Regression:

→ Predicting House Price, Predicting credit card score, Predict subs w.r.t. d..

Input Var : (Features/ independent var / covariate / predictors)
Output Var :- (Response)

Fixed eq forms to relate y & x.

$$y = f(x) + \epsilon \quad \leftarrow \text{error (irreducable)}$$
$$f(x) = B_0 + B_1 X$$

Linear reg estimate:
$$\hat{y} = \hat{f}(x)$$
$$\hat{y} = \hat{B_0} + \hat{B_1} X$$

Derivation
$$e = y - \hat{y}$$
$$= B_0 + B_1 X + \epsilon - (\hat{B_0} + \hat{B_1} X)$$
$$= (B_0 - \hat{B_0}) + (B_1 - \hat{B_1}) X + \epsilon$$

Residual error of sample i, $e_i = y_i - \hat{y}_i$

Sum of squared residual (RSS)
$$RSS = \sum_{i=1}^{n} e_i^2$$

$$\text{Avg min}_{B_0, B_1} : \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$\text{Avg min}_{B_0, B_1} : \sum_{i=1}^{n} (y_i - \hat{B_0} - \hat{B_1} x_i)^2 \quad - (1)$$

Differentiate w.r.t. B_0
$$2 \sum_{i=1}^{n} (y_i - \hat{B_0} - \hat{B_1} x_i)(-1) = 0$$
$$\sum (y_i - \hat{B_0} - \hat{B_1} x_i) = 0$$
$$\hat{B_0} = \sum_{i=1}^{n} (y_i - \hat{B_1} x_i)/n$$
$$= \sum \left(\frac{y_i}{n} - \hat{B_1} \frac{x_i}{n} \right)$$
$$\hat{B_0} = \bar{y} - \hat{B_1} \bar{x}$$

Diff (1) w.r.t B_1: $2 \sum_i (y_i - \hat{B_0} - \hat{B_1} x_i)(-x_i) = 0$

$$\Sigma (y_i - \hat{B_0} - \hat{B_1} x_i)(x_i)$$

$$\Sigma x_i y_i - \hat{B_0} \Sigma x_i - \hat{B_1} \Sigma x_i^2$$

$$\hat{B_0} = \Sigma \left[\frac{y_i}{n} - \hat{B_i} \frac{x_i}{n}\right] \quad \Sigma x_i y_i - \left[\frac{\Sigma y_i}{n} - \hat{B_1}\frac{\Sigma x_i}{n}\right]\Sigma x_i$$

$$\hat{B_0} = \bar{y} - \hat{B_1}\bar{x} \qquad - \hat{B_1}\Sigma x_i^2 = 0$$

$$\Rightarrow \sum_{i=}^{n} x_i y_i - \sum_{i=1}^{n} \frac{y_i}{n} \Sigma x_i + \hat{B_1} \frac{\Sigma X_i}{n} \sum_{i=1}^{n} x_i - \hat{B_1}\Sigma x_i^2 = 0$$

$$\hat{B_1} = \frac{\Sigma x_i y_i - \frac{1}{n}\Sigma y_i \Sigma x_i}{\Sigma x_i^2 - \frac{1}{n}\Sigma (x_i)^2}$$

Now, $\hat{B_0} = \bar{y} - \hat{B_1}\bar{x}$

$$\hat{B_1} = \frac{\Sigma x_i y_i - n\bar{x}\bar{y}}{\Sigma x_i^2 - n\bar{x}^2}$$

Multiple Regression

$$y = f(x) + e$$

$$f(x) = B_0 + B_1 X_1 + B_2 X_2 + \dots B_p X_p \qquad p - \text{deg of param}$$

$$\hat{y} = \hat{B_0} + \hat{B_1} X_1 + \hat{B_2} X_2 + \dots \hat{B_p} X_p$$

For n sample, no of operations $= n \times (p-1)^2$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}_{n\times1}, \quad x = \begin{bmatrix} 1 & x_{1,1} & x_{1,2} & \dots & x_{1,p} \\ 1 & x_{2,1} & x_{2,2} & \dots & x_{2,p} \\ 1 & x_{3,1} & x_{3,2} & \dots & x_{3,p} \\ 1 & x_{4,1} & x_{4,2} & & x_{4,p} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n,1} & x_{n,2} & \dots & x_{n,p} \end{bmatrix}_{n\times(p+1)} \quad B = \begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ \vdots \\ B_p \end{bmatrix}_{(p+1)\times1} \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \vdots \\ \epsilon_n \end{bmatrix}_{n\times1}$$

$$y = XB + \epsilon \quad \text{—(1)}$$

$$\hat{y} = X\hat{B}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} B_0 + B_1 x_{11} + B_2 X_{12} + \dots + B_p X_{1p} + \epsilon_1 \\ B_0 + B_1 x_{21} + B_2 X_{22} + \dots + B_p X_{2p} + \epsilon_2 \\ \\ B_0 + B_1 X_{n1} + B_2 X_{n2} + \dots + B_p X_{np} + \epsilon_n \end{bmatrix} \quad (1)$$

$$e = y - \hat{y}$$

$$e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} y_1 - \hat{y_1} \\ y_2 - \hat{y_2} \\ \vdots \\ y_n - \hat{y_n} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} \hat{y_1} \\ \hat{y_2} \\ \vdots \\ \hat{y_n} \end{bmatrix}$$

classmate
Date
Page

$(n\times1)\times(n\times1)$ ✗
$(1\times n)\times(n\times1)$ ✓

$RSS = \sum_{i=1}^{n} e_i^2 \Rightarrow RSS = e^T e$

$RSS = (y-\hat{y})^T(y-\hat{y}) = (y - X\hat{B})^T (y - X\hat{B})$

$= (y^T - \hat{B}^T X^T)(y - X\hat{B})$

$= y^T y - y^T X\hat{B} - \hat{B}^T X^T y + \hat{B}^T X^T X \hat{B}$ [order is imp.]

Matrix Diff

- $y = A \rightarrow \frac{\partial y}{\partial x} = 0$ • $y = AX \rightarrow \frac{\partial y}{\partial x} = A$
- $y = XA \rightarrow \frac{\partial y}{\partial x} = A^T$ • $y = X^T A X \rightarrow \frac{\partial y}{\partial x} = 2X^T A$

$\frac{\delta(RSS)}{\delta \hat{B}} = \frac{\delta(y^T y - y^T X\hat{B} - \hat{B} X^T y + \hat{B}^T X^T X\hat{B})}{\delta \hat{B}} = 0$

$= 0 - y^T X - (X^T y)^T + 2\hat{B} X^T X$

$= 0 - y^T X - y^T X + 2\hat{B} X^T X$

$\Rightarrow \cancel{2}\hat{B} X^T X = \cancel{2} y^T X$

$\hat{B}^T = y^T X (X^T X)^{-1}$

$\hat{B} = (X^T X)^{-1}(X^T y)$

Q. Multiple Reg:

X_1 (IQ)	X_2 (study)	Y (Score)
110	40	100
120	30	90
100	20	80
90	0	70
80	10	60

$y = \hat{b}_0 + \hat{b}_1 x_1 + \hat{b}_2 x_2$

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

$X = \begin{bmatrix} 1 & 110 & 40 \\ 1 & 120 & 30 \\ 1 & 100 & 20 \\ 1 & 90 & 0 \\ 1 & 80 & 10 \end{bmatrix}$ $X^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 110 & 120 & 100 & 90 & 80 \\ 40 & 30 & 20 & 0 & 10 \end{bmatrix}$

$$X^TX = \begin{bmatrix} 5 & 500 & 100 \\ 500 & 51000 & 10800 \\ 100 & 10800 & 3000 \end{bmatrix}; (X^TX)^{-1} = \begin{bmatrix} 101/5 & -7/30 & 1/6 \\ -7/30 & 1/360 & -1/450 \\ 1/6 & -1/450 & 1/360 \end{bmatrix}$$

$$\hat{b} = (X^TX)^{-1} X^Ty$$

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 101/5 & -7/30 & 1/6 \\ -7/30 & 1/360 & -1/450 \\ 1/6 & -1/450 & 1/360 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 110 & 120 & 100 & 90 & 80 \\ 40 & 30 & 20 & 0 & 10 \end{bmatrix} \begin{bmatrix} 100 \\ 90 \\ 80 \\ 70 \\ 60 \end{bmatrix}$$

$$\Rightarrow \hat{b} = \begin{bmatrix} 20 \\ 0.5 \\ 0.5 \end{bmatrix}$$

$$\therefore \hat{b} = 20 + 0.5X_1 + 0.5X_2$$

Q: Use the following data to fit the linear regression model.

	wt (X)	hgt (Y)	XY	X^2	Y^2
	140	60	8400	19600	3600
	155	62	9610	24025	3844
	159	67	10653	25281	4489
	179	70	12530	32041	4900
	192	71	13632	36864	5041
	200	72	14400	40000	5184
	212	75	15900	44944	5625
Σ	1237	477	85125	222755	32683

$$\hat{y} = b_0 + b_1 X$$

$$b_0 = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$$

$$= \frac{477 \times 222755 - 1237 \times 85125}{7 \times 222755 - 1237^2} = \frac{954510}{29116} = 32.78$$

$$b_1 = \frac{n(\Sigma xy) - \Sigma x \Sigma y}{n(\Sigma x^2) - (\Sigma x)^2}$$

$$= \frac{7 \times (85125) - 1237 \times 477}{7 \times 222755 - 1237^2} = 0.2001$$

$$\hat{y} = 32.78 + 0.2001X.$$

classmate
Date
Page

Q.

	0	1	2	3	4
x(year)	2005	2006	2007	2008	2009
y (sales)	12	19	29	37	45

The sales of a cmpy (in million dollars) for each year.

a) Find the least sq. regression line $y = ax + b$

b) Use the least " " " as a model to estimate the sales of the cmpy in 2012.

Sol. For simplification we can take years as $t = x - 2005$

$\therefore$

t	y	ty	t^2
0	12	0	0
1	19	19	1
2	29	58	4
3	37	111	9
4	45	180	16
Σ 10	142	368	30

$$b_1 = \frac{n\Sigma ty - \Sigma t \Sigma y}{n\Sigma t^2 - (\Sigma t)^2}$$

$$\therefore \frac{5 \times 368 - 10 \times 142}{5 \times 30 - 100} = \frac{420}{50} = 8.4$$

$$b_0 = \frac{1}{n}(\Sigma y - b_0 \Sigma x)$$

$$= \frac{1}{5} \times (142 - 8.4 \times 10) = 11.6$$

$\hat{y} = b + ax$

$$\hat{y} = \cancel{8.4 + 11.6x}$$
$$= 8.4 \times 7 + 11.6$$

$t = 2012 - 2005 = 7$

Q Estimate the line fit for multiple regression.

y	x_1	x_2
140	60	22
155	62	25
159	67	24
179	70	20
192	71	15
200	72	14
212	75	14
215	78	11

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2$$

$$b_1 = \frac{(\Sigma x_2^2)(\Sigma x_1 y) - (\Sigma x_1 x_2)(\Sigma x_2 y)}{(\Sigma x_1^2)(\Sigma x_2^2) - (\Sigma x_1 x_2)^2}$$

$$b_2 = \frac{\Sigma x_1^2 \Sigma x_2 y - \Sigma x_1 x_2 \Sigma x_1 y}{\Sigma x_1^2 \Sigma x_2^2 - (\Sigma x_1 x_2)^2}$$

$$b_0 = \bar{y} - b_1 \bar{x}_1 - b_2 \bar{x}_2$$