

棄却サンプリング連鎖を用いた場合の周辺尤度の推定 —誤差項に t 分布を仮定した自己回帰モデルを例に—

砂 田 洋 志
(人文学部 法経政策学科)

1. はじめに

1.1 はじめに

経済学で広く利用されてきたモデルは線形回帰モデルであり、その推定には最小二乗法が利用されてきた。しかし、現在では多様なモデルが考案され、その推定方法も多様なものが利用されている。本論文で扱うマルコフ・チェーン・モンテ・カルロ法（以後、MCMC法と略記する）を用いたベイズ推定も近年、利用される機会が広がった推定方法である。MCMCを用いたベイズ推定といっても、アルゴリズムによっていくつかに分かれる。広く利用されているのはギブズ・サンプリング(Gibbs Sampling)法、メトロポリス=ヘイスティング(Metropolis=Hastings)法である。しかし、それ以外の推定方法も考案されている。たとえば、データ拡大法や棄却サンプリング連鎖などである。

ところで、複数のモデル間でどのモデルが適切であるかを議論する場合、線形回帰モデルを最小二乗推定した場合は決定係数によって比較する一方、ベイズ推定した場合は周辺尤度を推定し、それを用いて計算されるベイズファクターによってモデルを比較する。周辺尤度の推定方法であるが、ギブズ・サンプリングを用いた場合の周辺尤度の推定方法はChib(1995)で、メトロポリス=ヘイスティング法を用いた場合の周辺尤度の推定方法はChib=Jeliazkov(2001)で説明されている。¹

ファイナンス研究では資産収益率が正規分布よりも裾の厚い分布をしていると考えられており、誤差項として t 分布などを仮定することが多い。 t 分布を仮定した場合には、棄却サンプリング連鎖を利用してベイズ推定される。² その場合の周辺尤度の推定方法は前述した2つの論文で解説されている推定方法とは異なるので、後からChib=Jeliazkov(2005)によって推定方法が提案された。彼らは一般化されたモデルの周辺尤度の推定方法を紹介した後で、ロジット・モデルなどを例として挙げて、それらの周辺尤度の推定方法を説明している。

MCMC法を用いたベイズ推定法に関する邦語文献は、現在のところ数えるほどであり、周辺

¹ ギブズ・サンプリングを用いた場合の周辺尤度の計算方法と、メトロポリス=ヘイスティング法を用いた場合の周辺尤度の計算方法については、砂田(2006)でも詳細な説明がなされている。

² 中妻(2003)や大森(2001)では棄却サンプリング連鎖と呼んでいるが、AR-MH(Accept-Reject Metropolis-Hasting Sampling)という呼び方もある。本論文では推定方法であることを明示するために、この後から最後に「法」を付け加える。

尤度に関する記述も少ない。³そこで、本論文では、棄却サンプリング連鎖法でパラメータをベイズ推定した場合に、周辺尤度を推定する方法を紹介する。

本論文では、理解を深めるために、以下の二点を工夫している。第一に、簡単な具体例を用いて推定方法を紹介すること、第二に、周辺尤度を実際に推定してその結果を示すこと、である。この詳細は以下の通りである。

棄却サンプリング連鎖法でパラメータをベイズ推定した場合の周辺尤度を推定する方法を紹介するに当たっては、まず、Chib=Jeliazkov(2005)を参考にしながら、一般的なモデルを用いて紹介してある。その後で、誤差項に*t*分布を仮定した自己回帰モデルを具体例として取り上げて、周辺尤度の推定方法を紹介してある。このモデルを選んだ理由は、誤差項に*t*分布を仮定したモデルが、金融データの実証分析で広く利用されており、自己回帰モデルがその中で、最も単純なモデルの一つだからである。

周辺尤度の推定方法を紹介した後に、パラメータを与えて、誤差項が*t*分布に従う自己回帰モデルからシミュレーション・データを生成した。そのデータに対して、棄却サンプリング連鎖法を用いてパラメータをベイズ推定し、決定係数、対数尤度、対数周辺尤度も計算した。

また、モデル選択が適切に行われるかを確認するために、先ほど生成したシミュレーション・データに対して、誤差項が正規分布に従う1階の自己回帰モデルを当てはめて、決定係数、対数尤度と対数周辺尤度を計算して、比較した。シミュレーション・データを利用しているのは、真のモデルが分かっているので、モデル選択が適切に行われているかを確認できるからである。

本論文の構成は以下の通りである。第2節で、棄却サンプリング連鎖法の説明を行う。第3節で、誤差項に*t*分布を仮定した自己回帰モデルにおいて、パラメータをベイズ推定する方法について説明する。第4節で、棄却サンプリング連鎖法を用いて推定した場合の周辺尤度の計算方法をChib=Jeliazkov(2005)に基づいて説明する。第5節で、誤差項に*t*分布を仮定した自己回帰モデルにおける周辺尤度の推定方法を説明する。その後で、同モデルからのシミュレーション・データを用いてパラメータと周辺尤度を推定する。最後の第6節で結論を述べる。

1.2 周辺尤度の計算

周辺尤度 $m(y)$ の計算に利用する式を紹介しよう。データ、モデルのパラメータ、事前分布、尤度(LikelihoodあるいはSampling density)、事後分布をそれぞれ $y, \theta = \{\theta_1, \theta_2, \dots, \theta_B\}, \pi(\theta), f(y|\theta), \pi(\theta|y)$ と記述すると、モデルの周辺尤度は

$$m(y) = \frac{f(y|\theta)\pi(\theta)}{\pi(\theta|y)} \quad \dots(1)$$

³ MCMC法を用いたベイズ推定法に関する邦語文献は大森(2001)、中妻(2003)、伊庭他(2005)、和合(2005)などであり、周辺尤度の説明も簡潔に行われている。和合(2005)にはMCMC法を用いた実証分析も多数掲載されており、その中の幾つかで周辺尤度が計算されている。

棄却サンプリング連鎖を用いた場合の周辺尤度の推定—誤差項に t 分布を仮定した自己回帰モデルを例に—砂田

と記述できる。(1)の分子は尤度と事前分布の積であり、分母は θ の事後密度である。この式は任意の θ について成立つ式で、基本周辺尤度方程式(BMI, Basic Marginal likelihood Identity)と呼ばれる。

周辺尤度 $m(\mathbf{y})$ は θ に依存しないし、(1)は任意の θ について成立つので、周辺尤度 $m(\mathbf{y})$ を実際に計算する上では、計算に都合の良い $\theta=\theta^*$ における事後密度 $\pi(\theta^*|\mathbf{y})$ を利用して、 $m(\mathbf{y})$ を計算して構わない。 θ^* における事後密度の推定量を $\hat{\pi}(\theta^*|\mathbf{y})$ 、周辺尤度 $m(\mathbf{y})$ の推定量を $\hat{m}(\mathbf{y})$ と記述すれば、周辺尤度 $m(\mathbf{y})$ の推定量は、

$$\hat{m}(\mathbf{y}) = \frac{f(\mathbf{y}|\theta^*)\pi(\theta^*)}{\hat{\pi}(\theta^*|\mathbf{y})}$$

である。一般に周辺尤度を実際に計算する場合、上式の両辺を対数変換した

$$\ln \hat{m}(\mathbf{y}) = \ln f(\mathbf{y}|\theta^*) + \ln \pi(\theta^*) - \ln \hat{\pi}(\theta^*|\mathbf{y}) \quad \dots(2)$$

を利用する。(2)の推定では右辺第1項と第2項は尤度関数と事前分布なので容易に計算可能であるが、 $\hat{\pi}(\theta^*|\mathbf{y})$ の計算は容易でない。

$\hat{\pi}(\theta^*|\mathbf{y})$ の推定方法は、棄却サンプリング連鎖法を用いた場合、Chib(1995)やChib=Jeliazkov(2001)と異なる。したがって、本論文ではこの項の推定方法を中心に説明することとなる。

2. 棄却サンプリング連鎖法の概略

本節では確率密度関数 $f(\theta)$ に従う θ のサンプルを棄却サンプリング連鎖法によって得る方法を説明する。⁴

棄却サンプリング連鎖法は、 $f(\theta)$ という確率密度関数からの乱数を生成したいにも係わらず、 $f(\theta)$ から直接生成することが不可能であり、 $f(\theta)$ と似た確率密度関数 $h(\theta)$ からしか乱数を生成できない場合に利用される。そのような場合であっても、もし適当な大きさの K に対して $f(\theta) \leq K \cdot h(\theta)$ が全ての θ で満たされるならば、棄却採択法を用いることによって $f(\theta)$ という確率密度関数からの乱数を生成できる。しかし、 K を大きくしすぎると、効率的な乱数生成を行えないという問題点がある。また、生成したい乱数の確率密度関数が（一部でなく）完全に判っていなければならない。これは $f(\theta)$ の基準化定数が既知であることも意味する。棄却サンプリング連鎖法はこ

⁴ 棄却サンプリング連鎖法については大森(2001)、中妻(2003)、和合(2005)、砂田(2005)でも説明している。

これらの問題を避けたい場合に利用できる。

棄却サンプリング連鎖法は $f(\theta) \leq K \times h(\theta)$ が全ての θ で満たされる場合でなくても利用できる。例えば、 $f(\theta)$ の一部で $K \times h(\theta)$ が上回っているように K を選択すれば済む。また、基準化定数が分からなくても、後の節で説明するようにうまく相殺されてしまうので、乱数生成に問題が生じない。

以下では、生成したい乱数の確率密度関数を $f(\theta)$ と記述して、棄却サンプリング連鎖法の手順を示す。棄却サンプリング連鎖法はAR部分とM=H部分に大別されるので、順番に説明する。

AR部分

(1) θ' を提案分布 $h(\theta)$ からサンプリングする。

(2) θ' をサンプルとして採択するか、直前にサンプリングした値 θ を再びサンプルとして採択するか、を以下の採択確率で決める。

$$\alpha_{AR}(\theta, \theta') = \alpha_{AR}(\theta') = \min \left\{ 1, \frac{f(\theta')}{K \times h(\theta')} \right\}$$

M=H部分

AR部分で採択された θ' について、 θ' を採択するか、直前にサンプリングした値 θ を再び採択するかを、別の採択確率 $\alpha_{MH}(\theta, \theta')$ によってもう一度選択する。採択確率は θ' と θ の満たす条件によって以下の様に変化する。

$$\text{a. } f(\theta') \leq K \times h(\theta') \text{ ならば, } \alpha_{MH}(\theta, \theta') = \min \left\{ 1, \frac{K \times h(\theta)}{f(\theta)} \right\}$$

$$\text{b. } f(\theta') > K \times h(\theta') \text{ ならば, } \alpha_{MH}(\theta, \theta') = \min \left\{ 1, \frac{f(\theta') h(\theta)}{f(\theta) h(\theta')} \right\}$$

まず、棄却サンプリング連鎖法のAR部分を説明してみよう。

全ての θ が $f(\theta) \leq K \times h(\theta)$ という条件を必ずしも満たさない場合であっても、 $\min\{f(\theta), K \times h(\theta)\} \leq K \times h(\theta)$ という条件ならば全ての θ が満たす。したがって、 $\min\{f(\theta), K \times h(\theta)\}$ の基準化定数を $d(\theta) = [\min\{f(\theta), K \times h(\theta)\}] / d$ と記述すれば、全ての θ について次の不等式が成り立つ。

$$\frac{\min\{f(\theta), K \times h(\theta)\}}{d(\theta)} \leq \frac{K}{d(\theta)} \times h(\theta)$$

ではAR部分で行っているように、 $h(\theta)$ から θ のサンプルを生成して、採択確率

$$\min\left\{1, \frac{f(\theta^*)}{K \times h(\theta^*)}\right\} = \frac{\min\{f(\theta^*), K \times h(\theta^*)\}}{K \times h(\theta^*)}$$

で取捨選択した場合、いかなる確率分布に従うサ

ンプルを得られるのであろうか。

全ての θ について不等式

$$\frac{\min\{f(\theta), K \times h(\theta)\}}{d(\theta)} \leq \frac{K}{d(\theta)} \times h(\theta)$$

が成り立つので、このサンプリングは $h(\theta)$ から生成した θ のサンプルを採択確率

$$\frac{\min\{f(\theta^*), K \times h(\theta^*)\} / d(\theta^*)}{K / d(\theta^*) \times h(\theta^*)}$$

で取捨選択することと同じである。したがって、以下の $q(\theta)$ に従

う θ のサンプルが得られることとなる。

$$q(\theta) = \frac{\min\{f(\theta), K \times h(\theta)\}}{d(\theta)} \quad \dots(3)$$

ちなみに、 $\min\{f(\theta), K \times h(\theta)\} = K \times h(\theta) \times \alpha_{AR}(\theta)$ であることを用いると、 $q(\theta)$ は以下の様に記述することもできる。

$$q(\theta) = \frac{\alpha_{AR}(\theta) \times K \times h(\theta)}{d(\theta)} \quad \dots(3')$$

次に、M=H部分を説明してみよう。AR部分で生成した乱数の従う $q(\theta)$ は、目的とする $f(\theta)$ と異なる確率分布である。しかし、AR部分で生成した乱数に対して、さらにM=H法を利用することによって $f(\theta)$ からの乱数を生成できる。つまり、AR部分の $q(\theta)$ を提案密度関数と考えて θ を生成した後で、本来生成したい確率分布 $f(\theta)$ と提案分布 $q(\theta)$ から採択確率 $\alpha_{MH}(\theta, \theta^*)$ を計算し、その確率に従って採択するのである。

この採択確率は確率分布 $f(\theta)$ と提案分布 $q(\theta)$ を用いて以下のように記述できる。

$$\alpha_{MH}(\theta, \theta^*) = \min\left\{1, \frac{f(\theta^*)q(\theta)}{f(\theta)q(\theta^*)}\right\} = \min\left\{1, \frac{f(\theta^*) \times \min\{f(\theta), K \times h(\theta)\}}{f(\theta) \times \min\{f(\theta^*), K \times h(\theta^*)\}}\right\}$$

この採択確率を、 $f(\theta^*)$ と $K \times h(\theta^*)$ の大小で場合分けすれば、以下ようになる。⁵

$$\text{a. } f(\theta^*) < K \times h(\theta^*) \text{ ならば, } \alpha_{MH}(\theta, \theta^*) = \min\left\{1, \frac{K \times h(\theta)}{f(\theta)}\right\}$$

⁵ 詳細は砂田(2005)の付録 B を参考にされたい。

$$b. f(\theta^*) > K \times h(\theta^*) \text{ ならば、 } \alpha_{MH}(\theta, \theta^*) = \min \left\{ 1, \frac{f(\theta_i^*)h(\theta)}{f(\theta_i)h(\theta^*)} \right\}$$

以上から、 $h(\theta)$ と2種類の採択確率を利用して、 $f(\theta)$ に従うサンプルを生成できることが分かる。

3. 誤差項が t 分布に従う自己回帰モデルのパラメータ推定

本節では誤差項が t 分布に従う自己回帰モデル(以下では t -ARモデルと略記する)を対象にして、パラメータをベイズ推定する方法を解説する。⁶ベイズ推定するに当たっては情報事前分布(informed prior)を仮定する。

3.1節では、誤差項が t 分布に従う自己回帰モデルと誤差項が不均一分散に従う自己回帰モデルとの関係を述べる。3.2節では、誤差項が t 分布に従う自己回帰モデルのパラメータをベイズ推定する方法について述べる。

3.1 t -ARモデル

t -ARモデルを行列表記する際に用いる Y と X であるが、以下の通りである。

$$Y = \begin{pmatrix} y_{p+1} \\ y_{p+2} \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & y_p & y_{p-1} & \cdots & y_1 \\ 1 & y_{p+1} & y_p & \cdots & y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & y_{n-2} & y_{n-3} & \cdots & y_{n-p-1} \\ 1 & y_{n-1} & y_{n-2} & \cdots & y_{n-p} \end{pmatrix}$$

Y と X のデータを合わせて y と記述することにする。

誤差項が正規分布に従うものの、その分散が不均一であると仮定した回帰モデルをモデル1と呼ぼう。数式を用いると、以下のように記述できる。ただし、対角行列 Ω の対角要素を纏めて n 次元ベクトル $\omega = (\omega_1, \omega_2, \dots, \omega_{n-1}, \omega_n)'$ と記述する。

モデル1 : $Y = X\gamma + u \quad u \sim N(0, \sigma^2 \Omega) \quad \text{diag}(\Omega) = (\omega_1, \omega_2, \dots, \omega_{n-1}, \omega_n)$

ベイズ推定に必要な事前分布として、 γ には正規分布 $N(\gamma_0, A_0)$ 、 σ^2 には自然共役な逆ガンマ分布 $IG(v_0/2, v_0 \times s_0^2/2)$ 、 ω の各要素 ω_i には逆ガンマ分布 $IG\left(\frac{v_\lambda}{2}, \frac{v_\lambda}{2}\right)$ を共通に仮定する。そし

⁶ t 分布の自由度を棄却サンプリング連鎖で推定する方法については Watanabe(2001) や 中妻(2003) を参照されたい。本論文でも参考になっている。

棄却サンプリング連鎖を用いた場合の周辺尤度の推定—誤差項に t 分布を仮定した自己回帰モデルを例に—砂田

て尤度関数は以下の通りである。

$$L(y|\gamma, \sigma^2, \Omega) = \frac{1}{(2\sigma^2)^{\frac{n}{2}}} |\Omega|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2} (Y - X\gamma)' \Omega^{-1} (Y - X\gamma)\right\}$$

次に、誤差項が自由度 ν_2 の t 分布に従う均一分散の回帰モデルをモデル2と呼ぼう。数式を用いると、以下のように記述できる。

モデル2 : $Y = X\gamma + u \quad u \sim t(0, \sigma^2, \nu_2)$

つまり、モデル1と同じ回帰係数 γ 、同じ分散 σ^2 を有する回帰モデルであるが、誤差項は平均が0で分散が σ^2 の t 分布に従う。ただし、 t 分布の自由度はモデル1の Ω のハイパーパラメータ ν_1 と同じである。

モデル1では γ 、 σ^2 と ω に上述したような情報事前分布を仮定したが、その場合の事後密度と、モデル2で γ と σ^2 にモデル1と同じ情報事前分布を仮定した場合の事後密度は、Koop(2003)によれば同一である。

したがって、 ν_2 を所与として誤差項の分散が不均一な回帰モデル（モデル1）のパラメータをベイズ推定することは、自由度 ν_2 を所与とした t 分布に誤差項が従う回帰モデル（モデル2）のパラメータを推定したことに同一である。

本論文では、自由度 ν_2 も未知パラメータとしてモデル2のパラメータをベイズ推定したい。そこで、モデル2の γ, σ^2 と ν_2 という3つの未知パラメータをベイズ推定する際の方針を以下に示す。

モデル2で ν_2 を未知パラメータとした場合の事後密度は、上述した両モデルに共通な事後密度に ν_2 の事前密度をかけた値である。ゆえに、その事後密度からモデル1のパラメータ $\gamma, \sigma^2, \omega, \nu_1$ のフルコンディショナルな事後分布を導出して、ベイズ推定することにより、モデル2の γ, σ^2, ν_2 を推定できる。ただし、この場合は、 ν_2 の事前密度をかけても事後密度の中で γ, σ^2 に関する部分は変化しないから、 ν_2 を未知パラメータとしたモデル1のパラメータ γ, σ^2 のフルコンディショナルな事後分布は、 ν_2 を所与としたモデル1の γ, σ^2 のフルコンディショナルな事後分布と同一である。そこで、 ν_2 を未知パラメータとしたモデル2の γ, σ^2 は、 ν_2 を所与としたモデル1の γ, σ^2 のフルコンディショナルな事後分布をそのまま利用してベイズ推定する。 ω のフルコンディショナルな事後分布はモデル2で ν_2 を未知パラメータとした場合の事後密度から導出する。

最後に、未知パラメータである自由度 ν_1 の推定である。 ν_1 はフルコンディショナルな事後分布から簡単にサンプリングできない。そこで、棄却サンプリング連鎖法という効率的なサンプリング方法を用いて推定する。Geweke(1993)は、棄却サンプリング連鎖法と比べて効率性の低い採択棄却法を用いて、自由度を含めたパラメータの推定を行っている。

以上のように、不均一分散モデルの推定方法を利用しながら、誤差項が t 分布に従うモデル

の推定を行う。

本論文におけるベイズ推定では、 $\omega, \gamma, \sigma^2, v_\lambda$ をサンプリングして、その平均を推定量とする。そこで、以後の小節でこれらのパラメータのサンプリング方法について詳しく説明しよう。

3.2 パラメータの推定

3.2.1 ω のサンプリング

分散共分散行列 Ω は $n \times n$ の対角行列であり、その対角要素の ω は n 個の要素から構成され、

$\{\omega_1, \omega_2, \dots, \omega_n\}$ と記述される。前述した通り $\omega_i (i=1, 2, \dots, n)$ の事前分布として $IG\left(\frac{v_\lambda}{2}, \frac{v_\lambda}{2}\right)$ という

逆ガンマ分布を共通に仮定すると、 ω_i のフルコンディショナルな事後分布は以下のような分布となる。

$$p(\omega_i | y, \sigma^2, \gamma, v_\lambda) = IG\left(\frac{v_\lambda + 1}{2}, \frac{v_\lambda + \varepsilon_i^2 / \sigma^2}{2}\right) \quad (i=1, 2, \dots, n)$$

サンプリングされた v_λ をパラメータとする、この逆ガンマ分布から ω_i をサンプリングする。

3.2.2 γ のサンプリング

事前分布として $N(\gamma_0, A_0)$ を仮定すると、 γ のフルコンディショナルな事後分布は以下の通りである。この正規分布から γ をサンプリングする。

$$p(\gamma | y, \sigma^2, \Omega) = N\left(\left(\frac{X' \Omega^{-1} X}{\sigma^2} + A_0^{-1}\right)^{-1} \left(\frac{X' \Omega^{-1} Y}{\sigma^2} + A_0^{-1} \gamma_0\right), \left(\frac{X' \Omega^{-1} X}{\sigma^2} + A_0^{-1}\right)^{-1}\right)$$

3.2.3 σ^2 のサンプリング

事前分布として逆ガンマ分布 $IG(v_0/2, v_0 \times s_0^2/2)$ を仮定すると、 σ^2 のフルコンディショナルな事後分布は以下の通りである。

$$p(\sigma^2 | y, \gamma, \Omega) \sim IG\left(\frac{v_0 + n}{2}, \frac{s^2 + v_0 s_0^2}{2}\right)$$

ただし、 $s^2 = (Y - \hat{Y}) \Omega^{-1} (Y - \hat{Y})$ 、 $\hat{Y} = X\gamma$ である。この逆ガンマ分布から σ^2 をサンプリングする。

3.2.4 v_λ のサンプリング⁷

v_λ の事前分布としてガンマ分布 $\text{GAM}(1, \theta_0)$ を仮定する。この場合の v_λ のフルコンディショナルな事後分布は以下のように示せる。

$$p(v_\lambda | \mathbf{y}, \gamma, \sigma^2, \omega) \propto \left(\frac{v_\lambda}{2}\right)^{\frac{nv_\lambda}{2}} \Gamma\left(\frac{v_\lambda}{2}\right)^{-n} \exp(-\eta v_\lambda), \text{ ただし } \eta = \frac{1}{\theta_0} + \frac{1}{2} \sum_{i=1}^n \left(\ln \frac{1}{\lambda_i} + \lambda_i\right) \text{ である。}$$

しかし、この分布からサンプリングするのは困難なので、棄却サンプリング連鎖法を用いて v_λ をサンプリングする。

$$f(v_\lambda) = \left(\frac{v_\lambda}{2}\right)^{\frac{nv_\lambda}{2}} \Gamma\left(\frac{v_\lambda}{2}\right)^{-n} \exp(-\eta v_\lambda)$$

と記述し、テイラー展開を利用すれば、 $f(v_\lambda)$ は以下の様に近似できる。⁸

$$f(v_\lambda) \approx f(v_\lambda^*) \times \exp\left\{-\frac{(v_\lambda - v_\lambda^*)^2}{2\sigma_*^2}\right\} \approx K^* \times h(v_\lambda)$$

$$\text{ただし、} K^* = \sqrt{2\pi\sigma_*^2} \Phi\left(\frac{v_\lambda^*}{\sigma_*}\right) f(v_\lambda^*), \quad h(v_\lambda) = \frac{1}{\sqrt{2\pi\sigma_*^2} \Phi\left(\frac{v_\lambda^*}{\sigma_*}\right)} \exp\left\{-\frac{(v_\lambda - v_\lambda^*)^2}{2\sigma_*^2}\right\} \text{ である。}$$

して、モード v_λ^* は $\ln f(v_\lambda)$ のモードであり、以下に示す一階の条件を満たす。

$$\frac{N}{2} \left(\ln\left(\frac{v_\lambda^*}{2}\right) + 1 - \Psi\left(\frac{v_\lambda^*}{2}\right) \right) - \eta = 0$$

棄却サンプリング連鎖法を用いてサンプリングするに当たっては、第2節の $f(\theta)$ 、 K 、 $h(\theta)$ の代わりに、上述した $f(v_\lambda)$ 、 K^* 、 $h(v_\lambda)$ を用いる。

以下に確率密度関数でない $f(v_\lambda)$ をサンプリングに利用する理由と棄却サンプリング連鎖法を利用する理由を述べよう。

$p(v_\lambda | \mathbf{y}, \gamma, \sigma^2, \omega) f(v_\lambda)$ であるから、 $f(v_\lambda)$ は v_λ のサンプルが従う確率密度関数でない。 v_λ のサンプルが従う確率密度関数は $f(v_\lambda)$ を基準化定数 (d) で割った $f(v_\lambda)/d$ である。 $f(v_\lambda)/d$ からのサンプリングでは、切断正規分布 $N^+(v_\lambda^*, \sigma_*^2)$ である $h(v_\lambda)$ を用いて v_λ をサンプリングし、その後で $f(v_\lambda)/d$ と K^*/d を用いて計算した採択確率 α_{AR} ならびに α_{MH} を利用して取捨選択する。⁹ ここで、 $f(v_\lambda)/d$ と

⁷ 本小節の前半部分の中妻(2003)を参考にしている。

⁸ $f(v_\lambda)$ の近似については砂田(2005)でも詳しく説明している。

⁹ $v > 0$ であるから、正規分布でなく、正值のみを考える切断正規分布を利用する。

K^*/d を用いて採択確率 α_{AR} ならびに α_{MH} を計算してみると、基準化定数(d)は相殺されてしまう。したがって、 $f(v_\lambda)/d$ と K/d の代わりに $f(v_\lambda)$ と K を代用することで問題は生じないのである。

つぎに棄却サンプリング法ではなく、棄却サンプリング連鎖法を利用する理由である。 $f(v_\lambda)$ は K^* と $h(v_\lambda)$ の積によって近似できることは既に説明してある。したがって、

$$f(v_\lambda)/d \approx K^* \times h(v_\lambda)/d$$

でもある。もし、確率密度関数 $f(v_\lambda)/d$ 、定数 K^*/d と $h(v_\lambda)$ の間に、

$$f(v_\lambda)/d = K^* \times h(v_\lambda)/d、\text{つまり} f(v_\lambda) = K^* \times h(v_\lambda)$$

という関係が全ての v_λ に対して成り立っているならば、棄却サンプリング法で v_λ を推定できる。しかし、全ての v_λ に対して成り立っているとは限らない。そこで、棄却サンプリング連鎖法を利用するのである。

4. 棄却サンプリング連鎖法における周辺尤度の計算¹⁰

4.1節ではパラメータがBブロックに分けられるという、一般的なモデルにおける周辺尤度の推定方法を説明する。第5節でt分布を仮定した自己回帰モデルの推定を行うので、それに備えてパラメータが3つで潜在変数が1つのモデルで周辺尤度を推定する方法を4.2節で示す。¹¹

4.1 パラメータがBブロックに分けられるモデルにおける周辺尤度の推定

4.1.1 $\pi(\theta^*_i | y, \theta^*_{-i})$ の導出

D個のパラメータがB個のブロックにグループ分けされるモデルの周辺尤度 $m(y)$ の推定を考えよう。この場合も以下に示す基本周辺尤度方程式を用いて議論を始める。

$$m(y) = \frac{f(y | \theta^*_1, \dots, \theta^*_B) \pi(\theta^*_1, \dots, \theta^*_B)}{\pi(\theta^*_1, \dots, \theta^*_B | y)}$$

分子は事前分布と尤度なので簡単に計算できる。問題は分母の計算である。分母の $\pi(\theta^*_1, \dots, \theta^*_B | y)$ は以下の様に分割できる。

$$\pi(\theta^*_1, \dots, \theta^*_B | y) = \prod_{i=1}^B \pi(\theta^*_i | y, \theta^*_1, \dots, \theta^*_{i-1})$$

したがって、 $\pi(\theta^*_1, \dots, \theta^*_B | y)$ の推定ではその要素である $\pi(\theta^*_i | y, \theta^*_1, \dots, \theta^*_{i-1}), i=1, 2, \dots, B$ を個別に推定することが求められる。上式の右辺の積の要素の中で、フルコンディショナルな事後密度をすぐに利用できる $\pi(\theta^*_B | y, \theta^*_1, \dots, \theta^*_{B-1})$ を除いた $\pi(\theta^*_i | y, \theta^*_1, \dots, \theta^*_{i-1}), i=1, 2, \dots, B-1$ の推定方法を以下で考えてみよう。

説明の都合上、まず、パラメータ θ_i を棄却サンプリング連鎖法で推定する手順を以下で簡単

¹⁰ 本節の説明はChib=Jeliazkov(2005)を参考している。

¹¹ 周辺尤度の計算における潜在変数の扱いはChib(1995)を参考している。

に紹介しておこう。

最初に、 $\psi_{i-1}^* = (\theta_{i-1}^*, \dots, \theta_{i-1}^*), \psi^{i+1} = (\theta_{i+1}, \theta_{i+2}, \dots, \theta_B)$ を用いて、 $h(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1})$ と記述される提案分布から θ_i をサンプリングする。その値を以下の採択確率で取捨選択する。

$$\alpha_{AR}(\theta_i, \theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) = \alpha_{AR}(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) = \min \left\{ 1, \frac{f(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1})}{K \times h(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1})} \right\},$$

その結果、AR部分では以下の確率分布に従う θ_i のサンプルを生成できる。

$$q(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) = \frac{\min \left\{ f(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}), K \times h(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \right\}}{d(\theta_i)}$$

$$= \frac{\alpha_{AR}(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \times K \times h(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1})}{d(\theta_i)}$$

次に、AR部分で生成し、採択された θ_i のサンプルを、さらに採択確率 $\alpha_{MH}(\theta_i, \theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1})$ で、もう一度、取捨選択する。こうして得た θ_i のサンプルの平均が推定量である。

この小節の目的である、 $\pi(\theta_i^* | \mathbf{y}, \theta_{i-1}^*, \dots, \theta_{i-1}^*), i=1, 2, \dots, B-1$ の推定方法について、以下で説明する。 θ_i に対する部分核はM=H部分の採択確率と提案密度を用いて、

$$p(\theta_i, \theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) = \alpha_{MH}(\theta_i, \theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \times q(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \quad \dots(4)$$

と記述できる。この部分核は以下の局所反転条件

$$p(\theta_i, \theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) = p(\theta_i^*, \theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \quad \dots(5)$$

を満たす。(5)に $\pi(\psi^{i+1} | \mathbf{y}, \psi_{i-1}^*)$ をかけてから ψ^i で積分すると、

$$\int p(\theta_i, \theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\psi^{i+1} | \mathbf{y}, \psi_{i-1}^*) d\psi^i$$

$$= \int p(\theta_i^*, \theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\psi^{i+1} | \mathbf{y}, \psi_{i-1}^*) d\psi^i$$

である。さらに、式変形すると、

$$\int p(\theta_i, \theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\theta_i, \psi^{i+1} | \mathbf{y}, \psi_{i-1}^*) d\psi^i$$

$$= \int p(\theta_i^*, \theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\theta_i^*, \psi_{i-1}^*) \pi(\psi^{i+1} | \mathbf{y}, \theta_i^*, \psi_{i-1}^*) d\psi^i$$

次に、 $\pi(\theta_i^* | \mathbf{y}, \psi_{i-1}^*)$ について解いた後で、(4)を代入すると、以下の通りである。

$$\pi(\theta_i^* | \mathbf{y}, \psi_{i-1}^*) = \frac{\int \alpha_{MH}(\theta_i, \theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) q(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\psi^i | \mathbf{y}, \psi_{i-1}^*) d\psi^i}{\int \alpha_{MH}(\theta_i^*, \theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) q(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\psi^{i+1} | \mathbf{y}, \psi_{i-1}^*) d\psi^i}$$

さらに、 $q(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1})$ を代入すると、基準化定数 $d(\theta_i)$ は相殺されて、次式が導出される。

$$\pi(\theta_i^* | \mathbf{y}, \psi_{i-1}^*) = \frac{\int \alpha_{MH}(\theta_i, \theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \alpha_{AR}(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) h(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\psi^i | \mathbf{y}, \psi_{i-1}^*) d\psi^i}{\int \alpha_{MH}(\theta_i^*, \theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \alpha_{AR}(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) h(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\psi^{i+1} | \mathbf{y}, \psi_i^*) d\psi^i} \dots (6)$$

(6)の推定では、モンテカルロ積分を分子と分母に別々に適用する。分子については、 $\pi(\psi^i | \mathbf{y}, \psi_{i-1}^*)$ に従う ψ^i を用いて $\alpha_{MH}(\theta_i, \theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \alpha_{AR}(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) h(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1})$ をモンテカルロ積分して、

$$\int \alpha_{MH}(\theta_i, \theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \alpha_{AR}(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) h(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\theta^i | \mathbf{y}, \psi_{i-1}^*) d\psi^i$$

を推定する。分母については、 $\pi(\psi^{i+1} | \mathbf{y}, \psi_i^*) h(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1})$ に従う (ψ^{i+1}, θ_i) を用いて $\alpha_{MH}(\theta_i^*, \theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \alpha_{AR}(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1})$ をモンテカルロ積分して、

$$\int \alpha_{MH}(\theta_i^*, \theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \alpha_{AR}(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\psi^{i+1} | \mathbf{y}, \psi_i^*) h(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) d\psi^i$$

を推定する。以下の小節で、推定の手順を詳細に説明しよう。

4.1.2 $\pi(\theta_i^* | \mathbf{y}, \psi_{i-1}^*)$ の推定

まず、(6)の分子を推定してみよう。分子では $\pi(\psi^i | \mathbf{y}, \psi_{i-1}^*)$ に従う ψ^i を用いて積分するので、モンテカルロ積分では、 $\pi(\psi^i | \mathbf{y}, \psi_{i-1}^*)$ に従う ψ^i のサンプルを利用する。

$\pi(\psi^i | \mathbf{y}, \psi_{i-1}^*)$ からのサンプルは、 $\psi_{i-1} = \psi_{i-1}^*$ とおいた $\{\theta_i, \dots, \theta_B\}$ のフルコンディショナルな事後密度 $\pi(\theta_k | \mathbf{y}, \psi_{i-1}^*, \theta_i, \dots, \theta_{k-1}, \theta_{k+1}, \dots, \theta_B), k=i, i+1, \dots, B$ に従う θ_k を棄却サンプリング連鎖法により生成することで得られる。g回目に得られたサンプルを $\psi^{i,(g)} = (\theta_i^{(g)}, \theta_{i+1}^{(g)}, \dots, \theta_B^{(g)})$ と記述すれば、G回のサン

プリングで得られたG個のサンプルは $\{\psi^{i,(g)}\}_{g=1}^G$ と記述できる。このG個のサンプルを

$\alpha_{MH}(\theta_i, \theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \alpha_{AR}(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) h(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1})$ の ψ^i に代入して平均した値が(6)の分子

$$\int \alpha_{MH}(\theta_i^*, \theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \alpha_{AR}(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\psi^{i+1} | \mathbf{y}, \psi_i^*) h(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) d\psi^i$$

の推定量である。

次に、(6)の分母を推定してみよう。分母では $\pi(\psi^{i+1} | \mathbf{y}, \psi_i^*) h(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1})$ に従う (ψ^{i+1}, θ_i) を用いて積分するので、モンテカルロ積分では、 $\pi(\psi^{i+1} | \mathbf{y}, \psi_i^*) h(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1})$ に従う (ψ^{i+1}, θ_i) のサンプルを利用する。

ψ^{i+1} のサンプルは、 $\psi_i = \psi_i^*$ とおいた $\{\theta_i, \theta_{i+1}, \dots, \theta_B\}$ のフルコンディショナルな事後密度 $\pi(\theta_k | \mathbf{y}, \psi_{i-1}^*, \theta_i, \dots, \theta_{k-1}, \theta_{k+1}, \dots, \theta_B), k=i+1, \dots, B$ に従う θ_k を棄却サンプリング連鎖法により生成することで得られる。 θ_i は $\theta_{i+1}, \theta_{i+2}, \dots, \theta_B$ のサンプリング結果を用いた $h(\theta_i | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1})$ からのサンプリングで得られる。これが1回のサンプリングであり、この作業を何度も繰り返して $\psi^i = (\theta_i, \theta_{i+1}, \dots, \theta_B)$ のサンプルを得る。j回目のサンプリングで得られた ψ^i のサンプルを $\psi^{i,(j)} = (\theta_i^{(j)}, \theta_{i+1}^{(j)}, \dots, \theta_B^{(j)})$ と記述し、J回

棄却サンプリング連鎖を用いた場合の周辺尤度の推定—誤差項に t 分布を仮定した自己回帰モデルを例に—砂田

のサンプリングで得られた J 個のサンプルを $\{\psi^{i,(j)}\}_{j=1}^J$ と記述する。この J 個のサンプルを

$\alpha_{MH}(\theta_i^*, \theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \alpha_{AR}(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1})$ の θ_i^* と ψ^{i+1} に代入し平均した値が(6)の分母、

$$\int \alpha_{MH}(\theta_i^*, \theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \alpha_{AR}(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) \pi(\psi^{i+1} | \mathbf{y}, \psi_i^*) h(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1}) d\psi^i$$

の推定量である。

上述したように $\pi(\theta_i^* | \mathbf{y}, \psi_{i-1}^*)$ の推定は、縮尺されたMCMCサンプリングによって得られたサンプルを利用したモンテカルロ積分によって得られる。 $\pi(\theta_i^* | \mathbf{y}, \psi_{i-1}^*)$ の推定量を数式で表せば次式の通りである。

$$\hat{\pi}(\theta_i^* | \mathbf{y}, \psi_{i-1}^*) = \frac{G^{-1} \sum_{g=1}^G \alpha_{MH}(\theta_i^{(g)}, \theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1,(g)}) \alpha_{AR}(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1,(g)}) h(\theta_i^* | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1,(g)})}{J^{-1} \sum_{j=1}^J \alpha_{MH}(\theta_i^*, \theta_i^{(j)} | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1,(j)}) \alpha_{AR}(\theta_i^{(j)} | \mathbf{y}, \psi_{i-1}^*, \psi^{i+1,(j)})} \quad \dots(7)$$

同じ作業を繰り返して、 $\pi(\theta_1^* | \mathbf{y}), \pi(\theta_2^* | \mathbf{y}, \theta_1^*), \dots, \pi(\theta_{B-1}^* | \mathbf{y}, \theta_1^*, \theta_2^*, \dots, \theta_{B-2}^*)$ の推定量も計算する。それを用いて、 $\pi(\theta_1^*, \dots, \theta_B^* | \mathbf{y}) = \prod_{i=1}^B \pi(\theta_i^* | \mathbf{y}, \theta_1^*, \dots, \theta_{i-1}^*)$ の推定量を計算する。

対数周辺尤度の推定量は次式の様になるので、上述した $\pi(\theta_1^*, \dots, \theta_B^* | \mathbf{y})$ の推定量を利用して計算する。

$$\ln \hat{m}(\mathbf{y}) = \ln f(\mathbf{y} | \theta^*) + \ln \pi(\theta^*) - \sum_{i=1}^B \ln \hat{\pi}(\theta_i^* | \mathbf{y}, \theta_1^*, \dots, \theta_{i-1}^*) \quad \dots(8)$$

4.2 パラメータが3つのモデルにおける周辺尤度の推定

1つのパラメータ (θ_1) だけは棄却サンプリング連鎖法を利用しないとサンプリングできないが、それ以外の2つのパラメータ (θ_2, θ_3) と1つの潜在変数 (z) はフルコンディショナルな事後分布からギブズ・サンプリングできるケースを考えてみる。

まず、 θ_1 の棄却サンプリング連鎖法によるサンプリングであるが、提案分布 $h(\theta_1 | \mathbf{y}, \theta_2, \theta_3, z)$ から θ_1 をサンプリングし、以下の採択確率に基づいて θ_1' の採否を決める。

$$\alpha_{AR}(\theta_1, \theta_1' | \mathbf{y}, \theta_2, \theta_3, z) = \alpha_{AR}(\theta_1' | \mathbf{y}, \theta_2, \theta_3, z) = \min \left\{ 1, \frac{f(\theta_1' | \mathbf{y}, \theta_2, \theta_3, z)}{K \times h(\theta_1' | \mathbf{y}, \theta_2, \theta_3, z)} \right\}$$

次のM=H部分で、新たな採択確率 $\alpha_{MH}(\theta_1, \theta_1' | \mathbf{y}, \theta_2, \theta_3, z)$ を用いて θ_1' の採否をさらに決める。採択確率は θ_1' と θ_1 の満たす条件によって以下の様に変化する。

- a. $f(\theta_1' | \mathbf{y}, \theta_2, \theta_3, z) > K \times h(\theta_1' | \mathbf{y}, \theta_2, \theta_3, z)$ ならば

$$\alpha_{MH}(\theta_1, \theta_1^* | \mathbf{y}, \theta_2, \theta_3, z) = \min \left\{ 1, \frac{K \times h(\theta_1 | \mathbf{y}, \theta_2, \theta_3, z)}{f(\theta_1 | \mathbf{y}, \theta_2, \theta_3, z)} \right\}$$

b. $f(\theta_1^* | \mathbf{y}, \theta_2, \theta_3, z) > K \times h(\theta_1^* | \mathbf{y}, \theta_2, \theta_3, z)$ ならば、

$$\alpha_{MH}(\theta_1, \theta_1^* | \mathbf{y}, \theta_2, \theta_3, z) = \min \left\{ 1, \frac{f(\theta_1^* | \mathbf{y}, \theta_2, \theta_3, z) h(\theta_1 | \mathbf{y}, \theta_2, \theta_3, z)}{f(\theta_1 | \mathbf{y}, \theta_2, \theta_3, z) h(\theta_1^* | \mathbf{y}, \theta_2, \theta_3, z)} \right\}$$

次に、 $\pi(\theta_1^* | \mathbf{y})$ の推定方法を記述しよう。 $\pi(\theta_1^* | \mathbf{y})$ はAR部分の採択確率、M=H部分の採択確率、AR部分の提案密度関数などを用いて、(9)の様に示される。多重積分が行えるならば簡単に推定できるが、現実には計算ができないので、モンテカルロ積分を利用する。

$$\pi(\theta_1^* | \mathbf{y}) = \frac{\int \alpha_{MH}(\theta_1, \theta_1^* | \mathbf{y}, \theta_2, \theta_3, z) \alpha_{AR}(\theta_1^* | \mathbf{y}, \theta_2, \theta_3, z) h(\theta_1^* | \mathbf{y}, \theta_2, \theta_3, z) \pi(\theta_1, \theta_2, \theta_3, z | \mathbf{y}) d\theta_1 d\theta_2 d\theta_3 dz}{\int \alpha_{MH}(\theta_1, \theta_1^* | \mathbf{y}, \theta_2, \theta_3, z) \alpha_{AR}(\theta_1 | \mathbf{y}, \theta_2, \theta_3, z) h(\theta_1 | \mathbf{y}, \theta_2, \theta_3, z) \pi(\theta_1, \theta_2, \theta_3, z | \mathbf{y}, \theta_1^*) d\theta_1 d\theta_2 d\theta_3 dz} \dots (9)$$

まず、(9)の分子を推定しよう。 θ_1 を棄却サンプリング連鎖法でサンプリングする。その結果を利用して、 θ_2, θ_3, z をフルコンディショナルな事後密度 $\pi(\theta_2 | \mathbf{y}, \theta_1, \theta_3, z)$ 、 $\pi(\theta_3 | \mathbf{y}, \theta_1, \theta_2, z)$ 、 $\pi(z | \mathbf{y}, \theta_1, \theta_2, \theta_3)$ からサンプリングする。この作業をG回行う。g回目のサンプリングの結果を $(\theta_1^{(g)}, \theta_2^{(g)}, \theta_3^{(g)}, z^{(g)})$

と記述すれば、得られるデータは $\{\theta_1^{(g)}, \theta_2^{(g)}, \theta_3^{(g)}, z^{(g)}\}_{g=1}^G$ である。このサンプルを

$\alpha_{MH}(\theta_1, \theta_1^* | \mathbf{y}, \theta_2, \theta_3, z) \alpha_{AR}(\theta_1, \theta_1^* | \mathbf{y}, \theta_2, \theta_3, z) h(\theta_1^* | \mathbf{y}, \theta_2, \theta_3, z)$ に代入した値が分子の推定量である。

次に、(9)の分母を推定しよう。 θ_2, θ_3, z のフルコンディショナルな事後分布 $\pi(\theta_2 | \mathbf{y}, \theta_1, \theta_3, z)$ 、 $\pi(\theta_3 | \mathbf{y}, \theta_1, \theta_2, z)$ 、 $\pi(z | \mathbf{y}, \theta_1, \theta_2, \theta_3)$ の θ_1 に θ_1^* を代入して θ_2, θ_3, z をサンプリングする。つまり、 $\theta_2, \theta_3, \theta_4$ をそれぞれ $\pi(\theta_2 | \mathbf{y}, \theta_3, z, \theta_1^*)$ 、 $\pi(\theta_3 | \mathbf{y}, \theta_2, z, \theta_1^*)$ 、 $\pi(z | \mathbf{y}, \theta_2, \theta_3, \theta_1^*)$ からサンプリングする。その後で θ_1 を $h(\theta_1 | \mathbf{y}, \theta_2, \theta_3, z)$ からサンプリングする。これが1回のサンプリングであり、この作業をJ回行う。j回目のサンプリングの結果を $(\theta_1^{(j)}, \theta_2^{(j)}, \theta_3^{(j)}, z^{(j)})$ と記述すれば、得られるサンプルは

$\{\theta_1^{(j)}, \theta_2^{(j)}, \theta_3^{(j)}, z^{(j)}\}_{j=1}^J$ である。このサンプルを $\alpha_{MH}(\theta_1, \theta_1^* | \mathbf{y}, \theta_2, \theta_3, z) \alpha_{AR}(\theta_1^* | \mathbf{y}, \theta_2, \theta_3, z)$ に代入した値が分母の推定量である。

$\pi(\theta_1^* | \mathbf{y})$ の推定には、得られたサンプル $\{\theta_1^{(g)}, \theta_2^{(g)}, \theta_3^{(g)}, z^{(g)}\}_{g=1}^G$ と $\{\theta_1^{(j)}, \theta_2^{(j)}, \theta_3^{(j)}, z^{(j)}\}_{j=1}^J$ を

利用してモンテカルロ積分を行う。 $\pi(\theta_1^* | \mathbf{y})$ の推定量 $\hat{\pi}(\theta_1^* | \mathbf{y})$ は次式の通りである。

$$\hat{\pi}(\theta_1^* | \mathbf{y}) = \frac{\frac{1}{G} \sum_{g=1}^G \alpha_{MH}(\theta_1^{(g)}, \theta_1^* | \mathbf{y}, \theta_2^{(g)}, \theta_3^{(g)}, z^{(g)}) \alpha_{AR}(\theta_1^* | \mathbf{y}, \theta_2^{(g)}, \theta_3^{(g)}, z^{(g)}) h(\theta_1^* | \mathbf{y}, \theta_2^{(g)}, \theta_3^{(g)}, z^{(g)})}{\frac{1}{J} \sum_{j=1}^J \alpha_{MH}(\theta_1^*, \theta_1^{(j)} | \mathbf{y}, \theta_2^{(j)}, \theta_3^{(j)}, z^{(j)}) \alpha_{AR}(\theta_1^{(j)} | \mathbf{y}, \theta_2^{(j)}, \theta_3^{(j)}, z^{(j)})} \dots (10)$$

θ_2, θ_3, z はフルコンディショナルな事後分布を導出できると仮定しているので、縮尺されたギブズ・サンプリングの手法を用いながら、 $\pi(\theta_3^* | \mathbf{y}, \theta_1^*, \theta_2^*)$ 、 $\pi(\theta_2^* | \mathbf{y}, \theta_1^*)$ の推定量を計算する。

$\pi(\theta_2^* | \mathbf{y}, \theta_1^*)$ の推定は、 $p(\theta_3 | \mathbf{y}, z, \theta_2, \theta_1^*)$ 、 $p(z | \mathbf{y}, \theta_2, \theta_3, \theta_1^*)$ 、 $p(\theta_2 | \mathbf{y}, z, \theta_3, \theta_1^*)$ を用いて $\{z, \theta_2, \theta_3\}$ をサンプリングする。得られた $\{\theta_2^{(g)}, \theta_3^{(g)}, z^{(g)}\}_{g=1}^G$ の中の $\{\theta_3^{(g)}, z^{(g)}\}_{g=1}^G$ を次式に代入すれば、 $\pi(\theta_2^* | \mathbf{y}, \theta_1^*)$ を推定できる。

$$\hat{\pi}(\theta_2^* | \mathbf{y}, \theta_1^*) = \frac{1}{G} \sum_{g=1}^G \pi(\theta_2^* | \mathbf{y}, \theta_1^*, \theta_3^{(g)}, z^{(g)})$$

$\pi(\theta_3^* | \mathbf{y}, \theta_1^*, \theta_2^*)$ の推定は $\pi(\theta_3^* | \mathbf{y}, \theta_1^*, \theta_2^*)$ は $\pi(z | \mathbf{y}, \theta_1^*, \theta_2^*, \theta_3)$ 、 $\pi(\theta_3 | \mathbf{y}, \theta_1^*, \theta_2^*, z)$ を用いて θ_3 と z を交互にサンプリングして、 (θ_3, z) のサンプルを得る。得られた $\{\theta_3^{(g)}, z^{(g)}\}_{g=1}^G$ の中の $\{z^{(g)}\}_{g=1}^G$ を次式に代入すれば、 $\pi(\theta_3^* | \mathbf{y}, \theta_1^*, \theta_2^*)$ を推定できる。

$$\hat{\pi}(\theta_3^* | \mathbf{y}, \theta_2^*, \theta_1^*) = \frac{1}{G} \sum_{g=1}^G \pi(\theta_3^* | \mathbf{y}, \theta_1^*, \theta_2^*, z^{(g)})$$

$\pi(\theta_1^* | \mathbf{y})$ 、 $\pi(\theta_2^* | \mathbf{y}, \theta_1^*)$ 、 $\pi(\theta_3^* | \mathbf{y}, \theta_1^*, \theta_2^*)$ の推定量を上述した方法に従って計算した後、次式に代入して対数周辺尤度 $\ln m(\mathbf{y})$ の推定を行う。

$$\ln \hat{m}(\mathbf{y}) = \ln f(\mathbf{y} | \theta_1^*, \theta_2^*, \theta_3^*) + \ln \pi(\theta_1^*, \theta_2^*, \theta_3^*) - \ln \hat{\pi}(\theta_1^* | \mathbf{y}) - \ln \hat{\pi}(\theta_2^* | \mathbf{y}, \theta_1^*) - \ln \hat{\pi}(\theta_3^* | \mathbf{y}, \theta_1^*, \theta_2^*) \dots (11)$$

5. 誤差項が t 分布に従う自己回帰モデルの周辺尤度の推定

5.1節では誤差項が t 分布に従う自己回帰モデルの周辺尤度の推定を行う。5.2節ではデータをシミュレーションによって生成した後、これまでの節で紹介してきた方法を利用して、パラメータと周辺尤度を推定した結果を示す。

5.1 周辺尤度の推定

$v_i, \sigma^2, \gamma, \omega$ を順番に $\theta_1, \theta_2, \theta_3, z$ と名づけて、4.2節の推定方法を適用すればよい。つまり、 $\pi(v_i^* | \mathbf{y})$ は棄却サンプリング連鎖法を用いた場合の推定方法、 $\pi(\gamma^* | \mathbf{y}, v_i^*, \sigma^2)$ 、 $\pi(\sigma^2 | \mathbf{y}, v_i^*)$ はギブズ・サン

プリングを用いた場合の推定方法を利用して、周辺尤度 $m(\mathbf{y})$ を推定すれば良い。¹²

5.1.1 $\pi(\sigma^2 | \mathbf{y}, v_\lambda^*)$ の計算

$\pi(\sigma^2 | \mathbf{y}, v_\lambda^*)$ の計算には $(\sigma^2, \boldsymbol{\gamma}, \boldsymbol{\omega})$ のサンプルが必要である。そこで、

$$p(\sigma^2 | \mathbf{y}, \boldsymbol{\gamma}, \boldsymbol{\omega}, v_\lambda^*) = IG\left(\frac{v_0 + n}{2}, \frac{s^2 + v_0 s_0^2}{2}\right),$$

$$p(\omega_i | \mathbf{y}, \boldsymbol{\gamma}, \sigma^2, v_\lambda^*) = IG\left(\frac{v_\lambda^* + 1}{2}, \frac{v_\lambda^* + \varepsilon_i^2 / \sigma^2}{2}\right),$$

$$p(\boldsymbol{\gamma} | \mathbf{y}, \sigma^2, \boldsymbol{\omega}, v_\lambda^*) = N\left(\left(\frac{X' \boldsymbol{\Omega}^{-1} X}{\sigma^2} + A_0^{-1}\right)^{-1} \left(\frac{X' \boldsymbol{\Omega}^{-1} Y}{\sigma^2} + A_0^{-1} \boldsymbol{\gamma}_0\right), \left(\frac{X' \boldsymbol{\Omega}^{-1} X}{\sigma^2} + A_0^{-1}\right)^{-1}\right)$$

を用いて $(\sigma^2, \boldsymbol{\gamma}, \boldsymbol{\omega})$ のサンプル $\{\sigma^{2(g)}, \boldsymbol{\gamma}^{(g)}, \boldsymbol{\omega}^{(g)}\}_{g=1}^G$ を生成する。その中の $(\boldsymbol{\gamma}, \boldsymbol{\omega})$ のサンプル

$\{\boldsymbol{\gamma}^{(g)}, \boldsymbol{\omega}^{(g)}\}_{g=1}^G$ と v_λ^*, σ^{2*} を密度関数 $p(\sigma^2 | \mathbf{y}, \boldsymbol{\gamma}, \boldsymbol{\omega}, v_\lambda^*)$ に代入、及び平均して $\pi(\sigma^2 | \mathbf{y}, v_\lambda^*)$ を推定する。

5.1.2 $\pi(\boldsymbol{\gamma}^* | \mathbf{y}, v_\lambda^*, \sigma^{2*})$ の計算

$\pi(\boldsymbol{\gamma}^* | \mathbf{y}, v_\lambda^*, \sigma^{2*})$ の計算には $(\boldsymbol{\gamma}, \boldsymbol{\omega})$ のサンプルが必要である。まず

$$p(\boldsymbol{\gamma} | \mathbf{y}, \sigma^{2*}, \boldsymbol{\omega}, v_\lambda^*) = N\left(\left(\frac{X' \boldsymbol{\Omega}^{-1} X}{\sigma^{2*}} + A_0^{-1}\right)^{-1} \left(\frac{X' \boldsymbol{\Omega}^{-1} Y}{\sigma^{2*}} + A_0^{-1} \boldsymbol{\gamma}_0\right), \left(\frac{X' \boldsymbol{\Omega}^{-1} X}{\sigma^{2*}} + A_0^{-1}\right)^{-1}\right),$$

$$p(\omega_i | \mathbf{y}, \boldsymbol{\gamma}, \sigma^{2*}, v_\lambda^*) = IG\left(\frac{v_\lambda^* + 1}{2}, \frac{v_\lambda^* + \varepsilon_i^2 / \sigma^{2*}}{2}\right)$$

を用いて $(\boldsymbol{\gamma}, \boldsymbol{\omega})$ のサンプル $\{\boldsymbol{\gamma}^{(g)}, \boldsymbol{\omega}^{(g)}\}_{g=1}^G$ を生成する。その中の $\boldsymbol{\omega}$ のサンプル $\{\boldsymbol{\omega}^{(g)}\}_{g=1}^G$ と

$v_\lambda^*, \sigma^{2*}, \boldsymbol{\gamma}^*$ を密度関数 $p(\boldsymbol{\gamma} | \mathbf{y}, \sigma^{2*}, \boldsymbol{\omega}, v_\lambda^*)$ に代入、及び平均して $\pi(\boldsymbol{\gamma}^* | \mathbf{y}, v_\lambda^*, \sigma^{2*})$ を推定する。

5.1.3 $\pi(v_\lambda^* | \mathbf{y})$ の計算

$\pi(v_\lambda^* | \mathbf{y})$ の計算を4.2節の(10)と対比させながら行う。(10)は以下に示す(12)に対応する。

¹² ギブズ・サンプリング法でパラメータを推定した場合の周辺尤度の推定方法についてはChib(1995)や砂田(2006)を参照されたい。

棄却サンプリング連鎖を用いた場合の周辺尤度の推定—誤差項に t 分布を仮定した自己回帰モデルを例に—砂田

$$\hat{\pi}(v_\lambda^* | \mathbf{y}) = \frac{\frac{1}{G} \sum_{g=1}^G \alpha_{MH}(v_\lambda^{(g)}, v_\lambda^* | \mathbf{y}, \sigma^{2(g)}, \boldsymbol{\gamma}^{(g)}, \boldsymbol{\omega}^{(g)}) \alpha_{AR}(v_\lambda^* | \mathbf{y}, \sigma^{2(g)}, \boldsymbol{\gamma}^{(g)}, \boldsymbol{\omega}^{(g)}) h(v_\lambda^* | \mathbf{y}, \sigma^{2(g)}, \boldsymbol{\gamma}^{(g)}, \boldsymbol{\omega}^{(g)})}{\frac{1}{J} \sum_{j=1}^J \alpha_{MH}(v_\lambda^*, v_\lambda^{(j)} | \mathbf{y}, \sigma^{2(j)}, \boldsymbol{\gamma}^{(j)}, \boldsymbol{\omega}^{(j)}) \alpha_{AR}(v_\lambda^{(j)} | \mathbf{y}, \sigma^{2(j)}, \boldsymbol{\gamma}^{(j)}, \boldsymbol{\omega}^{(j)})} \quad \dots(12)$$

まず、(12)の分子の計算を考えてみよう。分子の計算には $(v_\lambda, \sigma^2, \boldsymbol{\gamma}, \boldsymbol{\omega})$ のサンプルが必要である。そこで以下の4つの確率分布を用いてサンプルを生成する。

$$p(v_\lambda | \mathbf{y}, \boldsymbol{\gamma}, \sigma^2, \boldsymbol{\omega}) \propto \left(\frac{v_\lambda}{2}\right)^{\frac{nv_\lambda}{2}} \Gamma\left(\frac{v_\lambda}{2}\right)^{-n} \exp(-\eta v_\lambda), \text{ ただし } \eta = \frac{1}{\theta_0} + \frac{1}{2} \sum_{i=1}^n \left(\ln \frac{1}{\lambda_i} + \lambda_i\right) \text{ である。}$$

$$p(\boldsymbol{\gamma} | \mathbf{y}, \sigma^2, \boldsymbol{\omega}, v_\lambda) = N\left(\left(\frac{X' \boldsymbol{\Omega}^{-1} X}{\sigma^2} + A_0^{-1}\right)^{-1} \left(\frac{X' \boldsymbol{\Omega}^{-1} Y}{\sigma^2} + A_0^{-1} \boldsymbol{\gamma}_0\right), \left(\frac{X' \boldsymbol{\Omega}^{-1} X}{\sigma^2} + A_0^{-1}\right)^{-1}\right),$$

$$p(\sigma^2 | \mathbf{y}, \boldsymbol{\gamma}, \boldsymbol{\omega}, v_\lambda) = IG\left(\frac{v_0 + n}{2}, \frac{s^2 + v_0 s_0^2}{2}\right),$$

$$p(\boldsymbol{\omega}_i | \mathbf{y}, \boldsymbol{\gamma}, \sigma^2, v_\lambda) = IG\left(\frac{v_\lambda + 1}{2}, \frac{v_\lambda + \boldsymbol{\varepsilon}_i^2 / \sigma^2}{2}\right)$$

これらを用いて生成した $\{v_\lambda^{(g)}, \boldsymbol{\omega}^{(g)}, \boldsymbol{\gamma}^{(g)}, \sigma^{2(g)}\}_{g=1}^G$ と v_λ^* を(12)の分子に代入する。

次に(12)の分母の計算を考えてみよう。分母の計算には $(v_\lambda, \sigma^2, \boldsymbol{\gamma}, \boldsymbol{\omega})$ のサンプルが必要である。そこで以下の3つの確率分布を用いて $(\sigma^2, \boldsymbol{\gamma}, \boldsymbol{\omega})$ のサンプルを生成する。

$$p(\boldsymbol{\gamma} | \mathbf{y}, \sigma^2, \boldsymbol{\omega}, v_\lambda^*) = N\left(\left(\frac{X' \boldsymbol{\Omega}^{-1} X}{\sigma^2} + A_0^{-1}\right)^{-1} \left(\frac{X' \boldsymbol{\Omega}^{-1} Y}{\sigma^2} + A_0^{-1} \boldsymbol{\gamma}_0\right), \left(\frac{X' \boldsymbol{\Omega}^{-1} X}{\sigma^2} + A_0^{-1}\right)^{-1}\right),$$

$$p(\sigma^2 | \mathbf{y}, \boldsymbol{\gamma}, \boldsymbol{\omega}, v_\lambda^*) = IG\left(\frac{v_0 + n}{2}, \frac{s^2 + v_0 s_0^2}{2}\right),$$

$$p(\boldsymbol{\omega}_i | \mathbf{y}, \boldsymbol{\gamma}, \sigma^2, v_\lambda^*) = IG\left(\frac{v_\lambda^* + 1}{2}, \frac{v_\lambda^* + \boldsymbol{\varepsilon}_i^2 / \sigma^2}{2}\right)$$

である。そのサンプルを $h(v_\lambda)$ 、つまり切断正規分布 $N^+(v_\lambda^*, \sigma^{*2})$ の条件部分に代入して v_λ のサン

プルを生成する。こうして生成した $\{v_\lambda^{(j)}, \omega^{(j)}, \gamma^{(j)}, \sigma^{2(j)}\}_{j=1}^J$ を(12)の分母に代入する。

こうして(12)の分子と分母にサンプルを代入して $\pi(v_\lambda^* | \mathbf{y})$ を推定する。ちなみに、(12)の $\alpha_{AR}(v_\lambda, v_\lambda^* | \mathbf{y}, \sigma^2, \gamma, \omega)$ は

$$\alpha_{AR}(v_\lambda, v_\lambda^* | \mathbf{y}, \sigma^2, \gamma, \omega) = \min \left\{ 1, \frac{f(v_\lambda^* | \mathbf{y}, \sigma^2, \gamma, \omega)}{K^* \times h(v_\lambda^* | \mathbf{y}, \sigma^2, \gamma, \omega)} \right\}$$

である。ただし、(12)の $\alpha_{MH}(v_\lambda, v_\lambda^* | \mathbf{y}, \sigma^2, \gamma, \omega)$ は $\pi(v_\lambda) \times f(\mathbf{y} | v_\lambda, \sigma^2, \gamma, \omega)$ と $K^* \times h(v_\lambda | \mathbf{y}, \sigma^2, \gamma, \omega)$ の大小関係で以下のように変化する。

a. $f(v_\lambda^* | \mathbf{y}, \sigma^2, \gamma, \omega) < K^* \times h(v_\lambda^* | \mathbf{y}, \sigma^2, \gamma, \omega)$ ならば、

$$\alpha_{MH}(v_\lambda, v_\lambda^* | \mathbf{y}, \sigma^2, \gamma, \omega) = \min \left\{ 1, \frac{K^* \times h(v_\lambda^* | \mathbf{y}, \sigma^2, \gamma, \omega)}{f(v_\lambda^* | \mathbf{y}, \sigma^2, \gamma, \omega)} \right\}$$

b. $f(v_\lambda^* | \mathbf{y}, \sigma^2, \gamma, \omega) > K^* \times h(v_\lambda^* | \mathbf{y}, \sigma^2, \gamma, \omega)$ ならば、

$$\alpha_{MH}(v_\lambda, v_\lambda^* | \mathbf{y}, \sigma^2, \gamma, \omega) = \min \left\{ 1, \frac{f(v_\lambda^* | \mathbf{y}, \sigma^2, \gamma, \omega) \times h(v_\lambda | \mathbf{y}, \sigma^2, \gamma, \omega)}{f(v_\lambda | \mathbf{y}, \sigma^2, \gamma, \omega) \times h(v_\lambda^* | \mathbf{y}, \sigma^2, \gamma, \omega)} \right\}$$

こうして $\hat{\pi}(v_\lambda^* | \mathbf{y}), \hat{\pi}(\sigma^{2*} | \mathbf{y}, v_\lambda^*), \hat{\pi}(\gamma^* | \mathbf{y}, v_\lambda^*, \sigma^{2*})$ を計算して、

$$\ln \hat{m}(\mathbf{y}) = \ln f(\mathbf{y} | v_\lambda^*, \sigma^{2*}, \gamma^*) + \ln \pi(v_\lambda^*, \sigma^{2*}, \gamma^*) - \ln \hat{\pi}(v_\lambda^* | \mathbf{y}) - \ln \hat{\pi}(\sigma^{2*} | \mathbf{y}, v_\lambda^*) - \ln \hat{\pi}(\gamma^* | \mathbf{y}, v_\lambda^*, \sigma^{2*})$$

に代入すれば、対数周辺尤度は推定できる。ただし、右辺第1項の対数尤度は t 分布の密度関数で構成されている。

5.2 シミュレーション・データを用いた推定

パラメータを与えた上で生成したデータに対して、上述した推定方法を適用してパラメータを推定してみる。その後で、これまでに説明してきた推定方法に基づいて周辺尤度を推定してみる。

モデルとしては自己回帰モデルのラグ次数が1という単純なモデルを利用する。ただし、誤差項は t 分布に従う。パラメータは $\gamma_0=3.0, \gamma_1=0.5, \sigma^2=5, v_\lambda=6$ においてデータを99個生成した。その平均は5.109、標準偏差は2.983であった。このデータに対して、誤差項が t 分布に従う一階の自己回帰モデル (t -AR(1)) を当てはめて、パラメータをベイズ推定した結果を表1に示す。各パラメータの事前

分布であるが、 γ_0 と γ_1 は $N\left(\begin{pmatrix} 3.0 \\ 0.5 \end{pmatrix}, \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix}\right)$ 、 σ^2 の事前分布は $IG\left(\frac{6}{2}, \frac{6 \times 4}{2}\right)$ 、 v_λ の事前

分布は $GAM(1,6)$ とそれぞれ設定した。表1から、推定値が真の値と近い値となっていることが分かる。推定に当たっては、6千回サンプリングして、最初の2千個を捨てて、残りの4千個を利用している。参考のために、誤差項が正規分布に従う一階の自己回帰モデル(N-AR(1))を最小二乗推定した結果、つまり $\gamma_0, \gamma_1, \sigma^2$ の最小二乗推定値を示しておく。

表 1 : t -AR(1)の推定結果

パラメータ	OLSE	真の値	事後平均	事後標準偏差	事後自己相関
γ_0	3.15271	3.0	3.19598	0.53795	0.02026
γ_1	0.48825	0.5	0.48978	0.07311	-0.02386
σ^2	11.6352	5	6.3793	1.62291	0.36947
v_λ		6	4.91121	2.70593	0.69444

$y_t = \gamma_0 + \gamma_1 y_{t-1} + u_t$ $u_t \sim t(0, \sigma^2, v_\lambda)$, v_λ の採択率:0.79633, OLSによる(N-AR(1))決定係数=0.26228

表 2 : 決定係数と尤度の比較(MCMC の推定値を利用)

項目	決定係数	対数尤度	対数周辺尤度
t -AR(1)	0.26410	-256.755	-257.985
N-AR(1)	0.26227	-261.452	-265.214

表 3 : 密度の推定値

項目	$\ln\pi(v_\lambda^* y)$	$\ln\pi(\sigma^{2*} y, v_\lambda^*)$	$\ln\pi(\gamma^* y, v_\lambda^*, \sigma^{2*})$
t -AR(1)	-3.58401	-1.06211	6.24203
項目		$\ln\pi(\sigma^{2*} y)$	$\ln\pi(\gamma^* y, \sigma^{2*})$
N-AR(1)		-1.37375	1.75595

表2に、MCMCによる推定値を利用して計算した決定係数、対数尤度、対数周辺尤度を示してある。真のデータ生成過程である、誤差項が t 分布に従う自己回帰モデルの決定係数、対数尤度、対数周辺尤度が、誤差項が正規分布に従う自己回帰モデルのそれと比べて大きいことがわかった。これはモデルと整合的である。ベイズファクターは3.13967であった。

6. 結論

本論文ではChib=Jeliazkov(2005)の提案した棄却サンプリング連鎖法を用いた場合の周辺尤度の推定方法を紹介した。その後で、誤差項が t 分布に従う自己回帰モデルを例に挙げ、棄却サンプリング連鎖法を用いてパラメータ推定した場合の周辺尤度の推定方法を示した。

その後でパラメータを与えて、誤差項が t 分布に従う自己回帰モデルからシミュレーション・データを生成した。そのデータに対して、棄却サンプリング連鎖法を用いてパラメータをベイズ推定し、決定係数、対数尤度、対数周辺尤度も計算してみた。

またモデル選択と比較のために、 t 分布に従う自己回帰モデルの推定に用いたシミュレーション・データに対して、誤差項が正規分布に従う1階の自己回帰モデルを当てはめて、決定係数、対数尤度と対数周辺尤度を計算してみた。

その結果、真のデータ生成過程である、誤差項が t 分布に従う自己回帰モデルの決定係数、対数尤度、対数周辺尤度は誤差項が正規分布に従う自己回帰過程のそれと比べて大きいことがわかった。これはモデルと整合的であり、周辺尤度の有用性が示された。

参考文献

- [1] Chib,S.,1995, "Marginal Likelihood From the Gibbs Output," *Journal of the American Statistical Association*, Vol.90,pp1313-1321.
- [2] Chib,S. and I.,Jeliazkov,2001, "Marginal Likelihood From the Metropolis - Hastings Output," *Journal of the American Statistical Association*, Vol.96,pp270-281.
- [3] Chib,S. and I.,Jeliazkov,2005, "Accept-reject Metropolis-Hastings Sampling and Marginal Likelihood Estimation," *Statistica Neerlandica*, Vol.59,pp30-44.
- [4] Geweke,J.,1993, "Bayesian Treatment of the Independent Student-t Linear Model," *Journal of Applied Econometrics*, Vol.8,ppS19-S40.
- [5] 伊庭幸人他,2005,『計算統計Ⅱ マルコフ連鎖モンテカルロ法とその周辺』,岩波書店.
- [6] 刈屋武昭・矢島美寛・田中勝人・竹内啓,2003,『経済時系列の統計』,岩波書店
- [7] Koop,G.,2003,*Bayesian Econometrics*, John Wiley & Sons.
- [8] 中妻照雄,2003,『ファイナンスのためのMCMC法によるベイズ分析』,三菱経済研究所.
- [9] 大森裕浩,2001,『マルコフ連鎖モンテカルロ法の最近の展開』,日本統計学会誌,第31巻4号,pp305-344.
- [10] 砂田洋志,2005, "誤差項に t 分布を仮定した閾値自己回帰モデルのベイズ推定,"山形大学紀要,第36巻第1号,pp77-96.
- [11] 砂田洋志,2006, "周辺尤度の推定理論とその応用—ギブズ・サンプリング法、及びM=H法で推定した場合—,"山形大学紀要,第36巻第2号.
- [12] 和合肇,2005,『ベイズ計量経済分析』,東洋経済新報社.
- [13] Watanabe,T.,2001, "On Sampling the Degree of Freedom of Student's t Disturbances," *Statistics & Probability Letters*, Vol.52,pp177-181.
- [14] 山本拓,1988,『経済の時系列分析』,創文社.

Estimation of Marginal Likelihood, When We Use Accept-reject Metropolis-Hastings Sampling

— Auto Regressive Model with t Distribution Error Term Case —

Hiroshi SUNADA

(Department of Law, Economics and Public Policy)

we introduce Accept-Reject Metropolis-Hastings(AR-MH) Sampling and explain how to estimate the parameters of the auto regressive model, whose error terms obey t distribution, by the AR-MH sampling.

Then we introduce how to estimate marginal likelihood, when parameters are estimated by AR-MH Sampling. As example, we simulate the data from the auto regressive model, whose error terms obey t distribution, and estimate both the parameters and marginal likelihood.