論文題目

確率的ボラティィリティィモデル

京都大学大学院経済学研究科修士課程

2000 年 経済動態分析専攻　　入学

氏名　　金谷　　太郎[1]

提出年　　　2002 年 1 月

目 次

1　はじめに

確率的ボラティリティ（stochastic volatility;SV）モデルは，数理ファイナンスとファイナンス計量の両方に由来し，いくつかのSVモデルはまったく異なった分析対象から考え出された．例えば，Clark(1973)は，資産の収益率をinformation arrival(市場に入ってくる情報量または取引回数）に依存させてモデル化し(分布混合仮説)，これが，時間によって変動するボラティリティモデルを生んだ．また，Tauchen and Pitts(1983)はこのモデルを洗練し，資産の収益率をinformation arrivalによって同時決定される変数(例えば，取引高)とランダムな確率変数の混合分布としてモデル化した(2変量分布混合仮説)．Hull and White(1987)は直接，information arrivalと資産の収益率を結びつけるのではなく，原資産が連続時間SVモデルに従うと考えたときのヨーロピアンコールオプションの価格を考察した．彼らは原資産のボラティリティを正の拡散過程に従うと仮定している．さらに別のアプローチとして，Taylor(1986)は，ARCH(Autoregressive Conditional Heteroskedasticity)型モデルの代替モデルとして離散時間SVモデルを定式化した．Taylorのモデルやその他のSVモデルはほとんど推定が不可能だったが，最近の計量経済学の理論の発展によってSVモデルの推定が可能になった．それゆえ，現在ではSVモデルは魅力的なモデルになってきており，ARCH型モデルのような他のモデルに取って代わろうとしている．

SVモデルの文献は数理ファイナンスと計量経済学の両方の分野で見ることができ，多様性に富んでいる．ARCH型モデルについては，

たくさんの優れたサーベイがある (例えば， Bollerslev, Engle and Nelson(1994) など).

また，オプションの価格付けは，ボラティリティの理論と密接な関係があるので，最低限のことは触れなければならない．実際，2 節ではボラティリティの定義を与えるが，そこで Black-Scholes(B-S) インプライドボラティリティについて言及する．また，2 節では現実の金融市場のデータの特徴を要約して，ボラティリティのモデル化について一般的に述べる．このボラティリティの一般的なモデル化によって，ARCH 型モデルと SV モデルの関係を明らかにしていく．3 節では SV モデルの推定について，マルコフ連鎖モンテカルロ (Markov-chain Monte Carlo;MCMC) 法によるベイズ推定を中心に述べる．4 節は結論と今後の課題である．

2 金融市場におけるボラティリティ

ボラティリティは派生証券の価格付けにおいて重要な役割を果たす．ヨーロピアンオプションの B-S 公式は，その前提となる仮定が侵されている場合でさえも，最も広範に使用される．それゆえ,2.1 節ではボラティリティの複数の表現を議論するために B-S モデルを取り上げる.2.2 節はボラティリティとオプション価格に関する実際のデータについて述べる. 両節ともに 2.3 節と 2．4 節で述べるボラティリティのモデル化の準備になっている.

2.1 B-S モデルとインプライドボラティリティ

連続時間確率過程で資産価格の振るまいを記述するのが標準的になったのは, Bachelier の研究から半世紀以上も後のことであった.Black and Scholes(1973) や Merton(1990) は特に影響力があった.2.1.1 節では拡散過程によって資産価格をモデル化する際に必要な仮定を見ていく. 特に, 瞬間的なボラティリティの概念について述べる.2.1.2 節ではオプションの価格付けモデルといくつかのインプライドボラティリティの概念について述べる.

2.1.1 瞬間的なボラティリティ

ここでは株式のような金融資産を考えて, 現在 (時点 t) の市場価格を S_t と書くことにする. 時点 t での利用可能な情報を I_t として, 期間 $[t, t+h]$ において資産を保有したときの収益率 $\dfrac{S_{t+h}-S_t}{S_t}$ の I_t を所与とした条件付き分布を考える. 本論文を通じての仮定は，資産の収益率は有限な I_t を所与とした条件付き期待値をもつ, すなわち,

$$E_t\left(\frac{S_{t+h}-S_t}{S_t}\right) = \frac{1}{S_t}E_tS_{t+h} - 1 < +\infty$$

である. ただし,$E_t(\cdot) \equiv E(\cdot|I_t)$ である. また同様に, 資産の収益率は有限な I_t を所与とした条件付き分散をもつ, すなわち,

$$V_t\left(\frac{S_{t+h}-S_t}{S_t}\right) = \frac{1}{S_t^2}V_tS_{t+h} < +\infty$$

と仮定しておく.

仮定 2.1 期待収益率 $\dfrac{1}{h}E_t\left(\dfrac{S_{t+h}-S_t}{S_t}\right)$ は $h \to +0$ のとき有限な値 $\mu_S(I_t)$ に概収束する.

この仮定から $E_t S_{t+h} - S_t \sim h\mu_S(I_t)S_t$ であるが, これを微分表現すると

$$\frac{d}{d\tau}E_t(S_\tau)\Big|_{\tau=t} = \mu_S(I_t)S_t, \quad a.s. \tag{1}$$

である．ここでの導関数は右側からとっている.(1) 式はしばしば $E_t(dS_t) = \mu_S(I_t)S_t dt$ と書かれる.

仮定 2.2 収益率の条件付き分散 $\dfrac{1}{h}V_t\left(\dfrac{S_{t+h}-S_t}{S_t}\right)$ は $h \to +0$ のとき有限な値 $\sigma_S^2(I_t)$ に概収束する.

この仮定から $V_t S_{t+h} - V_t S_t \sim h\sigma_S^2(I_t)S_t^2$(ただし，$V_t(S_t)=0$) であるが, 微分表現で書くと,

$$\frac{d}{d\tau}V_t(S_\tau)\Big|_{\tau=t} = \sigma_S^2(I_t)S_t^2, \quad a.s. \tag{2}$$

となる．これはまた,$V_t(dS_t) = \sigma_S^2(I_t)S_t^2 dt$ と表現される.

この仮定のもとで，$\sigma_S(I_t)$ を瞬間的なボラティィリティと定義する．すなわち，

$$\sigma_S(I_t) \equiv \left[\lim_{h\to+0}\frac{1}{h}V_t\left(\frac{S_{t+h}-S_t}{S_t}\right)\right]^{\frac{1}{2}} \tag{3}$$

である．

仮定 2.1 と 2.2 を満たす確率微分方程式として,

$$dS_t = \mu_S(I_t)S_t dt + \sigma_S(I_t)S_t dW_t \tag{4}$$

を考えることができる. ただし,W_t は標準ブラウン運動である.(4) 式は最も標準的な資産価格過程の定式化であり, 金融理論の発展上で大きな役割を果たしている．最後に，この特殊型について少し触れてお

く.$\mu_S(I_t)$ と $\sigma_S(I_t)$ が全ての時点 t に関して一定であるとき, 資産価格は幾何ブラウン運動に従う. このプロセスは Black and Scholes(1973) で, 有名なヨーロピアンオプションの価格公式を導出するのに使われている. 明らかに, 瞬間的なボラティィリティ $\sigma_S(I_t)$ を定数にすることで, たくさんのことを単純化している. もう 1 つ触れておかなければならないのは, 仮定 2.1 と 2.2 は資産価格のプロセスのジャンプの可能性を残しているということである. 例えば Merton(1976) では，株価のプロセスはポアソン過程をつかって,

$$\frac{dS_t}{S_t} = (\mu - \lambda k)dt + \sigma dW_t + dq_t \tag{5}$$

のようになっている．ただし, ここでは有限なジャンプによる変化率を Y(確率変数) として,

$$dq_t = \begin{cases} Y-1 & (\text{確率}\ ;\lambda dt) \\ 0 & (\text{確率}\ ;1-\lambda dt) \end{cases}$$

である.λ はポアソンパラメータ,dW_t と dq_t は互いに独立,$k = E_Y[Y-1]$(E_Y は Y に関する期待値) である. このとき (5) 式は仮定 2.1 と 2.2 を満たしている.[2] この 2 つの仮定は原理的にはジャンプを許すが, 本論文ではサンプルパスは連続であると仮定してジャンプの可能性を排除する.

2.1.2　オプション価格とインプライドボラティリティ

イントロダクションで SV モデルはオプションの価格付けの文献にも由来すると述べた. ここ 25 年間でオプションやその他の派生証券の市場は著しく発展している. オプション市場はしばしば,“ ボラ

ティリティを取引する ”市場として特徴づけられる. この節では, このようなことが言われる理由を考え, さらに, オプションとインプライドボラティリティの関係, 原資産の収益の過程の瞬間的なボラティリティと平均ボラティリティの概念について考えていく.

B-S オプションプライシングモデルは原資産の価格過程が幾何ブラウン運動に従うモデルである. すなわち,

$$dS_t = \mu_S S_t dt + \sigma_S S_t dW_t \tag{6}$$

ただし,μ_S と σ_S は固定されたパラメータである. 行使価格 K のヨーロピアンコールオプションの満期時点 $t+h$ におけるペイオフは,

$$[S_{t+h} - K]_+ = \begin{cases} S_{t+h} - K, & S_{t+h} \geq K \\ 0, & S_{t+h} < K \end{cases} \tag{7}$$

Black and Scholes(1973) 以来, この契約の価格公式の導出の様々な方法が提案されてきた. 本論文でこれらの方法の詳細について述べないが, ボラティリティに関連する概念の議論のために最小限のことは述べる.

市場での取引が連続的で取引費用 0 であると仮定すれば, Black-Scholes の仮定する世界ではコールの買いと原資産の空売りという戦略によってリスクを完全に排除するポートフォリオを組むことができる. これによって, 無裁定の議論さえ用いれば, コールオプションを含むリスクのないポートフォリオの収益率とリスクフリーレートを一致させることで, 偏微分方程式が得られる. この偏微分方程式を境界条件 (7) で解けば, オプションの価格を計算することができる.

B-S 公式の導出は“ リスク中立な世界 ”を経由すれば最も簡単に

なる. この方法をリスク中立化法とよぶ. リスク中立の世界では, 資産価格過程は修正された確率測度で特徴付けられ, その測度をリスク中立測度と言い, Q と書くことにする. この仮想的な確率測度はオプションの価格付けのみに使われ, 一般に, データを生成するプロセスの観測確率とリスク中立確率は一致しない. リスク中立な世界では

$$
\begin{aligned}
\frac{dS_t}{S_t} &= r_t dt + \sigma_S dW_t^Q & (8) \\
C_t &= C(S_t, K, h, t) = v(t, t+h) E_t^Q [S_{t+h} - K]_+ & (9)
\end{aligned}
$$

である. W_t^Q は Q のもとで標準ブラウン運動, E_t^Q は Q のもとでの期待値, $v(t, t+h)$ は満期 $t+h$ の割引債の t 時点での価格, さらに r_t は無リスク資産の瞬間的な利子率とする. イールドを $Y(t, t+h)$ と書くことにすると[3]

$$
\begin{aligned}
r_t &= \lim_{h \to 0} Y(t, t+h) = -\lim_{h \to 0} \frac{1}{h} \log v(t, t+h) \\
&= -\lim_{h \to 0} \frac{1}{h} (\log v(t, t+h) - \log v(t, t)) \\
&= -\frac{d}{d\tau} v(t, \tau) \Big|_{\tau = t} & (10)
\end{aligned}
$$

いま, 利子率が確定的 (不確実性のソースは W_t^Q のみ) であるとすると

$$
v(t, t+h) = \exp\left[-\int_t^{t+h} r_\tau d\tau \right] \tag{11}
$$

リスク中立な世界ではリスクプレミアムは存在しないので,r_t と株式の期待収益率は一致していて,(9) 式が述べるとおり, コールオプションの価格 C_t は満期時点のペイオフ $[S_{t+h} - K]_+$ を現在価値に割り引いたものになっている.

S_t を所与とした S_{t+h} の条件付き分布が対数正規分布であること

がわかるので,(9) 式の期待値を計算することによって時点 t でのコールオプションの公式が次のように得られる.

$$
\begin{aligned}
C_t &= S_t[\Phi(d_t) - e^{-x_t}\Phi(d_t - \sigma_S\sqrt{h})] && (12) \\
d_t &= \frac{x_t}{\sigma_S\sqrt{h}} + \frac{\sigma_S\sqrt{h}}{2} && (13) \\
x_t &= \log\frac{S_t}{Kv(t,t+h)} && (14)
\end{aligned}
$$

ただし $\Phi(\cdot)$ は標準正規分布の累積分布関数である, (12) 式を Black-Scholes の公式といい, このオプション価格 C_t は株価 S_t や行使価格 K や割引関数 $v(t,t+h)$ に依存している.x_t はオプションの moneyness とよばれ, これを基準にオプションの状態は以下の 3 つに分けられる.

- $x_t = 0$ のとき, 現在の株価が行使価格の割引現在価値に一致している. この状態 を at the money という.
- $x_t > 0$ のとき, in the money という.
- $x_t < 0$ のとき, out of the money という.

B-S 公式はその仮定が侵されているときでさえ, 一般に使われていることは前にも述べたが, 特に, ボラティリティが一定であるという仮定は非現実的である. このことは 2.2 節の現実の金融市場のデータを見ればわかる. このことが Hull and White(1987) がオプションの価格付けモデルに確率的なボラティリティを導入した誘因である. そこではボラティリティは W_t と独立な状態変数と仮定されている.

$$
\begin{cases}
\dfrac{dS_t}{S_t} = r_t dt + \sigma_{St} dW_t \\
\{\sigma_{St}\}_{t\in[0,T]}, \{W_t\}_{t\in[0,T]}, \text{互いに独立なマルコフ過程}
\end{cases}
\tag{15}
$$

ここでは, リスク中立的な世界を考えているので, 株式の瞬間的な期待収益率は r_t に一致している. もし,(S_t, σ_{St}) を所与とした (S, σ_S) の同時分布の期待値を計算できれば, オプション価格は (12) 式で与えられる.(9) 式は次のように書き直せる.

$$
\begin{aligned}
C_t &= v(t, t+h) E_t[S_{t+h} - K]_+ \\
&= v(t, t+h) E_t\{E_t[(S_{t+h} - K)_+ | \{\sigma_{S\tau}\}_{t \le \tau \le t+h}]\}
\end{aligned}
\tag{16}
$$

ここでは, 内側の期待値は I_t とボラティリティのパス $\{\sigma_{S\tau}\}_{t \le \tau \le t+h}$ を所与とした S_{t+h} の期待値になっている. しかし, ボラティリティのプロセスと W_t は独立なので,(12) 式を使うと

$$
\begin{aligned}
&v(t, t+h) E_t[(S_{t+h} - K)_+ | \{\sigma_{S\tau}\}_{t \le \tau \le t+h}] \\
&\qquad = S_t E_t[\Phi(d_{1t}) - e^{-x_t} \Phi(d_{2t})]
\end{aligned}
\tag{17}
$$

ただし,

$$
\begin{aligned}
&d_{1t} = \frac{x_t}{\gamma(t, t+h)\sqrt{h}} + \frac{\gamma(t, t+h)\sqrt{h}}{2} \\
&d_{2t} = d_{1t} - \gamma(t, t+h)\sqrt{h} \\
&\gamma^2(t, t+h) = \frac{1}{h} \int_t^{t+h} \sigma_{S\tau}^2 d\tau
\end{aligned}
\tag{18}
$$

これらより, Hull and White のオプション価格公式が得られて

$$
C_t = S_t E_t[\Phi(d_{1t}) - e^{-x_t} \Phi(d_{2t})] \tag{19}
$$

ここでの期待値は $\gamma(t, t+h)$ についての (リスク中立測度での) 期待値である.

以後この節ではオプションの市場価格が Hull and White(H-W) 公式 (19) に従うと仮定する. そうすると,2 つのインプライドボラティリティが考えられる.(1) 瞬間的なインプライドボラティリティ (2) 平均インプライドボラティリティである. このことを詳しく議論するために, リスク中立確率分布はパラメータ族 $P_\theta, \theta \in \Theta$ に属していると仮定する. すると,H-W 公式は以下のように表現できる.

$$C_t = S_t F[\sigma_{St}, x_t, \theta_0] \tag{20}$$

ここで θ_0 は真の未知パラメータである. 公式 (19) から"オプション市場はバラティィリティを取引する市場と考えられる"と主張される理由がわかる. 実際, 任意の (x_t, θ) に対して, $F[\cdot, x_t, \theta]$ は 1 対 1 の関数であり,$F[\cdot, x_t, \theta]$ の逆関数をとることによって (19) から瞬間的なインプライドボラティリティ

$$\sigma_t^{imp} = G[S_t, C_t, x_t, \theta] \tag{21}$$

を求めることができる. Bajeux and Rochet はオプション価格と瞬間的なボラティリティの 1 対 1 の対応関係を証明してオプション市場を使ったボラティリティリスクのヘッジを定式化した. もし, 真の未知パラメータ θ_0 がわかっているか, 少なくとも, よい推定量があるならば, 瞬間的なボラティリティ (20) は有用な道具である.

しかしながら, SV モデル推定の難しさにより, 長い間, 瞬間的なボラティリティの応用上の利用は妨げられてきた. このことが, もう 1 つのインプライドボラティリティ, すなわち Black-Scholes インプライドボラティリティ, が広く使われてきた理由である.B-S インプライ

ドボラティィリティは以下の $\omega^{imp}(t,t+h)$ のように定義される．

$$
\begin{aligned}
C_t &= S_t[\Phi(d_{1t}) - e^{-x_t}\Phi(d_{2t})] \\
d_{1t} &= \frac{x_t}{\omega^{imp}(t,t+h)\sqrt{h}} + \frac{\omega^{imp}(t,t+h)\sqrt{h}}{2} \\
d_{2t} &= d_{1t} - \omega^{imp}(t,t+h)\sqrt{h} \\
x_t &= \log\frac{S_t}{Kv(t,t+h)}
\end{aligned}
\tag{22}
$$

ここでは，C_t は市場価格である．Hull and White のオプション価格公式はこの B-S インプライドボラティィリティを使う根拠になっている．(19) と (22) を比較すると，B-S インプライドボラティィリティ $\omega^{imp}(t,t+h)$ が $\gamma(t,t+h)$ の期待値のようなものなので，$\omega^{imp}(t,t+h)$ は平均的なボラティィリティであると解釈できる (もちろん，オプションの市場価格が H-W 公式に一致していると仮定しているからである)．このことをより詳しくみるために，最も簡単な at the money のケースについて考える．at the money では $x_t = 0$ なので $d_{2t} = -d_{1t}$ である．したがって，$\Phi(d_{1t}) - e^{-x_t}\Phi(d_{2t}) = 2\Phi(d_{1t}) - 1$ である．よって，(19) と (22) より．at the money での B-S インプライドボラティィリティ $\omega^{imp}(t,t+h)$ は

$$
\Phi\left(\frac{\omega^{imp}(t,t+h)\sqrt{h}}{2}\right) = E_t\Phi\left(\frac{\gamma(t,t+h)\sqrt{h}}{2}\right)
$$

を満たす．ここで，標準正規分布の分布関数は 0 の近傍でほぼ線形なので，もし h が小さければ

$$
\omega^{imp}(t,t+h) \approx E_t\gamma(t,t+h) = E_t\left[\frac{1}{h}\int_t^{t+h}\sigma_{S\tau}d\tau\right]^{\frac{1}{2}}
$$

とすることができる．このことから，B-S インプライドボラティィリティ $\omega^{imp}(t,t+h)$ は平均的なインプライドボラティィリティであると

解釈できる.

2.2　現実の金融データの特徴

　モデルの定式化と選択は, 現実のデータに基づいて行われる. そこでこの節では, 現実の金融市場のデータがどのような特徴をもっているのかをみていく.

(a) Thick tails

1960 年代前半以降, Mandelbrot や Fama などが資産収益は正規分布よりも急尖的な分布を持つということが報告している.

(b) Volatility clustering

金融時系列データは大抵, ボラティィリティの大きい期間と小さい期間がひとかたまりになっている.volatility clustering と thick tails には, 密接な関係がある. 後者は前者の静学的説明になっている.ARCH モデルの示す重要な特徴は動学的な (条件付き) ボラティィリティの振るまいと (無条件の) 分布の裾の厚さとの関係である.ARCH モデルは, Engle(1982) によって導入され多くの拡張がなされてきたが, 同様に SV モデルも volatility clustering を表現できるようにつくられている.

(c) Leverage effects

Black(1976) はボラティィリティと株価の負の相関のことを leverage effect と名づけた. 株価の下落は企業の leverage一自己資本の収益力を高めるために借入金などの他人資本を活用すること一の増加を意味

するので, 不確実性すなわちボラティリティが大きくなると考えられる. しかし,Schwert(1989) などの実証分析によると leverage だけで株価の非対称性を説明するのは難しいと報告されている.

(d) Long memory and persistence

一般的にボラティリティは persistent―過去のボラティリティの影響が残りつづける―であることが多くの実証研究で報告されている.

2.3　情報集合

これまでは情報集合の定義を明らかにしてこなかったが，次節以降で，SV モデルと ARCH モデルの関係を明らかにするために情報を明確に定義することは必要である．そのために Granger の因果関係を定義しておく．I_t を t 時点において利用可能な全ての情報の集合，$\{x_t\}, \{z_t\}$ を確率過程とする．また，$I_t - \{z_s\}_{s=0}^{t}$ を I_t から $\{z_s\}_{s=0}^{t}$ を除いた情報とする.

定義 2.1 (強 Granger の因果関係) [4]

$$E(x_t^n | I_{t-1}) = E(x_t^n | I_{t-1} - \{z_s\}_{s=0}^{t-1}), \quad n = 1, 2$$

となるとき，強 *Granger* の意味で z_t から x_t への因果関係がないという.

2.3.1 状態変数と情報集合

いま，観測されない状態変数によって株価のダイナミクスが決まるモデルを考える．すなわち，

$$\begin{cases} \dfrac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t \\ dU_t = \gamma_t dt + \delta_t dW_t^U \\ Cov(dW_t, dW_t^U) = \rho_t dt \end{cases} \tag{23}$$

ただし，以下の仮定を置く．

仮定 2.3 確率過程 $\mu_t, \sigma_t, \gamma_t, \delta_t, \rho_t$ は $I_t^U = \sigma[U_s, s \leq t]$ 適合的[5] である．

時点で利用可能な情報は

$$I_t = \sigma[U_\tau, S_\tau; \tau \leq t] \tag{24}$$

である．

2.3.2 離散時間モデルと Granger の因果関係

$R_t = \dfrac{S_t - S_{t-1}}{S_{t-1}}$ とし，(23) 式の離散時間のアナロジーとして次の式を考える．

$$R_{t+1} = \mu(U_t) + \sigma(U_t) z_{t+1} \tag{25}$$

z_t になんらかの制約を課すことによって, 例えば leverage effects のような現象を表現することができる．

仮定 2.4 z_t は $i.i.d.$ かつ強 $Granger$ の意味で U_t から z_t への因果関係がない

仮定 2.5 強 $Granger$ の意味で z_t から U_t への因果関係がない.

仮定 2.5 は仮定 2.3 の離散時間のアナロジーになっている. また, 仮定 2.4 から $E[z_{t+1}|(U_\tau, z_\tau; \tau \leq t)] = E[z_{t+1}|z_\tau; \tau \leq t] = 0$ なのでこれを使って,

$$
\begin{aligned}
&E[R_{t+1}|(R_\tau; \tau \leq t)] \\
&= E[E[R_{t+1}|(U_\tau, R_\tau; \tau \leq t)]|(R_\tau; \tau \leq t)] \\
&= E[E[R_{t+1}|(U_\tau, z_\tau; \tau \leq t)]|(R_\tau; \tau \leq t)] \\
&= E[\mu(U_t)|(R_\tau; \tau \leq t)]
\end{aligned} \tag{26}
$$

同様に, 仮定 2.4 から $E[z_{t+1}^2|(U_\tau, z_\tau; \tau \leq t)] = E[z_{t+1}^2|z_\tau; \tau \leq t] = 1$ なのでこれを使って,

$$
\begin{aligned}
&Var[R_{t+1}|(R_\tau; \tau \leq t)] \\
&= E[R_{t+1}^2|(R_\tau; \tau \leq t)] - (E[R_{t+1}|(R_\tau; \tau \leq t)])^2 \\
&= E[\sigma^2(U_t) z_{t+1}^2|(R_\tau; \tau \leq t)] \\
&= E[\sigma^2(U_t) E[z_{t+1}^2|(U_\tau, R_\tau; \tau \leq t)]|(R_\tau; \tau \leq t)] \\
&= E[\sigma^2(U_t) E[z_{t+1}^2|(U_\tau, z_\tau; \tau \leq t)]|(R_\tau; \tau \leq t)] \\
&= E[\sigma^2(U_t)|(R_\tau; \tau \leq t)]
\end{aligned} \tag{27}
$$

ここで,

$$
I_t^R \equiv (R_\tau; \tau \leq t) \tag{28}
$$

と定義しておく. さらにもう 1 つの仮定を置く.

仮定 2.6 $\mu(U_t)$ は I_t^R-可測である.

このとき，(26),(27) は次のようになる.

$$E[R_{t+1}|I_t^R] = \mu(U_t) \tag{29}$$

$$Var[R_{t+1}|I_t^R] = E[\sigma^2(U_t)|I_t^R] \tag{30}$$

2.4　確率的ボラティリティの統計モデル

この節では離散時間モデルの一般型について議論する.2.4.1 節では“ lagged autoregressive random variance models ”と“ contemporaneous autoregressive random variance models ”の 2 つのモデルについて述べる．2.4.2 節では stochastic autoregressive volatility(SARV) モデルを導入し,2.4.3 節ではパラメータの識別とそのために必要な制約について述べる.

2.4.1　2 つのモデルの同値性

いま，簡単化のために $\mu(U_t) = 0$ とする．“ lagged autoregressive random variance models ”は以下のように定義される.

$$R_{t+1} = \sigma_t z_{t+1} \tag{31}$$

$$Cov(\sigma_{t+1}, z_{t+1}|I_t^R) \neq 0 \tag{32}$$

(32) が負の共分散であるなら leverage effect をあらわしている. また, “ contemporaneous autoregressive random variance models ”は以下のように定義される.

$$R_t = \sigma_t z_t \tag{33}$$

$$Cov(\sigma_{t+1}, z_t|I_t^R) \neq 0 \tag{34}$$

ボラティリティのプロセス σ_t は観測できないのであるから,(32) と (32) の組み合わせと (33) と (34) の組み合わせは同値である. すなわち，(32) と (33) のどちらを使ってもよいが，leverage effect を考えるときだけ組み合わせを気をつければよい.

2.4.2 SARV モデル

$\epsilon_t = \sigma_t z_t$ と定義すると,“ contemporaneous autoregressive random variance models ”は

$$R_t = \mu_t + \epsilon_t \tag{35}$$

と書くことができる．仮定 2.6 より μ_t は I^R_{t-1}-可測であるので,$t-1$ 時点でわかる値であり, 一方,ϵ_t は $t-1$ 時点でわからない値である. ここで,

$$\begin{aligned} E(R_t|I^R_{t-1}) &= \mu_t && (36) \\ Var(R_t|I^R_{t-1}) &= E(\sigma_t^2|I^R_{t-1}) && (37) \end{aligned}$$

である. 収益率 R_t の条件付き分散 (37) のモデル化を考える.

(1) ボラティリティクラスタリングは (37) の条件付き期待値を自己回帰モデルにすることによってとらえることができる.

(2) ファットテールを捉えるには次の 2 つの方法がある. (a) ホワイトノイズ z_t に裾の厚い分布を仮定する. (b)z_t が標準正規分布であっても, ボラティリティσ_t^2 が変動するならば,ϵ_t の尖度は 3 を上回る (付録 5.1 参照).

ボラティリティのプロセスに定常性と 1 次のマルコフ性を仮定し

ても，かならずしも σ_t 自身が線形な AR(1) プロセスである必要はない．Andersen(1994) は多くのモデルを含む以下のような一般的なモデルを考案した．

$$\sigma_t^q = g(K_t), \qquad q \in \{1,2\}$$
$$K_t = w + \beta K_{t-1} + [\gamma + \alpha K_{t-1}]u_t \tag{38}$$

ただし，$\tilde{u}_t = u_t - 1$ は平均 0，分散 1 のホワイトノイズである．例えば，GARCH(1,1) モデルを得たいならば,(38) に $K_t = \sigma_t^2, \gamma = 0, u_t = z_{t-1}^2$ を代入することにより，

$$\sigma_t^2 = w + \beta\sigma_{t-1}^2 + \alpha\sigma_{t-1}^2 z_{t-1}^2$$

とすればよい.(35) と (38) のランダム項の間には $u_t = z_{t-1}^2$ という関係があるので,GARCH モデルではボラティリティのプロセスに固有の不確実性はない. これが ARCH 型モデルの特徴である. また,SV モデルを得たいならば,(38) に $K_t = \log\sigma_t^2, \alpha = 0, \gamma\tilde{u}_{t+1} = \eta_{t+1}, w+\gamma = \omega, \beta = \phi$ を代入することにより，

$$\log\sigma_{t+1}^2 = \omega + \phi\log\sigma_t^2 + \eta_{t+1} \tag{39}$$

とすればよい. ここで, 撹乱項 η_t はホワイトノイズである.$Cov(\eta_{t+1}, \epsilon_t) \neq 0$ とすれば leverage effects をあらわすことができる．

2.4.3　パラメータの識別

この節ではSARVモデルのパラメータの識別について述べる. 考えるモデルは

$$
\begin{aligned}
\epsilon_t &= \sigma_t z_t \\
\sigma_t^q &= g(K_t), \qquad q \in \{1,2\} \\
K_t &= w + \beta K_{t-1} + [\gamma + \alpha K_{t-1}]u_t
\end{aligned} \tag{40}
$$

であるが, 平均0のホワイトノイズ $\tilde{u}_t = u_t - 1$ で書くと,

$$
K_t = (w+\gamma) + (\alpha+\beta)K_{t-1} + [\gamma + \alpha K_{t-1}]\tilde{u}_t \tag{41}
$$

である. しかし,w と γ を別々に識別することは困難である. 同様に α と β の識別も困難である. これらを識別するために普通は何らかの制約が課される.

識別問題を詳しく見るために,Andersen(1994) にしたがって以下のようにパラメータを変換する (ただし,$\alpha = 0$).

$$
\begin{aligned}
K &= \frac{w+\gamma}{1-\alpha-\beta} \\
\rho &= \alpha + \beta \\
\delta &= \frac{\gamma}{\alpha}
\end{aligned} \tag{42}
$$

(41) は以下のように書き換えられる.

$$
K_t = K + \rho(K_{t-1} - K) + (\delta + K_{t-1})\bar{U}_t
$$

ただし,$\bar{U}_t = \alpha\tilde{u}_t$ である. この式より, 例えば K_t の3つのモーメント $E(K_t), E(K_t^2), E(K_t K_{t-1})$ から,3つのパラメータ K, ρ, δ しか識別で

きないことがわかる.

実証分析の観点から考えると, 以下のことを考える必要がある.

(1) 観測されるデータ Y_t のモーメントをボラティリティσ_t のモーメントであらわす.

(2) ボラティリティσ_t のモーメントを状態変数 K_t のモーメントであらわす.

(1) については z_t に仮定を置いて対処する. 例えば,z_t が正規分布に従うと仮定するならば,

$$
\begin{aligned}
E(|y_t|) &= \sqrt{\frac{2}{\pi}} E(\sigma_t) \\
E(|y_t||y_{t-j}|) &= \frac{2}{\pi} E(\sigma_t \sigma_{t-j}) \\
E(|y_t^2||y_{t-j}|) &= \sqrt{\frac{2}{\pi}} E(\sigma_t^2 \sigma_{t-j})
\end{aligned} \tag{43}
$$

となる.(2) については一般に g の関数型と $\tilde{u}_t$ の分布の仮定が必要である. 例えば,(39) の SV モデルならば,$\sigma_t^2 = e^{K_t}$ であり, さらに $\tilde{u}_t$ を正規分布に従うとすると,

$$
\begin{aligned}
E(\sigma_t^n) &= \exp\left[\frac{n}{2} E(K_t) + \frac{n^2 Var(k_t)}{8}\right] \\
E(\sigma_t^m \sigma_{t-j}^n) &= E(\sigma_t^m) E(\sigma_{t-j}^n) \exp\left[\frac{mnCov(K_t, K_{t-j})}{4}\right] \\
Cov(K_t, K_{t-j}) &= \phi^j Var(K_t)
\end{aligned} \tag{44}
$$

となる.

3　SV モデルの推定

2.4.3 節の SV モデル

$$\begin{aligned} \epsilon_t &= \sigma_t z_t, \qquad z_t \sim i.i.d.N(0,1) & (45) \\ \log \sigma_t^2 &= \omega + \phi \log \sigma_{t-1} + \eta_t, \quad \eta_t \sim i.i.d.N(0,\sigma_\eta^2) & (46) \end{aligned}$$

を推定することを考える．ただし, ここでは z_t と η_t は独立, すなわち,leverage effects はないものとする.[6] SV モデルの尤度関数は ARCH 型モデルの尤度関数のように簡単に書くことができないので，最尤法は困難である．そこで，尤度に基づかない推定方法が提案されている．3.1 節ではその代表的な方法である一般化積率法 (generalized method of moments;GMM) と擬似最尤法 (quasi-maximum likelihood estimation;QMLE) について簡単に触れる．3.2 節では MCMC 法によるベイズ推定の方法について述べる.3.3 節では MCMC 法によるベイズ推定のシミュレーション実験の結果である.

3.1　GMM と QMLE

Jacquier,Polson and Rossi(1994) の GMM は以下のとおりである.(45),(46) 式の SV モデルにおいて推定すべきパラメータをまとめて $\theta = (\omega, \phi, \sigma_\eta^2)'$ とする. サンプル数 T 個のとき,m 個の理論的なモーメント $E|\epsilon_t^c||\epsilon_{t-p}^d|$ とサンプルモーメントとの差を $m \times 1$ ベクトル $g_T(\theta)$ とし, ウェイト行列を W_T としておくと,

$$\hat{\theta}_T = \arg\max_\theta g_T(\theta)' W_T g_T(\theta)$$

を求めればよい. しかし, このような GMM による推定は効率性が低いことが知られている.Anderson,Chung and Sorensen(1999) によると, モーメントやウェイと行列の選択を工夫しても効率性はさほど上昇しない.

SV モデルの QMLE による推定は,Nelson(1988) によって提案されたが, その方法は以下のとおりである.(45) 式の両辺を 2 乗し, 対数をとると,

$$\log \epsilon_t^2 = \log \sigma_t^2 + \log z_t^2 \tag{47}$$

となるが,(47) 式と (46) 式からなる線形状態空間モデルにカルマン・フィルタを適用することによって尤度を計算し, この尤度を最大化するようなパラメータを選択する。ただし,(47) 式の $\log z_t^2$ の分布は正規分布でないため, このカルマン・フィルタを使って計算された尤度は正しい尤度ではない．それゆえ，この尤度は擬似尤度とよばれる．このため，この方法もまた，推定値の効率性は低い.

3.2 MCMC によるベイズ推定

この節では，MCMC によるベイズ推定の方法について説明する．まず，新たな変数を

$$h_t \equiv \log \sigma_t^2 - \frac{\omega}{1-\phi}$$
$$\sigma_r \equiv \exp\left(\frac{\omega}{2(1-\phi)}\right)$$

と定義し，(45),(46) 式を次のように書き換えておく．

$$
\begin{aligned}
\epsilon_t &= \sigma_r \exp\left(\frac{h_t}{2}\right) z_t, \qquad z_t \sim i.i.d.N(0,1) & (48)\\
h_t &= \phi h_{t-1} + \eta_t, \qquad \eta_t \sim i.i.d.N(0,\sigma_\eta^2) & (49)
\end{aligned}
$$

この書き換えによって，推定すべきパラメータは $(\phi, \sigma_\eta^2, \sigma_r^2)$ になる．これらのパラメータをベイズ推定するには，事前分布を設定する必要があるが，ここでは noninformative な事前分布[7]

$$
f(\phi) \propto I[-1,1], \quad f(\sigma_r^2) \propto \frac{1}{\sigma_r^2}, \quad f(\sigma_\eta^2) \propto \frac{1}{\sigma_\eta^2} \tag{50}
$$

を用いる．ただし，$I[-1,1]$ は $[-1,1]$ の区間では 1，それ以外では 0 をとる関数である．

3.2.1 Gibbs sampler(single-move sampler の場合)

$(\phi, \sigma_r^2, \sigma_\eta^2)$ を事後分布 $f(\phi, \sigma_r^2, \sigma_\eta^2 | \{\epsilon_t\}_{t=1}^T)$ からサンプリングしたいが，それができない．このとき，Gibbs sampler(付録 5.2 参照) を使うためには,

$$
\begin{aligned}
&f(\phi | \sigma_r^2, \sigma_\eta^2, \{\epsilon_t\}_{t=1}^T)\\
&f(\sigma_r^2 | \phi, \sigma_\eta^2, \{\epsilon_t\}_{t=1}^T)\\
&f(\sigma_\eta^2 | \phi, \sigma_r^2, \{\epsilon_t\}_{t=1}^T)
\end{aligned}
$$

がサンプリングできる分布にならなければならないが，これらは求めることができない．そこで，条件の中に $\{h_t\}_{t=1}^T$ を含め，[8]

$$f(\phi|\sigma_r^2, \sigma_\eta^2, \{h_t\}_{t=1}^T, \{\epsilon_t\}_{t=1}^T) \tag{51}$$

$$f(\sigma_r^2|\phi, \sigma_\eta^2, \{h_t\}_{t=1}^T, \{\epsilon_t\}_{t=1}^T) \tag{52}$$

$$f(\sigma_\eta^2|\phi, \sigma_r^2, \{h_t\}_{t=1}^T, \{\epsilon_t\}_{t=1}^T) \tag{53}$$

$$f(h_t|\phi, \sigma_r^2, \sigma_\eta^2, \{h_s\}_{s\neq t}, \{\epsilon_t\}_{t=1}^T), \quad (t = 1, \cdots, T) \tag{54}$$

とすると，(51),(52),(53) は計算できて (付録 5.5 参照)

$$\phi|\cdot \sim N\left(\frac{\sum_{t=2}^T h_t h_{t-1}}{\sum_{t=2}^T h_{t-1}^2}, \frac{\sigma_\eta^2}{\sum_{t=2}^T h_{t-1}^2}\right) I[-1,1] \tag{55}$$

$$\sigma_r^{-2}|\cdot \sim Gamma\left(\frac{T}{2}, \frac{2}{\sum_{t=1}^T \epsilon_t^2 \exp(-h_t)}\right) \tag{56}$$

$$\sigma_\eta^{-2}|\cdot \sim Gamma\left(\frac{T-1}{2}, \frac{2}{\sum_{t=1}^T (h_t - \phi h_{t-1})^2}\right) \tag{57}$$

となる．(54) から h_t をサンプリングする方法は次節で説明する．

Gibbs sampler を実行するために初期値

$$\sigma_r^{2(0)}, \sigma_\eta^{2(0)}, \{h_t^{(0)}\}_{t=1}^T$$

を決める．これらの値を使って以下のループを繰り返す.

第 1 ループ

$f(\phi|\sigma_r^{2(0)},\sigma_\eta^{2(0)},\{h_t^{(0)}\}_{t=1}^T,\{\epsilon_t\}_{t=1}^T)$ から $sampling \to \phi^{(1)}$

$f(\sigma_r^2|\phi^{(1)},\sigma_\eta^{2(0)},\{h_t^{(0)}\}_{t=1}^T,\{\epsilon_t\}_{t=1}^T)$ から $sampling \to \sigma_r^{2(1)}$

$f(\sigma_\eta^2|\phi^{(1)},\sigma_r^{2(1)},\{h_t^{(0)}\}_{t=1}^T,\{\epsilon_t\}_{t=1}^T)$ から $sampling \to \sigma_\eta^{2(1)}$

$f(h_1|\phi^{(1)},\sigma_r^{2(1)},\sigma_\eta^{2(1)},\{h_t^{(0)}\}_{t=2}^T,\{\epsilon_t\}_{t=1}^T)$ から $sampling \to h_1^{(1)}$

$f(h_2|\phi^{(1)},\sigma_r^{2(1)},\sigma_\eta^{2(1)},h_1^{(1)},\{h_t^{(0)}\}_{t=3}^T,\{\epsilon_t\}_{t=1}^T)$ から $sampling \to h_2^{(1)}$

$\vdots$

$f(h_T|\phi^{(1)},\sigma_r^{2(1)},\sigma_\eta^{2(1)},\{h_t^{(1)}\}_{t=1}^{T-1},\{\epsilon_t\}_{t=1}^T)$ から $sampling \to h_T^{(1)}$

$\vdots$

第 n ループ

$f(\phi|\sigma_r^{2(n-1)},\sigma_\eta^{2(n-1)},\{h_t^{(n-1)}\}_{t=1}^T,\{\epsilon_t\}_{t=1}^T)$ から $sampling \to \phi^{(n)}$

$f(\sigma_r^2|\phi^{(n)},\sigma_\eta^{2(n-1)},\{h_t^{(n-1)}\}_{t=1}^T,\{\epsilon_t\}_{t=1}^T)$ から $sampling \to \sigma_r^{2(n)}$

$f(\sigma_\eta^2|\phi^{(n)},\sigma_r^{2(n)},\{h_t^{(n-1)}\}_{t=1}^T,\{\epsilon_t\}_{t=1}^T)$ から $sampling \to \sigma_\eta^{2(n)}$

$f(h_1|\phi^{(n)},\sigma_r^{2(n)},\sigma_\eta^{2(n)},\{h_t^{(n-1)}\}_{t=2}^T,\{\epsilon_t\}_{t=1}^T)$ から $sampling \to h_1^{(n)}$

$f(h_2|\phi^{(n)},\sigma_r^{2(n)},\sigma_\eta^{2(n)},h_1^{(n)},\{h_t^{(n-1)}\}_{t=3}^T,\{\epsilon_t\}_{t=1}^T)$ から $sampling \to h_2^{(n)}$

$\vdots$

$f(h_T|\phi^{(n)},\sigma_r^{2(n)},\sigma_\eta^{2(n)},\{h_t^{(n)}\}_{t=1}^{T-1},\{\epsilon_t\}_{t=1}^T)$ から $sampling \to h_T^{(n)}$

$n \to \infty$ のとき, $(\phi^{(n)},\sigma_r^{2(n)},\sigma_\eta^{2(n)},\{h_t^{(n)}\}_{t=1}^T)$ は同時密度

$f(\phi,\sigma_r^2,\sigma_\eta^2,\{h_t\}_{t=1}^T|\{\epsilon_t\}_{t=1}^T)$ からサンプリングされた確率変数に分布

収束するので，n が十分に大きければ，$(\phi^{(n)}, \sigma_r^{2(n)}, \sigma_\eta^{2(n)}, \{h_t^{(n)}\}_{t=1}^T)$ は事後分布 $f(\phi, \sigma_r^2, \sigma_\eta^2, \{h_t\}_{t=1}^T | \{\epsilon_t\}_{t=1}^T)$ からサンプリングされた値であると考えてよい．そこで，上記のループを十分大きな回数 (M 回) 繰り返したあと，さらに N 回のループをおこなって得られる

$$
\begin{gathered}
(\phi^{(M+1)}, \sigma_r^{2(M+1)}, \sigma_\eta^{2(M+1)}, \{h_t^{(M+1)}\}_{t=1}^T) \\
\vdots \\
(\phi^{(M+N)}, \sigma_r^{2(M+N)}, \sigma_\eta^{2(M+N)}, \{h_t^{(M+N)}\}_{t=1}^T)
\end{gathered}
$$

を事後分布 $f(\phi, \sigma_r^2, \sigma_\eta^2, \{h_t\}_{t=1}^T | \{\epsilon_t\}_{t=1}^T)$ からサンプリングされたものと考え，それらを用いて事後平均をとればよい．すなわち，

$$
\begin{aligned}
\hat{\phi} &= \frac{1}{N} \sum_{i=M+1}^{M+N} \phi^{(i)} \\
\hat{\sigma_r^2} &= \frac{1}{N} \sum_{i=M+1}^{M+N} \sigma_r^{2(i)} \\
\hat{\sigma_\eta^2} &= \frac{1}{N} \sum_{i=M+1}^{M+N} \sigma_\eta^{2(i)}
\end{aligned}
$$

が推定値である．

3.2.2 $\{h_t\}_{t=1}^T$ の single-move sampler による sampling

(54) は次のように計算できる (付録 5.6 参照)．

$$
\begin{aligned}
&f(h_t | \theta, \{h_s\}_{s \neq t}, \{\epsilon_t\}_{t=1}^T) \\
&\propto \exp\left(-\frac{h_t}{2}\right) \exp\left(-\frac{\epsilon_t^2}{2\sigma_r^2} \exp(-h_t)\right) \exp\left(-\frac{(h_t - \mu_t)^2}{2\tilde{\sigma_t}^2}\right)
\end{aligned}
\tag{58}
$$

ただし，

$$\mu_t = \begin{cases} \dfrac{h_2}{\phi}, & t = 1 \\ \dfrac{\phi(h_{t+1} - h_{t-1})}{1+\phi^2}, & 2 \le t \le T-1 \\ \phi h_{T-1}, & t = T \end{cases} \tag{59}$$

$$\tilde{\sigma_t}^2 = \begin{cases} \dfrac{\sigma_\eta^2}{\phi^2}, & t = 1 \\ \dfrac{\sigma_\eta^2}{1+\phi^2}, & 2 \le t \le T-1 \\ \sigma_\eta^2, & t = T \end{cases} \tag{60}$$

この条件付密度からは直接サンプリングできないが，ある正の整数 c のもとで，すべての h_t にたいして $f(h_t|\cdot) \le cg(h_t)$ となるような $g(\cdot)$ がみつかれば A-R 法 (付録 5.3 参照) によって $f(\cdot)$ からサンプリングできる．C, C' を定数とすると $\exp(-h_t)$ を h_t^* でテイラー展開することによって

$$\begin{aligned} &\log f(h_t|\cdot) = \log C - \frac{h_t}{2} - \frac{(h_t - \mu_t)^2}{2\tilde{\sigma_t}^2} - \frac{\epsilon_t^2}{2\sigma_r^2}\exp(-h_t) \\ &\le \log C - \frac{h_t}{2} - \frac{(h_t - \mu_t)^2}{2\tilde{\sigma_t}^2} - \frac{\epsilon_t^2}{2\sigma_r^2}\exp(-h_t^*)(1 - (h_t - h_t^*)) \\ &= \log C + \log C' + \log \frac{1}{\sqrt{2\pi\tilde{\sigma_t}^2}} - \frac{(h_t - \mu_t^*)^2}{2\tilde{\sigma_t}^2} \end{aligned} \tag{61}$$

となる. ただし,

$$\begin{aligned}
\mu_t^* &= \mu_t + \tilde{\sigma}_t{}^2 \left(\frac{\epsilon_t^2}{2\sigma_r^2} \exp(-h_t^*) - 1 \right) \\
f^*(h_t) &= \exp\left(-\frac{h_t}{2}\right) \exp\left(-\frac{\epsilon_t^2}{2\sigma_r^2} \exp(-h_t)\right) \exp\left(-\frac{(h_t-\mu_t)^2}{2\tilde{\sigma}_t{}^2}\right) \\
g^*(h_t) &= \exp\left(-\frac{h_t}{2} - \frac{(h_t-\mu_t)^2}{2\tilde{\sigma}_t{}^2} - \frac{\epsilon_t^2}{2\sigma_r^2} \exp(-h_t^*)(1-(h_t-h_t^*))\right) \\
&= \frac{C'}{\sqrt{2\pi\tilde{\sigma}_t{}^2}} \exp\left(-\frac{(h_t-\mu_t^*)^2}{2\tilde{\sigma}_t{}^2}\right) \\
g(h_t) &= \frac{1}{\sqrt{2\pi\tilde{\sigma}_t{}^2}} \exp\left(-\frac{(h_t-\mu_t^*)^2}{2\tilde{\sigma}_t{}^2}\right) \\
c &= CC'
\end{aligned}$$

である.A-R 法の採択確率は

$$\frac{f(h_t|\cdot)}{cg(h_t)} = \frac{f^*(h_t)}{g^*(h_t)}$$

である.

ここで, 平均 μ_t^*, 分散 $\tilde{\sigma}_t{}^2$ の正規分布からサンプリングして確率 $\frac{f^*(h_t)}{g^*(h_t)}$ で採択し, 確率 $1-\frac{f^*(h_t)}{g^*(h_t)}$ で棄却すれば, 条件付き密度 $f(h_t|\cdot)$ から h_t がサンプリングできたことになる.

また, テイラー展開をする点 h_t^* の選択については, 効率的にサンプリングを行うために,$g(h_t)$ のモード近辺で $\frac{f^*(h_t)}{g^*(h_t)}$ がなるべく大きくなるように選択すべきである. そのためには, なんらかの方法により $f^*(h_t)$ のピークにほぼ対応する h_t を求めて, それと $g(h_t)$ のピークが一致するように h_t^* を求めればよい. すなわち,

$$\hat{h}_t \approx \arg\max_{h_t} f^*(h_t) \tag{62}$$

とすると $\hat{h_t} = \mu_t^*$ より

$$h_t^* = -\log \frac{2\sigma_r^2}{\epsilon_t^2}\left(\frac{\hat{h_t} - \mu_t}{\tilde{\sigma_t}^2} + \frac{1}{2}\right) \tag{63}$$

とすればよい.

3.2.3 Gibbs sampler(multi-move sampler の場合)

Shephard and Pitt(1997) によれば, $\{h_t\}_{t=1}^T$ のサンプリングに single-move sampler を用いると ϕ が 1 に近い場合 Gibbs sampler の収束速度が遅い．ふつうボラティリティは persistent，すなわち ϕ が 1 に近いので，この問題は無視できない．そこで，Shephard and Pitt(1997) は，相関の高い変数をひとまとめにサンプリングすることによって，Gibbs sampler の収束速度をあげる方法を提案した．いま，$\{h_t\}_{t=1}^T$ を L 個の区切りで分けることを考えると，Gibbs sampler は

第 1 ループ

$f(\phi|\sigma_r^{2(0)}, \sigma_\eta^{2(0)}, \{h_s^{(0)}\}_{s=1}^T, \{\epsilon_s\}_{s=1}^T)$ から $sampling \to \phi^{(1)}$

$f(\sigma_r^2|\phi^{(1)}, \sigma_\eta^{2(0)}, \{h_s^{(0)}\}_{s=1}^T, \{\epsilon_s\}_{s=1}^T)$ から $sampling \to \sigma_r^{2(1)}$

$f(\sigma_\eta^2|\phi^{(1)}, \sigma_r^{2(1)}, \{h_s^{(0)}\}_{s=1}^T, \{\epsilon_s\}_{s=1}^T)$ から $sampling \to \sigma_\eta^{2(1)}$

$f(\{h_s\}_{s=1}^{k_1}|\theta^{(1)}, \{h_s^{(0)}\}_{s=k_1+1}^T, \{\epsilon_s\}_{s=1}^T)$ から $sampling \to \{h_s^{(1)}\}_{s=1}^{k_1}$

$f(\{h_s\}_{s=k_1+1}^{k_2}|\theta^{(1)}, \{h_s^{(1)}\}_{s=1}^{k_1}, \{h_s^{(0)}\}_{s=k_2+1}^T, \{\epsilon_s\}_{s=1}^T)$ から $sampling$

$\to \{h_s^{(1)}\}_{s=k_1+1}^{k_2}$

$\vdots$

$f(\{h_s\}_{s=k_L+1}^{T}|\theta^{(1)}, \{h_s^{(1)}\}_{s=1}^{k_L}, \{\epsilon_s\}_{s=1}^T)$ から $sampling \to \{h_s^{(1)}\}_{s=k_L+1}^{T}$

$\vdots$

第 n ループ

$f(\phi|\sigma_r^{2(n-1)}, \sigma_\eta^{2(n-1)}, \{h_s^{(n-1)}\}_{s=1}^T, \{\epsilon_s\}_{s=1}^T)$ から $sampling \to \phi^{(n)}$

$f(\sigma_r^2|\phi^{(n)}, \sigma_\eta^{2(n-1)}, \{h_s^{(n-1)}\}_{s=1}^T, \{\epsilon_s\}_{s=1}^T)$ から $sampling \to \sigma_r^{2(n)}$

$f(\sigma_\eta^2|\phi^{(n)}, \sigma_r^{2(n)}, \{h_s^{(n-1)}\}_{s=1}^T, \{\epsilon_s\}_{s=1}^T)$ から $sampling \to \sigma_\eta^{2(n)}$

$f(\{h_s\}_{s=1}^{k_1}|\theta^{(n)}, \{h_s^{(n-1)}\}_{s=k_1+1}^T, \{\epsilon_s\}_{s=1}^T)$ から $sampling \to \{h_s^{(n)}\}_{s=1}^{k_1}$

$f(\{h_s\}_{s=k_1+1}^{k_2}|\theta^{(n)}, \{h_s^{(n)}\}_{s=1}^{k_1}, \{h_s^{(n-1)}\}_{s=k_2+1}^T, \{\epsilon_s\}_{s=1}^T)$ から $sampling$

$\to \{h_s^{(n)}\}_{s=k_1+1}^{k_2}$

$\vdots$

$f(\{h_s\}_{s=k_L+1}^T|\theta^{(n)}, \{h_s^{(n)}\}_{s=1}^{k_L}, \{\epsilon_s\}_{s=1}^T)$ から $sampling \to \{h_s^{(n)}\}_{s=k_L+1}^T$

とすればよい．ただし，ここで $\theta^{(i)} = (\phi^{(i)}, \sigma_r^{2(i)}, \sigma_\eta^{2(i)})$ である. あとは,3.2.1 節と同様に事後平均をとればよい.

3.2.4 $\{h_t\}_{t=1}^T$ の multi-move sampler による sampling

いま，$\{h_s\}_{s=t}^{t+k}$ を 1 つのブロックと考えるとし，その 1 ブロックを

$$f(\{h_s\}_{s=t}^{t+k}|\{h_s\}_{s=1}^{t-1}, \{h_s\}_{s=t+k+1}^T, \{\epsilon_s\}_{s=1}^T, \theta)$$

から同時に sampling したいが，$\{h_s\}_{s=t}^{t+k}$ は時間を通じてそれぞれ依存しているので，これは容易ではない．そこで，Shephard and Pitt(1997) は $\{h_s\}_{s=t}^{t+k}$ ではなく $\{\eta_s\}_{s=t}^{t+k}$ を

$$f(\{\eta_s\}_{s=t}^{t+k}|\{h_s\}_{s=1}^{t-1}, \{h_s\}_{s=t+k+1}^T, \{\epsilon_s\}_{s=1}^T, \theta) \tag{64}$$

からサンプリングするという方法を提案している．h_{t-1} がわかっていれば，$\{\eta_s\}_{s=t}^{t+k}$ がサンプリングされれば，(49) から $\{h_s\}_{s=t}^{t+k}$ が計算

できる．条件付き密度 (64) の対数をとって計算すると
$\left(ここで, l(h_s) \equiv - \left[\frac{h_s}{2} + \frac{\epsilon_s^2}{2\sigma_r^2} \exp(-h_s) \right] \right)$

$$
\begin{aligned}
&\log f(\{\eta_s\}_{s=t}^{t+k} | \{h_s\}_{s=1}^{t-1}, \{h_s\}_{s=t+k+1}^{T}, \{\epsilon_s\}_{s=1}^{T}, \theta) \\
&= 定数 + \log f(\{\eta_s\}_{s=t}^{t+k}) + \log f(\{\epsilon_s\}_{s=t}^{t+k} | h_s{}_{s=t}^{t+k}, \theta) \\
&= 定数 + \sum_{s=t}^{t+k} \log f(\eta_s) + \sum_{s=t}^{t+k} \log f(\epsilon_s | h_s, \sigma_r^2) \\
&= 定数 - \frac{1}{2\sigma_\eta^2} \sum_{s=t}^{t+k} \eta_s^2 - \sum_{s=t}^{t+k} \left[\frac{h_s}{2} + \frac{\epsilon_s^2}{2\sigma_r^2} \exp(-h_s) \right] \\
&= 定数 - \frac{1}{2\sigma_\eta^2} \sum_{s=t}^{t+k} \eta_s^2 + \sum_{s=t}^{t+k} l(h_s) \qquad (65)
\end{aligned}
$$

(65) の最後の項を $\{\hat{h_s}\}_{s=t}^{t+k}$ まわりで 2 次のテイラー展開をして ($\{\hat{h_s}\}_{s=t}^{t+k}$ の選択についてはあとで説明する)，さらに g を以下のように定義する．

$$
\begin{aligned}
&\approx 定数 - \frac{1}{2\sigma_\eta^2} \sum_{s=t}^{t+k} \eta_s^2 + \sum_{s=t}^{t+k} \left[l(\hat{h_s}) + (h_s - \hat{h_s}) l'(\hat{h_s}) + \frac{1}{2} (h_s - \hat{h_s})^2 l''(\hat{h_s}) \right] \\
&= 定数 - \frac{1}{2\sigma_\eta^2} \sum_{s=t}^{t+k} \eta_s^2 + \sum_{s=t}^{t+k} \frac{1}{2} l''(\hat{h_s}) \left[\hat{h_s} - \frac{l'(\hat{h_s})}{l''(\hat{h_s})} - h_s \right]^2 \\
&\equiv \log g
\end{aligned}
$$

ここで,

$$
\begin{aligned}
l'(\hat{h_s}) &= \frac{1}{2} \left(\frac{\epsilon_s^2}{2\sigma_r^2} \exp(-\hat{h_s}) - 1 \right) \\
l''(\hat{h_s}) &= -\frac{\epsilon_s^2}{2\sigma_r^2} \exp(-\hat{h_s})
\end{aligned}
$$

である．また,

$$
y_s = \hat{h_s} - \frac{l'(\hat{h_s})}{l''(\hat{h_s})}, \qquad s = t, \cdots, t+k \qquad (66)
$$

と定義して，この y_s を観測される変数，h_s を状態変数とする次のような線型状態空間モデルを考える．

$$y_s = h_s + \xi_s, \qquad \xi_s \sim N\left(0, -\frac{1}{l''(\hat{h}_s)}\right) \tag{67}$$

$$h_s = \phi h_{s-1} + \eta_s, \qquad \eta_s \sim N(0, \sigma_\eta^2) \tag{68}$$

これに，de Jong and Shephard(1995) によって提案された simulation smoother を適用すると，$\{\eta_s\}_{s=t}^{t+k}$ を g からサンプリングすることができる．その手順は以下のとおりである．

まず，$a_t = \phi h_{t-1}, P_t = \sigma_\eta^2$ からスタートして次のフィルタを $i = 0, \cdots, k$ の順に逐次的に解くことにより，$e_{t+i}, D_{t+i}, K_{t+i}, (i = 0, \cdots, k)$ を計算する．

$$e_{t+i} = y_{t+i} - a_{t+i} \tag{69}$$

$$D_{t+i} = P_{t+i} - \frac{1}{l''(\hat{h}_{t+i})} \tag{70}$$

$$K_{t+i} = \frac{\phi P_{t+i}}{D_{t+i}} \tag{71}$$

$$L_{t+i} = \phi - K_{t+i} \tag{72}$$

$$a_{t+i+1} = \phi a_{t+i} + K_{t+i} e_{t+i} \tag{73}$$

$$P_{t+i+1} = \phi P_{t+i} + \sigma_\eta^2 \tag{74}$$

次に，$U_{t+k} = 0, r_{t+k} = 0$ からスタートして，上のフィルタで計算された $e_{t+i}, D_{t+i}, K_{t+i}, (i = 0, \cdots, k)$ を使って，次のような平滑化を

$i=k,\cdots,0$ の順に逐次的に解けば，$\{\eta_s\}_{s=t}^{t+k}$ をサンプリングできる．

$$\begin{aligned}
C_{t+i} &= 1-U_{t+i} && (75)\\
\xi_{t+i} &\sim N(0,C_{t+i}) && (76)\\
\eta_{t+i} &= r_{t+i}+\xi_{t+i} && (77)\\
V_{t+i} &= U_{t+i}L_{t+i} && (78)\\
r_{t+i-1} &= \frac{e_{t+i}}{D_{t+i}}+L_{t+i}r_{t+i}-\frac{V_{t+i}\xi_{t+i}}{C_{t+i}} && (79)\\
U_{t+i-1} &= \frac{1}{D_{t+i}}+L_{t+i}^2U_{t+i}+\frac{V_{t+i}^2}{C_{t+i}} && (80)
\end{aligned}$$

ただし，(76) は ξ_{t+i} を平均 0, 分散 C_{t+i} の正規分布からサンプリングするという意味である．

このようにして g からサンプリングされた $\{\eta_s\}_{s=t}^{t+k}$ は，M-R/A-R 法を用いることによって，正しい密度 f からサンプリングされたことになる．その手順は以下のとおりである．

(1) 適当な初期値 $\{\eta_s^0\}_{s=t}^{t+k}$ を用意する．

(2)g からサンプリングされた $\{\eta_s\}_{s=t}^{t+k}$ を確率 $\min\left[\frac{f(\{\eta_s^0\}_{s=t}^{t+k})}{g(\{\eta_s^0\}_{s=t}^{t+k})},1\right]$ で採択し，確率 $1-\min\left[\frac{f(\{\eta_s^0\}_{s=t}^{t+k})}{g(\{\eta_s^0\}_{s=t}^{t+k})},1\right]$ で棄却する．採択されれば (3) に進み，棄却されればやり直す．

(3)

$$\alpha(\{\eta_s^0\}_{s=t}^{t+k},\{\eta_s\}_{s=t}^{t+k})=$$
$$\begin{cases}
1, & f(\{\eta_s^0\}_{s=t}^{t+k})\le g(\{\eta_s^0\}_{s=t}^{t+k})\\
\dfrac{g(\{\eta_s^0\}_{s=t}^{t+k})}{f(\{\eta_s^0\}_{s=t}^{t+k})}, & \begin{cases} f(\{\eta_s^0\}_{s=t}^{t+k})>g(\{\eta_s^0\}_{s=t}^{t+k})\\ f(\{\eta_s\}_{s=t}^{t+k})\ge g(\{\eta_s\}_{s=t}^{t+k})\end{cases}\\
\min\left[\dfrac{f(\{\eta_s\}_{s=t}^{t+k})g(\{\eta_s^0\}_{s=t}^{t+k})}{f(\{\eta_s^0\}_{s=t}^{t+k})g(\{\eta_s\}_{s=t}^{t+k})},1\right], & \begin{cases} f(\{\eta_s^0\}_{s=t}^{t+k})>g(\{\eta_s^0\}_{s=t}^{t+k})\\ f(\{\eta_s\}_{s=t}^{t+k})<g(\{\eta_s\}_{s=t}^{t+k})\end{cases}
\end{cases}$$

を計算する．

(4)(2) で得られた $\{\eta_s\}_{s=t}^{t+k}$ を確率 $\alpha(\{\eta_s^0\}_{s=t}^{t+k}, \{\eta_s\}_{s=t}^{t+k})$ で採択し，確率 $1-\alpha(\{\eta_s^0\}_{s=t}^{t+k}, \{\eta_s\}_{s=t}^{t+k})$ で棄却する．

次に，テイラー展開を行う点 $\{\hat{h_s}\}_{s=t}^{t+k}$ の選択について説明する．$\{\hat{h_s}\}_{s=t}^{t+k}$ に適当な初期値を選ぶと (76) から $\{y_s\}_{s=t}^{t+k}$ が計算できる．それを使って (67)，(68) の線型状態空間モデルにカルマンフィルタを適用して状態変数 $\{h_s\}_{s=t}^{t+k}$ の平滑化推定量を求める．すなわち，$h_{t|t-1} = \phi h_{t-1}, P_{t|t-1} = \sigma_\eta^2$ からスタートして次のフィルタを $i = 0, \cdots, k$ の順に逐次的に解くことにより，$h_{t+i|t+i-1}, P_{t+i|t+i-1}, h_{t+i|t+i}, P_{t+i|t+i}, (i = 0, \cdots, k)$ を計算して，

$$
\begin{aligned}
h_{t+i|t+i-1} &= \phi h_{t+i-1|t+i-1} && (81)\\
P_{t+i|t+i-1} &= \phi^2 P_{t+i-1|t+i-1} + \sigma_\eta^2 && (82)\\
h_{t+i|t+i} &= h_{t+i|t+i-1} + \frac{P_{t+i|t+i-1}}{F_{t+i}} \nu_{t+i} && (83)\\
P_{t+i|t+i} &= P_{t+i|t+i-1} - \frac{P_{t+i|t+i-1}^2}{F_{t+i}} && (84)\\
\nu_{t+i} &= y_{t+i} - h_{t+i|t+i-1} && (85)\\
F_{t+i} &= P_{t+i|t+i-1}^2 - \frac{1}{l''(\hat{h}_{t+i})} && (86)
\end{aligned}
$$

上のフィルタで計算された $h_{t+i|t+i-1}, P_{t+i|t+i-1}, h_{t+i|t+i}, P_{t+i|t+i}, (i = 0, \cdots, k)$ を使って，次のような平滑化を $i = k, \cdots, 0$ の順に逐次的に解けば，平滑化推定量 $h_{t+i|t+k}, (i = 0, \cdots, k)$ を計算することがで

きる．

$$
\begin{aligned}
h_{t+i|t+k} &= h_{t+i|t+i} + P^*_{t+i}(h_{t+i+1|t+k} - h_{t+i+1|t+i}) && (87)\\
P_{t+i|t+k} &= P_{t+i|t+i} + P^{*2}_{t+i}(P_{t+i+1|t+k} - P_{t+i+1|t+i}) && (88)\\
P^*_{t+i} &= \phi \frac{P_{t+i|t+i}}{P_{t+i+1|t+i}} && (89)
\end{aligned}
$$

この $h_{t+i|t+k}, (i = 0, \cdots, k)$ を (76) に代入して新たな $\{y_s\}_{s=t}^{t+k}$ が計算できる．それを使って再び (67)，(68) の線型状態空間モデルにカルマンフィルタを適用して状態変数 $\{h_s\}_{s=t}^{t+k}$ の新たな平滑化推定量を求める．これを数回繰り返して得られる $\{h_s\}_{s=t}^{t+k}$ の平滑化推定量を $\{\hat{h}_s\}_{s=t}^{t+k}$ にする．

最後に，ブロッキングを行う際，どこで区切るかであるが，Shephard and Pitt(1997) は，区切る点をランダムに選んでいる．具体的には U_i を $[0,1]$ の一様乱数とし，

$$
k_i = int\left[T \times \frac{i + U_i}{L + 2}\right], \qquad i = 1, \cdots, L
$$

としている．ただし，$int[x]$ は x に最も近い整数である．このようにループ毎にブロッキングの区切りをランダムに選ぶと，あるループのサンプリングで非常に多く棄却されるブロックがあったとしても，次のループでのサンプリングでは異なるブロックが選ばれるので，棄却が続いてサンプリングが行き詰まるということを排除できる．

3.3 シミュレーション実験

前節で説明した MCMC によるベイズ推定のシミュレーション実験をおこなった[9]．まず, パラメータの真の値を設定し,(48),(49) から

データ $\{\epsilon_s\}_{s=1}^T$ を発生させてそれを使って，事後平均を計算した.$\{h_s\}_{s=1}^T$ のサンプリングには single-move sampler を用いている[10]．データ数は 50，Gibbs sampler のループの回数は 60000 回で，最初の 50000 回は捨てて，最後の 10000 回のサンプリングで得られた値の平均を推定値としている．Gibbs sampler の収束判定のために CD(convergence diagnostic) 統計量[11] を使っている．結果は表 1 である. いずれのパラメータも収束判定をパスしている.

表 1: 推定結果

パラメータ	真の値	平均	標準誤差	90 %信頼区間	CD
ϕ	0.90	0.9227	0.0302	$[0.8717, 0.9722]$	0.0137
σ_η^2	0.01	0.0084	0.0017	$[0.0059, 0.0117]$	0.4741
σ_r^2	1.00	1.2376	0.2598	$[0.8799, 1.7051]$	0.4180

4　おわりに

最後に今後の課題について述べる.SV モデルではもっぱらその推定方法が研究対象として注目されてきて,ARCH 型モデルのようにモデルの改良はあまりおこなわれてこなかった. これは SV モデルの推定が ARCH 型モデルの推定のように簡単でないことが原因であるが,推定方法の改善と普及によって今後はモデル自体の改良も進むと考えられる. また,SV モデルによる実証分析も同じ理由で今後は増えてくるであろう. また, 紹介した方法は GMM のような尤度に基づかない方法,QMLE のような最尤法ではあるが正しい尤度に基づいていない方法，MCMC によるベイズ推定であったが，コンピュータの処理

能力の改善にともない, 正確な尤度を計算し, それに基づいて SV モデルの推定をおこなおうとする方法も提案されている. これらの推定方法を比較検討し，さらに推定のパフォーマンスと使いやすさを改善するのが今後の課題である.

5　付録

5.1　ボラティリティの変動と尖度

ここでは簡単化のために σ_t と z_t が互いに独立である場合について考える. まず,

$$E(\epsilon_t^2|\sigma_t^2) = \sigma_t^2 E(z_t^2|\sigma_t^2) = \sigma_t^2 E(z_t^2) = \sigma_t^2$$

より,

$$E(\sigma_t^2) = E(E(\epsilon_t^2|\sigma_t^2)) = E(\epsilon_t^2)$$

である. これと Jensen の不等式より

$$\begin{aligned} E(\epsilon_t^4) &= E(\sigma_t^4)E(z_t^4) \\ &= 3E((\sigma_t^2)^2) \geq 3(E(\sigma_t^2))^2 \\ &= 3(E(\epsilon_t^2))^2 \end{aligned}$$

よって,

$$\frac{E(\epsilon_t^4)}{(E(\epsilon_t^2))^2} \geq 3$$

である.

5.2　Gibbs sampler

$(\theta_1, \theta_2, \theta_3)$ という 3 つの確率変数を同時密度 $f(\theta_1, \theta_2, \theta_3)$ からサンプリングしたいが，$f(\theta_1, \theta_2, \theta_3)$ からは直接サンプリングできないという状況を考える．ただし，$f(\theta_1|\theta_2, \theta_3)$，$f(\theta_2|\theta_3, \theta_1)$，$f(\theta_3|\theta_1, \theta_2)$ か

らはサンプリングできるとする．このような時，$(\theta_1, \theta_2, \theta_3)$ を同時密度 $f(\theta_1, \theta_2, \theta_3)$ からサンプリングする方法が Gibbs sampler である．その手順は以下のとおりである．

まず，適当な初期値 $(\theta_2^{(0)}, \theta_3^{(0)})$ を決めて $f(\theta_1|\theta_2^{(0)}, \theta_3^{(0)})$ からサンプリングし，その値を $\theta_1^{(1)}$ とする．次に，新たに得られた $\theta_1^{(1)}$ を使って $f(\theta_2|\theta_3^{(0)}, \theta_1^{(1)})$ からサンプリングし，その値を $\theta_2^{(1)}$ とする．さらに $\theta_1^{(1)}$，$\theta_2^{(1)}$ を使って $f(\theta_3|\theta_1^{(1)}, \theta_2^{(1)})$ からサンプリングし，その値を $\theta_3^{(1)}$ とする．

以上を第 1 ループをよび，同様に $(\theta_2^{(1)}, \theta_3^{(1)})$ を使って第 2 ループをはじめる．第 2 ループが終われば，$(\theta_1^{(2)}, \theta_2^{(2)}, \theta_3^{(2)})$ が得られる．このようなループを繰り返すと第 n ループでは，$(\theta_1^{(n)}, \theta_2^{(n)}, \theta_3^{(n)})$ が得られる．$n \to \infty$ のとき，$(\theta_1^{(n)}, \theta_2^{(n)}, \theta_3^{(n)})$ は同時密度 $f(\theta_1, \theta_2, \theta_3)$ からサンプリングされた確率変数に分布収束することが知られている．

5.3 A-R 法

密度関数 $f(x)$ からサンプリングしたいが，直接サンプリングできない状況を考える．このようなときに，正規分布などサンプリングが簡単な確率分布 (密度関数を $g(x)$ とする) からのサンプリングを利用して，$f(x)$ からサンプリングする方法が A-R(acceptance-rejection; 受容・棄却) 法である．もし，“ $g(x)$ がすべての x について $f(x) \leq cg(x)$(ただし c は定数) ”であるならば，以下の手順で $f(x)$ からサンプリングできる．

(1)$g(x)$ から乱数 x^*，$[0, 1]$ 区間上の一様乱数 u を発生させる．

(2) 確率 $\dfrac{f(x^*)}{cg(x^*)}$ で x^* を $f(x)$ からのサンプルとして採択 (受容) し，確率 $1-\dfrac{f(x^*)}{cg(x^*)}$ で棄却して (1) に戻る．すなわち，$u \leq \dfrac{f(x^*)}{cg(x^*)}$ ならば x^* を採択し，$u > \dfrac{f(x^*)}{cg(x^*)}$ ならば棄却して (1) に戻る．

この方法は,

$$
\begin{aligned}
Pr\left(x \middle| u \leq \frac{f(x)}{cg(x)}\right) &= \frac{Pr\left(u \leq \dfrac{f(x)}{cg(x)} \middle| x\right) g(x)}{\displaystyle\int Pr\left(u \leq \dfrac{f(x)}{cg(x)} \middle| x\right) g(x)dx} \\
&= \frac{\dfrac{f(x)}{cg(x)} g(x)}{\displaystyle\int \dfrac{f(x)}{cg(x)} g(x)dx} = \frac{f(x)}{\int f(x)dx} \\
&= f(x)
\end{aligned}
$$

であることによる．また，効率的にサンプリングをおこなうには $g(x)$ のモード近辺で採択確率 $\dfrac{f(x)}{cg(x)}$ がなるべく大きくなるように $g(x)$ と c を選択する必要がある．

5.4 M-H/A-R 法

付録 5.3 の A-R 法では,“ すべての x について $f(x) \leq cg(x)$ ” を満たす $g(x)$ を探さなければならなかった．そのような条件が満たされないときに，それでも直接サンプリングできない $f(x)$ からサンプリングする方法が M-H/A-R(Metropolis-Hasting/acceptance-rejection) 法である．付録 5.3 と同様に，直接サンプリングできる確率分布の密度関数を $g(x)$ とする．このとき，M-H/A-R 法の手順は以下のとおりである．ここでは，$f(x)$ から N 個の値 $(X_1, \cdots, X_N)$ をサンプリングすることを考える．

(1)$n=1$ のときは適当な初期値を決めそれを x とする．$n>1$ であれば，$x=X_{n-1}$ とする．

(2)A-R 法を採択確率 $\min\left[\dfrac{f(x)}{cg(x)},1\right]$ で実行し，サンプリングされた値 y を X_n の候補にする．

(3)

$$\alpha(x,y)=\begin{cases} 1, & f(x)\leq cg(x) \\ \dfrac{cg(x)}{f(x)}, & \begin{cases} f(x)>cg(x) \\ f(y)\geq cg(y) \end{cases} \\ \min\left[\dfrac{f(y)g(x)}{f(x)g(y)},1\right], & \begin{cases} f(x)>cg(x) \\ f(y)<cg(y) \end{cases} \end{cases}$$

を計算する．

(4)(2) で得られた y を確率 $\alpha(x,y)$ で採択し，確率 $1-\alpha(x,y)$ で棄却する．採択された場合は $X_n=y$，棄却された場合は $X_n=x$ とする．

(5)$n<N$ であれば，$n=n+1$ として (2) に戻る．$n=N$ ならば終了．

すべての値で $f(x)\leq cg(x)$ が成り立っているとき，$\alpha(x,y)=1$ であるから，通常の A-R 法である．

5.5　SV モデルのパラメータの条件付き分布

ここでは,h_1 の事前分布には一様分布を仮定し,$f(h_1|\theta) \propto c$(正の定数) とする.

$$
\begin{aligned}
& f(\phi|\cdot) \propto f(\theta, \{h_s\}_{s=1}^T, \{\epsilon_s\}_{s=1}^T) \\
& = f(\theta) f(\{h_s\}_{s=1}^T|\theta) f(\{\epsilon_s\}_{s=1}^T|\{h_s\}_{s=1}^T, \theta) \\
& = f(\phi) f(\sigma_\eta^2) f(\sigma_r^2) f(h_T|h_{T-1}, \theta) \cdots f(h_1|\theta) f(\{\epsilon_s\}_{s=1}^T|\{h_s\}_{s=1}^T, \sigma_r^2) \\
& \propto f(\phi) \prod_{t=2}^T \exp\left(-\frac{(h_t - \phi h_{t-1})^2}{2\sigma_\eta^2}\right) \\
& \propto \exp\left(-\frac{\left(\phi - \frac{\sum_{t=2}^T h_t h_{t-1}}{\sum_{t=2}^T h_t^2}\right)^2}{\frac{2\sigma_\eta^2}{\sum_{t=2}^T h_t^2}}\right) I[-1, 1]
\end{aligned}
$$

$$
\begin{aligned}
& f(\sigma_\eta^{-2}|\cdot) = f(\sigma_\eta^2|\cdot)\left|\frac{d\sigma_\eta^2}{d\sigma_\eta^{-2}}\right| \\
& \propto f(\theta, \{h_s\}_{s=1}^T, \{\epsilon_s\}_{s=1}^T)\sigma_\eta^4 \\
& = f(\theta) f(\{h_s\}_{s=1}^T|\theta) f(\{\epsilon_s\}_{s=1}^T|\{h_s\}_{s=1}^T, \theta)\sigma_\eta^4 \\
& = f(\phi) f(\sigma_\eta^2) f(\sigma_r^2) f(h_T|h_{T-1}, \theta) \cdots f(h_1|\theta) f(\{\epsilon_s\}_{s=1}^T|\{h_s\}_{s=1}^T, \sigma_r^2)\sigma_\eta^4 \\
& \propto f(\sigma_\eta^2)\sigma_\eta^4 \prod_{t=2}^T \frac{1}{\sigma_\eta} \exp\left(-\frac{(h_t - \phi h_{t-1})^2}{2\sigma_\eta^2}\right) \\
& = (\sigma_\eta^{-2})^{\frac{T-1}{2}-1} \exp\left(-\frac{\sigma_\eta^{-2}}{\frac{2}{\sum_{t=2}^T (h_t - \phi h_{t-1})^2}}\right)
\end{aligned}
$$

$$
\begin{aligned}
&f(\sigma_r^{-2}|\cdot) = f(\sigma_r^2|\cdot)\left|\frac{d\sigma_r^2}{d\sigma_r^{-2}}\right| \\
&\propto f(\theta, \{h_s\}_{s=1}^T, \{\epsilon_s\}_{s=1}^T)\sigma_r^4 \\
&= f(\theta)f(\{h_s\}_{s=1}^T|\theta)f(\{\epsilon_s\}_{s=1}^T|\{h_s\}_{s=1}^T, \theta)\sigma_r^4 \\
&= f(\phi)f(\sigma_\eta^2)f(\sigma_r^2)f(h_T|h_{T-1}, \phi, \sigma_\eta^2)\cdots f(h_1|\theta)f(\{\epsilon_s\}_{s=1}^T|\{h_s\}_{s=1}^T, \sigma_r^2)\sigma_r^4 \\
&\propto f(\sigma_r^2)\sigma_r^4 \prod_{t=1}^T \frac{1}{\sigma_r \exp(h_t/2)} \exp\left(-\frac{\epsilon_t^2}{2\sigma_r^2 \exp(h_t)}\right) \\
&\propto (\sigma_r^{-2})^{\frac{T}{2}-1} \exp\left(-\frac{\sigma_r^{-2}}{\frac{2}{\sum_{t=1}^T \epsilon_t^2 \exp(-h_t)}}\right)
\end{aligned}
$$

5.6 SV モデルの潜在変数 h_t の条件付き分布

付録 5.5 同様,h_1 の事前分布には一様分布を仮定し,$f(h_1|\theta) \propto c$(正の定数) である.

$$
\begin{aligned}
&f(h_t|\theta, \{h_s\}_{s\neq t}, \{\epsilon_s\}_{s=1}^T) \\
&= f(\{\epsilon_s\}_{s=1}^T|\theta, \{h_s\}_{s=1}^T)\frac{f(\theta, \{h_s\}_{s=1}^T)}{f(\theta, \{h_s\}_{s\neq t}, \{\epsilon_s\}_{s=1}^T)} \\
&= f(\{\epsilon_s\}_{s=1}^T|\theta, \{h_s\}_{s=1}^T)f(h_t|\theta, \{h_s\}_{s\neq t})\frac{f(\theta, \{h_s\}_{s\neq t})}{f(\theta, \{h_s\}_{s\neq t}, \{\epsilon_s\}_{s=1}^T)} \\
&\propto f(\epsilon_1|\theta, h_1)\cdots f(\epsilon_T|\theta, h_T)\frac{f(\{h_s\}_{s=1}^T|\theta)}{f(\{h_s\}_{s\neq t}|\theta)} \\
&\propto f(\epsilon_t|\theta, h_t)f(\{h_s\}_{s=1}^T|\theta) \\
&= f(\epsilon_t|\theta, h_t)f(h_T|h_{T-1}, \theta)\cdots f(h_1|\theta) \\
&\propto \begin{cases} f(\epsilon_1|\theta, h_1)f(h_2|h_1, \theta), & t = 1 \\ f(\epsilon_t|\theta, h_t)f(h_t|h_{t-1}, \theta)f(h_{t+1}|h_t, \theta), & 2 \leq t \leq T-1 \\ f(\epsilon_T|\theta, h_T)f(h_T|h_{T-1}, \theta), & t = T \end{cases}
\end{aligned}
$$

$t=1$ のとき

$$
\begin{aligned}
&f(h_t|\theta,\{h_s\}_{s\neq t},\{\epsilon_s\}_{s=1}^T)\\
&\propto \frac{1}{\sigma_r\exp(h_1/2)}\exp\left(-\frac{\epsilon_1^2}{2\sigma_r^2\exp(h_1)}\right)\frac{1}{\sigma_\eta}\exp\left(-\frac{(h_2-\phi h_1)^2}{2\sigma_\eta^2}\right)\\
&\propto \exp\left(-\frac{h_1}{2}\right)\exp\left(-\frac{\epsilon_1^2}{2\sigma_r^2}\exp(-h_1)\right)\exp\left(-\frac{(h_1-\frac{h_2}{\phi})^2}{\frac{2\sigma_\eta^2}{\phi^2}}\right)
\end{aligned}
$$

$2\le t\le T-1$ のとき

$$
\begin{aligned}
&f(h_t|\theta,\{h_s\}_{s\neq t},\{\epsilon_s\}_{s=1}^T)\\
&\propto \exp\left(-\frac{h_t}{2}\right)\exp\left(-\frac{\epsilon_t^2}{2\sigma_r^2\exp(h_t)}\right)\exp\left(-\frac{(h_t-\phi h_{t-1})^2}{2\sigma_\eta^2}\right)\exp\left(-\frac{(h_{t+1}-\phi h_t)^2}{2\sigma_\eta^2}\right)\\
&\propto \exp\left(-\frac{h_1}{2}\right)\exp\left(-\frac{\epsilon_t^2}{2\sigma_r^2}\exp(-h_t)\right)\exp\left(-\frac{\left(h_t-\frac{\phi(h_{t-1}+h_{t+1})}{1+\phi^2}\right)^2}{\frac{2\sigma_\eta^2}{1+\phi^2}}\right)
\end{aligned}
$$

$t=T$ のとき

$$
\begin{aligned}
&f(h_t|\theta,\{h_s\}_{s\neq t},\{\epsilon_s\}_{s=1}^T)\\
&\propto \frac{1}{\sigma_r\exp(h_T/2)}\exp\left(-\frac{\epsilon_T^2}{2\sigma_r^2\exp(h_T)}\right)\frac{1}{\sigma_\eta}\exp\left(-\frac{(h_T-\phi h_{T-1})^2}{2\sigma_\eta^2}\right)\\
&\propto \exp\left(-\frac{h_T}{2}\right)\exp\left(-\frac{\epsilon_T^2}{2\sigma_r^2}\exp(-h_T)\right)\exp\left(-\frac{(h_T-\phi h_{T-1})^2}{2\sigma_\eta^2}\right)
\end{aligned}
$$

6　脚注

1　本論文に対して，京都大学大学院経済学研究科森棟公夫教授及び京都大学大学院経済学研究科鍵原理人氏より有益なアドバイスを頂いた．ここに記して感謝の意を表したい．

2

有限の期待収益率と条件付き分散

$$\begin{aligned} E_{WqY}\left(\frac{dS}{S}\right) &= E_{qY}[(\alpha-\lambda k)dt + dq] \\ &= E_Y[(\alpha-\lambda k)dt + (Y-1)\lambda dt] = \alpha dt \\ V_{WqY}\left(\frac{dS}{S}\right) &= E_{WqY}[-\lambda k dt + dW + dq]^2 \\ &\approx E_{qY}[\sigma^2 dt + (dq)^2 - 2\lambda k dt dq] \\ &\approx (\sigma^2 + E_Y(Y-1)^2)dt \end{aligned}$$

をもつ. ここで E_{WqY} は I_t を所与とした dW_t, dq_t, Y に関する期待値である.

3

イールドとは $[t, t+h]$ における平均の利子率で

$$v(t, t+h) = \exp(-hY(t, t+h))$$

すなわち,

$$Y(t, t+h) = -\frac{1}{h}\log v(t, t+h)$$

である.

4

ここでの定義は，山本 (1988) の定義の十分条件になっているので，強 Granger の因果関係とよぶことにした．山本 (1988) の Granger の因果関係の定義は以下のとおりである．ある情報集合を用いて x_t の予測を行ったときに，その予測の平均 2 乗誤差 (predicton mean squared error:PMSE) が最小になるものを最適予測とよび，そのときの平均 2 乗誤差を $\sigma^2(x_t|I_{t-1})$ や $\sigma^2(x_t|I_{t-1} - \{z_s\}_{s=0}^{t-1})$ などと書くことにする．すなわち，$\sigma^2(x_t|I_{t-1})$ は $t-1$ 時点までの利用可能な全ての情報を使って最適予測を行ったときの PMSE であり，$\sigma^2(x_t|I_{t-1} - \{z_s\}_{s=0}^{t-1})$ は $t-1$ 時点までの $\{z_s\}_{s=0}^{t-1}$ 以外の利用可能な全ての情報を用いて最適予測を行ったときの PMSE である．

定義 6.1 (Granger の因果関係)

$$\sigma^2(x_t|I_{t-1}) = \sigma^2(x_t|I_{t-1} - \{z_s\}_{s=0}^{t-1})$$

となるとき，*Granger* の意味で z_t から x_t への因果関係がないという．

このように Granger の因果関係は x_t の最適予測を行うとき過去の情報 z_t が PMSE の減少に貢献するかどうかを因果関係の判断基準にするものである.

5

すべての t にたいして，$\mu_t, \sigma_t, \gamma_t, \delta_t, \rho_t$ は I_t^U-可測であるということである.

6

このように z_t と η_t は独立を独立にしても,(46) 式に z_{t-1} を入れ

$$\begin{aligned} \epsilon_t &= \sigma_t z_t \\ \log \sigma_t^2 &= \omega + \phi \log \sigma_{t-1} + \lambda z_{t-1} + \eta_t \end{aligned}$$

とすることによって leverage effects をとらえることができる.

7

定常性を仮定して，ϕ の事前分布は，$[-1, 1]$ の間の一様分布を考えている．σ_η^2，σ_r^2 に関しては正の値をとることはわかっているので，対数をとった $\log \sigma_\eta^2$，$\log \sigma_r^2$ が $[-\infty, \infty]$ の間で一様分布すると考える．そのとき，σ_η^2，σ_r^2 の事前分布は

$$\begin{aligned} f(\sigma_\eta^2) = f(\log \sigma_r^2)\left|\frac{d \log \sigma_\eta^2}{d\sigma_\eta^2}\right| \propto \frac{1}{\sigma_\eta^2} \\ f(\sigma_r^2) = f(\log \sigma_r^2)\left|\frac{d \log \sigma_r^2}{d\sigma_r^2}\right| \propto \frac{1}{\sigma_r^2} \end{aligned}$$

である.

8

事後分布は $f(\phi, \sigma_r^2, \sigma_\eta^2, \{h_t\}_{t=1}^T | \{\epsilon_t\}_{t=1}^T)$ であり，パラメータに $\{h_t\}_{t=1}^T$ を含めた $(\phi, \sigma_r^2, \sigma_\eta^2, \{h_t\}_{t=1}^T)$ を Gibbs sampler でサンプリングする.

9

行列言語 GAUSS でプログラムをかいた.

10

(62) の最大化問題を解く際,$f^*(h_t)$ の対数をとった $\log f^*(h_t)$ を最大化している. 最大化の方法は,2 次導関数 $(\log f^*(h_t))''$ が常に負なので 2 分割法を用いた.

11

CD 統計量は，平均をとる 10000 個のサンプリングされた値 $\{\theta^{(i)}\}_{i=1}^{10000}$ のうちはじめの 3000 個 $\{\theta^{(i)}\}_{i=1}^{3000}$ と最後の 3000 個 $\{\theta^{(i)}\}_{i=7001}^{10000}$ の平均に有意な差があるかどうか調べるもので，$\{\theta^{(i)}\}_{i=1}^{3000}$ の平均を $\bar{\theta}_1$，標準偏差を $\hat{\sigma}_1$，$\{\theta^{(i)}\}_{i=7001}^{10000}$ の平均を $\bar{\theta}_2$，標準偏差を $\hat{\sigma}_2$ とする．CD 統計量は

$$CD = \frac{\bar{\theta}_1 - \bar{\theta}_2}{\sqrt{\frac{\hat{\sigma}_1^2}{3000} + \frac{\hat{\theta}_2^2}{3000}}}$$

で定義され, はじめの 3000 個 $\{\theta^{(i)}\}_{i=1}^{3000}$ と最後の 3000 個 $\{\theta^{(i)}\}_{i=7001}^{10000}$ の平均に有意な差はないという帰無仮説のもとで，標準正規分布に従う.

参考文献

[1] 山本拓, 『経済の時系列分析』, 創文社, (1988)

[2] 渡部敏明, 『ボラティリティ変動モデル』, 朝倉書店, (2000)

[3] Andersen, T.G. " Stochastic autoregressive volatility: A framework for volatility modeling "*Math.Finance*, Vol.4, (1994), pp75-102

[4] Andersen, T.G., Chung, H.J. and Sorensen, B.E. " Efficient Method of Moments Estimation of a Stochastic Volatility Model: A Monte Carlo Study "*Journal of Econometrics*, Vol.91, (1999), pp61-87

[5] Black, F. " Studies in stock price volatility changes "*Proceeding of the 1976 Business Meeting of the Business and Economic Statistics Section, Amer. Statist. Assoc.*, (1976), pp177-181

[6] Black, F. and Scholes, M." The pricing of options and corporate liabilities "*J.Politic.Econom.*, Vol.81, (1973), pp637-654

[7] Bollerslev, T., Engle, R.F., and Nelson, D.B. " ARCH Models "in R.F.Engle and D.L.McFadden(eds), *The Handbook of Econometrics*, Vol.4, North-Holland, (1994), pp2959-3038

[8] Chib, S., and Greenberg, E. " Understanding the Metropolis-Hastings Algorithim "*American Statistician*, Vol.49, (1995), pp327-335

[9] Clark, P. " A Subordinated Stochastic Process Model with Finite Variance for Speculative Process "*Econometrica*, Vol.41, (1973), pp135-155

[10] de Jong, P. and Shephard, G. " The Simulation Smoother for Time Series "*Biometrika*, Vol.82, (1995), pp339-350

[11] Engle, R.F." Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation "*Econometrica*, Vol.55, (1982), pp987-1007

[12] Ghysels, Harvey and Renault " Stochastic Volatility "in G.S.Maddala and D.L.Rao(eds), *The Handbook of Econometrics*, North-Holland, (1996), pp119-191

[13] Hull, J. and White, A. " The pricing of options on assets with stochastic volatilities "*J.Finance*, Vol.42, (1987), pp281-300

[14] Jacquier, E., Polson, N.G., and Rossi, P.E. " Bayesian analysis of stochastic volatility models (with discussion) "*J.Business Econom.Statist.*, Vol.12, (1994), pp371-417

[15] Merton, R.C. " Option pricing when underlying stock returns are discontinuous "*J.Financ.Econom.*, Vol.3, (1976), pp125-144

[16] Merton, R.C. *Continuous Time Finance*, Basil Blackwell, Oxford, (1990)

[17] Nelson, D.B. " The Time Series Behavior of Stock Market Volatility and Returns "unpublished doctoral dissertation, Department of Economics, M.I.T.

[18] Schwert, G.W. " Business cycles, financial cycles and stock volatility "*Carnegie-Rochester Conference Series on Public Policy*, Vol.39, (1989), pp83-126

[19] Shephard, G. and Pitt, M.K. " Likelihood Analysis of Non-Gaussian Measurement Time Series "*Biometrika*, Vol.84, (1997), pp653-667

[20] Tauchen, G. and Pitts, M." The Price Variability-Volume Relationship on Speculative Markets "*Econometrica*, Vol.51, (1983), pp485-505

[21] Taylor, S.J. *Modeling Financial Time Series*, John Wiley & Sons, (1986)