

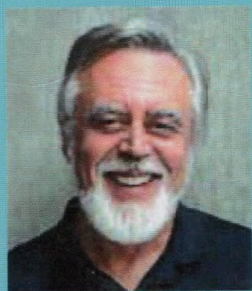
This is a thoroughly rewritten and substantially expanded edition of our best-selling nonlinear programming book. The treatment focuses on iterative algorithms for constrained and unconstrained optimization, Lagrange multipliers and duality, large scale problems, and the interface between continuous and discrete optimization.

“This is a beautifully written book by a prolific author...who has taken painstaking care in making the presentation extremely lucid... The style is unhurried and intuitive yet mathematically rigorous...The numerous figures in the book are extremely well thought out and are used in a very effective way to elucidate the text. The detailed and self-explanatory long captions accompanying each figure are extremely helpful.”

—From the review of the 1st edition by Olvi Mangasarian (Optima, March 1997).

Among its special features, this edition of the book:

- Provides extensive coverage of iterative optimization methods within a unifying framework
- Covers in depth duality theory from both a variational and a geometric/convex analysis point of view
- Provides a detailed treatment of interior point methods for linear programming
- Includes much new material on a number of topics, such as neural network training, conic programming, proximal algorithms, and alternating direction methods of multipliers
- Focuses on large-scale optimization topics of much current interest, such as incremental methods, and parallel and distributed computation, and their applications in machine learning and signal processing
- Includes a large number of examples and exercises
- Was developed through extensive classroom use in first-year graduate courses.



DIMITRI P. BERTSEKAS, a member of the U.S. National Academy of Engineering, is McAfee Professor of Engineering at the Massachusetts Institute of Technology. Among others, he has received the 2001 AACC John R. Ragazzini Education Award, the 2009 INFORMS Expository Writing Award, the 2014 AACC Richard Bellman Heritage Award, the 2014 Khachiyan Prize, and the 2015 SIAM/MOS George B. Dantzig award.

Related Athena Scientific books by the same author:

Convex Optimization Theory, 2009

Convex Optimization Algorithms, 2015

Dynamic Programming and Optimal Control, 2012

Network Optimization: Continuous and Discrete Models, 1998

Constrained Optimization and Lagrange Multiplier Methods, 1982

Visit Athena Scientific online at:
www.athenasc.com

ISBN 978-1-886529-05-2



9 781886 529052

Contents

- 1. Unconstrained Optimization: Basic Methods p. 1
 - 1.1. Optimality Conditions p. 5
 - 1.1.1. Variational Ideas p. 5
 - 1.1.2. Main Optimality Conditions p. 15
 - 1.2. Gradient Methods – Convergence p. 28
 - 1.2.1. Descent Directions and Stepsize Rules p. 28
 - 1.2.2. Convergence Results p. 49
 - 1.3. Gradient Methods – Rate of Convergence p. 67
 - 1.3.1. The Local Analysis Approach p. 69
 - 1.3.2. The Role of the Condition Number p. 70
 - 1.3.3. Convergence Rate Results p. 82
 - 1.4. Newton’s Method and Variations p. 95
 - 1.4.1. Modified Cholesky Factorization p. 101
 - 1.4.2. Trust Region Methods p. 103
 - 1.4.3. Variants of Newton’s Method p. 105
 - 1.4.4. Least Squares and the Gauss-Newton Method p. 107
 - 1.5. Notes and Sources p. 117
- 2. Unconstrained Optimization: Additional Methods . . . p. 119
 - 2.1. Conjugate Direction Methods p. 120
 - 2.1.1. The Conjugate Gradient Method p. 125
 - 2.1.2. Convergence Rate of Conjugate Gradient Method p. 132
 - 2.2. Quasi-Newton Methods p. 138
 - 2.3. Nonderivative Methods p. 148
 - 2.3.1. Coordinate Descent p. 149
 - 2.3.2. Direct Search Methods p. 154
 - 2.4. Incremental Methods p. 158
 - 2.4.1. Incremental Gradient Methods p. 161
 - 2.4.2. Incremental Aggregated Gradient Methods p. 172
 - 2.4.3. Incremental Gauss-Newton Methods p. 178
 - 2.4.3. Incremental Newton Methods p. 185
 - 2.5. Distributed Asynchronous Algorithms p. 194

2.5.1. Totally and Partially Asynchronous Algorithms	p. 197
2.5.2. Totally Asynchronous Convergence	p. 198
2.5.3. Partially Asynchronous Gradient-Like Algorithms	p. 203
2.5.4. Convergence Rate of Asynchronous Algorithms	p. 204
2.6. Discrete-Time Optimal Control Problems	p. 210
2.6.1. Gradient and Conjugate Gradient Methods for Optimal Control	p. 221
2.6.2. Newton's Method for Optimal Control	p. 222
2.7. Solving Nonlinear Programming Problems - Some Practical Guidelines	p. 227
2.8. Notes and Sources	p. 232
3. Optimization Over a Convex Set	p. 235
3.1. Constrained Optimization Problems	p. 236
3.1.1. Necessary and Sufficient Conditions for Optimality	p. 236
3.1.2. Existence of Optimal Solutions	p. 246
3.2. Feasible Directions - Conditional Gradient Method	p. 257
3.2.1. Descent Directions and Stepsize Rules	p. 257
3.2.2. The Conditional Gradient Method	p. 262
3.3. Gradient Projection Methods	p. 272
3.3.1. Feasible Directions and Stepsize Rules Based on Projection	p. 272
3.3.2. Convergence Analysis	p. 283
3.4. Two-Metric Projection Methods	p. 292
3.5. Manifold Suboptimization Methods	p. 298
3.6. Proximal Algorithms	p. 307
3.6.1. Rate of Convergence	p. 312
3.6.2. Variants of the Proximal Algorithm	p. 318
3.7. Block Coordinate Descent Methods	p. 323
3.7.1. Variants of Coordinate Descent	p. 327
3.8. Network Optimization Algorithms	p. 331
3.9. Notes and Sources	p. 338
4. Lagrange Multiplier Theory	p. 343
4.1. Necessary Conditions for Equality Constraints	p. 345
4.1.1. The Penalty Approach	p. 349
4.1.2. The Elimination Approach	p. 352
4.1.3. The Lagrangian Function	p. 356
4.2. Sufficient Conditions and Sensitivity Analysis	p. 364
4.2.1. The Augmented Lagrangian Approach	p. 365
4.2.2. The Feasible Direction Approach	p. 369
4.2.3. Sensitivity	p. 370
4.3. Inequality Constraints	p. 376
4.3.1. Karush-Kuhn-Tucker Necessary Conditions	p. 378

4.3.2. Sufficient Conditions and Sensitivity	p. 383
4.3.3. Fritz John Optimality Conditions	p. 386
4.3.4. Constraint Qualifications and Pseudonormality	p. 392
4.3.5. Abstract Set Constraints and the Tangent Cone	p. 399
4.3.6. Abstract Set Constraints, Equality, and Inequality Constraints	p. 415
4.4. Linear Constraints and Duality	p. 429
4.4.1. Convex Cost Function and Linear Constraints	p. 429
4.4.2. Duality Theory: A Simple Form for Linear Constraints	p. 432
4.5. Notes and Sources	p. 441
5. Lagrange Multiplier Algorithms	p. 445
5.1. Barrier and Interior Point Methods	p. 446
5.1.1. Path Following Methods for Linear Programming	p. 450
5.1.2. Primal-Dual Methods for Linear Programming	p. 458
5.2. Penalty and Augmented Lagrangian Methods	p. 469
5.2.1. The Quadratic Penalty Function Method	p. 471
5.2.2. Multiplier Methods – Main Ideas	p. 479
5.2.3. Convergence Analysis of Multiplier Methods	p. 488
5.2.4. Duality and Second Order Multiplier Methods	p. 492
5.2.5. Nonquadratic Augmented Lagrangians - The Exponential Method of Multipliers	p. 494
5.3. Exact Penalties – Sequential Quadratic Programming	p. 502
5.3.1. Nondifferentiable Exact Penalty Functions	p. 503
5.3.2. Sequential Quadratic Programming	p. 513
5.3.3. Differentiable Exact Penalty Functions	p. 520
5.4. Lagrangian Methods	p. 527
5.4.1. First-Order Lagrangian Methods	p. 528
5.4.2. Newton-Like Methods for Equality Constraints	p. 535
5.4.3. Global Convergence	p. 545
5.4.4. A Comparison of Various Methods	p. 548
5.5. Notes and Sources	p. 550
6. Duality and Convex Programming	p. 553
6.1. Duality and Dual Problems	p. 554
6.1.1. Geometric Multipliers	p. 556
6.1.2. The Weak Duality Theorem	p. 561
6.1.3. Primal and Dual Optimal Solutions	p. 566
6.1.4. Treatment of Equality Constraints	p. 568
6.1.5. Separable Problems and their Geometry	p. 570
6.1.6. Additional Issues About Duality	p. 575
6.2. Convex Cost – Linear Constraints	p. 582
6.3. Convex Cost – Convex Constraints	p. 589

6.4. Conjugate Functions and Fenchel Duality	p. 598
6.4.1. Conic Programming	p. 604
6.4.2. Monotropic Programming	p. 612
6.4.3. Network Optimization	p. 617
6.4.4. Games and the Minimax Theorem	p. 620
6.4.5. The Primal Function and Sensitivity Analysis	p. 623
6.5. Discrete Optimization and Duality	p. 630
6.5.1. Examples of Discrete Optimization Problems	p. 631
6.5.2. Branch-and-Bound	p. 639
6.5.3. Lagrangian Relaxation	p. 648
6.6. Notes and Sources	p. 660
7. Dual Methods	p. 663
7.1. Dual Derivatives and Subgradients	p. 666
7.2. Dual Ascent Methods for Differentiable Dual Problems	p. 673
7.2.1. Coordinate Ascent for Quadratic Programming	p. 673
7.2.2. Separable Problems and Primal Strict Convexity	p. 675
7.2.3. Partitioning and Dual Strict Concavity	p. 677
7.3. Proximal and Augmented Lagrangian Methods	p. 682
7.3.1. The Method of Multipliers as a Dual	
Proximal Algorithm	p. 682
7.3.2. Entropy Minimization and Exponential	
Method of Multipliers	p. 686
7.3.3. Incremental Augmented Lagrangian Methods	p. 687
7.4. Alternating Direction Methods of Multipliers	p. 691
7.4.1. ADMM Applied to Separable Problems	p. 699
7.4.2. Connections Between Augmented Lagrangian-	
Related Methods	p. 703
7.5. Subgradient-Based Optimization Methods	p. 709
7.5.1. Subgradient Methods	p. 709
7.5.2. Approximate and Incremental Subgradient Methods	p. 714
7.5.3. Cutting Plane Methods	p. 717
7.5.4. Ascent and Approximate Ascent Methods	p. 724
7.6. Decomposition Methods	p. 735
7.6.1. Lagrangian Relaxation of the Coupling Constraints	p. 736
7.6.2. Decomposition by Right-Hand Side Allocation	p. 739
7.7. Notes and Sources	p. 742
Appendix A: Mathematical Background	p. 745
A.1. Vectors and Matrices	p. 746
A.2. Norms, Sequences, Limits, and Continuity	p. 749
A.3. Square Matrices and Eigenvalues	p. 757
A.4. Symmetric and Positive Definite Matrices	p. 760
A.5. Derivatives	p. 765

A.6. Convergence Theorems	p. 770
-------------------------------------	--------

Appendix B: Convex Analysis p. 783

B.1. Convex Sets and Functions	p. 783
--	--------

B.2. Hyperplanes	p. 793
----------------------------	--------

B.3. Cones and Polyhedral Convexity	p. 796
---	--------

B.4. Extreme Points and Linear Programming	p. 798
--	--------

B.5. Differentiability Issues	p. 803
---	--------

Appendix C: Line Search Methods p. 809

C.1. Cubic Interpolation	p. 809
------------------------------------	--------

C.2. Quadratic Interpolation	p. 810
--	--------

C.3. The Golden Section Method	p. 812
--	--------

Appendix D: Implementation of Newton's Method . . . p. 815

D.1. Cholesky Factorization	p. 815
---------------------------------------	--------

D.2. Application to a Modified Newton Method	p. 817
--	--------

References	p. 821
----------------------	--------

Index	p. 857
-----------------	--------