L13: Estimation when X does not have full column rank

1. $\frac{SSE}{\sigma^2} \sim \chi^2(n-r)$ and $AX^+y \sim N(A\beta,\, \sigma^2 A(X'X)^+A')$ are independent.

   $y = X\beta + e$, $e \sim N(0,\, \sigma^2 I_n)$, $X \in R^{n\times p}$ with rank$(X) = r$.

   (1) $\frac{\text{SSE}}{\sigma^2} \sim \chi^2(n-r)$.
   $\widehat{\beta}$ is a LSE for $\beta \iff \|y - X\widehat{\beta}\|^2 \leq \|y - X\beta\|^2$ for all $\beta \iff \widehat{\beta} \in X^+y + \mathcal{N}(X)$.
   With $\widehat{\beta}$, LSE of β, SSE $= \|y - X\widehat{\beta}\|^2 = y'(I - XX^+)y$. Then $\frac{\text{SSE}}{\sigma^2} \sim \chi^2(n-r)$.
   Proof $\frac{\text{SSE}}{\sigma} = y' \frac{I-XX^+}{\sigma^2} y$. But $\frac{I-XX^+}{\sigma^2}\left(\sigma^2 I_n\right)\frac{I-XX^+}{\sigma^2} = \frac{I-XX^+}{\sigma^2}$,
   $(X\beta)'\frac{I-XX^+}{\sigma^2}(X\beta) = 0$ and tr $\left[\frac{I-XX^+}{\sigma^2}\left(\sigma^2 I_n\right)\right] = n - r$. So $\frac{\text{SSE}}{\sigma^2} \sim \chi^2(n-r)$.

   (2) $\theta = A\beta$ is estimable.
   Then the BLUE for $A\beta$, $\widehat{\theta} = A[\text{LSE}(\beta)] = AX^+y \sim N\left(\theta,\, \sigma^2 A(X'X)^+A'\right)$.
   Proof $A\beta$ is estimable$\iff A = LX$ for some L.
   $$\begin{aligned} A[\text{LSE}(\beta)] &= AX^+y + LX\mathcal{N}(X) = AX^+y + 0 = AX^+y \\ &\sim N\left(AX^+X\beta,\, \sigma^2(AX^+)(AX^+)'\right) = N(A\beta,\, \sigma^2 A(X'X)^+A'). \end{aligned}$$

   (3) SSE and X^+y are independent
   Proof $(X^+)\left(\sigma^2 I_n\right)(I - XX^+) = 0$.

   Ex1: $MSE = \frac{SSE}{n-r}$ is an UE for σ^2.
   $$E\left(\frac{SSE}{n-r}\right) = E\left(\frac{SSE}{\sigma^2}\,\frac{\sigma^2}{n-r}\right) = E[\chi^2(n-r)]\,\frac{\sigma^2}{n-r} = (n-r)\,\frac{\sigma^2}{n-r} = \sigma^2.$$

2. t-intervals

   (1) With estimable $\theta = A\beta$, let $l'\theta = l'A\beta$, $l'\widehat{\theta} = l'AX^+y$ and
   $s_{l'\theta} = \sqrt{\text{MSE}\; l'A(X'X)^+A'l}$. Then $\frac{l'\theta - l'\theta}{s_{l'\widehat{\theta}}} \sim t(n-r)$.
   Proof $l'\widehat{\theta} \sim N(l'\theta,\, \sigma^2_{l'\widehat{\theta}})$ where $\sigma^2_{l'\widehat{\theta}} = \sigma^2 l'A(X'X)^+A'l$ is estimated by
   $s^2_{l'\widehat{\theta}} = \text{MSE}\; l'A(X'X)^+A'l$.
   $\frac{l'\widehat{\theta} - l'\theta}{\sqrt{\sigma^2_{l'\widehat{\theta}}}} \sim N(0, 1^2)$ is independent to $\frac{SSE}{\sigma^2} \sim \chi^2(n-r)$.
   So $\frac{l'\widehat{\theta} - l'\theta}{\sqrt{\sigma^2_{l'\widehat{\theta}}\text{SSE}/[(n-r)\sigma^2]}} \sim t(n-r)$, i.e., $\frac{l'\theta - l'\theta}{s_{l'\widehat{\theta}}} \sim t(n-r)$.

   (2) t-interval for $l'\theta = l'A\beta$
   From the random variable in (1) with t-distribution, it can be shown that
   $l'\widehat{\theta} \pm t_{\alpha/2}(n-r)s_{l'\widehat{\theta}} = l'A\beta \pm t_{\alpha/2}(n-r)s_{l'\widehat{\theta}}$ is a $1-\alpha$ CI for $l'\theta = l'A\beta$.
   Comments: $s^2_{l'\widehat{\theta}} = \text{MSE}\; l'A(X'X)^+Al$.
   When $p = r$ and $A = I_p$, the CI becomes
   $l'X^+y \pm t_{\alpha/2}(n-p)s_{l'X^+y}$ is a $1-\alpha$ CI for $l'\beta$ where $s^2_{l'X^+y} = \text{MSE}\; l'(X'X)^{-1}l$.

   (3) Bonferroni intervals
   $l_i'AX^+y \pm t_{\alpha/(2k)}(n-r)s_{l_i'AX^+y}$, $i = 1, ..., k$, are k simultaneous CIs for $l_i'A\beta$, $i = 1, ..., k$,
   with overall confidence coefficient $1-\alpha$.

3. Confidence region and Schefee's intervals

(1) With estimable $\theta = A\beta$ and $\widehat{\theta} = AX^{+}y$, $\frac{(\widehat{\theta}-\theta)'[A(X'X)^{+}A']^{-1}(\widehat{\theta}-\theta)}{q\text{MSE}} \sim F(q,\, n-r)$.

Proof $\widehat{\theta} \sim N(\theta,\, \sigma^2 A(X'X)^{+}A')$. So $[\sigma^2 A(X'X)^{+}A']^{-1/2}(\widehat{\theta}-\theta) \sim N(0,\, I_q)$. Thus $(\widehat{\theta}-\theta)'[A(X'X)^{+}A']^{-1}(\widehat{\theta}-\theta)/\sigma^2 \sim \chi^2(q)$ and it is independent to $\frac{SSE}{\sigma^2} \sim \chi^2(n-r)$. Therefore $\frac{(\widehat{\theta}-\theta)'[A(X'X)^{+}A']^{-1}(\widehat{\theta}-\theta)}{q\,\text{MSE}} \sim F(q,\, n-r)$.

(2) $1-\alpha$ CR for $\theta = A\beta$.

$$\left[\theta \in R^q : \frac{(\widehat{\theta}-\theta)'[A(X'X)^{+}A']^{-1}(\widehat{\theta}-\theta)}{q\,\text{MSE}} \le F_\alpha(q,\, n-r)\right]$$

is a $1-\alpha$ CR for $\theta = A\beta \in R^q$.

Proof Based on the distribution of the random variable in (1),

$$P\left(\frac{(\widehat{\theta}-\theta)'[A(X'X)^{+}A']^{-1}(\widehat{\theta}-\theta)}{q\,\text{MSE}} \le F_\alpha(q,\, n-r)\right) = 1-\alpha.$$

Comment: The CR is an ellipsoid in R^q with center $\widehat{\theta} = AX^{+}y$.
When X has full column rank, $r = p$. With $A = I_p$ and $\widehat{\theta} = X^{+}y$, we obtain $1-\alpha$ CR for $\theta = \beta \in R^p$

$$\frac{(X^{+}y-\beta)'(X'X)(X^{+}y-\beta)}{p\,\text{MSE}} \le F_\alpha(p,\, n-p).$$

(3) Schefee's intervals
With estimable $A\beta$, $l_i'AX^{+}y \pm \sqrt{qF_\alpha(q,\, n-r)}\, s_{l_i'AX^{+}y}$, $i = 1, 2, ...$, are simultaneous CIs for $l_i'A\beta$, $i = 1,\, 2, ...$ with overall confidence coefficient $1-\alpha$.
Here $s^2_{l_i'AX^{+}y} = \text{MSE}\ \ l_i'A(X'X)^{+}A'l_i$

Proof Extended Cauchy-Schwartz inequatlity $(u'v)^2 \le (u'B^{-1}u)(v'Bv)$ where $B > 0$ implies that $u'B^{-1}u \le c \Longrightarrow -\sqrt{c\,(v'Bv)} \le u'v \le \sqrt{c\,(v'Bv)}$.
Let $u = AX^{+}y - A\beta$, $v = l_i$, $c = qF_\alpha(q,\, n-r)$ and $B = \text{MSE}\ A(X'X)^{+}A'$. Then

$$\begin{aligned} E_0 &= [u'B^{-1}u \le c] \\ &= \left[(AX^{+}y - A\beta)'\left(\text{MSE}\ A(X'X)^{+}A'\right)^{-1}(AX^{+}y - A\beta) \le qF_\alpha(q,\, n-r)\right] \\ &\subset E_i = [-\sqrt{c,\ (v'Bv)} \le u'v \le \sqrt{c\ (v'Bv)}] \\ &= [l_i'A\beta \in l_i'AX^{+}y \pm \sqrt{qF_\alpha(q,\, n-r)}\, s_{l_i'AX^{+}y}]. \end{aligned}$$

Thus $E_0 \subset \cap E_i \Longrightarrow 1-\alpha = P(E_0) \le P(\cap_i E_i)$.