POLYTECHNIC UNIVERSITY
Department of Computer and Information Science

METROPOLIS MONTE CARLO SIMULATION

K. Ming Leung

Abstract: The Metropolis method is another method of generating random deviates that are distributed according to an arbitrary probability density function. The basic ideas will be illustrated by applying the method to evaluate an indefinite integral.

Directory

Last Revision Date: March 25, 2003

Table of Contents

1. Introduction

The Metropolis method introduced by Metropolis, Metropolis, Rosenbluth, Rosenbluth, Teller and Teller in 1953[1] is another method of generating random deviates distributed according to a probability density function, $f(x)$. It is a general method that applies to any arbitrary $f(x)$, which need not be normalized. No prior knowledge of the form of the distribution is required.

In the Metropolis method, we imagine a particle located initially at x_0 performs a kind of random walk within the domain $[a, b]$. Time is discretized into finite steps. In each time step the particle attempts to jump to a new randomly chosen position. There is a certain criterion, which depends on $f(x)$, that determines whether or not the jump is allowed. The jump is made if it is allowed to do so, otherwise it is not moved. The process is repeated numerous times. After many time steps the particle will have forgotten where its starting position was. The final N positions visited by the particles should be distributed according to $f(x)$. The above dynamics of this particle is totally fictitious, and it is executed in a fictitious time step known as a Monte

Carlo time step.

2. The Metropolis Method

To completely specify the algorithm of generating these random positions, we have to specify the random walk process and the criterion for accepting or rejecting a certain move. The Metropolis method is therefore a special importance sampling procedure that involves accepting or rejecting certain possible sampling attempts.

Suppose the particle is currently located at x. We can consider it jumping to a new position x' given by $x' = x + \Delta x$, where Δx is randomly chosen in the interval $[-\delta, \delta]$, and δ is a parameter of the simulation. If the domain, $[a, b]$, does not cover the entire real line, additional care must be exercised to make sure that the particle does not attempt to move outside the domain.

We also need to define a transition probability $T(x \to x')$ for accepting this trial move from x to x' in such a way that the ultimate distribution of the positions visited by the particle will be distributed according to $f(x)$. It can be shown that a sufficient (but not necessary)

condition for the transition probability $T(x \to x')$ for moving from x to x' is the detailed balance condition:

$$f(x)T(x \to x') = f(x')T(x' \to x), \tag{1}$$

where $T(x' \to x)$ is the transition probability for moving from x' to x. The detailed balance condition does not specify the transition probability uniquely. There are still many ways of choosing $T(x \to x')$ and $T(x' \to x)$. A simple choice of $T(x \to x')$ that is consistent with the detailed balance condition is:

$$T(x \to x') = \min\left[1, \frac{f(x')}{f(x)}\right]. \tag{2}$$

Therefore if this trial position, x', is such that $f(x')/f(x)$ is greater than 1, meaning that this new position would increase the value of f, then this move is accepted (with probability 1). However, if $f(x')/f(x)$ is less than 1, and so this move would actually decrease the value of f, then this move is still accepted, although with a lower probability of $f(x')/f(x)$.

It is clear that the situation is totally symmetric between the po-

sitions x and x', and so the transition probability for moving from x' to x is:

$$T(x' \to x) = \min\left[1, \frac{f(x)}{f(x')}\right]. \tag{3}$$

To see that the detailed balance condition holds for this choice of transition probabilities, we assume first that $f(x')/f(x)$ is greater than 1. Then according to Eq.(2), $T(x \to x')$ is given by 1 and $T(x' \to x)$ is given by $f(x')/f(x)$. Consequently

$$f(x)T(x \to x') = f(x) = f(x')\frac{f(x)}{f(x')} = f(x')T(x' \to x), \tag{4}$$

and so the detailed balance condition is obeyed.

On the other hand if $f(x')/f(x)$ is less than 1, then according to Eq.(2), $T(x \to x')$ is given by $f(x')/f(x)$ and $T(x' \to x)$ is given by 1. Consequently

$$f(x)T(x \to x') = f(x') = f(x')T(x' \to x), \tag{5}$$

and so the detailed balance condition is also obeyed.

2.1. Algorithm for the Metropolis Method

The Metropolis algorithm of generating points distributed according to an arbitrary pdf $f(x)$:

1. Choose parameters $N_{\max}$ and δ, and set initial position x, and N to 0.
2. Go through the following loop until N equals $N_{\max}$:
   (a) Let trial position $x' = x + (2u_1 - 1)\delta$, where u_1 is a uniform deviate.
   (b) Compute the ratio $r = f(x')/f(x)$.
   (c) If $r \geq 1$, the move is accepted, x is replaced by x', and N is incremented by 1, otherwise
      i. Choose another uniform deviate, u_2.
      ii. If $u_2 \leq r$, the move is accepted, x is replaced by x', and N is incremented by 1

2.2. Choice of parameters of the method

We discuss here some of the finer details of the Metropolis algorithm that have to do with the choices of parameters used in the simulation. First consider the choice of the starting position for the random walk. Notice that if the chosen initial position lies in a region where $f(x)$ is exceedingly small, it will take the random walker many steps before reaching a region where $f(x)$ is large. Therefore all the earlier positions visited by the walker will all have small f and the probability distribution of these points is far from that given by $f(x)$. These points are the results of this unfortunate choice of initial position and thus are often discarded. In such a case the number of accepted points generated by the random walk, $N_{\max}$, which is another parameter of the simulation, has to be larger than the number of points we actually want to obtain. Therefore the method is more efficient if the walker starts from a position where $f(x)$ is large. The set of points obtained then more quickly approaches the asymptotic distribution $f(x)$.

Another parameter we have to choose in the simulation is the maximum step size δ. First suppose that δ is large. Most of the points

visited in the random walk are located where the corresponding values of $f(x)$ are relatively large, because the visited points are supposed to be distributed according to $f(x)$. Taking a big step from these positions will very likely end up with trial positions, x', in regions where $f(x')$ are very small compared with $f(x)$. According to the Metropolis algorithm these points are accepted with probability a very small probability $f(x')/f(x)$. So most of the trials points end up have to be rejected.

On the other hand if the maximum step size is very small, the trials points x' are then very close to the current point x. Of course if the ratios $r = f(x')/f(x)$ are larger than 1, then the trial positions are accepted. However even if r turn out to be smaller than 1, since x' are so close to x, r must be nearly 1, and these trial positions will still be accepted with overwhelming probability. The sampling of the entire probability distribution is therefore also not efficient. As a rule of thumb, δ should be chosen so that roughly one third to a half of the trials positions should be accepted.

References

[1] Metropolis, Rosenbluth, Rosenbluth, Teller and Teller.

[2] J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods*, (Methuen, London, 1964).

[3] M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods*, (Willey, New York, 1986).

[4] H. Gould and J. Tobochnik, *An Introduction to Computer Simulation Methods: Applications to Physical Systems*, Second Edition, Ch. 11, Sec. 8, and Ch. 17, Sec. 2. (Addison Wesley, New York, 1996).

3