Monte Carlo Algorithm and its Application

Applied Math Tutorial

Tianlu Zhu

Shanghaitech Unversity

February 28, 2023

Outline

MC Integration

Suppose we want to evaluate an integral

$$\int_D \phi(x)dx$$

for which there is no closed analytic solution. If the integration has the form

$$\phi(x) = \tilde{\phi}(x)f(x)$$

for some density function f, then the integral has the form:

$$\int_D \phi(x)dx = \int_D \tilde{\phi}(x)f(x)dx = E[\tilde{\phi}(X)]$$

where X is an RV with $\mathrm{PDF} f$.

MC Integration

If we know how to simulate realisations of X, say $x^{(1)}, \ldots, x^{(n)}$, then we have an estimate

$$\int_D \phi(x) dx = E[\tilde{\phi}(X)] \approx \frac{1}{n} \sum_{i=1}^{n} \tilde{\phi}\left(x^{(i)}\right) = \hat{I}$$

MC Integration

If we know how to simulate realisations of X, say $x^{(1)}, \ldots, x^{(n)}$, then we have an estimate

$$\int_D \phi(x)dx = E[\tilde{\phi}(X)] \approx \frac{1}{n}\sum_{i=1}^{n} \tilde{\phi}\left(x^{(i)}\right) = \hat{I}$$

Also, the variance should be

$$\begin{aligned} \mathrm{Var}(I) &= \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}\left(\tilde{\phi}\left(x\right)\right) \\ &\approx \frac{1}{n(n-1)}\sum_{i=1}^{n}\left(\tilde{\phi}\left(x^{(i)}\right) - \hat{I}\right) \end{aligned}$$

So it is crucial to generate sample from a specific distribution.

Outline

Derivation of the Algorithm

The easiest distribution for computer to generate is $\mathrm{Unif}[0,1]$. If X owns CDF $F(\cdot)$, Then $F(X) \sim \mathrm{Unif}[0.1]$. By *inversion*, setting $X \sim F^{-1}(U)$, which $U \sim \mathrm{Unif}[0,1]$. Then From the uniform distribution sample we can generate sample follows $F(\cdot)$.

Derivation of the Algorithm

The easiest distribution for computer to generate is $\mathrm{Unif}[0,1]$. If X owns CDF $F(\cdot)$, Then $F(X) \sim \mathrm{Unif}[0.1]$. By *inversion*, setting $X \sim F^{-1}(U)$, which $U \sim \mathrm{Unif}[0,1]$. Then From the uniform distribution sample we can generate sample follows $F(\cdot)$.

Consider discrete random variable, simply left

$$F^{-1}(u) = \min\{x : F(x) \geq u\}$$

And the algorithm goes on as continuous situation.

Pros and Cons

Pros

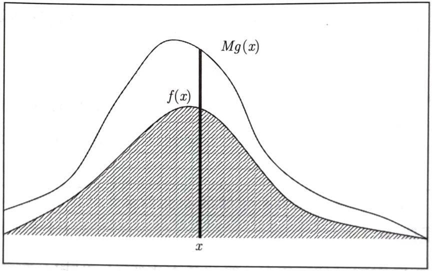
- Naive way to generate a group of sample

Cons

- The CDF is not invertible,i.e. PDF $f(\cdot) = 0$ somewhere.

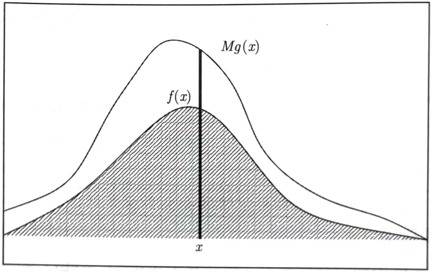
Outline

Intuition



- $f(x)$: object function
- $g(x)$: proposal function
- M: auxiliary constant
- Make sure that $Mg(x) \geq f(x)$ for all $x \in \Omega$

Algorithm



1. Produce a sample y from $g(\cdot)$
2. Produce a sample u from $\mathrm{Unif}(0, 1)$
3. Do comparasion. If $u \leq \frac{f(y)}{Mg(y)}$, accept the sample. Otherwise, reject the sample.
4. Back to the first step unless we already have enough sample.

Proof of Algorithm

Proof of Algorithm

Let $U \sim \mathrm{Unif}[0,1]$ and Y owns PDF $g(\omega) = F'(\omega)$, one obtain:

$$P\left(U \leq \frac{f(Y)}{Mg(Y)} \middle| Y = y\right) = \frac{f(y)}{Mg(y)}$$

Proof of Algorithm

Let $U \sim \mathrm{Unif}[0,1]$ and Y owns PDF $g(\omega) = F'(\omega)$, one obtain:

$$P\left(U \leq \frac{f(Y)}{Mg(Y)} \middle| Y = y\right) = \frac{f(y)}{Mg(y)}$$

The total probability is:

$$P\left(U \leq \frac{f(Y)}{Mg(Y)}\right) = \int_{-\infty}^{\infty} \frac{f(y)}{Mg(y)} g(y) dy = \frac{1}{M}$$

Proof of Algorithm

Let $U \sim \mathrm{Unif}[0,1]$ and Y owns PDF $g(\omega) = F'(\omega)$, one obtain:

$$P\left(U \leq \frac{f(Y)}{Mg(Y)}\middle| Y = y\right) = \frac{f(y)}{Mg(y)}$$

The total probability is:

$$P\left(U \leq \frac{f(Y)}{Mg(Y)}\right) = \int_{-\infty}^{\infty} \frac{f(y)}{Mg(y)} g(y)dy = \frac{1}{M}$$

By Bayesian's formula,

$$P\left(Y = y \middle| U \leq \frac{f(Y)}{Mg(Y)}\right) = P(A|B) = \frac{p(B|A)P(A)}{P(B)}$$

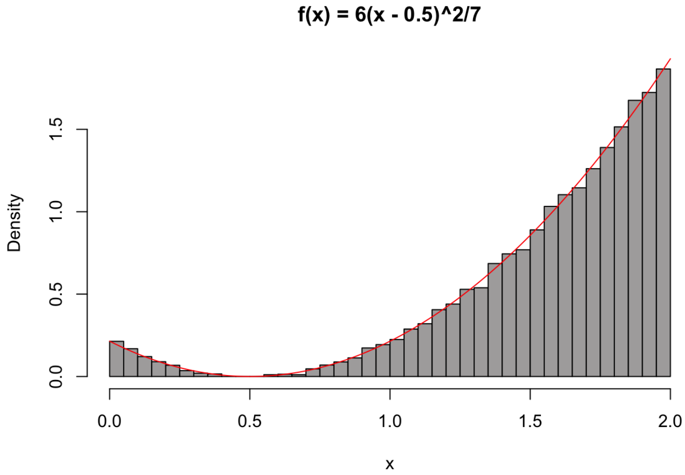
$$= M\int_{-\infty}^{y} P\left(U \leq \frac{f(Y)}{Mg(Y)}\middle| Y = \omega \leq y\right) g(\omega)d\omega$$

$$= M\int_{-\infty}^{y} \frac{f(\omega)}{Mg(\omega)} g(\omega)d\omega = F(y)$$

An Implementation of the Algorithm in R

```r
N <- 50000
M <- 3.858
y <- runif(N, min = 0, max = 2)
u <- runif(N, min = 0, max = 1)
gy <- 0.5
fy <- 6 * (y - 0.5)^2 / 7
x <- y[u < fy / gy / M]
sample <- length(x)
hist(x, breaks = 50, freq = FALSE, col = "#adabab",
main = "f(x) = 6(x - 0.5)^2/7")
curve(6 * (x - 0.5)^2 / 7, from = 0,
      to = 2, add = TRUE, col = "red")
```

Output



Pros and Cons

Pros

- Deal with almost every distribution.
- Easy to handle.

Cons

- Lack of efficient.
- Do not perform perfectly with unbounded PDF, i.e. Beta distribution.
- Do not work well with PMF.

Outline

Background

Property of Markov Chain:

- $P(X_n = x_n | X_{n-1} = x_{n-1} \cdots X_1 = x_1) = P(X_n = x_n | X_{n-1} = x_{n-1})$ Next state independent of the past states and only depends on the present state.

Background

Property of Markov Chain:

- $P(X_n = x_n | X_{n-1} = x_{n-1} \cdots X_1 = x_1) = P(X_n = x_n | X_{n-1} = x_{n-1})$ Next state independent of the past states and only depends on the present state.
- Homogeneous: Transition matrix is independent with time.
- Stationary distribution: $\pi = \pi P$. Every finite positive homogeneous chain have unique Stationary distribution, which is given by *Perron-Frobenius* theorem. The stationary distribution does not depend on initial distribution.

MCMC Sampling: Metropolis Algorithm

Suppose we want to generate a group of sample from the object PMF $f(\cdot)$

1. Initialization: Choose arbitrary value x_0, proposal markov chain with transition probability $g(\cdot|\cdot)$ and set $t = 1$.
2. Generate $x_t^* \sim P(x_t^*|x_{t-1})$ and $u \sim \text{Unif}[0,1]$.
3. Compute

$$h(x_{t-1}, x_t^*) = \min\left\{1, \frac{f(x_t^*)g(x_{t-1}|x_t^*)}{f(x_{t-1})g(x_t^*|x_{t-1})}\right\}$$

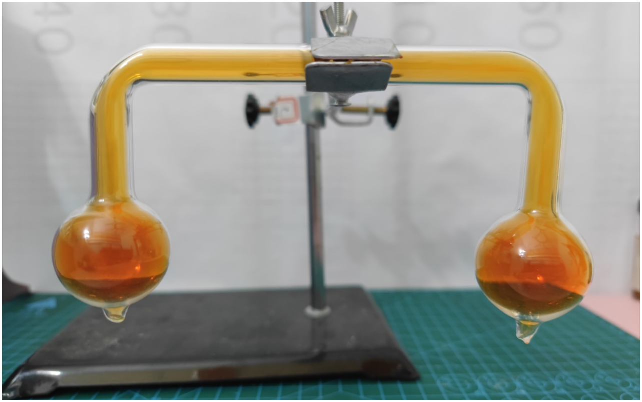
4. If $u < h(x_{t-1}, x_t^*)$, let $x_t = x_t^*$; Otherwise, let $x_t = x_{t-1}$
5. Let $t++$. Back to the second step unless reach stationary distribution.

Intuition

Immigration problem.

- To keep the population ratio stable.
- Randomly reject some visa with specific probability.
- Finally, reach a balance.
- Aribtrary start does not affect.
- Dynamic balance.

Another Example



Proof of Metropolis Algorithm

Detailed balance condition:

$$\delta_{xy} = m(x)P(x \to y) - m(y)P(y \to x) = 0$$

One need to show that $m(x) \propto f(x)$.

$$\frac{m(x)}{m(y)} = \frac{P(y \to x)}{P(x \to y)} = \frac{h(x,y)}{h(y,x)} = \frac{f(x)}{f(y)}$$

Simulation: Throwing Two Dice

The object PMF is:

x	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	1	2	3	4	5	6	5	4	3	2	1

We apply minimum neighborhood method:

$$G = \begin{pmatrix} 1/2 & 1/2 & 0 & \cdots & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & \cdots & 0 & 0 & 0 \\ 0 & 1/2 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1/2 & 0 \\ 0 & 0 & 0 & \cdots & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & \cdots & 0 & 1/2 & 1/2 \end{pmatrix}.$$

Thanks