

Монте Карло Интеграција

4.час

22. март 2016.

Монте Карло методе

Ове нумеричке методе код којих се употребљавају низови случајних бројева за извршење симулације. Првобитно су познати као "статистичка упрошћавања", али назив "Монте Карло", популаризован од стране првих истраживача у овој области је проистекао из назива чувеног казина у Монаку.

Како су за добијање довољно тачне оцене тражене величине, потребна израчунавања за веома велики број посебних случајева и одговарајућа статистичка обрада огромног нумеричког материјала, то је ефективна примена методе Монте Карло омогућена тек појавом електронских рачунара.

Монте Карло метода захтева да се физички систем опише густинама вероватноћа. Када су познате ове функције, Монте Карло симулација се наставља случајним избором вредности функција. Потом се изврше многе симулације, а за решење се узима просечан резултат свих симулација.

При решавању различитих проблема код којих је тешко доћи до аналитичких израза користе се рачунске методе, помоћу моделирања случајних величина и статистичког оцењивања карактеристика тих величина.

Неке од области примене методе Монте Карло су:

- биологија
- генетика
- екологија
- хидрологија
- статистика
- физика
- системи масовног опслуживања

Проблеми који се срећу у разним областима се могу свести на математичке проблеме:

- решавање система линеарних једначина или неједначина
- рачунање интеграла
- решавање диференцијалних једначина
- решавање парцијалних диференцијалних једначина

Сваки од наведених математичких задатака се може решити и методом Монте Карло, што се нарочито користи кад је теоријско решење сувише компликовано или не може да се одреди, иако се зна да постоји.

Методама Монте Карло се могу решити и неки задаци у којима се класичне методе нумеричке математике не могу применити.

Оно што је такође значајно за алгоритме Монте Карло јесте да су обично једноставни и лаки за програмирање.

Интеграција Монте Карло

Интеграција Монте Карло је Монте Карло метода којом се (приближно) нумерички израчунава дати интеграл. Примењује се када је интеграл сложен и аналитички тежак или немогућ за израчунавање.

За приближно израчунавање интеграла

$$I = \int_a^b g(x)dx$$

користе се две Монте Карло методе:

- 1 **Монте Карло метода погодака и промашаја** (*The Hit and Miss Monte Carlo Method*) која је заснована на геометријској интерпретацији интеграла као површине;
- 2 **Монте Карло метода узорачке средине** (*The Sample Mean Monte Carlo Method*) која је заснована на интерпретацији интеграла као средње вредности.

Монте Карло метода погодака и промашаја (The Hit and Miss Monte Carlo Method)

Нека је функција $g(x)$ ограничена $0 \leq g(x) \leq c$, $a \leq x \leq b$ и означимо са D област

$$D = \{(x, y) : a \leq x \leq b, 0 \leq y \leq c\}.$$

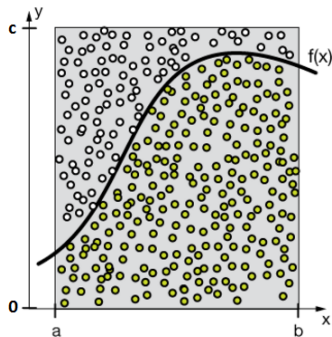
Нека је (X, Y) случајни вектор са униформном расподелом на правоугаонику D са густином расподеле

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{c(b-a)}, & (x, y) \in D \\ 0, & (x, y) \notin D \end{cases}$$

Монте Карло метода погодака и промашаја (The Hit and Miss Monte Carlo Method)

Означимо са S површ испод криве $g(x)$, $S = \{(x, y) : y \leq g(x)\}$.
Вероватноћа да се случајни вектор нађе испод криве $g(x)$ је тада једнака

$$p = \frac{\int_a^b g(x) dx}{c(b-a)} = \frac{I}{c(b-a)}$$



Монте Карло метода погодака и промашаја (The Hit and Miss Monte Carlo Method)

Ако претпоставимо да је генерисано N независних случајних вектора $(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$, на основу закона великих бројева, вероватноћа p се може оценити са:

$$\hat{p} = \frac{N_H}{N},$$

где је:

N_H број погодака тј. број случајева када је $Y_i \leq g(X_i)$, $i = 1, 2, \dots, N$
 $N - N_H$ број промашаја тј. број случајева када је $Y_i > g(X_i)$, $i = 1, 2, \dots, N$.

Из претходног закључујемо да се интеграл I може оценити као:

$$I \approx \theta = c(b - a) \frac{N_H}{N},$$

Монте Карло метода погодака и промашаја (The Hit and Miss Monte Carlo Method)

Сваки од N покушаја има Бернулијеву расподелу са вероватноћом p , па случајна величина N_H има Биномну расподелу са параметрима (N, p) .

Математичко очекивање и дисперзија оцене θ су:

$$E(\theta) = \frac{c(b-a)}{N} E(N_H) = c(b-a)p = I$$

$$D(\theta) = \frac{c^2(b-a)^2}{N^2} D(N_H) = c^2(b-a)^2 \frac{p(1-p)}{N}$$

Треба приметити да је оцена θ непристрасна, а пошто $D(\theta) \rightarrow 0$, када $N \rightarrow \infty$, оцена θ је и постојана.

Монте Карло метода узорачке средине (The Sample Mean Monte Carlo Method)

Број потребних експеримената N се може одредити тако да је:

$$P\{|\theta - I| \leq \epsilon\} = \alpha,$$

где је:

ϵ – грешка апроксимације,

α – ниво поверења, (обично се узима $\alpha = 95\%$ или $\alpha = 99\%$).

За довољно велико N , применом централне граничне теореме имамо:

$$\theta^* = \frac{\theta - I}{\sigma_\theta} \sim \mathbf{N}(0, 1).$$

Тада се добија $P\{|\theta^*| \leq z_\alpha\} = \alpha$ где z_α налазимо тако да је $F(z_\alpha) = \frac{1+\alpha}{2}$, где је F функција нормалне расподеле.

Користећи апроксимацију $p(1-p) \approx \frac{1}{4}$, добијамо:

$$N \geq \frac{c^2(b-a)^2 z_\alpha^2}{4\epsilon^2}.$$

Алгоритам за Монте Карло методу погодака и промашаја

- 1 Генерисати низ $(U_j)_{j=1}^{2N}$ од $2N$ псеудослучајних бројева.
- 2 Поређати бројеве у N парова $(U_1, U'_1), (U_2, U'_2), \dots, (U_N, U'_N)$ на било који начин тако да се сваки број U_i појави тачно једанпут.
- 3 Израчунати $X_i = a + (b - a)U_i$, $g(X_i)$ и $Y_i = cU'_i$ за $i = 1, 2, \dots, N$.
- 4 Пребројати број погодака N_H за које важи $Y_i \leq g(X_i)$.
- 5 Оценити интеграл I са

$$\theta = c(b - a) \frac{N_H}{N}$$

Монте Карло метода погодака и промашаја - пример

Користећи Монте Карло методу погодака и промашаја израчунати интеграл

$$I = \int_0^2 e^{-x^2} dx.$$

За грешку апроксимације узети $\epsilon = 0,001$ и ниво поверења $\alpha = 0,95$.

Монте Карло метода погодака и промашаја - пример

```
integral_e1 <- function(){  
  c<-1 # jer je podintegralna funkcija manja od 1  
  a<-0  
  b<-2 # granice integrala  
  epsilon <- 0.001 # greska aproksimacije  
  alfa <- 0.95 # nivo poverenja  
  z <- qnorm((1+alfa)/2)  
  N <- trunc(((c*(b-a)*z)^2)/(4*epsilon^2))+1 # obim uzorka  
  U <- runif(N) #generisemo niz od N clanova  
  V <- runif(N) #generisemo niz od N clanova  
  X <- a+(b-a)*U  
  Y <- c*V  
  Nh <- sum(Y<=exp(-X^2))  
  teta1 <- c*(b-a)*Nh/N  
  return (teta1)  
}
```

```
> integral_e1()
```

```
[1] 0.8829455
```

Монте Карло метода узорачке средине (The Sample Mean Monte Carlo Method)

Интеграл $I = \int_a^b g(x)dx$ се може представити као очекивана вредност неке случајне величине. Прво напишимо интеграл на следећи начин:

$$I = \int_a^b \frac{g(x)}{f_X(x)} f_X(x) dx.$$

Затим претпоставимо да је $f_X(x)$ произвољна густина расподеле таква да је $f_X(x) > 0$ када $g(x) \neq 0$. Тада је:

$$I = E \left(\frac{g(X)}{f_X(X)} \right),$$

где је X случајна величина са густином расподеле $f_X(x)$.

Монте Карло метода узорачке средине (The Sample Mean Monte Carlo Method)

Узмимо, једноставности ради, да

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & x \in (a, b) \\ 0, & x \notin (a, b) \end{cases}$$

Тада имамо да је $E(g(x)) = \frac{I}{b-a}$, односно

$$I = (b - a)E(g(X)).$$

Оцена за интеграл I је узорачка средина

$$\theta = (b - a) \cdot \frac{1}{N} \sum_{i=1}^N g(X_i).$$

Монте Карло метода узорачке средине (The Sample Mean Monte Carlo Method)

Математичко очекивање и дисперзија оцене θ су:

$$E(\theta) = I$$

$$D(\theta) = \dots = \frac{1}{N} \left((b-a) \int_a^b g^2(x) dx - I^2 \right)$$

Треба приметити да је оцена θ непристрасна, а пошто $D(\theta) \rightarrow 0$, када $N \rightarrow \infty$, оцена θ је и постојана.

Алгоритам за Монте Карло методу узорачке средине

- 1 Генерисати низ $(U_j)_{j=1}^N$ од N псеудослучајних бројева.
- 2 Израчунати $X_i = a + (b - a)U_i$, $i = 1, 2, \dots, N$.
- 3 Израчунати $g(X_i)$, $i = 1, 2, \dots, N$.
- 4 Оценити интеграл узорачком средином

$$\theta = (b - a) \frac{1}{N} \sum_{i=1}^N g(X_i).$$

Монте Карло метода узорачке средине - пример

Користећи Монте Карло методу узорачке средине израчунати интеграл

$$I = \int_0^2 e^{-x^2} dx.$$

Монте Карло метода узорачке средине - пример

```
integral_e2 <- function(){  
  a<-0  
  b<-2 # granice integrala  
  N <- 10000 #obim uzorka  
  U <- runif(N) #generisemo N slucajnih brojeva  
  X <- a+(b-a)*U  
  teta2 <- (b-a)*1/N*sum(exp(-X^2))  
  return(teta2)  
}  
  
> integral_e2()  
[1] 0.8824376
```

Интеграција у R-у

R нуди две сличне функције које рачунају интеграле, `area` (у пакету `MASS`) и `integrate`.

Међутим, `area` не ради за интеграле са бесконачном границом и стога захтева одређено предзнање из интеграције. Друга функција, `integrate`, прихвата и ради са бесконачним границама али је веома осетљива и може да произведе непоуздан излаз.

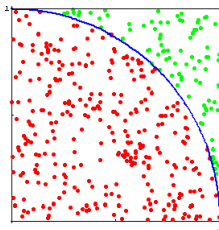
```
> area(function(x)exp(-x^2), 0, 2)
[1] 0.8820814
```

```
> integrate(function(x)exp(-x^2), 0, 2)
0.8820814 with absolute error < 9.8e-15
```

Апроксимација броја π

Ако насумице изаберемо тачку у равни унутар јединичног квадрата $[0, 1] \times [0, 1]$, вероватноћа да ће се она наћи унутар јединичне кружнице је однос површине четвртине круга и укупне површине квадрата.

$$p = \frac{P_{\text{deoKrug}}}{P_{\text{kvadrata}}} = \frac{\frac{1^2\pi}{4}}{1^2} = \frac{\pi}{4}$$



Закон великих бројева нам говори да ће однос броја тачака које су завршиле унутар јединичне кружнице и укупног броја тачака бити приближно једнака траженој вероватноћи.

Задаци

1.(2 поена) Користећи Монте Карло методе (методу погодака и промашаја и методу узорачке средине) израчунати интеграл

$$I = \int_0^1 (\sin(20x) + \cos(50x))^2 dx$$

и добијени резултат упоредити са резултатом добијемо помоћу неке од функција у R-у.

2.(3 поена) Написати функцију у R-у за одређивање децимала броја π .