

---

## Contents

<i>Foreword</i>	<i>page xi</i>
<i>Preface</i>	xiii
<b>Appetizer Using Probability to Cover a Geometric Set</b>	1
0.1 Notes	4
<b>1 Preliminaries on Random Variables</b>	5
1.1 Basic Quantities Associated with Random Variables	5
1.2 Some Classical Inequalities	6
1.3 Limit Theorems	8
1.4 Notes	10
<b>2 Concentration of Sums of Independent Random Variables</b>	11
2.1 Why Concentration Inequalities?	11
2.2 Hoeffding's Inequality	13
2.3 Chernoff's Inequality	17
2.4 Application: Degrees of Random Graphs	19
2.5 Sub-Gaussian Distributions	21
2.6 General Hoeffding and Khintchine Inequalities	26
2.7 Sub-Exponential Distributions	28
2.8 Bernstein's Inequality	33
2.9 Notes	36
<b>3 Random Vectors in High Dimensions</b>	38
3.1 Concentration of the Norm	39
3.2 Covariance Matrices and Principal Component Analysis	41
3.3 Examples of High-Dimensional Distributions	45
3.4 Sub-Gaussian Distributions in Higher Dimensions	51
3.5 Application: Grothendieck's Inequality and Semidefinite Programming	55
3.6 Application: Maximum Cut for Graphs	60
3.7 Kernel Trick, and Tightening of Grothendieck's Inequality	64
3.8 Notes	68
<b>4 Random Matrices</b>	70
4.1 Preliminaries on Matrices	70

4.2	Nets, Covering Numbers, and Packing Numbers	75
4.3	Application: Error Correcting Codes	79
4.4	Upper Bounds on Random Sub-Gaussian Matrices	83
4.5	Application: Community Detection in Networks	87
4.6	Two-Sided Bounds on Sub-Gaussian Matrices	91
4.7	Application: Covariance Estimation and Clustering	93
4.8	Notes	97
<b>5</b>	<b>Concentration Without Independence</b>	<b>98</b>
5.1	Concentration of Lipschitz Functions for the Sphere	98
5.2	Concentration for Other Metric Measure Spaces	104
5.3	Application: Johnson–Lindenstrauss Lemma	110
5.4	Matrix Bernstein Inequality	113
5.5	Application: Community Detection in Sparse Networks	121
5.6	Application: Covariance Estimation for General Distributions	122
5.7	Notes	125
<b>6</b>	<b>Quadratic Forms, Symmetrization, and Contraction</b>	<b>127</b>
6.1	Decoupling	127
6.2	Hanson–Wright Inequality	130
6.3	Concentration for Anisotropic Random Vectors	134
6.4	Symmetrization	136
6.5	Random Matrices With Non-I.I.D. Entries	138
6.6	Application: Matrix Completion	140
6.7	Contraction Principle	143
6.8	Notes	145
<b>7</b>	<b>Random Processes</b>	<b>147</b>
7.1	Basic Concepts and Examples	147
7.2	Slepian’s Inequality	151
7.3	Sharp Bounds on Gaussian Matrices	157
7.4	Sudakov’s Minoration Inequality	160
7.5	Gaussian Width	162
7.6	Stable Dimension, Stable Rank, and Gaussian Complexity	167
7.7	Random Projections of Sets	170
7.8	Notes	174
<b>8</b>	<b>Chaining</b>	<b>176</b>
8.1	Dudley’s Inequality	176
8.2	Application: Empirical Processes	183
8.3	VC Dimension	188
8.4	Application: Statistical Learning Theory	200
8.5	Generic Chaining	206
8.6	Talagrand’s Majorizing Measure and Comparison Theorems	210
8.7	Chevet’s Inequality	212
8.8	Notes	214

*Contents*

ix

<b>9</b>	<b>Deviations of Random Matrices and Geometric Consequences</b>	216
9.1	Matrix Deviation Inequality	216
9.2	Random Matrices, Random Projections, and Covariance Estimation	222
9.3	The Johnson–Lindenstrauss Lemma for Infinite Sets	225
9.4	Random Sections: $M^*$ Bound and Escape Theorem	227
9.5	Notes	231
<b>10</b>	<b>Sparse Recovery</b>	232
10.1	High-Dimensional Signal Recovery Problems	232
10.2	Signal Recovery Based on $M^*$ Bound	234
10.3	Recovery of Sparse Signals	236
10.4	Low-Rank Matrix Recovery	239
10.5	Exact Recovery and the Restricted Isometry Property	241
10.6	Lasso Algorithm for Sparse Regression	247
10.7	Notes	252
<b>11</b>	<b>Dvoretzky–Milman Theorem</b>	254
11.1	Deviations of Random Matrices with respect to General Norms	254
11.2	Johnson–Lindenstrauss Embeddings and Sharper Chevet Inequality	257
11.3	Dvoretzky–Milman Theorem	259
11.4	Notes	264
	<i>Hints for Exercises</i>	265
	<i>References</i>	272
	<i>Index</i>	281