

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

# Numerical Optimization

Theoretical and Practical Aspects

With 26 Figures



Springer

J. Frédéric Bonnans  
INRIA Rocquencourt  
BP 105  
78153 Le Chesnay  
France  
*e-mail*: Frederic.Bonnans@inria.fr

J. Charles Gilbert  
INRIA Rocquencourt  
BP 105  
78153 Le Chesnay  
France  
*e-mail*: Jean-Charles.Gilbert@inria.fr

Claude Lemaréchal  
INRIA Rhône-Alpes  
655, avenue de l'Europe  
Montbonnot  
38334 Saint Ismier  
France  
*e-mail*: Claude.Lemarechal@inria.fr

Claudia A. Sagastizábal  
IMPA  
110, Estrada Dona Castorina  
22460-320 Jardim Botânico  
Rio de Janeiro-RJ  
Brazil  
*e-mail*: sagastiz@impa.br

The original French edition “Optimisation Numérique” was published by Springer-Verlag Berlin Heidelberg, 1997.

**Cover figure:** The figure illustrates the behaviour of a bundle method on MAXQUAD, a standard test problem for nonsmooth optimization; see page 131.

Catalog-in-Publication Data applied for

A catalog record for this book is available from the Library of Congress.

Bibliographic information published by Die Deutsche Bibliothek Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.

ISBN 978-3-540-00191-1 ISBN 978-3-662-05078-1 (eBook)  
DOI 10.1007/978-3-662-05078-1

---

Mathematics Subject Classification (2000): 65K10, 90-01, 90-08, 90CXX

---

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag Berlin Heidelberg GmbH. Violations are liable for prosecution under the German Copyright Law.

<http://www.springer.de>

© Springer-Verlag Berlin Heidelberg 2003  
Originally published by Springer-Verlag Berlin Heidelberg New York in 2003

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Cover design: *design & production* GmbH, Heidelberg  
Typeset by the authors using a  $\text{\LaTeX}$  macro package  
Production: LE- $\text{\TeX}$  Jelonek, Schmidt & Vöckler GbR, Leipzig

Printed on acid-free paper

SPIN 10836974

41/3142YL - 5 4 3 2 1 0

# Table of Contents

---

## Preliminaries

---

<b>1</b>	<b>General Introduction</b> .....	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	General Convergence Theorem: Zangwill .....	12
1.5	Generalities on Convergence .....	13
1.6	Computing the Gradient .....	16

---

## Part I Unconstrained Problems

---

<b>2</b>	<b>Basic Methods</b> .....	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 $\ell_1$ -Norm .....	29
2.4.2	Choosing the $\ell_2$ -Norm .....	30
2.5	Steepest-Descent Method .....	30
2.6	Implementation .....	34
<b>3</b>	<b>Line-Searches</b> .....	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
<b>4</b>	<b>Newtonian Methods</b> . . . . .	<b>51</b>
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
<b>5</b>	<b>Conjugate Gradient</b> . . . . .	<b>67</b>
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
<b>6</b>	<b>Special Methods</b> . . . . .	<b>77</b>
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

---

## Part II Nonsmooth Optimization

---

<b>7</b>	<b>Some Theory of Nonsmooth Optimization</b> . . . . .	<b>95</b>
7.1	First Elements of Convex Analysis . . . . .	95
7.2	Lagrangian Relaxation and Duality . . . . .	97
7.3	Two Convex Nondifferentiable Functions . . . . .	98
7.3.1	Finite Minimax Problems . . . . .	99
7.3.2	Dual Functions in Lagrangian Duality . . . . .	99

<b>8</b>	<b>Some Methods in Nonsmooth Optimization</b> . . . . .	101
8.1	Why Special Methods? . . . . .	101
8.2	Descent Methods . . . . .	102
8.2.1	Steepest-Descent Method . . . . .	103
8.2.2	Stabilization. A Dual Approach. The $\varepsilon$ -subdifferential . . . . .	106
8.3	Two Black-Box Methods . . . . .	108
8.3.1	Subgradient Methods . . . . .	108
8.3.2	Cutting-Planes Method . . . . .	111
<b>9</b>	<b>Bundle Methods. The Quest of Descent</b> . . . . .	117
9.1	Stabilization. A Primal Approach . . . . .	117
9.2	Some Examples of Stabilized Problems . . . . .	120
9.3	Penalized Bundle Methods . . . . .	121
9.3.1	A Trip to the Dual Space . . . . .	122
9.3.2	Managing the Bundle. Aggregation . . . . .	126
9.3.3	Updating the Penalization Parameter. Reversal Forms . . . . .	129
9.3.4	Relation with Proximal Methods for Generalized Equations . . . . .	131
9.3.5	Convergence Analysis . . . . .	133
<b>10</b>	<b>Decomposition and Duality</b> . . . . .	137
10.1	Primal-Dual Relations . . . . .	138
10.2	Back to the Primal. Recovering Primal Solutions . . . . .	139
10.3	Price Decomposition . . . . .	142
10.3.1	General Algorithm . . . . .	142
10.3.2	Relation with Dantzig-Wolfe Decomposition . . . . .	144
10.4	Resource Decomposition . . . . .	145
10.5	Other Decomposition Techniques . . . . .	147

---

### Part III Newton's Methods in Constrained Optimization

---

<b>11</b>	<b>Background</b> . . . . .	157
11.1	Differential Calculus . . . . .	157
11.2	Existence and Uniqueness of Solutions . . . . .	159
11.3	First-Order Optimality Conditions . . . . .	160
11.4	Second-Order Optimality Conditions . . . . .	162
11.5	Speed of Convergence . . . . .	163
11.6	Projection onto a Closed Convex Set . . . . .	164
11.7	The Newton Method . . . . .	165
	Exercises . . . . .	167

**12 Local Methods for Problems with Equality Constraints** ... 169

- 12.1 Newton's Method ..... 170
- 12.2 Adapted Decompositions of  $\mathbb{R}^n$  ..... 176
- 12.3 Local Analysis of Newton's Method ..... 181
- 12.4 Computation of the Newton Step ..... 184
- 12.5 Reduced Hessian Algorithm ..... 190
- 12.6 A Comparison of the Algorithms ..... 197
- Notes ..... 199
- Exercises ..... 200

**13 Local Methods for Problems with Equality and Inequality Constraints** ..... 203

- 13.1 The SQP Algorithm ..... 204
- 13.2 Primal-Dual Quadratic Convergence ..... 207
- 13.3 Primal Superlinear Convergence ..... 212
- Notes ..... 214

**14 Exact Penalization** ..... 217

- 14.1 Overview ..... 217
- 14.2 The Lagrangian ..... 220
- 14.3 The Augmented Lagrangian ..... 221
- 14.4 Nondifferentiable Augmented Function ..... 224
- Notes ..... 230
- Exercises ..... 230

**15 Globalization by Line-Search** ..... 235

- 15.1 Line-Search SQP Algorithms ..... 237
- 15.2 Truncated SQP ..... 244
- 15.3 From Global to Local ..... 253
- Notes ..... 262
- Exercises ..... 263

**16 Quasi-Newton Versions** ..... 265

- 16.1 Principles ..... 265
- 16.2 Quasi-Newton SQP ..... 269
- 16.3 Reduced Quasi-Newton Algorithm ..... 273

---

**Part IV Interior-Point Algorithms for Linear and Quadratic Optimization**

---

**17 Linearly Constrained Optimization and Simplex Algorithm** 291

- 17.1 Existence of Solutions ..... 291
  - 17.1.1 Existence Result ..... 291
  - 17.1.2 Basic Points and Extensions ..... 293

17.2	Duality	294
17.2.1	Introducing the Dual Problem	294
17.2.2	Concept of Saddle-Point	296
17.2.3	Other Formulations	300
17.2.4	Strict Complementarity	301
17.3	The Simplex Algorithm	302
17.3.1	Computing the Descent Direction	302
17.3.2	Stating the algorithm	303
17.4	Comments	305
<b>18</b>	<b>Linear Monotone Complementarity and Associated Vector Fields</b>	<b>307</b>
18.1	Logarithmic Penalty and Central Path	307
18.1.1	Logarithmic Penalty	307
18.1.2	Central Path	308
18.2	Linear Monotone Complementarity	309
18.2.1	General Framework	309
18.2.2	A Group of Transformations	313
18.2.3	Standard Form	314
18.2.4	Partition of Variables and Canonical Form	315
18.2.5	Magnitudes in a Neighborhood of the Central Path	315
18.3	Vector Fields Associated with the Central Path	317
18.3.1	General Framework	318
18.3.2	Scaling the Problem	318
18.3.3	Analysis of the Directions	319
18.3.4	Modified Field	322
18.4	Continuous Trajectories	324
18.4.1	Limit Points of Continuous Trajectories	324
18.4.2	Developing Affine Trajectories and Directions	326
18.4.3	Mizuno's Lemma	327
18.5	Comments	328
<b>19</b>	<b>Predictor-Corrector Algorithms</b>	<b>329</b>
19.1	Overview	329
19.2	Statement of the Methods	330
19.2.1	General Framework for Primal-Dual Algorithms	330
19.2.2	Weighting After Displacement	331
19.2.3	The Predictor-Corrector Method	331
19.3	A small-Neighborhood Algorithm	332
19.3.1	Statement of the Algorithm. Main Result	332
19.3.2	Analysis of the Centralization Move	333
19.3.3	Analysis of the Affine Step and Global Convergence	334
19.3.4	Asymptotic Speed of Convergence	335
19.4	A Predictor-Corrector Algorithm with Modified Field	336
19.4.1	Principle	336

19.4.2	Statement of the Algorithm. Main Result . . . . .	338
19.4.3	Complexity Analysis . . . . .	339
19.4.4	Asymptotic Analysis . . . . .	340
19.5	A Large-Neighborhood Algorithm . . . . .	340
19.5.1	Statement of the Algorithm. Main Result . . . . .	340
19.5.2	Analysis of the Centering Step . . . . .	341
19.5.3	Analysis of the Affine Step . . . . .	342
19.5.4	Asymptotic Convergence . . . . .	343
19.6	Practical Aspects . . . . .	343
19.7	Comments . . . . .	344
<b>20</b>	<b>Non-Feasible Algorithms</b> . . . . .	<b>345</b>
20.1	Overview . . . . .	345
20.2	Principle of the Non-Feasible Path Following . . . . .	345
20.2.1	Non-Feasible Central Path . . . . .	345
20.2.2	Directions of Move . . . . .	346
20.2.3	Orders of Magnitude of Approximately Centered Points . . . . .	347
20.2.4	Analysis of Directions . . . . .	349
20.2.5	Modified Field . . . . .	352
20.3	Non-Feasible Predictor-Corrector Algorithm . . . . .	353
20.3.1	Complexity Analysis . . . . .	354
20.3.2	Asymptotic Analysis . . . . .	356
20.4	Comments . . . . .	356
<b>21</b>	<b>Self-Duality</b> . . . . .	<b>357</b>
21.1	Overview . . . . .	357
21.2	Linear Problems with Inequality Constraints . . . . .	357
21.2.1	A Family of Self-Dual Linear Problems . . . . .	357
21.2.2	Embedding in a Self-Dual Problem . . . . .	359
21.3	Linear Problems in Standard Form . . . . .	361
21.3.1	The Associated Self-Dual Homogeneous System . . . . .	361
21.3.2	Embedding in a Feasible Self-Dual Problem . . . . .	362
21.4	Practical Aspects . . . . .	363
21.5	Extension to Linear Monotone Complementarity Problems . . . . .	365
21.6	Comments . . . . .	366
<b>22</b>	<b>One-Step Methods</b> . . . . .	<b>367</b>
22.1	Overview . . . . .	367
22.2	The Largest-Step Method . . . . .	368
22.2.1	Largest-Step Algorithm . . . . .	368
22.2.2	Largest-Step Algorithm with Safeguard . . . . .	368
22.3	Centralization in the Space of Large Variables . . . . .	369
22.3.1	One-Sided Distance . . . . .	369
22.3.2	Convergence with Strict Complementarity . . . . .	373
22.3.3	Convergence without Strict Complementarity . . . . .	375



22.3.4	Relative Distance in the Space of Large Variables . . . . .	376
22.4	Convergence Analysis . . . . .	377
22.4.1	Global Convergence of the Largest-Step Algorithm . . .	377
22.4.2	Local Convergence of the Largest-Step Algorithm . . . .	378
22.4.3	Convergence of the Largest-Step Algorithm with Safeguard . . . . .	379
22.5	Comments . . . . .	382
<b>23</b>	<b>Complexity of Linear Optimization Problems with Integer Data . . . . .</b>	<b>383</b>
23.1	Overview . . . . .	383
23.2	Main Results . . . . .	384
23.2.1	General Hypotheses . . . . .	384
23.2.2	Statement of the Results . . . . .	384
23.2.3	Application . . . . .	385
23.3	Solving a System of Linear Equations . . . . .	385
23.4	Proofs of the Main Results . . . . .	387
23.4.1	Proof of Theorem 23.1 . . . . .	387
23.4.2	Proof of Theorem 23.2 . . . . .	387
23.5	Comments . . . . .	388
<b>24</b>	<b>Karmarkar's Algorithm . . . . .</b>	<b>389</b>
24.1	Overview . . . . .	389
24.2	Linear Problem in Projective Form . . . . .	389
24.2.1	Projective Form and Karmarkar Potential . . . . .	389
24.2.2	Minimizing the Potential and Solving ( <i>PLP</i> ) . . . . .	390
24.3	Statement of Karmarkar's Algorithm . . . . .	391
24.4	Analysis of the Algorithm . . . . .	392
24.4.1	Complexity Analysis . . . . .	392
24.4.2	Analysis of the Potential Decrease . . . . .	392
24.4.3	Estimating the Optimal Cost . . . . .	393
24.4.4	Practical Aspects . . . . .	394
24.5	Comments . . . . .	395
	<b>References . . . . .</b>	<b>397</b>
	<b>Index . . . . .</b>	<b>415</b>