Counting faces of random polytopes and applications

Patrick Tardivel

Let us consider the following high-dimensional linear regression model

$$Y = X\beta + \varepsilon,$$

where X is a $n \times p$ matrix with $n < p$, $\beta \in \mathbb{R}^p$ is an unknown parameter and $\varepsilon \in \mathbb{R}^n$ is a centered random noise. Our purpose is to recover the unknown parameter β.

Even in the noiseless case when $\varepsilon = \mathbf{0}$ (and thus $Y = X\beta$) recovering the parameter β is not obvious since β is a solution among many of the linear system $Y = X\gamma$. However, under the assumption that lot of components β are null, one can recover β by solving the convex optimization problem

$$\text{argmin } \|\gamma\|_1 \text{ subject to } X\gamma = Y. \tag{1}$$

When $X = (X_1|\dots|X_p)$ is a random $n \times p$ matrix having i.i.d $\mathcal{N}(0,1)$ entries, the phase transition curve provides theoretical guaranties so that solving problem (1) allows to recover β. Indeed, when n and p are both very large and $\varepsilon = \mathbf{0}$, this curve provides a bound k, depending on n/p, so that β can be recovered by solving (1) under the assumption that $|\{i \in \{1,\dots,p\} \mid \beta_i \neq 0\}| < k$.

In this talk, I will show that the phase transition curve is related to counting faces of the random polytope $\text{conv}(\pm X_1, \dots, \pm X_p)$. Finally, I will introduce open questions related to the phase transition curve which I would like to investigate in my future research.