

INTRODUCTION TO NONLINEAR OPTIMIZATION



MOS-SIAM Series on Optimization

This series is published jointly by the Mathematical Optimization 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 for the intended audience.

Editor-in-Chief

Katya Scheinberg
Lehigh University

Editorial Board

Santanu S. Dey, *Georgia Institute of Technology*
Maryam Fazel, *University of Washington*
Andrea Lodi, *University of Bologna*
Arkadi Nemirovski, *Georgia Institute of Technology*
Stefan Ulbrich, *Technische Universität Darmstadt*
Luis Nunes Vicente, *University of Coimbra*
David Williamson, *Cornell University*
Stephen J. Wright, *University of Wisconsin*

Series Volumes

Terlaky, Tamás, Anjos, Miguel F., and Ahmed, Shabbir, editors, *Advances and Trends in Optimization with Engineering Applications*
Todd, Michael J., *Minimum-Volume Ellipsoids: Theory and Algorithms*
Bienstock, Daniel, *Electrical Transmission System Cascades and Vulnerability: An Operations Research Viewpoint*
Koch, Thorsten, Hiller, Benjamin, Pfetsch, Marc E., and Schewe, Lars, editors, *Evaluating Gas Network Capacities*
Corberán, Ángel, and Laporte, Gilbert, *Arc Routing: Problems, Methods, and Applications*
Toth, Paolo, and Vigo, Daniele, *Vehicle Routing: Problems, Methods, and Applications, Second Edition*
Beck, Amir, *Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MATLAB*
Attouch, Hedy, Buttazzo, Giuseppe, and Michaille, Gérard, *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization, Second Edition*
Shapiro, Alexander, Dentcheva, Darinka, and Ruszczyński, Andrzej, *Lectures on Stochastic Programming: Modeling and Theory, Second Edition*
Locatelli, Marco and Schoen, Fabio, *Global Optimization: Theory, Algorithms, and Applications*
De Loera, Jesús A., Hemmecke, Raymond, and Köppe, Matthias, *Algebraic and Geometric Ideas in the Theory of Discrete Optimization*
Blekherman, Grigoriy, Parrilo, Pablo A., and Thomas, Rekha R., editors, *Semidefinite Optimization and Convex Algebraic Geometry*
Delfour, M. C., *Introduction to Optimization and Semidifferential Calculus*
Ulbrich, Michael, *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*
Biegler, Lorenz T., *Nonlinear Programming: Concepts, Algorithms, and Applications to Chemical Processes*
Shapiro, Alexander, Dentcheva, Darinka, and Ruszczyński, Andrzej, *Lectures on Stochastic Programming: Modeling and Theory*
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, Philippe L., *Trust-Region Methods*

INTRODUCTION TO NONLINEAR OPTIMIZATION

Theory, Algorithms, and Applications with MATLAB

Amir Beck

Technion–Israel Institute of Technology
Kfar Saba, Israel



Society for Industrial and Applied Mathematics
Philadelphia



Mathematical
Optimization Society

Mathematical Optimization Society
Philadelphia

Copyright © 2014 by the Society for Industrial and Applied Mathematics and the Mathematical Optimization Society

10 9 8 7 6 5 4 3 2

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: 508-647-7001, info@mathworks.com, www.mathworks.com.

Library of Congress Cataloging-in-Publication Data

Beck, Amir, author.

Introduction to nonlinear optimization : theory, algorithms, and applications with MATLAB / Amir Beck, Technion-Israel Institute of Technology, Kfar Saba, Israel.

pages cm. -- (MOS-SIAM series on optimization)

Includes bibliographical references and index.

ISBN 978-1-611973-64-8

1. Mathematical optimization. 2. Nonlinear theories. 3. MATLAB. I. Title.

QA402.5.B4224 2014

519.6-dc23

2014029493



For
My wife Nili
My daughters Noy and Vered
My parents Nili and Itzhak



Contents

Preface	xi
1 Mathematical Preliminaries	1
1.1 The Space \mathbb{R}^n	1
1.2 The Space $\mathbb{R}^{m \times n}$	2
1.3 Inner Products and Norms	2
1.4 Eigenvalues and Eigenvectors	5
1.5 Basic Topological Concepts	6
Exercises	10
2 Optimality Conditions for Unconstrained Optimization	13
2.1 Global and Local Optima	13
2.2 Classification of Matrices	17
2.3 Second Order Optimality Conditions	23
2.4 Global Optimality Conditions	30
2.5 Quadratic Functions	32
Exercises	34
3 Least Squares	37
3.1 “Solution” of Overdetermined Systems	37
3.2 Data Fitting	39
3.3 Regularized Least Squares	41
3.4 Denoising	42
3.5 Nonlinear Least Squares	45
3.6 Circle Fitting	45
Exercises	47
4 The Gradient Method	49
4.1 Descent Directions Methods	49
4.2 The Gradient Method	52
4.3 The Condition Number	58
4.4 Diagonal Scaling	63
4.5 The Gauss–Newton Method	67
4.6 The Fermat–Weber Problem	68
4.7 Convergence Analysis of the Gradient Method	73
Exercises	79
5 Newton’s Method	83
5.1 Pure Newton’s Method	83

5.2	Damped Newton's Method	88
5.3	The Cholesky Factorization	89
	Exercises	94
6	Convex Sets	97
6.1	Definition and Examples	97
6.2	Algebraic Operations with Convex Sets	100
6.3	The Convex Hull	101
6.4	Convex Cones	104
6.5	Topological Properties of Convex Sets	108
6.6	Extreme Points	111
	Exercises	113
7	Convex Functions	117
7.1	Definition and Examples	117
7.2	First Order Characterizations of Convex Functions	119
7.3	Second Order Characterization of Convex Functions	123
7.4	Operations Preserving Convexity	125
7.5	Level Sets of Convex Functions	130
7.6	Continuity and Differentiability of Convex Functions	132
7.7	Extended Real-Valued Functions	135
7.8	Maxima of Convex Functions	137
7.9	Convexity and Inequalities	139
	Exercises	141
8	Convex Optimization	147
8.1	Definition	147
8.2	Examples	149
8.3	The Orthogonal Projection Operator	156
8.4	CVX	158
	Exercises	166
9	Optimization over a Convex Set	169
9.1	Stationarity	169
9.2	Stationarity in Convex Problems	173
9.3	The Orthogonal Projection Revisited	173
9.4	The Gradient Projection Method	175
9.5	Sparsity Constrained Problems	183
	Exercises	189
10	Optimality Conditions for Linearly Constrained Problems	191
10.1	Separation and Alternative Theorems	191
10.2	The KKT conditions	195
10.3	Orthogonal Regression	203
	Exercises	205
11	The KKT Conditions	207
11.1	Inequality Constrained Problems	207
11.2	Inequality and Equality Constrained Problems	210
11.3	The Convex Case	213
11.4	Constrained Least Squares	218

11.5	Second Order Optimality Conditions	222
11.6	Optimality Conditions for the Trust Region Subproblem	227
11.7	Total Least Squares	230
	Exercises	233
12	Duality	237
12.1	Motivation and Definition	237
12.2	Strong Duality in the Convex Case	241
12.3	Examples	247
	Exercises	270
	Bibliographic Notes	275
	Bibliography	277
	Index	281

Preface

This book emerged from the idea that an optimization training should include three basic components: a strong theoretical and algorithmic foundation, familiarity with various applications, and the ability to apply the theory and algorithms on actual “real-life” problems. The book is intended to be the basis of such an extensive training. The mathematical development of the main concepts in nonlinear optimization is done rigorously, where a special effort was made to keep the proofs as simple as possible. The results are presented gradually and accompanied with many illustrative examples. Since the aim is not to give an encyclopedic overview, the focus is on the most useful and important concepts. The theory is complemented by numerous discussions on applications from various scientific fields such as signal processing, economics and localization. Some basic algorithms are also presented and studied to provide some flavor of this important aspect of optimization. Many topics are demonstrated by MATLAB programs, and ideally, the interested reader will find satisfaction in the ability of actually solving problems on his or her own. The book contains several topics that, compared to other classical textbooks, are treated differently. The following are some examples of the less common issues.

- The treatment of stationarity is comprehensive and discusses this important notion in the presence of sparsity constraints.
- The concept of “hidden convexity” is discussed and illustrated in the context of the trust region subproblem.
- The MATLAB toolbox CVX is explored and used.
- The gradient mapping and its properties are studied and used in the analysis of the gradient projection method.
- Second order necessary optimality conditions are treated using a descent direction approach.
- Applications such as circle fitting, Chebyshev center, the Fermat–Weber problem, denoising, clustering, total least squares, and orthogonal regression are studied both theoretically and algorithmically, illustrating concepts such as duality. MATLAB programs are used to show how the theory can be implemented.

The book is intended for students and researchers with a basic background in advanced calculus and linear algebra, but no prior knowledge of optimization theory is assumed. The book contains more than 170 exercises, which can be used to deepen the understanding of the material. The MATLAB functions described throughout the book can be found at

www.siam.org/books/mo19.

The outline of the book is as follows. **Chapter 1** recalls some of the important concepts in linear algebra and calculus that are essential for the understanding of the mathematical developments in the book. **Chapter 2** focuses on local and global optimality conditions for smooth unconstrained problems. Quadratic functions are also introduced along with their basic properties. Linear and nonlinear least squares problems are introduced and studied in **Chapter 3**. Several applications such as data fitting, denoising, and circle fitting are discussed. The gradient method is introduced and studied in **Chapter 4**. The chapter also contains a discussion on descent direction methods and various stepsize strategies. Extensions such as the scaled gradient method and damped Gauss–Newton are considered. The connection between the gradient method and Weiszfeld’s method for solving the Fermat–Weber problem is established. Newton’s method is discussed in **Chapter 5**. Convex sets and functions along with their basic properties are the subjects of **Chapters 6 and 7**. Convex optimization problems are introduced in **Chapter 8**, which also includes a variety of applications as well as CVX demonstrations. **Chapter 9** focuses on several important topics related to optimization problems over convex sets: stationarity, gradient mappings, and the gradient projection method. The chapter ends with results on sparsity constrained problems, illuminating the different type of results obtained when the underlying set is not convex. The derivation of the KKT optimality conditions from the separation and alternative theorems is the subject of **Chapter 10**, where only linearly constrained problems are considered. The extension of the KKT conditions to problems with nonlinear constraints is discussed in **Chapter 11**, which also considers the second order necessary conditions. Applications of the conditions to the trust region and total least squares problems are studied. The book ends with a discussion on duality in **Chapter 12**. Strong duality under convexity assumptions is established. This chapter also includes a large amount of examples, applications, and MATLAB illustrations.

I would like to thank Dror Pan and Luba Tetruashvili for reading the book and for their helpful remarks. It has been a pleasure working with the SIAM staff, namely with Bruce Bailey, Elizabeth Greenspan, Sara Murphy, Gina Rinelli, and Kelly Thomas.