

# 最尤推定法

## Maximum-Likelihood Method

名古屋工業大学セラミックス基盤工学研究センター  
井田 隆

計数法で強度測定をするときに、強度が弱ければ、実際に測定されるカウント数がゼロになることはありえます。

最小二乗法では、誤差の二乗に反比例する重みをつけて残差の二乗和が最小になるパラメータを求めます。計数法で測定されたカウント数の統計的な変動はカウント数の平方根に近似的に比例するので、重みとしてはカウント数の逆数を用いる場合が多いようです。しかしこの場合、ゼロカウントの強度の誤差が形式的にゼロになり、重みが無限大になってしまうという不都合があります。

最小二乗法によるデータの解析を行う場合に、強度ゼロのデータを除外してしまうというやり方もあるようですが、「カウント数がゼロ」ということは、「強度が低い」という情報を含んでいるので、この情報を捨ててしまうのはいかにも「もったいない」ことのように思われます。

粉末回折測定では、回折ピークからはずれた場所で観測されるバックグラウンド強度は、回折ピークの裾や試料中のアモルファス（非晶質）成分からの散乱、迷光、電気的な雑音の影響を受けています。試料中の結晶成分の構造を解析するためには、回折強度データに含まれるバックグラウンドの低い方が質の高いデータだと言えるので、回折強度測定を行う前に試料の結晶性を高めたり、回折ピークの裾の強度を低減したり迷光や電気的な雑音を減らすために装置や測定条件に工夫をする人もいるでしょう。質の高いデータを得るための努力の結果として出現しうる「ゼロカウント」のデータを捨てられてしまっは、まじめに実験に取り組もうとする人にとっては、やりきれないだろうと思います。

この問題は、最尤推定法 Maximum-likelihood method の考え方 [Antoniadis, A., Berruyer, J. and Filhol, A. (1990) *Acta Cryst. A* **46**, 692-711] を使えば（部分的には）解決するように思われます。

まず、カウント数が仮想的に平均  $\mu$ 、標準偏差  $\sigma$  の正規分布に従う場合について考えます（実際にはカウント数は正の整数値しかとりませんが、仮想的に任意の実数値をとるとします）。このとき、 $Y$  カウントされる確率（確率密度関数）は

$$f_{normal}(Y; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(Y - \mu)^2}{2\sigma^2}\right] \quad (1)$$

で表されます。標本点の数が  $N$ 、実測のカウント数が  $\{Y_1, Y_2, \dots, Y_N\}$  であるとして、それぞれの標本点での強度データが相関を持たない（独立である）とします。そのようなカウ

ント数の組み合わせが実現する確率は、各標本点でのそれぞれのカウント数を実現される確率の積に等しいとみなすことができ、各標本点での期待値が  $\{\mu_1, \mu_2, \dots, \mu_N\}$ 、標準偏差が  $\{\sigma_1, \sigma_2, \dots, \sigma_N\}$  であるとすれば、

$$P_{normal}\{Y_1, Y_2, \dots, Y_N\} = \prod_{j=1}^N f_{normal}(Y_j; \mu_j, \sigma_j) = \prod_{j=1}^N \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left[-\frac{(Y_j - \mu_j)^2}{2\sigma_j^2}\right]$$

$$= \left(\prod_{j=1}^N \frac{1}{\sqrt{2\pi}\sigma_j}\right) \exp\left[-\sum_{j=1}^N \frac{(Y_j - \mu_j)^2}{2\sigma_j^2}\right] \quad (2)$$

と表されます。ただし、ここで各標本点でのカウント数は独立であるとします。

式(2)から、標準偏差  $\{\sigma_1, \sigma_2, \dots, \sigma_N\}$  が既知の場合、確率最大となるのは  $\sum_{j=1}^N \frac{(Y_j - \mu_j)^2}{\sigma_j^2}$  が最小となる場合であることはすぐにわかります。つまり、観測強度の統計的な分布が「標準偏差既知の正規分布」に従う場合には、最小二乗法により得られる解（重み付きの残差二乗和が最小になる解）は、最尤推定解に一致します。また、期待値  $\mu_j$  がパラメータ  $\{a_1, a_2, \dots, a_M\}$  によって滑らかに変化する関数だとすると、パラメータが最小二乗条件を満たす場合に、

$$\frac{\partial}{\partial a_i} \sum_{j=1}^N \frac{(Y_j - \mu_j)^2}{2\sigma_j^2} = -\sum_{j=1}^N \frac{Y_j - \mu_j}{\sigma_j^2} \cdot \frac{\partial \mu_j}{\partial a_i} = 0 \quad (3)$$

の関係が成り立ちます。

一方、計数法により観測されるカウント数は、ポアソン Poisson 分布に従うとされています。カウント数はゼロ以上の整数値しかとりえないので、任意の実数値をとりうる正規分布とは、本来はまったく違うものだという点には注意するべきでしょう。

期待値  $\lambda$  のポアソン分布を仮定すると、 $Y$  カウントされる確率は

$$f_{Poisson}(Y; \lambda) = \frac{\lambda^Y}{Y!} e^{-\lambda} = \frac{\lambda^Y}{\Gamma(Y+1)} e^{-\lambda} \quad (4)$$

で表されます。かりに  $N$  標本点で実測のカウント数が  $\{Y_1, Y_2, \dots, Y_N\}$  であったとすると、これが実現する確率は、各標本点での期待値  $\{\lambda_1, \lambda_2, \dots, \lambda_N\}$  に対して

$$P_{Poisson}\{Y_1, Y_2, \dots, Y_N\} = \prod_{j=1}^N \frac{\lambda_j^{Y_j}}{\Gamma(Y_j+1)} e^{-\lambda_j} = \exp\left[\sum_{j=1}^N (Y_j \ln \lambda_j - \ln \Gamma(Y_j+1) - \lambda_j)\right] \quad (5)$$

ですから、パラメータ  $\{\lambda_1, \lambda_2, \dots, \lambda_N\}$  を変化させたときに、確率最大となる（＝最も尤もらしい）のは  $\sum_{j=1}^N (Y_j \ln \lambda_j - \lambda_j)$  が最大となる場合です。  $\lambda_j$  がパラメータ  $\{a_1, a_2, \dots, a_M\}$  によって滑らかに変化する関数だとすると、

$$\frac{\partial}{\partial a_i} \sum_{j=1}^N (Y_j \ln \lambda_j - \lambda_j) = \sum_{j=1}^N \left( \frac{Y_j}{\lambda_j} \cdot \frac{\partial \lambda_j}{\partial a_i} - \frac{\partial \lambda_j}{\partial a_i} \right) = \sum_{j=1}^N \left( \frac{Y_j}{\lambda_j} - 1 \right) \frac{\partial \lambda_j}{\partial a_i}$$

$$= \sum_{j=1}^N \frac{Y_j - \lambda_j}{\lambda_j} \cdot \frac{\partial \lambda_j}{\partial a_i} = 0 \quad (6)$$

の関係が成り立ちます。

式 (6) の関係を式 (3) と比較すると、ポアソン分布に関する最尤推定解は、 $\lambda_j = \mu_j = \sigma_j^2$  のときの最小二乗解とまったく同じになっていることがわかります。つまり、ポアソン分布に従うデータに対する最尤推定法は、「計算強度の平方根を誤差とみなす最小二乗法」と厳密に一致します。

期待値が大きい場合にポアソン分布が正規分布に近い分布になることは良く知られていますが、「ポアソン分布の最尤推定解が計算強度の平方根を誤差とみなす最小二乗解と一致する」という関係は、意外なことに期待値が大きい場合だけでなく、「期待値が小さく分布の形が正規分布から大きく異なっている場合」でさえも常に成立するのです。

なお、普通の最小二乗法は「計算強度の平方根」ではなく「実測強度の平方根」を誤差とみなし  $\sigma_j^2 = Y_j$  とするのに対して、最尤推定法では「計算強度の平方根」を誤差とみなし  $\sigma_j^2 = \lambda_j$  とします。これらの違いはわずかなことに見えるかもしれませんが、最尤推定法の考え方によれば、計算強度が正の値をとりさえすれば、実測の強度がゼロだったとしてもそれを考慮に入れた解析ができるので、バックグラウンド強度が低い場合には、大きな違いになりえます。そこで、Antoniadis らは、実測強度の平方根を誤差とみなして重みをつける方法を「最小二乗法」、計算強度の平方根を誤差とみなして重みをつける方法を「最尤推定法」と呼んで区別しています。

最尤推定法を実現するための方法としては、例えばはじめに標準偏差  $\{\sigma_1, \sigma_2, \dots, \sigma_N\} = \{\sqrt{Y_1}, \sqrt{Y_2}, \dots, \sqrt{Y_N}\}$  あるいは  $\{\sigma_1, \sigma_2, \dots, \sigma_N\} = \{\sqrt{Y_1+1}, \sqrt{Y_2+1}, \dots, \sqrt{Y_N+1}\}$  として最小二乗解を求めます。この段階で  $\{\sigma_j\} = \{\sqrt{Y_j}\}$  の関係を使うとすると強度ゼロのデータは捨てるしかありませんが、後から考慮に入れることができるので、この段階では我慢することにします。一度最小二乗解  $\{\lambda_1, \lambda_2, \dots, \lambda_N\}$  が求まれば、次回からは標準偏差  $\{\sigma_1, \sigma_2, \dots, \sigma_N\} = \{\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_N}\}$  として最小二乗法を繰り返していけば最尤推定解が得られるはずで

す。

ただし、計算の途中でたまたま計算強度がゼロまたは負になってしまうと、やはり不都合が生じてしまいます。はじめの段階では通常最小二乗法で最適化を行って、収束が十分に進んで計算強度が常に正の値になることが確認できたら、最終段階で計算強度の平方根を重みとする最尤推定法に進むという方法が現実的でしょう。

ところで、観測されるカウント数は本当にポアソン分布に従うと仮定して良いでしょうか？これは必ずしも自明な問題ではありません。計数法でX線の強度を測定する場合に、X線の光子は可視光パルスに変換され、さらに電気的なパルスに変換されるのですが、この後にさらに電気的なパルスの波形が整形されパルス高分析器 (pulse height analyzer) により信号パルスが抽出されます。この過程で、検出器も電子回路も、光子あるいは電気信号に反応するのに有限の時間がかかるので、必ず数え落としがあるはずです。特に高計数率の場合には数え落としの影響が無視できないことが良く知られています。検出系の数え落としが無視できない場合には、検出されるカウント数がポアソン分布に従うことは期待できません。

したがって、計数法により測定された実測の強度データに対して最尤推定法を適用するためには、数え落としがあることを前提としてカウント数の統計的な分布を知る必要があ

ります。極端な言い方をすれば、現実の検出システムの反応時間 response time が有限である以上、計数法により測定された強度データに対してポアソン分布を仮定した最小二乗法を適用しても、最尤推定解は得られないはずです。