Homework Assignment #7

1. (a) Argue that if a matrix A has rank q then $A(X'X)^{-}A'$ is invertible.

COMMENT: Here are two proposed solutions, followed by a heuristic argument. Since this problem turned out to be a lot harder than anticipated, it was graded based on "effort" rather than correctness.

Solution: Let r be the rank of $X \in \mathbb{R}^{n\times p}$. Without loss of generality, we can assume that $X = (X_1, X_2)$ where $X_1 \in \mathbb{R}^{n\times r}$ consists of r linearly independent columns and $X_2 \in \mathbb{R}^{n\times(p-r)}$. We have a generalized inverse

$$(X'X)^{-} = \begin{bmatrix} (X_1'X_1)^{-1} & 0 \\ 0 & 0 \end{bmatrix}.$$

Let $A = [A_1, A_2] \in \mathbb{R}^{q\times p}$ where $A_1 \in \mathbb{R}^{q\times r}$ and $A_2 \in \mathbb{R}^{q\times(p-r)}$. Then we have

$$\begin{aligned} A(X'X)^{-}A' &= [A_1, A_2]\begin{bmatrix} (X_1'X_1)^{-1} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} A_1' \\ A_2' \end{bmatrix} \\ &= A_1(X_1'X_1)^{-1}A_1'. \end{aligned}$$

Note that $(X_1'X_1)^{-1}$ is positive definite. Note also that if $q \le r$, A_1 is of full rank. In this case, $A_1(X_1'X_1)^{-1}A_1'$ is positive definite and hence $A(X'X)^{-}A'$ is invertible. Since $H : A\beta = 0$ is testable, there exists some matrix M such that $A = MX$. Thus, using the projection matrix P we can view $A(X'X)^{-}A' = (MX)A(X'X)^{-}A'(MX)' = MPM'$. Since P is unique, the choice of $(X'X)^{-}$ above does not affect our argument. Furthermore, the rank of $A = MX$ is q and this is less than or equal to the rank of X so that $q \le r$. This completes the proof.

Another Solution (from Last Year): Since A has rank q then each row of A corresponds to an estimable function of β. Thus $A = MX$ for some $M(q \times n)$ of rank q. Suppose X has rank $r > q$. Then

$$A(X'X)^{-}A = MX(X'X)^{-}X'M' = MPM'$$

where P is a projection matrix or rank r. Also,

$$MPM' = MPP'M' = (MP)(MP)' \Rightarrow \text{rank}(A(X'X)^{-}A) = \text{rank}(MP)$$

Finally,

$$\text{rank}(MP) = \text{rank}(MX(X'X)^{-}X') = \text{rank}(MX) = \text{rank}(A) = q.$$

(I am not sure if the equality holds in $\text{rank}(MX(X'X)^-X') = \text{rank}(MX)$. It seems to me that equality should be replaced with "$\leq$.")

Thus since $A(X'X)^-A$ is a $q \times q$ matrix of full rank, it is invertible.

Heuristically, note $A(X'X) - A'$ is proportional to the covariance matrix for $A\beta$. Since the rows of A are linearly independent, $A(X'X)^-A'$ should be a non-singular covariance matrix.

(b) Show the result on page 8 of Lecture 13 that if A has rank q then

$$E[RSS_H - RSS] = \sigma^2 q + (A\beta)'[A(X'X)^-A']^{-1}(A\beta)$$

Solution: From theorem 13.2.1, we know

$$RSS_H - RSS = (A\hat{\beta})'[A(X'X)^-A']^{-1}(A\hat{\beta})$$

and noting that

$$\begin{aligned} E[A\hat{\beta}] &= A\beta \equiv \mu \\ Var[A\hat{\beta}] &= AVar[\hat{\beta}]A' = \sigma^2 A(X'X)^-A \equiv \Sigma \end{aligned}$$

we can apply the results about expectations of quadratic forms from Lecture 3:

$$\begin{aligned} E[(A\hat{\beta})'[A(X'X)^-A']^{-1}(A\hat{\beta})] &= \text{tr}([A(X'X)^-A']^{-1}\Sigma) + \mu'[A(X'X)^-A']^{-1}\mu \\ &= \text{tr}(\sigma^2[A(X'X)^-A']^{-1}A(X'X)^-A) + (A\beta)'[A(X'X)^-A']^{-1}(A\beta) \\ &= \sigma^2\text{tr}(I_{(q\times q)}) + (A\beta)'[A(X'X)^-A']^{-1}(A\beta) \\ &= \sigma^2 q + (A\beta)'[A(X'X)^-A']^{-1}(A\beta) \end{aligned}$$

2. Let

$$\begin{aligned} Y_1 &= \theta_1 + \theta_2 + \epsilon_1 \\ Y_2 &= 2\theta_2 + \epsilon_2 \\ Y_3 &= -\theta_1 + \theta_2 + \epsilon_3 \end{aligned}$$

where $\epsilon_i \sim_{iid} N(0, \theta^2), i = 1, 2, 3$. Derive an F-statistic for testing the hypothesis $H : \theta_1 = 2\theta_2$.

Solution: The model can be written in matrix form as:

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix} \equiv Y = X\beta + \epsilon$$

and thus the hypothesis H is equivalent to $H: a'\beta = 0$ where $a = (1 \;\; -2)'$. The least squares estimate of β is

$$(X'X)^{-1}X'Y = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/6 \end{pmatrix} \begin{pmatrix} Y_1 - Y_3 \\ Y_1 + 2Y_2 + Y_3 \end{pmatrix} = \begin{pmatrix} \frac{Y_1 - Y_3}{2} \\ \frac{Y_1 + 2Y_2 + Y_3}{6} \end{pmatrix}$$

and hence

$$\hat{Y} = \begin{pmatrix} \frac{4Y_1 + 2Y_2 - 2Y_3}{6} \\ \frac{Y_1 + 2Y_2 + Y_3}{3} \\ \frac{-2Y_1 + 2Y_2 + 4Y_3}{6} \end{pmatrix}$$

Thus the unrestricted sum of squares is

$$\begin{aligned} RSS = (Y - \hat{Y})'(Y - \hat{Y}) &= \frac{1}{36}(2Y_1 - 2Y_2 + 2Y_3)^2 + \frac{1}{9}(-Y_1 + Y_2 - Y_3)^2 + \frac{1}{36}(2Y_1 - 2Y_2 + 2Y_3)^2 \\ &= \frac{1}{9}(Y_1 - Y_2 + Y_3)^2 + \frac{(-1)^2}{9}(Y_1 - Y_2 + Y_3)^2 + \frac{1}{9}(Y_1 - Y_2 + Y_3)^2 \\ &= \frac{1}{3}(Y_1 - Y_2 + Y_3)^2. \end{aligned}$$

Now we compute the difference of the sum of squares

$$\begin{aligned} RSS_H - RSS &= (a'\hat{\beta})'[a(X'X)^{-1}a']^{-1}(a'\hat{\beta}) \\ &= (\hat{\beta}_1 - 2\hat{\beta}_2)\left(\frac{6}{7}\right)(\hat{\beta}_1 - 2\hat{\beta}_2) \\ &= \left(\frac{6}{7}\right)(\hat{\beta}_1 - 2\hat{\beta}_2)^2. \end{aligned}$$

Therefore the F-statistic is

$$\frac{RSS_H - RSS/1}{RSS/(3-2)} = \frac{\left(\frac{6}{7}\right)(\hat{\beta}_1 - 2\hat{\beta}_2)^2}{\frac{1}{3}(Y_1 - Y_2 + Y_3)^2}$$

Note that there is a typo in the solution given in Seber and Lee. They claim that

$$RSS = S^2 = Y_1^2 + Y_2^2 + Y_3^2 = 2\hat{\beta}_1^2 = 6\hat{\beta}_2^2$$

The "=" signs above should be "−" signs. Then the solutions are equivalent, which can be seen after some simplification:

$$
\begin{aligned}
Y_1^2+Y_2^2+Y_3^2-2\hat\beta_1^2-6\hat\beta_2^2 &= Y_1^2+Y_2^2+Y_3^2+\frac{-Y_1^2-Y_3^2+2Y_1Y_3}{2} \\
&\quad +\frac{-Y_1^2-4Y_2^2-Y_3^2-4Y_1Y_2-2Y_1Y_3-4Y_2Y_3}{6} \\
&= \frac{1}{6}[6Y_1^2+6Y_2^2+6Y_3^2-3Y_1^2-3Y_3^2+6Y_1Y_3 \\
&\quad -Y_1^2-4Y_2^2-Y_3^2-4Y_1Y_2-2Y_1Y_3-4Y_2Y_3] \\
&= \frac{1}{6}[2Y_1^2+2Y_2^2+2Y_3^2-4Y_1Y_2+4Y_1Y_3-4Y_2Y_3] \\
&= \frac{1}{3}(Y_1-Y_2+Y_3)^2
\end{aligned}
$$

3. Given the two regression lines

$$Y_{ki}=\beta_k x_i+\epsilon_{ki},\ k=1,2;\ i=1,\ldots,n$$

show that the F-statistic for testing $H:\beta_1=\beta_2$ can be put in the form

$$F=\frac{(\hat\beta_1-\hat\beta_2)^2}{2S^2(\sum_i x_i^2)^{-1}}$$

Obtain RSS and RSS_H and verify that

$$RSS_H-RSS=\frac{1}{2}\sum_i x_i^2(\hat\beta_1-\hat\beta_2)^2.$$

Solution: The model can be written in matrix form as

$$Y_{(2n\times 1)}=\begin{pmatrix} x_1 & 0 \\ \vdots & \vdots \\ x_n & 0 \\ 0 & x_1 \\ \vdots & \vdots \\ 0 & x_n \end{pmatrix}\begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix}+\epsilon\equiv X\beta+\epsilon$$

and the hypothesis is equivalent to $H:a\beta=0$ where $a=(1\ -1)$.

$$(X'X)^{-1}=\begin{pmatrix}\sum x_i^2 & 0\\ 0 & \sum x_i^2\end{pmatrix}^{-1}=\begin{pmatrix}\frac{1}{\sum x_i^2} & 0\\ 0 & \frac{1}{\sum x_i^2}\end{pmatrix}$$

thus

$$\hat{\beta} = (X'X)^{-1}X'Y = \begin{pmatrix} \frac{1}{\sum x_i^2} & 0 \\ 0 & \frac{1}{\sum x_i^2} \end{pmatrix} \begin{pmatrix} \sum_i x_i y_{1i} \\ \sum_i x_i y_{2i} \end{pmatrix} = \begin{pmatrix} \frac{\sum_i x_i y_{1i}}{\sum x_i^2} \\ \frac{\sum_i x_i y_{2i}}{\sum x_i^2} \end{pmatrix}$$

and therefore the unrestricted sum of squares is

$$\begin{aligned} RSS &= (Y - X\hat{\beta})'(Y - X\hat{\beta}) = \sum_i [(y_{1i} - x_i\hat{\beta}_1)^2 + (y_{2i} - x_i\hat{\beta}_2)^2] \\ &= \sum_i [y_{1i}^2 + y_{2i}^2 - 2y_{1i}x_i\hat{\beta}_1 - 2y_{2i}x_i\hat{\beta}_2 + x_i^2(\hat{\beta}_1^2 + \hat{\beta}_2^2)] \end{aligned}$$

Now under H, the model is

$$Y = \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \beta_{H(1\times 1)} + \epsilon$$

and the least squares estimate is

$$\hat{\beta_H} = \frac{\sum_i x_i(y_{1i} + y_{2i})}{2\sum_i x_i^2} = \frac{\hat{\beta}_1 + \hat{\beta}_2}{2}$$

Therefore the restricted sum of squares is

$$\begin{aligned} RSS_H &= \sum_i [(y_{1i} - x_i(\hat{\beta}_1 + \hat{\beta}_2)/2)^2 + (y_{2i} - x_i(\hat{\beta}_1 + \hat{\beta}_2)/2)^2] \\ &= \sum_i [y_{i1}^2 + y_{i2}^2 - y_{1i}x_i(\hat{\beta}_1 + \hat{\beta}_2) - y_{2i}x_i(\hat{\beta}_1 + \hat{\beta}_2) + x_i^2(\hat{\beta}_1 + \hat{\beta}_2)^2/2] \end{aligned}$$

Hence,

$$\begin{aligned} RSS - RSS_H &= \sum_i \left\{ x_i^2(\hat{\beta}_1^2 + \hat{\beta}_2^2) - 2y_{1i}x_i\hat{\beta}_1 - 2y_{2i}x_i\hat{\beta}_2 + y_{1i}x_i(\hat{\beta}_1 + \hat{\beta}_2) + y_{2i}x_i(\hat{\beta}_1 + \hat{\beta}_2) - x_i^2(\hat{\beta}_1 + \hat{\beta}_2)^2/2 \right\} \\ &= \sum_i \left\{ x_i^2\hat{\beta}_1^2 + x_i^2\hat{\beta}_2^2 - x_i^2\hat{\beta}_1^2/2 - x_i^2\hat{\beta}_2^2/2 - x_i^2\hat{\beta}_1\hat{\beta}_2 \right\} \\ &= \frac{1}{2}\sum_i \left\{ x_i^2[\hat{\beta}_1^2 + \hat{\beta}_2^2 - 2\hat{\beta}_1\hat{\beta}_2] \right\} = \frac{1}{2}\sum_i x_i^2(\hat{\beta}_1 - \hat{\beta}_2)^2 \end{aligned}$$

So combining these results we have

$$F = \frac{RSS_H - RSS}{RSS/(n-1)} = \frac{(\hat{\beta}_1 - \hat{\beta}_2)^2}{2S^2(\sum_i x_i^2)^{-1}}$$

Note that here $S^2 = RSS/(n-p)$, *not* the sample variance.

4. A series of $n+1$ observaions $Y_i (i = 1, 2, \ldots, n+1)$ are taken from a normal distribution with unknown variance σ^2. After the first n observations it is suspected that there is a sudden change in the mean of the distribution. Derive a test statistic for testing the hypothesis that the $(n+1)^{th}$ observation has the same mean as the previous observations.

Solution: The linear model is

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_n \\ Y_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \vdots & \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{n+1} \end{bmatrix}$$

or $Y = X\beta + \epsilon$. Let $A = [1, -1]$. Note that $H : A\beta = 0$ is testable and the rank of A is 1. The least squares estimate of β is $\hat{\beta} = (\frac{1}{n}\sum_{i=1}^n Y_i, Y_{n+1}) \equiv (\overline{Y_n}, Y_{n+1})$. Also,

$$\begin{aligned} RSS_H - RSS &= (A\hat{\beta})'(A(X'X)^{-1}A')^{-1})(A\hat{\beta}) \\ &= (\overline{Y_n} - Y_{n+1})(A\text{diag}(n^{-1}, 1)A')^{-1}(\overline{Y_n} - Y_{n+1}) \\ &= (\overline{Y_n} - Y_{n+1})(1 + 1/n)^{-1}(\overline{Y_n} - Y_{n+1}) \\ &= \frac{n}{n+1}(\overline{Y_n} - Y_{n+1})^2 \end{aligned}$$

and

$$\begin{aligned} RSS &= Y'Y - \hat{\beta}X'X\hat{\beta} \\ &= \sum_{i=1}^{n+1} Y_i - \hat{\beta}\text{diag}(n, 1)\hat{\beta} \\ &= \sum_{i=1}^{n+1} Y_i - n\overline{Y_n}^2 - Y_{n+1}^2 \\ &= \sum_{i=1}^{n} Y_i - n\overline{Y_n}^2 \\ &= \sum_{i=1}^{n} (Y_i - \overline{Y_n})^2 \\ &\equiv (n-1)S_n^2. \end{aligned}$$

Thus the test statistic is

$$F = \frac{(RSS_H - RSS)/1}{RSS/(n+1-2)} = \frac{(\overline{Y_n} - Y_{n+1})^2}{S_n^2(1+1/n)},$$

which is distributed as F distribution with degrees of freedom 1 and $n-1$ under the null hypothesis.

5. Prove the result on the page 6 of Lecture 14 that $\frac{1}{\sigma^2}(RSS_H - RSS)$ has non-central chisquared distribution with noncentrality parameter $\lambda = \frac{1}{\sigma^2}\boldsymbol{\mu}'(P_\Omega - P_\omega)\boldsymbol{\mu}$.

Solution: Note that $RSS_H - RSS = Y'(P_\Omega - P_\omega)Y$ and that $P_\Omega - P_\omega = P_{\omega^\perp \cup \Omega}$ is the projection matrix by the lemma on page 8 of the Lecture note 13. Since $Y \sim N(\mu, \sigma^2 I)$, the claim follows by the Theorem on page 6 of the Lecture note 6.