

研究ノート

SCDモデルの理論と実証 －東京商品取引所における金先物取引への応用－

砂 田 洋 志

(人文社会科学部教授)

1 はじめに

証券市場における1日の取引の中には一定の傾向のあることが分かっている。たとえば、寄り付きと大引けに取引が集中して、取引間隔が短くなること、出来高が多くなること、ボラティリティーが高くなること、一方、昼間は取引が少なくなり、取引間隔が広がること、出来高が少ないこと、ボラティリティーが小さいことなどが挙げられる。こうした現象は日中効果と呼ばれている。本稿ではこの中から取引時間の間隔が時間帯によって変化することを取り上げ、その推移を時系列分析の立場からモデル化した研究を紹介する。取引時間の間隔を時系列分析の立場からモデル化して数量的に分析した研究としては、Engle=Russel(1997, 1998)のACD (Autoregressive Conditional Duration)モデル、Zhang他(2001)の非線形ACDモデル、Bauwens = Giot(2000)のLog-ACDモデル、Russel=Engle(2005)のACM-ACDモデル、Bauwens=Veredas(2004)のSCD(Stochastic Conditional Duration)モデルを用いた研究がある。本稿で取り上げるSCDモデルの研究としては擬似最尤法で推定するBauwens=Veredas(2004)の他に、レバレッジ効果を導入したSCDモデルをモンテカルロ最尤法で推定したFeng他(2004)、ベイズ統計学の立場からSCDモデルを推定する方法を示したStrickland他(2006)、レバレッジ効果を導入したSCDモデルを粒子フィルターとベイズ統計学を用いて推定したMen他(2015)などがある。日本の市場を対象とした研究としては森保(2006)がある他は、幾つかの研究が行われているに過ぎない。森保(2006)では、Log-ACDモデルとACM-ACDモデルを用いて東京金先物市場が計量的に分析されている。杉浦=中妻(2016)では、株式市場データへSCDモデルが適用され、粒子フィルターを用いてパラメータが推定された後、約定時間の予測へ応用された。

本稿の目的は、取引の時間間隔をモデル化して分析する際に利用されるモデルの中で発展的なモデルと呼べるSCDモデルを紹介することである。以下の構成でSCDモデルを紹介する。まず、第2節では点過程まで戻ってSCDモデルを説明する。第3節ではデータの説明とデータを平滑化する方法について説明する。Strickland他(2006)に基づいて、第4節では状態変数、第5節では未知パラメータの推定方法を紹介する。第6節では、東京商品取引所に上場されてい

る金先物データを対象にSCDモデルを適用してバイズ推定した結果を示す。その際には同じデータにACDモデルを用いて推定した結果と比較する。第7節では本稿の内容をまとめる。

2 SCDモデル

以下では、点過程、ポアソン過程とACDモデルを説明した後に、SCDモデルを説明する。こうして、基礎からSCDモデルを説明する。

2.1 点過程

尾崎（1998）によれば、時間軸上に発生時点を示す点が置かれている確率的現象を考え、これを点過程と呼ぶと記されている。つまり、事象がランダムな時点で発生する状況を想定し、時間を示す半直線 $[0, \infty)$ 上にその発生時点を記述した点の系列 $\{t_i; 0 < t_1 < t_2 < \dots\}$ が点過程である。ただし、 t_i は事象が i 番目に発生する時点を表す。後述するが、点の系列の確率的な構造を記述するために点の発生間隔、つまり事象の発生間隔 $x_i = t_i - t_{i-1}$ を用いる。発生間隔 x_i はデュレーションと呼ばれる。

事象の発生時点 t_i だけの一変量の系列 $\{t_i\}$ ではなく、発生時点に加えてそれ以外のデータを伴う多変量の系列の場合、マーク付き点過程と呼ばれる。たとえば、株式取引の発生時点 t_i に加えて、その取引の価格 p_i や取引量 vol_i も含めた3変量の系列 $\{t_i, p_i, vol_i\}$ はマーク付き点過程となる。

時系列データは一般に等間隔で観測されたデータであり、各時点において観測される数量の系列 $\{\dots, y_{t-1}, y_t, y_{t+1}, \dots\}$ であって発生時点のデータ $\{t_i\}$ ではない。時系列分析ではそうしたデータをモデル化して分析する。

点過程を基にして時点 t までに事象が発生した回数を得ることができる。この回数を $N(t)$ と記述すると、 $\{N(t)\}$ は計数過程（counting process）と呼ばれる過程に従う。計数過程 $\{N(t)\}$ によって、事象の発生時点が記録された点過程 $\{t_i\}$ を定義することができる¹。

尾形（1998）によれば、点過程では条件付き強度関数 $\lambda(t|H_t)$ が以下のように定義され、 $\lambda(t|H_t)$ によって点過程は完全に特定される²。

$$P\{N(t+\delta) - N(t) = 1 | H_t\} = \lambda(t|H_t) \times \delta + o(\delta)$$

¹ 林（2012）の473ページを参照されたい。

² 尾形（1998）では、条件付き強度関数は点過程を完全に記述すると述べられている。さらに、尾形（2008）では、条件付き強度関数と点過程は一对一の関係に対応すると述べられている。これこそ条件付き強度関数が重要な理由である。分析対象となる点過程の条件付き強度関数を導出することは有益である。

2.2 ポアソン過程

基本的な点過程としてポアソン過程が挙げられるので、以下で紹介する。ポアソン過程では発生間隔（デュレーション）を表す確率変数の系列 $x_i (i=1, 2, 3, \dots, n)$ が共通にパラメータ $\frac{1}{\lambda}$ の指数分布に従い、相互に独立である。

$$x_i \sim i.i.d. \text{Exp} \left(\frac{1}{\lambda} \right) \quad (1)$$

x_i の確率密度関数は $f(x) = \lambda \exp(-\lambda x)$ であるから、相互に独立で同一分布に従う x_1, x_2, \dots, x_n の和である $S_n = x_1 + x_2 + \dots + x_n$ はガンマ分布 $GAM \left(n, \frac{1}{\lambda} \right)$ に従う³。

$$S_n = x_1 + x_2 + \dots + x_n \sim GAM \left(n, \frac{1}{\lambda} \right) \quad (2)$$

このことを利用して、 $\Pr(N(t)=k)$ を計算すると、

$$\begin{aligned} \Pr(N(t)=k) &= \Pr(N(t) \geq k) - \Pr(N(t) \geq k+1) = \Pr(S_k \leq t) - \Pr(S_{k+1} \leq t) \\ &= \left\{ 1 - e^{-\lambda t} \sum_{r=0}^{k-1} \frac{(\lambda t)^r}{r!} \right\} - \left\{ 1 - e^{-\lambda t} \sum_{r=0}^k \frac{(\lambda t)^r}{r!} \right\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \end{aligned} \quad (3)$$

となる。したがって、確率変数 $N(t)$ の確率分布はパラメータが λt のポアソン分布である。このとき、計数過程 $\{N(t)\}$ はパラメータ λt のポアソン過程に従うと呼ばれる。なお、計数過程 $\{N(t)\}$ がポアソン過程に従っている場合の必要十分条件は以下の2点である。

(a) 定常かつ独立な増分を有している

$$(b) \Pr(N(t)=k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

上述した計数過程 $\{N(t)\}$ は、この2つの条件を満すポアソン過程である。パラメータ λt のポアソン過程の場合、その条件付き強度関数 $\lambda(t|H_t)$ は以下のとおり一定値の λ である⁴。

³ 相互に独立な確率変数 X_i が $GAM(\alpha_i, \beta)$ に従う場合、 $S_n = \sum_{i=1}^n X_i$ は再生性によって $GAM\left(\sum_{i=1}^n \alpha_i, \beta\right)$ に従う。したがって、この確率変数 $x_i (i=1, 2, 3, \dots)$ が共通のパラメータ $1/\lambda$ の指数分布、つまり、 $GAM(1, 1/\lambda)$ に従うならば、 $\sum_{i=1}^n X_i$ は再生性によって $GAM(n, 1/\lambda)$ に従う。さらに、 $GAM(n, 1/\lambda)$ の確率密度関数と分布関数は以下のとおりであることを用いて(3)を導出する。

$$f(x) = \frac{x^{n-1} \exp(-x\lambda)}{\lambda^{-n} \Gamma(n)}, \quad F(x) = 1 - \left\{ \exp(-x\lambda) \times \sum_{r=0}^{n-1} \frac{(x\lambda)^r}{r!} \right\}$$

⁴ 伏見 (1987) の定理 5-3 とマクローリン展開を利用して強度関数を導出してある。

$$\begin{aligned} \Pr(N(t+\delta)-N(t)=1) &= \exp(-\lambda\delta) \times \lambda\delta = \left(1 - \frac{\lambda\delta}{1!} + \frac{(\lambda\delta)^2}{2!} - \frac{(\lambda\delta)^3}{3!} + \dots\right) \times \lambda\delta \\ &= \lambda\delta - (\lambda\delta)^2 + \frac{(\lambda\delta)^3}{2!} - \frac{(\lambda\delta)^4}{3!} + \dots \doteq \lambda\delta + o(\delta) \\ \lambda(t|H_t) &= \frac{d}{d\delta} (\Pr(N(t+\delta)-N(t)=1)) = \frac{d}{d\delta} (\lambda\delta + o(\delta)) \doteq \lambda \end{aligned}$$

2.3 ACDモデル

前述したとおり、2つ事象が発生する時点 t_i と t_{i-1} の間隔は $x_i = t_i - t_{i-1}$ と記述され、デュレーションと呼ばれる。ACDモデルでは、観測される変数 x_i が x_i の条件付き期待値 Ψ_i (状態変数) と誤差項 ε_i を用いて以下のように定式化される。

$$x_i = \Psi_i \varepsilon_i \quad i = 1, 2, \dots, n \quad (4)$$

$$\varepsilon_i \sim i.i.d. \quad (5)$$

Ψ_i 、あるいは Ψ_i を対数変換した $\psi_i (= \ln \Psi_i)$ の定式化は様々なものを提案することが可能である。期待デュレーション Ψ_i を定式化する基本モデルとして Ψ_i と x_i の一次のラグ付き変数である Ψ_{i-1} と x_{i-1} を説明変数とする一次式を仮定する。

$$\Psi_i = \omega + \alpha \Psi_{i-1} + \beta x_{i-1} \quad i = 1, 2, \dots, n \quad (6a)$$

$$\psi_1 \sim N\left(0, \frac{\omega}{1 - \alpha - \beta}\right) \quad (6b)$$

ここで、誤差項 ε_i の確率分布としてパラメータが γ のガンマ分布 $GAM\left(\gamma, \frac{1}{\gamma}\right)$ を仮定する。

ε_i の確率密度関数は以下のとおりである。

$$\varepsilon_i \sim i.i.d. GAM\left(\gamma, \frac{1}{\gamma}\right) \quad f(\varepsilon_i|\gamma) = \frac{\varepsilon_i^{\gamma-1} \exp(-\varepsilon_i \gamma)}{\gamma^{-\gamma} \Gamma(\gamma)} \quad (7)$$

(4)と(7)から、 x_i の確率密度関数として次式を得る。

$$g(x_i|\Psi_i, \gamma) = \frac{1}{x_i} \left(\frac{x_i}{\Psi_i}\right)^\gamma \exp\left(-\frac{x_i \gamma}{\Psi_i}\right) \frac{\gamma^\gamma}{\Gamma(\gamma)}$$

$GAM\left(\gamma, \frac{1}{\gamma}\right)$ において $\gamma=1$ とおけば、パラメータが1の指数分布となる。そこで、 ε_i の確率分布

としてパラメータが1の指数分布を仮定する。

$$\varepsilon_i \sim i.i.d. Exp(1) \quad f(\varepsilon_i) = \exp(-\varepsilon_i) \quad (8)$$

(4)と(8)から、 x_i の確率密度関数として次式を得る⁵⁾。

$$g(x_i|\Psi_i, \gamma) = \frac{1}{\Psi_i} \exp\left(-\frac{x_i}{\Psi_i}\right) \frac{1}{\Gamma(1)} = \frac{1}{\Psi_i} \exp\left(-\frac{x_i}{\Psi_i}\right)$$

したがって、対数尤度は以下のとおりである。

$$\ln L(\Psi_1, \dots, \Psi_N, \omega, \alpha, \beta, |x_1, x_2, \dots, x_n) = -\sum_{i=1}^n \ln \Psi_i - \sum_{i=1}^n \frac{x_i}{\Psi_i}$$

ただし、 Ψ_i は (6a) と (6b) で定めたとおりである。

なお、林・佐藤 (2016) によれば、ACDモデルの条件付き強度関数は以下のとおりである。

$$\lambda(t|N(t), t_1, t_2, \dots, t_{N(t)}) = \lambda_0 \left(\frac{t - t_{N(t)}}{\Psi_{N(t)-1}} \right) \frac{1}{\Psi_{N(t)-1}}$$

ただし、 λ_0 はベースライン・ハザード関数である。

2.4 SCDモデル

SCDモデルでは、観測される変数 x_i が x_i の条件付き期待値 Ψ_i (状態変数) と誤差項 ε_i を用いて以下のように定式化される。

$$x_i = \Psi_i \varepsilon_i \quad i = 1, 2, \dots, n \quad (9)$$

$$\varepsilon_i \sim i.i.d. \quad (10)$$

$$E[x_i | \Psi_{i-1}, \Psi_{i-2}, \dots, \Psi_{i-k}, \eta_i] = \Psi(\Psi_{i-1}, \Psi_{i-2}, \dots, \Psi_{i-k}, \eta_i) \equiv \Psi_i$$

このモデルではデュレーション x_i の条件付き期待値 Ψ_i が自らの過去の値 $\Psi_{i-1}, \Psi_{i-2}, \Psi_{i-3}, \dots$ と誤差項 η_i に依存している。

期待デュレーション Ψ_i の定式化は自由である。また、 ε_i の確率分布も 0 以上の値だけを実現する確率分布という条件は必要であるものの、細かい条件はない。つまり、 Ψ_i の定式化と ε_i の確率分布の特定化によって様々なSCDモデルを定義できる。

前述したポアソン過程では x_i が独立かつ同一の指数分布に従っている。その結果、計数過程 $\{N(t)\}$ がポアソン過程に従っている。一方、SCDモデルでは、 ε_i が独立かつ同一の指数分布に従っていたとしても、 Ψ_i が変動するので、 x_i の確率分布は変動する。

⁵ $\varepsilon_i \sim \text{Exp}(\beta)$ であれば、 $f(\varepsilon_i) = \frac{1}{\beta} \exp\left(-\frac{\varepsilon_i}{\beta}\right)$ であるから、畳み込みによって x_i の確率密度関数は以下のとおり
に導出される。

$$g(x_i) = \frac{1}{\beta} \exp\left(-\frac{x_i/\Psi_i}{\beta}\right) \times \left| \frac{d\varepsilon_i}{dx_i} \right| = \frac{1}{\beta} \exp\left(-\frac{x_i}{\Psi_i \beta}\right) \times \left| \frac{1}{\Psi_i} \right|$$

ここで、 $\beta = 1$ を代入すると次式が得られる。

$$g(x_i) = \frac{1}{\Psi_i} \exp\left(-\frac{x_i}{\Psi_i}\right)$$

前述したとおり、 Ψ_i 、あるいは Ψ_i を対数変換した $\psi_i (= \ln \Psi_i)$ の定式化は様々なものを提案することが可能である。本稿では単純なモデルとして、期待デューレーション Ψ_i を対数変換した ψ_i の一階の自己回帰モデルを仮定する。なお、 W_i は k 次の列ベクトルであり、その要素は定数項に加えて Ψ_{i-1} と x_{i-1} 以外で Ψ_i に影響を与える変数である。

$$\psi_i = W_i \delta + \phi \psi_{i-1} + \eta_i \quad \eta_i \sim N(0, \sigma_\eta^2) \quad i = 1, 2, \dots, n \quad (11a)$$

$$\psi_1 \sim N\left(0, \frac{\sigma_\eta^2}{1 - \phi^2}\right) \quad (11b)$$

このようにSCDモデルでは誤差項を含むように ψ_i が定式化されるので、デューレーションに加えてボラティリティーにも確率的な変動が含まれるようにACDモデルが拡張されている⁶。

Bauwens=Veredas (2004)によれば、誤差項がワイブル分布に従うSCDモデルのハザード関数(=強度関数)は以下のとおりである。⁷

$$h(d_i | \psi_{i-1}) = \frac{\gamma \int_{-\infty}^{\infty} \Psi_i^{-\gamma} d_i^{\gamma-1} \exp\left[-\left(\frac{d_i}{\Psi_i}\right)^\gamma - \left(\frac{u_i^2}{2\sigma^2}\right)\right] du_i}{\int_{-\infty}^{\infty} \exp\left[-\left(\frac{d_i}{\Psi_i}\right)^\gamma - \left(\frac{u_i^2}{2\sigma^2}\right)\right] du_i} \quad (12)$$

3 データとその加工方法について

3.1 分析するデータについて

本稿では東京商品取引所で取引されている金先物（現物受渡用）のティックデータを用いてSCDモデルを推定した。試験的に2016年4月11日を選んでいる。金先物も含めて、日本の商品先物市場では一般に期先物の出来高が最も多い。そこで、今回の分析でも4月11日における期先物である2017年2月限のデータを利用した。金先物の取引時間は午前9時～15時15分が日中の取引時間である。9時に板合せ取引が行なわれて9時1分に始値が付く。その後、ザラバ取引が15時15分まで行なわれる⁸。これ以外の時間帯に時間外取引も行われているが、本稿の分析では9時1分から15時15分までの日中のデータだけを利用してパラメータの推定を行った。分析対象とする取引は4,017個であった。

本稿で分析するティックデータとは、1日の中で取引される一つ一つの取引に関するデータを集めたものである。したがって、ティックデータには、取引時刻（時分秒）、取引価格（円）、

⁶ ACDモデルでは Ψ_i を定式化する際に誤差項を含めない。たとえば、 $\Psi_i = \omega + \alpha x_i + \beta \Psi_{i-1}$ と定式化する。

⁷ Bauwens=Veredas (2004)の389ページの(15)を参照されたい。ガンマ分布に従う場合のハザード関数は解析的に導出されないと記述されている。

⁸ 現在は取引システムが変更されて、8時45分から15時15分までが日中の取引時間であり、寄り付きと大引けで板合せ取引を行なう。

取引数量（枚）が含まれている。東京商品取引所から購入したティックデータには、日付（年月日）、取引種別（先物かオプションか）、商品番号、限月（年月）、約定時刻（時分秒）、約定価格（円）、出来高（枚）が入力されている。

ティックデータを用いて分析するので、1秒間に複数の取引が執行される場合もある。この場合は表1に示すとおり、同じ時点に複数の取引情報が表示される。取引間隔を計算する場合には同じ時点（秒）で執行された取引は1つにまとめ、1つのデータと考える。こうして、取引が執行された時点（秒）ごとに出来高（取引枚数）と約定価格が2,088個得られた。

表1 取引データの集計例

取引データ集計前			取引データ集計後		
取引時刻	取引価格	出来高	取引時刻	取引価格	出来高
9:05:10	1000	1	9:05:10	1000	4
9:05:10	1000	3	9:05:11	1002	5
9:05:11	1002	2	9:05:13	1003	2
9:05:11	1002	3	⋮	⋮	⋮
9:05:13	1003	2	⋮	⋮	⋮

3.2 データの加工方法について

全ての取引が含まれているティックデータから、時間間隔データを作成する。そのデータにSCDモデルを直接適用して、パラメータを推定するのではなく、時間間隔データを元に計算した平滑化曲線から得られた値でデータを割った後、そのデータを用いてSCDモデルのパラメータを推定する。

平滑化の方法は様々である。全データに1つの直線や曲線を当て嵌める方法、多項式を当て嵌める方法、移動平均を当て嵌める方法がある。これに加えて、データの存在する領域を節点で幾つかに区切って区分毎に曲線や直線を当て嵌める方法もある。⁹たとえば、区分的一次回帰により線を当て嵌める方法である。さらに一般化して、節点において微分した曲線が連続となるスプライン関数を用いて平滑化する方法もある。その中には、スプライン関数を3次のB-スプライン基底の一次結合に限定して平滑化する方法もある。

横軸の値に対し、縦軸の値を1つ1つ定める、つまり点に点に対応させて平滑化する方法としては、カーネル関数を重みとして区間内のデータを加重平均する、ナガラヤ・ワトソン推定

⁹ 上述した様々な手法に対する評価はいろいろであると思われる。竹澤（2010）はハット行列を用いて、最小二乗法で直線を当て嵌める方法、移動平均を用いて当て嵌める方法、3次スプライン関数、平滑化スプライン関数を当て嵌める方法について評価している。その結果、平滑化スプライン関数を当て嵌める方法が最も良いと評価している。スプライン関数の次数としてさまざまな次数が考えられるが、多くの場合3次のスプライン関数が利用される。

量がある。また、カーネル関数を重みとして区間内のデータに最適な多項式を当て嵌める、局所多項式回帰もある。これらの方法では、カーネル関数を利用する際のバンド幅が全データで共通の値となっている。

前述したとおり、局所多項式回帰ではバンド幅が一定である。しかし、現実のデータはデータの密度の濃さに違いがある。そこで、局所多項式回帰においてバンド幅を可変としたLOESS (LOcally weighted ReGrESSion) が考案された。LOESSをさらに発展させたものにスーパースムーザがある。さらに、外れ値の影響を弱めて平滑化する方法が考案され、LOWESS (LOcally WEighted Scatter plot Smoother) と呼ばれている。なお、本稿ではスーパースムーザを用いてデータを平滑化する。

4 状態変数の推定方法

SCDモデルの状態変数を推定する手順をまとめると以下のとおりである。本節の中で、この一つ一つの手順を詳しく説明する。

- 1) 観測される変数 $\{x_i\}_{i=1}^n$ と状態変数 $\{\psi_i\}_{i=1}^n$ の間にある関数関係 ((9), (10), (11a,11b)) を線形正規状態空間モデルで近似するため、 $\{x_i\}_{i=1}^n$ を変形した変数 $\{\tilde{x}_i\}_{i=1}^n$ が観測変数、 $\{\psi_i\}_{i=1}^n$ を状態変数とする線形正規状態空間モデル ((22), (11a,11b)) を導出する。
- 2) 線形正規状態変数モデル ((22), (11a,11b)) に対して、ブロック・サンプリング、カルマン・フィルターとシミュレーション・スムーザを適用して、 $\{\eta_s\}_{s=i}^{i+k}$ の候補を生成する。
- 3) SCDモデル ((9), (10), (11a,11b)) における $\{x_i\}_{i=1}^n$ と $\{\psi_i\}_{i=1}^n$ の対数結合密度関数 (14) と、近似した線形正規状態空間モデルにおける同様の対数結合密度関数 (21) を用いて採択確率を計算する。この採択確率を用いて、近似モデルから生成した $\{\eta_s\}_{s=i}^{i+k}$ のサンプルの採否を決める。
- 4) (11a) を通じて間接的に $\{\psi_s\}_{s=i}^{i+k}$ を生成する。

4.1 SCDモデルにおける x と ψ の結合密度の計算

$\psi = \{\psi_i\}_{i=1}^n$ と $x = \{x_i\}_{i=1}^n$ の結合分布 $p(\psi, x | \phi, W, \delta, \gamma, \sigma_\eta)$ は $p(\psi | \phi, W, \delta, \gamma, \sigma_\eta)$ と $g(x | \psi, \gamma)$ の積である。

$$p(\psi, x | \phi, W, \delta, \gamma, \sigma_\eta) = g(x | \psi, \gamma) \times p(\psi | \phi, W, \delta, \gamma, \sigma_\eta) \quad (13)$$

上式へ (11a,11b) を代入して対数変換すると、次式が得られる。

$$\begin{aligned} & \ln p(\psi, x | \phi, W, \delta, \gamma, \sigma_\eta) \\ &= -\frac{1}{2\sigma_\eta^2} \sum_{i=2}^n (\psi_i - W_i \delta - \phi \psi_{i-1})^2 - \frac{1-\phi^2}{2\sigma_\eta^2} \left(\psi_1 - \frac{W_1 \delta}{1-\phi} \right)^2 + \sum_{i=1}^n \ln g(x_i | \psi_i, \gamma) \end{aligned} \quad (14)$$

誤差項 ε_i の確率分布としてパラメータが γ のガンマ分布 $GAM\left(\gamma, \frac{1}{\gamma}\right)$ を仮定すると、次式が得られる。

$$\begin{aligned} \ln p(\psi, x | \phi, W, \delta, \gamma, \sigma_\eta) = & n\gamma \ln \gamma - \gamma \sum_{i=1}^n \psi_i + (\gamma - 1) \sum_{i=1}^n \ln x_i - \sum_{i=1}^n x_i \gamma \times \exp(-\psi_i) \\ & - n \ln \Gamma(\gamma) - \frac{1}{2\sigma_\eta^2} \sum_{i=2}^n (\psi_i - W_i \delta - \phi \psi_{i-1})^2 - \frac{1-\phi^2}{2\sigma_\eta^2} \left(\psi_1 - \frac{W_1 \delta}{1-\phi} \right)^2 \end{aligned} \quad (15)$$

SCDモデルは、観測方程式の誤差項が正規分布に従う線形正規状態空間モデルでないから、(15)の右辺第1項から第4項が線形正規状態空間モデルの場合とは異なる形になる。そこで、次の小節において、SCDモデルを近似する線形正規状態空間モデルの対数結合密度関数を導出する。

4.2 線形正規状態空間モデルによる近似

(9), (10), (11a, 11b) で構成されるSCDモデルは非線形モデルである。そこで、最初に (9) を対数変換することによって線形化する。

$$\ln x_i = \ln \Psi_i + \ln \varepsilon_i = \psi_i + \ln \varepsilon_i \quad (9')$$

$\ln \varepsilon_i$ は平均が0ではない上、正規分布に従わないので、変数を対数変換するだけでは推定できない。SCDモデルは非線形モデルであるとともに、誤差項が正規分布に従っていないので、パラメータや状態変数の系列 $\{\psi_i\}$ の推定は簡単でない。本稿では、SVモデルの推定と同様に、(9'), (10), (11a, 11b) で構成されるSCDモデルを線形正規状態空間モデルで近似した上で、カルマン・フィルター、シミュレーション・スモザ (シミュレーション平滑法) と M-H 法 (メトロポリス・ヘイスティング法) を利用して状態変数とパラメータを推定する。¹⁰

ψ と x の結合分布 $p(\psi, x | \phi, W, \delta, \gamma, \sigma_\eta)$ は $p(\psi | \phi, W, \delta, \gamma, \sigma_\eta)$ と $g(x | \psi, \gamma)$ の積であり、(13) で記述される。(13) へ (11a, 11b) を代入して対数変換すると、(14) が得られる。

(14) を線形正規状態空間モデルで近似するために、観測方程式の対数尤度に相当する (14) の $\sum_{i=1}^n \ln g(x_i | \psi_i, \gamma)$ を ψ に関する式と捉え直し、上式のモードである ψ_i^* の近傍でテイラー展開して、線形近似する。なお、 $h(x_i | \psi_i, \lambda) = -\ln g(x_i | \psi_i, \lambda)$ である。

$$\begin{aligned} -\ln g(x_i | \psi_i, \gamma) = & h(x_i | \psi_i, \gamma) \doteq h(x_i^* | \psi_i, \gamma) + \frac{\partial h(x_i | \psi_i, \gamma)}{\partial \psi_i} \Big|_{\psi_i = \psi_i^*} \times (\psi_i - \psi_i^*) \\ & + \frac{\partial^2 h(x_i | \psi_i, \gamma)}{\partial \psi_i^2} \Big|_{\psi_i = \psi_i^*} \times \frac{(\psi_i - \psi_i^*)^2}{2} \end{aligned} \quad (16)$$

$$\dot{h}_i \equiv \frac{\partial h(x_i | \psi_i, \gamma)}{\partial \psi_i} \Big|_{\psi_i = \psi_i^*} = - \frac{\partial \ln g(x_i | \psi_i, \gamma)}{\partial \psi_i} \Big|_{\psi_i = \psi_i^*} \quad (17)$$

¹⁰ シミュレーション・スモザは de Jong=Shephard (1995) で提案された。

$$\ddot{h}_i \equiv \left. \frac{\partial^2 h(x_i|\psi_i, \gamma)}{\partial \psi_i^2} \right|_{\psi_i=\psi_i^*} = - \left. \frac{\partial^2 \ln g(x_i|\psi_i, \gamma)}{\partial \psi_i^2} \right|_{\psi_i=\psi_i^*} \quad (18)$$

以上のように \dot{h}_i と \ddot{h}_i を定義すれば、上式を以下のように記述できる。

$$\begin{aligned} h(x_i|\psi_i, \gamma) &\doteq h(x_i|\psi_i^*, \gamma) + \dot{h}_i (\psi_i - \psi_i^*) + \ddot{h}_i \times \frac{(\psi_i - \psi_i^*)^2}{2} \\ &\doteq h(x_i|\psi_i^*, \gamma) + \frac{\dot{h}_i}{2} \left\{ (\psi_i - \psi_i^*)^2 + \frac{2\dot{h}_i}{\ddot{h}_i} (\psi_i - \psi_i^*) + \frac{\dot{h}_i^2}{\ddot{h}_i^2} \right\} - \frac{\dot{h}_i^2}{2\ddot{h}_i} \\ &\doteq \frac{\dot{h}_i}{2} \left\{ \psi_i^* - \frac{\dot{h}_i}{\ddot{h}_i} - \psi_i \right\}^2 + h(x_i|\psi_i^*, \gamma) - \frac{\dot{h}_i^2}{2\ddot{h}_i} \end{aligned}$$

ここで、 \tilde{x}_i と \tilde{H}_i を以下のように定義する。

$$\tilde{H}_i = \left[\left. \frac{\partial^2 h(x_i|\psi_i, \lambda)}{\partial \psi_i^2} \right|_{\psi_i=\psi_i^*} \right]^{-1}, \quad \dot{h}_i = \frac{1}{\tilde{H}_i} \quad (19)$$

$$\tilde{x}_i = \psi_i^* - \tilde{H}_i \left(\left. \frac{\partial h(x_i|\psi_i, \lambda)}{\partial \psi_i} \right|_{\psi_i=\psi_i^*} \right) = \psi_i^* - \frac{\dot{h}_i}{\ddot{h}_i} \quad (20)$$

その結果、 $h(x_i|\psi_i, \gamma)$ 、さらに $\sum_{i=1}^n \ln g(x_i|\psi_i, \gamma)$ は以下のとおりになる。

$$\begin{aligned} h(x_i|\psi_i, \gamma) &\doteq \frac{1}{2} \frac{(\tilde{x}_i - \psi_i)^2}{\tilde{H}_i} + h(x_i|\psi_i^*, \gamma) - \frac{\dot{h}_i^2}{2\ddot{h}_i} \\ \sum_{i=1}^n \ln g(x_i|\psi_i, \gamma) &\doteq -\frac{1}{2} \sum_{i=1}^n \frac{(\tilde{x}_i - \psi_i)^2}{\tilde{H}_i} - \sum_{i=1}^n h(x_i|\psi_i^*, \gamma) + \sum_{i=1}^n \frac{\dot{h}_i^2}{2\ddot{h}_i} \end{aligned}$$

以上から ψ と x の結合密度関数を対数変換した (14) は以下のように近似できる。

$$\begin{aligned} \ln p(\psi, x|\phi, W, \delta, \gamma, \sigma_\eta) \\ \doteq -\frac{1}{2\sigma_\eta^2} \sum_{i=2}^n (\psi_i - W_i \delta - \phi \psi_{i-1})^2 - \frac{1-\phi^2}{2\sigma_\eta^2} \left(\psi_1 - \frac{W_1 \delta}{1-\phi} \right)^2 - \frac{1}{2} \sum_{i=1}^n \frac{(\tilde{x}_i - \psi_i)^2}{\tilde{H}_i} \quad (21) \end{aligned}$$

(21) の右辺は、 $\tilde{x}_i = \psi_i^* - \frac{\dot{h}_i}{\ddot{h}_i}$ を観測変数、 ψ_i を状態変数、以下に示す (22) を観測方程式、(11a, 11b) を遷移方程式とする線形正規状態空間モデルにおいて $\{\tilde{x}_i\}_{i=1}^n$ と $\{\psi_i\}_{i=1}^n$ の結合密度関数を対数変換した式である。つまり、SCDモデルは、(22) を観測方程式、(11a, 11b) を遷移方程式とする線形正規状態空間モデルで近似することができる。

$$\tilde{x}_i = \psi_i + \tilde{\varepsilon}_i \quad \tilde{\varepsilon}_i \sim N(0, \tilde{H}_i) \quad (22)$$

ここで、誤差項 ε_i の確率分布としてパラメータが γ のガンマ分布 $GAM\left(\gamma, \frac{1}{\gamma}\right)$ を仮定すると、

(19) の \tilde{H}_i と (20) の \tilde{x}_i は以下のとおり記述される。

$$\dot{h}_i \equiv -\frac{\partial g(x_i|\psi_i, \gamma)}{\partial \psi_i} \Big|_{\psi_i=\psi_i^*} = \gamma - \exp\{\ln(x_i\gamma) - \psi_i^*\} \quad (17')$$

$$\ddot{h}_i \equiv -\frac{\partial^2 g(x_i|\psi_i, \gamma)}{\partial \psi^2} \Big|_{\psi_i=\psi_i^*} = \exp\{\ln(x_i\gamma) - \psi_i^*\} \quad (18')$$

$$\tilde{H}_i = \frac{1}{\dot{h}_i} = (x_i\gamma)^{-1} \exp(\psi_i^*) \quad (19')$$

$$\tilde{x}_i = \psi_i^* - \frac{h_i}{\ddot{h}_i} = \psi_i^* - x_i^{-1} \exp(\psi_i^*) + 1 \quad (20')$$

また、誤差項 ε_i にパラメータが 1 の指数分布を仮定すると、(19) の \tilde{H}_i と (20) の \tilde{x}_i は以下のとおり記述される。

$$\tilde{H}_i = x_i^{-1} \exp(\psi_i^*) \quad (19'')$$

$$\tilde{x}_i = \psi_i^* - x_i^{-1} \exp(\psi_i^*) + 1 \quad (20'')$$

なお、 ε_i に別の確率分布を仮定したならば、上式とは異なる式で記述される。¹¹

4.3 カルマン・フィルター

本稿で扱う以下のSCDモデルに対してカルマン・フィルターを適用して、状態変数 ψ_i や観測変数 \tilde{x}_i の平均、分散と確率分布を導出する。¹²

$$\tilde{x}_i = \psi_i + \tilde{\varepsilon}_i \quad \tilde{\varepsilon}_i \sim N(0, \tilde{H}_i) \quad i = 1, 2, \dots, n \quad (22)$$

$$\psi_i = W_i \delta + \phi \psi_{i-1} + \eta_i \quad \eta_i \sim N(0, \sigma_\eta^2) \quad (11a)$$

$$\psi_1 \sim N\left(0, \frac{\sigma_\eta^2}{1 - \phi^2}\right) \quad (11b)$$

観測方程式を (22)、遷移方程式を (11a, 11b) に定めた線形状態空間モデルについて、状態変数 ψ_i と観測変数 \tilde{x}_i の一期先予測分布は以下のとおりである。

$$\psi_i | I_{i-1} \sim N(\hat{\psi}_i, P_i) \quad (23)$$

$$\tilde{x}_i | I_{i-1} \sim N(\hat{\psi}_i, D_i) \quad (24)$$

$$\text{ただし、} \hat{\psi}_{i+1} = \begin{cases} \phi \psi_0 & i=0 \\ W_i \delta + \phi \hat{\psi}_i + K_i e_i & i \geq 1 \end{cases} \quad P_{i+1} = \begin{cases} \sigma_\eta^2 I & i=0 \\ \phi P_i L_i' + \sigma_\eta J_i' & i \geq 1 \end{cases}$$

¹¹ 他の確率分布を誤差項に仮定した場合の変形方法が Strickland 他 (2006) の表 6 に記載されている。

¹² 一般的な線形正規状態空間モデルにカルマン・フィルターを適用する場合の説明を付録に記述しておく。

$$e_i = \tilde{x}_i - \hat{\psi}_i, \quad K_i = \phi P_i D_i^{-1}$$

$$D_i = P_i + \tilde{H}_i \quad L_i = \phi - K_i \quad J_i = \sigma_\eta - K_i \tilde{H}_i^{\frac{1}{2}}$$

さらに、状態変数 ψ_i の事後分布は以下のとおりである。

$$\psi_i | I_i \sim N(\hat{\mu}_i, \hat{\Sigma}_i) \tag{25}$$

ただし、 $\hat{\mu}_i = \hat{\psi}_i + P_i' D_i^{-1} e_i$, $\hat{\Sigma}_i = P_i - P_i D_i^{-1} P_i$ である。

(23)～(25)を繰り返し利用することで、逐次的に $\psi_i | I_{i-1}$, $\tilde{x}_i | I_{i-1}$, $\psi_i | I_i$ を計算することができる。さらに、全データを利用した場合の状態変数の確率分布を導出することもできる（平滑化）。こうした方法に従い、状態変数 ψ_i を生成して平均した値を状態変数の推定値と考える。

4.4 シミュレーション・スムーザ

状態空間モデルの次元が大きい場合に状態変数を生成する際の計算効率は、カルマン・フィルターの結果から直接生成するよりも、シミュレーション・スムーザを用いて生成する方が高いことが知られている¹³。そこで、本稿においてもシミュレーション・スムーザを用いて状態変数を生成する。本稿で扱うSCDモデルにシミュレーション・スムーザを適用する際の手順は以下のとおりである。¹⁴

以下の(i)～(v)の順番で作業を進める。最初に C_i と V_i を計算する。計算した C_i を用いて z_i を生成し、さらに、 C_i と V_i と z_i を用いて r_i を計算する。最終的には r_i と z_i を用いて y_i を生成する。この計算にはカルマン・フィルターで得た結果も利用する。この作業を $i = n-1$ から $i=1$ に向けて継続して、 $i=1, 2, \dots, n-1$ の各時点において、 $C_i, V_i, z_i, r_i, U_i, y_i$ を計算する。

$i=n$ では C_n と V_n を計算した結果を用いて z_n, y_n を計算するが、 r_n と U_n については $r_n=0, U_n=0$ とおき、計算しない。

(i) カルマン・フィルターの結果を利用して、 C_i, V_i を計算する。

$$C_i = \tilde{H}_i - \tilde{H}_i' U_i \tilde{H}_i, \quad V_i = \tilde{H}_i U_i L_i$$

(ii) 直前に計算した C_i を分散共分散行列、平均が0の正規分布を利用して、 z_i を生成する。

$$z_i \sim N(0, C_i)$$

(iii) C_i, V_i と z_i を利用して、以下の式に従って r_i を計算する。

$$r_i = D_{i+1}^{-1} e_{i+1} + L_{i+1} r_{i+1} + V_{i+1}' C_{i+1}^{-1} z_{i+1}$$

(iv) C_i, V_i と D_i を利用して、以下の式に従って U_i を計算する。

$$U_i = D_{i+1}^{-1} + L_{i+1}' U_{i+1} L_{i+1} + V_{i+1}' C_{i+1}^{-1} V_{i+1}$$

(v) r_i と z_i を利用して、以下の式に従って y_i を計算する。

¹³ 中妻（2003）の268ページを参照されたい。

¹⁴ 一般的な線形正規状態空間モデルにシミュレーション・スムーザを適用する場合の説明を付録に記述しておく。

$$y_i = \tilde{H}_i r_i + z_i$$

(v)で生成された y_i はフルコンディショナルな事後分布から生成された η_i の乱数になる。そこで生成された η_i の乱数と遷移方程式 (11a,11b) を用いて ψ_i を計算する。こうして計算された $\{\psi_i\}_{i=1}^n$ は $f(\psi_1, \psi_2, \dots, \psi_n | \tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)$ から生成された乱数となる。

上述したとおり、(11a,11b) と (22) という線形正規線形関係にシミュレーション・スムーザを適用して生成された ψ_i は、(9) と (11a,11b) から生成した ψ_i の乱数とはならない。そこで、(11a,11b) と (22) から生成した乱数を M-H 法によって採択することによって、(9) と (11a,11b) から生成した ψ_i の乱数を得る。

シミュレーション・スムーザを適用する際、採択確率が低くなり過ぎないためにブロック・サンプリングを適用する。そのためにカルマン・フィルターとシミュレーション・スムーザの両方の過程を実行する際にはデータを一定数に限定した上で実行することにした。したがって、シミュレーション・スムーザにより生成したサンプルの採択もデータを一定数に限定した上で実行することにした。こうしたブロック・サンプリングについての詳細は付録を参照にされたい。

5 パラメータの推定

本稿では、行列 W が n 行 1 列で要素が全て 1 の n 次元列ベクトルと仮定する。つまり、遷移方程式 (11a) は定数項つきの 1 階の自己回帰モデルとする。

5.1 ϕ のサンプリング

ϕ の定常性を確保するために、 ϕ の事前分布として -1 よりも大きく $+1$ よりも小さな値をとりうる対象とする一様分布を仮定する。

(11a,11b) から対数尤度は以下のとおりである。

$$\begin{aligned} \ln L(\sigma_\eta, \phi, \delta | \psi_1, \dots, \psi_n) &= \sum_{i=2}^n \ln(2\pi\sigma_\eta^2)^{\frac{1}{2}} - \frac{(\psi_i - \delta - \phi\psi_{i-1})^2}{2\sigma_\eta^2} \\ &\quad + \ln\left(\frac{2\pi\sigma_\eta^2}{1-\phi^2}\right)^{\frac{1}{2}} - \frac{\left(\psi_1 - \frac{W_1'\delta}{1-\phi}\right)^2(1-\phi^2)}{2\sigma_\eta^2} \end{aligned} \quad (26)$$

ϕ の事前分布として一様分布を仮定するので、対数尤度の (26) を事後密度とみなして、 ϕ の提案分布を以下のとおりに導出する。

$$\phi | \sigma_\eta^2, \delta \sim N\left(\hat{\phi}, \frac{\sigma_\eta^2}{\sum_{i=1}^T (\psi_i - \bar{\psi})^2}\right) \quad (27)$$

ただし、 ϕ は $[-1, +1]$ の範囲内の値しかとらないと仮定する。また、 $\hat{\phi} = \frac{\sum_{i=2}^n (\psi_i - \bar{\psi})(\psi_{i-1} - \bar{\psi})}{\sum_{i=1}^{n-1} (\psi_i - \bar{\psi})^2}$

である。(27) の提案分布から ϕ のサンプルを生成して、採択確率に従って選択する。

5.2 δ のサンプリング

前述したとおり、遷移方程式 (11a) は定数項つきの1階の自己回帰モデルとする。 δ の事前分布として一様分布を仮定する。事前分布として一様分布を仮定するので、対数尤度の (26) が事後密度である。そこで、(26) を δ の適当な確率分布で近似する、つまり、 δ の提案分布を以下のとおりに導出する。

$$\delta | \sigma_\eta^2, \phi \sim N \left(\hat{\delta}, \sigma_\eta^2 \left(\frac{1}{n} + \frac{\bar{\psi}^2}{\sum_{i=1}^{n-1} (\psi_i - \bar{\psi})^2} \right) \right) \quad \text{ただし、} \hat{\delta} = \bar{\psi} - \phi \bar{\psi} \text{ である。} \quad (28)$$

(28) の提案分布から δ のサンプルを生成して、採択確率に従って選択する。

5.3 σ_η^2 のサンプリング

σ_η^2 が正值だけをとる値とすることを確保するために、 σ_η^2 の事前分布として0よりも大きな値をとる対象とする確率分布を仮定する。対数尤度は (26) なので、 σ_η^2 の事前分布の確率密度として $p(\sigma_\eta) \propto \frac{1}{\sigma_\eta}$ を仮定した場合の事後密度は以下のとおりである¹⁵。

事後密度 $\propto \ln L(\sigma_\eta, \phi, \delta | \psi_1, \dots, \psi_n) \times p(\sigma_\eta)$

$$\begin{aligned} &= \ln(\sigma_\eta^2)^{\frac{n-1}{2}} - \frac{1}{\sigma_\eta^2} \sum_{i=2}^n \frac{(\psi_i - \delta - \phi \psi_{i-1})^2}{2} + \ln \left(\frac{2\pi\sigma_\eta^2}{1-\phi^2} \right)^{-\frac{1}{2}} - \frac{1-\phi^2}{2\sigma_\eta^2} \left(\psi_1 - \frac{W_1' \delta}{1-\phi} \right)^2 + \ln(\sigma_\eta^2)^{-\frac{1}{2}} \\ &\propto \ln(\sigma_\eta^2)^{\left(\frac{n-1}{2} + \frac{1}{2} + \frac{1}{2}\right)} - \frac{1}{\sigma_\eta^2} \frac{\left(\psi_1 - \frac{W_1' \delta}{1-\phi} \right)^2 (1-\phi^2) + \sum_{i=2}^n (\psi_i - \delta - \phi \psi_{i-1})^2}{2} \\ &\propto \ln(\sigma_\eta^2)^{-(n+1)} - \frac{1}{\sigma_\eta^2} \frac{\left(\psi_1 - \frac{W_1' \delta}{1-\phi} \right)^2 (1-\phi^2) + \sum_{i=2}^n (\psi_i - \delta - \phi \psi_{i-1})^2}{2} \end{aligned}$$

したがって、 σ_η のフルコンディショナルな事後分布は以下に示す平方根逆ガンマ分布に従う。

$$\sigma_\eta | \phi, \psi_1, \psi_2, \psi_3, \dots, \psi_n \sim IG^{\frac{1}{2}} \left(\frac{n}{2}, \frac{\left(\psi_1 - \frac{W_1' \delta}{1-\phi} \right)^2 (1-\phi^2) + \sum_{i=2}^n (\psi_i - \delta - \phi \psi_{i-1})^2}{2} \right)$$

ゆえに、2乗した σ_η^2 のフルコンディショナルな事後分布は以下に示す逆ガンマ分布に従う。

$$\sigma_\eta^2 | \phi, \psi_1, \psi_2, \psi_3, \dots, \psi_n \sim IG \left(\frac{n}{2}, \frac{\left(\psi_1 - \frac{W_1' \delta}{1-\phi} \right)^2 (1-\phi^2) + \sum_{i=2}^n (\psi_i - \delta - \phi \psi_{i-1})^2}{2} \right) \quad (29)$$

したがって、 σ_η^2 のフルコンディショナルな事後分布である(29)を用いて σ_η^2 をサンプリングすれば良い。

6 推定結果

本稿では試験的に2016年4月11日に東京商品取引所で取引された金先物(現物受渡用)のティックデータを用いてSCDモデルを推定した。具体的には、誤差項に指数分布を仮定したSCDモデル、つまり、(9)、(11a,11b)と(8)式に基づいたモデルのパラメータをベイズ推定した。ベイズ推定に当たっては、全部で5,000回のサンプリングを行い、後半の3,000回のサンプリングの結果を用いて事後平均などを計算した。

$$x_i = \Psi_i \varepsilon_i \quad i = 1, 2, \dots, n \quad (9)$$

$$\psi_i = W_i \delta + \phi \psi_{i-1} + \eta_i \quad \eta_i \sim N(0, \sigma_\eta^2) \quad (11a)$$

$$\psi_1 \sim N\left(0, \frac{\sigma_\eta^2}{1 - \phi^2}\right) \quad (11b)$$

$$\varepsilon_i \sim i.i.d. \text{Exp}(1) \quad f(\varepsilon_i) = \exp(-\varepsilon_i) \quad (8)$$

推定を行なう前に取引間隔(デュレーション)のデータを平滑化によって加工しなくてはならない。平滑化の方法は幾つかあるが、森保(2006)に従ってスーパースムーザを利用して、取引間隔のデータを平滑化した。原データを平滑化した値で割ることによって確定的な日中の傾向を予め除去しておく¹⁶。取引間隔を平滑化した結果が図1である。横軸は取引開始時刻近く(9時)からの経過時間(単位は秒)である。午前中は最初の1時間半、次の1時間、午後には山があり、大きく3つの山があることが読み取れる。

平滑化した値で取引間隔 x_i を割った値を新たにデータ x_i と定めてSCDモデルを当てはめる。パラメータを推定した結果が表2である。状態変数の系列 $\{\psi_i\}_{i=1}^n$ について推定した結果を図2に示した。こちらは、横軸として取引された順番を選んでいる。

表2 SCDモデルの推定結果

パラメータ	事後平均	事後標準偏差	事後自己相関
δ	0.004955	0.005017	0.220067
ϕ	0.855059	0.022434	0.409312
σ_η^2	0.027147	0.002642	0.712333

¹⁵ あるいは、 $p(\sigma_\eta^2) \propto \frac{1}{\sigma_\eta^2}$ を仮定しても同じである。

¹⁶ Engle=Russell(1998)の1137ページに期待デュレーションは確定的な成分と確率的な成分に分けられると記述されている。

定数項 δ は0.00496であり、有意に0であることを棄却できない。しかし、回帰係数 ϕ は0.855であり有意に0と異なると判定された。採択率は約97.7%であった。 $\{\psi_i\}_{i=1}^n$ の採択率は約56.7%であった。また、DICは4472.84、状態変数の系列 $\{\psi_i\}_{i=1}^n$ の事後平均を用いて計算した対数尤度は-2117.84であった。

状態変数の系列 $\{\psi_i\}_{i=1}^n$ を3,000回生成した結果を平均した値の推移が図2である。横軸は取引された順番である。図2を見ると、状態変数の系列 $\{\psi_i\}_{i=1}^n$ の事後平均の系列は0の周囲を変動していることが分かる。 $\psi_i = \ln \Psi_i$ であるから Ψ_i が1に近い値となっていると考えられる。

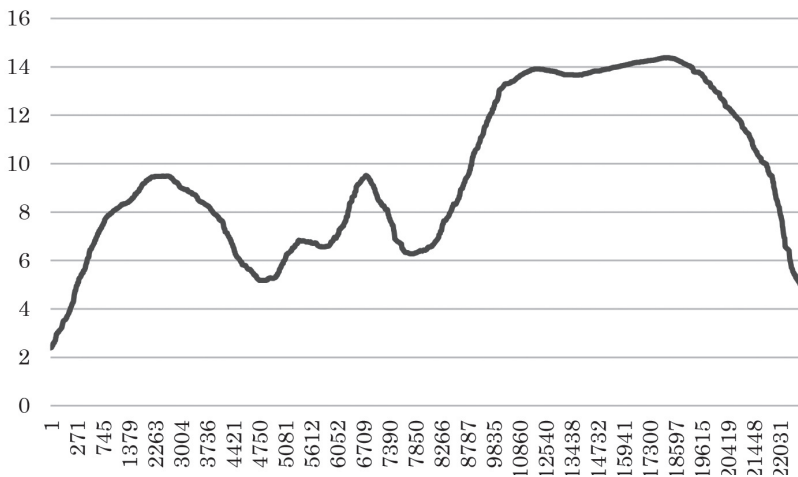


図1 デュレーション ξ_i を平滑した結果（縦軸は取引間隔、横軸は経過時間）

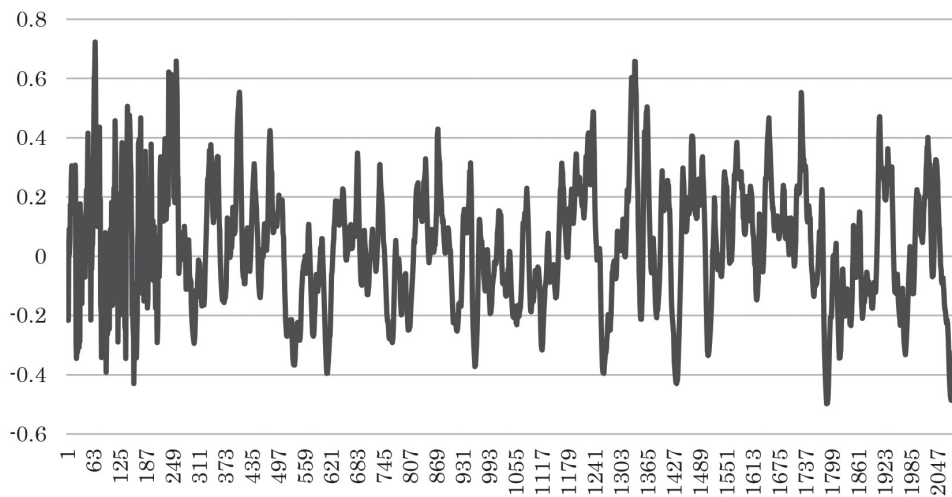


図2 ψ_i の事後平均の推移（縦軸は ψ_i 、横軸は取引された順番）

比較のため、ACDモデルを用いてパラメータを推定した結果が表3である。ベイズ推定に当たっては、全部で10,000回のサンプリングを行い、後半の5,000回のサンプリングの結果を用いて事後平均などを計算した。

表3 ACDモデルの推定結果

パラメータ	事後平均	事後標準偏差	事後自己相関
δ	0.262043	0.008933	0.009021
α	0.123917	0.002643	-0.016410
β	0.637249	0.000766	-0.014000

DICは4484.977、状態変数の系列 $\{\psi_i\}_{i=1}^n$ の事後平均を用いて計算した対数尤度は-2241.036であった。DICを計算してモデルの適合性を検討すると、このデータにはACDモデルよりもSCDモデルの方が適切であると考えられる。

7 結論

本稿の主たる目的はSCDモデルとその推定方法を紹介することである。第2節において、SCDモデルを基礎から説明した。具体的には、点過程、ポアソン過程とACDモデルを説明した後でSCDモデルについて説明した。第3節では、分析する取引間隔（デュレーション）データについて説明した後、平滑化について説明した。平滑化の手法はいろいろあるが、本稿では森保(2006)と同じスーパースムーザを利用して平滑化したデータを利用して推定した。第4節と第5節では状態変数 ψ_i とパラメータの推定方法についてSCDモデルの先行研究であるStrickland他(2006)の推定方法を紹介した。第6節では、東京商品取引所の金先物を例にSCDモデルのパラメータをベイズ推定した。さらに、ACDモデルのパラメータをベイズ推定した結果と比較した。DICに従うと実証分析に用いたデータの場合、ACDモデルよりもSCDモデルの方が相応しいことがわかった。

SCDモデルの拡張については今後の研究で行なう予定である。

参考文献

- [1] Bauwens, L., and Giot, P., 2000, "The logarithmic ACD model : An application to the bid-ask quote process of three NYSE stocks," *Annales d'Economie et de Statistique*, 60, pp.117-149.
- [2] Bauwens, L., and Veredas, D., 2004, "The stochastic conditional duration model: a latent variable

- model for the analysis of financial durations,” *Journal of Econometrics*, 119, pp.381-412.
- [3] de Jong, P., and Shephard, N., 1995, “The simulation smoother for time series models,” *Biometrika*, 82 (2) , pp.339-350.
- [4] Engle, R.F., and Russel, J.R., 1997, “Forecasting the frequency of changes in quoted foreign exchange price with the ACD model,” *Journal of Empirical Finance*, 4 ,pp.187 - 212.
- [5] Engle, R.F. and Russel, J.R., 1998, “Autoregressive conditional duration: a new model for irregularly spaced transaction data ,” *Econometrica*, 66 (5) , pp.1127-1162.
- [6] Feng, D., Jiang, G. J., and Song, P., 2004, “Stochastic conditional duration models with “Leverage Effect” for financial transaction data ,” *Journal of Financial Econometrics*, 2 (3) , pp.390 - 421.
- [7] 伏見正則, 1987, 『確率と確率過程』, 講談社.
- [8] 林高樹・佐藤彰洋, 2016, 『金融市場の高頻度データ分析』, 朝倉書店.
- [9] 林高樹, 2012, 「高頻度データの分析(2):不等間隔データ分析」, 刈屋武昭他編, 『経済時系列ハンドブック』の第6章4節に所収, 朝倉書店.
- [10] Men, Z., and Kolkiewicz, A. W., Wirjanto, T. S., 2015, “Bayesian analysis of asymmetric stochastic conditional duration model,” *Journal of Forecasting*, 34, pp.36 - 56.
- [11] 森保洋, 「金先物市場の日中価格変動と取引時間間隔」, 『経営と経済』, 長崎大学経済学会, 2006, p p. 408 - 430.
- [12] O'Hara, M., *Market Microstructure Theory*, Blackwell, 1995, ; 大村・宇野・宗近共訳, 『マーケット・マイクロストラクチャー』, きんざい, 1996.
- [13] 浪花貞夫=森保洋, 2010, 「フィナンシャル・エコノミクス」, 蓑谷・牧編, 『応用計量経済学ハンドブック』の第6章に所収, 朝倉書店.
- [14] 中妻照雄, 2003, 『ファイナンスのためのMCMCによるベイズ分析』, 三菱経済研究所.
- [15] 尾形良彦, 2008, 「地震活動の統計科学」, 小西・國友編, 『21世紀の統計科学Ⅱ』の第1章に所収, 東京大学出版会.
- [16] 尾形良彦, 1998, 「点過程モデル」, 尾崎・北川編, 『時系列解析の方法』の第15章に所収, 朝倉書店.
- [17] 大森・渡部, 2008, 「MCMCとその確率的ボラティリティー変動モデルへの応用」, 國友・山本編, 『21世紀の統計科学Ⅰ』の第9章に所収, 東京大学出版会.
- [18] Russel, J.R., and Engle, R.F., 2005, “A discrete-state continuous-time model of financial transactions prices and times : the autoregressive conditional multinomial - autoregressive conditional duration model,” *Journal of Business & Economic Statistics*, 23 (2) , pp.166-180.
- [19] Strickland, C.M., Forbes, C.S., and Martin, G.M., 2006, “Bayesian analysis of stochastic conditional duration model,” *Computational Statistics & Data Analysis*, 50, pp.2247-2267.

- [20] 杉浦航・中妻照雄, 2016, 「ティックデータを用いた株式市場における約定予測」, 日本金融・証券計量・工学学会編『ファイナンスにおける数値計算手法の新展開』の第4章に所収, 朝倉書店.
- [21] 竹内啓編集代表, 1989, 『統計学辞典』, 東洋経済新報社.
- [22] 竹澤邦夫, 2007, 『みんなのためのノンパラメトリック回帰 (上), 第3版』, 吉岡書店.
- [23] Zhang, M., Russel, J.R. and Tsay, R., 2001, "A nonlinear autoregressive conditional duration model with applications to financial transaction data," *Journal of Econometrics*, 104 (1), pp.179-207.

付 録¹⁷

[1] カルマン・フィルター

カルマン・フィルターは過去または現在までの情報を利用して状態変数 α_t や観測される変数 y_t の平均、分散と確率分布を計算する方法である。以下のような線形正規状態空間モデルについて考えてみる。

$$\begin{aligned} y_t &= Z\alpha_t + G_t u_t & G_t u_t &\sim N(0, G_t G_t') \text{ なお, } G_t G_t' = \Sigma_t \\ \alpha_{t+1} &= T\alpha_t + H u_t & H u_t &\sim N(0, H_t H_t') \text{ なお, } H_t H_t' = \Omega_t \\ u_t &\sim N(0, I) \quad t=1, 2, \dots, T \end{aligned}$$

この場合、過去または現在までの情報を利用して導出される状態変数や観測される変数の確率分布は以下のとおりである。

$$\begin{aligned} \alpha_t | I_{t-1} &\sim N(\hat{\alpha}_t, P_t) \cdots \text{予測} \\ y_t | I_{t-1} &\sim N(Z\hat{\alpha}_t, D_t) \cdots \text{予測} \\ \alpha_t | I_t &\sim N(\hat{\mu}_t, \hat{\Sigma}_t) \cdots \text{フィルタリング} \end{aligned}$$

$$\text{ただし, } \hat{\alpha}_{t+1} = \begin{cases} T_0 \alpha_0 & t=0 \\ T_t \hat{\alpha}_t + K_t e_t & t \geq 1 \end{cases}, \quad P_{t+1} = \begin{cases} H_0 H_0' & t=0 \\ T_t P_t L_t' + H_t J_t' & t \geq 1 \end{cases}$$

$$\hat{\mu}_t = \hat{\alpha}_t + P_t Z' D_t^{-1} e_t, \quad \hat{\Sigma}_t = P_t - P_t Z' D_t^{-1} Z P_t$$

$$e_t = y_t - Z\hat{\alpha}_t, \quad K_t = (T P_t Z' + H_t G_t') D_t^{-1}$$

$$D_t = Z P_t Z' + G_t G_t', \quad L_t = T_t - K_t Z, \quad J_t = H_t - K_t G_t$$

$t=1$ において以下の順番で状態変数 α_1 の2種類の条件付き平均 $\hat{\alpha}_1$ と $\hat{\mu}_1$ 、2種類の条件付き分散 P_1 と $\hat{\Sigma}_1$ を計算する。

(i) $\hat{\alpha}_1$ と P_1 と e_1 の計算

¹⁷ 付録の記述に関しては中妻(2003)と大森・渡部(2008)に負うところが多い。

なお、カルマン・フィルターだけは下付き添字を t にしてある。

(ii) K_t と D_t の計算

(iii) $\hat{\mu}_t$ と $\hat{\Sigma}_t$ の計算

この作業を $t=2, 3, \dots, T$ と継続することで $t=2, \dots, T$ の各時点において状態変数 α_t の2種類の条件付き平均 $\hat{\alpha}_t$ と $\hat{\mu}_t$ 、2種類の条件付き分散 P_t と $\hat{\Sigma}_t$ を計算する。こうして $\{\hat{\alpha}_t\}_{t=1}^T$ 、 $\{P_t\}_{t=1}^T$ 、 $\{\hat{\mu}_t\}_{t=1}^T$ と $\{\hat{\Sigma}_t\}_{t=1}^T$ を得る。

[2] シミュレーション・スムーザ

シミュレーション・スムーザとは、カルマン・フィルターの結果を利用して、観測方程式の誤差項 $\varepsilon_t = G_t u_t$ の候補、あるいは遷移方程式の誤差項 $\eta_t = H_t u_t$ の候補を生成する方法である。カルマン・フィルターで推定した結果を保存しておき、上記の誤差項の候補を生成する際に利用する。

$t=T$ において以下の (i) で C_T と V_T を計算し、(ii) で z_T 、(v) で x_T を計算する。つまり、時点 T では (iii) と (iv) を飛ばし、 r_T と U_T を計算せずに $r_{T=0}$ と $U_{T=0}$ とおく。 $t=T-1$ では (i) から (v) の順番で C_{T-1} 、 V_{T-1} 、 z_{T-1} 、 r_{T-1} 、 U_{T-1} 、 x_{T-1} を計算する。この作業を $t=T-2, T-3, \dots, 2, 1$ と継続することで、 $t=1, 2, \dots, T-1$ の各時点における C_t 、 V_t 、 z_t 、 r_t 、 U_t 、 x_t を計算する。

(i) カルマン・フィルターの結果を利用して、 C_t, V_t を計算する。

$$C_t = \begin{cases} \Sigma_t - \Sigma_t' N_t \Sigma_t & \varepsilon_t = G_t u_t \text{ を生成する場合} \\ \Omega_t - \Omega_t' U_t \Omega_t & \eta_t = H_t u_t \text{ を生成する場合} \end{cases}$$

$$V_t = \begin{cases} \Sigma_t (N_t Z - K_t' U_t T) & \varepsilon_t = G_t u_t \text{ を生成する場合} \\ \Omega_t U_t L_t & \eta_t = H_t u_t \text{ を生成する場合} \end{cases} \quad \text{ただし、} N_t = D_t^{-1} + K_t' U_t K_t$$

(ii) 直前に計算した C_t を分散共分散行列、平均が0の正規分布を利用して、 z_t を生成する。

$$z_t \sim N(0, C_t)$$

(iii) C_t 、 V_t と z_t を利用して、以下の式に従って r_t を計算する。

$$r_t = Z' D_{t+1}^{-1} e_{t+1} + L_{t+1}' r_{t+1} + V_{t+1}' C_{t+1}^{-1} z_{t+1}$$

(iv) C_t 、 V_t と D_t を利用して、以下の式に従って U_t を計算する。

$$U_t = Z' D_{t+1}^{-1} Z + L_{t+1}' U_{t+1} L_{t+1} + V_{t+1}' C_{t+1}^{-1} V_{t+1}$$

(v) r_t と z_t を利用して、以下の式に従って x_t を計算する。

$$x_t = \begin{cases} \Sigma_t (D_t^{-1} e_t - K_t' r_t) + z_t & \varepsilon_t = G_t u_t \text{ を生成する場合} \\ \Omega_t r_t + z_t & \eta_t = H_t u_t \text{ を生成する場合} \end{cases}$$

ε_t を生成する場合という式に従って (v) で生成された x_t はフルコンディショナルな事後分布から生成された ε_t の乱数になる。一方、 η_t を生成する場合という式に従って (v) で生成された x_t はフルコンディショナルな事後分布から生成された η_t の乱数になる。

以下の式に従い、 ε_t の乱数を生成した場合、 ε_t の乱数と観測方程式を用いて α_t を計算する。一方、 η_t の乱数を生成した場合、 η_t の乱数と遷移方程式を用いて α_t を計算する。計算された $\{\alpha_1, \alpha_2, \dots, \alpha_T\}$ は $f(\alpha_1, \alpha_2, \dots, \alpha_T | y_1, y_2, \dots, y_T)$ から生成された乱数となる。

$$\alpha_t = \begin{cases} Z^{-1}(y_t - x_t) & \varepsilon_t = G_t u_t \text{ を生成する場合} \\ T\alpha_{t-1} + x_{t-1} & \eta_t = H_t u_t \text{ を生成する場合} \end{cases}$$

[3] ブロック・サンプリング

状態変数の候補を生成する際に一度に生成するデータの大きさが極端に大きいと、採択確率がかなり低くなることが知られている。そこで、推定に利用するデータを細かく区分することによって、一度に生成するデータが多くなり過ぎないようにする。なお、データを区分する際の境目は確率的に決める。

本稿の場合、状態変数 ψ_i を生成する際に利用する。(9)、(10)、(11a,11b)で構成されるSCDモデルの推定を例にとりて説明する。

$$x_i = \Psi_i \varepsilon_i \tag{9}$$

$$\varepsilon_i \sim i.i.d. \tag{10}$$

$$\psi_i = W_i \delta + \phi \psi_{i-1} + \eta_i \quad \eta_i \sim N(0, \sigma_\eta^2) \tag{11a}$$

$$\psi_1 \sim N\left(0, \frac{\sigma_\eta^2}{1 - \phi^2}\right) \tag{11b}$$

以下の確率分布に従う $k+1$ 個の状態変数 $\psi_i, \psi_{i+1}, \dots, \psi_{i+k}$ を生成することを考えてみる。

$$p(\psi_i, \psi_{i+1}, \dots, \psi_{i+k-1}, \psi_{i+k} | \psi_{i+k+1}, \psi_{i-1}, \phi, \delta, \sigma_\eta, W) = \frac{p(\psi_i, \psi_{i+1}, \dots, \psi_{i+k}, \psi_{i+k+1} | \psi_{i-1}, \phi, \delta, \sigma_\eta, W)}{p(\psi_{i+k+1} | \phi, \delta, \sigma_\eta, W)}$$

上式の分子にある結合分布の密度を条件付分布の積に変形すると、

$$\begin{aligned} & p(\psi_i, \psi_{i+1}, \dots, \psi_{i+k-1}, \psi_{i+k}, \psi_{i+k+1} | \psi_{i-1}, \phi, \delta, \sigma_\eta, W) \\ &= p(\psi_i | \psi_{i-1}, \phi, \delta, \sigma_\eta, W) \times p(\psi_{i+1} | \psi_i, \phi, \delta, \sigma_\eta, W) \times \dots \times p(\psi_{i+k+1} | \psi_{i+k}, \phi, \delta, \sigma_\eta, W) \end{aligned}$$

である。この式を元の式へ代入すると、

$$\begin{aligned} &= \frac{p(\psi_i | \psi_{i-1}, \phi, \delta, \sigma_\eta, W) p(\psi_{i+1} | \psi_i, \phi, \delta, \sigma_\eta, W) \times \dots \times p(\psi_{i+k+1} | \psi_{i+k}, \phi, \delta, \sigma_\eta, W)}{p(\psi_{i+k+1} | \phi, \delta, \sigma_\eta, W)} \\ &\propto p(\psi_i | \psi_{i-1}, \phi, \delta, \sigma_\eta, W) \times p(\psi_{i+1} | \psi_i, \phi, \delta, \sigma_\eta, W) \times \dots \times p(\psi_{i+k} | \psi_{i+k-1}, \phi, \delta, \sigma_\eta, W) \end{aligned}$$

が得られる。 ψ_i の動きを記述する(11a)から $\psi_i | \psi_{i-1}, \phi, \delta, \sigma_\eta$ から ψ_i を生成することと、 $\eta_i | \sigma_\eta$ から η_i を生成することは同じであることがわかる。同様に、 $\{\psi_i, \psi_{i+1}, \dots, \psi_{i+k-1}, \psi_{i+k}\}$ を生成することは $\{\eta_i, \eta_{i+1}, \dots, \eta_{i+k-1}, \eta_{i+k}\}$ を生成することと同じであることがわかる。

$$\begin{aligned}
 & p(\psi_i, \psi_{i+1}, \dots, \psi_{i+k-1}, \psi_{i+k} | \psi_{i+k+1}, \psi_{i-1}, \phi, \delta, \sigma_\eta, W) \\
 & \propto p(\eta_i | \sigma_\eta) \times p(\eta_{i+1} | \sigma_\eta) \times \dots \times p(\eta_{i+k} | \sigma_\eta) \times p(\psi_{i+k+1} | \psi_{i+k}, \phi, \delta, \sigma_\eta, W) \\
 & \propto p(\eta_i, \eta_{i+1}, \dots, \eta_{i+k-1}, \eta_{i+k} | \sigma_\eta) \times p(\psi_{i+k+1} | \psi_{i+k}, \phi, \delta, \sigma_\eta, W) \tag{A1}
 \end{aligned}$$

本稿では、サンプラーが相関を持たないようにするため、シミュレーション・スムーザを用いて、 $\{\eta_s\}_{s=i}^{i+k}$ を生成し、(11a) を通じて間接的に $\{\psi_s\}_{s=i}^{i+k}$ を生成する。したがって、上記の式変形が必要である。ここで、 $\{\psi_i, \psi_{i+1}, \dots, \psi_{i+k-1}, \psi_{i+k}\}$ の事後密度を計算する。

$$\begin{aligned}
 & p(\psi_i, \psi_{i+1}, \dots, \psi_{i+k-1}, \psi_{i+k} | x_i, x_{i+1}, \dots, x_{i+k}, \psi_{i+k+1}, \psi_{i-1}, \phi, \delta, \sigma_\eta, W) \\
 & = \frac{p(x_i, x_{i+1}, \dots, x_{i+k} | \psi_i, \psi_{i+1}, \dots, \psi_{i+k}) p(\psi_i, \psi_{i+1}, \dots, \psi_{i+k} | \psi_{i-1}, \psi_{i+k+1}, \phi, \delta, \sigma_\eta, W)}{p(x_i, x_{i+1}, \dots, x_{i+k})} \tag{A2}
 \end{aligned}$$

(A1) を (A2) へ代入すると、次式が得られる。

$$= \frac{p(x_i, x_{i+1}, \dots, x_{i+k} | \psi_i, \psi_{i+1}, \dots, \psi_{i+k}) p(\eta_i, \eta_{i+1}, \dots, \eta_{i+k} | \sigma_\eta) p(\psi_{i+k+1} | \psi_{i+k+1}, \phi, \delta, \sigma_\eta, W)}{p(x_i, x_{i+1}, \dots, x_{i+k})} \tag{A3}$$

さらに、(A3) を対数変換すると次式が得られる。

$$\begin{aligned}
 & \ln p(\psi_i, \psi_{i+1}, \dots, \psi_{i+k-1}, \psi_{i+k} | x_i, x_{i+1}, \dots, x_{i+k}, \psi_{i+k+1}, \psi_{i-1}, \phi, \delta, \sigma_\eta, W) \\
 & \propto \ln p(x_i, x_{i+1}, \dots, x_{i+k} | \psi_i, \psi_{i+1}, \dots, \psi_{i+k}) + \ln p(\eta_i, \eta_{i+1}, \dots, \eta_{i+k} | \sigma_\eta) + \ln p(\psi_{i+k+1} | \psi_{i+k+1}, \phi, \delta, \sigma_\eta, W) \\
 & \propto \text{定数} + \sum_{s=i}^{i+k} \ln p(x_s | \psi_s) + \sum_{s=i}^{i+k} \ln p(\eta_s | \sigma_\eta) + \ln p(\psi_{i+k+1} | \psi_{i+k+1}, \phi, \delta, \sigma_\eta, W) \tag{A4}
 \end{aligned}$$

SCDモデルから導出される $\{\psi_s\}_{s=i}^{i+k}$ の事後分布から $\{\psi_s\}_{s=i}^{i+k}$ を生成する際には、(A4) を利用する。

\tilde{x}_i を観測変数、 ψ_i を状態変数、(22) を観測方程式、(11a,11b) を遷移方程式とする線形正規状態空間モデルにおける ψ_i の事後密度は、SCDモデルにおける ψ_i の提案分布に対応する。SCDモデルの対数事後密度は (14)、近似した線形正規状態空間モデルの事後密度は (21) である。遷移方程式は (11a,11b) で共通であるから、両者の違いは観測方程式である (9') と (22) の第1項と第3項に反映されるが、それ以外は共通である。

\tilde{x}_i を観測変数とする線形状態空間モデルにおける ψ_i の事後密度は以下のとおりである。

$$\begin{aligned}
 & p(\psi_i, \psi_{i+1}, \dots, \psi_{i+k} | \tilde{x}_i, \tilde{x}_{i+1}, \dots, \tilde{x}_{i+k}, \psi_{i-1}, \psi_{i+k+1}, \phi, \delta, \sigma_\eta, W) \\
 & = \frac{p(\tilde{x}_i, \tilde{x}_{i+1}, \dots, \tilde{x}_{i+k} | \psi_i, \psi_{i+1}, \dots, \psi_{i+k}) p(\psi_i, \psi_{i+1}, \dots, \psi_{i+k} | \psi_{i-1}, \psi_{i+k+1}, \phi, \delta, \sigma_\eta, W)}{p(\tilde{x}_i, \tilde{x}_{i+1}, \dots, \tilde{x}_{i+k})} \tag{A2'}
 \end{aligned}$$

(A1) を (A2') へ代入すると、次式が得られる。

$$= \frac{p(\tilde{x}_i, \tilde{x}_{i+1}, \dots, \tilde{x}_{i+k} | \psi_i, \psi_{i+1}, \dots, \psi_{i+k}) p(\eta_i, \eta_{i+1}, \dots, \eta_{i+k} | \sigma_\eta) p(\psi_{i+k+1} | \psi_{i+k+1}, \phi, \delta, \sigma_\eta, W)}{p(\tilde{x}_i, \tilde{x}_{i+1}, \dots, \tilde{x}_{i+k})} \tag{A3'}$$

さらに、(A3') を対数変換すると次式が得られる。

$$\begin{aligned}
 & \ln p(\psi_i, \psi_{i+1}, \dots, \psi_{i+k} | \tilde{x}_i, \tilde{x}_{i+1}, \dots, \tilde{x}_{i+k}, \psi_{i+k+1}, \psi_{i-1}, \phi, \delta, \sigma_\eta^2, W) \\
 & \propto \ln p(\tilde{x}_i, \tilde{x}_{i+1}, \dots, \tilde{x}_{i+k} | \psi_i, \psi_{i+1}, \dots, \psi_{i+k}, \tilde{H}_i, \tilde{H}_{i+1}, \dots, \tilde{H}_{i+k}) \\
 & \quad + \ln p(\eta_i, \eta_{i+1}, \eta_{i+2}, \dots, \eta_{i+k} | \sigma_\eta^2) + \ln p(\psi_{i+k+1} | \psi_{i+k}, \phi, \delta, \sigma_\eta^2, W) \tag{A4'} \\
 & = \sum_{s=i}^{i+k} \ln p(\tilde{x}_s | \psi_s, \tilde{H}_s) + \sum_{s=i}^{i+k} \ln f(\eta_s | \sigma_\eta^2) + \ln f(\psi_{i+k+1} | \psi_{i+k}, \phi, \delta, \sigma_\eta^2, W)
 \end{aligned}$$

シミュレーション・スモザを用いて $\{\psi_i, \psi_{i+1}, \dots, \psi_{i+k}\}$ をサンプリングした後、(A4) と (A4') を用いて計算される採択確率を用いて採否を決定する。こうして $\{\psi_i, \psi_{i+1}, \dots, \psi_{i+k}\}$ を生成する。