Lecture Notes in Mathematics 2033

Editors:
J.-M. Morel, Cachan
B. Teissier, Paris

Subseries:
École d'Été de Probabilités de Saint-Flour

For further volumes:
http://www.springer.com/series/304

Saint-Flour Probability Summer School

The Saint-Flour volumes are reflections of the courses given at the Saint-Flour Probability Summer School. Founded in 1971, this school is organised every year by the Laboratoire de Mathématiques (CNRS and Université Blaise Pascal, Clermont-Ferrand, France). It is intended for PhD students, teachers and researchers who are interested in probability theory, statistics, and in their applications.

The duration of each school is 13 days (it was 17 days up to 2005), and up to 70 participants can attend it. The aim is to provide, in three high-level courses, a comprehensive study of some fields in probability theory or Statistics. The lecturers are chosen by an international scientific board. The participants themselves also have the opportunity to give short lectures about their research work.

Participants are lodged and work in the same building, a former seminary built in the 18th century in the city of Saint-Flour, at an altitude of 900 m. The pleasant surroundings facilitate scientific discussion and exchange.

The Saint-Flour Probability Summer School is supported by:

- Université Blaise Pascal
- Centre National de la Recherche Scientifique (C.N.R.S.)
- Ministère délégué à l'Enseignement supérieur et à la Recherche

For more information, see back pages of the book and
http://math.univ-bpclermont.fr/stflour/

Jean Picard
Summer School Chairman
Laboratoire de Mathématiques
Université Blaise Pascal
63177 Aubière Cedex
France

Vladimir Koltchinskii

Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems

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

Springer

Prof. Vladimir Koltchinskii
Georgia Institute of Technology
School of Mathematics
Atlanta
USA
vlad@math.gatech.edu

ISBN 978-3-642-22146-0 e-ISBN 978-3-642-22147-7
DOI 10.1007/978-3-642-22147-7
Springer Heidelberg Dordrecht London New York

Lecture Notes in Mathematics ISSN print edition: 0075-8434
ISSN electronic edition: 1617-9692

Library of Congress Control Number: 2011934366

Mathematics Subject Classification (2010): 62J99, 62H12, 60B20, 60G99

Cover design: deblik, Berlin

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The purpose of these lecture notes is to provide an introduction to the general theory of empirical risk minimization with an emphasis on excess risk bounds and oracle inequalities in penalized problems. In the recent years, there have been new developments in this area motivated by the study of new classes of methods in machine learning such as large margin classification methods (boosting, kernel machines). The main probabilistic tools involved in the analysis of these problems are concentration and deviation inequalities by Talagrand along with other methods of empirical processes theory (symmetrization inequalities, contraction inequality for Rademacher sums, entropy and generic chaining bounds). Sparse recovery based on ℓ_1-type penalization and low rank matrix recovery based on the nuclear norm penalization are other active areas of research, where the main problems can be stated in the framework of penalized empirical risk minimization, and concentration inequalities and empirical processes tools proved to be very useful.

My interest in empirical processes started in the late 1970s and early 1980s. It was largely influenced by the work of Vapnik and Chervonenkis on Glivenko–Cantelli problem and on empirical risk minimization in pattern recognition, and, especially, by the results of Dudley on uniform central limit theorems. Talagrand's concentration inequality proved in the 1990s was a major result with deep consequences in the theory of empirical processes and related areas of statistics, and it inspired many new approaches in analysis of empirical risk minimization problems.

Over the last years, the work of many people have had a profound impact on my own research and on my view of the subject of these notes. I was lucky to work together with several of them and to have numerous conversations and email exchanges with many others. I am especially thankful to Peter Bartlett, Lucien Birgé, Gilles Blanchard, Stephane Boucheron, Olivier Bousquet, Richard Dudley, Sara van de Geer, Evarist Giné, Gabor Lugosi, Pascal Massart, David Mason, Shahar Mendelson, Dmitry Panchenko, Alexandre Tsybakov, Aad van der Vaart, Jon Wellner and Joel Zinn.

I am thankful to the School of Mathematics, Georgia Institute of Technology and to the Department of Mathematics and Statistics, University of New Mexico where most of my work for the past years have taken place.

The research described in these notes has been supported in part by NSF grants DMS-0304861, MSPA-MPS-0624841, DMS-0906880 and CCF-0808863.

I was working on the initial draft while visiting the Isaac Newton Institute for Mathematical Sciences in Cambridge in 2008. I am thankful to the Institute for its hospitality.

I am also very thankful to Jean-Yves Audibert, Kai Ni and Jon Wellner who provided me with the lists of typos and inconsistencies in the first draft.

I am very grateful to all the students and colleagues who attended the lectures in 2008 and whose questions motivated many of the changes I have made since then.

I am especially thankful to Jean Picard for all his efforts that make The Saint-Flour School truly unique.

Atlanta *Vladimir Koltchinskii*
February 2011

Contents