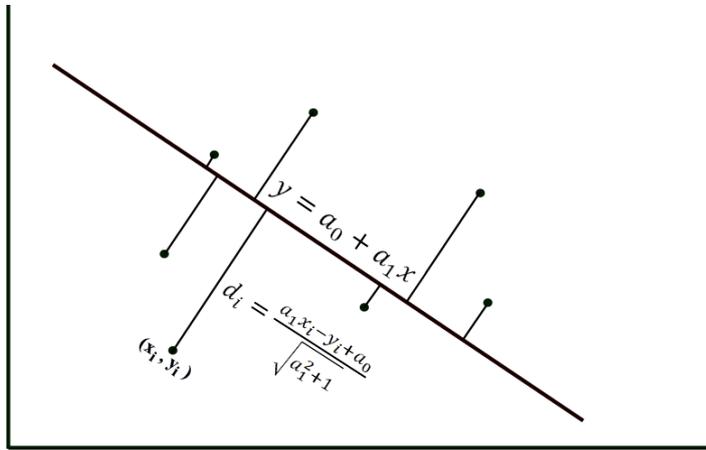


Optimal Spaceway

Solution Sketch:

Problem type: Modified Regression, Partial Differentiation, fundamentals of coordinate geometry.



This problem boils down to: Given some points, find the straight line for which the summation of square of distance from all the points to this line is minimum. This problem is a lot similar to finding the least square regression line (Linear regression), but it is a lot harder. In linear regression the distance is measured along y-axis but here the distance is measured along the perpendicular of the line found. But here we also need to use partial

differentiation like we needed in linear regression to find the desired line.

Suppose the coordinates of the given points are: $(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)$ and let's assume that the equation of the desired line is $y = a_0 + a_1 x \Rightarrow a_1 x - y + a_0 = 0$.

From a point (x_i, y_i) the distance of this straight line is $= \frac{(a_1 x_i - y_i + a_0)}{\sqrt{a_1^2 + 1}}$

So summation of square of distance from all the points is $= \sum_{i=1}^n \left(\frac{(a_1 x_i - y_i + a_0)}{\sqrt{a_1^2 + 1}} \right)^2 = \sum_{i=1}^n \frac{(a_1 x_i - y_i + a_0)^2}{a_1^2 + 1}$

So here we have to find the value of a_0 and a_1 , for which $\sum_{i=1}^n \frac{(a_1 x_i - y_i + a_0)^2}{a_1^2 + 1}$ is minimum. To do this we have to use partial differentiation.

Lets assume that $S_r = \sum_{i=1}^n \frac{(a_1 x_i - y_i + a_0)^2}{a_1^2 + 1}$

Now $\frac{\partial S_r}{\partial a_0} = \frac{2 \sum_{i=1}^n (a_1 x_i - y_i + a_0) * 1}{a_1^2 + 1} = 0$ //First derivative is zero at minima

Or, $na_0 + a_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \dots (i)$ //Skipped some steps

Multiplying both sides of (i) with $a_0 a_1$ we get:

$na_0^2 a_1 + a_1^2 a_0 \sum_{i=1}^n x_i - a_0 a_1 \sum_{i=1}^n y_i = 0 \dots (ii)$

Now again, $S_r = \sum_{i=1}^n \frac{(a_1 x_i - y_i + a_0)^2}{a_1^2 + 1} \dots (iii)$

Or, $\frac{\partial S_r}{\partial a_1} = \frac{(a_1^2 + 1)2 \sum_{i=1}^n (a_1 x_i - y_i + a_0)(x_i) - \sum_{i=1}^n (a_1 x_i - y_i + a_0)^2 * 2a_1}{a_1^2 + 1} = 0$

//First derivative is zero at minima

Or, $(a_1^2 + 1)2 \sum_{i=1}^n (a_1 x_i - y_i + a_0)(x_i) - \sum_{i=1}^n (a_1 x_i - y_i + a_0)^2 * 2a_1 = 0$

After some manipulation we get from this:

$\sum_{i=1}^n x_i^2 a_1 + \sum_{i=1}^n x_i y_i (a_1^2 - 1) - a_1 \sum_{i=1}^n y_i^2 + \sum_{i=1}^n x_i (a_0 - \underline{a_1^2 a_0}) + \underline{2a_0 a_1} \sum_{i=1}^n y_i = \underline{na_0^2 a_1} \dots (iv)$

But from (ii) we know that $na_0^2 a_1 + a_1^2 a_0 \sum_{i=1}^n x_i - a_0 a_1 \sum_{i=1}^n y_i = 0$

But these terms are wholly or partially present in (iv) (underlined), so we can remove them. After that we get:

$\sum_{i=1}^n x_i^2 a_1 + \sum_{i=1}^n x_i y_i (a_1^2 - 1) - a_1 \sum_{i=1}^n y_i^2 + \sum_{i=1}^n x_i a_0 + a_0 a_1 \sum_{i=1}^n y_i = 0 \dots (v)$

From these relations we can get the expressions to find the value of a_0 and a_1