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

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

a) Recall that the integral of a probability density function is 1. Also, $\mathbb{E}[X] = \int xp(x)dx$ for a probability density function $p(x)$ and random variable X.

$\mathbb{E}[Y] = \mathbb{E}[A\mathbf{x} + \mathbf{b}] = \int(A\mathbf{x} + \mathbf{b})p(x) = A\int \mathbf{x}p(x)dx + \mathbf{b}\int p(x)dx$
$= A\mathbb{E}[X] + \mathbf{b}$, as desired.

b) Recall $Cov(\mathbf{x}) = \Sigma = \mathbb{E}[(\mathbf{x} - \mathbb{E}[x])(\mathbf{x} - \mathbb{E}[x])^T$.

By linearity of expectation, constants can be taken outside of the expectation.

Here, $Cov(\mathbf{y}) = Cov(A\mathbf{x} + \mathbf{b})$
$= \mathbb{E}[(A\mathbf{x} + \mathbf{b} - \mathbb{E}[A\mathbf{x} + \mathbf{b}])(A\mathbf{x} + \mathbf{b} - \mathbb{E}[A\mathbf{x} + \mathbf{b}])^T]$.
Using what we proved in part a, the above equals $\mathbb{E}[(A\mathbf{x} + \mathbf{b} - A\mathbb{E}[\mathbf{x}] - \mathbf{b})(A\mathbf{x} + \mathbf{b} - A\mathbb{E}[\mathbf{x}] - \mathbf{b})^T]$
$= \mathbb{E}[(A\mathbf{x} - A\mathbb{E}[\mathbf{x}])(A\mathbf{x} - A\mathbb{E}[\mathbf{x}])^T]$
$= \mathbb{E}[A(\mathbf{x} - \mathbb{E}[\mathbf{x}])(A(\mathbf{x} - \mathbb{E}[\mathbf{x}]))^T]$
$= A\mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^T]A^T$
$= Acov(\mathbf{x})A^T$
$= A\Sigma A^T$, 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) Let $X = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}$. We give X a first column of all 1's to give it the correct dimensions for future matrix multiplication.

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

Calculate $X^T \mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 18 \\ 56 \end{bmatrix}$

Recall the normal equation $X^T X \theta^* = X^T \mathbf{y}$.

This equation can be thought of as $\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$.

This will help us in using Cramer's Rule, where

$\theta_0^* = \dfrac{\begin{bmatrix} c_1 & b_1 \\ c_2 & b_2 \end{bmatrix}}{\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}} = \dfrac{\begin{vmatrix} 18 & 9 \\ 56 & 29 \end{vmatrix}}{\begin{vmatrix} 4 & 9 \\ 9 & 29 \end{vmatrix}} = \frac{18}{35}$ and $\theta_1^* = \dfrac{\begin{bmatrix} a_1 & c_1 \\ a_2 & c_2 \end{bmatrix}}{\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}} = \dfrac{\begin{vmatrix} 4 & 18 \\ 9 & 56 \end{vmatrix}}{\begin{vmatrix} 4 & 9 \\ 9 & 29 \end{vmatrix}} = \frac{62}{35}$.

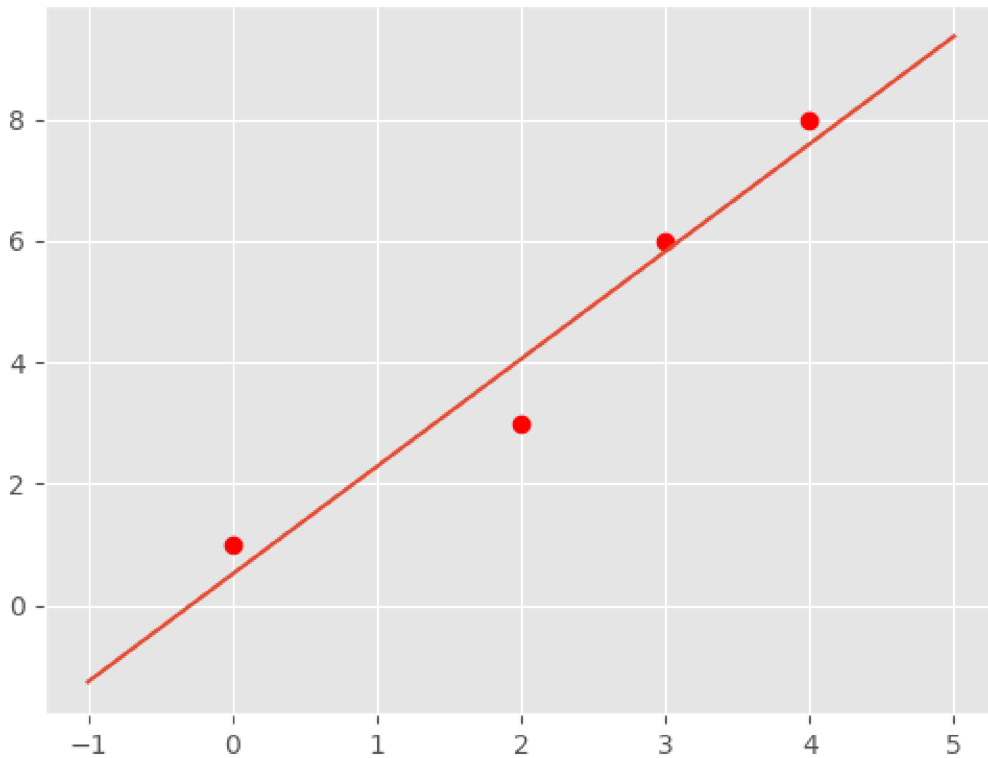
Since $y = \theta_0 + \theta_1 x$, our least squares estimate is $y = \frac{18}{35} + \frac{62}{35} x$.

b) From the normal equation, $\theta^* = (X^T X)^{-1} X^T \mathbf{y} = \begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}$

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

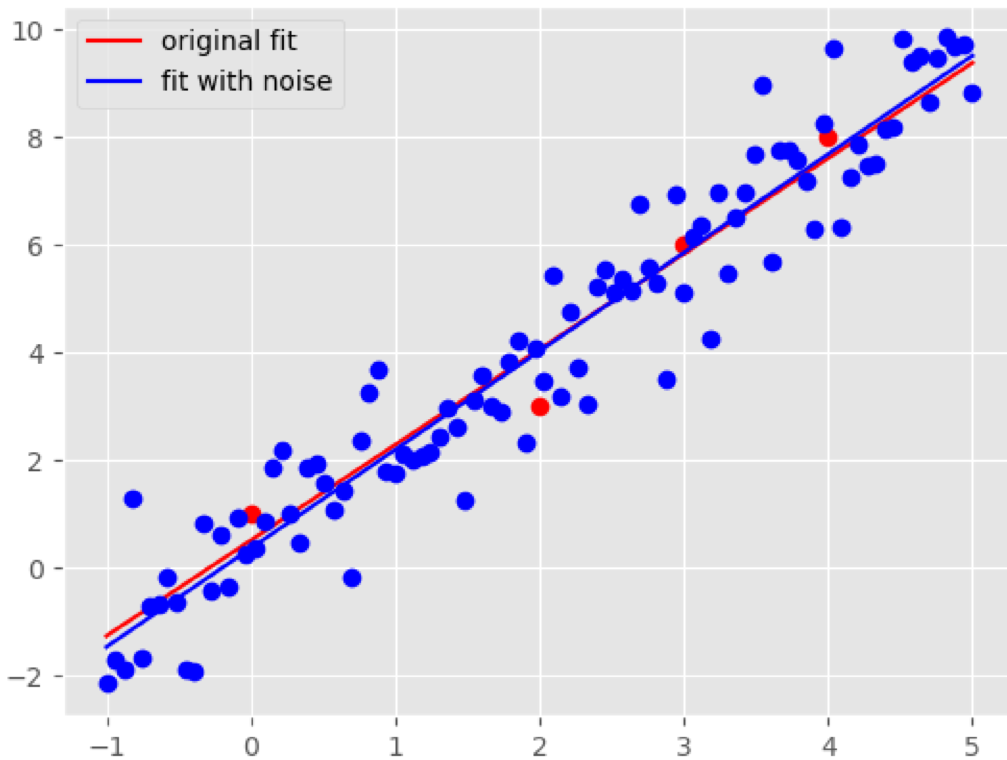
$$= \frac{1}{35}\begin{bmatrix} 29 & 11 & 2 & -7 \\ -9 & -1 & 3 & 7 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix} = \frac{1}{35}\begin{bmatrix} 18 \\ 62 \end{bmatrix} = \begin{bmatrix} \frac{18}{35} \\ \frac{62}{35} \end{bmatrix}.$$

Again, we get $\theta_0^* = \frac{18}{35}$, and we get $\theta_1^* = \frac{62}{35}$, leading to $y = \frac{18}{35} + \frac{62}{35}x$.



c)

d) The new fitted line with noise is close to the original fit.


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

■