

DEPARTMENT OF STATISTICS AND BIOSTATISTICS

Alexandre TsybakovLaboratoire de Statistique, CREST
ENSAE, (Paris)*Optimal and adaptive variable selection***October 31, 2017****1:30 – 2:30pm**

Light refreshments will be served

**110 Frelinghuysen Road
Hill Center, Room 552**

Abstract: We consider the problem of variable selection based on n observations from a high-dimensional linear regression model. The unknown parameter of the model is assumed to belong to the class \mathcal{S}_s of all s -sparse vectors in \mathbb{R}^p whose non-zero components are greater than $a > 0$. Variable selection in this context is an extensively studied problem and various methods of recovering sparsity pattern have been suggested. However, in the theory not much is known beyond the consistency of selection. For Gaussian design, which is of major importance in the context of compressed sensing, necessary and sufficient conditions of consistency for some configurations of n, p, s, a are available. They are known to be achieved by the exhaustive search selector, which is not realizable in polynomial time and requires the knowledge of \mathcal{S}_s . This talk will focus on the issue of optimality in variable selection based on the Hamming risk criterion. We first consider a general setting of variable selection problem and we derive the explicit expression for the minimax Hamming risk on \mathcal{S}_s . Then, we specify it for the Gaussian sequence model and for high-dimensional linear regression with Gaussian design. In the latter model, we propose an adaptive algorithm independent of s, a , and of the noise level that nearly attains the value of the minimax risk. This algorithm is the first method, which is both realizable in polynomial time and is consistent under almost the same (minimal) sufficient conditions as the exhaustive search selector. This talk is based on a joint work with C. Butucea, M. Ndaoud and N. Stepanova.

Bio: Alexandre B. Tsybakov is Professor and Head of Statistics Laboratory at CREST-ENSAE (Paris). He also holds a Professor position at the University Pierre and Marie Curie (Paris 6). His main research interests are in nonparametric estimation, high-dimensional statistics, and machine learning. He is an author of three monographs and more than 150 journal papers. Alexandre Tsybakov is a Fellow of the Institute of Mathematical Statistics and belongs to Highly Cited in Mathematics. He was awarded Le Cam's Lecture by the French Statistical Society, Miller Professorship by the University of California (Berkeley), Medallion Lecture by the Institute of Mathematical Statistics, Humboldt-Gay-Lussac Prize, and an Invited Lecture at the International Congress of Mathematicians.

