UDC 519.71

Egamberdiyev N.A.
PhD
Associate Professor
Tashkent University of Information Technologies
named after Muhammad Al-Khwarizmi
Uzbekistan, Tashkent

Toshtemirov Z.N.
master's student
Tashkent University of Information Technologies
named after Muhammad Al-Khwarizmi
Uzbekistan, Tashkent

Akbarov S.U.
master's student
Tashkent University of Information Technologies
named after Muhammad Al-Khwarizmi
Uzbekistan, Tashkent

ANALYSIS FUZZY APPROACH OF CLASSIFICATION WEAKLY FORMALIZED PROCESSES

Annotation: Problems and shortcomings arising during the process of classification of weakly formed processes and construction of fuzzy logical models of regression problems were analyzed. As a result of the analysis, it is important to determine to what extent modeling of the current information about the research object is used correctly, that is, to what extent the model is adequate. Based on this, the main problems of developing classification models of poorly formed processes can be formulated.

Keywords: fuzzy sets, fuzzy approach, neural network, classification, weakly formalized processes, TSK fuzzy model.

Fuzzy sets are widely used in modeling complex nonlinear, uncertain objects and approximation of functions. The idea behind fuzzy modeling is that mathematical relationships between object parameters are replaced by qualitative relationships, usually expressed in terms of linguistic rules.

Fuzzy modeling is a method of describing system descriptions using fuzzy rules. Fuzzy models are mainly divided into three classes:

linguistic fuzzy models;

fuzzy relational models;

Sugeno-Takahi-Kang (TSK) models.

The construction of fuzzy models is refined using the input and output data of the object.

Linguistic fuzzy models. The structure of linguistic fuzzy models (MIMO-model) in relation to multiple-input (n) and multiple-output objects or systems:

If $X_1 = A_{11}\, end\, X_2 = A_{12}\, end \ldots X_n = A_{1n}$, in that

$Y_1 = B_{11}\, end\, Y_2 = B_{12}\, end \ldots Y_m = B_{1m}$,

and,

if $X_1 = A_{21}\, end\, X_2 = A_{22}\, end \ldots X_n = A_{2n}$, in that

$Y_1 = B_{21}\, end\, Y_2 = B_{22}\, end \ldots Y_m = B_{2m}$, (1)

also,

if $X_1 = A_{r1}\, end\, X_2 = A_{r2}\, end \ldots X_n = A_{rn}$, in that

$Y_1 = B_{r1}\, and\, Y_2 = B_{r2}\, and \ldots Y_m = B_{rm}$.

For objects with one input (x) and one output (u) (SISO-model), the r-order rule of the model is expressed in the following form:

If $X = A_r$, in that B_i, $i = \overline{1, n}$. (2)

Fuzzy models (1) and (2) can be structurally described in the form of a fuzzy graph, look-up-table with interpolation and arbitrary accentuation of inputs and outputs (structure-free). The construction of models (1) and (2) includes the following steps:

1. Determination of the number of rules. Applying fuzzy clustering to both the input space and the output space creates fuzzy networks that divide the input

and output spaces. The number of fuzzy networks generated after splitting determines the number of rules in models (1) or (2).

2. Selection of relevance function. The main search parameters in the construction of fuzzy models are the center and shape of the relevance function. These parameters are often used to adapt fuzzy models. It is theoretically possible to achieve an arbitrary accuracy of approximation by applying empirical data about the inputs and outputs of the object being modeled, using (1) or (2) in both the antecedent and consequent parts of the linguistic rules of the model. It should be noted that there is no direct way to solve this problem, that is, to choose alternative relevance functions. Recently, this problem is often put as a problem of numerical alternativeization, education. Using neural networks, genetic algorithms, the parameters of the relevance function are selected, based on which the accuracy and correctness of the models are ensured.

The considered type of fuzzy models should be capable of good semantic interpretation, controlled interpolation, which is an important factor in the accuracy and reliability of the model, and the ability to add heuristics to the generalization.

Fuzzy model of relationship. Building a fuzzy relational model is considered as a problem of solving a system of relational equations with unknown fuzzy relation:

$$B = A \circ R \tag{3}$$

$A \in F(x)$ and $B \in F(x)$ - fuzzy sets defined in the space of universes and $R \in F(X * Y)$ - a fuzzy relationship defined by the Cartesian multiplication, which represents the relationship between the inputs and outputs of the modeled object or system, $\circ$ - submin operator of composition.

The problem of identification of the fuzzy system given by (3) is evaluated based on the input and output data represented in the form of fuzzy networks A_i end B_i, $i = \overline{1, n}$:

A_i end B_i, $i=\overline{1,n}$ R is evaluated based on incoming and outgoing data, which can be represented in the form of fuzzy networks:

$R_i = A_i \circ R$, $i=\overline{1,n}$.

TSK fuzzy model. We consider a fuzzy modeling system with N inputs and one output. In this case, the TSK-model is compared to such systems.

R^i: if $X_1 = A_1^i \; end \; X_2 = A_2^i \ldots X_n = A_n^i$,

in that $Y^i = f^i(X_1, X_2, \ldots, X_n)$

looks like, here R^i, $i=\overline{1,m}$. $i-$implication, m- the number of rules of the fuzzy model, $x_1, x_2, \ldots, x_n$- model input variables, A_j^i, $i=\overline{1,n}$- fuzzy subsets of input variables, y^i- i-output defined as a nonlinear (or linear) function of the inputs.

A TSK fuzzy model can generally describe a complex nonlinear functional relationship using several rules. The rules for building such a model provide for determining the alternative structure of the conditional part (that is, the relevance function of the incoming variables), the consequent of the structure (that is, what terms are included in the nonlinear or linear equation and the value of their parameters).

It should be noted that building a TSK fuzzy model in practice is not an easy matter. There is no strict systematic procedure for building such models. In addition, it is difficult to collect input and output data for structural and parametric off-line identification in industrial process modeling.

Taking into account the advantages and disadvantages of one or another model, a new approach is proposed that takes into account the law of complementary fuzzy division based on the division of the fuzzy linear model.

Hybrid learning algorithm. This algorithm can be applied to both structures described above, which we will consider in relation to the TSK network. The hybrid algorithm for training fuzzy networks can be considered as a version of the hybrid algorithm for training radial networks.

The algorithm is performed alternately in two stages:

1. The values of the linear p_{ij} parameters of the third layer of the network are searched for the fixed values of the nonlinear parameters c_{ij}, s_{ij} end b_{ij} in the first layer of neurons;

2. At the fixed values of the linear p_{ij} parameters of the third layer, the nonlinear parameters c_{ij}, s_{ij} end b_{ij} of the first layer of the network are determined.

Stage 1. Non-linear parameters are fixed at this stage of training. The output signal is defined as:

$$y(X)=\sum_{i=1}^{M}(w^{'}*(p_{i0}+\sum_{j=1}^{N}(p_{ij}*x_j)))$$

$$w^{'}=v_i=prod[j=1:N](\mu_{ij}(x_j))/\sum_{l=1}^{N}\prod_{j=1}^{N}(\mu_{lj}(x_j))=const\,.$$

here for study choices $<X^k,d^k>, k=1,2,...,K$, K we form a system of linear equations:

$$A*P=D$$

here $P=[p_{10},p_{11},...,p_{1N},p_{M0},p_{M1},...,p_{MN}]^T$ - we form a system of linear equations, $D=[d^1,d^2,...,d^k]^T$ - is a vector of expected values constructed from all K training samples. The matrix A is given below:

$$v_1^1\ v_1^1*x_1^1....v_1^1*x_N^1....v_M^1v_M^1\ *x_1^1...v_M^1\ *x_N^1,$$
$$v_1^2\ v_1^2*x_1^2...v_1^2*x_N^2...v_M^2v_M^2\ *x_1^2...v_M^2\ *x_N^2,$$
....,
$$v_1^k\ v_1^k*x_1^k...v_1^k*x_N^k...v_M^kv_M^k\ *x_1^k...v_M^k\ *x_N^k.$$

the number of rows of the matrix is significantly greater than the number of its columns. The solution of this system of linear algebraic equations can be obtained in one step as follows:

$$P=A^{-1}*D$$

in this A^{-1} - A inverse matrix.

Stage 2. Here, the values of the polynomial coefficients of the third layer are fixed, and for the first layer of the network, the Gaussian function coefficients are refined (usually multiple) using the standard gradient method:

$$c_{ij}^{k+1} = c_{ij}^{k} - v_c * \partial E^k / \partial c_{ij}^{k},$$

$$s_{ij}^{k+1} = s_{ij}^{k} - v_s * \partial E^k / \partial s_{ij}^{k},$$

$$b_{ij}^{k+1} = b_{ij}^{k} - v_b * \partial E^k / \partial b_{ij}^{k}.$$

where k is the order of the next cycle of education (in the "online" mode, it corresponds to the order of educational selection). From a technical point of view, it is not a problem to obtain analytical solutions for the derivatives of the objective function with respect to nonlinear parameters. However, these expressions will not be presented here.

Since the step of determining the nonlinear parameters of the Gaussian function is relatively slow in the sequence of steps, the second step is usually performed several times during the first step of training.

Fuzzy self-organization algorithm C-means. In this algorithm, the next learner is passed to the input X^k vector u_i^k, $0 < u_i^k < 1$ level belong to different clusters, subject to the following condition:

$$\sum_{i=1}^{M}(u_i^k) = 1$$

In this case, the larger the value of uuuuu, the closer X^k is to C_i.

For all trained p vectors, training X^k vectors and C_i centering errors can be expressed as:

$$E = \sum_{i=1}^{M}(\sum_{k=1}^{p}(u_i^k)^m * | X^k - C_i |_2)$$

here m - 1,2,... an index selected from an array.

The goal of training is to choose such values of centers C_i that provide the minimum value of error E, which simultaneously satisfies the following condition:

$$\sum_{i=1}^{M}(u_i^k)=1.$$

The solution to this problem can be reduced to the minimization of the Lagrange function in the following form:

$$LE=\sum_{i=1}^{M}(\sum_{k=1}^{p}((u_i^k)^m*|X^k-C_i|))+\sum_{k=1}^{p}(L_k*(\sum_{i=1}^{M}(u_i^k)-1))$$

here $L_k,\ k=1,2,...,p$ - Lagrange multipliers.

It is proved that the solution of this problem can be expressed in the following form:

$$C_i=\sum_{k=1}^{p}((u_i^k)^m*X^k)/\sum_{k=1}^{p}((u_i^k)^m),$$

$$u_i^k=1/\sum_{l=1}^{M}(((d_i^k)^2/(d_i^k)^2))^{1/(m-1)}$$

here $d_i^k=|X^k-C_i|$ - X^k and C_i distance between

The learning algorithm that implements the above idea is called C-means. It has an iterative nature and is described as follows:

1. $\sum_{i=1}^{M}(u_i^k)=1$ subject to the condition $[0,1]$ from the interval u_i^k perform a random selection of coefficients.

2. Yuqorida keltirilgan bo‘yicha C_i markazlarning barcha hisoblash.

2. All calculation of C_i centers according to the above.

3. calculate the value of the error. If this value is less than the set threshold or has changed insignificantly since the previous iteration, then the calculation is terminated. Otherwise, go to step 4.

4. Calculate the new values of according to the above formula and proceed to step 2.

The iterative algorithm described above causes the error to reach a minimum, but not necessarily a global minimum. The probability of finding the global minimum is affected by the choice of initial values of . special formation processes have been developed to select "good" initial values of centers. Great progress has been made in solving classification problems, but research has shown that there are certain problems in solving this problem.

References:

1. Bekmuratov T,F., Muhamedieva D.T., Bobomuradov O.J. Fuzzy inference system for forecasting problems. IJUCI - International Journal of Ubiquitous Computing and Internationalization. V.3, No. 1, April 2011. ITIRC. - pp. 1-6.

2. Camazine S, Sneyd J.: A model of collective nectar source by honey bees: Self-organization through simple rules. J. Theor. Biol. 149 (1991) 547-571

3. Egamberdiyev N.A., FUZZY REGRESSION ALGORITHM FOR CLASSIFICATION OF WEAKLY FORMED PROCESSES // SCIENCE AND PRACTICE: IMPLEMENTATION TO MODERN SOCIETY MANCHESTER, GREAT BRITAIN 26-28.12.2020

4. Egamberdiyev N.A., Hasanov U., Neyron to’rlari yordamida boshqaruv // Материалы республиканской научно-технической конференции “Инновационные идеи в областе ИКТ и программного обеспечения”, Самарқанд, 16-17 апрель, 2019 г.-С.278-281

5. Muhamediyeva D.T., Egamberdiyev N.A., Xushboqov I.U., FORMULATION OF THE PROBLEM PARTICLE SWARM METHOD FOR SOLVING THE GLOBAL OPTIMIZATION // SCIENTIFIC HORIZON IN THE CONTEXT OF SOCIAL CRISES TOKYO, JAPAN 26-28.02.2021

6. Muhamediyeva D.T., Jurayev Z.Sh., Egamberdiyev N.A., Qualitative analysis of mathematical models based on Z-number // Proceedings of the Joint International Conference STEMM: Science – Technology – Education – Mathematics – Medicine. May 16-17, 2019, Tashkent, pp.42-43.

7. Muhamediyeva D.T., Egamberdiyev N.A. Algorithm and the Program of Construction of the Fuzzy Logical Model // International Conference on Information Science and Communications Technologies (ICISCT), DOI: 10.1109/ICISCT47635.2019.9011960, Tashkent 2019

8. Zadeh L. A. Fuzzy Sets // Information and Control. 1965. -Vol. 8, №3.-P. 338-353.

9. Круглов В.В., Дли М.И., Голунов Р.Ю. Нечеткая логика и искусственные нейронные сети. –М.: Физматлит. 2001. - 224 с.

10. D Muhamedieva, N Egamberdiev, O Xolmuminov. APPLICATION OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES FOR CREDIT RISK ASSESSMENT. Science and innovation. 2022, №6. 388-395 pp.