Justin Jiang
Math189R SU20
Homework 1
Wednesday, 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

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

a) To show that A$\mathbb{E}$[**x**] + **b**is linear, we first need to integrate the given statement $\mathbb{E}$[**y**] = $\mathbb{E}$[A**x**+ **b**] such that $\mathbb{E}$[**y**] = $\int_S \mathbf{y}P(x)dx$. Since $\mathbf{y} = A\mathbf{x} + \mathbf{b}$, $\mathbb{E}$[**y**] = $\int_S (A\mathbf{x} + \mathbf{b})P(x)dx$.

Now, factor in the P(x), then split the intergral across the addition, such that:
$\int_S (A\mathbf{x})P(x)dx + \int_S \mathbf{b}P(x)dx$
Next, factor out the constants A and **b**, such that:
A$\int_S (\mathbf{x})P(x)dx + \mathbf{b} \int_S P(x)dx$
This simplifies to: A$\mathbb{E}$[**x**] + **b**, since the integral $\int_S (\mathbf{x})P(x)dx$ = $\mathbb{E}$[**x**]

b) Since $\mathbf{y} = A\mathbf{x} + \mathbf{b}$, cov[**y**] = cov[A**x**+ **b**].
Per the definition of covariance, $\mathrm{cov}[\mathbf{x}] = \mathbb{E}[(x - \mathbb{E}[\mathbf{x}])(x - \mathbb{E}[\mathbf{x}])^\top]$
Thus, $\mathrm{cov}[\mathbf{y}] = \mathrm{cov}[A\mathbf{x} + \mathbf{b}] = \mathbb{E}[(Ax + b - \mathbb{E}[Ax + b])(Ax + b - \mathbb{E}[Ax + b])^\top$
Per a), we can factor out the internal $\mathbb{E}$[Ax+b], such that: $\mathbb{E}[(Ax + b - A\mathbb{E}[x] - b)(Ax + b - A\mathbb{E}[x] - b])^\top$
Combine terms to get: $\mathbb{E}[(Ax - A\mathbb{E}[x])(Ax - A\mathbb{E}[x]^\top]$
Factor out A: $A\mathbb{E}[(x - \mathbb{E}[x])(x - \mathbb{E}[x]^\top]A^\top$
Per the definition of the covariance, $\mathbb{E}[(x - \mathbb{E}[x])(x - \mathbb{E}[x]^\top] = \mathrm{cov}[\mathbf{x}]$
Therefore, the above statement is equal to: $A\mathrm{cov}[\mathbf{x}]A^\top$
Which is equal to $A\mathbf{\Sigma}A^\top$, per definition of the covariance. QED.

■

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) X = $\begin{pmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix}$ $\mathbf{y} = \begin{pmatrix} 1 \\ 3 \\ 6 \\ 8 \end{pmatrix}$ Therefore, $X^\top X = \begin{pmatrix} 4 & 9 \\ 9 & 29 \end{pmatrix}$ ∎

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

Now, to find the 2x1 θ matrix, we need to use Cramer's rule, such that:

$\theta_0 = \det[(18, 9)(56, 29)] / \det[(4, 9)(9, 29)] = 18/35$
$\theta_1 = \det[(4, 18)(9, 56)] / \det[(4, 9)(9, 29)] = 62/35$

b) The normal equation, per the lecture notes, is $\theta^* = inv((X^\top X))X^\top \mathbf{y}$
Using the matricies from part a, and plugging them into Matlab (sorry, linalg triggers me), we get that $\theta^* = \begin{pmatrix} 18/35 \\ 62/35 \end{pmatrix}$, as per part a.

For part c and d, I included the graphs as part of my submission to github.