知の物理学研究センター

ipi seminar

Institute for Physics of Intelligence

Denny Wu

（University of Toronto & Vector Institute for Artificial Intelligence）

「Generalization of two-layer neural networks : an asymptotic view」

２０19年8月8日(木)　１０時３０分～１２時００分

東京大学本郷キャンパス理学部１号館9階 913 号室

We investigate the generalization properties of two-layer neural networks in a high-dimensional setting, i.e. when the number of samples n, features d and neurons h tend to infinity with their ratios converge to a non-zero constant. Specifically, we consider the unregularized least squares regression problem with two-layer neural networks trained with gradient flow from small initialization, in which either the first layer or the second layer is trained. When the first layer is random, and second layer coefficients are optimized, we recover the double descent phenomenon: a cusp in the population risk appears at $h \approx n$ and further overparameterization decreases the risk. In contrast, when the first layer weights are optimized, we highlight the difference between two choices of weight initialization, and show that the risk is independent of overparameterization in both cases. Our theoretical and experimental results suggest that the previously studied problem structure that provably gives rise to double descent might not translate to optimizing two-layer neural networks.