THE METROPOLIS–HASTINGS ALGORITHM

BENEDICT MORRISSEY

This is a brief note explaining the Metropolis–Hastings algorithm[3, 2] and giving a brief example. This is a note for part of Math 312 Summer 2019 lecture 13. It is primarily based off the excellent exposition in [1].

Aim 0.1. *Given a vector $\vec{p} = (p_1, ..., p_n)^T$, satisfying $0 \leq p_i \leq 1$ for each i, $1 \leq i \leq n$, and satisfying $\sum_{i=1}^{n} = 1$, we want to create a Markov chain with the property that $\vec{p}$ is its (only) steady state (equilibrium) probability distribution.*

Warning 0.2. *We continue to use the convention that the* transition matrix *of a Markov chain is the matrix T such that T_{ij} (the entry in i^{th} row and j^{th} column of T) is the probability of transitioning from state j to state i in a given time step*[1]*. Some authors instead work with T^T.*

1. Metropolis–Hastings Algorithm

The Metropolis Hastings Algorithm consists of the following two steps, and produces a Markov chain as wanted in aim 0.1.

(1) Pick an arbitrary Markov chain on n states, with the property that
   - For any pair of states i, and j, if we are in state i there is a non-zero probability that we will be in state j in some future time step.

(2) Let T' be the transition matrix for the Markov chain in the first step of this algorithm. We form a new Markov chain by forming the new transition matrix:

$$T_{ij} = \begin{cases} T'_{ij} \min\{1, \frac{T'_{ji} p_i}{T'_{ij} p_j}\} \text{ if } i \neq j \\ 1 - \sum_{l \neq j} T_{lj} \text{ if } i = j \end{cases}$$

Note firstly that the entries in the columns of T are all positive, and for each column the entries sum to one. Hence we indeed have the transition matrix for a Markov chain.

Note that it is not a problem if T'_{ij} or p_i is zero – as we will simply take 1 as the minimum.

Proposition 1.1. *The vector $\vec{p}$ is a steady state probability distribution for the Markov chain produced by the Metropolis–Hastings Algorithm.*

Proof. First note that $T_{ij} p_j = T_{ji} p_i$.

Hence the i^{th} entry of $T\vec{p}$ is

$$\begin{aligned}(T\vec{p})_i &= \sum_{j=1}^{n} T_{ij} p_j \\ &= \sum_{j=1}^{n} T_{ji} p_i \\ &= 1\end{aligned}$$

where the final line follows as the entries in each column of T sum to one.

Hence $\vec{p}$ is a steady state (equilibrium) distribution of the Markov chain produced by the Metropolis–Hastings algorithm. □

[1] Assuming that we are in state j.

2. Example

Suppose we want a Markov chain with the steady state (equilibrium) probability distribution $\vec{p} = \begin{pmatrix} 3/4 \\ 1/4 \end{pmatrix}$.

We implement the Metropolis–Hastings algorithm as follows:

- We first pick the Markov chain with transition matrix

$$T' = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

- We now have $T_{12} = \frac{1}{2}\min\{1, \frac{3/4}{1/4}\} = 1/2$. Hence $T_{22} = 1 - \frac{1}{2} = \frac{1}{2}$.
  We have $T_{21} = \frac{1}{2}min\{1, \frac{1/4}{3/4}\} = \frac{1}{6}$. Hence $T_{11} = 5/6$.
  The Markov chain we get hence has transition matrix

$$T = \begin{pmatrix} 5/6 & 1/2 \\ 1/6 & 1/2 \end{pmatrix}$$

Solving for eigenvectors of T corresponding to the eigenvalue 1, gives us $s\begin{pmatrix} 3 \\ 1 \end{pmatrix}$. Normalizing tells us that $\begin{pmatrix} 3/4 \\ 1/4 \end{pmatrix}$ is a steady state probability distribution for T.

References

[1] Leonardo Ferreira Guilhoto. Applying markov chains to monte carlo integration.
[2] W Keith Hastings. Monte carlo sampling methods using markov chains and their applications. 1970.
[3] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.