

1. Record Nr.	UNISALENTO991004367530607536
Autore	Powell, Warren B., 1955-
Titolo	Reinforcement learning and stochastic optimization : a unified framework for sequential decisions / Warren B Powell
ISBN	9781119815037
Descrizione fisica	XXXIV, 1099 ; 24 cm
Disciplina	006.31
Soggetti	Reinforcement learning Decision making - Statistical methods Mathematical optimization Stochastic analysis
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di bibliografia	Includes bibliographical references and index
Nota di contenuto	Preface xxv -- Acknowledgments xxxi -- Part I -- Introduction 1 -- 1 Sequential Decision Problems 3 -- 1.1 The Audience 7 -- 1.2 The Communities of Sequential Decision Problems 8 -- 1.3 Our Universal Modeling Framework 10 -- 1.4 Designing Policies for Sequential Decision Problems 15 -- 1.5 Learning 20 -- 1.6 Themes 21 -- 1.7 Our Modeling Approach 27 -- 1.8 How to Read this Book 27 -- 1.9 Bibliographic Notes 33 -- Exercises 34 -- Bibliography 38 -- 2 Canonical Problems and Applications 39 -- 2.1 Canonical Problems 39 -- 2.2 A Universal Modeling Framework for Sequential Decision Problems 64 -- 2.3 Applications 69 -- 2.4 Bibliographic Notes 85 -- Exercises 90 -- Bibliography 93 -- 3 Online Learning 101 -- 3.1 Machine Learning for Sequential Decisions 102 -- 3.2 Adaptive Learning Using Exponential Smoothing 110 -- 3.3 Lookup Tables with Frequentist Updating 111 -- 3.4 Lookup Tables with Bayesian Updating 112 -- 3.5 Computing Bias and Variance* 118 -- 3.6 Lookup Tables and Aggregation* 121 -- 3.7 Linear Parametric Models 131 -- 3.8 Recursive Least Squares for Linear Models 136 -- 3.9 Nonlinear Parametric Models 140 -- 3.10 Nonparametric Models* 149 -- 3.11 Nonstationary Learning* 159 -- 3.12 The Curse of Dimensionality 162 -- 3.13 Designing Approximation Architectures in Adaptive Learning 165 -- 3.14 Why Does It Work?** 166 -- 3.15 Bibliographic Notes 174

-- Exercises 176 -- Bibliography 180 -- 4 Introduction to Stochastic Search 183 -- 4.1 Illustrations of the Basic Stochastic Optimization Problem 185 -- 4.2 Deterministic Methods 188 -- 4.3 Sampled Models 193 -- 4.4 Adaptive Learning Algorithms 202 -- 4.5 Closing Remarks 210 -- 4.6 Bibliographic Notes 210 -- Exercises 212 -- Bibliography 218 -- Part II – Stochastic Search 221 -- 5 Derivative-Based Stochastic Search 223 -- 5.1 Some Sample Applications 225 -- 5.2 Modeling Uncertainty 228 -- 5.3 Stochastic Gradient Methods 231 -- 5.4 Styles of Gradients 237 -- 5.5 Parameter Optimization for Neural Networks* 242 -- 5.6 Stochastic Gradient Algorithm as a Sequential Decision Problem 247 -- 5.7 Empirical Issues 248 -- 5.8 Transient Problems* 249 -- 5.9 Theoretical Performance* 250 -- 5.10 Why Does it Work? 250 -- 5.11 Bibliographic Notes 263 -- Exercises 264 -- Bibliography 270 -- 6 Stepsize Policies 273 -- 6.1 Deterministic Stepsize Policies 276 -- 6.2 Adaptive Stepsize Policies 282 -- 6.3 Optimal Stepsize Policies* 289 -- 6.4 Optimal Step sizes for Approximate Value Iteration* 297 -- 6.5 Convergence 300 -- 6.6 Guidelines for Choosing Stepsize Policies 301 -- 6.7 Why Does it Work* 303 -- 6.8 Bibliographic Notes 306 -- Exercises 307 -- Bibliography 314 -- 7 Derivative-Free Stochastic Search 317 -- 7.1 Overview of Derivative-free Stochastic Search 319 -- 7.2 Modeling Derivative-free Stochastic Search 325 -- 7.3 Designing Policies 330 -- 7.4 Policy Function Approximations 333 -- 7.5 Cost Function Approximations 335 -- 7.6 VFA-based Policies 338 -- 7.7 Direct Lookahead Policies 348 -- 7.8 The Knowledge Gradient (Continued)* 362 -- 7.9 Learning in Batches 380 -- 7.10 Simulation Optimization* 382 -- 7.11 Evaluating Policies 385 -- 7.12 Designing Policies 394 -- 7.13 Extensions* 398 -- 7.14 Bibliographic Notes 409 -- Exercises 412 -- Bibliography 424 -- Part III – State-dependent Problems 429 -- 8 State-dependent Problems 431 -- 8.1 Graph Problems 433 -- 8.2 Inventory Problems 439 -- 8.3 Complex Resource Allocation Problems 446 -- 8.4 State-dependent Learning Problems 456 -- 8.5 A Sequence of Problem Classes 460 -- 8.6 Bibliographic Notes 461 -- Exercises 462 -- Bibliography 466 -- 9 Modeling Sequential Decision Problems 467 -- 9.1 A Simple Modeling Illustration 471 -- 9.2 Notational Style 476 -- 9.3 Modeling Time 478 -- 9.4 The States of Our System 481 -- 9.5 Modeling Decisions 500 -- 9.6 The Exogenous Information Process 506 -- 9.7 The Transition Function 515 -- 9.8 The Objective Function 518 -- 9.9 Illustration: An Energy Storage Model 523 -- 9.10 Base Models and Lookahead Models 528 -- 9.11 A Classification of Problems* 529 -- 9.12 Policy Evaluation* 532 -- 9.13 Advanced Probabilistic Modeling Concepts** 534 -- 9.14 Looking Forward 540 -- 9.15 Bibliographic Notes 542 -- Exercises 544 -- Bibliography 557 -- 10 Uncertainty Modeling 559 -- 10.1 Sources of Uncertainty 560 -- 10.2 A Modeling Case Study: The COVID Pandemic 575 -- 10.3 Stochastic Modeling 575 -- 10.4 Monte Carlo Simulation 581 -- 10.5 Case Study: Modeling Electricity Prices 589 -- 10.6 Sampling vs. Sampled Models 595 -- 10.7 Closing Notes 597 -- 10.8 Bibliographic Notes 597 -- Exercises 598 -- Bibliography 601 -- 11 Designing Policies 603 -- 11.1 From Optimization to Machine Learning to Sequential Decision Problems 605 -- 11.2 The Classes of Policies 606 -- 11.3 Policy Function Approximations 610 -- 11.4 Cost Function Approximations 613 -- 11.5 Value Function Approximations 614 -- 11.6 Direct Lookahead Approximations 616 -- 11.7 Hybrid Strategies 620 -- 11.8 Randomized Policies 626 -- 11.9 Illustration: An Energy Storage Model Revisited 627 -- 11.10 Choosing the Policy Class 631 -- 11.11 Policy Evaluation 641 -- 11.12 Parameter Tuning 642 -- 11.13 Bibliographic Notes 646 -- Exercises 646 -- Bibliography 651 -- Part IV – Policy Search 653 -- 12 Policy Function

Approximations and Policy Search 655 -- 12.1 Policy Search as a Sequential Decision Problem 657 -- 12.2 Classes of Policy Function Approximations 658 -- 12.3 Problem Characteristics 665 -- 12.4 Flavors of Policy Search 666 -- 12.5 Policy Search with Numerical Derivatives 669 -- 12.6 Derivative-Free Methods for Policy Search 670 -- 12.7 Exact Derivatives for Continuous Sequential Problems* 677 -- 12.8 Exact Derivatives for Discrete Dynamic Programs** 680 -- 12.9 Supervised Learning 686 -- 12.10 Why Does it Work? 687 -- 12.11 Bibliographic Notes 690 -- Exercises 691 -- Bibliography 698 -- 13 Cost Function Approximations 701 -- 13.1 General Formulation for Parametric CFA 703 -- 13.2 Objective-Modified CFAs 704 -- 13.3 Constraint-Modified CFAs 714 -- 13.4 Bibliographic Notes 725 -- Exercises 726 -- Bibliography 729 -- Part V – Lookahead Policies 731 -- 14 Exact Dynamic Programming 737 -- 14.1 Discrete Dynamic Programming 738 -- 14.2 The Optimality Equations 740 -- 14.3 Finite Horizon Problems 747 -- 14.4 Continuous Problems with Exact Solutions 750 -- 14.5 Infinite Horizon Problems* 755 -- 14.6 Value Iteration for Infinite Horizon Problems* 757 -- 14.7 Policy Iteration for Infinite Horizon Problems* 762 -- 14.8 Hybrid Value-Policy Iteration* 764 -- 14.9 Average Reward Dynamic Programming* 765 -- 14.10 The Linear Programming Method for Dynamic Programs** 766 -- 14.11 Linear Quadratic Regulation 767 -- 14.12 Why Does it Work?*** 770 -- 14.13 Bibliographic Notes 783 -- Exercises 783 -- Bibliography 793 -- 15 Backward Approximate Dynamic Programming 795 -- 15.1 Backward Approximate Dynamic Programming for Finite Horizon Problems 797 -- 15.2 Fitted Value Iteration for Infinite Horizon Problems 804 -- 15.3 Value Function Approximation Strategies 805 -- 15.4 Computational Observations 810 -- 15.5 Bibliographic Notes 816 -- Exercises 816 -- Bibliography 821 -- 16 Forward ADP I: The Value of a Policy 823 -- 16.1 Sampling the Value of a Policy 824 -- 16.2 Stochastic Approximation Methods 835 -- 16.3 Bellman's Equation Using a Linear Model* 837 -- 16.4 Analysis of TD(0), LSTD, and LSPE Using a Single State* 842 -- 16.5 Gradient-based Methods for Approximate Value Iteration* 845 -- 16.6 Value Function Approximations Based on Bayesian Learning* 852 -- 16.7 Learning Algorithms and Atepsizes 855 -- 16.8 Bibliographic Notes 860 -- Exercises 862 -- Bibliography 864 -- 17 Forward ADP II: Policy Optimization 867 -- 17.1 Overview of Algorithmic Strategies 869 -- 17.2 Approximate Value Iteration and Q-Learning Using Lookup Tables 871 -- 17.3 Styles of Learning 881 -- 17.4 Approximate Value Iteration Using Linear Models 886 -- 17.5 On-policy vs. off-policy learning and the exploration–exploitation problem 888 -- 17.6 Applications 894 -- 17.7 Approximate Policy Iteration 900 -- 17.8 The Actor–Critic Paradigm 907 -- 17.9 Statistical Bias in the Max Operator* 909 -- 17.10 The Linear Programming Method Using Linear Models* 912 -- 17.11 Finite Horizon Approximations for Steady-State Applications 915 -- 17.12 Bibliographic Notes 917 -- Exercises 918 -- Bibliography 924 -- 18 Forward ADP III: Convex Resource Allocation Problems 927 -- 18.1 Resource Allocation Problems 930 -- 18.2 Values Versus Marginal Values 937 -- 18.3 Piecewise Linear Approximations for Scalar Functions 938 -- 18.4 Regression Methods 941 -- 18.5 Separable Piecewise Linear Approximations 944 -- 18.6 Benders Decomposition for Nonseparable Approximations** 946 -- 18.7 Linear Approximations for High-Dimensional Applications 956 -- 18.8 Resource Allocation with Exogenous Information State 958 -- 18.9 Closing Notes 959 -- 18.10 Bibliographic Notes 960 -- Exercises 962 -- Bibliography 967 -- 19 Direct Lookahead Policies 971 -- 19.1 Optimal Policies Using Lookahead Models 974 -- 19.2 Creating an

Approximate Lookahead Model 978 -- 19.3 Modified Objectives in
Lookahead Models 985 -- 19.4 Evaluating DLA Policies 992 -- 19.5
Why Use a DLA? 997 -- 19.6 Deterministic Lookaheads 999 -- 19.7 A
Tour of Stochastic Lookahead Policies 1005 -- 19.8 Monte Carlo Tree
Search for Discrete Decisions 1009 -- 19.9 Two-Stage Stochastic
Programming for Vector Decisions* 1018 -- 19.10 Observations on
DLA Policies 1024 -- 19.11 Bibliographic Notes 1025 -- Exercises
1027 -- Bibliography 1031 -- Part VI – Multiagent Systems 1033 -- 20
Multiagent Modeling and Learning 1035 -- 20.1 Overview of Multiagent
Systems 1036 -- 20.2 A Learning Problem – Flu Mitigation 1044 --
20.3 The POMDP Perspective* 1059 -- 20.4 The Two-Agent
Newsvendor Problem 1062 -- 20.5 Multipl ...
