JANOSCH ORTMANN, UQAM

Para-Gaussian Approximations of Random Deep Neural Networks

It is known that randomly initialised neural networks have approximate Gaussian behaviour in the infinite-width limit. Unfortunately, this approximation can be poor for finite width and gets worse as one goes deeper into the network. In this talk, I will introduce a more accurate approximation that is slightly non-Gaussian. Constructed from the Hermite polynomials, we call it the para-Gaussian distribution. I will discuss how this para-Gaussian approximation can lead to a practical recursive method to approximate the deeper layers of the network. I will also explain how to apply this theory to linear neural networks (where things can be solved exactly) and to ReLU neural networks (where things can be approximated).