Least Squares Assignment 5

Sharon Morris

2/26/2017

Problem Set 1 In this problem set we'll work out some properties of the least squares solution that wereviewed in the weekly readings. Consider the unsolvable system $Ax = b$ as given below:

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix}$$

Write R markdown script to compute $A^T A$ and $A^T b$.

```
#Define matrix
A <- matrix(c(1,1,1,1,0,1,3,4), nrow = 4)
b <- matrix(c(0,8,8,20), nrow = 4)

#Compute A^TA and A^Tb
ATA <- t(A)%*%A
ATA
```

```
##      [,1] [,2]
## [1,]    4    8
## [2,]    8   26
```

```
ATb <- t(A)%*%b
ATb
```

```
##      [,1]
## [1,]   36
## [2,]  112
```

Solve for $\hat{x}$ in R using the above two computed matrices.

```
xHat <- solve(ATA)%*%ATb #solves for x in the equation Ax=b-- finds inverse
xHat
```

```
##      [,1]
## [1,]    1
## [2,]    4
```

What is the squared error of this solution?

```
p <- A %*% xHat #vector error
p
```

```
##      [,1]
## [1,]    1
## [2,]    5
## [3,]   13
## [4,]   17
```

```
e <- b-p #test
e
```

```
##      [,1]
```

```
## [1,]   -1
## [2,]    3
## [3,]   -5
## [4,]    3
```

```
lsfit(A, b, intercept = FALSE)$residuals #find least square fit - results
```

```
## [1] -1  3 -5  3
```

```
sError <- sum(e^2)  #find squared error
sError
```

```
## [1] 44
```

Instead of $b = [0; 8; 8; 20]$, start with $p = [1; 5; 13; 17]$ and find the exact solution (i.e. show that this system is solvable as all equations are consistent with each other. This should result in an error vector $e = 0$).

```
p <-matrix(c(1,5,13,17), nrow = 4)
ATp <- t(A) %*% p
xHatp <- solve(ATA) %*% ATp
p2 <- A %*% xHatp
p2
```

```
##      [,1]
## [1,]    1
## [2,]    5
## [3,]   13
## [4,]   17
```

```
e2 <- p2-p
e2
```

```
##              [,1]
## [1,] 0.000000e+00
## [2,] 8.881784e-16
## [3,] 3.552714e-15
## [4,] 3.552714e-15
```

```
sError2 <- round(sum(e2^2)) #is e2 zero check
sError2
```

```
## [1] 0
```

Show that the error $e = bp = [1; 3; 5; 3]$.

```
b-p  #residual error
```

```
##      [,1]
## [1,]   -1
## [2,]    3
## [3,]   -5
## [4,]    3
```

Show that the error e is orthogonal to p and to each of the columns of A.

```
sum(e*p) # is e*p equal to 0
```

```
## [1] -1.030287e-13
```

```
sum(e*A[,1]) #column 1 matrix A --e and the columns of A are orthogonal
```

```
## [1] -7.993606e-15
```

```r
sum(e*A[,2]) #column 2 matrix A
```

```
## [1] -2.575717e-14
```

Problem Set 2 Consider the modified auto-mpg data (obtained from the UC Irvine Machine Learning dataset). This dataset contains 5 columns: displacement, horsepower, weight, acceleration, mpg. We are going to model mpg as a function of the other four variables.

Write an R markdown script that takes in the auto-mpg data, extracts an A matrix from the first 4 columns and b vector from the fifth (mpg) column. Using the least squares approach, your code should compute the best fitting solution. That is, find the best fitting equation that expresses mpg in terms of the other 4 variables. Finally, calculate the fitting error between the predicted mpg of your model and the actual mpg. Your script should be able to load in the 5 column data set, extract A and b, and perform the rest of the calculations.

```r
# read data into a matrix and add variable names
data <- as.matrix(read.table("auto-mpg.data", header = FALSE, as.is = TRUE))
colnames(data) <- c("displacement", "horsepower","weight","acceleration","mpg")

A <- data.matrix(data[,1:4]) #Separate matrix
b <- data.matrix(data[,5])   #Fifth column as vector
intercept <- rep(1, nrow(A))  #A a column to be used for intercept
A <- cbind(intercept, A)

#Least Squares Approach
ATa <- t(A) %*% A
ATb <- t(A) %*% b
xHats <- solve(ATa) %*% ATb #Calculate xhat
p <- A %*% xHats #Scared error
error <- b-p  #Error vector
errorSq <- sum(e^2)
errorSq
```

```
## [1] 44
```

```r
lsfit(A, b, intercept=FALSE)$coefficients #Compare outcome against xhat
```

```
##    intercept displacement   horsepower       weight acceleration
## 45.251139699 -0.006000871 -0.043607731 -0.005280508 -0.023147999
```

References

https://www.youtube.com/watch?v=eJqreNZPS64

https://www.r-bloggers.com/quick-review-of-matrix-algebra-in-r/

http://www.endmemo.com/program/R/solve.php