Available online at www.sciencedirect.com
ScienceDirect
IFAC-PapersOnLine 48-10 (2015) 129–134

Comparison of Approaches for Identification of All-data Cloud-based Evolving Systems

Sašo Blažič * **Plamen Angelov** **,*** **Igor Škrjanc** ****

* *Faculty of Electrical Engineering, University of Ljubljana, Tržaška 25, Ljubljana, Slovenia (e-mail: saso.blazic@fe.uni-lj.si).*
** *School of Computing and Communications, Lancaster University, United Kingdom, (e-mail: p.angelov@lancaster.ac.uk)*
*** *Chair of Excellence, Carlos III University, Madrid, Spain.*
**** *Faculty of Electrical Engineering, University of Ljubljana, Tržaška 25, Ljubljana, Slovenia (e-mail: igor.skrjanc).*

Abstract: In this paper we deal with identification of nonlinear systems which are modelled by fuzzy rule-based models that do not assume fixed partitioning of the space of antecedent variables. We first present an alternative way of describing local density in the cloud-based evolving systems. The Mahalanobis distance among the data samples is used which leads to the density that is more suitable when the data are scattered around the input-output surface. All the algorithms for the identification of the cloud parameters are given in a recursive form which is necessary for the implementation of an evolving system. It is also shown that a simple linearised model can be obtained without identification of the consequent parameters. All the proposed algorithms are illustrated on a simple simulation model of a static system.

Keywords: Identification, evolving systems, Mahalanobis distance, Takagi-Sugeno model, clusters

1. INTRODUCTION

Real systems are inherently nonlinear. Under certain circumstances it is possible and reasonable to assume linearity in relatively small operating regions. Thus, the analysis and control design of the system become considerably less problematic. Sometimes, the assumption of system linearity oversimplifies the system which leads to reduced control performance, poor identification model, non-robust fault detection system etc.

In this paper we deal with nonlinear systems which are modelled by fuzzy rule-based (FRB) models. The paper is focused on the identification issues of the FRB models. Traditionally, FRB systems often assumed fixed partitioning of the space of antecedent variables. This means that only the consequent models' parameters need to be estimated. Identification of the Takagi-Sugeno fuzzy model is the one that arguably received the most attention; the early works date to the 1980s (Sugeno and Kang, 1988). Later many works followed and the area is still alive. When the model needs to be estimated on line, recursive algorithms are needed. Often a version of recursive least squares algorithm has been applied. Global and local approaches to estimate consequent models' parameters were presented by Angelov and Filev (2004). The problem of identification of dynamic systems is that with the local approach the local models have more appropriate local behaviour while the fuzzy model is less accurate globally (Yen et al., 1998; Sorensen, 1993; Abonyi et al., 2001). Nevertheless, different local identification approaches are presented by Soleimani-B et al. (2010), Dovžan and Škrjanc (2011b), Dovžan et al. (2012), Blažič et al. (2014), Precup and Preitl (2006), Blažič et al. (2009) etc.

The second problem in nonlinear system identification is to properly partition the space of antecedent variables. The methods are based on learning algorithms for neural networks (Werbos, 1974), evolving clustering (Kasabov and Song, 2002), subtractive clustering (Angelov and Filev, 2004), fuzzy c-means clustering (Dovžan and Škrjanc, 2011b), Gustafson-Kessel clustering (Dovžan and Škrjanc, 2011a) and others (Johanyák and Papp, 2012; Vaščák, 2012; Rădac et al., 2011).

Recently, a special type of fuzzy FRB systems with non-parametric antecedents has been proposed by Angelov and Yager (2010). Unlike traditional Mamdani and Takagi-Sugeno FRB systems, the approach does not require an explicit definition of fuzzy sets (and their corresponding membership functions) for each input variable. It introduced the concept of granules in Angelov and Yager (2010) and later clouds (Angelov and Yager, 2011) that rely on relative data density to define antecedents. Data clouds are subsets of previous data samples with common properties. In the original works (Angelov and Yager, 2010, 2011) data closeness has been used as a similarity measure. The approach itself is not limited to any particular similarity measure to classify data into clouds. In identification of dynamical systems it is very important to distinguish among the operating regions that represent different system dynamics. Those regions could be seen as natural

2405-8963 © 2015, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2015.08.120

clouds. Even if we choose to select the framework of cloud based system identification, there are still a number of subtasks that have to executed. There are also some possible changes that can be introduced to the original method while still keeping the general methodology.

The relative density in the original papers (Angelov and Yager, 2010, 2011) was based on Euclidean distance among the data samples in the cloud although it was stated that any other distance could be used. In the current paper two distance metrics are compared: the original Euclidean distance and Mahalanobis distance where we introduced some versions for calculating actual density.

We limit ourselves to static systems that map the multi-dimensional input space to the real numbers in this work. This simplifies the problem of identification because the problems related to identification of dynamic systems are omitted.

In Section 2 the description of Takagi-Sugeno fuzzy model used in this paper is given, Section 3 treats the identification of the antecedent part, Section 4 shows some simulation examples, and conclusions are given in 5.

2. TAKAGI-SUGENO FUZZY MODEL OF A NONLINEAR SYSTEM

A typical Takagi-Sugeno fuzzy model (Takagi and Sugeno, 1985) is given in the form of rules:

$$\textbf{if } z_1 \textbf{ is } \mathbf{A}_{1,k_1} \ \dots \ \textbf{and } z_q \textbf{ is } \mathbf{A}_{q,k_q} \textbf{ then } y = \phi_j(\mathbf{x})$$
$$j = 1, \dots, m \quad k_1 = 1, \dots, f_1 \qquad k_q = 1, \dots, f_q \quad (1)$$

The q-element vector $\mathbf{z}^T = [z_1, ..., z_q]$ denotes the input or variables in the antecedent part of the rules, and variable y is the output of the model. With each variable in the antecedent z_i $(i = 1, \dots, q)$, f_i fuzzy sets $(\mathbf{A}_{i,1}, \dots, \mathbf{A}_{i,f_i})$ are associated, and each fuzzy set $\mathbf{A}_{i,k_i}$ $(k_i = 1, \dots, f_i)$ is associated with a real-valued function $\mu_{A_{i,k_i}}(z_i) : \mathbb{R} \to [0, 1]$, that produces membership grade of the variable z_i with respect to the fuzzy set $\mathbf{A}_{i,k_i}$. To make the list of fuzzy rules complete, all possible variations of fuzzy sets are given in Eq. (1), yielding the number of fuzzy rules $m = f_1 \times f_2 \times \dots \times f_q$. The variables z_i are not the only inputs of the fuzzy system. Implicitly, the n-element vector $\mathbf{x}^T = [x_1, ..., x_n]$ also represents the input to the system. It is usually referred to as the consequence vector. The functions $\phi_j(\cdot)$ can be arbitrary smooth functions in general, although linear or affine functions are usually used.

The system in Eq. (1) is easily described in the closed form in the case of a product-sum Takagi-Sugeno fuzzy model

$$y = \frac{\sum_{k_1=1}^{f_1} \cdots \sum_{k_q=1}^{f_q} \mu_{A_{1,k_1}}(z_1) \dots \mu_{A_{q,k_q}}(z_q)\ \phi_j(\mathbf{x})}{\sum_{k_1=1}^{f_1} \cdots \sum_{k_q=1}^{f_q} \mu_{A_{1,k_1}}(z_1) \dots \mu_{A_{q,k_q}}(z_q)} \quad (2)$$

Note a slight abuse of notation in Eq. (2) since j is not explicitly defined as a running index. From Eq. (1) it is evident that each j corresponds to a specific variation of indexes k_i, $i = 1, \dots, q$.

To simplify Eq. (2), a partition of unity is considered where functions $\beta_j(\mathbf{z})$ defined as

$$\beta_j(\mathbf{z}) = \frac{\mu_{A_{1,k_1}}(z_1) \dots \mu_{A_{q,k_q}}(z_q)}{\sum_{k_1=1}^{f_1} \cdots \sum_{k_q=1}^{f_q} \mu_{A_{1,k_1}}(z_1) \dots \mu_{A_{q,k_q}}(z_q)}$$
$$j = 1, \dots, m \quad (3)$$

give information about the fulfilment of the respective fuzzy rule in the normalized form. It is obvious that $\sum_{j=1}^{m} \beta_j(\mathbf{z}) = 1$ irrespective of $\mathbf{z}$ as long as the denominator of $\beta_j(\mathbf{z})$ is not equal to zero (this can be easily prevented by stretching the membership functions over the whole potential area of $\mathbf{z}$). Combining Eqs. (2) and (3) and changing summation over k_i by summation over j we arrive to the following equation:

$$y = \sum_{j=1}^{m} \beta_j(\mathbf{z}) \phi_j(\mathbf{x}) \quad (4)$$

From Eq. (4) it is evident that the output of a fuzzy system is a function of the antecedent vector $\mathbf{z}$ (q-dimensional) and the consequence vector $\mathbf{x}$ (n-dimensional). The dimension of the input space d may be and usually is lower than $(q + n)$ since it is very usual to have the same variables present in vectors $\mathbf{z}$ and $\mathbf{x}$.

The class of fuzzy models have the form of linear models, this refers to $\{\beta_j\}$ as a set of basis functions. The use of membership functions in input space with overlapping receptive fields provides interpolation and extrapolation. It is very common to define the output value as a linear combination of consequence variables $\mathbf{x}$

$$\phi_j(\mathbf{x}) = \boldsymbol{\theta}_j^T \mathbf{x}, \quad j = 1, \dots, m, \quad \boldsymbol{\theta}_j^T = [\theta_{j1}, \dots, \theta_{jn}] \quad (5)$$

If the matrix of the coefficients for the whole set of rules is denoted as $\boldsymbol{\Theta}^T = [\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_m]$ and the vector of membership values as $\boldsymbol{\beta}^T(\mathbf{z}) = [\beta_1(\mathbf{z}), \dots, \beta_m(\mathbf{z})]$, then Eq. (4) can be rewritten in the matrix form

$$y = \boldsymbol{\beta}^T(\mathbf{z}) \boldsymbol{\Theta} \mathbf{x} = \sum_{j=1}^{m} \beta_j(\mathbf{z}) \boldsymbol{\theta}_j^T \mathbf{x} \quad (6)$$

A fuzzy model in the form given in Eq. (6) is referred to as an affine Takagi-Sugeno model and can be used to approximate any arbitrary function that maps any compact set $\mathbf{C} \subset \mathbb{R}^d$ from the input space (the input space is the space of the union of variables in $\mathbf{x}$ and $\mathbf{z}$) to $\mathbb{R}$ with any desired degree of accuracy.

3. IDENTIFICATION OF THE ANTECEDENT PART

The local density γ_k^j is defined by a suitable kernel over the distances between the current sample $\mathbf{z}(k)$ and all the previous samples that have already been classified to a particular cloud (j-th in this case) (Angelov and Yager, 2011):

$$\gamma_k^j = \frac{1}{1 + \rho \frac{\sum_{i=1}^{M^j} d_{ki}^j}{M^j}} \quad j = 1, ..., m \quad (7)$$

where d_{ki}^j denotes the square of the (Euclidean) distance between the current data sample $\mathbf{z}(k)$ and the i-th sample of the j-th cloud $\mathbf{z}_i^j$, while M^j is the number of input data samples associated with the j-th cloud. Note the factor ρ which is not present in Angelov and Yager (2011) and will be discussed later.