

# AICによるウェーブレット基底関数の選択

大正製薬株式会社 松 嶋 優 貴

大阪大学 大学院基礎工学研究科 白 旗 慎 吾

大阪大学 大学院基礎工学研究科 坂 本 亘

**要 旨** ウェーブレットは不規則、不連続な関数(または信号や画像)を効率的に表現するための道具であり、スケーリング関数、ウェーブレット関数と呼ばれる2種類の関数から作られる基底関数の線形結合によって関数などを表現する。スケーリング関数、ウェーブレット関数の組(この組もウェーブレットと呼ばれる)については、すでに数種類が開発されており適宜選択することができる。しきい値法(Thresholding)は、ウェーブレットによるノイズ除去のための代表的な手法であり、観測値にもとづく関数の推定を基底関数の線形結合で表現するための係数を「しきい値」と比較し、しきい値よりも絶対値が小さな係数の値を0で置き換える。この置き換えによって小さな変動を無視することによりノイズを取り除く。しかし、しきい値法によって得られるノイズ除去の結果はウェーブレットの種類に依存するので、しきい値法で用いるウェーブレットの種類によってしきい値法の結果も異なる。それゆえに観測値に応じてウェーブレットを選択する必要がある。本論文では種々のウェーブレットから観測値に適するものを数値的に選択するための基準を示す。選択基準としてはAIC(赤池情報量基準)を用い、選択の対象となるウェーブレットの中で最小のAICをとるものを最良として選択する。小さなAICに対応するウェーブレットほどしきい値法における考えをよく満たすことが、AICを選択基準として採用した根拠である。AICによるウェーブレットの選択法について、その適用結果を図によって例示し、シミュレーションを行うことにより、AICが適切なウェーブレットを選択することを示す。また、ウェーブレットを選択するための他の方法(最小自乗法、MDL法)との比較を行う。

## 1. はじめに

ウェーブレットは不規則、不連続な挙動を示す関数(または信号や画像)を効率的に表現するための強力な道具である。ウェーブレットはスケーリング関数、ウェーブレット関数と呼ばれる2種類の関数から作られる基底関数を持ち、それらの基底関数は $L^2(\mathbb{R})$ の正規直交基底を構成する。そして基底関数の線形結合によって $L^2(\mathbb{R})$ 関数、信号、画像を表現する。関数を基底関数の線形結合で表現する点はフーリエ級数と同じである。しかしながらフーリエ級数では観測値の局所的な変動がすべての係数に影響する。一方、ウェーブレットにおける基底関数は解像度(細かさの程度)が高くなるにつれてサポートが狭くなり、そのような欠点がない。それゆえにウェーブレット

は局所的な変化や不連続点をもつ関数の表現に優れている。

基底関数のもととなるスケーリング関数、ウェーブレット関数に対しては、サポートが有限区間で、かつ得られる近似関数が滑らかであるという2つの条件が重要である。最も簡単な関数の組(ウェーブレット)としては、階段関数によって構成される Haar ウェーブレットがあるが、Haar ウェーブレットは前者のみを満たしている。これら2つの条件を満たすウェーブレットとしては、1988年に Daubechies によって開発された Daubechies ウェーブレットがあり、その後には Symlet, Coiflet が開発された。

ウェーブレットを用いて信号や画像からノイズを除去する方法としては、しきい値法(Thresholding)が代表的である。しきい値法の手順としては、まず観測された信号(画像)を基底関数の線形結合で表現するために線形結合における係数を計算する。次に線形結合における係数をしきい値と比較し、しきい値よりも絶対値が小さな係数の値を0で置き換える。この置き換えによって小さな変動を無視することによりノイズを取り除く。

しきい値法においては、ノイズ除去の結果として得られる結果は基底関数の線形結合で表現される。それゆえに、しきい値法の結果は、(基底関数のもととなる)ウェーブレットの種類に依存する(後掲の図2,3を参照されたい)。この点で、滑らかな関数を核関数で近似するカーネル型推定とは事情が異なる。実際、カーネル型推定では核関数の種類の選択よりもウィンドウ幅の選択が重要となる。また、スケーリング関数、ウェーブレット関数の形はウェーブレットの種類ごとに異なり、滑らかさ、サポートの広さも異なる。すなわちしきい値法においては、それぞれの観測値の表現に適した(適さない)ウェーブレットが存在する。以上の事柄より、しきい値法の際に用いるウェーブレットを選択することには価値があると考えられる。

以上の背景より、本論文では、しきい値法適用の際に用いるウェーブレットを数値的に選択することを試みる。選択のための基準としては AIC を用いる。選択の対象となるウェーブレットを用いたしきい値法の結果のそれぞれに対して AIC を計算し、最小の AIC に対応する種類を最良のウェーブレットとして選択する。AIC は計算が容易であり、本論文の手法に関してはほとんど計算時間を要しない。

ウェーブレットの選択基準としては、すでに MDL(minimum description length)による方法が Saito (1994) によって提案されている(今後はこの方法を MDL 法と略して呼ぶ)。MDL 法は情報理論によってもたらされた基準であり、データを表現するためのコードの長さが最小になるウェーブレットを最良として選択する。

補足であるが、ウェーブレット・パケットにおける最良基底アルゴリズム(Coifman & Wickerhauser, 1992)は本論文で議論されるウェーブレットの選択とは異なる。ウェーブレット・パケットにおいては、ウェーブレットを決定し、それらからウェーブレット・パケット関数  $\{W_{j,b,k}, j \geq j_0, b \in \mathbb{Z}, b \geq 0, k \in \mathbb{Z}\}$  を生成する。ウェーブレット・パケット関数は2節で述べられる正規直交基底  $\{\phi_{j_0,k}, \psi_{j,k}, j = j_0, \dots, \infty, k \in \mathbb{Z}\}$  を包括する冗長な基底関数であり、全ての関数  $\{W_{j,b,k}, j \geq j_0, b \in \mathbb{Z}, b \geq 0, k \in \mathbb{Z}\}$  を用いずとも  $L^2(\mathbb{R})$  関数を表現できる。ウェーブレット・パケットにおける最良基底アルゴリズムは、ウェーブレットを決定し、それらから得られるウェーブレット・パケット関数の中から最適な組み合わせ  $\{(j, b, k)\}$  を観測値に照らし合わせて選ぶ手法であり、ウェーブレットを選択する手法ではない。また、Donoho (1997) は CART(分類回帰樹木)を用いて基底関数を選択する手法を提案したが、この手法は不均等 Haar 基底関数と称する冗長な基底関数の中から最良の組み合わせを選ぶ手法であり、やはりウェーブレットを選

択する手法ではない。

本論文の構成は以下のとおりである。2節では、関数(または信号)を基底関数の線形結合で表現するための方法や基底関数の構成方法に関して簡単に述べる。3節では、ノイズ除去のためのしきい値法について述べる。4節では、しきい値法の際に用いるウェーブレットをAICによって選択するための方法とその根拠となる考え方について説明する。5節では、AICによって選択されたウェーブレットは観測値に対して適切であるかどうかを図示によって確認する。6節では、シミュレーションを実行することにより、AICは真の関数との差が小さくなる推定関数を選ぶかどうかを調べる。7節では、ウェーブレットを選択するための基準の比較を行う。比較に用いられる基準は、AIC, MDL, 最小自乗法の3種類である。8節が結びである。

## 2. ウェーブレットによる関数およびデータの表現

ウェーブレットは、スケーリング関数、ウェーブレット関数と呼ばれる2種類の関数から生成される基底関数を持ち、 $L^2(\mathbb{R})$  関数を基底関数の線形結合で表現する手法である。

基底関数は多重解像度解析から作られる。以下の条件を満たす空間の列  $\{V_j\}_{j \in \mathbb{Z}}$  を  $L^2(\mathbb{R})$  の多重解像度解析という (Ogden, 1997) :

1.  $\dots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \dots$ ,
2.  $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$ ,  $\overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R})$ ,
3.  $f \in V_j$  の時に限って  $f(2 \cdot) \in V_{j+1}$ ,
4.  $f \in V_0$  ならば,  $k \in \mathbb{Z}$  に対して  $f(\cdot - k) \in V_0$ ,
5. 集合  $\{\phi_{0,k} = \phi(\cdot - k), k \in \mathbb{Z}\}$  が  $V_0$  の正規直交基底を構成するような  $\phi \in V_0$  が存在する。

ここで、 $\bar{V}$  は  $V$  の閉包を表す。添え字  $j$  は解像度を表し、値が大きいほど解像度が高くなる、すなわち細かくなる。ここで

$$\phi_{j,k}(x) = 2^{j/2} \phi(2^j x - k)$$

と定義する。このとき  $\{\phi_{j,k}, k \in \mathbb{Z}\}$  は  $V_j$  の正規直交基底を構成する。関数  $\phi_{j,k}$  を生成する関数  $\phi$  をスケーリング関数と呼ぶ。次に

$$V_{j+1} = V_j \oplus W_j \tag{1}$$

となる空間  $W_j$  を考える。そして  $W_0$  の正規直交基底を構成する関数を  $\{\psi_{0,k} = \psi(\cdot - k), k \in \mathbb{Z}\}$  で表す。関数  $\phi$  の場合と同様に

$$\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k)$$

と定義する。このとき  $\{\psi_{j,k}, k \in \mathbb{Z}\}$  は  $W_j$  の正規直交基底を構成する。関数  $\psi_{j,k}$  を生成する関数  $\psi$  をウェーブレット関数と呼ぶ。このようにして  $\phi_{j,k}$  と  $\psi_{j,k}$  が作られる。

式(1)における分解の繰り返しと多重解像度解析の条件から次式が分かる：

$$L^2(\mathbb{R}) = \overline{V_{j_0} \oplus \bigoplus_{j=j_0}^{\infty} W_j}.$$

これより  $\{\phi_{j_0,k}, \psi_{j,k}, j = j_0, \dots, \infty, k \in \mathbb{Z}\}$  は  $L^2(\mathbb{R})$  の正規直交基底を構成する。関数  $f \in L^2(\mathbb{R})$

のウェーブレット表現は次式で表される：

$$f(x) = \sum_k c_{j_0,k} \phi_{j_0,k}(x) + \sum_{j=j_0}^{\infty} \sum_k d_{j,k} \psi_{j,k}(x), \quad (2)$$

$$c_{j_0,k} = \int_{\mathbb{R}} f(x) \cdot \phi_{j_0,k}(x) dx, \quad d_{j,k} = \int_{\mathbb{R}} f(x) \cdot \psi_{j,k}(x) dx.$$

係数  $c_{j_0,k}$  はスケーリング係数,  $d_{j,k}$  はウェーブレット係数と呼ばれる. 関数  $f$  からスケーリング係数, ウェーブレット係数  $\{c_{j_0,k}, d_{j,k}, j = j_0, \dots, \infty, k \in \mathbb{Z}\}$  への変換をウェーブレット変換と呼ぶ.

現実の適用においては, 離散点における観測値に対してウェーブレット変換が行われる. 位置  $x = 0$  と  $x = 1$  の間に等間隔に並んだ有限個の観測値  $y_1, \dots, y_n$  に対する経験スケーリング係数, 経験ウェーブレット係数は次式で計算される：

$$c_{j_0,k} = \sum_{i=1}^n y_i \cdot \phi_{j_0,k}(i/n), \quad d_{j,k} = \sum_{i=1}^n y_i \cdot \psi_{j,k}(i/n). \quad (3)$$

なお本論文では, 観測値が  $x = 0$  と  $x = 1$  の間に等間隔に並ぶことを今後も仮定する.

スケーリング関数  $\phi$  とウェーブレット関数  $\psi$  の具体的な関数形については, さまざまな関数の組(ウェーブレット)が開発されている. 代表的なウェーブレットを紹介し, それらウェーブレットのいくつかを図 1 に示す：

**Haar** ウェーブレット： Haar によって 1909 年頃作られた階段関数である. ただし, その当時はウェーブレットの概念は存在しなかった.

**Daubechies** ウェーブレット： Daubechies (1988) によって開発されたウェーブレットであり, スケーリング関数  $\phi$ , ウェーブレット関数  $\psi$  は有限なサポートをもつ. Daubechies ウェーブレットはカスケードアルゴリズムによる反復操作を無限回行った極限関数として構成されている(芦野・山本, 1997).

Daubechies ウェーブレットはインデックス  $N$  ( $N$ : 整数) によって分類される. インデックス  $N$  が大きくなると, スケーリング関数  $\phi$ , ウェーブレット関数  $\psi$  は滑らかになり, サポートは広がる. かつ, Daubechies ウェーブレットにおけるウェーブレット関数  $\psi$  の  $N - 1$  次までのモーメントが 0 になる：

$$\int_{\mathbb{R}} x^l \psi(x) dx = 0, \quad l = 0, 1, \dots, N - 1.$$

また,  $N = 1$  の Daubechies ウェーブレットは Haar ウェーブレットに一致することも知られている(榊原, 1995).

**Symlet**： Daubechies ウェーブレットは図 1 からも分かるように, 対称な関数ではない. この点を改良するために Daubechies (1992) によって作られたウェーブレットが Symlet である. Daubechies ウェーブレットを構成するためのツー・スケール数列の選び方を変えることによって, スケーリング関数  $\phi$ , ウェーブレット関数  $\psi$  ができるべく対称に近くなるウェーブレットが構成された.

Symlet も Daubechies ウェーブレットと同様, インデックス  $N$  で分類される. インデックス

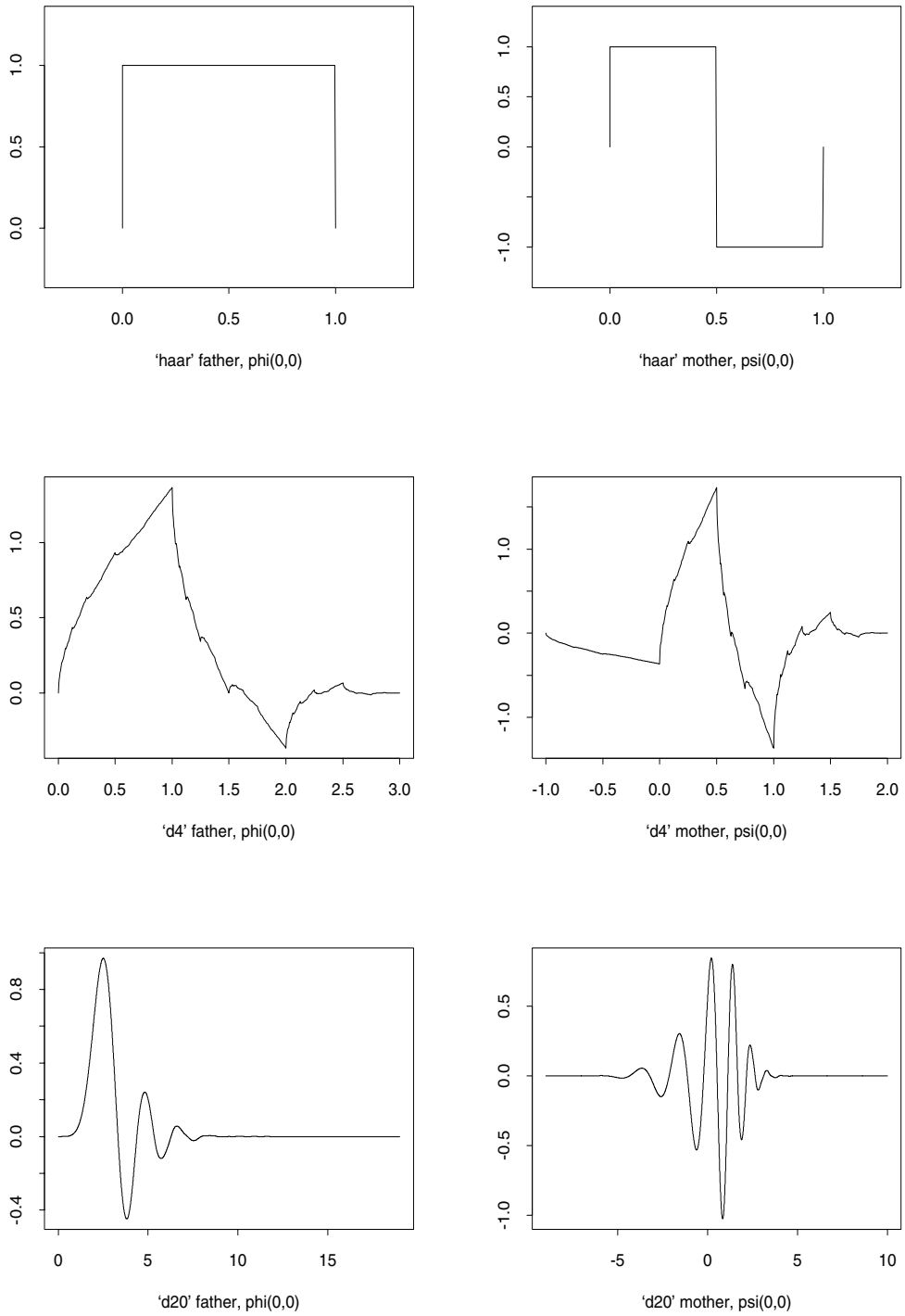


図 1. さまざまな種類のウェーブレット: Haar ウェーブレット  $\phi$  (左上), Haar ウェーブレット  $\psi$  (右上), Daubechies (N=2)  $\phi$  (左中), Daubechies (N=2)  $\psi$  (右中), Daubechies (N=10)  $\phi$  (左下), Daubechies (N=10)  $\psi$  (右下)

AICによるウェーブレット基底関数の選択

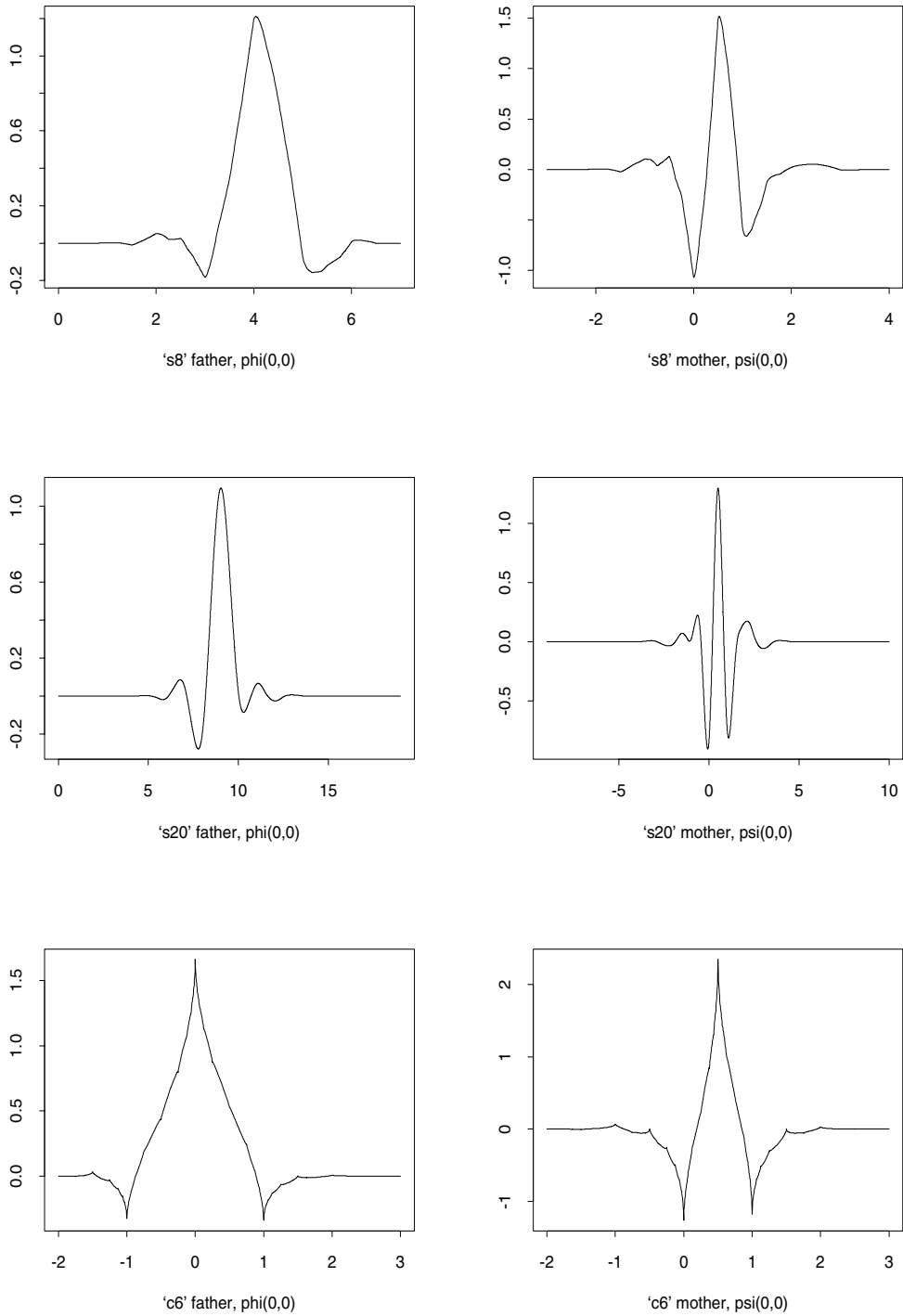


図 1. さまざまな種類のウェーブレット(続き) : Symlet (N=4)  $\phi$ (左上), Symlet (N=4)  $\psi$ (右上), Symlet (N=10)  $\phi$ (左中), Symlet (N=10)  $\psi$ (右中), Coiflet (N=2)  $\phi$ (左下), Coiflet (N=2)  $\psi$ (右下)

$N$  についての性質は Daubechies ウェーブレットの場合と同じである。また  $N = 0, 1, 2$  の場合はツェー・スケール数列の選び方が一意に決定するので、 $N = 0, 1, 2$  の Symlet は Daubechies ウェーブレットと同一なものとなる。

**Coiflet:** Coiflet も Daubechies によって開発された (Daubechies, 1993)。上で述べられた Daubechies ウェーブレット、Symlet はウェーブレット関数  $\psi$  に関して  $N-1$  次までのモーメントが 0 であるが、Coiflet ではスケーリング関数  $\phi$  に関しても  $N-1$  次までのモーメントが 0 である。Coiflet という名前は、スケーリング関数のモーメントも 0 にすることを提案した Ronald Coifman にちなんで付けられた。また Coiflet は数値解析の分野で特に利用されている。

### 3. ウェーブレットによるノイズ除去(しきい値法)

#### 3.1. しきい値法におけるモデル

0 から 1 の間に等間隔に並んだ有限個の値  $y_1, \dots, y_n$  が観測されているとき、次の回帰モデルを考える：

$$y_i = f(i/n) + \eta_i, \quad i = 1, \dots, n. \quad (4)$$

$f$  は真の関数、 $\eta_i$  はノイズ(誤差)で  $\eta_i \sim \text{i.i.d. } N(0, \sigma^2)$  とする。式 (4) をベクトルで表現する：

$$\mathbf{y} = \mathbf{f} + \boldsymbol{\eta}. \quad (5)$$

式 (5) をウェーブレット変換することによって次のモデルが得られる(ここではスケーリング係数とウェーブレット係数を合わせて一つのベクトルと考える)：

$$\mathbf{w} = \mathbf{v} + \boldsymbol{\omega}. \quad (6)$$

ここで  $\mathbf{w}$  は  $\mathbf{y}$  に対応する係数(観測値に対応する係数)であり、式 (3) で計算される  $\{c_{j_0,k}, d_{j,k}\}$  を要素としてもつ。ベクトル  $\mathbf{v}$  は  $\mathbf{f}$  に対応する係数(真の係数)、 $\boldsymbol{\omega}$  は  $\boldsymbol{\eta}$  に対応する係数(ノイズ)を表す。ベクトル  $\mathbf{w}$  はやはり  $n$  個の点から成り立つ。ウェーブレット変換は直交変換なので  $\boldsymbol{\omega}$  の共分散行列は  $\sigma^2 \mathbf{I}$  ( $\mathbf{I}$ : 単位行列) のままである。

観測値に対応する係数の中で、その絶対値が大きな係数は大きな真値をもつと式 (6) より考えられる。また、絶対値が小さな係数はほとんどノイズで占められていると考えられる。それゆえに、ノイズの混ざった観測値から真値を抽出しノイズを除くために、観測値に対応する係数の絶対値を考慮する。

#### 3.2. しきい値法

ウェーブレットによるノイズ除去のための代表的な手法はしきい値法である。しきい値法は Donoho & Johnstone (1994) によって開発された手法であり、推定される真の関数は少数個の基底関数の線形結合で表現できるという考えに基づく。しきい値法では、観測値から計算された係数の絶対値がしきい値よりも小さい場合はその係数の値を 0 で置き換える。この置き換えによって関数の特徴を保ちつつノイズを除去する。ノイズ除去の結果は以下の手順で得られる：

1. 観測値  $y_1, \dots, y_n$  から係数  $\{c_{j_0,k}, d_{j,k}, j = j_0, \dots, J, k \in \mathbb{Z}\}$  を計算する(ウェーブレット変換を行う)。有限個の観測値に対しては、それを表現するための解像度も有限であり、最高の解像度を  $J$  とする。

2. しきい値よりも絶対値が小さい係数  $d_{j,k}$  については、その値を 0 に置き換える．しきい値よりも絶対値が大きい係数  $d_{j,k}$  については、その値を保存する．係数  $c_{j_0,k}$  は観測値のおおまかな形を表現するための係数なので、0 への置き換えは行われない．
3. ステップ 2 で置き換えられた係数を用いて関数  $f(x)$  を推定する．

これらの手順と式 (2) より、真の関数に対する推定関数  $\hat{f}$  は次式で表される：

$$\hat{f}_\lambda(x) = \sum_k c_{j_0,k} \phi_{j_0,k}(x) + \sum_{j=j_0}^J \sum_k I_{\{|d_{j,k}| > \lambda\}} \cdot d_{j,k} \psi_{j,k}(x). \quad (7)$$

関数  $I_{\{\cdot\}}$  は指標関数を表し、 $\lambda$  はしきい値を表す．

しきい値法においては、しきい値  $\lambda$  は平滑化スプラインにおける平滑化パラメータと同じ働きを示す．しきい値を大きくすれば、多くの係数が 0 に置き換えられて、結果として推定関数(式 (7))における線形結合の数が少なくなり滑らかな推定関数が与えられる．しかしバイアスが大きくなる恐れがある．またしきい値を小さくすれば、0 に置き換えられる係数の数が減り、結果としてギザギザとした推定関数もたらされる恐れがある．しきい値はこれらのバランスに配慮して選ばれる必要があり、しきい値の決定方法に関してはさまざまな提案がなされている．

#### 4. AIC によるウェーブレットの選択

本節では、しきい値法を適用する際に、観測値に対するウェーブレット変換で用いるウェーブレット(スケーリング関数  $\phi$ 、ウェーブレット関数  $\psi$ )を AIC によって選択することを試みる．しきい値法においては推定関数は基底関数の線形結合で表される(式 (7))．それゆえに、しきい値法で用いるウェーブレットが異なると(ウェーブレットの種類については図 1 を参照)、推定関数も異なる．また、ウェーブレットはおのおの特徴的な形の関数であるので、ある観測値を表現する場合を考えると、適切なウェーブレット、不適切なウェーブレットが存在する．それゆえに観測値に対して適切なウェーブレットを選択することを試みる．

AIC によるウェーブレットの選択は以下の手順で実行される：

1. 選択の対象となるそれぞれのウェーブレット(例えば、Haar ウェーブレット、Daubechies ウェーブレットなど)を用いて、観測値  $y_1, \dots, y_n$  に対するウェーブレット変換を行う．ウェーブレットはそれぞれ異なった形状をもつので、同じ観測値を表現する場合においても、ウェーブレットごとに値の異なった係数が計算される．
2. それぞれのウェーブレットを用いることによって計算された係数に対して、しきい値法を適用し、それぞれのウェーブレットに対応する推定関数  $\hat{f}_\lambda(x)$  を求める．しきい値は適当な方法によって決定する(本論文における適用例、シミュレーションにおいては Universal 法(Donoho & Johnstone, 1994)を用いて、しきい値を決定した)．
3. それぞれのウェーブレットに対応する推定関数  $\hat{f}_\lambda(x)$  に対して AIC を計算する．実際の計算には、離散点における推定関数の値  $\hat{f}_\lambda(1/n), \dots, \hat{f}_\lambda(1)$  を用いる：

$$\text{AIC} = n \log \left\{ \sum_{i=1}^n (\hat{f}_\lambda(i/n) - y_i)^2 \right\} + 2 \times (\text{推定関数の作成に含めた係数の個数}). \quad (8)$$



式 (8) における (推定関数の作成に含めた係数の個数) とは, 式 (7) における係数  $\{c_{j_0,k}, k \in \mathbb{Z}\}$  の個数と, 係数  $\{d_{j,k}, j = j_0, \dots, J, k \in \mathbb{Z}\}$  の中でしきい値  $\lambda$  よりも大きな絶対値をもつ係数の個数の和を意味する. ウェーブレットが異なると, 変換の結果求められる係数の値も異なるので, 式 (8) における係数の個数も異なる.

式 (7) は,  $\phi_{j_0,k}$  と  $\psi_{j,k}$  が説明変数である重回帰式とみなすことができる. それゆえに, 式 (8) における AIC は重回帰分析における変数選択で用いられるものに準じた.

4. 計算された中で, 最も小さな AIC に対応するウェーブレットを, その観測値に対して最適として選択し, 最適なウェーブレットを用いたしきい値法の結果を最適な推定関数とする.

AIC を用いてウェーブレットを選択することについての根拠となる考え方を述べる.

少数個の係数のみを用いて観測値における情報の多くを表現できる場合に AIC が小さな値をとることは, 式 (8) より分かる. 本論文では, ノイズの分布として独立同一な正規分布を仮定しているので, 観測値における情報の多くを表現することは, すなわち真の関数における情報の多くを表現することにつながる. つまり小さな AIC に対応するウェーブレットは, 少数個の係数のみを用いて真の関数における多くの情報を表現することができるといえる. このことはしきい値法が基とする考えに対応し, それゆえに小さな AIC に対応するウェーブレットはしきい値法が基とする考えをよく満たすウェーブレット, すなわちしきい値法の適用において適切なウェーブレットということができる.

以上の理由より, 最も小さな AIC に対応するウェーブレットを最適として選択する.

なお本論文では, しきい値法の際に用いるウェーブレットを選択するために AIC を用いているが, 本論文における AIC の利用方法は AIC の本来の利用方法と異なることを補足する. AIC は本来, 最良なモデルを含む包括的なモデルの族からのモデル選択のための基準である.

つまり本論文における AIC の利用は, 最少個数の係数のみを用いて観測値における最大限の情報を表現する場合に値が小さくなるという AIC の性質のみに基づくものである.

## 5. 図による例示

4 節では, AIC によってウェーブレットを選択する方法と考え方を述べた. 本節では, AIC によって選択されたウェーブレットは, 観測値に対して適切であるかを図によって視覚的に確認する.

### Blocks 関数による例示

Donoho & Johnstone (1994) の Blocks 関数を用いて例示を行う. ただし Blocks 関数を定数倍することにより, 最大値と最小値の差を 100 に変更してある. 例示は以下の要領で行った.

Blocks 関数に対して, 独立同一な正規分布に従う乱数をノイズとして加えた値を観測値とする. ノイズにおける標準偏差  $\sigma$  は 5, 観測値のサンプルサイズ  $n$  は 256 とする. その観測値に対して, さまざまなウェーブレットを用いて, しきい値法を適用する. しきい値の決定方法としては Universal 法を用いる (今後の適用例, シミュレーションにおいても常に Universal 法を用いる). 式 (7) における  $j_0$  の値に関して, 本論文の適用例, シミュレーションにおいては, 使用したソフトウェア (S+ Wavelets version1.1) における初期設定を用いる. この場合,  $j_0$  の値は, サンプル

サイズ  $n$  と用いられるウェーブレットの種類(厳密にはウェーブレットにおけるフィルターの長さ)に応じて自動的に計算される. Universal 法を適用するためには, ノイズの標準偏差  $\sigma$  を推定する必要があるが, その推定には最高の解像度  $J$  をもつ係数  $\{d_{J,k}, k \in \mathbb{Z}\}$  の Median of Absolute Deviation (MAD):

$$\hat{\sigma} = \frac{\text{median}(|d_{J,k} - \text{median}(d_{J,k})|)}{0.6745}$$

を用いた. それぞれのウェーブレットを用いたしきい値法で得られた推定関数から AIC を計算し, またそれぞれの推定関数を図示する. 小さな AIC に対応するウェーブレットは, 真の関数と視覚的に類似する推定関数を導いているかを図によって確認する.

比較に用いるウェーブレットは, Haar ウェーブレット, Daubechies ウェーブレット ( $N = 2, N = 6, N = 10$ ), Symlet ( $N = 4, N = 7, N = 10$ ), Coiflet ( $N = 2, N = 6, N = 10$ ) の 10 種類である. スケーリング関数  $\phi$ , ウェーブレット関数  $\psi$  の形については一部を図 1 に示してある. これらの関数はいずれも有限区間からなるサポートをもつ点で優れており, 頻繁に用いられる. また Daubechies ウェーブレットなどはインデックス  $N$  により異なるウェーブレットを構成するので, インデックス  $N$  についても 3 通りずつを設定した.

Blocks 関数に対しては, Haar ウェーブレットが最適であることは視覚的に明らかであろう. AIC によるウェーブレットの選択においても, Haar ウェーブレットが最適として選択されるか (Haar ウェーブレットに対する AIC が最小となるか) も確認したい.

Blocks 関数に対する例示の結果(しきい値法の結果の図示と計算された AIC)を図 2 に示す. 図 2 では, 最も上の段において Blocks 関数とノイズの混ざった観測値を示す. 2 段目以降では, それぞれ表示されたウェーブレットを用いて得られたしきい値法の結果を示す. 図を見たところ, やはり Haar ウェーブレットによる推定関数が真の関数と視覚的に最も類似しており, なおかつ AIC が最小となっている.

### Bumps 関数による例示

Donoho & Johnstone (1994) の Bumps 関数を用いて例示を行う. Bumps 関数には多少の変更を加えてあり, また最大値と最小値の差を 100 に変更してある.

例示の要領, 比較に用いるウェーブレットは Blocks 関数による例示と同じである.

Bumps 関数に対する例示の結果(しきい値法の結果の図示と計算された AIC)を図 3 に示す. Bumps 関数の場合, Daubechies ウェーブレットなどにおいてインデックス  $N$  が大きくなるにしたがって推定関数の精度が悪くなり, AIC の値も大きくなる傾向が見られた. インデックス  $N$  が大きくなるにしたがって, 関数  $\phi, \psi$  のサポートは広がる. 急激な尖りをもつ Bumps 関数へはサポートの狭いウェーブレットが適するので, この傾向が現れた.

Bumps 関数による例示においても, 全般的な傾向として, 小さな AIC に対応するウェーブレットが見栄えがよく, 真の関数と視覚的に類似する推定関数を導いているといえそうである.

## 6. シミュレーションによる検証

本節では, AIC によって選択されたウェーブレットによる推定関数は, 適切な(真の関数に近い)推定関数であるかどうかをシミュレーションによって検証する. シミュレーションは以下の要

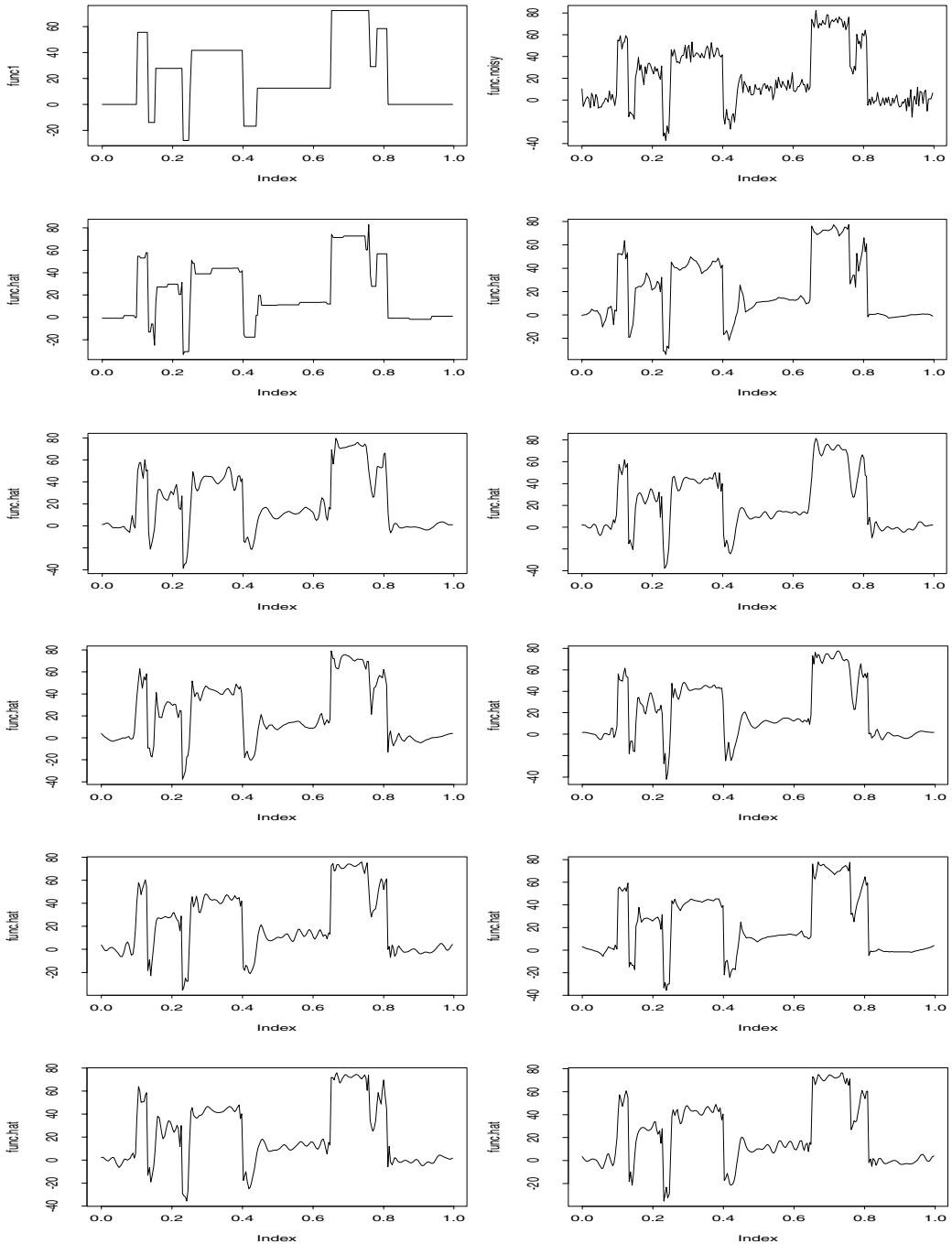


図 2. Blocks 関数に対する例示の結果：真の関数(1 段目左)，ノイズの入った観測値(1 段目右)，Haar ウェーブレット(AIC = 2313) (2 段目左)，Daubechies N=2(AIC = 2411) (2 段目右)，Daubechies N=6(AIC = 2454) (3 段目左)，Daubechies N=10(AIC = 2499) (3 段目右)，Symlet N=4(AIC = 2454) (4 段目左)，Symlet N=7 (AIC = 2415) (4 段目右)，Symlet N=10(AIC = 2443) (5 段目左)，Coiflet N=2(AIC = 2344) (5 段目右)，Coiflet N=6(AIC = 2399) (6 段目左)，Coiflet N=10(AIC = 2442) (6 段目右)

AIC によるウェーブレット基底関数の選択

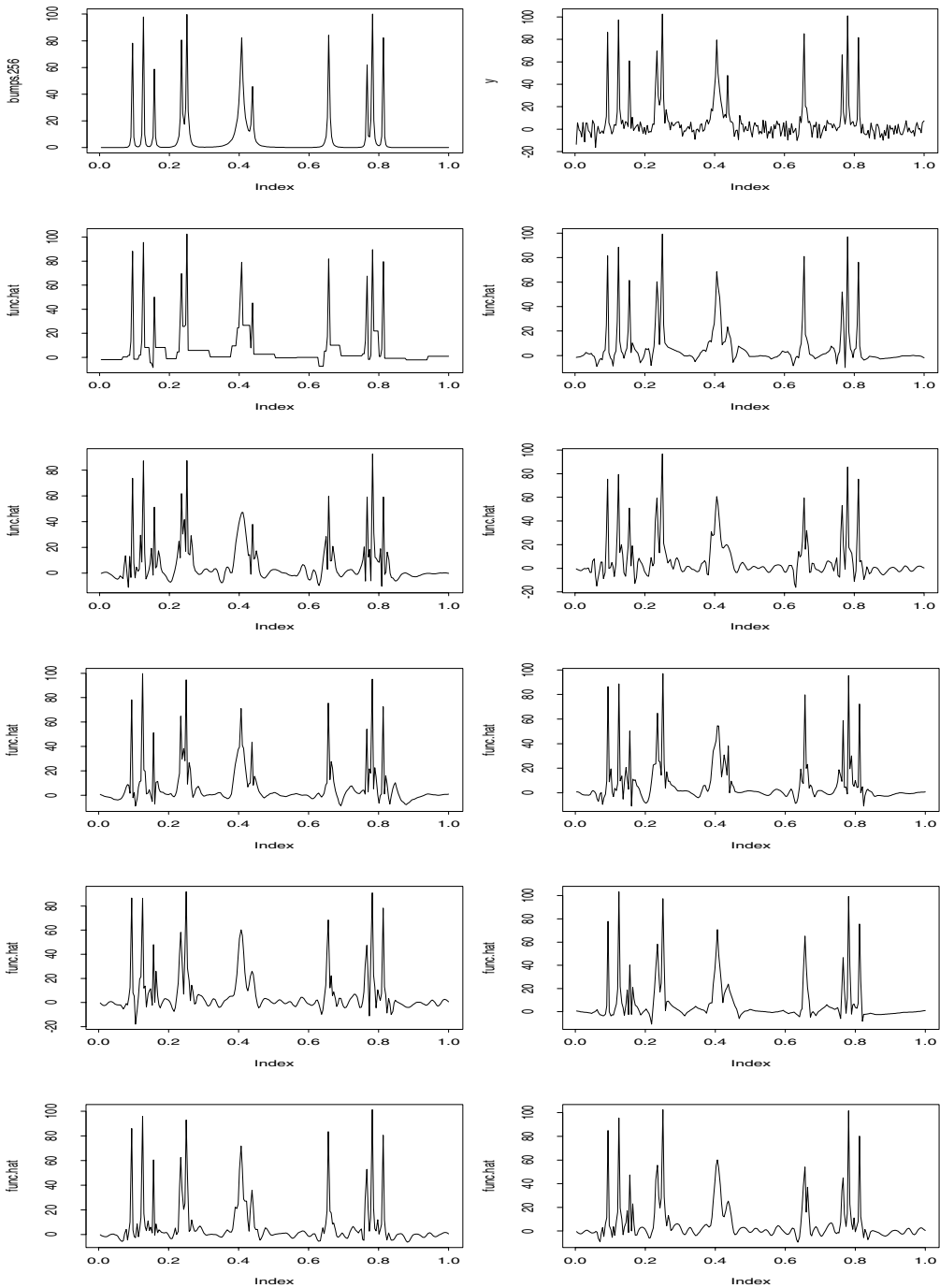


図 3. **Bumps** 関数に対する例示の結果：真の関数(1 段目左), ノイズの入った観測値(1 段目右), Haar ウェーブレット (AIC = 2433) (2 段目左), Daubechies N=2 (AIC = 2440) (2 段目右), Daubechies N=6 (AIC = 2613) (3 段目左), Daubechies N=10 (AIC = 2607) (3 段目右), Symlet N=4 (AIC = 2484) (4 段目左), Symlet N=7 (AIC = 2525) (4 段目右), Symlet N=10 (AIC = 2547) (5 段目左), Coiflet N=2 (AIC = 2473) (5 段目右), Coiflet N=6 (AIC = 2432) (6 段目左), Coiflet N=10 (AIC = 2525) (6 段目右)

領で行う。

シミュレーションに用いたテスト関数は、Donoho & Johnstone (1994) の4つのテスト関数に多少の変更を加えて作成した関数である。変更が加えられた4つのテスト関数を図4に示す(左上: Blocks関数, 右上: Bumps関数, 左下: Doppler関数, 右下: Heavisine関数)。等間隔な離散点における各テスト関数の値に、独立同一な正規分布に従う乱数をノイズとして加えた値を観測値とする。観測値のサンプルサイズ  $n$ , 正規乱数における標準偏差  $\sigma$  に関しては6通りの組み合わせ  $(n, \sigma) = (64, 5), (256, 5), (1024, 5), (256, 2.5), (256, 10)$  を用いる。各テスト関数の最大値と最小値の差を100に揃えているので、例えば  $\sigma = 5$  はその5%の大きさにあたる。

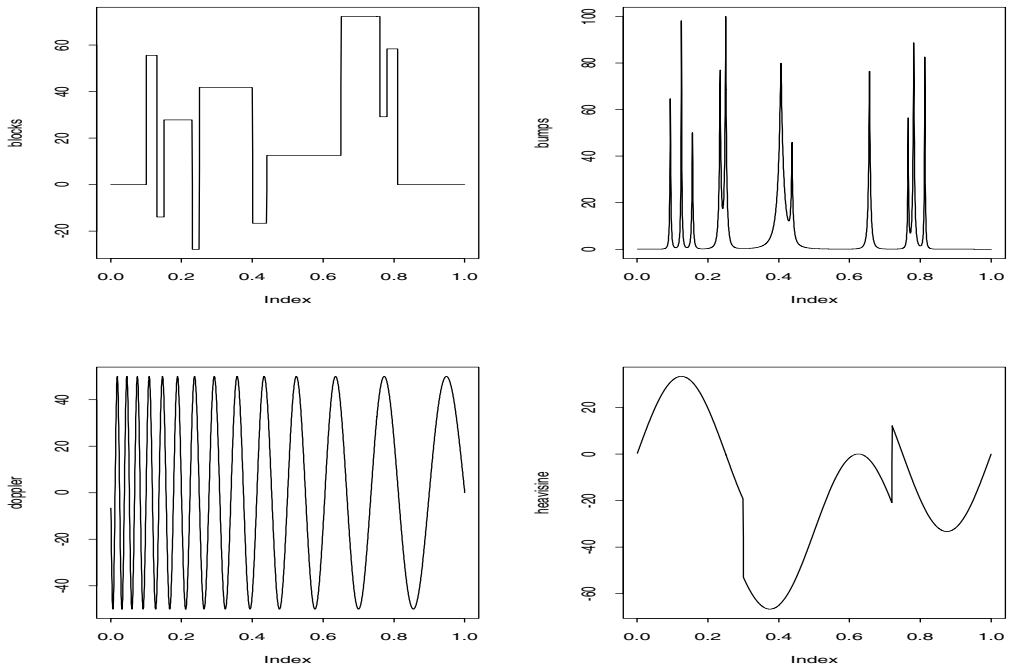


図4. シミュレーションで用いるテスト関数: Blocks関数(左上), Bumps関数(右上), Doppler関数(左下), Heavisine関数(右下)

それらの設定の下で生成した観測値に対して、5節における例示と同様の10種類のウェーブレットを用いて、5節における例示と同様の設定でしきい値法を行い、それぞれのウェーブレットに対応する推定関数  $\hat{f}_\lambda(x)$  からAICを計算する。同時に、それぞれのウェーブレットに対応する推定関数  $\hat{f}_\lambda(x)$  と真の関数  $f(x)$  の差を求める。実際の計算には、離散値における関数の値  $\hat{f}_\lambda(1/n), \dots, \hat{f}_\lambda(1), f(1/n), \dots, f(1)$  を用いて、Average Squared Error (ASE)を計算する:

$$ASE = \frac{1}{n} \sum_{i=1}^n (\hat{f}_\lambda(i/n) - f(i/n))^2. \tag{9}$$

次にAICによって選択された(AICが最小となる)ウェーブレットに対して、ASE比を計算する:

$$\text{ASE 比} = \frac{\text{あるウェーブレットについての ASE}}{10 \text{ 種類のウェーブレットについての ASE における最小値}}$$

AICによって選択されたウェーブレットが、観測値に対して最適であるならば、そのウェーブレットについての ASE の値は ASE の最小値と一致する、すなわち ASE 比は 1 となるはずである。

比較のために、2 番目、3 番目、5 番目に小さな AIC をとるウェーブレットについても ASE 比を計算する。小さな AIC をとるウェーブレットほど、真の関数との差が小さくなる傾向が存在するならば、2 番目、3 番目、5 番目と AIC が大きくなるに従って、ASE 比は大きな値となるはずである。

シミュレーションでは、設定された条件の下で、観測値の生成と推定を 1000 回繰り返す。そして 1000 回それぞれにおいて ASE 比を計算する。シミュレーションの結果は ASE 比の中央値で与えられる。

シミュレーションの結果を表 1 に示す。AIC が小さいほど ASE 比が 1 に近くなる傾向が、ほとんどの条件に対して表 1 より確認できる。それゆえに、小さな AIC に対応するウェーブレットほど真の関数に近い推定関数を導くことが、さまざまな条件のもとで示された。Heavisine 関数 ( $n = 64, \sigma = 5$ ) と Heavisine 関数 ( $n = 256, \sigma = 10$ ) の 2 つの条件に関しては、AIC と ASE 比についての傾向が逆転した箇所もあるが、逆転における数値幅は小さく、実用においては問題にな

表 1. シミュレーションの結果：AIC によるウェーブレットの選択法の検証

テスト関数	$n$	$\sigma$	最小	2 番目	3 番目	5 番目
Blocks	64	5	1	1.3548	1.6967	2.2697
	256	5	1	1.8528	2.0642	2.2490
	1024	5	1	1.7006	1.8685	2.4459
	256	2.5	1	2.0631	2.4811	3.7410
	256	5	1	1.8528	2.0642	2.2490
	256	10	1	1.2442	1.2535	1.3119
Bumps	64	5	1	1.8313	2.3415	3.4075
	256	5	1	1.1477	1.2853	1.5373
	1024	5	1	1.1614	1.3504	1.5463
	256	2.5	1	1.3576	1.5812	1.8139
	256	5	1	1.1477	1.2853	1.5373
	256	10	1	1.1474	1.2511	1.4071
Doppler	64	5	1	1.1748	1.3214	1.9405
	256	5	1.0699	1.1176	1.1707	1.3344
	1024	5	1.0441	1.0801	1.1670	1.3892
	256	2.5	1.0256	1.1112	1.1644	1.3234
	256	5	1.0699	1.1176	1.1707	1.3344
	256	10	1.0945	1.1152	1.1444	1.2330
Heavisine	64	5	1.2284	1.2305	1.2279	1.2454
	256	5	1.1442	1.2171	1.2662	1.3478
	1024	5	1.0931	1.1068	1.1426	1.2107
	256	2.5	1.1630	1.2025	1.2489	1.3115
	256	5	1.1442	1.2171	1.2662	1.3478
	256	10	1.2317	1.2109	1.2187	1.2481

らない。

用いたテスト関数と結果との関係について述べる。Blocks 関数, Bumps 関数に対しては AIC が小さいほど ASE 比が小さくなる傾向が顕著であり, Doppler 関数, Heavisine 関数に対してはその傾向はさほどに顕著ではない。Blocks 関数, Bumps 関数は極端な不連続点などの著しい特徴をもつ関数であり, それらの特徴の表現に関しては適切なウェーブレットが明白に存在した。逆に Doppler 関数, Heavisine 関数は比較的滑らかな関数であり, これらのテスト関数に対しては適切なウェーブレットが明白には存在しない場合もあった。テスト関数のこの性質の違いが結果に影響を与えたのではないかと考えられる。

## 7. 最小自乗法, MDL 法との比較

前節のシミュレーションにより, 与えられた観測値に対して最適なウェーブレットを AIC によって選択できることが示された。本節では, ウェーブレットの選択を行う他の方法, すなわち最小自乗法, MDL 法を 7.1 節において紹介し, 7.2 節においては, AIC による方法, 最小自乗法, MDL 法の 3 者の比較をシミュレーションによって行う。

### 7.1. 最小自乗法, MDL 法によるウェーブレットの選択

最小自乗法, MDL 法を用いてウェーブレットを選択する手順を簡単に説明する。最小自乗法では, しきい値法を適用して得られた推定関数  $\hat{f}_\lambda$  に対して残差平方和を計算し, それが最小となるウェーブレットを最良として選択する:

$$\text{残差平方和} = \sum_{i=1}^n (\hat{f}_\lambda(i/n) - y_i)^2.$$

残差平方和による選択は, 推定関数の作成に含めた係数の個数を考慮しない点で AIC による選択方法と異なる。

AIC による選択法では, 観測値に対して, 選択の対象となるウェーブレットごとに, 既存のしきい値法による推定関数を作成し, その後に AIC を用いて最良なウェーブレットの選択を行う。しかし MDL 法では推定関数の作成とウェーブレットの選択を一つの基準で行う。MDL 法では, まず選択の対象となるそれぞれのウェーブレットに対して, AMDL (Approximate Minimum Description Length) を最小にする  $k$  (推定関数の作成に含めた係数の個数) を選択する:

$$\text{AMDL}(k) = \frac{n}{2} \log_2 \left\{ \sum_{i=1}^n (w_i - I_{\{|w_i| \geq w_{(k)}\}} w_i)^2 \right\} + \frac{3}{2} k \log_2 n. \quad (10)$$

ここで  $w_i$  は式 (6) における  $w$  の要素を表し,  $w_{(k)}$  は  $w_i, i = 1, \dots, n$  の中で  $k$  番目に大きな絶対値をもつ係数を表す。それぞれのウェーブレットに対応する推定関数は, AMDL を最小にする  $k$  を選択することによって作成できる。最後に, それぞれのウェーブレットに対応する推定関数の中で最も小さな AMDL をもつ推定関数を最良として選択する。MDL 法においては, 最良なウェーブレットと推定に含める係数の個数を同時に推定する点に関して AIC とは異なる。かつ, MDL 法では解像度の大小に関わらず全ての係数が取捨選択の対象となる点で, しきい値法と異なる。

ウェーブレットの選択についての他の方法としては, Nason によるクロスバリデーション法(例えば, Ogden (1997) を参照)を応用することなどによっても可能であるが, 本論文では考慮しない。

## 7.2. シミュレーションによる比較

本小節では、AIC による選択方法、最小自乗法、MDL 法の 3 つの方法をシミュレーションによって比較する。シミュレーションの実行条件は以下の通りである。用いたテスト関数は図 4 に示された 4 つの関数、観測値のサンプルサイズとノイズの標準偏差は  $(n, \sigma) = (256, 5)$  とする。この条件に対して、観測値を 6 節のシミュレーションと同じ方法により生成する。観測値に対して、5 節の要領でしきい値法を適用し、AIC によるウェーブレットの選択を行い、同時に最小自乗法による選択を行う。選択の対象としたウェーブレットは、5 節における例示と同様の 10 種類のウェーブレットとする。それぞれの方法によって選ばれたウェーブレットに対応する推定関数に対して、真の関数との距離を測るために、式 (9) の ASE を計算する。また、MDL 法による推定関数も同時に作成し、MDL 法による推定関数に対しても ASE を計算する。MDL 法においても、選択の対象としたウェーブレットは、5 節における例示と同様の 10 種類とする (MDL 法では、ウェーブレット・パケット基底も選択の対象となることが Saito (1994) で指摘されているが、本論文ではウェーブレット・パケット基底は選択肢には含めなかった。補足であるが、AIC による選択法もウェーブレット・パケットへの拡張は MDL 法と同様に可能である)。ノイズの標準偏差  $\sigma$  について、AIC、最小自乗法による選択では 5 節と同様に MAD で推定した (MDL 法では、 $\sigma$  の推定は不要である)。シミュレーションにおいては、観測値の生成と推定関数の作成、ASE の計算を 500 回繰り返す。シミュレーションの結果は、得られた ASE の平均値で与えられる。得られた ASE が小さいほど良い選択方法といえるであろう。

シミュレーションの結果を表 2 に示す。この表から、AIC による選択は最小自乗法による選択よりも、全体的にみれば ASE の平均値が小さい (真の関数との差が小さい) ことが分かり、ウェーブレットを選択する上で、推定関数の作成に含める係数の個数を考慮することの有効性が示唆された。最小自乗法による選択が AIC による選択に劣った理由としては、最小自乗法では観測値に当てはまりすぎる推定関数を導くウェーブレットが選択され、それゆえに推定関数はギザギザとした関数となり、結果として真の関数との距離 (ASE) が大きくなったことが考えられる。特に、Haar ウェーブレットを用いれば少数個の係数で表現できる Blocks 関数に対しては、係数の個数を考慮する価値は大きいだろう。しかし、Heavisine 関数に対しては最小自乗法がより小さな ASE を達成したことより、Heavisine 関数に対しては係数の個数を増やしてでも観測値との差を小さくすることが、真の関数との差を小さくする上で有効であったといえる。

MDL 法では ASE の平均値が大きな値となったが、これは MDL 法がしきい値法を用いず、AMDL を最小にするように係数を取捨選択することに起因したかと思われる。しきい値法においては、解像度の低い係数は推定対象の関数の大まかな形状を表すと考えられることより、しきい値と比較され 0 に置き換えられる解像度には制限が設けられている。具体的には式 (7) の  $j_0$  以下

表 2. シミュレーションの結果 2: AIC, 最小自乗法, MDL 法の比較

$n = 256, \sigma = 5$	AIC	最小自乗法	MDL 法
Blocks	9.946	11.410	29.157
Bumps	21.704	23.116	199.509
Doppler	26.524	27.364	100.674
Heavisine	12.727	11.739	111.922



の解像度においては置き換えは行われぬ。しかし、MDL 法においては、全ての解像度の係数が取捨選択の対象となる。このことが、MDL 法に対する ASE が大きくなったことと関連するかもしれない。AIC と MDL はどちらも、少数個の係数のみを用いつつ、観測値との差が小さい推定関数を探索するための基準である。AIC と MDL は開発された背景こそ違おうが、考え方は類似しており、それゆえに今回のシミュレーションにおける差は、AIC と MDL の基準の定義(式 (8) と式 (10))の違いではなく、しきい値法の適用の有無に起因すると考える方が自然であろう。

ところで、AIC による選択は MDL 法よりも常に優れているわけではない。例えば、 $(n, \sigma) = (1024, 1)$  の条件のもとで、Blocks 関数を用いて上述のシミュレーションを行ったところ、ASE の平均値は AIC による選択に対しては 0.1016、MDL 法に対しては 0.0754 となり、MDL 法がより真の関数に近い推定関数を導いた。これらの方法の間には絶対的な優劣は存在せず、状況に応じた方法を適宜選択することが現実的であると思われる。

## 8. 結 び

本論文では、しきい値法を適用する際に、観測値に対するウェーブレット変換で使用するウェーブレットを AIC によって選択することを試みた。より少ない個数の係数のみを用いて観測値におけるより多くの情報を表現できるウェーブレット、すなわち小さな AIC をとるウェーブレットはその観測値に対して適するという考え方から AIC によるウェーブレットの選択を試みた。

5 節の図による例示、6 節のシミュレーションによる検証を通して AIC によってウェーブレットを選択することの効果を調べたが、結果としては小さな AIC に対応するウェーブレットほど見栄えの良く、かつ真の関数に近い推定関数を導くことが判明した。7 節の他方法との比較では、AIC による選択法と最小自乗法と比較することにより、推定関数の作成に含める係数の個数を考慮することの有効性が示された。また MDL 法とも比較を行ったが、どのような条件に対しても常に一方が優れているという結果は得られず、状況に応じて手法を選択して適用する必要が示唆された。

謝 辞 本論文の作成にあたり、貴重な助言を頂きました査読者の先生方に感謝いたします。

## 参 考 文 献

- Coifman, R. and Wickerhauser, V. (1992): Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory* **38**, 713–718.
- Daubechies, I. (1988): Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics* **41**, 909–996.
- Daubechies, I. (1992): *Ten Lectures on Wavelets*. SIAM.
- Daubechies, I. (1993): Orthonormal bases of compactly supported wavelets II. Variations on a theme. *SIAM Journal on Mathematical Analysis* **24**, 499–519.
- Donoho, D.L. (1997): CART and best-ortho-basis: A connection. *Annals of Statistics* **25**, 1870–1911.
- Donoho, D.L. and Johnstone, I.M. (1994): Ideal spatial adaptation by wavelet shrinkage. *Biometrika* **81**, 425–455.
- Ogden, R.T. (1997): *Essential Wavelets for Statistical Applications and Data Analysis*. Birkhäuser.
- Saito, N. (1994): Simultaneous noise suppression and signal compression using a library of orthonormal bases and the minimum description length criterion. In *Wavelets in Geophysics*, Fofoula-Georgiou, E. and Kumar, P. (eds.), Academic Press. pp.299–324.
- 芦野 隆一・山本 鎮男 (1997): ウェーブレット解析—誕生・発展・応用. 共立出版.

AIC によるウェーブレット基底関数の選択

榑原 進 (1995): ウェーブレット ビギナーズガイド, 東京電機大学出版局.

(2003 年 3 月 31 日受付 2004 年 4 月 12 日最終修正 6 月 22 日採択)

著者連絡先 : 〒 170-8633 東京都豊島区高田 3-24-1

大正製薬株式会社 医薬開発部

e-mail: yuuki.matsushima@po.rd.taisho.co.jp

松嶋優貴

〒 560-8531 大阪府豊中市待兼山町 1-3

大阪大学 大学院基礎工学研究科 システム創成専攻 数理科学領域

e-mail: sirahata@sigmath.es.osaka-u.ac.jp

白旗慎吾

〒 560-8531 大阪府豊中市待兼山町 1-3

大阪大学 大学院基礎工学研究科 システム創成専攻 数理科学領域

e-mail: sakamoto@sigmath.es.osaka-u.ac.jp

坂本 亘

## Selecting Wavelet Basis Functions by AIC

Yuuki Matsushima<sup>1,\*</sup>, Shingo Shirahata<sup>2</sup> and Wataru Sakamoto<sup>2</sup>

<sup>1</sup> Medical Development Division,  
Taisho Pharmaceutical Co., Ltd.,  
24–1, Takata 3-chome, Toshima-ku, Tokyo, 170–8633, Japan

<sup>2</sup> Division of Mathematical Science,  
Department of Informatics and Mathematical Science,  
Graduate School of Engineering Science,  
Osaka University,  
Machikaneyama-cho 1–3, Toyonaka, Osaka 560–8531, Japan

### Abstract

The wavelet is an effective tool for representing an irregular or discontinuous function (or signal and image), which uses a linear combination of two types of basis functions: a scaling function and a wavelet function. The pair is also called a “wavelet”, and some types of wavelet have already been developed. Thresholding is the most common method of noise reduction for wavelets. With this method, coefficients, which represent the observed data, are compared with a “threshold” value, and are replaced with zero if their absolute values are less than the threshold. This replacement removes noise since it disregards small changes in the data. However, estimates obtained by thresholding depend on the type of wavelet. Therefore, we need to adopt a specific wavelet according to the observed data. We consider a numerical criterion, AIC (Akaike Information Criterion), and adopt the wavelet that corresponds to the least AIC. We illustrate our method with figures and simulations.

**Key words:** multiresolution analysis, noise reduction, orthonormal basis, thresholding.

\* Corresponding author

E-mail address: [yuuki.matsushima@po.rd.taisho.co.jp](mailto:yuuki.matsushima@po.rd.taisho.co.jp)

Received March 31, 2003; Received in final form April 12, 2004; Accepted June 22, 2004.