Séminaire de Probabilités et Statistiques

09h30
Salle de Conférences

Karim Lounici

(Georgia Tech)

Concentration bounds and asymptotic distribution for the empirical spectral projectors of sample covariance operators

Let $X, X_1, \dots, X_n$ be i.i.d. Gaussian random variables in a separable Hilbert space $\mathbb{H}$ with zero mean and covariance operator $\Sigma = \mathbb{E}(X \otimes X)$, and let $\hat{\Sigma} := n^{-1} \sum_{j=1}^{n} (X_j \otimes X_j)$ be the sample (empirical) covariance operator based on $(X_1, \dots, X_n)$. Denote by P_r the spectral projector of Σ corresponding to its r-th eigenvalue μ_r and by $\hat{P}_r$ the empirical counterpart of P_r. Our goal is to obtain tight bounds on

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left\{ \frac{\|\hat{P}_r - P_r\|_2^2 - \mathbb{E}\|\hat{P}_r - P_r\|_2^2}{\operatorname{Var}^{1/2}(\|\hat{P}_r - P_r\|_2^2)} \leq x \right\} - \Phi(x) \right|,$$

where $\|\cdot\|_2$ denotes the Hilbert–Schmidt norm and Φ is the standard normal distribution function. Such accuracy of normal approximation of the distribution of squared Hilbert–Schmidt error is characterized in terms of so called effective rank of Σ defined as $\mathbf{r}(\Sigma) = \frac{\operatorname{tr}(\Sigma)}{\|\Sigma\|_\infty}$, where $\operatorname{tr}(\Sigma)$ is the trace of Σ and $\|\Sigma\|_\infty$ is its operator norm, as well as another parameter characterizing the size of $\operatorname{Var}(\|\hat{P}_r - P_r\|_2^2)$.

(Joint work with Vladimir Koltchinskii)