

Title: High Dimensional Probability: On the Connections between Geometry and Statistics

Course Description:

The goal of this course is to explore the rather surprising connections between Statistical Learning Theory, Empirical Processes Theory and Asymptotic Geometric Analysis.

Those connections play an instrumental role in modern Data Science.

At the heart of these connections are two important ideas: that randomness can be used to expose "hidden structures"; and that structure often appears in "extremal situations"

(both these rather vague statements will be made clear during the course). As it happens, exposing hidden structures is the key to solutions of (high dimensional) statistical recovery problems.

The main topics that the course will cover are:

- 1) Empirical Processes Theory from a geometric viewpoint; Talagrand's generic chaining mechanism; Bernoulli processes.
- 2) The small ball methods and some of its applications (including: noise free recovery; Gelfand widths of convex bodies).
- 3) Mean estimation, covariance estimation and related questions.
- 4) Real valued VC-theory and its applications (in Statistical Learning Theory and in Asymptotic Geometric Analysis)
- 5) The Dvoretzky-Milman Theorem and its connection to lower bounds on sample complexity estimates.
- 6) Extremal singular values of random matrices.

Average Time Dedicated Per Week: 3 hours