

Springer Series in Operations Research and Financial Engineering

Editors:

Thomas V. Mikosch Sidney I. Resnick Stephen M. Robinson

Springer Series in Operation Research and Financial Engineering

- Altiok*: Performance Analysis of Manufacturing Systems
- Birge and Louveaux*: Introduction to Stochastic Programming
- Bonnans and Shapiro*: Perturbation Analysis of Optimization Problems
- Bramel, Chen, and Simchi-Levi*: The Logic of Logistics: Theory, Algorithms, and Applications for Logistics and Supply Chain Management (second edition)
- Dantzig and Thapa*: Linear Programming 1: Introduction
- Dantzig and Thapa*: Linear Programming 2: Theory and Extensions
- Drezner (Editor)*: Facility Location: A Survey of Applications and Methods
- Facchinei and Pang*: Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume I
- Facchinei and Pang*: Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume II
- Fishman*: Discrete-Event Simulation: Modeling, Programming, and Analysis
- Fishman*: Monte Carlo: Concepts, Algorithms, and Applications
- Haas*: Stochastic Petri Nets: Modeling, Stability, Simulation
- Klamroth*: Single-Facility Location Problems with Barriers
- Muckstadt*: Analysis and Algorithms for Service Parts Supply Chains
- Nocedal and Wright*: Numerical Optimization
- Olson*: Decision Aids for Selection Problems
- Pinedo*: Planning and Scheduling in Manufacturing and Services
- Pochet and Wolsey*: Production Planning by Mixed Integer Programming
- Whitt*: Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues
- Yao (Editor)*: Stochastic Modeling and Analysis of Manufacturing Systems
- Yao and Zheng*: Dynamic Control of Quality in Production-Inventory Systems: Coordination and Optimization
- Yeung and Petrosyan*: Cooperative Stochastic Differential Games

Jorge Nocedal Stephen J. Wright

Numerical Optimization

Second Edition

 Springer

Jorge Nocedal
EECS Department
Northwestern University
Evanston, IL 60208-3118
USA
nocedal@eecs.northwestern.edu

Stephen J. Wright
Computer Sciences Department
University of Wisconsin
1210 West Dayton Street
Madison, WI 53706-1613
USA
swright@cs.wisc.edu

Series Editors:

Thomas V. Mikosch
University of Copenhagen
Laboratory of Actuarial Mathematics
DK-1017 Copenhagen
Denmark
mikosch@act.ku.dk

Stephen M. Robinson
Department of Industrial and Systems
Engineering
University of Wisconsin
1513 University Avenue
Madison, WI 53706-1539
USA
smrobins@facstaff.wise.edu

Sidney I. Resnick
Cornell University
School of Operations Research and
Industrial Engineering
Ithaca, NY 14853
USA
sirl@cornell.edu

Mathematics Subject Classification (2000): 90B30, 90C11, 90-01, 90-02

Library of Congress Control Number: 2006923897

ISBN-10: 0-387-30303-0

ISBN-13: 978-0387-30303-1

Printed on acid-free paper.

© 2006 Springer Science+Business Media, LLC.

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 in the United States of America. (TB/HAM)

9 8 7 6 5 4 3 2 1

springer.com

To Sue, Isabel and Martin
and
To Mum and Dad

Contents

Preface	xvii
Preface to the Second Edition	xxi
1 Introduction	1
Mathematical Formulation	2
Example: A Transportation Problem	4
Continuous versus Discrete Optimization	5
Constrained and Unconstrained Optimization	6
Global and Local Optimization	6
Stochastic and Deterministic Optimization	7
Convexity	7
Optimization Algorithms	8
Notes and References	9
2 Fundamentals of Unconstrained Optimization	10
2.1 What Is a Solution?	12

	Recognizing a Local Minimum	14
	Nonsmooth Problems	17
2.2	Overview of Algorithms	18
	Two Strategies: Line Search and Trust Region	19
	Search Directions for Line Search Methods	20
	Models for Trust-Region Methods	25
	Scaling	26
	Exercises	27
3	Line Search Methods	30
3.1	Step Length	31
	The Wolfe Conditions	33
	The Goldstein Conditions	36
	Sufficient Decrease and Backtracking	37
3.2	Convergence of Line Search Methods	37
3.3	Rate of Convergence	41
	Convergence Rate of Steepest Descent	42
	Newton's Method	44
	Quasi-Newton Methods	46
3.4	Newton's Method with Hessian Modification	48
	Eigenvalue Modification	49
	Adding a Multiple of the Identity	51
	Modified Cholesky Factorization	52
	Modified Symmetric Indefinite Factorization	54
3.5	Step-Length Selection Algorithms	56
	Interpolation	57
	Initial Step Length	59
	A Line Search Algorithm for the Wolfe Conditions	60
	Notes and References	62
	Exercises	63
4	Trust-Region Methods	66
	Outline of the Trust-Region Approach	68
4.1	Algorithms Based on the Cauchy Point	71
	The Cauchy Point	71
	Improving on the Cauchy Point	73
	The Dogleg Method	73
	Two-Dimensional Subspace Minimization	76
4.2	Global Convergence	77
	Reduction Obtained by the Cauchy Point	77
	Convergence to Stationary Points	79
4.3	Iterative Solution of the Subproblem	83

	The Hard Case	87
	Proof of Theorem 4.1	89
	Convergence of Algorithms Based on Nearly Exact Solutions	91
4.4	Local Convergence of Trust-Region Newton Methods	92
4.5	Other Enhancements	95
	Scaling	95
	Trust Regions in Other Norms	97
	Notes and References	98
	Exercises	98
5	Conjugate Gradient Methods	101
5.1	The Linear Conjugate Gradient Method	102
	Conjugate Direction Methods	102
	Basic Properties of the Conjugate Gradient Method	107
	A Practical Form of the Conjugate Gradient Method	111
	Rate of Convergence	112
	Preconditioning	118
	Practical Preconditioners	120
5.2	Nonlinear Conjugate Gradient Methods	121
	The Fletcher–Reeves Method	121
	The Polak–Ribière Method and Variants	122
	Quadratic Termination and Restarts	124
	Behavior of the Fletcher–Reeves Method	125
	Global Convergence	127
	Numerical Performance	131
	Notes and References	132
	Exercises	133
6	Quasi-Newton Methods	135
6.1	The BFGS Method	136
	Properties of the BFGS Method	141
	Implementation	142
6.2	The SR1 Method	144
	Properties of SR1 Updating	147
6.3	The Broyden Class	149
6.4	Convergence Analysis	153
	Global Convergence of the BFGS Method	153
	Superlinear Convergence of the BFGS Method	156
	Convergence Analysis of the SR1 Method	160
	Notes and References	161
	Exercises	162

7	Large-Scale Unconstrained Optimization	164
7.1	Inexact Newton Methods	165
	Local Convergence of Inexact Newton Methods	166
	Line Search Newton–CG Method	168
	Trust-Region Newton–CG Method	170
	Preconditioning the Trust-Region Newton–CG Method	174
	Trust-Region Newton–Lanczos Method	175
7.2	Limited-Memory Quasi-Newton Methods	176
	Limited-Memory BFGS	177
	Relationship with Conjugate Gradient Methods	180
	General Limited-Memory Updating	181
	Compact Representation of BFGS Updating	181
	Unrolling the Update	184
7.3	Sparse Quasi-Newton Updates	185
7.4	Algorithms for Partially Separable Functions	186
7.5	Perspectives and Software	189
	Notes and References	190
	Exercises	191
8	Calculating Derivatives	193
8.1	Finite-Difference Derivative Approximations	194
	Approximating the Gradient	195
	Approximating a Sparse Jacobian	197
	Approximating the Hessian	201
	Approximating a Sparse Hessian	202
8.2	Automatic Differentiation	204
	An Example	205
	The Forward Mode	206
	The Reverse Mode	207
	Vector Functions and Partial Separability	210
	Calculating Jacobians of Vector Functions	212
	Calculating Hessians: Forward Mode	213
	Calculating Hessians: Reverse Mode	215
	Current Limitations	216
	Notes and References	217
	Exercises	217
9	Derivative-Free Optimization	220
9.1	Finite Differences and Noise	221
9.2	Model-Based Methods	223
	Interpolation and Polynomial Bases	226
	Updating the Interpolation Set	227

	A Method Based on Minimum-Change Updating	228
9.3	Coordinate and Pattern-Search Methods	229
	Coordinate Search Method	230
	Pattern-Search Methods	231
9.4	A Conjugate-Direction Method	234
9.5	Nelder–Mead Method	238
9.6	Implicit Filtering	240
	Notes and References	242
	Exercises	242
10	Least-Squares Problems	245
10.1	Background	247
10.2	Linear Least-Squares Problems	250
10.3	Algorithms for Nonlinear Least-Squares Problems	254
	The Gauss–Newton Method	254
	Convergence of the Gauss–Newton Method	255
	The Levenberg–Marquardt Method	258
	Implementation of the Levenberg–Marquardt Method	259
	Convergence of the Levenberg–Marquardt Method	261
	Methods for Large-Residual Problems	262
10.4	Orthogonal Distance Regression	265
	Notes and References	267
	Exercises	269
11	Nonlinear Equations	270
11.1	Local Algorithms	274
	Newton’s Method for Nonlinear Equations	274
	Inexact Newton Methods	277
	Broyden’s Method	279
	Tensor Methods	283
11.2	Practical Methods	285
	Merit Functions	285
	Line Search Methods	287
	Trust-Region Methods	290
11.3	Continuation/Homotopy Methods	296
	Motivation	296
	Practical Continuation Methods	297
	Notes and References	302
	Exercises	302
12	Theory of Constrained Optimization	304
	Local and Global Solutions	305

	Smoothness	306
12.1	Examples	307
	A Single Equality Constraint	308
	A Single Inequality Constraint	310
	Two Inequality Constraints	313
12.2	Tangent Cone and Constraint Qualifications	315
12.3	First-Order Optimality Conditions	320
12.4	First-Order Optimality Conditions: Proof	323
	Relating the Tangent Cone and the First-Order Feasible Direction Set	323
	A Fundamental Necessary Condition	325
	Farkas' Lemma	326
	Proof of Theorem 12.1	329
12.5	Second-Order Conditions	330
	Second-Order Conditions and Projected Hessians	337
12.6	Other Constraint Qualifications	338
12.7	A Geometric Viewpoint	340
12.8	Lagrange Multipliers and Sensitivity	341
12.9	Duality	343
	Notes and References	349
	Exercises	351

13	Linear Programming: The Simplex Method	355
	Linear Programming	356
13.1	Optimality and Duality	358
	Optimality Conditions	358
	The Dual Problem	359
13.2	Geometry of the Feasible Set	362
	Bases and Basic Feasible Points	362
	Vertices of the Feasible Polytope	365
13.3	The Simplex Method	366
	Outline	366
	A Single Step of the Method	370
13.4	Linear Algebra in the Simplex Method	372
13.5	Other Important Details	375
	Pricing and Selection of the Entering Index	375
	Starting the Simplex Method	378
	Degenerate Steps and Cycling	381
13.6	The Dual Simplex Method	382
13.7	Presolving	385
13.8	Where Does the Simplex Method Fit?	388
	Notes and References	389
	Exercises	389

14 Linear Programming: Interior-Point Methods	392
14.1 Primal-Dual Methods	393
Outline	393
The Central Path	397
Central Path Neighborhoods and Path-Following Methods	399
14.2 Practical Primal-Dual Algorithms	407
Corrector and Centering Steps	407
Step Lengths	409
Starting Point	410
A Practical Algorithm	411
Solving the Linear Systems	411
14.3 Other Primal-Dual Algorithms and Extensions	413
Other Path-Following Methods	413
Potential-Reduction Methods	414
Extensions	415
14.4 Perspectives and Software	416
Notes and References	417
Exercises	418
15 Fundamentals of Algorithms for Nonlinear Constrained Optimization	421
15.1 Categorizing Optimization Algorithms	422
15.2 The Combinatorial Difficulty of Inequality-Constrained Problems	424
15.3 Elimination of Variables	426
Simple Elimination using Linear Constraints	428
General Reduction Strategies for Linear Constraints	431
Effect of Inequality Constraints	434
15.4 Merit Functions and Filters	435
Merit Functions	435
Filters	437
15.5 The Maratos Effect	440
15.6 Second-Order Correction and Nonmonotone Techniques	443
Nonmonotone (Watchdog) Strategy	444
Notes and References	446
Exercises	446
16 Quadratic Programming	448
16.1 Equality-Constrained Quadratic Programs	451
Properties of Equality-Constrained QPs	451
16.2 Direct Solution of the KKT System	454
Factoring the Full KKT System	454
Schur-Complement Method	455
Null-Space Method	457

16.3	Iterative Solution of the KKT System	459
	CG Applied to the Reduced System	459
	The Projected CG Method	461
16.4	Inequality-Constrained Problems	463
	Optimality Conditions for Inequality-Constrained Problems	464
	Degeneracy	465
16.5	Active-Set Methods for Convex QPs	467
	Specification of the Active-Set Method for Convex QP	472
	Further Remarks on the Active-Set Method	476
	Finite Termination of Active-Set Algorithm on Strictly Convex QPs	477
	Updating Factorizations	478
16.6	Interior-Point Methods	480
	Solving the Primal-Dual System	482
	Step Length Selection	483
	A Practical Primal-Dual Method	484
16.7	The Gradient Projection Method	485
	Cauchy Point Computation	486
	Subspace Minimization	488
16.8	Perspectives and Software	490
	Notes and References	492
	Exercises	492
17	Penalty and Augmented Lagrangian Methods	497
17.1	The Quadratic Penalty Method	498
	Motivation	498
	Algorithmic Framework	501
	Convergence of the Quadratic Penalty Method	502
	Ill Conditioning and Reformulations	505
17.2	Nonsmooth Penalty Functions	507
	A Practical ℓ_1 Penalty Method	511
	A General Class of Nonsmooth Penalty Methods	513
17.3	Augmented Lagrangian Method: Equality Constraints	514
	Motivation and Algorithmic Framework	514
	Properties of the Augmented Lagrangian	517
17.4	Practical Augmented Lagrangian Methods	519
	Bound-Constrained Formulation	519
	Linearly Constrained Formulation	522
	Unconstrained Formulation	523
17.5	Perspectives and Software	525
	Notes and References	526
	Exercises	527

18 Sequential Quadratic Programming	529
18.1 Local SQP Method	530
SQP Framework	531
Inequality Constraints	532
18.2 Preview of Practical SQP Methods	533
IQP and EQP	533
Enforcing Convergence	534
18.3 Algorithmic Development	535
Handling Inconsistent Linearizations	535
Full Quasi-Newton Approximations	536
Reduced-Hessian Quasi-Newton Approximations	538
Merit Functions	540
Second-Order Correction	543
18.4 A Practical Line Search SQP Method	545
18.5 Trust-Region SQP Methods	546
A Relaxation Method for Equality-Constrained Optimization	547
$S\ell_1$ QP (Sequential ℓ_1 Quadratic Programming)	549
Sequential Linear-Quadratic Programming (SLQP)	551
A Technique for Updating the Penalty Parameter	553
18.6 Nonlinear Gradient Projection	554
18.7 Convergence Analysis	556
Rate of Convergence	557
18.8 Perspectives and Software	560
Notes and References	561
Exercises	561
19 Interior-Point Methods for Nonlinear Programming	563
19.1 Two Interpretations	564
19.2 A Basic Interior-Point Algorithm	566
19.3 Algorithmic Development	569
Primal vs. Primal-Dual System	570
Solving the Primal-Dual System	570
Updating the Barrier Parameter	572
Handling Nonconvexity and Singularity	573
Step Acceptance: Merit Functions and Filters	575
Quasi-Newton Approximations	575
Feasible Interior-Point Methods	576
19.4 A Line Search Interior-Point Method	577
19.5 A Trust-Region Interior-Point Method	578
An Algorithm for Solving the Barrier Problem	578
Step Computation	580
Lagrange Multipliers Estimates and Step Acceptance	581

	Description of a Trust-Region Interior-Point Method	582
19.6	The Primal Log-Barrier Method	583
19.7	Global Convergence Properties	587
	Failure of the Line Search Approach	587
	Modified Line Search Methods	589
	Global Convergence of the Trust-Region Approach	589
19.8	Superlinear Convergence	591
19.9	Perspectives and Software	592
	Notes and References	593
	Exercises	594
A	Background Material	598
A.1	Elements of Linear Algebra	598
	Vectors and Matrices	598
	Norms	600
	Subspaces	602
	Eigenvalues, Eigenvectors, and the Singular-Value Decomposition	603
	Determinant and Trace	605
	Matrix Factorizations: Cholesky, LU, QR	606
	Symmetric Indefinite Factorization	610
	Sherman–Morrison–Woodbury Formula	612
	Interlacing Eigenvalue Theorem	613
	Error Analysis and Floating-Point Arithmetic	613
	Conditioning and Stability	616
A.2	Elements of Analysis, Geometry, Topology	617
	Sequences	617
	Rates of Convergence	619
	Topology of the Euclidean Space \mathbb{R}^n	620
	Convex Sets in \mathbb{R}^n	621
	Continuity and Limits	623
	Derivatives	625
	Directional Derivatives	628
	Mean Value Theorem	629
	Implicit Function Theorem	630
	Order Notation	631
	Root-Finding for Scalar Equations	633
B	A Regularization Procedure	635
	References	637
	Index	653

Preface

This is a book for people interested in solving optimization problems. Because of the wide (and growing) use of optimization in science, engineering, economics, and industry, it is essential for students and practitioners alike to develop an understanding of optimization algorithms. Knowledge of the capabilities and limitations of these algorithms leads to a better understanding of their impact on various applications, and points the way to future research on improving and extending optimization algorithms and software. Our goal in this book is to give a comprehensive description of the most powerful, state-of-the-art, techniques for solving continuous optimization problems. By presenting the motivating ideas for each algorithm, we try to stimulate the reader's intuition and make the technical details easier to follow. Formal mathematical requirements are kept to a minimum.

Because of our focus on continuous problems, we have omitted discussion of important optimization topics such as discrete and stochastic optimization. However, there are a great many applications that can be formulated as continuous optimization problems; for instance,

- finding the optimal trajectory for an aircraft or a robot arm;

- identifying the seismic properties of a piece of the earth's crust by fitting a model of the region under study to a set of readings from a network of recording stations;

designing a portfolio of investments to maximize expected return while maintaining an acceptable level of risk;

controlling a chemical process or a mechanical device to optimize performance or meet standards of robustness;

computing the optimal shape of an automobile or aircraft component.

Every year optimization algorithms are being called on to handle problems that are much larger and complex than in the past. Accordingly, the book emphasizes large-scale optimization techniques, such as interior-point methods, inexact Newton methods, limited-memory methods, and the role of partially separable functions and automatic differentiation. It treats important topics such as trust-region methods and sequential quadratic programming more thoroughly than existing texts, and includes comprehensive discussion of such “core curriculum” topics as constrained optimization theory, Newton and quasi-Newton methods, nonlinear least squares and nonlinear equations, the simplex method, and penalty and barrier methods for nonlinear programming.

The Audience

We intend that this book will be used in graduate-level courses in optimization, as offered in engineering, operations research, computer science, and mathematics departments. There is enough material here for a two-semester (or three-quarter) sequence of courses. We hope, too, that this book will be used by practitioners in engineering, basic science, and industry, and our presentation style is intended to facilitate self-study. Since the book treats a number of new algorithms and ideas that have not been described in earlier textbooks, we hope that this book will also be a useful reference for optimization researchers.

Prerequisites for this book include some knowledge of linear algebra (including numerical linear algebra) and the standard sequence of calculus courses. To make the book as self-contained as possible, we have summarized much of the relevant material from these areas in the Appendix. Our experience in teaching engineering students has shown us that the material is best assimilated when combined with computer programming projects in which the student gains a good feeling for the algorithms—their complexity, memory demands, and elegance—and for the applications. In most chapters we provide simple computer exercises that require only minimal programming proficiency.

Emphasis and Writing Style

We have used a conversational style to motivate the ideas and present the numerical algorithms. Rather than being as concise as possible, our aim is to make the discussion flow in a natural way. As a result, the book is comparatively long, but we believe that it can be read relatively rapidly. The instructor can assign substantial reading assignments from the text and focus in class only on the main ideas.

A typical chapter begins with a nonrigorous discussion of the topic at hand, including figures and diagrams and excluding technical details as far as possible. In subsequent sections,

the algorithms are motivated and discussed, and then stated explicitly. The major theoretical results are stated, and in many cases proved, in a rigorous fashion. These proofs can be skipped by readers who wish to avoid technical details.

The practice of optimization depends not only on efficient and robust algorithms, but also on good modeling techniques, careful interpretation of results, and user-friendly software. In this book we discuss the various aspects of the optimization process—modeling, optimality conditions, algorithms, implementation, and interpretation of results—but not with equal weight. Examples throughout the book show how practical problems are formulated as optimization problems, but our treatment of modeling is light and serves mainly to set the stage for algorithmic developments. We refer the reader to Dantzig [86] and Fourer, Gay, and Kernighan [112] for more comprehensive discussion of this issue. Our treatment of optimality conditions is thorough but not exhaustive; some concepts are discussed more extensively in Mangasarian [198] and Clarke [62]. As mentioned above, we are quite comprehensive in discussing optimization algorithms.

Topics Not Covered

We omit some important topics, such as network optimization, integer programming, stochastic programming, nonsmooth optimization, and global optimization. Network and integer optimization are described in some excellent texts: for instance, Ahuja, Magnanti, and Orlin [1] in the case of network optimization and Nemhauser and Wolsey [224], Papadimitriou and Steiglitz [235], and Wolsey [312] in the case of integer programming. Books on stochastic optimization are only now appearing; we mention those of Kall and Wallace [174], Birge and Louveaux [22]. Nonsmooth optimization comes in many flavors. The relatively simple structures that arise in robust data fitting (which is sometimes based on the ℓ_1 norm) are treated by Osborne [232] and Fletcher [101]. The latter book also discusses algorithms for nonsmooth penalty functions that arise in constrained optimization; we discuss these briefly, too, in Chapter 18. A more analytical treatment of nonsmooth optimization is given by Hiriart-Urruty and Lemaréchal [170]. We omit detailed treatment of some important topics that are the focus of intense current research, including interior-point methods for nonlinear programming and algorithms for complementarity problems.

Additional Resource

The material in the book is complemented by an online resource called the NEOS Guide, which can be found on the World-Wide Web at

<http://www.mcs.anl.gov/otc/Guide/>

The Guide contains information about most areas of optimization, and presents a number of case studies that describe applications of various optimization algorithms to real-world problems such as portfolio optimization and optimal dieting. Some of this material is interactive in nature and has been used extensively for class exercises.

For the most part, we have omitted detailed discussions of specific software packages, and refer the reader to Moré and Wright [217] or to the Software Guide section of the NEOS Guide, which can be found at

<http://www.mcs.anl.gov/otc/Guide/SoftwareGuide/>

Users of optimization software refer in great numbers to this web site, which is being constantly updated to reflect new packages and changes to existing software.

Acknowledgments

We are most grateful to the following colleagues for their input and feedback on various sections of this work: Chris Bischof, Richard Byrd, George Corliss, Bob Fourer, David Gay, Jean-Charles Gilbert, Phillip Gill, Jean-Pierre Goux, Don Goldfarb, Nick Gould, Andreas Griewank, Matthias Heinkenschloss, Marcelo Marazzi, Hans Mittelmann, Jorge Moré, Will Naylor, Michael Overton, Bob Plemmons, Hugo Scolnik, David Stewart, Philippe Toint, Luis Vicente, Andreas Wächter, and Ya-xiang Yuan. We thank Guanghui Liu, who provided help with many of the exercises, and Jill Lavelle who assisted us in preparing the figures. We also express our gratitude to our sponsors at the Department of Energy and the National Science Foundation, who have strongly supported our research efforts in optimization over the years.

One of us (JN) would like to express his deep gratitude to Richard Byrd, who has taught him so much about optimization and who has helped him in very many ways throughout the course of his career.

Final Remark

In the preface to his 1987 book [101], Roger Fletcher described the field of optimization as a “fascinating blend of theory and computation, heuristics and rigor.” The ever-growing realm of applications and the explosion in computing power is driving optimization research in new and exciting directions, and the ingredients identified by Fletcher will continue to play important roles for many years to come.

Jorge Nocedal Stephen J. Wright
Evanston, IL Argonne, IL

Preface to the Second Edition

During the six years since the first edition of this book appeared, the field of continuous optimization has continued to grow and evolve. This new edition reflects a better understanding of constrained optimization at both the algorithmic and theoretical levels, and of the demands imposed by practical applications. Perhaps most notably, new chapters have been added on two important topics: derivative-free optimization (Chapter 9) and interior-point methods for nonlinear programming (Chapter 19). The former topic has proved to be of great interest in applications, while the latter topic has come into its own in recent years and now forms the basis of successful codes for nonlinear programming.

Apart from the new chapters, we have revised and updated throughout the book, de-emphasizing or omitting less important topics, enhancing the treatment of subjects of evident interest, and adding new material in many places. The first part (unconstrained optimization) has been comprehensively reorganized to improve clarity. Discussion of Newton's method—the touchstone method for unconstrained problems—is distributed more naturally throughout this part rather than being isolated in a single chapter. An expanded discussion of large-scale problems appears in Chapter 7.

Some reorganization has taken place also in the second part (constrained optimization), with material common to sequential quadratic programming and interior-point methods now appearing in the chapter on fundamentals of nonlinear programming

algorithms (Chapter 15) and the discussion of primal barrier methods moved to the new interior-point chapter. There is much new material in this part, including a treatment of nonlinear programming duality, an expanded discussion of algorithms for inequality constrained quadratic programming, a discussion of dual simplex and presolving in linear programming, a summary of practical issues in the implementation of interior-point linear programming algorithms, a description of conjugate-gradient methods for quadratic programming, and a discussion of filter methods and nonsmooth penalty methods in nonlinear programming algorithms.

In many chapters we have added a Perspectives and Software section near the end, to place the preceding discussion in context and discuss the state of the art in software. The appendix has been rearranged with some additional topics added, so that it can be used in a more stand-alone fashion to cover some of the mathematical background required for the rest of the book. The exercises have been revised in most chapters. After these many additions, deletions, and changes, the second edition is only slightly longer than the first, reflecting our belief that careful selection of the material to include and exclude is an important responsibility for authors of books of this type.

A manual containing solutions for selected problems will be available to bona fide instructors through the publisher. A list of typos will be maintained on the book's web site, which is accessible from the web pages of both authors.

We acknowledge with gratitude the comments and suggestions of many readers of the first edition, who sent corrections to many errors and provided valuable perspectives on the material, which led often to substantial changes. We mention in particular Frank Curtis, Michael Ferris, Andreas Griewank, Jacek Gondzio, Sven Leyffer, Philip Loewen, Rembert Reemtsen, and David Stewart.

Our special thanks goes to Michael Overton, who taught from a draft of the second edition and sent many detailed and excellent suggestions. We also thank colleagues who read various chapters of the new edition carefully during development, including Richard Byrd, Nick Gould, Paul Hovland, Gabo López-Calva, Long Hei, Katya Scheinberg, Andreas Wächter, and Richard Waltz. We thank Jill Wright for improving some of the figures and for the new cover graphic.

We mentioned in the original preface several areas of optimization that are not covered in this book. During the past six years, this list has only grown longer, as the field has continued to expand in new directions. In this regard, the following areas are particularly noteworthy: optimization problems with complementarity constraints, second-order cone and semidefinite programming, simulation-based optimization, robust optimization, and mixed-integer nonlinear programming. All these areas have seen theoretical and algorithmic advances in recent years, and in many cases developments are being driven by new classes of applications. Although this book does not cover any of these areas directly, it provides a foundation from which they can be studied.