Homework 1

Hoang Chu

February 5, 2024

Problem 1.

Solution:

1. Since X, Y contain elements that are continuous variables within the sample space S, by definition:

$$E[\mathbf{x}] = \int_S \mathbf{x} P(\mathbf{x}) dx$$

By the Law of Unconscious Statistician:

$$\begin{aligned}\mathbb{E}[\mathbf{y}] &= \int_S (A\mathbf{x} + \mathbf{b}) P(x) dx \\ &= A \int_S \mathbf{x} P(\mathbf{x}) dx + \mathbf{b} \int_S P(x) dx \\ &= A\mathbb{E}[\mathbf{x}] + \mathbf{b}\end{aligned}$$

2. By definition:

$$\operatorname{cov}[\mathbf{x}] = \mathbf{\Sigma} = \mathbb{E}\left[(\mathbf{x} - \mathbb{E}[x])(\mathbf{x} - \mathbb{E}[x])^\top\right]$$

Hence:

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

Applying Linearity of Expectation proven in Problem 1:

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

Utilizing Matrix Algebra and definition:

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

Problem 2.

Solution:

1. From the problem statement:

$$\mathbf{x} = \begin{bmatrix} 0 \\ 2 \\ 3 \\ 4 \end{bmatrix} \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}$$

Since we don't know the output of $\mathbf{y}$ when $\mathbf{x} = \mathbf{0}$, we need a bias term i.e adding a column vector of $\mathbf{1}$ to $\mathbf{x}$:

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$$

We can build the equation:

$$X^\top X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}$$

and

$$X^\top \mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 18 \\ 56 \end{bmatrix}$$

Since X's dimension is 4×2 and $\mathbf{y}$'s dimension is 4×1 we need to find $\boldsymbol{\theta}$:

$$\begin{bmatrix} \boldsymbol{\theta_0} \\ \boldsymbol{\theta_1} \end{bmatrix}$$

such that:

$$X^\top X \boldsymbol{\theta}^\star = X^\top \mathbf{y}$$

Using Cramer's Rule, we have:

$$\boldsymbol{\theta}_0^\star = \frac{\begin{vmatrix} 18 & 9 \\ 56 & 29 \end{vmatrix}}{\begin{vmatrix} 4 & 9 \\ 9 & 29 \end{vmatrix}} = \frac{18}{35} \quad \text{and} \quad \boldsymbol{\theta}_1^\star = \frac{\begin{vmatrix} 4 & 18 \\ 9 & 56 \end{vmatrix}}{\begin{vmatrix} 4 & 9 \\ 9 & 29 \end{vmatrix}} = \frac{62}{35}$$

2. Using the normal equation, we have

$$\begin{aligned}
\boldsymbol{\theta}^\star &= \left(X^\top X\right)^{-1} X^\top \mathbf{y} \\
&= \begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix} \\
&= \frac{1}{35} \begin{bmatrix} 29 & -9 \\ -9 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix} \\
&= \frac{1}{35} \begin{bmatrix} 18 \\ 62 \end{bmatrix} = \begin{bmatrix} \frac{18}{35} \\ \frac{62}{35} \end{bmatrix}
\end{aligned}$$

which matches (1)

3. See the repo code and image (p2c.png)
4. See the repo code and image (p2d.png)