Monte Carlo Algorithm and its Application

Applied Math Tutorial

Tianlu Zhu

Shanghaitech Unversity

February 23, 2023

Outline

Introduction

In the presentation, I would first introduce Monte Carlo Integration, which requires enormous random sample. Then we introduce common ways of sampling from a given distribution, including Transformation method, Accept-rejection method, and Markov Chain Monte Carlo Algorithm. In each method I would show code implementation, analyse the pros and cons, and a brief proof.

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

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

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 $F(\cdot)$.

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

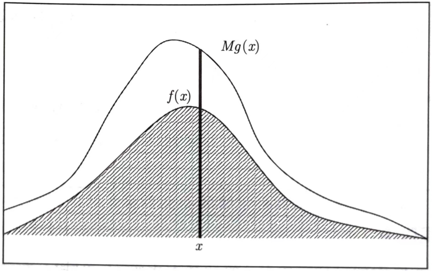
- Naive way to generate a group of sample

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 $x \in \Omega$

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,

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

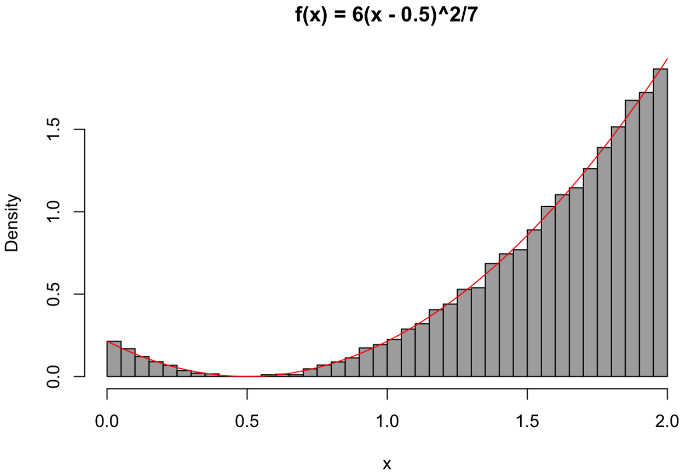
$$= M \int_{-\infty}^{y} P\left(U \leq \frac{f(Y)}{Mg(Y)}\middle| Y = \omega \leq y\right) g(\omega) d\omega$$

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

An Implementation of the Algorithm in R

```
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
- ?

Cons

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

Outline

Background

Property of Markov Chain:

- $P(X_n = x_n | X_{n-1} = x_{n-1} \cdots X_1 = x_1) = P(X_n = x_n | X_{n-1} = x_{n-1})$ Only depends on last state.
- Invertible: If $X_1, X_2 \cdots X_n$ is MC, so do $X_n, X_{n-1} \cdots X_1$.
-

Columns

This is a text in first column.

$$E = mc^2$$

- First item
- Second item

first block

columns achieves splitting the screen

second block

stack block in columns

Create Tables

first	second	third
1	2	3
4	5	6
7	8	9

Equation1

A matrix in text must be set smaller: $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ to not increase leading in a portion of text.

$$f(n) = \begin{cases} n/2 & \text{if } n \text{ is even} \\ -(n+1)/2 & \text{if } n \text{ is odd} \end{cases}$$

$$50 apples \times 100 apples = lots of apples^2$$

Equation2

$$\sum_{\substack{0<i<m \\ 0<j<n}} P(i,j) = \int_a^b \prod P(i,j)$$

$$P\left(A = 2 \middle| \frac{A^2}{B} > 4\right)$$

$$(a), [b], \{c\}, |d|, \|e\|, \langle f \rangle, \lfloor g \rfloor, \lceil h \rceil, \ulcorner i \urcorner$$

Equation3

$$Q(\alpha) = \alpha_i \alpha_j y_i y_j (x_i \cdot x_j)$$

$$Q(\alpha) = \alpha^i \alpha^j y^{(i)} y^{(j)} (x^i \cdot x^j)$$

$$\Gamma = \beta + \alpha + \gamma + \rho$$

Outline

End

The last page.