Math 342W/642/742W

Recitation – Day #11 (3.18.25)

Problem #1

Consider the following `R` code from the class demos and labs. The numbers on the right are line numbers that will be referred to later. They are not part of the code.

```
y = MASS::Cars93$Price
var(y)
[1] 93.30458
X = as.matrix(cbind(1, MASS::Cars93[,c(7,8,12,13,14,15,17,18)]))
n = nrow(X)
n
[1] 93
Xt = t(X)
XtX = Xt %*% X
XtXinv = solve(XtX)
XtXinvXt = XtXinv %*% Xt
b = XtXinvXt %*% y
H = X %*% XtXinvXt
I_minus_H = diag(n) - H
yhat = H %*% y
e = I_minus_H %*% y
var(e)
          [,1]
[1,] 32.70302
```

(*This problem is adapted from Problem #2 of Professor Kapelner's Spring 2024 Midterm I.*)

- What is returned by `R` when evaluating `length(b)`?

- What is returned by `R` when evaluating `ncol(XtX)`?

- What is returned by `R` when evaluating `ncol(H)`?

- What is returned by `R` when evaluating when evaluating `Matrix::rankMatrix(I_minus_H)`?

- Compute SST, SSE, SSR, and R^2 to the nearest two decimal places.

- What is the number of random predictors that can be added before the code throws an error and halts?

Problem #2

(a) Assume that $\mathbf{X} \in \mathbb{R}^{n \times (p+1)}$ with $p > 1$ composed of random realizations from iid standard normal random variables, full rank. Let $\mathbf{Q}$, $\mathbf{R}$ be the matrix results of the QR-decomposition procedure run on $\mathbf{X}$. Let $\boldsymbol{y} \in \mathbb{R}^n$ which represents a vector of measurements of a phenomenon of interest.

Prove that

$$\frac{\|\mathbf{Q}\mathbf{Q}^T\boldsymbol{y}\| + \|(I - \mathbf{Q}\mathbf{Q}^T)\boldsymbol{y}\|}{\|\boldsymbol{y}\|} \geq 1$$

(b) Claim: $\exists c \neq 0$ such that $\mathbf{X}_{\cdot 3} = c\,\mathbf{Q}_{\cdot 3}$.

Show whether the claim is true or not.

(*This problem is adapted from Problem #4 of Professor Kapelner's Spring 2024 Midterm I.*)