Markov Chain Monte Carlo using Hamiltonian Dynamics

Joshua James MacDonald

August 22, 2018

Performance of MCMC

- MCMC is used to simulate a dependent samples from intractable distributions
- The Performance of MCMC in high dimensions has become important with the rise of Big Data
- How far can we move with each proposal while keeping a reasonable acceptance rate?

Motivation for HMC

Algorithm	Scaling	Optimal Acceptance Rate ($d \to \infty$)
RWM	$d^{-1/2}$	23.4% (Roberts & Rosenthal 2001)
MALA	$d^{-1/3}$	57.4% (Roberts & Rosenthal 2001)
HMC	$d^{-1/4}$	64% (Beskos et. al. 2010)

Extended Parameter Space

- A ball rolling around a potential surface
- At any time the ball has displacement x and momentum p
- Also has a mass m

Hamiltonian Dynamics

The Hamiltonian

Potential Energy,

$$U(x) = -\log\{\pi(x)\}$$

Kinetic Energy,

$$K(p;m) = \frac{1}{2}\frac{p^2}{m}$$

The Hamiltonian,

$$H(x,p;m) = U(x) + K(p;m)$$

Hamiltonian Equations

$$\frac{dx}{dt} = \frac{p}{m}$$

$$\frac{dp}{dt} = -\frac{dU}{dx}$$

Can be generalised to higher dimensions

$$\frac{dx}{dt} = M^{-1}p$$

$$\frac{dp}{dt} = -\nabla U(x)$$

1-dimensional Gaussian Example

Approximating Hamiltonian Dynamics

- Hamiltonian Dynamics are approximated using a numerical discretisation.
- Simulate over time T, using L steps and stepsize ε
- When used in MCMC, gives acceptance probability

$$\alpha(x^*, p^*; x, p) = \min\left\{1, \exp(-(H(x^*, p^*; M) - H(x, p; M))\right\}$$

HMC Algorithm

- We have some position x.
- Draw a new momentum $p \sim \mathcal{N}(0, M)$.
- Approximate Hamiltonian Dynamics for time T in L steps, using stepsize ε.
- Proposes new position and momentum (x^*, p^*).
- Accept proposal with probability

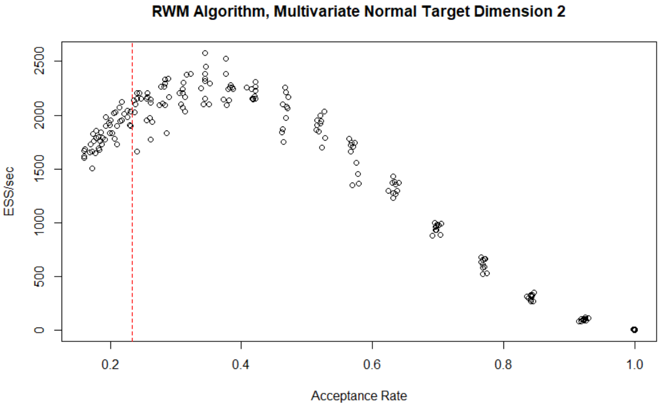
$$\alpha(x^*, p^*; x, p) = \min\left\{1, \exp(-(H(x^*, p^*; M) - H(x, p; M))\right\}$$

Choosing a Stepsize

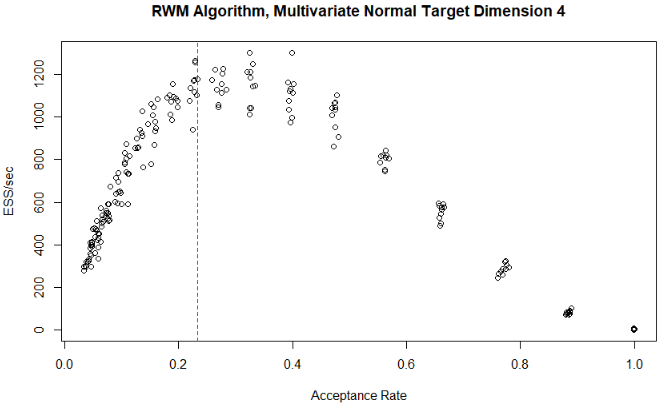
- Approximation can become unstable if the chosen stepsize is too large
- On the other hand, if the chosen stepsize is too smaller, the approximation becomes intensive.
- Find a range of stepsizes which perform suitably

Choosing the Integration Time

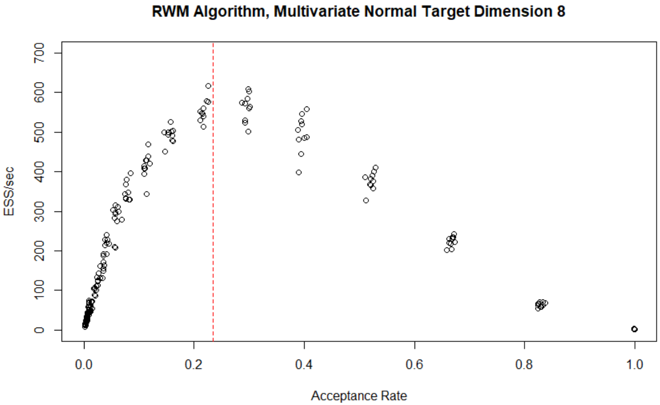
Optimal Acceptance Rate: RWM



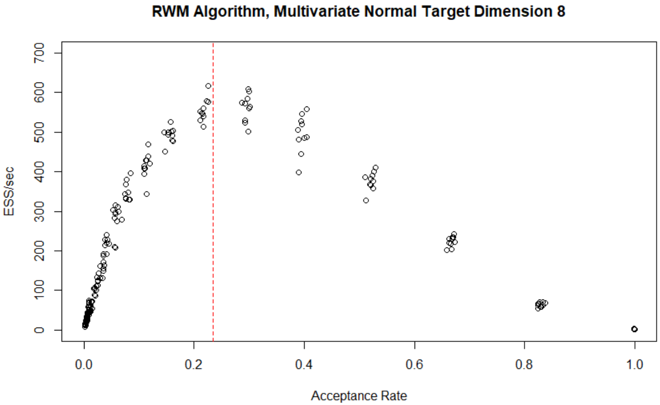
Optimal Acceptance Rate: RWM



Optimal Acceptance Rate: RWM



Optimal Acceptance Rate: RWM



Questions

Thank you for listening.

Any Questions?