# 非線形・非正規状態空間モデルの推定について

谷崎, 久志

(Citation)
国民経済雑誌,193(4):37-52

(Issue Date)
2006-04

(Resource Type)
departmental bulletin paper

(Version)
Version of Record

(JaLCDOI)
https://doi.org/10.24546/00056064

(URL)
https://hdl.handle.net/20.500.14094/00056064

# 非線形・非正規状態空間モデルの推定について*

谷　崎　久　志

本稿では非線形・非正規分布の状態空間モデルの推定を紹介する。非線形・非正規分布の状態空間モデルの推定には大きく分けて2つの種類が考えられる。一つは分布関数による逐次アルゴリズムに基づいたもの，もう一つは全部の状態変数の同時密度関数に基づいたものである。非線形・非正規状態空間モデルの場合，前者は乱数の逐次アルゴリズムを得ることになり，後者は MCMC (Gibbs サンプラー，Metropolis-Hasting アルゴリズムを含む) と呼ばれる乱数生成法を用いることになる。本稿では，特に，分布関数に基づいた逐次アルゴリズムのみに焦点を当てる。

キーワード　状態空間モデル，分布関数，逐次アルゴリズム，乱数生成法

## 1 はじめに

フィルタリング (filtering) 理論は1960年代に工学の分野で発展してきた。経済学に応用されはじめたのは1970年代に入ってからである。可変パラメータ・モデル，自己回帰移動平均モデル，季節調整モデル，経済変数の予測問題，確率的ボラティリティ (Stochastic Volatility, SV) 変動の推定等，観測されない変数を推定するのに状態空間モデルは有効である。経済学に関連したものは渡部 (2000) がある。最近は Markov Chain Monte Carlo (MCMC) という乱数生成の手法がが利用されるようになって以来多くの実証研究ができるようになってきた。谷崎 (2006) では線形・正規性のもとでの状態空間モデルの解説を行ったのに対して，本稿では非線形・非正規性のもとでの状態空間モデルの推定とその問題点を紹介することを目的とする。非線形・非正規分布に基づく状態空間モデルの推定には大きく分けて2つの種類が考えられる。一つは分布関数による逐次アルゴリズムに基づいたもの，もう一つは全部の状態変数の同時密度関数に基づいたものである。非線形・非正規状態空間モデルの場合，前者は乱数の逐次アルゴリズムを得ることになり，後者は MCMC (Gibbs サンプラー，Metropolis-Hasting アルゴリズムを含む) と呼ばれる乱数生成法を用いることになる。本稿では，特に，分布関数に基づいた逐次アルゴリズムのみに焦点を当てる。

## 2　状態空間モデルの定義

非線形・非正規分布の状態空間モデルでは一般的に次のように書かれる。

（観測方程式）　$y_t=h_t(\alpha_t,\varepsilon_t)$　　(1)

（遷移方程式）　$\alpha_t=q_t(\alpha_{t-1},\eta_t)$　　$t=1,2,\cdots,T$　　(2)

$y_t$は観測可能な変数，$\varepsilon_t$と$\eta_t$は撹乱項とする。$T$は標本数を表す。$\alpha_t$は状態変数と呼ばれ推定されるべき変数である。(1)式は観測方程式と呼ばれ，(2)式は遷移方程式と呼ばれる。外生変数は関数$h_t(\cdot,\cdot)$や$q_t(\cdot,\cdot)$の中に既に含まれているものとする。以下の(3)式，(4)式等の$\alpha_t$の関数に関する期待値を得ることが目的である。このモデルでは観測される変数$y_t$を用いて，観測されない変数$\alpha_t$を推定しようというものである。

$h_t(\cdot,\cdot)$，$q_t(\cdot,\cdot)$，$\varepsilon_t$，$\eta_t$は未知パラメータ（例えば，$\theta$）に依存してもよい。この場合には，未知パラメータ$\theta$は状態変数$\alpha_t$と共に推定されなければならない。$\alpha_t$の推定問題については3節に述べられ，$\theta$については5節で触れる。観測方程式に含まれる撹乱項$\varepsilon_t$と遷移方程式の撹乱項$\eta_t$は互いに無相関として，本稿では議論を進める。しかし，この仮定を緩めることも可能である。

## 3　状態変数の推定問題

2節で状態空間モデルを定義した。本節では状態変数の推定問題を考える。まず，$\mathrm{E}(\cdot|\cdot)$，$\mathrm{Var}(\cdot|\cdot)$をそれぞれ数学的条件付き期待値，条件付き分散共分散行列とする。$Y_s$を$s$期に利用可能な情報集合と定義する。すなわち，$Y_s=\{y_s, y_{s-1}, \cdots\cdots, y_1\}$である。(1)式や(2)式の中に含まれるすべての外生変数もまた情報集合$Y_s$の中に含まれていることに注意せよ。$a_{r|s}$，$\Sigma_{r|s}$を次のように定義する。

$$a_{r|s}=\mathrm{E}(\alpha_r|Y_s)=\int \alpha_r f(\alpha_r|Y_s)\,\mathrm{d}\alpha_r \qquad (3)$$

$$\Sigma_{r|s}=\mathrm{Var}(\alpha_r|Y_s)=\int (\alpha_r-a_{r|s})(\alpha_r-a_{r|s})' f(\alpha_r|Y_s)\mathrm{d}\alpha_r \qquad (4)$$

ただし，$f(\cdot|\cdot)$を条件付き分布関数とする。状態ベクトル（state-vector）の推定問題として，情報量の多さによって次の3種類を考えることができる。

$r>s$のとき，予測（プレディクション，prediction）

$r=s$のとき，濾波（フィルタリング，filtering）

$r<s$のとき，平滑（スムージング，smoothing）

プレディクション，フィルタリング，スムージングのアルゴリズム（algorithm）はそれぞれ3.1節，3.2節，3.3節で取り扱われる。分布関数に基づいてそれぞれアルゴリズムが導出さ

れる。詳しくは，Tanizaki (1996, 2003, 2004a), Kitagawa and Gersch (1996), 片山 (2000) 等を参照せよ。

### 3.1 プレディクション（予測推定）

本節では予測（プレディクション）問題を考える。そこでは，(3)，(4)の添字記号 $r$, $s$ をそれぞれ $t+L$, $t$ に置き換える。分布関数に基づいたプレディクション・アルゴリズムは以下のように表される。

$$f(\alpha_{t+L}|Y_t)=\int f_\alpha(\alpha_{t+L}|\alpha_{t+L-1})f(\alpha_{t+L-1}|Y_t)\,d\alpha_{t+L-1} \tag{5}$$

$L=1, 2, \cdots,$ に関する逐次アルゴリズムとなっている。上に記したプレディクション・アルゴリズムでは $f(\alpha_t|Y_t)$ を必要とする。$f(\alpha_t|Y_t)$ はフィルタリングの密度関数であり，3.2節で解説される。本節では，今のところ $f(\alpha_t|Y_t)$ を既知の分布関数と考えることにする。分布関数 $f_\alpha(\alpha_{t+L}|\alpha_{t+L-1})$, $L=1, 2, \cdots,$ は(2)式で表される遷移方程式とその中の撹乱項 $\eta_{t+L}$ の分布関数をもとにして，$\eta_{t+L}$ から $\alpha_{t+L}$ への変数変換によって計算される。それゆえに，$f_\alpha(\alpha_{t+L}|\alpha_{t+L-1})$, $L=1, 2, \cdots,$ の関数形もまた既知である。したがって，$f(\alpha_t|Y_t)$ が与えられると，$f_\alpha(\alpha_{t+1}|\alpha_t)$ を掛け合わせて，$\alpha_t$ について積分すると，$f(\alpha_{t+1}|Y_t)$ が得られる。同様に，$f(\alpha_{t+1}|Y_t)$ と $f_\alpha(\alpha_{t+2}|\alpha_{t+1})$ から $f(\alpha_{t+2}|Y_t)$ を得ることができる。このように，$f(\alpha_{t+L}|Y_t)$, $L=1, 2, \cdots,$ が逐次的 (recursive) に計算される。分布関数が得られると(3)，(4)のように期待値や分散が求められる。

### 3.2 フィルタリング(濾波推定)

予測問題では，現在（$t$期）の情報をもとにして将来（$L$期先）の状態変数を推定するものであるが，濾波（フィルタリング）問題では，現在（$t$期）に利用可能な情報をもとに現在（$t$期）の状態変数を推定するものである。よって，フィルタリング問題では，(3)，(4)の添字記号 $r$, $s$ をそれぞれ $t$, $t$ に置き換える。分布に基づいたフィルタリング・アルゴリズムは以下のように示される。

$$f(\alpha_t|Y_{t-1})=\int f_\alpha(\alpha_t|\alpha_{t-1})f(\alpha_{t-1}|Y_{t-1})d\alpha_{t-1} \tag{6}$$

$$f(\alpha_t|Y_t)=\frac{f_y(y_t|\alpha_t)f(\alpha_t|Y_{t-1})}{\int f_y(y_t|\alpha_t)f(\alpha_t|Y_{t-1})d\alpha_t} \tag{7}$$

$t=1, 2, \cdots, T$ に関する逐次アルゴリズムになっている。(6)式は予測方程式と呼ばれる。それは，(5)式において，$L=1$，かつ，$t$ を $t-1$ に置き換えたものに一致する。分布関数 $f_y(y_t|\alpha_t)$ は(1)式で表される観測方程式とその中の撹乱項 $\varepsilon_t$ の分布関数をもとにして，$\varepsilon_t$ か

ら$y_t$への変数変換によって求められる。(7)式は更新方程式と呼ばれる。それは，過去の情報$Y_{t-1}$と$t$期に得られたデータ$y_t$とを結び付ける役割を果たす。

(6)式と(7)式で与えられるアルゴリズムでは，まず初期分布$f(\alpha_0|Y_0)$を既知とすると，(2)式の遷移方程式を通して(6)式から$f(\alpha_1|Y_0)$を得ることができる。次に，(1)式の観測方程式から得られる分布関数$f_y(y_1|\alpha_1)$と(6)式から得られた$f(\alpha_1|Y_0)$とを結び付けて$f(\alpha_1|Y_1)$が導かれる。さらに，(6)式によって$f(\alpha_1|Y_1)$から$f(\alpha_2|Y_1)$，(7)式によって$f(\alpha_2|Y_1)$から$f(\alpha_2|Y_2)$がそれぞれ得られる。このように，初期分布$f(\alpha_0|Y_0)$が与えられると，それ以降のすべての分布$f(\alpha_t|Y_{t-1})$，$f(\alpha_t|Y_t)$，$t=1, 2, \cdots, T$が逐次的に計算される。

初期分布に関連して$\alpha_0$が確率変数か非確率変数かで$t=1$の場合の予測方程式(6)は異なることに注意せよ。

$$f(\alpha_1|Y_0)=\begin{cases} f_\alpha(\alpha_1|\alpha_0) & \alpha_0\text{が非確率変数のとき} \\ \int f_\alpha(\alpha_1|\alpha_0)f(\alpha_0)\mathrm{d}\alpha_0 & \alpha_0\text{が確率変数のとき} \end{cases}$$

ただし，$f(\alpha_0)$は$\alpha_0$が確率変数のときの分布関数である。初期値に近いほどフィルタリング推定値は初期値の値に影響され推定値の動きは不安定である。初期値に近いほど状態変数を推定するための情報量は少ないからである。

### 3.3　スムージング（平滑推定）

状態変数の推定問題の中のスムージングについて述べる。$L$をある固定された値，$T$を標本数としたとき，スムージングには次の3つの種類がある。

・固定点(fixed-point) スムージング：　$a_{L|t}$　$\Sigma_{L|t}$　$t=L+1, L+2, \cdots$

・固定ラグ(fixed-lag) スムージング：　$a_{t|t+L}$　$\Sigma_{t|t+L}$　$t=1, 2, \cdots, T-L$

・固定区間(fixed-interval) スムージング：　$a_{t|T}$　$\Sigma_{t|T}$　$t=1, 2, \cdots, T$

の3種類である。初期状態をデータから推定するのに固定点スムージングは有効であり，一定時間の遅れを伴う場合には固定ラグ・スムージングが有効とされ，過去に起こったことをデータから分析する場合には固定区間スムージングが適している。経済学においては今までに利用可能なデータを用いて過去の経済状況を分析する場合が多いので，本稿ではスムージングの中でも経済学で最も有用な固定区間スムージングのみを考える。すなわち，(3)，(4)の添字記号$r$, $s$をそれぞれ$t$, $T$に置き換えることになる。分布に基づいた固定区間スムージング・アルゴリズムは以下のように表される。

$$f(\alpha_t|Y_T)=f(\alpha_t|Y_t)\int\frac{f(\alpha_{t+1}|Y_T)f_\alpha(\alpha_{t+1}|\alpha_t)}{f(\alpha_{t+1}|Y_t)}\mathrm{d}\alpha_{t+1} \qquad (8)$$

これは$t=T-1, T-2, \cdots, 1$に関する逆向きの逐次アルゴリズム（Backward Recursive

Algorithm) である。フィルタリング・アルゴリズムの(6)式と(7)式から得られる分布関数 $f(\alpha_t|Y_t)$, $f(\alpha_{t+1}|Y_t)$ と遷移方程式から得られる $f_a(\alpha_{t+1}|\alpha_t)$ をもとにして，スムージングの分布関数は(8)式によって $f(\alpha_{t+1}|Y_t)$ から $f(\alpha_t|Y_t)$ へと逐次的に求められる。固定区間スムージングはすべての $t$ について同じ情報量 $Y_T$ で $t$ 期の状態変数 $\alpha_t$ を推定する。そのため，フィルタリングは初期値に近い状態変数の推定値はばらつきが大きく不安定であるが，スムージングは全期間について安定的であるといえる。

## 4　非線形・非正規状態空間モデルの推定問題

(3)式，(4)式等の期待値を得ることが目的である。非線形・非正規状態空間モデルの場合は $\alpha_t$ の乱数を発生させて期待値を評価することを考える。まず，乱数生成の方法について解説を行う。

### 4.1　乱数生成法

連続型確率変数の密度関数 $f(x)$ から乱数を生成したいが，$f(x)$ から乱数を生成することが難しい場合を想定する。$f(x)$ はターゲット密度関数 (target density) と呼ばれる。一方，乱数の生成が簡単にできる別の密度関数 $f_*(x)$ を選ぶ。$f_*(x)$ はサンプリング密度関数 (sampling density) と呼ばれる。このとき，$f_*(x)$ から生成される乱数を利用して，$f(x)$ からの乱数を発生させることができる。$x_i$ を $f(x)$ から生成された $i$ 番目の乱数とする。$\omega(x)$ をターゲット密度関数とサンプリング密度関数の比に比例するものとする。すなわち，$\omega(x) \propto f(x)/f_*(x)$ と表されるものとする。このとき，ターゲット密度関数は $f(x) \propto \omega(x) f_*(x)$ として書き直される。$\omega(x)$ は受容確率 (Acceptance Probability) に密接に関連している。受容確率の形によって，棄却法 (Rejection Sampling，以下 RS と略)，重点的リサンプリング法 (Importance Resampling，以下 IR と略)，Metropolis-Hasting アルゴリズム (以下 MH と略) の3つの乱数生成法が考えられる。Liu (1996)，Tanizaki (2004a) では，3つのサンプリング方法の紹介と比較が行われている。$f(x)$ から乱数を生成できるには，$\omega(x)$ の関数形が既知であることと $f_*(x)$ から乱数の生成が簡単であることが必要となる。以下に3つの乱数生成法を記しておく (ただし，証明は省略する)。

**棄却法 (RS)**：$x^*$ をサンプリング密度関数 $f_*(x)$ から生成された $x$ の乱数とする。任意の $x$ について $\omega(x)$ は有限であるとする。このとき，受容確率は $w(x)=\omega(x)/\sup_z \omega(z)$ によって表される。次のようにして，ターゲット密度関数 $f(x)$ から $x$ の乱数を $N$ 個生成することができる。

(i)　サンプリング密度関数 $f_*(x)$ から $x$ の乱数 $x^*$ を生成し，受容確率 $w(x^*)$ を計算する。

(ii)　$w(x^*)$ の確率で $x_i=x^*$ とし，$1-w(x^*)$ の確率で(i)に戻る。すなわち，$u$ を区間(0,

1) での一様乱数としたとき，$u \le w(x^*)$ のとき $x_i = x^*$，$u > w(x^*)$ のとき（i）に戻る。

（iii） $i=1, 2, \cdots, N$について（i）と（ii）を繰り返す。

生成された$N$個の乱数$x_1, x_2, \cdots, x_N$は互いに独立となっているので，乱数の精度という観点ではRSはIRやMHよりも最もよいサンプリング方法である。しかしながら，問題はRSを適用するためには$\omega(x)$の上限を求める必要があることである。もし上限がなければ受容確率$w(x)$がほとんどの$x$でゼロになり，したがって，$x^*$は（i），（ii）で$x_i$として採択されないことになる。$N_R$を乱数の平均棄却回数としたとき，ターゲット密度関数$f(x)$から一つの乱数を得るためには，平均で$1+N_R$個の乱数を必要とする（言い換えると，採択率は平均で$1/(1+N_R)$である）。$\omega(x) \to \infty$は$N_R \to \infty$を意味する。$f(x)$から$N$個の乱数を取り出すためには，$f_*(x)$から$N(1+N_R)$個の乱数を生成する必要がある。RSについては，例えば，Boswell *et al.* (1993)，O'Hagan (1994)，Geweke (1996) 等を参照せよ。

$\omega(x)$に上限があるという条件を調べるために，$f(x)$ と $f_*(x)$ は共に正規分布でそれぞれ$N(\mu, \sigma^2)$と$N(\mu_*, \sigma_*^2)$の分布の場合を考える。このとき，$\omega(x)$が有限であるためには，$\sigma_*^2 > \sigma^2$が必要となる。すなわち，サンプリング密度関数$f_*(x)$がターゲット密度関数$f(x)$より幅広く分布しなければならないということになる。

**重点的リサンプリング法（IR）**：$x_i^*$をサンプリング密度関数$f_*(x)$から$i$番目に生成された$x$の乱数とする。受容確率を$w(x_i^*) = \omega(x_i^*) / \sum_{j=1}^{N} \omega(x_j^*)$とする。ターゲット密度関数$f(x)$から$N$個の乱数を生成するには次の通りである。

（i） サンプリング密度関数$f_*(x)$から$x_j^*, j=1, 2, \cdots, N$を生成し，すべての$j=1, 2, \cdots, N$について$w(x_j^*)$を計算する。

（ii） $w(x_j^*)$の確率で$x_i = x_j^*$とする。

（iii） $i=1, 2, \cdots, N$について（ii）を繰り返す。

ステップ（ii）では，具体的には，まず$W_j$を$j=1, 2, \cdots, N$について$W_j \equiv W_{j-1} + w(x_j^*)$として計算する（ただし，$W_0 \equiv 0$である）。そして，区間$(0, 1)$で一様乱数（$u$とする）を発生させ，$W_{j-1} \le u < W_j$のとき$x_i = x_j^*$とする。IRについては，例えば，Smith and Gelfand (1992) を参照せよ。

乱数の精度という面についてIRはRSより劣る。IRはサンプリング密度関数から$N$個の乱数を生成しておき，それぞれ対応する確率でそれらのうち一つを取り出す方法である。よって，IRではターゲット密度関数から最終的に得られた$N$個の乱数の中でいくつかは同じ数値となる。一方，RSでは生成された$N$個の乱数はすべて異なった値が得られる。ターゲット密度関数$f(x)$から$N$個の乱数を得るためには，IRはちょうど$N$個の乱数がサンプリング密度関数$f_*(x)$から生成されればよいが，RSは$N$個以上（より正確に言えば，$N(1+N_R)$個）

の乱数をサンプリング密度関数 $f_*(x)$ から発生させる必要がある。

**Metropolis-Hasting アルゴリズム (MH)**：$x^*$ を $f_*(x)$ から生成された乱数とする。受容確率を $w(x_{i-1}, x^*)=\min(\omega(x^*)/\omega(x_{i-1}), 1)$ とする。ターゲット密度関数 $f(x)$ から $N$ 個の乱数を生成するためには次のようにすればよい。

(i) $x$ の初期値を $x_{-M}$ とする。

(ii) サンプリング密度関数 $f_*(x)$ から $x^*$ を生成し，$x^*$ と $x_{i-1}$ を与えたもとで受容確率 $w(x_{i-1}, x^*)$ を計算する。

(iii) $w(x_{i-1}, x^*)$ の確率で $x_i=x^*$，それ以外は $x_i=x_{i-1}$ とする。言い換えると，$u$ を区間 $(0, 1)$ の一様乱数としたとき，$u \le w(x_{i-1}, x^*)$ のとき $x_i=x^*$，$u>w(x_{i-1}, x^*)$ のとき $x_i=x_{i-1}$ とする。

(iv) $i=-M+1, -M+2, \cdots, N$ について(ii)と(iii)を繰り返す。

以上のようなアルゴリズムで問題となるのは $f_*(x)$ と $M$ の選択の仕方である。サンプリング密度関数 $f_*(x)$ はターゲット密度関数 $f(x)$ と同じような分布を選ぶべきである。サンプリング密度関数 $f_*(x)$ の分散がターゲット密度関数 $f(x)$ の分散に比べて，大き過ぎたり小さ過ぎたりすると正しい乱数が得られない。また，$x_{i-1}$ に依存して $f_*(x)=f_*(x|x_{i-1})$ となるサンプリング密度関数を選んでもよい。MH については，例えば，Chib and Greenberg (1995), Geweke (1996), Geweke and Tanizaki (2003) を参照せよ。

MH では十分に大きな $M$ について $x_1$ がターゲット密度関数 $f(x)$ から生成された $x$ の乱数とみなされる。$x_1$ が $f(x)$ からの乱数であれば，$x_2$ もまた $f(x)$ からの乱数となっている。そのため，$x_1, x_2, \cdots, x_N$ はすべて $f(x)$ から生成された乱数となっている。このように，$N$ 個の乱数を得るためにはサンプリング密度関数 $f_*(x)$ から $M+N$ 個の乱数を生成する必要がある。ただし，明らかに $\mathrm{Cov}(x_{i-1}, x_i)>0$ である。なぜなら，アルゴリズムのステップ(ii), (iii) で，$x_i$ は $x_{i-1}$ をもとにして生成されているからである。したがって，乱数の精度という面では MH は 3 つの乱数生成方法の中で最も悪い方法であると言える。$M$ の選択については Geweke (1992) が CD (Convergence Diagnostic) 検定を提案した。$x_1, x_2, \cdots, x_N$ の乱数の最初と最後の 2，3 割程度のサンプルを取り出してそれぞれの平均に差があるかどうかの検定を行い，差がないという結果であれば $x_1, x_2, \cdots, x_N$ はすべて $f(x)$ から生成された乱数となっているとみなす。

### 4.2 フィルタリング

状態変数の推定に上述のサンプリング法を適用する。$\alpha_{i,t|s}$ を密度関数 $f(\alpha_t|Y_s)$ から生成された $\alpha_t$ の $i$ 番目の乱数とする。もし乱数 $\alpha_{1,t|s}, \alpha_{2,t|s}, \cdots, \alpha_{N,t|s}$ $(s=t, T,$ かつ $t=1,2,\cdots,T)$ が生

成されたとすると，(3)式，(4)式は $a_{t|s}\equiv\bar{a}_{t|s}\approx(1/N)\sum_{i=1}^{N}\alpha_{i,t|s}$, $\Sigma_{t|s}\approx(1/N)\sum_{i=1}^{N}(\alpha_{i,t|s}-\bar{a}_{t|s})(\alpha_{i,t|s}-\bar{a}_{t|s})'$ として求められる。今のところ，$\alpha_{i,t-1|t-1}$, $i=1, 2, \cdots, N$は既に生成されていて利用可能であると考える。このときに，$\alpha_{i,t|t}$, $i=1, 2, \cdots, N$を生成する方法を考える。初期値$\alpha_0$が確率変数かどうかで$\alpha_{i,0|0}$, $i=1, 2, \cdots, N$は$f_\alpha(\alpha_0)$から生成されることになるか，または，すべての$i$についてある一定の値をとることになる。

フィルタリング密度関数(7)式について2通りの表し方がある。一つは(7)式から$f(\alpha_t|Y_t)$は次のように表される。

$$f(\alpha_t|Y_t)\propto f_y(y_t|\alpha_t)f(\alpha_t|Y_{t-1})\propto\omega_1(\alpha_t)f(\alpha_t|Y_{t-1}) \qquad (9)$$

ただし，$\omega_1(\alpha_t)\propto f_y(y_t|\alpha_t)$とする。この場合，4.1節の$f_*(x)$と$\omega(x)$はそれぞれ$f(\alpha_t|Y_{t-1})$と$\omega_1(\alpha_t)$に対応する。$f_y(y_t|\alpha_t)$は(1)式から変数変換によって得られるので，$\omega_1(\alpha_t)$の関数形は既知である。$\alpha_{i,t-1|t-1}$, $i=1, 2, \cdots, N$を与えたもとで，$f(\alpha_t|Y_{t-1})$から生成される$\alpha_t$の乱数は，遷移方程式(2)式から簡単に得られる。したがって，$\alpha_{i,t-1|t-1}$を与えたもとで，$\alpha_{i,t|t}$の乱数生成方法は次のようになる。

(i)　$\alpha_0$を確率変数と仮定するなら$f_\alpha(\alpha_0)$から$\alpha_0$の乱数（$i$番目の乱数を$\alpha_{i,0|0}$とする）を$N$個発生させる。または，$\alpha_0$を非確率変数と仮定するなら$\alpha_0$に何らかの値を与える（すべての$i$について$\alpha_{i,0|0}=\alpha_0$とする）。

(ii)　$f(\alpha_t|Y_{t-1})$から$\alpha_t$の乱数（$\alpha_t^*$とする）を生成する。$\alpha_t^*$は$f(\alpha_t|Y_t)$から生成される$\alpha_t$の$i$番目の乱数（$\alpha_{i,t|t}$とする）の候補となる。

(iii)　$\alpha_t^*$に基づいて，$f(\alpha_t|Y_t)$から生成された$\alpha_t$の$i$番目の乱数（$\alpha_{i,t|t}$とする）が得られる。ここで乱数生成の方法は，上述のRS，IR，MH等のいづれかを用いればよい。

(iv)　$i=1, 2, \cdots, N$について，(ii)と(iii)を繰り返す。

(v)　$t=1, 2, \cdots, T$について，(ii)～(iv)を繰り返す。

ステップ(ii)で，$\alpha_t^*$はサンプリング密度関数$f(\alpha_t|Y_{t-1})$から生成される乱数であり，この場合は$\alpha_{i,t}^*=\alpha_{i,t|t-1}$となる。ただし，$\alpha_{i,t|t-1}$は，$\eta_t$の乱数（$\eta_{i,t}$とする）を生成して，遷移方程式$\alpha_{i,t|t-1}=q_t(\alpha_{i,t-1|t-1}, \eta_{i,t})$から得られる。Gordon *et al.* (1993)，Kitagawa (1996, 1998)，Kitagawa and Gersch (1996)は，(9)式にIRを適用してフィルタリングを提案した。すなわち，ステップ(iii)の乱数生成でIRを用いた。Tanizaki (1996, 1999)，Hürzeler and Künsch (1998)は(9)式にRSを適用した。

$t$時点に構造変化や異常値があった場合，フィルタリング密度関数$f(\alpha_t|Y_t)$はプレディクション密度関数$f(\alpha_t|Y_{t-1})$から乖離する。この状況下では，(9)式に基づいたIRやMHによる$f(\alpha_t|Y_t)$から$\alpha_t$の乱数は非現実的なものとなる。また，RSについては，受容確率が非常に小さくなるため，計算時間が極端に長くなる。さらに，$\eta_t$から乱数が生成するのが難しい場合，$f(\alpha_t|Y_{t-1})$から$\alpha_t$の乱数を生成するのも難しくなる。よって，フィルタリング密度

関数の二つ目の表し方として，より良い乱数を発生させるために$\alpha_t$のサンプリング密度関数$f_*(\alpha_t|\alpha_{t-1})$を明示的に導入する。プレディクション密度関数(6)式をフィルタリング密度関数(9)式に代入して，$f(\alpha_t|Y_t)\propto f_y(y_t|\alpha_t)f(\alpha_t|Y_{t-1})=\int f_y(y_t|\alpha_t)f_\alpha(\alpha_t|\alpha_{t-1})f(\alpha_{t-1}|Y_{t-1})d\alpha_{t-1}\propto\int f(\alpha_t,\alpha_{t-1}|Y_t)d\alpha_{t-1}$のように変形される。$\alpha_{t-1}$に関する積分を取り除くと，$Y_t$を与えたもとで$\alpha_t$と$\alpha_{t-1}$の条件付密度関数$f(\alpha_t,\alpha_{t-1}|Y_t)\propto f_y(y_t|\alpha_t)f_\alpha(\alpha_t|\alpha_{t-1})f(\alpha_{t-1}|Y_{t-1})$は，$\omega_2(\alpha_t,\alpha_{t-1})\propto f_y(y_t|\alpha_t)f_\alpha(\alpha_t|\alpha_{t-1})/f_*(\alpha_t|\alpha_{t-1})$を利用して，以下のように書き直される。

$$f(\alpha_t,\alpha_{t-1}|Y_t)\propto\omega_2(\alpha_t,\alpha_{t-1})f_*(\alpha_t|\alpha_{t-1})f(\alpha_{t-1}|Y_{t-1}) \tag{10}$$

(10)式では，$f_*(\alpha_t|\alpha_{t-1})f(\alpha_{t-1}|Y_{t-1})$がサンプリング密度関数となる。$f(\alpha_{t-1}|Y_{t-1})$から生成された$\alpha_{t-1}$の$N$個の乱数$\alpha_{i,t-1|t-1}$，$i=1,2,\cdots,N$が利用可能とする。$f(\alpha_{t-1}|Y_{t-1})$から$\alpha_{t-1}$の乱数を生成することは，$\alpha_{1,t-1|t-1}$，$\alpha_{2,t-1|t-1}$，$\cdots$，$\alpha_{N,t-1|t-1}$から1つを同じ確率（すなわち，$1/N$）で選ぶことである。$\alpha_{i,t-1|t-1}$を与えて，$f_*(\alpha_t|\alpha_{i,t-1|t-1})$から$\alpha_t$の$i$番目の乱数$\alpha^*_{i,t}$を生成する。このように，$\omega_2(\alpha_t,\alpha_{t-1})$の関数形は既知で，しかも，$(\alpha_t,\alpha_{t-1})$の乱数は$f_*(\alpha_t|\alpha_{t-1})f(\alpha_{t-1}|Y_{t-1})$から生成できるので，$f(\alpha_t,\alpha_{t-1}|Y_t)$からの$(\alpha_t,\alpha_{t-1})$の乱数はRS，IR，MHを通して生成可能である。$f(\alpha_t,\alpha_{t-1}|Y_t)$からの$(\alpha_t,\alpha_{t-1})$の$i$組目の乱数を$(\alpha_{i,t|t},\alpha_{i,t-1|t})$とする。$\alpha_{i,t|t}$と$\alpha_{i,t-1|t}$の2つの乱数が同時に得られるが，フィルタリングに関する乱数は$\alpha_{i,t|t}$である。$\alpha_{i,t-1|t}$は固定ラグスムージング(3.3節の固定ラグスムージングについて，$L=1$，かつ，$t$を$t-1$で置き換えたものに対応する）である。$f(\alpha_t,\alpha_{t-1}|Y_t)$から生成される$\alpha_t$の乱数は，$f(\alpha_t|Y_t)$から生成される$\alpha_t$の乱数と同じ性質を持つ。$f_*(\alpha_t|\alpha_{t-1})$を，$\alpha_{t-1}$に依存しないで，$f_*(\alpha_t)$としてもよい。RSを適用する場合，$\alpha_t$と$\alpha_{t-1}$の任意の値に対して$\omega_2(\alpha_t,\alpha_{t-1})$が有限でなければならない。(10)式に基づくと，時には，RSは適用不可能な場合も起こる。よって，(10)式を選んだ場合は，乱数生成の容易さという意味で，IRやMHがより良い乱数生成法と言えるであろう。

以上をまとめると，$\alpha_{i,t-1|t-1}$を与えたもとで，$\alpha_{i,t|t}$の生成方法は，(i)，(iv)，(v)は(9)式に基づいたものと同じアルゴリズムで，(ii)と(iii)を以下のものに代えればよい。

(ii)　サンプリング密度関数$f_*(\alpha_t|\alpha_{t-1})f(\alpha_{t-1}|Y_{t-1})$から$(\alpha_t,\alpha_{t-1})$の乱数$(\alpha^*_t,\alpha^*_{t-1})$を生成する。

(iii)　$(\alpha^*_t,\alpha^*_{t-1})$に基づき，RS，IRやMHを用いて，$Y_t$を与えたもとで$(\alpha_t,\alpha_{t-1})$の$i$組目の乱数$(\alpha_{i,t|t},\alpha_{i,t-1|t})$を得る。$\alpha_{i,t|t}$が$f(\alpha_t|Y_t)$生成された乱数となる。

$N$個の乱数$\alpha_{1,t-1|t-1}$，$\alpha_{2,t-1|t-1}$，$\cdots$，$\alpha_{N,t-1|t-1}$が利用可能であるとき，ステップ(ii)で，$f(\alpha_{t-1}|Y_{t-1})$から$\alpha_{t-1}$の乱数を生成させるには，$N$個の乱数から同じ確率（すなわち，$1/N$）で一つを選べばよい。すなわち，$1,2,\cdots,N$から等確率で選び出された$i$について，$\alpha^*_{t-1}=\alpha_{i,t-1|t-1}$とする。

### 4.3 固定区間スムージング

スムージングでは，線形アルゴリズムと同様に，分布関数に基づいて後ろ向きの逐次アルゴリズムとなっている。そのため，$\alpha_{i,t+1|T}$を与えたもとで，$\alpha_{i,t|T}$を生成することを考える。$T$時点において，スムージングの乱数とフィルタリング乱数は同じものであり，$\alpha_{i,T|T}$によって表される。$\alpha_{i,T|T}$を与えたもとで，$\alpha_{i,t-1|T}$が得られ，最後に初期値のスムージング乱数$\alpha_{i,0|T}$が生成される。

(8)式に基づいて，スムージング密度関数に関する3種類の表現方法(後述の(11)式，(13)式，(14)式の3種類) がある。(8)式から，$\alpha_{t+1}$の積分を取り除くと，一つ目のスムージング密度関数$f(\alpha_{t+1}, \alpha_t|Y_T)$の表し方が次のように得られる。

$$f(\alpha_{t+1}, \alpha_t|Y_T) \propto \omega_3(\alpha_{t+1}, \alpha_t) f(\alpha_t|Y_t) f(\alpha_{t+1}|Y_T) \tag{11}$$

ただし，$\omega_3(\alpha_{t+1}, \alpha_t) \propto f_a(\alpha_{t+1}, \alpha_t)/f(\alpha_{t+1}|Y_t)$である。(11)式の$\omega_3(\alpha_{t+1}, \alpha_t)$に含まれる$f(\alpha_{t+1}|Y_t)$を評価するためには，$L=1$のときのプレディクション(5)式を利用して，次のようにモンテ・カルロ積分を行えばよい。

$$f(\alpha_{t+1}|Y_t) = \int f_a(\alpha_{t+1}, \alpha_t) f(\alpha_t|Y_t)\, d\alpha_t \approx \frac{1}{N'} \sum_{j=1}^{N'} f_a(\alpha_{t+1}|\alpha_{j,t|t}) \tag{12}$$

$N'$は必ずしも$N$に等しくする必要はない。計算負担を減らすためには，$N'$を$N$より小さくすればよい。$N$個の乱数$\alpha_{1,t|t}, \alpha_{2,t|t}, \cdots, \alpha_{N,t|t}$から，ランダムに$N'$個の乱数を取り出し，(12)式の積分を計算すればよい。(12)式によると，スムージングはフィルタリングより$N'$倍の計算を要するということになる。すなわち，各$t$時点で，計算量はスムージングで$N \times N'$，フィルタリングで$N$のオーダーとなる。

(11)式では，サンプリング密度関数は$f(\alpha_{t+1}|Y_T) f(\alpha_t|Y_t)$となる。$\alpha_t$の乱数は$f(\alpha_t|Y_t)$から生成され，$\alpha_{t+1}$の乱数は$f(\alpha_{t+1}|Y_T)$から生成される。$f(\alpha_t|Y_t)$からの$\alpha_t$のサンプリングは，$\alpha_{i,t|t}$, $i=1, 2, \cdots, N$からランダムに等確率 (すなわち，$1/N$) で一つ選ぶことになる。$f(\alpha_{t+1}|Y_T)$から$\alpha_{t+1}$のサンプリングもまた，$\alpha_{i,t+1|T}$, $i=1, 2, , \cdots, N$からランダムに等確率 (すなわち，$1/N$) で一つ選べばよい。このように，$(\alpha_t, \alpha_{t+1})$の乱数の候補 ($\alpha_t^*$, $\alpha_{t+1}^*$とする) を生成することができる。$(\alpha_t^*, \alpha_{t+1}^*)$に基づき，RS, IRやMHを用いて，$Y_T$を与えたもとで $(\alpha_t, \alpha_{t+1})$ の$i$組目の乱数$(\alpha_{i,t|T}, \alpha_{i,t+1|T})$を得ることができる。この場合，$\alpha_{i,t+1|T}$は必要としない (なぜなら，$\alpha_{i,t+1|T}$は前段階で既に得られているからである)。必要なのは$\alpha_{i,t|T}$であり，$\alpha_{i,t+1|T}$は捨てられることになる。以上をまとめると，$\alpha_{i,t+1|T}$を与えたもとで，$\alpha_{i,t|T}$の生成方法は次のようになる。

(i) あらかじめ，フィルタリング密度関数から乱数$\alpha_{i,t|t}$, $i=1, 2, \cdots, N$, $t=0, 1, \cdots, T$を生成しておく。$T$時点 (最終時点) でのスムージングの乱数$\alpha_{i,T|T}$はフィルタリング乱数は同じものであり，これをスムージング乱数の出発点にする。

(ii)　サンプリング密度関数$f(\alpha_t|Y_t)f(\alpha_{t+1}|Y_T)$から$(\alpha_t, \alpha_{t+1})$の乱数$(\alpha_t^*, \alpha_{t+1}^*)$を生成する。

(iii)　$(\alpha_t^*, \alpha_{t+1}^*)$に基づき，RS, IRやMHを用いて，$Y_T$を与えたもとで$(\alpha_t, \alpha_{t+1})$の$i$組目の乱数$(\alpha_{i,t|T}, \alpha_{i,t+1|T})$を得る。$\alpha_{i,t|T}$が$f(\alpha_t|Y_T)$から生成された乱数となる。

(iv)　$i=1, 2, \cdots, N$について，(ii)と(iii)を繰り返す。

(v)　$t=T-1, T-2, \cdots, 0$について，(ii)～(iv)を繰り返す（逆向きの逐次アルゴリズム）。

ステップ(i)で$(\alpha_t^*, \alpha_{t+1}^*)$を生成するためには，それぞれ次のようにする。$\alpha_t^*$については，$\alpha_{i,t|t}$, $i=1, 2, \cdots, N$からランダムに等確率で一つを選び出し，それを$\alpha_t^*$とする。$\alpha_{t+1}^*$については，$\alpha_{i,t+1|T}$, $i=1, 2, \cdots, N$から等確率で一つを選び出し，それを$\alpha_{t+1}^*$とする。

さらに，(11)式の$f(\alpha_t|Y_t)$を(9)式に置き換えると，次のようにスムージング密度関数$f(\alpha_{t+1}, \alpha_t|Y_T)$の二つ目の表現を得る。

$$f(\alpha_{t+1}, \alpha_t|Y_T) \propto \omega_4(\alpha_{t+1}, \alpha_t) f(\alpha_t|Y_{t-1}) f(\alpha_{t+1}|Y_T) \tag{13}$$

ただし，$\omega_4(\alpha_{t+1}, \alpha_t) \propto f_y(y_t|\alpha_t) f_\alpha(\alpha_{t+1}|\alpha_t) / f(\alpha_{t+1}|Y_T) \propto \omega_1(\alpha_t)\omega_3(\alpha_{t+1}, \alpha_t)$とする。(13)式では，$f(\alpha_{t+1}|Y_T)f(\alpha_t|Y_{t-1})$がサンプリング密度関数として採用される。(11)式と(13)式との違いは$\alpha_t$のサンプリング密度関数にある。前者は$f(\alpha_t|Y_t)$に基づくが，後者は$f(\alpha_t|Y_{t-1})$を利用する。(11)式や(13)式によって，$f(\alpha_{t+1}, \alpha_t|Y_T)$から$(\alpha_{t+1}, \alpha_t)$の乱数$(\alpha_{i,t+1|T}, \alpha_{i,t|T}$とする）を生成することができる。$t=T-1, T-2, \cdots, 0$について繰り返して，$\alpha_{i,t|T}$を後ろ向きの逐次計算によって得ることができる。

以上をまとめると，$\alpha_{i,t+1|T}$を与えたもとで，$\alpha_{i,t|T}$の生成方法は，(i)，(iii)～(v)は(11)に基づいたものと同じアルゴリズムで，(ii)を以下のものに入れ替えればよい。

(ii)　サンプリング密度関数$f(\alpha_t|Y_{t-1})f(\alpha_{t+1}|Y_T)$から$(\alpha_t, \alpha_{t+1})$の乱数$(\alpha_t^*, \alpha_{t+1}^*)$を生成する。

$\alpha_t$の乱数$\alpha_{i,t|t-1}$は遷移方程式$\alpha_{i,t|t-1}=q_t(\alpha_{j,t-1|t-1}, \eta_{i,t})$から生成できる。ただし，$\eta_{i,t}$は$\eta_t$の$i$番目の乱数であり，$j$は$1, 2, \cdots, N$から無作為に一つ選ばれる。また，$f(\alpha_{t+1}|Y_T)$からの$\alpha_{t+1}$の乱数は，$\alpha_{i,t+1|T}$, $i=1, 2, \cdots, N$, から無作為に一つ選べばよい。

$t$が$T$に近づくにつれて，$Y_t$は$Y_T$と同じになる（すなわち，フィルタリングは固定区間スムージングに近づく）。そのため，$t$が$T$に近いときにスムージング密度関数からの乱数を得るためには，(11)式では$f(\alpha_t|Y_t)$，(13)式では$f(\alpha_t|Y_{t-1})$を$\alpha_t$のサンプリング密度関数として選べばよい。しかし，$t$が初期値に近い場合，$f(\alpha_t|Y_t)$や$f(\alpha_t|Y_{t-1})$は$f(\alpha_t|Y_T)$から乖離する場合が起こる。したがって，この場合の問題を改善するためには，スムージング密度関数の三つ目の表現として，$\alpha_t$に関するサンプリング密度関数$f_*(\alpha_t|\alpha_{t-1}, \alpha_{t+1})$を明示的に取り入れる必要がある。(13)式に(6)式を代入して，$\alpha_{t-1}$に関する積分を取り除くと，$Y_T$を与

えたもとでの $(\alpha_{t+1}, \alpha_t, \alpha_{t-1})$ の条件付密度関数，すなわち，スムージング密度関数 $f(\alpha_{t+1}, \alpha_t, \alpha_{t-1}|Y_T)$ が次のように得られる。

$$f(\alpha_{t+1}, \alpha_t, \alpha_{t-1}|Y_T) \propto \omega_5(\alpha_{t+1}, \alpha_t, \alpha_{t-1}) f(\alpha_{t-1}|Y_{t-1}) f_*(\alpha_t|\alpha_{t-1}, \alpha_{t+1}) f(\alpha_{t+1}|Y_T) \qquad (14)$$

ただし，$\omega_5(\alpha_{t+1}, \alpha_t, \alpha_{t-1}) \propto \omega_4(\alpha_{t+1}, \alpha_t) f_a(\alpha_t|\alpha_{t-1}) / f_*(\alpha_t|\alpha_{t-1}, \alpha_{t+1})$ である。(14)式では，サンプリング密度関数は $f(\alpha_{t+1}|Y_T) f_*(\alpha_t|\alpha_{t-1}, \alpha_{t+1}) f(\alpha_{t-1}|Y_{t-1})$ となる。$f(\alpha_{t+1}|Y_T)$ と $f(\alpha_{t-1}|Y_{t-1})$ から $\alpha_{t+1}$ と $\alpha_{t-1}$ の乱数が独立に生成されて後，$\alpha_t$ の乱数が別のサンプリング密度関数 $f_*(\alpha_t|\alpha_{t-1}, \alpha_{t+1})$ から生成される。

$\alpha_{i,t+1|T}$ を与えたもとで，$\alpha_{i,t|T}$ の生成方法は，(i)，(iii)～(v)は(11)に基づいたものと同じアルゴリズムで，(ii)を以下のものに入れ替えればよい。

(ii) サンプリング密度関数 $f(\alpha_{t+1}|Y_T) f_*(\alpha_t|\alpha_{t-1}, \alpha_{t+1}) f(\alpha_{t-1}|Y_{t-1})$ から $(\alpha_{t+1}, \alpha_t, \alpha_{t-1})$ の乱数 $(\alpha^*_{t+1}, \alpha^*_t, \alpha^*_{t-1})$ を生成する。

このようにして，$(\alpha_{i,t+1|T}, \alpha_{i,t|T}, \alpha_{i,t-1|T})$ が(14)式から得られる。ここで必要とする乱数は $\alpha_{i,t|T}$ であり，$\alpha_{i,t+1|T}$ と $\alpha_{i,t-1|T}$ は捨てられる。$\alpha_{i,t+1|T}$ は前段階で既に得られており，$\alpha_{i,t-1|T}$ は次の段階でもう一度得られる乱数である。また，$f_*(\alpha_t|\alpha_{t-1}, \alpha_{t+1}) = f_*(\alpha_t)$ もまた $\alpha_t$ のサンプリング密度関数の一つの候補となりうる。

(9)式に基づく MH フィルタ，(10)式に基づく IR, RS, MH フィルタ，(11)式，(13)式，(14)式をもとにした IR, RS, MH スムーザ (smooter) は Tanizaki (2001, 2004a) によって議論された。

乱数の精度の面から，RS が MH より優れている。しかし，RS は適用範囲は MH より狭い。RS の性質は(i)$\omega(\cdot)$の上限が存在するときにのみ RS を適用できる，(ii)乱数の精度はサンプリング密度関数の選択には依存しない（計算時間はサンプリング密度関数の選択に大いに依存する)。$\omega(\cdot)$の上限が存在しなければ，RS を用いることはできない。例え，上限が存在していたとしても，特殊な場合を除いて，一般的には，$\omega_2$, $\omega_3$, $\omega_4$, $\omega_5$ の上限を求めることは難しい。それでも，$\alpha_t$ に関して $f_y(y_t|\alpha_t)$ の上限が存在する場合が多いので，$\omega_1$ の上限は存在する場合は多い。よって，RS を用いる場合には，(10)式，(11)式，(13)式，(14)式よりむしろ，フィルタリングの(9)式に限って用いる方がよい。

フィルタリングとスムージングとで異なる乱数生成方法を採用することも可能である。例えば，(9)式に基づいて RS をフィルタリングでは使い，(11)式をもとにして IR をスムージングでは採用するといったことも可能である。さらに，フィルタリングでは異なる時点で(9)式，(10)式を併用したり，スムージングについても(11)式，(13)式，(14)式を組み合わせることもできる。一つの例を示すと，ある時点 $t'$ で構造変化，または，異常値があったとしよう。この場合，プレディクション密度関数 $f(\alpha_{t'}|Y_{t'-1})$ とフィルタリング密度関数 $f(\alpha_{t'}|Y_{t'})$ は大きく異なることが十分に考えられる。この状況下で，フィルタリングで(9)式を用いる

と，IR や MH については $f(\alpha_{t'}|Y_{t'})$ から生成された $\alpha_{t'}$ の乱数は尤もらしい値とはならない。また，RS については $f(\alpha_{t'}|Y_{t'})$ から $\alpha_{t'}$ の乱数を生成するために極端に計算時間が増大することになる。よって，(10)式で用いられる $\alpha_{t'}$ のサンプリング密度関数 $f_*(\alpha_{t'}|\alpha_{t'-1})$ を取り入れると，より効率的な乱数の生成が可能になることが予想される。このように状況に応じて，時点 $t'$ では(10)式を使い，それ以外の時点では(9)式を使うようにすることも可能である。

$f(\alpha_t|Y_{t-1})$ から $\alpha_t$ の乱数生成が難しいとき，すなわち，$\eta_t$ の密度関数は既知であるが乱数を生成するのが難しい場合にも，サンプリング密度関数 $f_*(\alpha_t|Y_{t-1})$ を利用して $\alpha_t$ の乱数を生成することができる。

## 5　最尤法によるパラメータ推定

(1)式と(2)式の観測・遷移方程式に未知パラメータ（例えば，$\theta$）が含まれる場合，未知パラメータと状態変数を同時に推定する必要がある。未知パラメータの推定に通常よく用いられる方法は最尤法 (maximum likelihood estimation) である。これは尤度関数 (likelihood function) を最大にするような未知パラメータの値をその推定値とするものである。尤度関数 $f(Y_T)$ は，

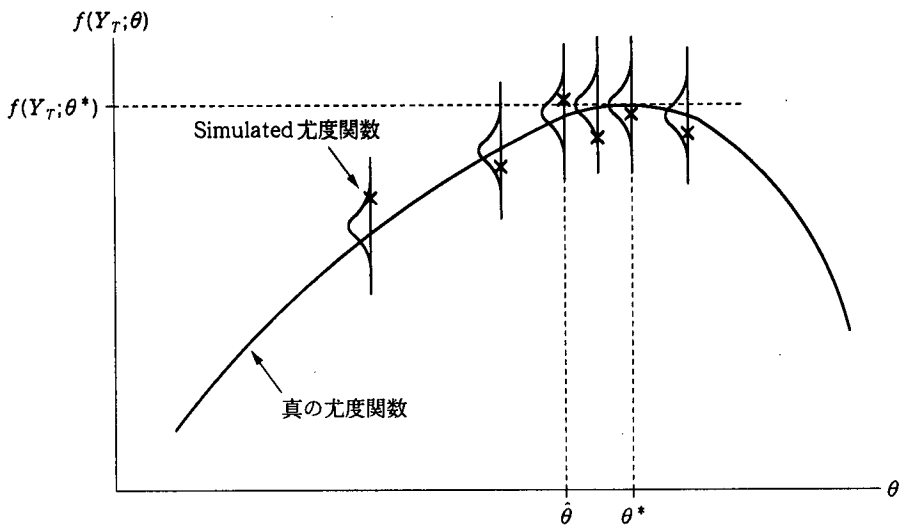
$$f(Y_T)=\prod_{t=1}^{T}f(y_t|Y_{t-1})=\prod_{t=1}^{T}\int f_y(y_t|\alpha_t)f(\alpha_t|Y_{t-1})\mathrm{d}\alpha_t\approx\prod_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N}f_y(y_t|\alpha_{i,t|t-1})\right) \tag{15}$$

と近似できる。ここで $f(y_1)=f(y_1|Y_0)$ みなされる。$f(y_t|Y_{t-1})$ は(7)式の分母に当たり，$f(y_t|Y_{t-1})=\int f_y(y_t|\alpha_t)f(\alpha_t|Y_{t-1})\mathrm{d}\alpha_t$ として表される。$\eta_{i,t}$ は $\eta_t$ の $i$ 番目の乱数とする。$\alpha_{i,t|t-1}$ を $f(\alpha_t|Y_{t-1})$ から生成された $\alpha_t$ の乱数は，$j$ を $1,2,\cdots,N$ からランダムに一つ選んで，$\alpha_{i,t|t-1}=q_t(\alpha_{j,t-1|t-1},\ \eta_{i,t})$ として生成される。(15)式は Simulated 尤度関数と呼ばれ，$\theta$ について最大化される。最大化の方法は単純サーチ法やスコアリング法が用いられる。

しかし，Simulated 尤度関数を最大化する場合，$\theta$ のある値に対応して，$\alpha_{i,t|t-1}$ の値が大きく異なってしまう。これを示したものが図1ある。データ $y_t$ を与えたもとで，真の尤度関数を最大にする $\theta$ を $\theta^*$ とする。図中の×で表された値は，$\theta$ を与えたときの Simulated 尤度関数の値を示している。すなわち，Simulated 尤度関数は真の尤度関数の値の周りで分布する確率変数であり，図中の×はその実現値を表す。6通りの $\theta$ の値について，それぞれに対応する6通りの実現値が×として得られたとしよう。この場合，$\theta^*$ が尤度関数の最大値であるにもかかわらず，Simulated 尤度関数の実現値の最大値となっている $\hat{\theta}$ が $\theta$ の最尤推定値として選ばれることになる。

このような状況を回避するために，Kitagawa (1998) は Self-Organizing フィルタを提案した。これは未知パラメータも状態変数（すなわち，$\theta_t=\theta_{t-1}$ とする）とみなし，状態変数と

図1 真の尤度関数とSimulated尤度関数




同時にパラメータも推定しようというものである。その他では，Markov Chain Monte Carlo (MCMC) と呼ばれる乱数生成法である Gibbs sampler と MH アルゴリズムを組み合わせて，状態空間モデルに応用される。Gibbs sampler とは，条件付き分布から順番に生成した乱数の収束先は無条件分布から生成した乱数に等しくなるという性質を利用した乱数生成法である。ベイズ統計学の分野で近年非常によく用いられる。例えば，$\theta$の事前分布を flat prior (または，improper prior や diffuse prior とも呼ばれる) と仮定したとき，真の尤度関数を$\theta$の密度関数に比例するとみなして$\theta$の乱数を生成し，その平均値を推定値とする。紙面の都合上このあたりの解説は省略するが，参考文献は Jacquier *et al.* (1994), Watanabe (2000), Watanabe and Omori (2004), Tanizaki (2004b)，渡部 (2000) を参照せよ。

## 6 おわりに

本稿では非線形・非正規分布のもとでの状態空間モデルにおいて，状態変数の推定を紹介した。このモデルでは大きく分けて2つの状態変数の推定方法が考えられている。一つは分布関数による逐次アルゴリズムに基づいたもの，もう一つは全部の状態変数の同時密度関数に基づいたものである。非線形・非正規状態空間モデルの場合，前者は乱数の逐次アルゴリズムを得ることになり，後者は MCMC (Gibbs サンプラー, Metropolis-Hasting アルゴリズムを含む) と呼ばれる乱数生成法を用いることになる。本稿では前者の乱数を逐次アルゴリズムによって生成するという方法を紹介した。そして，状態空間モデルに未知パラメータが含まれている場合の未知パラメータの推定問題を紹介した。状態空間モデルのような観測で

きない変数（すなわち，状態変数）が含まれていて，かつ，未知パラメータも同時に推定する必要がある場合，Simulated 尤度関数を未知パラメータに関して最大化する必要がある。しかし，Simulated 尤度関数は確率変数となるため，得られた Simulated 尤度関数の実現値を最大にする未知パラメータは，必ずしも真の尤度関数は最大化されていないという状況が出てくるという問題点がある。本稿では以上のような状況が起こることを紹介した。この問題を打開するためには今のところ Kitagawa (1998) によって示された Self-Organizing フィルタを採用するか，または，MCMC を利用したベイズ推定を行うかのどちらかである。観測されない変数を含む場合の未知パラメータの推定問題は今後も研究対象となり得る分野であると言える。

＊本稿は谷崎 (2006) の続編として執筆し，Tanizaki (1996, 2003, 2004a) に基づいて加筆・修正したものである。本研究は科学研究費 (C) (2) 14530033 (2002年～2005年) と21世紀 COE プログラムからの助成を受けた研究の一部である。

## 参 考 文 献

Boswell, M.T., Gore, S.D., Patil, G.P. and Taillie, C. (1993). "The Art of Computer Generation of Random Variables," *Handbook of Statistics*, Vol. 9 (ed. C.R. Rao), pp. 661-721, North-Holland.

Chib, S. and Greenberg, E. (1995). "Understanding the Metropolis-Hastings Algorithm," *The American Statistician*, Vol. 49, pp. 327-335.

Geweke, J. (1992). "Evaluating the Accuracy of Sampling-based Approaches to the Calculation of Posterior Moments," in *Bayesian Statistics 4*, edited by Bernardo, J.M., Berger, J.O., Dawid, A.P. and Smith, A.F.M., pp. 169-193, Oxford University Press.

Geweke, J. (1996). "Monte Carlo Simulation and Numerical Integration," *Handbook of Computational Economics*, Vol. 1 (ed. H.M. Amman, D.A. Kendrick and J. Rust), pp. 731-800, North-Holland.

Geweke, J. and Tanizaki, H. (2003). "Note on the Sampling Density for the Metropolis-Hastings Algorithm," *Communications in Statistics, Theory and Methods*, Vol. 32, No. 4, pp. 775-789.

Gordon, N.J., Salmond, D.J. and Smith, A.F.M. (1993). "Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation," *IEE Proceedings-F*, Vol. 140, No. 2, pp. 107-113.

Hürzeler, M. and Künsch, H.R. (1998). "Monte Carlo Approximations for General State-Space Models," *Journal of Computational and Graphical Statistics*, Vol. 7, pp. 175-193.

Jacquier, E., Polson, N.G. and Rossi, P.E. (1994). "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business and Economic Statistics*, Vol. 14, pp. 371-417 (with discussion).

Kitagawa, G. (1996). "Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State-Space Models," *Journal of Computational and Graphical Statistics*, Vol. 5, pp. 1-25.

Kitagawa, G. (1998). "A Self-Organizing State-Space Model," *Journal of the American Statistical Association*, Vol. 93, pp. 1203-1215.

Kitagawa, G. and Gersch,W. (1996). *Smoothness Priors Analysis of Time Series* (Lecture Notes in Statistics, No. 116), Springer-Verlag.

Liu, J.S. (1996). "Metropolized Independent Sampling with Comparisons to Rejection Sampling and Importance Sampling," *Statistics and Computing*, Vol. 6, pp. 113-119.

O'Hagan, A. (1994). *Kendall's Advanced Theory of Statistics*, Vol. 2B, Edward Arnold.

Smith, A.F.M. and Gelfand, A.E. (1992). "Bayesian Statistics without Tears: A Sampling-Resampling Perspective," *The American Statistician*, Vol. 46, pp. 84-88.

Tanizaki, H. (1996). *Nonlinear Filters: Estimation and Applications* (Second, Revised and Enlarged Edition), Springer-Verlag.

Tanizaki, H. (1999). "On the Nonlinear and Nonnormal Filter Using Rejection Sampling," *IEEE Transactions on Automatic Control*, Vol. 44, No. 2, pp. 314-319.

Tanizaki, H. (2001). "Nonlinear and Non-Gaussian State Space Modeling Using Sampling Techniques," *Annals of the Institute of Statistical Mathematics*, Vol. 53, No. 1, pp. 63-81.

Tanizaki, H. (2003). "Nonlinear and Non-Gaussian State-Space Modeling with Monte Carlo Techniques: A Survey and Comparative Study," in *Handbook of Statistics*, Vol. 21: *Stochastic Processes: Modeling and Simulation*, Chap. 22, pp. 871-929 (C.R. Rao and D.N. Shanbhag, Eds.), North-Holland.

Tanizaki, H. (2004a). *Computational Methods in Statistics and Econometrics* (STATISTICS: textbooks and monographs, Vol. 172), Mercel Dekker.

Tanizaki, H. (2004b). "On Asymmetry, Holiday and Day-of-the-Week Effects in Volatility of Daily Stock Returns: The Case of Japan," *Journal of the Japan Statistical Society*, Vol. 34, No. 2, pp. 129-152.

Tanizaki, H. and Mariano, R.S. (1998). "Nonlinear and Non-Gaussian State-Space Modeling with Monte-Carlo Simulations," *Journal of Econometrics*, Vol. 83, No. 1 & 2, pp. 263-290.

Watanabe, T. (2000). "Bayesian Analysis of Dynamic Bivariate Mixture Models: Can They Explain the Behavior of Returns and Trading Volume?," *Journal of Business and Economic Statistics*, Vol. 18, pp. 199-210.

Watanabe, T. and Omori, Y. (2004). "A Multi-move Sampler for Estimating Non-Gaussian Time Series Models: Comments on Shephard & Pitt (1997)," *Biometrika*, Vol. 91, pp. 246-248.

片山徹 (2000).『新版応用カルマンフィルタ』朝倉書店。

谷崎久志 (2006).「状態空間モデルの紹介と最近の展開」『計量経済学ハンドブック』(箕谷千鳳彦, 縄田和満, 和合肇編), 17章, 朝倉書店。

渡部敏明 (2000).『ボラティリティ変動モデル』(シリーズ<現代金融工学>4)朝倉書店。