Numerical Analysis and Scientific Computing Seminar

Bayesian Sparse Learning With Preconditioned Stochastic Gradient MCMC and its Applications

Guang Lin
Purdue University

Abstract: Deep neural networks have been successfully employed in an extensive variety of research areas, including solving partial differential equations. Despite its significant success, there are some challenges in effectively training DNN, such as avoiding over-fitting in over-parameterized DNNs and accelerating the optimization in DNNs with pathological curvature. In this work, we propose a Bayesian type sparse deep leaning algorithm. The algorithm utilizes a set of spike-and-slab priors for the parameters in deep neural network. The hierarchical Bayesian mixture will be trained using an adaptive empirical method. That is, one will alternatively sample from the posterior using appropriate stochastic gradient Markov Chain Monte Carlo method (SG-MCMC), and optimize the latent variables using stochastic approximation. The sparsity of the network is achieved while optimizing the hyperparameters with adaptive searching and penalizing. A popular SG-MCMC approach is Stochastic gradient Langevin dynamics (SGLD). However, considering the complex geometry in the model parameter space in non-convex learning, updating parameters using a universal step size in each component as in SGLD may cause slow mixing. To address this issue, we apply computational manageable preconditioner in the updating rule, which provides step size adapt to local geometric properties. Moreover, by smoothly optimizing the hyperparameter in the preconditioning matrix, our proposed algorithm ensures a decreasing bias, which is introduced by ignoring the correction term in preconditioned SGLD. According to existing theoretical framework, we show that the proposed method can asymptotically converge to the correct distribution with a controllable bias under mild conditions. Numerical tests are performed on both synthetic regression problems and learning the solutions of elliptic PDE, which demonstrate the accuracy and efficiency of present work.

Friday, October 23, 2020, 2:40 pm
https://emory.zoom.us/j/95900585494

Mathematics
Emory University