1. One way to obtain bootstrap estimates for the coefficients \( \beta \) from the regression

\[
y = X\beta + e, \quad E[e|X] = 0, \quad Var[e|X] = \sigma^2 I_n,
\]

is based on sampling the residuals as in the procedure below:

Step 1. Obtain the LSE \( \hat{\beta} \) and residuals \( \hat{e} \) by regressing \( y \) on \( X \) as in (*)

Step 2. Bootstrap the residuals: sample with replacement and equally likely each of the residuals in \( \hat{e} \) to obtain \( \hat{e}^* \). Note that we can write \( \hat{e}^* = B\hat{e} \) where \( B \) is a matrix where the \( i\)-th row “selects” the \( i\)-th sampled residual, that is, if the \( j\)-th residual was sampled at the \( i\)-th time then the \( i\)-th row of \( B \) has zeros in every position but \( j \), which has one \( (B_{ij} = 1) \)

Step 3. Define bootstrap responses \( y^* = X\hat{\beta} + \hat{e}^* \) using the residuals in the previous step. Note that \( \hat{\beta} \) is fixed.

Step 4. Finally, obtain a bootstrap estimate for \( \beta \) as the LSE \( \hat{\beta}^* \) from regressing \( y^* \) on \( X \).

Let \( H \) and \( \hat{y} \) be the hat matrix and the fitted values from (*).

(a) Show that \( y^* \) is a linear combination of \( y \), that is, find a matrix \( A \) that depends on \( H \) and \( B \) such that \( y^* = Ay \). Is it possible for \( y^* \) to be \( y \)? Explain.

(b) Show that (i) \( \text{Var}[\hat{e}^*|X,B] = \sigma^2 B(I-H)B^T \) and so, using the result from the previous item, that

\[
\text{Var}[y^*|X,B] = \text{Var}[\hat{y}|X] + \text{Var}[\hat{e}^*|X,B].
\]

(ii) What can you conclude from (ii) about the correlation between \( \hat{y} \) and \( \hat{e}^* \)? How would you explain this result in light of the correlation between \( \hat{y} \) and the original residuals \( \hat{e} \)?

(c) Suppose you regress the bootstrap residuals \( \hat{e}^* \) on \( X \) with mean function \( E[\hat{e}^*|X] = X\gamma \) to obtain the LSE \( \hat{\gamma} \). Now show, using the fact that regressing the fitted values \( \hat{y} \) on \( X \) yields the same LSE \( \hat{\beta} \), that \( \hat{\beta}^* = \hat{\beta} + \hat{\gamma} \). Is \( \hat{\beta}^* \) unbiased for \( \beta \)? Explain.

(d) Show that when \( \beta \) includes an intercept then, on average, the bootstrap estimate for \( \beta \) is the LSE \( \hat{\beta} \), that is, show that

\[
E_B[\hat{\beta}^*|X] = \hat{\beta}
\]

where the expectation is taken over bootstrap samples.
2. Suppose that in an experimental study you suspect that many observations were tainted by a technician and now you want to test them jointly for being outliers. To this end, you organize the suspected observations as the last $q$ observations from a total of $n$ and adopt a mean shift outlier model (MSOM) on these last observations:

$$
y_1 = x^T_1 \beta + e_1
$$

$$
\vdots
$$

$$
y_{n-q} = x^T_{n-q} \beta + e_{n-q}
$$

$$
y_{n-q+1} = x^T_{n-q+1} \beta + \delta_1 + e_{n-q+1}
$$

$$
\vdots
$$

$$
y_n = x^T_n \beta + \delta_q + e_n
$$

This model can be specified in matrix form by

$$
\begin{bmatrix}
y_1 \\
y_2 \\
y
\end{bmatrix} =
\begin{bmatrix}
X_1 & 0 \\
X_2 & I_q \\
X
\end{bmatrix}
\begin{bmatrix}
\beta \\
\delta
\end{bmatrix} + e
$$

where $E[e|X] = 0$ and $\text{Var}[e|X] = \sigma^2 I_n$ (as usual) and $\delta = [\delta_1 \cdots \delta_q]^T$. After some algebra, we can show that

$$
(X^T X)^{-1} = \begin{bmatrix}
(X_1^T X_1)^{-1} & -(X_1^T X_1)^{-1} X_2^T \\
-X_2(X_1^T X_1)^{-1} & I_q + X_2(X_1^T X_1)^{-1} X_2^T
\end{bmatrix}.
$$

Now consider $\hat{\beta}$ and $\hat{\delta}$, the LSE for $\beta$ and $\delta$ under this model, and $\hat{\beta}_1$, the LSE for $\beta$ when regressing only $y_1$ on $X_1$, that is, when ignoring the last $q$ observations.

(a) Show that (i) $\hat{\beta} = \hat{\beta}_1$ and (ii) $\hat{\delta} = y_2 - X_2 \hat{\beta}_1$, that is, the LSE for $\delta$ is the difference between the (removed) observed values and the fitted values for $X_2$ in the model without the last $q$ suspected observations.

(b) Show that the last $q$ observations are perfectly fit by the MSOM: $\hat{y}_2 = X_2 \hat{\beta} + \hat{\delta} = y_2$. What can you say about the relation between the LSE $\hat{\sigma}^2$ for $\sigma^2$ under the MSOM and the LSE $\hat{\sigma}_1^2$ for $\sigma^2$ under the model without the last $q$ observations?

(c) Find the hat matrix for the MSOM and comment on the leverage for the suspected data points in light of the results from the previous item.

(d) Conduct a joint outlier test by testing $\delta_1 = \cdots = \delta_q = 0$. State the test statistic and its distribution under the null.