Question 1.

a.

$$\begin{aligned}\mathcal{L}(\mu,\sigma|x_1,x_2,\ldots,x_n) &= \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2}\\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\sum_{i=1}^{n}\frac{1}{2\sigma^2}(x_i-\mu)^2}\\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}\end{aligned}$$

Thus

$$l(\mu,\sigma|x_1,x_2,\ldots,x_n) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

and

$$\frac{dl}{d\mu} = \frac{1}{2\sigma^2}\sum_{i=1}^{n}2(x_i-\mu) = \frac{1}{\sigma^2}\left(-n\mu+\sum_{i=1}^{n}x_i\right) = 0 \implies \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}x_i$$

b.

We also have that

$$\frac{dl}{d\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\hat{\mu})^2 = 0 \implies \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2$$

2.

a.

$$P_X^2 = X(X^TX)^{-1}\underbrace{X^TX(X^TX)^{-1}}_{I_p}X^T = X(X^TX)^{-1}X^T = P_X$$

For any $u = X\beta$,

$$P_Xu = P_X\underbrace{X\beta = X(X^TX)^{-1}X^TX}_{I_p}\beta = X\beta = \mu$$

b.

The modified version does not have part b

3.

a.

$$\hat{y} = X\hat{\beta} = X(X^T X)^{-1} X^T y = P_X y$$

b.

$$X^T y - X^T X\hat{\beta} = X^T y - X^T X(X^T X)^{-1} X^T y = X^T y - X^T y = 0$$

c.

Since $X^T y = X^T \hat{y}$, we proved that there is no better solution that can be represented using the column space of X because $X^T \hat{y}$ already has the same value.