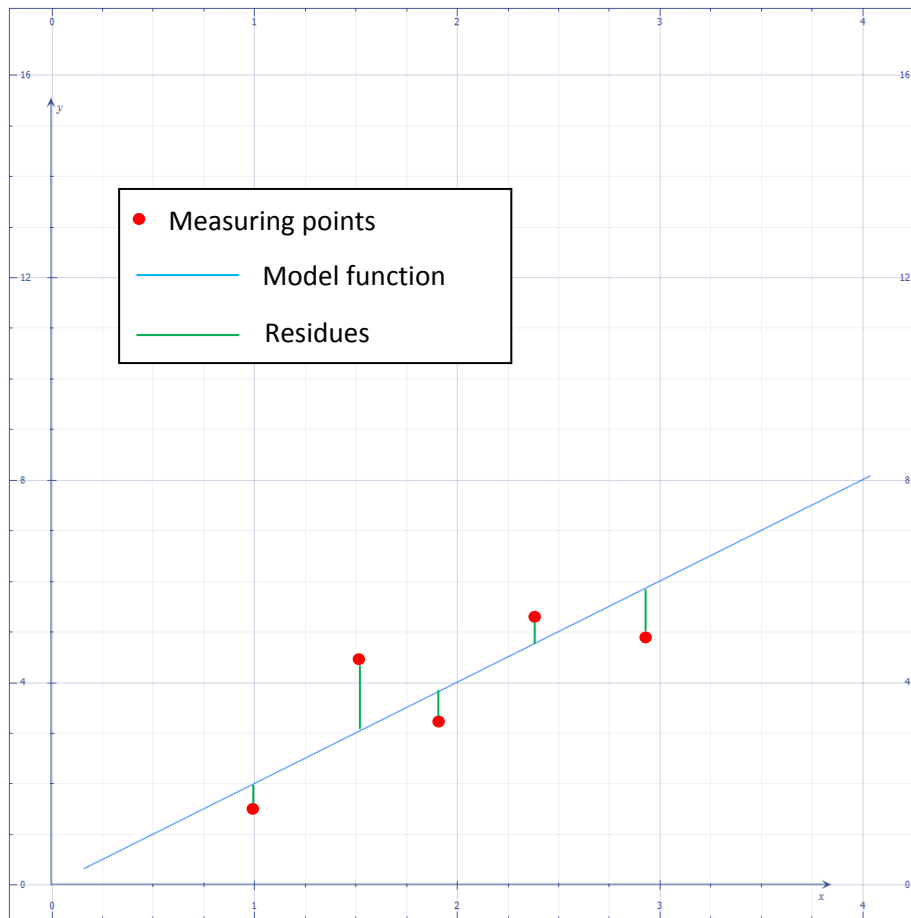


Supervised machine learning

In supervised machine learning, each example is a pair consisting of an input value x and a desired output value $f(x)$. A supervised learning algorithm analyzes the training data and produces an inferred function. The training data are given by a 'teacher'.

Least squares - special case first order polynomial

The graph below shows measuring points (learned input and output values) and its distance (residues) to a function computed by the method of least squares (line).



The model function with two linear parameters is given by

$$f(x) = \alpha_0 + \alpha_1 \cdot x$$

For the n learned input and output value pairs $(x_1, y_1), \dots, (x_n, y_n)$ we are searching now for the parameters α_0 and α_1 of the best fitting line. The according residues between the wanted line and the input and output value pairs are computed by:

$$\begin{aligned} r_1 &= \alpha_0 + \alpha_1 \cdot x_1 - y_1 \\ r_2 &= \alpha_0 + \alpha_1 \cdot x_2 - y_2 \\ &\vdots \\ &\vdots \\ &\vdots \\ r_n &= \alpha_0 + \alpha_1 \cdot x_n - y_n \end{aligned}$$

Squaring and summation of the residues yields to:

$$r_1^2 + r_2^2 + \dots + r_n^2 = (\alpha_0 + \alpha_1 \cdot x_1 - y_1)^2 + (\alpha_0 + \alpha_1 \cdot x_2 - y_2)^2 + \dots + (\alpha_0 + \alpha_1 \cdot x_n - y_n)^2$$

$$\sum_{i=1}^n r_i^2 = (\alpha_0 + \alpha_1 \cdot x_1 - y_1)^2 + (\alpha_0 + \alpha_1 \cdot x_2 - y_2)^2 + \dots + (\alpha_0 + \alpha_1 \cdot x_n - y_n)^2$$

The square of a trinomial can be written as:

$$(a + b + c)^2 = a^2 + b^2 + c^2 + 2 \cdot a \cdot b + 2 \cdot a \cdot c + 2 \cdot b \cdot c$$

Therefore:

$$(\alpha_0 + \alpha_1 \cdot x_i - y_i)^2 = \alpha_0^2 + \alpha_1^2 \cdot x_i^2 + y_i^2 + 2 \cdot \alpha_0 \cdot \alpha_1 \cdot x_i - 2 \cdot \alpha_0 \cdot y_i - 2 \cdot \alpha_1 \cdot x_i \cdot y_i$$

$$\sum_{i=1}^n r_i^2 = \sum_{i=1}^n (\alpha_0^2 + \alpha_1^2 \cdot x_i^2 + y_i^2 + 2 \cdot \alpha_0 \cdot \alpha_1 \cdot x_i - 2 \cdot \alpha_0 \cdot y_i - 2 \cdot \alpha_1 \cdot x_i \cdot y_i)$$

We consider the sum function now as a function of the two variables α_0 and α_1 (the values of the learned data set are just considered as constants) and compute the partial derivative of the two variables and its zeros.

For α_0 :

$$\sum_{i=1}^n r_i^2 (\alpha_0) = \sum_{i=1}^n (2 \cdot \alpha_0 + 2 \cdot \alpha_1 \cdot x_i - 2 \cdot y_i)$$

$$\sum_{i=1}^n (2 \cdot \alpha_0 + 2 \cdot \alpha_1 \cdot x_i - 2 \cdot y_i) = 0$$

$$\sum_{i=1}^n (\alpha_0 + \alpha_1 \cdot x_i - y_i) = 0$$

$$n \cdot \alpha_0 + \sum_{i=1}^n (\alpha_1 \cdot x_i - y_i) = 0$$

$$n \cdot \alpha_0 + \alpha_1 \cdot \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$

For α_1 :

$$\sum_{i=1}^n r_i^2 (\alpha_1) = \sum_{i=1}^n (2 \cdot \alpha_1 \cdot x_i^2 + 2 \cdot \alpha_0 \cdot x_i - 2 \cdot x_i \cdot y_i)$$

$$\sum_{i=1}^n (2 \cdot \alpha_1 \cdot x_i^2 + 2 \cdot \alpha_0 \cdot x_i - 2 \cdot x_i \cdot y_i) = 0$$

$$\sum_{i=1}^n (\alpha_1 \cdot x_i^2 + \alpha_0 \cdot x_i - x_i \cdot y_i) = 0$$

$$\alpha_0 \cdot \sum_{i=1}^n x_i + \alpha_1 \cdot \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i \cdot y_i$$

Solving the system of linear equations:

$$n \cdot \alpha_0 + \alpha_1 \cdot \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$

$$\alpha_0 = \frac{1}{n} \cdot \left(\sum_{i=1}^n y_i \right) - \alpha_1 \cdot \frac{1}{n} \cdot \sum_{i=1}^n x_i$$

Substituting $\frac{1}{n} \cdot \sum_{i=1}^n y_i = \bar{y}$ as the arithmetic mean of the y -values and $\frac{1}{n} \cdot \sum_{i=1}^n x_i = \bar{x}$ of the x -values:

$$\alpha_0 = \bar{y} - \alpha_1 \cdot \bar{x}$$

Replacing α_0 in following equation

$$\alpha_0 \cdot \left(\sum_{i=1}^n x_i \right) + \alpha_1 \cdot \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i \cdot y_i$$

yields to

$$(\bar{y} - \alpha_1 \cdot \bar{x}) \cdot \left(\sum_{i=1}^n x_i \right) + \alpha_1 \cdot \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i \cdot y_i$$

$$\bar{y} \cdot \left(\sum_{i=1}^n x_i \right) + \alpha_1 \cdot \left(\sum_{i=1}^n x_i^2 \right) - \alpha_1 \cdot \bar{x} \cdot \sum_{i=1}^n x_i = \sum_{i=1}^n x_i \cdot y_i$$

$$\alpha_1 \cdot \left(\sum_{i=1}^n x_i^2 \right) - \alpha_1 \cdot \bar{x} \cdot \sum_{i=1}^n x_i = \left(\sum_{i=1}^n x_i \cdot y_i \right) - \bar{y} \cdot \sum_{i=1}^n x_i$$

$$\alpha_1 = \frac{(\sum_{i=1}^n x_i \cdot y_i) - \bar{y} \cdot \sum_{i=1}^n x_i}{(\sum_{i=1}^n x_i^2) - \bar{x} \cdot \sum_{i=1}^n x_i}$$

Because $\frac{1}{n} \cdot \sum_{i=1}^n x_i = \bar{x}$ and $\sum_{i=1}^n x_i = n \cdot \bar{x}$:

$$\alpha_1 = \frac{(\sum_{i=1}^n x_i \cdot y_i) - n \cdot \bar{x} \cdot \bar{y}}{(\sum_{i=1}^n x_i^2) - n \cdot (\bar{x})^2}$$

For a new input value $x_n + 1$ the learned output value is

$$y(x_{n+1}) = \bar{y} - \alpha_1 \cdot \bar{x} + \alpha_1 \cdot x_{n+1}$$

$$y(x_{n+1}) = \bar{y} - \alpha_1 \cdot (\bar{x} - x_{n+1})$$

$$y(x_{n+1}) = \bar{y} - \frac{(\sum_{i=1}^n x_i \cdot y_i) - n \cdot \bar{x} \cdot \bar{y}}{(\sum_{i=1}^n x_i^2) - n \cdot (\bar{x})^2} \cdot (\bar{x} - x_{n+1})$$