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

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}....., \begin{pmatrix} x_n \\ y_n \end{pmatrix} \tag{1}$$

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

$$y = a_0 + a_1 x \tag{2}$$

$$e_i = y_i - (a_0 + a_1 x_i) \tag{3}$$

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ y3 \\ . \\ . \\ y_{n-1} \\ y_n \end{pmatrix}, \mathbf{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ . & \\ . & \\ 1 & x_{n-1} \\ 1 & x_n \end{pmatrix}, \mathbf{A} = \begin{pmatrix} a_1 \\ a_0 \end{pmatrix}, \mathbf{E} = \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ . \\ . \\ e_{n-1} \\ e_n \end{pmatrix} \tag{4}$$

This gives us the equation

$$\mathbf{Y} = \mathbf{AX} + \mathbf{E} \tag{5}$$

SSE is sum of square of errors ,for minimizing use gradient wrt to A, substitue E from above equation

$$SSE = \mathbf{E}^\top \mathbf{E} \tag{6}$$

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

After minimizing SSE value we will get

$$\mathbf{A} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y} \tag{8}$$

If we want in line form after finding A

$$\mathbf{n}^\top \mathbf{x} = c, \mathbf{n} = \begin{pmatrix} \mathbf{A}^\top \begin{pmatrix} -1 \\ 0 \end{pmatrix} \\ 1 \end{pmatrix}, c = \mathbf{A}^\top \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{9}$$

For this problem

$$\begin{pmatrix} x \\ y \end{pmatrix} : \begin{pmatrix} 1 \\ 7 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 4 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 3 \end{pmatrix} \tag{10}$$

$$\mathbf{Y} = \begin{pmatrix} 7 \\ 6 \\ 5 \\ 4 \\ 3 \end{pmatrix}, \mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{pmatrix} \tag{11}$$

Substitue these Y and X and get A

$$A = \begin{pmatrix} -1 \\ 8 \end{pmatrix} \tag{12}$$

After using the line form stated before

$$y = 8 - x \tag{13}$$

$$x + y = 8 \tag{14}$$

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

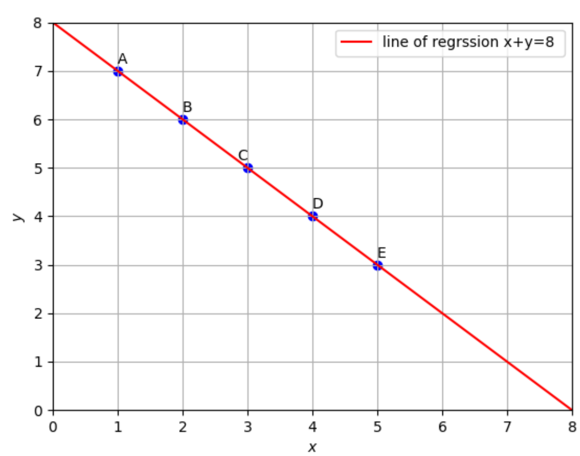

Fig. 0: plot of all points