Solutions to Problem Set 3: Regression Analysis

Jura Ivanković

MIT Financial Mathematics course website: `https://ocw.mit.edu/courses/mathematics/18-s096-topics-in-mathematics-with-applications-in-finance-fall-2013/`

Problem sets: `https://ocw.mit.edu/courses/mathematics/18-s096-topics-in-mathematics-with-applications-in-finance-fall-2013/assignments/`

Problem set 3: `https://ocw.mit.edu/courses/mathematics/18-s096-topics-in-mathematics-with-applications-in-finance-fall-2013/assignments/MIT18_S096F13_pset3.pdf`

1. (a)

$$\begin{aligned} H^T &= \left(X\left(X^TX\right)^{-1}X^T\right)^T \\ &= X^{T^T}\left(X\left(X^TX\right)^{-1}\right)^T \\ &= X\left(X^TX\right)^{-1^T}X^T \\ &= X\left(X^TX\right)^{T^{-1}}X^T \\ &= X\left(X^TX\right)^{-1}X^T \\ &= H \end{aligned}$$

$$H \times H = X\left(X^TX\right)^{-1}X^TX\left(X^TX\right)^{-1}X^T = X\left(X^TX\right)^{-1}X^T = H$$

(b) Let $H_{i,*}$ denote the ith row of H. Then,

$$\frac{d\hat{y}_i}{dy_i} = \frac{d(H_{i,*}y)}{dy_i} = \frac{d\left(\sum_{j=1}^n H_{i,j}y_j\right)}{dy_i} = \sum_{j=1}^n H_{i,j}\frac{dy_j}{dy_i} = H_{i,i} \qquad \left(\because \frac{dy_j}{dy_i} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}\right)$$

(c)

$$Average(H_{i,i}) = \frac{tr(H)}{n} = \frac{tr\left(X\left(X^TX\right)^{-1}X^T\right)}{n} = \frac{tr\left(\left(X^TX\right)^{-1}X^TX\right)}{n} = \frac{tr(I_p)}{n} = \frac{p}{n}$$

(d)

$$X'\left(X'^TX'\right)^{-1}X'^T = XG\left((XG)^TXG\right)^{-1}(XG)^T$$

$$
\begin{aligned}
&= XG\left(G^T X^T X G\right)^{-1} G^T X^T \\
&= XGG^{-1}\left(G^T X^T X\right)^{-1} G^T X^T \\
&= X\left(X^T X\right)^{-1} G^{T^{-1}} G^T X^T \\
&= X\left(X^T X\right)^{-1} X^T
\end{aligned}
$$

(e)

$$X'\beta' = XGG^{-1}\beta = X\beta$$

$\therefore$ If β is the regression parameter for X, then β' is the regression parameter for X'.

$$\begin{pmatrix} 1 & -\bar{x}_1 & -\bar{x}_2 & \cdots & -\bar{x}_p \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \bar{x}_1 & \bar{x}_2 & \cdots & \bar{x}_p \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} = I_{p+1}$$

$$\therefore G^{-1} = \begin{pmatrix} 1 & \bar{x}_1 & \bar{x}_2 & \cdots & \bar{x}_p \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

$$\therefore \beta' = G^{-1}\beta = \begin{pmatrix} \beta_0 + \bar{x}_1\beta_1 + \bar{x}_2\beta_2 + \cdots + \bar{x}_p\beta_p \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}$$

$$X'_{*,1} = XG_{*,1} = X \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = X_{*,1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

Let $k, l > 1$.

$$G_{i,k} = \begin{cases} -\bar{x}_{k-1} & i = 1 \\ 1 & i = k \\ 0 & \text{otherwise} \end{cases}$$

$$X_{i,*}G_{*,k} = \sum_{j=1}^{n} X_{i,j}G_{j,k} = X_{i,1}(-\bar{x}_{k-1}) + X_{i,k} = x_{i,k-1} - \bar{x}_{k-1} \quad \forall i \in \{1, 2, \ldots, n\}$$

$$X'_{*,k} = \begin{pmatrix} X_{1,*}G_{*,k} \\ X_{2,*}G_{*,k} \\ \vdots \\ X_{n,*}G_{*,k} \end{pmatrix} = \begin{pmatrix} x_{1,k-1} - \bar{x}_{k-1} \\ x_{2,k-1} - \bar{x}_{k-1} \\ \vdots \\ x_{n,k-1} - \bar{x}_{k-1} \end{pmatrix}$$

$$\left(X'^T X'\right)_{1,1} = (X'_{*,1})^T X'_{*,1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}^T \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \sum_{i=1}^{n} 1 \times 1 = \sum_{i=1}^{n} 1 = n$$

$$\begin{aligned}
\left(X'^T X'\right)_{k,1} &= (X'_{*,k})^T X'_{*,1} \\
&= \begin{pmatrix} x_{1,k-1} - \bar{x}_{k-1} \\ x_{2,k-1} - \bar{x}_{k-1} \\ \vdots \\ x_{n,k-1} - \bar{x}_{k-1} \end{pmatrix}^T \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \\
&= \sum_{i=1}^{n} (x_{i,k-1} - \bar{x}_{k-1}) \\
&= \sum_{i=1}^{n} x_{i,k-1} - n\bar{x}_{k-1} \\
&= n\bar{x}_{k-1} - n\bar{x}_{k-1} \\
&= 0
\end{aligned}$$

$$\left(X'^T X'\right)_{1,k} = (X'_{*,1})^T X'_{*,k} = (X'_{*,k})^T X'_{*,1} = 0$$

$$\begin{aligned}
\left(X'^T X'\right)_{k,l} &= (X'_{*,k})^T X'_{*,l} \\
&= \begin{pmatrix} x_{1,k-1} - \bar{x}_{k-1} \\ x_{2,k-1} - \bar{x}_{k-1} \\ \vdots \\ x_{n,k-1} - \bar{x}_{k-1} \end{pmatrix}^T \begin{pmatrix} x_{1,l-1} - \bar{x}_{l-1} \\ x_{2,l-1} - \bar{x}_{l-1} \\ \vdots \\ x_{n,l-1} - \bar{x}_{l-1} \end{pmatrix} \\
&= (\mathcal{X}_{*,k-1})^T \mathcal{X}_{*,l-1} \\
&= \left(\mathcal{X}^T \mathcal{X}\right)_{k-1,l-1}
\end{aligned}$$

$$\therefore X'^T X' = \begin{pmatrix} n & \mathbf{0}_p^T \\ \mathbf{0}_p & \mathcal{X}^T \mathcal{X} \end{pmatrix}$$

$$\begin{pmatrix} n & \mathbf{0}_p^T \\ \mathbf{0}_p & \mathcal{X}^T \mathcal{X} \end{pmatrix} \begin{pmatrix} \frac{1}{n} & \mathbf{0}_p^T \\ \mathbf{0}_p & \left(\mathcal{X}^T \mathcal{X}\right)^{-1} \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0}_p^T \\ \mathbf{0}_p & I_p \end{pmatrix} \implies \left(X'^T X'\right)^{-1} = \begin{pmatrix} \frac{1}{n} & \mathbf{0}_p^T \\ \mathbf{0}_p & \left(\mathcal{X}^T \mathcal{X}\right)^{-1} \end{pmatrix}$$

$$\therefore H_{i,j} = X'_{i,*} \left(X'^T X'\right)^{-1} (X'_{j,*})^T$$

$$
\begin{aligned}
&= \begin{pmatrix} 1 & x_{i,1} - \bar{x}_1 & \cdots & x_{i,p} - \bar{x}_p \end{pmatrix} \begin{pmatrix} \frac{1}{n} & \mathbf{0}_p^T \\ \mathbf{0}_p & \left(\mathcal{X}^T\mathcal{X}\right)^{-1} \end{pmatrix} \begin{pmatrix} 1 \\ x_{j,1} - \bar{x}_1 \\ \vdots \\ x_{j,p} - \bar{x}_p \end{pmatrix} \\
&= \begin{pmatrix} 1 & (x_i - \bar{x})^T \end{pmatrix} \begin{pmatrix} \frac{1}{n} & \mathbf{0}_p^T \\ \mathbf{0}_p & \left(\mathcal{X}^T\mathcal{X}\right)^{-1} \end{pmatrix} \begin{pmatrix} 1 \\ x_j - \bar{x} \end{pmatrix} \\
&= \begin{pmatrix} 1 & (x_i - \bar{x}_1)^T \end{pmatrix} \begin{pmatrix} \frac{1}{n} \\ \left(\mathcal{X}^T\mathcal{X}\right)^{-1}(x_j - \bar{x}) \end{pmatrix} \\
&= \frac{1}{n} + (x_i - \bar{x}_1)^T \left(\mathcal{X}^T\mathcal{X}\right)^{-1}(x_j - \bar{x})
\end{aligned}
$$

2. (a)

$$
\begin{aligned}
&\left(A + a^T b\right)\left(A^{-1} - A^{-1}a^T\left(I_q + bA^{-1}a^T\right)^{-1}bA^{-1}\right) \\
&= I_p - a^T\left(I_q + bA^{-1}a^T\right)^{-1}bA^{-1} + a^T bA^{-1} - a^T bA^{-1}a^T\left(I_q + bA^{-1}a^T\right)^{-1}bA^{-1} \\
&= I_p + a^T bA^{-1} - a^T\left(I_q + bA^{-1}a^T\right)\left(I_q + bA^{-1}a^T\right)^{-1}bA^{-1} \\
&= I_p + a^T bA^{-1} - a^T bA^{-1} \\
&= I_p
\end{aligned}
$$

$$
\therefore \left(A + a^T b\right)^{-1} = \left(A^{-1} - A^{-1}a^T\left(I_q + bA^{-1}a^T\right)^{-1}bA^{-1}\right)
$$

(b) Let

$$
X_{(i)} = \begin{pmatrix} x_1^T \\ \vdots \\ x_{i-1}^T \\ x_{i+1}^T \\ \vdots \\ x_n^T \end{pmatrix} \qquad y_{(i)} = \begin{pmatrix} y_1 \\ \vdots \\ y_{i-1} \\ y_{i+1} \\ \vdots \\ y_n \end{pmatrix}
$$

$$
\left(X_{(i)}^T X_{(i)}\right)_{j,k} = \left(X_{(i)_{*,j}}\right)^T X_{(i)_{*,k}} = \left(X^T X\right)_{j,k} - x_{ij}x_{ik}
$$

$$
X_{(i)}^T X_{(i)} = X^T X - x_i x_i^T
$$

$$
\left(X_{(i)}^T y_{(i)}\right)_j = \left(X_{(i)_{*,j}}\right)^T y_{(i)} = X_{*,j}y - x_{ij}y_i
$$

$$
X'^T y' = X^T y - x_i y_i
$$

$$
\begin{aligned}
\therefore \beta_{(i)} &= \left(X^T X - x_i x_i^T\right)^{-1}\left(X^T y - x_i y_i\right) \\
&= \left(\left(X^T X\right)^{-1} + \left(X^T X\right)^{-1} x_i \left(1 - x_i^T\left(X^T X\right)^{-1}x_i\right)^{-1} x_i^T\left(X^T X\right)^{-1}\right)\left(X^T y - x_i y_i\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \left(\left(X^TX\right)^{-1} + \frac{\left(X^TX\right)^{-1} x_i x_i^T \left(X^TX\right)^{-1}}{1-H_{i,i}}\right)\left(X^Ty - x_iy_i\right) \\
&= \left(X^TX\right)^{-1} X^Ty - \left(X^TX\right)^{-1} x_iy_i + \frac{\left(X^TX\right)^{-1} x_i x_i^T \left(X^TX\right)^{-1} \left(X^Ty - x_iy_i\right)}{1-H_{i,i}} \\
&= \hat{\beta} - \frac{\left(X^TX\right)^{-1} x_iy_i\left(1-H_{i,i}\right)}{1-H_{i,i}} + \frac{\left(X^TX\right)^{-1} x_i \left(x_i^T \left(X^TX\right)^{-1} X^Ty - x_i^T \left(X^TX\right)^{-1} x_iy_i\right)}{1-H_{i,i}} \\
&= \hat{\beta} - \frac{\left(X^TX\right)^{-1} x_iy_i(1-H_{i,i})}{1-H_{i,i}} + \frac{\left(X^TX\right)^{-1} x_i \left(x_i^T\hat{\beta} - H_{i,i}y_i\right)}{1-H_{i,i}} \\
&= \hat{\beta} - \frac{\left(X^TX\right)^{-1} x_i \left(y_i(1-H_{i,i}) - x_i^T\hat{\beta} + H_{i,i}y_i\right)}{1-H_{i,i}} \\
&= \hat{\beta} - \frac{\left(X^TX\right)^{-1} x_i \left(y_i - x_i^T\hat{\beta}\right)}{1-H_{i,i}} \\
&= \hat{\beta} - \frac{\left(X^TX\right)^{-1} x_i\hat{\epsilon}_i}{1-H_{i,i}}
\end{aligned}
$$

(c)

$$
\begin{aligned}
CD_i &= \left(\frac{1}{p\hat{\sigma}^2}\right)|\hat{y} - \hat{y}_{(i)}|^2 \\
&= \left(\frac{1}{p\hat{\sigma}^2}\right)|X(\hat{\beta} - \hat{\beta}_{(i)})|^2 \\
&= \left(\frac{1}{p\hat{\sigma}^2}\right)\left|X\frac{\left(X^TX\right)^{-1} x_i\hat{\epsilon}_i}{1-H_{i,i}}\right|^2 \\
&= \left(\frac{\hat{\epsilon}_i^2}{p\hat{\sigma}^2}\right)\frac{\left|X(X^TX)^{-1}x_i\right|^2}{(1-H_{i,i})^2} \\
&= \left(\frac{\hat{\epsilon}_i^2}{p\hat{\sigma}^2}\right)\frac{\left(X\left(X^TX\right)^{-1}x_i\right)^T\left(X\left(X^TX\right)^{-1}x_i\right)}{(1-H_{i,i})^2} \\
&= \left(\frac{\hat{\epsilon}_i^2}{p\hat{\sigma}^2}\right)\frac{x_i^T{\left(X^TX\right)^{-1}}^T X^TX\left(X^TX\right)^{-1}x_i}{(1-H_{i,i})^2} \\
&= \left(\frac{\hat{\epsilon}_i^2}{p\hat{\sigma}^2}\right)\frac{x_i^T\left(X^TX\right)^{T^{-1}}x_i}{(1-H_{i,i})^2} \\
&= \left(\frac{\hat{\epsilon}_i^2}{p\hat{\sigma}^2}\right)\frac{x_i^T\left(X^TX\right)^{-1}x_i}{(1-H_{i,i})^2} \\
&= \left(\frac{\hat{\epsilon}_i^2}{p\hat{\sigma}^2}\right)\frac{H_{i,i}}{(1-H_{i,i})^2}
\end{aligned}
$$

(d)

$$(H_{*,i})^T = H_{i,*} \quad (\because H^T = H)$$

$$|H_{*,i}|^2 = (H_{*,i})^T H_{*,i} = H_{i,*} H_{*,i} = H_{i,i} \quad (\because H^2 = H)$$

$$(H_{*,i})^T \hat{\epsilon} = H_{i,*}\hat{\epsilon} = 0 \quad (\because H\hat{\epsilon} = Hy - H\hat{y} = Hy - HHy = Hy - Hy = \mathbf{0}_n)$$

$$\begin{aligned}
\therefore \hat{\sigma}^2_{(i)} &= \frac{|y - X\hat{\beta}_{(i)}|^2 - \left|y_i - x_i^T\hat{\beta}_{(i)}\right|^2}{n-p-1} \\
&= \frac{\left|y - X\left(\hat{\beta} - \frac{(X^TX)^{-1}x_i\hat{\epsilon}_i}{1-H_{i,i}}\right)\right|^2 - \left(y_i - x_i^T\left(\hat{\beta} - \frac{(X^TX)^{-1}x_i\hat{\epsilon}_i}{1-H_{i,i}}\right)\right)^2}{n-p-1} \\
&= \frac{\left|y - X\hat{\beta} + \frac{X(X^TX)^{-1}x_i\hat{\epsilon}_i}{1-H_{i,i}}\right|^2 - \left(y_i - x_i^T\hat{\beta} + \frac{x_i^T(X^TX)^{-1}x_i\hat{\epsilon}_i}{1-H_{i,i}}\right)^2}{n-p-1} \\
&= \frac{\left|\hat{\epsilon} + \frac{H_{*,i}\hat{\epsilon}_i}{1-H_{i,i}}\right|^2 - \left(\hat{\epsilon}_i + \frac{H_{i,i}\hat{\epsilon}_i}{1-H_{i,i}}\right)^2}{n-p-1} \\
&= \frac{|\hat{\epsilon}|^2 + 2\frac{\hat{\epsilon}_i}{1-H_{i,i}}(H_{*,i})^T\hat{\epsilon} + \left(\frac{\hat{\epsilon}_i}{1-H_{i,i}}\right)^2|H_{*,i}|^2 - \left(\frac{\hat{\epsilon}_i}{1-H_{i,i}}\right)^2}{n-p-1} \\
&= \frac{(n-p)|\hat{\epsilon}|^2}{(n-p)(n-p-1)} + \frac{\left(\frac{\hat{\epsilon}_i}{1-H_{i,i}}\right)^2 H_{i,i} - \left(\frac{\hat{\epsilon}_i}{1-H_{i,i}}\right)^2}{n-p-1} \\
&= \frac{((n-p-1)+1)\hat{\sigma}^2}{n-p-1} - \frac{\frac{\hat{\epsilon}_i^2}{1-H_{i,i}}}{n-p-1} \\
&= \hat{\sigma}^2 + \frac{\hat{\sigma}^2 - \frac{\hat{\epsilon}_i^2}{1-H_{i,i}}}{n-p-1}
\end{aligned}$$

3. (a) Let $\beta_{k+1} = \beta_{k+2} = \cdots = \beta_p = 0$. Then,

$$\begin{aligned}
SS(\beta) &= (y - X\beta)^T(y - X\beta) \\
&= \sum_{i=1}^n (y_i - X_i\beta)^2 \\
&= \sum_{i=1}^n (y_i - X_{i,1}\beta_2 - X_{i,1}\beta_2 - \cdots - X_{i,p}\beta_p)^2 \\
&= \sum_{i=1}^n (y_i - X_{i,1}\beta_1 - X_{i,2}\beta_2 - \cdots - X_{i,k}\beta_k)^2
\end{aligned}$$

If $1 \le j \le k$, then

$$
\begin{aligned}
\frac{\partial SS}{\partial \beta_j} &= \frac{\sum_{i=1}^{n} \partial\left((y_i - X_{i,1}\beta_1 - \cdots - X_{i,k}\beta_k)^2\right)}{\partial \beta_j} \\
&= \sum_{i=1}^{n} \frac{d\left((y_i - X_{i,1}\beta_1 - \cdots - X_{i,k}\beta_k)^2\right)}{d(y_i - X_{i,1}\beta_1 - \cdots - X_{i,k}\beta_k)} \frac{\partial(y_i - X_{i,1}\beta_1 - \cdots - X_{i,k}\beta_k)}{\partial \beta_j} \\
&= \sum_{i=1}^{n} 2(y_i - X_{i,1}\beta_1 - \cdots - X_{i,k}\beta_k)(-X_{i,j}) \\
&= -2\sum_{i=1}^{n} X_{i,j} \left(y_i - \begin{pmatrix} X_{i,1} & \cdots & X_{i,k} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} \right) \\
&= -2(X_{[j]})^T \left(y - X_I \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} \right)
\end{aligned}
$$

$$
\frac{\partial^2 SS}{\partial \beta_j^2} = \sum_{i=1}^{n} 2(-X_{i,j}) \frac{\partial(y_i - X_{i,1}\beta_1 - \cdots - X_{i,k}\beta_k)}{\partial \beta_j} = \sum_{i=1}^{n} 2(-X_{i,j})^2 \ge 0
$$

$$
\frac{\partial SS}{\partial \beta_j}(\beta) = 0 \implies (X_{[j]})^T y = (X_{[j]})^T X_I \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}
$$

$$
\begin{pmatrix} \frac{\partial SS}{\partial \beta_1}(\beta) \\ \vdots \\ \frac{\partial SS}{\partial \beta_k}(\beta) \end{pmatrix} = \mathbf{0}_k \implies X_I^T y = X_I^T X_I \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} \implies \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} = \left(X_I^T X_I\right)^{-1} X_I^T y = \hat{\beta}_I
$$

$$
\nabla SS(\beta) = \begin{pmatrix} \frac{\partial SS}{\partial \beta_1}(\beta) \\ \vdots \\ \frac{\partial SS}{\partial \beta_p}(\beta) \end{pmatrix} = \mathbf{0}_p \implies \beta = \begin{pmatrix} \hat{\beta}_I \\ \mathbf{0}_{p-k} \end{pmatrix} = \hat{\beta}_0
$$

$\therefore \hat{\beta}_0$ minimizes $SS(\beta)$ subject to $\hat{\beta}_{k+1} = \hat{\beta}_{k+2} = \cdots = \hat{\beta}_p = 0$.

(b) For any $j \in \{1, 2, \ldots, k\}$,

$$
\begin{aligned}
X_{[i]} &= Q_{[1]}R_{1,i} + Q_{[2]}R_{2,i} + \cdots + Q_{[i]}R_{i,i} \\
&= Q_{[1]}R_{1,i} + Q_{[2]}R_{2,i} + \cdots + Q_{[k]}R_{k,i} \quad (\because R_{j,i} = 0 \quad \forall j > i) \\
&= \begin{pmatrix} Q_{[1]} & Q_{[2]} & \cdots & Q_{[k]} \end{pmatrix} \begin{pmatrix} R_{1,i} \\ R_{2,i} \\ \vdots \\ R_{k,i} \end{pmatrix} \\
&= Q_I R_{I[i]}
\end{aligned}
$$

$$\therefore X_I = \begin{pmatrix} X_{[1]} & X_{[2]} & \cdots & X_{[k]} \end{pmatrix} = Q_I \begin{pmatrix} R_{I[1]} & R_{I[2]} & \cdots & R_{I[k]} \end{pmatrix} = Q_I R_I$$

$$\begin{aligned}
\hat{\beta}_I &= \left(X_I^T X_I\right)^{-1} X_I^T y \\
&= \left((Q_I R_I)^T Q_I R_I\right)^{-1} (Q_I R_I)^T y \\
&= \left(R_I^T Q_I^T Q_I R_I\right)^{-1} R_I^T Q_I^T y \\
&= \left(R_I^T R_I\right)^{-1} R_I^T Q_I^T y \\
&= R_I^{-1} R_I^{T^{-1}} R_I^T Q_I^T y \\
&= R_I^{-1} Q_I^T y
\end{aligned}$$

$$\hat{y}_I = X_I \hat{\beta}_I = Q_I R_I R_I^{-1} Q_I^T y = Q_I Q_I^T y = H_I y$$

(c)

$$y_1^2 + y_2^2 + \cdots + y_n^2 = y^T y = y^T A^T A y = (Ay)^T A y = z^T z = z_1^2 + z_2^2 + \cdots + z_n^2$$

$$\begin{aligned}
\hat{y}_1^2 + \hat{y}_2^2 + \cdots + \hat{y}_n^2 &= \hat{y}^T \hat{y} \\
&= (Hy)^T Hy \\
&= \left(QQ^T y\right)^T QQ^T y \\
&= y^T QQ^T QQ^T y \\
&= y^T QQ^T y \\
&= \left(Q^T y\right)^T Q^T y \\
&= z_Q^T z_Q \\
&= z_1^2 + z_2^2 + \cdots + z_p^2
\end{aligned}$$

$$\begin{aligned}
\hat{\epsilon}_1^2 + \hat{\epsilon}_2^2 + \cdots + \hat{\epsilon}_n^2 &= \hat{\epsilon}^T \hat{\epsilon} \\
&= (y - \hat{y})^T (y - \hat{y}) \\
&= \left(y - QQ^T y\right)^T \left(y - QQ^T y\right) \\
&= \left(y^T - y^T QQ^T\right) \left(y - QQ^T y\right) \\
&= y^T y - y^T QQ^T y - y^T QQ^T y + y^T QQ^T QQ^T y \\
&= y^T y - 2 y^T QQ^T y + y^T QQ^T y \\
&= y^T y - y^T QQ^T y \\
&= \left(z_1^2 + z_2^2 + \cdots + z_n^2\right) - \left(z_1^2 + z_2^2 + \cdots + z_p^2\right) \\
&= z_{p+1}^2 + z_{p+2}^2 + \cdots + z_n^2
\end{aligned}$$

Let $W_I = \begin{pmatrix} Q_{[k+1]} & Q_{[k+2]} & \cdots & Q_{[p]} & W \end{pmatrix}$. Then,

$$\begin{pmatrix} Q_I^T \\ W_I^T \end{pmatrix} = \begin{pmatrix} Q^T \\ W^T \end{pmatrix} = A$$

$$y_1^2 + y_2^2 + \cdots + y_n^2 = y^T y = y^T A^T A y = (Ay)^T A y = z^T z = z_1^2 + z_2^2 + \cdots + z_n^2$$

$$\begin{aligned}
(\hat{y}_I)_1^2 + (\hat{y}_I)_2^2 + \cdots + (\hat{y}_I)_n^2 &= \hat{y}_I^T \hat{y}_I \\
&= (H_I y)^T H_I y \\
&= \left(Q_I Q_I^T y\right)^T Q_I Q_I^T y \\
&= y^T Q_I Q_I^T Q_I Q_I^T y \\
&= y^T Q_I Q_I^T y \\
&= \left(Q_I^T y\right)^T Q_I^T y \\
&= z_1^2 + z_2^2 + \cdots + z_k^2
\end{aligned}$$

$$\begin{aligned}
(\hat{\epsilon}_I)_1^2 + (\hat{\epsilon}_I)_2^2 + \cdots + (\hat{\epsilon}_I)_n^2 &= \hat{\epsilon}_I^T \hat{\epsilon}_I \\
&= (y - \hat{y}_I)^T (y - \hat{y}_I) \\
&= \left(y - Q_I Q_I^T y\right)^T \left(y - Q_I Q_I^T y\right) \\
&= \left(y^T - y^T Q_I Q_I^T\right)\left(y - Q_I Q_I^T y\right) \\
&= y^T y - y^T Q_I Q_I^T y - y^T Q_I Q_I^T y + y^T Q_I Q_I^T Q_I Q_I^T y \\
&= y^T y - 2 y^T Q_I Q_I^T y + y^T Q_I Q_I^T y \\
&= y^T y - y^T Q_I Q_I^T y \\
&= \left(z_1^2 + z_2^2 + \cdots + z_n^2\right) - \left(z_1^2 + z_2^2 + \cdots + z_k^2\right) \\
&= z_{k+1}^2 + z_{k+2}^2 + \cdots + z_n^2
\end{aligned}$$

(d) Since z has a multivariate normal distribution with a diagonal covariance matrix, $z_{k+1}, z_{k+2}, \cdots, z_n$ are mutually independent and normally distributed, Hence, $z_1, z_2, \ldots, z_n \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$. Hence,

$$SS_{ERROR} = \hat{\epsilon}^T \hat{\epsilon} = z_{p+1} + z_{p+2} + \cdots + z_n \sim \sigma^2 \times \chi_{n-p}^2$$

For the constrained model,

$$SS_{REG(k+1,\ldots,p|1,\ldots,k)} = \hat{y}^T \hat{y} - \hat{y}_I^T \hat{y}_I = \hat{\epsilon}_I^T \hat{\epsilon}_I - \hat{\epsilon}^T \hat{\epsilon} = z_{k+1} + \cdots + z_p \sim \sigma^2 \times \chi_{n-k}^2$$

Since $SS_{REG(k+1,\ldots,p|1,\ldots,k)}$ and SS_{ERROR} do not contain any of the same terms, they are independent. Hence,

$$\hat{F} = \frac{SS_{REG(k+1,\ldots,p|1,\ldots,k)}/(n-k)}{SS_{ERROR}/(n-p)} \sim F_{n-k,n-p}$$