Available online at www.sciencedirect.com
SCIENCE DIRECT®

automatica

Automatica 41 (2005) 1313–1321

www.elsevier.com/locate/automatica

Data compression for estimation of the physical parameters of stable and unstable linear systems☆

Peter J. Gawthrop[a,*], Liuping Wang[b]

[a]*Centre for Systems and Control and Department of Mechanical Engineering, University of Glasgow, Glasgow G12 8QQ, Scotland*
[b]*Discipline of Electrical Energy and Control Systems, School of Electrical and Computer Engineering, RMIT University, Melbourne, Victoria 3000, Australia*

Received 19 January 2004; received in revised form 24 February 2005; accepted 4 March 2005
Available online 23 May 2005

Abstract

A two-stage method for the identification of physical system parameters from experimental data is presented. The first stage *compresses* the data as an empirical model which encapsulates the data content at frequencies of interest. The second stage then uses data extracted from the empirical model of the first stage within a nonlinear estimation scheme to estimate the unknown physical parameters. Furthermore, the paper proposes use of exponential data weighting in the identification of partially unknown, unstable systems so that they can be treated in the same framework as stable systems. Experimental data are used to demonstrate the efficacy of the proposed approach.

Keywords: Parameter estimation; Partially known systems; Basis function

1. Introduction

Many engineering systems of interest to the control engineer are *partially known* in the sense that the system structure, together with some system parameters are known, but some system parameters are unknown. This gives rise to a problem of *parameter estimation* when values for the unknown parameters are to be determined from experimental data comprising measurements of system inputs and outputs. There is a considerable literature in the area (An, Atkeson, & Hollerbach, 1988; Canudas de Wit, 1988; Dasgupta, Anderson, & Kaye, 1986; Gawthrop, Jones, & MacKenzie, 1992; Gawthrop, 2000a,b; Nagy & Ljung, 1991). Although, in special cases, such identification may be *linear*-in-the parameters (An et al., 1988) or *polynomial*-in-the parameters (Gawthrop et al., 1992), in general the problem is *nonlinear*-in-the parameters. This means that, in general, the resultant optimisation problem is not quadratic or polynomial, and may even be non-convex. In such cases, the optimisation task is eased by knowing (rather than deducing numerically) the derivative of the error function with respect to the unknown system parameters. The generation of such sensitivity information is aided by the symbolic methods for nonlinear systems modelling, analysis and optimisation which are currently strong research areas (Munro, 1999) driven by the ready availability of symbolic computational tools. In particular, the bond graph approach (Gawthrop & Smith, 1996; Karnopp, Margolis, & Rosenberg, 1990; Ljung & Glad, 1994) has been used to generate models both applicable to control design (Gawthrop, 1995; Gawthrop & Ronco, 2000) and partially-known system identification (Gawthrop, 2000a, 2003; Nagy & Ljung, 1991). Bond graph models are used in all the examples of this paper, but are not discussed further here.

Data acquisition systems typically yield large amounts of discrete-time data. On the other hand, the aforementioned

☆This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor A. Vicino under the direction of Editor T. Soederstroem.

* Corresponding author. Tel.: +44 141 330 4960; fax: +44 141 330 4343.
E-mail address: p.gawthrop@eng.gla.ac.uk (P.J. Gawthrop).

0005-1098/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2005.03.013

partially known systems are usually best expressed in continuous-time differential equation form and, even with these sensitivity function enhancements, use of the raw data may lead to unacceptable computational times. Thus, although it is, in principle, possible to use algorithms for partially-known system identification directly on the raw data, it is not practically useful. In addition, the raw data may contain complex system disturbance information which may require a sophisticated optimisation algorithm to achieve desirable results.

In this paper, the authors propose a two-stage identification procedure to extract physical parameters from discrete-time data pertaining to partially known systems. The first stage (which we call data compression) analyses the raw data to obtain a parameter vector θ describing an empirical model obtained from the data. The second stage uses this empirical model (parameterised by θ) to generate continuous-time data suitable for identifying physical parameters. Because the first stage is essentially a linear-in the parameter problem, not only can large amounts of data be processed rapidly, but also established system identification tools can be used to obtain data-quality models (Ljung, 1999). Because the second stage uses a relatively short length of relatively noise free data, the iteration time and convergence properties are much improved compared to using the raw data directly. The continuous time step response is used as the empirical model as it has a transparent representation in terms of gain, time delay and time constant and thus it is widely accepted by engineers and practitioners. Other forms of empirical model are also possible within this context. The basic idea of a two-stage method is not new, see for example, Ljung (1999, Section 10.4) and Wang, Gawthrop, Chessari, Podsiadly, and Giles (2004); our method is new in so far as it uses the frequency sampling filter (FSF) approach for the first stage and a physical model-based approach for the second.

In order for the same framework to be applicable to unstable systems, this paper proposes the use of exponential data weighting in the data compression procedure. This exponential weighting converts an unstable system into a stable system with the same unknown parameters to which the two-stage approach is applicable.

The motivation for this work is to generate models suitable for model-based predictive control (Mayne, Rawlings, Rao, & Scokaert, 2000; Rawlings, 2000), in particular models suitable for continuous time methods, such as those of Wang (2001) and Gawthrop and Ronco (2000, 2002).

The outline of the paper is as follows. Section 2 considers the FSF approach to data compression and extends the procedure to cope with unstable systems. Section 3 considers physical parameter estimation and Section 4 considers frequency-domain approaches. Section 5 gives illustrative experimental results using data obtained from both electrical and electro-mechanical systems. Section 6 concludes the paper.

2. Data compression

The first stage of the two-stage process is data compression: encapsulating the important features of the measured data into a few parameters within an empirical system model. There are many possible empirical models available including ARX (Ljung, 1999) and general basis-function approaches (Ninness & Gustafsson, 1997; Wahlberg, 1991; Wang & Cluett, 2000). In a fast-sampling environment, it is known that discrete-time ARX models encounter numerical ill-conditioning (Åström, Hagander, & Sternby, 1980) as the sampling rate increases and the problem is worse for unstable systems. On the contrary, the FSF approach of Bitmead and Anderson (1981), Wang and Cluett (1997, 2000), lies between the continuous and discrete-time domains and the coefficients converge to sampling-rate independent values. This latter approach is discussed in Section 2.1 and extended to unstable systems in Section 2.2.

2.1. Stable systems

The book by Wang and Cluett (2000) gives a comprehensive discussion of the FSF approach (including its relation to the discrete Fourier transform); this section provides a brief discussion of the material required for this paper. We consider linear time-invariant continuous-time systems with output $y(t)$ and input $u(t)$ uniformly sampled with time interval Δ to give input and output sequences $y_i = y(\mathrm{i}\Delta)$ and $u_i = u(\mathrm{i}\Delta)$. In the time-domain, the input and output sequences are related by $y_i = g_i * u_i$ where g_i is the discrete-time system impulse response and $*$ is the convolution operator. In the z-transform domain, $\bar{Y}(z) = \bar{G}(z)\bar{U}(z)$ where $\bar{Y}$ and $\bar{U}$ are the z-transforms of y_i and u_i, respectively and $\bar{G}$ the corresponding transfer function. In this section, it is assumed that the system is *stable* and can be associated with a settling time $T = N\Delta$; the time after which the system impulse response is sufficiently small: $|g_i|\langle\varepsilon\ \forall i\rangle N$.

The FSF approach approximates the transfer function $\bar{G}(z)$ as

$$\bar{G}_{\mathrm{fsf}}(z) = \sum_{k=-\frac{n-1}{2}}^{\frac{n-1}{2}} \theta_k \bar{H}_k(z), \tag{1}$$

$$\bar{H}_k(z) = \frac{1}{N}\frac{1 - z^{-N}}{1 - \mathrm{e}^{\mathrm{j}\Omega k} z^{-1}}, \tag{2}$$

where n is odd and the *frequency sample interval* $\Omega = 2\pi/T$, $\bar{H}_k(z)$ is the kth FSF and θ_k the corresponding (complex) parameter. The name arises because the kth FSF has a frequency response with a peak at $\omega = k\Omega$. Fig. 1(a) shows the superimposed frequency responses of $\bar{H}_k(z)$ for $0 \leqslant k \leqslant 4$ when $T = 5$ implying $\Omega = 1.26$ for a frequency range $0 \leqslant \omega \leqslant 10$. The symbol “x” marks the frequency samples which coincide with the peaks of the FSFs. The kth