Nonlinear Analysis: Hybrid Systems 5 (2011) 242–253

Contents lists available at ScienceDirect

Nonlinear Analysis: Hybrid Systems

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

A recursive identification algorithm for switched linear/affine models

Laurent Bako *, Khaled Boukharouba, Eric Duviella, Stéphane Lecoeuche

Univ Lille Nord de France, F-59000 Lille, France
EMDouai, IA, F-59500 Douai, France

ARTICLE INFO

Article history:
Received 30 November 2009
Accepted 4 May 2010

Keywords:
Switched systems
Piecewise affine systems
System identification
Recursive identification
Open channel systems

ABSTRACT

In this work, a recursive procedure is derived for the identification of switched linear models from input–output data. Starting from some initial values of the parameter vectors that represent the different submodels, the proposed algorithm alternates between data assignment to submodels and parameter update. At each time instant, the discrete state is determined as the index of the submodel that, in terms of the prediction error (or the posterior error), appears to have most likely generated the regressor vector observed at that instant. Given the estimated discrete state, the associated parameter vector is updated based on recursive least squares or any fast adaptive linear identifier. Convergence of the whole procedure although not theoretically proved, seems to be easily achieved when enough rich data are available. It has been also observed that by appropriately choosing the data assignment criterion, the proposed on-line method can be extended to deal also with the identification of piecewise affine models. Finally, performance is tested through some computer simulations and the modeling of an open channel system.

1. Introduction

Given a mixture of data generated by a set of interacting linear/affine discrete-time submodels, the switched affine regression problem refers to the problem of estimating the parameter vectors (PVs) associated with each of these submodels. This is known to be a challenging identification problem as it involves both data assignment to submodels and parameter estimation. It can be fairly observed that a common weakness to the majority of the existing contributions is the computational complexity. For example, feasibility of the optimization approach reported in [1] is restricted to small size problems due to its considerable cost. The Minimum Partition into Feasible Subsystems (Min-PFS) solution proposed in [2] is a multi-step greedy algorithm that is likely to be computationally demanding. The algebraic–geometric method [3–5] embeds the regression data into a higher-dimensional space whose dimension increases exponentially with respect to the dimensions of the considered switched system. The clustering based identification algorithm [6] constructs a neighboring set for each regressor and so, when the number of training data is large, a computational difficulty may arise. In the Bayesian approach [7], each parameter vector is in principle regarded as a random variable and therefore substituted in the identification procedure for its probability density function, which in turn is approximated through a particle filtering model. However this latter approximation comes with a significant increase of computational complexity because fine precision requires the number of particles to be large.

Additionally, most of the published methods for switched systems' identification are batch mode algorithms (see also the survey paper [8] and some more recent techniques presented in [9–13]) except the algebraic algorithm derived in [14,15] and further extended in [16]. However, this latter method inherits from its batch version which first appeared in [3],

* Corresponding author at: Univ Lille Nord de France, F-59000 Lille, France. Tel.: +33 327 712 127.
E-mail addresses: laurent.bako@mines-douai.fr (L. Bako), khaled.boukharouba@mines-douai.fr (K. Boukharouba), eric.duviella@mines-douai.fr (E. Duviella), stephane.lecoeuche@mines-douai.fr (S. Lecoeuche).

1751-570X/$ – see front matter
doi:10.1016/j.nahs.2010.05.003

a problem of dimensionality induced by the polynomial embedding. Beyond the elements of complexity that have been observed so far, note that in general, a batch algorithm has a computational cost that depends more than linearly on the number of training data. Moreover, batch methods are not convenient for real-time applications in which the data need to be processed on-line. It is therefore of interest to develop some recursive identification algorithms for hybrid systems.

In this paper, we present a simple recursive approach to the switched affine regression problem. The data are assumed to be sequentially acquired. Then, starting from some initial values, we proceed alternately to data classification and parameter update on-line. At a given time, the discrete state is inferred based on the information available up to that time and the PVs are accordingly refined via, for example, recursive least squares. In comparison with the batch mode methods mentioned above, it is fair to say that our method, though possibly less effective than some of them, makes it easier to effectively handle higher-dimensional data or larger amounts of data. Furthermore, it can be used for on-line identification of hybrid systems which, to the best of our knowledge, has not received much attention yet. In addition to the references [14–16] mentioned above, the only works which are relevant to the subject of recursive identification for hybrid systems, are the ones reported in [17,18]. The former runs a bank of parallel linear identifiers and uses a set of decision rules for updating, creating and removing submodels. The latter reference deals with the problem of recursively separating observations generated by a mixture of Gaussian distributions. In contrast to this work, we consider here a set of observations stemming from a number of interacting linear ARX submodels. Furthermore, we show that with a simple modification of the data assignment criterion, the proposed method is also applicable without much additional cost to the estimation of continuous piecewise affine maps.

The outline of the paper is as follows. Problem statement is given in Section 2. We describe in Section 3 the proposed algorithm for recursive identification of switched systems with arbitrary switches. In Section 4 we study the particular case when the switched model is intended to approximate a continuous nonlinear system. In this scenario the switching mechanism takes a particular form that can also be inferred from data along with the submodels parameters. Finally, some numerical results are shown and discussed in Section 5.

2. Problem statement

We consider a switched linear model defined by

$$y(t) = \theta_{\lambda_t}^\top x(t) + e_{\lambda_t}(t), \tag{1}$$

where $x(t) \in \mathbb{R}^n$ is referred to as the regression vector, $y(t) \in \mathbb{R}$ is designated as the output of the model, $\lambda_t \in \mathcal{S} = \{1, \ldots, s\}$ is the discrete state and $\theta_{\lambda_t} \in \mathbb{R}^n$ is the associated parameter vector (PV). The sequences of errors $\{e_j(t)\}$, $j = 1, \ldots, s$, are assumed to be zero-mean Gaussian, i.i.d random variables. Here, sequences of errors related to two different submodels are not required to have the same variances. When dealing with the input–output behavior of hybrid dynamical systems, the vector $x(t)$ takes sometimes the form

$$x(t) = \begin{bmatrix} y(t-1) & \cdots & y(t-n_a) & u(t-1)^\top & \cdots & u(t-n_b)^\top \end{bmatrix}^\top \tag{2}$$

where $u(t) \in \mathbb{R}^{n_u}$ and $y(t) \in \mathbb{R}$ are respectively the input and output of the considered system, n_a and n_b are the orders. The model (1) is then designated as a Switched Auto-Regressive eXogenous (SARX) model.

Problem. Given observations $\{x(t), y(t)\}_{t=1}^N$ generated by a switched linear model of the form (1), with $x(t)$ defined as in (2), we are interested here in estimating the parameter vectors (PVs) $\{\theta_j\}_{j=1}^s$ under the assumption that the discrete mode sequence $\{\lambda_t\}_{t=1}^N$ is not known.

We start by recalling from Ref. [19] that the problem of inferring a switched model such as (1) from data, admits multiple solutions so that the identification problem is not well posed. If the structural indexes n_a and n_b are not fixed, then we can find for example a trivial switched model consisting of one single submodel with sufficiently large orders that fits all the finite set of measurements. Even if we assign finite and fixed values to n_a and n_b, there are still infinitely many switched models that explain the data. For example, it can be verified that there is a switched model with $s = N$ submodels that can reproduce the data. In order to alleviate the identifiability issue, we will assume in this paper that the orders n_a and n_b are finite, equal for all submodels and known a priori. We will also assume (whenever s is unknown) that an upper bound $\bar{s} \ll N$ on the number of submodels is available.

3. Algorithm description

In this section we present the main contribution of the paper. From a conceptual viewpoint, the approach to be presented below is comparable to the Bayesian learning algorithm of Juloski et al. [7]. The method derived therein was fairly argued by the authors to allow for the incorporation of prior physical knowledge, when available, in the identification procedure. We will see that the same argument holds also, at least to some extent, for the method suggested in the present paper but with much less computational load. In the Bayesian approach, each parameter vector θ_j is treated as a random variable which is characterized by its probability density function $p_{\theta_j}(\cdot)$. Therefore, instead of estimating directly each θ_j as a single vector, one tries to estimate the density function $p_{\theta_j}(\cdot)$, which is done by approximating it with a particle filtering model. Based on