Andy Liu (Github: al1729)
Math189R
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

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

(a) Note that, by definition, $\mathbb{E}[\mathbf{y}] = \mathbb{E}[A\mathbf{x} + \mathbf{b}] = \int (A\mathbf{x} + \mathbf{b})P(x)$, where the integral is taken over all possible values of x. By the properties of integrals, this is equal to

$$A\int \mathbf{x}P(x)\mathrm{dx} + \int \mathbf{b}P(x)\mathrm{dx}$$

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

(using the property that b is constant to simplify the right-hand side), as desired. Thus, expectation is linear.

(b) Note that, by definition,
$\mathrm{cov}[\mathbf{y}] = \mathrm{cov}[A\mathbf{x} + \mathbf{b}] = \mathbb{E}[(A\mathbf{x} + \mathbf{b} - \mathbb{E}[A\mathbf{x} + \mathbf{b}])(A\mathbf{x} + \mathbf{b} - \mathbb{E}[A\mathbf{x} + \mathbf{b}])^T]$. Since expectation is linear, the innermost expectation expressions can be simplified to become

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

We can then factor out the A and A^T from the left- and right-hand sides of the expression, yielding

$$(A\mathbb{E}[\mathbf{x} - \mathbb{E}[\mathbf{x}]]\mathbb{E}[\mathbf{x} - \mathbb{E}[\mathbf{x}]]^T A^T)$$

However, the middle expression is just $\mathbf{\Sigma}$, meaning the entire expression is equivalent to $A\mathbf{\Sigma}A^T$, as desired.

■

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.

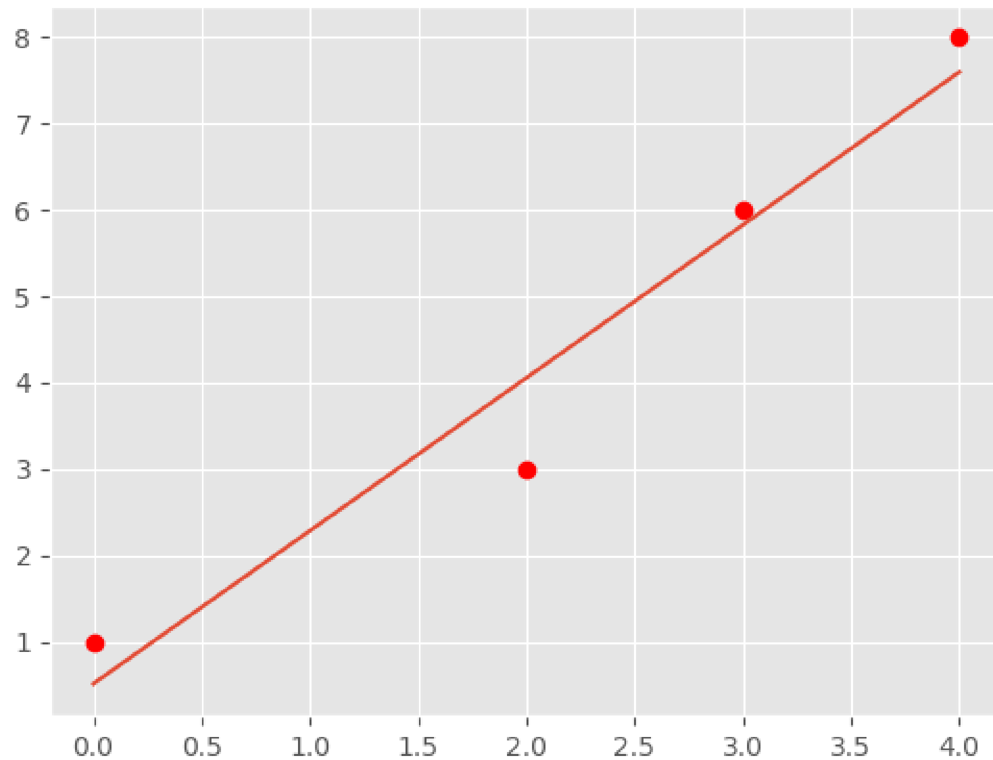
(d) Find randomly generate 100 points near the line with white Gaussian noise and then compute the least squares estimate (using a computer). Verify that this new line is close to the original and plot the new dataset, the old line, and the new line.

(a) It is known that, by Cramer's Rule, $y = mx + b$, where $m = \frac{n\sum x_i y_i - (\sum x_i)(\sum y_i)}{n\sum x_i^2 - (\sum x_i)^2}$ and $b = \frac{(\sum x_i^2)(\sum y_i) - (\sum x_i)(\sum x_i y_i)}{n\sum x_i^2 - (\sum x_i)^2}$. Since we are given $\mathcal{D}$, we know that $n = 4$, and we can manually compute $\sum x_i = 9, \sum y_i = 18, \sum x_i y_i = 56, \sum x_i^2 = 29$.

Substituting these values into our equation, we find that $m = \frac{4(56)-9(18)}{4(29)-9^2} = \frac{62}{35}$, and $b = \frac{29(18)-9(56)}{4(29)-9^2} = \frac{18}{35}$. Thus, we conclude that $y = \frac{62x}{35} + \frac{18}{35}$.

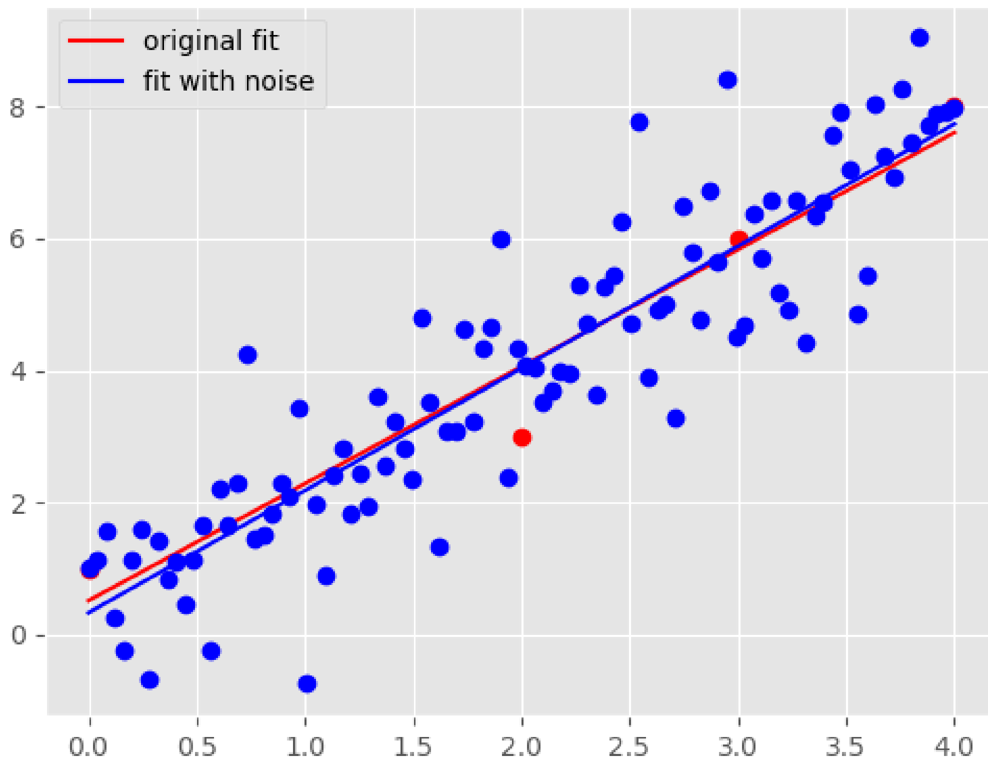
(b) Note that we are given that $X = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, Y = \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}$. Thus, $X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}$. Additionally, $X^T Y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 18 \\ 56 \end{bmatrix}$.

By the normal equations, we know that $\boldsymbol{\theta} = (X^T X)^{-1}(X^T y) \Rightarrow X^T X \boldsymbol{\theta} = X^T y \Rightarrow \begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix} \boldsymbol{\theta} = \begin{bmatrix} 18 \\ 56 \end{bmatrix}$. Solving for θ using Cramer's Rule, we find that $\theta = \begin{bmatrix} \frac{18}{35} \\ \frac{62}{35} \end{bmatrix}$, thus confirming that $y = \frac{62x}{35} + \frac{18}{35}$.



(c)

(Github: al1729)



(d)

(Github: al1729)

■