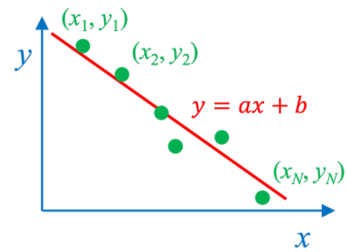
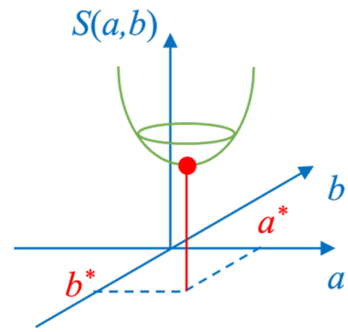
Least Square Fit of a Line

Problem: Given a set of N pairs of numbers, $\{(x_i, y_i) \mid i = 1,\ldots, N\}$, 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 $a \to \pm\infty$ or $b \to \pm\infty$. 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.*,



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

This is a set of linear equations,

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

which, in the matrix notation, becomes

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

The solution is

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

$$= \frac{1}{\sum_{i=1}^{N} x_i^2 N - \left(\sum_{i=1}^{N} x_i\right)^2} \begin{bmatrix} N \sum_{i=1}^{N} x_i y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i \\ -\sum_{i=1}^{N} x_i \sum_{i=1}^{N} x_i y_i + \sum_{i=1}^{N} x_i^2 \sum_{i=1}^{N} y_i \end{bmatrix}$$

or

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