1. Consider a regression model $y = X\beta + \epsilon$, where it is assumed that $E(\epsilon|X) = 0$ and $E(\epsilon\epsilon'|X) = \Sigma$. The OLS estimator of $\beta$ is $b = (X'X)^{-1}X'y$.

(a) First, suppose that you allow for heteroskedasticity in $\epsilon$, but assume there is no autocorrelation.

(i) What do we know about the numerical values of $\Sigma$?

(ii) Describe how to compute the heteroskedasticity-robust variance-covariance matrix estimator (HCCME) of $\text{Var}(b)$.

(b) Next, suppose that you allow for heteroskedasticity and autocorrelation in $\epsilon$.

(i) What do we know about the numerical values of $\Sigma$?

(ii) Describe how to compute the Newey-West “HAC” variance-covariance matrix estimator of $\text{Var}(b)$.

2. The White heteroskedasticity-consistent covariance estimator estimates the matrix $(X'X)^{-1}X'(\sigma^2\Omega)X(X'X)^{-1}$ by replacing the $(i,i)^{th}$ element of $\sigma^2\Omega$ (this element is $E(\epsilon_i^2|x_i)$ with $\epsilon_i^2$, where $\epsilon_i$ is from the OLS residual vector $\epsilon = y - Xb$, and it replaces the $(i,j)^{th}$ element of $\sigma^2\Omega$ with zero.

The Newey-West autocorrelation-consistent covariance estimator attempts to deal with autocorrelation as well as heteroskedasticity. By analogy with the White estimator, one might expect that the Newey-West estimator would replace the $(i,j)^{th}$ element of $\sigma^2\Omega$ (which is $E(\epsilon_i\epsilon_j|x_i)$) with $\epsilon_i\epsilon_j$. But instead, the Newey-West estimator does something more complicated. Why doesn’t the simpler method work?

3. Suppose $\epsilon_t$ follows a stationary AR(1) process:

$$
\epsilon_t = \rho \epsilon_{t-1} + u_t, \quad t = 1, 2, 3
$$

where $u_t$ is white noise. Let $\epsilon = [\epsilon_1 \ \epsilon_2 \ \epsilon_3]'$.

(a) Defining $E(\epsilon\epsilon') = \sigma^2_{\epsilon}\Omega$, where $\sigma^2_{\epsilon} = E(\epsilon_1^2)$, express the $3 \times 3$ matrix $\Omega$ as a function of $\rho$.

(b) For this $\Omega$, write a $3 \times 3$ matrix $P$ for which $P\epsilon$ is not autocorrelated.
(c) Calculate every element of the following $3 \times 3$ matrices as functions of $\rho$ only, using standard matrix multiplication:

(i) $P\Omega$
(ii) $(P\Omega)P'$
(iii) $P'P$
(iv) $(P'P)\Omega$

(d) Suppose that you have obtained an estimate of $\rho$ for this model as $\hat{\rho} = 0.40$. Calculate the OLS, the Prais-Winsten, and the Cochrane-Orcutt estimators of $\beta$ in the model $y = x\beta + \epsilon$ for the data:

<table>
<thead>
<tr>
<th>$i$</th>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>2.4</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

Prais-Winsten is FGLS. Cochrane-Orcutt is like FGLS, but it omits the first observation of the transformed data $y_*$ and $X_*$ from the calculation.

4. Consider a regression model: $y = X\beta + \epsilon$ where $E(\epsilon|X) = 0$ and $E(\epsilon\epsilon'|X) = \Sigma$.

(a) Describe how to compute the heteroskedasticity-robust variance-covariance matrix estimator (HCCME) of $\text{Var}(b)$.
(b) Describe how to compute the Newey-West “HAC” variance-covariance matrix estimator of $\text{Var}(b)$, which is valid when $\epsilon$ has autocorrelation and heteroskedasticity.

5. One way to write the GLS estimator of $\beta$ in the model

$$y = X\beta + \epsilon, \quad E(\epsilon|X) = 0, \quad E(\epsilon\epsilon'|X) = \Sigma$$

is $\hat{\beta} = (X_s'X_s)^{-1}X_s'y_s$, where $X_s = PX$ and $y_s = Py$ for a certain $n \times n$ matrix $P$.

(a) Write a mathematical relation involving $P$ and $\Sigma$ that must hold for $\hat{\beta}$ to be the GLS estimator.
(b) Suppose $\epsilon_t$ follows a stationary AR(1) process:

$$\epsilon_t = \rho\epsilon_{t-1} + u_t, \quad t = 1, 2, \ldots, n$$

where $u_t$ is white noise.

(i) Describe the $\Sigma$ and $P$ matrices in this case.
(ii) Describe how the vector $y_*$ is related to the original dependent variable vector $y$ in this case.

Answers

1.  
   (a) (i) off-diagonal elements equal zero 
        diagonal elements are $\geq 0$, not necessarily equal 
   (ii) Compute OLS residual vector $e = y - Xb$.
        
        \[
        S = \text{diag}(e_i) = \begin{bmatrix}
        e_1^2 & 0 & \ldots & 0 \\
        0 & e_2^2 & 0 & \ldots & 0 \\
        0 & 0 & \ddots & 0 & 0 \\
        0 & \ldots & 0 & \ddots & 0 \\
        0 & \ldots & \ldots & 0 & e_n^2
        \end{bmatrix}
        \]
        
        Then the HCCME is $(X'X)^{-1}X'SX(X'X)^{-1}$.

   (b) (i) $\Sigma$ is symmetric and positive semidefinite 
        diagonal elements are $\geq 0$, not necessarily equal 
   (ii) Like in the answer to q.8(a)(ii) except now the $(i,j)^{th}$ element of $S$ is $S_{ij} = w_{ij}e_i e_j$ 
        where 
        \[
        w_{ij} = \begin{cases}
        1 - \frac{|i-j|}{L} & \text{if } |i-j| < L \\
        0 & \text{if } |i-j| \geq L
        \end{cases}
        \]
        $L$ is an increasing function of $n$. A common choice is $L = n^{1/4}$.

2. Replacing the $(i,j)^{th}$ element of $\sigma^2 \Omega$ with $e_ie_j$, is the same as replacing the matrix $\sigma^2 \Omega$ with the matrix $ee'$, where $e$ is the vector of OLS residuals. Then the covariance matrix estimator would be

   \[
   (X'X)^{-1}X'(ee')X(X'X)^{-1} = (X'X)^{-1}(X'e)(e'X)(X'X)^{-1} = (X'X)^{-1}(X'e)(X'e)'(X'X)^{-1}
   \]
   
   But since $X'e = 0$, this covariance matrix estimator would always consist of a matrix of zeroes.

3.  
   (a) 
   \[
   \Omega = \begin{bmatrix}
   1 & \rho & \rho^2 \\
   \rho & 1 & \rho \\
   \rho^2 & \rho & 1
   \end{bmatrix}
   \]
   (b)
\[ P = \begin{bmatrix} \sqrt{1 - \rho^2} & 0 & 0 \\ -\rho & 1 & 0 \\ 0 & -\rho & 1 \end{bmatrix} \]

(c) (i)

\[ P\Omega = \begin{bmatrix} \sqrt{1 - \rho^2} \rho \sqrt{1 - \rho^2} \rho^2 \sqrt{1 - \rho^2} \\ 0 & 1 - \rho^2 & \rho^2 \\ 0 & 0 & 1 - \rho^2 \end{bmatrix} \]

(ii)

\[ (P\Omega)P' = \begin{bmatrix} \sqrt{1 - \rho^2} & 0 & 0 \\ 0 & 1 - \rho^2 & \rho^2 \\ 0 & 0 & 1 - \rho^2 \end{bmatrix} \begin{bmatrix} \sqrt{1 - \rho^2} & -\rho & 0 \\ 0 & 1 - \rho^2 & 0 \\ 0 & 0 & 1 - \rho^2 \end{bmatrix} = (1 - \rho^2)I \]

(iii)

\[ P'P = \begin{bmatrix} \sqrt{1 - \rho^2} & -\rho & 0 \\ 0 & 1 - \rho^2 & \rho \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{1 - \rho^2} & 0 & 0 \\ -\rho & 1 - \rho^2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 - \rho^2 & 0 & 0 \\ 0 & 1 - \rho^2 & 0 \\ 0 & 0 & 1 - \rho^2 \end{bmatrix} \]

(iv)

\[ (P'P)\Omega = \begin{bmatrix} 1 & -\rho & 0 \\ -\rho & 1 + \rho^2 & -\rho \\ 0 & -\rho & 1 \end{bmatrix} \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \end{bmatrix} \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \end{bmatrix} = (1 - \rho^2)I \]

(d)

\[ \text{OLS} = \frac{\sum_{i=1}^{3} x_i y_i}{\sum_{i=1}^{5} x_i^2} = \frac{94.2}{42.76} = 2.20 \]

\[ \text{FGLS (Prais-Winsten)} = \frac{\sum_{i=1}^{3} x_i y_i}{\sum_{i=1}^{5} x_i^2} = \frac{60.492}{30.252} \approx 2.00 \]
where
\[
\begin{array}{c|c|c}
  i & x_{si} & y_{si} \\
  1 & 5.50 & 11.00 \\
  2 & 0 & 3.2 \\
  3 & 0.04 & -0.2 \\
\end{array}
\]

and
\[
\text{Cochrane-Orcutt} = \frac{\sum_{i=2}^{3} x_{si} y_{si}}{\sum_{i=2}^{3} x_{si}^2} = \frac{-0.008}{0.0016} = -5.0
\]

4. (a) Compute the OLS residual vector \( e = y - Xb \).

Construct \( S = \text{diag}(e_i) = \begin{bmatrix} e_1^2 & 0 & \ldots & 0 \\ 0 & e_2^2 & 0 & \ldots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & e_n^2 \end{bmatrix} \).

Then the HCCME is \( (X'X)^{-1}X'SX(X'X)^{-1} \).

(b) Let the \((i,j)\)th element of \( S \) be \( S_{ij} = w_{ij} e_i e_j \) where
\[
w_{ij} = \begin{cases} 
1 - \frac{|i-j|}{L} & \text{if } |i-j| < L \\
0 & \text{if } |i-j| \geq L
\end{cases}
\]

\( L \) is an increasing function of \( n \). A common choice is \( L = n^{1/4} \).

5. (a) \( P \Sigma P' = I \), or \( P'P = \Sigma^{-1} \)

(b) (i)
\[
\Sigma = \sigma^2 \begin{bmatrix}
  1 & \rho & \rho^2 & \ldots & \rho^{n-1} \\
  \rho & 1 & \rho & \ldots & \rho^{n-2} \\
  \rho & \rho & \ldots & \rho & \rho^{n-2} \\
  \vdots & \vdots & \ddots & \ddots & \vdots \\
  \rho^{n-2} & \ldots & \rho & 1 & \rho \\
  \rho^{n-1} & \ldots & \rho^2 & \rho & 1
\end{bmatrix}
\]

or describe as: the \((i,j)\)th element of \( \Sigma \) is \( \sigma^2 \rho^{|i-j|} \).
\[ P = \frac{1}{\sigma_u} \begin{bmatrix} \sqrt{1 - \rho^2} & 0 & 0 & \ldots & 0 \\ -\rho & 1 & 0 & \ldots & 0 \\ 0 & -\rho & 1 & \ldots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & -\rho & 1 \end{bmatrix} \]

(ii) Letting \( y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \) then \( y_* = \begin{bmatrix} \sqrt{1 - \rho^2}y_1 \\ y_2 - \rho y_1 \\ \vdots \\ y_n - \rho y_{n-1} \end{bmatrix} \)