

Contents

Part I Introduction to Stochastic Recursive Algorithms

1	Introduction	3
	1.1 Introduction	3
	1.2 Overview of the Remaining Chapters	7
	1.3 Concluding Remarks	11
	References	11
2	Deterministic Algorithms for Local Search	13
	2.1 Introduction	13
	2.2 Deterministic Algorithms for Local Search	14
	References	15
3	Stochastic Approximation Algorithms	17
	3.1 Introduction	17
	3.2 The Robbins-Monro Algorithm	18
	3.2.1 Convergence of the Robbins-Monro Algorithm	19
	3.3 Multi-timescale Stochastic Approximation	23
	3.3.1 Convergence of the Multi-timescale Algorithm	24
	3.4 Concluding Remarks	26
	References	27

Part II Gradient Estimation Schemes

4	Kiefer-Wolfowitz Algorithm	31
	4.1 Introduction	31
	4.2 The Basic Algorithm	31
	4.2.1 Extension to Multi-dimensional Parameter	35
	4.3 Variants of the Kiefer-Wolfowitz Algorithm	36
	4.3.1 Fixed Perturbation Parameter	36
	4.3.2 One-Sided Variants	37

4.4	Concluding Remarks	38
	References	38
5	Gradient Schemes with Simultaneous Perturbation Stochastic Approximation	41
5.1	Introduction	41
5.2	The Basic SPSA Algorithm	41
5.2.1	Gradient Estimate Using Simultaneous Perturbation	42
5.2.2	The Algorithm	43
5.2.3	Convergence Analysis	44
5.3	Variants of the Basic SPSA Algorithm	47
5.3.1	One-Measurement SPSA Algorithm	47
5.3.2	One-Sided SPSA Algorithm	49
5.3.3	Fixed Perturbation Parameter	49
5.4	General Remarks on SPSA Algorithms	51
5.5	SPSA Algorithms with Deterministic Perturbations	52
5.5.1	Properties of Deterministic Perturbation Sequences	52
5.5.2	Hadamard Matrix Based Construction	54
5.5.3	Two-Sided SPSA with Hadamard Matrix Perturbations	56
5.5.4	One-Sided SPSA with Hadamard Matrix Perturbations	62
5.5.5	One-Measurement SPSA with Hadamard Matrix Perturbations	63
5.6	SPSA Algorithms for Long-Run Average Cost Objective	65
5.6.1	The Framework	65
5.6.2	The Two-Simulation SPSA Algorithm	65
5.6.3	Assumptions	66
5.6.4	Convergence Analysis	68
5.6.5	Projected SPSA Algorithm	73
5.7	Concluding Remarks	75
	References	75
6	Smoothed Functional Gradient Schemes	77
6.1	Introduction	77
6.2	Gaussian Based SF Algorithm	79
6.2.1	Gradient Estimation via Smoothing	79
6.2.2	The Basic Gaussian SF Algorithm	81
6.2.3	Convergence Analysis of Gaussian SF Algorithm	82
6.2.4	Two-Measurement Gaussian SF Algorithm	88
6.3	General Conditions for a Candidate Smoothing Function	91
6.4	Cauchy Variant of the SF Algorithm	92
6.4.1	Gradient Estimate	92
6.4.2	Cauchy SF Algorithm	94
6.5	SF Algorithms for the Long-Run Average Cost Objective	94
6.5.1	The G-SF1 Algorithm	95
6.5.2	The G-SF2 Algorithm	99

6.5.3	Projected SF Algorithms	100
6.6	Concluding Remarks	101
	References	101

Part III Hessian Estimation Schemes

7	Newton-Based Simultaneous Perturbation Stochastic Approximation	105
7.1	Introduction	105
7.2	The Framework	106
7.3	Newton SPSA Algorithms	106
7.3.1	Four-Simulation Newton SPSA (N-SPSA4)	107
7.3.2	Three-Simulation Newton SPSA (N-SPSA3)	109
7.3.3	Two-Simulation Newton SPSA (N-SPSA2)	110
7.3.4	One-Simulation Newton SPSA (N-SPSA1)	111
7.4	Woodbury's Identity Based Newton SPSA Algorithms	113
7.5	Convergence Analysis	114
7.5.1	Assumptions	114
7.5.2	Convergence Analysis of N-SPSA4	117
7.5.3	Convergence Analysis of N-SPSA3	125
7.5.4	Convergence Analysis of N-SPSA2	126
7.5.5	Convergence Analysis of N-SPSA1	128
7.5.6	Convergence Analysis of W-SPSA Algorithms	130
7.6	Concluding Remarks	130
	References	131
8	Newton-Based Smoothed Functional Algorithms	133
8.1	Introduction	133
8.2	The Hessian Estimates	134
8.2.1	One-Simulation Hessian SF Estimate	134
8.2.2	Two-Simulation Hessian SF Estimate	136
8.3	The Newton SF Algorithms	137
8.3.1	The One-Simulation Newton SF Algorithm (N-SF1)	137
8.3.2	The Two-Simulation Newton SF Algorithm (N-SF2)	138
8.4	Convergence Analysis of Newton SF Algorithms	139
8.4.1	Convergence of N-SF1	139
8.4.2	Convergence of N-SF2	146
8.5	Concluding Remarks	147
	References	147

Part IV Variations to the Basic Scheme

9	Discrete Parameter Optimization	151
9.1	Introduction	151
9.2	The Framework	152
9.2.1	The Deterministic Projection Operator	153
9.2.2	The Random Projection Operator	154

9.2.3	A Generalized Projection Operator	155
9.2.4	Regular Projection Operator to \bar{C}	157
9.2.5	Basic Results for the Generalized Projection Operator Case ..	157
9.3	The Algorithms	160
9.3.1	The SPSA Algorithm	160
9.3.2	The SFA Algorithm	161
9.3.3	Convergence Analysis	162
9.4	Concluding Remarks	164
	References	166
10	Algorithms for Constrained Optimization	167
10.1	Introduction	167
10.2	The Framework	168
10.3	Algorithms	171
10.3.1	Constrained Gradient-Based SPSA Algorithm (CG-SPSA) ..	172
10.3.2	Constrained Newton-Based SPSA Algorithm (CN-SPSA) ..	173
10.3.3	Constrained Gradient-Based SF Algorithm (CG-SF)	175
10.3.4	Constrained Newton-Based SF Algorithm (CN-SF)	176
10.4	A Sketch of the Convergence	178
10.5	Concluding Remarks	185
	References	186
11	Reinforcement Learning	187
11.1	Introduction	187
11.2	Markov Decision Processes	188
11.3	Numerical Procedures for MDPs	191
11.3.1	Numerical Procedures for Discounted Cost MDPs	192
11.3.2	Numerical Procedures for Long-Run Average Cost MDPs ..	193
11.4	Reinforcement Learning Algorithms for Look-up Table Case	194
11.4.1	An Actor-Critic Algorithm for Infinite Horizon Discounted Cost MDPs	195
11.4.2	The Q-Learning Algorithm and a Simultaneous Perturbation Variant for Infinite Horizon Discounted Cost MDPs	198
11.4.3	Actor-Critic Algorithms for Long-Run Average Cost MDPs	202
11.5	Reinforcement Learning Algorithms with Function Approximation ..	206
11.5.1	Temporal Difference (TD) Learning with Discounted Cost ..	206
11.5.2	An Actor-Critic Algorithm with a Temporal Difference Critic for Discounted Cost MDPs	210
11.5.3	Function Approximation Based Q-learning Algorithm and a Simultaneous Perturbation Variant for Infinite Horizon Discounted Cost MDPs	213
11.6	Concluding Remarks	218
	References	218

Part V Applications

12	Service Systems	225
12.1	Introduction	225
12.2	Service System Framework	226
12.3	Problem Formulation	228
12.4	Solution Methodology	232
12.5	First Order Methods	234
12.5.1	SASOC-SPSA	234
12.5.2	SASOC-SF-N	235
12.5.3	SASOC-SF-C	236
12.6	Second Order Methods	236
12.6.1	SASOC-H	237
12.6.2	SASOC-W	237
12.7	Notes on Convergence	238
12.8	Summary of Experiments	239
12.9	Concluding Remarks	240
	References	240
13	Road Traffic Control	243
13.1	Introduction	243
13.2	Q-Learning for Traffic Light Control	245
13.2.1	Traffic Control Problem as an MDP	245
13.2.2	The TLC Algorithm	246
13.2.3	Summary of Experimental Results	248
13.3	Threshold Tuning Using SPSA	248
13.3.1	The Threshold Tuning Algorithm	249
13.3.2	Traffic Light Control with Threshold Tuning	250
13.3.3	Summary of Experimental Results	254
13.4	Concluding Remarks	255
	References	255
14	Communication Networks	257
14.1	Introduction	257
14.2	The Random Early Detection (RED) Scheme for the Internet	258
14.2.1	Introduction to RED Flow Control	258
14.2.2	The Framework	259
14.2.3	The B-RED and P-RED Stochastic Approximation Algorithms	263
14.2.4	Summary of Experimental Results	266
14.3	Optimal Policies for the Retransmission Probabilities in Slotted Aloha	267
14.3.1	Introduction to the Slotted Aloha Multiaccess Communication Protocol	267
14.3.2	The SDE Framework	268

14.3.3	The Algorithm	270
14.3.4	Summary of Experimental Results	272
14.4	Dynamic Multi-layered Pricing Schemes for the Internet	272
14.4.1	Introduction to Dynamic Pricing Schemes	273
14.4.2	The Pricing Framework	273
14.4.3	The Price Feed-Back Policies and the Algorithms	275
14.4.4	Summary of Experimental Results	277
14.5	Concluding Remarks	278
	References	279

Part VI Appendix

A	Convergence Notions for a Sequence of Random Vectors	283
	Reference	285
B	Martingales	287
	References	289
C	Ordinary Differential Equations	291
	References	294
D	The Borkar-Meyn Theorem for Stability and Convergence of Stochastic Approximation	295
	References	296
E	The Kushner-Clark Theorem for Convergence of Projected Stochastic Approximation	297
	References	300
	Index	301