

# Nonlinear Least Squares for Inverse Problems

# Scientific Computation

---

## **Editorial Board**

J.-J. Chattot, Davis, CA, USA  
P. Colella, Berkeley, CA, USA  
Weinan E, Princeton, NJ, USA  
R. Glowinski, Houston, TX, USA  
M. Holt, Berkeley, CA, USA  
Y. Hussaini, Tallahassee, FL, USA  
P. Joly, Le Chesnay, France  
H. B. Keller, Pasadena, CA, USA  
J. E. Marsden, Pasadena, CA, USA  
D. I. Meiron, Pasadena, CA, USA  
O. Pironneau, Paris, France  
A. Quarteroni, Lausanne, Switzerland  
and Politecnico of Milan, Italy  
J. Rappaz, Lausanne, Switzerland  
R. Rosner, Chicago, IL, USA  
P. Sagaut, Paris, France  
J. H. Seinfeld, Pasadena, CA, USA  
A. Szepessy, Stockholm, Sweden  
M. F. Wheeler, Austin, TX, USA

G. Chavent

# Nonlinear Least Squares for Inverse Problems

Theoretical Foundations and  
Step-by-Step Guide for Applications

With 25 Figures

 Springer

Guy Chavent  
Ceremade, Université Paris-Dauphine  
75775 Paris Cedex 16  
France

and

Inria-Rocquencourt  
BP 105, 78153 Le Chesnay Cedex  
France  
Guy.Chavent@inria.fr

ISSN 1434-8322  
ISBN 978-90-481-2784-9 e-ISBN 978-90-481-2785-6  
DOI 10.1007/978-90-481-2785-6  
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2009927706

©Springer Science+Business Media B.V. 2009

No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

*Cover design:* eStudio Calamar S.L.

Printed on acid-free paper

Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

To my wife Annette

# Preface

The domain of inverse problems has experienced a rapid expansion, driven by the increase in computing power and the progress in numerical modeling. When I started working on this domain years ago, I became somehow frustrated to see that my friends working on modeling were producing existence, uniqueness, and stability results for the solution of their equations, but that I was most of the time limited, because of the nonlinearity of the problem, to prove that my least squares objective function was differentiable. . . . But with my experience growing, I became convinced that, after the inverse problem has been properly trimmed, the *final least squares problem*, the one solved on the computer, should be *Quadratically (Q)-wellposed*, that is, both *well-posed* and *optimizable*: optimizability ensures that a global minimizer of the least squares function can actually be found using efficient local optimization algorithms, and wellposedness that this minimizer is stable with respect to perturbation of the data.

But the vast majority of inverse problems are nonlinear, and the classical mathematical tools available for their analysis fail to bring answers to these crucial questions: for example, compactness will ensure existence, but provides no uniqueness results, and brings no information on the presence or absence of parasitic local minima or stationary points . . . .

This book is partly a consequence of this early frustration: a first objective is to present a *geometrical theory* for the analysis of NLS problems from the point of view of *Q-wellposedness*: for an attainable set with finite curvature, this theory provides an estimation of the size of the admissible parameter set and of the error level on the data for which Q-wellposedness

holds. The various regularization techniques used to trim the inverse problem can then be checked against their ability to produce the desirable Q-wellposed problems.

The second objective of the book is to give a detailed presentation of important practical issues for the resolution of NLS problems: sensitivity functions and adjoint state methods for the computations of derivatives, choice of optimization parameters (calibration, sensitivity analysis, multiscale and/or adaptive parameterization), organization of the inversion code, and choice of the descent step for the minimization algorithm. Most of this material is seldom presented in detail, because it is quite elementary from the mathematical point of view, and has usually to be rediscovered by trial-and-error!

As one can see from these objectives, this book does not pretend to give an exhaustive panorama of nonlinear inverse problems, but merely to present the author's view and experience on the subject. Alternative approaches, when known, are mentioned and referenced, but not developed. The book is organized in two parts, which can be read independently:

Part I (Chaps. 1–5) is devoted to the step-by-step resolution and analysis of NLS inverse problems. It should be of interest to scientists of various application fields interested in the practical resolution of inverse problems, as well as to applied mathematicians interested also in their analysis. The required background is a good knowledge of Hilbert spaces, and some notions of functional analysis if one is interested in the infinite dimensional examples. The elements of the geometrical theory of Part II, which are necessary for the Q-wellposedness analysis, are presented without demonstration, but in an as-intuitive-as-possible way, at the beginning of Chap. 4, so that it is not necessary to read Part II, which is quite technical.

Part II (Chaps. 6–8) presents the geometric theory of quasi-convex and strictly quasi-convex sets, which are the basis for the results of Chaps. 4 and 5. These sets possess a neighborhood where the projection is well-behaved, and can be recognized by their finite curvature and limited deflection. This part should be of interest to those more interested in the theory of projection on nonconvex sets. It requires familiarity with Hilbert spaces and functional analysis. The material of Part II was scattered in various papers with different notations. It is presented for the first time in this book in a progressive and coherent approach, which benefits from substantial enhancements and simplifications in the definition of strictly quasi-convex sets.

To facilitate a top-to-bottom approach of the subject, each chapter starts with an overview of the concepts and results developed herein – at the price of

some repetition between the overview and the main corpus of the chapter. . . . Also, we have tried to make the index more user-friendly, all indexed words or expressions are emphasized in the text (but not all emphasized words are indexed!).

I express my thanks to my colleagues, and in particular to François Clement, Karl Kunisch, and Hend Benameur for the stimulating discussions we had over all these years, and for the pleasure I found interacting with them.

*March 2009*  
*Lyon*

GUY CHAVENT



# Contents

<b>Preface</b>	<b>vii</b>
<b>I Nonlinear Least Squares</b>	<b>1</b>
<b>1 Nonlinear Inverse Problems: Examples and Difficulties</b>	<b>5</b>
1.1 Example 1: Inversion of Knott–Zoeppritz Equations . . . . .	6
1.2 An Abstract NLS Inverse Problem . . . . .	9
1.3 Analysis of NLS Problems . . . . .	10
1.3.1 Wellposedness . . . . .	10
1.3.2 Optimizability . . . . .	12
1.3.3 Output Least Squares Identifiability and Quadratically Wellposed Problems . . . . .	12
1.3.4 Regularization . . . . .	14
1.3.5 Derivation . . . . .	20
1.4 Example 2: 1D Elliptic Parameter Estimation Problem . . . . .	21
1.5 Example 3: 2D Elliptic Nonlinear Source Estimation Problem	24
1.6 Example 4: 2D Elliptic Parameter Estimation Problem . . . . .	26
<b>2 Computing Derivatives</b>	<b>29</b>
2.1 Setting the Scene . . . . .	30
2.2 The Sensitivity Functions Approach . . . . .	33
2.3 The Adjoint Approach . . . . .	33
2.4 Implementation of the Adjoint Approach . . . . .	38
2.5 Example 1: The Adjoint Knott–Zoeppritz Equations . . . . .	41

2.6	Examples 3 and 4: Discrete Adjoint Equations . . . . .	46
2.6.1	Discretization Step 1: Choice of a Discretized Forward Map . . . . .	47
2.6.2	Discretization Step 2: Choice of a Discretized Objective Function . . . . .	52
2.6.3	Derivation Step 0: Forward Map and Objective Function	52
2.6.4	Derivation Step 1: State-Space Decomposition . . . . .	53
2.6.5	Derivation Step 2: Lagrangian . . . . .	54
2.6.6	Derivation Step 3: Adjoint Equation . . . . .	56
2.6.7	Derivation Step 4: Gradient Equation . . . . .	58
2.7	Examples 3 and 4: Continuous Adjoint Equations . . . . .	59
2.8	Example 5: Differential Equations, Discretized Versus Discrete Gradient . . . . .	65
2.8.1	Implementing the Discretized Gradient . . . . .	68
2.8.2	Implementing the Discrete Gradient . . . . .	68
2.9	Example 6: Discrete Marching Problems . . . . .	73
<b>3</b>	<b>Choosing a Parameterization</b>	<b>79</b>
3.1	Calibration . . . . .	80
3.1.1	On the Parameter Side . . . . .	80
3.1.2	On the Data Side . . . . .	83
3.1.3	Conclusion . . . . .	84
3.2	How Many Parameters Can be Retrieved from the Data? . . . .	84
3.3	Simulation Versus Optimization Parameters . . . . .	88
3.4	Parameterization by a Closed Form Formula . . . . .	90
3.5	Decomposition on the Singular Basis . . . . .	91
3.6	Multiscale Parameterization . . . . .	93
3.6.1	Simulation Parameters for a Distributed Parameter . . . . .	93
3.6.2	Optimization Parameters at Scale $k$ . . . . .	94
3.6.3	Scale-By-Scale Optimization . . . . .	95
3.6.4	Examples of Multiscale Bases . . . . .	105
3.6.5	Summary for Multiscale Parameterization . . . . .	108
3.7	Adaptive Parameterization: Refinement Indicators . . . . .	108
3.7.1	Definition of Refinement Indicators . . . . .	109
3.7.2	Multiscale Refinement Indicators . . . . .	116
3.7.3	Application to Image Segmentation . . . . .	121
3.7.4	Coarsening Indicators . . . . .	122
3.7.5	A Refinement/Coarsening Indicators Algorithm . . . . .	124

3.8	Implementation of the Inversion . . . . .	126
3.8.1	Constraints and Optimization Parameters . . . . .	126
3.8.2	Gradient with Respect to Optimization Parameters . . . . .	129
3.9	Maximum Projected Curvature: A Descent Step for Nonlinear Least Squares . . . . .	135
3.9.1	Descent Algorithms . . . . .	135
3.9.2	Maximum Projected Curvature (MPC) Step . . . . .	137
3.9.3	Convergence Properties for the Theoretical MPC Step . . . . .	143
3.9.4	Implementation of the MPC Step . . . . .	144
3.9.5	Performance of the MPC Step . . . . .	148
<b>4</b>	<b>Output Least Squares Identifiability and Quadratically Wellposed NLS Problems</b>	<b>161</b>
4.1	The Linear Case . . . . .	163
4.2	Finite Curvature/Limited Deflection Problems . . . . .	165
4.3	Identifiability and Stability of the Linearized Problems . . . . .	174
4.4	A Sufficient Condition for OLS-Identifiability . . . . .	176
4.5	The Case of Finite Dimensional Parameters . . . . .	179
4.6	Four Questions to Q-Wellposedness . . . . .	182
4.6.1	Case of Finite Dimensional Parameters . . . . .	183
4.6.2	Case of Infinite Dimensional Parameters . . . . .	184
4.7	Answering the Four Questions . . . . .	184
4.8	Application to Example 2: 1D Parameter Estimation with $H^1$ Observation . . . . .	191
4.8.1	Linear Stability . . . . .	193
4.8.2	Deflection Estimate . . . . .	198
4.8.3	Curvature Estimate . . . . .	199
4.8.4	Conclusion: OLS-Identifiability . . . . .	200
4.9	Application to Example 4: 2D Parameter Estimation, with $H^1$ Observation . . . . .	200
<b>5</b>	<b>Regularization of Nonlinear Least Squares Problems</b>	<b>209</b>
5.1	Levenberg–Marquardt–Tychonov (LMT) Regularization . . . . .	209
5.1.1	Linear Problems . . . . .	211
5.1.2	Finite Curvature/Limited Deflection (FC/LD) Problems . . . . .	219
5.1.3	General Nonlinear Problems . . . . .	231

5.2	Application to the Nonlinear 2D Source Problem . . . . .	237
5.3	State-Space Regularization . . . . .	246
5.3.1	Dense Observation: Geometric Approach . . . . .	248
5.3.2	Incomplete Observation: Soft Analysis . . . . .	256
5.4	Adapted Regularization for Example 4: 2D Parameter Estimation with $H^1$ Observation . . . . .	259
5.4.1	Which Part of $a$ is Constrained by the Data? . . . . .	260
5.4.2	How to Control the Unconstrained Part? . . . . .	262
5.4.3	The Adapted-Regularized Problem . . . . .	264
5.4.4	Infinite Dimensional Linear Stability and Deflection Estimates . . . . .	265
5.4.5	Finite Curvature Estimate . . . . .	267
5.4.6	OLS-Identifiability for the Adapted Regularized Problem . . . . .	268
<b>II A Generalization of Convex Sets</b>		<b>271</b>
<b>6</b>	<b>Quasi-Convex Sets</b>	<b>275</b>
6.1	Equipping the Set $D$ with Paths . . . . .	277
6.2	Definition and Main Properties of q.c. Sets . . . . .	281
<b>7</b>	<b>Strictly Quasi-Convex Sets</b>	<b>299</b>
7.1	Definition and Main Properties of s.q.c. Sets . . . . .	300
7.2	Characterization by the Global Radius of Curvature . . . . .	304
7.3	Formula for the Global Radius of Curvature . . . . .	316
<b>8</b>	<b>Deflection Conditions for the Strict Quasi-convexity of Sets</b>	<b>321</b>
8.1	The General Case: $D \subset F$ . . . . .	327
8.2	The Case of an Attainable Set $D = \varphi (C)$ . . . . .	337
<b>Bibliography</b>		<b>345</b>
<b>Index</b>		<b>353</b>