

# Contents

<b>Preface</b>	<b>xiii</b>
<b>How to Get the Software</b>	<b>xv</b>
<b>I Optimization of Smooth Functions</b>	<b>1</b>
<b>1 Basic Concepts</b>	<b>3</b>
1.1 The Problem . . . . .	3
1.2 Notation . . . . .	4
1.3 Necessary Conditions . . . . .	5
1.4 Sufficient Conditions . . . . .	6
1.5 Quadratic Objective Functions . . . . .	6
1.5.1 Positive Definite Hessian . . . . .	7
1.5.2 Indefinite Hessian . . . . .	9
1.6 Examples . . . . .	9
1.6.1 Discrete Optimal Control . . . . .	9
1.6.2 Parameter Identification . . . . .	11
1.6.3 Convex Quadratics . . . . .	12
1.7 Exercises on Basic Concepts . . . . .	12
<b>2 Local Convergence of Newton's Method</b>	<b>13</b>
2.1 Types of Convergence . . . . .	13
2.2 The Standard Assumptions . . . . .	14
2.3 Newton's Method . . . . .	14
2.3.1 Errors in Functions, Gradients, and Hessians . . . . .	17
2.3.2 Termination of the Iteration . . . . .	21
2.4 Nonlinear Least Squares . . . . .	22
2.4.1 Gauss–Newton Iteration . . . . .	23
2.4.2 Overdetermined Problems . . . . .	24
2.4.3 Underdetermined Problems . . . . .	26
2.5 Inexact Newton Methods . . . . .	28
2.5.1 Convergence Rates . . . . .	29
2.5.2 Implementation of Newton–CG . . . . .	30
2.6 Examples . . . . .	33
2.6.1 Parameter Identification . . . . .	33
2.6.2 Discrete Control Problem . . . . .	35
2.7 Exercises on Local Convergence . . . . .	36

<b>3</b>	<b>Global Convergence</b>	<b>39</b>
3.1	The Method of Steepest Descent . . . . .	39
3.2	Line Search Methods and the Armijo Rule . . . . .	40
3.2.1	Stepsize Control with Polynomial Models . . . . .	43
3.2.2	Slow Convergence of Steepest Descent . . . . .	45
3.2.3	Damped Gauss–Newton Iteration . . . . .	47
3.2.4	Nonlinear Conjugate Gradient Methods . . . . .	48
3.3	Trust Region Methods . . . . .	50
3.3.1	Changing the Trust Region and the Step . . . . .	51
3.3.2	Global Convergence of Trust Region Algorithms . . . . .	52
3.3.3	A Unidirectional Trust Region Algorithm . . . . .	54
3.3.4	The Exact Solution of the Trust Region Problem . . . . .	55
3.3.5	The Levenberg–Marquardt Parameter . . . . .	56
3.3.6	Superlinear Convergence: The Dogleg . . . . .	58
3.3.7	A Trust Region Method for Newton–CG . . . . .	63
3.4	Examples . . . . .	65
3.4.1	Parameter Identification . . . . .	67
3.4.2	Discrete Control Problem . . . . .	68
3.5	Exercises on Global Convergence . . . . .	68
<b>4</b>	<b>The BFGS Method</b>	<b>71</b>
4.1	Analysis . . . . .	72
4.1.1	Local Theory . . . . .	72
4.1.2	Global Theory . . . . .	77
4.2	Implementation . . . . .	78
4.2.1	Storage . . . . .	78
4.2.2	A BFGS–Armijo Algorithm . . . . .	80
4.3	Other Quasi-Newton Methods . . . . .	81
4.4	Examples . . . . .	83
4.4.1	Parameter ID Problem . . . . .	83
4.4.2	Discrete Control Problem . . . . .	83
4.5	Exercises on BFGS . . . . .	85
<b>5</b>	<b>Simple Bound Constraints</b>	<b>87</b>
5.1	Problem Statement . . . . .	87
5.2	Necessary Conditions for Optimality . . . . .	87
5.3	Sufficient Conditions . . . . .	89
5.4	The Gradient Projection Algorithm . . . . .	91
5.4.1	Termination of the Iteration . . . . .	91
5.4.2	Convergence Analysis . . . . .	93
5.4.3	Identification of the Active Set . . . . .	95
5.4.4	A Proof of Theorem 5.2.4 . . . . .	96
5.5	Superlinear Convergence . . . . .	96
5.5.1	The Scaled Gradient Projection Algorithm . . . . .	96
5.5.2	The Projected Newton Method . . . . .	100
5.5.3	A Projected BFGS–Armijo Algorithm . . . . .	102
5.6	Other Approaches . . . . .	104
5.6.1	Infinite-Dimensional Problems . . . . .	106
5.7	Examples . . . . .	106
5.7.1	Parameter ID Problem . . . . .	106
5.7.2	Discrete Control Problem . . . . .	106

5.8	Exercises on Bound Constrained Optimization . . . . .	108
<b>II Optimization of Noisy Functions</b>		<b>109</b>
<b>6</b>	<b>Basic Concepts and Goals</b>	<b>111</b>
6.1	Problem Statement . . . . .	112
6.2	The Simplex Gradient . . . . .	112
6.2.1	Forward Difference Simplex Gradient . . . . .	113
6.2.2	Centered Difference Simplex Gradient . . . . .	115
6.3	Examples . . . . .	118
6.3.1	Weber's Problem . . . . .	118
6.3.2	Perturbed Convex Quadratics . . . . .	119
6.3.3	Lennard–Jones Problem . . . . .	120
6.4	Exercises on Basic Concepts . . . . .	121
<b>7</b>	<b>Implicit Filtering</b>	<b>123</b>
7.1	Description and Analysis of Implicit Filtering . . . . .	123
7.2	Quasi-Newton Methods and Implicit Filtering . . . . .	124
7.3	Implementation Considerations . . . . .	125
7.4	Implicit Filtering for Bound Constrained Problems . . . . .	126
7.5	Restarting and Minima at All Scales . . . . .	127
7.6	Examples . . . . .	127
7.6.1	Weber's Problem . . . . .	127
7.6.2	Parameter ID . . . . .	129
7.6.3	Convex Quadratics . . . . .	129
7.7	Exercises on Implicit Filtering . . . . .	133
<b>8</b>	<b>Direct Search Algorithms</b>	<b>135</b>
8.1	The Nelder–Mead Algorithm . . . . .	135
8.1.1	Description and Implementation . . . . .	135
8.1.2	Sufficient Decrease and the Simplex Gradient . . . . .	137
8.1.3	McKinnon's Examples . . . . .	139
8.1.4	Restarting the Nelder–Mead Algorithm . . . . .	141
8.2	Multidirectional Search . . . . .	143
8.2.1	Description and Implementation . . . . .	143
8.2.2	Convergence and the Simplex Gradient . . . . .	144
8.3	The Hooke–Jeeves Algorithm . . . . .	145
8.3.1	Description and Implementation . . . . .	145
8.3.2	Convergence and the Simplex Gradient . . . . .	148
8.4	Other Approaches . . . . .	148
8.4.1	Surrogate Models . . . . .	148
8.4.2	The DIRECT Algorithm . . . . .	149
8.5	Examples . . . . .	152
8.5.1	Weber's Problem . . . . .	152
8.5.2	Parameter ID . . . . .	153
8.5.3	Convex Quadratics . . . . .	153
8.6	Exercises on Search Algorithms . . . . .	159

**Bibliography****161****Index****177**