1 Abstract

- Using Monte Carlo to approximate π;
- Introduce Monte Carlo basics.

2 Problem

Approximate the value π.

3 Analysis

Consider the following question:

- You shoot a square $(-1,1)^2$. Suppose your shot is uniform in this square, then what is the probability you have a successful shot? We say "your shot is successful", if your shot belongs to the unit ball B_1.

The answer is

$$\text{Prob of succesful shot } = \frac{\text{Area of } B_1}{\text{Area of } (-1,1)^2} = \frac{\pi}{4}.$$

This means that, as long as one can approximate probability of successful shot, one can approximate π by multiplying 4. This can be done by computer:

Algorithm 1 MC estimation of π

procedure MCPI(N) ▷ N is total number of samples
  $n \leftarrow 0$ ▷ n is number of hits
  for $i = 1...N$ **do**
    generate two numbers X, Y from $U(-1,1)$
    if $X^2 + Y^2 < 1$ **then** $n \leftarrow n + 1$
  return $\frac{4n}{N}$

Example 1 • *Using Algo1, design estimator* $\hat{\pi}(N)$ *and compute* $\hat{\pi}(10000)$.

4 Monte Carlo basics

4.1 Bias and MSE

One can implement above approximation multiple times and observe that

- (random estimator) Target value π is deterministic, but each implementation gives different outcome $\hat{\pi}$;
- (Convergence) Each obtained outcome, as long as N is large enough, gives some close approximation.

We are going to generalize our observations in this below.

- A random estimator $\hat{\alpha}$ to a deterministic value α is called as Monte Carlo (MC) approximation.
- Moreover, we define
$$Bias = \mathbb{E}[\hat{\alpha}] - \alpha$$
and
$$MSE = \mathbb{E}[(\hat{\alpha} - \alpha)^2].$$
- (def) If Bias is zero, then we call this as *unbiased* MC.

Proposition 1 *$MSE(\hat{\alpha}) = |Bias(\hat{\alpha})|^2 + Var(\hat{\alpha})$. In particular, if $\hat{\alpha}$ is unbiased, then MSE is Variance.*

PROOF: ... □

Although seemingly absurd, we consider the above estimator with $N = 1$, which is equivalent to

- Consider
$$\hat{\alpha} = 4I(X_1^2 + Y_1^2 < 1), \ X_1, Y_1 \sim U(-1, 1)$$
as MC for π. Then the outcome is either 0 or 4. In any case, it is a bad approximation.
- However, we can show that it's an unbiased MC. (why?)
- Find MSE?

4.2 Ordinary Monte Carlo

Unbiased MC is very desirable, because one can employ crude (ordinary) MC to make it more accurate: [1]

- Suppose $\hat{\alpha}$ is a square integrable unbiased MC;
- Obtain N independent replicates
$$\{\hat{\alpha}_i : i = 1, \dots, N\}.$$
- Taking their average, it gives a new MC:
$$\beta_N = \frac{1}{N} \sum_{i=1}^{N} \hat{\alpha}_i.$$
- β_N is unbiased again. (why?)
- $MSE(\beta_N) = \frac{1}{N} MSE(\hat{\alpha}) \to 0$. (why?)
- β_N is *almost surely consistent*, (why?) i.e.
$$\beta_N \to \alpha, \ \text{almost surely or } \mathbb{P}(\lim_N \beta_N = \alpha) = 1.$$

[1] We say a random variable X is in L^p, if its pth moment exists, i.e. $\mathbb{E}|X|^p < \infty$. If $X \in L^2$, then we say it's square integrable.

- β_N is L^2-consistent, (why?) i.e.

$$\beta_N \to \alpha \text{ in } L^2 \text{ or } \mathbb{E}(\beta_N - \alpha)^2 \to 0.$$

As a conclusion, one can always use crude MC to make better approximation provided there exists an unbiased MC $\hat{\alpha}$, which is obtainable in sacrifice of higher computational cost. Given a fixed amount of computational cost, to improve the efficiency, it is essential to reduce $Var(\hat{\alpha})$ as much as possible.

Proposition 2 *Prove that both almost sure and L^2 consistency implies consistency in probability.*

Example 2 *Consider α_n is a sequence of estimators to the value α. Prove that, if $MSE(\alpha_n) \to 0$, then α_n is L^2 consistent to α.*

Example 3 *Given i.i.d $\{\alpha_i : i \in 1, 2, \dots, M\}$, we use*

$$\bar{\alpha}_M = \frac{1}{M} \sum_{i=1}^{M} \alpha_i$$

as its estimator of the mean $\mathbb{E}[\alpha_1]$ and use

$$\beta_M = \frac{1}{M} \sum_{i=1}^{M} (\alpha_i - \bar{\alpha}_M)^2$$

as the estimator of $Var(\alpha_1)$. Suppose $\alpha_1 \in L^4$, then

- *Prove β_M is biased.*
- *(optional) Prove that β_M is consistent in L^2.*
- *Can you propose an unbiased estimator?*

Example 4 *Our goal is to estimate $MSE(\hat{\pi}_N)$ for $\hat{\pi}_N$ of Example 1. Since $\hat{\pi}_N$ is unbiased, its MSE is equal to its variance $Var(\hat{\pi}_N)$.*

- *Use β_{100} of Example 3 to estimate $MSE(\hat{\pi}_N)$ after collecting $\{\hat{\pi}_{N,i} : i = 1, \dots, 100\}$. One must write both pseudocode and python code.*
- *Repeat above estimation of $MSE(\hat{\pi}_N)$ for $N = 2^i : i = 5, ...10$ and plot log-log chart.*