I Probabilistic Models

Review of Basic Probability
0.1 Discrete Probability . . . .
0.2 Conditional Probability, Bayes' Rule, and Independence . . . .
0.3 Random Variables on Discrete Probability Spaces . . . .
0.4 Discrete Distributions I . . . .
0.5 Discrete Distributions II . . . .
0.6 Continuous Random Variables . . . .
Exercises . . . .

More Probability
1.1 Modes of Convergence . . . .
1.2 Moments and Characteristic Functions . . . .
1.3 Important Inequalities . . . .
1.4 The Central Limit Theorem . . . .
1.5 Functions (Transformations) of Random Variables . . . .
1.6 Conditioning of Random Variables . . . .
1.7 Information and Entropy . . . .

Markov Processes
2.1 Markov Processes . . . .
Exercises . . . .

II Basic Statistical Modeling

Classical Inference
3.1 Nonparametric Estimation: CDFs . . . .
3.2 Nonparametric Estimation: Histograms and KDE . . . .

3.3 Parametric Estimation . . . .
3.4 *Key Properties of the MLE and Method of Moments . . . .
3.5 Fisher Information . . . .
3.6 Bootstrap Methods . . . .
3.7 Hypothesis Testing . . . .
3.8 Bayesian Statistics I . . . .
3.9 Bayesian Statistics II: Priors and Model Selection . . . .
3.10 Classical Sampling Methods . . . .

Regression
4.1 Linear Regression and GLMs . . . .
4.2 Linear Regression . . . .
4.3 Logistic Regression . . . .
4.4 Model Selection . . . .
4.5 Regularization . . . .
Exercises . . . .

III Adaptive Methods

Graphical Models
5.1 Graphical Models . . . .
5.2 Latent Variable Models and Expectation Maximization . . . .
5.3 Mixture Models . . . .
5.4 Discrete Hidden Markov Models . . . .
5.5 CDHMMs . . . .
5.6 MCMC: Metropolis-Hastings . . . .
5.7 MCMC: Gibbs . . . .
5.8 LDA . . . .
Exercises . . . .

Estimation in State-Space Models
6.1 State-Space Modeling . . . .
6.2 Estimating State Variables . . . .
6.3 Recursive Least Squares . . . .
6.4 Kalman Filtering I . . . .
6.5 Kalman Filtering II . . . .
6.6 Time Series Models . . . .
6.7 Training ARMA Models . . . .
6.8 Forecasting ARMA Models . . . .
Exercises . . . .

IV Interlude

Review of Multivariate Calculus
7.1 Gradient and Hessian and Chain Rule . . . .
7.2 Matrix Derivatives . . . .

Review of Optimization
8.1 Stochastic Gradient Descent . . . .
8.2 Second-order Methods . . . .
8.3 Duality in convex optimization . . . .

V Machine Learning

ML basics
9.1 Fundamental Concepts . . . .
Exercises . . . .

Unsupervised Methods
10.1 Linear Dimensionality Reduction . . . .
10.2 Randomized PCA . . . .
10.3 Nonnegative Matrix Factorization . . . .
10.4 Clustering . . . .
10.5 The Johnson–Lindenstrauss Theorem . . . .
10.6 Embeddings . . . .
10.7 Word Representations . . . .
10.8 Compressed Sensing . . . .
Exercises . . . .

Linear Models
11.1 Binary Classification . . . .
11.2 Logistic and Softmax Regression . . . .
11.3 Algorithms for Logistic Regression . . . .
11.4 Maximum Entropy Classifiers . . . .
11.5 Hard Margin Support Vector Machines . . . .
11.6 Soft Margin Support Vector Machines . . . .
11.7 Kernels and Their Applications . . . .
Exercises . . . .

Decision Trees
12.1 Classification Trees . . . .
12.2 Regression Trees . . . .
12.3 Ensembles: Condorcet's Theorem . . . .
12.4 Random Forests . . . .
12.5 Gradient Boosted Trees . . . .
12.6 Differentiable Trees . . . .
12.7 Convergence and Expressiveness of Trees . . . .
Exercises . . . .

Neural Networks
13.1 Basic Neural Network Structure . . . .
13.2 Back Propagation and Automatic Differentiation . . . .
13.3 Regularization Methods . . . .
Exercises . . . .

Deep Learning
14.1 Convolutional Neural Networks . . . . . .
14.2 Residual Networks . . . . . .
14.3 Long and Short Term Memory Networks . . . . . .
14.4 Gated Rectified Units . . . . . .
14.5 Sequence to Sequence . . . . . .

VI Data Generation, Online Learning, and Reinforcement Learning

Data Augmentation and Generation
15.1 Augmentation . . . . . .
15.2 Generation by White-box Models . . . . . .
15.3 Generative Adversarial Networks . . . . . .

Reinforcement Learning
16.1 Review of MDPs . . . . . .
16.2 Q-learning . . . . . .
16.3 Deep Reinforcement Learning . . . . . .