Question 4: (Appendix C)

(i) $Z = \frac{Y}{X} \Rightarrow E(Z|X) = \frac{1}{X} E(Y|X) = \frac{\theta X}{X} = \theta$

$$E(Z) = E\left(E(Z|X)\right) = E(\theta) = \theta$$

(ii) $E(W_1 | X_1 \cdots X_n) = \frac{1}{n} \sum_{i=1}^{n} E(Y_i / X_i) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{X_i} \theta X_i$

$$= \theta$$

(iii) $W_2 = \dfrac{\frac{1}{n} \sum_{i=1}^{n} Y_i}{\frac{1}{n} \sum_{i=1}^{n} X_i} = \dfrac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} X_i} \neq \sum_{i=1}^{n} \dfrac{Y_i}{X_i}$

$$E(W_2 | X_1 \ldots X_n) = \frac{1}{\frac{1}{n} \sum_{i=1}^{n} X_i} \frac{1}{n} \sum_{i=1}^{n} E(Y_i | X_i)$$

$$= \frac{1}{\frac{1}{n} \sum_{i=1}^{n} X_i} \frac{1}{n} \sum_{i=1}^{n} \theta X_i = \theta .$$

Question 1 (Appendix D)

(i) $$AB = \begin{pmatrix} 2 & -1 & 7 \\ -4 & 5 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 6 \\ 1 & 8 & 0 \\ 3 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 20 & -6 & 12 \\ 5 & 36 & -24 \end{pmatrix}$$

(ii) BA is not defined.

QUESTION 2 (Appendix D)

$$A = \begin{bmatrix} A_{11} & 0 & \cdots & 0 \\ 0 & A_{22} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & A_{nn} \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & 0 & \cdots & 0 \\ 0 & B_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & B_{nn} \end{bmatrix}.$$

$$AB = \begin{pmatrix} A_{11}B_{11} & 0 & \cdots & 0 \\ 0 & A_{22}B_{22} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & A_{nn}B_{nn} \end{pmatrix}$$ and

$$BA = \begin{pmatrix} A_{11}B_{11} & 0 & \cdots & 0 \\ 0 & A_{22}B_{22} & & \vdots \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & A_{nn}B_{nn} \end{pmatrix}$$

Question 3:

$\underset{K\times K}{A}$ is symmetric if $A_{ij} = A_{ji}$ for all $i = 1, \ldots, K$ and $j = 1, 2, \ldots, K$. Let $\underset{n\times K}{X} = [X_{ij}]_{i=1, j=1}^{n, K}$

$$X^T X = \begin{bmatrix} \sum_{i=1}^n X_{i1}^2 & \sum_{i=1}^n X_{i1}X_{i2} & \cdots & \sum_{i=1}^n X_{i1}X_{iK} \\ \sum_{i=1}^n X_{i2}X_{i1} & \sum_{i=1}^n X_{i2}^2 & \cdots & \sum_{i=1}^n X_{i2}X_{iK} \\ \vdots & \vdots & & \vdots \\ \sum_{i=1}^n X_{iK}X_{i1} & \sum_{i=1}^n X_{iK}X_{i2} & \cdots & \sum_{i=1}^n X_{iK}^2 \end{bmatrix}_{K\times K}$$

But the (k, ℓ) position in this matrix is equal to the (ℓ, k) position, i.e., $\sum_{i=1}^n X_{ik} X_{i\ell} = \sum_{i=1}^n X_{i\ell} X_{ik}$.

Question 4 (Appendix D)

(i) The trace of a square matrix is the sum of the elements in the main diagonal. If A is of dimension $n \times m$, $A^T A$ is $m \times m$, and from Question 3,

$$tr(A^T A) = \sum_{j=1}^{m} \sum_{i=1}^{n} X_{ij}^2$$

$A A^T$ is $n \times n$ and

$$tr(A A^T) = \sum_{i=1}^{n} \sum_{j=1}^{m} X_{ij}^2 .$$

(ii) $tr(A^T A) = 14 = tr(A^T A)$

2. (a) Let $S_n(\beta) = \sum_{i=1}^{n} (Y_i - \beta X_i)^2$.

$$\frac{dS_n(\beta)}{d\beta} = \sum_{i=1}^{n} 2(Y_i - \beta X_i)(-X_i) = -2\sum_{i=1}^{n}(Y_i - \beta X_i)X_i$$

Setting $\frac{dS_n}{d\beta}(\hat\beta) = 0$, we have $\sum_{i=1}^{n}(Y_iX_i - \hat\beta X_i^2) = 0$.

The last equality implies $\hat\beta = \dfrac{\sum_{i=1}^{n} X_iY_i}{\sum_{i=1}^{n} X_i^2}$.

$$E(\hat\beta \mid X_1, \ldots, X_n) = \frac{1}{\sum_{i=1}^{n} X_i^2}\sum_{i=1}^{n} X_i E(Y_i \mid X_1 \cdots X_n)$$

$$= \frac{1}{\sum_{i=1}^{n} X_i^2} \sum_{i=1}^{n} X_i^2 \beta = \beta.$$

Hence, by the Law of Iterated expectations

$$E(\hat\beta) = E(E(\hat\beta \mid X_1, \ldots, X_n)) = E(\beta) = \beta.$$

Now,

$$\hat\beta = \frac{1}{\sum_{i=1}^{n} X_i^2}\sum_{i=1}^{n} X_i(X_i\beta + U_i) = \beta + \frac{\sum_{i=1}^{n} X_iU_i}{\sum_{i=1}^{n} X_i^2}, \quad \text{i.e.,}$$

$$\hat\beta - \beta = \frac{\sum_{i=1}^{n} X_iU_i}{\sum_{i=1}^{n} X_i^2}$$

$$V(\hat\beta - \beta \mid X_1 \cdots X_n) = \frac{1}{\left(\sum_{i=1}^{n} X_i^2\right)^2}\sum_{i=1}^{n} X_i^2 V(U_i \mid X_1, \ldots, X_n)$$

by the fact that $\{U_i\}_{i=1}^{n}$ is independent

$$= \sigma^2 \Big/ \sum_{i=1}^{n} X_i^2$$

Since as $n \to \infty$, $\sum_{i=1}^n X_i^2 \to \infty$

$$V(\hat{\beta} - \beta \mid X_1, \ldots, X_n) = V(\hat{\beta} \mid X_1, \ldots, X_n) \to 0$$

This is a desirable property. As the sample size grows the distribution of $\hat{\beta}$ conditional on $X_1 \ldots X_n$ gets closer to a degenerate distribution centered at β.

(b) Let $\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \hat{\beta} X_i)^2$ and consider

$$\begin{aligned}
\sum_{i=1}^n (Y_i - \hat{\beta} X_i)^2 &= \sum_{i=1}^n (\beta X_i + U_i - \hat{\beta} X_i)^2 \\
&= \sum_{i=1}^n \left(-(\hat{\beta} - \beta) X_i + U_i\right)^2 \\
&= \sum_{i=1}^n \left((\hat{\beta} - \beta)^2 X_i^2 + U_i^2 - 2(\hat{\beta} - \beta) X_i U_i\right) \\
&= (\hat{\beta} - \beta)^2 \sum_{i=1}^n X_i^2 + \sum_{i=1}^n U_i^2 - 2(\hat{\beta} - \beta) \sum_{i=1}^n X_i U_i
\end{aligned}$$

But since $\sum_{i=1}^n X_i U_i = (\hat{\beta} - \beta) \sum_{i=1}^n X_i^2$,

$$\sum_{i=1}^n (Y_i - \hat{\beta} X_i)^2 = -(\hat{\beta} - \beta)^2 \sum_{i=1}^n X_i^2 + \sum_{i=1}^n U_i^2.$$

Then, taking expectations conditional on $X_1, \ldots, X_n$

$$\begin{aligned}
E\left(\sum_{i=1}^n (Y_i - \hat{\beta} X_i)^2 \mid X_1, \ldots, X_n\right) &= -\sum_{i=1}^n X_i^2 \frac{\sigma^2}{\sum_{i=1}^n X_i^2} + n\sigma^2 \\
&= \sigma^2 (n-1)
\end{aligned}$$

Hence,

$$E(\hat{\sigma}^2 \mid X_1, \ldots, X_n) = \frac{1}{n-1} \sigma^2 (n-1) = \sigma^2.$$

and

$$E(\hat{\sigma}^2) = E\left(E(\hat{\sigma}^2 \mid X_1 \ldots X_n)\right) = \sigma^2$$