

## A I C 法による不適切問題の解法と信号復元問題への応用について

戸水 誠、小川 重義 (金沢大学自然科学研究科)

### 1 はじめに

信号は伝送路を通る段階でブレを生じ、我々が観測するときには、さらに雑音が付加した状態で観測される。

仮に、雑音を無視すれば、逆処理によって復元は可能に思えるが、後に示すように、これは不適切逆問題となり簡単には解決できない。

今回、こうした関数への A I C 法の応用について考える。既にこうした研究は多くの著者によってなされている (cf 田辺氏 [2])。本稿では、第一に A I C 法によるブレ信号の復元と、第二に新たな展開として、ブレ幅そのものの推定 (熱方程式で考えれば、信号の発生時期の推定) を A I C 法の応用によって行い、それぞれの実験結果などを報告する。

### 2 背景

ガウス型のブレを受けた信号を、

$$u(x, \sigma^2) \quad : \text{ブレ信号} (\sigma^2 : \text{ブレ幅}) \quad (1)$$

$$u(x, 0) = f(\xi) \quad : \text{元信号} \quad (2)$$

とする。すなわち、ガウス型関数

$$G_t(x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}} \quad (3)$$

に対して、 $u(x, \sigma^2)$  は  $f(x)$  と  $G_{\sigma^2}(x)$  の合成積で表されるものである。

$$u(x, \sigma^2) = (f * G_{\sigma^2})(x) = \frac{1}{\sqrt{4\pi\sigma}} \int_{-\infty}^{\infty} f(\xi) e^{-\frac{(x-\xi)^2}{4\sigma^2}} d\xi \quad (4)$$

別の観点から見れば、この  $u(x, \sigma^2)$  はブレ幅  $\sigma^2$  を拡散時間  $t$  と考えると、次の熱方程式に対する初期値問題

$$\begin{cases} \frac{\partial}{\partial t} u(x, t) = \frac{\partial^2}{\partial x^2} u(x, t) \\ u(x, 0) = f(x) \end{cases} \quad (5)$$

の解として考えることができる。

熱力学第 2 法則からもわかるように、熱方程式を時間逆向きに解くことは不可能であり、従ってブレ信号の復元が一般に直接的には解けないことが容易に予想される。

ところで我々が取り扱う信号はブレ信号にさらに雑音が付加したもので、観測信号  $u(x, \sigma^2)$  は、

$$v(x, \sigma^2) = u(x, \sigma^2) + Z(x) \quad (6)$$

と表される。ここで、 $Z(x)$  は雑音である。

すなわち、我々の問題は  $v(x, \sigma^2)$  の  $n$  個の観測点  $x_i$  におけるデータ

$$\{v(x_i, \sigma^2)\}_{1 \leq i \leq n} \quad (7)$$

を得て、 $f(x)$  を推定することである。

### 3 統計的モデル化

元信号  $f(x)$  は以下のように表現されているものとしよう。

$$f(x) = \sum_{i=1}^N a_i \varphi_i(x) \quad (8)$$

ここに  $\varphi_i(x)$  は  $L^2(0, 1)$  における基底関数系である。実際には、我々は Schauder 基底あるいは三角関数系を採用する。

ところで  $T_t$  を初期値関数 (2) において、初期データ  $f(x)$  を時刻  $t$  における像  $u(x, t)$  と対応させる写像

$$(T_t f)(x) \equiv (f * G_t)(x) \quad (9)$$

とすれば、この時

$$u(x, t) = \sum_{i=1}^N a_i (T_t \varphi_i)(x) = \sum_{i=1}^N a_i \psi_i(x, t) \quad (10)$$

$$\psi_i(x, t) \equiv (T_t \varphi_i)(x), \quad (11)$$

観測信号  $v(x)$  は

$$v(x) = u(x, t) + Z(x) = \sum_{i=1}^N a_i \psi_i(x, t) + Z(x) \quad (12)$$

と表される。

すなわち、観測データ  $\{v(x_i, \sigma^2)\}$  から  $\mathbf{a}$ ,  $N$  を決定すれば、 $f(x)$  を推定できる。このうち係数ベクトル  $\mathbf{a}$  は、観測データ  $\{v(x_i, \sigma^2)\}$  より、最小二乗法で推定する。しかし解像度  $N$  を厳密に推定する方法はなく、従来解析者の主観にまかされてきた。

本稿では 田辺氏 ([1]) の手法に従い、 $N$  を観測データを基に客観的に推定する方法として AIC 法の応用を考える。

## 4 基底関数系

本稿では基底関数系  $\varphi_i(x)$  として、Schauder 基底を使う。Schauder 関数は次のように定義される。

$$S_{2^n, i}(x) \equiv \int_0^x H_{2^n, i}(\tau) d\tau \quad (13)$$

Harr 関数は区間  $[0, 1]$  で以下のように定義される。 ( $n \geq 0$ )

$$H_{0,1}(x) = 1 \quad 0 \leq x \leq 1, \quad (14)$$

$$H_{2^n, 1}(x) = \begin{cases} 2^{n/2} & 0 \leq x < 2^{-(n+1)}, \\ -2^{n/2} & 2^{-(n+1)} \leq x \leq 2^{-n}, \\ 0 & \text{otherwise,} \end{cases} \quad (15)$$

$$H_{2^n, i}(x) = H_{2^n, 1}\left(x - \frac{i-1}{2^n}\right) \quad i = 1, \dots, 2^n. \quad (16)$$

Schauder 関数は一つ一つの区分が例えば三角関数とは異なり、他の区分にまで影響を与えることがないために、三角関数列を用いた直交分解より、収束が速いことが期待できる。しかし実際には基底と復元結果との因果関係は未だ未解決の問題である。

## 5 統計的モデルの決定

基底関数系を Schauder 関数とすれば、モデルは以下のように書ける。

$$f(x) = a_{0,1} S_{0,1}(x) + \sum_{n=0}^N \sum_{i=1}^{2^n} a_{2^n, i} S_{2^n, i}(x) \quad (17)$$

本稿では、係数ベクトル  $\mathbf{a}$  を最小二乗法により決定する。観測信号はブレを生じた状態であるので、最小二乗法を使うためにはモデル  $f(x)$  もブレを生じた状態にしなければならない。

すなわち式 (17) を式 (4) によってブレ信号モデル  $g(x, \sigma^2)$  にする。

$$g(x, \sigma^2) = \frac{1}{\sqrt{4\pi\sigma}} \int_{-\infty}^{\infty} f(\xi) e^{-\frac{(x-\xi)^2}{4\sigma^2}} d\xi \quad (18)$$

$\mathbf{a}$  はこのブレ信号モデル  $g(x, \sigma^2)$  と観測データ  $\{v(x_i, \sigma^2)\}$  から最小二乗法で推定することになる。

モデルの式 (17) は  $\mathbf{a}$  が決まれば、あとは  $N$  を決めることで推定出来る。ここで  $N$  を解像度と呼ぶことにする。

解像度が高ければ、より細かく復元出来ることになるが、十分なデータが揃わなければ余計なところを目立たせることになり、かえって信号本来の特徴をつかみづらくなる。そのため、適切な解像度を求める必要がある。

本稿では解像度  $N$  を AIC 法により決定する。AIC 法はモデルと真の分布との近さを真の分布からの観測データのみから与えるものであり、真の分布は不明でも構わない。

具体的には、各  $N$  ごとに AIC と呼ばれる数値を計算する。そして、AIC が最小となる時の  $N$  を採用することにより、最も真に近いモデルを選択できる。

## 6 AIC 法

文献 [1] を参考に、AIC 法によるブレ信号の復元方法を説明する。

従来、真の分布の密度関数とモデルの密度関数  $g(x)$  との近さを客観的に求めるために Kulback-Leibler 情報量 (以下 K-L 情報量) がある。

$$I(f; g) = \int_{-\infty}^{\infty} \log \left\{ \frac{f(x)}{g(x)} \right\} f(x) dx \quad (19)$$

$$= \int_{-\infty}^{\infty} \{\log f(x)\} f(x) dx - \int_{-\infty}^{\infty} \{\log g(x)\} f(x) dx \quad (20)$$

K-L 情報量は分布が近いほど小さい値となり、全く同じ場合 0 となる数である。よってモデルの優劣をつけるだけなら、(20) の右辺第二項の比較だけで十分である。

すなわち、

$$\int_{-\infty}^{\infty} \{\log g(x)\} f(x) dx \quad (21)$$

が大きい程、そのモデルは良いということになる。これを平均対数尤度と呼ぶ。

ここで注意しなければならないのが、平均対数尤度は真の分布がわかっていないと求めることが出来ない量であるということである。よって、平均対数尤度を真の分布からの観測データのみから推定することが必要となる。そこで以下のように考える。

$$\int_{-\infty}^{\infty} \{\log g(x)\} f(x) dx = E_X [\log g(X)] \quad (22)$$

よって大数の法則から、 $n \rightarrow \infty$  のとき、

$$\frac{1}{n} \sum_{i=1}^n \log g(X_i) \rightarrow E_X [\log g(X)] \quad (23)$$

となるので、

$$l(\theta) = \sum_{i=1}^n \log g(X_i) \quad (24)$$

が平均対数尤度の推定量とり、この値を対数尤度と呼ぶ。ここで  $\theta$  はそのモデルを決定するパラメータである。

以上から観測データにもとづいてK-L情報量の大小を比較するためには、対数尤度を比較すれば良いということになる。よってモデルがいくつかの調整できるパラメータを含む場合は、対数尤度を最大とするようにパラメータの値を選ぶことによって良いモデルが選択できる。

次に、対数尤度を最大にするようにパラメータを定めた（最尤パラメータ）モデルが複数ある場合を考える。これらのモデルは平均対数尤度を用いて比較することになる。しかし最尤パラメータは観測データを元に決めるので、この時の平均対数尤度は観測データに関する確率変量ということになる。よって、平均対数尤度の期待値を使ってモデルの比較を行うのが最も望ましいと考え、この値を期待平均対数尤度と呼ぶ。

$$l_n^*(k) = E_X \left[ l^*(\hat{\theta}) \right] = \int_{-\infty}^{\infty} l^*(\hat{\theta}(x_1, \dots, x_n)) \prod_{i=1}^n g(x_i) dx \quad (25)$$

AICは期待平均対数尤度の普遍推定量であり、ただ一度の観測データのみから求められ、以下のように定義される。

$$\begin{aligned} AIC(K) &\equiv -2 \times (\text{モデルの最大対数尤度}) \\ &\quad + 2 \times (\text{モデルの自由パラメータ数}) \end{aligned}$$

この定義より、AICが小さい程、良いモデルと評価することが出来る。

### (信号復元実験)

まず、ブレ信号を得る。

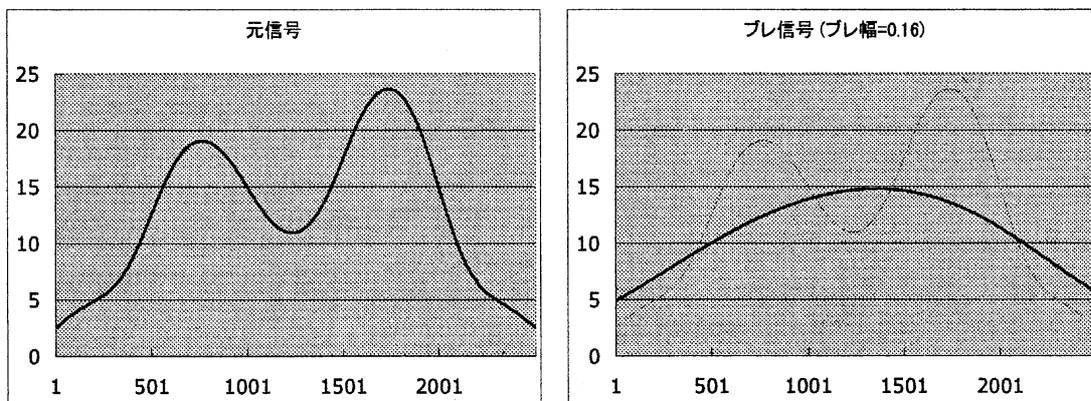


図 1: 実験で用いた元信号とブレ信号

次に既知であるブレ幅(今実験では0.16)を用いて、復元信号の各パラメータを推定する。

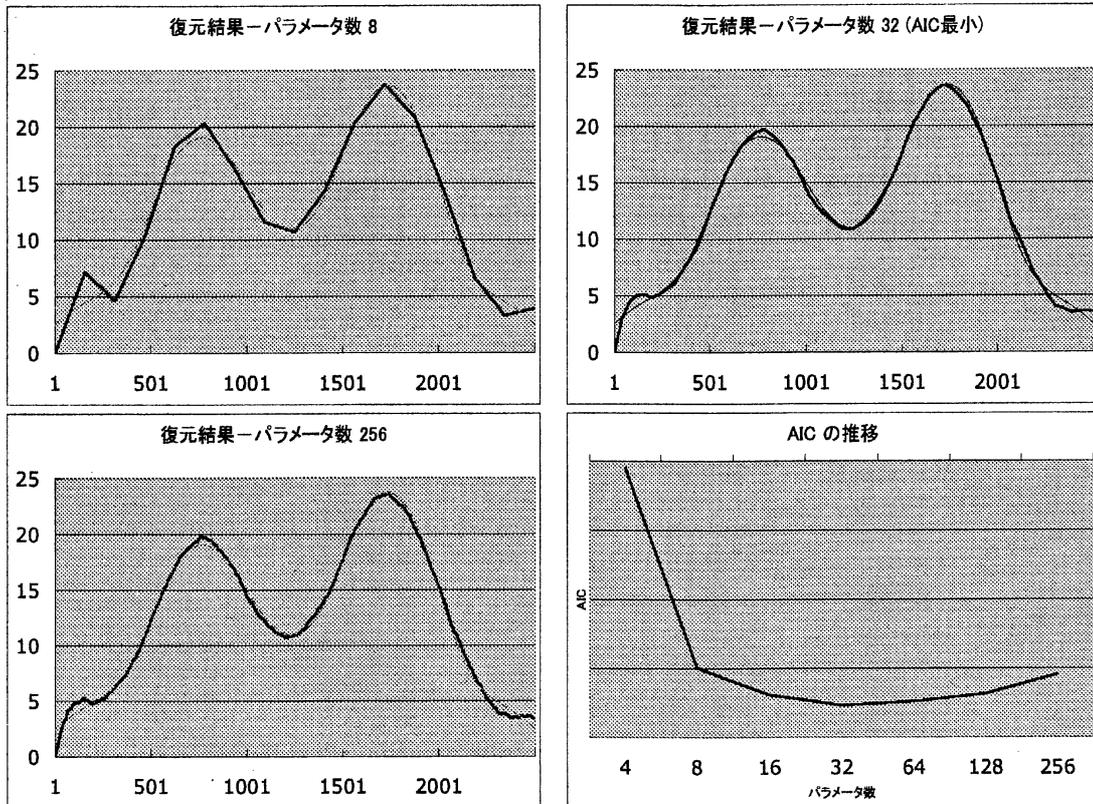


図 2: 各パラメータ数における復元結果と、AIC の推移

AIC の推移から、パラメータ数が 32 の時を復元結果として採用する。

## 7 ブレ幅の推定について

この方法で AIC 法により信号復元をする際に、重要なことはブレ幅が既知でないと復元できないということである。というのも、復元の過程においてモデルをブレさせる必要があり、以下のブレの式にはブレ幅 ( $= \sigma^2$ ) が含まれているからである。

$$u(x, \sigma^2) = \frac{1}{\sqrt{4\pi\sigma}} \int_{-\infty}^{\infty} f(\xi) e^{-\frac{(x-\xi)^2}{4\sigma^2}} d\xi \quad (26)$$

そこで本稿では AIC 法により、信号復元と同時にブレ幅も推定出来ないかを考えた。そのため次の実験を行う。ブレ幅の推定が可能ということは、式 (5) の熱方程式の初期値問題で言えば拡散時間が推定できるということになる。

以下の手順で実験を行う。

1. 用意した元信号にブレや雑音を付加した観測信号から、観測データを得る。以降、ここで用いたブレを真のブレ幅と呼ぶ。

2. 真のブレ幅を想定ブレ幅とした複数のモデルを、各基底関数ごとに作成する。
3. 他のブレ幅（複数）を想定ブレ幅とした複数のモデルを、各基底関数ごとに作成する。
4. 前実験と同様に、各モデルのパラメータを決定し  $AIC$  を計算する。

最後に、 $AIC$  が最小のものを選択するのだが、もし  $AIC$  法によりブレ幅の推定が出来るならば真のブレ幅を用いたモデルが選択されると考えられる。

モデル	ブレ過程	推定量	最終推定量
$g(x)$	→	$A^1$ 真のブレ幅	→ $g^1$
$g(x)$	→	$A^2$ 偽のブレ幅	→ $g^2$
$g(x)$	→	⋮	⋮
$g(x)$	→	$A^l$ 偽のブレ幅	→ $g^l$

$\left. \begin{array}{l} \phantom{g(x)} \\ \phantom{g(x)} \\ \phantom{g(x)} \\ \phantom{g(x)} \end{array} \right\} g^*$

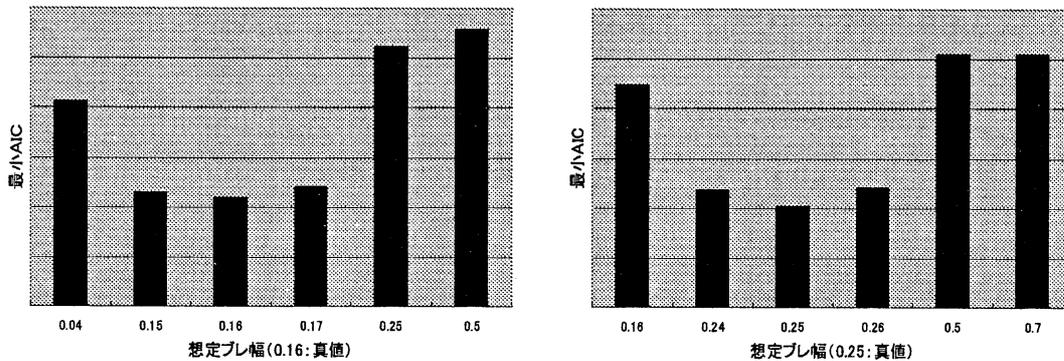


図 3: 想定ブレ幅ごとの最小  $AIC$

そして実験結果は、真のブレ幅が 0.16、0.25 共に、モデルの想定しているブレ幅が真のブレ幅と一致している時、 $AIC$  が最小になった。この結果から、既知でなければならぬブレ幅の推定にも  $AIC$  法の有効性が示され、ブレ信号の完全自動復元が可能になったと言える。

## 参考文献

- [1] 坂元慶行, 石黒真木夫, 北川源四郎: 情報量統計学, 共立出版 (1983).
- [2] 田辺國士: 不適切問題への統計的アプローチ, 数理科学 No.153, pp. 60-64 (1976).
- [3] 河田竜夫: 応用数学概論 II, 岩波全書.
- [4] The Entropy Concept in Probability Theory  
(Uspekhi Matematicheskikh Nauk, vol. 3, 1953, pp. 3-20)
- [5] 樋口浩一: ガウシアンブレ画像の多段階表現による復元について, 金沢大学工学部電気・情報工学科卒業論文 (1998).