Least Square Method (최소 자승법)

개요 :

Measurement : (x_1,y_1), (x_2,y_2), (x_3,y_3) …. (x_n,y_n)

↓ Fit

$y = ax + b$

Principle

Error between calculated and measured variables

$$e_i = y_i - ax_i - b_i$$

Objective of the fit is to minimize sum of squared errors

$$S_r = \sum_{i=1}^{n} e_i^{\,2} = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

At the minimum, two derivatives with respect to a and b have to be zero.

$$\frac{\partial S_r}{\partial a} = 0 \quad \text{and} \quad \frac{\partial S_r}{\partial b} = 0$$

$$\frac{\partial S_r}{\partial a} = -2\sum_{i=1}^{n} (y_i - ax_i - b)x_i = 0$$

$$\frac{\partial S_r}{\partial b} = -2\sum_{i=1}^{n} (y_i - ax_i - b) = 0$$

Unknown a and b can be solved using two equations above .

$$\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} ax_i - \sum_{i=1}^{n} b = 0$$

$$\sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{n} ax_i^{\,2} - \sum_{i=1}^{n} bx_i = 0$$

Final Solution:

$$a = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^{\,2} - \left(\sum x_i\right)^2}, \quad b = \bar{y} - a\bar{x}$$

$$\bar{y} = \frac{\sum y_i}{n}, \quad \bar{x} = \frac{\sum x_i}{n}$$

Using the notation in the text book,

$$s_x = \frac{\sum x_i}{n} \qquad s_y = \frac{\sum y_i}{n}$$

$$s_{xx} = \frac{\sum x_i^2}{n} \qquad s_{xy} = \frac{\sum x_i y_i}{n}$$

$$a = \frac{s_{xy} - s_x s_y}{s_{xx} - (s_x)^2}, \quad b = \frac{s_{xx} s_y - s_{xy} s_x}{s_{xx} - (s_x)^2}$$

<u>Representation for goodness of fit</u>

(1) Sum of squares : Total sum of squared error

$$S_r = \sum_{i=1}^{n} e_i^{\,2} = \sum_{i=1}^{n} (y_i - a x_i - b)^2$$

(2) Standard Deviations

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}}$$

(3) Total sum of squared error around mean y

$$S_t = \sum_{i=1}^{n} e_i^{\,2} = \sum_{i=1}^{n} (\bar{y} - a x_i - b)^2$$

(4) Correlation coefficient

$$r^2 = \frac{S_t - S_r}{S_t}$$

r→1 for good fit (Sr →0)

r→0 for bad fit (St > Sr)