

Data-Driven Science and Engineering

Machine Learning, Dynamical Systems, and Control

STEVEN L. BRUNTON

University of Washington

J. NATHAN KUTZ

University of Washington



CAMBRIDGE
UNIVERSITY PRESS

Contents

<i>Preface</i>	<i>page</i> ix
<i>Common Optimization Techniques, Equations, Symbols, and Acronyms</i>	xiii

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	10
1.4 Pseudo-Inverse, Least-Squares, and Regression	15
1.5 Principal Component Analysis (PCA)	21
1.6 Eigenfaces Example	25
1.7 Truncation and Alignment	30
1.8 Randomized Singular Value Decomposition	37
1.9 Tensor Decompositions and N -Way Data Arrays	41
2 Fourier and Wavelet Transforms	47
2.1 Fourier Series and Fourier Transforms	47
2.2 Discrete Fourier Transform (DFT) and Fast Fourier Transform (FFT)	56
2.3 Transforming Partial Differential Equations	63
2.4 Gabor Transform and the Spectrogram	69
2.5 Wavelets and Multi-Resolution Analysis	75
2.6 2D Transforms and Image Processing	77
3 Sparsity and Compressed Sensing	84
3.1 Sparsity and Compression	84
3.2 Compressed Sensing	88
3.3 Compressed Sensing Examples	92
3.4 The Geometry of Compression	95
3.5 Sparse Regression	98
3.6 Sparse Representation	103
3.7 Robust Principal Component Analysis (RPCA)	107
3.8 Sparse Sensor Placement	110

Part II	Machine Learning and Data Analysis	115
4	Regression and Model Selection	117
4.1	Classic Curve Fitting	118
4.2	Nonlinear Regression and Gradient Descent	123
4.3	Regression and $\mathbf{Ax} = \mathbf{b}$: Over- and Under-Determined Systems	130
4.4	Optimization as the Cornerstone of Regression	136
4.5	The Pareto Front and <i>Lex Parsimoniae</i>	140
4.6	Model Selection: Cross-Validation	143
4.7	Model Selection: Information Criteria	148
5	Clustering and Classification	154
5.1	Feature Selection and Data Mining	155
5.2	Supervised versus Unsupervised Learning	160
5.3	Unsupervised Learning: k -means Clustering	164
5.4	Unsupervised Hierarchical Clustering: Dendrogram	168
5.5	Mixture Models and the Expectation-Maximization Algorithm	172
5.6	Supervised Learning and Linear Discriminants	176
5.7	Support Vector Machines (SVM)	180
5.8	Classification Trees and Random Forest	185
5.9	Top 10 Algorithms in Data Mining 2008	190
6	Neural Networks and Deep Learning	195
6.1	Neural Networks: 1-Layer Networks	196
6.2	Multi-Layer Networks and Activation Functions	199
6.3	The Backpropagation Algorithm	204
6.4	The Stochastic Gradient Descent Algorithm	209
6.5	Deep Convolutional Neural Networks	212
6.6	Neural Networks for Dynamical Systems	216
6.7	The Diversity of Neural Networks	220
Part III	Dynamics and Control	227
7	Data-Driven Dynamical Systems	229
7.1	Overview, Motivations, and Challenges	230
7.2	Dynamic Mode Decomposition (DMD)	235
7.3	Sparse Identification of Nonlinear Dynamics (SINDy)	247
7.4	Koopman Operator Theory	257
7.5	Data-Driven Koopman Analysis	266
8	Linear Control Theory	276
8.1	Closed-Loop Feedback Control	277
8.2	Linear Time-Invariant Systems	281
8.3	Controllability and Observability	287
8.4	Optimal Full-State Control: Linear Quadratic Regulator (LQR)	292

8.5	Optimal Full-State Estimation: The Kalman Filter	296
8.6	Optimal Sensor-Based Control: Linear Quadratic Gaussian (LQG)	299
8.7	Case Study: Inverted Pendulum on a Cart	300
8.8	Robust Control and Frequency Domain Techniques	308
9	Balanced Models for Control	321
9.1	Model Reduction and System Identification	321
9.2	Balanced Model Reduction	322
9.3	System identification	336
10	Data-Driven Control	345
10.1	Nonlinear System Identification for Control	346
10.2	Machine Learning Control	352
10.3	Adaptive Extremum-Seeking Control	362
Part IV	Reduced Order Models	373
11	Reduced Order Models (ROMs)	375
11.1	POD for Partial Differential Equations	375
11.2	Optimal Basis Elements: The POD Expansion	381
11.3	POD and Soliton Dynamics	387
11.4	Continuous Formulation of POD	391
11.5	POD with Symmetries: Rotations and Translations	396
12	Interpolation for Parametric ROMs	403
12.1	Gappy POD	403
12.2	Error and Convergence of Gappy POD	409
12.3	Gappy Measurements: Minimize Condition Number	413
12.4	Gappy Measurements: Maximal Variance	418
12.5	POD and the Discrete Empirical Interpolation Method (DEIM)	423
12.6	DEIM Algorithm Implementation	426
12.7	Machine Learning ROMs	429
	<i>Glossary</i>	436
	<i>Bibliography</i>	443
	<i>Index</i>	471