

① Linear Regression

$$f(x) = ax + b \quad a, b \text{ constants}$$

$$\min_{a, b} E = \sum_{i=1}^n (y_i - ax_i - b)^2$$

From calculus:

$$\frac{dE}{da} = \sum_{i=1}^n 2(y_i - ax_i - b)(-x_i) = 0$$

$$\Rightarrow -2 \sum_{i=1}^n (x_i y_i - ax_i^2 - bx_i) = 0$$

$$\Rightarrow \sum_{i=1}^n x_i y_i - a \sum_{i=1}^n x_i^2 - b \sum_{i=1}^n x_i = 0$$

① — $a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i$

$$\frac{dE}{db} = \sum_{i=1}^n 2(y_i - ax_i - b)(-1) = 0$$

$$\frac{df}{db} = \sum_{i=1}^n 2(y_i - ax_i - b)(-1) = 0$$

$$= -2 \left[\sum_{i=1}^n y_i - a \sum_{i=1}^n x_i - bn \right] = 0$$

$$\textcircled{2} - a \sum_{i=1}^n x_i + bn = \sum_{i=1}^n y_i$$

Solve for a, b from $\textcircled{1}, \textcircled{2}$

$$b = \frac{(\sum x_i y_i)(\sum x_i) - (\sum x_i^2)(\sum y_i)}{(\sum x_i)^2 - n(\sum x_i^2)}$$

$$a = \frac{\sum x_i \sum y_i - n \sum x_i y_i}{(\sum x_i)^2 - n(\sum x_i^2)}$$

Fit a straight line to the data shown on the right

	x	y
/	-2.0000	-2.7664
/	-1.6000	-1.8747
/	-1.2000	-2.7127
/	-0.8000	-1.7671
/	-0.4000	0.1816
/	0.0000	1.0231
/	0.4000	2.0852
/	0.8000	1.3998
/	1.2000	3.6906
/	1.6000	3.0207
/	2.0000	4.9748

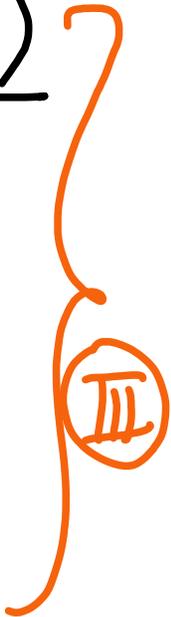
Solution

$$y = f(x) = ax + b$$

$$b = \frac{(\sum x_i y_i)(\sum x_i) - (\sum x_i^2)(\sum y_i)}{(\sum x_i)^2 - n(\sum x_i^2)}$$

$$a = \frac{\sum x_i \sum y_i - n \sum x_i y_i}{(\sum x_i)^2 - n(\sum x_i^2)}$$

$$n = 11 \text{ (# data points)}$$



x_i	y_i	x_i^2	$x_i y_i$
<u>-2</u>	-2.7664	4	5.5328
-1.6	-1.8747	2.56	2.9995
-1.2	-2.7127	1.44	3.2552
-0.8	-1.7671	0.64	1.4137
-0.4	0.1816	0.16	-0.0726
0	1.0231	0	0
0.4	2.0852	0.16	0.8341
0.8	1.3998	0.64	1.1199
1.2	3.6906	1.44	4.4287
1.6	3.0207	2.56	4.8331
<u>2.0</u>	4.9748	4.0	9.9496
0	7.255	17.6	34.2939

$$\sum x_i = 0$$

$$\sum y_i = 7.255$$

$$\sum x_i y_i = 34.2939$$

$$\sum x_i^2 = 17.6$$

$$n = 11$$

Substitute in (III)

$$a = 1.9485$$

$$b = 0.6595$$

$$y = ax + b = 1.9485x + 0.6595$$