Nick Waalkes
Math189R SU20
Homework 1
Wednesday, June 10, 2020

Feel free to work with other students, but make sure you write up the homework and code on your own (no copying homework *or* code; no pair programming). Feel free to ask students or instructors for help debugging code or whatever else, though. The starter code for problem 2 part c and d can be found under the Resource tab on course website.

Note: You need to create a Github account for submission of the coding part of the homework. Please create a repository on Github to hold all your code and include your Github account username as part of the answer to problem 2.

1 (**Linear Transformation**) Let $\mathbf{y} = A\mathbf{x} + \mathbf{b}$ be a random vector. show that expectation is linear:

$$\mathbb{E}[\mathbf{y}] = \mathbb{E}[A\mathbf{x} + \mathbf{b}] = A\mathbb{E}[\mathbf{x}] + \mathbf{b}.$$

Also show that

$$\text{cov}[\mathbf{y}] = \text{cov}[A\mathbf{x} + \mathbf{b}] = A\text{cov}[\mathbf{x}]A^\top = A\mathbf{\Sigma}A^\top.$$

We know that

$$\mathbb{E}[\mathbf{y}] = \int \mathbf{y}p(x)dx = \int (Ax+b)p(x)dx$$

$$\mathbb{E}[\mathbf{y}] = A\int xp(x)dx + b\int p(x)dx$$

$$\mathbb{E}[\mathbf{y}] = A\mathbb{E}[\mathbf{x}] + b$$

Now, let's solve the second part of the problem.
We know that:

$$cov[\mathbf{y}] = \mathbb{E}[(Ax+b-\mathbb{E}[Ax+b])(Ax+b-\mathbb{E}[Ax+b])^\top]$$

Using the linearity we proved above:

$$cov[\mathbf{y}] = \mathbb{E}[(Ax+b-A\mathbb{E}[x]-b)(Ax+b-A\mathbb{E}[x]-b)^\top]$$

$$cov[\mathbf{y}] = \mathbb{E}[(Ax-A\mathbb{E}[x])(Ax-A\mathbb{E}[x])^\top]$$

$$cov[\mathbf{y}] = \mathbb{E}[A(x-\mathbb{E}[x])(x-\mathbb{E}[x])^\top A^\top]$$

$$cov[\mathbf{y}] = A\mathbb{E}[(x-\mathbb{E}[x])(x-\mathbb{E}[x])^\top]A^\top$$

$$cov[\mathbf{y}] = Acov[x]A^\top = A\mathbf{\Sigma}A^\top$$

■

2 Given the dataset $\mathcal{D} = \{(x, y)\} = \{(0,1), (2,3), (3,6), (4,8)\}$

(a) Find the least squares estimate $y = \boldsymbol{\theta}^\top \mathbf{x}$ by hand using Cramer's Rule.

(b) Use the normal equations to find the same solution and verify it is the same as part (a).

(c) Plot the data and the optimal linear fit you found.

(d) Find randomly generate 100 points near the line with white Gaussian noise and then compute the least squares estimate (using a computer). Verify that this new line is close to the original and plot the new dataset, the old line, and the new line.

A) We know that:

$$X^\top X\boldsymbol{\theta} = X^\top y$$

Using the given data, we can construct the following matrices:

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \; y = \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}$$

Thus we have the following equation:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \boldsymbol{\theta} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix} \boldsymbol{\theta} = \begin{bmatrix} 18 \\ 56 \end{bmatrix}$$

Using Cramer's Rule:

$$\boldsymbol{\theta}_0 = det(\begin{bmatrix} 18 & 9 \\ 56 & 29 \end{bmatrix})/det(\begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}) = 18/35$$

$$\boldsymbol{\theta}_1 = det(\begin{bmatrix} 4 & 18 \\ 9 & 56 \end{bmatrix})/det(\begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}) = 62/35$$

Thus, the least square estimate is:

$$y = \begin{bmatrix} 18/35 & 62/35 \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix}$$

B)
Using the normal equation:

$$\boldsymbol{\theta} = (X^\top X)^{-1} X^\top y$$

We can also calculate the least square estimate using the same matrices:

$$\boldsymbol{\theta} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix})^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}$$

$$\boldsymbol{\theta} = \begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}^{-1} \begin{bmatrix} 18 \\ 56 \end{bmatrix}$$

$$\boldsymbol{\theta} = \begin{bmatrix} 29/35 & -9/35 \\ -9/35 & 4/35 \end{bmatrix} \begin{bmatrix} 18 \\ 56 \end{bmatrix}$$

$$\boldsymbol{\theta} = \begin{bmatrix} 18/35 \\ 62/35 \end{bmatrix}$$

Thus, the two solutions are equivalent!

$\blacksquare$