Some Remarks on Random Matrix Theory and Statistics

Noureddine El Karoui
University of California, Berkeley

Statisticians are nowadays working with datasets of larger and larger sizes, both in the number of observations collected (the sample size, n) and the number of measurements made per observation (the dimension of the observations, p). In this setting of large n and large p, many classical results in asymptotic statistics fail to hold. They underlie some statistical practice, hence creating potential problems for practitioners.

Consequently, there is currently quite a bit of activity in high-dimensional inference, trying to develop a statistical theory for large datasets and high-dimensional data. Random matrix theory is playing a part in this development. I'll discuss some of the relevant results, positive aspects and potential limitations of random matrix models and will hint at alternatives currently considered.