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

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

1a) The question is essentially asking to show that the expectation of our linear system is linear and to do so we can remember what the Expectation of a random variable is. In the case that X is a continuous random variable, we have that $\mathbb{E}$[X] is the integral of x and a probability function of x, f(x). Which can be written as $\int_{\mathbb{R}} xf(x)dx$. Knowing this we can rewrite our formula as follows.

$$\begin{aligned}
\mathbb{E}[\mathbf{y}] = \mathbb{E}[A\mathbf{x} + \mathbf{b}] = A\mathbb{E}[\mathbf{x}] + \mathbf{b}. \\
= \int_S (Ax + B) f(\mathbf{x}) dx \\
= \int_s A\mathbf{x} f(\mathbf{x}) dx \int_s \mathbf{b} f(\mathbf{x}) dx \\
= A \int_s \mathbf{x} f(\mathbf{x}) dx * \mathbf{b} \int_s f(\mathbf{x}) dx \\
= A\mathbb{E}[\mathbf{x}] + \mathbf{b}
\end{aligned}$$

The result above comes from knowing that the integral of a probability function is 1 yielding $\mathbf{b}$ by itself. □
1b) To show this we first remember that cov$[\mathbf{x}]$ is:

$$\text{cov}[\mathbf{x}] = \mathbb{E}[\mathbf{x} - \mathbb{E}[X])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^T]$$

If we combine this fact along with our findings from part a we can write our problem as such

$$
\begin{aligned}
\mathrm{cov}[\mathbf{y}] &= \mathrm{cov}[A\mathbf{x} + \mathbf{b}] \\
&= \mathbb{E}[(A\mathbf{x} + b - \mathbb{E}[A\mathbf{x} + B])(A\mathbf{x} + b - \mathbb{E}[A\mathbf{x} + B])^T] \\
&= \mathbb{E}[(A\mathbf{x} + b - A\mathbb{E}[\mathbf{x}] + b)(A\mathbf{x} + b - A\mathbb{E}[\mathbf{x}] + b)^T] \\
&= \mathbb{E}[(A\mathbf{x} - A\mathbb{E}[\mathbf{x}])(A\mathbf{x} - A\mathbb{E}[\mathbf{x}])^T] \\
&= \mathbb{E}[A(\mathbf{x} - \mathbb{E}[\mathbf{x}])A^T(\mathbf{x} - \mathbb{E}[\mathbf{x}])^T] \\
&= A\mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^T]A^T \\
&= A\mathrm{cov}(\mathbf{x})A^T \\
&= A\Sigma A^T \square
\end{aligned}
$$

■

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

(a) Find the least squares estimate $y = \boldsymbol{\theta}^\top \mathbf{x}$ 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.

2a) Since we are given that the least squares estimate is in the form $y = \boldsymbol{\theta}^\top \mathbf{x}$ we can write our points in the form

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

Therefore we can find that $X^T X = \begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}$ and $X^T Y = \begin{bmatrix} 18 \\ 56 \end{bmatrix}$

Now with these values we can apply the normal equation to use with Crammer's Rule. The normal equation is $X^T X\theta^* = X^T Y$. So seeing that this is linear we can rearrange our equation to find θ^* using Crammers Rule. So we find that $\theta_0^* = \frac{\begin{bmatrix} 18 & 9 \\ 56 & 29 \end{bmatrix}}{\begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}}$ and $\theta_1^* = \frac{\begin{bmatrix} 4 & 18 \\ 9 & 56 \end{bmatrix}}{\begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}}$ Thus we find $\theta_0^* = \frac{18}{35}$ and $\theta_1^* = \frac{62}{35}$ with them being represented in the formula as $y = \theta_0^* + \theta_1^* x$

2B) The Normal Equation is $\theta = (X^T X)^{-1} X^T \vec{y}$ So we can find out solution in the form

$$\begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix} = \begin{bmatrix} \frac{18}{35} \\ \\ \frac{62}{35} \end{bmatrix}$$

and this is the same solution as found in part a.

2C

2D

■