Stanford University
Department of Statistics

DEPARTMENTAL SEMINAR

4:30pm, Tuesday, December 5, 2017
Main Quad Building 01-370 Room 370
Cookies served at 4pm, Sequoia 1st floor Lounge.

Speaker: Weijie Su
Wharton Statistics Department,
University of Pennsylvania

Title: HiGrad: Statistical Inference for Stochastic Approximation and Online Learning

Abstract:

Stochastic gradient descent (SGD) is an immensely popular approach for online learning in settings where data arrives in a stream or data sizes are very large. However, despite an ever-increasing volume of works on SGD, much less is known about the statistical inferential properties of predictions based on SGD solutions. In this talk, we introduce a novel procedure termed HiGrad to conduct statistical inference for online learning, without incurring additional computational cost compared with the vanilla SGD. The HiGrad procedure begins by performing SGD iterations for a while and then splits the single thread into a few; this procedure hierarchically operates in this fashion along each thread. With predictions provided by multiple threads in place, a \( t \)-based confidence interval is constructed by decorrelating predictions using covariance structures given by the Ruppert–Polyak averaging scheme. Under certain regularity conditions, the HiGrad confidence interval is shown to attain asymptotically exact coverage probability. Finally, the performance of HiGrad is evaluated through extensive simulation studies and a real data example.

This is joint work with Yuancheng Zhu.