BHao_Assign5

Problem Set 1

```r
A = matrix(c(1,1,1,1,0,1,3,4), nrow = 4)
b = matrix(c(0,8,8,20), nrow = 4)
# for verification purposes
df = data.frame(y = b, x = A[,2])
model = lm(y ~ ., data = df)

#AtA
t(A) %*% A
```

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

```r
#Atb
t(A) %*% b
```

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

```r
#x hat
x_hat = solve(t(A) %*% A) %*% t(A) %*% b
x_hat
```

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

```r
model$coefficients
```

```
## (Intercept)           x
##           1           4
```

```r
#squared error
sum((b - A %*% x_hat)^2)
```

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

```r
sum(model$residuals^2)
```

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

```r
#exact solution using p
p = matrix(c(1,5,13,17), nrow = 4)
x = solve(t(A) %*% A) %*% t(A) %*% p
sum((p - A %*% x)^2)
```

```
## [1] 1.755216e-29
```

```r
#show the error
e = b - p
e
```

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

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

```r
#show that e is orthogonal to p and to columns of A
#dot products equal zero when vectors are orthogonal
t(p) %*% e
```

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

```r
t(A[,1]) %*% e
```

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

```r
t(A[,2]) %*% e
```

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

Problem Set 2

```r
df = read.delim('~/Google Drive/CUNY/git/DATA605/HW5/assign5/auto-mpg.data', header = F, sep = "")
names(df) = c('displacement', 'horsepower', 'weight', 'acceleration', 'mpg')
# for verification purposes
model = lm(mpg ~ ., data = df)

# convert data frame to A and b matrices
A = data.matrix(df[,1:4])
# add column of 1s for intercept term
A = cbind(rep(1, nrow(A)), A)
b = data.matrix(df[,5])

# calculate x_hat
x_hat = solve(t(A) %*% A) %*% t(A) %*% b
x_hat
```

```
##                      [,1]
##               45.251139699
## displacement  -0.006000871
## horsepower    -0.043607731
## weight        -0.005280508
## acceleration  -0.023147999
```

```r
model$coefficients
```

```
##  (Intercept) displacement   horsepower       weight acceleration
## 45.251139699 -0.006000871 -0.043607731 -0.005280508 -0.023147999
```

```r
SSE = sum((b - A %*% x_hat)^2)
SSE
```

```
## [1] 6979.413
```

```
anova(model)
```

```
## Analysis of Variance Table
##
## Response: mpg
##               Df  Sum Sq Mean Sq F value     Pr(>F)
## displacement   1 15440.2 15440.2 856.139 < 2.2e-16 ***
## horsepower     1   383.5   383.5  21.264 5.447e-06 ***
## weight         1  1015.3  1015.3  56.298 4.315e-13 ***
## acceleration   1     0.6     0.6   0.034    0.8539
## Residuals    387  6979.4    18.0
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```