For favour of Posting

THE HONG KONG UNIVERSITY OF SCIENCE & TECHNOLOGY

Department of Mathematics

SEMINAR ON STATISTICS AND DATA SCIENCE

Generalization error of linearized neural networks: staircase and double-descent

By

Prof. Song MEI
UC Berkeley

Abstract

Deep learning methods operate in regimes that defy the traditional statistical mindset. Despite the non-convexity of empirical risks and the huge complexity of neural network architectures, stochastic gradient algorithms can often find the global minimizer of the training loss and achieve small generalization error on test data. As one possible explanation to the training efficiency of neural networks, tangent kernel theory shows that a multi-layers neural network — in a proper large width limit — can be well approximated by its linearization. As a consequence, the gradient flow of the empirical risk turns into a linear dynamics and converges to a global minimizer. Since last year, linearization has become a popular approach in analyzing training dynamics of neural networks. However, this naturally raises the question of whether the linearization perspective can also explain the observed generalization efficacy. In this talk, I will discuss the generalization error of linearized neural networks, which reveals two interesting phenomena: the staircase decay and the double-descent curve. Through the lens of these phenomena, I will also address the benefits and limitations of the linearization approach for neural networks.

Biography: *Song Mei is an incoming assistant professor in the Department of Statistics at UC Berkeley. He recently finished his Ph.D. at Stanford University. His research is motivated by data science, and lies at the intersection of statistics, machine learning, information theory, and computer science. His work often build on insights that originated within the statistical physics literature. His recent research interests include theory of deep learning, high dimensional geometry, approximate Bayesian inferences, and applied random matrix theory.*

Date : 24 July 2020 (Friday)
Time : 11:00am – 12:00noon
Zoom Meeting : https://hkust.zoom.us/j/5616960008

All are Welcome!