New Monte Carlo Methods Based on Hamiltonian Dynamics

Radford M. Neal
University of Toronto
(radford@utstat.toronto.edu)

Abstract

Hamiltonian dynamics has been used to sample complex distributions for almost as long as the Metropolis algorithm has. Only from the 1980s, however, with the development of the HMC algorithm, has it been applied to distributions other than systems of molecules, such as in my work on Bayesian neural networks. The big advantages of HMC are its suppression of random walk behaviour, which can greatly speed exploration of the state space, and its superior scaling with dimensionality. After reviewing HMC, I will introduce two new Monte Carlo methods based on Hamiltonian dynamics. In "billiard HMC", the quadratic kinetic energy used in standard HMC is replaced by a piecewise-constant kinetic energy function. Hamiltonian dynamics can then be simulated exactly by solving equations for the times when the trajectory "bounces" off the locations of the discontinuities in the kinetic energy. With exact dynamics, trajectories are never rejected. Furthermore, one can define the kinetic energy in a way that leads to only one state variable changing at a time, which for many distributions allows fast methods for incremental computation to be used. Standard HMC can have difficulty handling multiple modes, and lacks a way of estimating the normalizing constant for the distribution sampled. My "Hamiltonian importance sampling" method aims to address these problems, by exploiting the volume preservation property of Hamiltonian dynamics to create an importance sampling distribution that can closely approximate a complex distribution while still having an easily-computable probability density function, which is necessary for re-weighting the sample points. Several tricks are required to get this idea to work. In particular, for applications to Bayesian inference, it needs to be combined with "slice sampling" for the prior. Athough this methods works very naturally with Hamiltonian dynamics, it can also be adapted for use with other MCMC methods, such as simple Metropolis updates, and so provides a general alternative to methods such as simulated tempering and multicanonical sampling.