

ICSE class 12 paper 2018

1 QUESTION 20(A)

Find the line of regression of y on x from the following table. Hence, estimate the y value when x=6.

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

TABLE I

Solution. Given observations

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \dots, \begin{pmatrix} x_n \\ y_n \end{pmatrix} \quad (1.0.1)$$

Best fit a straight line to it, e_i are the corresponding residual errors, coefficients a_0 and a_1

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix}, \mathbf{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_{n-1} \\ 1 & x_n \end{pmatrix} \quad (1.0.2)$$

$$\mathbf{A} = \begin{pmatrix} a_1 \\ a_0 \end{pmatrix}, \mathbf{E} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_{n-1} \\ e_n \end{pmatrix} \quad (1.0.3)$$

$$\mathbf{E} = \mathbf{Y} - \mathbf{XA} \quad (1.0.4)$$

SSE is sum of square of errors, substitute E from above equation and get A

$$SSE = \|\mathbf{E}^\top \mathbf{E}\| \quad (1.0.5)$$

$$= (\mathbf{Y} - \mathbf{XA})^\top (\mathbf{Y} - \mathbf{XA}) \quad (1.0.6)$$

$$= (\mathbf{Y}^\top - \mathbf{A}^\top \mathbf{X}^\top) (\mathbf{Y} - \mathbf{XA}) \quad (1.0.7)$$

$$= (\mathbf{Y}^\top \mathbf{Y} - 2\mathbf{A}^\top \mathbf{X}^\top \mathbf{Y} + \mathbf{A}^\top \mathbf{X}^\top \mathbf{XA}) \quad (1.0.8)$$

for minimizing use gradient wrt A

$$\nabla SSE = (\nabla \mathbf{Y}^\top \mathbf{Y} - 2\nabla \mathbf{A}^\top \mathbf{X}^\top \mathbf{Y} + \nabla \mathbf{A}^\top \mathbf{X}^\top \mathbf{XA}) \quad (1.0.9)$$

$$= 2(\mathbf{X}^\top \mathbf{XA} - \mathbf{X}^\top \mathbf{Y}) \quad (1.0.10)$$

We now set this to zero at the optimum

$$(\mathbf{X}^\top \mathbf{XA} - \mathbf{X}^\top \mathbf{Y}) = 0 \quad (1.0.11)$$

$$\mathbf{A} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y} \quad (1.0.12)$$

If we want to give line form after finding A

$$\mathbf{n}^\top \mathbf{x} = c, \mathbf{n} = \begin{pmatrix} \mathbf{A}^\top \begin{pmatrix} -1 \\ 0 \end{pmatrix} \\ 1 \end{pmatrix}, c = \mathbf{A}^\top \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (1.0.13)$$

For this problem

$$\begin{pmatrix} x \\ y \end{pmatrix} : \begin{pmatrix} 1 \\ 7 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 4 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 3 \end{pmatrix} \quad (1.0.14)$$

$$\mathbf{Y} = \begin{pmatrix} 7 \\ 6 \\ 5 \\ 4 \\ 3 \end{pmatrix}, \mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{pmatrix} \quad (1.0.15)$$

Substitute these Y and X and get A

$$\mathbf{A} = \begin{pmatrix} -1 \\ 8 \end{pmatrix} \quad (1.0.16)$$

After using equations before we will get normal vector and parameter then we can write the line form

$$\mathbf{n} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, c = 8 \quad (1.0.17)$$

$$(1 \ 1) \begin{pmatrix} x \\ y \end{pmatrix} = 8 \quad (1.0.18)$$

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

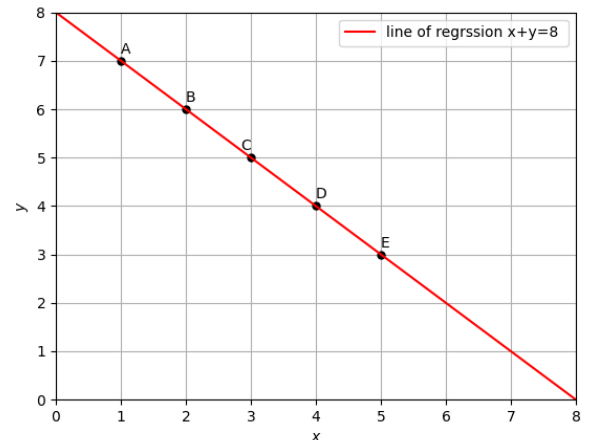


Fig. 0. plot of all points