


Jorge Nocedal Stephen J. Wright

Numerical Optimization

Second Edition

 Springer

Contents

Preface	xvii
Preface to the Second Edition	xxi
1 Introduction	1
Mathematical Formulation	2
Example: A Transportation Problem	4
Continuous versus Discrete Optimization	5
Constrained and Unconstrained Optimization	6
Global and Local Optimization	6
Stochastic and Deterministic Optimization	7
Convexity	7
Optimization Algorithms	8
Notes and References	9
2 Fundamentals of Unconstrained Optimization	10
2.1 What Is a Solution?	12

	Recognizing a Local Minimum	14
	Nonsmooth Problems	17
2.2	Overview of Algorithms	18
	Two Strategies: Line Search and Trust Region	19
	Search Directions for Line Search Methods	20
	Models for Trust-Region Methods	25
	Scaling	26
	Exercises	27
3	Line Search Methods	30
3.1	Step Length	31
	The Wolfe Conditions	33
	The Goldstein Conditions	36
	Sufficient Decrease and Backtracking	37
3.2	Convergence of Line Search Methods	37
3.3	Rate of Convergence	41
	Convergence Rate of Steepest Descent	42
	Newton's Method	44
	Quasi-Newton Methods	46
3.4	Newton's Method with Hessian Modification	48
	Eigenvalue Modification	49
	Adding a Multiple of the Identity	51
	Modified Cholesky Factorization	52
	Modified Symmetric Indefinite Factorization	54
3.5	Step-Length Selection Algorithms	56
	Interpolation	57
	Initial Step Length	59
	A Line Search Algorithm for the Wolfe Conditions	60
	Notes and References	62
	Exercises	63
4	Trust-Region Methods	66
	Outline of the Trust-Region Approach	68
4.1	Algorithms Based on the Cauchy Point	71
	The Cauchy Point	71
	Improving on the Cauchy Point	73
	The Dogleg Method	73
	Two-Dimensional Subspace Minimization	76
4.2	Global Convergence	77
	Reduction Obtained by the Cauchy Point	77
	Convergence to Stationary Points	79
4.3	Iterative Solution of the Subproblem	83

	The Hard Case	87
	Proof of Theorem 4.1	89
	Convergence of Algorithms Based on Nearly Exact Solutions	91
4.4	Local Convergence of Trust-Region Newton Methods	92
4.5	Other Enhancements	95
	Scaling	95
	Trust Regions in Other Norms	97
	Notes and References	98
	Exercises	98
5	Conjugate Gradient Methods	101
5.1	The Linear Conjugate Gradient Method	102
	Conjugate Direction Methods	102
	Basic Properties of the Conjugate Gradient Method	107
	A Practical Form of the Conjugate Gradient Method	111
	Rate of Convergence	112
	Preconditioning	118
	Practical Preconditioners	120
5.2	Nonlinear Conjugate Gradient Methods	121
	The Fletcher–Reeves Method	121
	The Polak–Ribière Method and Variants	122
	Quadratic Termination and Restarts	124
	Behavior of the Fletcher–Reeves Method	125
	Global Convergence	127
	Numerical Performance	131
	Notes and References	132
	Exercises	133
6	Quasi-Newton Methods	135
6.1	The BFGS Method	136
	Properties of the BFGS Method	141
	Implementation	142
6.2	The SR1 Method	144
	Properties of SR1 Updating	147
6.3	The Broyden Class	149
6.4	Convergence Analysis	153
	Global Convergence of the BFGS Method	153
	Superlinear Convergence of the BFGS Method	156
	Convergence Analysis of the SR1 Method	160
	Notes and References	161
	Exercises	162

7	Large-Scale Unconstrained Optimization	164
7.1	Inexact Newton Methods	165
	Local Convergence of Inexact Newton Methods	166
	Line Search Newton–CG Method	168
	Trust-Region Newton–CG Method	170
	Preconditioning the Trust-Region Newton–CG Method	174
	Trust-Region Newton–Lanczos Method	175
7.2	Limited-Memory Quasi-Newton Methods	176
	Limited-Memory BFGS	177
	Relationship with Conjugate Gradient Methods	180
	General Limited-Memory Updating	181
	Compact Representation of BFGS Updating	181
	Unrolling the Update	184
7.3	Sparse Quasi-Newton Updates	185
7.4	Algorithms for Partially Separable Functions	186
7.5	Perspectives and Software	189
	Notes and References	190
	Exercises	191
8	Calculating Derivatives	193
8.1	Finite-Difference Derivative Approximations	194
	Approximating the Gradient	195
	Approximating a Sparse Jacobian	197
	Approximating the Hessian	201
	Approximating a Sparse Hessian	202
8.2	Automatic Differentiation	204
	An Example	205
	The Forward Mode	206
	The Reverse Mode	207
	Vector Functions and Partial Separability	210
	Calculating Jacobians of Vector Functions	212
	Calculating Hessians: Forward Mode	213
	Calculating Hessians: Reverse Mode	215
	Current Limitations	216
	Notes and References	217
	Exercises	217
9	Derivative-Free Optimization	220
9.1	Finite Differences and Noise	221
9.2	Model-Based Methods	223
	Interpolation and Polynomial Bases	226
	Updating the Interpolation Set	227

	A Method Based on Minimum-Change Updating	228
9.3	Coordinate and Pattern-Search Methods	229
	Coordinate Search Method	230
	Pattern-Search Methods	231
9.4	A Conjugate-Direction Method	234
9.5	Nelder-Mead Method	238
9.6	Implicit Filtering	240
	Notes and References	242
	Exercises	242
10	Least-Squares Problems	245
10.1	Background	247
10.2	Linear Least-Squares Problems	250
10.3	Algorithms for Nonlinear Least-Squares Problems	254
	The Gauss-Newton Method	254
	Convergence of the Gauss-Newton Method	255
	The Levenberg-Marquardt Method	258
	Implementation of the Levenberg-Marquardt Method	259
	Convergence of the Levenberg-Marquardt Method	261
	Methods for Large-Residual Problems	262
10.4	Orthogonal Distance Regression	265
	Notes and References	267
	Exercises	269
11	Nonlinear Equations	270
11.1	Local Algorithms	274
	Newton's Method for Nonlinear Equations	274
	Inexact Newton Methods	277
	Broyden's Method	279
	Tensor Methods	283
11.2	Practical Methods	285
	Merit Functions	285
	Line Search Methods	287
	Trust-Region Methods	290
11.3	Continuation/Homotopy Methods	296
	Motivation	296
	Practical Continuation Methods	297
	Notes and References	302
	Exercises	302
12	Theory of Constrained Optimization	304
	Local and Global Solutions	305

	Smoothness	306
12.1	Examples	307
	A Single Equality Constraint	308
	A Single Inequality Constraint	310
	Two Inequality Constraints	313
12.2	Tangent Cone and Constraint Qualifications	315
12.3	First-Order Optimality Conditions	320
12.4	First-Order Optimality Conditions: Proof	323
	Relating the Tangent Cone and the First-Order Feasible Direction Set	323
	A Fundamental Necessary Condition	325
	Farkas' Lemma	326
	Proof of Theorem 12.1	329
12.5	Second-Order Conditions	330
	Second-Order Conditions and Projected Hessians	337
12.6	Other Constraint Qualifications	338
12.7	A Geometric Viewpoint	340
12.8	Lagrange Multipliers and Sensitivity	341
12.9	Duality	343
	Notes and References	349
	Exercises	351
13	Linear Programming: The Simplex Method	355
	Linear Programming	356
13.1	Optimality and Duality	358
	Optimality Conditions	358
	The Dual Problem	359
13.2	Geometry of the Feasible Set	362
	Bases and Basic Feasible Points	362
	Vertices of the Feasible Polytope	365
13.3	The Simplex Method	366
	Outline	366
	A Single Step of the Method	370
13.4	Linear Algebra in the Simplex Method	372
13.5	Other Important Details	375
	Pricing and Selection of the Entering Index	375
	Starting the Simplex Method	378
	Degenerate Steps and Cycling	381
13.6	The Dual Simplex Method	382
13.7	Presolving	385
13.8	Where Does the Simplex Method Fit?	388
	Notes and References	389
	Exercises	389

14 Linear Programming: Interior-Point Methods	392
14.1 Primal-Dual Methods	393
Outline	393
The Central Path	397
Central Path Neighborhoods and Path-Following Methods	399
14.2 Practical Primal-Dual Algorithms	407
Corrector and Centering Steps	407
Step Lengths	409
Starting Point	410
A Practical Algorithm	411
Solving the Linear Systems	411
14.3 Other Primal-Dual Algorithms and Extensions	413
Other Path-Following Methods	413
Potential-Reduction Methods	414
Extensions	415
14.4 Perspectives and Software	416
Notes and References	417
Exercises	418
15 Fundamentals of Algorithms for Nonlinear Constrained Optimization	421
15.1 Categorizing Optimization Algorithms	422
15.2 The Combinatorial Difficulty of Inequality-Constrained Problems	424
15.3 Elimination of Variables	426
Simple Elimination using Linear Constraints	428
General Reduction Strategies for Linear Constraints	431
Effect of Inequality Constraints	434
15.4 Merit Functions and Filters	435
Merit Functions	435
Filters	437
15.5 The Maratos Effect	440
15.6 Second-Order Correction and Nonmonotone Techniques	443
Nonmonotone (Watchdog) Strategy	444
Notes and References	446
Exercises	446
16 Quadratic Programming	448
16.1 Equality-Constrained Quadratic Programs	451
Properties of Equality-Constrained QPs	451
16.2 Direct Solution of the KKT System	454
Factoring the Full KKT System	454
Schur-Complement Method	455
Null-Space Method	457

16.3	Iterative Solution of the KKT System	459
	CG Applied to the Reduced System	459
	The Projected CG Method	461
16.4	Inequality-Constrained Problems	463
	Optimality Conditions for Inequality-Constrained Problems	464
	Degeneracy	465
16.5	Active-Set Methods for Convex QPs	467
	Specification of the Active-Set Method for Convex QP	472
	Further Remarks on the Active-Set Method	476
	Finite Termination of Active-Set Algorithm on Strictly Convex QPs	477
	Updating Factorizations	478
16.6	Interior-Point Methods	480
	Solving the Primal-Dual System	482
	Step Length Selection	483
	A Practical Primal-Dual Method	484
16.7	The Gradient Projection Method	485
	Cauchy Point Computation	486
	Subspace Minimization	488
16.8	Perspectives and Software	490
	Notes and References	492
	Exercises	492
17	Penalty and Augmented Lagrangian Methods	497
17.1	The Quadratic Penalty Method	498
	Motivation	498
	Algorithmic Framework	501
	Convergence of the Quadratic Penalty Method	502
	Ill Conditioning and Reformulations	505
17.2	Nonsmooth Penalty Functions	507
	A Practical ℓ_1 Penalty Method	511
	A General Class of Nonsmooth Penalty Methods	513
17.3	Augmented Lagrangian Method: Equality Constraints	514
	Motivation and Algorithmic Framework	514
	Properties of the Augmented Lagrangian	517
17.4	Practical Augmented Lagrangian Methods	519
	Bound-Constrained Formulation	519
	Linearly Constrained Formulation	522
	Unconstrained Formulation	523
17.5	Perspectives and Software	525
	Notes and References	526
	Exercises	527

18 Sequential Quadratic Programming	529
18.1 Local SQP Method	530
SQP Framework	531
Inequality Constraints	532
18.2 Preview of Practical SQP Methods	533
IQP and EQP	533
Enforcing Convergence	534
18.3 Algorithmic Development	535
Handling Inconsistent Linearizations	535
Full Quasi-Newton Approximations	536
Reduced-Hessian Quasi-Newton Approximations	538
Merit Functions	540
Second-Order Correction	543
18.4 A Practical Line Search SQP Method	545
18.5 Trust-Region SQP Methods	546
A Relaxation Method for Equality-Constrained Optimization	547
Sl_1 QP (Sequential l_1 Quadratic Programming)	549
Sequential Linear-Quadratic Programming (SLQP)	551
A Technique for Updating the Penalty Parameter	553
18.6 Nonlinear Gradient Projection	554
18.7 Convergence Analysis	556
Rate of Convergence	557
18.8 Perspectives and Software	560
Notes and References	561
Exercises	561
19 Interior-Point Methods for Nonlinear Programming	563
19.1 Two Interpretations	564
19.2 A Basic Interior-Point Algorithm	566
19.3 Algorithmic Development	569
Primal vs. Primal-Dual System	570
Solving the Primal-Dual System	570
Updating the Barrier Parameter	572
Handling Nonconvexity and Singularity	573
Step Acceptance: Merit Functions and Filters	575
Quasi-Newton Approximations	575
Feasible Interior-Point Methods	576
19.4 A Line Search Interior-Point Method	577
19.5 A Trust-Region Interior-Point Method	578
An Algorithm for Solving the Barrier Problem	578
Step Computation	580
Lagrange Multipliers Estimates and Step Acceptance	581

	Description of a Trust-Region Interior-Point Method	582
19.6	The Primal Log-Barrier Method	583
19.7	Global Convergence Properties	587
	Failure of the Line Search Approach	587
	Modified Line Search Methods	589
	Global Convergence of the Trust-Region Approach	589
19.8	Superlinear Convergence	591
19.9	Perspectives and Software	592
	Notes and References	593
	Exercises	594
A	Background Material	598
A.1	Elements of Linear Algebra	598
	Vectors and Matrices	598
	Norms	600
	Subspaces	602
	Eigenvalues, Eigenvectors, and the Singular-Value Decomposition	603
	Determinant and Trace	605
	Matrix Factorizations: Cholesky, LU, QR	606
	Symmetric Indefinite Factorization	610
	Sherman–Morrison–Woodbury Formula	612
	Interlacing Eigenvalue Theorem	613
	Error Analysis and Floating-Point Arithmetic	613
	Conditioning and Stability	616
A.2	Elements of Analysis, Geometry, Topology	617
	Sequences	617
	Rates of Convergence	619
	Topology of the Euclidean Space \mathbf{R}^n	620
	Convex Sets in \mathbf{R}^n	621
	Continuity and Limits	623
	Derivatives	625
	Directional Derivatives	628
	Mean Value Theorem	629
	Implicit Function Theorem	630
	Order Notation	631
	Root-Finding for Scalar Equations	633
B	A Regularization Procedure	635
	References	637
	Index	653