Linear Least Squares Fit and Error of Fitting

- **Least squares fit**

We assume a linear least squares line as follows:

$$y = a + bx\,. \qquad \text{(Eq0)}$$

The quantitative procedure to find the least squares is to minimize the deviation between the expected line and data points. Therefore, we have

$$Q = \sum_{i=1}^{m} w_i \{y_i - (a + bx_i)\}^2 \qquad \text{(Eq1)}$$

where w_i is weight $\left(= \sum_i 1/\sigma_i\right)$. Here, we use 1 for w_i. y_i and x_i are the data set. The following two equations will be the necessary condition for Q to be a minimum:

$$\frac{\partial Q}{\partial a} = 0, \qquad \frac{\partial Q}{\partial b} = 0. \qquad \text{(Eq2)}$$

We can calculate (Eq2) with (Eq1), which can be cast in the form of a single matrix equation:

$$\mathbf{MP} = \mathbf{Y} \qquad \text{(Eq3)}$$

where $\mathbf{M} = \begin{bmatrix} \sum_{i=1}^{\mathbf{m}} 1 & \sum_{i=1}^{m} x_i \\ \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} x_i^2 \end{bmatrix}$, $\mathbf{P} = \begin{bmatrix} a \\ b \end{bmatrix}$, and $\mathbf{Y} = \begin{bmatrix} \sum_{i=1}^{m} y_i \\ \sum_{i=1}^{m} x_i y_i \end{bmatrix}$. To solve for $\mathbf{P}$, we use the inversion of the matrix $\mathbf{M}$. Thus,

$$\mathbf{P} = \mathbf{M}^{-1}\mathbf{MP} = \mathbf{M}^{-1}\mathbf{P}\,. \qquad \text{(Eq4)}$$

We therefore obtain a and b for the least squares fit:

$$a = \frac{1}{\Delta}\left\{\left(\sum_{i=1}^{m} x_i^2\right)\left(\sum_{i=1}^{m} y_i\right) - \left(\sum_{i=1}^{m} x_i\right)\left(\sum_{i=1}^{m} x_i y_i\right)\right\} \qquad \text{(Eq5)}$$

$$b = \frac{1}{\Delta}\left\{\left(\sum_{i=1}^{m} 1\right)\left(\sum_{i=1}^{m} x_i y_i\right) - \left(\sum_{i=1}^{m} x_i\right)\left(\sum_{i=1}^{m} y_i\right)\right\} \qquad \text{(Eq6)}$$

where $\Delta = \left(\sum_{i=1}^{m} 1\right)\left(\sum_{i=1}^{m} x_i^2\right) - \left(\sum_{i=1}^{m} x_i\right)^2$. This is the systematic way to obtain (Eq0).

- **Error of fitting (χ-square)**

The error analysis of this calculation will be given as follows:

$$\chi^2 = \frac{1}{m-n-1}\sum_{i=1}^{m}\{y_i - (a+bx_i)\}^2 \qquad \text{(Eq7)}$$

where n is the number of parameters (only a and b here). This is an analogy from the method of standard deviation. If you have a good fit, χ^2 will approach 1 in the limit of large m. The values y, a, and b have deviations due to the error, so each variant will be given as follows:

$$\sigma_y^2 = \frac{1}{m-3}\sum_{i=1}^{m}\{y_i - a - bx_i\}^2 \qquad \text{(Eq8)}$$

$$\sigma_a^2 = \frac{\sigma_y^2}{\Delta}\sum_{i=1}^{m}x_i^2 \qquad \text{(Eq9)}$$

$$\sigma_b^2 = \frac{m\sigma_y^2}{\Delta} \qquad \text{(Eq10)}$$

where $\Delta = \left(\sum_{i=1}^{m}1\right)\left(\sum_{i=1}^{m}x_i^2\right) - \left(\sum_{i=1}^{m}x_i\right)^2$.