

Random matrices and dynamics of optimization in very high dimensions 3/3

mardi 18 juin 2024 16:45 (1h 30m)

Machine learning and Data science algorithms include the need for efficient optimization of topologically complex random functions in very high dimensions. Surprisingly, simple algorithms like Stochastic Gradient Descent (with small batches) are used very effectively.

I will concentrate on trying to understand why these simple tools can still work in these complex and very over-parametrized regimes.

I will first introduce the whole framework for non-experts, from the structure of the typical tasks to the natural structures of simple neural nets used in standard contexts. I will then cover briefly the classical and usual context of SGD in finite dimensions.

I will then survey recent work with Reza Gheissari (Northwestern), Aukosh Jagannath (Waterloo) giving a general view for the existence of projected "effective dynamics" for "summary statistics" in much smaller dimensions, which still rule the performance of very high dimensional systems, as well. These effective dynamics define a dynamical system in finite dimensions which may be quite complex, and rules the performance of the learning algorithm.

The next step will be to understand how the system finds these low dimensional "summary statistics".

RMT enters the game for this next step (which is done in the next works with the same authors and with Jiaoyang Huang (Wharton, U-Penn)).

This is based on a dynamical spectral transition: along the trajectory of the optimization path, the Gram matrix or the Hessian matrix develop BBP outliers which carry these effective dynamics.

I will illustrate the use of this point of view on a few central examples of ML: multilayer neural nets for classification (of Gaussian mixtures), and the XOR examples, for instance.

Orateur: BEN AROUS, Gérard