Automatica 50 (2014) 2597–2605

Contents lists available at ScienceDirect

Automatica

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

Brief paper

# Extended accuracy analysis of a covariance matching approach for identifying errors-in-variables systems[☆]

Torsten Söderström[a,1], David Kreiberg[b,c], Magnus Mossberg[d]

[a] Division of Systems and Control, Department of Information Technology, Uppsala University, P O Box 337, SE-75105 Uppsala, Sweden

[b] Department of Statistics, Uppsala University, P O Box 513, SE-75120 Uppsala, Sweden

[c] Department of Economics, BI Norwegian Business School, 0442 Oslo, Norway

[d] Department of Engineering and Physics, Karlstad University, SE-651 88, Karlstad, Sweden

ARTICLE INFO

*Article history:*
Received 8 August 2013
Received in revised form
21 January 2014
Accepted 21 May 2014
Available online 16 September 2014

*Keywords:*
System identification
Errors-in-variables models
Linear systems
Covariance functions
Covariance matrix

ABSTRACT

A covariance matching approach for identifying errors-in-variables systems is analyzed for the general case. The asymptotic covariance matrix of the jointly estimated system parameters, noise variances and auxiliary parameters is derived. An algorithm for how to compute this covariance matrix from given system descriptions is also provided. The results generalize previous known special cases. Using Monte Carlo analysis, we illustrate the proposed algorithm. The results suggest close agreement between the theoretical and empirical accuracy.



## 1. Introduction

Within system identification, the errors-in-variables problem is known to contain several particular difficulties due to the presence of measurement noise on both inputs and outputs (Söderström, 2007, 2012). A number of different estimator classes have been proposed in the literature. Of these classes, the covariance matching (CM) approach introduced in Söderström, Mossberg, and Hong (2009) has an attractive tradeoff between computational complexity and statistical performance. It was first formulated for discrete-time models, but later generalized to continuous-time models (Mossberg & Söderström, 2011b). It has been shown to be closely related to structural equation modeling techniques, which are developed in multivariate statistics for static problems (Bartholomew, Knott, & Moustaki, 2011; Jöreskog, 1970). For a description of such relations, see Kreiberg, Söderström, and Wallentin (2013). Other methods, also based on a finite number of covariances obtained from the measured data, include the Frisch method (Beghelli, Guidorzi, & Soverini, 1990; Guidorzi, Diversi, & Soverini, 2008) as well as extensions of the instrumental variable method. The development and description of such methods can be found in Ekman (2005) and Söderström (2011).

The accuracy of the system parameter estimates for the CM approach is analyzed in Söderström and Mossberg (2011), where the asymptotic covariance matrix of the parameter estimates is derived. In this paper, we generalize the analysis to also include the accuracy of the estimated noise variances as well as the estimated auxiliary parameters. An explicit algorithm for how to compute the theoretical covariance matrix of the joint parameters is provided. It turns out that the new result is not only more general but also neater than the previous special case presented in Söderström and Mossberg (2011).

There are good reasons for extending the analysis to also include the noise variances and other auxiliary parameters. For instance, it is of interest to consider the accuracy of the estimated noise variances, as these estimates can be used to obtain the signal-to-noise ratios (that is, to assess the relative magnitude of the undisturbed signal component and the measurement noise component in the measurements. For example, for the output signal the SNR can be computed as $\mathsf{E}\left\{y_0^2(t)\right\}/\mathsf{E}\left\{\tilde{y}^2(t)\right\}$ where $\mathsf{E}$ denotes the expectation operator, cf. (3) below. Estimates of the variances involved in this

---

[☆] This research was partially supported by The Swedish Research Council, contract 621-2007-6364. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Tongwen Chen under the direction of Editor Ian R. Petersen.

*E-mail addresses:* David.Kreiberg@statistics.uu.se (T. Söderström), David.Kreiberg@statistics.uu.se, David.Kreiberg@bi.no (D. Kreiberg), Magnus.Mossberg@kau.se (M. Mossberg).

[1] Tel.: +46 184713075; fax: +46 18511925.

http://dx.doi.org/10.1016/j.automatica.2014.08.020
0005-1098/

expression for SNR can be computed from $\boldsymbol{\theta}$ and $\mathbf{r}_z$, here given in (4) and (7), respectively). Moreover, auxiliary parameter estimates are of interest, not only as vehicle for obtaining the system parameter estimates, but also for testing the physical relevance of these estimates. As an example, one may require that the estimates of $r_0(\tau)$ in (7), (9) below form a positive definite sequence (although this is not included in implementations for simplicity reasons). A test for relevance would take the uncertainties of the estimated elements of $r_0(\tau)$ into account when testing for positive definiteness.

## 2. Background

Consider the following errors-in-variables problem. The noise-free inputs $u_0(t)$ and outputs $y_0(t)$ are linked through a dynamic system of the form

$$A(q^{-1})y_0(t) = B(q^{-1})u_0(t), \tag{1}$$

where $q^{-1}$ is the backward shift operator, and

$$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}, \\ B(q^{-1}) &= b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b}. \end{aligned} \tag{2}$$

The observed variables are

$$\begin{matrix} y(t) = y_0(t) + \tilde{y}(t) \\ u(t) = u_0(t) + \tilde{u}(t) \end{matrix} \quad t = 1, \ldots, N. \tag{3}$$

It is assumed that the measurement noises $\tilde{y}(t), \tilde{u}(t)$ are both white, and that $\tilde{y}(t)$, $\tilde{u}(t)$ and $u_0(t)$ are all mutually independent. The noise variances will be denoted $\lambda_y = \mathsf{E}\left\{\tilde{y}^2(t)\right\}$ and $\lambda_u = \mathsf{E}\left\{\tilde{u}^2(t)\right\}$, respectively. The aim of the identification is to estimate the unknown parameter vector

$$\boldsymbol{\theta} = \begin{pmatrix} a_1 & \cdots & a_{n_a} & b_1 & \cdots & b_{n_b} \end{pmatrix}^T. \tag{4}$$

The covariance matching (CM) method for errors-in-variables identification is now well known. It is described and analyzed in Söderström et al. (2009). The asymptotic covariance matrix of the parameter estimates is derived in Söderström and Mossberg (2011).

The underlying principles can be summarized as follows. Introduce

$$\mathbf{r} \triangleq \begin{pmatrix} \mathbf{r}_y \\ \mathbf{r}_u \\ \mathbf{r}_{yu} \end{pmatrix}, \tag{5}$$

where

$$\mathbf{r}_y = \begin{pmatrix} r_y(0) \\ \vdots \\ r_y(p_y) \end{pmatrix}, \quad \mathbf{r}_u = \begin{pmatrix} r_u(0) \\ \vdots \\ r_u(p_u) \end{pmatrix}, \quad \mathbf{r}_{yu} = \begin{pmatrix} r_{yu}(p_1) \\ \vdots \\ r_{yu}(p_2) \end{pmatrix}, \tag{6}$$

contain the covariance elements $r_y(\tau) = \mathsf{E}\{y(t+\tau)y(t)\}$, etc. Further, by defining

$$\mathbf{r}_z = \begin{pmatrix} r_0(0) \\ \vdots \\ r_0(k) \\ \lambda_y \\ \lambda_u \end{pmatrix}, \tag{7}$$

the system of equations

$$\mathbf{r} = \mathbf{F}(\boldsymbol{\theta})\mathbf{r}_z = \begin{pmatrix} \mathbf{F}_y(\boldsymbol{\theta}) \\ \mathbf{F}_u(\boldsymbol{\theta}) \\ \mathbf{F}_{yu}(\boldsymbol{\theta}) \end{pmatrix} \mathbf{r}_z \tag{8}$$

can be derived. In (7),

$$r_0(\tau) = \mathsf{E}\left\{\frac{1}{A(q^{-1})}u_0(t+\tau)\frac{1}{A(q^{-1})}u_0(t)\right\}. \tag{9}$$

For specific details on how the matrix $\mathbf{F}(\boldsymbol{\theta})$ in (8) depends on $\boldsymbol{\theta}$, see Söderström et al. (2009). Note that, compared to the presentation in Söderström et al. (2009) and Söderström and Mossberg (2011), we have 0 as the lowest index in $\mathbf{r}_y$ and $\mathbf{r}_u$, and not 1. Further, we have included the noise variances $\lambda_y$ and $\lambda_u$ in the vector $\mathbf{r}_z$.

The estimator

$$\{\hat{\boldsymbol{\theta}}, \hat{\mathbf{r}}_z\} = \arg\min_{\boldsymbol{\theta}, \mathbf{r}_z} J(\boldsymbol{\theta}, \mathbf{r}_z), \tag{10}$$

$$J(\boldsymbol{\theta}, \mathbf{r}_z) = \|\hat{\mathbf{r}} - \mathbf{F}(\boldsymbol{\theta})\mathbf{r}_z\|^2_{\mathbf{Q}} \tag{11}$$

for $\boldsymbol{\theta}$ and $\mathbf{r}_z$, based on (8), is suggested in Söderström et al. (2009). In (11), $\hat{\mathbf{r}}$ is an estimate of $\mathbf{r}$ while $\mathbf{Q}$ is a symmetric weighting matrix. From (10),

$$\hat{\mathbf{r}}_z = \left(\mathbf{F}^T(\boldsymbol{\theta})\mathbf{Q}\mathbf{F}(\boldsymbol{\theta})\right)^{-1}\mathbf{F}^T(\boldsymbol{\theta})\mathbf{Q}\hat{\mathbf{r}}, \tag{12}$$

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} V(\boldsymbol{\theta}), \tag{13}$$

and

$$V(\boldsymbol{\theta}) = \hat{\mathbf{r}}^T\left[\mathbf{Q} - \mathbf{Q}\mathbf{F}(\boldsymbol{\theta})\left(\mathbf{F}^T(\boldsymbol{\theta})\mathbf{Q}\mathbf{F}(\boldsymbol{\theta})\right)^{-1}\mathbf{F}^T(\boldsymbol{\theta})\mathbf{Q}\right]\hat{\mathbf{r}}. \tag{14}$$

**Remark.** The weighting matrix $\mathbf{Q}$ must be non-negative definite, but not necessarily positive definite. There are important cases where $\mathbf{Q}$ is positive semidefinite and thus singular. What matters is that the matrix product $\mathbf{F}^T\mathbf{Q}\mathbf{F}$ is positive definite and hence invertible, see (12). Specifically, if the matrix $\mathbf{Q}$ is chosen so that the rows and columns corresponding to $r_y(0)$ and $r_u(0)$ are set to zero, the matrix $\mathbf{Q}$ becomes singular. In that case, the CM approach does not include estimation of the noise variances $\lambda_y$ and $\lambda_u$, and is the case originally treated in the paper Söderström et al. (2009). It is hence a special case of the general formulation here. □

**Remark.** Structural equation modeling (SEM) can be seen as closely related to the CM approach, see Kreiberg et al. (2013) for some details. SEM includes several possible estimation criteria. Some of these (specific examples include *unweighted least squares* (ULS) and *generalized least squares* (GLS)) can indeed be transformed into the general case in (11) by specific selection of the weighting matrix $\mathbf{Q}$. There are though also other possible SEM criteria (for example the one often called maximum likelihood, and labeled $V_1$ in Kreiberg et al. (2013)) that cannot be transformed into the form in (11). □

The following results are derived in Söderström and Mossberg (2011).

**Theorem 2.1.** *The asymptotic normalized covariance matrix of the parameter estimates fulfils*

$$\begin{aligned} \mathbf{C}_{\boldsymbol{\theta}} &\triangleq \lim_{N\to\infty} N\mathrm{cov}(\hat{\boldsymbol{\theta}}) \\ &= [\mathbf{S}^T\mathbf{P}\mathbf{S}]^{-1}\mathbf{S}^T\mathbf{P}\mathbf{R}\mathbf{P}\mathbf{S}[\mathbf{S}^T\mathbf{P}\mathbf{S}]^{-1}. \end{aligned} \tag{15}$$

*Here,*

$$\mathbf{P} = \mathbf{Q} - \mathbf{Q}\mathbf{F}\left(\mathbf{F}^T\mathbf{Q}\mathbf{F}\right)^{-1}\mathbf{F}^T\mathbf{Q}, \tag{16}$$

$$\mathbf{R} = \lim_{N\to\infty} N\mathsf{E}\left\{(\hat{\mathbf{r}} - \mathbf{r})(\hat{\mathbf{r}} - \mathbf{r})^T\right\}, \tag{17}$$

$$\mathbf{S} = \begin{pmatrix} \mathbf{s}_1 & \cdots & \mathbf{s}_{n_a+n_b} \end{pmatrix}, \tag{18}$$

$$\mathbf{s}_j = \mathbf{F}_j\mathbf{r}_z, \tag{19}$$

$$\mathbf{F}_j = \frac{\partial \mathbf{F}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_j}, \tag{20}$$