Sidhant Rastogi
Math189R SP19
Homework 1
Monday, Feb 4, 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

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

Let $\mathbf{y} = A\mathbf{x} + \mathbf{b}$ be a random vector with probability distribution $p(\mathbf{x})$, where X and $\mathbf{b}$ are constant vectors. Then the expectation can be calculated as

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

■

Similarly, for covariance, we calculate as follows:

$$\begin{aligned}\text{cov}[\mathbf{y}] &= \mathbb{E}[(\mathbf{y} - \mathbb{E}[\mathbf{y}])(\mathbf{y} - \mathbb{E}[\mathbf{y}])^T] = \mathbb{E}[(A\mathbf{x} + \mathbf{b} - (A\mathbb{E}[\mathbf{x}] + \mathbf{b}))(A\mathbf{x} + \mathbf{b} - (A\mathbb{E}[\mathbf{x}] + \mathbf{b}))^T] \\ &= \mathbb{E}[A(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^T A^T] = A\text{cov}[\mathbf{x}]A^T\end{aligned}$$

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) Given the equations found by setting the partial derivatives of the Cost function equal to 0,

$$\begin{aligned} \theta_0 \sum x_i + \theta_1 \sum x_i^2 &= \sum x_i y_i \\ \theta_0 n + \theta_1 \sum x_i &= \sum y_i \end{aligned}$$

Cramer's rule gives us that

$$\theta_0 = \frac{\begin{vmatrix} \sum x_i y_i & \sum x_i^2 \\ \sum y_i & \sum x_i \end{vmatrix}}{\begin{vmatrix} \sum x_i & \sum x_i^2 \\ n & \sum x_i \end{vmatrix}} = \frac{\sum x_i y_i \sum x_i - \sum x_i^2 \sum y_i}{(\sum x_i)^2 - n \sum x_i^2}$$

$$\theta_1 = \frac{\begin{vmatrix} \sum x_i & \sum x_i y_i \\ n & \sum y_i \end{vmatrix}}{\begin{vmatrix} \sum x_i & \sum x_i^2 \\ n & \sum x_i \end{vmatrix}} = \frac{\sum x_i \sum y_i - n \sum x_i y_i}{(\sum x_i)^2 - n \sum x_i^2}$$

Now, we simply calculate the necessary quantities:

$$\begin{aligned} \sum x_i &= 9 \\ \sum y_i &= 18 \\ \sum x_i^2 &= 29 \\ \sum x_i y_i &= 56 \end{aligned}$$

Plugging this in, we get

$$\begin{aligned} \theta_0 &= \frac{(56)(9) - (29)(18)}{(9)^2 - 4(29)} = \frac{18}{35} \\ \theta_1 &= \frac{(9)(18) - 4(56)}{(9)^2 - 4(29)} = \frac{62}{35} \end{aligned}$$

So we get the line $y = \frac{18}{35} + \frac{62}{35} x$.

(b) The normal equation is $\boldsymbol{\theta} = (X^T X)^{-1} X^T \mathbf{y}$. We write

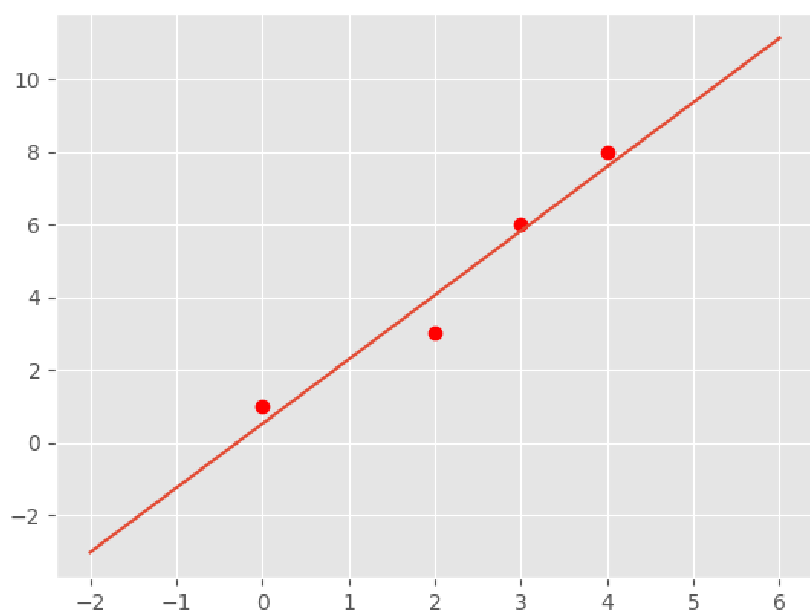
$$X = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}$$

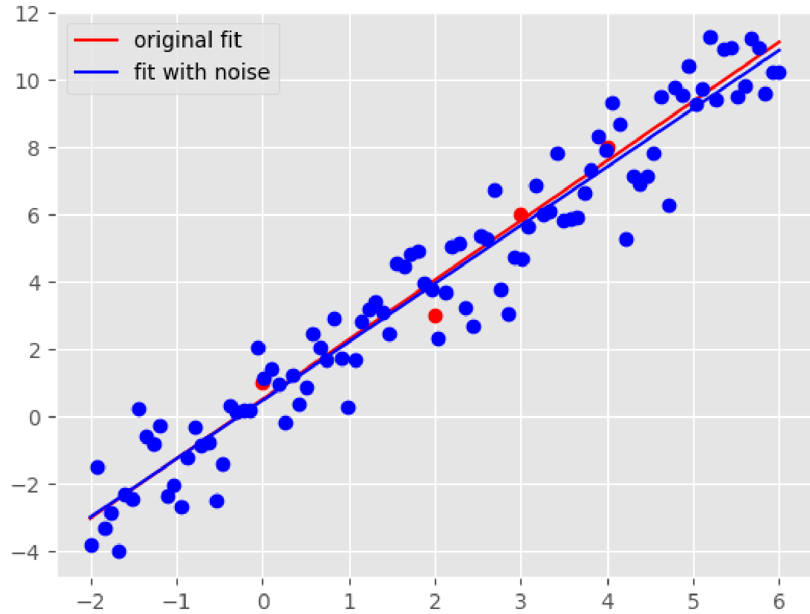
where the first column of X are all ones by definition of the data design matrix. Plugging this in, we get

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

so we get the line $y = \frac{18}{35} + \frac{62}{35}x$, which is the same as part (a).

(c)





(d)

■