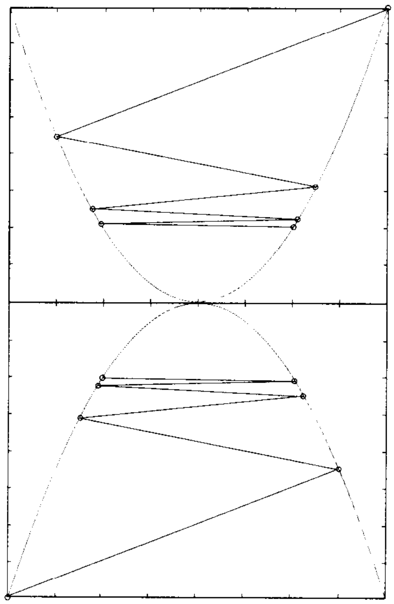
Computational Methods for INVERSE PROBLEMS

Curtis R. Vogel

siam

FRONTIERS IN APPLIED MATHEMATICS

Computational Methods for Inverse Problems



Curtis R. Vogel
Montana State University
Bozeman, Montana

Society for Industrial and Applied Mathematics
Philadelphia

10 9 8 7 6 5 4 3 2 1

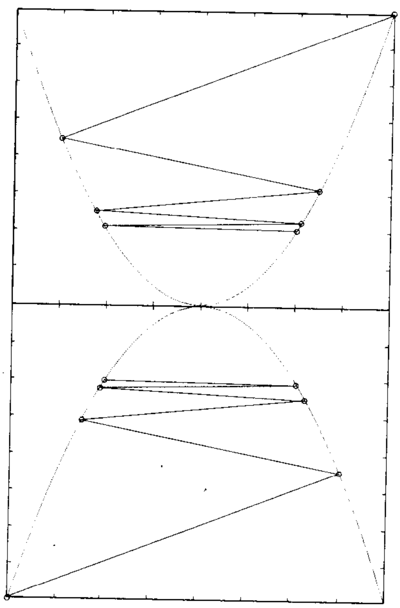
Library of Congress Cataloging-in-Publication Data

Vogel, Curtis R.
Computational methods for inverse problems / Curtis R. Vogel.
p. cm.— (Frontiers in applied mathematics)
Includes bibliographical references and index.
ISBN 0-89871-507-5
1. Inverse problems (Differential equations)—Numerical solutions. I. Title.
II. Series.

QA377 .V575 2002
515'.35 — dc21

2002022386

Computational Methods for Inverse Problems



FRONTIERS IN APPLIED MATHEMATICS

The SIAM series on Frontiers in Applied Mathematics publishes monographs dealing with creative work in a substantive field involving applied mathematics or scientific computation. All works focus on emerging or rapidly developing research areas that report on new techniques to solve mainstream problems in science or engineering.

The goal of the series is to promote, through short, inexpensive, expertly written monographs, cutting edge research poised to have a substantial impact on the solutions of problems that advance science and technology. The volumes encompass a broad spectrum of topics important to the applied mathematical areas of education, government, and industry.

BOOKS PUBLISHED IN FRONTIERS IN APPLIED MATHEMATICS

Vogel, Curtis R., *Computational Methods for Inverse Problems*

Lewis, F. L.; Campos, J.; and Selmic, R., *Neuro-Fuzzy Control of Industrial Systems with Actuator Nonlinearities*

Bao, Gang; Cowsar, Lawrence; and Masters, Wen, editors, *Mathematical Modeling in Optical Science*

Banks, H. T.; Buksas, M. W.; and Lin, T., *Electromagnetic Material Interrogation Using Conductive Interfaces and Acoustic Wavefronts*

Oostveen, Job, *Strongly Stabilizable Distributed Parameter Systems*

Griewank, Andreas, *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*

Kelley, C. T., *Iterative Methods for Optimization*

Greenbaum, Anne, *Iterative Methods for Solving Linear Systems*

Kelley, C. T., *Iterative Methods for Linear and Nonlinear Equations*

Bank, Randolph E., *PLTMG: A Software Package for Solving Elliptic Partial Differential Equations. Users' Guide 7.0*

Moré, Jorge J. and Wright, Stephen J., *Optimization Software Guide*

Rüde, Ulrich, *Mathematical and Computational Techniques for Multilevel Adaptive Methods*

Cook, L. Pamela, *Transonic Aerodynamics: Problems in Asymptotic Theory*

Banks, H. T., *Control and Estimation in Distributed Parameter Systems*

Van Loan, Charles, *Computational Frameworks for the Fast Fourier Transform*

Van Huffel, Sabine and Vandewalle, Joos, *The Total Least Squares Problem: Computational Aspects and Analysis*

Castillo, José E., *Mathematical Aspects of Numerical Grid Generation*

Bank, R. E., *PLTMG: A Software Package for Solving Elliptic Partial Differential Equations. Users' Guide 6.0*

McCormick, Stephen F., *Multilevel Adaptive Methods for Partial Differential Equations*

Grossman, Robert, *Symbolic Computation: Applications to Scientific Computing*

Coleman, Thomas F. and Van Loan, Charles, *Handbook for Matrix Computations*

McCormick, Stephen F., *Multigrid Methods*

Buckmaster, John D., *The Mathematics of Combustion*

Ewing, Richard E., *The Mathematics of Reservoir Simulation*

To my father,

Fred Nickey Vogel

Foreword

Inverse problems are ubiquitous in science and engineering and have rightfully received a great deal of attention by applied mathematicians, statisticians, and engineers. Since most inverse problems cannot be solved analytically, computational methods play a fundamental role. The present volume is a research level introduction to a large class of techniques developed over the past several decades to treat inverse problems primarily formulated in the context of convolution-type Fredholm integral equations Kf=d which must be inverted. Discretization (desirable for solution on digital computers) leads to an undesirable ill-posedness in matrix equation inversions. Motivated by an image reconstruction example, the author treats both deterministic and statistical aspects of computational methods (a wide range of numerical optimization techniques are included) with a strong focus on regularization. Statistical aspects are treated in terms of model uncertainty (in K) and measurement error (noisy data d).

It is not surprising that there is a large mathematical literature on inverse problem methods. What might be surprising is that this literature is significantly divided along deterministic/nondeterministic lines. Methods abound in the statistics literature, where generally the models are assumed quite simple (and often even analytically known!) and the emphasis is on treating statistical aspects of fitting models to data. On the other hand, the applied mathematical literature has a plethora of increasingly complex parameterized models (nonlinear ordinary differential equations, partial differential equations, and delay equations) which are treated theoretically and computationally in a deterministic framework with little or no attention to inherent uncertainty in either the modeled mechanisms or the data used to validate the models. The present monograph is a successful attempt to treat certain probabilistic aspects of a class of inverse problems. It is a research monograph and as such is not meant to be a comprehensive treatment of statistical methods in inverse problems. For example, it does not treat models with random parameters in complex systems, mixed or random effects, mixing distributions, etc. (e.g., see M. Davidian and D. Giltinan, Nonlinear Models for Repeated Measurement Data, Monographs on Statistics and Applied Probability 62 (1998), Chapman & Hall/CRC, Boca Raton, FL), or statistical methods (e.g., ANOVA) associated with model validation. It is, however, a most welcome addition and just the first of what the editors hope will be several volumes treating randomness and uncertainty in computational aspects of inverse or parameter estimation problems.

H. T. Banks
Center for Research in Scientific Computation
North Carolina State University
Raleigh, N. C.

Preface

The field of inverse problems has experienced explosive growth in the last few decades. This is due in part to the importance of applications, like biomedical and seismic imaging, that require the practical solution of inverse problems. It is also due to the recent development of powerful computers and fast, reliable numerical methods with which to carry out the solution process. This monograph will provide the reader with a working understanding of these numerical methods. The intended audience includes graduate students and researchers in applied mathematics, engineering, and the physical sciences who may encounter inverse problems in their work.

Inverse problems typically involve the estimation of certain quantities based on indirect measurements of these quantities. For example, seismic exploration yields measurements of vibrations recorded on the earth's surface. These measurements are only indirectly related to the subsurface geological formations that are to be determined. The estimation process is often ill-posed in the sense that noise in the data may give rise to significant errors in the estimate. Techniques known as regularization methods have been developed to deal with this ill-posedness.

The first four chapters contain background material related to inverse problems, regularization, and numerical solution techniques. Chapter 1 provides an informal overview of a variety of regularization methods. Chapter 2 is guided by the philosophy that sensible numerical solutions to discretized problems follow from a thorough understanding of the underlying continuous problems. This chapter contains relevant functional analysis and infinite-dimensional optimization theory. Chapter 3 contains a review of relevant numerical optimization methods. Chapter 4 presents statistics material that pertains to inverse problems. This includes topics like maximum likelihood estimation and Bayesian estimation.

The remaining five chapters address more specialized topics. Emphasis is placed on the two-dimensional image reconstruction problem, which is introduced in Chapter 5. While this problem is quite simple to formulate and to visualize, its solution draws on a variety of fairly sophisticated tools, including mathematical modeling, statistical estimation theory, Fourier transform methods, and large-scale optimization and numerical linear algebra. This problem also provides motivation and a test case for more specialized techniques like total variation regularization (see Chapter 8) and nonnegativity constraints (Chapter 9).

Chapter 6 contains a brief introduction to parameter (i.e., coefficient) identification for differential equations. This topic serves to introduce nonlinear optimization techniques like the Gauss–Newton and Levenberg–Marquardt methods for nonlinear least squares. It also provides motivation for adjoint, or costate, methods for the efficient computation of gradients and higher order derivatives.

Chapter 7 covers the important topic of regularization parameter selection from a statistical perspective. This chapter includes practical implementations as well as a theoretical analysis of several of the more popular regularization parameter selection methods.

Several web-based resources are available (@http://www.siam.org/books/fr23) to make this monograph somewhat interactive. One of these resources is a collection of MATLAB* m-files used to generate many of the examples and figures. This enables readers to conduct their own computational experiments to gain insight and build intuition. It also provides templates for the implementation of regularization methods and numerical solution techniques for other inverse problems. Moreover, it provides some realistic test problems to be used to further develop and test various numerical methods.

Also available on the web are a list of errata and comments and a collection of solutions to some of the exercises.

I would like to express special gratitude to my colleagues and close friends Martin Hanke, Jim Nagy, Bob Plemmons, and Brian Thelen. These individuals have strongly influenced my research career, and they have made many helpful comments and corrections to preliminary versions of the manuscript. I also owe thanks to a number of other colleagues who reviewed parts of the manuscript. These include John Bardsley, Warren Esty, Luc Gilles, Eldad Haber, Per Christian Hansen, Thomas Scofield, and Lisa Stanley. Finally, I would like to thank Tom Banks and the editorial staff at SIAM for their patience and assistance.

Curt Vogel
Bozeman, Montana

*MATLAB is a registered trademark of The MathWorks, Inc., 3 Apple Hill Drive, Natick, MA 01760, USA, tel. 508-647-7000, fax 508-647-7001, info@mathworks.com, http://www.mathworks.com.

Contents

Chapter 1

Introduction

Inverse problems arise in a variety of important applications in science and industry. These range from biomedical and geophysical imaging to groundwater flow modeling. See, for example, [6, 7, 35, 70, 87, 90, 107, 108] and the references therein. In these applications the goal is to estimate some unknown attributes of interest, given measurements that are only indirectly related to these attributes. For instance, in medical computerized tomography, one wishes to image structures within the body from measurements of X-rays that have passed through the body. In groundwater flow modeling, one estimates material parameters of an aquifer from measurements of pressure of a fluid that immerses the aquifer. Unfortunately, a small amount of noise in the data can lead to enormous errors in the estimates. This instability phenomenon is called ill-posedness. Mathematical techniques known as regularization methods have been developed to deal with ill-posedness. This chapter introduces the reader to the concepts ill-posedness and regularization. Precise definitions are given in the next chapter.

1.1 An Illustrative Example

Consider the Fredholm first kind integral equation of convolution type in one space dimension:

$$g(x) = \int_0^1 k(x - x')\, f(x')\, dx' \stackrel{\text{def}}{=} (\mathcal{K} f)(x), \qquad 0 < x < 1. \tag{1.1}$$

This is a one-dimensional version of a model that occurs in two-dimensional optical imaging and is discussed in more detail in Chapter 5. In this application, f represents light source intensity as a function of spatial position, and g represents image intensity. The kernel k characterizes blurring effects that occur during image formation. A kernel that models the long-time average effects of atmospheric turbulence on light propagation is the Gaussian [7, 99]. Its one-dimensional version is

$$k(x) = C\, \exp(-x^2/2\gamma^2), \tag{1.2}$$

where C and γ are positive parameters.

The direct problem, or forward problem, associated with the model equation (1.1) is the following: Given the source f and the kernel k, determine the blurred image g. Figure 1.1 shows the blurred image corresponding to a piecewise smooth source. Since k is a smooth function, the accurate approximation of $g = \mathcal{K}f$ using standard numerical quadrature is straightforward.

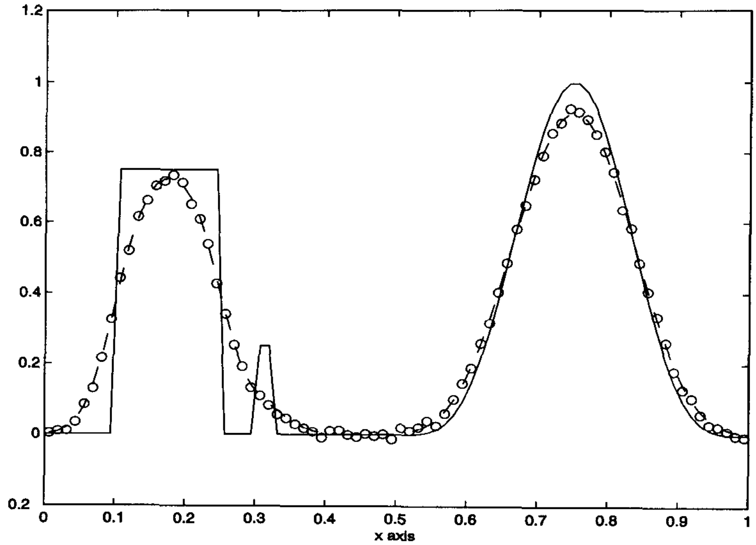


Figure 1.1. *One-dimensional image data. The source function f is represented by the solid line, the blurred image $g = \mathcal{K}f$ is represented by the dashed line, and the discrete noisy data $\mathbf{d}$ is represented by circles. The data were generated according to* (1.1)–(1.4) *with parameters $\gamma = 0.05$ and $C = 1/(\gamma\sqrt{2\pi})$. Midpoint quadrature was used to approximate integrals.*

An associated inverse problem of practical interest is as follows: Given the kernel k and the blurred image g, determine the source f. At first glance, the approximate solution to this inverse problem seems straightforward. One may simply discretize equation (1.1), e.g., using collocation in the independent variable x and quadrature in x', to obtain a discrete linear system $K\mathbf{f} = \mathbf{d}$. For instance, if midpoint quadrature is applied, then K has entries

$$[K]_{ij} = h\ C\ \exp\left(-\frac{((i-j)h)^2}{2\gamma^2}\right), \qquad 1 \le i, j \le n, \tag{1.3}$$

where $h = 1/n$. If the matrix K is nonsingular, one may then compute the discrete approximation $K^{-1}\mathbf{d}$ to f. To obtain an accurate quadrature approximation, n must be relatively large. Unfortunately, the matrix K becomes increasingly ill-conditioned as n becomes large, so errors in $\mathbf{d}$ may be greatly amplified. Certain errors, like those due to quadrature, can be controlled. Others, like the noise in the image recording device, cannot be controlled in a practical setting. Consequently, this straightforward solution approach is likely to fail.

1.2 Regularization by Filtering

Despite ill-conditioning, one can extract some useful information from the discrete linear system $K\mathbf{f} = \mathbf{d}$. To simplify the presentation, consider a discrete data model

$$\mathbf{d} = K\mathbf{f}_{\text{true}} + \eta \tag{1.4}$$

with

$$\delta \stackrel{\text{def}}{=} ||\eta|| > 0. \tag{1.5}$$

Here $||\cdot||$ denotes standard Euclidean norm, $\mathbf{f}_{\text{true}}$ represents the true discretized source, and η represents error in the data. The parameter δ is called the error level. For further simplicity, assume K is an invertible, real-valued matrix. It then has a singular value decomposition (SVD) [46],

$$K = U \operatorname{diag}(s_i)\, V^T, \tag{1.6}$$

with strictly positive decreasing singular values s_i. The SVD and its connection with inverse problems are discussed in more detail in the next chapter; cf. Definition 2.15, and see [58]. At this point we require the following facts: The column vectors $\mathbf{v}_i$ of V, which are called right singular vectors, and the column vectors $\mathbf{u}_i$ of U, which are the left singular vectors, satisfy

$$\mathbf{u}_i^T \mathbf{u}_j = \delta_{ij}, \qquad \mathbf{v}_i^T \mathbf{v}_j = \delta_{ij}, \tag{1.7}$$

$$K\mathbf{v}_i = s_i \mathbf{u}_i, \qquad K^T \mathbf{u}_i = s_i \mathbf{v}_i. \tag{1.8}$$

Here δ_{ij} denotes the Kronecker delta (equation (2.2)), and $U^T = U^{-1}$ and $V^T = V^{-1}$. Note that if K is symmetric and positive definite, then the singular values s_i are the eigenvalues of K, and $U = V$ has columns consisting of orthonormalized eigenvectors. The singular values and vectors for our discretized one-dimensional imaging problem are represented graphically in Figure 1.2.

Using properties (1.7)–(1.8),

$$K^{-1}\mathbf{d} \;=\; V \operatorname{diag}(s_i^{-1})\, U^T \mathbf{d} \;=\; \mathbf{f}_{\text{true}} + \sum_{i=1}^{n} s_i^{-1} (\mathbf{u}_i^T \eta) \mathbf{v}_i. \tag{1.9}$$

Instability arises due to division by small singular values. One way to overcome this instability is to modify the s_i^{-1}'s in (1.9), e.g., by multiplying them by a regularizing filter function $w_\alpha(s_i^2)$ for which the product $w_\alpha(s^2)s^{-1} \to 0$ as $s \to 0$. This filters out singular components of $K^{-1}\mathbf{d}$ corresponding to small singular values and yields an approximation to $\mathbf{f}_{\text{true}}$ with a representation

$$\begin{aligned} \mathbf{f}_\alpha &= V \operatorname{diag}(w_\alpha(s_i^2) s_i^{-1})\, U^T \mathbf{d} \\ &= \sum_{i=1}^{n} w_\alpha(s_i^2) s_i^{-1} (\mathbf{u}_i^T \mathbf{d}) \mathbf{v}_i. \end{aligned} \tag{1.10}$$

To obtain some degree of accuracy, one must retain singular components corresponding to large singular values. This is done by taking $w_\alpha(s^2) \approx 1$ for large values of s^2. An example of such a filter function is

$$w_\alpha(s^2) = \begin{cases} 1 & \text{if} \quad s^2 > \alpha, \\ 0 & \text{if} \quad s^2 \le \alpha. \end{cases} \tag{1.11}$$

The approximation (1.10) then takes the form

$$\mathbf{f}_\alpha = \sum_{s_i^2 > \alpha} s_i^{-1} (\mathbf{u}_i^T \mathbf{d}) \mathbf{v}_i \tag{1.12}$$