Lecture 14 Oct 25

Q orthogonality of the residuals and the column space $C(X)$

Model: $\underline{Y} = X\beta + \underline{\varepsilon}$

where $\underline{\varepsilon} \sim N(\underline{0}, I \otimes \sigma^2)$

$\underline{\mu} = E\underline{Y} = E(X\underline{\beta} + \underline{\varepsilon}) = X\underline{\beta}$.

Estimate $\underline{\beta}$ or $\underline{\mu}$

$\underline{\mu} = X\underline{\beta}$

$\underline{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} \qquad X = [\underline{X}_1 | \dots | \underline{X}_p]$

$\underline{\mu} = \underline{X}_1 \beta_1 + \underline{X}_2 \beta_2 + \dots + \underline{X}_p \beta_p$

Define $C(X)$ = space spanned by the columns of X

In particular, $\underline{\mu} \in C(X)$.

A candidate estimate of $\underline{\mu}$ must also belong to $C(X)$,

$$\tilde{\underline{\mu}}(\underline{b}) = \underline{X}_1 b_1 + \cdots + \underline{X}_p b_p$$
$$= X\underline{b}$$

The squared error between $\underline{Y}$ and $\tilde{\underline{\mu}}(\underline{b})$ is:

$$\|\underline{Q}\|_2 = \sum_{i=1}^{N} q_i^2$$

$$S(\underline{b}) = \| \underline{Y} - X\underline{b} \|_2$$

$$= (\underline{Y} - X\underline{b})^T (\underline{Y} - X\underline{b})$$
$$= (\underline{Y} - (\underline{X}_1 b_1 + \cdots + \underline{X}_p b_p))^T (\qquad)$$

The LSE of $\underline{\mu}$ (or $\underline{\beta}$):

$\hat{\underline{\mu}} = X\hat{\underline{\beta}}$ satisfies:

$$\hat{\underline{\beta}} = \arg\min_{\underline{b}} S(\underline{b})$$

$$\frac{\partial S(\underline{b})}{\partial b_1} = -2 \ X_1^T (\underbrace{\underline{Y} - X\underline{b}})$$

$$\vdots$$

$$\frac{\partial S(\underline{b})}{\partial b_p} = -2 \cdot X_p^T (\underline{Y} - X\underline{b})$$

Summary $\frac{\partial S(\underline{b})}{\partial \underline{b}} = -2 \ X^T (\underline{Y} - X\underline{b})$

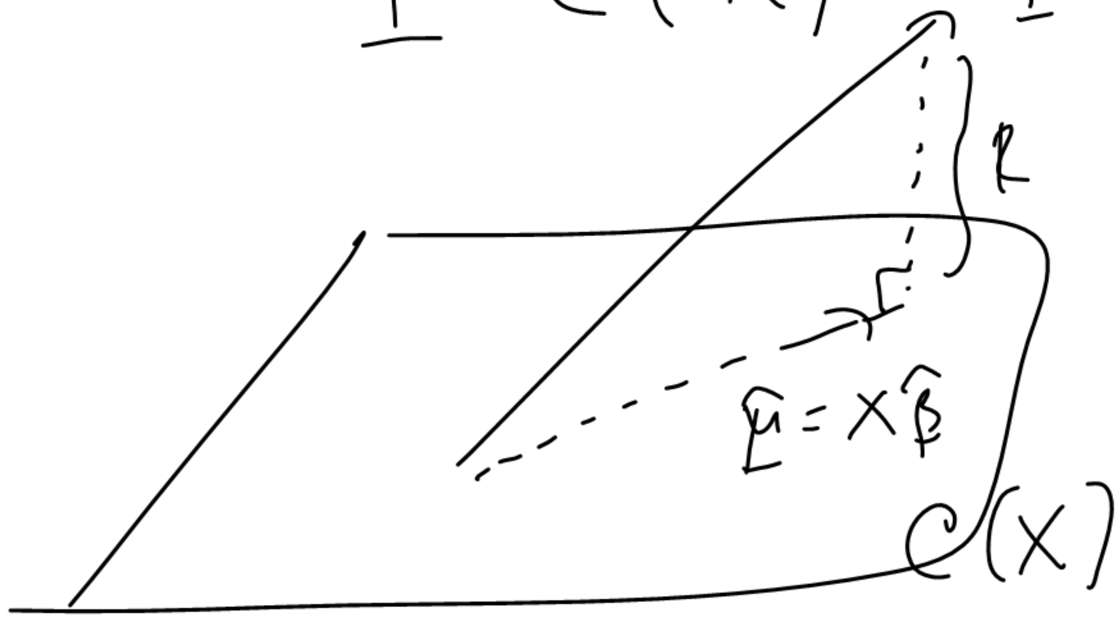
$$\left.\frac{\partial S(\underline{b})}{\partial \underline{b}}\right|_{\hat{\beta}} = \underline{0}$$

$$\Rightarrow \left.\begin{aligned} X_1^T (\underline{Y} - X\hat{\underline{\beta}}) &= 0 \\ &\vdots \\ X_p^T (\underline{Y} - X\hat{\underline{\beta}}) &= 0 \end{aligned}\right\}$$

$$\underline{Y} - X\hat{\underline{\beta}} = \underline{R}$$

$$\Rightarrow \underline{R} \perp \{\underline{X}_1, \dots, \underline{X}_p\}$$

$$\perp \mathcal{C}(X)$$



$$\sum R_i = \sum R_i (1)$$

$$= \langle \underline{R}, \underline{\mathbb{1}} \rangle$$

$\therefore \sum R_i = 0$ if $\underline{\mathbb{1}}$ is contained in X

Q2 $\hat{\beta}_1$

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

$$\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

The LSE of β_1:

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x})(Y_i - \bar{Y})}{\sum (x_i - \bar{x})(x_i - \bar{x})} = \frac{S_{xY}}{S_{xx}}$$

$$= \sum w_i Y_i$$

where $w_i = \dfrac{(x_i - \bar{x})}{\sum_{j=1}^{n} (x_j - \bar{x})^2}$

Q.

Pop 1. $N(\mu_1, \sigma^2)$ $\quad Y^1_1 \dots Y^1_n$ iid $N(\mu_1, \sigma^2)$

Pop 2. $N(\mu_2, \sigma^2)$ $\quad Y^2_1 \dots Y^2_n$ iid $N(\mu_2, \sigma^2)$

$$\delta = \mu_2 - \mu_1$$

Test H_0: $\delta = 0 \iff \mu_2 = \mu_1$

Full Model

DATA: $\{ (Y_i, x_i),\ i = 1, \dots, n, \dots, 2n \}$

- $x_i = \begin{cases} 1, & \text{if } i^{th} \text{ group } 1 \\ 2, & i^{th} \quad 2 \end{cases}$

- $G_{2i} = \begin{cases} 1, & \text{if } i^{th} \text{ group } 2 \\ 0, & \text{o/w} \end{cases}$

$\mu(x_i)$

$$\mu(G_{2i}) = \mu_1 + \delta . G_{2i}$$

if $i^{\underline{th}} \in Gp1 \Rightarrow G_{2i} = 0$

$$\Rightarrow \mu(G_{2i}) = \mu_1$$

$i^{\underline{th}} \in Gp2 \Rightarrow G_{2i} = 1$

$$\mu(G_{2i}) = \underbrace{\mu_1 + \delta}_{\mu_2}$$

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \\ \hline Y_{n+1} \\ \vdots \\ Y_{2n} \end{pmatrix} = \left[\begin{array}{c:c} \underline{\mathbb{1}} & \underline{0} \\ \hline \underline{\mathbb{1}} & \underline{\mathbb{1}} \end{array}\right] \begin{pmatrix} \mu_1 \\ \delta \end{pmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{2n} \end{bmatrix}$$

$$\underline{Y} = X \underline{\beta} + \underline{\varepsilon}$$

$$\hat{\underline{\beta}} = (X'X)^{-1} X' \underline{Y}$$

$$\hat{\beta} = \begin{pmatrix} \hat{\mu}_1 \\ \hat{\delta} \end{pmatrix} = \begin{pmatrix} \bar{Y}^1 \\ \bar{Y}^2 - \bar{Y}^1 \end{pmatrix}$$ CONFIRM

$$\hat{\underline{Y}} = X \hat{\beta} = \left(\begin{array}{c} \underline{\mathbb{1}} \cdot \bar{Y}^1 \\ \hline \underline{\mathbb{1}} \, \bar{Y}^2 \end{array} \right)$$ CONFIRM

$$SSE(1) = \| \underline{Y} - \hat{\underline{Y}} \|_2$$

$$= \sum_{i=1}^{n} (Y_i - \bar{Y}^1)^2 + \sum_{i=n+1}^{2n} (Y_i - \bar{Y}^2)^2$$

$$df(1) = 2n - 2$$

M_0 Reduced Model $H_0: \delta = 0 \quad (\mu_2 = \mu_1)$

$$\underline{Y} = X^0 \beta^0 + \underline{\varepsilon}$$

$$= \left(\begin{array}{c} \underline{\mathbb{1}} \\ \hline \underline{\mathbb{1}} \end{array} \right) (\mu_1) + \underline{\varepsilon}$$

$$\hat{\beta}^0 = \hat{\mu}_1 = \bar{Y} = \frac{n}{2n}\bar{Y}_1 + \frac{n}{2n}\bar{Y}_2$$

$$\hat{\underline{Y}}^0 = X^0 \hat{\beta}^0 = \begin{pmatrix} \mathbb{1} \\ \mathbb{1} \end{pmatrix} \bar{Y}$$

$$\underline{R}^0 = \underline{Y} - \hat{\underline{Y}}^0$$

$$SSE(0) = \|\underline{R}^0\|_2$$

$$= (\underline{Y} - \hat{\underline{Y}}^{(0)})^T (\underline{Y} - \hat{\underline{Y}}^{(0)})$$

$$= \sum_{i=1}^{n} (Y_i - \bar{Y})^2 + \sum_{i=n+1}^{2n} (Y_i - \bar{Y})^2$$

$$df_0 = 2n - 1$$

$$F_{STAT} = \frac{\frac{(SSE(0) - SSE(1))}{1}}{\frac{SSE(1)}{2n-2}} \quad \star$$

$$\overset{H_0}{\sim} F(1, 2n-2)$$

$SSE(0) - SSE(1)$

$$= n(\bar{Y}^1 - \bar{Y})^2 + n(\bar{Y}^2 - \bar{Y})^2$$

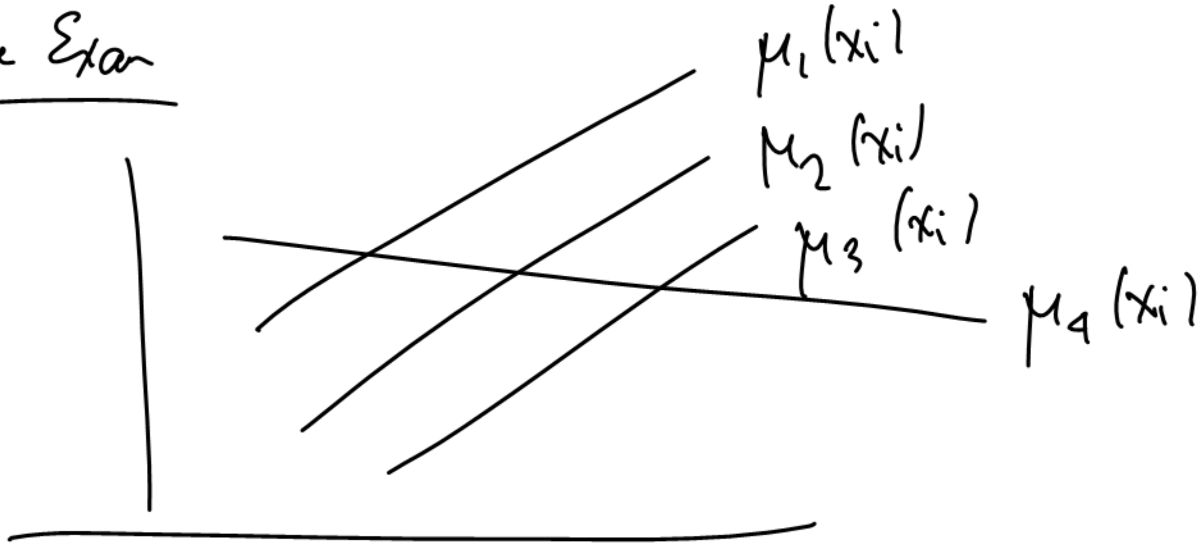
<u>Now</u> if H_0 true (i.e; $\mu_1 = \mu_2$ or $\delta = 0$)

then we expect $\bar{Y}^1$ & $\bar{Y}^2$ to be "close"

$\therefore$ $\bar{Y}$ "close" to both $\bar{Y}^1$ and $\bar{Y}^2$

$\therefore$ $(\bar{Y}^1 - \bar{Y})^2$ & $(\bar{Y}^2 - \bar{Y})^2$ "small"

Sample Exam



$$\mu_1(x_i) = \beta_0^1 + \beta_1^1 x_i$$

$$\mu_2(x_i) = \beta_0^2 + \beta_1^2 x_i$$
$$= (\beta_0^1 + \Delta_0^2) + (\beta_1^1 + \Delta_1^2) x_i$$

$$\mu_3(x_i) = \beta_0^3 + \beta_1^3 x_i$$
$$= (\beta_0^1 + \Delta_0^3) + (\beta_1^1 + \Delta_1^3) x_i$$

$$\mu_4(x_i) = (\beta_0^1 + \Delta_0^4) + (\beta_1^1 + \Delta_1^4) x_i$$

H_0: $\mu_1(x_i)$, $\mu_2(x_i)$, $\mu_3(x_i)$ parallel

$\iff \beta_1^1 = \beta_1^2 = \beta_1^3$

$\iff \Delta_1^2 = 0$ and $\Delta_1^3 = 0$

M_1 General (full model)

$$
\begin{pmatrix} \underline{Y}^{(1)} \\ \hline \underline{Y}^{(2)} \\ \hline \underline{Y}^{(3)} \\ \hline \underline{Y}^{(4)} \end{pmatrix}
=
\left[\begin{array}{cc|cc|cc|cc}
\underline{\mathbb{1}} & \begin{matrix} x_1 \\ \vdots \\ x_{n_1} \end{matrix} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{0} \\ \hline
\underline{\mathbb{1}} & \begin{matrix} x_{n_1+1} \\ \vdots \\ x_{n_1+n_2} \end{matrix} & \underline{\mathbb{1}} & \begin{matrix} x_{n_1+1} \\ \\ x_{n_1+n_2} \end{matrix} & \underline{0} & \underline{0} & \underline{0} & \underline{0} \\ \hline
 & & & & & & & \\ \hline
 & & & & & & &
\end{array}\right]
\begin{pmatrix} \beta_0^1 \\ \beta_1^1 \\ \Delta_0^2 \\ \Delta_1^2 \\ \Delta_0^3 \\ \Delta_1^3 \\ \Delta_0^4 \\ \Delta_1^4 \end{pmatrix}
+
\begin{pmatrix} \underline{\varepsilon}^{(1)} \\ \vdots \\ \hline \underline{\varepsilon}^{(4)} \end{pmatrix}
$$

$$\underline{Y} = X \underline{\beta} + \underline{\varepsilon}$$

$$\hat{\underline{\beta}}, \quad \underline{R}, \quad SSE(1) = \|\underline{R}\|_2$$

$$df_1 = n_1 + n_2 + n_3 + n_4 - 8$$

M_0 (reduced model)

$$\mu_1(x_i) = \beta_0^1 + \underline{\underline{\beta_1^1}}\, x_i$$

$$\mu_2(x_i) = (\beta_0^1 + \Delta_0^2) + \underline{\underline{\beta_1^1}}\, x_i$$

$$\mu_3(x_i) = (\beta_0^1 + \Delta_0^3) + \beta_1^1\, x_i$$

$$\mu_4(x_i) = (\beta_0^1 + \Delta_0^4) + (\beta_1^1 + \Delta_1^4)\, x_i$$

$$\begin{bmatrix} \underline{Y}_1 \\ \underline{Y}_2 \\ \underline{Y}_3 \\ \underline{Y}_4 \end{bmatrix} = \underbrace{\begin{bmatrix} \underline{1} & \underline{x}'_s & \underline{0} & \underline{0} & \underline{0} & \underline{0} \\ \underline{1} & \underline{x}'_s & \underline{1} & \underline{0} & \underline{0} & \underline{0} \\ \underline{1} & \underline{x}'_s & \underline{0} & \underline{1} & \underline{0} & \underline{0} \\ \underline{1} & \underline{x}'_s & \underline{0} & \underline{0} & \underline{1} & \underline{x}'_1 \end{bmatrix}}_{X^0} \underbrace{\begin{bmatrix} \beta_0^1 \\ \beta_1^1 \\ \Delta_0^2 \\ \Delta_0^3 \\ \Delta_0^4 \\ \Delta_1^4 \end{bmatrix}}_{\beta^0} + \underline{\varepsilon}$$

$$SSE(0) = \|\underline{R}^0\|_2$$

$$df_0 = (n_1 + n_2 + n_3 + n_4) - 6$$

$$F_{STAT} = ----$$

Suppose $F_{STAT,obs} > F(.95;\ 2,\ n_1+n_2+n_3+n_4-8)$

Post- hoc