Zhi Chao Peng (ZP50)

Problem 1.

From OLS: $Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n + \varepsilon$, $E[\varepsilon] = 0$

we know: $E[Y \mid X_1 = a] = \beta_0 + \beta_1 a + \cdots \beta_n X_n$

From Conditional distribution of Multivariate Normal:

$$\bar{\mu} = \mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (a - \mu_1) \quad , \text{given } X_1 = a$$

$$\bar{\Sigma} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}$$

$$Y \sim N(\bar{\mu}, \bar{\Sigma})$$

$\Rightarrow$ Intuitively, this mean Given $X_1 = a$, Y is normally distributed with mean $= \bar{\mu}$, and variance $= \bar{\Sigma}$

This implies that Given $X_1 = a$, the expected value of Y is $\bar{\mu}$ $\Rightarrow E[Y \mid X] = \bar{\mu}$

Therefore, we know that

$$\boxed{E[Y \mid X_1 = a]} = \beta_0 + \beta_1 a + \cdots + \beta_n X_n = \boxed{\bar{\mu}} = \mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (a - \mu_1)$$

Thus,

The value of conditional distribution of the Multivariate Normal and the value of OLS equation are the same

They are actually the same thing.

I also proved this in the coding part, Letting $X = 5$, And the value generated for $E[Y \mid X = 5]$ is equal to the value of $\bar{\mu}$, given $X = 5$.