Math189R SU17 Homework 1 Wednesday, May 17, 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

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

- **2** Given the dataset $\mathcal{D} = \{(x, y)\} = \{(0, 1), (2, 3), (3, 6), (4, 8)\}$
 - (a) Find the least squares estimate $y = \theta^{\top} \mathbf{x}$ 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.