Michelle Li
Math189R SP20
Homework 1
Monday, Feb 3, 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

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

(a) We know by definition

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

where $p(x)$ is the probability density function.
Then we have

$$\begin{aligned}\mathbb{E}[\mathbf{y}] &= \int_S (A\mathbf{x} + \mathbf{b})p(x)dx \\ &= A\int_S \mathbf{x}p(x)dx + \mathbf{b}\int_S p(x)dx \\ &= A\mathbb{E}[\mathbf{x}] + \mathbf{b}\end{aligned}$$

$\therefore$ Expected value is linear.

(b) We know by definition

$$\text{cov}[\mathbf{x}] = \mathbf{\Sigma} = \mathbb{E}\left[(\mathbf{x} - \mathbb{E}[x])(\mathbf{x} - \mathbb{E}[x])^\top\right]$$

Plugging $\mathbf{y} = A\mathbf{x} + \mathbf{b}$ into the definition, we simplify as follows

$$\begin{aligned}
\mathrm{cov}[\mathbf{y}] &= \mathrm{cov}[A\mathbf{x} + \mathbf{b}] \\
&= \mathbb{E}\left[\left(A\mathbf{x} + \mathbf{b} - \mathbb{E}[A\mathbf{x} + \mathbf{b}]\right)\left(A\mathbf{x} + \mathbf{b} - \mathbb{E}[A\mathbf{x} + \mathbf{b}]\right)^\top\right] \\
&= \mathbb{E}\left[\left(A\mathbf{x} + \mathbf{b} - A\mathbb{E}[\mathbf{x}] - \mathbf{b}\right)\left(A\mathbf{x} + \mathbf{b} - A\mathbb{E}[\mathbf{x}] - \mathbf{b}\right)^\top\right] \quad \text{(linearity)} \\
&= \mathbb{E}\left[\left(A\mathbf{x} - A\mathbb{E}[\mathbf{x}]\right)\left(A\mathbf{x} - A\mathbb{E}[\mathbf{x}]\right)^\top\right] \\
&= \mathbb{E}\left[A\left(\mathbf{x} - \mathbb{E}[\mathbf{x}]\right)\left(\mathbf{x} - \mathbb{E}[\mathbf{x}]\right)^\top A^\top\right] \\
&= A\mathbb{E}\left[\left(\mathbf{x} - \mathbb{E}[\mathbf{x}]\right)\left(\mathbf{x} - \mathbb{E}[\mathbf{x}]\right)^\top\right] A^\top \\
&= A\boldsymbol{\Sigma}A^\top
\end{aligned}$$

■

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) Let

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}.$$

We know that $X^\top X \boldsymbol{\theta}^* = X^\top \mathbf{y}$ from normal equations. Solve for $X^\top X$ and $X^\top \mathbf{y}$ as follows

$$X^\top X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}, \quad X^\top \mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 18 \\ 56 \end{bmatrix}$$

We are estimating a line $y = \theta_0 + \theta_1 x$. Use Cramer's rule to find θ_0 and θ_1.

$$\boldsymbol{\theta}_0^* = \frac{\begin{vmatrix} 18 & 9 \\ 56 & 29 \end{vmatrix}}{\begin{vmatrix} 4 & 9 \\ 9 & 29 \end{vmatrix}} = \frac{18}{35} \qquad \boldsymbol{\theta}_1^* = \frac{\begin{vmatrix} 4 & 18 \\ 9 & 56 \end{vmatrix}}{\begin{vmatrix} 4 & 9 \\ 9 & 29 \end{vmatrix}} = \frac{62}{35}$$

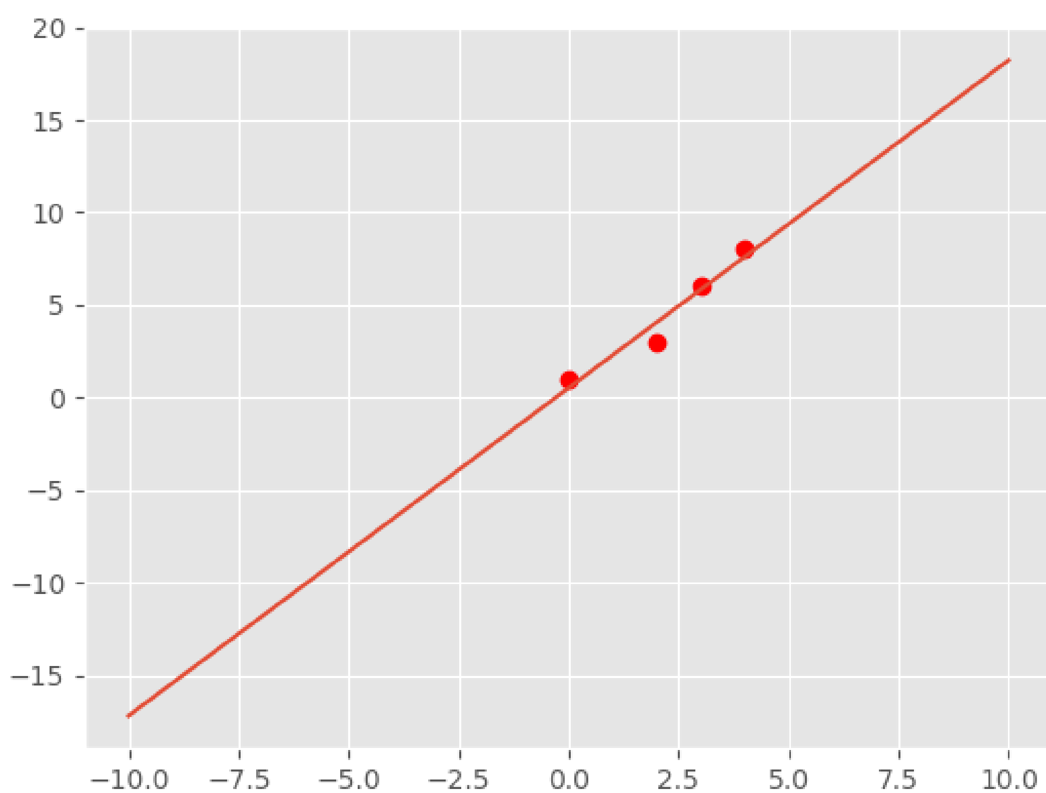
Thus, our least squares estimate is

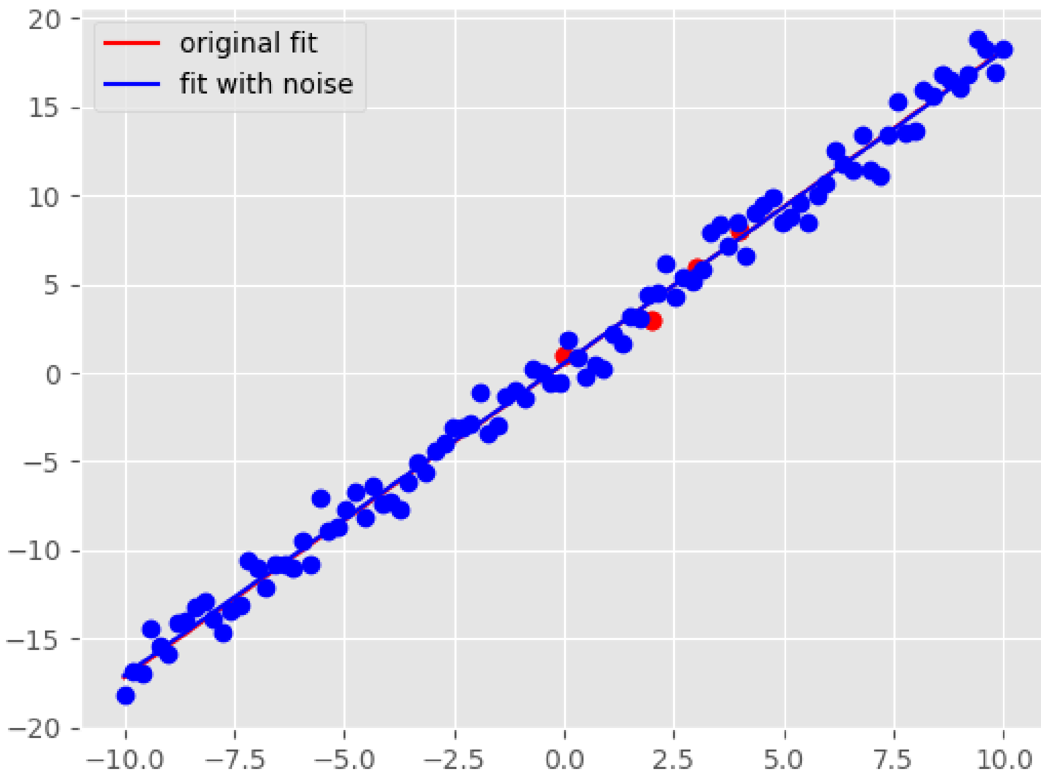
$$y = \frac{18}{35} + \frac{62}{35} x.$$

(b) From the normal equation, we know $\boldsymbol{\theta}^* = (X^\top X)^{-1} X^\top \mathbf{y}$. Thus,

$$
\begin{aligned}
\boldsymbol{\theta}^* &= \left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix} \\
&= \begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}^{-1} \begin{bmatrix} 18 \\ 56 \end{bmatrix} \qquad \text{(from (a))} \\
&= \frac{1}{35} \begin{bmatrix} 29 & -9 \\ -9 & 4 \end{bmatrix} \begin{bmatrix} 18 \\ 56 \end{bmatrix} \\
&= \frac{1}{35} \begin{bmatrix} 18 \\ 62 \end{bmatrix} \\
&= \begin{bmatrix} 18/35 \\ 62/35 \end{bmatrix} \qquad \text{(same as (a))}
\end{aligned}
$$

(c) Plot shown below





(d) Plot above. The new line and original line are very close.

■