

“Non-asymptotic, high-dimensional theory is critical for modern statistics and machine learning. This book is unique in providing a crystal clear, complete, and unified treatment of the area. With topics ranging from concentration of measure to graphical models, the author weaves together probability theory and its applications to statistics. Ideal for graduate students and researchers. This will surely be the standard reference on the topic for many years.”

— Larry Wasserman, *Carnegie Mellon University*

“Martin Wainwright brings his large box of analytical power tools to bear on the problems of the day—the analysis of models for wide data. A broad knowledge of this new area combines with his powerful analytical skills to deliver this impressive and intimidating work—bound to be an essential reference for all the brave souls that try their hand.”

— Trevor Hastie, *Stanford University*

“This book provides an excellent treatment of perhaps the fastest growing area within high-dimensional theoretical statistics—non-asymptotic theory that seeks to provide probabilistic bounds on estimators as a function of sample size and dimension. It offers the most thorough, clear, and engaging coverage of this area to date, and is thus poised to become the definitive reference and textbook on this topic.”

— Genevera Allen, *Rice University*

“Statistical theory and practice have undergone a renaissance in the past two decades, with intensive study of high-dimensional data analysis. No researcher has deepened our understanding of high-dimensional statistics more than Martin Wainwright. This book brings the signature clarity and incisiveness of his published research into book form. It will be a fantastic resource for both beginning students and seasoned researchers, as the field continues to make exciting breakthroughs.”

— John Lafferty, *Yale University*

“This is an outstanding book on high-dimensional statistics, written by a creative and celebrated researcher in the field. It gives comprehensive treatments of many important topics in statistical machine learning and, furthermore, is self-contained, from introductory material to the most up-to-date results on various research frontiers. This book is a must-read for those who wish to learn and to develop modern statistical machine theory, methods and algorithms.”

— Jianqing Fan, *Princeton University*

“This book provides an in-depth mathematical treatment and methodological intuition for high-dimensional statistics. The main technical tools from probability theory are carefully developed and the construction and analysis of statistical methods and algorithms for high-dimensional problems are presented in an outstandingly clear way. Martin Wainwright has written a truly exceptional, inspiring, and beautiful masterpiece!”

— Peter Bühlmann, *ETH Zurich*

“This new book by Martin Wainwright covers modern topics in high-dimensional statistical inference, and focuses primarily on explicit non-asymptotic results related to sparsity and non-parametric estimation. This is a must-read for all graduate students in mathematical statistics and theoretical machine learning, both for the breadth of recent advances it covers and the depth of results which are presented. The exposition is outstandingly clear, starting from the first introductory chapters on the necessary probabilistic tools. Then, the book covers state-of-the-art advances in high-dimensional statistics, with always a clever choice of results which have the perfect mix of significance and mathematical depth.”

— Francis Bach, *INRIA Paris*

“Wainwright’s book on those parts of probability theory and mathematical statistics critical to understanding of the new phenomena encountered in high dimensions is marked by the clarity of its presentation and the depth to which it travels. In every chapter he starts with intuitive examples and simulations which are systematically developed either into powerful mathematical tools or complete answers to fundamental questions of inference. It is not easy, but elegant and rewarding whether read systematically or dipped into as a reference.”

— Peter Bickel, *UC Berkeley*

High-Dimensional Statistics

Recent years have witnessed an explosion in the volume and variety of data collected in all scientific disciplines and industrial settings. Such massive data sets present a number of challenges to researchers in statistics and machine learning. This book provides a self-contained introduction to the area of high-dimensional statistics, aimed at the first-year graduate level. It includes chapters that are focused on core methodology and theory—including tail bounds, concentration inequalities, uniform laws and empirical process, and random matrices—as well as chapters devoted to in-depth exploration of particular model classes—including sparse linear models, matrix models with rank constraints, graphical models, and various types of non-parametric models.

With hundreds of worked examples and exercises, this text is intended both for courses and for self-study by graduate students and researchers in statistics, machine learning, and related fields who must understand, apply, and adapt modern statistical methods suited to large-scale data.

MARTIN J. WAINWRIGHT is a Chancellor's Professor at the University of California, Berkeley, with a joint appointment between the Department of Statistics and the Department of Electrical Engineering and Computer Sciences. His research lies at the nexus of statistics, machine learning, optimization, and information theory, and he has published widely in all of these disciplines. He has written two other books, one on graphical models together with Michael I. Jordan, and one on sparse learning together with Trevor Hastie and Robert Tibshirani. Among other awards, he has received the COPSS Presidents' Award, has been a Medallion Lecturer and Blackwell Lecturer for the Institute of Mathematical Statistics, and has received Best Paper Awards from the NIPS, ICML, and UAI conferences, as well as from the IEEE Information Theory Society.

CAMBRIDGE SERIES IN STATISTICAL AND
 PROBABILISTIC MATHEMATICS

Editorial Board

- Z. Ghahramani (Department of Engineering, University of Cambridge)
 R. Gill (Mathematical Institute, Leiden University)
 F. P. Kelly (Department of Pure Mathematics and Mathematical Statistics,
 University of Cambridge)
 B. D. Ripley (Department of Statistics, University of Oxford)
 S. Ross (Department of Industrial and Systems Engineering,
 University of Southern California)
 M. Stein (Department of Statistics, University of Chicago)

This series of high-quality upper-division textbooks and expository monographs covers all aspects of stochastic applicable mathematics. The topics range from pure and applied statistics to probability theory, operations research, optimization, and mathematical programming. The books contain clear presentations of new developments in the field and also of the state of the art in classical methods. While emphasizing rigorous treatment of theoretical methods, the books also contain applications and discussions of new techniques made possible by advances in computational practice.

A complete list of books in the series can be found at www.cambridge.org/statistics.
 Recent titles include the following:

23. *Applied Asymptotics*, by A. R. Brazzale, A. C. Davison and N. Reid
24. *Random Networks for Communication*, by Massimo Franceschetti and Ronald Meester
25. *Design of Comparative Experiments*, by R. A. Bailey
26. *Symmetry Studies*, by Marlos A. G. Viana
27. *Model Selection and Model Averaging*, by Gerda Claeskens and Nils Lid Hjort
28. *Bayesian Nonparametrics*, edited by Nils Lid Hjort *et al.*
29. *From Finite Sample to Asymptotic Methods in Statistics*, by Pranab K. Sen, Julio M. Singer and Antonio C. Pedrosa de Lima
30. *Brownian Motion*, by Peter Mörters and Yuval Peres
31. *Probability: Theory and Examples (Fourth Edition)*, by Rick Durrett
33. *Stochastic Processes*, by Richard F. Bass
34. *Regression for Categorical Data*, by Gerhard Tutz
35. *Exercises in Probability (Second Edition)*, by Loïc Chaumont and Marc Yor
36. *Statistical Principles for the Design of Experiments*, by R. Mead, S. G. Gilmour and A. Mead
37. *Quantum Stochastics*, by Mou-Hsiung Chang
38. *Nonparametric Estimation under Shape Constraints*, by Piet Groeneboom and Geurt Jongbloed
39. *Large Sample Covariance Matrices and High-Dimensional Data Analysis*, by Jianfeng Yao, Shurong Zheng and Zhidong Bai
40. *Mathematical Foundations of Infinite-Dimensional Statistical Models*, by Evarist Giné and Richard Nickl
41. *Confidence, Likelihood, Probability*, by Tore Schweder and Nils Lid Hjort
42. *Probability on Trees and Networks*, by Russell Lyons and Yuval Peres
43. *Random Graphs and Complex Networks (Volume 1)*, by Remco van der Hofstad
44. *Fundamentals of Nonparametric Bayesian Inference*, by Subhashis Ghosal and Aad van der Vaart
45. *Long-Range Dependence and Self-Similarity*, by Vlasos Pipiras and Murad S. Taqqu
46. *Predictive Statistics*, by Bertrand S. Clarke and Jennifer L. Clarke
47. *High-Dimensional Probability*, by Roman Vershynin
48. *High-Dimensional Statistics*, by Martin J. Wainwright
49. *Probability: Theory and Examples (Fifth Edition)*, by Rick Durrett

High-Dimensional Statistics

A Non-Asymptotic Viewpoint

Martin J. Wainwright
University of California, Berkeley



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom
One Liberty Plaza, 20th Floor, New York, NY 10006, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India
79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781108498029

DOI: 10.1017/9781108627771

© Martin J. Wainwright 2019

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2019

Printed in the United Kingdom by TJ International Ltd. Padstow Cornwall

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data

Names: Wainwright, Martin (Martin J.), author.

Title: High-dimensional statistics : a non-asymptotic viewpoint / Martin J. Wainwright (University of California, Berkeley).

Description: Cambridge ; New York, NY : Cambridge University Press, 2019. |

Series: Cambridge series in statistical and probabilistic mathematics ; 48 |

Includes bibliographical references and indexes.

Identifiers: LCCN 2018043475 | ISBN 9781108498029 (hardback)

Subjects: LCSH: Mathematical statistics—Textbooks. | Big data.

Classification: LCC QA276.18 .W35 2019 | DDC 519.5—dc23

LC record available at <https://lcn.loc.gov/2018043475>

ISBN 978-1-108-49802-9 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

List of chapters

<i>List of illustrations</i>	xv
1 Introduction	1
2 Basic tail and concentration bounds	21
3 Concentration of measure	58
4 Uniform laws of large numbers	98
5 Metric entropy and its uses	121
6 Random matrices and covariance estimation	159
7 Sparse linear models in high dimensions	194
8 Principal component analysis in high dimensions	236
9 Decomposability and restricted strong convexity	259
10 Matrix estimation with rank constraints	312
11 Graphical models for high-dimensional data	347
12 Reproducing kernel Hilbert spaces	383
13 Nonparametric least squares	416
14 Localization and uniform laws	453
15 Minimax lower bounds	485
<i>References</i>	524
<i>Subject index</i>	540
<i>Author index</i>	548

Cambridge University Press
978-1-108-49802-9 — High-Dimensional Statistics
Martin J. Wainwright
Frontmatter
[More Information](#)

Contents

<i>List of illustrations</i>	xv
1 Introduction	1
1.1 Classical versus high-dimensional theory	1
1.2 What can go wrong in high dimensions?	2
1.2.1 Linear discriminant analysis	2
1.2.2 Covariance estimation	5
1.2.3 Nonparametric regression	7
1.3 What can help us in high dimensions?	9
1.3.1 Sparsity in vectors	10
1.3.2 Structure in covariance matrices	11
1.3.3 Structured forms of regression	12
1.4 What is the non-asymptotic viewpoint?	14
1.5 Overview of the book	15
1.5.1 Chapter structure and synopses	15
1.5.2 Recommended background	17
1.5.3 Teaching possibilities and a flow diagram	17
1.6 Bibliographic details and background	19
2 Basic tail and concentration bounds	21
2.1 Classical bounds	21
2.1.1 From Markov to Chernoff	21
2.1.2 Sub-Gaussian variables and Hoeffding bounds	22
2.1.3 Sub-exponential variables and Bernstein bounds	25
2.1.4 Some one-sided results	31
2.2 Martingale-based methods	32
2.2.1 Background	33
2.2.2 Concentration bounds for martingale difference sequences	35
2.3 Lipschitz functions of Gaussian variables	40
2.4 Appendix A: Equivalent versions of sub-Gaussian variables	45
2.5 Appendix B: Equivalent versions of sub-exponential variables	48
2.6 Bibliographic details and background	49
2.7 Exercises	50
3 Concentration of measure	58
3.1 Concentration by entropic techniques	58
3.1.1 Entropy and its properties	58
3.1.2 Herbst argument and its extensions	60

x	<i>Contents</i>	
	3.1.3 Separately convex functions and the entropic method	62
	3.1.4 Tensorization and separately convex functions	64
3.2	A geometric perspective on concentration	67
	3.2.1 Concentration functions	67
	3.2.2 Connection to Lipschitz functions	70
	3.2.3 From geometry to concentration	72
3.3	Wasserstein distances and information inequalities	76
	3.3.1 Wasserstein distances	76
	3.3.2 Transportation cost and concentration inequalities	78
	3.3.3 Tensorization for transportation cost	80
	3.3.4 Transportation cost inequalities for Markov chains	82
	3.3.5 Asymmetric coupling cost	84
3.4	Tail bounds for empirical processes	87
	3.4.1 A functional Hoeffding inequality	87
	3.4.2 A functional Bernstein inequality	89
3.5	Bibliographic details and background	91
3.6	Exercises	92
4	Uniform laws of large numbers	98
4.1	Motivation	98
	4.1.1 Uniform convergence of cumulative distribution functions	98
	4.1.2 Uniform laws for more general function classes	101
4.2	A uniform law via Rademacher complexity	104
	4.2.1 Necessary conditions with Rademacher complexity	107
4.3	Upper bounds on the Rademacher complexity	109
	4.3.1 Classes with polynomial discrimination	109
	4.3.2 Vapnik–Chervonenkis dimension	111
	4.3.3 Controlling the VC dimension	115
4.4	Bibliographic details and background	117
4.5	Exercises	117
5	Metric entropy and its uses	121
5.1	Covering and packing	121
5.2	Gaussian and Rademacher complexity	132
5.3	Metric entropy and sub-Gaussian processes	134
	5.3.1 Upper bound by one-step discretization	135
	5.3.2 Some examples of discretization bounds	137
	5.3.3 Chaining and Dudley’s entropy integral	139
5.4	Some Gaussian comparison inequalities	143
	5.4.1 A general comparison result	143
	5.4.2 Slepian and Sudakov–Fernique inequalities	145
	5.4.3 Gaussian contraction inequality	146
5.5	Sudakov’s lower bound	148
5.6	Chaining and Orlicz processes	150
5.7	Bibliographic details and background	153
5.8	Exercises	154
6	Random matrices and covariance estimation	159
6.1	Some preliminaries	159

Contents

xi

	6.1.1 Notation and basic facts	159
	6.1.2 Set-up of covariance estimation	160
6.2	Wishart matrices and their behavior	161
6.3	Covariance matrices from sub-Gaussian ensembles	165
6.4	Bounds for general matrices	168
	6.4.1 Background on matrix analysis	168
	6.4.2 Tail conditions for matrices	169
	6.4.3 Matrix Chernoff approach and independent decompositions	172
	6.4.4 Upper tail bounds for random matrices	174
	6.4.5 Consequences for covariance matrices	179
6.5	Bounds for structured covariance matrices	180
	6.5.1 Unknown sparsity and thresholding	180
	6.5.2 Approximate sparsity	183
6.6	Appendix: Proof of Theorem 6.1	185
6.7	Bibliographic details and background	188
6.8	Exercises	189
7	Sparse linear models in high dimensions	194
7.1	Problem formulation and applications	194
	7.1.1 Different sparsity models	194
	7.1.2 Applications of sparse linear models	196
7.2	Recovery in the noiseless setting	199
	7.2.1 ℓ_1 -based relaxation	200
	7.2.2 Exact recovery and restricted nullspace	200
	7.2.3 Sufficient conditions for restricted nullspace	202
7.3	Estimation in noisy settings	206
	7.3.1 Restricted eigenvalue condition	207
	7.3.2 Bounds on ℓ_2 -error for hard sparse models	209
	7.3.3 Restricted nullspace and eigenvalues for random designs	213
7.4	Bounds on prediction error	216
7.5	Variable or subset selection	218
	7.5.1 Variable selection consistency for the Lasso	219
	7.5.2 Proof of Theorem 7.21	222
7.6	Appendix: Proof of Theorem 7.16	224
7.7	Bibliographic details and background	227
7.8	Exercises	229
8	Principal component analysis in high dimensions	236
8.1	Principal components and dimension reduction	236
	8.1.1 Interpretations and uses of PCA	237
	8.1.2 Perturbations of eigenvalues and eigenspaces	241
8.2	Bounds for generic eigenvectors	242
	8.2.1 A general deterministic result	242
	8.2.2 Consequences for a spiked ensemble	245
8.3	Sparse principal component analysis	248
	8.3.1 A general deterministic result	249
	8.3.2 Consequences for the spiked model with sparsity	252
8.4	Bibliographic details and background	255
8.5	Exercises	256

9	Decomposability and restricted strong convexity	259
9.1	A general regularized M -estimator	259
9.2	Decomposable regularizers and their utility	269
	9.2.1 Definition and some examples	269
	9.2.2 A key consequence of decomposability	272
9.3	Restricted curvature conditions	276
	9.3.1 Restricted strong convexity	277
9.4	Some general theorems	279
	9.4.1 Guarantees under restricted strong convexity	280
	9.4.2 Bounds under Φ^* -curvature	284
9.5	Bounds for sparse vector regression	286
	9.5.1 Generalized linear models with sparsity	286
	9.5.2 Bounds under restricted strong convexity	287
	9.5.3 Bounds under ℓ_∞ -curvature conditions	288
9.6	Bounds for group-structured sparsity	290
9.7	Bounds for overlapping decomposition-based norms	293
9.8	Techniques for proving restricted strong convexity	297
	9.8.1 Lipschitz cost functions and Rademacher complexity	298
	9.8.2 A one-sided bound via truncation	302
9.9	Appendix: Star-shaped property	306
9.10	Bibliographic details and background	306
9.11	Exercises	307
10	Matrix estimation with rank constraints	312
10.1	Matrix regression and applications	312
10.2	Analysis of nuclear norm regularization	317
	10.2.1 Decomposability and subspaces	317
	10.2.2 Restricted strong convexity and error bounds	319
	10.2.3 Bounds under operator norm curvature	320
10.3	Matrix compressed sensing	321
10.4	Bounds for phase retrieval	326
10.5	Multivariate regression with low-rank constraints	329
10.6	Matrix completion	330
10.7	Additive matrix decompositions	337
10.8	Bibliographic details and background	341
10.9	Exercises	343
11	Graphical models for high-dimensional data	347
11.1	Some basics	347
	11.1.1 Factorization	347
	11.1.2 Conditional independence	350
	11.1.3 Hammersley–Clifford equivalence	351
	11.1.4 Estimation of graphical models	352
11.2	Estimation of Gaussian graphical models	352
	11.2.1 Graphical Lasso: ℓ_1 -regularized maximum likelihood	353
	11.2.2 Neighborhood-based methods	359
11.3	Graphical models in exponential form	365
	11.3.1 A general form of neighborhood regression	366
	11.3.2 Graph selection for Ising models	367

Contents

xiii

11.4	Graphs with corrupted or hidden variables	368
11.4.1	Gaussian graph estimation with corrupted data	368
11.4.2	Gaussian graph selection with hidden variables	373
11.5	Bibliographic details and background	376
11.6	Exercises	378
12	Reproducing kernel Hilbert spaces	383
12.1	Basics of Hilbert spaces	383
12.2	Reproducing kernel Hilbert spaces	385
12.2.1	Positive semidefinite kernel functions	386
12.2.2	Feature maps in $\ell^2(\mathbb{N})$	387
12.2.3	Constructing an RKHS from a kernel	388
12.2.4	A more abstract viewpoint and further examples	390
12.3	Mercer's theorem and its consequences	394
12.4	Operations on reproducing kernel Hilbert spaces	400
12.4.1	Sums of reproducing kernels	400
12.4.2	Tensor products	403
12.5	Interpolation and fitting	405
12.5.1	Function interpolation	405
12.5.2	Fitting via kernel ridge regression	407
12.6	Distances between probability measures	409
12.7	Bibliographic details and background	411
12.8	Exercises	412
13	Nonparametric least squares	416
13.1	Problem set-up	416
13.1.1	Different measures of quality	416
13.1.2	Estimation via constrained least squares	417
13.1.3	Some examples	418
13.2	Bounding the prediction error	420
13.2.1	Bounds via metric entropy	425
13.2.2	Bounds for high-dimensional parametric problems	427
13.2.3	Bounds for nonparametric problems	429
13.2.4	Proof of Theorem 13.5	430
13.3	Oracle inequalities	432
13.3.1	Some examples of oracle inequalities	434
13.3.2	Proof of Theorem 13.13	437
13.4	Regularized estimators	439
13.4.1	Oracle inequalities for regularized estimators	439
13.4.2	Consequences for kernel ridge regression	439
13.4.3	Proof of Corollary 13.18	443
13.4.4	Proof of Theorem 13.17	444
13.5	Bibliographic details and background	448
13.6	Exercises	449
14	Localization and uniform laws	453
14.1	Population and empirical L^2 -norms	453
14.1.1	A uniform law with localization	454
14.1.2	Specialization to kernel classes	458

xiv	<i>Contents</i>	
	14.1.3 Proof of Theorem 14.1	460
14.2	A one-sided uniform law	462
	14.2.1 Consequences for nonparametric least squares	466
	14.2.2 Proof of Theorem 14.12	468
14.3	A uniform law for Lipschitz cost functions	469
	14.3.1 General prediction problems	469
	14.3.2 Uniform law for Lipschitz cost functions	472
14.4	Some consequences for nonparametric density estimation	475
	14.4.1 Density estimation via the nonparametric maximum likelihood estimate	475
	14.4.2 Density estimation via projections	477
14.5	Appendix: Population and empirical Rademacher complexities	480
14.6	Bibliographic details and background	481
14.7	Exercises	482
15	Minimax lower bounds	485
15.1	Basic framework	485
	15.1.1 Minimax risks	486
	15.1.2 From estimation to testing	487
	15.1.3 Some divergence measures	489
15.2	Binary testing and Le Cam's method	491
	15.2.1 Bayes error and total variation distance	491
	15.2.2 Le Cam's convex hull method	497
15.3	Fano's method	500
	15.3.1 Kullback–Leibler divergence and mutual information	501
	15.3.2 Fano lower bound on minimax risk	501
	15.3.3 Bounds based on local packings	503
	15.3.4 Local packings with Gaussian entropy bounds	506
	15.3.5 Yang–Barron version of Fano's method	512
15.4	Appendix: Basic background in information theory	515
15.5	Bibliographic details and background	518
15.6	Exercises	519
	<i>References</i>	524
	<i>Subject index</i>	540
	<i>Author index</i>	548

Illustrations

1.1	High-dimensional phenomena in linear discriminant analysis	4
1.2	Illustration of the Marčenko–Pastur law	6
1.3	Geometry of sampling in a high-dimensional cube	8
1.4	Hard- and soft-thresholding estimators for sparse vectors	11
1.5	Hard-thresholded estimators of a sparse covariance matrix	12
1.6	Flow diagram of chapters and their interdependence	18
3.1	Concentration of measure on the sphere S^{n-1}	69
4.1	Population and empirical cumulative distribution functions	99
5.1	Packings and coverings of a set	122
5.2	Metric entropy of Lipschitz function class	128
5.3	Illustration of Kolmogorov chaining procedure	141
6.1	Covariance estimation and graphical structure	182
7.1	Weak models of sparsity using ℓ_q -balls	195
7.2	Geometry of the tangent cone and restricted nullspace property	201
7.3	Phase transitions for basis pursuit linear program	206
7.4	Basis pursuit performance for correlated designs	207
7.5	From strong convexity to bounds on estimation error	208
7.6	Failure of strong convexity in high dimensions	209
7.7	Phase transitions for Lasso variable selection	222
8.1	Principal component analysis for low-rank matrix approximation	238
8.2	PCA and eigenfaces for face recognition	239
8.3	Estimating Gaussian mixture models via PCA	240
8.4	Performance of PCA as a function of signal-to-noise ratio	246
8.5	Sparse versions of eigenfaces	249
9.1	Unit balls of various decomposable norms	261
9.2	Standard versus overlapping group Lasso	263
9.3	Group Lasso with disjoint and overlapping groups	264
9.4	Multivariate regression models and matrix recovery	266
9.5	Geometry of the nuclear norm ball	268
9.6	Geometry of decomposable regularizers	270
9.7	Illustration of the cone $\mathbb{C}_{\theta^*}(\mathbb{M}, \mathbb{M}^\perp)$	274
9.8	Cost functions and convexity in high dimensions	276
9.9	Illustration of proof technique in Lemma 9.21	282
9.10	Decomposable norms for additive matrix decomposition	293
10.1	Low-rank matrix completion and recommender systems	314
10.2	Consistency of estimator based on nuclear norm regularization	324
11.1	Maximal cliques and vertex cutsets in graphs	348
11.2	Hammersley–Clifford theorem for Gaussian graphical models	353

12.1	Exact interpolation using RKHS methods	405
12.2	Kernel ridge regression with polynomial and Sobolev kernels	408
13.1	Convex and monotonic regression estimates	421
13.2	Localized complexity measures and the critical radius	422
13.3	Star-shaped sets and their properties	424
13.4	Trade-offs between approximation and estimation error	435
13.5	Eigenspectra of empirical kernel matrices	441
14.1	Plots of the behavior of the histogram density estimate	478
14.2	Plots of the behavior of the orthogonal series density estimate	479
15.1	Reduction from estimation to testing	488
15.2	Lower bounds for Lipschitz classes via hat functions	495
15.3	Perturbed densities for Le Cam lower bound	497
15.4	Construction of mixture distribution for lower bounds	500

Acknowledgements

This book would not exist without the help of many people in my life. I thank my parents, John and Patricia, for their nurture and support over the years, and my family, Haruko, Hana, Mina, and Kento for bringing daily joy. In developing my own understanding of and perspective on the broad area of high-dimensional statistics, I have been fortunate to interact with and learn from many wonderful colleagues, both here at Berkeley and elsewhere. For valuable discussions, insights and feedback, I would like to thank, among others, Bryon Aragam, Alex d'Asprémont, Francis Bach, Peter Bickel, Peter Bühlmann, Tony Cai, Emmanuel Candès, Constantine Caramanis, David Donoho, Nouredine El Karoui, Jianqing Fan, Aditya Guntuboyina, Trevor Hastie, Iain Johnstone, Michael Jordan, John Lafferty, Eliza Levina, Zongming Ma, Nicolai Meinshausen, Andrea Montanari, Axel Munk, Richard Nickl, Eric Price, Philippe Rigollet, Alessandro Rinaldo, Richard Samworth, Robert Tibshirani, Ryan Tibshirani, Alexander Tsybakov, Sara Van de Geer, Larry Wasserman, Frank Werner, Bin Yu, Ming Yuan, and Harry Zhou. The Statistics and EECS departments, staff, faculty and students at UC Berkeley have provided a wonderful intellectual environment for this work; I also thank the Statistics Group at ETH Zurich, as well as the Laboratory for Information and Decision Sciences (LIDS) at MIT for their support during my visiting professor stays.

Over the years, I have also had the pleasure of working with many outstanding students and postdoctoral associates. Our discussions and research together as well as their feedback have been instrumental in shaping and improving this book. I would like to thank my current and former students, postdocs and visting fellows—many of whom are now current colleagues—including Alekh Agarwal, Arash Amini, Sivaraman Balakrishnan, Merle Behr, Joseph Bradley, Yudong Chen, Alex Dimakis, John Duchi, Reinhard Heckel, Nhat Ho, Johannes Lederer, Po-Ling Loh, Sahand Negahban, Xuanlong Nguyen, Nima Noorshams, Jonas Peters, Mert Pilanci, Aaditya Ramdas, Garvesh Raskutti, Pradeep Ravikumar, Prasad Santhanam, Nihar Shah, Yuting Wei, Fanny Yang, Yun Yang, and Yuchen Zhang. Yuting Wei put her artistic skills to excellent use in designing the cover for this book. Last but not least, I would like to thank Cambridge University Press, and in particular the support and encouragement throughout the process of Senior Editor Diana Gillooly. My apologies to anyone whose help I may have failed inadvertently to acknowledge.

Cambridge University Press
978-1-108-49802-9 — High-Dimensional Statistics
Martin J. Wainwright
Frontmatter
[More Information](#)
