

最小二乗法

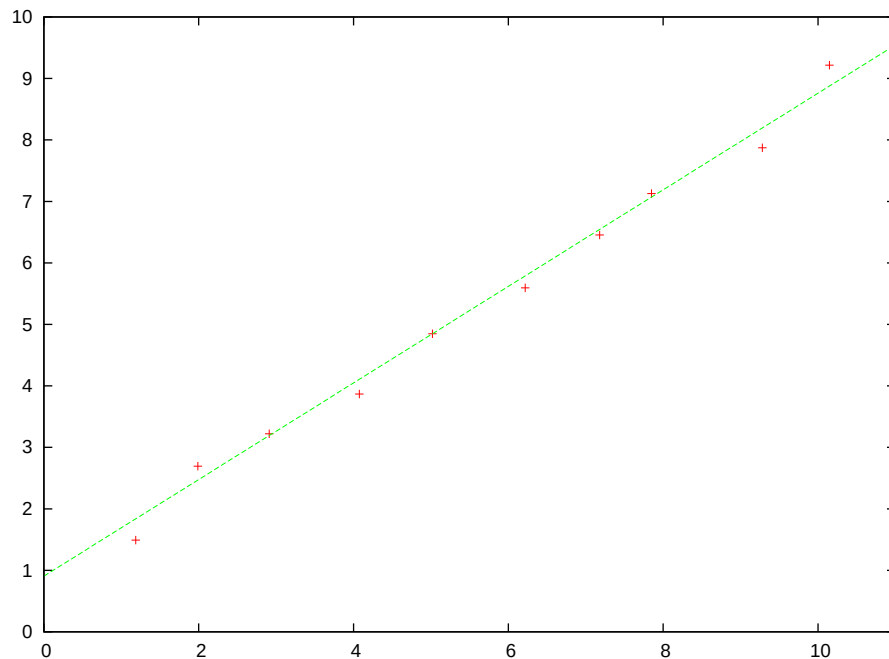
次のようなデータの組 (x_i, y_i) が与えられたとする。

x	1.186	1.989	2.911	4.074	5.019	6.217	7.18	7.849	9.28	10.148
y	1.492	2.693	3.221	3.867	4.847	5.595	6.454	7.127	7.873	9.217

例えば、 x が重りの重さ、 y がバネの長さ、とかを想像して欲しい。
これらは測定データなので誤差が含まれているが、大雑把に

$$y = ax + b$$

という関係を満たしていると推定されるとき、「最もふさわしい」 a と b を求める方法が、最小二乗法である。



x_i における点 (x_i, y_i) と直線 $y = ax + b$ との y 座標の差は、

$$|(ax_i + b) - y_i|$$

で表される。全ての点で総合的に見てこの値が小さくなるような a, b を求める必要があるので、

$$E = \sum_{i=1}^n (ax_i + b - y_i)^2$$

を最小にするような a, b を求めることにする。(ここで二乗してるのが名前の由来。絶対値のままだと計算が面倒)

E を a と b でそれぞれ偏微分する。

$$\begin{aligned}\frac{\partial E}{\partial a} &= \sum_{i=1}^n 2(ax_i + b - y_i)x_i \\ &= 2 \left(a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i - \sum_{i=1}^n x_i y_i \right) \\ \frac{\partial E}{\partial b} &= \sum_{i=1}^n 2(ax_i + b - y_i) \\ &= 2 \left(a \sum_{i=1}^n x_i + bn - \sum_{i=1}^n y_i \right)\end{aligned}$$

a, b が E を最小化するならこの偏微分が 0 になるはずなのでそれぞれ 0 とおいて、

$$\begin{aligned}a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i &= \sum_{i=1}^n x_i y_i \\ a \sum_{i=1}^n x_i + bn &= \sum_{i=1}^n y_i\end{aligned}$$

これを a と b に関する連立方程式と見て解けば良い。

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