Data science sur un plateau

Nous avons le plaisir de vous inviter pour cette troisième édition du séminaire à ENSAE ParisTech le lundi 22 janvier 2018 de 14:00 à 17:30 dans l'amphi 200. Lors de cette session nous pourrons écouter :

Learning from MOM's principles (Joint with M. Lerasle) *Slides*

Guillaume Lecué (ENSAE-CREST, CNRS)

We obtain estimation error rates for estimators obtained by aggregation of regularized median-of-means tests, following a construction of Le Cam. The results hold with exponentially large probability – as in the Gaussian framework with independent noise–under only weak moments assumptions on data and without assuming independence between noise and design. Any norm may be used for regularization. When it has some sparsity inducing power we recover sparse rates of convergence. The procedure is robust since a large part of data may be corrupted, these outliers have nothing to do with the oracle we want to reconstruct. Our general risk bound is of order

$$\max\left(\text{minimax rate in the i.i.d. setup}, \frac{\text{number of outliers}}{\text{number of observations}}\right) .$$

In particular, the number of outliers may be as large as (number of data) $\times$(minimax rate) without affecting this rate. The other data do not have to be identically distributed but should only have equivalent L^1 and L^2 moments. For example, the minimax rate $s\log(ed/s)/N$ of recovery of a s-sparse vector in $\mathbb{R}^d$ is achieved with exponentially large probability by a median-of-means version of the LASSO when the noise has q_0 moments for some $q_0 > 2$, the entries of the design matrix should have $C_0 \log(ed)$ moments and the dataset can be corrupted up to $C_1 s \log(ed/s)$ outliers.

Computational Optimal Transport for Imaging and Learning

Gabriel Peyré (CNRS and Ecole Normale Supérieure)

Optimal transport (OT) has become a fundamental mathematical tool at the interface between calculus of variations, partial differential equations and probability. It took however much more time for this notion to become mainstream in numerical applications. This situation is in large part due to the high computational cost of the underlying optimization problems. There is a recent wave of activity on the use of OT-related methods in fields as diverse as computer vision, computer graphics, statistical inference, machine learning and image processing. In this talk, I will review an emerging class of numerical approaches for the approximate resolution of OT-based optimization problems. This offers a new perspective for the application of OT in imaging sciences (to perform color transfer or shape and texture morphing) and machine learning (to perform clustering, classification and generative models in deep learning).
More information and references can be found on the website of our book "Computational Optimal Transport" `https://optimaltransport.github.io/`

Structures élémentaires de la Géométrie de l'Information : fonction caractéristique de Koszul-Souriau, métrique de Fisher-Souriau et densité à maximum d'entropie d'ordre supérieur *Slides*

Frédéric Barbaresco (Representative of Key Technology Domain PCC (Processing, Control & Cognition) Board, THALES LAND & AIR SYSTEMS)