Metropolis Hastings Explained

[see Hahn (2014), ch. 4, and Kruschke (2014), ch.?]

The Random Walk Metropolis Algorithm

Suppose **we wish to draw from density function,** $p(\theta)$, where we can write down (and hence evaluate) the function given values of θ, but it is not a known form that we can use a standard pseudo-random number generator for.

For example, let $p(\theta)$ be the kernel of a standard normal density, but pretend we do not recognize it as such, so we just have the function, and we also do not know the normalizing constant,

$$p(\theta) = c^{-1} \exp\left(-\frac{(\theta-\mu)^2}{2\,\sigma^2}\right), \tag{1}$$

where c is the normalizing constant.

We start at some arbitrary point, $\theta^{(0)}$, and draw a random value from some distribution. This can be *any* distribution provided it covers the entire support (set of possible values) of θ. The closer this "proposal" distribution is to the target distribution, $p(\theta)$, the better, but any distribution will work (eventually). For example, we can use a uniform over some finite interval, or a normal distribution with given precision, and draw our proposal values as

$$\theta^{(i)} = \theta^{(i-1)} + v^{(i)},\ i = 1,2, \dots, R \tag{2}$$

with $v^{(i)} \sim U[-s, s]$, or $v^{(i)} \sim N(0, s^2)$, where s is the "tuning parameter" that determines the average distance of the move between $i-1$ and i.

For each new value we obtain, $\theta^{(i)}$, Metropolis *et al.* (1953) showed that, by following specific rules on whether to accept the new proposed value or not, we can simulate draws from the target distribution (equation (1) in our example).

That is, we decide whether to accept the proposed new value according to the following rules:

1. *Always accept "uphill moves",* i.e. if the value of the density at the new value, $\theta^{(i)}$, is greater than the value of the density at $\theta^{(i-1)}$, then accept the new value.
2. *Accept a "downhill move" according to the ratio of the of the density values at* $\theta^{(i)}$ and $\theta^{(i-1)}$.

To implement step 2, we draw a random value, a, from a $U[0,1]$, and **if**

$$a \ \leq p(\theta^{(i)})/p(\theta^{(i-1)}),$$

we accept the new value, otherwise we keep the previous value, setting $\theta^{(i)} = \theta^{(i-1)}$.

We iterate through steps 1 and 2 R times. Metropolis *et al.* showed that, surprisingly, this procedure allows us to simulate draws from almost any continuous distribution.

Summarizing, the algorithm is

1. Select an initial value, $\theta^{(0)}$. While any value is valid, selecting a value that is consistent with the distribution (i.e. not too far from the highest density regions), will help speed convergence of the algorithm.
2. For iterations $i = 1, \ldots, R$, propose a move from $\theta^{(i-1)}$ to a new location θ^* by drawing from a symmetric jumping proposal distribution.
3. Draw an acceptance probability $a^{(i)}$ from a $U[0,1]$, and compare the ratio of the values of the target density at from $\theta^{(i-1)}$ and θ^* to $a^{(i)}$. If
$$a \leq \frac{p(\theta^{(i)})}{p(\theta^{(i-1)})},$$
accept the move, setting $\theta^{(i)} = \theta^*$, otherwise reject it, setting $\theta^{(i)} = \theta^{(i-1)}$. Always accept if the ratio is greater than one.

The normal density example is implemented using both the uniform and the normal jumping proposal distributions in **metropolis.R** and **metropolis2.R**.

Linear regression examples in:

MetropRegHahnp85.R – uses Gibbs step for tau

MetropRegHahnp85-metroptau.R – uses Metropolis step for tau also

IndMHexampleHahnp89.R – example of independent MH step

* Consider binomial with normal prior example using metropolis algorithm.

metropolispoissonnightmare.R – example of metropolis for poisson likelihood and nonconjugate prior

mcintex1.R – simple examples of MC integration and MCMC by Gibbs and Metropolis

____

Bayesregexample.R for Gibbs sampling for linear regression using bayesm package.

gibbsegch3.R – simpler version of above.

gibbsgraph.R - plots path of MCMC algorithm one step at a time for several steps.

triathlete.R – regression example both analytical and using Gibbs sampling for triathlon data.