

大規模ベイズモデルに基づく スピンノイズの除去法

統計数理研究所 樋口知之

(1993年7月 受付)

1. 研究の背景

人工衛星は、広範囲を一樣かつ継続的に観測できる唯一の手段として、地球環境モニタリングで重要な役割を果たしている。しかし、人工衛星データからの情報の抽出には、大きく2つの問題点がある。第一が、観測手段がリモートセンシングであるために、データには、単純な観測ノイズの他に、推測結果に破滅的な影響を及ぼす可能性のある分離が容易でない系統的なノイズが含まれること、すなわち、解くべき問題が性質の悪い逆問題になることである。特に、人工衛星データには、姿勢安定に必要な回転をもたらす人工的ノイズ——スピンノイズ——が程度の差はあれ必ず付随し、最悪の場合、情報がノイズからの偽情報に埋没する (Higuchi et al. (1988))。また、画像のような2次元情報 (イメージ) も、このスピンノイズの影響から避けられない。なぜならば、画像は普通計測機の2次元の走査を利用した1次元情報の組合せで構成されており、2次元の走査の内一つは通常スピン運動そのものによって行われるからである (Cracknell (1981), Kosiishi (1991))。

第二の問題点は、一般に同じ条件下でのデータの取得がほぼ不可能なことから、人間の経験と直感に基づく情報処理の介入を必要とし、データのバッチ的な高速大量処理が非常に難しいことである。スピンノイズの除去法についても、その出現形態の観測場所・時間などの様々な状況に応じた多様性により、同じ問題が生ずる。従来は、既存のノイズ除去法を数多くのデータに適用して形造られた直感と、経験豊富な指導者の洞察力とを判断基準として、一つ一つのデータセットを処理する (Mukai et al. (1990)) という人海戦術的なアプローチがされてきた (Higuchi (1992))。そのため、スピンノイズ除去の on-line 処理はもちろん、off-line での高速大量バッチ処理への可能性も妨げられている。

上記の問題には、データの持つ情報を最大限活かすようなモデリングとデータ解析法——大規模ベイズモデル——が有効である。簡単にその特徴をまとめると、自在にデータを表現できる柔軟性に富むモデルの採用、つまり数多くのパラメータを持つ大規模なモデルの採用がある。また、人間の知識獲得のプロセスそのものがある種の主観的要因を含んでいる事実をふまえて、データ解析の全体の枠組みを構成している点が挙げられる。考えうる多数のモデルからの最適なモデルの客観的同定には、AICの流に沿うベイズアプローチの枠組みを取り込んだ情報量規準を利用している (Akaike (1980))。

本研究の目的は、Higuchi et al. (1992) の提案したスピンノイズ除去法のモデルを一般化し、*ad-hoc* なパラメータを除くといった意味でのさらなる自動化を図ることである。提案したモデルの画像データへの適用も同時に示す。

2. データの説明

前述したように人工衛星には、衛星本体の中心軸のまわりの規則正しい自己回転運動（スピンの）が与えられている。この中心軸を普通スピン軸といい、その軸周りの角度 $\theta(n)$ を位相と呼ぶ。 n は時刻。 $\theta(n)$ は、普通太陽や明るい星などを reference として正確に計測できるので、測定値に含まれる誤差はほとんど考えなくて良いデータである。一般に人工衛星データの場合、 $\theta(n) - \theta(n-1)$ の値は定数だが、本研究の方法においてはそれを条件とはしていないことを付言しておく。極言すれば、ランダムであってもよい。

本研究の最終的な目的は、時刻 n での 1 変量観測値 $y(n)$ と $\theta(n)$ のセットからなるデータ $(y(n), \theta(n))$ ($n=1, \dots, N$) から、系統的ノイズ項である background noise $B(n)$ を推定し、それを除去することで、真の信号——つまり、自然現象による信号—— $x(n)$ を推測することである。 N はデータ数。

実際の観測量は多変量 $\mathbf{y}(n) = [y_1(n), \dots, y_m(n), \dots, y_M(n)]^T$ ($n=1, \dots, N$) であることが多い。本研究での応用データ例のように、各成分（通常チャンネルとかバンドと呼ぶものに相当）に含まれる background noise $B_m(n)$ ($m=1, \dots, M$) は、普通チャンネル毎に大きく変化するので、その場合データ $(y_m(1), \dots, y_m(N))$ を用いて $B_m(n)$ を推定することになる。 $B_m(n)$ が各チャンネルに共通の場合は、全てのデータを使って共通の background noise の形 $B(n)$ を推定すればよい。また analytic な関数で $B_m(n)$ の間の関係が書ける場合、つまり関数 $F(m)$ を用いて $B_m(n)$ が $B_m(n) = B(n) \cdot F(m)$ (もしくは $B_m(n) = B(n) + F(m)$) という形で表現できる場合は、 $F(m)$ の効果を考慮しながら全データを使って、提案する方法で $B(n)$ を推定することも可能である (Higuchi (1991, 1994))。

3. モデルの説明

3.1 観測モデル

実測値 $y(n)$ は、測定の様子が特殊でない限り、 $B(n)$ と $x(n)$ の和である。つまり、観測値 $y(n)$ は次のように分解できると仮定できる。

$$(3.1) \quad y(n) = B(n) + x(n)$$

この仮定の背後には、分解された各成分 $B(n)$, $x(n)$ がおのおの物理的意味を明確に持ち、和算が実際の観測メカニズムを反映するような意図があることを強調しておく。単に解法の単純さを狙ったものではない。

background noise は θ に依存しているので、以下 $B(n, \theta = \theta(n))$ と表記する。本研究で取り扱ったデータの $B(n, \theta)$ として次のようなものを考えた。

$$(3.2: \text{モデル } M) \quad B(n, \theta) = w(n)t(n)f(\theta)$$

ここで、 $t(n)$ はゆっくりと変動する background noise の強さ、つまり intensity のトレンド (envelope の様なもの) を表す。 $f(\theta)$ は、background noise の θ 依存性を表現する任意の関数、いかえれば background noise の位相応答関数——*phase response function*——である。 $w(n)$ は、時間相関の無い background noise の揺らぎの成分に相当する。物理的にはこのモデルは、 $t(n)f(\theta(n))$ で表現される background noise が、 $w(n)$ で modulate (モデル M の由来) されているイメージを具象化している。

本研究で取り扱う人工衛星データには不適当だが、一般には background noise が以下のよ

うに分解できるモデルも考えうる。

$$(3.2: \text{モデル } L) \quad B(n, \theta) = w(n) + t(n) + f(\theta)$$

L は, $w(n)$ が線形に付加されていることを表す。以下の文章で, L は linear model, M は modulation model を意味することを注意しておく。また式番号で L 及び M が付いているものは, 文字に対応したモデルのみに成立する式であることを明示する。

3.2 phase response function の数値的表現

$f(\theta)$ の数値的表現方法としては, phase response が人工衛星の位置変化にともない多彩な様相を呈することを踏まえると, 柔軟性の高いものが必要である。本研究では, θ 空間を $\theta = i\Delta\theta$ ($i=1, \dots, 360/\Delta\theta$) のように離散化し, $\Delta\theta$ を充分細かくすることにより, $\theta_i = [(i-1)\Delta\theta, i\Delta\theta)$ の範囲の $f(\theta)$ を f_i の値で代表させる方法を採用した。具体的には $\Delta\theta = 1$ とし, よって推定するのは, 360 個の f_i の値である。

3.3 事前情報の設定

多数のパラメーター, f_i ($i=1, \dots, 360$) と $t(n)$ ($n=1, \dots, N$), を一意に定めるには, 自然な事前情報の利用が必要である。 $t(n)$ は noise intensity の非常に緩やかに変動する成分であるから, smoothness prior を設定した。phase response function は, 周期関数でかつ局所的には滑らかなものであることが期待できるから, f_i には periodic smoothness constraint を課した。periodic smoothness constraint とは, f_{360} と f_1 との間の滑らかさも要求する拘束条件である。

さらに, $t(n)$ と $f(\theta = \theta(n))$ の分離を一意に定めるため, f_i の下限をモデル M の場合は 1 に, またモデル L の場合は 0 に設定している。

4. 推定方法の概要

方法自体がかなり複雑でこみいっているため, 本章で推定方法の概要を示す。

4.1 $x(n)$ の分布について

3章により, 我々は観測値を次のような式でモデル化したことになる。

$$(4.1M) \quad y(n, \theta(n)) = w(n)t(n)f(\theta = \theta(n)) + x(n)$$

$$(4.1L) \quad y(n, \theta(n)) = w(n) + t(n) + f(\theta = \theta(n)) + x(n)$$

本研究の目的は, $f(\theta)$ の推定と (4.1) に示されるような $y(n)$ の成分分解であるが, $x(n)$ の存在が問題を非常に難しくしている。真の信号は常時あらゆる方向から受信されているのではなく, ある一定の時間帯に, ある位相範囲に観測される。しかしながら, その信号出現を特徴付けるような n と θ 上での分布は全く解らない。ただ, 正值をとるという事だけが解っている。

4.2 データの選別

background noise の成分しか含まないデータの集まりを, グループ- B (以後 BG と略す), また $x(n)$ を含んでいるデータの集まりを, グループ- $B+x$ (以後 BX と略す) と呼ぶことにする。データが BG あるいは BX のどちらに帰属するかを選別するために, 次のような $y(n)$ から $t(n)$ と $f(\theta = \theta(n))$ の効果を除いた量, つまり系統的誤差除去量を考える。モデル L の場合

$$(4.2L) \quad \begin{aligned} e(n) &= y(n) - t(n) - f(\theta = \theta(n)) \\ &= w(n) + x(n) \end{aligned}$$

で定義する。BG の $e(n)$ がつくる分布は、 $w(n)$ に一致する。一方、BX の $e(n)$ は、比較的大きい正値をとる $x(n)$ により、 $w(n)$ の分布から大きく外れる。この点をデータを判別する基本的なアイデアとする。

モデル M の場合は、 $y(n)/t(n)f(\theta = \theta(n))$ のかわりに以下のような量

$$(4.2M) \quad \begin{aligned} e(n) &= \log y(n) - \log t(n) - \log f(\theta = \theta(n)) \\ &= \log w(n) + \log \left(1 + \frac{x(n)}{w(n)t(n)f(\theta = \theta(n))} \right) \end{aligned}$$

を考える。BG の $e(n)$ がつくる分布は、モデル M の場合 $\log w(n)$ の分布に一致する。BX の $e(n)$ は、 $\log w(n)$ の分布から大きく外れる。

4.3 推定のアルゴリズム

$x(n)$ の直接的な推定は困難なので、 $e(n)$ を考えデータを選別し、 $f(\theta = \theta(n))$, $t(n)$, そして $w(n)$ の分解を行う。つまり、式 (4.1) に示されるような、同時的な成分分解を行うのではなく、データの選別と平滑化法を組み合わせ、逐次的に成分の推定を行う。図1に解法のアルゴリズムを示す。

5. パラメーターの推定

5.1 ステップ0: $t(n)$ の初期推定

ある区間 θ_i に $\theta(n)$ が含まれる観測値 $y(n)$ のみに注目し、新しく構成された時系列、 $\{y(n) | \theta(n) \in \theta_i\}$ 、を適当に内挿・平滑化したものは、 $t(n)$ の最初の推定値として充分であろう。ただこの方法が有効であるためには、設定する区間 θ_i にはほとんど $x(n)$ が受信されてはならない、という条件がある。前述したように、前もってこのような特殊な θ_i は知る由もないので、本研究ではこの方法は採用できない。

ある時間内 T に観測されたデータに注目する。 $x(n)$ の正値性から、 $\theta(n)$ が $[0^\circ, 360^\circ)$ の範囲に満遍無く分布する程度に T を長くすれば、その時間内 T に測定された観測値の最小値は BG に属し、その時間内の $w(n)t(n)f(\theta)$ (モデル L の場合は、 $w(n) + t(n) + f(\theta)$) の下限を与えているものと見積っても差し支えない。この最小値がつくる時系列のトレンドをもって、 $t(n)$ の初期推定とする。

本研究では人工衛星が1回転する時間を T とする。取り扱ったデータの $\theta(n) - \theta(n-1)$ は約 6.9° なので、約52個毎に、 $\theta(n)$ は $[0^\circ, 360^\circ)$ の範囲を動く。この52個の中の最小値 $y(n)$ を、この時間内のトレンドの代表値、 $\hat{t}(n) = y(n)$ とする。最小値をとらなかった残り(約51個)のデータの $\hat{t}(n)$ は、欠損値とする。

$\hat{t}(n)$ を内挿・平滑化することで、 $t(n)$ の初期推定値 $t_0(n)$ を得る。ただし、モデル M の時は後での諸処の計算の都合上、 $\hat{t}(n)$ ではなく $\log \hat{t}(n)$ を平滑化する。この後、ステップ数 k を $k=1$ にセットする。内挿・平滑化の方法には、Akaike (1980) によって提案された、離散スプラインを用いる (Tanabe and Tanaka (1983))。このモデルは、システムノイズ、観測ノイズともガウスであり、またシステム及び観測モデルとも線形なので、カルマンフィルターと平滑化のアルゴリズムが適応できる (Gersch and Kitagawa (1988))。平滑化の程度を決めるパラメーター、つまり hyper-parameter の選択には AIC を使い、自動的にその値を決定する

解法のアゴリズム

モデル : $y(n) = w(n)t(n)f(\theta(n)) + x(n) \cdots$ モデル M

: $y(n) = w(n) + t(n) + f(\theta(n)) + x(n) \cdots$ モデル L

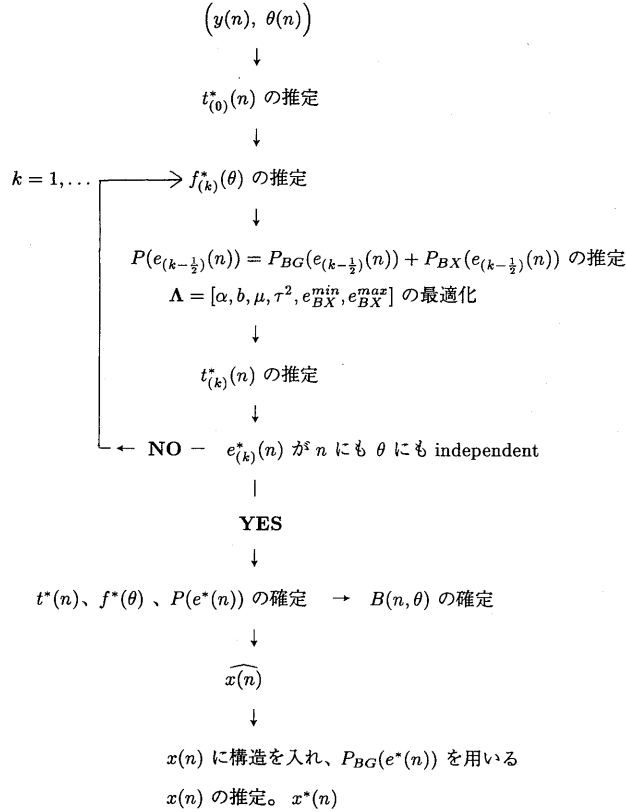


図 1. 解法のアゴリズム.

(Gersch and Kitagawa (1988)).

5.2 ステップ $k(i)$: $f(\theta)$ の推定値の更新

ステップ $k-1$ 回目の $t(n)$ の推定値, $t_{k-1}^*(n)$ を用いて以下のような時系列を構成する.

$$(5.1M) \quad u(n) = \log y(n) - \log t_{k-1}^*(n)$$

$$(5.1L) \quad u(n) = y(n) - t_{k-1}^*(n)$$

いま, $\theta(n)$ が θ_i の範囲にある $u(n)$ をとりだし, 小さい順に並べ換え, それを u_i^j ($j=1, \dots, M_i$) で表そう. M_i はこの位相区間内の $u(n)$ の数.

$u(n)$ は, モデル M の場合

$$(5.2M) \quad \begin{aligned} u(n) &= \log y(n) - \log t_{k-1}^*(n) \\ &\simeq \log y(n) - \log t(n) = e(n) + \log f(\theta = \theta(n)) \end{aligned}$$

モデル L の場合は

$$(5.2L) \quad \begin{aligned} u(n) &= y(n) - t_{(k-1)}^*(n) \\ &\simeq y(n) - t(n) = e(n) + f(\theta = \theta(n)) \end{aligned}$$

である。 u_i^j の $f(\theta = \theta(n))$ はすべて f_i となるから、 M_i 個の u_i^j のうち、異常値のように振舞う大きい K_i 個の u_i^j は BX に属する可能性が高いであろう。異常値の個数 K_i の推定には、Kitagawa (1981) 及び Kitagawa and Akaike (1982) の方法を基本的には採用した (Higuchi et al. (1992)). K_i の値が定まると、小さい $M_i - K_i$ 個の u_i^j の平均が、モデル M の場合は $\log f_i$ の粗い推定値 $\log \hat{f}_i$ を、モデル L の場合は f_i の粗い推定値 \hat{f}_i を与える。

得られた $\log \hat{f}_i$ (モデル L の場合は \hat{f}_i) に periodic smoothness constraint を加え、平滑化を施すことにより $\log f_i^*$ (モデル L の場合は \hat{f}_i^*) を得る。システム及び観測モデルとも線形ガウスモデルを採用する。この平滑化の際にも、平滑化の程度を決めるパラメーターは、AIC を使い自動的に決める。この $\log f_i^*$ から構成される、 $f(\theta)$ の第 k 回目の推定値を $f_{(k)}^*(\theta)$ で記す。

5.3 ステップ k (ii): $t(n)$ の推定値の更新

$t_{(k-1)}^*(n)$ と $f_{(k)}^*(\theta)$ を用いた $e(n)$ の推定値は、次式で与えられる。

$$(5.3M) \quad e_{(k-0.5)}^*(n) = \log y(n) - \log t_{(k-1)}^*(n) - \log f_{(k)}^*(\theta = \theta(n))$$

$$(5.3L) \quad e_{(k-0.5)}^*(n) = y(n) - t_{(k-1)}^*(n) - f_{(k)}^*(\theta = \theta(n))$$

$e_{(k-0.5)}^*(n)$ に長期的変動成分が残っている可能性がある場合は、 $e_{(k-0.5)}^*(n)$ の平滑化を行い、得られたトレンドを $\log t_{(k-1)}^*(n)$ (モデル L の場合は、 $t_{(k-1)}^*(n)$) に加える手続きを行う。

平滑化法には、5.1, 5.2 節同様、ベイズモデルを採用する。ただし、トレンドの周りの分布、つまり観測ノイズとしては、非ガウス分布を想定しているため、非ガウスベイズモデルによる平滑化法を用いる (Kitagawa (1987)). となりあうトレンド成分間のずれに対応するシステムノイズの分布としては、 $e_{(k-0.5)}^*(n)$ のトレンドに突然の jump など予想されない事を踏まえて、ガウスだけを考える。ベイズモデルによる平滑化の枠組みでは、システムノイズや観測ノイズの最適な分布型は AIC 最小化で同時に定められる。ただ、最適化の簡便化もあり、後述の方法で予め観測ノイズの分布を推測し、次に、求めた観測ノイズの分布を固定したままで、システムノイズの分散を変化させて、AIC を最小にする最適なシステムノイズモデルを捜す手続きをとる。

得られた、 $e_{(k-0.5)}^*(n)$ のトレンドを、 $\tilde{e}_{(k-0.5)}^*(n)$ で表す。 $t_{(k-1)}^*(n)$ の補正は、

$$(5.4M) \quad \log t_{(k)}^*(n) = \log t_{(k-1)}^*(n) + \tilde{e}_{(k-0.5)}^*(n)$$

$$(5.4L) \quad t_{(k)}^*(n) = t_{(k-1)}^*(n) + \tilde{e}_{(k-0.5)}^*(n)$$

で行う。

5.4 ステップ k (iii): 反復計算の継続、あるいは中止の判断

ステップ k の $e(n)$ の推定値は、 $t_{(k)}^*(n)$ と $f_{(k)}^*(\theta)$ を用いた

$$(5.5M) \quad e_{(k)}^*(n) = \log y(n) - \log t_{(k)}^*(n) - \log f_{(k)}^*(\theta = \theta(n))$$

$$(5.5L) \quad e_{(k)}^*(n) = y(n) - t_{(k)}^*(n) - f_{(k)}^*(\theta = \theta(n))$$

で定義される。この $e_{(k)}^*(n)$ にトレンドが無いと判断したなら、現在の推定値、 $t_{(k)}^*(n)$ と $f_{(k)}^*(\theta)$ を、 $t(n)$ と $f(\theta)$ に対する最終的な推定値、 $t^*(n)$ と $f^*(\theta)$ とする。同様に $e_{(k)}^*(n)$ を $e(n)$ の最終的な推定値 $e^*(n)$ とする。実際には、 $k=1$ で充分である。さもなければ、 $k := k+1$ として

ステップ $k(i)$ へ戻る.

6. $e(n)$ の密度関数の推定

本章では, ステップ $k(ii)$ での非ガウス平滑化法の適用に於て用いた観測ノイズ分布に相当する $e(n)$ の, その密度関数の推定法について示す.

6.1 $e(n)$ の密度関数

$e(n)$ の密度関数を, BG および BX のそれぞれの $e(n)$ の密度関数である, $P_{BG}(e(n))$ と $P_{BX}(e(n))$ の混合分布

$$(6.1) \quad P(e(n)) = \frac{N_{BG}}{N} P_{BG}(e(n)) + \frac{N_{BX}}{N} P_{BX}(e(n))$$

で表す. P_{BG} が $w(n)$ の分布に一致することは明らかである. N はデータ数. N_{BG}, N_{BX} は, BG 及び BX に属するデータ数. 以後簡単のために, $N_{BX}/N = \alpha$ とおく. N_{BG}/N は, $1 - \alpha$ とかける. α は, 一つのデータセット中の真の信号の受信率を意味する.

本研究において, $P_{BG}(e(n))$ としては, 対称単峰連続パラメトリック分布族である Pearson family

$$(6.2) \quad P(v | b, \mu, \tau^2) = \frac{\tau^{2b-1} \Gamma(b)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(b - \frac{1}{2}\right)} \frac{1}{\left(\tau^2 + (v - \mu)^2\right)^b} \quad \left(\text{for } \frac{1}{2} < b \leq +\infty\right)$$

を考える. この分布族は, ガウス分布 ($b = +\infty$) から, 裾のきわめて重いコーシー分布 ($b = 1$) までを含む表現力の高い分布族である.

4.1 節で述べたように, $x(n)$ の分布には正值をとるという事以外は特徴がないので, $P(x(n))$ の分布として, 一様分布 $U(e_x^{\min}, e_x^{\max})$ を考える. よって $P_{BX}(e(n))$ は, $P(\cdot)$ と $U(\cdot)$ の convolution で定義されるべきだが, 本研究の場合, $e_x^{\max} - e_x^{\min}$ が, τ に比較してかなり大きいので, ほとんど $U(\cdot)$ と同じになる. 従って, $P_{BX}(e(n))$ として一様分布 $U(e_{BX}^{\min}, e_{BX}^{\max})$ を考える.

モデル M の場合もモデル L の場合同様, $P(e(n))$ として (6.1) を考え, $P_{BG}(e(n))$ と $P_{BX}(e(n))$ もモデル L と同じモデルを考える. 4.2 節で $y(n)/t(n)f(\theta = \theta(n))$ を $e(n)$ として採用しなかった理由は, 本研究で取り扱った実際のデータの $\log w(n)$ の分布がほぼ対称であり, (6.2) で考える対称単峰分布族でうまく近似できる事による.

6.2 分布の推定

分布を記述するパラメーター Λ の最適値は, 以下に示す対数尤度を最大化して決定する.

$$(6.3) \quad l(\Lambda) = \sum_{n=1}^N \log P(e_{k-0.5}^*(n) | \Lambda)$$

$P(e(n))$ を記述するノイズモデルのパラメーター数は $\Lambda = [a, b, \mu, \tau^2, e_{BG}^{\min}, e_{BG}^{\max}]$ の 6 個であるが, 本研究では Λ の最適化を効率よく行うため, いくつかのパラメーターの値を固定している. まず, e_{BG}^{\max} を, 実際に出現した $e_{k-0.5}^*(n)$ の最大値, $e_{BG}^{\max} = e_{\max} = \max.\{e_{k-0.5}^*(n)\}$ で予め設定した.

つぎに, e_{BG}^{\min} について考察する. 実際の観測ノイズのレンジは, 多くの観測データに見受けられるように, ある限られた範囲内に収まっている事が多い. 本研究で取り扱うデータの $x(n)$ は, 比較的大きい正值をとるので, $e_{k-0.5}^*(n)$ の最小値, $e_{\min} = \min.\{e_{k-0.5}^*(n)\}$ を与えるデータ

は、 BG に属すると考えても良いであろう。その値を BG に属する $e(n)$ のとる下限と見積り、ノイズモデルの 0 対称性から、 $|e_{\min}|$ は、そのほはおおまかな上限となっているであろう。そこで、 $|e_{\min}|$ よりも大きい $e_{(k-0.5)}^*(n)$ の最小値をもって、 e_{BX}^* の推定値 \hat{e}_{BX}^* としてもよいし、簡単のために $\hat{e}_{BX}^* = |e_{\min}|$ でもよい。さらに、 $e(n)$ はトレンドからの観測値のずれであるから、 $\mu = 0$ の拘束を加えるべきである。

以上により本研究では、 Λ の 6 個のパラメーターを最適化する代わりに、 $\lambda = [\alpha, b, \tau^2]$ の 3 個のパラメーターについて (6.3) を最大化する。 λ の効率よい最適化のために、ステップ $k(i)$ (5.2 節) で用いた異常値の個数を求める方法できわめて粗い α の推定を行い、その値の周りでグリッド探索を行う。 b に関しては単純グリッド探索を、 τ^2 に関しては準ニュートン法を使う。 τ^2 の初期推定値には、ガウス分布 (つまり $b = +\infty$) の場合、異常値の個数を見積る方法の副産物として得られる正常値の分布の分散を、またガウス分布以外の時は、その正常値分布の下側 4 分位数の 2 乗 (コーシー分布の時は τ^2 に一致する) を使った。

7. $x(n)$ の推定

本章では、本来の目的である $x(n)$ の推定方法について示す。

7.1 データの判別法

まず $e^*(n)$ を用いて 6.2 節の手続きを行い、 $P(e(n))$ を表現するパラメーター λ の最終的な推定値、 λ^* を求める。この後 BX である可能性は、 $e(n)$ の関数

$$(7.1) \quad r(e(n)) = \frac{\alpha \cdot P_{BX}(e(n))}{(1-\alpha) \cdot P_{BG}(e(n))}$$

で評価される。 $P_{BG}(e(n))$ は対称単峰なので、 $e(n)$ が大きくなるにしたがい、 $P_{BG}(e(n))$ の値は単調に減少する。 $r(e(n)) = 1$ となる $e(n)$ が存在すれば、その値より大きい $e(n)$ の範囲では、 $\alpha \cdot P_{BX}(e(n)) \geq (1-\alpha) \cdot P_{BG}(e(n))$ となる。

本研究では、 $r(e(n))$ が適当な値 r_{SIN} より大きくなるような $e^*(n)$ をとるデータを BX 、それより小さい値をとるような $e^*(n)$ をもつデータを BG に分類する。本研究では、 $r_{SIN} = 1$ に設定した。

7.2 cleaned data の定義

最終的な目標である $x(n)$ の推定の前に、系統的ノイズである $t(n)$ と $f(\theta)$ の影響のみを除いた量 $\hat{x}(n)$ を、

$$(7.2M) \quad \hat{x}(n) = \begin{cases} y(n) - t^*(n) f^*(\theta = \theta(n)) & e^*(n) \in BX \\ 0 & e^*(n) \in BG \end{cases}$$

$$(7.2L) \quad \hat{x}(n) = \begin{cases} y(n) - t^*(n) - f^*(\theta = \theta(n)) & e^*(n) \in BX \\ 0 & e^*(n) \in BG \end{cases}$$

によって定義する。この $\hat{x}(n)$ は、明らかに $w(n)$ の影響を含んでいる。開発した人工衛星データ処理プログラム (Higuchi et al. (1992)) の中で、モデル M の場合特別に、観測されたデータ $y(n)$ から $f(\theta)$ の影響を除去し補正したデータ、 $t^*(n) + \hat{x}(n)$ 、を *cleaned data* として定義している。

7.3 $x^*(n)$ の定義

本研究の枠組みでは、 $x(n)$ に何も構造を想定していない。 $w(n)$ の影響までを除去した $x(n)$ の推定は、 $x(n)$ に構造を入れなければ不可能である。そこでいま、簡単な構造である smoothness constraint を $x(n)$ に課して議論を進める。具体的には、 $x(n)$ の 1 回差分の 0 からのずれが、適当な分布形（もちろん非ガウスも含める）から発生したシステムノイズで定められているといったシステムモデルを考える (Kitagawa (1987))。モデル L の場合は $P_{bc}(e(n)) = P(w(n))$ であるから、 BX に判別された $\hat{x}(n)$ に対して、7.1 節で定まった $P_{bc}(e(n))$ の観測ノイズをもつベイズモデルで平滑化を行えばよい。システムノイズ分布の選択には、AIC を用いる。 $\hat{x}(n)$ を平滑化して得られる値でもって、 $x(n)$ の推定値 $x^*(n)$ を定める。

モデル M の場合、システム方程式に関してはモデル L と同じでよいが、観測方程式は以下のように非線形である。

$$(7.3) \quad y(n) = x(n) + \exp(e(n))t^*(n)f^*(\theta = \theta(n))$$

$t^*(n)$ と $f^*(\theta = \theta(n))$ は、この段階では既に与えられていることに注意する。 $e(n)$ の分布は、モデル L の時と同じように $P_{bc}(e(n))$ で与えられる。最適なシステムノイズの選択にはやはり AIC を用いる。このベイズモデルによって得られた $x(n)$ の解を、モデル M の場合の $x^*(n)$ とする。

8. 結 果

8.1 人工衛星データへの応用

本研究で取り扱うデータは、米国の人工衛星 Pioneer Venus orbiter によって観測された電場の強さである (Higuchi et al. (1992))。単位は、 $(V/m)^2/Hz$ 。30 [kHz], 5.4 [kHz], 730 [Hz], 100 [Hz] の周波数帯における電場の強さが、4 チャンネルで同時に計測され、よって 2 章の M は 4 である。 $B_m(n)$ の様相が m ごとに著しく違い、チャンネル間の $B_m(n)$ の関係を表現できる関数 $F(m)$ も存在しないので、各チャンネル毎に $y_m(n)$ を使って background noise $B_m(n)$ を推定する。

4 チャンネルの中で、30 [kHz] の background noise の出現形態は非常に規則的で、また signal to noise ratio (S/N 比) も十分に大きい。従って、本研究で提案した方法を、5.4 [kHz], 730 [Hz], 100 [Hz] の 3 チャンネルに適応した。周波数が下がるにつれて S/N 比が下がり、100 [Hz] のチャンネルの background noise の除去が一番難しい。Higuchi et al. (1992) にくつつかの結果が示してあるので、ここでは、ごく簡単に結果の報告を行う。図 2(a) は、ある日の 5.4 [kHz] のデータセットのごく一部分である。一つのデータセットのデータ数は、20,394 個。図中に見える周期的な変動がスピンノイズに相当し、burst 的に観測されているものが real signal である。この時間内の background noise の強さ ($t(n)$ に対応する量) は、ほぼ一定であるが、データセットを通してみると変動している。得られた *cleaned data* と $f^*(\theta)$ を Panel (b), (c) に示す。Panel (a) に見られた明らかにスピンノイズに関連する系統的ノイズが、Panel (b) では除去されているのが解る。

図 3 に 5.4 [kHz] の $e^*(n)$ のヒストグラムと、その推定された密度関数、 $P(e(n))$ を示す。滑らかな実線が、 $P(e(n))$ に相当する。このデータセットの $P_{bc}(e(n))$ には、コーシー分布が最適であった。3 つのデータセットについて得られた最適な $P(e(n))$ のパラメーターを、表 1 に示す。この表からも、データセット毎に λ^* が変化するのがよく解る。

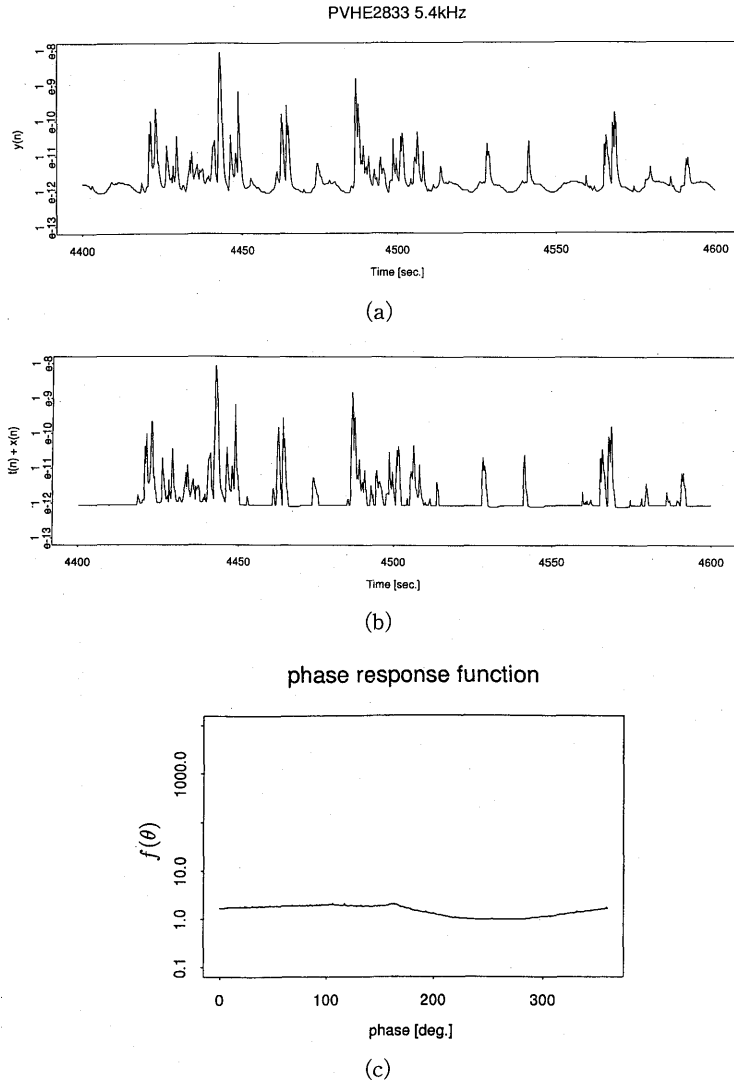


図2. (a) 5.4 [kHz] の $y(n)$. (b) 推定された cleaned data $t^*(n) + \hat{x}(n)$. (c) 推定された $f^*(\theta)$.

表1. 推定した λ^* .

Channel	推定した $P_{BG}(e(n))$		推定した $P_{BX}(e(n))$		
	b	τ^2	α	e_{BX}^{\min}	e_{BX}^{\max}
100 Hz	$+\infty$ (Gauss)	8.787×10^{-2}	0.06	1.358	3.891
730 Hz	$+\infty$ (Gauss)	1.972×10^{-3}	0.02	0.2131	1.861
5.4 kHz	1 (Cauchy)	2.122×10^{-4}	0.14	0.1795	3.871

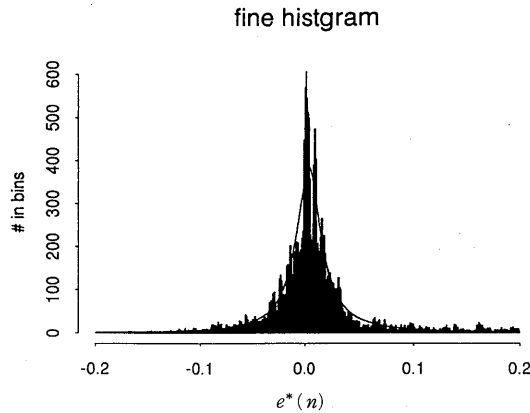


図3. 5.4 [kHz] の $e^*(n)$ のヒストグラムと、推定された密度関数, $P(e(n))$.

8.2 画像データへの応用

8.2.1 補正画像

1章に述べたように、画像データもスピンノイズの影響を受けてイメージが歪んでしまう。いま、人工衛星のスピンを利用して走査している画像の1方向を θ 、それと直交する方向を ϕ と表記しよう。(4.1) でモデル化されるようなデータが画像を構成すると、 $f(\theta)$ で表現されるような θ 方向の歪に加えて、走査中の観測時間に依存するレベルの変化 $t(n)$ が現れる。逆に、このような問題が生ずるような画像データには、本研究の方法を一部変更するだけで応用できる。いままでは、 $[0^\circ, 360^\circ)$ の範囲をとる $\theta(n)$ を考えてきたが、一方1画像を形成するデータの θ のレンジは限られている。つまり、ある範囲内 $[\theta_{\min}, \theta_{\max})$ を走査することで1画像は得られる。よって、 $f(\theta)$ の定義域は $[\theta_{\min}, \theta_{\max})$ であるから、5.2節で用いた periodic smoothness constraint は、単純な smoothness constraint に変える必要がある。図4(a) に、2次元データの構成法を模式的に示す。走査のパターンは図に示した限りではないが、 θ 方向の走査にスピンを利

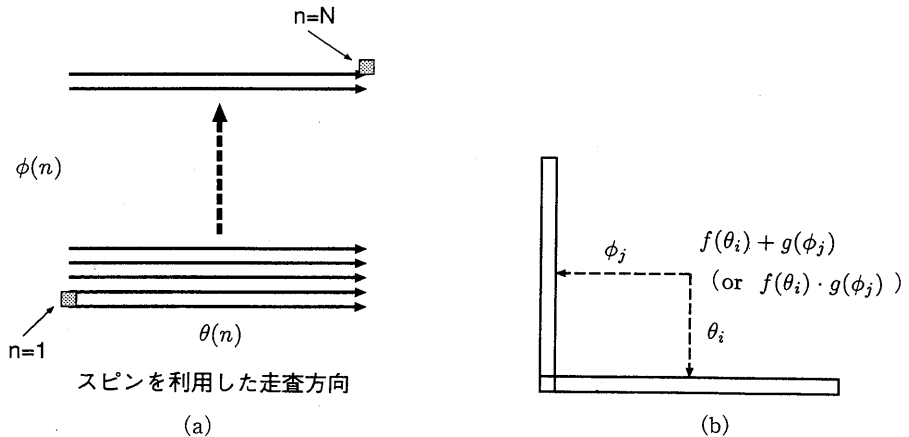


図4. (a) 人工衛星の画像が1次元情報から構成されていることを、模式的に表した図。走査のパターンは、図に示したのものには限らない。(b) 特殊な性質を持つ画像データの構成法。

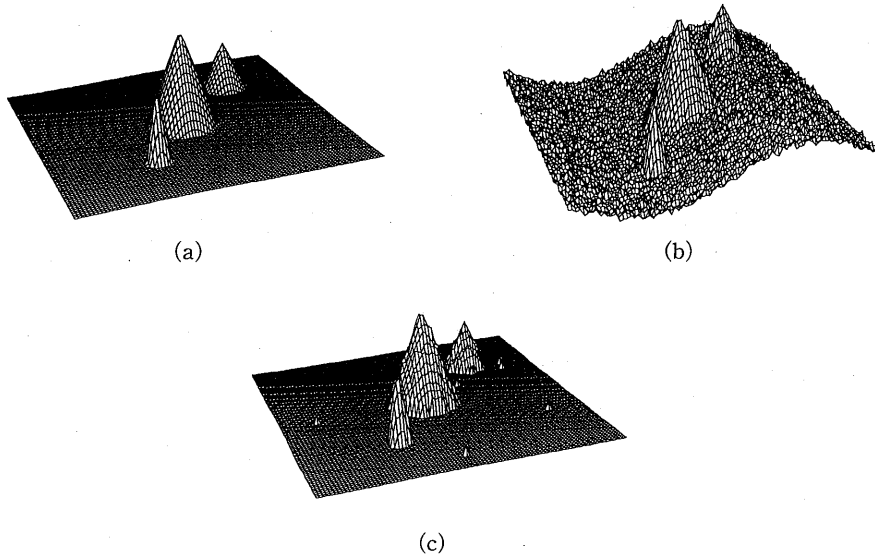


図5. (a) real signal $x(n)$. (b) ノイズを加えたデータ $y(n)$. (c) 系統的な background noise を除去したデータ $\hat{x}(n)$. $w(n)$ の影響は除去されていないので、山の表面は滑らかではない。

用している限り、普通である。

実際の画像データが手元にないので、模擬データで本研究の方法の有効性を確かめてみた。8.1節でモデル M を取り扱ったので、ここではモデル L で表現できるデータを模擬的に作製し、開発した方法を適用してみた。図5(a), (b) に、 $x(n)$, $y(n)$ をそれぞれ示す。時系列から画像を構成する方法は、図4(a) に示した走査パターンによるものとする。Panel(a) に見られる3つの山の高さは、2.0, 3.0, 4.0である。簡単のため、 $f(\theta)$ には $-\sin(\theta)$ を、 $t(n)$ には1次関数を採用した。ただし、本研究で開発した方法は、局所的に滑らかであればどのような $f(\theta)$, $t(n)$ でもよいことを強調しておく。加えたノイズの分布 $P(w(n))$ は、分散 $\tau^2=0.01$ のガウス分布である。background noise により画像は大きく歪み、Panel(b) からは、もともとの山の高さは単純には推察できない。Panel (c) に、系統的ノイズである $t(n)$ と $f(\theta)$ を除去したデータ、 $\hat{x}(n)$ を示す。前述したように $\hat{x}(n)$ には $w(n)$ の影響が含まれているので、画面の表面は滑らかではない。重要な点は、系統的ノイズ項の削除により、山の高さが補正されている点である。山の裾がどこまで低く復元されるかは、BG と BX を判別する数居値 $r_{s/N}$ をどの程度小さく設

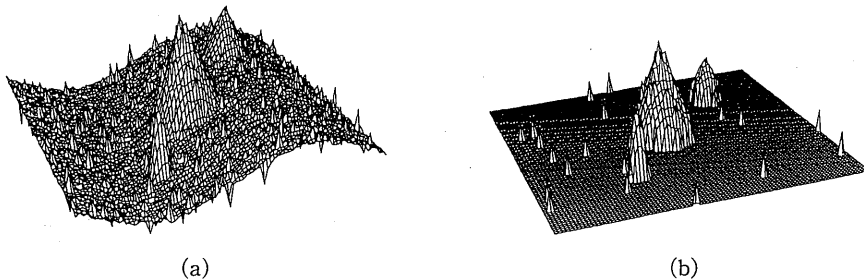


図6. (a) $y(n)$. (b) 系統的な background noise を除去したデータ $\hat{x}(n)$. $w(n)$ の影響は除去されていないので、山の表面は滑らかではない。

表 2. 推定した λ^* .

図番号	与えた $P(w(n))$		推定した λ^*		
	b	τ^2	a	b	τ^2
5	$+\infty$ (Gauss)	1.0×10^{-2}	0.06	$+\infty$ (Gauss)	1.058×10^{-2}
6	1 (Cauchy)	1.0×10^{-3}	0.04	1 (Cauchy)	1.489×10^{-3}
7	$+\infty$ (Gauss)	1.0×10^{-2}	0.08	$+\infty$ (Gauss)	1.172×10^{-2}

定するか完全に依存している。

図 6(a), (b) に, $w(n)$ の分布に $\tau^2=0.001$ のコーシー分布を使った模擬データ $y(n)$ と, 推定された $\hat{x}(n)$ を示す. 表 2 に, 図 5, 6 の各々の場合について, 与えた $P(w(n))$ と, 推定した $P(e(n))$ のパラメーター λ^* をまとめた.

8.2.2 平滑化された画像

7.3 節に示した, $w(n)$ の影響までを取り除いた量 $x^*(n)$ を推定する方法を, 画像データに適用してみる. 模擬データ $y(n)$ は, 図 7(a) に示す $x(n)$ に, 図 5, 6 に用いた $t(n)$ と $f(\theta)$ 及び, 分散 0.01 のガウスノイズ $w(n)$ を加えて作製する. Panel(b) に $y(n)$ を, Panel(c) に系統的ノイズ除去量 $\hat{x}(n)$ を示す. このデータにはモデル L をあてはめているので, BX の $\hat{x}(n)$ の平滑化によって得られるトレンドが, $x^*(n)$ になる. Panel(d) に $x^*(n)$ を示す. このデータに最適なシステムノイズ型は, AIC より $\tau^2=3.815 \times 10^{-6}$ のコーシー分布である. この実験の場合の $P(w(n))$ と, 推定した $P(e(n))$ のパラメーター λ^* の値も表 2 に与えた. $x^*(n)$ が構成する画像の θ 方向の滑らかさは, 時間軸に沿って smoothness constraint を加えている事で達成されている. 一方, モデルに明示的に ϕ 方向の構造は入っていないため, 画像に ϕ 方向の滑らかさは得られないことを注意しておく. よって, 採用した単純な smoothness constraint は, Panel(a)

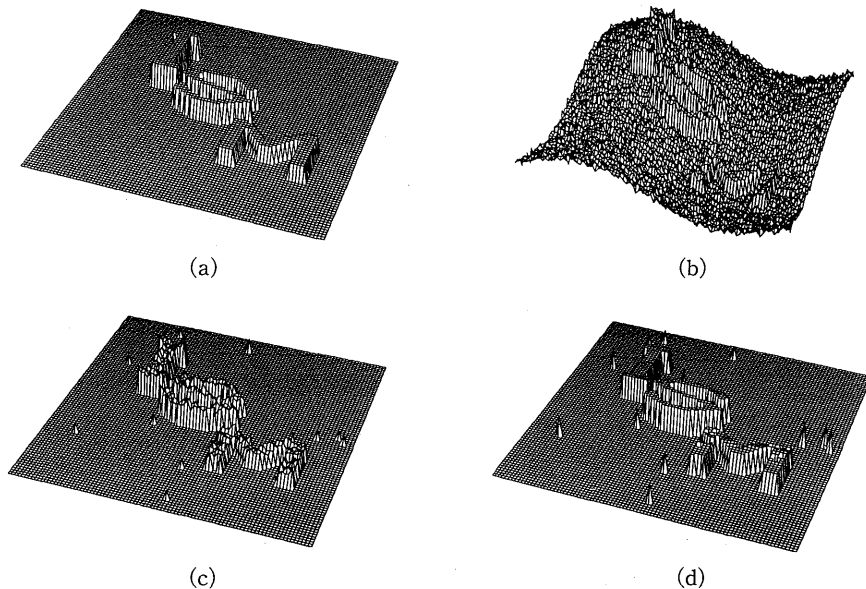


図 7. (a) real signal $x(n)$. (b) ノイズを加えたデータ $y(n)$. (c) 系統的な background noise を除去したデータ $\hat{x}(n)$. (d) $x(n)$ に滑らかさを仮定し, $w(n)$ の項も除去したデータ, $x^*(n)$.

の ISM の画像の事前情報としては不適當である。その単純な解決策には、1次元の平滑化を組合せ疑似的に2次元の平滑化を行うといったもの (Tiyip 他 (1992)) がすぐ考えられるが、便宜的であるため本稿ではあえて行わなかった。系統的な非ガウスノイズを含む画像回復の本質的な解法は、本稿で述べたようなアプローチとは全く違うモデリングが必要であることのみを付記するにとどめる。

8.2.3 直和・直積モデル

(3.2) の background noise モデル中の $t(n)$ のかわりに、 ϕ 方向の画像の歪みの関数 $g(\phi)$ を使って、画像の系統的歪みを表す background noise が、

$$(8.1M) \quad B(n, \theta, \phi) = w(n)g(\phi)f(\theta)$$

$$(8.1L) \quad B(n, \theta, \phi) = w(n) + g(\phi) + f(\theta)$$

で表現されるものを考える。 $w(\cdot)$ は、 θ にも ϕ にも依存しない、相互相関の無いノイズである。この時、画像データは

$$(8.2M) \quad y(n, \theta(n), \phi(n)) = w(n)g(\phi(n))f(\theta(n)) + x(n)$$

$$(8.2L) \quad y(n, \theta(n), \phi(n)) = w(n) + g(\phi(n)) + f(\theta(n)) + x(n)$$

と表現されている。画像データの構成法を図 4(b) に模式的に示す。 $g(\phi)$ としては、local に滑らかなものであることを想定する。

この特殊な画像モデルに於て $x(n)$ を推定する問題は、本研究で取り扱ったモデルの極めて簡単な例になるのは明らかである。 $t(n)$ の推定問題が $g(\phi)$ の推定問題に変わったことにより、 $g(\phi)$ の初期推定の際、同じ ϕ をもつ $[\theta_{\min}, \theta_{\max}]$ のデータからの情報を有益に使える。従って、その初期推定値も安定した良いものが得られ、 $x(n)$ を推定する最終的な目標達成も非常に楽になる。問題が簡単なので、具体的な画像データへの応用結果は割愛する。

9. ま と め

スピンノイズの様相は時間及び衛星の位置とともに大きく変動するが、提案されたモデルは、柔軟に状況に即応し、ノイズを自動的に取り除く事を可能にしている。*ad-hoc* なパラメータは、 r_{SIN} と反復計算の回数 k のみである事を特に強調しておく。この方法により、既存の方法で得られる情報の限界を越え、観測データに基づく新しい知見を得ることが可能になる。

謝 辞

本研究の一部はカリフォルニア大学ロサンゼルス校の C.T. Russell 教授、R.J. Strangeway 博士、大学院生 G.K. Crawford らとの共同で行われた。非ガウス平滑化 subroutine は、統計数理研究所・北川教授の開発されたプログラムのソースコードを基にしたことをここに記すと共に、快くソースプログラムを提供して下さった北川教授に感謝します。

この研究で使用した電場データは、人工衛星 Pioneer Venus orbiter によって得られ、アメリカ航空宇宙局の援助 (研究助成番号 NAG2-501 及び NAG2-485) のもとに処理されたものである。

また、原稿を精読し多くの貴重なコメントをいただいた査読者の方々に深く謝意を表したい。

参 考 文 献

- Akaike, H. (1980). Likelihood and the Bayes procedure (with discussion), *Bayesian Statistics*, 143-165, University Press, Valencia, Spain.
- Cracknell, A.P. (1981). Background: the physical basis of remote sensing, *Remote Sensing in Meteorology, Oceanography and Hydrology* (ed. A.P. Cracknell), Ellis Horwood, England.
- Gersch, W. and Kitagawa, G. (1988). Smoothness priors in time series, *Bayesian Analysis of Time Series and Dynamic Models* (ed. J.C. Spall), 431-476, Marcel Dekker, New York.
- Higuchi, T. (1991). Method to subtract an effect of the geocorona EUV radiation from the Low Energy Particle (LEP) data by the Akebono (EXOS-D) satellite, *Journal of Geomagnetism and Geoelectricity*, **43**, 957-978.
- Higuchi, T. (1992). Reply to comment by T. Mukai, *Journal of Geomagnetism and Geoelectricity*, **44**, 569-571.
- Higuchi, T. (1994). Separation of spin synchronized signals using a Bayesian approach, *Proceedings of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach* (ed. H. Bozdogan), 193-215, Kluwer, Netherlands.
- Higuchi, T., Kita, K. and Ogawa, T. (1988). Bayesian statistical inference to remove periodic noises in the optical observation aboard a spacecraft, *Applied Optics*, **27**, 4514-4519.
- Higuchi, T., Crawford, G.K., Strangeway, R.J. and Russell, C.T. (1992). Separation of spin synchronized signals, Research Memo., No. 430, The Institute of Statistical Mathematics, Tokyo.
- Kitagawa, G. (1981). 異常値解析ベイズモデル, *数理科学*, **213**, 62-66.
- Kitagawa, G. (1987). Non-Gaussian state space modeling of nonstationary time series (with discussion), *J. Amer. Statist. Assoc.*, **79**, 1032-1063.
- Kitagawa, G. and Akaike, H. (1982). A quasi Bayesian approach to outlier detection, *Ann. Inst. Statist. Math.*, **34**, 389-398.
- Kosiishi, H. (1991). 光学センサーの最新トレンド, *日本リモートセンシング学会誌*, **11**, 104-108.
- Mukai, T., Kaya, N., Sagawa, E., Hirahara, M., Miyake, W., Obara, T., Miyaoka, H., Machida, S., Yamagishi, H., Ejiri, M., Matsumoto, H. and Itoh, T. (1990). Low energy charged particle observations in the "auroral" magnetosphere first results from the Akebono (EXOS-D) satellite, *Journal of Geomagnetism and Geoelectricity*, **42**, 479-496.
- Tanabe, K. and Tanaka, T. (1983). ベイズモデルによる曲線・曲面のあてはめ, *月刊地球*, **5**, 179-186.
- Tiyip, T., 小島尚人, 大林成行 (1992). 衛星マルチスペクトル画像の画質改善手法の提案, *日本リモートセンシング学会誌*, **12**, 141-156.

A Method to Separate the Spin Synchronized Signals
Using a Bayesian Approach*

Tomoyuki Higuchi

(The Institute of Statistical Mathematics)

Data taken aboard a spinning spacecraft frequently suffers from an unexpected modulation synchronized with the rotation of the spacecraft. Higuchi et al. (1992) has proposed the method to separate natural emissions from the several possible sources of noise along the line of a Bayesian approach which has been recently applied to various inversion problems such as nonstationary time series modeling and image reconstruction. In the proposed approach, the observed data is modeled as $y(n) = w(n)t(n)f(\theta(n)) + x(n)$, where $\theta(n)$ is an angle (phase) between the sensor and the some view direction and $f(\theta)$ represents the phase response of the background noise. $t(n)$ and $w(n)$ are the long-term trend component and time- and phase-independent component of the intensity of the background noise, respectively. $x(n)$ is the estimated natural emissions. In this paper, we extend this model to a general case even where $e(n) = \log w(n)$ obeys a non-Gaussian distribution function. In addition, we apply it to an image restoration problem in which a given image is built up by a scanning like a TV picture.

Key words : Time series, Bayesian approach, outlier detection, smoothing, nonlinear modeling, spatial pattern analysis, image processing.

* The electric field data examined in this report were obtained by the Pioneer Venus orbiter and were processed with support furnished by the National Aeronautics and Space Administration (NASA) under research grants NAG2-501 and NAG2-485.