
INTRODUCTION TO NONLINEAR AND GLOBAL OPTIMIZATION

Springer Optimization and Its Applications

VOLUME 37

Managing Editor

Panos M. Pardalos (University of Florida)

Editor — Combinatorial Optimization

Ding-Zhu Du (University of Texas at Dallas)

Advisory Board

J. Birge (University of Chicago)

C.A. Floudas (Princeton University)

F. Giannessi (University of Pisa)

H.D. Sherali (Virginia Polytechnic and State University)

T. Terlaky (McMaster University)

Y. Ye (Stanford University)

Aims and Scope

Optimization has been expanding in all directions at an astonishing rate during the last few decades. New algorithmic and theoretical techniques have been developed, the diffusion into other disciplines has proceeded at a rapid pace, and our knowledge of all aspects of the field has grown even more profound. At the same time, one of the most striking trends in optimization is the constantly increasing emphasis on the interdisciplinary nature of the field. Optimization has been a basic tool in all areas of applied mathematics, engineering, medicine, economics, and other sciences.

The *Springer in Optimization and Its Applications* series publishes undergraduate and graduate textbooks, monographs, and state-of-the-art expository work that focus on algorithms for solving optimization problems and also study applications involving such problems. Some of the topics covered include nonlinear optimization (convex and nonconvex), network flow problems, stochastic optimization, optimal control, discrete optimization, multi-objective programming, description of software packages, approximation techniques and heuristic approaches.

INTRODUCTION TO NONLINEAR AND GLOBAL OPTIMIZATION

By

Eligius M.T. Hendrix
Málaga University, Spain

Boglárka G.-Tóth
Budapest University of Technology and Economics, Hungary

 Springer

Eligius M.T. Hendrix
Department of Computer Architecture
Málaga University
Málaga, Spain
Eligius.Hendrix@wur.nl

Boglárka G.-Tóth
Department of Differential Equations
Budapest University of Technology
and Economics
Budapest, Hungary
bog@math.bme.hu

ISSN 1931-6828
ISBN 978-0-387-88669-5 e-ISBN 978-0-387-88670-1
DOI 10.1007/978-0-387-88670-1
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2010925226

Mathematics Subject Classification (2010): 49-XX, 90-XX, 90C26

© Springer Science+Business Media, LLC 2010

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Preface	ix
1 Introduction	1
1.1 Optimization view on mathematical models	1
1.2 NLP models, black-box versus explicit expression	3
2 Mathematical modeling, cases	7
2.1 Introduction	7
2.2 Enclosing a set of points	7
2.3 Dynamic decision strategies	10
2.4 A black box design; a sugar centrifugal screen	13
2.5 Design and factorial or quadratic regression	15
2.6 Nonlinear optimization in economic models	17
2.6.1 Spatial economic-ecological model	18
2.6.2 Neoclassical dynamic investment model for cattle ranching	19
2.6.3 Several optima in environmental economics	19
2.7 Parameter estimation, model calibration, nonlinear regression .	20
2.7.1 Learning of neural nets seen as parameter estimation ..	24
2.8 Summary and discussion points	26
2.9 Exercises	27
3 NLP optimality conditions	31
3.1 Intuition with some examples	31
3.2 Derivative information	35
3.2.1 Derivatives	36
3.2.2 Directional derivative	36
3.2.3 Gradient	37
3.2.4 Second-order derivative	38
3.2.5 Taylor	40
3.3 Quadratic functions	41

3.4	Optimality conditions, no binding constraints	45
3.4.1	First-order conditions	45
3.4.2	Second-order conditions	46
3.5	Optimality conditions, binding constraints	48
3.5.1	Lagrange multiplier method	49
3.5.2	Karush–Kuhn–Tucker conditions	52
3.6	Convexity	54
3.6.1	First-order conditions are sufficient	56
3.6.2	Local minimum point is global minimum point	57
3.6.3	Maximum point at the boundary of the feasible area	59
3.7	Summary and discussion points	60
3.8	Exercises	60
3.9	Appendix: Solvers for Examples 3.2 and 3.3	64
4	Goodness of optimization algorithms	67
4.1	Effectiveness and efficiency of algorithms	67
4.1.1	Effectiveness	68
4.1.2	Efficiency	69
4.2	Some basic algorithms and their goodness	70
4.2.1	Introduction	70
4.2.2	NLP local optimization: Bisection and Newton	71
4.2.3	Deterministic GO: Grid search, Piyavskii–Shubert	74
4.2.4	Stochastic GO: PRS, Multistart, Simulated Annealing	78
4.3	Investigating algorithms	84
4.3.1	Characteristics	85
4.3.2	Comparison of algorithms	87
4.4	Summary and discussion points	88
4.5	Exercises	89
5	Nonlinear Programming algorithms	91
5.1	Introduction	91
5.1.1	General NLP problem	91
5.1.2	Algorithms	91
5.2	Minimizing functions of one variable	93
5.2.1	Bracketing	93
5.2.2	Bisection	94
5.2.3	Golden Section search	95
5.2.4	Quadratic interpolation	97
5.2.5	Cubic interpolation	99
5.2.6	Method of Newton	100
5.3	Algorithms not using derivative information	101
5.3.1	Method of Nelder and Mead	102
5.3.2	Method of Powell	105
5.4	Algorithms using derivative information	106
5.4.1	Steepest descent method	107

5.4.2	Newton method	108
5.4.3	Conjugate gradient method	109
5.4.4	Quasi-Newton method	111
5.4.5	Inexact line search	113
5.4.6	Trust region methods	115
5.5	Algorithms for nonlinear regression	118
5.5.1	Linear regression methods	118
5.5.2	Gauss–Newton and Levenberg–Marquardt	120
5.6	Algorithms for constrained optimization	121
5.6.1	Penalty and barrier function methods	121
5.6.2	Gradient projection method	125
5.6.3	Sequential quadratic programming	130
5.7	Summary and discussion points	131
5.8	Exercises	132
6	Deterministic GO algorithms	137
6.1	Introduction	137
6.2	Deterministic heuristic, DIRECT	138
6.2.1	Selection for refinement	139
6.2.2	Choice for sampling and updating rectangles	141
6.2.3	Algorithm and illustration	142
6.3	Stochastic models and response surfaces	144
6.4	Mathematical structures	147
6.4.1	Concavity	148
6.4.2	Difference of convex functions, d.c.	149
6.4.3	Lipschitz continuity and bounds on derivatives	150
6.4.4	Quadratic functions	154
6.4.5	Bilinear functions	155
6.4.6	Multiplicative and fractional functions	156
6.4.7	Interval arithmetic	158
6.5	Global Optimization branch and bound	159
6.6	Examples from nonconvex quadratic programming	161
6.6.1	Example concave quadratic programming	162
6.6.2	Example indefinite quadratic programming	163
6.7	Cutting planes	165
6.8	Summary and discussion points	168
6.9	Exercises	169
7	Stochastic GO algorithms	171
7.1	Introduction	171
7.2	Random sampling in higher dimensions	172
7.2.1	All volume to the boundary	172
7.2.2	Loneliness in high dimensions	173
7.3	PRS- and Multistart-based methods	174
7.3.1	Pure Random Search as benchmark	174

7.3.2	Multistart as benchmark	176
7.3.3	Clustering to save on local searches	178
7.3.4	Tunneling and filled functions	180
7.4	Ideal and real, PAS and Hit and Run	183
7.5	Population algorithms	187
7.5.1	Controlled Random Search and Raspberries	188
7.5.2	Genetic algorithms	191
7.5.3	Particle swarms	195
7.6	Summary and discussion points	197
7.7	Exercises	197
References		199
Index		205

Preface

This book provides a solid introduction for anyone who wants to study the ideas, concepts, and algorithms behind nonlinear and global optimization. In our experience instructing the topic, we have encountered applications of optimization methods based on easily accessible Internet software. In our classes, we find that students are more often scanning the Internet for information on concepts and methodologies and therefore a good understanding of the concepts and keywords is already necessary.

Many good books exist for teaching optimization that focus on theoretical properties and guidance in proving mathematical relations. The current text adds illustrations and simple examples and exercises, enhancing the reader's understanding of concepts. In fact, to enrich our didactical methods, this book contains approximately 40 algorithms that are illustrated by 80 examples and 95 figures. Additional comprehension and study is encouraged with numerous exercises. Furthermore, rather than providing rigorous mathematical proofs, we hope to evoke a critical approach toward the use of optimization algorithms. As an alternative to focusing on the background ideas often furnished on the Internet, we would like students to study pure pseudocode from a critical and systematic perspective.

Interesting models from an optimization perspective come from biology, engineering, finance, chemistry, economics, etc. Modeling optimization problems depends largely on the discipline and on the mathematical modeling courses that can be found in many curricula. In Chapter 2 we use several cases from our own experience and try to accustom the student to using intuition on questions of multimodality such as “is it natural that a problem has several local optima?” Examples are given and exercises follow. No formal methodology is presented other than using intuition and analytic skills.

In our experience, we have observed the application of optimization methods with an enormous trust in clicking buttons and accepting outcomes. It is often thought that what comes out of a computer program must be true. To have any practical value, the outcomes should at least fulfill optimality conditions. Therefore in Chapter 3, we focus on the criteria of optimality illustrated

with simple examples and referring further to the earlier mentioned books that present more mathematical rigor. Again, many exercises are provided.

The application and investigation of methods, with a nearly religious belief in concepts, like evolutionary programming and difference of convex programming, inspired us to explain such concepts briefly and then ask questions on the effectiveness and efficiency of these methods. Specifically, in Chapter 4 we pose questions and try to show how to investigate them in a systematic way. The style set in this chapter is then followed in subsequent chapters, where multiple algorithms are introduced and illustrated.

Books on nonlinear optimization often describe algorithms in a more or less explicit way discussing the ideas and their background. In Chapter 5, a uniform way of describing the algorithms can be found and each algorithm is illustrated with a simple numerical example. Methods cover one-dimensional optimization, derivative-free optimization, and methods for constrained and unconstrained optimization.

The ambition of global optimization algorithms is to find a global optimum point. Heuristic methods, as well as deterministic stochastic methods, often do not require or use specific characteristics of a problem to be solved. An interpretation of the so-called “no free lunch theorem” is that general-purpose methods habitually perform worse than dedicated algorithms that exploit the specific structure of the problem. Besides using heuristic methods, deterministic methods can be designed that give a guarantee to approach the optimum to an accuracy if structure information is available and used.

Many concepts exist which are popular in mathematical research on the structures of problems. For each structure at least one book exists and it was a challenge for us to describe these structures in a concise way. Chapter 6 explores deterministic global optimization algorithms. Each concept is introduced and illustrated with an example. Emphasis is also placed on how one can recognize structure when studying an optimization problem. The approach of branch and bound follows which aims to guarantee reaching a global solution while using the structure. Another approach that uses structure, the generation of cuts, is also illustrated. The main characteristic of deterministic methods is that no (pseudo-)random variable is used to find sample points. We start the chapter discussing heuristics that have this property. The main idea there is that function evaluations may be expensive. That means that it may require seconds, minutes, or even hours to find the objective function value of a suggested sample point.

Stochastic methods are extremely popular from an application perspective, as implementations of algorithms can be found easily. Although stochastic methods have been investigated thoroughly in the field of global optimization, one can observe a blind use of evolution-based concepts. Chapter 7 tries to summarize several concepts and to describe algorithms as basically and as dryly as possible, each illustrated. Focus is on a critical approach toward the results that can be obtained using algorithms by applying them to optimization problems.

We thank all the people who contributed to, commented on, and stimulated this work. The material was used and tested in master's and Ph.D. courses at the University of Almería where colleagues were very helpful in reading and commenting on the material. We thank the colleagues of the Computer Architecture Department and specifically its former director for helping to enhance the appearance of the book. Students and colleagues from Wageningen University and from Budapest University of Technology and Economics added useful commentary. Since 2008 the Spanish ministry of science has helped by funding a Ramón y Cajal contract at the Computer Architecture department of Málaga University. This enabled us to devote a lot of extra time to the book. The editorial department of Springer helped to shape the book and provided useful comments of the anonymous referees.

Eligius M.T. Hendrix
Boglárka G.-Tóth
September 2009