Math 189R problem set 1

Adam Guo 2020-02-03

1. **(Linear Transformation)** Let $\mathbf{y} = A\mathbf{x} + \mathbf{b}$ be a random vector. Show that the expectation is linear:

$$\mathbb{E}[y] = \mathbb{E}[A\mathbf{x} + \mathbf{b}] = A\mathbb{E}[\mathbf{x}] + \mathbf{b}$$

Also show that

$$cov[\mathbf{y}] = cov[A\mathbf{x} + \mathbf{b}] = Acov[\mathbf{x}]A^T = A\Sigma A^T$$

Solution:

$$\begin{aligned}
\mathbb{E}[y] &= \int_C (A\mathbf{x} + \mathbf{b})Pr(x)\,\mathrm{d}x \\
&= \int_C A\mathbf{x}Pr(x)\,\mathrm{d}x + \int_C \mathbf{b}Pr(x)\,\mathrm{d}x \qquad \text{by linearity of integration} \\
&= A\int_C \mathbf{x}Pr(x)\,\mathrm{d}x + \mathbf{b}\int_C Pr(x)\,\mathrm{d}x \\
&= A\mathbb{E}[\mathbf{x}] + \mathbf{b}
\end{aligned}$$

$$\begin{aligned}
cov[\mathbf{y}] &= \mathbb{E}[(\mathbf{y} - \mathbb{E}[\mathbf{y}])(\mathbf{y} - \mathbb{E}[\mathbf{y}])^T] \\
&= \mathbb{E}[(A\mathbf{x} + \mathbf{b} - \mathbb{E}[A\mathbf{x} + \mathbf{b}])(A\mathbf{x} + \mathbf{b} - \mathbb{E}[A\mathbf{x} + \mathbf{b}])^T] \\
&= \mathbb{E}[(A\mathbf{x} + \mathbf{b} - A\mathbb{E}[\mathbf{x}] - \mathbf{b})(A\mathbf{x} + \mathbf{b} - A\mathbb{E}[\mathbf{x}] - \mathbf{b})^T] \\
&= \mathbb{E}[(A(\mathbf{x} - \mathbb{E}[\mathbf{x}]))(A(\mathbf{x} - \mathbb{E}[\mathbf{x}]))^T] \\
&= \mathbb{E}[A(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^T A^T] \\
&= A\mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^T]A^T \\
&= Acov[\mathbf{x}]A^T \\
&= A\Sigma A^T, \quad \Sigma \equiv cov[\mathbf{x}]
\end{aligned}$$

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

(a) Find the least squares estimate $y = \theta^T \mathbf{x}$ by hand using Cramer's Rule.

Solution:

Let $\mathbf{y} = \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}.$

By normal equations, $\mathbf{X}^T\mathbf{X}\hat{\theta} = \mathbf{X}^T\mathbf{y}$ such that $\hat{\theta}$ is the least squares estimate of θ.

$\mathbf{X}^T\mathbf{X} = \begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}, \quad \mathbf{X}^T\mathbf{y} = \begin{bmatrix} 18 \\ 56 \end{bmatrix}$

Using Cramer's rule,

$$\hat{\theta}_1 = \frac{\begin{vmatrix} 18 & 9 \\ 56 & 29 \end{vmatrix}}{\begin{vmatrix} 4 & 9 \\ 9 & 29 \end{vmatrix}} = \frac{18}{35}, \quad \hat{\theta}_2 = \frac{\begin{vmatrix} 4 & 18 \\ 9 & 56 \end{vmatrix}}{\begin{vmatrix} 4 & 9 \\ 9 & 29 \end{vmatrix}} = \frac{62}{35}$$

Hence, $\hat{\theta} = \begin{bmatrix} \frac{18}{35} & \frac{62}{35} \end{bmatrix}$.

(b) Use the normal equations to find the same solution and verify it is the same as part (a).

Solution:

$\theta = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$

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

$$= \frac{1}{35} \begin{bmatrix} 29 & -9 \\ -9 & 4 \end{bmatrix} \begin{bmatrix} 18 \\ 56 \end{bmatrix}$$

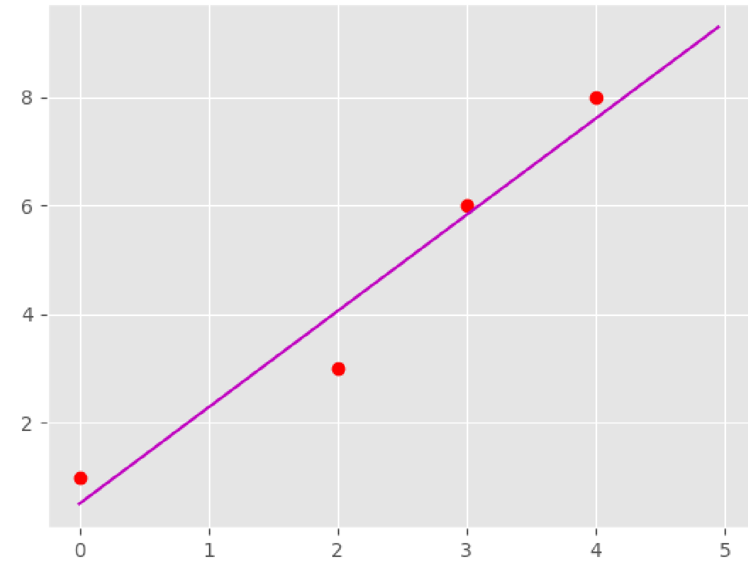
$$= \frac{1}{35} \begin{bmatrix} 18 \\ 62 \end{bmatrix}$$

$$= \begin{bmatrix} 18/35 \\ 62/35 \end{bmatrix}$$

Same solution as part a.

(c) Plot the data and the optimal linear fit you found.

Solution:



(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.

Solution:
The lines are very close.