STAT5030 Assignment 3 solution

1. (a) $E(\hat{b}) = (X^\top X)^{-1}X^\top(Xb + Z\gamma)$, this is a biased estimator and the bias is $(X^\top X)^{-1}X^\top Z\gamma$.

   (b) $Var(\hat{b}) = (X^\top X)^{-1}X^\top Var(Y)X(X^\top X)^{-1} = \sigma^2(X^\top X)^{-1}$.

   (c)
$$E(S^2) = E\left[\frac{Y^\top(I - X(X^\top X)^{-1}X^\top)Y}{n - r(X)}\right] = \sigma^2 + \frac{(Z\gamma)^\top(I - X(X^\top X)^{-1}X^\top)(Z\gamma)}{n - r(X)}.$$

   (d) $E(S^2) = \sigma^2 + \frac{(Z\gamma)^\top(I-H)(Z\gamma)}{n-r(X)}$ is overestimated.

   (e) $E(\hat{\varepsilon}) = (I - X(X^\top X)^{-1}X^\top)(Z\gamma)$.
   $Var(\hat{\varepsilon}) = \sigma^2(I - X(X^\top X)^{-1}X^\top)$

2. (a) $E(\hat{b}) = (X^\top X)^{-1}X^\top X_1 b_1$.

   (b) $E(S^2) = \sigma^2 + \frac{(X_1b_1)^\top(I - X(X^\top X)^{-1}X^\top)(X_1b_1)}{n-r(X)}$. Since $(I - H)X = (I - H)(X_1, X_2) = 0$, then $(I - H)X_1 = 0$. Therefore $E(S^2) = \sigma^2$.

   (c) $Var(\hat{b}) = \sigma^2(X^\top X)^{-1}$.

3.
$$\frac{1}{3}[(Y_1 - Y_2)^2 + (Y_2 - Y_3)^2 + (Y_3 - Y_1)^2] \sim \chi^2_2.$$

4. Suppose SVD decomposition of $X = UDV$, where $U'U = I$, $V'V = I$ and D is diagonal matrix. Denote $D = diag(\lambda_1, \lambda_2, \ldots, \lambda_p)$.
$$\frac{1}{p}\sum_{j=0}^{p-1} var(\hat{\beta}_j) = \frac{1}{p}\sigma^2 tr((X^\top X)^{-1}V'V) = \frac{1}{p}\sigma^2\sum_{i=1}^{p}\lambda_i^{-2} \geq \frac{1}{p}\sigma^2\frac{1}{\sum_{i=1}^{p}\lambda_i^{-2}}.$$
   The equality holds iff $\lambda_1^{-2} = \lambda_2^{-2} = \ldots = \lambda_p^{-2} = \lambda$. Therefore, $X^\top X = \lambda^2 V'V = I$. Therefore, columns of X are orthogonal.

5. (a)
$$\begin{aligned}
E(\hat{b}_k) &= E(X^\top X + kI)^{-1}X^\top Y \\
&= (X^\top X + kI)^{-1}X^\top EY \\
&= (X^\top X + kI)^{-1}X^\top Xb \\
&= (X^\top X + kI)^{-1}(X^\top X + kI - kI)b \\
&= b - k(X^\top X + kI)^{-1}b.
\end{aligned}$$

   (b) $Var(\hat{b}_k) = Var(E(X^\top X + kI)^{-1}X^\top Y) = (X^\top X + kI)^{-2}X^\top X\sigma^2$.

   (c) If $\hat{b}_k$ is unbiased, $E(\hat{b}_k) = b - (X^\top X + kI)^{-1}b = b$. Then $k(X^\top X + kI)^{-1}b = 0$. Since $(X^\top X + kI)^{-1} \neq 0$, then $k = 0$.

6. (a) $\hat{\theta}_i = Y_i - \frac{1}{4}\left(\sum_{i=1}^4 Y_i - 2\pi\right), \quad i = 1, \ldots, 4.$

   (b) $\hat{\sigma}^2 = \frac{1}{4}\left(\sum_{i=1}^4 Y_i - 2\pi\right)^2.$

   (c) Test for $K^\top\theta = 0$, where $K^\top\theta = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}$.

   $M^\top\hat{\theta} \sim \left(\begin{pmatrix} \theta_1 - \theta_3 \\ \theta_2 - \theta_4 \end{pmatrix}, 2\sigma^2 I\right).$

   Then $F = \frac{Q/2\sigma^2}{SSE/\sigma^2} \frac{(K^\top\theta)^\top(K^\top\theta)}{SSE} =\sim F_{(2,1)}$.

7. Denote $(y_1, \ldots, y_m, y_{m+1}, \ldots, y_{2m}, y_{2m+1}, \ldots, y_{2m+n})$ by the observations with the first m observations being type (a), the last n being type (c) and the rest being type (b).

$$\hat{\theta} = \frac{1}{m(m+13n)}[(m+4n)\sum_{i=1}^m y_i + 6n\sum_{i=m+1}^{2m} y_i + 3m\sum_{i=2m+1}^{2m+n} y_i].$$

$$\hat{\phi} = \frac{1}{m(m+13n)}[(2n-m)\sum_{i=1}^m y_i + (m+3n)\sum_{i=m+1}^{2m} y_i - 5m\sum_{i=2m+1}^{2m+n} y_i].$$

$Cor(\hat{\theta}, \hat{\phi}) = \sigma^2(X^\top X)^{-1} = \frac{\sigma^2}{m(m+13n)}\begin{pmatrix} 4n+m & 2n-m \\ 2n-m & 2m+n \end{pmatrix}$. So these estimates are uncorrelated if $m = 2n$.

8. (a)

$$\begin{aligned} MSE(\tilde{\beta}) &= E(c(X^\top X)^{-1}X^\top Y - \beta)^\top(c(X^\top X)^{-1}X^\top Y - \beta) \\ &= tr(c^2\sigma^2(X^\top X)^{-1}) + (c-1)^2\beta^\top\beta \\ &= c^2\sigma^2 tr((X^\top X)^{-1}) + (c-1)^2\beta^\top\beta. \end{aligned}$$

   (b) $MSE(\tilde{\beta}) = tr((X^\top X)^{-1}) + (c-1)^2\beta^\top\beta$. Then minimizer $c^* = \frac{\beta^\top\beta}{\sigma^2 tr((X^\top X)^{-1}) + \beta^\top\beta}$.

   (c) $c^* = \frac{3300}{3437}$.

9. The bias is

$$E(\hat{\beta}_0) - \beta_0 = 2\beta_2, \quad E(\hat{\beta}_1) - \beta_1 = 3.4\beta_3.$$

10. (a)

$$\begin{pmatrix} Y \\ u \end{pmatrix} = \begin{pmatrix} X \\ H \end{pmatrix}\beta + \begin{pmatrix} \varepsilon \\ \gamma \end{pmatrix}.$$

   Denote $V = cov\begin{pmatrix} \varepsilon \\ \gamma \end{pmatrix} = \begin{pmatrix} \sigma^2 I & 0 \\ 0 & W \end{pmatrix}$. Then $\hat{\beta}_{GLS} = (\frac{1}{\sigma^2}X^\top X + H^\top W^{-1}H)^{-1}(\frac{1}{\sigma^2}X^\top Y + H^\top W^{-1}u)$.

(b) $\hat{\beta}_{GLS} = w_1\hat{\beta} + w_2 = (\frac{1}{\sigma^2}X^\top X + H^\top W^{-1}H)^{-1}\frac{1}{\sigma^2}X^\top X\hat{\beta} + (\frac{1}{\sigma^2}X^\top X + H^\top W^{-1}H)^{-1}H^\top W^{-1}H\hat{\beta}_a$. Then $w_1 + w_2 = I$ and $|w_1 + w_2| = 1$.