Stats 401 Lab 3 Solutions and HW FAQ

401 GSI team

1/18/2018 and 1/19/2018

In Lab Activity

- Using the library(faraway) and data(sat). Let total sat score be the response and expend, ratio, and salary be explanatory variables

1. Construct the linear equation using vector and matrix notation.
2. Estimate the least squares estimate of $\mathbf{b}$ using the design matrix $\mathbb{X}$.
3. Check your estimate by using the lm function in R.
4. Suppose we are given that expend $= 5$, ratio $= 20$, salary $= 30$ for a certain state, estimate its average total SAT score.

Lab Activity Solution

1. Let $\mathbf{y} = (y_1, y_2, ..., y_50)$, $\mathbf{b} = (b_1, b_2, ..., b_7)$, and $\mathbf{e} = (e_1, e_2, ..., e_50)$ be the vector of SAT scores, predictor variables, and error terms. Then

$$\mathbf{y} = \mathbb{X}\mathbf{b} + \mathbf{e}$$

Lab Activity Solution (cont.)

2.

```
library(faraway)
data("sat")

sat_x <- cbind(intercept=rep(1, length(sat$expend)),
           sat$expend, sat$ratio, sat$salary)

solve(t(sat_x) %*% sat_x) %*% t(sat_x) %*% sat$total
```

```
##                    [,1]
## intercept 1069.234168
##              16.468866
##               6.330267
##              -8.822632
```

Lab Activity Solution (cont.)

3.

```
sat_lm = lm(total ~ expend + ratio + salary, data = sat)
sat_lm$coefficients
```

```
## (Intercept)      expend       ratio      salary
## 1069.234168   16.468866    6.330267   -8.822632
```

Lab Activity Solution (cont.)

4.

```
sat_ob <- data.frame(expend = c(5), ratio = c(20), salary =
sat_fitted = predict(sat_lm, sat_ob)
sat_fitted
```

```
##        1
## 1013.505
```