Metropolis-Hastings Algorithm

MCMC Seminar

Sachin Shanbhag

Department of Scientific Computing
Florida State University,
Tallahassee, FL 32306.

Fall 2014

Historical Notes

- 1970 paper generalized the original Metropolis algorithm to allow for non-symmetric proposal moves.
- From 1966 to 1971, Hastings was an Associate Professor in the Department of Mathematics at the University of Toronto.*

  When I returned to the University of Toronto, after my time at Bell Labs, I focused on Monte Carlo methods and at first on methods of sampling from probability distributions with no particular area of application in mind. [University of Toronto Chemistry professor] John Valleau and his associates consulted me concerning their work. They were using Metropolis's method to estimate the mean energy of a system of particles in a defined potential field. With 6 coordinates per particle, a system of just 100 particles involved a dimension of 600. When I learned how easy it was to generate samples from high dimensional distributions using Markov chains, I realised how important this was for Statistics, and I devoted all my time to this method and its variants which resulted in the 1970 paper.
- Wrote only 3 papers and a mentored a single graduate student.

*http://probability.ca/hastings/

Detailed Balance

Given states i and j, the condition of detailed balance requires that "at equilibrium" the net traffic between the two states be equal:

$$\pi_i q_{ij} \alpha_{ij} = \pi_j q_{ji} \alpha_{ji} \tag{1}$$

where,

- π_i is the probability of state i
- q_{ij} is the probability of proposing a MC move from state i to state j (sometimes written as $q(j|i)$ or $q(i \to j)$).
- α_{ij} is the probability of accepting such a move.

The equation above implies:

$$\frac{\alpha_{ij}}{\alpha_{ji}} = \frac{\pi_j q_{ji}}{\pi_i q_{ij}} = r_H \tag{2}$$

Metropolis MC

We read this last time. Proposals are symmetric:

$$q_{ij} = q_{ji}.$$

Thus,

$$\frac{\alpha_{ij}}{\alpha_{ji}} = \frac{\pi_j}{\pi_i} \tag{3}$$

There are many different ways in which can specify a form for α_{ij} so that it satisfies the equation above.

Metropolis and company suggested:

$$\alpha_{ij} = \begin{cases} \frac{\pi_j}{\pi_i} & \text{if } \frac{\pi_j}{\pi_i} < 1 \\ 1 & \text{otherwise} \end{cases} \tag{4}$$

Let us check if this indeed satisfies eqn. 3.

Metropolis

Suppose $\pi_j/\pi_i < 1$. Then, $\alpha_{ij} = \pi_j/\pi_i$ from eqn. 4.

Since, $\pi_i/\pi_j > 1$, $\alpha_{ji} = 1$, again, according to eqn. 4.

Hence,

$$\frac{\alpha_{ij}}{\alpha_{ji}} = \frac{\pi_j/\pi_i}{1} \tag{5}$$

Checks out!

Hastings suggested what seems like a not so big change. He suggested, for $q_{ij} \neq q_{ji}$:

$$\alpha_{ij} = \begin{cases} \frac{\pi_j q_{ji}}{\pi_i q_{ij}} & \text{if } \frac{\pi_j q_{ji}}{\pi_i q_{ij}} < 1 \\ 1 & \text{otherwise} \end{cases} \tag{6}$$

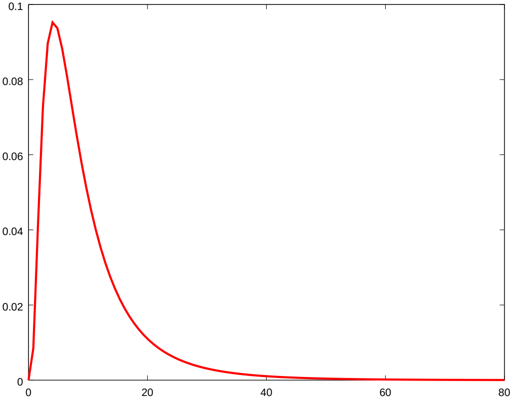
We can see that it satisfies eqn. 2, as required.

Example of Asymmetric Proposal

Suppose we want to sample a *wide and asymmetric* probability distribution like a log-normal distribution:

$$\pi(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}}\, e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, \quad x > 0 \tag{7}$$

Suppose, we set $\mu = 2.0$, and $\sigma = 0.75$.



Metropolis: Symmetric Move

Suppose we use standard Metropolis, with symmetric proposal

$$x_j = x_i + U(-\Delta, \Delta)$$

This implies that the probability of choosing something in the interval $[x_j, x_j + dx]$, when we are at x_i is:

$$q_{ij} = \frac{dx}{2\Delta}$$

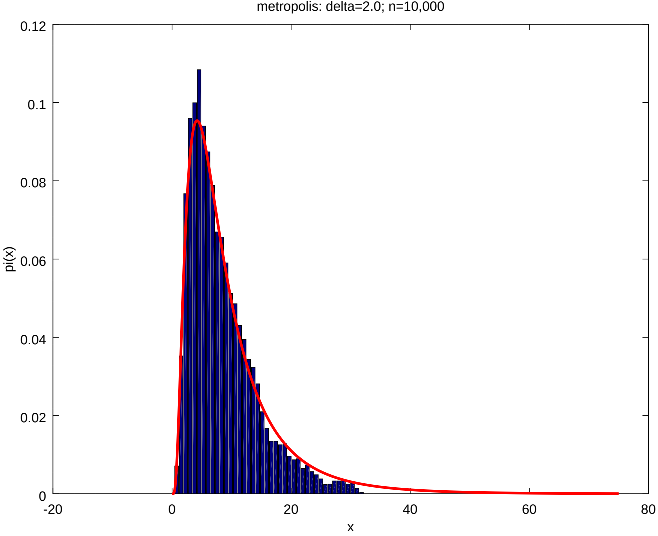
Similiarly, the probability of choosing something in the interval $[x_i, x_i + dx]$, when we are at x_j:

$$q_{ji} = \frac{dx}{2\Delta} = q_{ij}$$

Suppose I carry out 10,000 samples with $\Delta = 2.0$.

Metropolis: Symmetric Move

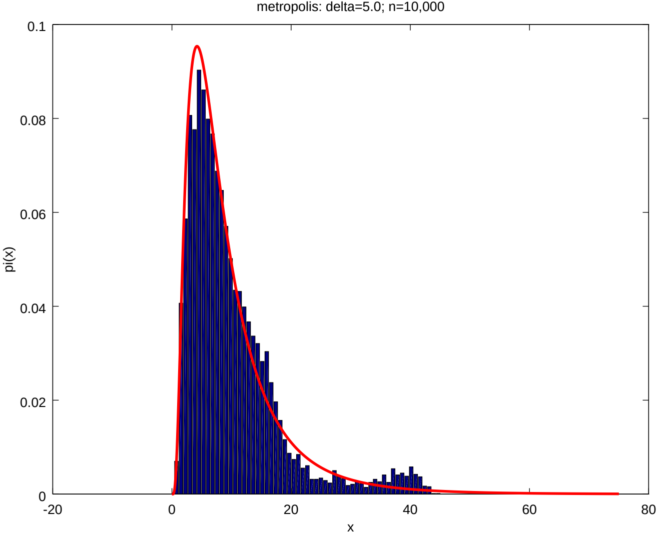
$$\Delta = 2.0$$



Tail doesn't seem to be sampled well.

Metropolis: Symmetric Move

$$\Delta = 5.0$$



Tail still doesn't seem to be sampled well.

Asymmetric Move

Suppose we decide to take a step in log-space, i.e., we set:

$$x_j = \beta x_i, \quad \beta = U[1/\rho, \rho]$$

Say $\rho = 1.5$, then we pick a random number between $\beta \sim U(2/3, 3/2)$ and set $x_j = \beta x_i$.

Note that this choice is asymmetric. You can just look at the support, for example.

This implies that the probability of choosing something in the interval $[x_j, x_j + dx]$, when we are at x_i is:

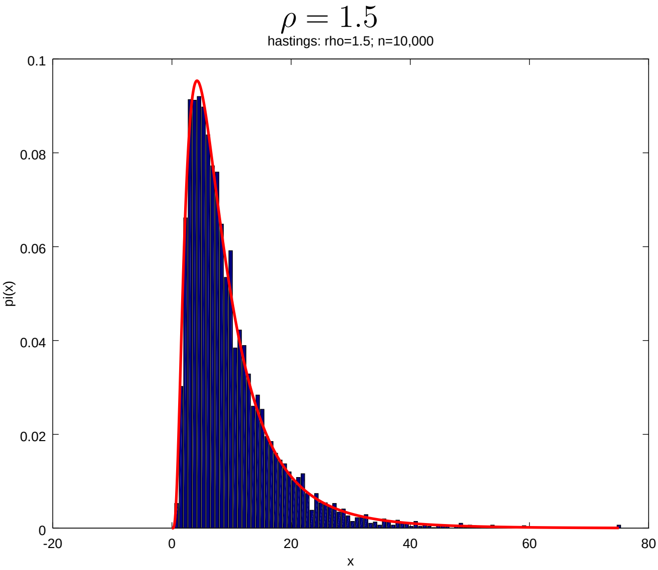
$$q_{ij} = \frac{dx}{x_i\,(\rho - 1/\rho)}$$

Similarly,

$$q_{ji} = \frac{dx}{x_j\,(\rho - 1/\rho)}$$

And, $q_{ij}/q_{ji} = x_j/x_i$.

Asymmetric Move



Tail now seems to be sampled well.

Why go beyond simple symmetric moves?

- Just saw an example of a wide asymmetric distribution
- Insight: Recall M-H criterion

$$\alpha_{ij} = \begin{cases} \frac{\pi_j q_{ji}}{\pi_i q_{ij}} & \text{if } \frac{\pi_j q_{ji}}{\pi_i q_{ij}} < 1 \\ 1 & \text{otherwise} \end{cases}$$

Ideally, I want to accept all my moves. From the equation above, if I can somehow get

$$\frac{\pi_j}{\pi_i} = \frac{q_{ij}}{q_{ji}},$$

then I am all set.

I can propose bold moves, and still have them all accepted.

Partial Illustration: Homework?

Consider a simple coupled Harmonic oscillator

$$U = \frac{1}{2}\sum_{i=0}^{N-1}(z_{i+1} - z_i)^2, \quad z_0 = z_N = 0; \tag{8}$$

The probability of a state $\mathbf{z} = (z_0, z_1, ..., z_N)$ is:

$$\pi(\mathbf{z}) \propto \exp\left(-U(\mathbf{z})\right)$$

Problem: Write a Metropolis MC to sample this distribution. Using $N =$ odd, report the distribution of the central bead $\pi(z_{cen})$.

Note: Coupled problem with a wide distribution. Try proposals of our usual form:

$$z_i^{'} = z_i + \Delta$$

This takes a long time to converge. Can we do better? Think about the distribution of z_i, for fixed z_{i-1} and z_{i+1}