Ans

1. (a). The fitted model will be

The taining error is

## [1] 10.96261

The test error is

## [1] 12.78361

The fitted model using predictors which are statically signicant will be

The taining error is

## [1] 11.16013

The test error is

## [1] 12.16591

(b). The fitted model of rigor regression will be

The taining error is

## [1] 11.53612

The test error is

## [1] 13.12343

(c). The fitted model of lasso regression will be

The taining error is

## [1] 11.17664

The test error is

## [1] 12.30693

(d). According to (a)-(c), multiple regression with significant variables has the smallest test error.

(e). Multiple regrssion bootstrap result

##   
## ORDINARY NONPARAMETRIC BOOTSTRAP  
##   
##   
## Call:  
## boot(data = Auto\_train, statistic = mul\_coef, R = 1000)  
##   
##   
## Bootstrap Statistics :  
## original bias std. error  
## t1\* -17.201371221 0.2056591419 6.564682038  
## t2\* -0.007486269 -0.0006769300 0.009319236  
## t3\* 0.031571242 -0.0009805104 0.016530374  
## t4\* -0.006863507 0.0001041661 0.001170817  
## t5\* 0.102209315 -0.0094372236 0.187127444  
## t6\* 0.758614461 -0.0018574626 0.068721270

rigor regrssion bootstrap result

##   
## ORDINARY NONPARAMETRIC BOOTSTRAP  
##   
##   
## Call:  
## boot(data = Auto\_train, statistic = rigor\_coef, R = 1000)  
##   
##   
## Bootstrap Statistics :  
## original bias std. error  
## t1\* -7.44231850 1.445685e-01 5.4196139147  
## t2\* -0.01781736 -2.775356e-04 0.0035586663  
## t3\* -0.01034566 -9.335460e-04 0.0076225139  
## t4\* -0.00404830 5.084114e-05 0.0003849608  
## t5\* -0.12856090 -2.013515e-03 0.1281383001  
## t6\* 0.65120273 -1.795618e-03 0.0605157788

lasso regrssion bootstrap result

##   
## ORDINARY NONPARAMETRIC BOOTSTRAP  
##   
##   
## Call:  
## boot(data = Auto\_train, statistic = lasso\_coef, R = 1000)  
##   
##   
## Bootstrap Statistics :  
## original bias std. error  
## t1\* -10.688448694 2.924819e-01 5.3323317017  
## t2\* -0.003747162 -2.517098e-03 0.0063875506  
## t3\* 0.000000000 6.094799e-05 0.0016033147  
## t4\* -0.006071524 2.429621e-04 0.0007322893  
## t5\* 0.000000000 -2.023654e-02 0.0855149766  
## t6\* 0.696516494 -2.979130e-03 0.0650959365

For displacement, weight and year, rigor regression have lowest standard error. For horsepower and acceleration, lasso regression have lowest standard error. For horsepower and acceleration lasso will tends to zeros, so the standard error will lower than rigor regression.

1. (a). The smallest k will be

## PC84   
## 84

(b). Fit LDA models. The training classification error will be

## Loading required package: MASS

## train\_Y  
## 0 1 2 3 4 5 6 7 8 9  
## 0 287 0 6 0 0 3 4 2 0 3  
## 1 0 338 6 6 3 2 2 5 13 2  
## 2 1 1 244 5 0 2 1 2 2 0  
## 3 1 1 7 272 0 12 0 1 8 4  
## 4 1 0 4 0 262 2 10 4 1 15  
## 5 5 5 1 16 0 224 10 1 16 0  
## 6 9 1 5 4 4 4 266 0 2 0  
## 7 0 0 4 5 0 0 0 258 0 8  
## 8 2 3 10 7 3 5 2 1 243 2  
## 9 0 2 4 4 23 10 1 26 8 251

## [1] 0.1183333

Test classification error will be

## test\_Y  
## 0 1 2 3 4 5 6 7 8 9  
## 0 272 0 1 3 0 5 1 0 3 3  
## 1 0 294 6 5 4 5 6 7 11 2  
## 2 2 2 224 11 0 1 0 3 5 1  
## 3 1 2 3 256 0 24 0 6 12 6  
## 4 1 0 6 1 227 3 2 4 0 19  
## 5 13 3 5 20 3 220 7 1 20 1  
## 6 2 0 13 3 0 5 287 0 3 0  
## 7 1 1 10 4 1 0 0 262 3 9  
## 8 5 12 9 9 1 8 1 2 247 0  
## 9 0 0 4 5 31 5 0 28 9 277

## [1] 0.1446667

(c). Fit QDA models. The training classification error will be

## train\_Y  
## 0 1 2 3 4 5 6 7 8 9  
## 0 306 0 0 0 0 0 0 0 0 0  
## 1 0 348 0 0 0 0 0 0 1 0  
## 2 0 1 291 1 0 0 0 0 0 0  
## 3 0 0 0 316 0 0 0 0 0 1  
## 4 0 1 0 0 295 0 0 0 0 2  
## 5 0 0 0 1 0 264 0 0 0 0  
## 6 0 1 0 0 0 0 296 0 0 0  
## 7 0 0 0 0 0 0 0 299 0 1  
## 8 0 0 0 1 0 0 0 1 292 0  
## 9 0 0 0 0 0 0 0 0 0 281

## [1] 0.004

Test classification error will be

## test\_Y  
## 0 1 2 3 4 5 6 7 8 9  
## 0 289 0 0 1 1 0 1 0 3 0  
## 1 0 278 0 0 0 0 0 0 0 0  
## 2 5 6 271 10 2 0 2 6 5 3  
## 3 1 0 1 296 0 13 0 4 9 4  
## 4 0 1 2 1 261 0 0 8 0 17  
## 5 0 0 0 5 0 257 5 2 2 1  
## 6 0 0 0 0 0 0 292 0 0 0  
## 7 0 0 1 1 0 0 0 279 0 3  
## 8 2 29 6 3 0 6 4 7 293 9  
## 9 0 0 0 0 3 0 0 7 1 281

## [1] 0.06766667

(d). Table of training and test errors will be

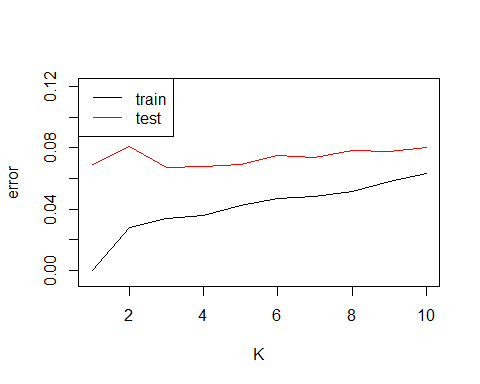
## Loading required package: class

## 1 2 3 4 5 6  
## train\_error 0.00000000 0.02833333 0.034 0.03600000 0.04266667 0.04733333  
## test\_error 0.06933333 0.08066667 0.067 0.06766667 0.06933333 0.07533333  
## 7 8 9 10  
## train\_error 0.048 0.05166667 0.05800000 0.063  
## test\_error 0.074 0.07833333 0.07733333 0.080

The best K value chosen by test calssification error will be

## [1] 3

The figure will be



(e). Fit multiple logistic regression

## Loading required package: iterators

## Loading required package: parallel

The training classification error will be

## train\_Y  
## mul\_log\_pred 0 1 2 3 4 5 6 7 8 9  
## 0 297 0 4 0 0 2 0 0 0 3  
## 1 0 346 2 1 2 2 0 2 3 2  
## 2 1 1 266 2 0 3 1 4 4 0  
## 3 0 0 3 291 0 7 0 0 6 4  
## 4 0 1 3 0 278 1 3 1 0 10  
## 5 1 0 0 10 0 237 5 1 7 1  
## 6 4 1 4 2 4 5 286 0 2 0  
## 7 0 1 4 2 0 1 0 283 1 8  
## 8 2 1 1 7 2 4 1 1 266 2  
## 9 1 0 4 4 9 2 0 8 4 255

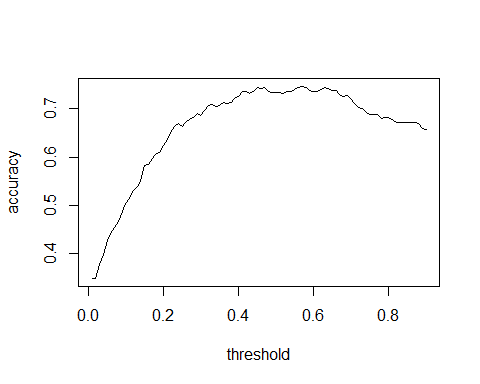
## [1] 0.065

The test classification error will be

## test\_Y  
## mul\_log\_pred\_test 0 1 2 3 4 5 6 7 8 9  
## 0 281 0 3 2 0 5 1 0 3 5  
## 1 0 302 2 2 1 2 1 5 7 2  
## 2 3 0 231 13 0 0 1 6 7 2  
## 3 2 2 3 267 1 16 0 4 13 6  
## 4 1 0 5 0 245 8 0 4 0 15  
## 5 6 4 4 22 0 232 0 1 7 2  
## 6 2 0 12 0 6 3 298 0 2 0  
## 7 1 1 13 4 1 1 1 277 1 11  
## 8 1 5 6 5 2 7 2 2 268 0  
## 9 0 0 2 2 11 2 0 14 5 275

## [1] 0.108

(f). knn with has the lowest test error.

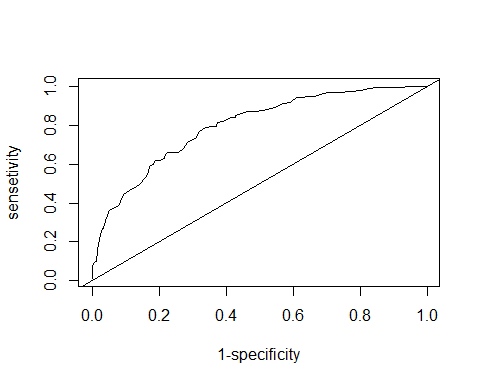
1. (a). The figure of accuracy will be  The threhold chosen by maximize accuracy will be

## [1] 0.57

The correponding confusion matrix will be

##   
## 0 1  
## FALSE 274 89  
## TRUE 28 71

1. The ROC curve will be



The threhold chosen by maximize accuracy will be

## [1] 0.32

The correponding confusion matrix will be

##   
## 0 1  
## FALSE 202 35  
## TRUE 100 125

(c). (b) is a better approach, because it has consider the proportion of the label.