Average	–	Average	1	1	1
High	Low	Above the average	1	1	0.773
High	Average	Below the average	1	0.884	0.586
High	High	Above the average	1	0.5	0.959

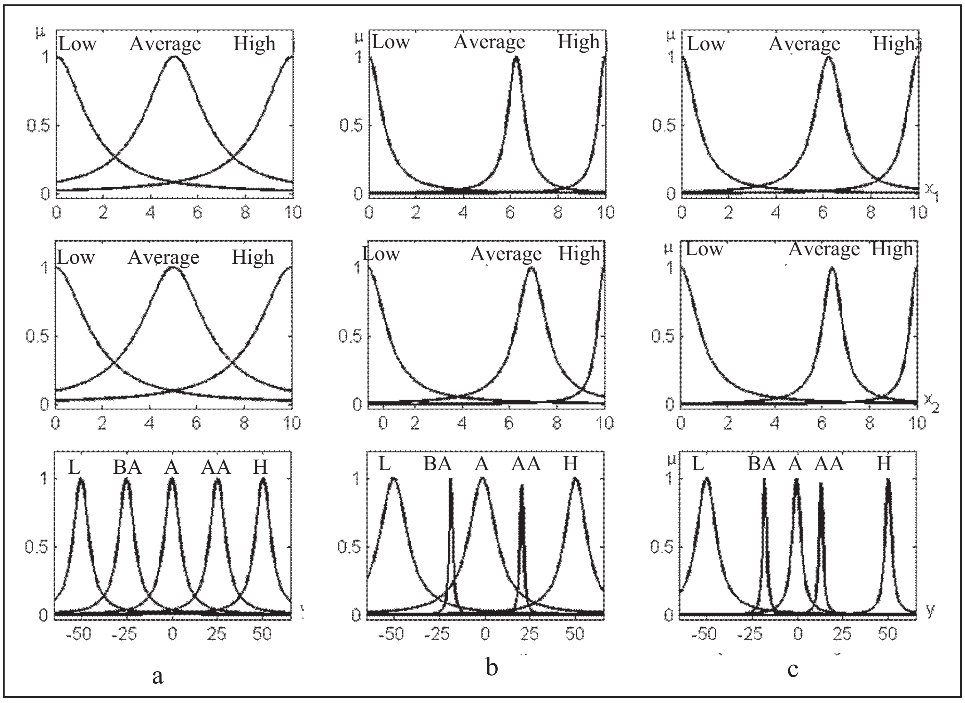

Fig. 4

models are given in Table 1. The results of testing the fuzzy models after tuning them with the help of the fuzzy and crisp sets (Figs. 5b and 5c) testify to an acceptable quality of identification of the nonlinear dependence (10).

To investigate the process of tuning the fuzzy model, training sets consisted of 10, 20, ... , 100 pairs "inputs–output" were used. The crisp sets were first generated in which the values of inputs were randomly chosen and the output was calculated by formula (10). In the fuzzy sets, one variable was specified by a number and the other was estimated by a linguistic term. A fragment of a fuzzy training set is given in Table 2. The fuzzy sets are obtained from the crisp ones by the following rules:

If $x^* \le 1.9$, then $x =$ "Low"

else if $x^* \ge 8.1$, then $x =$ "High,"

else if $x^* \in [3.8,\ 6.2]$, then $x =$ "average,"

else $x = x^*$.

Here, x^* and x are values of variables in the fuzzy and crisp sets.

In Fig. 6, the training curves of fuzzy models are shown. They reflect the dependences of identification errors (9) of training (*1*) and test (*2*) sets on the size (*M*) of the training set itself. The models were tuned by the quasi-Newtonian

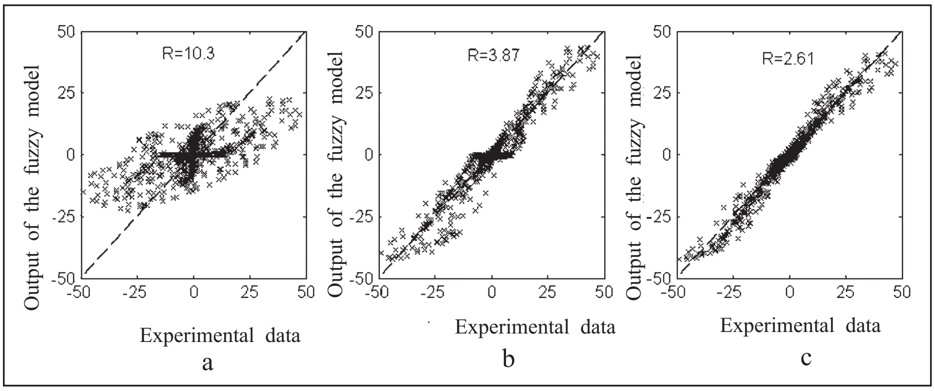

Fig. 5

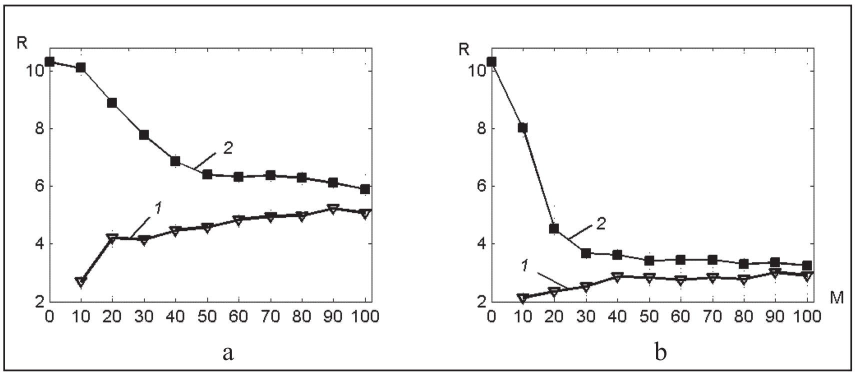

Fig. 6

TABLE 2

x_1	x_2	x_2
Low	9.8	26.46
5.38	High	–0.95
7.79	High	0.37
Average	5.46	–2.73
8.54	High	0.38
5.41	High	–0.69
Low	2.02	41.36
7.66	High	–0.21
Low	1.05	34.9
6.07	Low	0.71
2.19	Average	–21.1

Broiden–Fletcher–Golfarbd–Shanno method [11] in 15 iterations. Each point of the training curves was calculated as the average value of the results of experiments based on 10 different training sets. The discrepancy for a test set decreases and the difference between the discrepancies for training and test sets (Fig. 6a) also decreases with increasing the size of a fuzzy training set. The same phenomena also take place during the tuning based on the crisp training set (Fig. 6b).

CONCLUSIONS

A method of tuning fuzzy knowledgebases with the help of a fuzzy training set develops an earlier method of fuzzy identification of nonlinear dependences. The computer experiments made showed that the fuzziness in experimental data is no obstacle to identification. If a fuzzy set is wider than the corresponding crisp one by a factor of 3–4, then the results of tuning a fuzzy model on the basis of these sets practically coincide.

Owing to the use of fuzzy training sets, the proposed method can be applied in medicine, economy, sociology, political science, and other domains in which the experimental data used for identification of the “inputs–output” dependence being investigated are formed from expert judgements.

REFERENCES

1. Ya. Z. Tsypkin, Foundations of Information Theory of Identification [in Russian], Nauka, Moscow (1984).
2. A. Rotshtein, “Design and tuning of a fuzzy rule-based system for medical diagnosis,” in: N. H. Teodorescu (ed.), Fuzzy and Neuro-Fuzzy Systems in Medicine, CRC-Press, Boca-Raton (1998), pp. 243–289.

3. A. P. Rotshtein and D. I. Katelnikov, “Identification of nonlinear dependences by fuzzy knowledgebases,” Kibern. Sist. Anal., No. 5, 53–61 (1998).
4. A. P. Rotshtein, E. E. Loiko, and D. I. Katel’nikov, “Prediction of the number of diseases on the basis of expert-linguistic information,” Kibern. Sist. Anal., No. 2, 178–185 (1999).
5. A. P. Rotshtein, “Intellectual identification technologies: Fuzzy logic, genetic algorithms, and neural networks,” UNIVERSUM-Vinnytsya, Vinnitsa (1999).
6. A. P. Rotshtein and Yu. I. Mityushkin, “Extraction of fuzzy knowledge bases from experimental data by genetic algorithms,” Kibern. Sist. Anal., No. 3, 45–53 (2001).
7. A. Leonenkov, Fuzzy Modeling in the MATLAB and FuzzyTech Environments [in Russian], BHV, St. Petersburg (2003).
8. H. J. Zimmermann, Fuzzy Set Theory and Its Applications, Kluwer, Dordrecht (1996).
9. A. P. Rotshtein and S. D. Shtovba, “Influence of defuzzification methods on the rate of tuning a fuzzy model,” Kibern. Sist. Anal., No. 5, 169–176 (2002).
10. http://www.ksu.vstu.vinnica.ua/shtovba/benchmark.
11. Optimization Toolbox: User’s Guide, Version 2, The MathWorks, Inc. (2000).