Automatica 45 (2009) 1937–1942

Contents lists available at ScienceDirect

Automatica

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

Brief paper

Unifying some higher-order statistic-based methods for errors-in-variables model identification[☆]

Stéphane Thil[a,*], Wei Xing Zheng[b], Marion Gilson[c], Hugues Garnier[c]

[a] Laboratoire ELIAUS, Université de Perpignan Via Domitia, 52 avenue Paul Alduy, 66860 Perpignan, France
[b] School of Computing and Mathematics, University of Western Sydney, Penrith South DC NSW 1797, Australia
[c] Centre de Recherche en Automatique de Nancy (CRAN UMR 7039), Nancy-Université, CNRS, BP 239, 54506 Vandœuvre-lès-Nancy Cedex, France

ARTICLE INFO

Article history:
Received 13 May 2008
Received in revised form
23 March 2009
Accepted 17 April 2009
Available online 25 June 2009

Keywords:
System identification
Errors-in-variables
Instrumental variable
Higher-order statistics

ABSTRACT

In this paper, the problem of identifying linear discrete-time systems from noisy input and output data is addressed. Several existing methods based on higher-order statistics are presented. It is shown that they stem from the same set of equations and can thus be united from the viewpoint of extended instrumental variable methods. A numerical example is presented which confirms the theoretical results. Some possible extensions of the methods are then given.

1. Introduction

Identification of errors-in-variables (EIV) models has been a very active domain of research in the past few years (see *e.g.* Diversi, Guidorzi, and Soverini (2007), Hong, Söderström, and Zheng (2007), Mahata and Garnier (2006), Mahata (2007), Pintelon and Schoukens (2007), Söderström (2008), Thil, Garnier, Gilson, and Mahata (2007) and Thil, Gilson, and Garnier (2008)), and a survey paper gathering most of the known developments has been recently published (Söderström, 2007).

Most of the research has been concerned with estimating the parameters of discrete-time EIV models with the help of second-order statistics. Nonetheless, a recently published paper (Thil, Garnier, & Gilson, 2008) has shown that continuous-time EIV model identification can be successfully handled using higher-order statistics (HOS). Although much work has been conducted in the HOS field for EIV model identification in the 90's, it seems that several questions concerning the practical use of HOS for system identification remain to be answered.

The aim of this paper is to present some HOS-based methods for EIV model identification in a unified way. More precisely, the links between the methods developed in Inouye and Tsuchiya (1991) and Chen and Chen (1994) and the discrete-time version of an algorithm presented in Thil, Garnier et al. (2008) are explored. It is shown that these methods stem from the same set of equations. Simulation results support the theoretical analysis, and some possible extensions for future work are given.

2. Errors-in-variables framework

Consider a discrete-time, linear, time-invariant EIV system. The noise-free input/output signals are related by

$$y_o(t) = G_o(q)u_o(t) \tag{1}$$

where q is the forward operator and $G_o(\cdot)$ is the transfer operator of the 'true' system. The input and output signals are both contaminated by noise sequences, denoted as $\tilde{u}$ and $\tilde{y}$, respectively. The data-generating system is thus given by

$$\begin{cases} y_o(t) = G_o(q)u_o(t) \\ u(t) = u_o(t) + \tilde{u}(t) \\ y(t) = y_o(t) + \tilde{y}(t). \end{cases} \tag{2}$$

[☆] This work was supported by a research grant from the Australian Research Council. This paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Brett Ninness under the direction of Editor Torsten Söderström.

[*] Corresponding author. Tel.: +33 0 4 68 66 22 02; fax: +33 04 68 66 22 87.
E-mail addresses: stephane.thil@univ-perp.fr (S. Thil), w.zheng@uws.edu.au (W.X. Zheng), marion.gilson@cran.uhp-nancy.fr (M. Gilson), hugues.garnier@cran.uhp-nancy.fr (H. Garnier).

0005-1098/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2009.04.019

It is then parameterized as follows:

$$\begin{cases} y(t) = G(q,\boldsymbol{\theta})\left(u(t)-\tilde{u}(t)\right)+\tilde{y}(t) \\ G(q,\boldsymbol{\theta}) = B(q^{-1},\boldsymbol{\theta})/A(q^{-1},\boldsymbol{\theta}) \\ A(q^{-1},\boldsymbol{\theta}) = 1+a_1q^{-1}+\cdots+a_{n_a}q^{-n_a} \\ B(q^{-1},\boldsymbol{\theta}) = b_0+b_1q^{-1}+\cdots+b_{n_b}q^{-n_b} \end{cases} \tag{3}$$

with $n_a \geqslant n_b$ and $\boldsymbol{\theta}^{\mathrm{T}} = \left[a_1 \ \cdots \ a_{n_a} \ b_0 \ \cdots \ b_{n_b}\right]$. Equation (3) can be rewritten as

$$y(t) = \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta} + v(t,\boldsymbol{\theta}) \tag{4}$$

$$v(t,\boldsymbol{\theta}) = \tilde{y}(t) - \tilde{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\boldsymbol{\theta} \tag{5}$$

where the regression vector is given by

$$\boldsymbol{\varphi}^{\mathrm{T}}(t) = [-y(t-1) \ \cdots \ -y(t-n_a) \ u(t) \ \cdots \ u(t-n_b)] \tag{6}$$

and $\tilde{\boldsymbol{\varphi}}(t)$ is defined in a similar way to $\boldsymbol{\varphi}(t)$, but with u and y being replaced by $\tilde{u}$ and $\tilde{y}$, respectively. The problem of identifying this errors-in-variables model is concerned with consistently estimating the parameter vector $\boldsymbol{\theta}$ from the noisy input/output data $\{u(t), y(t)\}_{t=1}^{N}$.

2.1. Notations

As the input and output noises are additive, linear functions of the measured signals can be broken down into two parts: one part made up of the noise-free signal contribution (denoted with an 'o' subscript) and the other made up of the noises' contribution (denoted with the '˜' sign). For example, the regression vector $\boldsymbol{\varphi}$ can be decomposed into

$$\boldsymbol{\varphi}(t) = \boldsymbol{\varphi}_o(t) + \tilde{\boldsymbol{\varphi}}(t). \tag{7}$$

The following notations are used in what follows for the correlation vectors and matrices

$$\boldsymbol{R}_{\varphi\varphi} = \bar{E}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t), \quad \boldsymbol{r}_{\varphi y} = \bar{E}\boldsymbol{\varphi}(t)y(t) \tag{8}$$

where $\bar{E}\{\cdot\}$ stands for (see Ljung (1999))

$$\bar{E}\{f(t)\} = \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}E\{f(t)\}. \tag{9}$$

The notation used for the third-order cumulants is

$$C_{x_1x_2x_3}(\tau_1,\tau_2) = \bar{E}\{x_1(t)x_2(t+\tau_1)x_3(t+\tau_2)\} \tag{10}$$

$$= C_{x_1x_2x_3}(\boldsymbol{\tau}) \tag{11}$$

where, for the sake of conciseness, $\boldsymbol{\tau}$ denotes $[\tau_1, \tau_2]$.

2.2. Assumptions and elements of structure

The following assumptions are needed

A1. The system (1) is asymptotically stable, and all the system modes are observable and controllable;
A2. The signals u_o, $\tilde{u}$ and $\tilde{y}$ are stationary, ergodic and zero-mean;
A3. The signals $\tilde{u}$ and $\tilde{y}$ are assumed to be uncorrelated with the input u_o.

For methods based on second-order statistics to give unbiased estimates, it is usually assumed (and often implicitly) that the 'true' system belongs to the considered model set, a situation referred to as $\mathcal{S} \in \mathcal{M}^\star$ (Ljung, 1999). However, this notation has been introduced for systems with noise-free inputs, and – being too general – is not properly suited for errors-in-variables models. Indeed, more often than not, the input and output noises and the noise-free input must be modeled. Thus, some additional notations must be introduced. The whole 'true' system includes

(1) the 'true' process G_o and its associated model set $\mathcal{G}^\star = \{G(\cdot,\boldsymbol{\theta})\}$,
(2) the 'true' noise processes $H_o^{\tilde{u}}, H_o^{\tilde{y}}$ and their associated model set $\mathcal{H}^\star = \left\{H^{\tilde{u}}(\cdot,\boldsymbol{\eta}),\ H^{\tilde{y}}(\cdot,\boldsymbol{\eta})\right\}$,
(3) the 'true' noise-free input process $H_o^{u_o}$ and its associated model set $\mathcal{E}^\star = \{H^{u_o}(\cdot,\boldsymbol{\eta})\}$,

where $\boldsymbol{\eta}$ is a vector gathering the parameters of noise models and noise-free input models.

In addition to assuming that the 'true' process belongs to the model set, *i.e.*, $G_o \in \mathcal{G}^\star$, most methods based on second-order statistics require that the noise models belong to the model set, *i.e.*, $\left(H_o^{\tilde{u}}, H_o^{\tilde{y}}\right) \in \mathcal{H}^\star$. A few even require that the noise-free input is adequately modeled, and thus that $H_o^{u_o} \in \mathcal{E}^\star$. For example, the maximum likelihood and prediction error methods require such assumptions (Söderström, 1981, 2007).

On the contrary, methods based on higher-order statistics do not require *structural assumptions* on the input and output noises $\tilde{u}, \tilde{y}$, and on the noise-free input u_o. The only structural assumption needed is

A4. The true process belongs to the model set: $G_o \in \mathcal{G}^\star$.

The input and output noises can thus be arbitrarily coloured (and even mutually correlated), and there is no structural assumption on the noise-free input. However, for the higher-order cumulants of the noises to be zero and for the higher-order cumulants of the noise-free input not to be zero, *distributional assumptions* are needed. These distributional assumptions differ whether third- or fourth-order cumulants are used. For the third-order cumulants,

A5a. The input and output noises $\tilde{u}, \tilde{y}$ have symmetric probability density functions (pdfs),
A6a. The noise-free input u_o has a skewed pdf.

For the fourth-order cumulants,

A5b. The input and output noises $\tilde{u}, \tilde{y}$ have Gaussian pdfs,
A6b. The noise-free input u_o has a non-Gaussian pdf.

In what follows, the focus will be placed on the case of third-order cumulants, and consequently the assumptions A1–A4 and A5a–A6a are supposed to be satisfied.

3. Identification methods using HOS

3.1. Properties of HOS

The identification techniques presented in this paper are based on higher-order statistics (see *e.g.* Brillinger (1981) and Mendel (1991)). Here we recall a few of the numerous properties of higher-order cumulants.

P1. Multilinearity: cumulants are linear with respect to each of their arguments;
P2. Additivity: if two random vectors are *independent*, then the cumulant of their sum equals the sum of their cumulants;
P3. The third-order cumulant of a random variable with a symmetric pdf is equal to zero.

From assumptions A3, A5a, A6a and using properties P2, P3, the following holds

$$C_{uuy}(\boldsymbol{\tau}) = C_{u_0u_0y_0}(\boldsymbol{\tau}) + C_{\tilde{u}\tilde{u}\tilde{y}}(\boldsymbol{\tau}) = C_{u_0u_0y_0}(\boldsymbol{\tau})$$

$$C_{uuu}(\boldsymbol{\tau}) = C_{u_0u_0u_0}(\boldsymbol{\tau}) + C_{\tilde{u}\tilde{u}\tilde{u}}(\boldsymbol{\tau}) = C_{u_0u_0u_0}(\boldsymbol{\tau}).$$

The third-order (cross-)cumulants of the input and output signals are thus insensitive to symmetrically distributed noises. Note that this result is still valid for the third-order (cross-)cumulant of any combination of input and output signals.