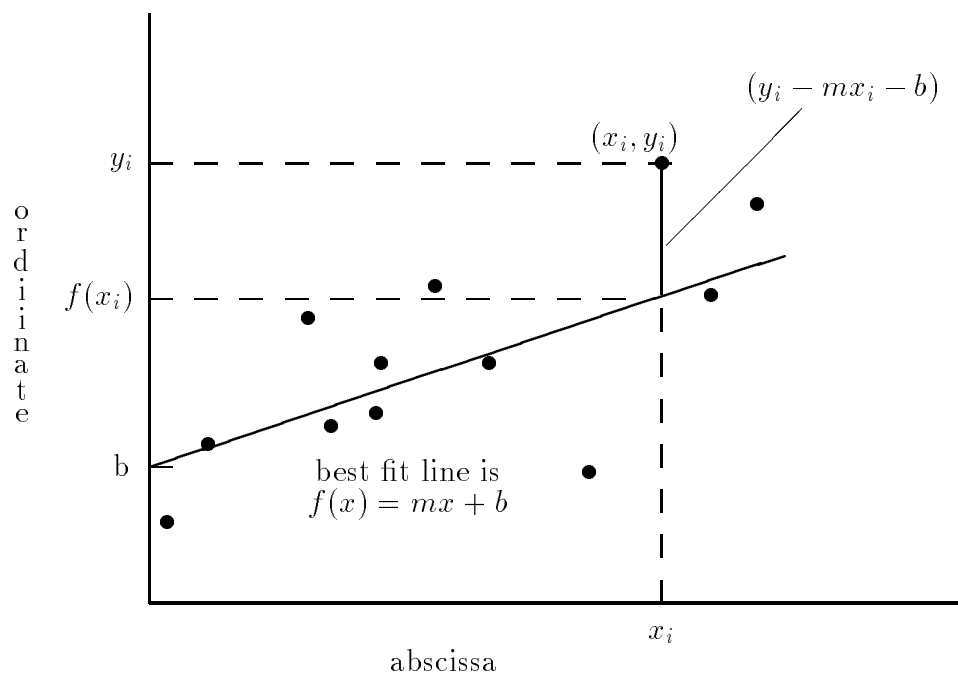


Taken from Calculus, eighth edition, by Varberg, Purcell, Rigdon, page 675.

Given n points $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ in the xy -plane, we wish to find the line $y = mx + b$ such that the sum of the squares of the vertical distances from the points to the line is a minimum; that is, we wish to minimize

$$E(m, b) = \sum_{i=1}^n (y_i - mx_i - b)^2. \quad (1)$$

Remember that the x_i 's and the y_i 's are fixed; they are not unknowns, they are the data values.



Taking derivatives and setting them equal to zero is how we minimize things. Setting $\partial E / \partial m$ and $\partial E / \partial b$ equal to zero leads to the system of two simultaneous equations for m and b —the two unknowns in $f(x) = mx + b$. This system is (the student should verify this):

$$m \sum x_i^2 + b \sum x_i = \sum y_i x_i \quad (2)$$

$$m \sum x_i + nb = \sum y_i. \quad (3)$$

The second partials test will show that f is minimized for this choice of m and b .

From Calculus, third edition, by Staruss, Bradley, and Smith, page 757.

$$m = \frac{n \sum x_i y_i - (\sum x_i) (\sum y_i)}{n \sum x_i^2 - (\sum x_i)^2} \quad (4)$$

and

$$b = \frac{\sum x_i^2 \sum y_i - (\sum x_i) (\sum x_i y_i)}{n \sum x_i^2 - (\sum x_i)^2} \quad (5)$$

It can be shown that these values of m and b yield an absolute minimum for $f(m, b)$.