CLASSICAL LINEAR MODEL:
ORDINARY LEAST SQUARES (OLS)

CONSIDER n OBSERVED PAIRS OF VARIABLES

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

& SOME UNOBSERVED STRAIGHT LINE

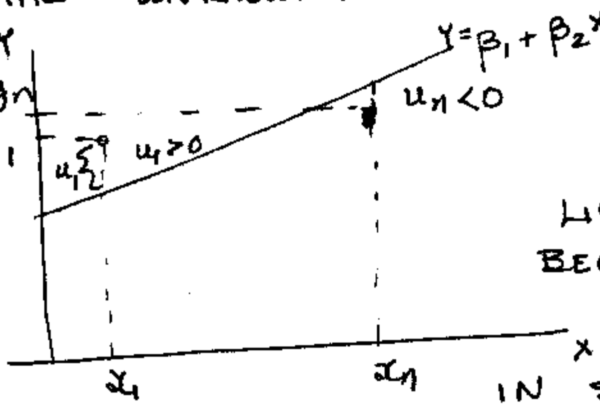
$$Y = \beta_1 + \beta_2 x .$$

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) $(i = 1, 2, \ldots, n)$ ARE OBSERVED. BUT THE PARAMETERS OF THE LINE (β_1, β_2) ARE UNKNOWN.



OUR OBJECTIVE IS TO RECOVER THE LINE FROM THE OBSERVED DATA. BECAUSE (β_1, β_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 $\sum_{i=1}^{n} e_i^2$.

BECAUSE EACH e_i IS A FUNCTION OF $(\hat{\beta}_1, \hat{\beta}_2)$, $S = \sum_i e_i^2$ IS ALSO A FUNCTION OF $(\hat{\beta}_1, \hat{\beta}_2)$

THE LEAST SQURES PROCEDURE

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

FOC.

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

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

BUT $e_i = Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i$

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

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

$$\sum e_i = 0 \quad (1B)$$

$$\sum e_i X_i = 0 \quad (2B)$$

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

(1B) $\Rightarrow$

$$\sum e_i = \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)$$

$$= \sum Y_i - n\hat{\beta}_1 - \hat{\beta}_2\left(\sum X_i\right) = 0$$

$$\Rightarrow \boxed{n\hat{\beta}_1 + \hat{\beta}_2\left(\sum X_i\right) = \left(\sum Y_i\right)} \quad (1C)$$

(2B) $\Rightarrow$

$$\sum X_i e_i = \sum X_i (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)$$

$$= \sum X_i Y_i - \hat{\beta}_1 \sum X_i - \hat{\beta}_2 \sum X_i^2 = 0$$

$$\Rightarrow \boxed{\left(\sum X_i\right)\hat{\beta}_1 + \left(\sum X_i^2\right)\hat{\beta}_2 = \left(\sum X_i Y_i\right)} \quad (2C)$$

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

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

THUS,

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

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

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

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

THUS, $$\boxed{\hat{\beta}_2 = \frac{S_{XY}}{S_X^2}} \quad (4A)$$

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

RETURNING TO (1c),

$$n\hat{\beta}_1 + \hat{\beta}_2 \Sigma X_i = \Sigma Y_i$$

$$\Rightarrow \hat{\beta}_1 + \hat{\beta}_2 \bar{X} = \bar{Y} \quad \Rightarrow \quad \boxed{\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}} \quad (4B)$$

THE 2-VARIABLE MODEL IN MATRIX FORM

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \hat\beta_1 \\ \hat\beta_2 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

$$y = X\hat\beta + e \quad \Rightarrow \quad e = y - X\hat\beta$$

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

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

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

THE FOCs FOR MINIMIZATION:

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

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

LET $y = a'x = a_1 x_1 + a_2 x_2$

THEN THE VECTOR OF PARTIAL DERIVATIVES IS

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

WE COULD ALSO WRITE

$$y = (x_1 \; x_2)\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = x'a$$

AGAIN $\left(\dfrac{\partial y}{\partial x}\right) = \boxed{\dfrac{\partial x'a}{\partial x} = a}$

NOW CONSIDER THE MATRIX

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

DEFINE

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

NEXT CONSIDER THE MATRIX

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

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

NEXT CONSIDER THE QUADRATIC FORM

$y = x'Ax$

WHERE A IS ASSUMED TO BE SYMMETRIC.

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

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

RETURN TO THE FOC.

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

$$\Rightarrow \quad -2X'y + 2X'X\hat\beta = 0 \quad \Rightarrow \quad (X'X)\hat\beta = X'y$$

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

RECALL THAT

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

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

$$X'y = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{bmatrix} \Sigma Y_i \\ \Sigma x_i Y_i \end{bmatrix}$$

THUS

$$X'X\hat\beta = X'y$$

$$\Leftrightarrow \begin{bmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{bmatrix} \begin{bmatrix} \hat\beta_1 \\ \hat\beta_2 \end{bmatrix} = \begin{bmatrix} \Sigma x_i Y_i \\ \Sigma x_i^2 \end{bmatrix}$$

THIS IS THE PAIR 1c - 2c

THAT WE SOLVED BEFORE

TO GET $(\hat\beta_1, \hat\beta_2)$.

$$y = x\hat{\beta} + e$$
$$\hat{\beta} = (x'x)^{-1}x'y$$
$$e = y - x\hat{\beta} = y - x(x'x)^{-1}x'y$$
$$e = \left[I - x(x'x)^{-1}x'\right] y$$

DEFINE THE MATRIX
$$M \equiv I - x(x'x)^{-1}x'$$

THEN $e = My$

THE MATRIX M IS SOMETIMES CALLED THE <u>RESIDUAL MAKER</u>. THE MATRIX
$$P \equiv x(x'x)^{-1}x'$$
IS CALLED THE <u>PROJECTION MATRIX</u>. THE PREDICTED y VECTOR IS
$$\hat{y} = x(x'x)^{-1}x'y = x\hat{\beta}.$$
$$\hat{y} = Py$$

<u>PROPERTIES OF M:</u>

① M IS SYMMETRIC.
$$M = I - x(x'x)^{-1}x'$$
$$M' = I - x(x'x)^{-1}x' = M.$$

② $$M'M = \left[I - x(x'x)^{-1}x'\right]\left[I - x(x'x)^{-1}x'\right]$$
$$= I - x(x'x)^{-1}x' - x(x'x)^{-1}x' + x(x'x)^{-1}x'x(x'x)^{-1}x'$$
$$= I - x(x'x)^{-1}x' = M$$

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

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

THUS $MP = M(I - M)$
$$= M - M.M = M - M = 0.$$

$$e = My$$
$$= M(x\beta + u) = Mx\beta + u$$
$$\Rightarrow e = Mu$$

$$e'e = (Mu)'(Mu)$$
$$= u'M'Mu = u'Mu$$