Theory and Application of Stochastic Gradient Algorithms

WEINAN E
Beijing Institute of Big Data Research and Princeton University
Email: weinan@math.princeton.edu

Stochastic gradient algorithms have been very popular in machine learning and have been THE algorithm of choice for very large data sets. We provide a detailed analysis of the dynamical behavior of this algorithm that captures both the details of the initial exponential convergence behavior as well as the final fluctuations. We also analyze the various acceleration schemes, compute the optimal learning rate and mini-batch size. We conclude by showing some novel new application areas for these algorithms. This is a joint work with Qianxiao Li and Cheng Tai.