In the name of God

Sharif University of Technology

Stochastic Process

Fall 2024

Hamid R. Rabiee

Quiz 10 (25)	Sampling	Release: Dey, 11th

1. Metropolis-Hastings Algorithm Pseudo-code

Find the errors in the following pseudo-codes and explain the reasons. Please note that x, x', and x_is in general all represent samples drawn from the distribution. These samples correspond to the states of the Markov Chain we are sampling from.

(a) **Initialize:**

- Choose an initial state x_0.
- Set $n = 0$.

(b) **While** $n < N$**:**

a. Propose a new state x' using the proposal distribution $Q(x' \mid x_n)$.

b. Calculate the acceptance ratio:

$$r = \frac{P(x') \cdot Q(x' \mid x_n)}{P(x_n) \cdot Q(x_n \mid x')}$$

c.
- Generate a random number $u \sim U(0, 1)$.
- If $u > \min(1, r)$:
  - Accept x', set $x_{n+1} = x'$.
- Otherwise:
  - Reject x', set $x_{n+1} = x_n$.

d. Increment n: $n = n + 1$.

(c) **Return** the set of samples $\{x_1, x_2, \ldots, x_N\}$.

Solution:

- The acceptance rate is: $r = \frac{P(x') \cdot Q(x_n|x')}{P(x_n) \cdot Q(x'|x_n)}$. This acceptance rate ensures the generated samples approximate the target distribution P(x) and converge to it in the long run. This rate is designed to satisfy the detailed balance condition, which guarantees reversibility of the Markov chain. As a result, P(x) becomes the stationary distribution of the chain.

  Let $P(x)$ be the target distribution and $Q(x'|x)$ be the proposal distribution. The transition probability from state x to x' is given by:

$$T(x'|x) = Q(x'|x) \min\left(1, \frac{P(x')Q(x|x')}{P(x)Q(x'|x)}\right)$$

To prove reversibility, we need to show that:

$$P(x)T(x'|x) = P(x')T(x|x')$$

We consider two cases:

(a) Case 1: $P(x')Q(x|x') \geq P(x)Q(x'|x)$

$$\begin{aligned} P(x)T(x'|x) &= P(x)q(x'|x) \cdot 1 \\ &= P(x)q(x'|x) \end{aligned}$$

$$\begin{aligned} P(x')T(x|x') &= P(x')q(x|x') \min\left(1, \frac{P(x)q(x'|x)}{P(x')q(x|x')}\right) \\ &= P(x')q(x|x') \cdot \frac{P(x)q(x'|x)}{P(x')q(x|x')} \\ &= P(x)q(x'|x) \end{aligned}$$

(b) Case 2: $P(x')q(x|x') < P(x)q(x'|x)$

$$\begin{aligned} P(x)T(x'|x) &= P(x)q(x'|x) \cdot \frac{P(x')q(x|x')}{P(x)q(x'|x)} \\ &= P(x')q(x|x') \end{aligned}$$

$$\begin{aligned} P(x')T(x|x') &= P(x')q(x|x') \cdot 1 \\ &= P(x')q(x|x') \end{aligned}$$

In both cases, we have shown that:

$$P(x)T(x'|x) = P(x')T(x|x')$$

This proves that the balance condition is satisfied, and thus the Markov chain is reversible with respect to the target distribution $P(x)$ and converges to the target distribution.

- If $u \leq \min(1, r)$, accept. The other condition leads an acceptance rate of $1 - r$ and therefore the above properties are not hold in that case.

2. **Gibbs Sampling vs Metropolis-Hastings Algorithm**

Gibbs sampling is a specialized variant of the Metropolis-Hastings algorithm with a key distinguishing feature. Taking this specific characteristic into account, let's compare Gibbs Sampling with the general Metropolis-Hastings algorithm in terms of the following aspects. Please provide brief explanations for your answers.

(a) **Proposal Distribution** (What are the specific conditions for the proposal distributions used in each of them?)

(b) **Acceptance Step** (Is the step the same for both of them?)

(c) **Complexity** (Is there any difference between their computational complexity in any step?)

(d) **Efficiency** (When both of the algorithms can be applied, which would provide the required samples faster?)

(e) **Use case** (In which situations / problems, Gibbs Sampling is preferred to the Metropolis-Hastings Algorithm?)

Solution:

(a) **Proposal Distribution**
In Gibbs Sampling, the conditional distributions of one variable given the others are used as the proposals. In Metropolis Hastings algorithm in general, any proposal distribution can be used as the proposal (Definitely, we would prefer a proposal that closely approximates the target distribution while remaining computationally efficient. This strategy typically results in fewer rejections, faster convergence, and overall improved efficiency of the sampling process).

(b) **Acceptance Step**
There is a acceptance rate in the general Metropolis-Hastings algorithm, while due to the specific proposal used in the Gibbs Sampling, the rate is 1 and all samples are accepted.

(c) **Complexity**
In Gibbs sampling, we sample for one variable given the values of the others at a time. For N random variables, we need to sample N times for each variable. Considering the complexity of the conditional distribution to be $O(C)$, the total complexity of generating each sample is of $O(NC)$.
In the general Metropolis-Hastings algorithm, considering no specific feature for the proposal, we have a N-dimensional joint distribution to sample from. Therefore, both sampling from the joint distribution and calculating the probabilities for the acceptance step is of $O(C^N)$ complexity.

(d) **Efficiency**
Considering the previous item, Gibbs Sampling is more efficient and would produce the samples faster. Additionally, there is no rejection and all samples are accepted leading to a faster convergence. Therefore it is more efficient.

(e) **Use case**
Gibbs Sampling, while effective in many scenarios, faces a significant limitation when conditional distributions are difficult to calculate or sample from. In such cases, the general Metropolis-Hastings algorithm may be more efficient, especially when using simpler proposal distributions with independence assumptions. Consequently, Gibbs Sampling is not

always the preferred choice over the general version. However, in situations where conditional distributions are easily calculated and sampled, Gibbs Sampling often outperforms the general Metropolis-Hastings algorithm in terms of efficiency. Thus, the choice between these methods depends largely on the specific characteristics of the problem at hand and the computational feasibility of sampling from conditional distributions.