Available online at www.sciencedirect.com

ScienceDirect

Automatica 43 (2007) 993–1008

automatica

www.elsevier.com/locate/automatica

Robust optimal experiment design for system identification☆

Cristian R. Rojas[a], James S. Welsh[a,*], Graham C. Goodwin[a], Arie Feuer[b]

[a]*School of Electrical Engineering and Computer Science, The University of Newcastle, Callaghan, NSW 2308, Australia*
[b]*Department of Electrical Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel*

Received 27 November 2005; received in revised form 3 August 2006; accepted 4 December 2006

Abstract

This paper develops the idea of min–max robust experiment design for dynamic system identification. The idea of min–max experiment design has been explored in the statistics literature. However, the technique is virtually unknown by the engineering community and, accordingly, there has been little prior work on examining its properties when applied to dynamic system identification. This paper initiates an exploration of these ideas. The paper considers linear systems with energy (or power) bounded inputs. We assume that the parameters lie in a given compact set and optimise the worst case over this set. We also provide a detailed analysis of the solution for an illustrative one parameter example and propose a convex optimisation algorithm that can be applied more generally to a discretised approximation to the design problem. We also examine the role played by different design criteria and present a simulation example illustrating the merits of the proposed approach.

Keywords: Robust experiment design; Optimal input design

1. Introduction

The goal of experiment design is to adjust the experimental conditions so that maximal information is gained from the experiment. Background to this problem can be found in early statistics literature (Cox, 1958; Fedorov, 1972; Karlin & Studden, 1966; Kempthorne, 1952; Kiefer & Wolfowitz, 1960; Wald, 1943; Whittle, 1973; Wynn, 1972) as well as in the engineering literature (Arimoto & Kimura, 1973; Gagliardi, 1967; Goodwin, Murdoch, & Payne, 1973; Goodwin & Payne, 1973; Goodwin, Payne, & Murdoch, 1973; Goodwin & Payne, 1977; Hildebrand & Gevers, 2003a; Levadi, 1966; Mehra, 1974; Zarrop, 1979). A recent survey is contained in Gevers (2005) where many additional references can be found. The focus in the engineering literature has been predominately on experiment design for dynamic system identification.

A key issue with experiment design for dynamic systems is that the model is typically nonlinearly parameterised. This means, amongst other things, that the Fisher information matrix (Goodwin & Payne, 1977, p. 6) which is typically used as the basis for experiment design, depends, inter alia, on the true system parameters (i.e. the nominal optimal experiment depends on the very thing that the experiment is aimed at finding).

This issue has been recognised in the statistics literature where several approaches have been explored. These include:

- Sequential design, where one iterates between parameter estimation, on the one hand, and experiment design using the current parameter estimates, on the other, see Chernoff (1975), Ford and Silvey (1980), Ford, Titterington, and Wu (1985), Müller and Pötscher (1992), Walter and Pronzato (1997), and Wu (1985).
- Bayesian design (Atkinson, Chaloner, Juritz, & Herzberg, 1993; Atkinson & Doner, 1992; Chaloner & Larntz, 1989; Chaloner & Verdinelli, 1995; El-Gamal & Palfrey, 1996; Sebastiani & Wynn, 2000). The Bayesian approach is characterised by the minimisation of the expected value (over the prior parameter distribution) of a local optimality criterion related to the information matrix.
- Min–max design (Biedermann & Dette, 2003; D'Argenio & Van Guilder, 1988; Dette, Melas, & Pepelyshev, 2003; Fedorov, 1980; Landaw, 1984; Melas, 1978; Pronzato & Walter, 1988).

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Johan Schoukens under the direction of Editor Torsten Söderström.

* Tel.: +61 2 49216087; fax: +61 2 49216993.
E-mail address: james.welsh@newcastle.edu.au (J.S. Welsh).

0005-1098/$ - see front matter
doi:10.1016/j.automatica.2006.12.013

However, there has been little work on robust experiment design for engineering problems. This has been highlighted in the recent survey paper (Hjalmarsson, 2005, p. 427) where it is stated that "…as usual in experiment design, in order to compute the optimal design the true system has to be known. Methods that are robust with respect to uncertainty about the system is a wide open research field."

Preliminary work in the engineering literature on robust experiment design includes substantial work on iterative design (Gevers, 2005; Hjalmarsson, 2005) and an insightful sub-optimal min–max solution for a one parameter problem (Walter & Pronzato, 1997, p. 339). Actually the latter problem will be discussed in detail in Section 3 of the current paper. Also, a number of very recent engineering papers refer to the idea of min–max optimal experiment design—see for example papers presented at SYSID'06, e.g., Gevers and Bombois (2006), Goodwin, Welsh, Feuer, and Derpich (2006), and Mårtensson and Hjalmarsson (2006).

Our goal in the current paper is to develop the idea of min–max optimal experiment design for dynamic system identification. To gain insight into this approach, we explore an illustrative example in depth.

We assume prior knowledge in the form that the system parameters, θ, are contained in a given compact set Θ. We then choose a design criterion $f(M(\theta),\theta)$ where $M(\theta)$ is the Fisher information matrix, evaluated at θ, and design the experiment to optimise the worst case of $f(M(\theta),\theta)$ over Θ. Notice that this differs from the usual approaches to experiment design *in the engineering literature* which typically optimise $f(M(\theta_0),\theta_0)$ for some given nominal value θ_0.

Our approach is more akin to the usual formulation of robust optimal control which typically considers the worst case (Zhou, Doyle, & Glover, 1996). Indeed, there are substantial links between the work presented here and continuous game theory (Başar & Bernhard, 1995; Başar & Olsder, 1995; Fudenberg & Tirole, 1991; Owen, 1995; Szép & Forgó, 1985). We explore some of these connections below.

The merits of the approach proposed in this paper are illustrated by an example (presented in Section 5) which shows, for a realistic second order system, that an order of magnitude improvement in the worst case performance in experiment design can be achieved at the expense of only a few percent degradation in the nominal performance.

The layout of the remainder of the paper is as follows: in Section 2 we give a general formulation of the min–max approach to robust optimal experiment design. Section 3 explores an illustrative one parameter example in considerable detail so as to give insight into the problem. In Section 4 we describe the extension to multi-parameter systems. In Section 5 we present several results illustrating the merits of the proposed approach. Finally, in Section 6 we draw conclusions.

2. Experiment design criteria

2.1. The information matrix

So as to be specific we first consider a single input single output linear discrete time system, with input $\{u_t\}$ and output $\{y_t\}$, of the form

$$y_t = G_1(q)u_t + G_2(q)w_t,$$

where G_1 and G_2 are rational transfer functions, q is the forward shift operator, $G_2(\infty)=1$, and $\{w_t\}$ is zero mean Gaussian white noise of variance Σ. We let $\beta \triangleq [\theta^{\mathrm{T}}, \gamma^{\mathrm{T}}, \Sigma]^{\mathrm{T}}$ where θ denotes the parameters in G_1 and γ denotes the parameters in G_2.

We recall that the log likelihood function (Goodwin & Payne, 1977, p. 130) for data Y given parameters β, is

$$\ln p(Y|\beta) = -\frac{N}{2}\ln 2\pi - \frac{N}{2}\ln\Sigma - \frac{1}{2\Sigma}\sum_{t=1}^{N}\varepsilon_t^2, \tag{1}$$

where

$$\varepsilon_t \triangleq G_2(q)^{-1}[y_t - G_1(q)u_t]. \tag{2}$$

Fisher's information matrix is obtained by taking the following expectation (Goodwin & Payne, 1977, p. 130):

$$M \triangleq \mathbb{E}_{Y|\beta}\left[\left(\frac{\partial \ln p(Y|\beta)}{\partial\beta}\right)\left(\frac{\partial \ln p(Y|\beta)}{\partial\beta}\right)^T\right], \tag{3}$$

where from (1)

$$\frac{\partial \ln p(Y|\beta)}{\partial\beta} = -\frac{1}{\Sigma}\sum_{t=1}^{N}\varepsilon_t\frac{\partial\varepsilon_t}{\partial\beta} - \frac{1}{2\Sigma}\frac{\partial\Sigma}{\partial\beta}\left[N - \frac{1}{\Sigma}\sum_{t=1}^{N}\varepsilon_t^2\right],$$

from (2)

$$\frac{\partial\varepsilon_t}{\partial\beta} = -G_2(q)^{-1}\left\{\frac{\partial G_2(q)}{\partial\beta}\varepsilon_t + \frac{\partial G_1(q)}{\partial\beta}u_t\right\}$$

and where $\mathbb{E}_{Y|\beta}$ denotes the expectation over the distribution of the data given β.

We assume an open-loop experiment so that w_t and u_t are uncorrelated. We also assume that G_1, G_2 and Σ are independently parameterised. Taking expectations, as in (3), M can be partitioned as

$$M = \begin{bmatrix} M_1 & 0 \\ 0 & M_2 \end{bmatrix}$$

where M_1 is the part of the information matrix related to θ, and M_2 is independent of the input. Thus,

$$M_1 \triangleq \frac{1}{\Sigma}\sum_{t=1}^{N}\left(\frac{\partial\varepsilon_t}{\partial\theta}\right)\left(\frac{\partial\varepsilon_t}{\partial\theta}\right)^T, \tag{4}$$

where $\partial\varepsilon_t/\partial\theta$ satisfies

$$\frac{\partial\varepsilon_t}{\partial\theta} = -G_2(q)^{-1}\frac{\partial G_1(q)}{\partial\theta}u_t.$$

Notice that M_1 depends on the full parameter vector β. Assuming N is large, it is more convenient to work with the scaled