1. Record Nr. UNINA9910830085103321

Autore	Hansson Anders
Titolo	Optimization for learning and control / / Anders Hansson and Martin Andersen
Pubbl/distr/stampa	Hoboken, New Jersey : , : John Wiley & Sons, Inc., , [2023] ©2023
ISBN	1-119-80918-5 1-119-80914-2
Descrizione fisica	1 online resource (435 pages)
Disciplina	519.3
Soggetti	System analysis - Mathematics Mathematical optimization Machine learning - Mathematics Signal processing - Mathematics
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	Cover -- Title Page -- Copyright -- Contents -- Preface -- Acknowledgments -- Glossary -- Acronyms -- About the Companion Website -- Part I Introductory Part -- Chapter 1 Introduction -- 1.1 Optimization -- 1.2 Unsupervised Learning -- 1.3 Supervised Learning -- 1.4 System Identification -- 1.5 Control -- 1.6 Reinforcement Learning -- 1.7 Outline -- Chapter 2 Linear Algebra -- 2.1 Vectors and Matrices -- 2.2 Linear Maps and Subspaces -- 2.2.1 Four Fundamental Subspaces -- 2.2.2 Square Matrices -- 2.2.3 Affine Sets -- 2.3 Norms -- 2.4 Algorithm Complexity -- 2.5 Matrices with Structure -- 2.5.1 Diagonal Matrices -- 2.5.2 Orthogonal Matrices -- 2.5.3 Triangular Matrices -- 2.5.4 Symmetric and SkewSymmetric Matrices -- 2.5.5 Toeplitz and Hankel Matrices -- 2.5.6 Sparse Matrices -- 2.5.7 Band Matrices -- 2.6 Quadratic Forms and Definiteness -- 2.7 Spectral Decomposition -- 2.8 Singular Value Decomposition -- 2.9 Moore-Penrose Pseudoinverse -- 2.10 Systems of Linear Equations -- 2.10.1 Gaussian Elimination -- 2.10.2 Block Elimination -- 2.11 Factorization Methods -- 2.11.1 LU Factorization -- 2.11.2 Cholesky Factorization -- 2.11.3 Indefinite LDL Factorization -- 2.11.4 QR Factorization --

2.11.5 Sparse Factorizations -- 2.11.6 Block Factorization -- 2.11.7 Positive Semidefinite Block Factorization -- 2.12 SaddlePoint Systems -- 2.12.1 H Positive Definite -- 2.12.2 H Positive Semidefinite -- 2.13 Vector and Matrix Calculus -- Exercises -- Chapter 3 Probability Theory -- 3.1 Probability Spaces -- 3.1.1 Probability Measure -- 3.1.2 Probability Function -- 3.1.3 Probability Density Function -- 3.2 Conditional Probability -- 3.3 Independence -- 3.4 Random Variables -- 3.4.1 VectorValued Random Variable -- 3.4.2 Marginal Distribution -- 3.4.3 Independence of Random Variables -- 3.4.4 Function of Random Variable -- 3.5 Conditional Distributions.
3.5.1 Conditional Probability Function -- 3.5.2 Conditional Probability Density Function -- 3.6 Expectations -- 3.6.1 Moments -- 3.6.2 Expected Value of Function of Random Variable -- 3.6.3 Covariance -- 3.7 Conditional Expectations -- 3.8 Convergence of Random Variables -- 3.9 Random Processes -- 3.10 Markov Processes -- 3.11 Hidden Markov Models -- 3.12 Gaussian Processes -- Exercises -- Part II Optimization -- Chapter 4 Optimization Theory -- 4.1 Basic Concepts and Terminology -- 4.1.1 Optimization Problems -- 4.1.2 Equivalent Problems -- 4.2 Convex Sets -- 4.2.1 ConvexityPreserving Operations -- 4.2.1.1 Intersection -- 4.2.1.2 Affine Transformation -- 4.2.1.3 Perspective Transformation -- 4.2.2 Examples of Convex Sets -- 4.2.2.1 Hyperplanes and Halfspaces -- 4.2.2.2 Polyhedral Sets -- 4.2.2.3 Norm Balls and Ellipsoids -- 4.2.2.4 Convex Cones -- 4.2.3 Generalized Inequalities -- 4.3 Convex Functions -- 4.3.1 First and SecondOrder Conditions for Convexity -- 4.3.2 ConvexityPreserving Operations -- 4.3.2.1 Scaling, Sums, and Integrals -- 4.3.2.2 Pointwise Maximum and Supremum -- 4.3.2.3 Affine Transformation -- 4.3.2.4 Perspective Transformation -- 4.3.2.5 Partial Infimum -- 4.3.2.6 Square of Nonnegative Convex Functions -- 4.3.3 Examples of Convex Functions -- 4.3.3.1 Norms -- 4.3.3.2 Indicator and Support Functions -- 4.3.4 Conjugation -- 4.3.5 Dual Norms -- 4.4 Subdifferentiability -- 4.4.1 Subdifferential Calculus -- 4.4.1.1 Nonnegative Scaling -- 4.4.1.2 Summation -- 4.4.1.3 Affine Transformation -- 4.4.1.4 Pointwise Maximum -- 4.4.1.5 Subgradients of Conjugate Functions -- 4.5 Convex Optimization Problems -- 4.5.1 Optimality Condition -- 4.5.2 Equality Constrained Convex Problems -- 4.6 Duality -- 4.6.1 Lagrangian Duality -- 4.6.2 Lagrange Dual Problem -- 4.6.3 Fenchel Duality -- 4.7 Optimality Conditions.
4.7.1 Convex Optimization Problems -- 4.7.2 Nonconvex Optimization Problems -- Exercises -- Chapter 5 Optimization Problems -- 5.1 LeastSquares Problems -- 5.2 Quadratic Programs -- 5.3 Conic Optimization -- 5.3.1 Conic Duality -- 5.3.2 Epigraphical Cones -- 5.4 Rank Optimization -- 5.5 Partially Separability -- 5.5.1 Minimization of Partially Separable Functions -- 5.5.2 Principle of Optimality -- 5.6 Multiparametric Optimization -- 5.7 Stochastic Optimization -- Exercises -- Chapter 6 Optimization Methods -- 6.1 Basic Principles -- 6.1.1 Smoothness -- 6.1.2 Descent Methods -- 6.1.3 Line Search Methods -- 6.1.3.1 Backtracking Line Search -- 6.1.3.2 Bisection Method for Wolfe Conditions -- 6.1.4 Surrogation Methods -- 6.1.4.1 TrustRegion Methods -- 6.1.4.2 Majorization Minimization -- 6.1.5 Convergence of Sequences -- 6.2 Gradient Descent -- 6.2.1 LSmooth Functions -- 6.2.2 Smooth and Convex Functions -- 6.2.3 Smooth and Strongly Convex Functions -- 6.3 Newton's Method -- 6.3.1 The Newton Decrement -- 6.3.2 Analysis of Newton's Method -- 6.3.2.1 Affine Invariance -- 6.3.2.2 Pure Newton Phase -- 6.3.2.3 Damped Newton Phase -- 6.3.3 Equality Constrained Minimization -- 6.4 Variable Metric Methods -- 6.4.1 QuasiNewton Updates -- 6.4.1.1 The BFGS Update -- 6.4.1.2 The DFP Update -- 6.4.1.3 The SR1 Update --

6.4.2 The Barzilai-Borwein Method -- 6.5 Proximal Gradient Method -- 6.5.1 Gradient Projection Method -- 6.5.2 Proximal QuasiNewton -- 6.5.3 Accelerated Proximal Gradient Method -- 6.6 Sequential Convex Optimization -- 6.7 Methods for Nonlinear LeastSquares -- 6.7.1 The LevenbergMarquardt Algorithm -- 6.7.2 The Variable Projection Method -- 6.8 Stochastic Optimization Methods -- 6.8.1 Smooth Functions -- 6.8.2 Smooth and Strongly Convex Functions -- 6.8.3 Incremental Methods -- 6.8.4 Adaptive Methods -- 6.8.4.1 AdaGrad -- 6.8.4.2 RMSprop.
6.8.4.3 Adam -- 6.9 Coordinate Descent Methods -- 6.10 InteriorPoint Methods -- 6.10.1 PathFollowing Method -- 6.10.2 Generalized Inequalities -- 6.11 Augmented Lagrangian Methods -- 6.11.1 Method of Multipliers -- 6.11.2 Alternating Direction Method of Multipliers -- 6.11.3 Variable Splitting -- Exercises -- Part III Optimal Control -- Chapter 7 Calculus of Variations -- 7.1 Extremum of Functionals -- 7.1.1 Necessary Condition for Extremum -- 7.1.2 Sufficient Condition for Optimality -- 7.1.3 Constrained Problem -- 7.1.4 Du Bois-Reymond Lemma -- 7.1.5 Generalizations -- 7.2 The Pontryagin Maximum Principle -- 7.2.1 Linear Quadratic Control -- 7.2.2 The Riccati Equation -- 7.3 The Euler-Lagrange Equations -- 7.3.1 Beltrami's Identity -- 7.4 Extensions -- 7.5 Numerical Solutions -- 7.5.1 The Gradient Method -- 7.5.2 The Shooting Method -- 7.5.3 The Discretization Method -- 7.5.4 The Multiple Shooting Method -- 7.5.5 The Collocation Method -- Exercises -- Chapter 8 Dynamic Programming -- 8.1 Finite Horizon Optimal Control -- 8.1.1 Standard Optimization Problem -- 8.1.2 Dynamic Programming -- 8.2 Parametric Approximations -- 8.2.1 FittedValue Iteration -- 8.3 Infinite Horizon Optimal Control -- 8.3.1 Bellman Equation -- 8.4 Value Iterations -- 8.5 Policy Iterations -- 8.5.1 Approximation -- 8.6 Linear Programming Formulation -- 8.6.1 Approximations -- 8.7 Model Predictive Control -- 8.7.1 Infinite Horizon Problem -- 8.7.2 Guessing the Value Function -- 8.7.3 Finite Horizon Approximation -- 8.7.4 Receding Horizon Approximation -- 8.8 Explicit MPC -- 8.9 Markov Decision Processes -- 8.9.1 Stochastic Dynamic Programming -- 8.9.2 Infinite Time Horizon -- 8.9.3 Stochastic Bellman Equation -- 8.10 Appendix -- 8.10.1 Stability and Optimality of Infinite Horizon Problem -- 8.10.2 Stability and Optimality of Stochastic Infinite Time Horizon Problem.
8.10.3 Stability of MPC -- Exercises -- Part IV Learning -- Chapter 9 Unsupervised Learning -- 9.1 Chebyshev Bounds -- 9.2 Entropy -- 9.2.1 Categorical Distribution -- 9.2.2 Ising Distribution -- 9.2.3 Normal Distribution -- 9.3 Prediction -- 9.3.1 Conditional Expectation Predictor -- 9.3.2 Affine Predictor -- 9.3.3 Linear Regression -- 9.4 The Viterbi Algorithm -- 9.5 Kalman Filter on Innovation Form -- 9.6 Viterbi Decoder -- 9.7 Graphical Models -- 9.7.1 Ising Distribution -- 9.7.2 Normal Distribution -- 9.7.3 Markov Random Field -- 9.8 Maximum Likelihood Estimation -- 9.8.1 Categorical Distribution -- 9.8.2 Ising Distribution -- 9.8.3 Normal Distribution -- 9.8.4 Generalizations -- 9.9 Relative Entropy and Cross Entropy -- 9.9.1 Gibbs' Inequality -- 9.9.2 Cross Entropy -- 9.10 The Expectation Maximization Algorithm -- 9.11 Mixture Models -- 9.12 Gibbs Sampling -- 9.13 Boltzmann Machine -- 9.14 Principal Component Analysis -- 9.14.1 Solution -- 9.14.2 Relation to RankConstrained Optimization -- 9.15 Mutual Information -- 9.15.1 Channel Model -- 9.15.2 Orthogonal Case -- 9.15.3 Nonorthogonal Case -- 9.15.4 Relationship to PCA -- 9.16 Cluster Analysis -- Exercises -- Chapter 10 Supervised Learning -- 10.1 Linear Regression -- 10.1.1 Least Squares Estimation -- 10.1.2 Maximum Likelihood Estimation --

10.1.3 Maximum a Posteriori Estimation -- 10.2 Regression in Hilbert Spaces -- 10.2.1 InfiniteDimensional LS Problem -- 10.2.2 The Kernel Trick -- 10.3 Gaussian Processes -- 10.3.1 Gaussian MAP Estimate -- 10.3.2 The Kernel Trick -- 10.4 Classification -- 10.4.1 Linear Regression -- 10.4.2 Logistic Regression -- 10.5 Support Vector Machines -- 10.5.1 Hebbian Learning -- 10.5.2 Quadratic Programming Formulation -- 10.5.3 Soft Margin Classification -- 10.5.4 The Dual Problem -- 10.5.5 Recovering the Primal Solution -- 10.5.6 The Kernel Trick.
10.6 Restricted Boltzmann Machine.

Sommario/riassunto

"Comprehensive resource providing a masters' level introduction to optimization theory and algorithms for learning and control Optimization for Learning and Control describes how optimization is used in these domains, giving a thorough introduction to both unsupervised learning, supervised learning, and reinforcement learning, with an emphasis on optimization methods for large-scale learning and control problems.A Several applications areas are also discussed, including signal processing, system identification, optimal control, and machine learning. Today, most of the material on the optimization aspects of deep learning that is accessible for students at a Masters' level is focused on surface-level computer programming; deeper knowledge about the optimization methods and the trade-offs that are behind these methods is not provided. The objective of this book is to make this scattered knowledge, currently mainly available in publications in academic journals, accessible for Masters' students in a coherent way"--