Question 1:

1: $\hat{\beta} = (X'X)^{-1}X'y$ least square estimate of β.

$$U = y - X(X'X)^{-1}X'y = (I-H)y$$

$$(I-H)X\beta = [X - X(X'X)^{-1}X'X]\beta = (X-X)\beta = 0$$

$$U = (I-H)(y - X\beta) = (I-H)\varepsilon.$$

2. $V = X(\hat{\beta}-\beta) = X[(X'X)^{-1}X'y - (X'X)^{-1}X'X\beta]$

$$= X(X'X)^{-1}X'y - X(X'X)^{-1}X'X\beta$$

$$= Hy - HX\beta = H(y - X\beta) = H\varepsilon.$$

3. $\mathrm{cov}[U,V] = \mathrm{cov}[(I-H)\varepsilon, H\varepsilon] = \sigma^2(I-H)H =$

$$= \sigma^2(H - H^2) = 0.$$

4. $(Y - X\beta)^T(Y - X\beta) = [Y - X\hat{\beta} + X(\hat{\beta}-\beta)]'[Y - X\hat{\beta} + X(\hat{\beta}-\beta)]$

$$= (Y - X\hat{\beta})'(Y - X\hat{\beta}) + 2(\hat{\beta}-\beta)'X'(Y - X\hat{\beta}) + (\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta)$$

$$= (Y - X\hat{\beta})'(Y - X\hat{\beta}) + (\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta)$$

since $(\hat{\beta}-\beta)'X'(Y - X\hat{\beta}) = (\hat{\beta}-\beta)(X'Y - X'X\hat{\beta}) = 0$

5. The left side is minimized uniquely when $\beta = \hat{\beta}$.

6. The left side is $\sum_{i=1}^{n} \varepsilon_i^2/\sigma^2 \sim \chi_n^2$.

The second term of the right side: $\hat{\beta} - \beta \sim N_p(0, \sigma^2 (X'X)^{-1})$

$$(\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta)/\sigma^2 = (\hat{\beta}-\beta)'[Var(\hat{\beta})]^{-1}(\hat{\beta}-\beta) \sim \chi_p^2$$

$\Rightarrow$ The first & second terms of the right side are independent $\Rightarrow (y - X\hat{\beta})^T(y - X\hat{\beta})/\sigma^2 \sim \chi^2_{n-p}$.

Question 2:

1. $X = 1_n$

$$\hat{\beta} = (1_n' 1_n)^{-1} 1_n' y = \frac{1}{n} 1_n' y = \bar{y}$$

2. $\hat{\beta}$ is the mean $\bar{y} = \frac{1}{n} \dots y$

3. $H = 1_n (1_n' 1_n)^{-1} 1_n' = \frac{1}{n}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \vdots & & & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} = \frac{1}{n} J_n$

$\rightarrow$ orthogonal projector onto $V = span(1_n)$

4. Let $y_c = [y_1 - \bar{y}, \dots, y_n - \bar{y}]' = (y - \hat{y})' = (y - Hy)'$
$= y'(I - H)$

$$I - H = \begin{bmatrix} 1-\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & \cdots & -\frac{1}{n} \\ \vdots & & & \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & 1-\frac{1}{n} \end{bmatrix}$$

$$\sum_{i=1}^{n} (y_i - \bar{y})^2 = y'(I-H)y$$

5. $\sum_{i=1}^{n} \bar{y}(y_i - \bar{y}) = y'H(I-H)y = 0 \longrightarrow cov(Hy, (I-H)y) = 0$ ③

6. Using $\bar{y} \sim N(\beta, \sigma^2/n)$ and 6 of question 1.

Question 3

1. $L(\beta, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\|y - X\beta\|^2\right\}$

Let $\vartheta = \sigma^2$, then $l(\beta, \vartheta) = \log L(\beta, \sigma^2)$ and ignoring constants

$$l(\beta, \vartheta) = -\frac{n}{2}\log\vartheta - \frac{1}{2\vartheta}\|y - X\beta\|^2$$

2. $\frac{\partial l}{\partial \beta} = -\frac{1}{2\vartheta}(-2X'y + 2X'X\beta)$

setting $\frac{\partial l}{\partial \beta} = 0 \longrightarrow$ we get the least squares estimate of β

It maximizes $l(\beta, \vartheta)$ for any $\vartheta > 0$.

3. Maximize $l(\hat{\beta}, \vartheta)$ with respect to ϑ

$$\frac{\partial l}{\partial \vartheta} = -\frac{n}{2\vartheta} + \frac{1}{2\vartheta^2}\|y - X\hat{\beta}\|^2$$

setting $\frac{\partial l}{\partial \vartheta} = 0$, we get $\frac{n}{2\vartheta} = \frac{1}{2\vartheta^2}\|y - X\hat{\beta}\|^2$

$$\Rightarrow \hat{\vartheta} = \frac{1}{n}\|y - X\hat{\beta}\|^2$$

4. $l(\hat{\beta}, \hat{\vartheta}) = -\frac{n}{2}\log\hat{\vartheta} - \frac{n}{2}$ or $L(\hat{\beta}, \hat{\vartheta}) = (2\pi\hat{\sigma}^2)^{-n/2} e^{-n/2}$

5. $$\frac{\partial^2 l}{\partial\beta\partial\beta'} = -\frac{1}{\vartheta}(X'X)$$

(4)

$$\frac{\partial^2 l}{\partial\vartheta^2} = \frac{n}{2\vartheta^2} - \frac{1}{\vartheta^3}\|y-X\beta\|^2 \qquad I = \begin{pmatrix} \frac{1}{\vartheta}(X'X) & 0 \\ 0 & \frac{n}{2\vartheta^2}\end{pmatrix}$$

$$\frac{\partial^2 l}{\partial\beta\partial\vartheta} = \frac{1}{\vartheta^2}(-2X'y+2X'X\beta) \qquad I^{-1} = \begin{pmatrix} \vartheta(X'X)^{-1} & 0 \\ 0 & \frac{2\vartheta^2}{n}\end{pmatrix}$$

6 since $Var(\hat\beta) = \vartheta(X'X)^{-1} \longrightarrow \hat\beta$ is the best unbiased estimator of β

since $(n-p)(y-X\hat\beta)'(y-X\hat\beta)/\vartheta \sim \chi^2_{n-p}$ and $Var(\chi^2_{n-p}) = 2(n-p)$

$Var[(y-X\hat\beta)'(y-X\hat\beta)] = \frac{2\vartheta^2}{n-p} \longrightarrow \frac{2\vartheta^2}{n}, n\to\infty$

$\Rightarrow (y-X\hat\beta)'(y-X\hat\beta)$ is also the minimum variance unbiased estimator of ϑ.

Question 4:

One Lagrange multiplier for each constraint

$$\lambda_i \rightsquigarrow a_i\beta = c_i \quad (i=1,2,\dots,q)$$

a_i : i^{th} row of

$$\sum_{i=1}^{q}\lambda_i'(a_i\beta - c_i) = \lambda'(A\beta - C) = (\beta'A' - C')\lambda$$

We consider the cost

$$L = (Y-X\beta)'(Y-X\beta) + (\beta'A'-C')\lambda$$

$$\frac{\partial L}{\partial\beta} = -2X'y + 2X'X\beta + A'\lambda = 0$$

$$\hat\beta_c = (X'X)^{-1}X'y - \frac{1}{2}(X'X)^{-1}A'\lambda$$

$$\hat\beta_c = \hat\beta_{LS} - \frac{1}{2}(X'X)^{-1}A'\lambda$$

The constraint gives

$$c = A\hat\beta_c = A\hat\beta_{LS} - \frac{1}{2}A(X'X)^{-1}A'\lambda$$

$$A(X'X)^{-1}A'\lambda = 2(A\hat\beta_{LS} - c)$$

$$\lambda = [A(X'X)^{-1}A']^{-1}2(A\hat\beta_{LS} - c)$$

A is rank q and $(X'X)^{-1}$ is p.d. $\longrightarrow A(X'X)^{-1}A'$ is p.d.

$$\hat\beta_C = \hat\beta_{LS} + (X'X)^{-1}A'[A(X'X)^{-1}A']^{-1}(c - A\hat\beta_{LS})$$

Question 5:

1. $X = S\begin{pmatrix}\Sigma_r & 0\\ 0 & 0\end{pmatrix}\Theta^T$, $\quad X^TX = \Theta\begin{pmatrix}\Sigma_r & 0\\ 0 & 0\end{pmatrix}S^TS\begin{pmatrix}\Sigma_r & 0\\ 0 & 0\end{pmatrix}\Theta^T$

$$= \Theta\begin{pmatrix}\Sigma_r^2 & 0\\ 0 & 0\end{pmatrix}\Theta^T$$

$$\beta = (X'X)^{-1}X^Ty = \Theta\begin{pmatrix}\Sigma_r^{-2} & 0\\ 0 & 0\end{pmatrix}\Theta^T\Theta\begin{pmatrix}\Sigma_r & 0\\ 0 & 0\end{pmatrix}S^Ty$$

$$= \Theta\begin{pmatrix}\Sigma_r^{-1} & 0\\ 0 & 0\end{pmatrix}S^Ty$$

2. $\beta = \sum_{i=1}^{r}\frac{s_i^Ty}{\sigma_i}q_i$

3. $\beta_p = \sum_{i=1}^{k}\frac{s_i^Ty}{\sigma_i}q_i$, β_p uses k vector of Θ instead of r

4. $R_k = \|y - X\hat\beta_p\|^2 = \|S^Ty - \begin{pmatrix}\Sigma\\ 0\end{pmatrix}z\|^2$, $\quad z = \begin{pmatrix}s_1^Ty/\sigma_1\\ \vdots\\ s_k^Ty/\sigma_k\\ 0\\ \vdots\\ 0\end{pmatrix}$

$$R_k = \sum_{i=k+1}^{n}(s_i^Ty)^2$$