Efficiency of Approximate MCMC

Aaron Smith
University of Ottawa

Abstract

It is widely known that the performance of MCMC algorithms can degrade quite quickly when targeting computationally expensive posterior distributions, including the posteriors associated with any large dataset. This has motivated the search for MCMC variants that scale well for large datasets. One general approach is to look at only a subsample of the data at every step. In this talk, I'll describe several simple lower bounds on the errors of such algorithms, providing basic limits on the performance of many such algorithms. I'll apply these generic results to realistic statistical problems and proposed algorithms, and also discuss some special examples that can avoid our limits.