Applied Soft Computing 25 (2014) 174–183

Contents lists available at ScienceDirect

Applied Soft Computing

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

Parameter identification of PWARX models using fuzzy distance weighted least squares method

Ankit K. Shah [a], Dipak M. Adhyaru [b,*]

[a] Department of Instrumentation & Control, SVIT, Vasad, Gujarat, India
[b] Department of Instrumentation & Control, Nirma University, Ahmedabad, Gujarat, India

ARTICLE INFO

Article history:
Received 19 July 2013
Received in revised form 4 July 2014
Accepted 9 September 2014
Available online 18 September 2014

Keywords:
PieceWise AutoRegressive eXogenous
Hybrid dynamical system
Fuzzy-c-means clustering
Fuzzy distance weight matrix
Weighted least squares

ABSTRACT

PieceWise AutoRegressive eXogenous (PWARX) models represent one of the broad classes of the hybrid dynamical systems (HDS). Among many classes of HDS, PWARX model used as an attractive modeling structure due to its equivalence to other classes. This paper presents a novel fuzzy distance weight matrix based parameter identification method for PWARX model. In the first phase of the proposed method estimation for the number of affine submodels present in the HDS is proposed using fuzzy clustering validation based algorithm. For the given set of input–output data points generated by predefined PWARX model fuzzy c-means (FCM) clustering procedure is used to classify the data set according to its affine submodels. The fuzzy distance weight matrix based weighted least squares (WLS) algorithm is proposed to identify the parameters for each PWARX submodel, which minimizes the effect of noise and classification error. In the final phase, fuzzy validity function based model selection method is applied to validate the identified PWARX model. The effectiveness of the proposed method is demonstrated using three benchmark examples. Simulation experiments show validation of the proposed method.

1. Introduction

Hybrid dynamical systems (HDS) are systems that involve nonlinear continuous dynamics and discrete events with some logical rules [1,2]. A mathematical model of the HDS must change its continuous dynamic as the discrete event occurred in the systems. PieceWise AutoRegressive eXogenous (PWARX) models have been widely adopted by researcher to represent various HDS due to their universal approximation properties of PieceWise Affine (PWA) maps and their equivalence to other classes of the HDS [3–5].

The PWARX models comprise a set of classical ARX submodels, each of which is partitioned with a finite number of the polyhedral partitions. The identification of PWARX models involves the estimation for the number of submodels present in the HDS, the parameters of each submodel and the regions of the submodel defining the polyhedral partition. In [6] methods for the identification of the PWA models of the HDS are classified. The identification problem becomes more challenging as it includes the estimation of the number of submodels along with data classification, parameters and method to verify the identified model parameters. According to [5], four procedures for the identification of PWARX models are algebraic procedure [7], clustering-based procedure [8], Bayesian procedure [9] and bounded-error procedure [10]. The PWA system identification using mixed-integer quadratic programming (MIQP) and mixed-integer linear programming (MILP) proposed in [11] which is NP-hard and required more computations. However, the estimation of the number of submodels and the selection of submodels has not been dealt properly in the reviewed literature.

The aim of the paper is to show how to estimate the number of submodels and identification of the parameters of the submodels associated with each discrete event along with the classification of the data points with reasonable computational complexity. Clustering algorithms have been widely used in data classification, pattern recognition, system modeling, machine learning and image processing [12]. In the past few years, many clustering algorithms have been developed and used for the classification of the data points for the PWARX models. Works based on clustering algorithms for identification of the PWARX models have been reported on k-means [8,13,14], support vector classifier [5] and fuzzy clustering [15–18]. The clustering based procedure, proposed by [8] and [13] consists of the k-means clustering approach with the assumption that the number of discrete modes given a priori. The statistical clustering based identification approach given in [14], proposed the algorithm for estimating the number of submodels, but it requires an exhaustive search between two predefined limits. However, the

* Corresponding author. Tel.: +91 9016805104.
E-mail addresses: shah_ankit2@gtu.edu.in (A.K. Shah), hetdip@rediffmail.com (D.M. Adhyaru).

http://dx.doi.org/10.1016/j.asoc.2014.09.005
1568-4946/

k-means clustering procedures suffer from the drawbacks of poor initialization and influence of outliers. In k-means clustering algorithms, data points are partitioned into some clusters such that each data point is assigned to exactly one cluster. However, in fuzzy clustering, a data point may belong to several clusters with degree of membership function. Therefore, the membership function values for a data point will represent the degree to which that point belongs to a particular cluster. The Fuzzy c-means (FCM) clustering is one of the most popular and efficient unsupervised partitioned algorithm used in several applications where the clustering is overlapping [19–21]. Moreover, the FCM can handle uncertainty and noise better when there is no information about the structure of the data. The FCM clustering based identification approach proposed by [15] estimates the number of submodels, but the nonlinear regression vector is used to identify the PWARX model. The fuzzy clustering based technique for PWA approximation proposed by [16,18] assumes that the number of submodels is given for the nonlinear systems. In this paper, the FCM clustering algorithm is used which enables an estimation of the number of submodels present in the PWARX models. We show how to estimate the number of submodels using fuzzy cluster validation approach. In this paper, we present new weighted least squares (WLS) approach for identification of parameters of the submodels based on fuzzy distance weight matrix. Once the data points have been classified according to its discrete events a linear regression is used to compute the parameters of the submodels. However, pure linear least squares do not give the optimal value of the parameter due to classification error and sensitive to outliers [8,22]. In order to solve this shortcoming, fuzzy distance weight matrix based weighted least squares (WLS) approach is proposed which suitably defined confidence measures and reducing the influence of outliers on the given data points by means of membership functions. We also verify the effectiveness of the proposed methodology on two benchmark PWARX systems and a Hammerstein model which is composed of static nonlinearity along with linear system.

This paper has been organized as follows: Section 2 represents the formulation of the identification problem for PWARX models. Section 3 provides details of the various phases of the proposed method for the given problem in Section 2. In the same section two illustrative benchmark mathematical examples and a Hammerstein model are used to support the proposed fuzzy weight matrix based parameter identification method. Finally, Section 4 summarizes the contributions of this paper and possible future research work.

2. Problem formulation for PWARX models

Consider the discrete-time HDS with input $u(t) \in R^{n_u}$ and output $y(t) \in R$ at time t having discontinuous dynamics where n_u is the number of inputs. Let $y(t-1)$ and $u(t-1)$ be past value of inputs and outputs, respectively, generated by the system up to time $t-1$. PWARX models are described by a relationship between past observations and future output $y(t)$ in the form given below:

$$y(t) = f(x(t)) + e(t) \tag{1}$$

where $x(t) \in R \subseteq R^p$ is referred to as the regression vector in terms of past observations of input and output, $e(t) \in R$ is a white noise added to the output, and f is a piecewise affine map function of the form

$$f(x(t)) = \theta_{\lambda t}^T \varphi(x(t)) \quad \text{if } x(t) \in \mathcal{R}_{\lambda t} \tag{2}$$

where the output $f(x(t)) \in R$ is a scalar, $\lambda t \in \{1, \ldots, s\}$ is the discrete events associated with the HDS, s is the number of submodels present in the PWARX model, $\theta_{\lambda t}^T \in R^n$ is the parameter vector defining for each λt in terms of matrixes, and $\varphi(x(t)) \in R^m$ is the linear extended regression vector which is a function of $x(t)$ given by $\varphi(x(t)) = [x(t)\, 1]^T$. The regression vector $x(t)$ with the fixed structure depending only on the past n_a outputs and n_b inputs are defined as

$$x(t) = [y(t-1) \ldots y(t-n_a) u(t-1)^T \ldots u(t-n_b)^T]^T \tag{3}$$

where $p = n_a + n_b n_u$.

The polyhedral regions $\mathcal{R}_{\lambda t}$ form a complete bounded polyhedron partition of $\mathcal{R}$ (i.e. $\mathcal{R} = \bigcup_{\lambda t=1}^{s} \lambda t$) and $\mathcal{R}_i \cap \mathcal{R}_j = \phi\, ; \forall\, i \neq j$. The each region $\mathcal{R}_{\lambda t}$ is assumed to be convex polyhedron. The switching law between the PWARX models is specified by the rule: if $x(t) \in \mathcal{R}_{\lambda t}$, the λtth submodel dynamic is active.

Assumption 1. The data points $z(t) = \{y(t), x(t)\}_{t=1}^{N}$ are generated from the PWARX model (1) with given number of discrete regions s, the parameter vectors $\{\theta_{\lambda t}\}_{\lambda t=1}^{s}$ for each submodels, the polyhedral regions $\mathcal{R}_{\lambda t}$ and the dimension p for the regression set with the model orders of the past n_a outputs and n_b inputs.

The data points $z(t) = \{y(t), x(t)\}_{t=1}^{N}$ are generated according to given Assumption 1 and the number of inputs $\{\mathcal{R}_{\lambda t}\}_{\lambda t=1}^{s}$ and $\mathcal{R}$ are known, the identification problem to solve s linear estimation problem for each partition. The regression vector $x(t)$ can be classified into λt sets of partitions $\mathcal{F}_{\lambda t}$ according to rule if $x(t) \in \mathcal{R}_{\lambda t}$, data points $z(t) = \{y(t), x(t)\}_{t=1}^{N} \in \mathcal{F}_{\lambda t}$.

2.1. Problem statement

Given a set of data points $z(t)$ which are generated by a hybrid or piecewise linear system of the form (1) and (2) with Assumption 1 hold, the objective is to estimate the number of submodels s present in the systems and the parameter vectors $\{\theta_{\lambda t}\}_{\lambda t=1}^{s}$ for each submodel along with switching condition based on regions $\mathcal{R}_{\lambda t}$.

The identification of a PWARX model of a finite of data points $z(t) = \{y(t), x(t)\}_{t=1}^{N}$ is a complex problem which requires techniques to estimate the number of discrete modes at the same time classification of the data points according to its discrete events. The considered identification problem consists in finding the PWARX model that involves the estimation for the number of submodels associated with HDS and parameter vector $\theta_{\lambda t}$ of the each affine submodel.

Example 1. Consider the following single-input-single-output (SISO) PWARX model [8]:

$$y(t) = \begin{cases} [1 \quad 2][u(t-1) 1]^T + \in(t) & \text{if } \ u(t-1) = x(t) \in X1 = [-4, \quad 0] \\ [-1 \quad 0][u(t-1) 1]^T + \in(t) & \text{if } \ u(t-1) = x(t) \in X2 = [-1, \quad 2] \\ [1 \quad 2][u(t-1) 1]^T + \in(t) & \text{if } \ u(t-1) = x(t) \in X3 = [2, \quad 4] \end{cases} \tag{4}$$

For the given PWARX system $n_a = 0$, $n_b = 1$ are given a priori and 50 input samples, $u(t) \in R$, are applied randomly according to the uniform distribution on $[-4,4]$ to generate the output samples. Here the regression vector is $x(t) = [u(t-1)]^T$ and data points $z(t) = \{y(t), u(t-1)\}_{t=1}^{50}$. The measured output is corrupted by Gaussian noise with variance 0.05. Fig. 1 shows 50 sample data points of the given system along with the region of partitions.

The problem becomes easy if the number of submodels s is fixed, and the switching conditions are known a priori. The procedure proposed in this paper assumed that the number of submodels present in the given PWARX models and switching mechanism between the submodels are not given a priori. The framework for the identification of the PWARX models is summarized as follows:

Phase 1: Estimation of the number of submodels.

In this phase, the measured inputs and output data points are used to estimate the number of submodels present in the HDS by using proposed fuzzy clustering based index. Initialize this phase with the assumption of maximum number of submodels and phase