Systems & Control Letters 62 (2013) 143–151

Contents lists available at SciVerse ScienceDirect

Systems & Control Letters

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

A sparse optimization approach to state observer design for switched linear systems

Laurent Bako*, Stéphane Lecoeuche

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

ARTICLE INFO

Article history:
Received 1 June 2011
Received in revised form
24 July 2012
Accepted 26 November 2012
Available online 27 December 2012

Keywords:
Linear switched systems
Observer design
Sparse optimization
Nonsmooth optimization

ABSTRACT

A new continuous state observer is derived for discrete-time linear switched systems under the assumptions that neither the continuous state nor the discrete state are known. A specificity of the proposed observer is that, in contrast to the state of the art, it does not require an explicit prior estimation of the discrete state. The key idea of the method consists in minimizing a non-smooth ℓ_2-norm-based weighted cost functional, constructed from the matrices of all the subsystems regardless of when each of them is active. In the light of recent development in the literature of compressed sensing, the minimized cost functional has the ability to promote sparsity in a way that makes the knowledge of the discrete mode sequence unnecessary.

1. Introduction

For a dynamic system, the state usually refers to a vector of signals that encodes at each time instant, from a modeling perspective, the full information about the past of that system. There are many practical engineering situations in which an accurate estimate of the state is desirable. Recovering the full state from partial observations has many quite obvious advantages. For example, this can help get around the necessity of instrumenting the system with possibly expensive state sensors. Another application of state estimation is in fault detection. In effect, comparing a model-based estimate of the state to its measured version can bring out model inconsistencies thereby enabling the detection of changes in the system whose nominal behavior is described by that model. Also, in state feedback control systems a complete knowledge of the state is required. Depending on whether past, present or future states are estimated, the terminologies of smoothing, filtering or prediction are respectively used.

Prior work. The state estimation problem has been extensively investigated for the classes of linear and nonlinear dynamic systems. For the class of hybrid systems, the interest of researchers in the state estimation problem is more recent, though a number of relevant approaches have already been published in the existing literature [1–7]. The earliest works on state estimation for hybrid systems were dedicated to the class of Jump Markov Linear Systems [8,6,9,4]. In these systems the switching mechanism is a first-order Markov chain and can be determined from the observations. For such systems the estimation problem is tackled in a stochastic framework using a particle filtering approach [4] or a bank of Kalman filters [6].

The main challenge of hybrid state estimation lies in the fact that the discrete mode is unknown and also needs to be inferred from the input–output measurements along with the continuous state. For switched systems in particular, there is not necessarily a model of the switching law that can be learned from data. For, the switches can in this case be exogenous, deterministic, state driven, event driven, time driven or even totally random. The work reported in [3] which is one of the first proposed approaches for switched linear systems, assumes that the discrete mode sequence is available so that the problem reduces to the synthesis of a classical Luenberger type of observer for each linear subsystem. In [2], the discrete mode is assumed unknown and a receding horizon procedure is presented. A moving horizon approach was also followed in [5,10,11] to derive a state smoother for the class of piecewise affine systems.

Our approach. In this paper we develop a new state observer for discrete-time linear switched systems. Since the discrete state is also unknown, a very natural approach would be to solve a mixed integer-continuous optimization problem for both the discrete mode and the continuous state. A great number of the existing hybrid state estimators for switched linear systems follow this idea. Unfortunately this may be computationally heavy and

* Corresponding author at: Univ Lille Nord de France, F-59000 Lille, France. Tel.: +33 327 712 127; fax: +33 327 712 980.
E-mail addresses: laurent.bako@mines-douai.fr (L. Bako), stephane.lecoeuche@mines-douai.fr (S. Lecoeuche).

0167-6911/$ – see front matter
doi:10.1016/j.sysconle.2012.11.017

sometimes intractable in practice. The main feature of the method introduced in this paper is that it approximates the mixed integer-continuous optimization problem with a non-smooth but still continuous optimization one. Moreover this last formulation of the estimation problem is convex and therefore solvable by efficient and well-documented techniques [12]. Thanks to the interesting properties of the proposed particular cost functional, we can get around the necessity of explicitly estimating the system discrete state. The proposed estimation method can be implemented in two different ways. A first possible implementation method is by resorting to a moving horizon strategy. The second, which will be further developed in the paper, hinges on a recursive and dynamic update of the state using the evidence gradually brought in by new input–output measurements. At each time t, the continuous state is computed in two steps: (1) a prediction step during which the estimate is predicted based on the system matrices and the previous estimate; this is performed through the minimization of an ℓ_2 norm criterion involving all the subsystems' matrices and (2) a correction step during which the measurements at time t are used to refine the predicted state.

Outline. We start with setting up the state estimation problem in Section 2. Section 3 presents the main concept behind the proposed solution in a batch estimation framework. In order to develop a recursive version of that method, we distinguish between two cases: the case when the modes are instantaneously discernible, is treated first in Section 4; the more general situation where the modes are discernible over a finite data-window, is dealt with in Section 5. Numerical experiments are depicted in Section 6. Some concluding remarks are provided in Section 7.

2. Problem formulation

We consider a discrete-time switched linear system described in state space form by

$$\begin{cases} x(t+1) = A_{\sigma(t)}x(t) + B_{\sigma(t)}u(t) \\ y(t) = C_{\sigma(t)}x(t) + D_{\sigma(t)}u(t) \end{cases} \tag{1}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^{n_u}$, $y(t) \in \mathbb{R}^{n_y}$ are respectively the state, the input and the output of the system at time $t \in \mathbb{N}$. $\sigma(t) \in Q$ is the discrete mode (state), that is, the index of the subsystem which is active at time t; $Q = \{1, \ldots, s\}$ is the finite set of discrete modes with cardinality $|Q| = s$. For each discrete mode $i \in Q$, A_i, B_i, C_i, D_i are the matrices (of appropriate dimensions) which are associated with that mode. The model (1) can be subject to possible disturbances. Although no special treatment of such disturbances will be made here, the method to be presented remains applicable to some extent.

Given the system matrices A_i, B_i, C_i, D_i, $i \in Q$, and the past input–output observations $(u(\tau), y(\tau))$, $\tau = 0, \ldots, t-1$, the problem discussed in this paper is concerned with the estimation of the continuous state $x(t)$ at time t under the assumption that the discrete state sequence is unknown. We more specifically try to answer the question of whether it is possible to reconstruct the continuous state without an explicit knowledge/estimate of the discrete mode sequence.

3. A batch mode estimation of the state

Most conventional approaches for hybrid observer design are based, at least in a first step, on an explicit attempt to recover the discrete state sequence. The motivation for such a treatment of the state estimation problem is that linear techniques can directly be carried over to hybrid systems once the discrete state is known. However, finding simultaneously both the discrete and continuous states is a problem that is partly combinatorial (mixed integer-continuous programming). Even worse, deciding the discrete state becomes much more problematic when the data are contaminated by noise.

This paper proposes a conceptually different approach in that it is able to overcome the need of explicitly recovering the discrete state before proceeding with the estimation of the continuous state. Instead, the designed observer relies on an appropriately weighted continuous and non-smooth optimization. To discuss clearly the foundation of the method, it is perhaps interesting to start with batch mode estimation.

We first consider a batch estimation of the state that is, an off-line estimation after a finite number of input–output measurements has been collected. The main idea is based on the following observation. If we denote by $\{x(t)\}_{t=0}^N$ the true state sequence, then the vector sequence $\{h_i(t), i = 1, \ldots, s, t = 0, \ldots, N-1\}$ with

$$h_i(t) = \begin{bmatrix} x(t+1) \\ y(t) \end{bmatrix} - \begin{bmatrix} A_i & B_i \\ C_i & D_i \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \tag{2}$$

forms (to some extent) a sparse sequence of vectors that is, a great number of elements of this sequence are equal to zero.[1] More precisely, if at any time t there is only one discrete mode $i \in \{1, \ldots, s\}$ satisfying the system Eq. (1), then exactly N elements of the sequence $\{h_i(t)\}_{i,t}$ are equal to zero. As a consequence of this observation, we can naturally search for the states by minimizing the number of nonzero vectors in the sequence $\{h_i(t)\}_{i,t}$. This translates into the so-called sparse optimization problem

$$\text{Minimize } \mathbf{card}\big\{(i,t) : h_i(t) \neq 0\big\}, \tag{3}$$

where **card** stands for cardinality. Minimizing directly the cardinality of a set is an NP-hard and nonconvex optimization problem [13] which is intractable in practice. It is hence necessary to relax the problem (3) into a convex one. To this end, we start by noting that sparsity of the vector sequence $\{h_i(t)\}_{i,t}$ is, in principle, equivalent to sparsity of the scalar sequence $\{\|h_i(t)\|\}_{i,t}$ for any norm $\|\cdot\|$. However, from an optimization standpoint, the nature of the norm might have an effect on the expected outcome. Since we want the entire vector $h_i(t)$ (and not only some of its components) to be equal to zero whenever possible, the ℓ_2-norm appears to be more suitable here. Capitalizing on this observation, we may relax the optimization problem (3) into the one of finding the estimate $\{\hat{x}(t)\}$ of the state sequence so as to minimize the cost functional

$$J\big(\hat{x}(0), \ldots, \hat{x}(N)\big) = \sum_{t=0}^{N-1} \sum_{i=1}^{s} w_i(t) \times \left\| \begin{bmatrix} \hat{x}(t+1) \\ y(t) \end{bmatrix} - \begin{bmatrix} A_i & B_i \\ C_i & D_i \end{bmatrix} \begin{bmatrix} \hat{x}(t) \\ u(t) \end{bmatrix} \right\|_2 \tag{4}$$

where the $w_i(t)$ are some (given) positive weights and N denotes the number of input–output samples. The choice of this criterion calls for some comments. First, note that the criterion is constructed as a sum-of-norms instead of a sum-of-squared-norms as is usually the case in estimation literature. This is motivated by the fact that the non-smooth sum-of-norms is recognized to promote the obtention of a solution $\{\hat{x}(t)\}$ such that the corresponding vector sequence $\{\hat{h}_i(t)\}_{i,t}$ is sparse [14–16]. While minimizing a sum-of-squared-norms cost function (which corresponds to the classical least squares) tends to make the average error vector small, the cost function in (4) can tolerate the presence of a few

[1] The notion of sparsity is very broad here; it does not mean necessarily that the majority of the vectors contained in the considered sequence is equal to zero.