ベイズの定理に基づく最大エントロピー法による エネルギー分散型 X 線スペクトル分析

新川 隆 朗*・中村奈津子*・加藤 裕 志*
渋谷 和 好**・中田 宗 隆***

*日本電子株式会社 昭島市武蔵野 3-1-2 (〒196-8558)

**日本電子エンジニアリング株式会社 昭島市武蔵野 3-1-2 (〒196-8558)

***東京農工大学 大学院生物システム応用科学研究所 小金井市中町 2-24-16 (〒184-8588)

(2005年 3 月 7 日 受理)

Applications of a Bayesian Based Maximum Entropy Method to Energy Dispersive X-ray Spectra

Takao SHINKAWA*, Natuko NAKAMURA*, Hiroshi KATO*,
Kazuyoshi SHIBUYA**, and Munetaka NAKATA***

*JEOL Ltd., Musashino, Akishima 196-8558

**JEOL ENGINEERING CO. Ltd., Musashino, Akishima 196-8558

***Graduate School of BASE, Tokyo University of Agriculture & Technology,
Naka-cho, Koganei 184-8558

(Received March 7, 2005)

An energy dispersive X-ray detector is widely used to measure intensity spectra of characteristic X-rays of component elements in a specimen. In a long-time measurement to obtain a low-noise spectrum, the spatial resolution of X-rays decreases partly due to the radiation of X-rays from the surrounding area larger than the probe size itself, and partly due to the drift of the electron beam and/or of the specimen stage. In a short-time measurement to avoid the above difficulties, the X-ray spectrum contains a much higher-level noise. It has been well known that usual smoothing techniques downgrade the sharpness of the spectrum. In the present study, a Bayesian Based Maximum Entropy Method (BBMEM) is applied to the spectrum obtained in a few seconds, and is proved to be an effective technique to reduce the noise in the spectrum.

Keywords: Bayesian Theory, Maximum Entropy Method, Energy Dispersive X-ray Spectrum

1. はじめに

電子顕微鏡で電子ビームを試料上の局所領域に照射すると、連続 X 線と特性 X 線が試料表面より放出される (Fig. 1). エネルギー分散型 X 線分光 (EDS; Energy Dispersive X-ray Spectroscopy) は、この X 線をエネルギー検出器によって検出し、エネルギー値ごとの X 線の強度を測定し、このスペクトルから特性 X 線のピークを検出して構成元素を特定する。通常、電子ビーム径は数 μm ～ 数十 μm という小さな大きさであり、電流量も数 nA 程度と微小であり、検出される X 線強度は低く、短時間の測定ではホワイトノイズを主成分とする種々のノイズを含むスペクトルとなる。一般に、このホワイトノイズを減少さ

せて十分な S/N 比を得るためには、長時間の測定が有効であると考えられる。しかし、電子顕微鏡で X 線スペクトルを測定する場合には、照射領域の周辺領域からの励起 X 線による分解能の劣化、ビームのドリフト、あるいは、試料台のドリフト等による空間分解能の劣化などによる再現性のない予測困難なノイズが含まれるようになる¹⁾、スペクトルの劣化が避けられない。できるならば、短時間でスペクトルを測定し、含まれるノイズをホワイトノイズが主たるノイズ成分とみなして、それらのノイズを減少させて質の良いスペクトルを得たい。

本研究では、短時間で測定したスペクトルに含まれるホワイトノイズを減少させるために、ベイズの定理に基づく最大エントロピー法を適用する。測定時間 5 s で測定した

Fig. 1 Various radiations and electrons induced by electron beam.

希土類含有試料のスペクトルに対して最大エントロピー法による変換を行い、測定時間30 sで測定したスペクトルのS/N比との比較から、最大エントロピー法のノイズ減少能力の評価を行う。また、5 s測定のスペクトルに対して、さらに人工的なホワイトノイズ（ランダムノイズ）を付加した後に最大エントロピー法による変換を行い、この変換が元のスペクトルをどの程度復元できるかを調べる。また、最大エントロピー法の収束条件を変更することによる計算結果への影響を調べ、妥当な収束条件を決定する方法を検討する。

2. 理 論

最大エントロピー法は様々な手法が提案されている。ここではベイズの定理に基づく方法（**BBMEM**法；**Bayesian Based Maximum Entropy Method**）²⁻⁵⁾を採用した。**BBMEM**法は識別関数としてスペクトルのエントロピーを用いたものである。はじめにベイズの定理および**BBMEM**法を導出する。ベイズの定理の式は、

$$P(f|D) = \frac{P(D|f)P(f)}{P(D)} \quad (1)$$

と定義される。左辺 $P(f|D)$ を観測値 D が与えられた時の仮説 f の事後確率と呼び、対する右辺の $P(f)$ を仮説 f の事前確率と呼ぶ。右辺の $P(D|f)$ は尤度と呼ばれ、ある仮説 f のもとで観察された値 D が生じる確率を意味する。分母 $P(D)$ は仮説 f に依存しない定数である。

ベイズの定理とは「ある仮説 f の事前確率と尤度の積が f の事後確率になる」ということを意味している。しかし、事前確率は一意的に与えられず、ある程度の主観的な推定によらざるを得ないという問題がある。そこで、本研

究では、以下のように述べる最大エントロピー法を適用することにした。

式(1)の中の尤度 $P(D|f)$ にガウス分布を仮定すると

$$\begin{aligned} P(D|f) &= \prod \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left[-\frac{(D_i - F_i(f_i))^2}{2\sigma_i^2} \right] \\ &= \frac{1}{\prod \sqrt{2\pi\sigma_i^2}} \exp \left[-\sum \frac{(D_i - F_i(f_i))^2}{2\sigma_i^2} \right] \\ &\propto \exp \left(-\frac{\chi^2}{2} \right) \end{aligned} \quad (2)$$

$$\chi^2 = \sum \frac{(D_i - F_i(f_i))^2}{2\sigma_i^2} \quad (3)$$

となる。 D_i と f_i はそれぞれが n 次のベクトル $D = (D_1, D_2, \dots, D_i, \dots, D_n)$ と $f = (f_1, f_2, \dots, f_i, \dots, f_n)$ であることを示し、 F_i は f_i に対する装置関数を意味する。 σ_i は i 番目の観測値 D_i が装置関数 $F_i(f_i)$ と誤差 σ_i の範囲で一致する条件として導入している。

同様に事前確率 $P(f)$ はpositive and additive distribution（positiveは事象が常に起きているので確率 $P(f)$ が負にならないこと、additiveは事象の生起が常に積算されることを示している）を仮定するとポアソン分布となり、

$$\begin{aligned} P(f) &= P(f_1, f_2, \dots | m_1, m_2, \dots) \\ &= \prod \frac{m_i^{f_i} e^{-m_i}}{f_i!} \approx \prod \frac{m_i^{f_i} e^{-m_i}}{\sqrt{2\pi f_i} f_i^{f_i} e^{-f_i}} \\ &\approx \prod \frac{1}{\sqrt{2\pi f_i}} \exp \left(\sum f_i - m_i - f_i \log \frac{f_i}{m_i} \right) \\ P(f) &\propto \exp(\alpha S) \end{aligned} \quad (4)$$

$$S = \sum \left(f_i - m_i - f_i \log \frac{f_i}{m_i} \right) \quad (5)$$

が得られる。ここで、 m は f の初期（またはデフォルト）モデルと呼ばれ、測定結果が存在していない段階（Prior）での f の値として想定されているものである。また、 α は正規化定数、 S は情報エントロピーと呼ばれている。

式(1)、(2)および(4)より

$$P(f|D) \propto P(D|f)P(f) = \exp \left[\alpha S - \frac{1}{2} \chi^2 \right] \quad (6)$$

となる。**BBMEM**法では、この仮説 f が真値であるものを求める。真値である根拠として、観測値 D が与えられたときの真値 f の確率 $P(f|D)$ が最大になることを挙げている。すなわち、 $P(f|D)$ の最大値を求めるために $\alpha S - (1/2)\chi^2$ の最大値を求めることが**BBMEM**の原理となる。なお、この式は $(1/2)\chi^2$ がラグランジェ未定乗数法の拘束条件となっている。

$$\frac{\partial}{\partial f} \left[\alpha S - \frac{1}{2} \chi^2 \right] = 0 \quad (7)$$

一般に、逆問題を解く解法では、数値解析的に逆関数が求まらない場合がある。一方、**BBMEM**法では、式(3)および式(7)から明らかなように、観測値 D に対して、真値 f との間に許容範囲である誤差 σ を想定し、逆関数 F^{-1} でなく装置関数 F を仮定することによって真値を推定することができる。なお、式(7)の χ^2 の項が誤差 σ に関する拘束条件に当たっている。

3. 結果と考察

3.1 EDS スペクトルへの応用

希土類として Nd を含有する試料に対して、5 s 間の EDS スペクトルの測定を行った (Fig. 2)。縦軸は比較のため、X 線計数率 (count/s) を対数で表示している。さらにこのスペクトルに対して、**BBMEM** 変換を行った (Fig. 3)。

Fig. 2 EDS spectrum of a specimen that includes rare metal Nd. Accelerating Voltage: 20 keV; Beam Current: 0.605 nA; Measuring Time: 5 s.

Fig. 3 Improved spectrum of the spectrum in Fig. 2 by BBMEM. The noise level is lower than that of Fig. 2 and degrading of the sharpness of each peak is avoided.

Fig. 2 と比較すると、Fig. 3 のピークは崩れずに、ノイズが減少していることがわかる。図中に示す定性分析 (ピーク位置より元素名を同定し、表示するソフトウェア) の結果も良好である。

BBMEM 計算の効果を検証するため、さらに測定時間 30 s のスペクトルとの比較を行った。Fig. 4 は、EDS で 5 s 間測定した結果のスペクトル (a) と、(a) に **BBMEM** 法を適用して得たスペクトル (b)、および同一サンプルを 30 s 測定したスペクトル (c) を比較したものである。エネルギーの表示範囲は Fig. 3 の 0~3.0 keV を拡大表示している。スペクトル (b) では、スペクトル (a) よりもノイズが少なく、30 s の測定のスペクトル (c) のノイズとほとんど変わらないことがわかる。また、スペクトル (c) がスペクトル (b) と異なる原因としては、測定時間が長いためにホワイトノイズ以外のノイズが含まれていることが考えられる。

Fig. 5 には、同じく Nd 含有試料で 5 s 測定結果と 30 s

Fig. 4 Qualitative analyses of a specimen that includes rare metal Nd (case 1). Accelerating Voltage: 20 keV; Beam Current: 0.605 nA. (a) Original spectrum of measuring time 5 s; (b) Improved spectrum for (a) by BBMEM; (c) Spectrum of measuring time 30 s.

る。したがって、BBMEM法はホワイトノイズを減少させ、ホワイトノイズによる擬似ピークを消し、また、隠れていたシグナルピークを見出す効果があることがわかった。

3.3 計算パラメータの評価

3.3.1 BBMEM法の特長

BBMEM法の特長を把握するために、計算のパラメータを変更して、結果を検討した。Fig. 7は観測値 D が2次元 (d_1, d_2) の場合のBBMEM法による計算の一般的な軌跡を示している。計算の起点 $(f_1, f_2) = (m_1, m_2)$ のときに、 $\alpha = \infty$ かつエントロピー $S = 0$ 、計算の終点 $(f_1, f_2) = (d_1, d_2)$ のときに、 $\alpha = 0$ かつエントロピー $S < 0$ である。ただし (m_1, m_2) は確率変数 (f_1, f_2) のデフォルト値、 (d_1, d_2) は観測値である。実際の計算は起点より出発し、後述する式(8)の収束条件のパラメータ（ここでは定数 A とする。）の値を満足したところで止まる。

BBMEM法では共役勾配法により、非線形最適化問題の近似解を求めている。この方法は反復法でありながら、有限回のステップで厳密解に到達する特長を持つが、ノイズが生じた場合に多くの回数を繰り返しても十分な精度が得られない場合がある。従って、以下で定義するパラメータ A はこの共役勾配法の収束条件を示すものであり、 A を1に近づけると $\alpha \approx 0$ となり (f_1, f_2) が最適値となるように計算される。しかし、同時に、後述するような発振等の問題点が生じる。従って最適な反復計算の収束条件のパラメータ値 A を定めることが重要となる。

パラメータ値 A の定義は³⁾,

$$A \geq \frac{-\alpha S}{\langle H \rangle c^2} \quad (8)$$

$$c^2 = \frac{2(L - \alpha S)}{N}$$

$$L = \frac{\alpha S \chi^2}{\chi^2 - N}, \quad N \text{ はデータの個数}$$

$$\langle H \rangle = \int \text{Pr}_1 \log(\text{Pr}_1 / \text{Pr}_0) d^L f / (\det[\mu])^{\frac{1}{2}}$$

$$\text{Pr}_0(f|D) = \text{const}$$

$$\times \exp \left\{ -\frac{\alpha}{2} (f - \hat{f})^T \left[\mu^{-\frac{1}{2}} \right] B \left[\mu^{-\frac{1}{2}} \right] (f - \hat{f}) \right\}$$

$$\text{Pr}_1(f|D) = \text{const}$$

$$\times \exp \left\{ -\frac{\alpha}{2} (f - \tilde{f})^T \left[\mu^{-\frac{1}{2}} \right] B \left[\mu^{-\frac{1}{2}} \right] (f - \tilde{f}) \right\}$$

である。ただし、

$$[\mu] = (-\nabla \nabla S)^{-1}, \quad B = I + J/\alpha,$$

$$J = \left[\mu^{-\frac{1}{2}} \right] \nabla \nabla L \left[\mu^{-\frac{1}{2}} \right]$$

である。ここで、 I はidentity matrixである。 $\text{Pr}_0(f|D)$

Fig. 7 Maximum entropy trajectory.

は最適な真値 $\hat{f}$ が求まった場合、その近くの f に対する事後確率の分布であり、 $\text{Pr}_1(h|D)$ は $\hat{f}$ から少しずれた値 $\tilde{f}$ を中心とした事後確率の分布を示す。 $\langle H \rangle$ はこの2つの確率分布の重なり度を示す値(cross-entropy)であり、この値は、計測系にノイズとして標準偏差 σ を仮定した場合に、

$$\langle H \rangle = \frac{1}{2} \text{trace}((\alpha B)^{-1} J) \quad (9)$$

となる。 $\langle H \rangle$ が(9)式で与えられるオーダの値となれば、真値 $\hat{f}$ と仮定した確率分布が真の分布に重なっているといえる。また c^2 は標準偏差 σ を仮定するもので、確率分布の値が $c^2 = (2(L - \alpha S)/N)$ で最大となる。計算は各繰返しにおいて、式(8)及び式(9)で求まる $(-\alpha S/\langle H \rangle c^2)$ を計算し、パラメータ A と比較し、 A 以下となった時点で計算を終了させる。

3.3.2 計算パラメータ A の変化

Fig. 2のNd含有試料のスペクトルを用いて、上記の計算パラメータ A を変化させたときのBBMEM計算結果をFig. 8に示す。 I は計算が収束するまでの繰返し回数である。

Fig. 8の(b)の値は、Fig. 7の計算の起点 $(f_1, f_2) = (m_1, m_2)$ の状態を示し、全ての値は初期値(一定値)となっている。 A の値が小さくなるとBBMEM計算結果のスペクトルは観測値(オリジナル・スペクトル)を反映したものとなってくるが、常にエントロピーの最も高い軌跡に沿って、エントロピーを減じながら、観測値 $(f_1, f_2) = (d_1, d_2)$ に近づいていく。この時に、誤差の係数 χ^2 を半径とし、中心を (d_1, d_2) とするFig. 7の破線の同心円の中心に近づくまで繰返し計算を行い、収束させる必要がある。しかしながら、ここで以下のような問題が生じる。

① 限りなく (d_1, d_2) に近づけることには意味がない。

Fig. 8 Transformation of the spectra with various values of the convergence parameter A in BBMEM. (a) Original spectrum; (b) $A=6$ and $I=1$; (c) $A=5.6$ and $I=5$; (d) $A=3$ and $I=9$; (e) $A=1$ and $I=14$.

なぜならば (d_1, d_2) は観測値であり、既知のものであるから。

- ② また A 値 $\ll 1$ とすると計算の繰り返し数 I が増加し、計算が収束しなくなる場合がある。
- ③ また A 値を減じていくと、スペクトル自身が発振し始め、疑似ピークを多発生するようになる（エントロピーの過減少）

従って、適正な繰り返し計算の収束条件である A 値を求めるために、Fig. 9 から Fig. 13 のような検討を行った。Fig. 9 の α の変化では、 α の値は $A < 3$ の範囲で十分に 0 に近づいていることが分かる。 α が 0 に近づくということは (d_1, d_2) に近づいていることを示している。

また、Fig. 10 では、エントロピー S は $3 > A > 1$ の区間で単調に減少し、Fig. 11 では χ^2 の値は $5 > A > 2$ の区間で小さい値であることが分かる。式 (7) から明らかなよう

Fig. 9 Transformation of α value against the convergence parameter A .

Fig. 10 Transformation of entropy S against the convergence parameter A .

Fig. 11 Transformation of χ^2 against the convergence parameter A .

に、エントロピー S があまり減少せずに、 χ^2 が小さい区間で計算を終了することが望ましい。

Fig. 12 の縦軸の値 $\text{Log } P(D|\alpha)$ は、 α の値に対する観測値 $D(d_1, d_2)$ の起生確率の対数値であるが、 $3 > A > 1$ の区

Fig. 12 Transformation of $\text{Log } P(D|\alpha)$ against the convergence parameter A . This value is expected to be maximized in the Bayesian theorem.

Fig. 13 Transformation of iteration number I against the convergence parameter A .

間がその値の大きな区間であることが分かる。ベイズの定理ではこの起生確率を高くすることが望まれる。また、 A 値と繰り返し計算の回数 I の関係を Fig. 13 に示す。

以上の結果から、妥当な A 値として次のように推定することができる。まず、 α の変化から $A < 3$ が、また、エントロピー S の変化から $3 > A > 1$ が期待される。一方、

χ^2 の変化では $5 > A > 2$ が期待される。したがって、今回の **BBMEM** 法の適用事例 (3.1) およびのノイズに対する適応性の評価 (3.2) では、これらの共通する範囲に含まれる $A = 2$ を仮定した。

4. ま と め

今回、電子顕微鏡における X 線分析システムに対して最大エントロピー法を適用し、その効果を検討した。実際に測定したスペクトルに人工的なホワイトノイズを加えて、S/N 比を劣化させた後に **BBMEM** 法を適用して定性分析を行った。その結果、重畳したホワイトノイズを取り除くことができ、擬似ピークを消し、ノイズに埋もれたピークを検出することができた。また、**BBMEM** 法の計算の収束条件を与えるパラメータ (A 値) を小さくするほど (α が 0 に近くなるまで繰り返し計算を行うほど) 結果は観測値に近づくが、あまり小さくなるまで計算すると、結果は発振しはじめ、リングングしはじめることがわかった。 A 値の変化に対する α 値の変化 (Fig. 9)、エントロピーの変化 (Fig. 10)、 χ^2 の変化 (Fig. 11) から A 値の最適値を求める方法を確認する必要があることがわかった。今後、エネルギー分散型 X 線スペクトル以外にも、波長分散型 X 線スペクトル (Wave Dispersive X-ray Spectrum) やオージェ電子スペクトル (Auger Electron Spectrum) にも **BBMEM** 法を応用する予定である。

参 考 文 献

- 1) 日本電子株式会社, 日本電子データ株式会社, 日本電子ハイテック株式会社: SEM の世界への招待.
- 2) S. F. Gull and J. Skilling: *Quantified Maximum Entropy Mem-Sys5 Users' Manual* (Maximum Entropy Data Consultants Ltd., 1999).
- 3) S. C. Lizamore: *Topics in Maximum Entropy Applications* (Master of Science in Statistics and Operations Research, Victoria Univ. of Wellington, 31 Aug. 1995).
- 4) D. S. Sivia: *Bayesian Inductive Inference Maximum Entropy & Neutron Scattering* (Los-Alamos Science Summer, 1990).
- 5) 高田昌樹, 坂田 誠: 電子顕微鏡 38, 104 (2003).