Learning by Random Features and Kernel Random Matrices: Sharp Asymptotics and Universality Laws

Yue Lu, Harvard University

Abstract: Many new random matrix ensembles arise in learning and modern signal processing. The spectral properties of these matrices help answer crucial questions regarding the training and generalization performance of neural networks, and the fundamental limits of high-dimensional signal recovery. As a result, there has been growing interest in precisely understanding the spectra and other asymptotic properties of these matrices. These new random matrices are often highly structured and are the result of nonlinear transformations. This combination of structure and nonlinearity leads to substantial technical challenges when applying existing tools from random matrix theory to these new random matrix ensembles.

We will consider learning by random feature models and the related problem of kernel ridge regression. In each case, a nonlinear random matrix plays a prominent role. We provide an exact characterization of the asymptotic training and generalization errors of these models. These results reveal the important roles played by the regularization, the loss function and the activation function in the mitigation of the "double descent phenomenon" in learning. The asymptotic analysis is made possible by a general universality theorem, which establishes the asymptotic equivalence between the nonlinear random matrices and a surrogate linear random matrix ensemble that is much easier to work with.

https://mathpicture.fas.harvard.edu/seminar