Data Science Seminar

Hosted by Department of Mathematical Sciences

- Date: Tuesday, November 14, 2017
- Time: 12:00pm – 1:00pm
- Room: WH-100E
- Speaker: Weijie Su (Wharton School of the University of Pennsylvania)
- Title: HiGrad: Statistical Inference for Stochastic Gradient Descent in Online Machine Learning

Abstract

Stochastic gradient descent (SGD) is an immensely popular approach for optimization in settings where data arrives in a stream or data sizes are very large. Despite an ever-increasing volume of works on SGD, much less is known about statistical inferential properties of predictions based on SGD solutions. In this talk, we introduce a novel procedure termed HiGrad to conduct inference on predictions, without incurring additional computational cost compared with the vanilla SGD. HiGrad begins by performing SGD iterations for a while and then split the single thread into a few, and it 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-Juditsky 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.

From:
http://www2.math.binghamton.edu/ - **Department of Mathematics and Statistics, Binghamton University**

Permanent link:
http://www2.math.binghamton.edu/p/seminars/datasci/171114

Last update: **2017/10/27 04:02**