CAMBRIDGE

Cambridge University Press
978-1-316-51888-5 — Mathematical Pictures at a Data Science Exhibition
Simon Foucart
Table of Contents
More Information

Contents

Preface *page* xiii
Notation xvii

PART ONE MACHINE LEARNING 1
Executive Summary 3

1 Rudiments of Statistical Learning Theory 4
1.1 True and Empirical Risks 4
1.2 PAC-Learnability 6
1.3 Validation 8
Exercises 9

2 Vapnik–Chervonenkis Dimension 10
2.1 Definitions 10
2.2 Examples 11
2.3 Sauer Lemma 13
Exercises 15

3 Learnability for Binary Classification 16
3.1 Uniform Convergence Property 16
3.2 Finite VC-Dimension Implies PAC-Learnability 17
3.3 No-Free-Lunch Theorem 20
Exercises 22

4 Support Vector Machines 23
4.1 Linear Separability 23
4.2 Hard and Soft SVM 25
4.3 Kernel Trick 27
Exercises 29

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-316-51888-5 — Mathematical Pictures at a Data Science Exhibition
Simon Foucart
Table of Contents
More Information

5 Reproducing Kernel Hilbert Spaces 31
5.1 Abstract Definition 31
5.2 Moore–Aronszajn Theorem 33
5.3 Mercer Theorem 37
Exercises 39

6 Regression and Regularization 41
6.1 Empirical Risk Minimization 41
6.2 Regularization 43
6.3 Classification via Regression 44
Exercises 45

7 Clustering 47
7.1 Single-Linkage Clustering 48
7.2 Center-Based Clustering 49
7.3 Spectral Clustering 51
Exercises 54

8 Dimension Reduction 56
8.1 Principal Component Analysis 56
8.2 Johnson–Lindenstrauss Lemma 58
8.3 Locally Linear Embedding 61
Exercises 64

PART TWO OPTIMAL RECOVERY 65
Executive Summary 67

9 Foundational Results of Optimal Recovery 68
9.1 Models, Errors, and Optimality 69
9.2 Linearity of Optimal Recovery Maps 71
9.3 An Extremal Property of Splines 73
Exercises 75

10 Approximability Models 76
10.1 The Model Set 76
10.2 Optimality in a Hilbert Setting 78
10.3 Optimality for Linear Functionals 82
Exercises 85

11 Ideal Selection of Observation Schemes 86
11.1 Hilbert Setting 86
11.2 Integration of Lipschitz Functions 88
11.3 Adaptivity Does Not Help Much 91
Exercises 92

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-316-51888-5 — Mathematical Pictures at a Data Science Exhibition
Simon Foucart
Table of Contents
More Information

12 Curse of Dimensionality 94
12.1 Notions of Tractability 94
12.2 Integration of Trigonometric Polynomials 96
12.3 Integration in Weighted Sobolev Spaces 98
Exercises 100

13 Quasi-Monte Carlo Integration 102
13.1 Variation and Discrepancy 103
13.2 Koksma–Hlawka Inequality 105
13.3 Low-Discrepancy Sets 107
Exercises 111

PART THREE COMPRESSIVE SENSING 113
Executive Summary 115

14 Sparse Recovery from Linear Observations 116
14.1 ℓ_0-Minimization 117
14.2 ℓ_1-Minimization 118
14.3 ℓ_1-Restricted Isometry Property 119
Exercises 122

15 The Complexity of Sparse Recovery 123
15.1 Limitations Imposed by Stability and Robustness 123
15.2 Gelfand Width of the ℓ_1-Ball 127
15.3 Irrelevance of ℓ_2-Stability 130
Exercises 130

16 Low-Rank Recovery from Linear Observations 132
16.1 Nuclear Norm Minimization 132
16.2 ℓ_1-Rank Restricted Isometry Property 134
16.3 Semidefinite Programming Formulation 136
Exercises 137

17 Sparse Recovery from One-Bit Observations 139
17.1 Estimating the Direction via Hard Thresholding 139
17.2 Estimating the Direction via Linear Programming 141
17.3 Estimating Both the Direction and the Magnitude 145
Exercises 147

18 Group Testing 149
18.1 Properties of the Test Matrix 149
18.2 Satisfying the Separabity Condition 151
18.3 Recovery via a Linear Feasibility Program 154
Exercises 155

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-316-51888-5 — Mathematical Pictures at a Data Science Exhibition
Simon Foucart
Table of Contents
More Information

PART FOUR OPTIMIZATION 157
Executive Summary 159

19 Basic Convex Optimization 160
19.1 Gradient Descent for Unconstrained Convex Programs 160
19.2 Rates of Convergence for Steepest Descent 162
19.3 Stochastic Gradient Descent 165
Exercises 168

20 Snippets of Linear Programming 169
20.1 Maximizers of a Convex Function 169
20.2 The Simplex Algorithm 171
20.3 Illustrative Linear Programs 173
Exercises 175

21 Duality Theory and Practice 177
21.1 Duality in Linear Programming 178
21.2 Examples of Robust Optimization 181
21.3 Duality in Conic Programming 183
Exercises 185

22 Semidefinite Programming in Action 187
22.1 Schur Complement 188
22.2 Sum-of-Squares Technique 189
22.3 Method of Moments 191
Exercises 193

23 Instances of Nonconvex Optimization 195
23.1 Quadratically Constrained Quadratic Programs 195
23.2 Dynamic Programming 198
23.3 Projected Gradient Descent 200
Exercises 203

PART FIVE NEURAL NETWORKS 205
Executive Summary 207

24 First Encounter with ReLU Networks 208
24.1 Some Terminology 209
24.2 Shallow ReLU Networks and CPwL Functions 210
24.3 Deep ReLU Networks and CPwL Functions 213
Exercises 215

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-316-51888-5 — Mathematical Pictures at a Data Science Exhibition
Simon Foucart
Table of Contents
More Information

25 Expressiveness of Shallow Networks 216
25.1 Activation Functions and Universal Approximation 216
25.2 Approximation Rate with ReLU: Upper Bound 219
25.3 Approximation Rate with ReLU: Lower Bound 221
Exercises 224

26 Various Advantages of Depth 226
26.1 Omnipotent Activation Functions 226
26.2 Compact Supports 229
26.3 Approximation Power 230
Exercises 237

27 Tidbits on Neural Network Training 239
27.1 Backpropagation 239
27.2 Overparametrized Empirical-Risk Landscapes 241
27.3 Convolutional Neural Networks 245
Exercises 246

APPENDICES 247

***Appendix A* High-Dimensional Geometry** 249
A.1 Volumes 250
A.2 Covering and Packing Numbers 252
A.3 Projected Cross-Polytope 255
Exercises 257

***Appendix B* Probability Theory** 259
B.1 Tails and Moment Generating Functions 259
B.2 Concentration Inequalities 262
B.3 Restricted Isometry Properties 267
Exercises 272

***Appendix C* Functional Analysis** 274
C.1 Completeness 274
C.2 Convexity 277
C.3 Extreme Points 281
Exercises 283

***Appendix D* Matrix Analysis** 285
D.1 Eigenvalues of Self-Adjoint Matrices 285
D.2 Singular Values 288
D.3 Matrix Norms 291
Exercises 295

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-316-51888-5 — Mathematical Pictures at a Data Science Exhibition
Simon Foucart
Table of Contents
More Information

Appendix E **Approximation Theory** 297
E.1 Classic Uniform Approximation Theorems 297
E.2 Riesz–Fejér and Carathéodory–Toeplitz Theorems 304
E.3 Kolmogorov Superposition Theorem 307
Exercises 310

References 311
Index 315

© in this web service Cambridge University Press www.cambridge.org