I M B M istanbul matematiksel bilimler merkezi
istanbul center for mathematical sciences

An Introduction to Stochastic Gradient Markov Chain Monte Carlo Methods

Umut Şimşekli

INRIA - ENS Paris

Abstract

Markov Chain Monte Carlo (MCMC) methods are one of the most important family of algorithms in Bayesian machine learning. These methods lie in the core of various applications such as the estimation of Bayesian predictive densities and Bayesian model selection. They can also provide important advantages over optimization-based point estimation methods, such as having better predictive accuracy and being more robust to over-fitting. Despite their well-known benefits, during the last decade, conventional MCMC methods have lost their charm since they are often criticized as being computationally very demanding. Indeed, classical approaches based on batch 'Metropolis-Hastings' algorithm would require passing over the whole data set at each iteration and the acceptance-rejection test makes the methods even more impractical for modern data sets. As a remedy, Welling & Teh (ICML 2011) developed one of the first scalable MCMC frameworks referred to as stochastic gradient Langevin dynamics (SGLD). Unlike conventional 'batch' MCMC methods, SGLD uses subsamples of the data per iteration similar to SGD. With this manner, SGLD is able to scale up to large datasets, while at the same time being a valid MCMC method that forms a Markov chain asymptotically sampling from the target density. Several extensions of SGLD have been proposed, which are coined under the term Stochastic Gradient MCMC (SG-MCMC). In this tutorial-like talk, I will cover the basics of MCMC and survey the recent advances in the field of SG-MCMC.

Date : Wednesday, March 23, 2022

Time: 14:00

Place: Boğaziçi University, South Campus