Computers and Chemical Engineering 77 (2015) 24–42

Contents lists available at ScienceDirect

Computers and Chemical Engineering

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

Nonlinear ill-posed problem analysis in model-based parameter estimation and experimental design

Diana C. López C.[a,*], Tilman Barz[b], Stefan Körkel[c], Günter Wozny[a]

[a] Technische Universität Berlin, School of Process Sciences, Chair of Process Dynamics and Operation, Sekr. KWT-9, Straße des 17. Juni 135, D-10623 Berlin, Germany
[b] AIT Austrian Institute of Technology GmbH, Giefinggasse 2, 1210 Wien, Austria
[c] Universität Heidelberg, Interdisciplinary Center for Scientific Computing, Im Neuenheimer Feld 368, 69120 Heidelberg, Germany

ARTICLE INFO

Article history:
Received 22 September 2014
Received in revised form 10 February 2015
Accepted 3 March 2015
Available online 12 March 2015

Keywords:
Ill-posed problems
Ill-conditioning analysis
Singular value decomposition
Identifiability problems
Parameter subset selection
Tikhonov regularization

ABSTRACT

Discrete ill-posed problems are often encountered in engineering applications. Still, their sound analysis is not yet common practice and difficulties arising in the determination of uncertain parameters are typically not assigned properly. This contribution provides a tutorial review on methods for identifiability analysis, regularization techniques and optimal experimental design. A guideline for the analysis and classification of nonlinear ill-posed problems to detect practical identifiability problems is given. Techniques for the regularization of experimental design problems resulting from ill-posed parameter estimations are discussed. Applications are presented for three different case studies of increasing complexity.

1. Introduction

Typical causes of poor or even no identifiability in nonlinear (and linear) system identification are: over-parameterized process models, model structures not matching the system, limited experimental information (either in quantity or quality), correlations in parameters or experimental data, parameters with little or null influence on observed variables and inappropriate initial parameter guesses (Bard, 1974; Bates and Watts, 1988; Franceschini and Macchietto, 2008; Vajda et al., 1989; Walter and Pronzato, 1996). This generates problems in parameter estimation (PE), e.g., multiple, meaningless, inaccurate or unstable solutions, and/or convergence and numerical problems in the solver. Those problems lead for instance, to a model with poor extrapolation properties or parameters with large variance or uncertainty (the so called non-identifiable parameters).

According to Hadamard (1923), a well-posed problem means existence and uniqueness of its solution, as well as continuous dependence of the solution on the data. In other words, if the solution does not exist, or if it is not unique, or if a small change of the experimental data produces a large perturbation of the solution, the problem is ill-posed (Bertero et al., 1988; Hansen, 1998, 2007; Hoerl and Kennard, 1970a; Marquardt, 1970). Continuity is related to the requirement of stability or robustness of the solution (Bertero et al., 1988). However, this condition is necessary for stability, but it is not sufficient as especially for discrete ill-posed problems (coming from discrete available data or discretization of ill-posed problems) lack of robustness against noise is usually evidenced (Bertero et al., 1988). Accordingly in PE, besides the strong relationship between identifiability and ill-posedness of a problem, the error propagation from data to the solution should also be addressed.

An analysis of the ill-posedness of a PE problem could be related to the ill-conditioning and thus, to the numerical stability and linear dependencies of a specific matrix, in the sense that noisy measurement data may lead to significant misinterpretations of the solution (Kaltenbacher et al., 2008). In linear estimation (e.g. the linear regression model $y = X\beta + \varepsilon$) the problem could be considered ill-conditioned and then ill-posed if the data matrix X is ill-conditioned (Belsley et al., 1980; Hansen, 1998, 2007; Hoerl and Kennard, 1970a; Marquardt, 1970). This ill-conditioning (also called collinearity) is of paramount importance to the efficacy of least-squares estimation though it is a non-statistical problem (Belsley et al., 1980). For instance, in the presence of collinearity the uniqueness of the least-squares estimator may not be guaranteed.

* Corresponding author. Tel.: +49 3031421634.
E-mail address: diana.lopez@mailbox.tu-berlin.de (D.C. López C.).

http://dx.doi.org/10.1016/j.compchemeng.2015.03.002

List of abbreviations

DAEs	differential and algebraic equations
FIM	Fisher information matrix
IG	initial guess
None	problem without regularization
ODEs	ordinary differential equations
OED	optimal experimental design
PE	parameter estimation
QRP	QR decomposition with column pivoting
Reg	regularization technique either SsS, TSVD or Tikh
SsS	identifiable parameter subset selection based-regularization
SVD	singular value decomposition
SVs	singular value spectrum
Tikh	Tikhonov based-regularization
TSVD	truncated singular value decomposition based-regularization

List of symbols

C	parameter variance-covariance matrix without regularization
C^{Reg}	Parameter variance-covariance matrix with regularization *Reg*
C_M	measurement error variance-covariance matrix
f	set of DAE representing the process model
h	set of relations between the measured variables and the dependent state variables
H_θ	simplified Hessian matrix (FIM)
H_θ^{Reg}	simplified Hessian matrix (FIM) after applying the regularization *Reg*
J_θ	Jacobian matrix
J_θ^{Reg}	Jacobian matrix after applying the regularization *Reg*
L	operator matrix in Tikhonov regularization, i.e., *Reg* = *Tikh*
Nm	number of experimental data sampling times
Nmu	number of input variable switching times
N_{mod}	number of available models
Np	number of parameters
Ny	number of measured response variables
Nx	number of dependent state variables
OF	objective function of PE
P	permutation matrix of QRP decomposition
$PD(Np)$	subspace of the positive definite matrices with dimension $Np \times Np$
$PSD(Np)$	subspace of positive semi-definite matrices with dimension $Np \times Np$
Q	orthogonal matrix of QRP
r_ϵ	numerical ϵ–rank
R	upper triangular matrix with decreasing diagonal elements of QRP
S	sensitivity matrix
$\tilde{S}$	scaled sensitivity matrix
$\tilde{S}^{Reg}$	scaled sensitivity matrix after applying regularization *Reg*
S_ν	rectangular diagonal matrix with singular values on the diagonal of SVD
$Sym(Np)$	space of symmetric matrices with dimension $Np \times Np$
t	independent variable time
t_k	discrete time instance kth
u	input action or experimental design vector
u_{IG}	initial guess of experimental design vector
$\hat{u}_{crit}$	optimal experimental design vector after computing OED with crit = {A, D, E}
U	real orthogonal matrix of SVD
$var(\hat{\theta}_k)$	parameter variance of parameter estimate $\hat{\theta}_k$
V	real orthogonal matrix of SVD
x	dependent state variable
x_0	initial conditions of the dependent state variable
y	predicted response variable
y^m	measured response variable
$\bar{y}$	normalized predicted response variable
Y^m	total experimental data vector
Y	total predicted response vector
Z	PE residual vector
α	scaling factor
$\gamma(A)$	collinearity index of matrix A = $\{\tilde{S}, \tilde{S}^{Reg}\}$
γ^{max}	maximum collinearity index
ε	measurement error
ϵ	singular value threshold
ϵ_κ	singular value threshold w.r.t. maximum condition number
ϵ_γ	singular value threshold w.r.t. maximum collinearity index
$\kappa(A)$	condition number of matrix A = $\{\tilde{S}, \tilde{S}^{Reg}\}$
$\kappa^{max}(A)$	maximum condition number of matrix A = $\{S, \tilde{S}\}$
θ	parameter vector
θ_{IG}	initial guess of parameter vector
$\hat{\theta}$	unbiased parameter estimate
$\tilde{\theta}^{(Np-r_\epsilon)}$	unidentifiable parameter vector after applying regularization *Reg* = *SsS*
$\tilde{\theta}^{(r_\epsilon)}$	identifiable parameter vector after applying regularization *Reg* = *SsS*
$\hat{\theta}^{(Reg)}$	regularized parameter estimate after applying regularization *Reg*
$\bar{\theta}$	normalized parameter vector
$\lambda(A)$	eigenvalue of matrix A = $\{C, H_\theta\}$
λ	Tikhonov regularization parameter
π_i	parameter variance-decomposition proportion associated with singular value ς_i
ρ	step size
σ	standard deviation of measurements
σ_i	standard deviation of measurement i
σ_i^2	variance of measurement i
ς_i	singular value ith
Σ_y	measurement error standard deviation matrix
υ	step direction
Φ^{LSQ}	weighted nonlinear least squares criterion
Φ^{Reg}	weighted nonlinear least squares criterion after applying regularization *Reg*
$\mathfrak{I}$	data sampling time grid
$\mathfrak{I}^u$	input action time grid
Ψ	information function in OED
Ψ^{crit}	information function in OED (alphabetic criteria in OED with crit = {A, D, E})
$\Omega(\theta)$	discrete smoothing norm functional in *Reg* = *Tikh*

In least-squares estimation for nonlinear systems ill-conditioning is locally analyzed based on the sensitivity matrix (S) of the predicted output trajectories with respect to the parameters (Asprey and Macchietto, 2000; Bard, 1974; Bates and Watts, 1988; Franceschini and Macchietto, 2008; Jacobson, 1985; Johansen, 1997; Marquardt, 1970; Vajda et al., 1989). If the sensitivity matrix is ill-conditioned, the PE problem is locally ill-posed, otherwise it is locally well-posed. For ill-conditioned problems a distinction is