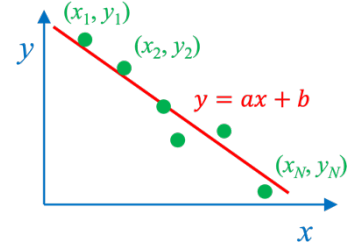


## Least Square Fit of a Line

**Problem:** Given a set of  $N$  pairs of numbers,  $\{(x_i, y_i) \mid i = 1, \dots, N\}$ , what is the best linear fit,  $y = ax + b$ , in the sense that it minimizes the square error,

$$S = \sum_{i=1}^N \left( \underbrace{ax_i + b}_{\text{prediction}} - \underbrace{y_i}_{\text{measured}} \right)^2 ?$$



**Answer:**  $S$  is a quadratic function of both  $a$  and  $b$ , and it becomes  $+\infty$  for  $a \rightarrow \pm\infty$  or  $b \rightarrow \pm\infty$ . There is a unique combination of  $a$  and  $b$ , at which  $S$  takes the minimum value and its derivatives with respect to  $a$  and  $b$  are zero, i.e.,

$$\begin{cases} \frac{\partial S}{\partial a} = 2 \sum_{i=1}^N (ax_i + b - y_i) x_i = 0 \\ \frac{\partial S}{\partial b} = 2 \sum_{i=1}^N (ax_i + b - y_i) = 0 \end{cases}$$

This is a set of linear equations,

$$\begin{cases} \left( \sum_{i=1}^N x_i^2 \right) a + \left( \sum_{i=1}^N x_i \right) b = \sum_{i=1}^N x_i y_i \\ \left( \sum_{i=1}^N x_i \right) a + Nb = \sum_{i=1}^N y_i \end{cases}$$

which, in the matrix notation, becomes

$$\begin{bmatrix} \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i \\ \sum_{i=1}^N x_i & N \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N x_i y_i \\ \sum_{i=1}^N y_i \end{bmatrix}$$

The solution is

$$\begin{aligned} \begin{bmatrix} a \\ b \end{bmatrix} &= \begin{bmatrix} \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i \\ \sum_{i=1}^N x_i & N \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^N x_i y_i \\ \sum_{i=1}^N y_i \end{bmatrix} = \frac{1}{\sum_{i=1}^N x_i^2 N - (\sum_{i=1}^N x_i)^2} \begin{bmatrix} N & -\sum_{i=1}^N x_i \\ -\sum_{i=1}^N x_i & \sum_{i=1}^N x_i^2 \end{bmatrix} \begin{bmatrix} \sum_{i=1}^N x_i y_i \\ \sum_{i=1}^N y_i \end{bmatrix} \\ &= \frac{1}{\sum_{i=1}^N x_i^2 N - (\sum_{i=1}^N x_i)^2} \begin{bmatrix} N \sum_{i=1}^N x_i y_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i \\ -\sum_{i=1}^N x_i \sum_{i=1}^N x_i y_i + \sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i \end{bmatrix} \end{aligned}$$

or

$$\begin{cases} a = \frac{N \sum_{i=1}^N x_i y_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i}{\sum_{i=1}^N x_i^2 N - (\sum_{i=1}^N x_i)^2} \\ b = \frac{-\sum_{i=1}^N x_i \sum_{i=1}^N x_i y_i - \sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i}{\sum_{i=1}^N x_i^2 N - (\sum_{i=1}^N x_i)^2} \end{cases}$$

