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

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

Note that we can write $E[y]$ as:

$$E[y] = E[Ax + b] = \int (Ax + b)P(x)dx$$

$$E[y] = \int AxP(x) + \int bP(x)$$

$$E[y] = A\int xP(x)dx + \int bP(x)$$

Note that $\int xP(x)dx = E[x]$.

$$E[y] = AE[x] + b\int P(x)dx$$

Also note that the integral of a probability density function is 1, so we have:

$$E[y] = AE[x] + b$$

as desired.
Remember $\mathrm{cov}[x] = \Sigma = E[(x - E[x])(x - E[x])^T]$
Thus we have:

$$\mathrm{cov}[y] = \mathrm{cov}[Ax + b] = E[(Ax + b - E[Ax + b])(Ax + b - E[Ax + b])^T]$$

$$\mathrm{cov}[y] = E[(Ax + b - AE[x] - b)(Ax + b - AE[x] - b])^T]$$

$$\mathrm{cov}[y] = E[(Ax - AE[x])(Ax - AE[x])^T]$$

$$\mathrm{cov}[y] = AE[(x - E[x])(x - E[x])^T]A^T$$

$$\mathrm{cov}[y] = A\mathrm{cov}[x]A^T$$

$$\mathrm{cov}[y] = A\Sigma A^T$$

∎

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. From the dataset $\mathcal{D}$, we have:

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \text{ and } Y = \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}$$

Thus we have:

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 6 & 8 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}$$

And:

$$X^T y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 6 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 18 \\ 56 \end{bmatrix}$$

Using Cramer's Rule we have:

$$\theta_0^* = \frac{18}{35}$$
$$\theta_1^* = \frac{62}{35}$$

b. From the normal equation we have,

$$\theta^* = (X^T X)^{-1} X^T y$$

$$\theta^* = \begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 6 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}$$

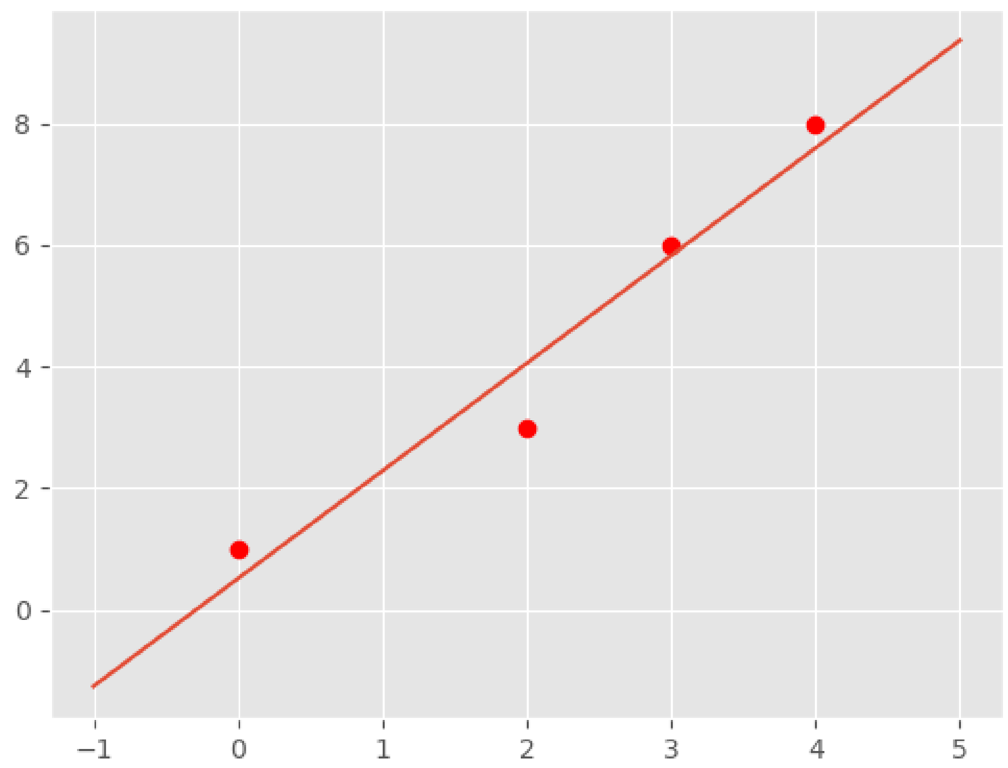
$$\theta^* = \frac{1}{35} \begin{bmatrix} 29 & -9 \\ -9 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 6 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}$$

$$\theta^* = \frac{1}{35}\begin{bmatrix} 29 & 11 & 2 & -7 \\ -9 & -1 & 3 & 7 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}$$

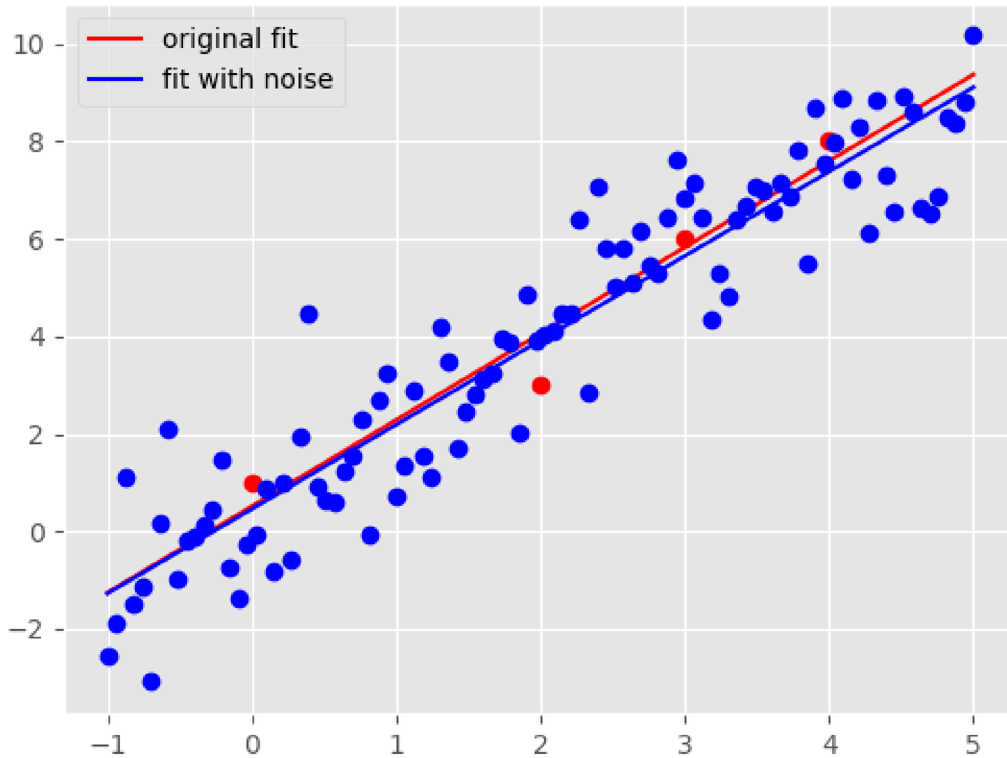
$$\theta^* = \frac{1}{35}\begin{bmatrix} 18 \\ 62 \end{bmatrix} = \begin{bmatrix} \frac{18}{35} \\ \frac{62}{35} \end{bmatrix}$$

Which is the same solution as part a.

c. Plot:



d. Plot:


original fit
fit with noise
10
8
6
4
2
0
−2
−1
0
1
2
3
4
5

■