

Contents

Preface to the Second Edition	vii
Preface to the First Edition	ix
1 Elementary Optimization	1
1.1 Introduction	1
1.2 Univariate Optimization	1
1.3 Multivariate Optimization	7
1.4 Constrained Optimization	10
1.5 Problems	17
2 The Seven C's of Analysis	23
2.1 Introduction	23
2.2 Vector and Matrix Norms	23
2.3 Convergence and Completeness	26
2.4 The Topology of $\mathbb{R}^n$	30
2.5 Continuous Functions	34
2.6 Semicontinuity	42
2.7 Connectedness	44
2.8 Uniform Convergence	46
2.9 Problems	47
3 The Gauge Integral	53
3.1 Introduction	53
3.2 Gauge Functions and δ -Fine Partitions	54

3.3	Definition and Basic Properties of the Integral	57
3.4	The Fundamental Theorem of Calculus	62
3.5	More Advanced Topics in Integration	66
3.6	Problems	71
4	Differentiation	75
4.1	Introduction	75
4.2	Univariate Derivatives	75
4.3	Partial Derivatives	79
4.4	Differentials	81
4.5	Multivariate Mean Value Theorem	88
4.6	Inverse and Implicit Function Theorems	89
4.7	Differentials of Matrix-Valued Functions	93
4.8	Problems	98
5	Karush-Kuhn-Tucker Theory	107
5.1	Introduction	107
5.2	The Multiplier Rule	108
5.3	Constraint Qualification	114
5.4	Taylor-Made Higher-Order Differentials	117
5.5	Applications of Second Differentials	123
5.6	Problems	128
6	Convexity	137
6.1	Introduction	137
6.2	Convex Sets	138
6.3	Convex Functions	142
6.4	Continuity, Differentiability, and Integrability	149
6.5	Minimization of Convex Functions	152
6.6	Moment Inequalities	159
6.7	Problems	162
7	Block Relaxation	171
7.1	Introduction	171
7.2	Examples of Block Relaxation	172
7.3	Problems	180
8	The MM Algorithm	185
8.1	Introduction	185
8.2	Philosophy of the MM Algorithm	186
8.3	Majorization and Minorization	187
8.4	Allele Frequency Estimation	189
8.5	Linear Regression	191
8.6	Bradley-Terry Model of Ranking	193
8.7	Linear Logistic Regression	194

8.8	Geometric and Signomial Programs	194
8.9	Poisson Processes	197
8.10	Transmission Tomography	198
8.11	Poisson Multigraphs	202
8.12	Problems	204
9	The EM Algorithm	221
9.1	Introduction	221
9.2	Definition of the EM Algorithm	222
9.3	Missing Data in the Ordinary Sense	224
9.4	Allele Frequency Estimation	225
9.5	Clustering by EM	226
9.6	Transmission Tomography	228
9.7	Factor Analysis	230
9.8	Hidden Markov Chains	234
9.9	Problems	236
10	Newton's Method and Scoring	245
10.1	Introduction	245
10.2	Newton's Method and Root Finding	246
10.3	Newton's Method and Optimization	248
10.4	MM Gradient Algorithm	250
10.5	Ad Hoc Approximations of $d^2 f(\theta)$	252
10.6	Scoring and Exponential Families	254
10.7	The Gauss-Newton Algorithm	257
10.8	Generalized Linear Models	258
10.9	Accelerated MM	259
10.10	Problems	262
11	Conjugate Gradient and Quasi-Newton	273
11.1	Introduction	273
11.2	Centers of Spheres and Centers of Ellipsoids	274
11.3	The Conjugate Gradient Algorithm	275
11.4	Line Search Methods	278
11.5	Stopping Criteria	280
11.6	Quasi-Newton Methods	281
11.7	Trust Regions	285
11.8	Problems	286
12	Analysis of Convergence	291
12.1	Introduction	291
12.2	Local Convergence	292
12.3	Coercive Functions	297
12.4	Global Convergence of the MM Algorithm	299
12.5	Global Convergence of Block Relaxation	302

12.6	Global Convergence of Gradient Algorithms	303
12.7	Problems	306
13	Penalty and Barrier Methods	313
13.1	Introduction	313
13.2	Rudiments of Barrier and Penalty Methods	314
13.3	An Adaptive Barrier Method	318
13.4	Imposition of a Prior in EM Clustering	325
13.5	Model Selection and the Lasso	327
13.6	Lasso Penalized ℓ_1 Regression	329
13.7	Lasso Penalized ℓ_2 Regression	330
13.8	Penalized Discriminant Analysis	333
13.9	Problems	334
14	Convex Calculus	341
14.1	Introduction	341
14.2	Notation	342
14.3	Fenchel Conjugates	342
14.4	Subdifferentials	351
14.5	The Rules of Convex Differentiation	358
14.6	Spectral Functions	365
14.7	A Convex Lagrange Multiplier Rule	372
14.8	Problems	375
15	Feasibility and Duality	383
15.1	Introduction	383
15.2	Dykstra's Algorithm	384
15.3	Contractive Maps	389
15.4	Dual Functions	393
15.5	Examples of Dual Programs	396
15.6	Practical Applications of Duality	402
15.7	Problems	406
16	Convex Minimization Algorithms	415
16.1	Introduction	415
16.2	Projected Gradient Algorithm	416
16.3	Exact Penalties and Lagrangians	421
16.4	Mechanics of Path Following	426
16.5	Bregman Iteration	432
16.6	Split Bregman Iteration	436
16.7	Convergence of Bregman Iteration	439
16.8	Problems	440
17	The Calculus of Variations	445
17.1	Introduction	445
17.2	Classical Problems	446

17.3	Normed Vector Spaces	448
17.4	Linear Operators and Functionals	451
17.5	Differentials	453
17.6	The Euler-Lagrange Equation	456
17.7	Applications of the Euler-Lagrange Equation	459
17.8	Lagrange's Lacuna	462
17.9	Variational Problems with Constraints	464
17.10	Natural Cubic Splines	466
17.11	Problems	467
Appendix: Mathematical Notes		473
A.1	Univariate Normal Random Variables	473
A.2	Multivariate Normal Random Vectors	475
A.3	Polyhedral Sets	477
A.4	Birkhoff's Theorem and Fan's Inequality	480
A.5	Singular Value Decomposition	485
A.6	Hadamard Semidifferentials	487
A.7	Problems	497
References		499
Index		519