

# **INTRODUCTION TO DERIVATIVE-FREE OPTIMIZATION**



## MPS-SIAM Series on Optimization

This series is published jointly by the Mathematical Programming Society and the Society for Industrial and Applied Mathematics. It includes research monographs, books on applications, textbooks at all levels, and tutorials. Besides being of high scientific quality, books in the series must advance the understanding and practice of optimization. They must also be written clearly and at an appropriate level.

### Editor-in-Chief

Philippe Toint, *University of Namur (FUNDP)*

### Editorial Board

Oktay Gunluk, *IBM T.J. Watson Research Center*

Matthias Heinkenschloss, *Rice University*

C.T. Kelley, *North Carolina State University*

Adrian S. Lewis, *Cornell University*

Pablo Parrilo, *Massachusetts Institute of Technology*

Daniel Ralph, *University of Cambridge*

Mike Todd, *Cornell University*

Laurence Wolsey, *Université Catholique de Louvain*

Yinyu Ye, *Stanford University*

### Series Volumes

Conn, Andrew R., Scheinberg, Katya, and Vicente, Luis N., *Introduction to Derivative-Free Optimization*

Ferris, Michael C., Mangasarian, Olvi L., and Wright, Stephen J., *Linear Programming with MATLAB*

Attouch, Hedy, Buttazzo, Giuseppe, and Michaille, Gérard, *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*

Wallace, Stein W. and Ziemba, William T., editors, *Applications of Stochastic Programming*

Grötschel, Martin, editor, *The Sharpest Cut: The Impact of Manfred Padberg and His Work*

Renegar, James, *A Mathematical View of Interior-Point Methods in Convex Optimization*

Ben-Tal, Aharon and Nemirovski, Arkadi, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*

Conn, Andrew R., Gould, Nicholas I. M., and Toint, Phillippe L., *Trust-Region Methods*

# INTRODUCTION TO DERIVATIVE-FREE OPTIMIZATION

**Andrew R. Conn**  
**Katya Scheinberg**

IBM  
Yorktown Heights, New York

**Luis N. Vicente**  
Universidade de Coimbra  
Coimbra, Portugal



Society for Industrial and Applied Mathematics  
Philadelphia



Mathematical Programming Society  
Philadelphia

Copyright © 2009 by the Society for Industrial and Applied Mathematics and the Mathematical Programming Society

10 9 8 7 6 5 4 3 2 1

All rights reserved. Printed in the United States of America. No part of this book may be reproduced, stored, or transmitted in any manner without the written permission of the publisher. For information, write to the Society for Industrial and Applied Mathematics, 3600 Market Street, 6th Floor, Philadelphia, PA 19104-2688 USA.

Trademarked names may be used in this book without the inclusion of a trademark symbol. These names are used in an editorial context only; no infringement of trademark is intended.

MATLAB is a registered trademark of The MathWorks, Inc. For MATLAB product information, please contact The MathWorks, Inc., 3 Apple Hill Drive, Natick, MA 01760-2098 USA, 508-647-7000, Fax: 508-647-7101, [info@mathworks.com](mailto:info@mathworks.com), [www.mathworks.com](http://www.mathworks.com).

### Library of Congress Cataloging-in-Publication Data

Conn, Andrew R.

Introduction to derivative-free optimization / Andrew R. Conn, Katya Scheinberg, Luis N. Vicente.

p. cm. — (MPS-SIAM series on optimization ; 8)

Includes bibliographical references and index.

ISBN 978-0-898716-68-9

1. Mathematical models. 2. Mathematical optimization—Industrial applications. 3. Engineering mathematics. 4. Industrial engineering—Mathematics. 5. Search theory. 6. Nonlinear theories. I. Scheinberg, Katya. II. Vicente, Luis N. III. Title.

TA342.C67 2008

519.6—dc22

2008038005

### About the Cover

Cover art © Masha Ryskin, coffee, graphite, ink on board, 61 cm x 61 cm (24" x 24"). Title: *Inside*. Private collection. Photographer: Masha Ryskin.

Masha Ryskin is a Russian-born painter, printmaker, and installation artist who lives and works in the United States. She holds degrees from the Rhode Island School of Design and the University of Michigan. Her work has been exhibited nationally and internationally, and she has worked in the United States, Europe, Costa Rica, and Indonesia. Her art deals with a sense of place, history, and memory through the use of landscape and its elements. *Inside* addresses the concept of landscape as an amalgamation of fragments fitting together to make an integral whole.

Cover art reprinted with permission from the artist.

**siam** is a registered trademark.



is a registered trademark.

(Andrew)

(Katya)

(Luis)

# Contents

<b>Preface</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Why derivative-free optimization . . . . .	1
1.2 Examples of problems where derivatives are unavailable . . . . .	3
1.3 Limitations of derivative-free optimization . . . . .	5
1.4 How derivative-free algorithms should work . . . . .	7
1.5 A short summary of the book . . . . .	11
<b>I Sampling and modeling</b>	<b>13</b>
<b>2 Sampling and linear models</b>	<b>15</b>
2.1 Positive spanning sets and positive bases . . . . .	15
2.2 Gradient estimates used in direct search . . . . .	21
2.3 Linear interpolation and regression models . . . . .	24
2.4 Error bounds for linear interpolation and regression . . . . .	26
2.5 Other geometrical concepts . . . . .	29
2.6 Simplex gradients . . . . .	32
2.7 Exercises . . . . .	33
<b>3 Interpolating nonlinear models</b>	<b>35</b>
3.1 Basic concepts in interpolation . . . . .	36
3.2 Lagrange polynomials . . . . .	39
3.3 $\Lambda$ -poisedness and other measures of well poisedness . . . . .	42
3.4 Condition number as a measure of well poisedness . . . . .	48
3.5 Exercises . . . . .	56
<b>4 Regression nonlinear models</b>	<b>57</b>
4.1 Basic concepts in polynomial least-squares regression . . . . .	58
4.2 Lagrange polynomials in the regression sense . . . . .	60
4.3 $\Lambda$ -poisedness in the regression sense . . . . .	62
4.4 Condition number as a measure of well poisedness . . . . .	68
4.5 Notes and references . . . . .	70
4.6 Exercises . . . . .	72

<b>5</b>	<b>Underdetermined interpolating models</b>	<b>73</b>
5.1	The choice of an underdetermined model . . . . .	74
5.2	Lagrange polynomials and $\Lambda$ -poisedness for underdetermined interpolation . . . . .	76
5.3	Minimum Frobenius norm models . . . . .	80
5.4	Notes and references . . . . .	87
5.5	Exercises . . . . .	87
<b>6</b>	<b>Ensuring well poisedness and suitable derivative-free models</b>	<b>89</b>
6.1	Fully linear and fully quadratic models . . . . .	90
6.2	Ensuring well poisedness using Lagrange polynomials . . . . .	93
6.3	Ensuring well poisedness using pivotal algorithms . . . . .	99
6.4	Practical considerations of geometry improvement algorithms . . . . .	107
6.5	Ensuring well poisedness for regression and minimum Frobenius norm models . . . . .	108
6.6	Other notes and references . . . . .	110
6.7	Exercises . . . . .	112
<b>II</b>	<b>Frameworks and algorithms</b>	<b>113</b>
<b>7</b>	<b>Directional direct-search methods</b>	<b>115</b>
7.1	The coordinate-search method . . . . .	115
7.2	A directional direct-search framework . . . . .	118
7.3	Global convergence in the continuously differentiable case . . . . .	120
7.4	Global convergence in the nonsmooth case . . . . .	124
7.5	Simple decrease with integer lattices . . . . .	127
7.6	The mesh adaptive direct-search method . . . . .	132
7.7	Imposing sufficient decrease . . . . .	134
7.8	Other notes and references . . . . .	135
7.9	Exercises . . . . .	139
<b>8</b>	<b>Simplicial direct-search methods</b>	<b>141</b>
8.1	The Nelder–Mead simplex method . . . . .	141
8.2	Properties of the Nelder–Mead simplex method . . . . .	148
8.3	A globally convergent variant of the Nelder–Mead method . . . . .	149
8.4	Other notes and references . . . . .	161
8.5	Exercises . . . . .	161
<b>9</b>	<b>Line-search methods based on simplex derivatives</b>	<b>163</b>
9.1	A line-search framework . . . . .	163
9.2	Global convergence for first-order critical points . . . . .	165
9.3	Analysis for noise . . . . .	167
9.4	The implicit-filtering algorithm . . . . .	168
9.5	Other simplex derivatives . . . . .	169
9.6	Other notes and references . . . . .	170
9.7	Exercises . . . . .	171

<b>10</b>	<b>Trust-region methods based on derivative-free models</b>	<b>173</b>
10.1	The trust-region framework basics . . . . .	173
10.2	Conditions on the trust-region models . . . . .	179
10.3	Derivative-free trust-region methods (first order) . . . . .	181
10.4	Global convergence for first-order critical points . . . . .	185
10.5	Derivative-free trust-region methods (second order) . . . . .	191
10.6	Global convergence for second-order critical points . . . . .	194
10.7	Model accuracy in larger concentric balls . . . . .	200
10.8	Trust-region subproblem . . . . .	202
10.9	Other notes and references . . . . .	204
10.10	Exercises . . . . .	205
<b>11</b>	<b>Trust-region interpolation-based methods</b>	<b>207</b>
11.1	Common features and considerations . . . . .	207
11.2	The “DFO” approach . . . . .	208
11.3	Powell’s methods . . . . .	211
11.4	Wedge methods . . . . .	215
11.5	Other notes and references . . . . .	225
<b>III</b>	<b>Review of other topics</b>	<b>227</b>
<b>12</b>	<b>Review of surrogate model management</b>	<b>229</b>
12.1	Surrogate modeling . . . . .	229
12.2	Rigorous optimization frameworks to handle surrogate models . . . . .	235
12.3	Exercises . . . . .	240
<b>13</b>	<b>Review of constrained and other extensions to derivative-free optimization</b>	<b>241</b>
13.1	Directional direct-search methods . . . . .	242
13.2	Trust-region interpolation-based methods . . . . .	248
13.3	Derivative-free approaches for global optimization, mixed-integer programming, and other problems . . . . .	249
<b>Appendix</b>	<b>Software for derivative-free optimization</b>	<b>251</b>
	<b>Bibliography</b>	<b>255</b>
	<b>Index</b>	<b>271</b>



# Preface

For many years all three of us have been interested in, and have tried to make contributions to, derivative-free optimization. Our motivation for writing this book resulted from various circumstances. We had the opportunity to work closely with the leading contributors to the field, including especially John Dennis, Michael Powell, Philippe Toint, and Virginia Torczon. We had some knowledge of various facets of the recent developments in the area, and yet felt there was a need for a unified view, and we hoped thereby to gain a better understanding of the field. In addition we were enticed by the growing number of applications. We also felt very strongly that there was a considerable need for a textbook on derivative-free optimization, especially since the foundations, algorithms, and applications have become significantly enriched in the past decade. Finally, although the subject depends upon much that is true for, and was developed for, optimization with derivatives, the issues that arise are new. The absence of computable derivatives naturally prohibits the use of Taylor models—so common in derivative-based optimization. The fact that typical applications involve expensive function evaluations shifts the emphasis from the cost of the linear algebra, or other contributors to the iteration complexity, to simply the number of function evaluations. Also, the noise in the function values affects the local convergence expectations. Thus, the area is both simpler, in the sense of diminished expectations, and harder, in the sense that one is trying to achieve something with considerably less information. It is definitely fun and challenging and, not incidentally, very useful.

Although we do make considerable effort to give a sense of the current state of the art, we do not attempt to present a comprehensive treatise on all the work in the area. This is in part because we think that the subject is not yet mature enough for such a treatise. For similar reasons our emphasis is on the unconstrained problem, although we include a review on the work done so far in constrained derivative-free optimization. The constrained problem is in fact very important for applications, but theoretical treatment of derivative-free methods for constrained problems is very limited in the literature published to date, and thus, for the present volume at least, we are content to concentrate on the unconstrained case.

The book is meant to be reasonably self-contained and is addressed to readers at the level of a graduate student or a senior undergraduate with a background in calculus, linear algebra, and numerical analysis. Some elementary notions of optimization would be helpful but are not necessary. It is certainly our intent that practitioners would find the material covered to be both accessible and reasonably complete for their needs, whether their emphasis is on the algorithms or the applications. We have also made an effort to include figures and exercises when appropriate. The major aims include giving any interested reader a good idea of the state of the art of derivative-free optimization, with a detailed de-

scription of the basic theory to the extent that the reader can well understand what is needed to ensure convergence, how this affects algorithm design, and what kind of success one can expect and where. Thus, it is certainly our goal that the material be of interest to those who want to do research in the area.

As we state in our introduction, due to the growing sophistication and efficiency of computer simulations as well as of other applications, there is an increasing number of instances where one wishes to perform optimization of a complex system and the derivative information of the resulting objective functions is not available. This book is intended to help the reader to study and select, if necessary, suitable approaches to do exactly that. It is also intended to extract and emphasize the common theoretical features used by modern derivative-free algorithms, as well as highlight the differences.

We would be remiss if we ended our preface without some indicator of what the future holds in this area. There is still much waiting to be discovered. Undoubtedly, researchers and practitioners, perhaps soon, will discover ways to tackle much larger problems, whether through the use of massively parallel architectures or through advances in hardware yet to be realized, or through breakthroughs in the theory, or likely all three. The theory of constrained derivative-free optimization is also likely to advance significantly in the near future. Certainly, and especially because of the broad availability of difficult and important applications, this promises to be an exciting, interesting, and challenging area for many years to come.

Aside from the colleagues we mentioned in the first paragraph of this preface, there are many others we would like to thank who have provided us their valuable feedback in specific parts of the book, including, in particular, Natalia Alexandrov, Charles Audet, and Paul Tseng. We are very grateful to Ana Luísa Custódio for her tireless proofreading of the manuscript. We would also like to thank the reviewers who have refereed the version originally submitted to SIAM, in particular Tim Kelley and Jorge Nocedal, for their many interesting comments and suggestions.