CAS MA575: Assignment #6

Hu Hengchang(U34497427)
hhc@bu.edu

Boston University — November 4, 2018

1.

$X \in \mathbf{R}^{n\times p}$ is a deterministic design matrix and rank$(X) = p$. $(XX^T)^T = XX^T$, so XX^T is a n by n real symmetric matrix.

It is obvious that rank(XX^T) = rank(X) = p, so the number of non-zero eigenvalues of matrix XX^T equals to rank(XX^T) which is p. Assume they are $\lambda_1, \lambda_2, ..., \lambda_p$.For symmetric matrix XX^T, we know that there is a set of orthonormal eigenvectors of XX^T *i.e.* $\vec{q_1}, \vec{q_2}, ..., \vec{q_n}$ and $A\vec{q_i} = \lambda_i \vec{q_i}$ for $i = 1, 2, ..., n$.

Therefore $QXX^TQ^T = diag(\lambda_1, \lambda_2, ..., \lambda_p, 0..., 0)$, Q is a matrix as $\begin{pmatrix} \vec{q_1}^T \\ \vec{q_2}^T \\ . \\ . \\ . \\ \vec{q_n}^T \end{pmatrix}$.

In particular, $X = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 2 & 0 \end{pmatrix}$. $XX^T = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$. There are two non-zero eigenvalues of XX^T, $\lambda_1 = 1, \lambda_2 = 4$. From these two eigenvalues, we can figure out that $\vec{q_1}^T = (0\ 1\ 0)$, $\vec{q_2}^T = (0\ 0\ 1)$, since eigenvalue $\lambda_3 = 0$, we set $\vec{q_3}^T = (1\ 0\ 0)$ so that $Q = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, and it is easy to prove that $QXX^TQ^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

2.

By Problem 1., we can know that $\exists$ orthogonal matrix Q, s.t $QXX^TQ^T = diag(\lambda_1, ..., \lambda_p, 0, ..., 0)$.

So we can get that: $\hat{\beta} = (X^TX)^{-1}X^TY = [(QX)^TQX]^{-1}(QX)^T(QY)$, we set $QX = A$, and $AA^T = diag(\lambda_1, ..., \lambda_p, 0, ..., 0)$, assume $A = \begin{pmatrix} \vec{a_1}^T \\ \vec{a_2}^T \\ \cdot \\ \cdot \\ \vec{a_n}^T \end{pmatrix} \Rightarrow$

$AA^T = (\vec{a_i}^T \vec{a_j})_{ij} \Rightarrow \vec{a_i}^T \vec{a_i} = 0$ if $i > p$, so $\vec{a_i} = 0$, if $i > p$. Matrix A has two parts, $A = \begin{pmatrix} A_p \\ 0 \end{pmatrix}$, A_p is a p by p full rank matrix, so $A^T A = A_p^T A_p$. By the way we can rewrite $\hat{\beta}$ as $\hat{\beta} = (A^T A)^{-1} A^T QY$.

$$\begin{aligned}
(n-p)\hat{\sigma}^2 &= (Y - X\hat{\beta})^T (Y - X\hat{\beta}) \\
&= (QY - QX\hat{\beta})^T (QY - QX\hat{\beta}) \\
&= \Big(Q(X\beta + e) - A(A^T A)^{-1} A^T Q(X\beta + e)\Big)^T \Big(Q(X\beta + e) - A(A^T A)^{-1} A^T Q(X\beta + e)\Big) \\
&= \Big((Qe - A(A^T A)^{-1} A^T Qe)^T (Qe - A(A^T A)^{-1} A^T Qe)\Big)
\end{aligned}$$

Assume $Qe = \begin{pmatrix} Qe_{(p)} \\ Qe_{(n-p)} \end{pmatrix}$, $A(A^T A)^{-1} A^T Qe = \begin{pmatrix} A_p \\ 0 \end{pmatrix} (A_p^T A_p)^{-1} \begin{pmatrix} A_p^T & 0 \end{pmatrix} \begin{pmatrix} Qe_{(p)} \\ Qe_{(n-p)} \end{pmatrix} = \begin{pmatrix} Qe_{(p)} \\ 0 \end{pmatrix}$, $\Rightarrow (n-p)\hat{\sigma}^2 = \begin{pmatrix} 0 \\ Qe_{(n-p)} \end{pmatrix}^T \begin{pmatrix} 0 \\ Qe_{(n-p)} \end{pmatrix} = (Qe_{(n-p)})^T (Qe_{(n-p)})$

Because Q is orthogonal matrix, so $\sigma^{-1} Qe$ is still standard multivariate normal distribution. Therefore, $\frac{(n-p)\hat{\sigma}^2}{\sigma^2} = (\sigma^{-1} Qe_{(n-p)})^T (\sigma^{-1} Qe_{(n-p)})$ follows a χ^2_{n-p}.

3.

Because $X^T X = \begin{pmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{pmatrix} \Rightarrow (X^T X)^{-1} = \frac{1}{nS_{XX}} \begin{pmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{pmatrix}$. And $X^T Y = \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{pmatrix}$.

$$\begin{aligned}
\frac{1}{nS_{XX}} \Big(\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i \Big) &= \frac{1}{nS_{XX}} \Big(\big(\sum (x_i - \bar{x}^2) + n\bar{x}^2 \big) \sum y_i - n\bar{x} \sum x_i y_i \Big) \\
&= \frac{1}{nS_{XX}} \Big(S_{XX} \sum y_i - (n\bar{x} \sum x_i y_i - n\bar{x}^2 \sum y_i) \Big) \\
&= \sum \Big(\frac{1}{n} - \frac{x_i - \bar{x}}{S_{XX}} \bar{x} \Big) y_i
\end{aligned}$$

$$
\begin{aligned}
\frac{1}{nS_{XX}}\left(n\sum x_i y_i - \sum x_i \sum y_i\right) &= \frac{1}{S_{XX}}\left(\sum x_i y_i - \bar{x}\sum y_i\right) \\
&= \sum\left(\frac{1}{S_{XX}}(x_i-\bar{x})y_i\right) \\
&= \sum\left(\frac{x_i-\bar{x}}{S_{XX}}(y_i-\bar{y})\right)
\end{aligned}
$$

Therefore, $(X^TX)^{-1}X^TY = \frac{1}{nS_{XX}}\begin{pmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{pmatrix}\begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{pmatrix}$
$= \begin{pmatrix} \sum\left(\frac{1}{n} - \frac{x_i-\bar{x}}{S_{XX}}\bar{x}\right)y_i \\ \sum\left(\frac{x_i-\bar{x}}{S_{XX}}(y_i-\bar{y})\right) \end{pmatrix}$.

It is obvious that $\sigma^2(X^TX)^{-1} = \frac{\sigma^2}{nS_{XX}}\begin{pmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{pmatrix} = \frac{\sigma^2}{nS_{XX}}\begin{pmatrix} S_{XX}+n\bar{x}^2 & -n\bar{x} \\ -n\bar{x} & n \end{pmatrix} =$
$\begin{pmatrix} \left(\frac{1}{n}+\frac{(\bar{x})^2}{S_{XX}}\right)\sigma^2 & -\frac{\bar{x}}{S_{XX}}\sigma^2 \\ -\frac{\bar{x}}{S_{XX}}\sigma^2 & \frac{1}{S_{XX}}\sigma^2 \end{pmatrix}$