MAT 653: Metropolis-Hastings Algorithm (discrete space)

Instructor: Dr. Wei Li Scribe: Peichen Yu

Nov 4th, 2021

Given a steady-state distribution π, how to find a M.C. that has π as its steady-state distribution?

Goal: to find the right transition kernel P_{ij}

Idea: Consider a M.C. that is "time reversible", i.e. whose direction of the time does not matter in the dynanic of the chain.

Farmally, $P(X^{(n)} = j | X^{(n+1)} = i) = P(X^{(n+1)} = j | X^{(n)} = i)$, for $\forall i, j$.

Suppose we are at steady-state π,

$$P(X^{(n)} = i | X^{(n+1)} = j) P(X^{(n+1)} = j) = P(X^{(n+1)} = j | X^{(n)} = i) P(X^{(n)} = i)$$

by time reversibility,

$$P(X^{(n)} = i | X^{(n+1)} = j) = P(X^{(n+1)} = i | X^{(n)} = j) = P_{ji}$$

imply that,

$$\pi_j P_{ji} = \pi_i P_{ij}, \forall i, j$$

(detailed balance condition)

This says that the flows of probabilities between two states i, j balance each other.

Theorem: A set of transition probabilities $P_{i,j}$, satisfying the "detailed balance condition" will have π as its steady state distribution.

Proof: $\sum_i \pi_i P_{ij} = \sum_i \pi_j P_{ji} = \pi_j \sum_i P_{ji} = \pi_j \times 1 = \pi_j$

Metropolis Algorithm

Goal: to find the transition kernel $\tilde{P}_{ij}$, s.t., the detailed balance condition holds, i.e., $\pi_j \tilde{P}_{ji} = \pi_i \tilde{P}_{ij}, \forall i, j$.

We can use some arbitrary transition kernel, called P_{ij}. By following the modification described below, we can change $P_{ij} \to \tilde{P}_{ij}$.

If $\pi_i P_{ij} < \pi_j P_{ji}$, we accept all the transition from i to j, but only some of the transition from j to i to bring back into balance.

If $\pi_i P_{ij} > \pi_j P_{ji}$, we are to bring back into balance by only accept some of the transition from i to j, but accept all of the transition from j to i.

Formally, Metropolis' idea works as follows:

Given π, the steady-state distribution (target distribution). For all i,

(1) Start from P_{ij}

(2) Define acceptance probability $\alpha_{ij} = min[\frac{\pi_j P_{ji}}{\pi_i P_{ij}}, 1]$ for each $j \neq i$. Let

$$\tilde{P}_{ij} = \alpha_{ij} P_{ij}, j \neq i$$

$$\tilde{P}_{ii} = 1 - \sum_{j:j \neq i} \tilde{P}_{ij}$$

Then π is the steady-state distribution from M.C. with transition probability given by this $\tilde{P}_{ij}$.

Proof:

Note: $\tilde{P}_{ii} = 1 - \sum_{j:j\neq i} \tilde{P}_{ij}$ is the probability that I reject any proposal that moves from i to any other state $j \neq i$, that is, it is the probability of moving from i to i itself. We can prove that our $\tilde{P}_{ij}$ is well-defined transition kernel.

Check: $1 \geq \tilde{P}_{ij} \geq 0$, $\sum_j \tilde{P}_{ij} = \sum_j \tilde{P}_{ij} = \tilde{P}_{ii} + \sum_{j:j\neq i} \tilde{P}_{ij} = 1 - \sum_{j:j\neq i} \tilde{P}_{ij} + \sum_{j:j\neq i} \tilde{P}_{ij} = 1$

The following algorithm implements above procedure.

Algorithm (Metropolis): $X^{(0)} \in S$, S is a discrete space.

1. draw $X^* \sim P(\cdot|X^{(t-1)})$
2. compute acceptance probability

$$\alpha(X^{(t-1)}, X^*) := min\left\{1, \frac{P(X^{(t-1)}|X^*)\pi_{X^*}}{P(X^*|X^{(t-1)})\pi_{X^{(t-1)}}}\right\}$$

   where π_{X^*} is the probability of taking value X^* under the steady-state distribution π, and $\pi_{X^{(t-1)}}$ is the probability of taking value $X^{(t-1)}$ under the steady-state distribution π.

3. with probability $\alpha(X^{(t-1)}, X)$, set $X^{(t)} \leftarrow X^*$; otherwise $X^{(t)} \leftarrow X^{(t-1)}$.

To implement step 3, one can implement: generate U~Unif(0,1), if $U \leq \alpha(X^{(t-1)}, X^*)$, then accept X^*, i.e., set $X^{(t)} \leftarrow X^*$; otherwise, reject X^*, set $X^{(t)} \leftarrow X^{(t-1)}$.