

Jean-Baptiste Hiriart-Urruty  
Claude Lemaréchal

# Convex Analysis and Minimization Algorithms II

Advanced Theory  
and Bundle Methods

With 64 Figures



Springer-Verlag  
Berlin Heidelberg New York  
London Paris Tokyo  
Hong Kong Barcelona  
Budapest

# Table of Contents Part II

Introduction . . . . .	XV
IX. Inner Construction of the Subdifferential . . . . .	1
1 The Elementary Mechanism . . . . .	2
2 Convergence Properties . . . . .	9
2.1 Convergence . . . . .	9
2.2 Speed of Convergence . . . . .	15
3 Putting the Mechanism in Perspective . . . . .	24
3.1 Bundling as a Substitute for Steepest Descent . . . . .	24
3.2 Bundling as an Emergency Device for Descent Methods . . . . .	27
3.3 Bundling as a Separation Algorithm . . . . .	29
X. Conjugacy in Convex Analysis . . . . .	35
1 The Convex Conjugate of a Function . . . . .	37
1.1 Definition and First Examples . . . . .	37
1.2 Interpretations . . . . .	40
1.3 First Properties . . . . .	42
1.4 Subdifferentials of Extended-Valued Functions . . . . .	47
1.5 Convexification and Subdifferentiability . . . . .	49
2 Calculus Rules on the Conjugacy Operation . . . . .	54
2.1 Image of a Function Under a Linear Mapping . . . . .	54
2.2 Pre-Composition with an Affine Mapping . . . . .	56
2.3 Sum of Two Functions . . . . .	61
2.4 Infima and Suprema . . . . .	65
2.5 Post-Composition with an Increasing Convex Function . . . . .	69
2.6 A Glimpse of Biconjugate Calculus . . . . .	71
3 Various Examples . . . . .	72
3.1 The Cramer Transformation . . . . .	72
3.2 Some Results on the Euclidean Distance to a Closed Set . . . . .	73
3.3 The Conjugate of Convex Partially Quadratic Functions . . . . .	75
3.4 Polyhedral Functions . . . . .	76
4 Differentiability of a Conjugate Function . . . . .	79
4.1 First-Order Differentiability . . . . .	79
4.2 Towards Second-Order Differentiability . . . . .	82

XI. Approximate Subdifferentials of Convex Functions . . . . .	91
1 The Approximate Subdifferential . . . . .	92
1.1 Definition, First Properties and Examples . . . . .	92
1.2 Characterization via the Conjugate Function . . . . .	95
1.3 Some Useful Properties . . . . .	98
2 The Approximate Directional Derivative . . . . .	102
2.1 The Support Function of the Approximate Subdifferential . . . . .	102
2.2 Properties of the Approximate Difference Quotient . . . . .	106
2.3 Behaviour of $f'_\varepsilon$ and $T_\varepsilon$ as Functions of $\varepsilon$ . . . . .	110
3 Calculus Rules on the Approximate Subdifferential . . . . .	113
3.1 Sum of Functions . . . . .	113
3.2 Pre-Composition with an Affine Mapping . . . . .	116
3.3 Image and Marginal Functions . . . . .	118
3.4 A Study of the Infimal Convolution . . . . .	119
3.5 Maximum of Functions . . . . .	123
3.6 Post-Composition with an Increasing Convex Function . . . . .	125
4 The Approximate Subdifferential as a Multifunction . . . . .	127
4.1 Continuity Properties of the Approximate Subdifferential . . . . .	127
4.2 Transportation of Approximate Subgradients . . . . .	129
XII. Abstract Duality for Practitioners . . . . .	137
1 The Problem and the General Approach . . . . .	137
1.1 The Rules of the Game . . . . .	137
1.2 Examples . . . . .	141
2 The Necessary Theory . . . . .	147
2.1 Preliminary Results: The Dual Problem . . . . .	147
2.2 First Properties of the Dual Problem . . . . .	150
2.3 Primal-Dual Optimality Characterizations . . . . .	154
2.4 Existence of Dual Solutions . . . . .	157
3 Illustrations . . . . .	161
3.1 The Minimax Point of View . . . . .	161
3.2 Inequality Constraints . . . . .	162
3.3 Dualization of Linear Programs . . . . .	165
3.4 Dualization of Quadratic Programs . . . . .	166
3.5 Steepest-Descent Directions . . . . .	168
4 Classical Dual Algorithms . . . . .	170
4.1 Subgradient Optimization . . . . .	171
4.2 The Cutting-Plane Algorithm . . . . .	174
5 Putting the Method in Perspective . . . . .	178
5.1 The Primal Function . . . . .	178
5.2 Augmented Lagrangians . . . . .	181
5.3 The Dualization Scheme in Various Situations . . . . .	185
5.4 Fenchel's Duality . . . . .	190

XIII. Methods of $\varepsilon$ -Descent . . . . .	195
1 Introduction. Identifying the Approximate Subdifferential . . . . .	195
1.1 The Problem and Its Solution . . . . .	195
1.2 The Line-Search Function . . . . .	199
1.3 The Schematic Algorithm . . . . .	203
2 A Direct Implementation: Algorithm of $\varepsilon$ -Descent . . . . .	206
2.1 Iterating the Line-Search . . . . .	206
2.2 Stopping the Line-Search . . . . .	209
2.3 The $\varepsilon$ -Descent Algorithm and Its Convergence . . . . .	212
3 Putting the Algorithm in Perspective . . . . .	216
3.1 A Pure Separation Form . . . . .	216
3.2 A Totally Static Minimization Algorithm . . . . .	219
XIV. Dynamic Construction of Approximate Subdifferentials:	
Dual Form of Bundle Methods . . . . .	223
1 Introduction: The Bundle of Information . . . . .	223
1.1 Motivation . . . . .	223
1.2 Constructing the Bundle of Information . . . . .	227
2 Computing the Direction . . . . .	233
2.1 The Quadratic Program . . . . .	233
2.2 Minimality Conditions . . . . .	236
2.3 Directional Derivatives Estimates . . . . .	241
2.4 The Role of the Cutting-Plane Function . . . . .	244
3 The Implementable Algorithm . . . . .	248
3.1 Derivation of the Line-Search . . . . .	248
3.2 The Implementable Line-Search and Its Convergence . . . . .	250
3.3 Derivation of the Descent Algorithm . . . . .	254
3.4 The Implementable Algorithm and Its Convergence . . . . .	257
4 Numerical Illustrations . . . . .	263
4.1 Typical Behaviour . . . . .	263
4.2 The Role of $\varepsilon$ . . . . .	266
4.3 A Variant with Infinite $\varepsilon$ : Conjugate Subgradients . . . . .	268
4.4 The Role of the Stopping Criterion . . . . .	269
4.5 The Role of Other Parameters . . . . .	271
4.6 General Conclusions . . . . .	273
XV. Acceleration of the Cutting-Plane Algorithm:	
Primal Forms of Bundle Methods . . . . .	275
1 Accelerating the Cutting-Plane Algorithm . . . . .	275
1.1 Instability of Cutting Planes . . . . .	276
1.2 Stabilizing Devices: Leading Principles . . . . .	279
1.3 A Digression: Step-Control Strategies . . . . .	283
2 A Variety of Stabilized Algorithms . . . . .	285
2.1 The Trust-Region Point of View . . . . .	286

2.2 The Penalization Point of View . . . . .	289
2.3 The Relaxation Point of View . . . . .	292
2.4 A Possible Dual Point of View . . . . .	295
2.5 Conclusion . . . . .	299
3 A Class of Primal Bundle Algorithms . . . . .	301
3.1 The General Method . . . . .	301
3.2 Convergence . . . . .	307
3.3 Appropriate Stepsize Values . . . . .	314
4 Bundle Methods as Regularizations . . . . .	317
4.1 Basic Properties of the Moreau-Yosida Regularization . . . . .	317
4.2 Minimizing the Moreau-Yosida Regularization . . . . .	322
4.3 Computing the Moreau-Yosida Regularization . . . . .	326
Bibliographical Comments . . . . .	331
References . . . . .	337
Index . . . . .	345