Automatica 92 (2018) 86–91

Contents lists available at ScienceDirect

Automatica

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

Brief paper

Optimal input design for parameter estimation in a bounded-error context for nonlinear dynamical systems ☆

Carine Jauberthie [a,*], Lilianne Denis-Vidal [b], Qiaochu Li [b], Zohra Cherfi-Boulanger [c]

[a] LAAS-CNRS, Université de Toulouse, UPS, Toulouse, France
[b] Sorbonne University, Université de technologie de Compiègne, LMAC Laboratory of Applied Mathematics of Compiègne - CS 60 319 - 60 203 Compiègne cedex, France
[c] Sorbonne University, Université de technologie de Compiègne, CNRS, UMR 7337 Roberval, Centre de recherche Royallieu - CS 60 319 - 60 203 Compiègne cedex, France

ARTICLE INFO

Article history:
Received 17 December 2015
Received in revised form 16 October 2017
Accepted 4 January 2018

Keywords:
Experiment design
Optimal input
Parameter estimation
State estimation
Continuous-time systems
Sensitivity functions
Nonlinear systems
Bounded noise
Interval analysis
Aerospace models

ABSTRACT

This paper deals with optimal input design for parameter estimation in a bounded-error context. Uncertain controlled nonlinear dynamical models, when the input can be parametrized by a finite number of parameters, are considered. The main contribution of this paper concerns criteria for obtaining optimal inputs in this context. Two input design criteria are proposed and analyzed. They involve sensitivity functions. The first criterion requires the inversion of the Gram matrix of sensitivity functions. The second one does not require this inversion and is then applied for parameter estimation of a model taken from the aeronautical domain. The estimation results obtained using an optimal input are compared with those obtained with an input optimized in a more classical context (Gaussian measurement noise and parameters a priori known to belong to some boxes). These results highlight the potential of optimal input design in a bounded-error context.

1. Introduction

In this paper, we are interested in parameter estimation of models describing industrial systems. Such systems are often prone to uncertainties that complicate the modeling task. Usually, uncertainties are described as realizations of random variables with known distributions, which is difficult to justify in practice. In the presented work, perturbations are only assumed to be bounded with known bounds. Thus, the set-membership framework is considered. In this framework, the set of all parameters consistent with the model structure, the measurements and the bounds on the perturbations can be defined as the set estimate for the parameters. Various techniques are then available to characterize this set estimate (see for example Jaulin, Kieffer, Dirit, & Walter, 2001 or Kieffer & Walter, 2011).

Set-membership estimation is an interesting alternative to classical least squares or maximum likelihood estimation. These methods have received a lot of attention in the last years, for example Jaulin (2009), Kieffer, Jaulin, and Walter (2002) and Rauh and Auer (2011).

Set-membership parameter estimation was first considered for models whose output is linear in their parameters, then models nonlinear in their parameters were considered. During the last decade, models described by nonlinear state equations have been considered in this context (Johnson & Tucker, 2008; Raïssi, Ramdani, & Candau, 2004).

Experiment design is important to identify more precisely mathematical models of complex systems. The overall goal is to design an experiment that produces data from which model parameters can be estimated accurately. The conventional approach for experiment design assumes stochastic models for uncertain parameters and measurement errors (see for example Rojas, Welsh, Goodwin, & Feuer, 2007). Several criteria for experiment design have been proposed involving a scalar function of the Fisher information matrix (FIM). For example the A-optimal experiment minimizes the trace of inverse of the FIM, which minimizes, in linear case, the average variance of the estimates. Another criterion widely used is the D-optimality. The D-optimal experiment

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Cristian R. Rojas under the direction of Editor Torsten Söderström.

* Corresponding author.
E-mail addresses: cjaubert@laas.fr (C. Jauberthie), lilianne.denis-vidal@utc.fr (L. Denis-Vidal), qiaochu.li@utc.fr (Q. Li), zohra.cherfi@utc.fr (Z. Cherfi-Boulanger).

https://doi.org/10.1016/j.automatica.2018.03.003

minimizes the volume of a confidence ellipsoid. However, some sources of uncertainty are better modeled as bounded uncertainty. This is the case of parameter uncertainties that generally arise from design tolerances and from aging (see for example Travé-Massuyès, Pons, Ribot, Pencolé, & Jauberthie, 2015). In a bounded-error context, the experiment design is much less studied. First result consists in designing experiments which minimize the volume of the estimate of parameter domain. Some works such as Belforte and Gay (2000) or Pronzato and Walter (1988) for models linear with respect to the input, consist in optimizing the worst possible performance of the experiment over the prior domain for the parameters. In Pronzato and Walter (1988), a minimax approach to synthetize the optimal experiment is described, using the Gram matrix of sensitivity functions and specific criteria are developed. These approaches take into account the bounds of the prior domain for the parameters into the search of the optimal experiment but do not take into account the set-membership estimation process which leads to the set estimate.

In this paper, to minimize the set of estimate of parameters, we exhibit an explicit expression linking this set of parameters with the Gram matrix of sensitivities. This work follows a study on the optimization of the initial conditions in the same context but with a different approach (Li, Jauberthie, Denis-Vidal, & Cherfi, 2016). To obtain an explicit expression of the set of parameters to be estimated, in Kieffer and Walter (2011), the authors have used a centered inclusion function (which is a set-membership extension of the equality obtained by the Mean Value Theorem) for the model output and they have built an operator of contraction for the set of parameter to be estimated based on sensitivity functions. Starting from this idea, we build explicitly some criteria to find an optimal experiment in the bounded-error context. In our work, we consider only the optimal input design. The proposed methodology requires a parametrization of the input using elementary functions with a reasonable number of parameters.

This paper is organized as follows. In Section 2, the problem statement is presented. Section 3 describes some basic tools of interval analysis. Section 4 introduces the proposed criteria for optimal input design. An aerospace application is given in Section 5. In Section 6, some conclusions and future research directions are discussed.

2. Problem formulation and notations

2.1. Notations

In what follows, boxes, i.e., cartesian products of intervals are denoted as $[\mathbf{x}] = [\underline{\mathbf{x}}, \overline{\mathbf{x}}]$ (see Jaulin et al., 2001). This paper deals with optimal input design for estimating the unknown parameters of a nonlinear dynamical model described by the following form:

$$\begin{cases} \dot{\mathbf{x}}(t,\mathbf{p},u) = \mathbf{f}(\mathbf{x}(t,\mathbf{p},u), u(t), \mathbf{p}), \\ \mathbf{y}_m(t,\mathbf{p},u) = \mathbf{h}(\mathbf{x}(t,\mathbf{p},u), \mathbf{p}), \end{cases} \tag{1}$$

where $\mathbf{x}(t,\mathbf{p},u) \in \mathbb{R}^{n_x}$ and $\mathbf{y}_m(t,\mathbf{p},u) \in \mathbb{R}^{n_y}$ denote respectively the vectors of state variables and the model output. The initial conditions for $\mathbf{x}(.)$ at $t = 0$ are supposed to belong to an initial bounded box $[\mathbf{x}_0]$. $u(t)$ represents the input, it is supposed to belong to an admissible set of inputs U_{ad} and the input is supposed to be composed of elementary functions.

The vector $\mathbf{p} \in \mathbb{R}^{n_p}$ is the vector of parameters to be estimated, which is supposed to belong to an a priori box $[\mathbf{p}_0]$.

The time t is assumed to belong to $[0, t_{\max}]$.

The functions $\mathbf{f}$ and $\mathbf{h}$ are nonlinear functions.

$\mathbf{f}$ is supposed analytic on M for every $\mathbf{p} \in [\mathbf{p}_0]$, where M is an open set of $\mathbb{R}^{n_x}$ such that $\mathbf{x}(t,\mathbf{p},u) \in M$ for every $\mathbf{p} \in [\mathbf{p}_0]$ and $t \in [0, t_{\max}]$.

The model output at the sample time t_k, with k from 1 to N is denoted $\mathbf{y}_m^k(\mathbf{p},u) = \mathbf{y}_m(t_k,\mathbf{p},u)$. $\mathbf{y}_m^k(\mathbf{p},u)$ is a vector with components $y_{m,i}^k(\mathbf{p},u) = y_{m,i}(t_k,\mathbf{p},u)$ for $i = 1,\ldots,n_y$, $k = 1,\ldots,N$.

Let $\mathbf{y}(t_k,u)$ be the vector of the measurements at the sample time t_k. Suppose that there exists a "true" value of parameters $\mathbf{p}^*$ such that we have:

$$\mathbf{y}(t_k,u) = \mathbf{y}_m(t_k,\mathbf{p}^*,u) + \mathbf{v}(t_k), \quad k = 1,\ldots,N, \tag{2}$$

where the measurement noise $\mathbf{v}(t_k)$ is supposed bounded by $\underline{\mathbf{v}}(t_k)$ and $\overline{\mathbf{v}}(t_k)$ which are known as lower and upper bounds. Such bounds may, for instance, correspond to a bounded measurement noise or tolerance on sensors.

2.2. Parameter estimation in a bounded error context

In a bounded-error estimation context, one is interested in estimating the set $\mathbb{P} \subset [\mathbf{p}_0]$ of all parameters $\mathbf{p}$ consistent with the model structure and the bounds on the measurement noise. In order to obtain the most accurate estimates, we choose to minimize a cost function, for example the volume, of the set $\mathbb{P}$ (or of an enclosure of $\mathbb{P}$). It may generally depend on the values of the input, the initial time, the sample times, among others. In this work, only the input is considered. Our aim is to design an input that minimizes the cost function. More formally, one has to find an input u^* such that:

$$u^* = \arg\min_{u \in U_{ad}} \Phi(\mathbb{P}). \tag{3}$$

Obtaining $\mathbb{P}$ is difficult in practice. Nevertheless there are efficient algorithms to obtain an outer-approximation $[\mathbf{p}]$ of $\mathbb{P}$. The Problem (3) is thus relaxed as follows:

$$u^* = \arg\min_{u \in U_{ad}} \Phi([\mathbf{p}]),$$

where $[\mathbf{p}]$ is an outer-approximation of $\mathbb{P}$ obtained by a bounded-error estimation algorithm from interval analysis.

The next section briefly describes the tools from interval analysis used to perform the set-membership estimation.

3. Basic tools of interval analysis

Interval analysis provides tools for computing with sets which are described using outer-approximations formed by union of non-overlapping boxes. The following results are mainly taken from Jaulin et al. (2001) and Moore (1966).

A real interval $[u] = [\underline{u}, \overline{u}]$ is a closed and connected subset of $\mathbb{R}$. The width of an interval $[u]$ is defined by $w([u]) = \overline{u} - \underline{u}$, and its midpoint by $m([u]) = (\overline{u} + \underline{u})/2$.

An interval vector (or box) $[\mathbf{x}]$ is a vector with interval components and may equivalently be seen as a cartesian product of scalar intervals $[\mathbf{x}] = [x_1] \times [x_2] \ldots \times [x_n]$. An interval matrix is a matrix with interval components. The set of $n \times m$ real interval matrices is denoted by $\mathbb{IR}^{n\times m}$. The width of an interval vector (or of an interval matrix) is the maximum of the widths of its interval components. The midpoint of an interval vector (resp. an interval matrix) is a vector (resp. a matrix) composed of the midpoints of its interval components.

Classical operations for interval vectors (resp. interval matrices) are direct extensions of the same operations for scalar vectors (resp. scalar matrices) (Moore, 1966).

The magnitude of an interval $[x]$, noted $|[x]|$ is given by the largest absolute value of $[x]$ that means the absolute value of the real with the largest value in $[x]$. The mignitude of an interval $[x]$, noted $\mathrm{mig}([x])$ is the smallest absolute value of the elements of $[x]$ that means the absolute value of the real with the smallest value in $[x]$. In this work, the notions of magnitude and mignitude of