Automatica 51 (2015) 233–242

Contents lists available at ScienceDirect

Automatica

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

A graph theoretical approach to input design for identification of nonlinear dynamical models ☆

Patricio E. Valenzuela, Cristian R. Rojas, Håkan Hjalmarsson

Department of Automatic Control and ACCESS Linnaeus Center, School of Electrical Engineering, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 13 November 2013
Received in revised form
6 June 2014
Accepted 26 September 2014
Available online 6 November 2014

Keywords:
System identification
Input design
Markov chains

ABSTRACT

In this paper the problem of optimal input design for model identification is studied. The optimal input signal is designed by maximizing a scalar cost function of the information matrix, where the input signal is a realization of a stationary process with finite memory, with its range being a finite set of values. It is shown that the feasible set for this problem can be associated with the prime cycles in the graph of possible values and transitions for the input signal. A realization of the optimal input signal is generated by running a Markov chain associated with the feasible set, where the transition matrix is built using a novel algorithm developed for de Bruijn graphs. The proposed method can be used to design inputs for nonlinear output-error systems, which are not covered in previous results. In particular, since the input is restricted to a finite alphabet, it can naturally handle amplitude constraints. Finally, our approach relies on convex optimization even for systems having a nonlinear structure. A numerical example shows that the algorithm can be successfully used to perform input design for nonlinear output-error models.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Input design considers the construction of an input signal to maximize the information obtained from an experiment. Some of the initial contributions in this line were presented in the works of Cox (1958), Fedorov (1972), and Goodwin and Payne (1977), where the latter contribution is concerned with input design for the identification of dynamic systems. Since then, several contributions in input design have been developed (see Gevers (2005), Hildebrand and Gevers (2003), Whittle (1973), and the references therein).

In the case of dynamic systems, input design maximizes the information related to the estimated parameters of the system. By maximizing a scalar function of the Fisher information matrix (Ljung, 1999) related to the accuracy of the estimated model for a particular application, we obtain an input signal that can be used to identify a good application model of the unknown system. The results in this area are mainly focused on input design for linear systems, where powerful tools can be applied to solve the problem (Goodwin, Murdoch, & Payne, 1973; Jansson & Hjalmarsson, 2005; Lindqvist & Hjalmarsson, 2000; Ljung, 1999; Rojas, Welsh, Goodwin, & Feuer, 2007). Several methods have been reported in the literature involving, e.g., linear matrix inequalities (LMI) (Jansson & Hjalmarsson, 2005; Lindqvist & Hjalmarsson, 2000; Sanchez, Rojas, Vandersteen, Bragos, & Schoukens, 2012; Wahlberg, Hjalmarsson, & Stoica, 2010), Markov chains (Brighenti, 2009; Brighenti, Wahlberg, & Rojas, 2009), and time domain gradient based schemes (Goodwin et al., 1973; Suzuki & Sugie, 2007), among others.

In recent years, the interest in input design has shifted from linear to nonlinear systems. Unfortunately, most of the tools used for input design for linear systems based on frequency domain techniques are no longer valid for the nonlinear case, which implies that new techniques need to be developed in this domain. One approach to input design for the identification of nonlinear systems is introduced in Hjalmarsson and Mårtensson (2007), where a linear systems perspective is considered. Extensions to a class of finite-impulse-response type systems are developed in Larsson, Hjalmarsson, and Rojas (2010), where a characterization of probability density functions is employed. Input design for structured nonlinear identification is introduced in Vincent, Novara, Hsu, and Poola (2009) and Vincent, Novara, Hsu, and Poolla (2010), where the

☆ This work was supported by the Swedish Research Council under contracts 621-2011-5890 and 621-2009-4017, and by the European Research Council under the advanced grant LEARN, contract 267381. The material in this paper was partially presented at the 52nd IEEE Conference on Decision and Control (CDC), December 10–13, 2013, Florence, Italy. This paper was recommended for publication in revised form by the Associate Editor Andrea Garulli under the direction of the Editor Torsten Söderström.

E-mail addresses: pva@kth.se (P.E. Valenzuela), crro@kth.se (C.R. Rojas), hjalmars@kth.se (H. Hjalmarsson).

http://dx.doi.org/10.1016/j.automatica.2014.10.097
0005-1098/© 2014 Elsevier Ltd. All rights reserved.

system is assumed to be an interconnection of known linear systems and unknown static nonlinearities. An input design method for a general class of nonlinear systems is presented in Gopaluni, Schön, and Wills (2011), based on a particle filter used to approximate the cost function, which is optimized over a particular class of input vectors using stochastic approximation. The methods previously mentioned (Gopaluni et al., 2011; Hjalmarsson & Mårtensson, 2007; Larsson et al., 2010) in general are highly complex (usually ending up with non-convex optimization formulations, (e.g., Gopaluni et al., 2011) and are restricted to particular model structures (e.g., Hjalmarsson and Mårtensson (2007) and Larsson et al. (2010)) and/or particular classes of input signals (e.g., white noise filtered through an ARX filter, Gopaluni et al., 2011). Moreover, except for the results in Brighenti (2009), Brighenti et al. (2009) and Larsson et al. (2010), the methods introduced cannot handle input design with amplitude constraints. Amplitude constraints can arise due to power and/or physical limitations in the system. Therefore, input design with amplitude constraints also requires further considerations.

As a first contribution, in this article we develop a novel approach for input design in nonlinear systems. This approach considers the design of an input sequence for models with additive white noise at the output, which extends the class of nonlinear systems considered in Larsson et al. (2010). The input is constrained to be a stationary process with a finite set of possible values, and where the associated probability mass function (pmf) has finite memory, i.e., a Markov chain of fixed order. Therefore, the optimization considers the design of an optimal pmf which maximizes the information obtained from the experiment, quantified as a scalar function of the information matrix. By using notions of graph theory, we can express the set of feasible pmfs as a convex combination of the pmfs of the prime cycles describing the vertices of the set. Since the prime cycles can be explicitly computed by known algorithms (Johnson, 1975; Zaman, 1983), the optimization problem becomes easy to pose. Furthermore, for standard choices of the cost function, the problem is convex even for nonlinear systems, which simplifies the problem formulation discussed in Brighenti (2009) and Brighenti et al. (2009). Finally, since the input is restricted to a finite set of possible values, the method naturally incorporates amplitude limitations.

Once the optimization problem is solved, we obtain the optimal stationary distribution over the possible states of the memory describing the pmf. To obtain an input with the desired stationary distribution, we must be able to design a feasible transition probability matrix satisfying the constraints of the graph associated with our problem. Unfortunately, due to the asymmetric structure of the graph, we cannot use standard Markov chain Monte Carlo (MCMC) methods (Boyd, Diaconis, & Xiao, 2004; Hastings, 1970) to determine a transition matrix for the graph. Therefore, and as a second contribution of this paper, we develop a method to design a valid transition probability matrix for graphs generated from stationary processes with finite memory.

The present article can be seen as an extension of the results in Brighenti (2009), Brighenti et al. (2009) and Larsson et al. (2010). The main difference with Brighenti (2009) and Brighenti et al. (2009) is that we optimize over the stationary pmf associated with the Markov chain, instead of directly optimizing over the transition probabilities. This approach results in a convex problem (which cannot be achieved in Brighenti (2009) and Brighenti et al. (2009), where optimization techniques guaranteeing local optima must be employed). In Larsson et al. (2010) a similar approach to the one presented in our article is discussed, but restricted to the analysis to nonlinear FIR systems. By using the finite memory property of nonlinear FIR models, the input design problem in Larsson et al. (2010) is solved in terms of an input realization of finite length. However, the results in Larsson et al. (2010) cannot be employed to design input sequences for identification of more general nonlinear output-error models, since the models will generally depend on the entire past input sequence. In this line, our article extends the analysis to more general nonlinear model structures, which includes nonlinear FIR systems (see Example 1 in Valenzuela, Rojas, and Hjalmarsson (2013) where the results are consistent with those introduced in Larsson et al. (2010)).

As with most optimal input design methods, the one proposed in this contribution relies on knowledge of the true system. This difficulty can be overcome by implementing a robust experiment design scheme on top of it (Rojas et al., 2007) or via an adaptive procedure, where the input signal is re-designed as more information is being collected from the system (Gerencsér, Hjalmarsson, & Mårtensson, 2009; Rojas, Hjalmarsson, Gerencsér, & Mårtensson, 2011). This issue goes beyond the scope of this article and it will not be addressed here.

A previous description of the proposed method has been presented in Valenzuela et al. (2013). In this paper we give a more detailed explanation of the input design technique, a method of the generation of the input signal from an optimal finite state Markov chain stationary distribution, and new numerical examples.

The rest of the paper is organized as follows. Section 2 introduces some background on graph theory. Section 3 presents the input design problem. In Section 4 we solve the input design problem using elements of graph theory. Section 5 presents a novel method to generate an input signal from the optimal stationary distribution obtained in Section 4. Section 6 illustrates the results with numerical examples. Finally, Section 7 presents conclusions.

Notation. In the sequel, we denote by $\mathbb{C}$ the complex set, by $\mathbb{Z}$ the integer set, by $\mathbb{R}$ the real set, by $\mathbb{R}^p$ the set of real p-dimensional vectors, and by $\mathbb{R}^{r\times s}$ the set of real $r \times s$ matrices. Given $z \in \mathbb{C}$, $|z|$ denotes its modulus. The expected value with respect to the random variable x and the probability measure are denoted by $\mathbf{E}_x\{\cdot\}$, and $\mathbf{P}\{\cdot\}$, respectively. det and tr stand for the determinant and the trace functions, respectively. Given a finite set T, $\#T$ denotes its cardinality.

2. Preliminaries on graph theory

In this section we provide a brief background on the concepts of graph theory used in the next sections. Our notation follows that of Johnson (1975, pp. 77).

A *directed graph* $\mathcal{G}_{\mathcal{V}} = (\mathcal{V}, \mathcal{X})$ consists of a nonempty and finite set of vertices (or nodes) $\mathcal{V}$ and a set $\mathcal{X}$ of ordered pairs of distinct vertices called *edges*. A *path* in $\mathcal{G}_{\mathcal{V}}$ is a sequence of vertices $p_{vu} = (v = v_1, v_2, \dots, v_k = u)$ such that $(v_i, v_{i+1}) \in \mathcal{X}$ for all $i \in \{1, \dots, k-1\}$. A *cycle* is a path in which the first and last vertices are identical. A cycle is *elementary* if no vertex but the first and last appears twice. Two elementary cycles are distinct if one is not a cyclic permutation of the other.

An n-dimensional *de Bruijn graph* of m symbols (de Bruijn & Erdos, 1946) is a directed graph representing overlaps between sequences of symbols (cf. Fig. 2). It has m^n vertices, consisting of all possible sequences of length n derived from the given symbols. The same symbol can appear multiple times in a sequence. If we have a set of symbols $\mathcal{C} = \{s_1, \dots, s_m\}$ then the set of n-dimensional vertices is

$$\begin{aligned}\mathcal{V} = \mathcal{C}^n = \{&(s_1, \dots, s_1, s_1), (s_1, \dots, s_1, s_2), \dots,\\ &(s_1, \dots, s_1, s_m), (s_1, \dots, s_2, s_1), \dots,\\ &(s_m, \dots, s_m, s_m)\}. \end{aligned} \tag{1}$$

If one of the vertices can be expressed as another vertex by shifting all its symbols one place to the left and adding a new symbol at the

Download English Version:

https://daneshyari.com/en/article/7110036

Download Persian Version:

https://daneshyari.com/article/7110036

Daneshyari.com