Slides used in Lab6

Raul Kangro

March 13, 2017

Variance reduction methods

- If we keep α fixed (say $\alpha = 0.05$), then the error estimate of MC goes to 0 as $\frac{C}{\sqrt{n}}$ for a constant $C = \Phi^{-1}(\alpha/2)\sigma_Y$ as n tends to infinity
- can not change the $\sqrt{n}$ part, but can try to change C
- Want to change Y to Z so that $\mathbb{E}Y = \mathbb{E}Z$, but $\text{Var}(Z) < \text{Var}(Y)$
- Effect on computation time: if we fix the accuracy ε, then the number of generated values is

$$n = \frac{C^2}{\varepsilon^2} = \frac{\Phi^{-1}(\alpha/2)^2 Var(Y)}{\varepsilon^2}$$

  so the computation time is proportional to variance of Y

Variance reduction methods (cont.)

- We'll consider 4 general methods
- Antithetic variates (this lab)
- Control variates (this lab)
- Importance sampling
- Stratified sampling

Antithetic variates

- Suppose we can generate together Y and $\tilde{Y}$ so that they have the same distribution
- Then $Z = \frac{Y+\tilde{Y}}{2}$ has the same expected value as Y
- Compute

$$\text{Var}(Z) = \frac{1}{4}(\text{Var}(Y) + \text{Var}(\tilde{Y}) + 2\,\text{cov}(Y, \tilde{Y})) = \text{Var}(Y)(1+\rho)/2$$

- If $\rho < 1$, then $\text{Var}(Z) < \text{Var}(Y)$, but in order to get one value of Z we have to generate Y and $\tilde{Y}$
- Can assume that generating a value of Z takes twice the time of generating Y only
- To gain in speed, we need $\text{Var}(Z) < \frac{\text{Var}(Y)}{2}$, so $\rho < 0$

Antithetic Variates (cont.)

- Y and $\tilde{Y}$ which have the same distribution, but are negatively correlated, are called *Antithetic variates*
- How to get dependent random variables from the same distribution ?
- If pdf of X is symmetric with respect to $x = \mu$, then X and $\tilde{X} = 2\mu - X$ have the same distribution
- Thus also $Y = g(X)$ and $\tilde{Y} = g(\tilde{X})$ have the same distribution
- Often (but not always) are they also negatively correlated
- if Y is computed by using $X_1, \ldots, X_m$, iid, from symmetric distribution, then $\tilde{Y}$ is computed by using $\tilde{X}_1 = 2\mu - X_1,\ \tilde{X}_1 = 2\mu - X_2, \ldots$

Implementing Antithetic Variates (AV)

- Suppose we have implemented MC for computing $\mathbb{E}Y$ for $Y = g(X)$
- We have generateor for X: for given n we get n values of X
- We have defined g(x)
- In order to implement AV we have to:
- change the generator so that for given n returns $n \times 2$ matrix, in the first column X, in the second column $\tilde{X}$
- define a function g_1:

```
g_1=function(x){
  return((g(x[,1])+g(x[,2]))/2)
}
```

- use MC1 or MC2 with the new generator and g_1

Implementing AV (example)

consider the problem of computing $E(\frac{X+X^3}{1-0.1\cdot\sqrt{X}})$ for $X \sim U(0,2)$ with MC with error less than 0.02 at the significance level 0.05 Define generator

```
gen=function(n){
  return(runif(n,min=0,max=2))
}
```

Define function g:

```
g=function(x){
  return((x+x^3)/(1-0.1* sqrt(x)))
}
```

Implementing AV (example)

Compute the result

```
set.seed(19)
MC2(gen,g,0.02,0.05)
```

```
## [1] 3.416022e+00 1.030000e+05
```

Implementing AV (example)

For AV:

```
gen2=function(n){
  X=runif(n,min=0,max=2)
  Xtilde=2-X
  return(cbind(X,Xtilde))
}
```

Define g_1 and compute the result

```
g_1=function(x){
      return((g(x[,1])+g(x[,2]))/2)
}
set.seed(19); MC2(gen2,g_1,0.02,0.05)
```

```
## [1]     3.406248 11000.000000
```

Implementing AV (example)

- results: without AV needed 103000 random variables
- with AV 2*11000=22000 random variables
- speedup factor 4.6818182
- Now it is your turn to practice!

Task 1 result

if `set.seed(10)` is used before computations, then the answer should be

```
## [1]     10.88179 325000.00000
```

Control variates (CV)

- Suppose we can use the same random variables, that were used to generate Y, to generate another random variable W, for which $\mathbb{E}W$ is known
- Then we can define
$$Z = Y - a \cdot (W - \mathbb{E}W)$$
- We have $\mathbb{E}Z = \mathbb{E}Y$ for all values of a
- The variance of Z can be written as
$$\text{Var}(Z) = \text{Var}(Y) + a^2\text{Var}(W) - 2a\text{Cov}(Y, W)$$
- The smallest variance corresponds to $a = \frac{\text{cov}(Y,W)}{\text{Var}(W)}$
- For this a, we have
$$\text{Var}(Z) = \text{Var}(Y)(1 - \text{Corr}(Y, W)^2).$$

Applying CV in R

- Find W that can be computed by the *same random variables* that are used for generating Y and for which $\mathbb{E}W$ *is known* (or can be easily found) and which has *high correlation with* Y
- Modify the generator for X so that it also generates W or $\tilde{X}$ from which W can be computed. So that the output of the generator has at least one additional column.
- Estimate the parameter a: use the generator to generate a moderate number (for example, 1000) of Y and W values, compute `a=cov(Y,W)/var(W)` or fit a linear model to Y with regressor W and take a to be the coefficient of W.
- define `g_cv=function(x){}` so that it computes values of $Z = Y - a \cdot (W - \mathbb{E}W)$ if x is the output of the generator
- use `MC1` or `MC2` to compute the answer

Applying CV (example)

- same problem: compute $E(\frac{X+X^3}{1-0.1\cdot\sqrt{X}})$ for $X \sim U(0,2)$ with MC with error less than 0.02 at the significance level 0.05
- It looks like $W = X + X^3$ could be a good predictor
- Modify the generator

```
gen_cv=function(n){
  X=runif(n, min=0,max=2)
  W=X+X^3
  return(cbind(X,W))
}
```

Applying CV (example)

- Estimate the coefficient a

```
tmp=gen_cv(1000)
Y=g(tmp[,1])
W=tmp[,2]
a=cov(Y,W)/var(W)
a
```

```
## [1] 1.161855
```

- The correlation is 0.9999345 so we may expect a very high speedup
- What is $\mathbb{E}Z$?

Applying CV (example)

- define the function `g_cv=function(x){}` so that it works correctly with the results of the new generator and apply `MC2`

```
g_cv=function(x){
  Y=g(x[,1])
  W=x[,2]
  EW=3
  return(Y-a*(W-EW))
}
MC2(g_cv,gen_cv,error=0.02,alpha=0.05)
```

```
## [1]    3.414766 1000.000000
```