

Vladimir Koltchinskii

Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems

École d'Été de Probabilités
de Saint-Flour XXXVIII-2008

Contents

1	Introduction	1
1.1	Abstract Empirical Risk Minimization	1
1.2	Excess Risk: Distribution Dependent Bounds	3
1.3	Rademacher Processes and Data Dependent Bounds on Excess Risk	5
1.4	Penalized Empirical Risk Minimization and Oracle Inequalities	7
1.5	Concrete Empirical Risk Minimization Problems	8
1.6	Sparse Recovery Problems	11
1.7	Recovering Low Rank Matrices	13
2	Empirical and Rademacher Processes	17
2.1	Symmetrization Inequalities	18
2.2	Comparison Inequalities for Rademacher Sums	19
2.3	Concentration Inequalities	23
2.4	Exponential Bounds for Sums of Independent Random Matrices	26
2.5	Further Comments	31
3	Bounding Expected Sup-Norms of Empirical and Rademacher Processes	33
3.1	Gaussian and Subgaussian Processes, Metric Entropies and Generic Chaining Complexities	33
3.2	Finite Classes of Functions	36
3.3	Shattering Numbers and VC-classes of Sets	38
3.4	Upper Entropy Bounds	42
3.5	Lower Entropy Bounds	48
3.6	Generic Chaining Complexities and Bounding Empirical Processes Indexed by \mathcal{F}^2	52
3.7	Function Classes in Hilbert Spaces	54
3.8	Further Comments	57

4	Excess Risk Bounds	59
4.1	Distribution Dependent Bounds and Ratio Bounds for Excess Risk	59
4.2	Rademacher Complexities and Data Dependent Bounds on Excess Risk	70
4.3	Further Comments	78
5	Examples of Excess Risk Bounds in Prediction Problems	81
5.1	Regression with Quadratic Loss	82
5.2	Empirical Risk Minimization with Convex Loss	88
5.3	Binary Classification Problems	91
5.4	Further Comments	97
6	Penalized Empirical Risk Minimization and Model Selection Problems	99
6.1	Penalization in Monotone Families \mathcal{F}_k	101
6.2	Penalization by Empirical Risk Minima	104
6.3	Linking Excess Risk and Variance in Penalization	110
6.4	Further Comments	118
7	Linear Programming in Sparse Recovery	121
7.1	Sparse Recovery and Neighborliness of Convex Polytopes	121
7.2	Geometric Properties of the Dictionary	123
7.2.1	Cones of Dominant Coordinates	123
7.2.2	Restricted Isometry Constants and Related Characteristics	126
7.2.3	Alignment Coefficients	130
7.3	Sparse Recovery in Noiseless Problems	133
7.4	The Dantzig Selector	140
7.5	Further Comments	149
8	Convex Penalization in Sparse Recovery	151
8.1	General Aspects of Convex Penalization	151
8.2	ℓ_1 -Penalization and Oracle Inequalities	157
8.3	Entropy Penalization and Sparse Recovery in Convex Hulls: Random Error Bounds	175
8.4	Approximation Error Bounds, Alignment and Oracle Inequalities	186
8.5	Further Comments	189
9	Low Rank Matrix Recovery: Nuclear Norm Penalization	191
9.1	Geometric Parameters of Low Rank Recovery and Other Preliminaries	191
9.2	Matrix Regression with Fixed Design	196
9.3	Matrix Regression with Subgaussian Design	203
9.4	Other Types of Design in Matrix Regression	219
9.5	Further Comments	233

A Auxiliary Material 235

 A.1 Orlicz Norms 235

 A.2 Classical Exponential Inequalities 236

 A.3 Properties of \sharp - and \flat -Transforms 237

 A.4 Some Notations and Facts in Linear Algebra 238

References 241

Index 249

Programme of the school 251

List of participants 253