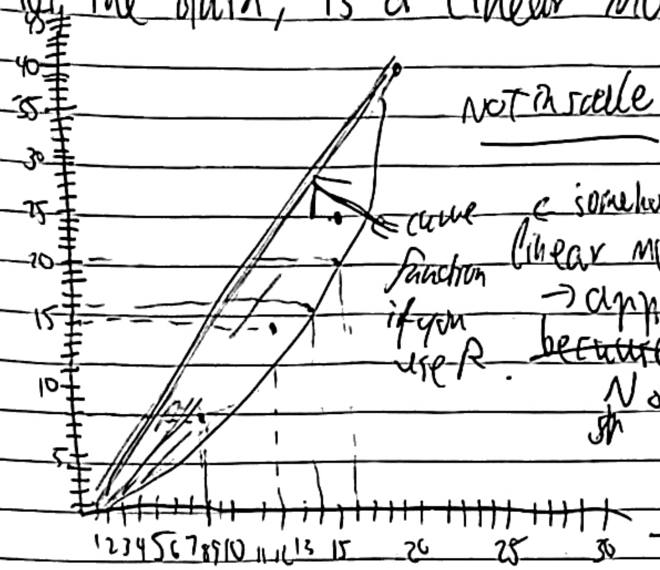
Week 4 Lab MAST30025: NOTE: This is the first tutorial exercises that uses both by hand & R

Q1)

Plot the data; is a linear model appropriate?



≤ somehow close to a linear model.

→ appropriate? No, ~~because~~ Non linear, ~ shape of a parabola,

→ There is no relationship b/w Income & no. of years of education.

Answer a bit vague!!!!???

there is a linear relationship between income & years of education.

Points have slight deviation from the linear model line.

By hand (1st Method)

2nd Method using R.

~~2. Write down the linear matrix form~~

Write down the linear model in matrix form.

$$y = \begin{bmatrix} 8 \\ 15 \\ 16 \\ 20 \\ 25 \\ 40 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 8 \\ 1 & 12 \\ 1 & 14 \\ 1 & 16 \\ 1 & 16 \\ 1 & 20 \end{bmatrix} \quad \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \beta \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_6 \end{bmatrix}$$

$Y = X\beta + \varepsilon$

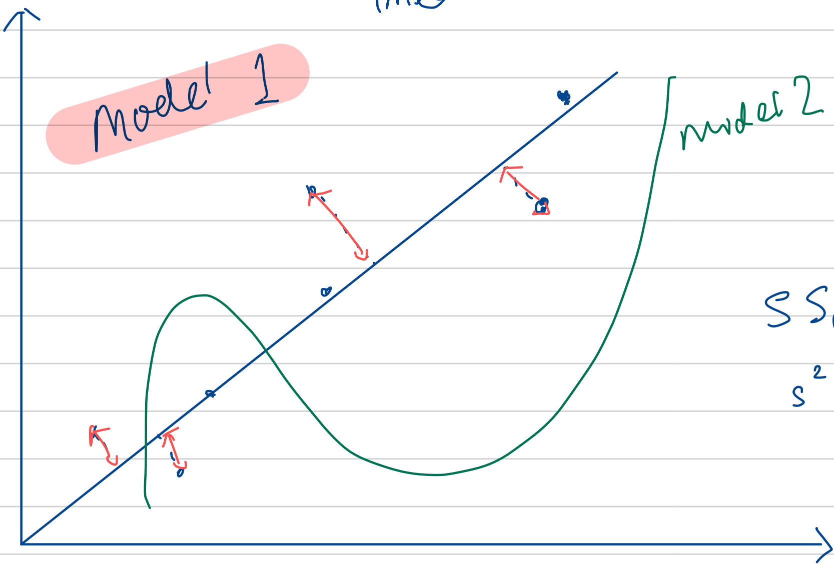
$\Rightarrow$ ~~$y = X\beta + \varepsilon$~~

$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ (simple linear regression)

$$\begin{bmatrix} 8 \\ 15 \\ 16 \\ 20 \\ 25 \\ 40 \end{bmatrix} = \begin{bmatrix} 1 & 8 \\ 1 & 12 \\ 1 & 14 \\ 1 & 16 \\ 1 & 16 \\ 1 & 20 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix}$$

$Y = X\beta + \varepsilon$

linear



$y = X\beta + \varepsilon$

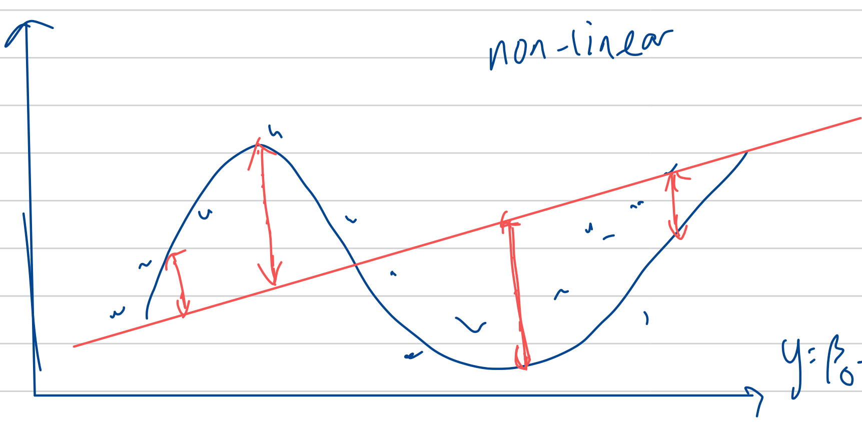
$\varepsilon \sim N(0, \sigma^2)$

$SS_{Res} = (y - Xb)^T (y - Xb)$

$s^2 = \frac{SS_{Res}}{n-p}$

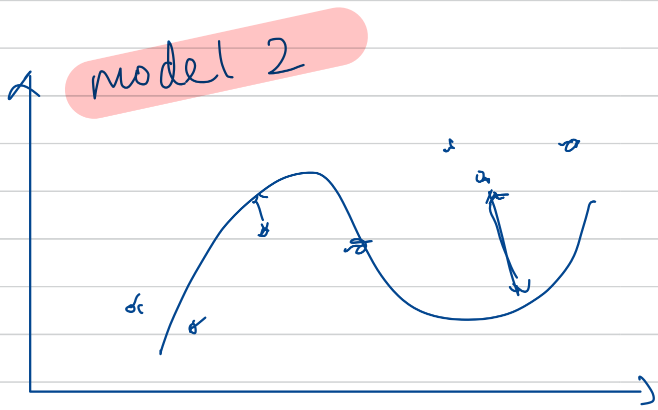
same datapoints.

non-linear



$\varepsilon \not\sim N(0, \sigma^2)$

$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 \dots$



$SS_{Res}^{(2)} > SS_{Res}^{(1)}$

3. Find the normal equations for this model.

from se. 15 (The full rank model).

$b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$, $e = \begin{bmatrix} e_0 \\ e_1 \end{bmatrix}$

Use $X^T X b = X^T y$ ← normal equations

$b = (X^T X)^{-1} X^T y$

$y = Xb + e$

So, $\sum_{i=1}^{n} e_i^2 = e^T e = (y - Xb)^T (y - Xb)$.

⇒ Choose b to minimize,

$$\begin{aligned} e^T e &= (y-Xb)^T (y-Xb) \\ &= y^T y - y^T X b - b^T X^T y + b^T X^T X b \\ &= y^T y - 2 y^T X b + b^T X^T X b \\ &= y^T y - 2(X^T y)^T b + b^T (X^T X) b. \end{aligned}$$

That is, we need,

$$\frac{\partial e^T e}{\partial b} = 0$$

Good practice!

since y does not depend on b.

$$\Rightarrow \frac{\partial}{\partial b}\left(y^T y - 2(X^T y)^T b + b^T (X^T X) b \right) = 0$$

$$\underbrace{\cancel{\frac{\partial}{\partial b}(y^T y)}}_{\substack{= 0 \\ \text{Case 1}}} - \underbrace{2 \frac{\partial}{\partial b}(X^T y)^T b}_{\substack{= ? \\ \text{Case 2}}} + \underbrace{\frac{\partial}{\partial b}(b^T (X^T X) b)}_{\substack{= ? \\ \text{Case 3}}} = 0$$

$$-2X^T y + \underline{(X^T X)} b + \underline{(X^T X)^T} b = 0$$

Both are symmetric I assume.

$$-2X^T y + 2(X^T X) b = 0$$

$y = \begin{bmatrix} 8 \\ 15 \\ 16 \\ 20 \\ 25 \\ 40 \end{bmatrix}$, $X = \begin{bmatrix} 1 & 8 \\ 1 & 12 \\ 1 & 14 \\ 1 & 16 \\ 1 & 16 \\ 1 & 20 \end{bmatrix}$, $b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$, $\varepsilon = \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \end{bmatrix}$

By hand

$\Rightarrow$ Normal equations,

$X^T X b = X^T y$

$$\underset{2\times6}{\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 8 & 12 & 14 & 16 & 16 & 20 \end{bmatrix}} \underset{6\times2}{\begin{bmatrix} 1 & 8 \\ 1 & 12 \\ 1 & 14 \\ 1 & 16 \\ 1 & 16 \\ 1 & 20 \end{bmatrix}} \underset{2\times1}{\begin{bmatrix} b_0 \\ b_1 \end{bmatrix}} = \underset{2\times6}{\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 8 & 12 & 14 & 16 & 16 & 20 \end{bmatrix}} \underset{6\times1}{\begin{bmatrix} 8 \\ 15 \\ 16 \\ 20 \\ 25 \\ 40 \end{bmatrix}}$$

$$\Rightarrow \underset{2\times2}{\overset{X^TX}{\begin{bmatrix} 6 & 86 \\ 86 & 1316 \end{bmatrix}}} \underset{2\times1}{\overset{b}{\begin{bmatrix} b_0 \\ b_1 \end{bmatrix}}} = \underset{2\times1}{\overset{X^Ty}{\begin{bmatrix} 124 \\ 1988 \end{bmatrix}}}$$

$b = (X^TX)^{-1} X^T y$

$$\begin{bmatrix} 6b_0 + 86b_1 \\ 86b_0 + 1316b_1 \end{bmatrix} = \begin{bmatrix} 124 \\ 1988 \end{bmatrix} \leftarrow t(X) \cancel{\%*\%} y$$

can solve by finding $(X^TX)^{-1}$ & then $(X^TX)^{-1}X^Ty$

QED

Q4). Solve the normal equations to obtain the least squares estimates of the parameters. Add the fitted regression line to your plot (using curve for example). By hand

$\Rightarrow$ from Q3 Solve b_0 & b_1 i.e. the b vector.

$6b_0 + 86b_1 = 124$ — (1)

$86b_0 + 1316b_1 = 1988$ — (2)

$516b_0 + 7396b_1 = 10664$ (1)×86 — (3)

$516b_0 + 7896b_1 = 11928$ (2)×6 — (4)

(4) − (3) $500b_1 = 1264$

$\Rightarrow b_1 = 2.528$

$\Rightarrow b_0 = \cancel{-93.408}$ 15.568

$\Rightarrow b = \begin{bmatrix} \cancel{-93.408} \; -15.568 \\ 2.528 \end{bmatrix} = \beta$

forgot to divide by 6.

oops!

→ So Q1 by hand!

2ND Method, by ~~hand~~ R check your mac!!

Q5) by hand,

This is a ~~model~~ simple linear regression model. Use the standard linear regression formulae

$$b_1 = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2}, \quad b_0 = \bar{y} - b_1\bar{x},$$

to estimate the parameters again (where the bar indicates the mean). Check that you have the same answers as above.

soln. (Ignore all of this / This is me slowly understanding the Q) in this page

we know,

$b_0 = -15.568$

$b_1 = 2.528$

$y = X\beta + \varepsilon$

$$= \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \varepsilon$$

$$\Rightarrow b_1 = 2.528 = \frac{Cov(X,Y)}{Var(X)}$$

$$\Rightarrow b_0 = -15.568 = \frac{\sum_i x_i^2 \sum_i y_i - \sum_i x_i y_i}{n\sum x_i^2 - (\sum_i x_i)^2}$$

Using Thm 4.2, where $\beta = (X^TX)^{-1}X^Ty$ is unbiased.

$$\Rightarrow E[b] = \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} -15.568 \\ 2.528 \end{bmatrix}$$

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$\Rightarrow Var\, b = (X^TX)^{-1}\sigma^2 = \begin{bmatrix} 6 & 86 \\ 86 & 1316 \end{bmatrix}^{-1}\sigma^2$$

$$= \frac{1}{6(1316) - 86^2} \begin{bmatrix} 1316 & -86 \\ -86 & 6 \end{bmatrix}\sigma^2$$

wait, next page

Attempt 2 (By hand)

$$\bar{y} = \frac{8+15+16+20+25+40}{6} = 20.67 \text{ (\$k)}$$

$$\bar{x} = \frac{8+12+14+16+16+20}{6} = \frac{43}{3} = 14.33 \text{ (in years)}$$

$$\overline{xy} = \frac{1}{n}\sum_i x_i y_i = \frac{1}{6}(64+180+224+320+400+800)$$
$$= 331.33$$

$$\Rightarrow \cancel{b_1 = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2}}$$

$$\bar{x}^2 = 205.444444$$

$$\overline{x^2} = \frac{1}{n}\sum_i x_i^2 = \frac{1}{6}(64+144+196+256+256+400)$$
$$= 219.333.$$

Now

$$b_1 = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2} = \cancel{\frac{331.33 - 205.44}{219.333 - 205}}$$

$$= \frac{331.33 - 14.33(20.67)}{219.333 - 205.4444} \approx 2.528$$

5 d.p.

$$b_0 = \bar{y} - b_1\bar{x},$$

$$b_0 = 20.67 - 2.528(14.33)$$
$$\approx -15.568$$

TIP: have / keep your significant figures to remain consistent.

$\Rightarrow$ Do I get conseq. marks? for small errors

Yes, but not much.

→ R code, see my macbook.

Q6) Calculate the sample ~~bias~~ variance s^2.

By hand,

Thm 4.6

deviance (model) *

$$s^2 = \frac{SS_{Res}}{n-p} = \frac{(y-Xb)^T(y-Xb)}{n-(k+1)} = \frac{1}{4}\left((y-Xb)^T(y-Xb)\right)$$

model $ df.residual

$$(y-Xb) = \begin{bmatrix} 8 \\ 15 \\ 16 \\ 20 \\ 25 \\ 40 \end{bmatrix} - \begin{bmatrix} 1 & 8 \\ 1 & 12 \\ 1 & 14 \\ 1 & 16 \\ 1 & 16 \\ 1 & 20 \end{bmatrix} \begin{bmatrix} -15.568 \\ 2.528 \end{bmatrix}$$

6×2   2×1

extra first column consists of all 1's.

no# of parameters β_0 & β_1. $p=2$

$$= \begin{bmatrix} 8 \\ 15 \\ 16 \\ 20 \\ 25 \\ 40 \end{bmatrix} - \begin{bmatrix} -15.568 + 8(2.528) \\ -15.568 + 12(2.528) \\ -15.568 + 14(2.528) \\ -15.568 + 16(2.528) \\ -15.568 + 16(2.528) \\ -15.568 + 20(2.528) \end{bmatrix}$$

$$= \begin{bmatrix} 3.344 \\ 0.232 \\ -3.824 \\ -4.88 \\ 0.12 \\ 5.008 \end{bmatrix}$$

$$(y-Xb)^T = [3.344 \quad 0.232 \quad -3.824 \quad -4.88 \quad 0.12 \quad 5.008]$$

$$\Rightarrow s^2 = \frac{1}{4}[3.344 \quad 0.232 \quad -3.824 \quad -4.88 \quad 0.12 \quad 5.008][\,\vdots\,]$$

$$s^2 = 0.25 \begin{bmatrix} 3.344 & 0.232 & -3.824 & -4.88 & 0.12 & 5.008 \end{bmatrix} \begin{bmatrix} 3.344 \\ 0.232 \\ -3.824 \\ -4.88 \\ 0.12 \\ 5.008 \end{bmatrix}$$

$$= \frac{1}{4}(74.768) = 18.692$$

QED

Q7) Estimate the average income of a person who has had 18 years of formal education.

$\cancel{y = X\beta + \varepsilon}$ By hand, $t^T = [1 \quad 18]$

Thm 4.5

$\cancel{E[y] = \beta_0 + \beta_1 x_1^* + \ldots}$

$$\overbrace{y}^{\text{income } (\$k)} = \beta_0 + \beta_1 \overbrace{x_1}^{\text{years}} + \varepsilon,$$ where $E[\varepsilon] = 0$ by independence

$$E[y] = \beta_0 + \beta_1 x_1^* = t^T \beta$$

$\boxed{\text{Income (in \$k)} = b_0 + b_1 (\text{no. of years})}$

where, $t = [1 \ x_1^* \ \cancel{x_2^*}]^T$.

$t^T b = [1 \ x_1^* \ \cancel{x_2^*}] b = b_0 + b_1 x_1^* \cancel{+ b_2 x_2^*}$ where b is the least squares estimator for β_1

$\cancel{= 212.528 - 15.568}$

$= -15.568 + 2.528 \times 18$

$= 29.936$ (in \$k)

$\Rightarrow$ \$29936

$\therefore$ The average income of a person who had 18 years of education is \$29936

Q8) - Not a good fit because there are 2 outliers, points 1&6 with Cook's distance greater than 1.

- OR: Good fit because points 1,6 have Cook's distance close to 1.

→ $\mathbb{E}[t^T b] = t^T \beta$ means that $t^T b$ is unbiased

Q9) We know that the least squares estimator b is an unbiased estimator for β. Show that $t^T b$ is an unbiased estimator for $t^T \beta$, where t is a vector of constants.

from Sl. 36/189.

(since t is a scalar)

$$E[t^T b] = t^T E[b]$$
$$= t^T \beta$$

← By the law of expectations.

By probability & statistics

Assumed / prior knowledge.

$$\mathbb{E}[t^T b] = t^T \mathbb{E}[b] \quad \text{since } t \text{ is a scalar}$$
$$= t^T \beta$$