Abel Sapirstein
Math189R SP19
Homework 1
Monday February 4th, 2019

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. Lets demystify the expectation. Assuming that $\mathbf{x}$ is a distribution and $\mathbb{P}(\mathbf{x})$ is the probability density for $\mathbf{x}$, $\mathbb{R}$ is the region in which $\mathbf{x}$ is defined;

$$\begin{aligned}
\mathbb{E}[\mathbf{y}] &= \int_{\mathbb{R}} (A\mathbf{x} + \mathbf{b})\mathbb{P}(\mathbf{x})dx & (1)\\
&= \int_{\mathbb{R}} (A\mathbf{x})\mathbb{P}(\mathbf{x})dx + \int_{\mathbb{R}} (\mathbf{b})\mathbb{P}(\mathbf{x})dx & (2)\\
&= A\int_{\mathbb{R}} \mathbf{x}\mathbb{P}(\mathbf{x})dx + \mathbf{b}\int_{\mathbb{R}} \mathbb{P}(\mathbf{x})dx & (3)\\
&= A\mathbb{E}(\mathbf{x}) + \mathbf{b} & (4)
\end{aligned}$$

B. Lets begin with a definition of an entry into a covariant matrix;

$$\mathrm{cov}(y_i, y_j) = \mathbb{E}\left[[y_i - \mathbb{E}[y_i]][y_j - \mathbb{E}[y_j]]\right]$$

It follows that the covariant matrix, is defined as;

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

■

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.

(a) Lets begin by finding minimum the sum of the error squared;

$$J(b, m) = \sum_{i}^{n} (y_i - (mx + b))^2$$

$$\frac{\delta J}{\delta m} = -2\sum(xy) + 2m\sum(x) + 2b\sum(x^2) = 0$$

$$\frac{\delta J}{\delta b} = -2\sum(y) + 2m + 2b\sum(x) = 0$$

When we plug in the values for our four points, we arrive at the following equation;

$$\begin{bmatrix} 9 & 29 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} b \\ m \end{bmatrix} = \begin{bmatrix} 56 \\ 18 \end{bmatrix}$$

By Cramer's Rule, we can solve this as follows;

$$m = \frac{\begin{vmatrix} 56 & 29 \\ 18 & 9 \end{vmatrix}}{\begin{vmatrix} 9 & 29 \\ 4 & 9 \end{vmatrix}} = \frac{62}{35} \qquad b = \frac{\begin{vmatrix} 9 & 56 \\ 4 & 18 \end{vmatrix}}{\begin{vmatrix} 9 & 29 \\ 4 & 9 \end{vmatrix}} = \frac{18}{35}$$

(b) Using the normal equation;

$$\mathbf{X} = \begin{bmatrix} 9 & 29 \\ 4 & 29 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 56 \\ 18 \end{bmatrix}, \theta = \begin{bmatrix} b \\ m \end{bmatrix}$$

$$
\begin{aligned}
\theta &= [\mathbf{X}^T\mathbf{X}]^{-1}\mathbf{X}^T\mathbf{y} \\
&= \left[\begin{bmatrix} 9 & 4 \\ 29 & 9 \end{bmatrix} \begin{bmatrix} 9 & 29 \\ 4 & 9 \end{bmatrix}\right]^{-1} \begin{bmatrix} 9 & 4 \\ 29 & 9 \end{bmatrix} \begin{bmatrix} 56 \\ 18 \end{bmatrix} \\
&= \begin{bmatrix} 97 & 297 \\ 297 & 922 \end{bmatrix}^{-1} \begin{bmatrix} 9 & 4 \\ 29 & 9 \end{bmatrix} \begin{bmatrix} 56 \\ 18 \end{bmatrix} \\
&= \frac{1}{1225} \begin{bmatrix} -135 & 1015 \\ 140 & -315 \end{bmatrix} \begin{bmatrix} 56 \\ 18 \end{bmatrix} \\
&= \frac{1}{1225} \begin{bmatrix} 630 \\ 2170 \end{bmatrix} \\
&= \begin{bmatrix} \frac{18}{35} \\ \frac{62}{35} \end{bmatrix}
\end{aligned}
$$

■