
LINEAR AND NONLINEAR PROGRAMMING

Stephen G. Nash and Ariela Sofer

George Mason University

The McGraw-Hill Companies, Inc.

New York St. Louis San Francisco Auckland Bogotá Caracas
Lisbon London Madrid Mexico City Milan Montreal
New Delhi San Juan Singapore Sydney Tokyo Toronto

Contents

Preface	xv
Acknowledgments	xxv

Part I Basics

1 Optimization Models	3
1.1 Introduction	3
1.2 Linear Equations	4
1.3 Linear Programming	6
1.4 Least Squares Data Fitting	9
1.5 Nonlinear Programming	11
2 Fundamentals of Optimization	15
2.1 Introduction	15
2.2 Feasibility and Optimality	16
2.3 Convexity	20
2.3.1 Derivatives and Convexity	23
2.4 The General Optimization Algorithm	25
2.5 Rates of Convergence	29
2.6 Taylor Series	33
2.7 Newton's Method for Nonlinear Equations	37
2.7.1 Systems of Nonlinear Equations	43
3 Representation of Linear Constraints	47
3.1 Basic Concepts	47
3.2 Null and Range Spaces	52
3.3 Generating Null-Space Matrices	56
3.3.1 Variable Reduction Method	56
3.3.2 Orthogonal Projection Matrix	59

3.3.3 Other Projections	61
3.3.4 The QR Factorization	61

Part II Linear Programming

4 Geometry of Linear Programming	67
4.1 Introduction	67
4.2 Standard Form	70
4.3 Basic Solutions and Extreme Points	76
4.4 Representation of Solutions; Optimality	86
5 The Simplex Method	94
5.1 Introduction	94
5.2 The Simplex Method	95
5.2.1 General Formulas	98
5.2.2 Unbounded Problems	103
5.3 The Simplex Method (Tableaus)	107
5.3.1 The Full Tableau	107
5.3.2 Deficiencies of the Full Tableau	110
5.3.3 The Revised Simplex Tableau	111
5.3.4 The Revised Tableau Versus the Full Tableau	114
5.4 The Simplex Method (Details)	117
5.4.1 Multiple Solutions	117
5.4.2 Feasible Directions and Edge Directions	118
5.5 Getting Started—Artificial Variables	121
5.5.1 The Two-Phase Method	123
5.5.2 The Big-M Method	128
5.6 Degeneracy and Termination	134
5.6.1 Resolving Degeneracy Using Perturbation	139
6 Duality and Sensitivity	144
6.1 The Dual Problem	144
6.2 Duality Theory	150
6.2.1 Complementary Slackness	153
6.2.2 Interpretation of the Dual	155
6.3 The Dual Simplex Method	159
6.4 Sensitivity	165
6.5 Parametric Linear Programming	173
7 Enhancements of the Simplex Method	181
7.1 Introduction	181
7.2 The Simplex Method Using the Product Form of the Inverse	182
7.3 Problems with Upper Bounds	191
7.4 Column Generation	200
7.5 The Decomposition Principle	204
7.6 Numerical Stability and Computational Efficiency	216
7.6.1 Representation of the Basis	218
7.6.2 Pricing	222
7.6.3 The Initial Basis	227

7.6.4	Tolerances; Degeneracy	227
7.6.5	Scaling	228
7.6.6	Preprocessing	229
7.6.7	Model Formats	230
8	Network Problems	234
8.1	Introduction	234
8.2	Basic Concepts and Examples	234
8.3	Representation of the Basis	244
8.4	The Network Simplex Method	251
8.5	Resolving Degeneracy	259
9	Computational Complexity of Linear Programming	265
9.1	Introduction	265
9.2	Computational Complexity	265
9.3	Worst-Case Behavior of the Simplex Method	269
9.4	The Ellipsoid Method	272
9.5	The Average Case Behavior of the Simplex Method	278
9.6	The Primal-Dual Interior-Point Method	280

Part III Unconstrained Optimization

10	Basics of Unconstrained Optimization	295
10.1	Introduction	295
10.2	Optimality Conditions	295
10.3	Newton's Method for Minimization	302
10.4	Guaranteeing Descent	308
10.5	Guaranteeing Convergence: Line-Search Methods	312
10.5.1	Other Line Searches	318
10.6	Guaranteeing Convergence: Trust-Region Methods	327
11	Methods for Unconstrained Optimization	338
11.1	Introduction	338
11.2	Steepest Descent	339
11.3	Quasi-Newton Methods	347
11.4	Automating and Avoiding Derivative Calculations	358
11.4.1	Finite Difference Derivative Estimates	358
11.4.2	Automatic Differentiation	362
11.4.3	Methods That Do Not Require Derivatives	365
11.5	Termination Rules	373
11.6	Historical Background	377
12	Low-Storage Methods For Unconstrained Problems	382
12.1	Introduction	382
12.2	Iterative Methods for Solving Linear Equations	383
12.3	Truncated-Newton Methods	391
12.4	Nonlinear Conjugate-Gradient Methods	396
12.5	Limited-Memory Quasi-Newton Methods	400
12.6	Preconditioning	404

13	Nonlinear Least-Squares Data Fitting	409
13.1	Introduction	409
13.2	Nonlinear Least-Squares Data Fitting	411
13.3	Statistical Tests	418
13.4	Orthogonal Distance Regression	420
 Part IV Nonlinear Programming		
14	Optimality Conditions for Constrained Problems	427
14.1	Introduction	427
14.2	Optimality Conditions for Linear Equality Constraints	428
14.3	The Lagrange Multipliers and the Lagrangian Function	435
14.4	Optimality Conditions for Linear Inequality Constraints	438
14.5	Optimality Conditions for Nonlinear Constraints	447
14.5.1	Statement of Optimality Conditions	447
14.6	Preview of Methods	454
14.7	Derivation of Optimality Conditions For Nonlinear Constraints	457
14.8	Duality	464
14.8.1	Games and Min-Max Duality	465
14.8.2	Lagrangian Duality	468
14.8.3	More on the Dual Function	475
14.9	Historical Background	481
15	Feasible-Point Methods	486
15.1	Introduction	486
15.2	Linear Equality Constraints	486
15.3	Computing the Lagrange Multipliers	494
15.4	Linear Inequality Constraints	500
15.4.1	Linear Programming	508
15.5	Sequential Quadratic Programming	511
15.6	Reduced-Gradient Methods	519
16	Penalty and Barrier Methods	527
16.1	Introduction	527
16.2	Classical Penalty and Barrier Methods	528
16.2.1	Barrier Methods	529
16.2.2	Penalty Methods	537
16.2.3	Convergence	539
16.3	Ill Conditioning	544
16.4	Stabilized Penalty and Barrier Methods	545
16.5	Exact Penalty Methods	549
16.6	Multiplier-Based Methods	552
16.6.1	Dual Interpretation	561
17	Interior-Point Methods for Linear and Convex Programming	567
17.1	Introduction	567
17.2	Interior-Point Methods for Linear Programming	567
17.3	Affine Scaling Methods	569

17.4	Path-Following Methods	578
17.5	Karmarkar's Projective Scaling Method	587
17.6	Interior-Point Methods for Convex Programming	593
17.6.1	Basic Ideas of Self Concordance	594
17.6.2	The Newton Decrement	598
17.6.3	Convergence of the Damped Newton Method	599
17.7	The Path-Following Method	604
17.7.1	Convergence and Complexity	606

Appendices

A	Topics from Linear Algebra	617
A.1	Introduction	617
A.2	Vector and Matrix Norms	617
A.3	Systems of Linear Equations	619
A.4	Solving Systems of Linear Equations by Elimination	621
A.5	Gaussian Elimination as a Matrix Factorization	624
A.5.1	Sparse Matrix Storage	631
A.6	Other Matrix Factorizations	632
A.6.1	Positive-Definite Matrices	632
A.6.2	The LDL^T and Cholesky Factorizations	634
A.6.3	An Orthogonal Matrix Factorization	636
A.7	Sensitivity (Conditioning)	638
A.7.1	Eigenvalues and Sensitivity	642
A.8	The Sherman-Morrison Formula	643
B	Other Fundamentals	646
B.1	Introduction	646
B.2	Computer Arithmetic	646
B.3	Big-O Notation, $O(\cdot)$	648
B.4	The Gradient, Hessian, and Jacobian	649
B.5	Gradient and Hessian of a Quadratic Function	651
B.6	Derivatives of a Product	652
B.7	The Chain Rule	653
B.8	Continuous Functions; Closed and Bounded Sets	654
B.9	The Implicit Function Theorem	655
C	Software	658
C.1	Software	658
D	Bibliography	661
	Index	679