ID de Contribution: **3** Type: **Non spécifié**

Estimation of (I_n - XDX^T)^{-1} with large dependence and concentration hypotheses.

vendredi 17 février 2023 15:30 (1 heure)

We provide in this presentation an estimation of the expectation of the matrix Q(D) = (I_p - XDX^T)^{-1} when the data matrix X = (x_1, ...,x_n) \in M_{p,n} has independent columns (but not identically distributed) and D is random, bounded, not independent with X but satisfies some constraints on the dependence on each x_i: for any i, there exists a random diagonal matrix D_i independent of x_i and sufficiently close to D. The formula giving the estimation of Q is a classical generalization of known deterministic equivalents, the difficulty mainly lies in the proof of the convergence. It is proven under concentration of the measure hypotheses on X and it relies in particular on a formula giving the concentration of the product of such random vectors. In a sense, the study of Q(D) is a perfect example to expose the efficiency of the concentration of measure framework to prove random matrix theory inferences.
We will also provide a machine learning application of the estimation of the matrix Q(D) that concerns the prediction of the performances of robust regularized regression (Ridge regression with a general convex loss replacing the squared loss).

Orateur: LOUART, Cosme