Lecture 06 Oct 04

Orthogonal vectors. Let $\underline{u} = (u_1, \dots, u_p)' \in \mathbb{R}^p$ and
$\underline{v} = (v_1, \dots v_p)' \in \mathbb{R}^p$

$\underline{u}$ and $\underline{v}$ are orthogonal if

$$\langle \underline{u}, \underline{v} \rangle = \sum_{i=1}^{p} u_i v_i = 0.$$

Linear Model: $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$

Recall that the LSE of β_0 and β_1:

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x})(Y_i - \bar{Y})}{\sum (x_i - \bar{x})^2} = \sum w_i Y_i \quad \text{where} \quad w_i = \frac{(x_i - \bar{x})}{\sum (x_j - \bar{x})^2}$$

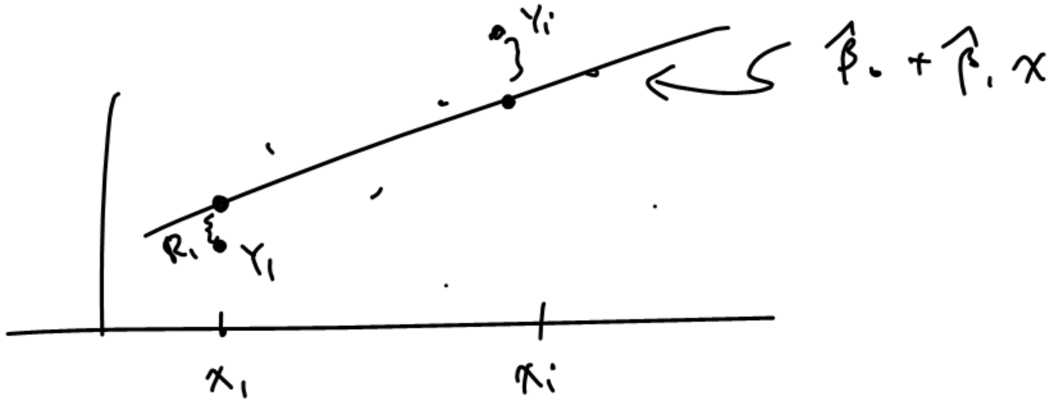
$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x}$$

Note that $(\hat{\beta}_0, \hat{\beta}_1)$ were derived from:

$$\sum_{i=1}^{n} \left(\overbrace{Y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)} \right) 1 = 0 \quad \leftarrow (*)$$

$$\sum \overbrace{\left(Y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)\right)} \underset{=}{x_i} = 0 \quad (**)$$

Define $\underline{R} = \begin{pmatrix} R_1 \\ \vdots \\ R_n \end{pmatrix} = \begin{pmatrix} Y_1 - (\hat{\beta}_0 + \hat{\beta}_1 x_1) \\ \vdots \\ Y_n - (\hat{\beta}_0 + \hat{\beta}_1 x_n) \end{pmatrix}$



Define $\mathbb{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ and $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

$$(*) : \quad \langle \underline{R}, \mathbb{1} \rangle = 0$$

$$\langle \underline{R}, \underline{x} \rangle = 0$$

$\underline{R}$ orthogonal to the colums of the design matrix

$$X = \begin{pmatrix} \underline{\mathbb{1}} & \begin{matrix} x_1 \\ \vdots \\ x_n \end{matrix} \end{pmatrix}$$

Geometric Interpretation of the LSE $(\hat\beta_0, \hat\beta_1)$:

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

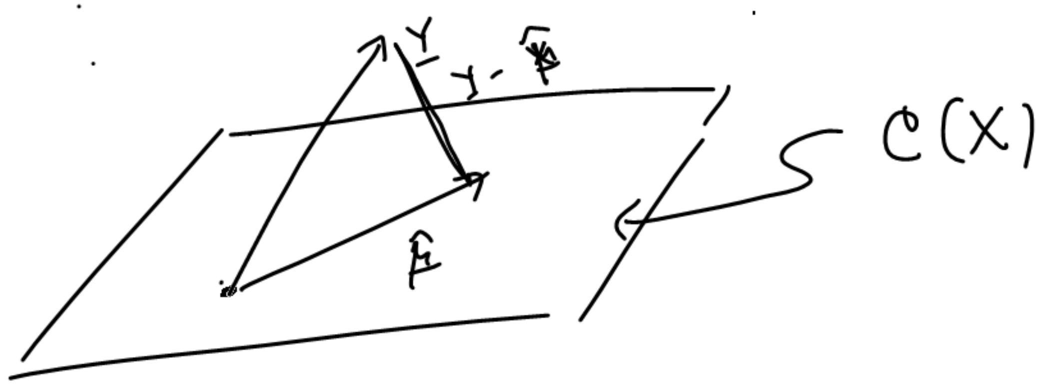
$$\underline{\varepsilon} \sim N(\underline{0}, \underbrace{I \otimes \sigma^2}_{\begin{pmatrix} \sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2 \end{pmatrix}})$$

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

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

$$\underline{\mu} = \mathbb{1}\cdot\beta_0 + \underline{x}\cdot\beta_1 \qquad X = (\mathbb{1} \,|\, \underline{x})$$

$\underline{\mu}$ is a linear combination of $\{\mathbb{1}, \underline{x}\}$



$C(\underline{X})$ = space spanned by the columns of X

$$= \{ \mathbb{1}\alpha_1 + \underline{x}\alpha_2, \quad \alpha_1, \alpha_2 \in \mathbb{R} \}$$

$$E(\underline{Y}) = \underline{\mu} = \mathbb{1}\beta_0 + \underline{x}\beta_1 \in C(X)$$

To estimate μ:

$$\hat{\underline{\mu}} \in C(X)$$

The estimator of $\underline{\mu}$: call $\hat{\underline{\mu}}$ is the

LSE if:

$\underbrace{(\underline{Y} - \hat{\underline{\mu}})}_{\underline{R}}$ is orthogonal to every column of X

In particular, if $X = (\underline{X}_1 | \dots | \underline{X}_p)$

then
$$\langle \underline{R}, \underline{X}_1 \rangle = 0 \iff \underline{X}_1' \underline{R} = 0$$
$$\vdots \qquad\qquad \vdots$$
$$\langle \underline{R}, \underline{X}_p \rangle = 0 \iff \underline{X}_p' \underline{R} = 0$$

$$\iff \quad X' \underbrace{(\underline{Y} - X\hat{\underline{\beta}})}_{R} = \underline{0}$$

$$\iff \quad X'Y - X'X\hat{\underline{\beta}} = \underline{0}$$

$$\iff \quad (X'X)\hat{\underline{\beta}} = X'\underline{Y} \qquad (X'X)^{-1} \text{ exists}$$

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

$(X'X)^{-1}$ exists $\iff$ columns of X are linearly independent

$$\left[\begin{array}{l} C(\underline{b}) = (\underline{Y} - X\underline{b})'(\underline{Y} - X\underline{b}) \\ \dfrac{\partial C(\underline{b})}{\partial \underline{b}} = -2\, X'(Y - Xb) \end{array}\right.$$

$$\left.\frac{\partial C(\underline{b})}{\partial \underline{b}}\right|_{\hat{\beta}} = 0 \Rightarrow \underline{\underline{X}}'\underbrace{(\underline{Y} - X\hat{\beta})}_{} = \underline{0}$$

BACK TO INFERENCE on β_1

The estimator of β_1 is:

$$\hat{\beta}_1 = \hat{\beta}_1(\underline{Y}) = \sum_{i=1}^{n} w_i Y_i \quad \text{where}$$

$$w_i = \frac{x_i - \bar{x}}{\Sigma (x_j - \bar{x})^2}$$

Derive the $E(\hat{\beta}_1)$, $Var(\hat{\beta}_1)$, distn of $\hat{\beta}_1$

$$E\,\hat{\beta}_1 = E\left(\sum_{i=1}^{n} w_i Y_i\right)$$

$$= \sum_{i=1}^{n} E(w_i Y_i) = \sum_{i=1}^{n} w_i\, E\, Y_i$$

$$= \sum_{i=1}^{n} \underline{w_i}\,(\beta_0 + \underline{\beta_1 x_i})$$

$$= \sum_{i=1}^{n} \frac{(x_i - \bar{x})}{\underbrace{\sum (x_j - \bar{x})^2}}\,\beta_0 + \beta_1 \sum_{i=1}^{n} \frac{x_i\,(x_i - \bar{x})}{\sum_j (x_j - \bar{x})^2}$$

$$= \frac{\beta_0}{\sum_j (x_j - \bar{x})^2} \cdot \cancel{\sum (x_i - \bar{x})}^{\,0} + \beta_1 \frac{\sum_i (x_i - \bar{x})(x_i - \bar{x})}{\sum_j (x_j - \bar{x})^2}$$

Note: $\sum_i (x_i - \bar{x}) = 0 \Rightarrow \bar{x} \sum_{i} (x_i - \bar{x}) = 0$

$$\sum x_i (x_i - \bar{x}) = \sum (x_i - \bar{x})(x_i - \bar{x}) = \sum (x_i - \bar{x})^2$$

$\therefore\ E\,\hat{\beta}_1 = \beta_1$ ✓

Idea: DATA 1: $\{(x_i, y_i^1), i = 1 \ldots, n\} \Rightarrow \hat{\beta}^1_{1,obs}$ data

DATA 2: $\{(x_i, y_i^2), i = 1 \ldots n\} \Rightarrow \hat{\beta}^2_{1,obs}$

⋮

many

$$\mathrm{Var}(\hat{\beta}_1) = \mathrm{Var}\left(\sum_{i=1}^{n} w_i Y_i\right)$$

$\{Y_i\text{'s indep}\}$

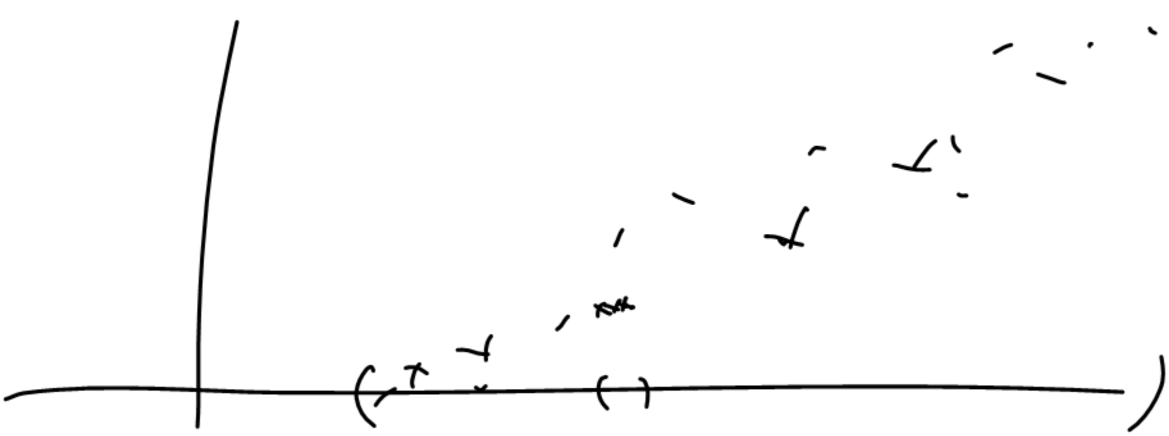
$$= \sum_{i=1}^{n} \mathrm{Var}(w_i Y_i)$$

$$= \sum_{i=1}^{n} w_i^2 \underbrace{\mathrm{Var}\, Y_i}_{\sigma^2}$$

$$= \sigma^2 \cdot \sum_{i=1}^{n} w_i^2 = \sigma^2 \cdot \sum \left(\frac{(x_i - \bar{x})}{\sum (x_j - \bar{x})^2} \right)^2$$

$$= \sigma^2 \cdot \frac{\sum (x_i - \bar{x})^2}{\left(\sum (x_j - \bar{x})^2\right)^2} = \frac{\sigma^2}{\sum (x_j - \bar{x})^2}$$

Rem: Uncertainty about $\hat{\beta}_1$ $\frac{1}{\alpha}$ Spread of $\{x_i\}$



Finally, since $\hat{\beta}_1$ is a linear combination of $\{Y_i\}$ which ~~has~~ are all normally distributed

MGF technique

$$\Rightarrow \quad \hat{\beta}_1 \sim N\left(\beta_1, \ \sigma^2 / \Sigma(x_i-\bar{x})^2\right)$$

$$\Rightarrow \quad \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\sigma^2 / \Sigma(x_i-\bar{x})^2}} \sim N(0,1)$$

Moreover:

$$\hat{\sigma}^2 = \frac{\Sigma R_i^2}{n-2} = \frac{\Sigma\left(Y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)\right)^2}{n-2}$$

$$\frac{(n-2)\hat{\sigma}^2}{\sigma^2} = \frac{\Sigma\left(Y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)\right)^2}{\sigma^2}$$

$$\sim \chi^2(df = n-2).$$

and $\hat{\underline{\beta}} \ \underset{\text{indep}}{\perp} \ (\underline{Y} - X\hat{\underline{\beta}})$

$$\hat{\beta}_1 \perp \hat{\sigma}^2$$

$$\Rightarrow \quad \frac{\dfrac{\hat{\beta}_1 - \beta_1}{\sqrt{\sigma^2 / \Sigma (x_i - \bar{x})^2}}}{\sqrt{\dfrac{\cancel{(n-2)}\hat{\sigma}^2}{\cancel{\sigma^2}} \Big/ \cancel{(n-2)}}} \sim t(df = n-2)$$

$$\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\hat{\sigma}^2 / \Sigma (x_i - \bar{x})^2}} \sim t(df = n-2)$$

$H_0: \quad \beta_1 = 0 \quad vs \quad H_1: \quad \beta_1 \neq 0$