Prelab5

Shu Zhou

2020/10/28

```
##This is the Prelab5 of STATS 413
##Author: Shu Zhou
##UMID: 19342932
```

(a.) Using the which function. Generate a vector Z of 100 iid standard normal variables. Use which to find the indices for which Zi > 1.8. Print the values of Zi for these indices. Use which to find the index of maxi Zi and maxi |Zi|.

```
set.seed(1)
Z=rnorm (100,0,1)
Z_I<-which(Z>1,8)#Zi > 1.8
Z_I;Z[Z_I]# Print the values of Zi for these indices
```

```
##  [1]  4 11 15 31 39 55 56 61 68 70 83 87 92 93 95
```

```
##  [1] 1.595281 1.511781 1.124931 1.358680 1.100025 1.433024 1.980400 2.401618
##  [9] 1.465555 2.172612 1.178087 1.063100 1.207868 1.160403 1.586833
```

```
which(Z==max(Z)) #maxi Zi
```

```
## [1] 61
```

```
which(abs(Z)==max(abs(Z)))#maxi |Zi|
```

```
## [1] 61
```

(b.) Converting data frames to matrices. Import the Carseats dataset from the ISLR package. Convert the variable Sales to a vector of length 400. Remove the three categorical variables, then convert the remaining continuous variables in the dataset to a matrix of size 400×7. Directly compute the estimate β for the regression of Sales on these seven continuous variables. (There is no need to add a column for the intercept.) (Hint: the functions data.matrix and crossprod may be helpful here.)

```
library(ISLR)
data(Carseats)
summary(Carseats)
```

```
##      Sales          CompPrice       Income        Advertising
##  Min.   : 0.000   Min.   : 77   Min.   : 21.00   Min.   : 0.000
##  1st Qu.: 5.390   1st Qu.:115   1st Qu.: 42.75   1st Qu.: 0.000
##  Median : 7.490   Median :125   Median : 69.00   Median : 5.000
##  Mean   : 7.496   Mean   :125   Mean   : 68.66   Mean   : 6.635
##  3rd Qu.: 9.320   3rd Qu.:135   3rd Qu.: 91.00   3rd Qu.:12.000
##  Max.   :16.270   Max.   :175   Max.   :120.00   Max.   :29.000
##    Population        Price        ShelveLoc        Age          Education
##  Min.   : 10.0   Min.   : 24.0   Bad   : 96   Min.   :25.00   Min.   :10.0
##  1st Qu.:139.0   1st Qu.:100.0   Good  : 85   1st Qu.:39.75   1st Qu.:12.0
##  Median :272.0   Median :117.0   Medium:219   Median :54.50   Median :14.0
```

```
##  Mean   :264.8   Mean   :115.8                Mean   :53.32   Mean   :13.9
##  3rd Qu.:398.5   3rd Qu.:131.0                3rd Qu.:66.00   3rd Qu.:16.0
##  Max.   :509.0   Max.   :191.0                Max.   :80.00   Max.   :18.0
##  Urban       US
##  No :118   No :142
##  Yes:282   Yes:258
##
##
##
##
```

```
y=Carseats$Sales
x=as.matrix(Carseats[,c(-1,-7,-10,-11)])
model_1<-lm(y~x)
solve(t(x) %*% x) %*% t(x) %*% y #directly calculate the esitmate hat(beta)
```

```
##                     [,1]
## CompPrice    0.127226700
## Income       0.020139922
## Advertising  0.133488157
## Population   0.001192478
## Price       -0.091783793
## Age         -0.029567078
## Education    0.083395449
```

```
summary(model_1) #Same as directly calculate the esitmate hat(beta)
```

```
##
## Call:
## lm(formula = y ~ x)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.0598 -1.3515 -0.1739  1.1331  4.8304
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   7.7076934  1.1176260   6.896 2.15e-11 ***
## xCompPrice    0.0939149  0.0078395  11.980  < 2e-16 ***
## xIncome       0.0128717  0.0034757   3.703 0.000243 ***
## xAdvertising  0.1308637  0.0151219   8.654  < 2e-16 ***
## xPopulation  -0.0001239  0.0006877  -0.180 0.857092
## xPrice       -0.0925226  0.0050521 -18.314  < 2e-16 ***
## xAge         -0.0449743  0.0060083  -7.485 4.75e-13 ***
## xEducation   -0.0399844  0.0371257  -1.077 0.282142
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.929 on 392 degrees of freedom
## Multiple R-squared:  0.5417, Adjusted R-squared:  0.5335
## F-statistic: 66.18 on 7 and 392 DF,  p-value: < 2.2e-16
```