

J. Frédéric Bonnans · J. Charles Gilbert
Claude Lemaréchal · Claudia A. Sagastizábal

Numerical Optimization

Theoretical and Practical Aspects

Second Edition

With 52 Figures

 Springer

Preface

This book is entirely devoted to numerical algorithms for optimization, their theoretical foundations and convergence properties, as well as their implementation, their use, and other practical aspects. The aim is to familiarize the reader with these numerical algorithms: understanding their behaviour in practice, properly using existing software libraries, adequately designing and implementing “home-made” methods, correctly diagnosing the causes of possible difficulties. Expected readers are engineers, Master or Ph.D. students, confirmed researchers, in applied mathematics or from various other disciplines where optimization is a need.

Our aim is therefore not to give most accurate results in optimization, nor to detail the latest refinements of such and such method. First of all, little is said concerning optimization theory itself (optimality conditions, constraint qualification, stability theory). As for algorithms, we limit ourselves most of the time to stable and well-established material. Throughout we keep as a leading thread the actual *practical value* of optimization methods, in terms of their efficiency to solve real-world problems. Nevertheless, serious attention is paid to the theoretical properties of optimization methods: this book is mainly based upon theorems. Besides, some new and promising results or approaches could not be completely discarded; they are also presented, generally in the form of special sections, mainly aimed at orienting the reader to the relevant bibliography.

An introductory chapter gives some generalities on optimization and iterative algorithms. It contains in particular motivating examples, ranging from meteorological forecast to power production management; they illustrate the large field of branches where optimization finds its applications. Then come four parts, rather independent of each other. The first one is devoted to algorithms for unconstrained optimization which, in addition to their direct usefulness, are a basis for more complex problems. The second part concerns rather special methods, applicable when the usual differentiability assumptions are not satisfied. Such methods appear in the decomposition of large-scale problems and the relaxation of combinatorial problems. Nonlinearly constrained optimization forms the third part, substantially more technical, as the subject is still in evolution. Finally, the fourth part gives a deep account of the more recent interior point methods, originally designed

for the simpler problems of linear and quadratic programming, and whose application to more general situations is the subject of active research.

This book is a translated and improved version of the monograph [43], written in French. The French monograph was used as the textbook of an intensive two week course given several times by the authors, both in France and abroad. Each topic was presented from a theoretical point of view in morning lectures. The afternoons were devoted to implementation issues and related computational work. The conception of such a course is due to J.-B. Hiriart-Urruty, to whom the authors are deeply indebted.

Finally, three of the authors express their warm gratitude to Claude Lemaréchal for having given the impetus to this new work by providing a first English version.

Notes on this revised edition. Besides minor corrections, the present version contains substantial changes with respect to the first edition. First of all, (simplified but) nontrivial application problems have been inserted. They involve the typical operations to be performed when one is faced with a real-life application: modelling, choice of methodology and some theoretical work to motivate it, computer implementation. Such computational exercises help getting a better understanding of optimization methods beyond their theoretical description, by addressing important features to be taken into account when passing to implementation of any numerical algorithm.

In addition, the theoretical background in Part I now includes a discussion on global convergence, and a section on the classical pivotal approach to quadratic programming. Part II has been completely reorganized and expanded. The introductory chapter, on basic subdifferential calculus and duality theory, has two examples of nonsmooth functions that appear often in practice and serve as motivation (pointwise maximum and dual functions). A new section on convergence results for bundle methods has been added. The chapter on applications of nonsmooth optimization, previously focusing on decomposition of complex problems via Lagrangian duality, describes also extensions of bundle methods for handling varying dimensions, for solving constrained problems, and for solving generalized equations. Also, a brief commented review of existing software for nonlinear optimization has been added in Part III.

Finally, the reader will find additional information at <http://www-rocq.inria.fr/~gilbert/bgls>. The page gathers the data for running the test problems, various optimization codes, including an SQP solver (in Matlab), and pieces of software that solve the computational exercises.

Paris, Grenoble, Rio de Janeiro,
May 2006

J. Frédéric Bonnans
J. Charles Gilbert
Claude Lemaréchal
Claudia A. Sagastizábal

Table of Contents

Preliminaries

1	General Introduction	3
1.1	Generalities on Optimization	3
1.1.1	The Problem	3
1.1.2	Classification	4
1.2	Motivation and Examples	5
1.2.1	Molecular Biology	5
1.2.2	Meteorology	6
1.2.3	Trajectory of a Deepwater Vehicle	8
1.2.4	Optimization of Power Management	9
1.3	General Principles of Resolution	10
1.4	Convergence: Global Aspects	12
1.5	Convergence: Local Aspects	14
1.6	Computing the Gradient	16
	Bibliographical Comments	19

Part I Unconstrained Problems

2	Basic Methods	25
2.1	Existence Questions	25
2.2	Optimality Conditions	26
2.3	First-Order Methods	27
2.3.1	Gauss-Seidel	27
2.3.2	Method of Successive Approximations, or Gradient Method	28
2.4	Link with the General Descent Scheme	28
2.4.1	Choosing the ℓ_1 -Norm	29
2.4.2	Choosing the ℓ_2 -Norm	30
2.5	Steepest-Descent Method	30
2.6	Implementation	34
	Bibliographical Comments	35

3	Line-Searches	37
	3.1 General Scheme	37
	3.2 Computing the New t	40
	3.3 Optimal Stepsize (for the record only)	42
	3.4 Modern Line-Search: Wolfe's Rule	43
	3.5 Other Line-Searches: Goldstein and Price, Armijo	47
	3.5.1 Goldstein and Price	47
	3.5.2 Armijo	47
	3.5.3 Remark on the Choice of Constants	48
	3.6 Implementation Considerations	49
	Bibliographical Comments	50
4	Newtonian Methods	51
	4.1 Preliminaries	51
	4.2 Forcing Global Convergence	52
	4.3 Alleviating the Method	53
	4.4 Quasi-Newton Methods	54
	4.5 Global Convergence	57
	4.6 Local Convergence: Generalities	59
	4.7 Local Convergence: BFGS	61
	Bibliographical Comments	65
5	Conjugate Gradient	67
	5.1 Outline of Conjugate Gradient	67
	5.2 Developing the Method	69
	5.3 Computing the Direction	70
	5.4 The Algorithm Seen as an Orthogonalization Process	70
	5.5 Application to Non-Quadratic Functions	72
	5.6 Relation with Quasi-Newton	74
	Bibliographical Comments	75
6	Special Methods	77
	6.1 Trust-Regions	77
	6.1.1 The Elementary Problem	78
	6.1.2 The Elementary Mechanism: Curvilinear Search	79
	6.1.3 Incidence on the Sequence x_k	81
	6.2 Least-Squares Problems: Gauss-Newton	82
	6.3 Large-Scale Problems: Limited-Memory Quasi-Newton	84
	6.4 Truncated Newton	86
	6.5 Quadratic Programming	88
	6.5.1 The basic mechanism	89
	6.5.2 The solution algorithm	90
	6.5.3 Convergence	92
	Bibliographical Comments	95

7	A Case Study: Seismic Reflection Tomography	97
7.1	Modelling	97
7.2	Computation of the Reflection Points	99
7.3	Gradient of the Traveltime	100
7.4	The Least-Squares Problem to Solve	101
7.5	Solving the Seismic Reflection Tomography Problem	102
	General Conclusion	103

Part II Nonsmooth Optimization

8	Introduction to Nonsmooth Optimization	109
8.1	First Elements of Convex Analysis	109
8.2	Lagrangian Relaxation and Duality	111
8.2.1	Primal-Dual Relations	111
8.2.2	Back to the Primal. Recovering Primal Solutions	113
8.3	Two Convex Nondifferentiable Functions	116
8.3.1	Finite Minimax Problems	116
8.3.2	Dual Functions in Lagrangian Duality	117
9	Some Methods in Nonsmooth Optimization	119
9.1	Why Special Methods?	119
9.2	Descent Methods	120
9.2.1	Steepest-Descent Method	121
9.2.2	Stabilization. A Dual Approach. The ε -subdifferential	124
9.3	Two Black-Box Methods	126
9.3.1	Subgradient Methods	127
9.3.2	Cutting-Planes Method	130
10	Bundle Methods. The Quest for Descent	137
10.1	Stabilization. A Primal Approach	137
10.2	Some Examples of Stabilized Problems	140
10.3	Penalized Bundle Methods	141
10.3.1	A Trip to the Dual Space	144
10.3.2	Managing the Bundle. Aggregation	147
10.3.3	Updating the Penalization Parameter. Reversal Forms	150
10.3.4	Convergence Analysis	154
11	Applications of Nonsmooth Optimization	161
11.1	Divide to conquer. Decomposition methods	161
11.1.1	Price Decomposition	163
11.1.2	Resource Decomposition	167
11.1.3	Variable Partitioning or Benders Decomposition	169
11.1.4	Other Decomposition Methods	171

11.2	Transpassing Frontiers	172
11.2.1	Dynamic Bundle Methods	173
11.2.2	Constrained Bundle Methods	177
11.2.3	Bundle Methods for Generalized Equations	180
12	Computational Exercises	183
12.1	Building Prototypical NSO Black Boxes	183
12.1.1	The Function MAXQUAD	183
12.1.2	The Function MAXANAL	184
12.2	Implementation of Some NSO Methods	185
12.3	Running the Codes	186
12.4	Improving the Bundle Implementation	187
12.5	Decomposition Application	187

Part III Newton's Methods in Constrained Optimization

13	Background	197
13.1	Differential Calculus	197
13.2	Existence and Uniqueness of Solutions	199
13.3	First-Order Optimality Conditions	200
13.4	Second-Order Optimality Conditions	202
13.5	Speed of Convergence	203
13.6	Projection onto a Closed Convex Set	205
13.7	The Newton Method	205
13.8	The Hanging Chain Project I	208
	Notes	213
	Exercises	214
14	Local Methods for Problems with Equality Constraints ..	215
14.1	Newton's Method	216
14.2	Adapted Decompositions of \mathbb{R}^n	222
14.3	Local Analysis of Newton's Method	227
14.4	Computation of the Newton Step	230
14.5	Reduced Hessian Algorithm	235
14.6	A Comparison of the Algorithms	243
14.7	The Hanging Chain Project II	245
	Notes	250
	Exercises	251
15	Local Methods for Problems with Equality and Inequality Constraints	255
15.1	The SQP Algorithm	256
15.2	Primal-Dual Quadratic Convergence	259
15.3	Primal Superlinear Convergence	264

15.4 The Hanging Chain Project III	267
Notes	270
Exercise	270
16 Exact Penalization	271
16.1 Overview	271
16.2 The Lagrangian	274
16.3 The Augmented Lagrangian	275
16.4 Nondifferentiable Augmented Function	279
Notes	284
Exercises	285
17 Globalization by Line-Search	289
17.1 Line-Search SQP Algorithms	291
17.2 Truncated SQP	298
17.3 From Global to Local	307
17.4 The Hanging Chain Project IV	316
Notes	320
Exercises	321
18 Quasi-Newton Versions	323
18.1 Principles	323
18.2 Quasi-Newton SQP	327
18.3 Reduced Quasi-Newton Algorithm	331
18.4 The Hanging Chain Project V	340

Part IV Interior-Point Algorithms for Linear and Quadratic Optimization

19 Linearly Constrained Optimization and Simplex Algorithm	353
19.1 Existence of Solutions	353
19.1.1 Existence Result	353
19.1.2 Basic Points and Extensions	355
19.2 Duality	356
19.2.1 Introducing the Dual Problem	357
19.2.2 Concept of Saddle-Point	358
19.2.3 Other Formulations	362
19.2.4 Strict Complementarity	363
19.3 The Simplex Algorithm	364
19.3.1 Computing the Descent Direction	364
19.3.2 Stating the algorithm	365
19.3.3 Dual simplex	367
19.4 Comments	368

20	Linear Monotone Complementarity and Associated Vector Fields	371
20.1	Logarithmic Penalty and Central Path	371
20.1.1	Logarithmic Penalty	371
20.1.2	Central Path	372
20.2	Linear Monotone Complementarity	373
20.2.1	General Framework	374
20.2.2	A Group of Transformations	377
20.2.3	Standard Form	378
20.2.4	Partition of Variables and Canonical Form	379
20.2.5	Magnitudes in a Neighborhood of the Central Path	380
20.3	Vector Fields Associated with the Central Path	382
20.3.1	General Framework	383
20.3.2	Scaling the Problem	383
20.3.3	Analysis of the Directions	384
20.3.4	Modified Field	387
20.4	Continuous Trajectories	389
20.4.1	Limit Points of Continuous Trajectories	389
20.4.2	Developing Affine Trajectories and Directions	391
20.4.3	Mizuno's Lemma	393
20.5	Comments	393
21	Predictor-Corrector Algorithms	395
21.1	Overview	395
21.2	Statement of the Methods	396
21.2.1	General Framework for Primal-Dual Algorithms	396
21.2.2	Weighting After Displacement	397
21.2.3	The Predictor-Corrector Method	397
21.3	A small-Neighborhood Algorithm	398
21.3.1	Statement of the Algorithm. Main Result	398
21.3.2	Analysis of the Centralization Move	398
21.3.3	Analysis of the Affine Step and Global Convergence	399
21.3.4	Asymptotic Speed of Convergence	401
21.4	A Predictor-Corrector Algorithm with Modified Field	402
21.4.1	Principle	402
21.4.2	Statement of the Algorithm. Main Result	404
21.4.3	Complexity Analysis	404
21.4.4	Asymptotic Analysis	405
21.5	A Large-Neighborhood Algorithm	406
21.5.1	Statement of the Algorithm. Main Result	406
21.5.2	Analysis of the Centering Step	407
21.5.3	Analysis of the Affine Step	408
21.5.4	Asymptotic Convergence	408
21.6	Practical Aspects	408
21.7	Comments	409

22 Non-Feasible Algorithms	411
22.1 Overview	411
22.2 Principle of the Non-Feasible Path Following	411
22.2.1 Non-Feasible Central Path	411
22.2.2 Directions of Move	412
22.2.3 Orders of Magnitude of Approximately Centered Points	413
22.2.4 Analysis of Directions	415
22.2.5 Modified Field	418
22.3 Non-Feasible Predictor-Corrector Algorithm	419
22.3.1 Complexity Analysis	420
22.3.2 Asymptotic Analysis	422
22.4 Comments	422
23 Self-Duality	425
23.1 Overview	425
23.2 Linear Problems with Inequality Constraints	425
23.2.1 A Family of Self-Dual Linear Problems	425
23.2.2 Embedding in a Self-Dual Problem	427
23.3 Linear Problems in Standard Form	429
23.3.1 The Associated Self-Dual Homogeneous System	429
23.3.2 Embedding in a Feasible Self-Dual Problem	430
23.4 Practical Aspects	431
23.5 Extension to Linear Monotone Complementarity Problems	433
23.6 Comments	434
24 One-Step Methods	435
24.1 Overview	435
24.2 The Largest-Step Method	436
24.2.1 Largest-Step Algorithm	436
24.2.2 Largest-Step Algorithm with Safeguard	436
24.3 Centralization in the Space of Large Variables	437
24.3.1 One-Sided Distance	437
24.3.2 Convergence with Strict Complementarity	441
24.3.3 Convergence without Strict Complementarity	443
24.3.4 Relative Distance in the Space of Large Variables	444
24.4 Convergence Analysis	445
24.4.1 Global Convergence of the Largest-Step Algorithm	445
24.4.2 Local Convergence of the Largest-Step Algorithm	446
24.4.3 Convergence of the Largest-Step Algorithm with Safeguard	447
24.5 Comments	450

25 Complexity of Linear Optimization Problems with Integer Data	451
25.1 Overview	451
25.2 Main Results	452
25.2.1 General Hypotheses	452
25.2.2 Statement of the Results	452
25.2.3 Application	453
25.3 Solving a System of Linear Equations	453
25.4 Proofs of the Main Results	455
25.4.1 Proof of Theorem 25.1	455
25.4.2 Proof of Theorem 25.2	455
25.5 Comments	456
26 Karmarkar's Algorithm	457
26.1 Overview	457
26.2 Linear Problem in Projective Form	457
26.2.1 Projective Form and Karmarkar Potential	457
26.2.2 Minimizing the Potential and Solving (<i>PLP</i>)	458
26.3 Statement of Karmarkar's Algorithm	459
26.4 Analysis of the Algorithm	460
26.4.1 Complexity Analysis	460
26.4.2 Analysis of the Potential Decrease	460
26.4.3 Estimating the Optimal Cost	461
26.4.4 Practical Aspects	462
26.5 Comments	463
References	465
Index	485