Abstract:

Introduction:

Hamiltonian Monte Carlo (HMC, Hybrid Monte Carlo) is an MCMC method for sampling from an intractable posterior. HMC improves upon the Metropolis-Hastings algorithm by reducing the autocorrelation of the Markov chain generated. As such, convergence is attained more quickly, reducing necessary computational time.

The Hamiltonian of a physical system is the following:

H = U + K

where K is the kinetic energy and U is the potential energy of the system, written in terms of momentum(p) and position (q). The following two first order differential equations are satisfied:

(INSERT 2 DE HERE)

With HMC, we transform the posterior f using the transform:

U = -log(f)

Note that intractable posteriors are simply probability densities multiplied by an unknown normalization constant. Therefore the log transform will not create any imaginary values in any location where there is a nonzero density, and the potential energy approaches infinity as the density approaches zero.

Now, given a starting location and momentum, it is possible to calculate the time evolution of the system (using the two first order differential equations). The system is deterministic, and after some time, we can calculate the new q and the new p.

However, according to Hamilton’s equations, the change of momentum depends on position, and the change of position depends on momentum. If we iterate this system using a certain number of time steps (See: leapfrog method - Appendix(SOME LETTER HERE)) an error is induced and the total energy may increase or decrease. If we make some minor adjustments\*(MAKE SOME SORT OF FOOTNOTE ABOUT HOW WE DON’T REALLY DO ANYTHING, BUT THE THEORY WORKS), we may make a Metropolis-Hastings step and accept or reject the new position as another sample from the intractable prior.

In order to ensure that the Metropolis-Hastings step is legitimate and our Markov chain converges to a stationary distribution, it is sufficient to show our Markov chain is reversible. To guarantee time reversibility of a physical system, the transition operator must reverse the momentum after the last step (Figure WHATEVER THE FIGURE IS THAT HAS A PHASE SPACE DIAGRAM). However, since our kinetic energy is a function of the square of p, rather than just p, no changes are necessary in the code.

Methods: