

University of Chicago and Toyota Technological Institute at Chicago
Machine Learning Seminar Series

PRESENTS:



Tengyuan Liang
University of Chicago Booth

Title: New Thoughts on Adaptivity, Generalization and Interpolation Motivated from Neural Networks.

Abstract: Consider the problem: given data pair (x, y) drawn from a population with $f_*(x) = E[y|x]$, specify a neural network and run gradient flow on the weights over time until reaching any stationarity. How does f_t , the function computed by the neural network at time t , relate to f_* , in terms of approximation and representation? What are the provable benefits of the adaptive representation by neural networks compared to the pre-specified fixed basis representation in the classical nonparametric literature? We answer the above questions via a dynamic reproducing kernel Hilbert space (RKHS) approach indexed by the training process of neural networks. We show that when reaching any local stationarity, gradient flow learns an adaptive RKHS representation, and performs the global least squares projection onto the adaptive RKHS, simultaneously. In addition, we prove that as the RKHS is data-adaptive and task-specific, the residual for f_* lies in a subspace that is smaller than the orthogonal complement of the RKHS, formalizing the representation and approximation benefits of neural networks. Then we will move to discuss generalization for interpolation methods in RKHS. In the absence of explicit regularization, Kernel “Ridgeless” Regression with nonlinear kernels has the potential to fit the training data perfectly. It has been observed empirically, however, that such interpolated solutions can still generalize well on test data. We isolate a phenomenon of implicit regularization for minimum-norm interpolated solutions which is due to a combination of high dimensionality of the input data, curvature of the kernel function, and favorable geometric properties of the data such as an eigenvalue decay of the empirical covariance and kernel matrices. In addition to deriving a data-dependent upper bound on the out-of-sample error, we present experimental evidence suggesting that the phenomenon occurs in the MNIST dataset.

Bio: Tengyuan Liang works on problems at the intersection of statistical inference, machine learning & optimization. His main research focuses on the mathematical and algorithmic aspects of inference and learning under computational budgets, especially for large-scale datasets. Specific topics include: computational difficulty and efficiency in statistical inference, statistical learning theory, and high-dimensional statistics. He is also interested in network science, online learning, stochastic optimization, and applied probability. His research has been published in top statistics journals, as well as in leading peer-reviewed machine learning conferences. Outside of academia, Liang has experience as a research scientist at Yahoo! Research at New York in 2016, collaborating on large-scale machine learning problems with industrial applications. Liang earned a PhD in statistics from the Wharton School at University of Pennsylvania in 2017, and a BS in mathematics and applied mathematics from Peking University in China in 2012. He joined the Chicago Booth faculty in 2017

Host: Rebecca Willett

Wednesday, May 1, 1:00 – 2:00 pm
Harper Center (Booth) Room 219
Pizza provided by UChicago CS Department