谷崎　久志
大阪大学大学院経済研究科

本稿は,

- 谷崎久志 (2012) 「カルマンフィルター」『経済時系列分析ハンドブック』 (刈屋武昭・前川功一・矢島美寛・川崎能典・福地純一郎編), 1 章 13 節, pp.190-203, 朝倉書店.
- 谷崎久志 (2007) 「状態空間モデル」『計量経済学ハンドブック』 (箕谷千凰彦，縄田和満，和合肇編), 20 章, pp.621-642, 朝倉書店.

の抜粋である。

1　はじめに

フィルタリング (filtering) 理論は，Kalman (1960)，Kalman and Bucy (1961) を 初めとして，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。本章では状態空間モデルの紹介を目的とし，導出方法等を中心として簡単にサーベイを行う（状態空間モデルの定義は次節参照）。本章は，Tanizaki (1993b, 1996，2003，2004a)，谷崎 (1993) に基づいて，加筆・修正したものである。フィルタリングに関する邦語文献としては，有本 (1977)，片山 (2000) など理工系のものが多い。経済学に関連したものは，渡部 (2000) があるがまだまだその数は少ない。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。状態空間モデルのいくつかの応用例を以下に紹介しておく。

可変パラメータ (Time-Varying Parameter)・モデル： 計量モデルの基本である最小自乗法 (OLS) の前提の一つに「パラメータは推定期間を通して一定である」がある。この根拠は「十分に長い期間をとれば，確かに経済構造は徐々に変化しているが，その一部分のごく短期的な期間をとれば十分に線形近似できる」というところにある。すなわち，通常の回帰モデルでは，$y_t = X_t\boldsymbol{\beta} + \epsilon_t$ と表され，固定パラメータである $\boldsymbol{\beta}$ を推定することになる。しかし，実際には近年特に，様々な外生的なショック (為替の変動相場制への移行，第一次・第二次石油ショック，貿易摩擦

による輸出規制，円高の進行，バブル等)，または他の政策的な要因等により経済構造は徐々に変化していると考えるのがより自然な考え方である。このように徐々に経済構造が変化している状況において，従来のOLS 等の固定パラメータ・モデルによる推定では，この経済構造の変化を表すことはできない。したがって，パラメータの変動を明示的に取り入れたモデル，すなわち，これは可変パラメータ・モデル (time-varying parameter model) を考える必要がある。可変パラメータを扱ったモデルにはいくつかの種類が考えてられいるが，パラメータの変動をランダムな確率的な観測できない変数として，状態空間モデルを応用することができる。可変パラメータ・モデルを扱ったものには，Cooper (1973), Belsley and Kuh (1973)，Sarris (1973)，Cooley and Prescott (1973，1976), Laumas and Mehra (1976)，Garbade (1977)，Cooley (1977)，Sant (1977), Pagan (1980)，谷崎 (1993)，Dziechciarz (1989)，Tanizaki (1989，1993a, 2000) 等数々ある。

この場合，回帰モデルは，

$$
\begin{aligned}
&(\text{観測方程式}) \qquad y_t = \boldsymbol{X}_t \boldsymbol{\beta}_t + \epsilon_t \\
&(\text{遷移方程式}) \qquad \boldsymbol{\beta}_t = \phi \boldsymbol{\beta}_{t-1} + \eta_t
\end{aligned}
$$

と修正される。観測方程式，遷移方程式については，次節の (1)，(2) に対応する。パラメータの変動をランダムな確率的な観測できない変数として，カルマンフィルター・モデルを応用することができる。谷崎 (1993, 2007)，Tanizaki (1989，1993a，2000) の参考文献を参照せよ。

自己回帰移動平均 (Autoregressive Moving Average) モデル： カルマンフィルター・モデルを用いて，任意の自己回帰移動平均過程 (AutoRegressive-Moving Average Process，ARMA Process)，すなわち，$y_t = a_1 y_{t-1} + \cdots + a_p y_{t-p} + \epsilon_t + b_1 \epsilon_{t-1} + \cdots + b_q \epsilon_{t-q}$ の係数パラメータを推定することができる。撹乱項 ϵ_t はすべての t についてホワイト・ノイズ (white noise) であるとする。この ARMA(p, q) モデルは，$m = \max(p, q+1)$ とおくと，$y_t = a_1 y_{t-1} + \cdots + a_m y_{t-m} + \epsilon_t + b_1 \epsilon_{t-1} + \cdots + b_{m-1} \epsilon_{t-m+1}$ のように ARMA$(m, m-1)$ モデルとして書き換えられ，

$$
\begin{aligned}
&(\text{観測方程式}) \qquad y_t = \boldsymbol{z} \boldsymbol{\alpha}_t \\
&(\text{遷移方程式}) \qquad \boldsymbol{\alpha}_t = \boldsymbol{A} \boldsymbol{\alpha}_{t-1} + \boldsymbol{B} \epsilon_t
\end{aligned}
$$

と表される。$\boldsymbol{\alpha}_t$ は $m\times 1$ ベクトル，$\boldsymbol{z}, \boldsymbol{A}, \boldsymbol{B}$ はそれぞれ以下の通りである。

$$\underset{1\times m}{\boldsymbol{z} = (1, 0, \cdots, 0)}, \qquad \boldsymbol{A} = \underset{m\times m}{\left(\begin{array}{c|c} a_1 & \\ \vdots & \boldsymbol{I}_{m-1} \\ a_{m-1} & \\ \hline a_m & \boldsymbol{0} \end{array}\right)}, \qquad \boldsymbol{B} = \underset{m\times 1}{\begin{pmatrix} 1 \\ b_1 \\ \vdots \\ b_{m-1} \end{pmatrix}}$$

ARMA モデルとカルマンフィルター・モデルとの関係については，例えば，Aoki (1987)，Burridge and Wallis (1988)，Harvey (1981，1989) 等数多くの文献がある。

季節調整 (Seasonal Adjustment) モデル： Pagan (1975)，Chow (1983) は，状態空間モデルを用いて，季節要素モデル (seasonal component model) の推定を提案した。カルマンフィルター・モデルを用いて，季節要素モデル (Seasonal Component Model) を推定することができる。時系列データは通常，循環的要素 (Cyclical Component)，季節的要素 (Seasonal Component)，撹乱的要素 (Irregular Component) から成る。それぞれの要素は観測されない変数であり，状態空間モデルを適用することによって，それらの要素を別々に推定することができる。

原系列データは，$y_t = y_t^c + y_t^s + v_t$ のように書き表される。y_t, y_t^c, y_t^s, v_t は，それぞれ原系列データ，循環的要素，季節的要素，撹乱的要素を表す。そして，循環的要素は，一期前の循環的要素と他の外生変数に依存すると仮定され，$y_t^c = Ay_{t-1}^c + \boldsymbol{C}\boldsymbol{X}_t + u_t$ となる。$A, \boldsymbol{C}$ は係数，$\boldsymbol{X}_t$ は $k\times 1$ の外生変数のベクトル，u_t は撹乱項である。さらに，季節的要素は一年前の同じ期のそれに依存すると考えられるので，$y_t^s = \boldsymbol{B}y_{t-m}^s + w_t$ と表される。$\boldsymbol{B}$ は係数，w_t は撹乱項である。また，m は四半期モデルのとき 4，月次モデルのとき 12 となる。以上をまとめると，次のモデルが得られる。

$$\begin{array}{ll} \text{(観測方程式)} & y_t = \boldsymbol{z}\boldsymbol{\alpha}_t + v_t \\ \text{(遷移方程式)} & \boldsymbol{\alpha}_t = \boldsymbol{M}\boldsymbol{\alpha}_{t-1} + \boldsymbol{N}\boldsymbol{X}_t + \boldsymbol{\epsilon}_t \end{array}$$

それぞれの記号は以下に示される。

$$\underset{1\times(m+1)}{\boldsymbol{z} = (1, 1, 0, \cdots, 0)}, \quad \boldsymbol{M} = \underset{(m+1)\times(m+1)}{\left(\begin{array}{c|cc} A & & \boldsymbol{0} \\ \hline & \boldsymbol{0} & \boldsymbol{I}_{m-1} \\ \boldsymbol{0} & B & \boldsymbol{0} \end{array}\right)}, \quad \boldsymbol{N} = \underset{(m+1)\times k}{\begin{pmatrix} \boldsymbol{C} \\ \boldsymbol{0} \end{pmatrix}}, \quad \boldsymbol{\epsilon}_t = \underset{(m+1)\times 1}{\begin{pmatrix} u_t \\ w_t \\ \boldsymbol{0} \end{pmatrix}}$$

$(m+1)\times 1$ ベクトル $\boldsymbol{\alpha}_t$ の第 1，第 2 要素は，それぞれ循環的要素，季節的要素である。それぞれの要素は観測されない変数であり，カルマンフィルター・モデルを適用することによって，それらの要素を別々に推定することができる。例えば，Chow (1983) を参照せよ。

確報値 (Final Data) の推定： 通常，経済データを得るとき，まず初めに速報値 (preliminary data) が得られる。次にしばらくしてから，確報値 (final data) または改訂値 (revised data) が公表される (ほとんどの経済データは改訂されるが，改訂が行われないデータとしては，金利，為替レート等のごく一部である)。速報値が利用可能であるときに，いかに確報値 (または改訂値) を推定するのかということが問題になる。例えば，国民所得統計のようなデータは最初の数年にわたって毎年改訂され，その後も一定期間 (5 年，10 年) ごとに改訂される (詳細な解説については,『国民経済計算年報』 (内閣府経済社会総合研究所編) の参考資料にある用語解説の「速報と確報」を参照せよ)。データは最初の 2，3 年間そして一定期間毎に改訂されるため，確報値は観測されない変数とみなされ得る。元来，速報値も確報値も同じデータであるので，速報値と確報値との間にはなんらかの関係が存在する。これを観測方程式 (次節の (1) 式参照) として表す。また，ある経済理論から導き出された関係式は，速報値よりむしろ，確報値に基づいたものであると考えるのが適当である。この関係式を遷移方程式 (次節の (2) 式参照) として扱う。「ある経済理論から導き出された関係式」を遷移方程式とするためには，方程式にラグ構造を含む必要がある。状態空間モデルの遷移方程式は観測不可能な変数の AR(1) 過程として表現される。一つの例としては，動学的最適化問題を解くことによって得られるオイラー方程式に基づいて，遷移方程式を導き出すことができる。このように，状態空間モデルを使って，この確報値の推定問題を考えることができる。Howrey (1978，1984)，Conrad and Corrado (1979)，Harvey (1989)，Tanizaki and Mariano (1994) 等の文献があげられる。Mariano and Tanizaki (1995) は，効用最大化問題を解くことによって得られる消費関数を遷移方程式として，消費の確報値を推定した。

この場合，モデルは，

$$
\begin{aligned}
&\text{(観測方程式)} \qquad y_t^p = \gamma y_t^f + \epsilon_t \\
&\text{(遷移方程式)} \qquad y_t^f = \theta y_{t-1}^f + \boldsymbol{X}_t \boldsymbol{\beta} + \eta_t
\end{aligned}
$$

となる。y_t^p, y_t^f は，それぞれ速報値，確報値を表し， ϵ_t と η_t は撹乱項である。$\boldsymbol{X}_t$ は他の外生変数であり，さらに， $\gamma, \theta, \boldsymbol{\beta}$ は推定されるべき未知パラメータである。速報値 $y_t^p, \boldsymbol{X}_t$ は観測可能な変数であるが，確報値 y_t^f

は観測できない変数である。カルマンフィルター・モデルを使って確報値の推定問題を考えることができる。Harvey (1989)，Mariano and Tanizaki (1995) 等の文献があげられる。

確率的ボラティリティ (Stochastic Volatility)・モデル： 最近では，Markov Chain Monte Carlo (MCMC) という乱数生成の方法が取り入れられ，より多くの実証分析が行われるようになった。その一つとして，確率的ボラティリティ (Stochastic Volatility，SV) モデルというものがある。ボラティリティが観測できない状態変数として扱われる。為替レートや株価の変動要因を調べるために用いられるモデルであり，Jacquier *et al.* (1994, 2004)，Tanizaki (2004b)，Watanabe (1999，2000)，渡部 (2000) 等の研究がある。他にも多くの研究が行われているが，詳しくは，これらの参考文献を参照せよ。

モデルは,

$$\begin{aligned} &(\text{観測方程式}) \qquad y_t = \boldsymbol{X}_t\boldsymbol{\beta} + \exp(\frac{1}{2}h_t)\epsilon_t \\ &(\text{遷移方程式}) \qquad h_t = \delta h_{t-1} + \eta_t \end{aligned}$$

と表される。ϵ_t は平均 0，分散 1 の誤差項であり，通常，正規分布が仮定される。$\exp(h_t)$ が回帰モデルの分散 (ボラティリティと呼ばれる) に対応し，時間に応じて変動するモデルである。ボラティリティを観測できない変数とみなす。このモデルは非線形カルマンフィルター・モデルであり本節では扱わない。為替レートや株価の変動要因を調べるために用いられるモデルであり，Tanizaki (2004b)，Tanizaki and Hamori (2009), Watanabe (1999，2000)，渡部 (2000) 等の研究がある。他にも多くの研究が行われているが，詳しくは，これらの参考文献を参照せよ。

マルコフ推移 (Markov Switching) モデル： マルコフ推移 (Markov switching) モデルは Hamilton (1989，1990，1991，1993，1994) によって発展された。マルコフ推移 (Markov Switching) モデルとは，経済の状態 (レジーム，regime) を離散型確率変数で，しかも，観測できない変数として扱うモデルである。このモデルでは経済の状態がある状態から別の状態へ確率的に変動すると考える。s_t を t 期の状態とするとき，回帰モデルは $y_t = \boldsymbol{X}_t\boldsymbol{\beta}_{s_t} + \epsilon_t$ となり，これが観測方程式に対応する。さらに，t 期の経済の状態 s_t は $t-1$ 期の経済の状態 s_{t-1} に依存するというマルコフ・プロセスに従う，すなわち，$\mathrm{P}(s_t = i|s_{t-1} = j) = p_{ij}$ と仮定する。すなわち，状態の変動の仕方は，前期の経済状態に応じて今期の経済状態が決まるというマルコフ・プロセスに従うとする。経済の状態の数を k 個とすれば，$i, j = 1, 2, \cdots, k$ となり，p_{ij} によって推移確率行列 ($k \times k$ 行列) が構成される。この s_t に関する推移確率行列が遷移方程式に対応する。ある

時点の経済がどの状態にあるかを調べることができる。マルコフ推移モデルは Hamilton (1994) によって発展された。このモデルは非正規カルマンフィルター・モデルでありこれも本節では扱わない。

いくつかの応用例を示したが，その他にも合理的期待変数の推定にも状態空間モデルは用いられる (McNelis and Neftci (1983)，Burmeister and Wall (1982)) ことを付け加えておく。次節ではこの状態空間モデルを紹介する。

2　状態空間モデルの定義

本節から 5 節までは線形・正規性の仮定のもとで状態空間モデルを考え，6 節で非線形・非正規状態空間モデルを扱う。状態空間モデル (state-space model) は，次のような観測方程式 (measurement equation) と遷移方程式 (transition equation) の 2 つの式によって表される。状態空間モデルの標準的なテキストとしては，Jazwinski (1970)，Gelb (1974)，Anderson and Moore (1979)，Harvey (1989)，Tanizaki (1993b，1996)，片山 (2000) 等が適当である。

$$\text{(観測方程式)} \qquad y_t = Z_t\alpha_t + d_t + S_t\epsilon_t \tag{1}$$

$$\text{(遷移方程式)} \qquad \alpha_t = T_t\alpha_{t-1} + c_t + R_t\eta_t \tag{2}$$

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} H_t & 0 \\ 0 & Q_t \end{pmatrix} \right) \qquad t = 1, 2, \cdots, T \tag{3}$$

$$\begin{array}{lllll} y_t:\ g\times 1 & Z_t:\ g\times k & d_t:\ g\times 1 & S_t:\ g\times g & \epsilon_t:\ g\times 1 \\ \alpha_t:\ k\times 1 & T_t:\ k\times k & c_t:\ k\times 1 & R_t:\ k\times k & \eta_t:\ k\times 1 \end{array}$$

y_t，Z_t，d_t，S_t，T_t，c_t，R_t は観測可能な変数，ϵ_t，η_t は攪乱項とする。T は標本数を表す。α_t は状態変数と呼ばれ，観測されない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。観測される変数を用いて，観測されない変数 α_t を推定することが目的である。以下の 2 つの仮定を必要とする。

(i) 初期値 α_0 は確率変数として扱う場合と非確率変数とする場合の両方がある。確率変数として扱う場合は，今のところ，$\alpha_0 \sim N(a_0, \Sigma_0)$ と仮定する。

(ii) 攪乱項 ϵ_t，η_s はすべての t，s について，互いに独立であり，初期値ベクトル α_0 とも無相関である (すなわち，すべての t，s について $\mathrm{E}(\epsilon_t\eta_s') = 0$，すべての t について $\mathrm{E}(\epsilon_t\alpha_0') = 0$，$\mathrm{E}(\eta_t\alpha_0') = 0$ が成り立つ)。

注意すべき点は以下の通りでる。

1) 仮定 (ii) は，ϵ_t と α_t の間の無相関，η_t と α_{t-1} との無相関を保証する。すなわち，$\mathrm{E}(\epsilon_t\alpha_t') = 0$，$\mathrm{E}(\eta_t\alpha_{t-1}') = 0$ となる。
2) Z_t，d_t，S_t，T_t，c_t，R_t は未知パラメータ (例えば，θ) に依存してもよい。この場合には，未知パラメータ θ は状態変数 α_t と共に推定されなければならない。α_t の推定問題については 3 節に述べられ，θ については 5 節で触れる。
3) 初期値 α_0 と撹乱項 ϵ_t，η_t には正規分布を仮定するのが一般的ではあるが，アルゴリズムの導出方法によっては，この仮定を必要としない。例えば，混合推定によるフィルタリング・アルゴリズムの導出は誤差項 ϵ_t，η_t に分布を仮定する必要はない。しかし，分布関数に基づいた線形アルゴリズムの導出には，正規性の仮定を必要とする (詳しくは，後述の 3 節を見よ)。
4) 観測方程式に含まれる撹乱項 ϵ_t と遷移方程式の撹乱項 η_t は互いに無相関 (これは (3) 式の分散の対角要素がゼロという仮定に相当する) として，本章では議論を進める。しかし，この仮定を緩めることも可能である。その場合の議論については，例えば，片山 (2000) や Harvey (1989) 等に述べられている。

以上のように，(1) 式と (2) 式の 2 つの式から成る状態空間モデルは観測されない変数 α_t と観測される変数 y_t，Z_t，d_t，S_t，T_t，c_t，R_t から構成される。そして，観測されない変数 α_t を推定する問題として，予測問題 (プレディクション，Prediction)，濾波問題 (フィルタリング，Filtering)，平滑問題 (スムージング，Smoothing) の 3 種類を考えることができる。次節ではこの 3 つの推定問題について述べる。そして，それぞれの 3 つの推定問題について，状態変数を算出するためのアルゴリズムが示される。

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，$Y_s = \{y_s, y_{s-1}, \cdots, y_1\}$ である。(1) 式や (2) 式の中で，Z_t，d_t，S_t，T_t，c_t，R_t，$t = 1, 2, \cdots, T$，のような，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。しかし，煩雑になるのを避ける

ため，ここではそれらを省略する。$a_{t|s}$，$\Sigma_{t|s}$ を次のように定義する。

$$a_{t|s} = \mathrm{E}(\alpha_t|Y_s) = \int \alpha_t f(\alpha_t|Y_s)\mathrm{d}\alpha_t \tag{4}$$

$$\Sigma_{t|s} = \mathrm{Var}(\alpha_t|Y_s) = \int (\alpha_t - a_{t|s})(\alpha_t - a_{t|s})' f(\alpha_t|Y_s)\mathrm{d}\alpha_t \tag{5}$$

このとき，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。

$t > s$ のとき，予測 (プレディクション，prediction)
$t = s$ のとき，濾波 (フィルタリング，filtering)
$t < s$ のとき，平滑 (スムージング，smoothing)

さらに，$f(\cdot|\cdot)$ を条件付き分布関数とする。このとき，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。

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，上記の添字記号 t，s をそれぞれ $(t + L)$，t に置き換えて述べられる。すなわち，以下の推定問題を考えることになる。

$$\mathrm{E}(\alpha_{t+L}|Y_t) = a_{t+L|t} \qquad \mathrm{Var}(\alpha_{t+L}|Y_t) = \Sigma_{t+L|t} \qquad L = 1, 2, \cdots$$

まず分布関数に基づいたプレディクション・アルゴリズムをあげておく。

$$\begin{aligned} f(\alpha_{t+L}|Y_t) &= \int f(\alpha_{t+L}, \alpha_{t+L-1}|Y_t)\mathrm{d}\alpha_{t+L-1} \\ &= \int f(\alpha_{t+L}|\alpha_{t+L-1}, Y_t) f(\alpha_{t+L-1}|Y_t)\mathrm{d}\alpha_{t+L-1} \\ &= \int f_\alpha(\alpha_{t+L}|\alpha_{t+L-1}) f(\alpha_{t+L-1}|Y_t)\mathrm{d}\alpha_{t+L-1} \end{aligned} \tag{6}$$

$L = 1, 2, \cdots$，に関する逐次アルゴリズムとなっている。3 つ目の等号が成り立つ理由 (すなわち，$f(\alpha_{t+L}|\alpha_{t+L-1}, Y_t) = f_\alpha(\alpha_{t+L}|\alpha_{t+L-1})$ が成り立つ理由) は，遷移方程式は t 期までの情報集合 Y_t に依存しないからである。分布関数に基づいたプレディクション・アルゴリズムについては，Kitagawa (1987)，Harvey (1989) 等を参照せよ。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，とりあえず $f(\alpha_t|Y_t)$ 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 (すなわち，$\alpha_{t+L}=T_{t+L}\alpha_{t+L-1}+c_{t+L}+R_{t+L}\eta_{t+L}$) とその中の撹乱項 η_{t+L} の分布関数をもとにして，η_{t+L} から α_{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，α_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期待値や分散が求められる

さらに，上の分布関数に基づく逐次アルゴリズム (recursive algorithm) について，遷移方程式の線形性と撹乱項の正規性の仮定を置くと，次の線形の逐次アルゴリズム (linear recursive algorithm) が得られる。ただし，導出方法によっては撹乱項の正規性の仮定は必要としない。遷移方程式が線形であれば，(2) 式の両辺に条件付き期待値と分散をとることによって，(7) 式と (8) 式に表される線形の逐次アルゴリズムが得られる。

$$a_{t+L|t}=T_{t+L}a_{t+L-1|t}+c_{t+L} \tag{7}$$

$$\Sigma_{t+L|t}=T_{t+L}\Sigma_{t+L-1|t}T'_{t+L}+R_{t+L}Q_{t+L}R'_{t+L} \qquad L=1,2,\cdots \tag{8}$$

$a_{t|t}$，$\Sigma_{t|t}$ が与えられているとき，(7) 式と (8) 式から次の期の予測値 $a_{t+1|t}$ とその分散 $\Sigma_{t+1|t}$ が得られる。同様にして，$a_{t+1|t}$，$\Sigma_{t+1|t}$ から $a_{t+2|t}$，$\Sigma_{t+2|t}$ が求められる。このように (7) 式，(8) 式は，逐次計算によって，$a_{t+L|t}$，$\Sigma_{t+L|t}$，$L=1,2,\cdots$，を求めるアルゴリズムになっている。ここで，$a_{t|t}$，$\Sigma_{t|t}$ は次節で述べるフィルタリング推定値とその分散である。

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。ゆえに，フィルタリング問題は次の数学的期待値を求めることに等しい。

$$\mathrm{E}(\alpha_t|Y_t)=a_{t|t} \qquad \mathrm{Var}(\alpha_t|Y_t)=\Sigma_{t|t} \qquad t=1,2,\cdots,T$$

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。

$$f(\alpha_t|Y_{t-1})=\int f_\alpha(\alpha_t|\alpha_{t-1})f(\alpha_{t-1}|Y_{t-1})\mathrm{d}\alpha_{t-1} \tag{9}$$

$$f(\alpha_t|Y_t)=f(\alpha_t|y_t,Y_{t-1})=\frac{f(\alpha_t,y_t|Y_{t-1})}{f(y_t|Y_{t-1})}$$

$$
\begin{aligned}
&= \frac{f(\alpha_t, y_t|Y_{t-1})}{\displaystyle\int f(\alpha_t, y_t|Y_{t-1})\mathrm{d}\alpha_t} = \frac{f_y(y_t|\alpha_t, Y_{t-1}) f(\alpha_t|Y_{t-1})}{\displaystyle\int f_y(y_t|\alpha_t, Y_{t-1}) f(\alpha_t|Y_{t-1})\mathrm{d}\alpha_t} \\
&= \frac{f_y(y_t|\alpha_t) f(\alpha_t|Y_{t-1})}{\displaystyle\int f_y(y_t|\alpha_t) f(\alpha_t|Y_{t-1})\mathrm{d}\alpha_t} \qquad (10)
\end{aligned}
$$

$t = 1, 2, \cdots, T$ に関する逐次アルゴリズムになっている。分布に基づいたフィルタリング・アルゴリズムについては，Kitagawa (1987)，Harvey (1989) 等を参照せよ。(9) 式は予測方程式と呼ばれる。それは，(6) 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。(10) 式については，最初の等号では $Y_t = \{y_t, Y_{t-1}\}$ であることが利用されている。また，5 つ目の等号について，$f(y_t|\alpha_t, Y_{t-1}) = f_y(y_t|\alpha_t)$ が成立するのは，(1) 式の観測方程式に過去の情報集合 Y_{t-1} が含まれていないためである。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 ϵ_t の分布関数をもとにして，ϵ_t から y_t への変数変換によって求められる。(10) 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。

(9) 式と (10) 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，(9) 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 (9) 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，(9) 式によって，$f(\alpha_1|Y_1)$ から $f(\alpha_2|Y_1)$，(10) 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，α_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 (9) は異なることに注意せよ。

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

ただし，$f(\alpha_0)$ は α_0 が確率変数のときの分布関数である。

線形の観測・遷移方程式と撹乱項の正規性の仮定のもとで，(9) 式，(10) 式から得られる 1 次，2 次の積率 (moment) から次の線形のフィルタリング・アルゴリズムが導出される。Kalman (1960)，Kalman and Bucy (1961) は線形確率システムの状態空間表現と最小分散推定の理論を組み合わせることにより，フィルタリング問題の定式化を行い，直交射影の定理を用いてフィルタリング・アルゴリズムを導出した。以下の (11) 式 (17) 式

によって表される線形の逐次アルゴリズムを，特に，カルマン・フィルタと呼ぶ。一般に，カルマン・フィルタ・アルゴリズムは攪乱項の正規性の仮定を必要としない。観測・遷移方程式の線形性の仮定のみで線形の逐次アルゴリズムが導出される。以下のアルゴリズムは 4 節で証明される。

$$a_{t|t-1} = T_t a_{t-1|t-1} + c_t \tag{11}$$

$$\Sigma_{t|t-1} = T_t \Sigma_{t-1|t-1} T_t^{'} + R_t Q_t R_t^{'} \tag{12}$$

$$y_{t|t-1} = Z_t a_{t|t-1} + d_t \tag{13}$$

$$F_{t|t-1} = Z_t \Sigma_{t-1|t-1} Z_t^{'} + S_t H_t S_t^{'} \tag{14}$$

$$K_t = \Sigma_{t|t-1} Z_t^{'} F_{t|t-1}^{-1} \tag{15}$$

$$a_{t|t} = a_{t|t-1} + K_t (y_t - y_{t|t-1}) \tag{16}$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - K_t F_{t|t-1} K_t^{'} \tag{17}$$

$t = 1, 2, \cdots, T$ の逐次アルゴリズムになっている。ただし，初期条件は $a_{0|0} = a_0$，$\Sigma_{0|0} = \Sigma_0$ である。α_0 を非確率変数と仮定した場合は $\Sigma_0 = 0$ となる。

(11) 式と (12) 式は，(9) 式の予測方程式 $f(\alpha_t|Y_{t-1})$ から得られ，α_t の $(t-1)$ 期からみた予測に相当する。(13) 式と (14) 式は，(10) 式の更新方程式の分母にあたる $f(y_t|Y_{t-1})$ から得られる。これは，y_t の $(t-1)$ 期までの情報を与えたもとでの予測である。(16) 式と (17) 式は，(10) 式の更新方程式 $f(\alpha_t|Y_t)$ から得られる。このように，(11) 式 〜 (14) 式は予測方程式と解釈され，(16) 式と (17) 式はフィルタリング推定値を与え，過去の情報をもとに現在に入手されたデータで予測値を更新する働きを持つ。さらに，(15) 式の K_t はカルマン・ゲイン (Kalman gain) と呼ばれ，y_t の予測誤差 $(y_t - y_{t|t-1})$ とフィルタリング推定値 $a_{t|t}$ との共分散がゼロという条件を満たしている。この条件のことを,「y_t の予測誤差 $y_t - y_{t|t-1}$ とフィルタリング推定値 $a_{t|t}$ とは直交する」ともいう。カルマン・フィルタを線形最小分散推定量として解釈した場合には，アルゴリズムの導出過程の中で K_t は $a_{t|t}$ が最小分散になるように求められる。このことから，K_t はカルマン・ゲインと呼ばれる。詳しくは，Burridge and Wallis (1988)，谷崎 (1993)，片山 (2000) 等を参照せよ。

(11) 式 〜 (17) 式によって表されるアルゴリズムによると，まず初期値 $a_{0|0}$，$\Sigma_{0|0}$ が与えられると予測方程式 (11) 式 〜 (14) 式によって，$a_{1|0}$，$\Sigma_{1|0}$，$y_{1|0}$，$F_{1|0}$ が得られる。そして，カルマン・ゲイン K_1 を通して，更新方程式 (16) 式，(17) 式から，$a_{1|1}$，$\Sigma_{1|1}$ が得られる。このようにして，$a_{t-1|t-1}$，$\Sigma_{t-1|t-1}$ から，(11) 式 〜 (14) 式によって $a_{t|t-1}$，$\Sigma_{t|t-1}$，$y_{t|t-1}$，$F_{t|t-1}$ が，さらに，(16) 式と (17) 式によって $a_{t|t}$，$\Sigma_{t|t}$ が逐次計算によって得ら

れる。$y_{t|t-1}$，$F_{t|t-1}$ は，y_t の予測，予測誤差分散をそれぞれ表す。

このように，(11) 式 〜 (17) 式の逐次アルゴリズムをみると，カルマン・フィルタ・モデルとはある新しい観測値が利用可能となる度ごとに状態変数を推定しなおすというモデルである。(2) 式において，$T_t = I_k$，$c_t = 0$，$R_t = 0$ のとき，(11) 式から (17) 式で表されるモデルは，逐次最小自乗法 (recursive least squares) に一致する。ただし，I_k は $k \times k$ の単位行列とする。逐次最小自乗法とは，t 期 $(t = k, k+1, \cdots, T)$ までのデータを用いてパラメータを推定するという方法であり，すなわち，データが追加される度毎にその都度パラメータを最小自乗法で推定する方法である。$R_t \neq 0$ のとき，カルマン・フィルタ・モデルは，t 期の状態変数を推定するのに，t 期に近いデータに分散共分散のウエイト (weight) を小さく置き，t 期から過去に離れれば離れるほどそのウエイトを増大させて，推定値を求めるという方法であり，いわゆる，逐次一般化最小自乗法となっている。詳しくは，逐次最小自乗法については Harvey (1981，1989) を参照せよ。またカルマン・フィルタ・モデルと GLS の同値性については Sant (1977)，Chow (1983) を参照せよ。

このように初期値を与えなければならないこと (カルマン・フィルタでは状態変数の初期値とその分散が与えられると，逐次的な計算によって，それ以降のすべての期の状態変数の推定値を求めることができる) が問題となる。そして，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 (Anderson and Moore (1979)，片山 (2000)，Harvey (1989))。

・固定点 (fixed-point) スムージング:

$$a_{L|t} \qquad \Sigma_{L|t} \qquad t = L+1, L+2, \cdots$$

・固定ラグ (fixed-lag) スムージング:

$$a_{t|t+L} \qquad \Sigma_{t|t+L} \qquad t = 1, 2, \cdots, T-L$$

・固定区間 (fixed-interval) スムージング:

$$a_{t|T} \qquad \Sigma_{t|T} \qquad t = 1, 2, \cdots, T$$

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，ここ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，本節では以下の期待値について述べる。

$$\mathrm{E}(\alpha_t|Y_T) = a_{t|T} \qquad \mathrm{Var}(\alpha_t|Y_T) = \Sigma_{t|T} \qquad t = 1, 2, \cdots, T$$

3.1 節，3.2 節と同様に，まず初めに，分布に基づいたスムージング・アルゴリズムについて説明する。アルゴリズムは以下の通りである。

$$\begin{aligned} f(\alpha_t|Y_T) &= \int f(\alpha_{t+1}, \alpha_t|Y_T)\mathrm{d}\alpha_{t+1} = \int f(\alpha_{t+1}|Y_T) f(\alpha_t|\alpha_{t+1}, Y_T)\mathrm{d}\alpha_{t+1} \\ &= \int f(\alpha_{t+1}|Y_T) f(\alpha_t|\alpha_{t+1}, Y_t)\mathrm{d}\alpha_{t+1} = \int f(\alpha_{t+1}|Y_T) \frac{f(\alpha_t, \alpha_{t+1}|Y_t)}{f(\alpha_{t+1}|Y_t)}\mathrm{d}\alpha_{t+1} \\ &= f(\alpha_t|Y_t) \int \frac{f(\alpha_{t+1}|Y_T) f_\alpha(\alpha_{t+1}|\alpha_t)}{f(\alpha_{t+1}|Y_t)}\mathrm{d}\alpha_{t+1} \end{aligned} \tag{18}$$

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。以上の分布に基づいたスムージング・アルゴリズムについては，Kitagawa (1987)，Harvey (1989) 等を参照せよ。フィルタリング・アルゴリズムの (9) 式と (10) 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 $f_\alpha(\alpha_{t+1}|\alpha_t)$ をもとにして，スムージングの分布関数は (18) 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。ただし，固定区間スムージングの T 期の分布関数は，フィルタリングの T 期のものに一致することに注意せよ。固定区間スムージングは，(9) 式 と (10) 式によって表される分布関数に基づいたフィルタリング・アルゴリズムと共に用いられる。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 α_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。

線形の観測・遷移方程式と撹乱項の正規性の仮定のもとで，(18) 式から α_t の条件付き期待値と分散を計算することによって

$$C_t = \Sigma_{t|t} T_{t+1}' \Sigma_{t+1|t}^{-1} \tag{19}$$

$$a_{t|T} = a_{t|t} + C_t(a_{t+1|T} - a_{t+1|t}) \tag{20}$$

$$\Sigma_{t|T} = \Sigma_{t|t} + C_t(\Sigma_{t+1|T} - \Sigma_{t+1|t})C_t' \tag{21}$$

の固定区間スムージング・アルゴリズムが導かれる (証明は省略する)。これは $t = T-1, T-2, \cdots, 0$ の逆向きの逐次アルゴリズムとなっている。プレディクション $\Sigma_{t+1|t}$ とフィルタリング $\Sigma_{t|t}$ が与えられると，(19) 式から，C_t が計算される。同様に，$\Sigma_{t+1|t}$，$\Sigma_{t|t}$，C_t，$a_{t+1|t}$，$a_{t|t}$，$\Sigma_{t+1|T}$，$a_{t+1|T}$ から，

(20) 式と (21) 式を通して，$\Sigma_{t|T}$ と $a_{t|T}$ が得られる。計算手順としては，まず (11) 式 〜 (17) 式を用いて $a_{t|t-1}$，$\Sigma_{t|t-1}$，$a_{t|t}$，$\Sigma_{t|t}$，$t = 1, 2, \cdots, T$ を求めておく。次に (19) 式 〜 (21) 式によって，$a_{t|T}$，$\Sigma_{t|T}$，$t = T, T-1, \cdots, 1$ が得られる。このように，スムージング・アルゴリズムは，プレディクションとフィルタリングをもとにして，逆向きの逐次アルゴリズム (backward recursive algorithm) となっている。

最後に，遷移方程式について，$T_t = I_k$，$c_t = 0$，$R_t = 0$ の場合を考える。この場合，(11) 式 〜 (17) 式のカルマン・フィルタ・アルゴリズムは逐次最小自乗法に等しいことが知られていることは既に述べたが，(19) 式 〜 (21) 式で与えられるスムージング・アルゴリズムの場合は，すべての t について，状態変数の推定値 $a_{t|T}$ とその分散 $\Sigma_{t|T}$ は同じ値になり，その値は T 個のすべてのデータを使って最小自乗法 (OLS) によって推定された推定値に等しい。

本節では，プレディクション，フィルタリング，スムージング 3 つの推定問題について，分布関数の逐次アルゴリズム (6) 式，(9) 式と (10) 式，(18) 式と線形の逐次アルゴリズム (7) 式と (8) 式，(11) 式 〜 (17) 式，(19) 式 〜 (21) 式の 2 つのタイプのアルゴリズムをそれぞれ紹介した。3 つの推定問題について，分布関数の逐次アルゴリズムから，観測・遷移方程式の線形性と攪乱項の正規性の仮定を置いて計算すると，線形の逐次アルゴリズムを導出することができる。しかし，線形の逐次アルゴリズムを求めるためには正規性の仮定を必ずしも必要としない。詳しくは，次節のカルマン・フィルタ・モデルのアルゴリズムの導出方法を参照せよ。分布関数の近似による方法は非線形関数や非正規分布を取り扱うことができるので，最近の流れとしては，この分布関数による逐次アルゴリズムについて，分布関数自体を数値積分やモンテ・カルロ積分で近似する方法，または，分布関数から直接乱数を生成する方法が主流となっている。次節では，プレディクション，フィルタリング，スムージングの 3 つのアルゴリズムのうち，特にフィルタリングのみに焦点をあてて議論を進める。

4　カルマン・フィルタ・モデルの導出と解釈

3 節では，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をそれぞれ紹介した。本節では，(11) 式 〜 (17) 式によって表されるカルマン・フィルタ・モデルのアルゴリズムを導出のみを考える。プレディクション，スムージング・アルゴリズムの導出については，Jazwinski (1970)，Anderson and Moore (1979)，片山 (2000)，Harvey (1989) 等の他の文献に譲る。考え方の最も簡単な導出方法 (計算自体は複雑である) は，

(6) 式と (18) 式で表される分布関数のプレディクション，スムージング・アルゴリズムをもとにして，観測・遷移方程式の線形性と撹乱項の正規性を仮定すると，プレディクション (7) 式，(8) 式とスムージング (19) 式 ～ (21) 式の線形の逐次アルゴリズムがそれぞれ導き出される。導出方法は他にもいくつか考案され，それぞれはカルマン・フィルタ・モデルの解釈と密接に関連している。分布関数に基づいて導出する方法 (Anderson and Moore (1979)，Kitagawa (1987)，Harvey (1989))，混合推定による導出 (Cooley (1977)，Harvey (1981)，Diderrich (1985)，Fomby *et al.* (1984))，線形最小分散推定量 (片山 (2000)，Burridge and Wallis (1988)) としての解釈がある。分布関数に基づく導出は線形性と正規性の両方の仮定を必要とするが，混合推定や線形最小分散推定量としての導出には線形性は必要な仮定であるが正規性の仮定は必要でない。混合推定量，線形最小分散推定量としての導出については，単に 1，2 次の積率を用いてアルゴリズムを導き出すことができる。その他にも，直行斜影の利用 (Anderson and Moore (1979)，片山 (2000)，Chow (1983)，Brockwell and Davis (1987))，一般化最小自乗法 (Sant (1977)，Chow (1983)) による証明等が考えられる。このように，同じフィルタリング・アルゴリズムが得られるにしても，導出方法は様々である。

本節では，正規分布を仮定して，(9) 式と (10) 式から，カルマン・フィルタ・モデルのアルゴリズム (11) 式 ～ (17) 式を導出する。もし，初期分布 $f(\alpha_0|Y_0)$ と撹乱項 ϵ_t，η_t が正規分布なら，$f(\alpha_t|Y_{t-1})$，$f(\alpha_t|Y_t)$ も正規分布となる。ゆえに，まず $f(\alpha_{t-1}|Y_{t-1})$ を平均 $a_{t-1|t-1}$，分散 $\Sigma_{t-1|t-1}$ の正規分布とする。$k\times 1$ の次元の確率変数 x が平均 μ，分散 Σ の正規分布 $f(x)$ に従うとするとき，$f(x) = \Phi(x-\mu,\ \Sigma)$ と書くことにする。すなわち，

$$\Phi(x-\mu,\Sigma) = (2\pi)^{-\frac{k}{2}}|\Sigma|^{-\frac{1}{2}}\exp(-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu))$$

と定義する。このとき，条件付き分布 $f(\alpha_{t-1}|Y_{t-1})$ は，次のように書き表される。

$$f(\alpha_{t-1}|Y_{t-1}) = \Phi(\alpha_{t-1} - a_{t-1|t-1}, \Sigma_{t-1|t-1})$$

同様に，α_t の分布 $f_\alpha(\alpha_t|\alpha_{t-1})$ と $f(\alpha_t|Y_{t-1})$ は，それぞれ，

$$f_\alpha(\alpha_t|\alpha_{t-1}) = \Phi(\alpha_t - T_t\alpha_{t-1} - c_t, R_tQ_tR'_{t-1})$$
$$f(\alpha_t|Y_{t-1}) = \Phi(\alpha_t - a_{t|t-1}, \Sigma_{t|1-1})$$

となる。一方，分布関数による一期先の予測方程式は上述の (9) 式で表されるので，

$$f(\alpha_t|Y_{t-1}) = \Phi(\alpha_t - a_{t|t-1}, \Sigma_{t|t-1}) = \int f_\alpha(\alpha_t|\alpha_{t-1})f(\alpha_{t-1}|Y_{t-1})\mathrm{d}\alpha_{t-1}$$

$$
\begin{aligned}
&= \int \Phi(\alpha_t - T_t\alpha_{t-1} - c_t, R_tQ_tR_t')\, \Phi(\alpha_{t-1} - a_{t-1|t-1}, \Sigma_{t-1|t-1})\mathrm{d}\alpha_{t-1} \\
&= \Phi(\alpha_t - T_t a_{t-1|t-1} - c_t,\ T_t\Sigma_{t-1|t-1}T_t' + R_tQ_tR_t')
\end{aligned}
$$

を得る。1 つ目の等号は定義による。読者は 4 つ目の等式が成り立つことを各自確認められたい。1 行目と 3 行目のそれぞれの $\Phi(\cdot,\cdot)$ の要素をそれぞれ比較して，(11) 式，(12) 式の予測方程式が得られる。

更新方程式については，(10) 式から，

$$
\begin{aligned}
f(\alpha_t|Y_t) = \Phi(\alpha_t - a_{t|t}, \Sigma_{t|t}) &= \frac{f_y(y_t|\alpha_t)f(\alpha_t|Y_{t-1})}{\int f_y(y_t|\alpha_t)f(\alpha_t|Y_{t-1})\mathrm{d}\alpha_t} \\
&= \frac{\Phi(y_t - Z_t\alpha_t - d_t, S_tH_tS_t')\, \Phi(\alpha_t - a_{t|t-1}, \Sigma_{t|t-1})}{\int \Phi(y_t - Z_t\alpha_t - d_t, S_tH_tS_t')\, \Phi(\alpha_t - a_{t|t-1}, \Sigma_{t|t-1})\mathrm{d}\alpha_t} \\
&= \Phi(\alpha_t - a_{t|t-1} - K_t(y_t - y_{t|t-1}),\ \Sigma_{t|t-1} - K_tF_{t|t-1}K_t')
\end{aligned}
$$

が得られる。1 つ目の等号は定義であり，それ以降は,

$$
\begin{aligned}
&\int f(y_t|\alpha_t)f(\alpha_t|Y_{t-1})\mathrm{d}\alpha_t = f(y_t|Y_{t-1}) = \Phi(y_t - y_{t|t-1}, F_{t|t-1}) \\
&f_y(y_t|\alpha_t)f(\alpha_t|Y_{t-1}) = \Phi(y_t - Z_t\alpha_t - d_t, S_tH_tS_t')\, \Phi(\alpha_t - a_{t|t-1}, \Sigma_{t|t-1}) \\
&\quad = \Phi(\alpha_t - a_{t|t-1} - K_t(y_t - y_{t|t-1}),\ \Sigma_{t|t-1} - K_tF_{t|t-1}K_t') \\
&\qquad \times\Phi(y_t - y_{t|t-1}, F_{t|t-1}) \qquad (22)
\end{aligned}
$$

が用いられている (読者は (22) 式の 2 つ目の等式が成立することを各自確認すること)。ただし，$y_{t|t-1}$，$F_{t|t-1}$，K_t は (13) 式 〜 (15) 式である。同様に，更新方程式についても，平均と分散をそれぞれ比較することにより，(16) 式と (17) 式が得られる。

以上のように，カルマン・フィルタの線形の逐次アルゴリズムにのみ焦点をあててその導出方法を考察した。本節では正規分布に基づいて導出を行ったが，分布には依存しない方法 (線形最小分散推定量や混合推定量に基づく方法) でも証明は可能である。また，線形最小分散推定量としての導出については，(1) 式と (2) 式の状態空間モデルの撹乱項 ϵ_t，η_t が互いに相関がある場合にでも，簡単にフィルタリング・アルゴリズムを求めることができる。撹乱項に相関がある場合のフィルタリング・アルゴリズムについては，Burridge and Wallis (1988)，Harvey (1989) を参照せよ。

次節では，未知パラメータ θ が状態空間モデルに含まれている場合の θ の推定問題について考える。すなわち，(1) 式と (2) 式の Z_t，d_t，S_t，H_t，T_t，c_t，R_t，Q_t が θ の関数である場合は，θ の推定値を使って，状態変数

の推定値 $a_{t+L|t}$，$a_{t|t}$，$a_{t|T}$ を対応するアルゴリズムから求めなければならない。

5 最尤法による未知パラメータの推定

前節では，観測・遷移方程式の中に未知パラメータが含まれていない状況を考えた。しかし，通常の推定には大抵の場合，方程式に未知パラメータが含まれる。本節では，(1) 式と (2) 式の観測・遷移方程式に未知パラメータ (例えば，θ) が含まれる場合，どのようにして未知パラメータと状態変数を推定するのかを考える。

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ので，この場合，撹乱項の分布関数の仮定を必要とする。4 節で述べたように，正規分布を仮定した場合は通常の線形の逐次アルゴリズムが導かれる。次の尤度関数はイノベーション・フォーム (innovation form) と呼ばれる。

$$\begin{aligned} f(Y_T) &= f(y_T|Y_{T-1})f(y_{T-1}|Y_{T-2})\cdots f(y_2|y_1)f(y_1) \\ &= \prod_{t=1}^{T} f(y_t|Y_{t-1}) \end{aligned} \tag{23}$$

ここで $f(y_1) = f(y_1|Y_0)$ とみなされる。$f(y_t|Y_{t-1})$ は (10) 式の分母に当たり次のように表される。

$$f(y_t|Y_{t-1}) = \int f_y(y_t|\alpha_t)f(\alpha_t|Y_{t-1})\mathrm{d}\alpha_t \tag{24}$$

もし観測・遷移方程式が線形で撹乱項が正規分布ならば，$f(y_t|Y_{t-1})$ は平均 (13)，分散 (14) の正規分布に従う。すなわち，(24) 式は，

$$\begin{aligned} f(y_t|Y_{t-1}) &= \Phi(y_t - y_{t|t-1}, F_{t|t-1}) \\ &= (2\pi)^{-\frac{g}{2}}|F_{t|t-1}|^{-\frac{1}{2}} \exp(-\frac{1}{2}(y_t - y_{t|t-1})'F_{t|t-1}^{-1}(y_t - y_{t|t-1})) \end{aligned}$$

となる。よって，線形性・正規性の仮定のもとで (23) 式のイノベーション・フォームによる対数尤度関数は次のように表される。

$$\begin{aligned} \log f(Y_T) &= \sum_{t=1}^{T} \log f(y_t|Y_{t-1}) = \sum_{t=1}^{T} \log \Phi(y_t - y_{t|t-1}, F_{t|t-1}) \\ &= -\frac{Tg}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log|F_{t|t-1}| \end{aligned}$$

$$-\frac{1}{2}\sum_{t=1}^{T}(y_t-y_{t|t-1})'F_{t|t-1}^{-1}(y_t-y_{t|t-1}) \tag{25}$$

Z_t，d_t，S_t，H_t，T_t，c_t，R_t，Q_t が未知パラメータ θ に依存しているとき，(25) 式の尤度関数が θ について最大化される (尤度関数 $f(Y_T)$ は未知パラメータ θ の関数であるので，$f(Y_T) \equiv f(Y_T;\theta)$ である)。明示的に θ の推定値が得られることは稀なので，単純サーチ法 (simple grid search)，スコアリング法 (scoring method) 等の方法で尤度関数が最大化される。最尤法で求められた未知パラメータの推定値を所与として，それぞれのアルゴリズムに代入し状態変数の推定値が求められるのである。この最大化方法の際には，収束計算が用いられる。すなわち，最初に適当な値を未知パラメータに与えておいて，(11) 式 〜 (17) 式のアルゴリズムから尤度関数の値を求める。次に，得られたプレディクション，フィルタリング推定値 $a_{t|t}$，$\Sigma_{t|t}$，$a_{t|t-1}$，$\Sigma_{t|t-1}$ を与えたもとで，尤度関数を最大にする未知パラメータの推定値を求める。この過程が繰り返され，未知パラメータの推定値の値が安定するまで続けられる。

このイノベーション・フォームを利用した最尤法によると，最大化に必要な $y_{t|t-1}$，$F_{t|t-1}$ は (11) 〜 (17) のカルマン・フィルタ・アルゴリズムの中で計算される。そこでは追加的な計算を必要としない。

このイノベーション・フォームの尤度最大化の他に，EM (Expectation-Maximization) アルゴリズムと呼ばれる方法で尤度関数を最大化する方法もある。この最大化の方法は，すべてのデータ $Y_T = \{y_1, y_2, \cdots, y_T\}$ を与えたもとで，対数尤度関数の期待値が最大にされる方法である。観測されない状態変数 α_t，$t = 1, 2, \cdots, T$，は，すべての情報 Y_T を与えたもとでの条件付き期待値 (すなわち，$a_{t|T}$，$\Sigma_{t|T}$) に置き換えられる。このように，EM アルゴリズムは対数尤度関数の最大化にスムージング推定値を必要とする。そのため，イノベーション・フォームを用いる方法よりも，余分の計算 (すなわち，(19) 式 〜 (21) 式のスムージング・アルゴリズム) が必要になる。イノベーション・フォームの尤度最大化によると未知パラメータの解は明示的には得られず単純サーチ法等によって行われるが，このEM アルゴリズムでは推定値はスムージング推定値の関数として明示的に得られる場合がある。一般に，EM アルゴリズムは真のパラメータの近傍をすばやく探し出すことができるが，収束は非常に遅いという欠点を持つ。Shumway and Stoffer (1982)，Watson and Engle (1983)，Tanizaki (1989) は状態空間モデルに EM アルゴリズムを応用した。また一般的な EM アルゴリズムの文献としては，Dempster *et al.* (1977)，Rund (1991) があげられる。

以上，これまでは，状態空間モデルの定義 (2 節)，状態変数の推定問

題 (3 節)，フィルタリング・アルゴリズムの導出 (4 節)，加えて，未知パラメータの推定 (5 節) について解説した。状態空間モデルを実際に適用する場合に問題となるのは，初期値 a_0，Σ_0 の扱いと未知パラメータ θ の推定である。フィルタリング・アルゴリズムについては，未知パラメータが状態空間モデルの中に含まれていないとき，初期値が与えられるとそれ以降のすべての期の状態変数が計算されるアルゴリズムになっている。また，未知パラメータの推定については，θ の推定値を明示的な形で得ることが難しい。したがって，(23) 式を最大化するには単純サーチ法等の方法に頼らざるをえない。未知パラメータの数が増えるとこの方法も難しくなる。また，スコアリング法も尤度関数を最大にするのに使われるが，この方法は尤度関数の 2 階微分を必要とするので時には複雑になる。

次節では，非線形・非正規分布のもとでのプレディクション，フィルタリング，スムージングの導出方法について述べる。

6　非線形・非正規状態空間モデル

モデルが非線形，または，撹乱項が非正規分布の場合には，プレディクション (7) 式と (8) 式，フィルタリング (11) 式 〜 (17) 式，スムージング (19) 式 〜 (21) 式のように，線形の逐次アルゴリズムを求めることは一般的には不可能である。

非線形・非正規分布の状態空間モデルでは，(1) 式，(2) 式は次のように書き直される。

$$\text{(観測方程式)} \qquad y_t = h_t(\alpha_t, \epsilon_t) \tag{26}$$

$$\text{(遷移方程式)} \qquad \alpha_t = q_t(\alpha_{t-1}, \eta_t) \tag{27}$$

(1) 式，(2) 式に含ま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。(4) 式，(5) 式等の期待値を得ることが最終的な目的であるので，α_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。

6.1　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。

以上のように，RS は $\omega(x)$ が有限でなければ使えないということが重要であり，時にはこの条件を満たさない場合もありうる。また，たとえ $\sup_z \omega(z)$ が有限であったとしても，$\sum_z \omega(z)$ が非常に大きな値をとる場合，受容確率 $w(x)$ はほとんどの x でゼロに近くなり，(i) と (ii) の繰り返し回数 N_R が非常に多くなり，乱数生成のための計算時間が膨大になる。

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，$u \leq w(x_{i-1}, x^*)$ 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。

x_{i-1} と x_i との間の正の相関をなくすための方法として，$N = 1$ として (i) 〜 (iv) を実行する。別の初期値 x_{-M} を取って，N 回繰り返す。これにとると，ターゲット密度関数 $f(x)$ から一つの乱数を生成するのに，サンプリング密度関数 $f_*(x)$ から $M+1$ 個の乱数を使うことになる。すなわち，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，サンプリング密度関数 $f_*(x)$ から $N(M+1)$ 個の乱数を生成する。このようにすると，N 個の互いに独立な乱数を作ることができる。乱数の精度としては，M の大きさによって，この方法は RS と同程度の精度が得られる。しかし，計算の負担が極端に多くなる。しかも，$M > N_R$ であれば RS よりも計算面での効率は悪くなる。したがって，この方法はあまり実際には用いられない。

6.2 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 α_t の i 番目の乱数とする。もし乱数 $\alpha_{1,t|s}$，$\alpha_{2,t|s}$，$\cdots$，$\alpha_{N,t|s}$ ($s = t, T$，$t = 1, 2, \cdots, T$) が生成されたとすると，(4) 式，(5) 式は $\mathrm{E}(\alpha_t|Y_s) \equiv \overline{a}_{t|s} \approx (1/N)\sum_{i=1}^{N}\alpha_{i,t|s}$，$\mathrm{Var}(\alpha_t|Y_s) \approx (1/N)\sum_{i=1}^{N}(\alpha_{i,t|s} - \overline{a}_{t|s})(\alpha_{i,t|s} - \overline{a}_{t|s})'$ として求められる。

とりあえず，$\alpha_{i,t-1|t-1}$，$i = 1, 2, \cdots, N$ 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 α_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 (10) 式について，2 通りの表し方がある。一つは，(10) 式から，$f(\alpha_t|Y_t)$ は次のように表される。

$$f(\alpha_t|Y_t) \propto f_y(y_t|\alpha_t)f(\alpha_t|Y_{t-1}) \propto \omega_1(\alpha_t)f(\alpha_t|Y_{t-1}) \tag{28}$$

ただし，$\omega_1(\alpha_t) \propto f_y(y_t|\alpha_t)$ とする。この場合，6.1 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 (26) 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 α_t の乱数は，遷移方程式 (27) 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) α_0 を確率変数と仮定するなら $f_\alpha(\alpha_0)$ から α_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，α_0 を非確率変数と仮定するなら α_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 α_t の乱数 (α_t^* とする) を生成する。α_t^* は $f(\alpha_t|Y_t)$ から生成される α_t の i 番目の乱数 ($\alpha_{i,t|t}$ とする) の候補となる。

(iii) α_t^* に基づいて，$f(\alpha_t|Y_t)$ から生成された α_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，α_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，η_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，(28) 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 (28) 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，(28) 式に基づいた IR や MH による $f(\alpha_t|Y_t)$ から α_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，η_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 α_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 α_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 (9) 式をフィルタリング密度関数 (28) 式に代入して，次のように変形される。

$$\begin{aligned} f(\alpha_t|Y_t) &\propto f_y(y_t|\alpha_t)f(\alpha_t|Y_{t-1}) = \int f_y(y_t|\alpha_t)f_\alpha(\alpha_t|\alpha_{t-1})f(\alpha_{t-1}|Y_{t-1})\,\mathrm{d}\alpha_{t-1} \\ &\propto \int f(\alpha_t,\alpha_{t-1}|Y_t)\,\mathrm{d}\alpha_{t-1} \end{aligned}$$

α_{t-1} に関する積分を取り除くと，Y_t を与えたもとで α_t と α_{t-1} の条件付密度関数 $f(\alpha_t,\alpha_{t-1}|Y_t)$ は以下のように書き直される。

$$\begin{aligned} f(\alpha_t,\alpha_{t-1}|Y_t) &\propto f_y(y_t|\alpha_t)f_\alpha(\alpha_t|\alpha_{t-1})f(\alpha_{t-1}|Y_{t-1})\,\mathrm{d}\alpha_{t-1} \\ &\propto \omega_2(\alpha_t,\alpha_{t-1})f_*(\alpha_t|\alpha_{t-1})f(\alpha_{t-1}|Y_{t-1}) \end{aligned} \tag{29}$$

ただし，$\omega_2(\alpha_t,\alpha_{t-1}) \propto f_y(y_t|\alpha_t)f_\alpha(\alpha_t|\alpha_{t-1})/f_*(\alpha_t|\alpha_{t-1})$ である。

(29) 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 α_{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 α_{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 α_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，(α_t,α_{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 (α_t,α_{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 (α_t,α_{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 α_t の乱数は，$f(\alpha_t|Y_t)$ から生成される α_t の乱数と同じ性質を

持つ。$f_*(\alpha_t|\alpha_{t-1})$ を，α_{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，α_t と α_{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。(29) 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，(29) 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 (28) 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 (α_t,α_{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 α_{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。

6.3 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。

(18) 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 (30) 式，(32) 式，(33) 式の 3 種類) がある。(18) 式から，α_{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。

$$f(\alpha_{t+1},\alpha_t|Y_T) \propto \omega_3(\alpha_{t+1},\alpha_t)f(\alpha_t|Y_t)f(\alpha_{t+1}|Y_T) \tag{30}$$

ただし，$\omega_3(\alpha_{t+1},\alpha_t) \propto f_\alpha(\alpha_{t+1}|\alpha_t)/f(\alpha_{t+1}|Y_t)$ である。(30) 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 (6) 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。

$$f(\alpha_{t+1}|Y_t) = \int f_\alpha(\alpha_{t+1}|\alpha_t)f(\alpha_t|Y_t)\,\mathrm{d}\alpha_t \approx \frac{1}{N'}\sum_{j=1}^{N'} f_\alpha(\alpha_{t+1}|\alpha_{j,t|t}) \tag{31}$$

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，(31) 式の積分を計算すればよい。(31) 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。

(30) 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。α_t の乱数は $f(\alpha_t|Y_t)$ から生成され，α_{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 α_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 α_{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，(α_t, α_{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 (α_t, α_{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 (α_t, α_{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 (α_t, α_{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。α_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 α_t^* とする。α_{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 α_{t+1}^* とする。

さらに，(30) 式の $f(\alpha_t|Y_t)$ を (28 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。

$$f(\alpha_{t+1}, \alpha_t|Y_T) \propto \omega_4(\alpha_{t+1}, \alpha_t) f(\alpha_t|Y_{t-1}) f(\alpha_{t+1}|Y_T) \tag{32}$$

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。(32) 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。(30) 式と (32) 式との違いは α_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。(30) 式や (32) 式によって，$f(\alpha_{t+1}, \alpha_t|Y_T)$ から (α_{t+1}, α_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 (30) 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 (α_t, α_{t+1}) の乱数 (α_t^*, α_{t+1}^*) を生成する。

α_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 η_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 α_{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，(30) 式では $f(\alpha_t|Y_t)$, (32) 式では $f(\alpha_t|Y_{t-1})$ を α_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，α_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。(32) 式に (9) 式を代入して，α_{t-1} に関する積分を取り除くと，Y_T を与えたもとでの (α_{t+1}, α_t,α_{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。

$$f(\alpha_{t+1}, \alpha_t, \alpha_{t-1}|Y_T) \propto \omega_5(\alpha_{t+1}, \alpha_t, \alpha_{t-1})f(\alpha_{t-1}|Y_{t-1})f_*(\alpha_t|\alpha_{t-1}, \alpha_{t+1})f(\alpha_{t+1}|Y_T) \tag{33}$$

ただし，$\omega_5(\alpha_{t+1}, \alpha_t, \alpha_{t-1}) \propto \omega_4(\alpha_{t+1}, \alpha_t)f_\alpha(\alpha_t|\alpha_{t-1})/f_*(\alpha_t|\alpha_{t-1}, \alpha_{t+1})$ である。(33) 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 α_{t+1} と α_{t-1} の乱数が独立に生成されて後，α_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 (30) 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 (33) 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 α_t のサンプリング密度関数の一つの候補となりうる。

(28) 式，(29) 式はフィルタリングであり，(30) 式，(32) 式，(33) 式はスムージングに対応する。これらの式をもとにして，α_t の乱数が生成される。6.1 節の $f(x)$，$f_*(x)$，$\omega(x)$ との対応を表 1 にまとめておく。ただし，x は確率変数，$f(x)$ はターゲット密度関数，$f_*(x)$ はサンプリング密度関数，$\omega(x)$ はターゲット密度関数とサンプリング密度関数の比で x に関わる部分をそれぞれ表すものとする。

(28) 式に基づいた IR フィルタは Gordon *et al.* (1993)，Kitagawa (1996, 1998)，Kitagawa and Gersch (1996) によって提案された。(28) 式をもとにして RS フィルタを Tanizaki (1996，1999，2001a)，Tanizaki and Mariano (1998)，Hürzeler and Künsch (1998) は提案した。(28) 式に基づく MH フィルタ，(29) 式に基づく IR，RS，MH フィルタ，(30) 式，(32) 式，(33) 式をもとにした IR，RS，MH スムーザ (smooter) は Tanizaki (2001a，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，ω_2，ω_3，ω_4，ω_5 の上限を求めることは難しい。それでも，α_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，ω_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，(29) 式，(30) 式，(32) 式，(33) 式よりむしろ，フィルタリングの (28) 式に限って用いる方がよい。

それぞれの乱数生成方法について，各時点での必要な乱数の平均生成数が表 2 に示されている。N_R と M はターゲット密度関数とサンプリング密度関数の関数形に依存する (N_R は RS における一つの乱数を生成するために必要な平均棄却回数を表し，M は MH における分析には用いられない最初の乱数の個数を表す)。IR，RS，MH について，計算時間のオーダーが表 2 に示されている。しかし，IR については，表 2 の値を N 倍して，計算時間のオーダーはフィルタリングでは N^2，スムージングでは $N^2 \times N'$ とされるべきである。なぜなら，IR では，一つの乱数を生成

表 1: (28) 式 ∼ (30) 式，(32) 式，(33) 式と x，$f(x)$，$f_*(x)$，$\omega(x)$ との関係

	確率変数 x	ターゲット密度 $f(x)$	サンプリング密度 $f_*(x)$	$\omega(x)$
(28) 式	α_t	$f(\alpha_t\vert Y_t)$	$f(\alpha_t\vert Y_{t-1})$	$f_y(y_t\vert\alpha_t)$
(29) 式	(α_t, α_{t-1})	$f(\alpha_t, \alpha_{t-1}\vert Y_t)$	$f_*(\alpha_t\vert\alpha_{t-1})f(\alpha_{t-1}\vert Y_{t-1})$	$\dfrac{f_y(y_t\vert\alpha_t)f_\alpha(\alpha_t\vert\alpha_{t-1})}{f_*(\alpha_t\vert\alpha_{t-1})}$
(30) 式	(α_{t+1}, α_t)	$f(\alpha_{t+1}, \alpha_t\vert Y_T)$	$f(\alpha_{t+1}\vert Y_T)f(\alpha_t\vert Y_{t-1})$	$\dfrac{f_\alpha(\alpha_{t+1}\vert\alpha_t)}{f(\alpha_{t+1}\vert Y_t)}$
(32) 式	(α_{t+1}, α_t)	$f(\alpha_{t+1}, \alpha_t\vert Y_T)$	$f(\alpha_{t+1}\vert Y_T)f(\alpha_t\vert Y_t)$	$\dfrac{f_y(y_t\vert\alpha_t)f_\alpha(\alpha_{t+1}\vert\alpha_t)}{f(\alpha_{t+1}\vert Y_t)}$
(33) 式	$(\alpha_{t+1}, \alpha_t, \alpha_{t-1})$	$f(\alpha_{t+1}, \alpha_t, \alpha_{t-1}\vert Y_T)$	$f(\alpha_{t+1}\vert Y_T)f_*(\alpha_t\vert\alpha_{t-1}, \alpha_{t+1})f(\alpha_{t-1}\vert Y_{t-1})$	$\dfrac{f_y(y_t\vert\alpha_t)f_\alpha(\alpha_t\vert\alpha_{t-1})f_\alpha(\alpha_{t+1}\vert\alpha_t)}{f_*(\alpha_t\vert\alpha_{t-1}, \alpha_{t+1})f(\alpha_{t+1}\vert Y_t)}$

表 2: t 時点における乱数の生成数

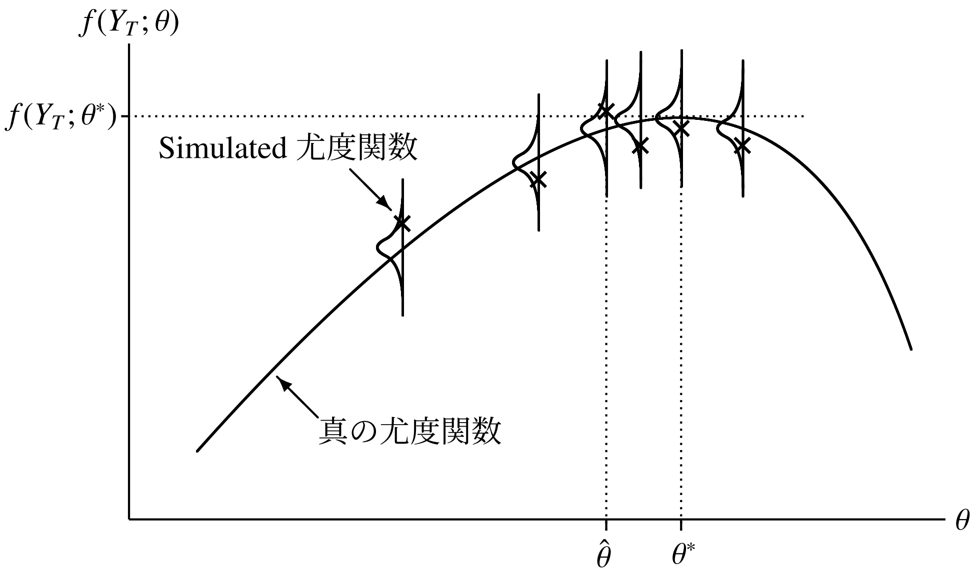
サンプリング 方法	フィルタリング (28) 式，(29) 式	スムージング (30) 式，(32) 式，(33) 式
RS	$N(1 + N_R)$	$N(1 + N_R) \times N'$
IR	N	$N \times N'$
MH	$N + M$	$(N + M) \times N'$

するためには，N 個のウエイトから一つを選ぶことになるので，N 倍の繰り返しを要するからである (これはプログラミングの工夫によって繰り返し回数を減らすことが出来るので，一概には，N 倍になるとは言えない)。

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，(28) 式に基づいて RS をフィルタリングでは使い，(30) 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 (28) 式，(29) 式を併用したり，スムージングについても (30) 式，(32) 式，(33) 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 (28) 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，(29) 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 (29) 式を使い，それ以外の時点では (28) 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 α_t の乱数生成が難しいとき，すなわち，η_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 $f_*(\alpha_t|\alpha_{t-1})$ を利用して α_t の乱数を生成することができる。

図 1: 真の尤度関数と Simulated 尤度関数



6.4　最尤法によるパラメータ推定

(26) 式，(27) 式に未知パラメータ (θ とする) が含まれている場合を考える。尤度関数 (23) 式は，(24) 式に注意すると，

$$
\begin{aligned}
f(Y_T) \equiv f(Y_T;\theta) &= \prod_{t=1}^{T} \int f_y(y_t|\alpha_t) f(\alpha_t|Y_{t-1}) \mathrm{d}\alpha_t \\
&\approx \prod_{t=1}^{T} \Big(\frac{1}{N} \sum_{i=1}^{N} f_y(y_t|\alpha_{i,t|t-1}) \Big) \qquad (34)
\end{aligned}
$$

と近似できる。$\eta_{i,t}$ は η_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 α_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。(34) 式は Simulated 尤度関数と呼ばれ，θ について最大化される。最大化の方法は，4 節で解説した線形・正規性の場合と同様に，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，θ 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 である。データ y_t を与えたもとで，真の尤度関数を最大にする θ を θ^* とする。図中の × で表された値は，θ 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 θ 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，θ^* 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 θ 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，θ 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 θ の密度関数に比例するとみなして θ 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 Geweke and Tanizaki (1999，2001)，Jacquier *et al.* (1994)，Watanabe (2000)，Watanabe and Omori (2004)，渡部 (2000) を参照せよ。

参考文献

外国語文献

Anderson, B.D.O. and Moore, J.B. (1979). *Optimal Filtering*, Prentice-Hall, New York.

Aoki, M. (1987). *State Space Modeling of Time Series*, Springer-Verlag.

Belsley, D.A. and Kuh, E. (1973). Time-Varying Parameter Structures: An Overview, *Annals of Economic and Social Measurement*, 2, pp.375-379.

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.

Brockwell, P.A. and Davis, R.A. (1987). *Time Series Theory and Models*, Springer-Verlag.

Burmeister, E. and Wall, K.D. (1982). Kalman Filtering Estimation of Unobserved Rational Expectations with an Application to the German Hyperinflation, *Journal of Econometrics*, Vol.20, pp.255-284.

Burridge, P. and Wallis, K.F. (1988). Prediction Theory for Autoregressive Moving Average Processes, *Econometric Reviews*, 7 (1). pp.65-95.

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

Chow, G.C. (1983). *Econometrics*, McGraw-Hill Book Company.

Conrad, W. and Corrado, C. (1979). Application of the Kalman Filter to Revisions in Monthly Retail Sales Estimates, *Journal of Economic Dynamic and Control*, Vol.1, pp.177-198.

Cooley, T.F. (1977). Generalized Least Squares Applied to Time Varying Parameter Models: A Comment, *Annals of Economic and Social Measurement*, 6/3, pp.313-314.

Cooley, T.F. and Prescott, E.C. (1973). Varying Parameter Regression, A Theory and Some Applications, *Annals of Economic and Social Measurement*, Vol.2, pp.463-474.

Cooley, T.F. and Prescott, E.C. (1976). Estimation in the presence of stochastic parameter variation, *Econometrica*, Vol.44, pp.167-183.

Cooper, J.P. (1973). Time-Varying Regression Coefficients: A Mixed Estimation Approach and Operational Limitations of the General Markov Structure, *Annals of Economic and Social Measurement*, 2/4, pp.525-530.

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society*, Ser.B, Vol.39, pp.1-38 (with discussiion) .

Diderrich, G.T. (1985). The Kalman Filter from the Perspective of Goldberger-Theil Estimators, *The American Statistician*, Vol.39, pp.193-198.

Dziechciarz, J. (1989). Changing and Random Coefficient Models. A Survey, in *Statistical Analysis and Forecasting of Economic Structural Change*, edited by P. Hackl, Springer-Verlag.

Fomby, T.B., Hill, R.C. and Johnson, S.R. (1984). *Advanced Econometric Methods*, Springer-Verlag.

Garbade, K. (1977). Two Methods for Examining the Stability of Regression Coefficients, *Journal of the American Statistical Association*, Vol.72, pp.54-63.

Gardner, G., Harvey, A.C. and Phillips, G.D.A. (1980). An Algorithm for Maximum Likelihood Estimation Autoregressive-Moving Average Models by means of Kalman Filtering, *Applied Statistics*, Vol.29, No.3, pp.311-322.

Gelb, A. (Ed.) (1974). *Applied Optimal Estimation*, MIT Press.

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. (1999). On Markov Chain Monte Carlo Methods for Nonlinear and Non-Gaussian State-Space Models, *Communications in Statistics, Simulation and Computation*, Vol.28, No.4, pp.867-894.

Geweke, J. and Tanizaki, H. (2001). Bayesian Estimation of State-Space Model Using the Metropolis-Hastings Algorithm within Gibbs Sampling, *Computational Statistics and Data Analysis*, Vol.37, No.2, pp.151-170.

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.

Hamilton, J.D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica*, Vol.57, pp.357-384.

Hamilton, J.D. (1990). Analysis of Time Series Subject to Changes in Regime, *Journal of Econometrics*, Vol.45, pp.39-70.

Hamilton, J.D. (1991). A Quasi-Bayesian Approach to Estimating Parameters for Mixtures of Normal Distributions, *Journal of Business and Economic Statistics*, Vol.9, pp.27-39.

Hamilton, J.D. (1993). Estimation, Inference and Forecasting of Time Series Subject to Changes in Regime, in *Handbook of Statistics, Vol.11*, edited by Maddala, G.S., Rao, C.R. and Vinod, H.D., pp.231-260, North-Holland.

Hamilton, J.D. (1994). *Time Series Analysis*, Princeton University Press.

Harvey, A.C. (1981). *Time Series Models*, Philip Allen Publishers Limited, Oxford.

Harvey, A.C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press.

Howrey, E.P. (1978). The Use of Preliminary Data in Econometric Forecasting, *The Review of Economics and Statistics*, Vol.60, pp.193-200.

Howrey, E.P. (1984). Data Revision, Reconstruction, and Prediction: An Application to Inventory Investment, *The Review of Economics and Statistics*, Vol.66, pp.386-393.

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).

Jacquier, E., Polson, N.G. and Rossi, P.E. (2004). Bayesian Analysis of Stochastic Volatility Models with Fat-tails and Correlated Errors, *Journal of Econometrics*, Vol.122, pp.185-212.

Jazwinski, A.H. (1970). *Stochastic Processes and Filtering Theory*, Academic Press, New York.

Kalman, R.E. (1960). A New Approach to Linear Filtering and Prediction Problems, *Journal of Basic Engineering, Transactions ASME*, Series D, 82, pp.35-45.

Kalman, R.E. and Bucy, R.S. (1961). New Results in Linear Filtering and Prediction Theory, *Journal of Basic Engineering, Transactions ASME*, Series D, 83, pp. 95-108.

Kitagawa, G. (1987). Non-Gaussian State-Space Modeling of Nonstationary Time Series, *Journal of the American Statistical Association*, Vol.82, pp.1032-1063 (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.

Laumas, G.S. and Mehra, Y.P. (1976). The Stability of the Demand for Money Function: The Evidence from Quarterly Data, *The Review of Economics and Statistics*, Vol.58, pp.464-468.

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

Mariano, R.S. and Tanizaki, H. (1995). Prediction of Final Data with Use of Preliminary and/or Revised Data " *Journal of Forecasting*, Vol.14, No.4, pp.351 - 380.

McNelis, P.D. and Neftci, S.N. (1983). Policy-dependent Parameters in the Presence of Optimal Learning: An Application of Kalman Filtering to the Fair and Sargent Supply-side Equations, *The Review of Economics and Statistics*, Vol.65, pp.296-306.

Pagan, A.R. (1975). A Note on the Extraction of Components from Time Series, *Econometrica*, Vol.43, pp.163-168.

Pagan, A.R. (1980). Some Identification and Estimation Results for Regression Models with Stochastically Varying, *Journal of Econometrics*, Vol.13, pp.343-363.

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

Rund, P.A. (1991). Extensions of Estimation Methods Using the EM Algorithm, *Journal of Econometrics*, Vol.49, pp.305-341.

Sant, D.T. (1977). Generalized Least Squares Applied to Time Varying Parameter Models, *Annals of Economic and Measurement*, 6/3, pp.301-311.

Sarris, A.H. (1973). A Bayesian Approach to Estimation of Time Varying Regression Coefficients, *Annals of Economic and Social Measurement*, Vol.2, pp.501-523.

Shumway, R.H. and Stoffer, D.S. (1982). An Approach to Time Series Smoothing and Forecasting Using the EM Algorithm, *Journal of Time Series Analysis*, Vol.3, PP.253-264.

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. (1989). The Kalman Filter Model under the Assumption of the First-Order Autoregressive Process in the Disturbance Terms, *Economics Letters*, Vol.31, No.2, pp.145-149.

Tanizaki, H. (1993a). Kalman Filter Model with Qualitative Dependent Variable, *The Review of Economics and Statistics*, Vol.75, No.4, pp.747-752.

Tanizaki, H. (1993b). *Nonlinear Filters: Estimation and Applications*, (Lecture Notes in Mathematical Economics and Systems, No.400), Springer-Verlag.

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. (2000). Time-Varying Parameter Model Revisited, *Kobe University Economic Review*, Vol.45, pp.41-57.

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. (1994). Prediction, Filtering and Smoothing in Nonlinear and Nonnormal Cases Using Monte-Carlo Integration, *Journal of Applied Econometrics*, Vol.9, No.2, pp.163-179 (in *Econometric Inference Using Simulation Techniques*, Chap.12, pp.245-261 (H.K. van DijK, A. Manfort and B.W. Brown, Eds., 1995), John Wiley & Sons).

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. (1999). A Non-linear Filtering Approach to Stochastic Volatility Models with an Application to Daily Stock Returns, *Journal of Applied Econometrics*, Vol.14, pp.101-121.

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.

Watson, M.W. and Engle, R.F. (1983). Alternative Algorithms for the Estimation of Dynamic Factor, MIMIC and Varying Coefficient Regression Models, *Journal of Econometrics*, Vol.23, pp.385-400.

日本語文献

青木正直 (1984). 『時系列解析と日本経済 − システム論的接近 −』東洋経済新報社。

有本卓 (1977). 『カルマン・フィルター』産業図書。

片山徹 (2000). 『新版 応用カルマンフィルタ』朝倉書店。

谷崎久志 (1993). 『状態空間モデルの経済学への応用』日本評論社。

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。