Communications and Control Engineering

Communications and Control Engineering is a high-level academic monograph series publishing research in control and systems theory, control engineering and communications. It has worldwide distribution to engineers, researchers, educators (several of the titles in this series find use as advanced textbooks although that is not their primary purpose), and libraries.

The series reflects the major technological and mathematical advances that have a great impact in the fields of communication and control. The range of areas to which control and systems theory is applied is broadening rapidly with particular growth being noticeable in the fields of finance and biologically-inspired control. Books in this series generally pull together many related research threads in more mature areas of the subject than the highly-specialised volumes of *Lecture Notes in Control and Information Sciences*. This series's mathematical and control-theoretic emphasis is complemented by *Advances in Industrial Control* which provides a much more applied, engineering-oriented outlook.

Publishing Ethics: Researchers should conduct their research from research proposal to publication in line with best practices and codes of conduct of relevant professional bodies and/or national and international regulatory bodies. For more details on individual ethics matters please see:
https://www.springer.com/gp/authors-editors/journal-author/journal-author-helpdesk/publishing-ethics/14214

More information about this series at http://www.springer.com/series/61

Torsten Söderström

Errors-in-Variables Methods in System Identification

Springer

Torsten Söderström
Division of Systems and Control,
Department of Information Technology
Uppsala University
Uppsala
Sweden

ISSN 0178-5354 ISSN 2197-7119 (electronic)
Communications and Control Engineering
ISBN 978-3-319-75000-2 ISBN 978-3-319-75001-9 (eBook)
https://doi.org/10.1007/978-3-319-75001-9

Library of Congress Control Number: 2018930366

Mathematics Subject Classification (2010): 62F10, 93E12

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To Andreas, Christer, David, Elisabet, Frida, Gunnar, Gustav, Hjalmar, Johanna, Klara, Marianne and Olof

Preface

This book is intended to give a comprehensive overview of errors-in-variables (EIV) problems in system identification. This problem is about modeling of dynamic systems when all measured variables and signals are noise-corrupted. A number of different approaches are described and analyzed. The area has been a central one in my own research for a long time, and this experience has influence my own way of thinking of how to describe and categorize the many proposed methods available in the literature. The area continues to be active today, and there is a steady inflow of articles on EIV for dynamic systems, to leading conferences as well as to journals.

As a proper background the reader is expected to have at least elementary knowledge of system identification. The textbooks (1999) and Söderström and Stoica (1989) can be recommended. They cover much more than what is required in this context.

This book starts with giving a background for the errors-in-variables (EIV) problem. First static systems are treated in some detail. The dominating part of this book copes with dynamic systems. The EIV problem as such is carefully analyzed, and it is demonstrated that some additional assumption(s) must be imposed if a unique solution is to be found. Several approaches and EIV methods are presented and analyzed. This book ends with a chapter on users' perspectives for applying EIV methods in practice. See also Sect. 1.2 for a more detailed description of this book.

It is a pleasure to express my gratitude to many colleagues with whom I over the years have discussed, learned from, and published work together with on errors-in-variables problems. These colleagues include Juan Carlos Agüero, Brian D. O. Anderson, Theodore Anderson, Keith Burnham, Mats Cedervall, Han-Fu Chen, Bart De Moor, Manfred Deistler, Roberto Diversi, Mats Ekman, Hugues Garnier, Marion Gilson, Tryphon Georgiou, Graham Goodwin, Roberto Guidorzi, Christiaan Heij, Håkan Hjalmarsson, Mei Hong Bjerstedt, Alireza Karimi, Erlendur Karlsson, David Kreiberg, Alexander Kukush, Tomas Larkowski, Erik K. Larsson, Jens Linden, Kaushik Mahata, Ivan Markovsky, Magnus Mossberg, Rik Pintelon, Agnes Rensfelt, Cristian Rojas, Wolfgang Scherrer, Johan Schoukens, Virginija Šimonytė, Joachim

Sorelius, Umberto Soverini, Petre Stoica, Stephane Thil, Klaske van Heusden, Sabine Van Huffel, Kiyoshi Wada, Fan Yang Wallentin, and Wei Xing Zheng.

I am truly indebted to many colleagues who have read earlier versions of the book, in whole or in part, and pointed out a large number of errors and unclear points, and also generously given me proposals for additional ideas. For example, the last chapter of the book was written and included based on such feedback. The reviewers who provided this indispensable help are Juan Carlos Agüero, Roberto Diversi, Håkan Hjalmarsson, David Kreiberg, Ivan Markovsky, Magnus Mossberg, Giorgio Picci, Rik Pintelon, Johan Schoukens, and Umberto Soverini.

I would also like to thank the personnel at Springer (Oliver Jackson, Meertinus Faber, Geethajayalaxmi Govindarjan, Komala Jaishankar, Ravikrishnan Karunanandam, and Balaganesh Sukumar) for a smooth cooperation in producing the book.

It happens sometimes that I have found manuscripts with ‘error-in-variables’, rather than ‘errors-in-variables’ in the title (yes, it has happened also in my own draft papers!). As is explained in the book, for EIV systems it is indeed a key aspect that there are errors on both input and output measurements, and therefore one must use plural! During the work with the book manuscript I have corrected quite a number of errors, and it is my sincere hope that not too many remain, even though there may still be more than a single one! During my scientific career I have mainly been active in the control community, and therefore I believe in feedback. In particular, I would welcome the readers’ comments on the text and possibly pointing out any remaining error. I can be reached on the e-mail address: `torsten.soderstrom@it.uu.se`.

Some years ago I told my family that I was planning to write another book. Some suggested that this time I should write a thriller. A plot was laid out about a murder that was detected at the opening ceremony of a major control conference. I quickly turned down this idea, that was not so serious anyway. It would demonstrate my inability to write something exciting from a fiction point of view. Furthermore, the theme does not match my general impressions from almost half a century with the control community. I have mainly found it to be characterized by friendly and helpful people. I have dedicated this book to my extended family (grandchildren are included): **A**ndreas,…, **O**lof. You are the A and O to me!

Vattholma and Uppsala, Sweden Torsten Söderström
December 2017

Contents

Abbreviations

AIC	Akaike's information criterion
AR	Autoregressive
AR(n)	AR of order n
arg	Argument
ARMA	Autoregressive moving average
ARMA(n, m)	ARMA where AR and MA parts have orders n and m, respectively
ARMAX	Autoregressive moving average with exogenous input
BELS	Bias-eliminating least squares
BIC	Bayesian information criterion
CFA	Confirmatory factor analysis
CM	Covariance matching
cov	Covariance matrix
CRB	Cramér-Rao lower bound
deg	Degree
DFT	Discrete Fourier transform
DLS	Data least squares
EIV	Errors-in-variables
EIV	Extended instrumental variable
ETFE	Empirical transfer function estimate
FD	Frequency domain
FIR	Finite impulse response
FML	Frequency domain maximum likelihood
FOH	First-order hold
FPE	Final prediction error
GIVE	Generalized instrumental variable estimate
iid	Independent and identically distributed
IIR	Infinite impulse response
Imag	Imaginary part
IV	Instrumental variable
LHS	Left-hand side

LS	Least squares
MA	Moving average
MA(n)	MA of order n
MC	Monte Carlo
MIMO	Multi-input, multi-output
ML	Maximum likelihood
pdf	Probability density function
PEM	Prediction error method
PLR	Pseudo-linear regression
Real	Real part
RHS	Right-hand side
RMS	Root-mean-square
SEM	Structural equation modeling
SISO	Single-input, single-output
SML	Sample maximum likelihood
SNR	Signal-to-noise ratio
STLS	Structured total least squares
SVD	Singular value decomposition
TD	Time domain
TLS	Total least squares
TML	Time domain maximum likelihood
tr	Trace (of a matrix)
vec	vec($\mathbf{A}$) is the long column vector obtained by stacking the columns of $\mathbf{A}$ on top of each other
YW	Yule Walker
ZOH	Zero-order hold

Notation

$\mathscr{CN}$	Complex Gaussian distribution
E	Expectation operator
$\mathsf{E}[\cdot\mid\cdot]$	Conditional expectation
e	Basis of natural logarithm
$e(t)$	White noise (a sequence of independent random variables)
$\mathbf{e}_i$	Unit vector (ith element equal to 1)
$\mathscr{GP}(\mu, K)$	Gaussian process with mean value function μ and covariance function K
$G(q^{-1})$	Transfer function operator
$\mathbf{g}_N$	Vector of dimension N with all elements equal to 1
$H(q^{-1})$	Transfer function operator, noise shaping filter
h	Sampling interval
h_k	Weighting function coefficient, $H(q^{-1}) = \sum_{k=0}^{\infty} h_k q^{-k}$
$\mathbf{I}$	Identity matrix
$\mathbf{I}_n$	$n \times n$ identity matrix
i	Imaginary unit
$\mathbf{J}$	Information matrix
L	Log likelihood function
$\mathscr{L}$	Laplace transform
$m \times n$	Matrix has dimension m by n
N	Number of data points
$\mathscr{N}(\mathbf{m},\mathbf{P})$	Normal (Gaussian) distribution with mean value $\mathbf{m}$ and covariance matrix $\mathbf{P}$
$\mathscr{N}$	Null space
$\mathbf{P}$	Covariance matrix of state or state prediction error
Pr	Probability
q	Shift operator, $qx(t) = x(t+1)$
q^{-1}	Backward shift operator, $q^{-1}x(t) = x(t-1)$
$\mathscr{R}$	Range space

$\mathscr{R}^n$	Euclidean n-dimensional space
$\mathscr{R}^{n\times m}$	Linear space of $n \times m$-dimensional matrices
$\mathbf{R}$	Covariance matrix
$\hat{\mathbf{R}}$	Sample covariance matrix (estimate of $\mathbf{R}$)
$\hat{\mathbf{R}}_N$	Sample covariance matrix based on N data points
$\mathbf{r}$	Covariance vector
r	Noise variance ratio
$\mathscr{S}(A, B)$	Sylvester matrix associated with polynomials A and B
t	Time variable (integer valued for discrete-time models)
tr	Trace (of a matrix)
$u(t)$	Input signal (possibly vector-valued)
$u_0(t)$	Noise-free input signal
$\tilde{u}(t)$	Input measurement noise
V	Loss function, performance index
V_N	Loss function based on N data points
V_∞	Asymptotic loss function, $V_\infty = \lim_{N\to\infty} V_N$
$v(t)$	White noise (a sequence of independent random variables)
$\mathbf{W}$	Weighting matrix
$\mathbf{x}(t)$	State vector
Y^t	All available output measurements at time t, $Y^t = \{y(t), y(t-1)\ldots\}$
$y(t)$	Output signal (possibly vector-valued)
$y_0(t)$	Noise-free output signal
$\hat{y}(t\|t-1)$	Optimal one-step predictor of $y(t)$ given Y^{t-1}
$\tilde{y}(t)$	Output measurement noise
$\mathscr{Z}$	z transform
$\mathbf{z}(t)$	Instrumental variable vector
$\gamma(\mathbf{x}; \mathbf{m}, \mathbf{P})$	Pdf of normal (Gaussian) distribution with mean value $\mathbf{m}$ and covariance matrix $\mathbf{P}$
Δ	Difference
$\delta_{t,s}$	Kronecker delta (=1 if $s = t$, else $= 0$)
$\varepsilon(t)$	Prediction error
Λ	Covariance matrix of innovations
$\lambda_{\min}(\mathbf{A})$	Smallest eigenvalue of the symmetric matrix $\mathbf{A}$
λ	Variance of white noise
λ_u	Variance of input noise
λ_y	Variance of output noise
$\phi(z)$	Spectrum
$\phi(\omega)$	Spectral density
$\phi_u(\omega)$	Spectral density of the signal $u(t)$
$\phi_{yu}(\omega)$	Cross-spectral density between the signals $y(t)$ and $u(t)$
$\Phi(z)$	Spectrum for a multivariate signal
$\boldsymbol{\rho}$	Noise parameter vector
$\boldsymbol{\tau}$	Vector of parameters describing effects of initial and final values
τ	Time lag (in covariance function)

ω	Angular frequency
$\varphi(t)$	Regressor vector
θ	Parameter vector
θ_0	True value of θ
$\hat{\theta}$	Estimate of θ
$\hat{\theta}_N$	Estimate $\hat{\theta}$ based on N data points
ϑ	Total parameter vector $\vartheta = (\theta^T\ \boldsymbol{\rho}^T)^T$
Θ	Extended parameter vector ($\Theta = (1\ \theta^T)^T$)
$\boldsymbol{\zeta}(t)$	Input–output data, $\boldsymbol{\zeta}(t) = (y(t)\ u(t)\ )^T$
$\oint$	Integration counterclockwise around the unit circle
$\mathbf{0}$	Zero matrix
$\mathbf{0}_{m\times n}$	Zero matrix of dimension $m \times n$
$O(x)$	Big ordo, $\lvert O(x)\rvert / \lvert x\rvert$ bounded when $x \to 0$
$o(x)$	Small ordo, $\lvert o(x)\rvert / \lvert x\rvert \ \to 0$ when $x \to 0$

Notational Conventions

$H^{-1}(q)$	$[H(q)]^{-1}$
$\mathbf{x}^T(t)$	$[\mathbf{x}(t)]^T$
$\mathbf{A}^{-T}$	$[\mathbf{A}^{-1}]^T$
$\mathbf{A} \geq \mathbf{B}$	The difference matrix $\mathbf{A} - \mathbf{B}$ is nonnegative definite
$\mathbf{A} > \mathbf{B}$	The difference matrix $\mathbf{A} - \mathbf{B}$ is positive definite
$\triangleq$	Defined as
$\sim$	Distributed as
$\overline{w}$	Complex conjugate of w
$\mathbf{w}^T$	Transpose of $\mathbf{w}$
$\mathbf{w}^*$	Conjugate transpose of $\mathbf{w}$, $\mathbf{w}^* = \overline{\mathbf{w}}^T$
$\mathbf{w}^H$	Conjugate transpose of $\mathbf{w}$, $\mathbf{w}^H = \overline{\mathbf{w}}^T$
w^{-*}	$(w^{-1})^* = (w^*)^{-1}$
$\mathbf{A}^\dagger$	Pseudo-inverse of the matrix $\mathbf{A}$, $\mathbf{A}^\dagger = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$
V'	Gradient of V
V''	Hessian of V
$\lVert\mathbf{A}\rVert_{\mathrm{F}}$	Frobenius norm of matrix $\mathbf{A}$
$\lVert\mathbf{x}\rVert_{\mathbf{W}}$	Weighted norm, $\lVert\mathbf{x}\rVert_{\mathbf{W}}^2 = \mathbf{x}^T\mathbf{W}\mathbf{x}$

Summary of Assumptions

The different general assumptions introduced in the book are summarized here for the reader's convenience. References are also provided to where the definitions are introduced.

Assumptions on the System

AS1. The system is linear and asymptotically stable [Chap. 3].
AS2. The system is causal, so $y_0(t)$ depends on $u_0(s)$ for $s \leq t$, but not on future values of $u_0(\cdot)$ [Chap. 3].
AS3. The noise-free input and output signals are linked by

$$A(q^{-1})y_0(t) = B(q^{-1})u_0(t).$$

All system modes are observable and controllable; *i.e.,* $A(z)$ and $B(z)$ have no common factor. The polynomial degrees n_a and n_b are known [Chap. 3].
AS4. The system transfer function $G(z)$ has no pair of zeros reflected in the unit circle; that is, if $G(z_1) = 0$, then $G(z_1^{-1}) \neq 0$ [Chap. 4].
AS5. If the system is non-causal, then $G(z)$ has no pair of poles reflected in the unit circle; that is, p_1 and p_1^{-1} cannot both be poles of $G(z)$ [Chap. 4].
AS6. The order of the transfer functions fulfills

$$\text{order}(GH) = \text{order}(G) + \text{order}(H)$$

[Chap. 5].

Assumptions on the Noise

AN1. The noise sequences $\tilde{u}(t)$, $\tilde{y}(t)$ are stationary random processes, with zero mean values and spectra $\phi_{\tilde{u}}(\omega)$ and $\phi_{\tilde{y}}(\omega)$, respectively. Further, $\tilde{u}(t)$ and $\tilde{y}(t)$ are mutually uncorrelated [Chap. 3].
AN2. The measurement noises are Gaussian distributed [Chap. 4].

Naturally, at most one of the next three noise assumptions applies at a given situation.
AN3a. Both $\tilde{y}(t)$ and $\tilde{u}(t)$ are ARMA processes, as in (4.25) and (4.26) [Chap. 4].
AN3b. The output noise $\tilde{y}(t)$ is an ARMA process, while the input noise $\tilde{u}(t)$ is white. This means that $n_k = n_m = 0$ in (4.26) [Chap. 4].
AN3c. Both $\tilde{y}(t)$ and $\tilde{u}(t)$ are white noise sequences. This means that $n_f = n_h = 0$ in (4.25) and $n_k = n_m = 0$ in (4.26) [Chap. 4].
AN4. Both $\tilde{y}(t)$ and $\tilde{u}$ are white noise sequences. The ratio of their variances, $r = \lambda_y/\lambda_u$, is known [Chap. 4].

Assumptions on the Noise-free Input

AI1. The true input $u_0(t)$ is a stationary process of zero mean, with spectral density $\phi_{u_0}(\omega)$. The input $u_0(t)$ is assumed to be persistently exciting of a suitable order, which means that $\phi_{u_0}(\omega) > 0$ for a sufficient number of frequencies [Chap. 3].
AI2. The input $u_0(t)$ is uncorrelated with the measurement noise sources $\tilde{u}(t)$ and $\tilde{y}(t)$ [Chap. 4].
AI3. The true input $u_0(t)$ is Gaussian distributed [Chap. 4].
AI4. The true input $u_0(t)$ is an ARMA process; that is, it can be modeled as

$$D(q^{-1})u_0(t) = C(q^{-1})e(t) \ ,$$

where $e(t)$ is a white noise signal [Chap. 4].
AI5. The noise-free signal $u_0(t)$ is a periodic function. The length of the period is denoted N. It is assumed that M periods of the data $u(t), y(t)$ are available. Hence the total data length is NM. In each period $u_0(t)$ is a stationary process [Chap. 12].
AI6. The measurement noise signals $\tilde{u}(t)$ and $\tilde{y}(t)$ are uncorrelated with the noise-free input $u_0(s)$ for all t and s. Further, the measurement noise signals within different periods are uncorrelated [Chap. 12].

Assumptions on the Experimental Conditions

AE1. The data comes from one (single) experiment [Chap. 3].
AE2a. There is more than one experiment. The spectrum of the noise-free input is different in the different experiments [Chap. 4].
AE2b. There is more than one experiment. The measurement noises $\tilde{u}(t)$, $\tilde{y}(t)$ are uncorrelated between different experiments. The true noise-free input $u_0(t)$ is correlated between the experiments [Chap. 4].
AE3. The noise-free signal $u_0(t)$ is a periodic function. It is assumed that M periods of the data $u(t), y(t)$ are available. In each period $u_0(t)$ is a stationary process [Chap. 12].

Applicability

Many assumptions are assumed to be generally valid throughout the text, while a few are valid only locally, when explicitly stated so. The table below summarizes the status of the assumptions.

Validity **Topic**	General	Default	Locally
System	**AS1**, **AS2**	**AS3**	**AS4**, **AS5**, **AS6**
Noise	**AN1**	**AN3c**	**AN2**, **AN3a**, **AN3b**, **AN4**
Noise-free input	**AI1**, **AI2**		**AI3**, **AI4**, **AI5**
Experiment		**AE1**	**AE2a**, **AE2b**, **AE3**