

Contents

<i>Preface</i>	v
1. Introduction	1
Preliminaries for Optimization Theory	13
2. Topology	15
2.1 Metric Spaces and Normed Vector Spaces	15
2.2 Topological Spaces	21
2.3 Subspace and Product Topologies	27
2.4 Continuous Functions	32
2.5 Limits and Continuity; Uniform Continuity	38
2.6 Continuous Linear and Multilinear Maps	43
2.7 Complete Metric Spaces and Banach Spaces	50
2.8 Completion of a Metric Space	50
2.9 Completion of a Normed Vector Space	59
2.10 The Contraction Mapping Theorem	60
2.11 Further Readings	62
2.12 Summary	62
2.13 Problems	63
3. Differential Calculus	65
3.1 Directional Derivatives, Total Derivatives	65
3.2 Properties of Derivatives	74
3.3 Jacobian Matrices	81
3.4 The Implicit and the Inverse Function Theorems	89

3.5	Second-Order and Higher-Order Derivatives	96
3.6	Taylor's Formula, Faà di Bruno's Formula	104
3.7	Further Readings	109
3.8	Summary	110
3.9	Problems	110
4.	Extrema of Real-Valued Functions	113
4.1	Local Extrema, Constrained Local Extrema, and Lagrange Multipliers	114
4.2	Using Second Derivatives to Find Extrema	127
4.3	Using Convexity to Find Extrema	131
4.4	Summary	141
4.5	Problems	142
5.	Newton's Method and Its Generalizations	145
5.1	Newton's Method for Real Functions of a Real Argument	145
5.2	Generalizations of Newton's Method	147
5.3	Summary	156
5.4	Problems	157
6.	Quadratic Optimization Problems	167
6.1	Quadratic Optimization: The Positive Definite Case	167
6.2	Quadratic Optimization: The General Case	178
6.3	Maximizing a Quadratic Function on the Unit Sphere	183
6.4	Summary	188
6.5	Problems	189
7.	Schur Complements and Applications	191
7.1	Schur Complements	191
7.2	Symmetric Positive Definite Matrices and Schur Complements	194
7.3	Symmetric Positive Semidefinite Matrices and Schur Complements	196
7.4	Summary	197
7.5	Problems	197

Linear Optimization	199
8. Convex Sets, Cones, \mathcal{H} -Polyhedra	201
8.1 What is Linear Programming?	201
8.2 Affine Subsets, Convex Sets, Affine Hyperplanes, Half-Spaces	204
8.3 Cones, Polyhedral Cones, and \mathcal{H} -Polyhedra	207
8.4 Summary	213
8.5 Problems	213
9. Linear Programs	215
9.1 Linear Programs, Feasible Solutions, Optimal Solutions	215
9.2 Basic Feasible Solutions and Vertices	222
9.3 Summary	231
9.4 Problems	231
10. The Simplex Algorithm	235
10.1 The Idea Behind the Simplex Algorithm	235
10.2 The Simplex Algorithm in General	245
10.3 How to Perform a Pivoting Step Efficiently	253
10.4 The Simplex Algorithm Using Tableaux	256
10.5 Computational Efficiency of the Simplex Method	267
10.6 Summary	268
10.7 Problems	268
11. Linear Programming and Duality	271
11.1 Variants of the Farkas Lemma	271
11.2 The Duality Theorem in Linear Programming	277
11.3 Complementary Slackness Conditions	287
11.4 Duality for Linear Programs in Standard Form	288
11.5 The Dual Simplex Algorithm	291
11.6 The Primal-Dual Algorithm	298
11.7 Summary	309
11.8 Problems	310

18.2	Solving SVM (SVM_{s_1}) Using ADMM	641
18.3	Soft Margin Support Vector Machines; (SVM_{s_2})	642
18.4	Solving SVM (SVM_{s_2}) Using ADMM	650
18.5	Soft Margin Support Vector Machines; ($SVM_{s_2'}$)	651
18.6	Classification of the Data Points in Terms of ν ($SVM_{s_2'}$)	662
18.7	Existence of Support Vectors for ($SVM_{s_2'}$)	665
18.8	Solving SVM ($SVM_{s_2'}$) Using ADMM	676
18.9	Soft Margin Support Vector Machines; (SVM_{s_3})	681
18.10	Classification of the Data Points in Terms of ν (SVM_{s_3})	689
18.11	Existence of Support Vectors for (SVM_{s_3})	690
18.12	Solving SVM (SVM_{s_3}) Using ADMM	692
18.13	Soft Margin SVM; (SVM_{s_4})	695
18.14	Solving SVM (SVM_{s_4}) Using ADMM	705
18.15	Soft Margin SVM; (SVM_{s_5})	707
18.16	Solving SVM (SVM_{s_5}) Using ADMM	712
18.17	Summary and Comparison of the SVM Methods	714
18.18	Problems	727
19.	Ridge Regression, Lasso, Elastic Net	733
19.1	Ridge Regression	734
19.2	Ridge Regression; Learning an Affine Function	738
19.3	Kernel Ridge Regression	747
19.4	Lasso Regression (ℓ^1 -Regularized Regression)	751
19.5	Lasso Regression; Learning an Affine Function	755
19.6	Elastic Net Regression	762
19.7	Summary	768
19.8	Problems	768
20.	ν -SV Regression	771
20.1	ν -SV Regression; Derivation of the Dual	771
20.2	Existence of Support Vectors	783
20.3	Solving ν -Regression Using ADMM	793
20.4	Kernel ν -SV Regression	800
20.5	ν -Regression Version 2; Penalizing b	803
20.6	Summary	811
20.7	Problems	811

Appendix A	Total Orthogonal Families in Hilbert Spaces	815
A.1	Total Orthogonal Families (Hilbert Bases), Fourier Coefficients	815
A.2	The Hilbert Space $\ell^2(K)$ and the Riesz–Fischer Theorem	825
A.3	Summary	835
A.4	Problems	835
Appendix B	Matlab Programs	837
B.1	Hard Margin (SVM $_{h2}$)	837
B.2	Soft Margin SVM (SVM $_{s2'}$)	842
B.3	Soft Margin SVM (SVM $_{s3}$)	850
B.4	ν -SV Regression	857
<i>Bibliography</i>		863
<i>Index</i>		867