ADVANCES IN EXPLORATION GEOPHYSICS

SERIES EDITOR: A.J. BERKHOUT

GLOBAL OPTIMIZATION METHODS IN GEOPHYSICAL INVERSION

M. SEN
AND
P.L. STOFFA

ELSEVIER

ADVANCES IN EXPLORATION GEOPHYSICS 4

GLOBAL OPTIMIZATION METHODS IN GEOPHYSICAL INVERSION

This Page Intentionally Left Blank

ADVANCES IN EXPLORATION GEOPHYSICS 4

GLOBAL OPTIMIZATION METHODS IN GEOPHYSICAL INVERSION

MRINAL SEN

AND

PAUL L. STOFFA

The University of Texas at Austin, Institute for Geophysics
8701 MoPac Expressway, Austin, Texas 78759-8397
USA

ELSEVIER
Amsterdam – Lausanne – New York – Oxford – Shannon – Tokyo 1995

ELSEVIER SCIENCE B.V.
P.O. Box 211,
1000 AE Amsterdam, The Netherlands

Library of Congress Cataloging-in-Publication Data

Sen, Mrinal K.
Global optimization methods in geophysical inversion / Mrinal K.
Sen and Paul L. Stoffa.
p. cm. -- (Advances in exploration geophysics ; 4)
Includes bibliographical references (p. -) and index.
ISBN 0-444-81767-0
1. Geological modeling. 2. Geophysics--Mathematical models.
3. Inverse problems (Differential equations) 4. Mathematical
optimization. I. Stoffa, Paul L., 1948- . II. Title.
III. Series.
QE43.S46 1995
550'.1'13--dc20 95-12651
CIP

ISBN: 0-444-81767-0

This book is printed on acid-free paper.

Printed in The Netherlands.

Preface

One of the major goals of geophysical inversion is to find earth models that explain the geophysical observations. Thus, the branch of mathematics known as *optimization* has found significant use in many geophysical applications. Geophysical inversion in this context involves finding an optimal value of a function of several variables. The function that we want to minimize (or maximize) is a misfit (or fitness) function that characterizes the differences (or similarities) between observed and synthetic data calculated by using an assumed earth model. The earth model is described by physical parameters that characterize the properties of rock layers, such as the compressional wave velocity, shear wave velocity, resistivity, etc.

Both *local* and *global* optimization methods are used in the estimation of material properties from geophysical data. As the title of the book suggests, our goal is to describe the application of several recently developed global optimization methods to geophysical problems. Although we emphasize the application aspects of these algorithms, we describe several parts of the theory in sufficient detail for the readers to understand the underlying fundamental principles upon which these algorithms are based. At this stage we take the opportunity to define some commonly used terms.

For many geophysical applications, the misfit surface as a function of the model parameters which is described by the mismatch between the predicted and observed geophysical data may be highly complicated and characterized by multiple hills and valleys. Thus such a function will have several minima and maxima; the minimum of all the minima is called the *global minimum* and all other minima are called *local minima*. Note that the global minimum is one of the local minima but the converse is not true and it is also possible to have several minima of nearly the same depth. Local optimization or search algorithms such as gradient descent methods typically attempt to find a local minimum in the close neighborhood of the starting solution. Almost all of the local search methods are deterministic algorithms. They use local properties of the misfit function to calculate an update to the current answer and search in the downhill direction. Thus these algorithms will miss the global minimum if the starting solution is nearer to one of the local minima than the global minimum. The local minimum syndrome has plagued geophysicists for over a decade now.

Recently (owing to the advent of powerful and relatively inexpensive computers), global optimization methods have been applied to several

geophysical problems. Unlike local optimization methods, these methods attempt to find the global minimum of the misfit function. Most of the global optimization algorithms are stochastic in nature and use more global information about the misfit surface to update their current position. The convergence of these methods to the globally optimal solution is not guaranteed for all the algorithms. Only for some of the *simulated annealing* algorithms under certain conditions is convergence to the globally optimal solution statistically guaranteed. Also, with real observational data it is never possible to know whether the derived solution corresponds to the global minimum or not. However, our experience indicates that we are able to find many good solutions starting with only poor initial models using global optimization methods.

These global optimization methods are computationally intensive but with the advent of vector computers, parallel computers and powerful desktop workstations, use of these methods is becoming increasingly practical. While finding the optimal solution will always be a goal and the global optimization methods described here are well suited for this purpose, they can also be used to obtain additional information about the nature of the solution. In particular, the description of a solution is not complete without assigning uncertainties to the derived answer. With noisy data it may not even be advisable to search for the so-called global minimum. In these situations a statistical formulation of the inverse problem is often appealing. Consequently, we also describe how global optimization methods can be applied in a statistical framework to estimate the uncertainties in the derived result.

This is not a book on inverse theory *per se;* several excellent texts already exist [e.g., *Menke* 1984; *Tarantola* 1987]. Our goal is to describe in sufficient detail the fundamentals of several optimization methods with application to geophysical inversion such that students, researchers and practitioners will be able to design practical algorithms to solve their specific geophysical inversion problems. We attempted to make this book virtually self-contained so that there are no pre-requisites, except for a fundamental mathematical background that includes a basic understanding of linear algebra and calculus. The material presented in the book can easily be covered in a one-semester graduate level course on geophysical inversion. We believe that after reviewing the materials presented in the book, the readers will be able to develop specific algorithms for their own applications. We will be happy to mail sample Fortran codes of proto-type Metropolis simulated annealing (SA), heat bath SA, very fast simulated annealing (VFSA) and a basic genetic algorithm, to those interested.

Much of our work on nonlinear inversion has been supported by grants from the National Science Foundation, the Office of Naval Research, Cray

Research, Inc., and the Texas Higher Education Coordinating Board. We acknowledge The University of Texas System Center for High Performance Computing for their support and computational resources. Adre Duijndam, Jacob Fokkema, Cliff Frohlich, John Goff, Lester Ingber, Tad Ulrych and Lian-She Zhao reviewed the manuscript and offered valuable suggestions. Milton Porsani who worked with us for one year as a visiting scientist along with several of our graduate students including Faruq Akbar, Carlos Calderon, Raghu Chunduru, Mike Jervis, Vik Sen, Mehmet Tanis, and Carlos Varela, participated in the research and reviewed the manuscript. Their contribution to the project has been extremely valuable. We thank Milo Backus for his many critical comments during the early stages of the work which helped tremendously in shaping our ideas on inverse theory in general. Charlene Palmer receives our thanks for painstakingly typesetting the manuscript. We also thank Gus Berkhout and Jacob Fokkema for inviting us to write the book for the series *Advances in Exploration Geophysics.*

Several figures and parts of the text in Chapter 7 are based on a paper presented at the 1994, EAEG 56th Annual Meeting and Exposition meeting in Vienna, Austria, which have been printed by permission of the copyright holders. The copyright of the paper belongs to the European Association of Geoscientists and Engineers.

Mrinal K. Sen wishes to thank his wife Alo and children Amrita and Ayon for their sacrifice and encouragement. He also thanks Neil Frazer for his suggestions following the 1990 Society of Exploration Geophysicists Meeting to write a book on this subject.

Paul L. Stoffa wishes to thank his wife Donna for her constant support and Gus Berkhout who motivated the writing of this book.

The University of Texas Institute for Geophysics contribution number 1075.

Mrinal K. Sen (mrinal@bullen.ig.utexas.edu)
Paul L. Stoffa (pauls@tau-p.ig.utexas.edu)
Institute for Geophysics
The University of Texas at Austin

This Page Intentionally Left Blank

Contents

This Page Intentionally Left Blank

Chapter 1

Preliminary Statistics

The solution of a geophysical inverse problem can be obtained by a combination of information from observed data, the theoretical relation between data and earth parameters (models) and prior information on data and models. Due to uncertainties in the data and model, probability theory can be used as a tool to describe the inverse problem. Excellent introductory books on the subject of probability theory are those of *Feller* [1968], *Papoulis* [1965] and *Ross* [1989]. In this chapter, we will review probability theory and stochastic processes, the concepts that will be used later to describe the global optimization methods used in geophysical inverse problems. Readers familiar with the subject can proceed directly to Chapter 2.

1.1. Random variables

In simple language, a random variable is a variable that is used to represent the outcome of a random experiment. Familiar random experiments are the tossing of a die and flipping of a coin. When a die is tossed, there are six possible outcomes and it is not certain which one will occur. Similarly when a coin is tossed, there are two possible outcomes and it is not certain which one will occur. The outcome of a random experiment is usually represented by a point, called a sample point s. The set which consists of all possible sample points is called the sample space S. Subsets of S represent certain events such that an event A consists of a certain collection of possible outcomes s. If two subsets contain no points s in common, they are said to be *disjoint*, and the corresponding events are said to be *mutually exclusive*. Formally, any single valued numerical function $X(s)$ defined on a sample space S is called a random variable and a unique real number is associated with each point s.

1.2. Random numbers

Most of the methods of geophysical inversion that we will discuss in this book will use probability or statistical theory that involves studying processes arising from random experiments. This means that we will need to simulate random processes on a computer. In practice this will require an algorithm to generate random numbers. Computer generated random numbers have been used extensively in several applications. Most commonly, sampling methods using computer generated random numbers have been used in situations where the mathematics become intractable. *Metropolis and Ulam* [1940] first proposed the use of random sampling to the solution of deterministic problems such as the evaluation of an integral of the type

$$I = \int_{x_{min}}^{x_{max}} f(x)\, dx \quad . \tag{1.1}$$

If this integral exists, it is given by the expected value of the function. For a random number X that has a uniform distribution over the interval (x_{min}, x_{max}), the above integral can be replaced by the following sum

$$I = \frac{1}{n} \sum_{i=1}^{i=n} f\left(X_i\right) \quad , \tag{1.2}$$

where X_i is a random sample from the uniform distribution. Techniques using this idea are referred to as *Monte Carlo* methods. In general, any method using a random walk is usually included in the category of Monte Carlo methods.

To apply any technique that involves random processes, computer generated random numbers are required. Although a computer is a machine that produces output which is always predictable, computer generated random numbers are used extensively and are called pseudo-random numbers. Most computers use a method called the congruential method to generate random samples from a uniform distribution [*Kennedy and Gentle*, 1980; *Rubinstein*, 1981]. The machines produces a sequence of *pseudo random* integers I_1, I_2, I_3, . . . between 0 and N by the recurrence relation:

$$I_{j+1} = \mathrm{mod}\left(a\, I_j + c, N\right) \quad , \tag{1.3}$$