Lecture 8 Exercises

Isabel Fulcher

8/20/2018

Install packages

```
library(tidyverse)
library(reshape2)
library(MASS)
```

Bootstrap exercise

We are going to use the bootstrap to get the standard errors from the previous function that you wrote, which likely looked something like,

```
negloglik = function(alpha, X, Y) {
  return(-sum(
    Y*(X%*%alpha)
    -log(1 + exp(X %*% alpha))
  )
  )
}
```

Recall this function corresponds to the following logistic model,

$$\text{logit}(Pr(Y = 1|X_1, X_2)) = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2$$

Let's create a dataset for this exercise.

```
set.seed(12345)
n <- 100
Sigma <- matrix(c(2,.2,.2,3),2,2)
alpha <- c(.5,.4,-.2)
X <- cbind(1,mvrnorm(n,rep(0,2),Sigma))
Y <- rbinom(n,1,plogis( X%*%alpha))

data <- data.frame(cbind(Y,X[,2:3]))
colnames(data) <- c("Y","X1","X2")
```

1. Perform the bootstrap using $B = 1,000$ replicates for $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2)$ and save the estimates from each replicate.
2. Using the results from 1, calculate the bootstrap standard error and 95% confidence intervals.
3. Compare the bootstrap standard error to that from the glm() function.

Note that you can also calculate the standard error for for $\hat{\boldsymbol{\alpha}}$ from the optim() function using standard likelihood theory. I have provided the code for this below.

```
fit.optim <- optim(runif(3,0,1),negloglik,Y=data$Y,X=X,method="BFGS",hessian=TRUE)
hess.optim <- fit.optim$hessian
se.optim <- sqrt(diag(solve(hess.optim)))

se.optim
```