

# AI1110: Assignment 2

SADINENI ABHINAY - CS21BTECH11055

## ICSE class 12 paper 2018

### I. QUESTION 20(A)

Find the line of regression of  $y$  on  $x$  from the following table.

x	1	2	3	4	5
y	7	6	5	4	3

Hence, estimate the  $y$  value when  $x=6$ .

### Solution.

Given observations

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \dots, \begin{pmatrix} x_n \\ y_n \end{pmatrix} \quad (1)$$

Best fit a straight line to it,  $e_i$  are the corresponding residual errors

$$y = a_0 + a_1 x \quad (2)$$

$$e_i = y_i - (a_0 + a_1 x_i) \quad (3)$$

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix}, \mathbf{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ \vdots & \vdots \\ 1 & x_{n-1} \\ 1 & x_n \end{pmatrix}, \mathbf{A} = \begin{pmatrix} a_1 \\ a_0 \end{pmatrix}, \mathbf{E} = \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_{n-1} \\ e_n \end{pmatrix} \quad (4)$$

$$\mathbf{Y} = \mathbf{AX} + \mathbf{E} \quad (5)$$

SSE is sum of square of errors, for minimizing use gradient wrt  $\mathbf{A}$ , substitute  $\mathbf{E}$  from above equation and get  $\mathbf{A}$

$$SSE = \mathbf{E}^T \mathbf{E} = (\mathbf{Y} - \mathbf{AX})^T (\mathbf{Y} - \mathbf{AX}) \quad (6)$$

$$\nabla SSE = \frac{2}{n} (\mathbf{X}^T \mathbf{XA} - \mathbf{X}^T \mathbf{Y}) = 0 \quad (7)$$

$$\Rightarrow (\mathbf{X}^T \mathbf{XA} - \mathbf{X}^T \mathbf{Y}) = 0 \quad (8)$$

$$\Rightarrow \mathbf{A} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (9)$$

If we want in line form after finding  $\mathbf{A}$

$$\mathbf{n}^T \mathbf{x} = c, \mathbf{n} = \begin{pmatrix} \mathbf{A}^T \begin{pmatrix} -1 \\ 0 \end{pmatrix} \\ 1 \end{pmatrix}, c = \mathbf{A}^T \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (10)$$

For this problem

$$\begin{pmatrix} x \\ y \end{pmatrix} : \begin{pmatrix} 1 \\ 7 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 4 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 3 \end{pmatrix} \quad (11)$$

$$\mathbf{Y} = \begin{pmatrix} 7 \\ 6 \\ 5 \\ 4 \\ 3 \end{pmatrix}, \mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{pmatrix} \quad (12)$$

Substitute these  $\mathbf{Y}$  and  $\mathbf{X}$  and get  $\mathbf{A}$

$$\mathbf{A} = \begin{pmatrix} -1 \\ 8 \end{pmatrix} \quad (13)$$

After using the line form stated before

$$y = 8 - x \quad (14)$$

$$x + y = 8 \quad (15)$$

When  $x = 6$  then  $y$  must be 2 from the line of regression.

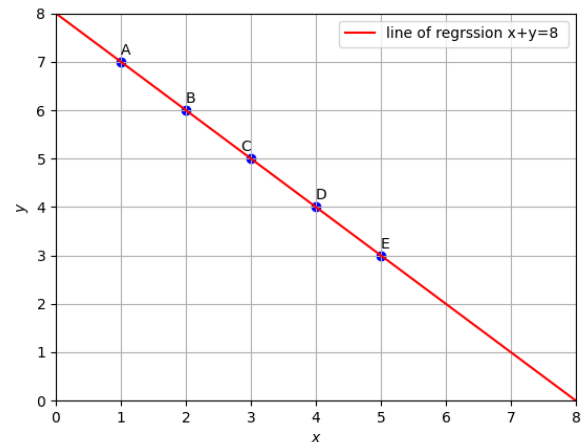


Fig. 0: plot of all points