Fast Algorithms for Estimating Covariance Matrices of Stochastic Gradient Descent Solutions

Wei Biao Wu

Department of Statistics, University of Chicago, U.S.A.

Abstract

Stochastic gradient descent (SGD), an important optimization method in machine learning, is widely used for parameter estimation especially in online setting where data comes in stream. While this recursive algorithm is popular for the computation and memory efficiency, it suffers from randomness of the solutions. In this talk we shall estimate the asymptotic covariance matrices of the averaged SGD iterates (ASGD) in a fully online fashion. Based on the recursive estimator and classic asymptotic normality results of ASGD, we can conduct online statistical inference of SGD estimators and construct asymptotically valid confidence intervals for model parameters. The algorithm for the recursive estimator is efficient and only uses SGD iterates: upon receiving new observations, we update the confidence intervals at the same time as updating the ASGD solutions without extra computational or memory cost. This approach fits in online setting even if the total number of data is unknown and takes the full advantage of SGD: computation and memory efficiency. This work is joint with Wanrong Zhu and Xi Chen.