



$$\begin{aligned} & \text{minimize} && f(x) \\ & \text{subject to} && g_i(x) = 0, \quad i \in \mathcal{E} \\ & && g_i(x) \geq 0, \quad i \in \mathcal{I}. \end{aligned}$$

$$\begin{aligned} x_{k+1} &= x_k + \alpha_P \Delta x_k \\ \lambda_{k+1} &= \lambda_k + \alpha_D \Delta \lambda_k \end{aligned}$$

Linear and Nonlinear Optimization

SECOND EDITION

Igor Griva ■ Stephen G. Nash ■ Ariela Sofer

Contents

Preface	xiii
I Basics	1
1 Optimization Models	3
1.1 Introduction	3
1.2 Optimization: An Informal Introduction	4
1.3 Linear Equations	7
1.4 Linear Optimization	10
Exercises	12
1.5 Least-Squares Data Fitting	12
Exercises	14
1.6 Nonlinear Optimization	14
1.7 Optimization Applications	18
1.7.1 Crew Scheduling and Fleet Scheduling	18
Exercises	22
1.7.2 Support Vector Machines	22
Exercises	24
1.7.3 Portfolio Optimization	25
Exercises	27
1.7.4 Intensity Modulated Radiation Treatment Planning	28
Exercises	31
1.7.5 Positron Emission Tomography Image Reconstruction	32
Exercises	34
1.7.6 Shape Optimization	35
1.8 Notes	40
2 Fundamentals of Optimization	43
2.1 Introduction	43
2.2 Feasibility and Optimality	43
Exercises	47
2.3 Convexity	48
2.3.1 Derivatives and Convexity	50

Exercises	52
2.4 The General Optimization Algorithm	54
Exercises	58
2.5 Rates of Convergence	58
Exercises	61
2.6 Taylor Series	62
Exercises	65
2.7 Newton's Method for Nonlinear Equations	67
2.7.1 Systems of Nonlinear Equations	72
Exercises	74
2.8 Notes	76
3 Representation of Linear Constraints	77
3.1 Basic Concepts	77
Exercises	82
3.2 Null and Range Spaces	82
Exercises	84
3.3 Generating Null-Space Matrices	86
3.3.1 Variable Reduction Method	86
3.3.2 Orthogonal Projection Matrix	89
3.3.3 Other Projections	90
3.3.4 The <i>QR</i> Factorization	90
Exercises	91
3.4 Notes	93
II Linear Programming	95
4 Geometry of Linear Programming	97
4.1 Introduction	97
Exercises	98
4.2 Standard Form	100
Exercises	105
4.3 Basic Solutions and Extreme Points	106
Exercises	114
4.4 Representation of Solutions; Optimality	117
Exercises	123
4.5 Notes	124
5 The Simplex Method	125
5.1 Introduction	125
5.2 The Simplex Method	126
5.2.1 General Formulas	129
5.2.2 Unbounded Problems	134
5.2.3 Notation for the Simplex Method (Tableaus)	135
5.2.4 Deficiencies of the Tableau	139

Exercises	141
5.3 The Simplex Method (Details)	144
5.3.1 Multiple Solutions	144
5.3.2 Feasible Directions and Edge Directions	145
Exercises	148
5.4 Getting Started—Artificial Variables	149
5.4.1 The Two-Phase Method	150
5.4.2 The Big-M Method	156
Exercises	159
5.5 Degeneracy and Termination	162
5.5.1 Resolving Degeneracy Using Perturbation	167
Exercises	170
5.6 Notes	171
6 Duality and Sensitivity	173
6.1 The Dual Problem	173
Exercises	177
6.2 Duality Theory	179
6.2.1 Complementary Slackness	182
6.2.2 Interpretation of the Dual	184
Exercises	185
6.3 The Dual Simplex Method	189
Exercises	194
6.4 Sensitivity	195
Exercises	201
6.5 Parametric Linear Programming	204
Exercises	210
6.6 Notes	211
7 Enhancements of the Simplex Method	213
7.1 Introduction	213
7.2 Problems with Upper Bounds	214
Exercises	221
7.3 Column Generation	222
Exercises	227
7.4 The Decomposition Principle	227
Exercises	238
7.5 Representation of the Basis	240
7.5.1 The Product Form of the Inverse	240
7.5.2 Representation of the Basis—The <i>LU</i> Factorization	248
Exercises	256
7.6 Numerical Stability and Computational Efficiency	259
7.6.1 Pricing	260
7.6.2 The Initial Basis	264
7.6.3 Tolerances; Degeneracy	265
7.6.4 Scaling	266

	7.6.5	Preprocessing	267
	7.6.6	Model Formats	268
	Exercises		269
	7.7	Notes	270
8	Network Problems		271
	8.1	Introduction	271
	8.2	Basic Concepts and Examples	271
	Exercises		280
	8.3	Representation of the Basis	280
	Exercises		287
	8.4	The Network Simplex Method	287
	Exercises		294
	8.5	Resolving Degeneracy	295
	Exercises		299
	8.6	Notes	299
9	Computational Complexity of Linear Programming		301
	9.1	Introduction	301
	9.2	Computational Complexity	302
	Exercises		304
	9.3	Worst-Case Behavior of the Simplex Method	305
	Exercises		308
	9.4	The Ellipsoid Method	308
	Exercises		313
	9.5	The Average-Case Behavior of the Simplex Method	314
	9.6	Notes	316
10	Interior-Point Methods for Linear Programming		319
	10.1	Introduction	319
	10.2	The Primal-Dual Interior-Point Method	321
		10.2.1 Computational Aspects of Interior-Point Methods	328
		10.2.2 The Predictor-Corrector Algorithm	329
	Exercises		330
	10.3	Feasibility and Self-Dual Formulations	331
	Exercises		334
	10.4	Some Concepts from Nonlinear Optimization	334
	10.5	Affine-Scaling Methods	336
	Exercises		343
	10.6	Path-Following Methods	344
	Exercises		352
	10.7	Notes	353

III	Unconstrained Optimization	355
11	Basics of Unconstrained Optimization	357
	11.1 Introduction	357
	11.2 Optimality Conditions	357
	Exercises	361
	11.3 Newton's Method for Minimization	364
	Exercises	369
	11.4 Guaranteeing Descent	371
	Exercises	374
	11.5 Guaranteeing Convergence: Line Search Methods	375
	11.5.1 Other Line Searches	381
	Exercises	385
	11.6 Guaranteeing Convergence: Trust-Region Methods	391
	Exercises	398
	11.7 Notes	399
12	Methods for Unconstrained Optimization	401
	12.1 Introduction	401
	12.2 Steepest-Descent Method	402
	Exercises	408
	12.3 Quasi-Newton Methods	411
	Exercises	420
	12.4 Automating Derivative Calculations	422
	12.4.1 Finite-Difference Derivative Estimates	422
	12.4.2 Automatic Differentiation	426
	Exercises	429
	12.5 Methods That Do Not Require Derivatives	431
	12.5.1 Simulation-Based Optimization	432
	12.5.2 Compass Search: A Derivative-Free Method	434
	12.5.3 Convergence of Compass Search	437
	Exercises	440
	12.6 Termination Rules	441
	Exercises	445
	12.7 Historical Background	446
	12.8 Notes	448
13	Low-Storage Methods for Unconstrained Problems	451
	13.1 Introduction	451
	13.2 The Conjugate-Gradient Method for Solving Linear Equations	452
	Exercises	459
	13.3 Truncated-Newton Methods	460
	Exercises	465
	13.4 Nonlinear Conjugate-Gradient Methods	466
	Exercises	469
	13.5 Limited-Memory Quasi-Newton Methods	470

Exercises	473
13.6 Preconditioning	474
Exercises	477
13.7 Notes	478
IV Nonlinear Optimization	481
14 Optimality Conditions for Constrained Problems	483
14.1 Introduction	483
14.2 Optimality Conditions for Linear Equality Constraints	484
Exercises	489
14.3 The Lagrange Multipliers and the Lagrangian Function	491
Exercises	493
14.4 Optimality Conditions for Linear Inequality Constraints	494
Exercises	501
14.5 Optimality Conditions for Nonlinear Constraints	502
14.5.1 Statement of Optimality Conditions	503
Exercises	508
14.6 Preview of Methods	510
Exercises	514
14.7 Derivation of Optimality Conditions for Nonlinear Constraints	515
Exercises	520
14.8 Duality	522
14.8.1 Games and Min-Max Duality	523
14.8.2 Lagrangian Duality	526
14.8.3 Wolfe Duality	532
14.8.4 More on the Dual Function	534
14.8.5 Duality in Support Vector Machines	538
Exercises	542
14.9 Historical Background	543
14.10 Notes	546
15 Feasible-Point Methods	549
15.1 Introduction	549
15.2 Linear Equality Constraints	549
Exercises	555
15.3 Computing the Lagrange Multipliers	556
Exercises	561
15.4 Linear Inequality Constraints	563
15.4.1 Linear Programming	570
Exercises	572
15.5 Sequential Quadratic Programming	573
Exercises	580
15.6 Reduced-Gradient Methods	581
Exercises	588

15.7	Filter Methods	588
	Exercises	597
15.8	Notes	598
16	Penalty and Barrier Methods	601
16.1	Introduction	601
16.2	Classical Penalty and Barrier Methods	602
16.2.1	Barrier Methods	603
16.2.2	Penalty Methods	610
16.2.3	Convergence	613
	Exercises	617
16.3	Ill-Conditioning	618
16.4	Stabilized Penalty and Barrier Methods	619
	Exercises	623
16.5	Exact Penalty Methods	623
	Exercises	626
16.6	Multiplier-Based Methods	626
16.6.1	Dual Interpretation	635
	Exercises	638
16.7	Nonlinear Primal-Dual Methods	640
16.7.1	Primal-Dual Interior-Point Methods	641
16.7.2	Convergence of the Primal-Dual Interior-Point Method	645
	Exercises	647
16.8	Semidefinite Programming	649
	Exercises	654
16.9	Notes	656
V	Appendices	659
A	Topics from Linear Algebra	661
A.1	Introduction	661
A.2	Eigenvalues	661
A.3	Vector and Matrix Norms	662
A.4	Systems of Linear Equations	664
A.5	<i>Solving Systems of Linear Equations by Elimination</i>	666
A.6	Gaussian Elimination as a Matrix Factorization	669
A.6.1	Sparse Matrix Storage	675
A.7	Other Matrix Factorizations	676
A.7.1	Positive-Definite Matrices	677
A.7.2	The LDL^T and Cholesky Factorizations	678
A.7.3	An Orthogonal Matrix Factorization	681
A.8	Sensitivity (Conditioning)	683
A.9	The Sherman–Morrison Formula	686
A.10	Notes	688

B	Other Fundamentals	691
B.1	Introduction	691
B.2	Computer Arithmetic	691
B.3	Big-O Notation, $\mathcal{O}(\cdot)$	693
B.4	The Gradient, Hessian, and Jacobian	694
B.5	Gradient and Hessian of a Quadratic Function	696
B.6	Derivatives of a Product	697
B.7	The Chain Rule	698
B.8	Continuous Functions; Closed and Bounded Sets	699
B.9	The Implicit Function Theorem	700
C	Software	703
C.1	Software	703
	Bibliography	707
	Index	727