Data-Driven Science and Engineering

Machine Learning, Dynamical Systems, and Control

Second Edition

STEVEN L. BRUNTON
University of Washington

J. NATHAN KUTZ
University of Washington

CAMBRIDGE
UNIVERSITY PRESS

Contents

Preface — *page* ix
Acknowledgments — xiv
Common Optimization Techniques, Equations, Symbols, and Acronyms — xv

Part I Dimensionality Reduction and Transforms — 1

1 Singular Value Decomposition (SVD) — 3
1.1 Overview — 3
1.2 Matrix Approximation — 7
1.3 Mathematical Properties and Manipulations — 12
1.4 Pseudo-Inverse, Least-Squares, and Regression — 16
1.5 Principal Component Analysis (PCA) — 23
1.6 Eigenfaces Example — 28
1.7 Truncation and Alignment — 35
1.8 Randomized Singular Value Decomposition — 40
1.9 Tensor Decompositions and N-Way Data Arrays — 46

2 Fourier and Wavelet Transforms — 53
2.1 Fourier Series and Fourier Transforms — 53
2.2 Discrete Fourier Transform (DFT) and Fast Fourier Transform (FFT) — 63
2.3 Transforming Partial Differential Equations — 70
2.4 Gabor Transform and the Spectrogram — 76
2.5 Laplace Transform — 81
2.6 Wavelets and Multi-Resolution Analysis — 85
2.7 Two-Dimensional Transforms and Image Processing — 87

3 Sparsity and Compressed Sensing — 97
3.1 Sparsity and Compression — 97
3.2 Compressed Sensing — 101
3.3 Compressed Sensing Examples — 105
3.4 The Geometry of Compression — 109
3.5 Sparse Regression — 113
3.6 Sparse Representation — 117
3.7 Robust Principal Component Analysis (RPCA) — 120
3.8 Sparse Sensor Placement — 123

Part II Machine Learning and Data Analysis 131

4 Regression and Model Selection 133
4.1 Classic Curve Fitting 134
4.2 Nonlinear Regression and Gradient Descent 140
4.3 Regression and Ax = b: Over- and Under-Determined Systems 145
4.4 Optimization as the Cornerstone of Regression 151
4.5 The Pareto Front and *Lex Parsimoniae* 155
4.6 Model Selection: Cross-Validation 158
4.7 Model Selection: Information Criteria 162

5 Clustering and Classification 168
5.1 Feature Selection and Data Mining 169
5.2 Supervised versus Unsupervised Learning 174
5.3 Unsupervised Learning: k-Means Clustering 178
5.4 Unsupervised Hierarchical Clustering: Dendrogram 182
5.5 Mixture Models and the Expectation-Maximization Algorithm 186
5.6 Supervised Learning and Linear Discriminants 189
5.7 Support Vector Machines (SVM) 193
5.8 Classification Trees and Random Forest 198
5.9 Top 10 Algorithms of Data Mining circa 2008 (Before the Deep Learning Revolution) 203

6 Neural Networks and Deep Learning 208
6.1 Neural Networks: Single-Layer Networks 209
6.2 Multi-Layer Networks and Activation Functions 214
6.3 The Backpropagation Algorithm 219
6.4 The Stochastic Gradient Descent Algorithm 222
6.5 Deep Convolutional Neural Networks 224
6.6 Neural Networks for Dynamical Systems 228
6.7 Recurrent Neural Networks 233
6.8 Autoencoders 236
6.9 Generative Adversarial Networks (GANs) 240
6.10 The Diversity of Neural Networks 242

Part III Dynamics and Control 251

7 Data-Driven Dynamical Systems 253
7.1 Overview, Motivations, and Challenges 254
7.2 Dynamic Mode Decomposition (DMD) 260
7.3 Sparse Identification of Nonlinear Dynamics (SINDy) 275
7.4 Koopman Operator Theory 286
7.5 Data-Driven Koopman Analysis 296

8 Linear Control Theory 311

8.1 Closed-Loop Feedback Control 312
8.2 Linear Time-Invariant Systems 317
8.3 Controllability and Observability 322
8.4 Optimal Full-State Control: Linear–Quadratic Regulator (LQR) 328
8.5 Optimal Full-State Estimation: the Kalman Filter 332
8.6 Optimal Sensor-Based Control: Linear–Quadratic Gaussian (LQG) 335
8.7 Case Study: Inverted Pendulum on a Cart 336
8.8 Robust Control and Frequency-Domain Techniques 346

9 Balanced Models for Control 360

9.1 Model Reduction and System Identification 360
9.2 Balanced Model Reduction 361
9.3 System Identification 375

Part IV Advanced Data-Driven Modeling and Control 387

10 Data-Driven Control 389

10.1 Model Predictive Control (MPC) 390
10.2 Nonlinear System Identification for Control 392
10.3 Machine Learning Control 398
10.4 Adaptive Extremum-Seeking Control 408

11 Reinforcement Learning 419

11.1 Overview and Mathematical Formulation 419
11.2 Model-Based Optimization and Control 426
11.3 Model-Free Reinforcement Learning and Q-Learning 429
11.4 Deep Reinforcement Learning 436
11.5 Applications and Environments 440
11.6 Optimal Nonlinear Control 444

12 Reduced-Order Models (ROMs) 449

12.1 Proper Orthogonal Decomposition (POD) for Partial Differential Equations 449
12.2 Optimal Basis Elements: the POD Expansion 455
12.3 POD and Soliton Dynamics 461
12.4 Continuous Formulation of POD 465
12.5 POD with Symmetries: Rotations and Translations 470
12.6 Neural Networks for Time-Stepping with POD 475
12.7 Leveraging DMD and SINDy for Galerkin–POD 479

13 Interpolation for Parametric Reduced-Order Models 485

13.1 Gappy POD 485
13.2 Error and Convergence of Gappy POD 490
13.3 Gappy Measurements: Minimize Condition Number 493
13.4 Gappy Measurements: Maximal Variance 497

13.5 POD and the Discrete Empirical Interpolation Method (DEIM) 500
13.6 DEIM Algorithm Implementation 504
13.7 Decoder Networks for Interpolation 508
13.8 Randomization and Compression for ROMs 512
13.9 Machine Learning ROMs 513

14 Physics-Informed Machine Learning 520
14.1 Mathematical Foundations 520
14.2 SINDy Autoencoder: Coordinates and Dynamics 523
14.3 Koopman Forecasting 526
14.4 Learning Nonlinear Operators 529
14.5 Physics-Informed Neural Networks (PINNs) 533
14.6 Learning Coarse-Graining for PDEs 535
14.7 Deep Learning and Boundary Value Problems 539

Glossary 542
References 552
Index 588