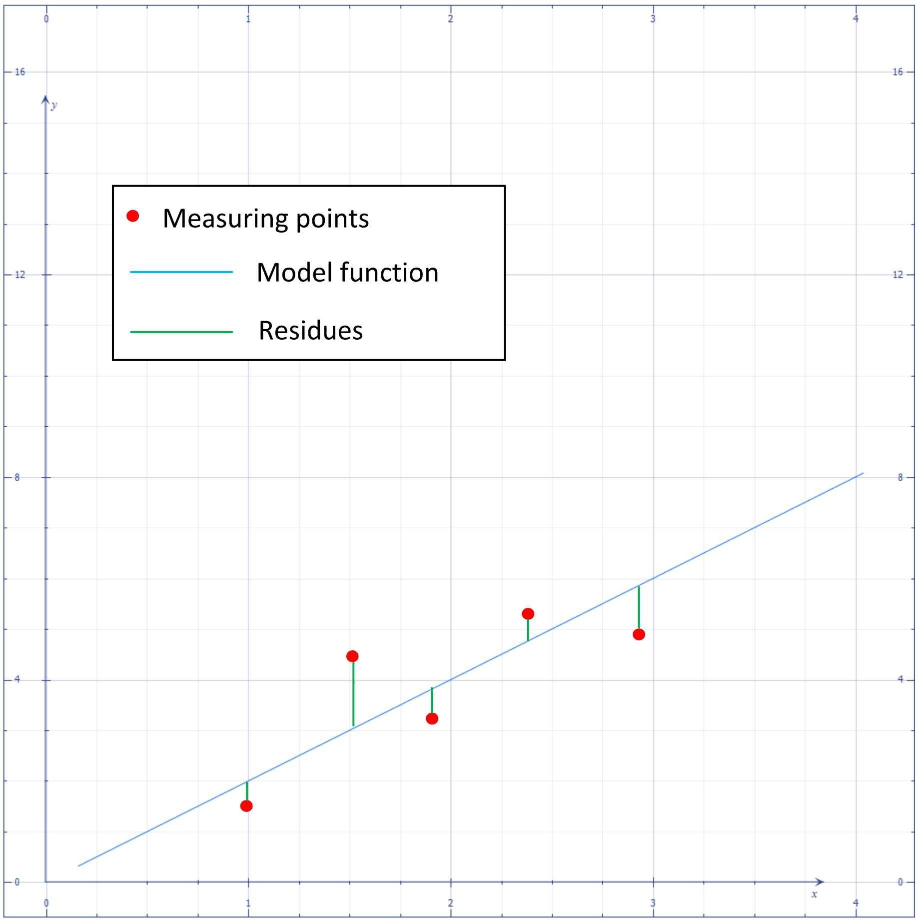
Author: M. Bindhammer, 2013

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:

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

Squaring and summation of the residues yields to:

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

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

The square of a trinomial can be written as:

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

Therefore:

$$(\alpha_0 + \alpha_1 \cdot x_\mathrm{i} - y_\mathrm{i})^2 = \alpha_0^2 + \alpha_1^2 \cdot x_i^2 + y_i^2 + 2 \cdot \alpha_0 \cdot \alpha_1 \cdot x_i - 2 \cdot \alpha_0 \cdot y_i - 2 \cdot \alpha_1 \cdot x_i \cdot y_i$$

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

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 a_0:

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

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

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

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

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

For a_1:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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