HW4_344

Zeqiu.Yu

2022-10-15

Homework 4

Question1

Suppose we want to estimate $E[X^2]$ where X has the density that is proportional to $q(x) = e^{\frac{-|x|^3}{3}}$.
Using adaptive squeezed rejection sampling by setting k=3 at x=-1,0, 1 for construction the upper and inner piece hull (envelope and squeezing functions). Attach your R code and output. Be sure to count and report your acceptance ratio.

Solution: Because f(x) is log-concave, continuous and diffrenciable, we have f(X) $\propto$ q(x), define f(x) and l(x) as follows:

```
f <- function(x){
  exp(-abs(x)^3/3)
}
l <- function(x){
  -abs(x)^3/3
}
```

Then we can get

$$e(x) = \begin{cases} e^{x+\frac{2}{3}} & x \in (-\infty, -\frac{2}{3}] \\ 1 & x \in (-\frac{2}{3}, \frac{2}{3}) \\ e^{-x+\frac{2}{3}} & x \in [\frac{2}{3}, \infty) \end{cases}$$

and

$$s(x) = e^{-\frac{1}{3}|x|} \qquad x \in [-1, 1]$$

```
fe <- function(x){
  ifelse(abs(x) >= 2/3, exp(-abs(x) +2/3 ), 1)
}
fs <- function(x){
  ifelse(abs(x)<=1, exp(-1/3*abs(x)), -1)
}
```

To sample from the inverse and get g(x), integrate e(x).

```
integrate(fe, lower=-Inf,upper=Inf)$value
```

```
## [1] 3.333334
```

So $g(x) = \frac{e(x)}{10/3}$ and

$$G(y) = \begin{cases} \frac{3}{10}e^{y+\frac{2}{3}} & y \in (-\infty, -\frac{2}{3}] \\ \frac{1}{2} + \frac{3}{10}y & y \in (-\frac{2}{3}, \frac{2}{3}) \\ 1 - \frac{3}{10}e^{-y+\frac{2}{3}} & y \in [\frac{2}{3}, \infty) \end{cases}$$

To sample g(y) we use the inverse of G(y):

```
G.inverse <- function(y){
  if((y>=0)&(y<=3/10)){
    log(10/3*y) - 2/3
  }else if((y>3/10)&(y<=7/10)){
    10/3*(y - 1/2)
  }else{2/3 - log(10/3*(1-y))}
  }
```

```
sSizes <- c(10000, 100000, 1000000)

for(sSize in sSizes){
  i = 0
  j = 0
  y <- c()
  while(i<=sSize){
    j = j + 1
    y_i.u <- runif(1)
    y_i <- G.inverse(y_i.u)
    ey_i <- fe(y_i)
    s_y_i <- fs(y_i)
    f_y_i <- exp(-abs(y_i)^3/3)
    criterion.u <- runif(1)
    if(criterion.u<=s_y_i/ey_i){
      y[i] <- y_i
      i = i+1
    }else if(criterion.u<=f_y_i/ey_i){
      y[i] <- y_i
      i = i+1
    }
    else{
      next
    }
  }
  estimation = sum(y^2)/sSize
  print(sprintf("After obtaining %d numbers from f(x), E[X^2] is estimated to be %.3f", sSize, estimati
  print(sprintf("The acceptance ratio is: %.2f", sSize/j))
}
```

```
## [1] "After obtaining 10000 numbers from f(x), E[X^2] is estimated to be 0.779"
## [1] "The acceptance ratio is: 0.77"
## [1] "After obtaining 100000 numbers from f(x), E[X^2] is estimated to be 0.776"
## [1] "The acceptance ratio is: 0.77"
## [1] "After obtaining 1000000 numbers from f(x), E[X^2] is estimated to be 0.775"
## [1] "The acceptance ratio is: 0.77"
```

Question2 (Master ~)

Consider to use a normal distribution (g(x) in lecture notes) with mean 0 and variance σ^2 for importance sampling to estimate $S = E[X^2]$. Try σ^2 on a grid from 0.5,1,1.5,. . . ,5 and plot the variance reduction as a function of σ^2 value compared to the regular Monte Carlo estimation. NOTE: since the target distribution only is proportional to q(x), you will need to use the normalized weight in the estimation. In addition, for each given , you could estimate the variance of importance sampling estimator by repeating the important sampling many times (say 50 times), and calculate the sample variance of the importance sampling estimators.

Please make sure to use different seeds in different runs.

Solution: Because $g(x) \sim N(0, \sigma^2)$, we sample x according to different sigma values. First, sample from q(x) and get the variance.

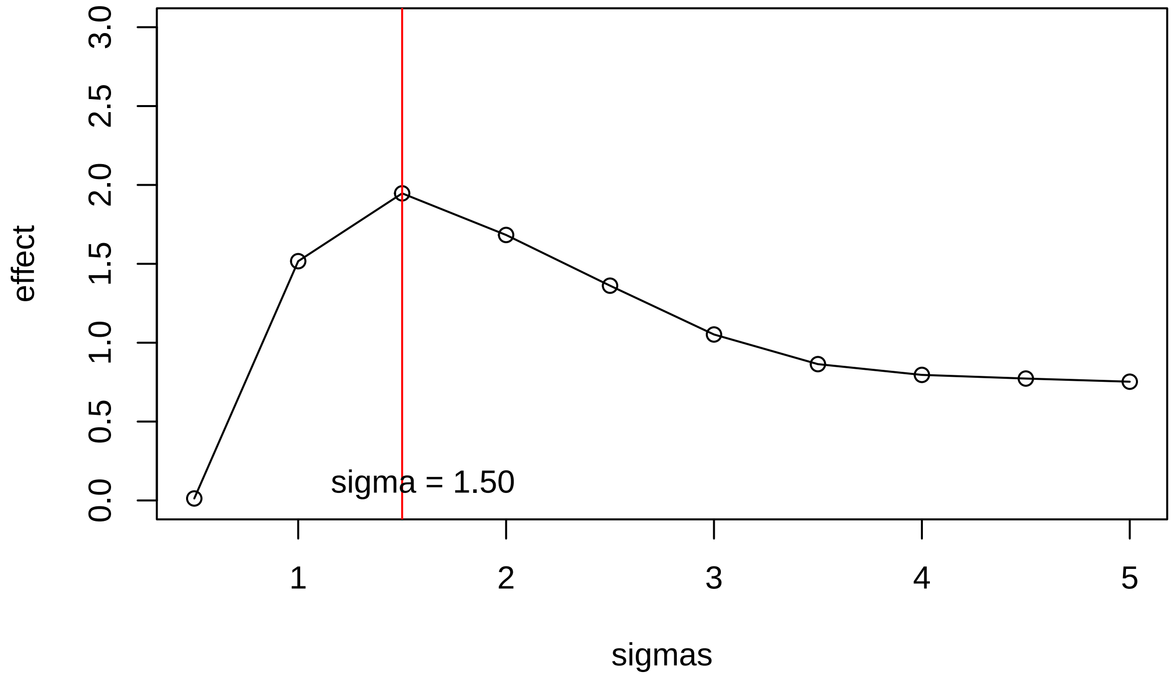
```r
deci <- function(){
  y_i <- rnorm(1)
  q_y <- exp(-abs(y_i)^3/3)
  e_y <- exp(-y_i^2/2 + 1/6)
  u <- runif(1)
  if(u<=q_y/e_y){
    return(y_i)
  }else{return(NULL)}
}

SampleSize <- 100000
i = 0
y <- c()
while(i<=SampleSize){
  y_i <- deci()
  if(is.null(y_i)){
    next
    }else{
      y[i] <- y_i
      i = i+1
    }
}
mu = sum(y^2)/SampleSize
v0 <- var(y^2)/SampleSize
```

```r
effect <- c()
sigmas <- seq(0.5,5,0.5)
for(i in 1:length(sigmas)){

  mus <- c()
  for(j in 1:50){
    set.seed(j)
    sigma <- sigmas[i]
    x <- rnorm(SampleSize, 0, sigma)
    wts <- exp(-abs(x)^3/3)/dnorm(x,0,sigma)
    weighted.wts <- wts/sum(wts)
    mu.is = sum(x^2*weighted.wts)
    mus[j] <- mu.is
  }
  v1 <- var(mus)
  effect[i] <- v0/v1
  }

plot(sigmas,effect,type="o",ylim=c(0,3))
abline(v=sigmas[which.max(effect)], col = "red")
text(sigmas[which.max(effect)] + 0.1 ,0.1,sprintf("sigma = %.2f",sigmas[which.max(effect)]))
```



Question3 (Master ~)

Can you comment how to calculate the theoretical variance of the importance sampling estimate of $S = E[X^2]$ for a given g(x) in 2? If so, choose one σ^2, repeat important sampling estimation for 50 times, calculate the sample variance of the importance sampling estimator and compare it to the theoretical value. Comment on your findings. (5 pts for STAT PhD students.)

Solution: We start from f(x). For $E[x^2]$, $\mu = \int_X h(x)\frac{f(x)}{g(x)}g(x)\ dx$. Then for q(x),we normalize it is actually $\mu = \int_X h(x)\frac{q(x)}{cg(x)}g(x)dx = \int_X h(x)\frac{q(x)}{\int_X \frac{q(x)}{g(x)}g(x)dxg(x)}g(x)dx$.That is,

$$E_f[c] = \frac{1}{n}\sum_{i=1}^{n_{sample}}\frac{q(x_i)}{g(x_i)} \quad (unbiased)$$

$$\sigma^2 = \int_X h(x)\frac{q(x)}{cg(x)}g(x)dx = \int_X [(h(x)\frac{q(x)}{\int_X \frac{q(x)}{g(x)}g(x)dxg(x)})^2 - \mu^2]g(x)dx = \int_X (\frac{h(x)q(x) - \mu\int_X \frac{q(x)}{g(x)}g(x)dxg(x)}{\int_X \frac{q(x)}{g(x)}g(x)dxg(x)})^2 g(x)dx$$

$$\hat{\mu} =$$

And for μ, it is also unbiased. Hence, we calculate the theoretical value by with respect to $\sigma = 1.5$: