Expert Systems with Applications 36 (2009) 12375–12382

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A fuzzy modeling method via Enhanced Objective Cluster Analysis for designing TSK model

Na Wang *, Yupu Yang

Department of Automation, Shanghai Jiao Tong University, 800 Dongchuan Road, Minhang Shanghai 200240, PR China

ARTICLE INFO

Keywords:
Fuzzy modeling
TSK model
Fuzzy clustering
Objective Cluster Analysis
Stable Kalman Filter

ABSTRACT

This paper proposes a fuzzy modeling method via Enhanced Objective Cluster Analysis to obtain the compact and robust approximate TSK fuzzy model. In our approach, the Objective Cluster Analysis algorithm is introduced. In order to obtain more compact and more robust fuzzy rule prototypes, this algorithm is enhanced by introducing the Relative Dissimilarity Measure and the new consistency criterion to represent the similarity degree between the clusters. By these additional criteria, the redundant clusters caused by iterations are avoided; the subjective influence from human judgment for clustering is weakened. Moreover the clustering results including the number of clusters and the cluster centers are considered as the initial condition of the premise parameters identification. Thus the traditional iteration modeling procedure for determining the number of rules and identifying parameters is changed into one-off modeling, which significantly reduces the burden of computation. Furthermore the decomposition errors and the approximation errors resulted from premise parameters identification by Fuzzy *c*-Means clustering are decreased. For the consequence parameters identification, the Stable Kalman Filter algorithm is adopted. The performance of the proposed modeling method is evaluated by the example of Box–Jenkins gas furnace. The simulation results demonstrate the power of our model.

1. Introduction

Fuzzy modeling (FM) – system modeling with fuzzy rule-based systems (FRBSs) (Alcala, Casillas, Cordon, & Herrera, 2001; Casillas, Cordon, & Herrera, 2002; Cordon & Herrera, 1997, 2001; Lee & Chen, 2008; Setnes, Babuska, & Verbruggen, 1998) – may be considered as an approach used to model a system by means of a descriptive language based on fuzzy logic with fuzzy predicates. Several types of modeling can be performed depending on the desired degree of interpretability and accuracy of the final model. Usually, both requirements are contradictory properties directly on the model structure and learning process. As for the former, there are at least two different kinds of fuzzy models in the literature, the Mamdani-Assilian (MA) (Mamdani, 1974) and Takagi–Sugeno–Kang (TSK) (Takagi & Sugeno, 1985) ones. Because of the nature of consequences, TSK fuzzy model is less readable but better approximator than MA one, being a good choice in the accuracy-oriented fuzzy modeling.

On the other hand, as a strong modeling tool for learning nonlinear systems, fuzzy clustering (Bezdek, 1981) can realize the direct fuzzy partition of input space and extract the efficient rules regardless of the number of input variables. Therefore, it is widely applied in the premises identification of TSK model, including the number of rules and premise parameters.

Generally, there are three kinds of fuzzy clustering methods to determine the number of rules. But it is difficult to obtain the accurate and reasonable result directly using them, which decreases the robustness of the model. Firstly, the number of rules is determined iteratively by increasing (Kim, Park, Ji, & Park, 1997; Tsekouras, 2005) or merging clusters (Kaymak & Babuska, 1995; Krishnapuram, 1994) from the initial clustering number. However, the iterative nature causes heavy burden in computation. Furthermore, for the case where the clusters increase, redundant rules are probably produced for reasons of noise, that means overfitting; while for those with clusters merging, some parameters require presetting, consequently the number of rules tends to be easily affected by subjective human judgment. Secondly, in order to avoid overfitting, the initial number of clusters is obtained by Mountain Clustering (Rickard, Yager, & Miller, 2005; Yager & Filev, 1994) or Subtractive Clustering (Chiu, 1994; Eftekhari, Katebi, Karimi, & Jahanmiri, 2008) methods. Then orthogonal transformation methods (Abonyi, Roubos, Babuska, & Szeifert, 2003; Yen & Wang, 1998) are proposed to simplify the redundant rules. Nevertheless, the number of rules is still easily influenced by human decision because whether the rule is redundant is determined by trial and error. Thirdly, the number of rules is produced through trade-off among several clustering validity indexes (Lee, 2008; Tsekourasa,

* Corresponding author.
E-mail addresses: wangna2004@sjtu.edu.cn, wnsh2008@yahoo.cn, wangna2008@foxmail.com (N. Wang).

0957-4174/$ - see front matter
doi:10.1016/j.eswa.2009.04.048

Sarimveis, Kavakli, & Bafas, 2005). Similarly, there is still the effect of human judgment.

For premise parameters, decomposition error and approximation error may occur in their identification procedure (Roubos & Setnes, 2000). Therefore, the estimation of fuzzy partition of input space is poor. Those errors come from two potential reasons: clustering projection to one-dimensional fuzzy sets and approximation of the point-wise defined membership functions by parametric ones. Those errors could be reduced by eigenvector projection (Babuska & Verbruggen, 1997), and/or by fine-tuning the parameterized membership functions. For the case of eigenvector projection, Gath–Geva (GG) clustering algorithm (Gath & Geva, 1989) and its modifiers (Abonyi, Babuska, & Szeifert, 2002) are proposed. Nevertheless, the number of rules still need to be specified in advance. For the case of fine-tuning, it tends to trapping into overfitting.

Hence, due to the poor performance of the fuzzy prototypes resulted from the above fuzzy clustering methods, the quality of the attained models is seriously reduced with regard to the compactness and robustness. In order to design the more compact and robust approximate TSK fuzzy model, a fuzzy modeling method via Enhanced Objective Cluster Analysis (EOCA) is proposed in this paper. Firstly, the Objective Cluster Analysis (OCA) (Muller, 1998) algorithm is introduced here. The clustering result is obtained by means of dipole partitioning and agglomerative hierarchical clustering procedure. However, the redundant clusters possibly occur in the OCA clustering for the reason of noise. Therefore, in order to acquire more robust fuzzy prototypes, OCA is enhanced by introducing the Relative Dissimilarity Measure (Mollineda & Vidal, 2000) into it and modifying its original consistency criterion. This avoids the redundant clusters, as well as decreases the influence from the human judgment. Furthermore, the clustering results are considered as the initial condition of the premise identification by means of the Fuzzy c-Means (FCM) algorithm (Bezdek, 1981). Therefore, not only the convergence of fuzzy clustering is improved correspondingly, but also the traditional iteration modeling procedure for determining the number of rules and identifying parameters is changed into one-off modeling, which significantly reduces the burden of computation. Simultaneously, the errors of decomposition and approximation from the premise parameters identification are also decreased. In addition, the Stable Kalman Filter algorithm (Takagi & Sugeno, 1985) is adopted to identify the consequence parameters for solving the problem of non-numerical solution.

The rest of this paper is organized as follows. In Section 2, the general TSK fuzzy model structure considered in this work is described. Section 3 presents the fuzzy prototypes extraction approach via EOCA including the relative definitions of EOCA, the principle of EOCA, the performance analysis of EOCA and the description of fuzzy prototypes extraction approach via EOCA. The fuzzy modeling method via EOCA, containing the identification of premise parameters via FCM, the estimation of consequence parameters via Stable Kalman Filter and the description of fuzzy modeling method via EOCA are presented in Section 4. Section 5 demonstrates the effectiveness of the proposed method by the famous example of Box–Jenkins gas systems (Box, Jekins, & Reinsel, 1994). Conclusions are drawn in Section 6.

2. TSK fuzzy model

The TSK fuzzy model proposed by Takagi–Sugeno–Kang (Takagi & Sugeno, 1985) is a set of fuzzy rules with the addition of fuzzy reasoning. Each rule of the model represents a local field of the system by a linear function, while the model itself can describe the global static input–output behaviors of the nonlinear system. Assume that the identified object is $\mathbf{P}(\mathbf{X},\mathbf{Y}), \mathbf{X} = (x_1, x_2, \ldots, x_r) \in \mathbf{R}^r$ is the input variable and $\mathbf{Y} = (y_1, y_2, \ldots, y_q)$ is the output variable of a MIMO nonlinear system. Then each fuzzy rule in the TSK fuzzy model is described in the formula (1):

$$
\begin{aligned}
&R^i: \text{ If } x_1 \text{ is } \mathbf{A}_1^1 \text{ and} \ldots \text{and } x_r \text{ is } \mathbf{A}_r^1. \text{ Then} \\
&y^i = p_0^i + p_1^i \cdot x_1 + \ldots + p_r^i \cdot x_r
\end{aligned}
\tag{1}
$$

where R^i is the ith fuzzy rule, $i = 1, 2 \ldots, n$, $\mathbf{A}_j^i$ is the jth fuzzy set in R^i, $j = 1, 2, \ldots, r$, p_j^i is the consequence parameter.

In this work, fuzzy clustering is proposed to identify the premise parameters; therefore, the expression of the premise parameters can be described as the following text.

For the sample set $\mathbf{Z} = \{z_1, z_2, \ldots, z_N\}$, where $z_k = (\mathbf{X}^k, \mathbf{Y}^k)$, $\mathbf{X}^k = (x_1^k, x_2^k, \ldots, x_r^k) \in \mathbf{R}^r$ is the input variable, $\mathbf{Y}^k \in \mathbf{R}$ is the output variable, $k = 1, \ldots, N$, and N is the sample number of $\mathbf{Z}$. Then the premise parameter of input $x_j^k \in \mathbf{X}^k$ is defined as the value of $\mathbf{A}_j^i(x_j^k)$.

3. Fuzzy prototypes extraction via EOCA

In this section, the EOCA algorithm is proposed to acquire more robust fuzzy prototypes – the number of clusters and the cluster centers. As the result of the partition of dipoles, the number of clusters, i.e., the rule number is determined by one-pass clustering, which decreases the huge computation in the clusters iterations. Furthermore, by introducing the Relative Dissimilarity Measure into the OCA algorithm and modifying its original consistency criterion to reevaluate the similarity degree between the clusters, the disturbance of noise for clustering is eliminated; simultaneously, the subjective judgment by the human is decreased. The definitions of dipole (Muller, 1998), Relative Dissimilarity Measure (Mollineda & Vidal, 2000) and enhanced consistency criterion is described in the next subsection respectively.

3.1. Related definitions of EOCA

Definition 1 (*Dipole*).
Given the sample set $\mathbf{Z} = (z_1, \ldots, z_N)$, and N is the number of samples. The sample pair $\mathbf{O}_{ij}^k = \begin{pmatrix} z_i \\ z_j \end{pmatrix}$ is called dipole, where $i, j = 1, 2, \ldots, N, i \neq j, k = 1, \ldots, \frac{N\times(N-1)}{2}$. The value of $\mathbf{O}_{ij}^k$ is defined as $d_{ij} = \|z_i - z_j\|$, and $\|\cdot\|$ is the Euclidean distance norm.

Definition 2 (*Relative Dissimilarity Measure*).
The Relative Dissimilarity Measure represents the relative similarity degree between different clusters in the same subset (refer to Section 3.2) before each clusters merging of hierarchical clustering, and is described in formulas (2) and (3):

$$D_{ij} = \frac{d_{ij}}{min(\overline{D}_{ij}, \overline{D}_{ji})} \tag{2}$$

$$\overline{D}_{ij} = \frac{\sum_{k\in c, k\neq j} d_{ik}}{c-2} \tag{3}$$

where D_{ij} is the relative similarity degree, $i \neq j, i, j = 1, \ldots, c$, c is the clustering number at each clusters merging. d_{ij} is the Euclidean distance of the ith cluster and the jth cluster. $\overline{D}_{ij}$ is the mean distance between the different clusters, i.e., the ith cluster and the kth cluster, where $k \neq j$, and $k = 1, \ldots, c$.

Definition 3 (*Enhanced Consistency Criterion*).
The Enhanced Consistency Criterion represents the similarity degree between the clusters in the different subsets (also refer to