Lecture Notes
in Control and Information Sciences 310

Editors: M. Thoma · M. Morari

Lecture Notes in Control and Information Sciences

Edited by M. Thoma and M. Morari

Further volumes of this series are listed at the end of the book or found on our homepage: springeronline.com

Vol. 309: Kumar, V.; Leonard, N.; Morse, A.S. (Eds.)
Cooperative Control
301 p. 2005 [3-540-22861-6]

Vol. 308: Tarbouriech, S.; Abdallah, C.T.; Chiasson, J. (Eds.)
Advances in Communication Control Networks
358 p. 2005 [3-540-22819-5]

Vol. 307: Kwon, S.J.; Chung, W.K.
Perturbation Compensator based Robust Tracking Control and State Estimation of Mechanical Systems
158 p. 2004 [3-540-22077-1]

Vol. 306: Bien, Z.Z.; Stefanov, D. (Eds.)
Advances in Rehabilitation
472 p. 2004 [3-540-21986-2]

Vol. 305: Nebylov, A.
Ensuring Control Accuracy
256 p. 2004 [3-540-21876-9]

Vol. 304: Margaris, N.I.
Theory of the Non-linear Analog Phase Locked Loop
303 p. 2004 [3-540-21339-2]

Vol. 303: Mahmoud, M.S.
Resilient Control of Uncertain Dynamical Systems
278 p. 2004 [3-540-21351-1]

Vol. 302: Filatov, N.M.; Unbehauen, H.
Adaptive Dual Control: Theory and Applications
237 p. 2004 [3-540-21373-2]

Vol. 301: de Queiroz, M.; Malisoff, M.; Wolenski, P. (Eds.)
Optimal Control, Stabilization and Nonsmooth Analysis
373 p. 2004 [3-540-21330-9]

Vol. 300: Nakamura, M.; Goto, S.; Kyura, N.; Zhang, T.
Mechatronic Servo System Control
Problems in Industries and their Theoretical Solutions
212 p. 2004 [3-540-21096-2]

Vol. 299: Tarn, T.-J.; Chen, S.-B.; Zhou, C. (Eds.)
Robotic Welding, Intelligence and Automation
214 p. 2004 [3-540-20804-6]

Vol. 298: Choi, Y.; Chung, W.K.
PID Trajectory Tracking Control for Mechanical Systems
127 p. 2004 [3-540-20567-5]

Vol. 297: Damm, T.
Rational Matrix Equations in Stochastic Control
219 p. 2004 [3-540-20516-0]

Vol. 296: Matsuo, T.; Hasegawa, Y.
Realization Theory of Discrete-Time Dynamical Systems
235 p. 2003 [3-540-40675-1]

Vol. 295: Kang, W.; Xiao, M.; Borges, C. (Eds)
New Trends in Nonlinear Dynamics and Control, and their Applications
365 p. 2003 [3-540-10474-0]

Vol. 294: Benvenuti, L.; De Santis, A.; Farina, L. (Eds)
Positive Systems: Theory and Applications (POSTA 2003)
414 p. 2003 [3-540-40342-6]

Vol. 293: Chen, G. and Hill, D.J.
Bifurcation Control
320 p. 2003 [3-540-40341-8]

Vol. 292: Chen, G. and Yu, X.
Chaos Control
380 p. 2003 [3-540-40405-8]

Vol. 291: Xu, J.-X. and Tan, Y.
Linear and Nonlinear Iterative Learning Control
189 p. 2003 [3-540-40173-3]

Vol. 290: Borrelli, F.
Constrained Optimal Control of Linear and Hybrid Systems
237 p. 2003 [3-540-00257-X]

Vol. 289: Giarré, L. and Bamieh, B.
Multidisciplinary Research in Control
237 p. 2003 [3-540-00917-5]

Vol. 288: Taware, A. and Tao, G.
Control of Sandwich Nonlinear Systems
393 p. 2003 [3-540-44115-8]

Vol. 287: Mahmoud, M.M.; Jiang, J. and Zhang, Y.
Active Fault Tolerant Control Systems
239 p. 2003 [3-540-00318-5]

Vol. 286: Rantzer, A. and Byrnes C.I. (Eds)
Directions in Mathematical Systems Theory and Optimization
399 p. 2003 [3-540-00065-8]

Vol. 285: Wang, Q.-G.
Decoupling Control
373 p. 2003 [3-540-44128-X]

Vol. 284: Johansson, M.
Piecewise Linear Control Systems
216 p. 2003 [3-540-44124-7]

Vol. 283: Fielding, Ch. et al. (Eds)
Advanced Techniques for Clearance of Flight Control Laws
480 p. 2003 [3-540-44054-2]

Vol. 282: Schröder, J.
Modelling, State Observation and Diagnosis of Quantised Systems
368 p. 2003 [3-540-44075-5]

A. Janczak

Identification of Nonlinear Systems Using Neural Networks and Polynomial Models

A Block-Oriented Approach

With 79 Figures and 22 Tables

Springer

Author

Prof. Andrzej Janczak
University of Zielona Góra
Institute of Control and Computation Engineering
ul. Podgorna 50
65-246 Zielona Góra
Poland

ISSN 0170-8643

ISBN 3-540-23185-4 **Springer Berlin Heidelberg New York**

Library of Congress Control Number: 2004097177

Springer is a part of Springer Science+Business Media

springeronline.com

Typesetting: Data conversion by the authors.
Final processing by PTP-Berlin Protago-TEX-Production GmbH, Germany
Cover-Design: design & production GmbH, Heidelberg
Printed on acid-free paper 62/3020Yu - 5 4 3 2 1 0

Preface

The identification of nonlinear systems using the block-oriented approach has been developed since the half of 1960s. A large amount of knowledge on this subject has been accumulated through literature. However, publications are scattered over many papers and there is no book which presents the subject in a unified framework. This has created an increasing need to systemize the existing identification methods and along with a presentation of some original results have been the main incentive to write this book. In writing the book, an attempt has been made at the presentation of some new ideas concerning the model parameter adjusting with gradient-based techniques.

Two types of models, considered in this book, use neural networks and polynomials as representations of Wiener and Hammerstein systems. The focus is placed on Wiener and Hammerstein models in which the nonlinear element is represented by a polynomial or a two-layer perceptron neural network with hyperbolic tangent hidden layer nodes and linear output nodes. Pulse transfer function models are common representations of system dynamics in both neural network and polynomial Wiener and Hammerstein models.

Neural network and polynomial models reveal different properties such as the approximation accuracy, computational complexity, available parameter and structure optimization methods, etc. All these differences make them complementary in solving many practical problems. For example, it is well known that the approximation of some nonlinear functions requires polynomials of a high order and this, in turn, results in a high parameter variance error. The approximation with neural network models is an interesting alternative in such cases.

The book results mainly from my research in the area of nonlinear system identification that have been performed since 1995. Two exceptions from this rule are Chapter 1, containing the introductory notes, and Chapter 5, which reviews the well-known Hammerstein system identification methods based on polynomial models of the nonlinearity. In writing the book, an emphasis has been put on presenting various identification methods, which are applicable to both neural network and polynomial models of Wiener and Hammerstein systems, in a unified framework.

The book starts with a survey of discrete-time models of time-invariant dynamic systems. Then the multilayer perceptron neural network is introduced and a brief review of the existing methods for the identification of Wiener and Hammerstein systems is presented. Two subsequent Chapters (2 and 3) introduce neural network models of Wiener and Hammerstein systems and present different algorithms for the calculation of the gradient or the approximate gradient of the model output w.r.t. model parameters. For both Wiener and Hammerstein models, the accuracy of gradient evaluation with the truncated backpropagation through time algorithm is analyzed. The discussion also includes advantages and disadvantages of the algorithms in terms of their approximation accuracy, computational requirements, and weight updating methods. Next, in Chapter 4, we present identification methods, which use polynomial models of Wiener systems. The parameters of the linear dynamic system and the inverse nonlinearity are estimated with the least squares method, and a combined least squares and instrumental variables approach. To estimate parameters of the noninverted nonlinearity, the recursive prediction error and the pseudolinear regression methods are proposed. Then the existing identification methods based on polynomial Hammerstein models are reviewed and presented in Chapter 5. Wiener and Hammerstein models are two examples of block-oriented models which have found numerous industrial applications. The most important of them, including nonlinear system modelling, control, and fault detection and isolation, are reviewed in Chapter 6. This chapters presents also two applications of Wiener and Hammerstein models – estimation of system parameter changes, and modelling vapor pressure dynamics in a five stage sugar evaporation station.

The book contains the results od research conducted by the author with the kind support of the State Committee for Scientific Research in Poland under the grant No 4T11A01425 and with the additional support of the European Union within the 5th Framework Programme under *DAMADICS* project No HPRN-CT-2000-00110.

A few people have contributed directly or indirectly to this book. First of all, I would like to express my sincere gratitude to Professor Józef Korbicz for his constant help, advice, encouragement, and support.
I am very grateful to Ms Agnieszka Rożewska for proofreading and linguistic advice on the text. I highly appreciate the help of my colleagues from the Institute of Control and Computational Engineering, University of Zielona Góra.

Zielona Góra, *Andrzej Janczak*
July 2004

Contents

Symbols and notation

a_k	kth parameter of the polynomial $A(q^{-1})$
$\hat{a}_k$	kth parameter of the polynomial $\hat{A}(q^{-1})$
$A(q^{-1})$	denominator polynomial of the system pulse transfer function
$\hat{A}(q^{-1})$	denominator polynomial of the model pulse transfer function
$\mathcal{A}(q^{-1})$	denominator polynomial of the faulty system pulse transfer function
$\mathbf{A}$	state space matrix, $(nx \times nx)$
$\hat{\mathbf{A}}^{(m)}$	mth parameter matrix of the MIMO dynamic model. The MIMO Wiener model, $(ns \times ns)$; the MIMO Hammerstein model, $(ny \times ny)$
b_k	kth parameter of the polynomial $B(q^{-1})$
$\hat{b}_k$	kth parameter of the polynomial $\hat{B}(q^{-1})$
$B(q^{-1})$	nominator polynomial of the system pulse transfer function
$\hat{B}(q^{-1})$	nominator polynomial of the model pulse transfer function
$\mathcal{B}(q^{-1})$	nominator polynomial of the faulty system pulse transfer function
$\mathbf{B}$	control matrix, $(nx \times nu)$
$\hat{\mathbf{B}}^{(m)}$	mth parameter matrix of the MIMO dynamic model. The MIMO Wiener model, $(ns \times nu)$; the MIMO Hammerstein model, $(ny \times nf)$
$\mathbf{C}$	observation matrix, $(ny \times nx)$
$\mathbf{D}$	matrix describing the effect of inputs on outputs, $(ny \times nu)$
$e(n)$	identification error (one-step-ahead prediction error)
E	mathematical expectation
$f(\cdot)$	nonlinear function of the system
$\hat{f}(\cdot)$	nonlinear function of the model
$\hat{f}_k(\cdot)$	kth nonlinear function of the MIMO nonlinear element model
$\hat{\mathbf{f}}(\cdot)$	nonlinear function of the MIMO nonlinear element model
$f^{-1}(\cdot)$	inverse nonlinear function of the system
$\hat{f}^{-1}(\cdot)$	inverse nonlinear function of the model
$\hat{g}_k(\cdot)$	kth nonlinear function of the MIMO inverse nonlinear element model
$\hat{\mathbf{g}}(\cdot)$	nonlinear function of the MIMO inverse nonlinear element model
$h(n)$	system impulse response
$h_1(n)$	impulse response of the sensitivity model
$h_2(n)$	impulse response of the linear dynamic model
$H_1(q^{-1})$	pulse transfer function of sensitivity models
$H_2(q^{-1})$	pulse transfer function of the linear dynamic model
J	global error function
$J(n)$	local error function
K	number of unfolded time steps
m_f	expected value of $\hat{f}(u(n), \mathbf{w})$

m_{w_c}	expected value of $\partial\hat{f}(u(n),\mathbf{w})/\partial w_c$
M	number of nonlinear nodes
n	discrete time
na	order of the polynomials $A(q^{-1})$ and $\hat{A}(q^{-1})$
nb	order of the polynomials $B(q^{-1})$ and $\hat{B}(q^{-1})$
nf	number of outputs of MIMO Hammerstein nonlinear element model
ns	number of outputs of MIMO Wiener linear dynamic model;
nu	number of system (model) inputs
ny	number of system (model) outputs
N	number of input-output measurements
q^{-1}	backward shift operator
$\mathbb{R}^d$	Euclidean d-dimensional space
$s(n)$	output of the linear dynamic part of Wiener system
$\hat{s}_k(n)$	kth output of the linear dynamic part of MIMO Wiener model
$\hat{s}(n)$	output of the linear dynamic part of Wiener model
$\hat{\mathbf{s}}(n)$	output of the linear dynamic part of MIMO Wiener model
$u(n)$	system input
$\mathbf{u}(n)$	MIMO system input
var	variance
$v_{ji}^{(1)}$	ith weight of the jth hidden layer node of the inverse nonlinear element model
$v_{kj}^{(2)}$	jth weight of the kth output node of the inverse nonlinear element model
$\hat{v}(n)$	output of the nonlinear element part of Hammerstein model
$\mathbf{v}$	weight vector of the inverse nonlinear element model
$\mathbf{v}_k$	kth path weight vector of the MIMO inverse nonlinear element model
$\mathbf{V}$	weight vector of the MIMO inverse nonlinear element model
$w_{ji}^{(1)}$	ith weight of the jth hidden layer node of the nonlinear element model
$w_{kj}^{(2)}$	jth weight of the kth output node of the nonlinear element model
$\mathbf{w}$	weight vector of the nonlinear element model
$\mathbf{w}_k$	kth path weight vector of the MIMO nonlinear element model,
$\mathbf{W}$	weight vector of the MIMO nonlinear element model
$x_j(n)$	activation of the jth nonlinear node
$\mathbf{x}(n)$	regression vector; system state
$\hat{\mathbf{x}}(n)$	model state
$y(n)$	system output
$\hat{y}(n)$	model output
$y_k(n)$	kth output of the MIMO Wiener (Hammerstein) system
$\hat{y}_k(n)$	kth output of the MIMO Wiener (Hammerstein) model
$\hat{y}(n\vert n-1)$	one-step-ahead predictor of $y(n)$
$\mathbf{y}(n)$	MIMO system output
$\hat{\mathbf{y}}(n)$	MIMO model output

$z_j(n)$	activation of the jth nonlinear node of the inverse nonlinear element model
$\mathbf{z}(n)$	instrumental variables vector
$\hat{\gamma}_k$	kth parameter of the polynomial of the inverse nonlinear element model
γ_k	kth parameter of the polynomial of the inverse nonlinear element
$\Delta A(q^{-1})$	change in the pulse transfer function denominator of the linear dynamic system
$\Delta f(\cdot)$	change in the nonlinear function of the nonlinear element
$\Delta f^{-1}(\cdot)$	change in the nonlinear function of the inverse nonlinear element
$\Delta \hat{s}_{\hat{a}_k}(n)$	computation error of $\partial \hat{s}(n)/\partial \hat{a}_k$
$\Delta \hat{s}_{\hat{b}_k}(n)$	computation error of $\partial \hat{s}(n)/\partial \hat{b}_k$
$\Delta \hat{y}_{\hat{a}_k}(n)$	computation error of $\partial \hat{y}(n)/\partial \hat{a}_k$
$\Delta \hat{y}_{\hat{b}_k}(n)$	computation error of $\partial \hat{y}(n)/\partial \hat{b}_k$
$\Delta \hat{y}_{w_c}(n)$	computation error of $\partial \hat{y}(n)/\partial w_c$
$\epsilon(n)$	discrete white noise disturbance
$\varepsilon(n)$	additive system output disturbance
η	learning rate
$\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}$	parameter vector
λ	exponential forgetting factor
μ_k	kth parameter of the polynomial of the nonlinear element,
$\hat{\mu}_k$	kth parameter of the polynomial of the nonlinear element model
$\xi_{a_k}(n)$	calculation accuracy degree of $\partial \hat{s}(n)/\partial \hat{a}_k$ – the Wiener model; $\partial \hat{y}(n)/\partial \hat{a}_k$ – the Hammerstein model
$\xi_{b_k}(n)$	calculation accuracy degree of $\partial \hat{s}(n)/\partial \hat{b}_k$ – the Wiener model; $\partial \hat{y}(n)/\partial \hat{b}_k$ – the Hammerstein model
$\xi_{w_c}(n)$	calculation accuracy degree of $\partial \hat{y}(n)/\partial w_c$
σ^2	variance of $u(n)$
σ_f^2	variance of $\hat{f}(u(n), \mathbf{w})$
$\sigma_{w_c}^2$	variance of $\partial \hat{f}(u(n), \mathbf{w})/\partial w_c$
$\varphi(\cdot)$	nonlinear activation function
$\boldsymbol{\psi}(n)$	gradient of the Wiener model output

List of abbreviations

i.i.d.	independent and identically distributed
r.h.s	right hand side
w.r.t	with respect to
AR	autoregressive
ARMA	autoregressive moving average
ARMAX	autoregressive moving average with exogenous input

ARX	autoregressive with exogenous input
BJ	Box-Jenkins
BP	back propagation
BPP	back propagation for parallel models
BPPT	back propagation through time
BPS	back propagation for series-parallel models
CSTR	continuous stirred tank reactor
DLOP	discrete Legendre orthogonal polynomial
FDI	fault detection and isolation
ELS	extended least squares
FIR	finite impulse response model
IMC	internal model control
IV	instrumental variables
MA	moving average model
MIMO	multiple-input multiple-output
MISO	multiple-input single-output
MLP	multilayer perceptron
MPC	model-based predictive control
MSE	mean square error
NAR	nonlinear autoregressive
NARMA	nonlinear autoregressive moving average
NARMAX	nonlinear autoregressive moving average with exogenous input
NARX	nonlinear autoregressive with exogenous input
NBJ	nonlinear Box-Jenkins
NFIR	nonlinear finite impulse response
NOBF	nonlinear orthonormal basis function
NOE	nonlinear output error
NMA	nonlinear moving average
OBFP	orthogonal basis with fixed poles
OE	output error
PE	prediction error
PI	proportional plus integral
PID	proportional plus integral plus derivative
PRBS	pseudorandom binary sequence
RIV	recursive instrumental variables
RLS	recursive least squares
RELS	recursive extended least squares
RMS	root mean square
RPE	recursive prediction error
RPLR	recursive pseudolinear regression
SIMO	single-input multiple-output
SISO	single-input single-output
SM	sensitivity method
WMPC	Wiener model-based predictive control