数理科学

特集 モデルとモデリング
モデル論の構築をめざして

モデルとは何か
予測とモデル
ニューラルネットワークによるモデル化と学習の理論
ロバスト制御とモデル
環境モデル 溫暖化統合評価モデル
経済学のモデル構築 経済成長理論におけるモデルビルディング
計算のモデル ソフトウェアのモデル
実世界カオスとモデリング

《政治の数理》
選挙制度と有権者の影響力
《連載》 微分積分 第14回
予測とモデル

北川源四郎・樋口知之

1. 統計的モデリングと予測

1.1 予測とモデルの評価

われわれをとりまく現実の世界は複雑であり、そこから得られる情報も基本的には不完全な、そして不十分なものである。その不確かさのなかでわれわれが利用するモデルは、所詮その近似に過ぎない。もちろん、モデルは複雑な現象を正確にかつ、なるべく簡潔に表現することが期待されている。しかしながらモデリングの目的が複雑な現象の正確な記述にあるならばモデルを簡潔に表すことの必要性は明らかではない。モデリングにおいてなぜ簡潔さが必要なのか？この問いに始めて予測の視点から明確な答えを与えたのが赤池13であった。

図1に示すように、現在までに得られたデータを用いて、将来現れるデータを推定するのが予測の問題である。予測にモデルを用いる場合には、そのモデルの選択が予測の精度を左右する。現象をより正確に再現することだけにとらわれると、モデルを果てしなく複雑化することになるが、このようなモデルが予測においても実際に良いパフォーマンスを示すとは限らないのである。

1.2 情報量規準とモデリング

赤池は統計的モデルの良さを、予測分布の良さで評価することを提案した。すなわち、統計的モデルの良さをモデルが定める予測分布と真の分布との近さで測ることとし、その近さをカルバックライブナー情報量で評価することを提案した。真の分布が未知の場合にはこの情報量は直接には計算できないが、赤池は対数尤度がその推定値とみなせるものを示した。この考え方から従来、矢崎法は予測分布の最適化を目指した推定法ということになる。赤池はさらに推定されたモデルを評価するときに生じる対数尤度のバイアスを補正し、情報量規準

\[ AIC = -2(対数尤度) + 2(パラメータ数) \] (1)

を提案した。この AIC を用いると、多数のモデルの良さを客観的に評価・比較でき、自然なモデリングの基盤が得られることとなる。AIC の導入を契機に1970年代からさまざまな分野において活発なモデリングが始まるようになった。

AIC の導出において真の分布の平均対数尤度の部分が共通項として除外されている13。この項は真のモデルが未知の場合にはその推定が困難であるが、モデルの比較には不要な項である。しかし、このために AIC は相対的な評価規準であり、絶対的な規準ではないことに注意する必要がある。これはモデリングにおいては、相対的な評価が可能であるが、絶対的な評価は困難であることを示唆している。AIC を用いるとき、考えられた候補のなかで最適なモデルを見つけることはできても、それよりも良いモデルはその候補の外にいくらでもあり得ることになる。したがって、モデリング
においては、固定したモデルの中での改良よりも予測分布の改良を目指してより良いモデルの候補を考えることが要求される（図2）。このモデル探索の姿勢を赤池は、「唯一無二の究極的モデルなどは存在せず、常により良いものを求めて前進することになる」と説明している3)。

2. 時系列の予測と状態空間モデル

2.1 時系列の予測と状態

図3に示すように、$Y_n$を時刻$n$までの観測値$y_1, y_2, \ldots, y_n$からなる空間、$Y_{n+1}$を時刻$n$後の将来的な観測値$y_n, y_{n+1}, \ldots$で張られる空間とする。$Y_n$と$Y_{n+1}$はともに現在の値$y_n$を含む共通部分を持つが、この共通部分を経由して過去の情報は将来に伝達される。$Y_n$と$Y_{n+1}$の共通部分が$y_n$だけの場合には時系列はマルコフ性を持つことになるが、一般にはより複雑な構造を持つ。赤池は時系列の予測においてはこの共通部分が時系列の状態に対応し、その最小次元の決定において最適なモデルが決定されることを示した2)。

2.2 状態の推定の問題

状態に関連する重要な問題は、観測値$y_n$に基づいて時刻$n$の状態$x_n$の推定を行うことである。この状態を推定できれば将来の予測に必要な情報が得られるばかりでなく、時系列の組み、分解、パラメータ推定などの時系列解析に関連する重要な問題の多くが統一的に解決できる。

時刻$j$までの観測値$Y_j = \{y_1, \ldots, y_j\}$に基づいて状態$x_n$の推定を行う場合を考えることにする。$j$と$n$の大小関係により状態推定の問題は以下のように3つの場合に分類される。

$j < n$の場合：予測
$j = n$の場合：フィルタ
$j > n$の場合：平滑化

観測値$Y_j$が与えられたとき、状態$x_n$を推定するためには、$x_n$の条件付き分布$p(x_n|Y_j)$を求めればよい。

2.3 状態空間モデル

この状態の推定のために導入されたモデルが状態空間モデルである。一般に状態空間モデルは、時間推移とともに状態の変化する様子をモデル化したシステムモデルと、状態と観測値の関係を表現した観測モデルの2つのモデルからなる。特に、線形・ガウス型の状態空間モデルはこのような構造を線形モデルと正規ノイズで近似したものである。時系列$y_n$に対して、以下のような2つのモデルをあわせて状態空間モデルと呼ぶ。

$$x_n = F_n x_{n-1} + G_n v_n \quad \text{【システムモデル】} \quad (2)$$
$$y_n = H_n x_n + w_n \quad \text{【観測モデル】} \quad (3)$$

ただし、$x_n$は状態ベクトル、$v_n \sim N(0, Q_n)$と$w_n \sim N(0, R_n)$はそれぞれ正規白色雑音で、システムノイズ、観測ノイズと呼ばれる。また、$F_n, G_n, H_n$はそれぞれ適当な次元の行列である。

時系列解析で用いられる多くのモデルは状態空間モデルの形で表現し統一的に取り扱うことができる1)。例えば観測モデルを時系列$y_n$が観測される仕組みを表す回帰モデルと考えると、状態$x_n$はその回帰係数となる。したがって、状態空間モデルでは回帰係数が時間的に変化する場合を考えることができ、システムモデルが回帰係数の時間的な変化の様子を表現するモデルとなる。また、状態ベクトル$x_n$が推定すべき信号と考えられる場合には、システムモデルは信号伝達のメカニズムを表すモデル、観測モデルは信号を観測
するときの機構を表現するモデルとなる。

2.4 カルマンフィルタと平滑化のアルゴリズム

この線形・ガウス型の状態空間モデルの場合には状態の分布 \( p(x_n|Y_j) \) は正規分布となるので, 状態推定のためには平均ベクトルと分散共分散行列だけを求めればよい。状態 \( x_n \) の条件付き平均と分散共分散行列をそれぞれ \( \mu_{n|i}, \Sigma_{n|i} \) と表すと, \( p(x_n|Y_j) \sim N(\mu_{n|i}, \Sigma_{n|i}) \)

一般に, 観測値 \( Y_j \) が与えられた下での状態 \( \{x_1, \ldots, x_N\} \) の条件付き同時分布を求めるためには多大な計算量を要するが, 状態空間モデルの場合, 逐次的な計算アルゴリズムによって条件付きの周辺分布 \( p(x_n|Y_j) \) はデータ数 \( N \) に比例する計算量で効率的に計算できる。特に線形・ガウス型の状態空間モデルの場合にはカルマンフィルタと呼ばれるアルゴリズムが利用できる。カルマンフィルタは以下のように前期先予測とフィルタの 2 つのステップから構成されている。

[前期先予測]

\[
\begin{align*}
x_{n|n-1} &= F_n x_{n-1|n-1}, \\
V_{n|n-1} &= F_n V_{n-1|n-1} F_n^T + G_n Q_n G_n^T.
\end{align*}
\]

[フィルタ]

\[
\begin{align*}
K_n &= V_{n|n-1} H_n^T (H_n V_{n-1|n-1} H_n^T + R_n)^{-1}, \\
x_{n|n} &= x_{n|n-1} + K_n (y_n - H_n x_{n|n-1}), \\
V_{n|n} &= (I - K_n H_n) V_{n|n-1}.
\end{align*}
\]

カルマンフィルタは時刻 \( n \) までの観測値だけを用いて状態 \( x_n \) を推定するが, 平滑化のアルゴリズムではより多くの観測値を用いて状態推定を行う。したがって, 平滑化を行えばフィルタよりも精度の良い状態推定が実現できる。平滑化に関しては以下のアルゴリズムが知られている。

[固定区間平滑化]

\[
\begin{align*}
A_n &= V_{n|n} F_n^T V_{n+1|n+1}^{-1}, \\
x_{n|N} &= x_{n|n} + A_n (x_{n+1|N} - x_{n+1|n}), \\
V_{n|N} &= V_{n|n} + A_n (V_{n+1|N} - V_{n+1|n}) A_n^T.
\end{align*}
\]

実際にはこの平滑化のアルゴリズムを利用する場合にはまずカルマンフィルタによって \( \{x_{n-1|n}, x_{n|n}, V_{n|n}, V_{n|n-1}, N \} \) を求め, 次に \( \{6\} \) によって時間的に逆方向に \( \{x_{N|N}, x_{N|N}^{\prime} \} \) を順次計算すればよい。

3. 一般化状態空間モデル

3.1 線形・ガウス型モデルの問題点とモデルの拡張

線形・ガウス型の状態空間モデルはいろいろな形の時系列モデルに対して適用でき, 非定常時系列のトレンドの推定, 季節調整, 時変スペクトルの推定などにも適用されている。しかしながら, 時系列には非急激な構造変化が見られる場合, 異常値が存在する場合, 非対称性, 非正規性がある場合, ポアソン過程, 二項過程などの離散過程の場合や非線形システムの場合には, 通常の線形・ガウス型の状態空間モデルでは良い結果は得られない。このような場合にも適切な非線形あるいは非ガウス型のモデルを利用すれば極めて自然にこれらの問題を取り扱うことができる。

システムモデルにおいて, \( F_n \) および \( H_n \) が行列でなく各々 \( x_{n-1} \) および \( x_n \) に関する非線形関数であるような状態空間モデルの拡張を考える。

\[
x_n = F_n(x_{n-1}) + G_n v_n \quad \text{[システムモデル]} (7)
\]
\[
y_n = H_n(x_n) + w_n \quad \text{[観測モデル]} (8)
\]

このような状態とノイズが分離した非線形状態空間モデルを取り扱う場合, 状態遷移を支配する \( F_n \) の非線形性にのみ研究上の多くの興味が集中し, システムノイズおよび観測ノイズの役割を重視しないような設定, 例えば伝統的なカオス時系列モデル \( \text{(9)} \) の場合が相当する。

これとは違った方向の状態空間モデルの拡張は, 時系列 \( y_n \) に対して以下のような線形・非ガウス型モデルを考えるものである。

\[
x_n = F_n x_{n-1} + G_n v_n \quad \text{[システムモデル]} (9)
\]
\[
y_n = H_n x_n + w_n \quad \text{[観測モデル]} (10)
\]

システムノイズ \( v_n \) と観測ノイズ \( w_n \) はそれぞれ, 非ガウス型の密度関数 \( q(v) \) および \( r(w) \) に従う白色雑音である。システム自体に関する情報はある程度十分得られている状況下で, 正規性から大きく外れるようなシステムへの干渉, あるいは異常値への適切な処理などのように, 期待の観測ノイズを取り扱うときにこのようなモデルを利用する。

3.2 一般化状態空間モデルの枠組

非線形性と非ガウス性の両方を組み込んだ非線形・非ガウス型状態空間モデルを考えるのは自然な流れであろう。
\[ x_n = F_n(x_{n-1}, u_n) \quad \text{[システムモデル]} \quad (11) \]
\[ y_n = H_n(x_n, w_n) \quad \text{[観測モデル]} \quad (12) \]

当然このモデルは上述した線形・ガウス型、非線形あるいは非ガウス型時系列モデルをすべて含んでいる。
非線形・非ガウス型状態空間モデルをさらに一般化したものとして、条件付き分布を用いて表現される以下のような一般化状態空間モデルを考えることができる。
\[ x_n \sim Q_n(-x_{n-1}) \quad \text{[システムモデル]} \quad (13) \]
\[ y_n \sim R_n(-x_n) \quad \text{[観測モデル]} \quad (14) \]

ただし、\( Q_n \) と \( R_n \) はそれぞれ状態 \( x_{n-1} \) と \( x_n \) が与えられたときの条件付き分布を表す。このモデルにより、観測値 \( y_n \) が二項分布やポアソン分布に従う count data のモデル化なども行うことができる。音声情報処理で利用されている隠れマルコフモデルは状態 \( x_n \) も離散値のみが許される一般化状態空間モデルの特殊なケースである。このように一般化状態空間モデルの利用は、離散データの発生メカニズムへの仮説を明示的にとりこんだ合理的情報処理を可能にし、シンボルの解析への応用が考えられる。

3.3 一般化状態空間モデルにおける状態の推定法

一般化状態空間モデルに対してもカルマンフィルタと同様に状態の分布 \( p(x_n|Y_{t-1}) \) が原理的には以下の式により逐次的に計算することができる。[一期先予測]
\[ p(x_n|Y_{n-1}) = \int_{-\infty}^{\infty} p(x_n|x_{n-1})p(x_{n-1}|Y_{n-1})dx_{n-1} \quad (15) \]

[フィルタ]
\[ p(x_n|Y_n) = \frac{p(y_n|x_n)p(x_n|Y_{n-1})}{p(y_n|Y_{n-1})} \quad (16) \]

ただし、\( p(y_n|Y_{n-1}) = \int p(y_n|x_n)p(x_n|Y_{n-1})dx_n \) である。

[固定区間平滑化]
\[ p(x_N|Y_N) = p(x_N|Y_N) \int_{-\infty}^{\infty} \frac{p(x_{n+1}|Y_N)p(x_{n+1}|x_n)}{p(x_{n+1}|Y_n)} dx_{n+1} \quad (17) \]

状態の条件付き分布 \( p(x_n|Y_t) \) は非ガウス分布となるので、その分布表現に何らかの形での近似が必要である。代表的な方法は、その近似（分布の表現）方法に応じて以下のようにまとめられる。

- 解析的表現を補助的に利用するもの。ただし適用可能なモデルに制限がある。
  1. 非線形フィルタ \(^4\)
  2. ガウス和フィルタ \(^10\)
  3. 非ガウス型カルマンフィルタ \(^16\)
  4. 積分利用のカルマンフィルタ
  5. Importance Sampling を利用する方法
- 数値的表現のみ、計算機を積極的に利用する。
  1. 非ガウス型フィルタ \(^9\)
  2. モンテカルロ・フィルタ \(^13\)
  3. 動的マルコフチェインを利用する方法

非ガウス型フィルタは、非ガウス分布を階段関数近似、折れ線近似、スプライン近似などを用いて数値的に表現するもので、少なくとも低次元の問題については、極めて有力な方法である。モンテカルロ・フィルタにおいてはさらに表現の簡素化を図るため、確率分布を多数の "粒子" を用いて近似し、上記の逐次式は各粒子の時間的変化とリサンプリングによって実現する。

4. 時系列モデルの尤度とパラメータ推定

1 節で述べたように、モデルの尤度は近似的にそのモデルの予測分布の良さを評価したものである。したがって、尤度を最大にすることにより近似的には予測能力最大のパラメータを求めることができる。時系列 \( y_1, \ldots, y_N \) に対して、パラメータ \( \Theta \) を持つ時系列モデルの尤度は \( L(\Theta) = f_N(y_1, \ldots, y_N|\Theta) \) によって定義される。ここで \( f_n(y_1, \ldots, y_n|\Theta) = f_{n-1}(y_1, \ldots, y_{n-1}|\Theta)p(y_n|y_1, \ldots, y_{n-1}, \Theta) \) を繰り返し適用すると、
\[ L(\Theta) = \prod_{n=1}^{N} p(y_n|y_1, \ldots, y_{n-1}, \Theta) \]
\[ = \prod_{n=1}^{N} p(y_n|Y_{n-1}, \Theta) \quad (18) \]
と表現できる。ただし、\( Y_0 = \phi, f_1(y_1|\Theta) = p(y_1|Y_0, \Theta) \) とする。したがって、時系列モデルの対数尤度は
\[ l(\Theta) = \log L(\Theta) = \sum_{n=1}^{N} \log p(y_n|Y_{n-1}, \Theta) \quad (19) \]
によって与えられる。時系列モデルのパラメータの最尤推定 \( \hat{\Theta} \) を求めるためには、数値的最適化の方法により、この対数尤度を最大化するパラメータを求めればよい。これからの一連の手続きにより、時系列モデル
5. 具体的なモデリングの例

本節では経済時系列を例として永続的にモデルの改良を進める作業、つまり統計モデルを利用した情報処理の過程を示す。よくにデータ中の長期的な増減傾向（トレンド成分）を，季節効果による成分（季節成分）を除去してから推定する手続き，いわゆる季節調整法をとりあげる。本節の例では $y_n$ は 1 値量のみを取り扱うが，多変量でも全く問題ないことを付記しておく。

5.1 状態空間モデル

5.1.1 トレンドモデル

トレンド成分推定のためのもっとも簡単なモデルは，

$$y_n = t_n + w_n, \quad w_n \sim N(0, \sigma^2) \tag{21}$$

のように，$y_n$ が時刻 $n$ のトレンド成分 $t_n$ と観測ノイズ $w_n$ の単純和で表現できると仮定するものである。

状態空間モデルの構成のために $t_n$ に対して，次のようなく確率差分方程式 $\Delta^k t_n = v_n, \quad v_n \sim N(0, \tau^2)$ でシステムモデルを定める。$k = 1$ の場合には $t_n = t_{n-1} + v_n$ となり $t_n$ は局所的には一定で $t_{n-1}$ を表現したランダムウォークモデルとなる。また $k = 2$ の場合には $t_n = 2t_{n-1} - t_{n-2} + v_n$ となり $t_n$ は局所的には直線的な変化を表現する。ノイズの分散 $\tau^2$ が小さな場合には，トレンドモデルの解は局所的には多項式に似た非常に滑らかな関数となる。しかし，多項式と異なり大規模には非常に柔軟な関数を表現できる。

図 4 は米国センスサス局から公表された休日における高の月次データ（BFRNWS データ）の対数値と（21）のモデル（ただし $k = 2$）によって推定したトレンドとノイズ成分を示す。AIC=3147 であった。非常に滑らかなトレンドが推定されているが，ノイズには周期的な変動が残っている。このモデルは，データ中の季節効果を全く無視しており，合理的なトレンド推定や予測は期待できない。

5.1.2 季節調整モデル

やはり季節調整のためには，季節成分と呼ばれる，時系列の変動のうち毎年繰り返して現れる変動パターン $s_n$ を，明示的に観測モデルにとりこも必要がある。時系列 $y_n$ の季節調整のためのモデルは

$$y_n = t_n + s_n + w_n, \quad w_n \sim N(0, \sigma^2) \tag{22}$$

により与えられる。周期 $p$ の季節変動は $s_n \approx s_{n-p}$ をみたし，季節成分モデル

$$s_n + \cdots + s_{n-p+1} = v_{n2}, \quad v_{n2} \sim N(0, \tau_{r}^2) \tag{23}$$

によって表現することができる。また，ここでは簡単に，$t_n$ は $\Delta^2 t_n = v_n, v_n \sim N(0, \tau^2)$ に従うものとする。

この季節調整モデルは状態空間モデルにより表現できる。例えば四半期データ（一年に 4 回データが得られる）の場合は以下のようにする。

$$x_n = [t_n, t_{n-1}, s_n, s_{n-1}, s_{n-2}]^T,$$

$$F = \begin{bmatrix} 2 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \end{bmatrix}, \quad G = \begin{bmatrix} 1 \ 0 \\ 0 \ 0 \end{bmatrix},$$

$$H = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \end{bmatrix}. \tag{24}$$

図 5 は BFRNWS データの季節調整の結果を示す。AIC=2859 で，図 4 の場合と比較して AIC の値が 300 近く改善されている。每年ほぼ同じパターンを繰り返す季節成分が抽出されているが，それに伴って何とか所がトレンドが下降しているところが検出されている。図 4 の場合はトレンドのまわりの変動はすべてノイズとして処理された。季節調整モデルの導入により，その
5.1.3 より複雑な季節調整モデル

季節調整モデルにおいてさらに以下のような拡張したモデルを考えることができる。

\[ y_n = t_n + s_n + p_n + t_{dn} + w_n \]  

ただし, \( p_n \) は定常 AR (自己回帰) 成分, \( t_{dn} \) は曜日効果項である。定常 AR 成分は低次の AR モデル \( p_n = \sum_{j=1}^{m} a_j p_{n-j} + v_n, v_n \sim N(0, \sigma^2) \) に従い, トレンドのようないろいろな傾向ではなく, 短期的な変動成分を表す。また, 曜日効果項は, それぞれの観測期間中に含まれる各曜日の数の違いによる影響を除去するため導入された項で, \( y_n \) が月別データの場合, \( t_{dn} = \sum_{i=1}^{7} \beta_i d_{in} \) と与えられる。ただし, \( d_{in} \) は \( n \) 番目の月に含まれる \( i \) 番目の曜日の数である。

図 6 は図 4, 5 のデータに対して, 曜日効果項を入れたモデルで解析した結果である。AIC の値は 7777 と, 図 5 の場合に比較してさらに 80 位減少している。曜日効果項自体は非常に小さいが季節成分 + 曜日効果項はデータの変動をよく表している。曜日効果項の導入が予測において有効であることを示している。

このようなモデル改良に関するさまざまなアイデアは状態空間モデルを利用すると容易に実現でき, 情報量

規準をとおして永続的なモデルの発展が可能となる。

5.2 一般化状態空間モデル

実際の季節調整を行う現場の立場からみて得られる積極的な成果が得られるまでモデルを発展させるためには, 一般化状態空間モデルの枠組みが必要となる。

5.2.1 非ガウス型モデル

状態空間モデルのシステムノイズの分布としてコーシー分布, ピアソン型分布, 混合ガウス分布などの場の重い非ガウス型分布を採用することにより急激な構造変化をモデル化することができる。例えば, 上記のトレンドモデルにおいてはトレンドのジャンプ状変化や急激な変曲点を, 季節調整モデルでは突然の季節変動パターンの変化を, 特殊な処理なしに自動的に検出することができる。

また, 現実のデータはその生成機構あるいはなんらかの変換の結果, 非対称な分布をしているものが多い。例えば, 觀測値に異常値が含まれる場合には, 觀測ノイズの分布 \( r(u) \) に求まる重い非ガウス分布を用いることによりロバストな推定を自動的に実現することができる。

5.3 非線形モデル

今まで, (22) にあるような和型の観測モデルを考
感でてきたが、積分あるいは混合型の季節調整
\[ y_n = f_n \times s_n \times w_n, \]
\[ y_n = f_n \times s_n + w_n \]
も必要ならば取り扱う。また \( y_n = e^{\lambda n + s_n} \) と
すると、分散季節変動がある時系列のモデルが得ら
れる。

5.4 非定常離散分布モデル

一般化状態空間モデルを用いると二項系列やポアソ
ン系列などの \textit{count data} も全く同じように取り扱
ることができる \(^7\)。観測モデルとして

\[ P(y_n | \lambda_n, m_n) = \left( \begin{array}{c} m_n \\ y_n \end{array} \right) \lambda_n^{y_n} (1 - \lambda_n)^{m_n - y_n} \]

[ポアソン分布]

\[ P(y_n | \lambda_n) = \frac{e^{-\lambda_n} \lambda_n^{y_n}}{y_n!} \]

などを想定し、平均値関数 \( \lambda_n \) の変化をシステムモ
デルで表現すればよい。また、一般化状態空間モデル
では非線形変換は簡単により処理できるので、対数変換
\( q_n = \log \lambda_n \)、ロジット変換 \( \alpha_n = \log \left( \frac{\lambda_n}{1 - \lambda_n} \right) \) など
を用いて \( \lambda_n \) の値に自由に制限を加えることができる。

さらに、変換後の \( q_n \) や \( \alpha_n \) をトレンドと季節
成分に分解することにより \textit{count data} の季節調整も
簡単に行うことができる \(^7\)。図 7 は平均値がトレンド
と季節変動を持つポアソン過程から人工的に生成され
たデータの解析結果である。上から順に、生成された
データ、推定されたトレンドと季節成分を示す。した
がってこの方法は、この例の場合よりもさらに平均が
小さく \( \theta \) が観測されるような場合にも問題なく適用で
きる。

6. より自由なモデリングをめざして：
 自己組織型の状態空間モデル

非ガウス型状態空間モデルは大きな計算量を伴うの
で数値的最適化の方法により最尤推定値 \( \hat{\theta} \) を求めること
とは困難な場合がある。また、モンテカルロ・フィルタ
を用いた場合には得られる尤度もモンテカルロ近似で
誤差を伴っており、正確な最尤推定値を求めることは
できない \(^6\)。このような場合、未知のパラメータを
状態の一部とみなして状態とパラメータを同時に推定
することによりこの問題を解決することができる \(^13\)。
すなわち、状態ベクトル \( x_n \) の下にパラメータベクト
ル \( \theta \) をあらかじめ、新しい状態空間ベクトル \( z_n \) を

\[ z_n = (x_n, \theta)^T \]

と定義する。このとき、新しい状態ベクトル \( z_n \) に関
する状態空間モデルは

\[ z_n = F_n(z_{n-1}, u_n), \quad y_n = H_n(z_n, w_n) \]

となる。ただし、非線形関数 \( F_n, H_n \) は \( F_n(z, v) = \]
\( [f_n(x, v|\theta)^T, \theta]^T, \quad H_n(z, w) = h_n(x, w|\theta) \) と定義
される。この状態空間モデルを用いると、本来の状態ベクト
ル \( x_n \) の事後分布だけでなくパラメータ \( \theta \) の事後分布
\( p(\theta | Y_n) \) も同時に求めることができる。この方法では、
最尤法と異なり一回の平滑化だけで最終的な結果が得
られる。

このように未知のパラメータを状態ベクトルの一部
とみなして同時推定を行うことは従来からも試みられ
ていたが、拡張カルマンフィルタなどの近似フィルタ
を用いているため良い結果が得られないことが多いと
考えられてきた。しかしながら、非ガウス型平滑化を
用いて事後分布を精確に計算することにより、良い結
果が得られることがわかった。また、状態ベクトルの
次元が高い場合には非ガウス型平滑化のアルゴリズム
は適用できないが、モンテカルロ・フィルタの利用に
よって、実用的な近似値を求めることもできてきた。

この方法では、状態ベクトルを \( z_n = (x_n, \theta_n)^T \) とす
ることにより、パラメータ \( \theta \) が時間変化する場合にも
対応することができる。

参考文献

1) Akaike, H. (1973), Proc. 2nd International Sym-
posium on Information Theory (B.N. Petrov and
F. Csaki eds.) Akademiai Kiado, Budapest, 267-

数理科学 NO. 423, SEPTEMBER 1998
281.
8) 片山 義 (1983) 応用カルマンフィルタ，朝倉書店.

（きたがわ・げんしろう，統計数理研究所）
（ひぐち・ともゆき，統計数理研究所）

好評発売中!!

新数理ライブラリ（数学編）

線形代数の基礎と応用 荒木不二洋 著 A5判 本体1550円
微分方程式の基礎と応用 佐藤 光 著 A5判 本体1600円
ベクトル解析の基礎と応用 渡辺 正 著 A5判 本体2136円
偏微分方程式の基礎と応用 矢吹治一 著 A5判 本体1900円

表示価格は税抜きです

[サイエンス社]