

# CLASSICAL LINEAR MODEL: ORDINARY LEAST SQUARES (OLS)

CONSIDER  $n$  OBSERVED PAIRS OF VARIABLES

$$(x_1, y_1) (x_2, y_2) \dots (x_n, y_n)$$

& SOME UNOBSERVED STRAIGHT LINE

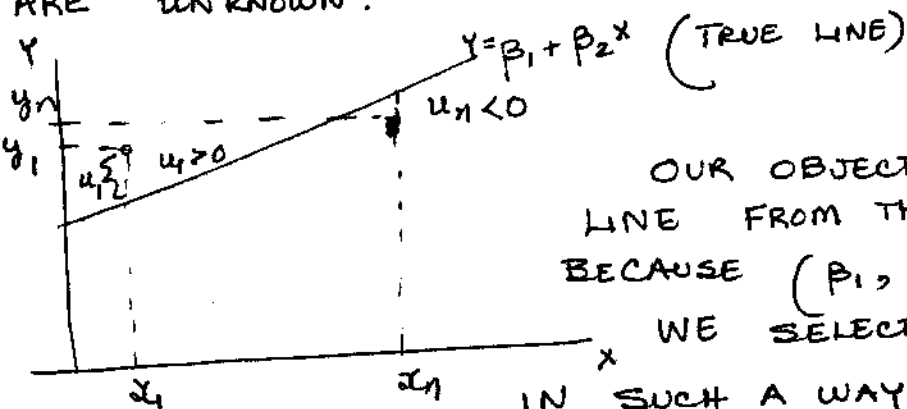
$$Y = \beta_1 + \beta_2 X.$$

THE OBSERVED POINTS WILL TYPICALLY NOT ALL LIE ON THE SAME STRAIGHT LINE. RATHER, THE  $Y$  VALUES HAVE BEEN GENERATED AS

$$y_i = \beta_1 + \beta_2 x_i + u_i$$

IF  $u_i$  IS POSITIVE,  $(x_i, y_i)$  LIES ABOVE THE LINE WHILE  $u_i < 0$  WHEN THE POINT LIES BELOW THE LINE.

NOTE THAT THE VALUES OF  $(x_i, y_i)$  ( $i=1, 2, \dots, n$ ) ARE OBSERVED. BUT THE PARAMETERS OF THE LINE  $(\beta_1, \beta_2)$  ARE UNKNOWN.



OUR OBJECTIVE IS TO RECOVER THE LINE FROM THE OBSERVED DATA. BECAUSE  $(\beta_1, \beta_2)$  ARE UNKNOWN, WE SELECT OUR ESTIMATES  $(\hat{\beta}_1, \hat{\beta}_2)$

IN SUCH A WAY THAT THE POINTS ARE "JOINTLY" CLOSEST TO THE FITTED LINE.

FOR ANY CHOSEN PAIR  $(\hat{\beta}_1, \hat{\beta}_2)$ ,

$$y_i = \hat{\beta}_1 + \hat{\beta}_2 x_i + e_i$$

WHERE  $e_i$  IS THE DEVIATION BETWEEN ACTUAL POINT & THE LINE. THAT IS

$$e_i = y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i$$

THE LEAST SQUARES PROCEDURE SELECTS  $(\hat{\beta}_1, \hat{\beta}_2)$

THAT MINIMIZES  $\sum_{i=1}^n e_i^2$ .

411 2-2

BECAUSE EACH  $e_i$  IS A FUNCTION OF  $\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$ ,  
 $S = \sum_i e_i^2$  IS ALSO A FUNCTION OF  $\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$ .

### THE LEAST SQUARES PROCEDURE

$$\min_{\hat{\beta}_1, \hat{\beta}_2} S = \sum e_i^2$$

FOC.  $\frac{\partial S}{\partial \hat{\beta}_1} = \sum 2e_i \left( \frac{\partial e_i}{\partial \hat{\beta}_1} \right) = 0$  (1)

$\frac{\partial S}{\partial \hat{\beta}_2} = \sum 2e_i \left( \frac{\partial e_i}{\partial \hat{\beta}_2} \right) = 0$  (2)

BUT  $e_i = y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i$

$\Rightarrow \frac{\partial e_i}{\partial \hat{\beta}_1} = -1$  &  $\frac{\partial e_i}{\partial \hat{\beta}_2} = -x_i$

THUS, (1)  $\Rightarrow \sum 2e_i (-1) = 0 \Rightarrow -2 \sum_i e_i = 0$  (1A)

(2)  $\Rightarrow \sum 2e_i (-x_i) = 0 \Rightarrow -2 \left( \sum_i e_i x_i \right) = 0$  (2A)

THUS, THE FIRST ORDER CONDITIONS FOR A MINIMUM ARE

$$\begin{aligned} \sum e_i &= 0 && \text{(1B)} \\ \sum e_i x_i &= 0 && \text{(2B)} \end{aligned}$$

WE NEED TO SOLVE (1B) - (2B) SIMULTANEOUSLY.

(1B)  $\Rightarrow \sum e_i = \sum (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i)$   
 $= \sum y_i - n \hat{\beta}_1 - \hat{\beta}_2 (\sum x_i) = 0$   
 $\Rightarrow \boxed{n \hat{\beta}_1 + \hat{\beta}_2 (\sum x_i) = (\sum y_i)}$  (1C)

(2B)  $\Rightarrow \sum x_i e_i = \sum x_i (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i)$   
 $= \sum x_i y_i - \hat{\beta}_1 \sum x_i - \hat{\beta}_2 \sum x_i^2 = 0$

$\Rightarrow \boxed{(\sum x_i) \hat{\beta}_1 + (\sum x_i^2) \hat{\beta}_2 = (\sum x_i y_i)}$  (2C)

USING CRAMER'S RULE TO SOLVE (1c) - (2c) 411 (2-3)

$$\hat{\beta}_1 = \frac{\begin{vmatrix} \sum Y_i & \sum X_i \\ \sum X_i Y_i & \sum X_i^2 \end{vmatrix}}{\begin{vmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{vmatrix}}; \quad \hat{\beta}_2 = \frac{\begin{vmatrix} n & \sum Y_i \\ \sum X_i & \sum X_i Y_i \end{vmatrix}}{\begin{vmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{vmatrix}}$$

THUS,

$$\hat{\beta}_1 = \frac{(\sum Y_i)(\sum X_i^2) - (\sum X_i)(\sum X_i Y_i)}{n(\sum X_i^2) - (\sum X_i)^2} \quad (3A)$$

$$\hat{\beta}_2 = \frac{n(\sum X_i Y_i) - (\sum X_i)(\sum Y_i)}{n(\sum X_i^2) - (\sum X_i)^2} \quad (3B)$$

FOCUS ON (3B). DIVIDING BOTH THE NUMERATOR & THE DENOMINATOR BY  $n^2$ ,

$$\hat{\beta}_2 = \frac{\frac{1}{n} \sum X_i Y_i - \left(\frac{1}{n} \sum X_i\right) \left(\frac{1}{n} \sum Y_i\right)}{\frac{1}{n} \sum X_i^2 - \left(\frac{1}{n} \sum X_i\right)^2} = \frac{\frac{1}{n} \sum X_i Y_i - \bar{X}\bar{Y}}{\frac{1}{n} \sum X_i^2 - \bar{X}^2}$$

THUS, 
$$\hat{\beta}_2 = \frac{S_{xy}}{S_x^2} \quad (4A)$$

$S_{xy}$  &  $S_x^2$  ARE, RESPECTIVELY, THE SAMPLE CO-VARIANCE BETWEEN (X, Y) & VARIANCE OF X.

RETURNING TO (1c),

$$\begin{aligned} n\hat{\beta}_1 + \hat{\beta}_2 \sum X_i &= \sum Y_i \\ \Rightarrow \hat{\beta}_1 + \hat{\beta}_2 \bar{X} &= \bar{Y} \quad \Rightarrow \end{aligned} \quad \boxed{\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}} \quad (4B)$$

# THE 2-VARIABLE MODEL IN MATRIX FORM

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

$$y = X\hat{\beta} + e \Rightarrow e = y - X\hat{\beta}$$

THUS  $S = \sum e_i^2 = e'e = (y - X\hat{\beta})'(y - X\hat{\beta})$

$$= y'y - \hat{\beta}'X'y - y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta}$$

$$= y'y - 2\hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta}$$

THE FOCs FOR MINIMIZATION:

$$\frac{\partial S}{\partial \hat{\beta}} = \frac{\partial (y'y - 2\hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta})}{\partial \hat{\beta}} = 0.$$

## DIGRESSION ON VECTOR-MATRIX CALCULUS

CONSIDER THE VECTORS  $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$  &  $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

THEN  $a'x = (a_1 \ a_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a_1x_1 + a_2x_2$

LET  $y = a'x = a_1x_1 + a_2x_2$

THEN THE VECTOR OF PARTIAL DERIVATIVES IS

$$\begin{pmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \end{pmatrix} \equiv \begin{pmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \Rightarrow \boxed{\frac{\partial (a'x)}{\partial x} = a}$$

WE COULD ALSO WRITE  $y = (x_1 \ x_2) \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = x'a$

AGAIN  $\begin{pmatrix} \frac{\partial y}{\partial x} \end{pmatrix} = \boxed{\frac{\partial x'a}{\partial x} = a}$

NOW CONSIDER THE MATRIX

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

DEFINE

$$\left. \begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 \\ y_2 &= a_{21}x_1 + a_{22}x_2 \end{aligned} \right\} \Rightarrow y = Ax$$

NEXT CONSIDER THE MATRIX

$$\begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} = A'$$

THUS

$$\begin{bmatrix} \frac{\partial y_1}{\partial x} & \frac{\partial y_2}{\partial x} \end{bmatrix} = \frac{\partial (Ax)}{\partial x} = A'$$

NEXT CONSIDER THE QUADRATIC FORM

WHERE  $y = x'Ax$  IS ASSUMED TO BE SYMMETRIC.

$$y = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

THUS

$$\frac{\partial y}{\partial x_1} = 2a_{11}x_1 + 2a_{12}x_2$$

$$\frac{\partial y}{\partial x_2} = 2a_{12}x_1 + 2a_{22}x_2$$

THUS

$$\frac{\partial y}{\partial x} = 2Ax$$

FURTHER

$$\frac{\partial^2 y}{\partial x \partial x'} = \frac{\partial (2Ax)}{\partial x} = 2A' = 2A$$

(BECAUSE A IS SYMMETRIC)

THUS,

$$\frac{\partial (x'Ax)}{\partial x} = 2Ax$$

$$\frac{\partial^2 (x'Ax)}{\partial x \partial x'} = 2A$$

RETURN TO THE FOC.

$$\frac{\partial}{\partial \beta} (y'y - 2\hat{\beta}'x'y + \hat{\beta}'x'x\hat{\beta}) = 0$$

$$\Rightarrow -2x'y + 2x'x\hat{\beta} = 0 \Rightarrow (x'x)\hat{\beta} = x'y$$

HENCE,

$$\hat{\beta} = (x'x)^{-1}x'y$$

RECALL THAT

$$X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$$

$$\Rightarrow X'X = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{bmatrix}$$

$$X'X = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}$$

$$X'y = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

THUS

$$\Leftrightarrow \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} \sum x_i y_i \\ \sum x_i^2 y_i \end{bmatrix}$$

THIS IS THE PAIR (1c) - (2c) THAT WE SOLVED BEFORE TO GET  $(\hat{\beta}_1, \hat{\beta}_2)$ .

$$y = X\hat{\beta} + e$$

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

$$e = y - X\hat{\beta} = y - X(X'X)^{-1}X'y$$

$$e = [I - X(X'X)^{-1}X']y$$

DEFINE THE MATRIX

$$M \equiv I - X(X'X)^{-1}X'$$

THEN

$$e = My$$

THE MATRIX  $M$  IS SOMETIMES CALLED THE RESIDUAL  
MAKER.

THE MATRIX

$$P \equiv X(X'X)^{-1}X'$$

IS CALLED THE PREDICTED  
VECTOR IS

PROJECTION MATRIX.

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

$$\hat{y} = Py$$

PROPERTIES OF  $M$ :

①  $M$  IS SYMMETRIC.

$$M = I - X(X'X)^{-1}X'$$

$$M' = I - X(X'X)^{-1}X' = M$$

$$\textcircled{2} \quad M'M = [I - X(X'X)^{-1}X'] [I - X(X'X)^{-1}X']$$

$$= I - X(X'X)^{-1}X' - X(X'X)^{-1}X' + X(X'X)^{-1}X'X(X'X)^{-1}X'$$

$$= I - X(X'X)^{-1}X' = M$$

$$M'M = M.M = M$$

THIS SHOWS THAT THE MATRIX  $M$  IS  
AN IDEMPOTENT MATRIX.

$$MX = \left[ I - X(X'X)^{-1}X' \right] X$$
$$= X - X(X'X)^{-1}X'X = X - X = 0.$$

THUS  $M$  &  $X$  ARE ORTHOGONAL.

NOTE THAT  $M = I - P \Rightarrow P = I - M$ .

$$\text{THUS } MP = M(I - M)$$
$$= M - M.M = M - M = 0.$$

$$e = My$$

$$= M(X\beta + u) = MX\beta + u$$

$$\Rightarrow e = Mu$$

$$e'e = (Mu)'(Mu)$$

$$= u'M'u = u'u$$