Function model fitting

January 10, 2013

1 Parameters

1.1 One dimensional function

A function $f(x, \mathbf{a})$ to be fitted to data (x_i, y_i) $(i = 1, 2, \cdots, n)$ is assumed to include parameters $\mathbf{a} = (a_1, a_2, \cdots, a_m)$, where $m < n$. In addition, it is assumed that the errors are not included in x_i or are small enough to be negligible when compared to the errors of y_i. If the population distribution of y_i is assumed to be Gaussian distribution $N(f(x_i, \mathbf{a}), \sigma_i^2)$ with mean $f(x_i, \mathbf{a})$ and variance σ_i^2, and y_i is a sample from the population, the likelihood function is expressed by

$$\mathcal{L} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{\{y_i - f(x_i, \mathbf{a})\}^2}{2\sigma_i^2} \right]. \tag{1}$$

If the population distribution of y_i is assumed to be Poisson distribution $P(y_i; f(x_i, \mathbf{a}))$ with mean $f(x_i, \mathbf{a})$, and y_i is a sample from the population, the likelihood function is expressed by

$$\mathcal{L} = \prod_{i=1}^{n} \frac{\{f(x_i, \mathbf{a})\}^{y_i}}{y_i!} \exp\{-f(x_i, \mathbf{a})\}. \tag{2}$$

In order to maximize the probability that the data $y_1, y_2, \cdots, y_n$ are measured, that is, to maximize the likelihood function $\mathcal{L}$, taking the logarithm of the function $\mathcal{L}$, differentiating by parameters a_j $(j = 1, 2, \cdots, m)$ and setting to zero,

$$\sum_{i=1}^{n} w_i \{y_i - f(x_i, \mathbf{a})\} \frac{\partial f(x_i, \mathbf{a})}{\partial a_j} = 0 \tag{3}$$

is obtained. Where $w_i \equiv 1/\sigma_i^2$ for the case that the data are obtained from Gaussian distribution or $w_i \equiv 1/f(x_i, \mathbf{a})$ for the case of Poisson distribution.

The equation (3) can be solved about parameters, a_j. This method is called the **method of maximum likelihood**[1, 2, 3]. If the data are samples from Gaussian distribution, the method of maximum likelihood is same as minimizing the argument in the exponential in the likelihood function (1), that is, minimizing the quantity,

$$\chi^2 = \sum_{i=1}^{n} \frac{\{y_i - f(x_i, \mathbf{a})\}^2}{\sigma_i^2}. \tag{4}$$

This method is called the **method of least squares**.

If the function model is nonlinear with respect to the parameters included in it, the equation (3) becomes nonlinear simultaneous equations and it can not be solved analytically. For that case, the following methods are taken in order to compute the optimal values of the parameters. When the parameters $\mathbf{a_0} = (a_{10}, a_{20}, \cdots, a_{m0})$ at the vicinity of the optimal parameters $\mathbf{a} = (a_1, a_2, \cdots, a_m)$ are given, the optimal parameters $\mathbf{a}$ can be expressed as $\mathbf{a} = \mathbf{a_0} + \delta\mathbf{a}$ using correction terms, $\delta\mathbf{a}$. Expanding the function $f(x, \mathbf{a})$ in a Taylor series about the vicinity of $\mathbf{a_0}$ to the first order in the parameter increments $\delta\mathbf{a}$, the function $f(x, \mathbf{a})$ is expressed by

$$f(x, \mathbf{a}) = f(x, \mathbf{a_0}) + \sum_{k=1}^{m} \frac{\partial f(x, \mathbf{a_0})}{\partial a_k} \delta a_k. \tag{5}$$

Substituting the equation (5) for (3) and ignoring second or higher order terms of δa_k,

$$\sum_{k=1}^{m} \alpha_{jk} \delta a_k = \beta_j \quad (j = 1, 2, \cdots, m) \tag{6}$$

is obtained. Where, α_{jk} and β_j are expressed by

$$\alpha_{jk} = \sum_{i=1}^{n} w_i \left[\frac{\partial f(x_i, \mathbf{a_0})}{\partial a_j} \frac{\partial f(x_i, \mathbf{a_0})}{\partial a_k} - \{y_i - f(x_i, \mathbf{a_0})\} \frac{\partial^2 f(x_i, \mathbf{a_0})}{\partial a_j \partial a_k} \right],$$

$$\beta_j = \sum_{i=1}^{n} w_i \{y_i - f(x_i, \mathbf{a_0})\} \frac{\partial f(x_i, \mathbf{a_0})}{\partial a_j},$$

respectively. This α_{jk} is called the **curvature matrix**. Here again, for the case that the data are the samples from Gauusian distribution, $w_i \equiv 1/\sigma_i^2$, and for the case of Poisson distribution, $w_i \equiv 1/f(x_i, \mathbf{a_0})$.

For the second term in [ ] of α_{jk}, the second order differentiation of $f(x_i, \mathbf{a_0})$ is zero, where $f(x_i, \mathbf{a_0})$ is linear function with respect to the parameters $\mathbf{a}$, or small enough to be negligible when compared to the term involving the first derivative. It also has an additional possibility of being ignorably small in practice: The term multiplying the second derivative is $\{y_i - f(x_i, \mathbf{a_0})\}$. For successful model, this term should just be the random measurement error of each point. This error can have either sign, and should in general be uncorrelated with the model. Therefore, the second derivative terms tend to cancel out when summed over i.

Computing the inverse matrix of α_{jk} which is expresses by $(\alpha^{-1})_{jk}$, the $\delta\mathbf{a}$ can be obtained from

$$\delta a_j = \sum_{k=1}^{m} (\alpha^{-1})_{jk} \beta_k \quad (j = 1, 2, \cdots, m). \tag{7}$$

Therefore, the optimal parameters can be obtained from $\mathbf{a} = \mathbf{a_0} + \delta\mathbf{a}$.

The almost right value can be acquired by the first computation of the equation (7) if the initial values of the parameters $\mathbf{a_0}$ given beforehand are close to the optimal values. However, it is difficult to give such initial values of the parameters beforehand. If the parameters is obtained by the first computation, they will be considered as new parameters and the same computation will be repeated. In such a repeated computation process, when the values of the correction terms $\delta\mathbf{a}$ become small enough, and further, for the case of the least-squares method, the value of χ^2 of the equation (4) is considered to be minimum, or for the case of the maximum likelihood method, the value of $\mathcal{L}$ in the equation (2) is considered to be maximum, the computation will be completed and the values of the parameters at that time will be considered as the values of the optimal parameters $\mathbf{a}$. (Refer to section3.)

The **modified Marquardt method**[2] is used in order to minimize χ^2 or maximize $\mathcal{L}$ and to converge the computations described above. Using the Marquardt method, near the starting point which gave the initial values of the parameters, to reduce the χ^2 or increase $\mathcal{L}$, the diagonal elements of the matrix α_{jk} are emphasized. When approaching a converging point, the optimal values of the parameters can be determined by the equation (7) with exact Taylor expansion of the function.

1.2 Two dimensional function

For the case that there are errors in both data, (X_i, Y_i) $(i = 1, 2, \cdots, n)$, the likelihood function corresponding to the population distribution must be considered.

For example, if both X_i and Y_i are samples from Gaussian distribution, $N(x_i, \sigma_{x_i}^2)$ and $N(y_i, \sigma_{y_i}^2)$, respectively, the likelihood function is expressed by

$$\mathcal{L} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_{x_i}} \exp\left[-\frac{(X_i - x_i)^2}{2\sigma_{x_i}^2} \right] \frac{1}{\sqrt{2\pi}\sigma_{y_i}} \exp\left[-\frac{(Y_i - y_i)^2}{2\sigma_{y_i}^2} \right]. \tag{8}$$

Or, if both X_i and Y_i are samples from Poisson distribution, $P(X_i; x_i)$ and $P(Y_i; y_i)$, respectively, the likelihood function is expressed by

$$\mathcal{L} = \prod_{i=1}^{n} \frac{x_i^{X_i}}{X_i!} \exp(-x_i) \frac{y_i^{Y_i}}{Y_i!} \exp(-y_i).$$

Where, x_i and y_i are the means, and $\sigma_{x_i}^2$ and $\sigma_{y_i}^2$ the variances in the distributions. Moreover, if there is a constraint,

$$f(x_i, y_i, \mathbf{a}) = 0,$$

between x_i and y_i, the paramerers, $\mathbf{a} = (a_1, a_2, \cdots, a_m)$ $(m < n)$, in the function model which maximize the likelihood function, can be obtained using the method of Lagrange multipliers[3].

For example, taking the logarithm of the likelihood function (8), the likelihood function with the multipliers λ_i $(i = 1, 2, \cdots, n)$ is expressed by

$$\mathcal{L} = \sum_{i=1}^{n} \left\{ -\frac{(X_i - x_i)^2}{2\sigma_{x_i}^2} - \ln \sigma_{x_i} - \frac{(Y_i - y_i)^2}{2\sigma_{y_i}^2} - \ln \sigma_{y_i} + \lambda_i f_i \right\} - n \ln(2\pi).$$

Maximizing this function with respect to x_i、y_i, λ_i and a_j $(i = 1, 2, \cdots, m)$, respectively,

$$\sum_{i=1}^{n} w_i \left\{ (X_i - x_i) \frac{\partial f_i}{\partial x_i} + (Y_i - y_i) \frac{\partial f_i}{\partial y_i} \right\} \frac{\partial f_i}{\partial a_j} = 0, \tag{9}$$

is obtained. Where, $f_i \equiv f(x_i, y_i, a_j)$, and $w_i = \left\{ \left(\frac{\partial f_i}{\partial x_i} \right)^2 \sigma_{x_i}^2 + \left(\frac{\partial f_i}{\partial y_i} \right)^2 \sigma_{y_i}^2 \right\}^{-1}$.

If the almost right values near the true values, x_i, y_i and a_j, which are expressed by x_{0i}, y_{0i} and a_{0j}, are given, the true values can be obtained from

$$x_i = x_{0i} + \delta x_i \tag{10}$$
$$y_i = y_{0i} + \delta y_i \tag{11}$$
$$a_j = a_{0j} - \delta a_j. \tag{12}$$

Where, the correction terms δx_i, δy_i and δa_j are small enough.

When expanding the function f_i in a Taylor series about the vicinity of x_{0i}, y_{0i} and a_{0j} to the first order, the function is expressed by

$$f_i = f_{0i} + \frac{\partial f_{0i}}{\partial x_i} \delta x_i + \frac{\partial f_{0i}}{\partial y_i} \delta y_i - \sum_{k=1}^{m} \frac{\partial f_{0i}}{\partial a_k} \delta a_k = 0. \tag{13}$$

Where, $f_{0i} \equiv f(x_{0i}, y_{0i}, a_{0j})$.

Substituting the equations (10) and (11) for (9) and using the relation (13),

$$\sum_{i=1}^{n} w_{0i} \left\{ f_{0i} + (X_i - x_{0i}) \frac{\partial f_{0i}}{\partial x_i} + (Y_i - y_{0i}) \frac{\partial f_{0i}}{\partial y_i} \right\} \frac{\partial f_{0i}}{\partial a_j} = \sum_{i=1}^{n} w_{0i} \frac{\partial f_{0i}}{\partial a_j} \sum_{k=1}^{m} \frac{\partial f_{0i}}{\partial a_k} \delta a_k$$

is obtained. Where, $w_{0i} = \left\{ \left(\frac{\partial f_{0i}}{\partial x_i} \right)^2 \sigma_{x_i}^2 + \left(\frac{\partial f_{0i}}{\partial y_i} \right)^2 \sigma_{y_i}^2 \right\}^{-1}$.

Here again, defining

$$\alpha_{jk} \equiv \sum_{i=1}^{n} w_{0i} \frac{\partial f_{0i}}{\partial a_j} \frac{\partial f_{0i}}{\partial a_k},$$
$$\beta_j \equiv \sum_{i=1}^{n} w_{0i} \left\{ f_{0i} + (X_i - x_{0i}) \frac{\partial f_{0i}}{\partial x_i} + (Y_i - y_{0i}) \frac{\partial f_{0i}}{\partial y_i} \right\} \frac{\partial f_{0i}}{\partial a_j},$$

the correction terms of the parameters, δa_j, can be obtained from

$$\delta a_j = \sum_{k=1}^{m} (\alpha^{-1})_{jk} \beta_k \quad (j = 1, 2, \cdots, m). \tag{14}$$

Therefore, the optimal parameters, a_j, can be obtained from the equation (12).

2 Errors of parameters[3]

For the case of the nonlinear least squares fitting, the χ^2 is expressed by

$$\chi^2 = \sum_{i=1}^{n} \left\{ \frac{y_i - f(x_i, \mathbf{a})}{\sigma_i} \right\}^2.$$

From the condition,

$$\frac{\partial \chi^2}{\partial a_j} = 0,$$

or from the equation (3),

$$\sum_{i=1}^{n} \frac{y_i - f(x_i, \mathbf{a})}{\sigma_i^2} \cdot \frac{\partial f(x_i, \mathbf{a})}{\partial a_j} = 0$$

is obtained. With $w_i \equiv \dfrac{1}{\sigma_i^2}$ and $\mathbf{y} = (y_1, y_2, \cdots, y_n)$, defining

$$F_j(\mathbf{a}, \mathbf{y}) \equiv \sum_{i=1}^{n} w_i \{y_i - f(x_i, \mathbf{a})\} \frac{\partial f(x_i, \mathbf{a})}{\partial a_j} = 0,$$

this F_j satisfies

$$\sum_{k=1}^{m} \frac{\partial F_j(\mathbf{a}, \mathbf{y})}{\partial a_k} da_k + \sum_{i=1}^{n} \frac{\partial F_j(\mathbf{a}, \mathbf{y})}{\partial y_i} dy_i = 0. \tag{15}$$

Differentiating F_j with respect to a_k,

$$\frac{\partial F_j(\mathbf{a}, \mathbf{y})}{\partial a_k} = -\sum_{i=1}^{n} w_i \left[\frac{\partial f(x_i, \mathbf{a})}{\partial a_k} \cdot \frac{\partial f(x_i, \mathbf{a})}{\partial a_j} - \{y_i - f(x_i, \mathbf{a})\} \frac{\partial^2 f(x_i, \mathbf{a})}{\partial a_j \partial a_k} \right]$$

is obtained. Here, if defining

$$\mathbf{D} = \begin{pmatrix} \frac{\partial F_1}{\partial a_1} & \frac{\partial F_1}{\partial a_2} & \cdots & \frac{\partial F_1}{\partial a_m} \\ \frac{\partial F_2}{\partial a_1} & \frac{\partial F_2}{\partial a_2} & \cdots & \frac{\partial F_2}{\partial a_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial a_1} & \frac{\partial F_m}{\partial a_2} & \cdots & \frac{\partial F_m}{\partial a_m} \end{pmatrix},$$

and from the equation (15),

$$da_k = -\sum_{j=1}^{m} (D^{-1})_{kj} \sum_{i=1}^{n} \frac{\partial F_j}{\partial y_i} dy_i \tag{16}$$

is obtained. While, considering a_k as a function of $y_1, y_2, \cdots, y_n$,

$$da_k = \sum_{i=1}^{n} \frac{\partial a_k}{\partial y_i} dy_i. \tag{17}$$

is obtained. Comparing the equation (16) with (17),

$$\frac{\partial a_k}{\partial y_i} = -\sum_{j=1}^{m} (D^{-1})_{kj} \frac{\partial F_j}{\partial y_i} = -w_i \sum_{j=1}^{m} (D^{-1})_{kj} \frac{\partial f(x_i, \mathbf{a})}{\partial a_j}$$

is obtained.

Therefore, expressing the errors of the parameters by s_{a_k}, from the propagation of errors,

$$s_{a_k}^2 = \sum_{i=1}^{n} \left(\frac{\partial a_k}{\partial y_i} \right)^2 s_i^2 = \sum_{i=1}^{n} w_i \left[\sum_{j=1}^{m} (D^{-1})_{kj} \frac{\partial f(x_i, \mathbf{a})}{\partial a_j} \right]^2$$

is obtained. Where, $s_i = \sigma_i$ is the error of y_i.

For the case of the maximum likelihood fitting, from the equation (3),

$$F_j(\mathbf{a}, \mathbf{y}) \equiv \sum_{i=1}^{n} \frac{y_i - f(x_i, \mathbf{a})}{f(x_i, \mathbf{a})} \cdot \frac{\partial f(x_i, \mathbf{a})}{\partial a_j} = 0$$

is defined and

$$\frac{\partial F_j}{\partial a_k} = -\sum_{i=1}^{n} \frac{1}{f(x_i, \mathbf{a})} \left[\frac{y_i}{f(x_i, \mathbf{a})} \cdot \frac{\partial f(x_i, \mathbf{a})}{\partial a_j} \cdot \frac{\partial f(x_i, \mathbf{a})}{\partial a_k} - \{y_i - f(x_i, \mathbf{a})\} \frac{\partial^2 f(x_i, \mathbf{a})}{\partial a_j \partial a_k} \right]$$

is obatined. Since this F_j satisfies

$$\sum_{k=1}^{m} \frac{\partial F_j(\mathbf{a}, \mathbf{y})}{\partial a_k} da_k + \sum_{i=1}^{n} \frac{\partial F_j(\mathbf{a}, \mathbf{y})}{\partial y_i} dy_i = 0,$$

$$da_k = -\sum_{j=1}^{m} (D^{-1})_{kj} \sum_{i=1}^{n} \frac{\partial F_j}{\partial y_i} dy_i \tag{18}$$

is obtained. Where,

$$D_{jk} = \frac{\partial F_j}{\partial a_k}.$$

While, considering a_k as a function of $y_1, y_2, \cdots, y_n$,

$$da_k = \sum_{i=1}^{n} \frac{\partial a_k}{\partial y_i} dy_i. \tag{19}$$

is obtained. Comparing the equation (18) with (19),

$$\frac{\partial a_k}{\partial y_i} = -\sum_{j=1}^{m} (D^{-1})_{kj} \frac{\partial F_j}{\partial y_i} = \frac{1}{f(x_i, \mathbf{a})} \sum_{j=1}^{m} (D^{-1})_{kj} \frac{\partial f(x_i, \mathbf{a})}{\partial a_j}$$

is obtained. Therefore, expressing the errors of the parameters by s_{a_k}, from the propagation of errors,

$$s_{a_k}^2 = \sum_{i=1}^{n} \left(\frac{\partial a_k}{\partial y_i} \right)^2 s_i^2 = \sum_{i=1}^{n} \frac{1}{f(x_i, \mathbf{a})} \left\{ \sum_{j=1}^{m} (D^{-1})_{kj} \frac{\partial f(x_i, \mathbf{a})}{\partial a_j} \right\}^2$$

is obtained. Where, $s_i = \sqrt{f(x_i, \mathbf{a})}$.

For the case of two dimensional fitting, from the equation (9), defining

$$F_j(\mathbf{a}, \mathbf{X}, \mathbf{Y}) \equiv \sum_{i=1}^{n} w_i \left\{ f_i + (X_i - x_i) \frac{\partial f_i}{\partial x_i} + (Y_i - y_i) \frac{\partial f_i}{\partial y_i} \right\} \frac{\partial f_i}{\partial a_j} = 0,$$

$$\begin{aligned}
\frac{\partial F_j}{\partial a_k} =& \sum_{i=1}^{n} \left[\left[-2 \times w_i^2 \sigma_{x_i}^2 \frac{\partial f_i}{\partial x_i} \left\{ f_i + (X_i - x_i) \frac{\partial f_i}{\partial x_i} + (Y_i - y_i) \frac{\partial f_i}{\partial y_i} \right\} + w_i (X_i - x_i) \right] \frac{\partial f_i}{\partial a_j} \cdot \frac{\partial^2 f_i}{\partial x_i \partial a_k} \right. \\
&+ \left[-2 \times w_i^2 \sigma_{y_i}^2 \frac{\partial f_i}{\partial y_i} \left\{ f_i + (X_i - x_i) \frac{\partial f_i}{\partial x_i} + (Y_i - y_i) \frac{\partial f_i}{\partial y_i} \right\} + w_i (Y_i - y_i) \right] \frac{\partial f_i}{\partial a_j} \cdot \frac{\partial^2 f_i}{\partial y_i \partial a_k} \\
&+ w_i \frac{\partial f_i}{\partial a_j} \frac{\partial f_i}{\partial a_k} \\
&\left. + w_i \left\{ f_i + (X_i - x_i) \frac{\partial f_i}{\partial x_i} + (Y_i - y_i) \frac{\partial f_i}{\partial y_i} \right\} \frac{\partial^2 f_i}{\partial a_j \partial a_k} \right]
\end{aligned}$$

is obtained. Where, $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ and $\mathbf{Y} = (Y_1, Y_2, \cdots, Y_n)$. Since this F_j satisfies

$$\sum_{k=1}^{m} \frac{\partial F_j}{\partial a_k} da_k + \sum_{i=1}^{n} \frac{\partial F_j}{\partial X_i} dX_i + \sum_{i=1}^{n} \frac{\partial F_j}{\partial Y_i} dY_i = 0,$$

$$da_k = -\sum_{j=1}^{m} (D^{-1})_{kj} \sum_{i=1}^{n} \left(\frac{\partial F_j}{\partial X_i} dX_i + \frac{\partial F_j}{\partial Y_i} dY_i \right) \tag{20}$$

is obtained. Where,

$$D_{jk} = \frac{\partial F_j}{\partial a_k}.$$

While, considering a_k as a function of X_i and Y_i,

$$da_k = \sum_{i=1}^{n} \left(\frac{\partial a_k}{\partial X_i} dX_i + \frac{\partial a_k}{\partial Y_i} dY_i \right) \tag{21}$$

is obtained. Comparing equation (20) with (21),

$$\frac{\partial a_k}{\partial X_i} = -\sum_{j=1}^{m} (D^{-1})_{kj} \frac{\partial F_j}{\partial X_i} = -w_i \frac{\partial f_i}{\partial x_i} \sum_{j=1}^{m} (D^{-1})_{kj} \frac{\partial f_i}{\partial a_j}$$

and

$$\frac{\partial a_k}{\partial Y_i} = -\sum_{j=1}^{m} (D^{-1})_{kj} \frac{\partial F_j}{\partial Y_i} = -w_i \frac{\partial f_i}{\partial y_i} \sum_{j=1}^{m} (D^{-1})_{kj} \frac{\partial f_i}{\partial a_j}$$

are obtained, respectively. Therefore, expressing the errors of the parameters by s_{a_k}, from the propagation of errors,

$$s_{a_k}^2 = \sum_{i=1}^{n} \left\{ \left(\frac{\partial a_k}{\partial X_i} \right)^2 s_{x_i}^2 + \left(\frac{\partial a_k}{\partial Y_i} \right)^2 s_{y_i}^2 \right\}$$

is obtained. Where, $s_{x_i} = \sigma_{x_i}$ and $s_{y_i} = \sigma_{y_i}$ for the case that x_i and y_i are samples from Gaussian distributions, or $s_{x_i} = \sqrt{x_i}$ and $s_{y_i} = \sqrt{y_i}$ for the case of Poisson distributions.

3 Convergence in the computation process for the correction terms $\delta\mathbf{a}$

For the least-squares method, when the values of the correction terms $\delta\mathbf{a}$ computed in the repeated computation process and the value of χ^2 of the equation (4) satisfy the following conditions[2],

$$\begin{cases} \dfrac{\triangle\chi^2}{1 + \chi^2(\mathbf{a}^{(k)})} \leqq \epsilon_{\chi^2} \\ \\ |\delta a_j| \leqq \epsilon_{a_j} \;\; (j = 1, 2, \cdots, m), \end{cases} \tag{22}$$

the computation is completed and the values of the parameters at that time are considered as the values of the optimal parameters $\mathbf{a}$. Where, the $\chi^2(\mathbf{a}^{(k)})$ is the value of the χ^2 computed in the kth computation process for the correction terms and the $\triangle\chi^2$ is defined by the following formula,

$$\triangle\chi^2 = \chi^2(\mathbf{a}^{(k-1)}) - \chi^2(\mathbf{a}^{(k)}),$$

and ϵ_{χ^2} and ϵ_{a_j} are defined by the following equations,

$$\begin{cases} \epsilon_{\chi^2} = max(\epsilon_M, \; \epsilon_u) \\ \\ \epsilon_{a_j} = max(\epsilon_{\chi^2} s_{a_j}, \; \epsilon_M |a_j|). \end{cases}$$

By default, setting $\epsilon_M = 2.22 \times 10^{-16}$ and $\epsilon_u = 1.0 \times 10^{-12}$, $\epsilon_{\chi^2} = 1.0 \times 10^{-12}$ is used. The s_{a_j} is the value of the error of the parameter a_j.

For the case of the maximum likelihood method, when the value of $\mathcal{L}$ in the equation (2) is computed and the above conditions (22) which χ^2 is replaced with $\mathcal{L}$ are satisfied, the computation is completed and the values of the parameters at that time are considered as the values of the optimal parameters $\mathbf{a}$.

References

[1] P.R.Bevington, D.K.Robinson, ”Data Reduction and Error Analysis for The Physical Sciences, Third Edition”, McGraw-Hill, 2003

[2] T.Nakagawa, Y.Oyanagi, ”Experimental Data Analysis Using Least Squares Method”, University of Tokyo Press, 1982 (Japanese) (中川徹、小柳義夫、”最小二乗法による実験データ解析”、 東京大学出版会、1982)

[3] T.Awaya, ”Data Analysis”, Gakkai Shuppan Center, 1991 (Japanese) (粟谷隆、”データ解析”、学会出版センター 、1991)