

CONSULTED SOLUTION

Math189R SP19
Homework 1
Monday, Feb 4, 2017

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

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

Let $f(\mathbf{x})$ be the corresponding density function for \mathbf{x} on some domain D . Then

$$\begin{aligned}\mathbb{E}[A\mathbf{x} + \mathbf{b}] &= \int_D (A\mathbf{x} + \mathbf{b})f(\mathbf{x})d\mathbf{x} \\ &= \int_D A\mathbf{x}f(\mathbf{x})d\mathbf{x} + \int_D \mathbf{b}f(\mathbf{x})d\mathbf{x} \\ &= A \int_D \mathbf{x}f(\mathbf{x})d\mathbf{x} + \mathbf{b} \\ &= A\mathbb{E}[\mathbf{x}] + \mathbf{b}.\end{aligned}$$

This follows because \mathbf{b} is not dependant on \mathbf{x} and can thus be factored out which then eliminates the relevant integral as $\int_D f(\mathbf{x})d\mathbf{x} = 1$ by definition. Furthermore, the simplification in the last step is because A is also not dependant on \mathbf{x} and because $\int_D \mathbf{x}f(\mathbf{x})d\mathbf{x} = \mathbb{E}[\mathbf{x}]$ by definition.

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

By the previous part we can substitute $\mathbb{E}[A\mathbf{x} + \mathbf{b}] = A\mathbb{E}[\mathbf{x}] + \mathbf{b}$, giving

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

Since A and A^\top are constants, we know that we can factor them out by the previous part, giving

$$\begin{aligned}&= A\mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^\top]A^\top \\ &= A\text{cov}[\mathbf{x}]A^\top \\ &= A\Sigma A^\top\end{aligned}$$

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

- (a) Find the least squares estimate $y = \theta^\top \mathbf{x}$ 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) Since we are using a linear regression we know that for this problem

$$\begin{pmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} b \\ m \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 6 \\ 8 \end{pmatrix}.$$

Since $X = \begin{pmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix}$ and $\vec{y} = \begin{pmatrix} 1 \\ 3 \\ 6 \\ 8 \end{pmatrix}$, it follows that

$$X^\top X = \begin{pmatrix} 4 & 9 \\ 9 & 29 \end{pmatrix}$$

$$X^\top \vec{y} = \begin{pmatrix} 18 \\ 56 \end{pmatrix}$$

By Cramer's rule it then follows that

$$b = \frac{\begin{vmatrix} 18 & 9 \\ 56 & 29 \end{vmatrix}}{\begin{vmatrix} 4 & 9 \\ 9 & 29 \end{vmatrix}} = \frac{18}{35},$$

and

$$m = \frac{\begin{vmatrix} 4 & 18 \\ 9 & 56 \end{vmatrix}}{\begin{vmatrix} 4 & 9 \\ 9 & 29 \end{vmatrix}} = \frac{62}{35},$$

Meaning that $m = \frac{62}{35}$ and $b = \frac{18}{35}$.

(b) We know from the previous section that we have the following formula

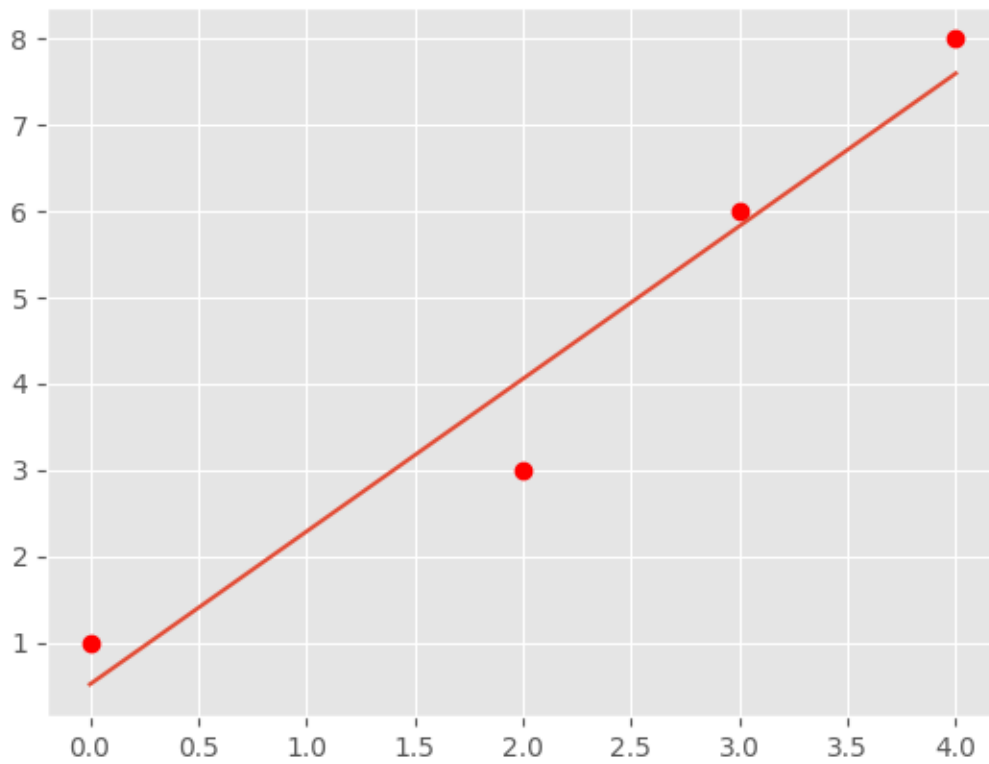
$$X^\top X \theta = X^\top \vec{y}.$$

We wish to solve for $\theta = \begin{pmatrix} m \\ b \end{pmatrix}$ which is equal to

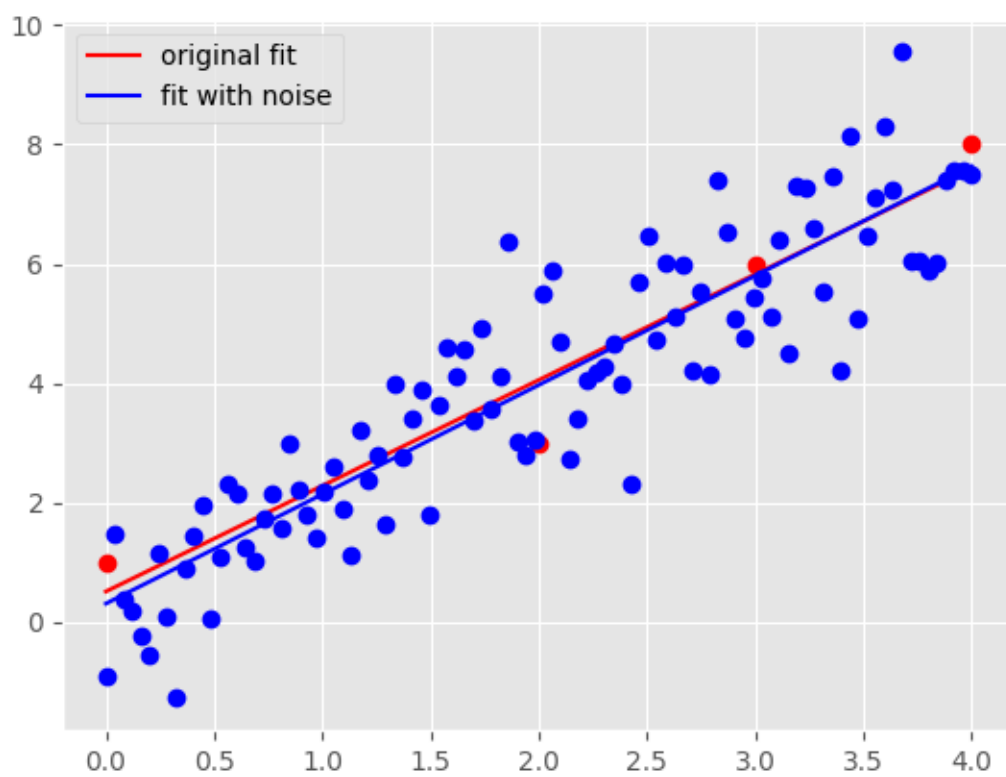
$$\begin{aligned}\theta &= (X^T X)^{-1} X^T \vec{y} \\ &= \left(\begin{pmatrix} 4 & 9 \\ 9 & 29 \end{pmatrix} \right)^{-1} \begin{pmatrix} 18 \\ 56 \end{pmatrix} \\ &= \frac{1}{35} \begin{pmatrix} 29 & -9 \\ -9 & 4 \end{pmatrix} \begin{pmatrix} 18 \\ 56 \end{pmatrix} \\ &= \frac{1}{35} \begin{pmatrix} 18 \\ 62 \end{pmatrix}\end{aligned}$$

showing that (a) and (b) have the same solution.

(c)



(d)



■