STAT 5030 assignment 3

1.(a) $E(\hat{b}) = (X^TX)^{-1}X^T(Xb+Z\gamma) = b + \underbrace{(X^TX)^{-1}X^TZ\gamma}_{\text{bias}},$

(b). $\mathrm{var}(\hat{b}) = (X^TX)^{-1}X^T\mathrm{Var}(Y)X(X^TX)^{-1} = \sigma^2(X^TX)^{-1}.$

(c) $$E(s^2) = E\left(\frac{Y^T(I-X(X^TX)^{-1}X^T)Y}{n-r(X)}\right)$$
$$= \sigma^2 + \frac{(Z\gamma)^T(I-X(X^TX)^{-1}X^T)Z\gamma}{n-r(X)}$$

(d) Since $I - X(X^TX)^{-1}X^T$ is idempotent and symmetric, it is positive semi-definite. Hence $Es^2 \geq \sigma^2$, i.e. s^2 overestimates σ^2.

(e) $\hat{\varepsilon} = (I - X(X^TX)^{-1}X)Y.$

$$E(\hat{\varepsilon}) = E((I - X(X^TX)^{-1}X^T)Y)$$
$$= I - X(X^TX)^{-1}X^T)Z\gamma$$

$$\mathrm{Var}(\hat{\varepsilon}) = (I - X(X^TX)^{-1}X^T)\mathrm{Var}(Y)(I - X(X^TX)^{-1}X^T)$$
$$= (I - X(X^TX)^{-1}X^T)\sigma^2$$

2.(a) $E(\hat{b}) = (X^TX)^{-1}X^TX_1b_1$

(b) $ES^2 = \frac{\sigma^2}{n-r(X)} E\left(Y^T\left(\frac{1}{\sigma^2}(I-H)\right)Y\right)$

$= \sigma^2 + \frac{(X_1 b_1)^T(I - X(X^TX)^{-1}X^T)(X_1 b_1)}{n-r(X)}$

and $(I-H)X=0$. Hence $(I-H)X_1=0$.

$ES^2 = \sigma^2$.

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

3. $\frac{1}{3}((Y_1-Y_2)^2+(Y_2-Y_3)^2+(Y_3-Y_1)^2)$

$= \frac{1}{3}(Y_1\ Y_2\ Y_3)\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}$

$= Y^T(I-\frac{1}{3}J)Y$

Since $I-\frac{1}{3}J$ is idempotent, $\mathrm{rank}(I-\frac{1}{3}J)=2$.

$Y^T(I-\frac{1}{3}J)Y \sim \chi_2^2$.

4. Consider the SVD decomposition of $X=U\Lambda V$, $\Lambda=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_p \\ & & & 0\end{pmatrix}$

$U^TU=I$, $V^TV=I$.

$\frac{1}{p}\sum_{j=1}^{p-1} Var(\hat{\beta}_j) = \frac{1}{p}\sigma^2 tr((X^TX)^{-1}V^TV)$

$= \frac{1}{p}\sigma^2\sum_{i=1}^{p}\lambda_i^{-2} \geq \frac{1}{p}\sigma^2\frac{1}{\sum_{i=1}^{p}\lambda_i^{2}}$

The equality holds iff $\lambda_1=\lambda_2=\cdots=\lambda_p=\lambda$ and

$\frac{1}{p}\sum_{j=0}^{p-1} Var(\hat{\beta}_j)$ is minimized.

Hence, $X^TX = \lambda^2 I$, X is orthogonal.

5. (a) $E(\hat{b}_k) = E(X^TX+kI)^{-1}X^TY$

$$= (X^TX+kI)^{-1}X^TXb$$

$$= b - k(X^TX+kI)^{-1}b.$$

(b) $var(\hat{b}_k) = var((X^TX+kI)^{-1}X^T\varepsilon)$

$$= \sigma^2(X^TX+kI)^{-1}X^TX(X^TX+kI)^{-1}$$

$$= \sigma^2((X^TX+kI)^{-1} - k(X^TX+kI)^{-2})$$

$$= \sigma^2 X^TX(X^TX+kI)^{-2}.$$

(c). $E(\hat{b}_k) = b - k(X^TX+kI)^{-1}b$

$$= (I - k(X^TX+kI)^{-1})b$$

($\Leftarrow$) If $k=0$, $E\hat{b}_k = b$.

($\Rightarrow$) If $E(\hat{b}_k) = b$, $k(X^TX+kI)^{-1}b = 0$

Assume $k \neq 0$, for $(X^TX+kI)^{-1}$, b is in the null space.

However, $(X^TX+kI)^{-1}$ is full rank, $b=0$. Contradiction.

Hence, $k=0$.

6. (a) $L(\theta) = (Y-\theta)^T(Y-\theta) - \lambda(1^T\theta - 2\pi)$

$$\begin{cases} \dfrac{\partial L}{\partial \theta} = -(Y-\theta) - \lambda 1 = 0 \\ \dfrac{\partial L}{\partial \lambda} = 1^T\theta - 2\pi = 0 \end{cases}$$

$$\hat{\theta} = Y + \frac{1}{4}(2\pi 1 - JY),$$

$$\hat{\theta}_i = Y_i + \frac{1}{4}\left(2\pi - \sum_{j=1}^{4} Y_j\right)$$

(b) Since $\varepsilon = Y - \theta \sim N(0, \sigma^2 I)$

$$1^T\varepsilon = 1^T(Y-\theta) = 1^TY - 2\pi \sim N(0, 4\sigma^2 I)$$

Since $4\sigma^2 = Var(1^T\varepsilon) = E(1^TY - 2\pi)^2$

$\frac{1}{4}(1^TY - 2\pi)^2 = \frac{1}{4}(\sum_{i=1}^{4} Y_i - 2\pi)^2$ is unbiased for σ^2.

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

$\hat{\theta} \sim N(\theta, \sigma^2(I - \frac{1}{4}J))$, where $\hat{\theta} = Y + \frac{1}{4}(2\pi 1 - JY)$

$K^T\hat{\theta} \sim N(0, 2\sigma^2 I)$

Hence $R = \frac{(K^T\hat{\theta})^T(K^T\hat{\theta})}{2\sigma^2} \sim \chi^2_2$ under H_0.

$Y - \hat{\theta} = \frac{1}{4}(JY - 2\pi 1) \sim N(0, \frac{\sigma^2}{4}J)$.

$E = \frac{(Y-\hat{\theta})^T(Y-\hat{\theta})}{\sigma^2} \sim \chi^2_1$.

$F(Y) = \frac{R/2}{E/1} \sim F(2,1)$. i.e.

$\frac{(K^T\hat{\theta})^T(K^T\hat{\theta})}{4(Y-\hat{\theta})^T(Y-\hat{\theta})} \sim F(2,1)$ under H_0.

7. Let $Y = (Y_1 \cdots Y_m \; Y_{m+1} \cdots Y_{2m} \; Y_{2m+1} \cdots Y_{2m+n})^T$

$= (\vec{Y}_a^T \quad \vec{Y}_b^T \quad \vec{Y}_c^T)^T$

$X = \begin{pmatrix} 1_m & 0 \\ 1_n & 1_m \\ 1_n & -21_n \end{pmatrix}$ $\quad \varepsilon \sim N(0, \sigma^2 I)$.

$\hat{\beta} = \begin{pmatrix} \hat{\theta} \\ \hat{\phi} \end{pmatrix} = (X^TX)^{-1}X^TY$ where

$(X^TX) = \begin{pmatrix} 2m+n & m-2n \\ m-2n & m+4n \end{pmatrix}$

$(X^TX)^{-1} = \frac{1}{m^2+13mn}\begin{pmatrix} m+4n & -m+2n \\ -m+4n & 2m+n \end{pmatrix}$

$$X^TY=\begin{pmatrix} 1_m^T & 1_m^T & 1_n^T \\ 0 & 1_m^T & -2 1_n^T \end{pmatrix}\begin{pmatrix} \vec{Y_a} \\ \vec{Y_b} \\ \vec{Y_c} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{2m+n} Y_i \\ \sum_{i=1}^{m} Y_{m+i} - 2\sum_{i=1}^{n} Y_{2m+i} \end{pmatrix}$$

$$\hat{\beta} = \begin{pmatrix} \hat{\theta} \\ \hat{\phi} \end{pmatrix} = \frac{1}{m^2+1)mn} \begin{pmatrix} (m+4n)\sum_{i=1}^{2m+n} Y_i + (-m+2n)\left(\sum_{i=1}^{m} Y_{m+i} - 2\sum_{i=1}^{n} Y_{2m+i}\right) \\ (-m+2n)\sum_{i=1}^{2m+n} Y_i + (2m+n)\left(\sum_{i=1}^{m} Y_{m+i} - 2\sum_{i=1}^{n} Y_{2m+i}\right) \end{pmatrix}$$

$\mathrm{cov}(\hat{\beta}) = \sigma^2(X^TX)^{-1}$.

$\mathrm{cov}(\hat{\theta}, \hat{\phi}) = 0$ when $m=2n$.

8. (a) $MSE(\tilde{\beta}) = E[(c(X^TX)^{-1}X^TY-\beta)^T(c(X^TX)^{-1}X^TY-\beta)]$

$= tr(c^2\sigma^2(X^TX)^{-1}) + (c-1)^2\beta^T\beta$

$= c^2\sigma^2 tr((X^TX)^{-1}) + (c-1)^2\beta^T\beta$

(b) $MSE(\tilde{\beta}) = tr(X^TX)^{-1}) + (c-1)^2\beta^T\beta$

$$\frac{\partial MSE(\hat{\beta})}{\partial c} = 2\sigma^2 c\, tr(X^TX)^{-1}) - 2(1-c)\beta^T\beta = 0.$$

$$\Rightarrow c^* = \frac{\beta^T\beta}{\sigma^2 tr((X^TX)^{-1}) + \beta^T\beta}$$

(c) Plug in. $c^* = \dfrac{\sum_{i=1}^{5} i^2}{\left(\sum_{i=1}^{5} \frac{1}{i}\right) + \left(\sum_{i=1}^{5} i^2\right)} \approx 0.96.$

9. Denote $X = (X^1\ X^2)$, $X^1 = \begin{pmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_5 \end{pmatrix}$. $X^2 = \begin{pmatrix} X_1^2 & X_1^3 \\ \vdots & \vdots \\ X_5^2 & X_5^3 \end{pmatrix}$

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = (X^{1T}X^1)^{-1}X^{1T}Y.$$

$$E\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + (X^{1T}X^1)^{-1}X^{1T}X^2\begin{pmatrix} \beta_2 \\ \beta_3 \end{pmatrix}$$

$$\text{bias} = (X^{1T}X^1)^{-1}X^{1T}X^2\begin{pmatrix} \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 2\beta_2 \\ 3.4\beta_3 \end{pmatrix}$$

10. (a)

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

Denote $\tilde{Y} = \begin{pmatrix} Y \\ u \end{pmatrix}$ $\tilde{X} = \begin{pmatrix} X \\ H \end{pmatrix} \beta$ $\tilde{\varepsilon} = \begin{pmatrix} \varepsilon \\ r \end{pmatrix}$ $\Sigma = var(\tilde{\varepsilon}) = \begin{pmatrix} \sigma^2 I & 0 \\ 0 & W \end{pmatrix}$

$$\hat{\beta}_k = (\tilde{X}^T \Sigma^{-1} \tilde{X})^{-1} \tilde{X}^T \Sigma^{-1} \tilde{Y}$$

$$= \left(\frac{1}{\sigma^2} X^T X + H^T W^{-1} H\right)^{-1} \left(\frac{1}{\sigma^2} X^T Y + H^T W^{-1} u\right).$$

(b) $\hat{\beta}_k = \left(\frac{1}{\sigma^2} X^T X + H^T W^{-1} H\right)^{-1} \cdot \frac{1}{\sigma^2} (X^T X)(X^T X)^{-1} X^T Y +$

$\left(\frac{1}{\sigma^2} X^T X + H^T W^{-1} H\right)^{-1} \cdot (H^T W^{-1} H)(H^T W^{-1} H)^{-1} H^T W^{-1} u$

$= W_1 \hat{\beta} + W_2 \hat{\beta}_a.$

$W_1 = \left(\frac{1}{\sigma^2} X^T X + H^T W^{-1} H\right)^{-1} \frac{1}{\sigma^2} (X^T X)$

$W_2 = \left(\frac{1}{\sigma^2} X^T X + H^T W^{-1} H\right)^{-1} (H^T W^{-1} H)$

$W_1 + W_2 = I.$ $|W_1 + W_2| = 1.$