lm( )   Find optimal $\hat{\beta}$

$$\min_{\hat{\beta}} \sum_{t=1}^{n} (Y_t - \hat{Y}_t)^2 = \min_{\boxed{\hat{\beta}}} \sum_{t=1}^{n} (Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_{1t} - \dots - \hat{\beta}_k X_{kt})^2$$

we observe the data

t	Y_t	X_t'			
1	Y_1	1	X_{11}	X_{21} ...	X_{k1}
2	Y_2	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮
n	Y_n	1	X_{1n}	X_{2n} ...	X_{kn}

X_t'

ES2: TE3

$$\min_{\hat{\beta}_0} \sum_{t=1}^{n} (Y_t - \hat{\beta}_0)^2$$ is a special case of

t	Y_t	X_t'
1	Y_1	1
2	Y_2	1
⋮	⋮	⋮
n	Y_n	1

$$\left(\frac{1}{n} \sum_{t=1}^{n} X_t X_t' \right)^{-1}$$

$$\frac{1}{n}\sum_{t=1}^{n} \underbrace{\overbrace{\begin{pmatrix} 1 \\ X_{1t} \\ \vdots \\ X_{kt} \end{pmatrix}}^{X_t}}_{(k+1)\times 1} \underbrace{\overbrace{\begin{pmatrix} 1 & X_{1t} & \dots & X_{kt} \end{pmatrix}}^{X_t'}}_{1\times(k+1)} = \frac{1}{n}\sum_{t=1}^{n} \overbrace{\begin{pmatrix} 1 & X_{1t} & X_{2t} & \dots & X_{kt} \\ X_{1t} & X_{1t}^2 & X_{1t}X_{2t} & \dots & X_{1t}X_{kt} \\ \vdots & \vdots & & \ddots & \vdots \\ X_{kt} & X_{kt}X_{1t} & X_{kt}X_{2t} & \dots & X_{kt}^2 \end{pmatrix}}^{(k+1)\times(k+1)}$$

NOTE

$$\underbrace{\begin{pmatrix} 1 \\ X_{1t} \\ \vdots \\ X_{kt} \end{pmatrix}}_{(k\times 1)\times 1} \underbrace{\begin{pmatrix} 1 & X_{1t} & \cdots & X_{kt} \end{pmatrix}}_{1\times(k+1)} \to (k+1)\times(k+1) \neq \underbrace{\cancel{\begin{pmatrix} 1 & X_{1t} & \cdots & X_{kt} \end{pmatrix}}}_{1\times(k+1)} \underbrace{\begin{pmatrix} 1 \\ X_{1t} \\ \vdots \\ X_{kt} \end{pmatrix}}_{(k+1)\times 1} = 1 + X_{1t}^2 + X_{2t}^2 + \dots + X_{kt}^2$$

1×1 scalar

$$\frac{1}{n}\sum_{t=1}^{n} \underbrace{\begin{pmatrix} 1 & X_{1t} & X_{2t} & \cdots & X_{kt} \\ X_{1t} & X_{1t}^2 & X_{1t}X_{2t} & \cdots & X_{1t}X_{kt} \\ \vdots & \vdots & & \ddots & \vdots \\ X_{kt} & X_{kt}X_{1t} & X_{kt}X_{2t} & \cdots & X_{kt}^2 \end{pmatrix}} = \underbrace{\begin{pmatrix} 1 & X_{11} & X_{21} & \cdots & X_{k1} \\ X_{11} & X_{11}^2 & X_{11}X_{21} & \cdots & X_{11}X_{k1} \\ \vdots & \vdots & & \ddots & \vdots \\ X_{k1} & X_{k1}X_{11} & X_{k1}X_{21} & \cdots & X_{k1}^2 \end{pmatrix}}_{\boxed{(k+1)} \times \boxed{(k+1)},\ t=1} + \underbrace{\begin{pmatrix} \ \ & & \\ & & \\ & & \end{pmatrix}}_{\boxed{(k+1)} \times \boxed{(k+1)},\ t=2} + \dots + \underbrace{\begin{pmatrix} \ \ & & \\ & & \\ & & \end{pmatrix}}_{\boxed{(k+1)} \times \boxed{(k+1)},\ t=n}$$

To add matrices together, make sure number of rows & columns of one matrix is the same as all other matrices.

an example of a symmetric matrix

$$= \frac{1}{n}\begin{pmatrix} n & \sum X_{1t} & \sum X_{2t} & \cdots & \sum X_{kt} \\ \sum X_{1t} & \sum X_{1t}^2 & \sum X_{1t}X_{2t} & \cdots & \sum X_{1t}X_{kt} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum X_{kt} & \sum X_{kt}X_{1t} & \sum X_{kt}X_{2t} & \cdots & \sum X_{kt}^2 \end{pmatrix}$$

A matrix A is symmetric when $A' = A$.
has to be a square matrix (number of rows = number of columns)

insight to when $\left(\frac{1}{n}\sum X_t X_t'\right)^{-1}$ does not exist

t	course eval		female
1		1	
2		1	
⋮		⋮	
n		1	

example. Quarto exercises course eval ~ female

course eval ~ female + male

t	course eval		female	male	$female^2$
		1	1	0	
→ 1		1	1	0	
2		⋮	0	1	
⋮					
n		1			

$X_t' = (1, female_t, male_t)$

n_f = total number of females
n_m = total " " males

$$\frac{1}{n}\sum X_t X_t' = \begin{pmatrix} n & \boxed{\sum_{t=1}^{n} female_t} & \sum_{t=1}^{n} male_t \\ \sum_{t=1}^{n} female_t & \sum_{t=1}^{n} female_t^2 & \sum_{t=1}^{n} female_t * male_t \\ \sum_{t=1}^{n} male_t & \sum_{t=1}^{n} female_t * male_t & \sum_{t=1}^{n} male_t^2 \end{pmatrix}$$

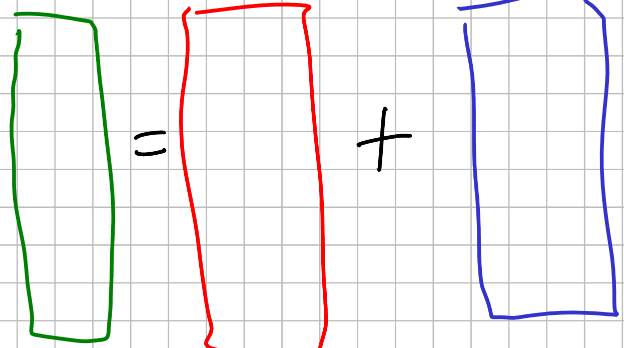
total number of females

total number of males

= Simplify this!

This matrix has no inverse!

$$= \begin{pmatrix} n & n_f & n_m \\ n_f & n_f & 0 \\ n_m & 0 & n_m \end{pmatrix}$$



Last time

$$\begin{pmatrix} 4 & 1 \\ 2 & 1/2 \end{pmatrix} \qquad \begin{pmatrix} 4 & 1 \\ 2 & 1/2 \end{pmatrix}$$

linearly combine vectors together

How do you resolve situations of perfect multicollinearity?

In $X_t' = (1, \text{female}_t, \text{male}_t)$ $\rightarrow$ $\hat{\beta}$ does not exist bec. of perfect multicollinearity

base category/group

either $X_t' = (1, \text{female}_t)$ in exercise

or $X_t' = (1, \text{male}_t)$ lm( )?

or $X_t' = (\text{female}_t, \text{male}_t)$ lm( )?

intercept	male
3.901	0.168
female	male
3.901	4.069

dummy variables → sometimes called fixed effects in economics

represent groups or individuals

example trade

i	j	X_{ij}	GDP_i	GDP_j
1	2	✓	•	–
1	3	✓	•	–
1	4	✓	•	–
1	5	✓	•	–
2	1	✓	☆	•
2	4	✓	☆	–
2	6	✓	☆	–

i, j represent countries — dyads

X_{ij} bilateral trade between the ith and jth country $i \neq j$

bootstrap

---

economics

simple linear regression

$$\hat{\beta}_1 = \frac{\sum (X_{1t} - \bar{X}_1)(Y_t - \bar{Y})}{\sum (X_t - \bar{X}_1)^2} = \frac{s_{X_1, Y}}{s^2_{X_1}}$$

$$\sum a X_t = a \sum X_t$$

$$\stackrel{TE}{=} \frac{\sum_{t=1}^{n} (X_{1t} - \bar{X}_1) Y_t}{\sum_{t=1}^{n} (X_{1t} - \bar{X}_1)^2} \rightarrow \text{just a number } c$$

$$\frac{\sum_{t=1}^{n} (X_{1t} - \bar{X}_1) Y_t}{c}$$

$$= \boxed{\sum_{t=1}^{n}} \left(\frac{X_{1t} - \bar{X}_1}{\sum (X_{1t} - \bar{X}_1)^2} \right) \boxed{Y_t}$$

$$\stackrel{TE}{=} \sum_{t=1}^{n} \left[\frac{1}{c} (X_{1t} - \bar{X}_1) \right] Y_t$$

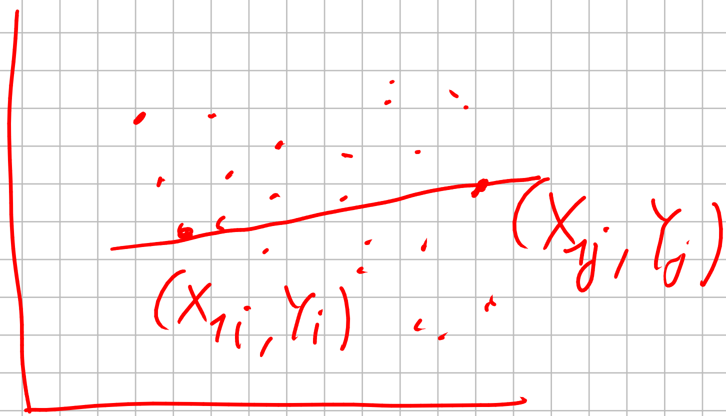
$$= \sum_{t=1}^{n} \boxed{w_t} Y_t \qquad w_t = \frac{X_{1t} - \bar{X}_1}{\sum (X_{1t} - \bar{X}_1)^2} \qquad \frac{\text{deviation from the mean of } X_1}{n * \text{variance of } X_1}$$

Recent applied research → weights

It can be shown that

$$\hat{\beta}_1 = \frac{\sum_i \sum_j (X_{1j} - X_{1i})^2 \boxed{\frac{Y_j - Y_i}{X_{1j} - X_{1i}}}}{\sum_i \sum_j (X_{1j} - X_{1i})^2}$$

→ remind you of the slope of a line passing through (X_{1i}, Y_i) & (X_{1j}, Y_j)



$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{pmatrix} \qquad \hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$$

e-16 ⇒ $\times 10^{-16}$

1.673e-04 ⇒ $1.673 \times 10^{-4} = 0.0001673$

$X_t' = (1, X_{1t} \dots X_{kt}) = (\boxed{1, X_{1t}, \dots, X_{(k-1)t}}, X_{kt})$

$\hat{\beta}_k$ obtained by FWL

$\Rightarrow$ simple linear regression of $\overbrace{Y_t - \text{fitted value of } Y_t \text{ using all regressors except the last one}}^{\text{residual}}$

on $X_{kt} -$ fitted value of X_{kt} using all other regressors except the last one

Some call this adjustment to Y + X_k as "controlling" or "holding other regressors constant"

BUT this is not accurate.

What you have seen (FWL) is pure algebra!

log10TotalComp <- log10(TotalCompMil * 10^6)

Similarly for log10NetSales

when log NetSales is zero

$\downarrow$

$\log_{10}$(Total compensation in dollars) = 2.75

predicted Total compensation in dollars = $10^{2.75}$ dollars

≈ 562

$\log_{10}$(Net Sales in dollars) = 0

$\Rightarrow$ Net Sales in dollars = 10^0 = 1