格子定数の精密化とその誤差

一般に n 個の観測値の組 $(X^1{}_1, X^2{}_1, ...X^m{}_1, 1, Z_1), (X^1{}_2, X^2{}_2, ...X^m{}_2, Z_2), \cdots, (X^1{}_n, X^2{}_n, ...X^m{}_n, Z_n)$ が与えられたとき、パラメータ $(a^1, \cdot\cdot, a^m)$ について線形的な観測方程式 $Z = a^1X^1 + a^2X^2 + ... + a^mX^m$ にフィッティングさせるためには、二乗残差を最小にすればいいということは、よく知られています。(上肩添え字は累乗ではなくパラメータの個数です。念のため。)

簡単に解き方を示すと、

$$a = \begin{pmatrix} a^1 \\ .. \\ .. \\ a^m \end{pmatrix},\ X = \begin{pmatrix} X^1{}_1 & X^2{}_1 & ... & X^m{}_1 \\ X^1{}_2 & X^2{}_2 & ... & ... \\ ... & ... & ... & ... \\ X^1{}_n & ... & ... & X^m{}_n \end{pmatrix},\ Y = \begin{pmatrix} Y_1 \\ .. \\ .. \\ Y_n \end{pmatrix},\ W = \begin{pmatrix} W_1 & 0 & ... & 0 \\ 0 & W_2 & ... & ... \\ ... & ... & ... & ... \\ 0 & ... & ... & W_n \end{pmatrix}$$

としたとき (ただし W は重みを対角成分に持つ対角行列) 、

$\Delta^2 = (\tilde{a} \cdot \tilde{X} - \tilde{Y}) \cdot W \cdot (X \cdot a - Y)$ を最小にすればいいから

$\delta\Delta^2 = \delta\tilde{a} \cdot \tilde{X} \cdot W \cdot (X \cdot a - Y) + (\tilde{a} \cdot \tilde{X} - \tilde{Y}) \cdot W \cdot X \cdot \delta a = 2\delta\tilde{a}(\tilde{X} \cdot W \cdot X \cdot a - \tilde{X} \cdot W \cdot Y) = 0$ 　よって

$a = (\tilde{X} \cdot W \cdot X)^{-1} \cdot \tilde{X} \cdot W \cdot Y$ 　となります。

格子定数の計算に際しては、結晶系によって観測方程式は変わりますが、たとえば三斜晶系のときは

$$a = \begin{pmatrix} a^1 \\ .. \\ .. \\ a^6 \end{pmatrix} = \begin{pmatrix} a^*a^* \\ b^*b^* \\ c^*c^* \\ 2b^*c^*\cos\alpha^* \\ 2c^*a^*\cos\beta^* \\ 2a^*b^*\cos\gamma^* \end{pmatrix},\ X = \begin{pmatrix} h_1h_1 & k_1k_1 & l_1l_1 & k_1l_1 & l_1h_1 & h_1k_1 \\ ... & ... & ... & ... & ... & ... \\ h_nh_n & k_nk_n & l_nl_n & k_nl_n & l_nh_n & h_nk_n \end{pmatrix},\ Y = \begin{pmatrix} (\frac{1}{d_1})^2 \\ .. \\ .. \\ (\frac{1}{d_n})^2 \end{pmatrix}$$

のようになります。ここで、a*,b*…などは逆格子定数を示しています。また重みは角度に依存していると考えられます。θ の測定にのみ誤差が生じていると考えると、θ の変化に対する $(\frac{1}{d})^2$ の応答は、ブラッグ条件から $G = (\frac{1}{d})^2 = \frac{4\sin^2(\theta)}{\lambda}$ としたとき $\frac{\partial G}{\partial \theta} = \frac{4\sin(2\theta)}{\lambda}$ となります。つまり、θ が $\delta\theta$ 変化すると

$(\frac{1}{d})^2$ は $\frac{4\sin(2\theta)}{\lambda}\delta\theta$ だけ変化するということです。結局$(1/d)^2$に対する重み行列対角成分は誤差の 2 乗の逆数の比である $1/\sin^2(2\theta)$ が適当であると考えられます。

$$W = \begin{pmatrix} 1/\sin^2(2\theta_1) & 0 & ... & 0 \\ 0 & 1/\sin^2(2\theta_2) & ... & ... \\ ... & ... & ... & ... \\ 0 & ... & ... & 1/\sin^2(2\theta_n) \end{pmatrix}$$

ただし $1/\sin^2(2\theta)$ は各$(1/d)^2$の分散の逆数の絶対値ではなくその比率をあらわしています。この場合でも上述の $a = (\tilde{X}\cdot W \cdot X)^{-1}\cdot \tilde{X}\cdot W \cdot Y$ という関係式で最適値を求めることはできます。これはWが 2 回作用して絶対値をキャンセルアウトするからです。

これで一応パラメータ a は求まるのですがその誤差についてはさらに計算が必要です。誤差の伝播式を解いていくと、正しい W を使っていれば、行列 $(\tilde{X}\cdot W \cdot X)^{-1}$ の対角成分が各パラメータの分散になることが分かりますが、上で述べたように W は割合しか決めていないので絶対値が分かりません。ところでそもそも分散の定義から、

$$N - P = \sum_{i=0}^{N}(\frac{\delta_i^{\,2}}{s_i})$$

(ただし N:データ個数, P: パラメータ数, δ_i : i 番目のデータの偏差, s_i : i 番目のデータの分散)

という関係があります。したがって求まったパラメータから以下の量を計算すれば各パラメータの分散の絶対値がわかるということになります。

$$s_i = \frac{\sum_{i=0}^{N}(\delta_i^{\,2} w_i)}{w_i(N-P)}$$

この分散の平方根が誤差になります。これらはあくまで a(すなわち逆格子定数)の誤差なので、格子定数の誤差に戻すためにはさらに誤差の伝播を解いていく必要がありますが、根気さえあればできるのでここでは省略します。

ピークのフィッティング方法

本ソフトではニュートン法に似たマルカール(Marquardt)法とよばれる逐次近似法を用いています。この手法は高速な収束速度と安定性を兼ね備えており、十分な精度で最適値を求めることができます。

マルカール法の要点を以下に述べます。n 個の観測点 $(X_1,Y_1),(X_2,Y_2),\cdots,(X_{n-1},Y_{n-1}),(X_n,Y_n)$ に対して m 個のパラメータを含む関数 $F=F(a_1,a_2,\cdots,a_{m-1},a_m,X)$ をフィッティングさせることを考えます。すなわち最も観測値に合うパラメータ a の組を探すことが目的です。関数 F はパラメータ a に対して非線形であってもかまいません。(線形的であれば上に述べた方法で簡単にもとまります。) まず適当な初期パラメータ $(a^0{}_1,a^0{}_2,\cdots,a^0{}_{m-1},a^0{}_m)$ を与えてそのときの二乗残差 R を求めます。

$$R=\sum_{i=1}^{n}[Yi-F(a^0{}_1,a^0{}_2,\cdots,a^0{}_{m-1},a^0{}_m,X_i)]^2$$

次に F を各パラメータで偏微分し、そのときの各パラメータを入力値とするような成分 $\frac{\partial}{\partial a_j}F(a^0{}_1,\cdots,a^0{}_m,X_i)$ を使って以下のような m 行 m 列対称行列 α をつくります。

$$\alpha=\begin{pmatrix} (1+\lambda)\sum_{i=1}^{m}\left[w_i\frac{\partial}{\partial a_1}F(a^0,X_i)\frac{\partial}{\partial a_1}F(a^0,X_i)\right] & .... & \sum_{i=1}^{m}\left[w_i\frac{\partial}{\partial a_1}F(a^0,X_i)\frac{\partial}{\partial a_m}F(a^0,X_i)\right] \\ .... & .... & .... \\ \sum_{i=1}^{m}\left[w_i\frac{\partial}{\partial a_m}F(a^0,X_i)\frac{\partial}{\partial a_1}F(a^0,X_i)\right] & .... & (1+\lambda)\sum_{i=1}^{m}\left[w_i\frac{\partial}{\partial a_m}F(a^0{}_1,X_i)\frac{\partial}{\partial a_m}F(a^0,X_i)\right] \end{pmatrix}$$

ここで λ は対角成分を強調させるための係数です。また w_i は観測点ごとの重み(誤差の二乗の逆数)です。

重みが分からない場合はすべて 1 でもかまいません。この行列 α は「誤差行列」と呼ばれています。その理由は線形の場合、対角要素がパラメータの分散になるからなのですが、詳しい説明は省略します。さらに残差に偏微分をかけたような m 行 *1* 列行列 β

$$\beta=\begin{pmatrix} \sum_{i=1}^{n}[w_i\{Yi-F(a^0,X_i)\}\frac{\partial}{\partial a_1}F(a^0,X_i)] \\ .... \\ \sum_{i=1}^{n}[w_i\{Yi-F(a^0,X_i)\}\frac{\partial}{\partial a_m}F(a^0,X_i)] \end{pmatrix}$$

を計算します。この二つの行列を使って新しいパラメータ $(a'_1,a'_2,\cdots,a'_{m-1},a'_m)$ を以下のように求めます。

$$\begin{pmatrix} a'_1 \\ \cdots \\ \cdots \\ a'_1 \end{pmatrix}=\begin{pmatrix} a^0{}_1 \\ \cdots \\ \cdots \\ a^0{}_m \end{pmatrix}+\alpha^{-1}\beta$$

次にこの新しいパラメータを用いて二乗残差R' を計算します。$R' < R$ であればa'を新しい初期値a^0とし、またR' を 新しいRとして採用し、さらにλを 0.1~0.5 倍します。$R' > R$ であれば新しいパラメータを破棄し、λを 2~10 倍します。そしてα, βを計算するところに戻ります。最後にRの変化が適当に小さくなってきたところで計算を打ち切れば終了です。

この手法の肝要な点は誤差行列の対角成分の重み付けを変化させながら次々と最適化を進めていく点です。F が非線形な関数でも、初期値が真の値に近いときには F のテイラー展開が一次で十分なため、線形的な取り扱いができます。このときは自動的にλの値が 0 に近くなるように設定され、対角項と非対角項を平等に扱います。一方初期値が真の値から遠いときは、線形的な取り扱いができないため、ベクトル解析でいうところの grad R に沿ってパラメータを動かすほうが懸命です。これはλの値を大きくして対角項を強調することを意味しています。つまり、上で述べたアルゴリズムは逐次計算に成功するとより大胆に(=線形的に)パラメータを動かし、失敗するとより小さく勾配(grad)に沿ってパラメータを動かす手法であるといえます。適切な初期値を選べばかなり高速に収束します。

ピークの関数形

フィッティング関数として現在のところガウス関数とローレンツ関数の混合である Symmetric Pseudo Voigt、確率密度分布関数のひとつである Symmetric PearsonVII、さらに PearsonVII を非対称に拡張した Split Pearson VII を利用することができます。計算速度と収束安定性の点からデフォルトでは Symmetric Pseudo Voigt 関数を採用しています。

具体的な関数の形は以下のとおりです。

・Symmetric Pseudo Voigt :

$$f(x,\eta,Hk)=\frac{2}{Hk\cdot\pi}\left(\eta\frac{1}{(1+4\{\frac{x}{Hk}\}^2)}+(1-\eta)\sqrt{\pi\ln 2}\cdot 2^{-4\left(\frac{x}{Hk}\right)^2}\right)$$

・Split Pseudo Voigt (Toraya 1990 改):

$$f(x,\eta_l,\eta_h,A,Hk)=\frac{2}{Hk\cdot\pi+\frac{Hk\cdot\pi\cdot(\eta_h-\eta_l)(Z-1)}{(1+e^{A})(\eta_h+Z(1-\eta_h))}}\left(\eta_l\frac{1}{1+(1+e^{-A})\left(\frac{x}{Hk}\right)^2}+(1-\eta_l)Z\cdot 2^{-4(1+e^{-A})\left(\frac{x}{Hk}\right)^2}\right)\quad (x<0)$$

$$f(x,\eta_l,\eta_h,A,Hk)=\frac{2}{Hk\cdot\pi+\frac{Hk\cdot\pi\cdot(\eta_l-\eta_h)(Z-1)}{(1+e^{-A})(\eta_l+Z(1-\eta_l))}}\left(\eta_h\frac{1}{1+(1+e^{A})\left(\frac{x}{Hk}\right)^2}+(1-\eta_h)Z\cdot 2^{-4(1+e^{A})\left(\frac{x}{Hk}\right)^2}\right)\quad (x>0)$$

・Symmetric Pearson VII: $$f(x,R,Hk)=\frac{2\sqrt{2^{1/R}-1}\Gamma(R)}{\sqrt{\pi}\Gamma(R-1/2)Hk}\left((1+4(2^{1/R}-1)\{\frac{x}{Hk}\}^2)\right)^{-R}$$

・Split Pearson VII (Toraya 1990 改)

$$f(x,R_l,R_h,A,Hk)=\frac{2(1+e^A)}{Hk\sqrt{\pi}}\left(e^A\frac{\Gamma(R_l-1/2)}{\Gamma(R_l)\sqrt{2^{1/R_l}-1}}+\frac{\Gamma(R_h-1/2)}{\Gamma(R_h)\sqrt{2^{1/R_h}-1}}\right)^{-1}\left(1+(1+e^{-A})^2(2^{1/R_l}-1)\left(\frac{x}{Hk}\right)^2\right)^{-R_l} \qquad (x<0)$$

$$f(x,Rl,Rh,A,Hk)=\frac{2(1+e^A)}{Hk\sqrt{\pi}}\left(e^A\frac{\Gamma(R_l-1/2)}{\Gamma(R_l)\sqrt{2^{1/R_l}-1}}+\frac{\Gamma(R_h-1/2)}{\Gamma(R_h)\sqrt{2^{1/R_h}-1}}\right)^{-1}\left(1+(1+e^{A})^2(2^{1/R_h}-1)\left(\frac{x}{Hk}\right)^2\right)^{-Rh} \qquad (x>0)$$

但しHk: 半値全幅, π: 円周率, η: ローレンツ関数とガウシアン関数の割合、Γ: ガンマ関数, R,Rl,Rh: ピアソン関数の指数部, A: 非対称パラメータ。

すべての関数の積分値は 1 に規格化してあります。式から自明なように前半の 2 つは x=0 を中心に対称な形をしていますが、最後の関数は x の符号によって形を変えて、非対称性を表現できるように拡張されています。実際のピーク形状は中心非対称で、低角側に尾を引くような形状を示すことが知られています。PearsonVII には係数部分にガンマ関数が含まれています。ガンマ関数は階乗を実数領域に拡張したような関数です。関数のフィッティングの良さ(残差が少なさ)は PeasonVII のほうが一般に優れているようですが、収束の安定度は PseudoVoigt のほうが優れているようです。

上記の関数群は、実際の計算ではバックグラウンドを考慮して、

$$F(\theta,\Theta,B_1,B_2)=I\cdot f(\theta-\Theta)+B_1+B_2(\theta-\Theta)$$

但し I:積分強度, B:バックグラウンド(一次関数)と拡張します。Θ はピークの中心位置で、θが観測値です。すなわち、あたえられた範囲内で $R=\sum(Y-F)^2$ が最小になるように各パラメータを変化させることでフィッティングを行います。

各関数の偏微分は複雑なので参考のため以下に解析結果を載せておきます。

Symmetric Pseudo Voigt: $$F(\theta,\Theta,\eta,Hk,I,B_1,B_2)=\frac{2}{Hk\cdot\pi}\left(\eta\frac{1}{(1+4\{\frac{\theta-\Theta}{Hk}\}^2)}+(1-\eta)\sqrt{\pi\ln2}\cdot e^{-4\ln2\left(\frac{\theta-\Theta}{Hk}\right)^2}\right)+B_1+B_2(\theta-\Theta)$$

$$\frac{\partial F}{\partial I}=\frac{2}{Hk\cdot\pi}\left(\eta\frac{1}{(1+4\{\frac{\theta-\Theta}{Hk}\}^2)}+(1-\eta)\sqrt{\pi\ln 2}\cdot e^{-4\ln 2\left(\frac{\theta-\Theta}{Hk}\right)^2}\right)$$

$$\frac{\partial F}{\partial \eta}=\frac{2I}{Hk\cdot\pi}\left(\frac{1}{(1+4\{\frac{\theta-\Theta}{Hk}\}^2)}-\sqrt{\pi\ln 2}\cdot e^{-4\ln 2\left(\frac{\theta-\Theta}{Hk}\right)^2}\right)$$

$$\frac{\partial F}{\partial Hk}=-\frac{2I}{\pi}\left(\eta\frac{\{Hk^2-4\{\theta-\Theta\}^2\}}{\{Hk^2+4\{\theta-\Theta\}^2\}^2}+(1-\eta)\frac{\sqrt{\pi\ln 2}}{Hk^4}e^{-4\ln 2\left(\frac{\theta-\Theta}{Hk}\right)^2}\right)$$

$$\frac{\partial F}{\partial \theta_{calc}}=\frac{16(\theta-\Theta)I}{Hk^3\pi}\left(\eta\frac{1}{\left(1+4\{\frac{\theta-\Theta}{Hk}\}^2\right)^2}+(1-\eta)\ln 2\sqrt{\pi\ln 2}\cdot e^{-4\ln 2\left(\frac{\theta-\Theta}{Hk}\right)^2}\right)+B_2$$

$$\frac{\partial F}{\partial B_1}=1$$

$$\frac{\partial F}{\partial B_2}=\theta-\Theta$$

Pearson VII は係数部分にガンマ関数が含まれており、その偏微分を真っ向から解くと計算速度、精度の点で問題が残ります。そこで、積分強度の規格化はとりあえず気にせず、係数項と積分強度の積をまとめて I とおいて最適化してしまいます。簡単な微分($I,B1,B2$)は省略します。

Symmetric Pearson VII: $f(\theta,\Theta,R,Hk,I,B_1,B_2)=I\left((1+4(2^{1/R}-1)\{\frac{\theta-\Theta}{Hk}\}^2)\right)^{-R}+B_1+B_2(\theta-\Theta)$

$$\frac{\partial F}{\partial \Theta}=\frac{8m\cdot I\cdot(\theta-\Theta)}{Hk^2}(2^{1/R}-1)\left(1+4(2^{1/R}-1)\left(\frac{\theta-\Theta}{Hk}\right)^2\right)^{-1-R}$$

$$\frac{\partial F}{\partial Hk}=\frac{8m\cdot I\cdot(\theta-\Theta)^2}{Hk^3}(2^{1/R}-1)\left(1+4(2^{1/R}-1)\left(\frac{\theta-\Theta}{Hk}\right)^2\right)^{-1-R}$$

$$\frac{\partial F}{\partial m}=\left(1+4(2^{1/R}-1)\left(\frac{\theta-\Theta}{Hk}\right)^2\right)^{-R}\left(\frac{4\cdot 2^{1/R}(\theta-\Theta)^2 Log(2)}{m(Hk^2+4(2^{1/R}-1)(\theta-\Theta)^2}-Log(1+4(2^{1/R}-1)\left(\frac{\theta-\Theta}{Hk}\right)^2)\right)$$

Split Pearson VII: $$Fl(\theta,\Theta,Rl,Rh,A,Hk,I,B_1,B_2)=I\left(1+(1+e^{-A})^2(2^{1/Rl}-1)\left(\frac{\theta-\Theta}{Hk}\right)^2\right)^{-Rl}+B_1+B_2(\theta-\Theta)\qquad (x<0)$$

$$Fl(\theta,\Theta,Rl,Rh,A,Hk,B_1,B_2)=I\left(1+(1+e^{A})^2(2^{1/Rh}-1)\left(\frac{\theta-\Theta}{Hk}\right)^2\right)^{-Rh}+B_1+B_2(\theta-\Theta)\qquad (x>0)$$

3 次スプラン関数の解き方

バックグラウンド曲線を引くためにPDIndexerでは3次スプライン曲線を利用しています。本当のバックグラウンド曲線の形はもちろん知ることはできませんが、本ソフトでは、ピークでない部分を自動で検出し、検出した点をスプラインで結んでバックグラウンド曲線とします。スプラインは，f(x)が十分に滑らかな場合，その導関数も含めて一様に近似し，データ点を密にすれば近似は上がることが知られています。

3次スプラインの定義は、n個の点$(X_1,Y_1),(X_2,Y_2),\cdots,(X_{n-1},Y_{n-1}),(X_n,Y_n)$が与えられたとき各区間は3次関数で、かつ各点で、値、傾き、曲率が等しくなるような曲線です。（ただし、区間$\{-\infty_1,X_1\},\{X_n,\infty\}$は1次関数）

区間$\{X_{m-1},X_m\}$の関数を$F_m=a_mX^3+b_mX^2+c_mX+d_m$とすると、

2 <= m <= n-1 のとき

$F_m(Xm)= a_mX_m^3+b_mX_m^2+c_mX_m+d_m =Y_m$

$F_{m+1}(X_m)= a_{m+1}X_m^3 + b_{m+1}X_m^2 + c_{m+1} X_m+d_{m+1} =Y_m$

$F_m'(X_m)= 3a_mX_m^2 + 2b_mX_m + c_m= F_{m+1}'(X_m)=3a_{m+1}X_m^2 + 2b_{m+1}X_m + c_{m+1}$

$F_m''(X_m)= 6a_mX_m + 2b_m = F_{m+1}''(X_m) = 6a_{m+1}X_m + 2b_{m+1}$

(4n-8 個の条件)

m=1 のときは

$F_1(X_1)= c_1X_1+d_1 = Y_1$

$F_2(X_1)= a_2X_1^3 + b_2X_1^2 + c_2X_1 + d_2 =Y_1$

$F_1'(X_1) = c_1= F_2'(X_1) = 3a_2X_1^2 + 2b_2X_1 + c_2$

$F_1''(X_1)= F_2''(X_1)=6a_2X_1+2b_2 = 0$

(4 個の条件)

m=nのときは
$F_n(X_n)= anXn^3 + bnXn^2 + cnXn + dn = Yn$
$F_{n+1}(X_n)= c_{n+1}X_n + d_{n+1} = \mathrm{Y}_\mathrm{n}$
$F_n'(X_n)= 3a_nX_n^2 + 2b_nX_n + c_n = F_{n+1}'(X_n)= c_{n+1}$
$F_n''(X_n)= 6a_nX_n + 2b_n = F_{n+1}''(X_n) = 0$
(4 個の条件)

結局、合計 4n 個の条件を使って 4n 個の変数を求める連立方程式問題に帰結します。行列で書き下して逆行列問題にすれば簡単に解くことが出来ます。

状態方程式

本ソフトでは標準的な物質についての状態方程式を内蔵しています。物性物理は難しくてよくわかりませんが、よく使われているのは Birch-Murnaghan 方程式や Mie-Gruneisen 方程式は解析的に解くようにしていますが、