

Contents

Preface		xi
1	Vector Spaces	1
1.1	Definition	1
1.2	Dimension	2
1.3	Norms	2
1.4	Inner Products	2
1.5	Affine Sets and Convex Sets	3
1.6	Euclidean Spaces	3
1.7	The Space \mathbb{R}^n	4
1.8	The Space $\mathbb{R}^{m \times n}$	6
1.9	Cartesian Product of Vector Spaces	7
1.10	Linear Transformations	8
1.11	The Dual Space	9
1.12	The Bidual Space	10
1.13	Adjoint Transformations	11
1.14	Norms of Linear Transformations	12
2	Extended Real-Valued Functions	13
2.1	Extended Real-Valued Functions and Closedness	13
2.2	Closedness versus Continuity	18
2.3	Convex Functions	21
2.4	Support Functions	26
3	Subgradients	35
3.1	Definitions and First Examples	35
3.2	Properties of the Subdifferential Set	39
3.3	Directional Derivatives	44
3.4	Computing Subgradients	53
3.5	The Value Function	67
3.6	Lipschitz Continuity and Boundedness of Subgradients	71
3.7	Optimality Conditions	72
3.8	Summary of Weak and Strong Subgradient Calculus Results	84
4	Conjugate Functions	87
4.1	Definition and Basic Properties	87
4.2	The Biconjugate	89

4.3	Conjugate Calculus Rules	91
4.4	Examples	93
4.5	Infimal Convolution and Conjugacy	102
4.6	Subdifferentials of Conjugate Functions	104
5	Smoothness and Strong Convexity	107
5.1	L -Smooth Functions	107
5.2	Strong Convexity	117
5.3	Smoothness and Strong Convexity Correspondence	123
6	The Proximal Operator	129
6.1	Definition, Existence, and Uniqueness	129
6.2	First Set of Examples of Proximal Mappings	131
6.3	Prox Calculus Rules	135
6.4	Prox of Indicators—Orthogonal Projections	146
6.5	The Second Prox Theorem	157
6.6	Moreau Decomposition	160
6.7	The Moreau Envelope	163
6.8	Miscellaneous Prox Computations	170
6.9	Summary of Prox Computations	177
7	Spectral Functions	179
7.1	Symmetric Functions	179
7.2	Symmetric Spectral Functions over \mathbb{S}^n	182
7.3	Symmetric Spectral Functions over $\mathbb{R}^{m \times n}$	188
8	Primal and Dual Projected Subgradient Methods	195
8.1	From Gradient Descent to Subgradient Descent	195
8.2	The Projected Subgradient Method	201
8.3	The Stochastic Projected Subgradient Method	221
8.4	The Incremental Projected Subgradient Method	229
8.5	The Dual Projected Subgradient Method	232
9	Mirror Descent	247
9.1	From Projected Subgradient to Mirror Descent	247
9.2	Convergence Analysis	252
9.3	Mirror Descent for the Composite Model	260
10	The Proximal Gradient Method	269
10.1	The Composite Model	269
10.2	The Proximal Gradient Method	270
10.3	Analysis of the Proximal Gradient Method—The Nonconvex Case	272
10.4	Analysis of the Proximal Gradient Method—The Convex Case	281
10.5	The Proximal Point Method	288
10.6	Convergence of the Proximal Gradient Method—The Strongly Convex Case	288
10.7	The Fast Proximal Gradient Method—FISTA	290
10.8	Smoothing	304

10.9	Non-Euclidean Proximal Gradient Methods	315
11	The Block Proximal Gradient Method	331
11.1	Decomposition Methods	331
11.2	Model and Assumptions	332
11.3	The Toolbox	333
11.4	The Cyclic Block Proximal Gradient Method	338
11.5	The Randomized Block Proximal Gradient Method	347
12	Dual-Based Proximal Gradient Methods	353
12.1	The Primal and Dual Models	353
12.2	The Dual Proximal Gradient Method	355
12.3	Fast Dual Proximal Gradient	358
12.4	Examples I	360
12.5	The Dual Block Proximal Gradient Method	369
12.6	Examples II	374
13	The Generalized Conditional Gradient Method	379
13.1	The Frank–Wolfe/Conditional Gradient Method	379
13.2	The Generalized Conditional Gradient Method	380
13.3	The Strongly Convex Case	391
13.4	The Randomized Generalized Block Conditional Gradient Method	400
14	Alternating Minimization	405
14.1	The Method	405
14.2	Coordinate-wise Minima	407
14.3	The Composite Model	411
14.4	Convergence in the Convex Case	413
14.5	Rate of Convergence in the Convex Case	415
15	ADMM	423
15.1	The Augmented Lagrangian Method	423
15.2	Alternating Direction Method of Multipliers (ADMM)	425
15.3	Convergence Analysis of AD-PMM	427
15.4	Minimizing $f_1(\mathbf{x}) + f_2(\mathbf{Ax})$	432
A	Strong Duality and Optimality Conditions	439
B	Tables	443
C	Symbols and Notation	451
D	Bibliographic Notes	457
	Bibliography	463
	Index	473